

Information and Communication Sciences E-book Series



Acts of the International
Colloquium on Open Access
Third edition, 2018

Open Access to Science: Foundations, issues and dynamics Le Libre Accès à la science : Fondements, enjeux et dynamiques

Editrice
Nozha Ibnlkhayat

Ecole des Sciences de l'Information
Maroc

Institut Supérieur de Documentation
Tunisie

Maghrebien Information Processing Society
MIPS - Maroc

ISBN: 978-9920-36-568-0

Dépôt légal: 2018MO4831

Open Access to Science: Foundations, Issues and Dynamics

حرية النفاذ إلى العلم: الأسس والرهانات والديناميكيات

Libre Accès à la Science : fondements, enjeux et dynamiques

Publication de la 1^{ère} version de l'E-book : Novembre 2018

Publication de la version complétée de l'E-book : Octobre 2019

Publication de la version papier : Décembre 2019

Les opinions exprimées dans la série n'engagent que leurs auteurs.

La série est protégée par la licence Creative Commons  BY-NC-SA :

- Toute reproduction ou traduction, même partielle, de l'œuvre ou de l'une de ses composantes, est conditionnée par la citation de la source et de(s) nom(s) d'auteur(s) [BY] ;
- Aucun profit commercial ne peut en être tiré sans autorisation des responsables de l'édition [NC] ;
- Tout partage devrait être fait à l'identique [SA].

Information and Communication Sciences

E-book Series

Editrice

Nozha Ibnlkhayat

ISBN : 978-9920-36-568-0

Dépôt légal : 2018MO4831

Information and Communication Sciences: E-book Series

3rd International Colloquium on Open Access



Open Access to science: Foundations, Issues and Dynamics

Le libre accès à la science : Fondements, enjeux et dynamiques

Rabat, 28-30 Novembre 2018

Organisateurs



Coordinateurs

Nozha Ibnlkhayat

Wahid Gdoura

Editrice : Nozha Ibnlkhayat

Editeurs associés : Ahmed Abdelilah Bachr

Adnane Benchakroun

Ounsa Roudiès

Comité scientifique
3e Colloque International sur le Libre Accès « icoa2018»

Abdallaoui Maan Najia, ESI, Maroc	GamouhNadjia, UC2, Algérie
Al Ghamidi Saad, URS, Arabie Saoudite	Gdoura Wahid, ISD, UM, Tunisie
AmrousNaila, ESI, Maroc	Ghaouti Loubna, UL, Canada
Anwar Adil, EMI, UM5R, Maroc	Guédon Jean Claude, UM, Canada
Bachr Ahmed Abdelilah, ESI, Maroc	Heather Joseph, Scholarly Publishing and Academic Resources Coalition (SPARC) U.S.
BamoflehFaten, URA, Arabie Saoudite	IbnlkhayatNozha, ESI, Maroc
Belhamel Kamel, DOAJ&U. Bejaïa, Algérie	Jabir Jamila, UIB, Liban
Ben Romdhane Mohamed ISD, UM, Tunisie	Jabr Naeema, USK, Oman, Sultanat Oman
Bensebti Abdelmalek, UC2, Algérie	Kadri Med Salah, ISD, UM, Tunisie
BouazzaAbdelmajid USK, Sultanat Oman	Ksibi Ahmed, ISD, UM, Tunisie
BouderbaneAzzeddine, UC2, Algérie	Kouassi Kouakou Sylvestre, EBAD, Sénégal
BoukacemChérifa, Lyon 1– ELICO, France	LarivièreVincent, UM, Québec, Canada
BounabatBouchaib, ENSIAS, UM5R/MIPS	Lrhoul Hanae, ESI, Maroc
Chartron Ghislaine CNAM-INTD, France	Mahmoud Saloua, ISD, UM, Tunisie
Cherif Walid, ESI, Maroc	Mikram Mounia, ESI, Maroc
Chiadmi Dalila, EMI, UM5-Rabat, Maroc	MkadaZghidiKhansa, ISD, UM, Tunisie
Daoudi Najima, ESI, Maroc	Olijhoek Tom, Editor-in-Chief, DOAJ
Diara Mamadou, EBAD, Sénégal	Ouazzani Khadija, ENSMR/MIPS. Maroc
DjeradNajoua, ISD, UM, Tunisie	Piron Florence, DIC, U. Laval, Canada
EleuldjMohcine, EMI, UM5R, Maroc	Rhanoui Meryem, ESI, Maroc
Elmqadde Nouredine, ESI, Maroc	RoudièsOunsa, EMI, UMVR, MIPS, Maroc
El Yousfi Chaminadas Hanane, ESI, Maroc	Sagna Olivier, EBAD, Sénégal
EttakiBadia, ESI, Maroc	Schopfel Joachim, U. Lille 3, France
FarajAbderahmane, U Beniswife, Egypte	SghirYousra, ISD, UM, Tunisie
Fayad El Mostafa, ESI, Maroc	SouissiNissrine, ENSMR/MIPS, Maroc
Fechtali Taoufiq, ESI, Maroc	Stephan Maud, FID, ULB, Liban
Fenniche-Daoues Rajae, ISD, UM, Tunisie	Zerouali Hayat, ESI, Maroc

Partenaires



UNIVERSITÉ
LAVAL



Information and Communication Sciences

E-book Series

Volume 1

Proceedings/Actes



Open Access to science:

Foundations, Issues and Dynamics

Impression HCP

Comité scientifique de l'E-book « icoa2018 »

- Abdallaoui Maan Najia, ESI, Maroc
- Bachr Ahmed Abdelilah, ESI, Maroc
- Bel Hamel Kemal, DOAJ, Nord de l'Afrique
- Ben Romdhane Mohamed ISD, UM, Tunisie
- Chartron Ghislaine CNAM INTD, France
- Gamouh Nadjia, UC2, Algérie
- Gdoura Wahid, ISD, UM, Tunisie
- Guédon Jean Claude, UM, Québec, Canada
- Heather Joseph, SPARC, USA
- Ibnlkhayat Nozha, ESI, Maroc
- Kouassi Kouakou Sylvestre, EBAD, Sénégal
- Olijhoek Tom, Editor-in-Chief, DOAJ
- Piron Florence, DIC, UL, Québec, Canada
- Roudiès Ounsa, EMI, UMVR & MIPS, Maroc
- Schöpfel Joachim, UL3, France
- Stephan Maud, UL, Liban

**Table des matières
(Communications en anglais et en français)**

Editorial: Nozha Ibnlkhayat v

Communications

Thème1 : Fondements théoriques et épistémologiques de l’Open Access

From Idea to Implementation: The State of “Open” 1
Heather Joseph

Géopolitique de l’open acces..... 7
Ghislaine Chartron

Articles, périodiques savants et plates-formes : repenser la publication savante dans un contexte numérique et dans un monde multipolaire..... 23
Jean-Claude Guédon

De l’économie du savoir à l’accès libre au savoir : Quelles limites ? 47
Mehdi Lahlou

Thème 2 : Aspects technologiques de l’Open Access

Mettre en pratique les recommandations sur les archives ouvertes de nouvelle génération de COAR pour l’archive ouverte institutionnelle, Horizon Pleins textes de l’IRD 67
Pascal Aventurier

Text data mining and data quality management for research information systems in the context of open data and open science 81
Otmane Azeroual, Gunter Saake, Mohammad Abuosba, Joachim Schöpfl

The role of the objects identification systems in ensuring permanent access to the digital content of the digital repositories:an analytical study 99
Behdja Boumarafi, Khaled Mettai

Thème 3 : Open Access et acteurs de l’IST

Open Access dans les bibliothèques universitaires arabes : Penser autrement l'accès à la science dans les pays du Sud..... 113
Wahid Gdoura

La naissance du DICAMES, l'archive ouverte du Conseil africain et malgache pour l'enseignement supérieur : récit de la mise en place d'une innovation sociale et technologique en Afrique francophone subsaharienne 129

Florence Piron, Djossè Roméo Tessy, Thomas Hervé Mboa Nkoudou, Antonin Benoit Diouf

Vers une stratégie de développement de la voie dorée marocaine : cas des revues en Sciences Humaines et Sociales 147

Hanae Lrhoul, Hayat Zerouali, Ahmed Abdelilah Bachr

Thème 4 : Open Access : évaluation et mesure

Analyse des articles scientifiques et libre accès à la littérature : Quels Bénéfices pour le Développement de la Science..... 173

Houda Sekkal, Naila Amrous, Samir Bennani

Le contrôle par les pairs ouvert : Où en sommes-nous ? 197

S. Hachani

Publication scientifique marocaine en computer science entre l'Open Access et l'accès payant : état des lieux et perspectives 231

Imane Rhassate, Naila Amrous, Samir Bennani

Open Access au Maroc : Vers un dispositif national d'Open Access collaboratif au profit de l'innovation..... 255

Majda Nabigh, Najima Daoudi

Indicateurs de consultation, indicateurs stratégiques : leur production à partir de l'analyse des consultations des documents d'une archive en libre accès..... 279

Pier Luigi Rossi

L'impact de la recherche en Tunisie entre webométrie et métriques alternatives 295

Hanen Gmara

Etude bibliométrique des publications sénégalaises sur les maladies non transmissibles : libre accès vs accès payant..... 323

Sylvestre Kouakou Kouassi, Djibril Diakhate

Publication scientifique en accès libre : Opinions et habitudes des chercheurs	349
Noureddine Elmqaddem	

Valoriser les publications d'un laboratoire universitaire dans l'environnement de la science ouverte Retour d'expérience de la collection GERüCO sur HAL	365
J. Schöpfel, H. Prost, A. Fraisse, S. Chaudiron	

Accès aux données ouvertes de la recherche : Vers l'application d'un système de recommandation optimisant le repérage de jeux de données libres	397
Amine Sennouni, Ahmed Abdelilah Bachr	

Le libre accès aux données de recherche : quelle politique de données pour les universités marocaines ?.....	415
Nadia Benaïssa, Saïda Belouali	

Le libre accès en Afrique de l'ouest francophone : état des lieux.....	435
Diéyi Diouf	

Thème 5 : Aspects didactiques, ressources pédagogiques ouvertes

Corrélation entre formation sur les ressources éducatives libres et culture du libre accès : retour d'expérience avec des enseignants universitaires en Tunisie	455
Sihem Zghidi, Haifa Belhadj Anane	

Analyse de l'offre des objets pédagogiques numériques dans les archives ouvertes.....	471
Mohamed Ben Romdhane, Rachid Zghibi	

Les MOOC dans le paysage universitaire tunisien : étude du cas de l'Institut supérieur de documentation	488
Rabii Djebbi, El-Khansa Mkada-Zghidi	

Annexes

Déclaration finale de « icoa2018 »	513
L'appel de Rabat en faveur de l'Open Access	517
The Rabat Call for Open Access	521

Editorial

Nozha Ibnlkhayat¹

La connaissance est intimement liée à la liberté ! Plus l'homme est libre dans la prospection de soi et de son environnement, plus se développe sa capacité à explorer la science et la technologie et à créer de nouvelles connaissances qui renforcent et/ou remplacent celles déjà existantes. De concert, plus la connaissance se développe et les espaces de liberté s'élargissent, plus les phénomènes de prohibition et de protection s'éclipsent. Ce sont là les fondements qui régissent les paradigmes renouvelés de l'« Open Access » et de l'« Open Science ».

Conjointement, si la science a été créée par l'homme pour contrôler et développer son être et son environnement, celle-ci le soumet à des processus, méthodologies et pratiques qui prennent la forme de « paradigmes » et de « lois » qui ne peuvent être changés que grâce à la création d'autres plus avancés. Et ce, en une dialectique menant à la génération de nouvelles connaissances, et appelant à la construction des équilibres nécessaires à la mobilité des sociétés et à la préservation de leurs droits à évoluer dans le respect de leurs spécificités culturelles.

Fait partie de ces processus la dialectique relative aux enjeux du monde virtuel qui bouleverse les parcours de la communauté scientifique et dont l'impact atteint les fondements épistémologiques de la production savante, de son évaluation et mesure, mais aussi de ses acteurs (décideurs, chercheurs, professionnels de l'information et de la communication, éditeurs, utilisateurs...).

Au-delà des acteurs, cet impact affecte les aspects éducatifs et renvoie à la formation à distance et aux mécanismes et ressources didactiques libres. Ce sont là quelques-uns des éléments clés qui régissent le nouveau paradigme

¹ Nozha Ibnlkhayat est professeure de l'enseignement supérieur en Sciences de l'Information. Mails: dr.nibnlkhayat@iulf.education
nibnlkhayat@yahoo.com

de l'accès à la science ; un accès libre certes, mais respectueux de la propriété intellectuelle et des licences qui la régissent.

Telle est la base des problématiques soulevées par le troisième colloque international sur le libre accès intitulé : "Le libre accès à la science : fondements, défis et dynamiques".

Inspiré par le paysage actuel de la société de la connaissance, cet intitulé s'inscrit dans un contexte caractérisé d'une part, par le contrôle exercé par la science et la technologie sur l'économie mondiale, et d'autre part, par le déplacement progressif des centres du savoir de l'Amérique du Nord et de l'Europe vers l'Asie du Sud-est [Chine, Japon, Corée du Sud], induisant la transformation des paradigmes de production et d'accès à la science, devenus ouverts et multilatéraux et engendrant des problématiques multiples :

- La première de ces problématiques est liée au volume de la production de l'information qui connaît une croissance phénoménale. IBM annonce en effet que l'humanité a produit plus d'informations entre 2014 et 2016 que ce qu'elle en a produit depuis sa naissance ! Cette information croît avec une cadence qui atteint les 45% par an, d'après IDC et se mesure par la zettaoctet. Cette dernière correspond à 10^{21} (autrement dit 10 suivie par 21 zéros!), ce qui induit des interrogations sur la maîtrise de ces flux. Quelles moissons, indexation, classification... en faire en vue de leur réutilisation ? D'où l'importance du métier/profession de l'informatiste dans sa version renouvelée ;
- La deuxième problématique concerne les fondements même de l'Open Access qui réfère, grosso modo, à la libre mise à disposition de contenus scientifiques numériques sans barrières financières, légales ou techniques. Mais, la mise à disposition engage ses propres concepts, approches et méthodologies qui envahissent le quotidien des communautés scientifiques induisant un bouleversement des processus de production et de communication de leurs œuvres. De nouveaux concepts tels que: la science ouverte, la société ouverte, la société de la connaissance, l'économie du savoir... apparaissent, se trouvent à

l'intersection de différentes disciplines et suscitent de nouveaux questionnements et de nouvelles dispositions.

Ces fondements se heurtent toutefois à deux courants de pensée : l'un, s'oppose à l'idée du libre accès à l'information et au savoir et l'autre, le soutient. S'appuyant sur une idéologie sociale qui prône l'égalité des droits et des devoirs, ses défenseurs estiment qu'il n'y a pas moyen de passer outre le libre accès pour contrecarrer l'exploitation des chercheurs, voire des utilisateurs par les éditeurs ;

- La troisième problématique est liée à la légalité des pratiques dans l'environnement virtuel : les violations des droits sont devenues inquiétantes pour leurs propriétaires et se répercutent sur l'arsenal juridique. Dès les années 1990 toutes les modifications introduites dans les processus et supports de lecture (directe ou en ligne), de l'écriture (individuelle ou collaborative) et d'accès (fermé ou partagé), ont donné lieu à des changements substantiels dans cet arsenal. Normes, lois, licences et sanctions se multiplient ainsi afin de préserver les droits et de lutter contre les violations. Ces efforts sont arbitrés aujourd'hui par les ayants droits, les avocats, les organisations et organismes régionaux et internationaux, et surtout par les maisons d'édition scientifiques, notamment celles commerciales...
- Ces maisons s'érigent en véritables monopoles de l'édition, et cela constitue la quatrième problématique soumise à la réflexion de *icoa2018*. Au nombre de quatre (RELX group, Spinger, Taylor & Francis et Wiley-Blackwell), toutes anglo-saxonnes, ces maisons détiennent 40 % du marché de l'édition, avec un revenu estimé à 15 milliards de dollars par an. Elles puisent leur force de leurs revenus colossaux, des prix qu'ils décrètent sur leurs ventes et abonnements, et de la primauté de la langue anglaise en tant que langue de production du savoir.

Le passage au numérique et au libre accès n'a pas diminué de la force de ces maisons qui se sont vite accaparées de la numérisation, des échanges sur le net et des logiciels libres pour éterniser leur domination sur la diffusion de la science.

La conscience quant à la démocratie liée à l'information et à la science et à leur accès libre, initiée par les appels des communautés scientifiques, semble ainsi ne pas être suffisamment ancrée. De cela atteste la remise en cause de la vision fondatrice de Google, qui se

trouve contredite par son classement comme deuxième plus grande capitalisation boursière depuis 2014 derrière Apple!

- Une autre problématique est liée cette fois-ci à l'impact économique du développement du libre accès et à ses modes de financement. Nombreuses sont les études qui attestent que celui-ci augmente l'efficacité et l'impact des auteurs et des institutions, avec des pourcentages variables, en fonction des disciplines et des contextes socioéconomiques. L'Open Access ferait ainsi réaliser un gain de 1,7 milliard de dollars à l'économie du Royaume-Uni et de 16 milliards de dollars à celle des États-Unis. Les difficultés constatées quant à son financement seraient transitoires, de l'avis des spécialistes, et pourraient être contournées par l'adoption des voies "verte" et "dorée" qui ont des ratios bénéfices/coûts positifs et potentiellement élevés ;
- La sixième et dernière problématique abordée concerne l'aspect éducatif : l'idée de l'Open Access s'est étendue pour couvrir la sphère de l'enseignement à travers le MOOC (Massive Open Online Courses) qui marque l'aube de la démocratisation de l'enseignement universitaire. Le Massachusetts Institute of Technology qui a créée en 2001, le MIT OpenCourseWare (MIT OCW) et proposé des ressources variées sous licence Creative Commons, a été suivie, en cela, par Stanford University puis par de nombreuses autres universités et instituts à travers le monde.

Les MOOCs se comptent actuellement par centaines de milliers dans différents pays, mais peu sont conçus au Sud, alors que ces pays disposent d'une masse importante de connaissances et que s'inscrire dans le mouvement, en tant qu'utilisateur ou producteur, ne requiert que peu de moyens (un matériel informatique et une initiation à son usage).

Il en résulte que les problématiques de l'Open Access sont nombreuses, variées et interreliées avec la révolution technologique et cognitive aussi bien qu'avec les droits publics, les questions économiques et celles didactiques... Tout cela dans le cadre de transformations politico-sociales exigeant davantage de droits pour les sociétés savantes certes, mais également pour le citoyen, appelé à contribuer concrètement à la vie sociale, politique et économique par le biais de l'accès à la connaissance et aux innovations.

Les pays des suds (le pluriel est de F. Piron) ne peuvent rester à l'écart de l'Open Access. Son statut, la conscience de son importance, les ressources mises à sa disposition et les modalités de son application varient selon le pays, la discipline et le secteur. Nombre de ses acteurs s'interrogent encore sur ses concepts de base ou sur ses pratiques! Il en est de même de la numérisation, de l'évaluation des publications, ou encore de la légitimité des taxes qui conditionnent son existence. D'autres questionnements restent en suspens pour ces pays, ils concernent les revues libres, leur rédaction, analyse, indexation... ainsi que les types de documents et le niveau de qualité nécessaire pour être placés dans des archives ouvertes, qu'elles soient personnelles ou institutionnelles.

Il s'agit là des problématiques et questionnements abordés par « icoa'18 » dans sa troisième édition, avec le but de promouvoir l'Open Access à la science par la mise en évidence de ses fondements épistémologiques, ses avancées, mais aussi des réticences auxquelles il fait face. L'ensemble de ces questions ont été traitées par d'éminents scientifiques venant de différentes disciplines et pays (Maghreb, Afrique subsaharienne, États-Unis, Canada, Europe, pays arabes).

Son programme s'est décliné quant à lui en six sessions thématiques et près d'une cinquantaine conférences et de communications. Ceci en plus d'une table ronde sur "Les spécificités des problématiques et processus de l'Open Access dans les pays des Suds" ; d'un atelier de formation sur "La création des revues électroniques en Open Access et leur indexation par le (DOAJ)" ; et d'un workshop sur "La Stratégie de publication en Open Access par les chercheurs".

L'approfondissement des débats sur ces thématiques, permettront, nous l'espérons, d'accompagner les décideurs, les institutions de recherche, les professionnels de l'information et de la communication des pays des Suds, dans la découverte des mécanismes leur permettant d'ouvrir vers une société de la connaissance humaine et intégrée.

Thème 1

Fondements théoriques et épistémologiques de l'Open Access

From Idea to Implementation: The State of “Open”

Presented by Heather Joseph, Executive Director, SPARC

Abstract

As the global movement promoting Open Access matures and the concept becomes more widely accepted, the conversation has shifted from whether or not it is a viable concept to how to best to implement it. With growing pressure from funders to move more quickly towards Open Access, there are many points of friction. This paper will focus on three immediate and important areas of current debate: ensuring that Open

Access truly provides equity in participation in scholarship; enabling the full vision of open scholarly communication, and catalyzing meaningful change in incentives and rewards for open scholarly communications practices.

Keywords. open access, geopolitics, international strategy, southern countries, genealogy, political, economic model, epistemic culture.

When the concept of Open Access was first articulated in 2002 in the aspirational vision of what was possible. They wrote: “An old tradition. An old tradition and a new technology have converged to make possible an unprecedented public good. The public good they make possible is the world-wide electronic distribution of the peer-reviewed journal literature and completely free and unrestricted access to it by all.”

The intention was clear from the outset that Open Access itself was not the end goal, but rather a means to an end. The Declaration went on to describe the vision of opening up access to this crucial layer of information in order to...”accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge...”

The authors rightly recognized that current system of scholarly communication has structural problems that extend far beyond unacceptably high subscription costs. Restrictive reuse policies, proprietary technical platforms, lack of transparency, bias in the review process, dominance of a single language – all are issues that the use of open as enabling strategy are designed to tackle. The Declaration’s emphasis on including all voices in this “common intellectual conversation, ” and in a bi-directional way, is particularly notable. The explicit intention of Open Access has always been to ensure not just equity in access to this information, but equity in contributing to it.

This idea is particularly important today, as research highlights the consistent dominance of voices from the global north in the scholarly journal literature. Of course this is not because there no worthwhile scholarship being produced in these regions; but rather the scholarship is not appearing in the journals that have come to dominate the market. That essentially means it is scholarship that is in written format - not oral stories, or other forms; it is generally written in English, and it addresses topics of interest to those who publish the dominant journals - primarily in Europe and North America. Topics that may be of deep interest and importance locally have little place in the “global” literature. This is deeply and dangerously reinforced by current systems of incentives and rewards,

which directly reward scholars for publication in “named” journals, “international” journals, and journals that carry high impact factors - and journals that also generally require high price tags to access.

This is highlighted in a particularly compelling way in a visualization of the “World Scaled by Number of Documents with Authors by Country,” prepared by Juan Pablo Alperin and colleagues in the Simon Fraser University ScholComm Lab, (<http://jalperin.github.io/d3-cartogram/>) which starkly illustrates the current imbalance in voices of scholars represented in the top journals today.

Addressing this challenge is one of the most immediately and important ones we face in working towards making the full vision of Open Access a reality, and its why the organizing committee of International Open Access Week very deliberately chose to highlight this issue as its central theme for 2018. By selecting “Designing Equitable Foundations for Open Knowledge” as the central focus, the committee intended to spark conversations about the for all stakeholders in the scholarly community to deliberately consider the implications on equity when we are making decisions on:

- What business models we choose to support scholarly communication
- What technology we choose for communications platforms
- What rights we choose to empower users and producers with
- What outputs and behaviors we choose to incentivize and reward

This is particularly critical in the context of global initiatives that are emerging, organized around the desire to accelerate the number of Open Access articles made available. Initiatives like “OA 2020,” promoted by the Max Planck Society (and to a certain degree, the original articulation of the European funder mandate, “Plan S”) focus on the concept of increasing Open Access article content with minimal disruption of the underlying system. They make the argument that by leaving all other elements unchanged and simply “flipping” the current model of payment from paying for access via subscription to paying for costs of publication through article processing charges (APC’s), the volume of OA articles can be quickly and dramatically increased.

This increase in OA articles has its appeal, to be sure. But it also begs the question: what do we actually gain if the same basic system - with the same inherent flaws – remains in place? We are still left with a market-based system centered on articles and journals as the primary (and perhaps, sole) unit of currency, and incentives and rewards built on the same. We will have the same largely commercial players in power, with the same geographic inequities institutionalized under a different business model. What would be - or would there be - the impetus for deeper change?

This begs the question: Do we want to simply “flip” the current system we have from one supported by subscriptions to one supported by Article Processing Charges (APCs)? Or do we want to truly transform the ways scholars produce and share their work? SPARC argues that the answer to this question should be the latter, and the global discussion should be centered on envisioning a system that moves beyond journals and articles as the only valued contributions, and pursuing a way to use the networked digital environment to foster and support an inclusive, ongoing conversation of scholarship.

For decades, we’ve been focusing on scholarly publishing - with the static, fixed object of an article as the ultimate expression of value to the exclusion of any other contributions to the conversation. Isn’t it time we move away from solely supporting scholarly publishing and into supporting true scholarly communication?

There are increasing signs that this kind of vision is beginning to take root among a very important constituency - research funders. They are mandating progressively more stringent Open Access policies, arguing that that communication of results is an essential, inextricable component of the research process itself, and asking what is the value of funding research if it isn’t shared as quickly and widely as possible?

They are supplementing these policies with investments in their own communications platforms – to host articles, yes - but also preprints, data sets annotations, visualizations, methodologies, algorithms and more a whole host of new contributions that can be accessed in real time by anyone, used...built upon. And ultimately, valued in new ways. From the Wellcome Trust, to the Gates Foundation the European Commission to the

African Academy of Sciences, these new platforms are coming on line in increasing number.

This development intersects quite nicely with the vision that is also being articulated in our libraries and in our institutions by proponents of “Next Generation Repositories, “ who are making the case for transforming existing institutional repositories from storage spaces into workspaces; into infrastructure that can be shared cost effectively across institutions and disciplines, widely enabling knowledge production and knowledge sharing. There is a growing will to start to provide this infrastructure upon which the conversation of scholarship can take place.

While some worry that these kinds of open platforms will result in a chaotic jumble of research objects of questionable value and validity. But perhaps this also provides us with an opportunity to reimagine the role of the communities that one would logically turn to help make sense of things – scholarly societies and similar disciplinary guilds. They are uniquely positioned to provide exactly the kind of valuable services they currently do in the journal environment, but on a much broader scale. Registration, validation, prioritization – the essential components of peer review that are now confined to being applied to only journal articles, could be applied to these new environments - and on different research outputs of all kinds, potentially in real time.

This also gives us the opportunity approach the question of what kind(s) of business models could be deployed to support scholarly communication. It would broaden the consideration from worrying about how not how to continue to support articles at current revenue/cost levels, and focus it on but how to facilitate a transition for societies to take their unique value-add, and create new, direct fee-for-service models that could be deployed in this environment. Societies have been wrestling with value proposition question in the digital era, and this seems to me to be a potentially very fruitful one.

It also opens up the opportunity to explore potentially powerful new stakeholder collaborations to underwrite such a transition among funders, institution/libraries, and scholarly societies. It would be useful to look to successful working examples of models from the era of transitioning society

publishing from the print to electronic environments for potential guidance. Models like BioOne, for example, were successful in raising Venture capital from libraries; providing risk mitigation strategies for both library investors and publishers, incentives for investors, and valid, viable exit strategies. Could be made the entire more powerful if funders are active partners in supporting the infrastructure; and institutions can be actively engaged.

Along with providing financial support, institutions have the opportunity - and the incentive - to be partners in critical culture change. Research dollars from funders is critical to their operations, and they want to make sure the channels facilitating that flow are as frictionless as possible. And if funders want researchers to put research outputs of all kind on these platforms, institutions will be incentivized to find ways to make it happen. That could mean revisiting the practices of relying on journal branding and impact factors as proxies for quality and impact.

None of these challenges are easy to address, but with the growing momentum towards open coming from so many diverse stakeholder communities, we have the unique opportunity to ask and to be deliberate about answering – with the full range of powerful outcomes that the Open Access originally challenged us to consider. It’s an opportunity we should not waste.

Géopolitique de l'open access

Geopolitics of open access

Ghislaine CHARTRON, Professeur du CNAM
(DICEN-IDF, EA 7339)
CNAM 2 rue Conté, 75003 Paris
Ghislaine.chartron@lecnam.net

Résumé

Cette communication commence par rappeler des repères de contextualisation du développement de l'open access. Elle insiste sur le brouillage qui s'est progressivement installé concernant les finalités du mouvement. La dimension internationale est ensuite considérée comme une entrée majeure, à la fois pour la consolidation de l'open access, pour analyser les déséquilibres potentiels et pour comprendre les stratégies de pouvoir sous-jacentes. Le texte appuie un renouveau du pilotage par les communautés scientifiques afin de redonner du sens au mouvement. Les principaux enjeux pour la recherche régionale des pays du Sud sont également débattus.

Mots-clés. Open access, géopolitique, international, stratégie, pays du sud, généalogie, politique, modèle économique, culture épistémique.

Abstract

This communication begins by giving context markers for the development of open access. We insist on the scrambling that has gradually taken place regarding the aims of the movement. The international dimension is then considered as a major entry, both for the consolidation of open access, to analyze potential imbalances and to understand the underlying power strategies. The text supports the control by the scientific communities in order to restore meaning to the movement. The main issues for regional research in Southern countries are also debated.

Keywords. Open access, geopolitics, international, strategy, southern countries, genealogy, political, economic model, epistemic culture.

1 L'Open access dans un cadre élargi

Le mouvement open access est bien sûr à comprendre dans le nouvel écosystème installé par Internet et le Web : travail coopératif, outils partagés potentiellement universels, communication mondiale facilitée, circulation directe de données et de connaissances. Cette dimension technologique s'est renforcée d'idéologies, de philosophiques politiques qui ont projeté des modèles de société radicalement différents. A la société industrielle, hiérarchisée, brutale devait succéder une société plus horizontale, plus transparente et collective installée par les TIC et couplée à la dimension du réseau (Curien et Muet, 2014). La société de l'information est alors caractérisée comme une société dominée par l'immatériel et le savoir, source d'innovation, de transformations des rapports sociaux et de croissance globale. « *Le réseau constitue la nouvelle morphologie sociale de nos sociétés* », Manuel Castells s'est attaché à reconstituer dans ses différents ouvrages, les nouvelles formes d'organisation des activités économiques, du travail, évoquant plus largement une révolution majeure de notre rapport à l'espace et au temps (Castells, 1998).

La transformation du capitalisme des sociétés industrielles en capitalisme informationnel insiste sur l'importance que revêtent aujourd'hui les mécanismes de droits de propriété attachés à l'information, aux logiciels, aux infrastructures de distribution numériques. Les logiciels libres, et plus largement l'open source prend alors ses racines en opposition à cette nouvelle forme de capitalisme, privilégiant le développement de biens communs numériques et informationnels, dans la vision d'une société plus solidaire et créatrice. « *Avec l'apparition du numérique, les 'créations' se détachent lentement de leurs supports matériels. Images, musique, mots et algorithmes sillonnent la planète jour et nuit, devant les yeux écarquillés des marchands. L'exode du savoir conduit à une terre promise à bien des bouleversements. Tandis que des armées de juristes s'interrogent sur la manière de pouvoir 'vendre des idées', une rumeur s'élève laissant entendre qu'elles doivent être libres comme l'eau, libres comme l'air, libre comme la connaissance.* » (Blondeau et Latrive, 2000).

L'open access scientifique co-émerge avec ces débats élargis qui ont lieu dès le début des années 90, annonçant de profondes transformations sociétales, transversales, impulsées par le numérique en réseau. Les discours, les militantismes se sont alors croisés et renforcés mutuellement entre différents champs d'activités. Outre ce contexte général, il faut souligner que l'open access, et l'open science, se sont également nourris d'autres registres plus spécifiques : le registre des valeurs éthiques de la science telles que définies par Merton (transparence de la science, partage des connaissances), celui de la science des données assimilé à un nouveau

paradigme de découverte (big data, fouille des données) et celui de l'innovation considéré comme porteur d'une croissance et d'un renouveau économique (Chartron, 2018).

2 Open Access : 30 ans et un brouillage progressif

Nous résumons ici quelques étapes marquantes argumentant la vision d'un brouillage qui, selon nous, s'est progressivement installé dans ce mouvement ; nous nous appuyons sur une précédente (Chartron, 2016).

L'archive ouverte ArXiv en 1991 mais aussi des projets antérieurs comme le projet Gutenberg lancé par Michael Hart en 1971, la revue électronique *Psychology* de Stevan Harnad en 1989, ou encore la revue *PACS-Review (Public-access Computer Systems Review)* lancé par Charles W. Bailey marquent les débuts de l'open access avec l'idée d'un accès ouvert aux savoirs, s'émancipant en grande partie des acteurs traditionnels de l'édition. On pourra se référer à la chronologie de Peter Suber¹ pour le contexte nord-américain et à celle de Hans Dillaerts et d'Hélène Bosc² pour le contexte français. Cette première période du début des années 90 est marquée par l'activisme de chercheurs pionniers, militants, visionnaires, de sensibilité libertaire (Harnad, Odlysko, Guédon...).

Parallèlement, la crise des coûts inflationnistes de certaines revues en sciences, techniques et médecine s'accélère, la révolte est en gestation et fortement pilotée par les bibliothèques américaines dont les dépenses documentaires sont les plus importantes au niveau mondial. La coalition SPARC (*Scholarly Publishing and Academic Resources Coalition*) est créée en 1998. Les bibliothèques entrent en scène en consolidant leurs actions au travers de nouvelles associations.

A partir des années 2000, le mouvement s'organise de façon politique à l'international. C'est l'étape des textes fondateurs, militants, partagés³ : PLOS en biomédecine (2001), Budapest (2002) définit les deux voies du libre accès, Berlin-Max-Planck (2003). L'affrontement du mouvement est alors de plus en plus vif avec les éditeurs commerciaux, qui parallèlement mûrissent une offre numérique plébiscitée par une majorité de chercheurs peu réceptifs à la question des coûts du système éditorial.

En 2015, après une période d'expérimentations foisonnantes mais souvent peu durables, une nouvelle étape s'annonce marquée par une reprise du

¹SUBER Peter, Timeline of Open Access Movement, <http://legacy.earlham.edu/~peters/fos/timeline.htm>

²DILLAERTS, Hans, Bosc, Hélène. *Chronologie/Historique du Libre Accès en France depuis 1999*. <http://open-access.infodocs.eu/chronologie-du-la-en-france/>

³ Textes de référence sur le blog de l'Inist : <http://openaccess.inist.fr/?-Textes-de-references->

pilotage par les gouvernements et leurs institutions, notamment la Commission européenne. L'OA est clairement identifié comme une brique du système d'innovation, les résultats de la recherche financée doivent circuler pour l'économie, et pour la société de l'information. La vision libertaire laisse place à une vision libérale, et les brouillages commencent à s'accroître. La 12^{ème} conférence de Berlin⁴, conférence emblématique qui avait consolidé le mouvement international par sa déclaration de 2003, se tient en décembre 2015 dans une configuration inédite, à savoir sur invitation (alors qu'elle rassemblait jusque-là les militants les plus engagés), elle réunit 90 représentants d'organisations scientifiques en provenance de 19 pays différents. L'objectif est d'accélérer le mouvement en prônant un système de communication scientifique nativement ouvert, le programme OA2020 est scellé par la signature d'une manifestation d'intérêt commune en faveur d'une implémentation à grande échelle des revues en accès libre immédiat. Pour parvenir à cet objectif, l'équation consiste à convertir les budgets actuellement dépensés pour les abonnements en fonds pour la publication en libre accès, engageant l'ensemble des institutions scientifiques et des financeurs de la recherche au niveau international, une sorte de modèle *gold international*. La vision développée s'est appuyée par une analyse chiffrée en avril 2015 par la Max Planck Digital Library (MPDL) (Schlimmer et al, 2015). La multiplication des négociations de type *offsetting* des dernières années (reconduction de budgets incluant à la fois l'abonnement aux revues et le coût des APC (*Article Processing Charge*) pour les chercheurs locaux) converge avec cette orientation mais l'adhésion des pays n'est pas totale et le raisonnement vaut essentiellement pour les sciences de la nature et bio-médicales.

Le pilotage au plus haut niveau politique s'est accéléré dans l'espace européen avec l'initiative de septembre 2018 concernant le « Plan S »⁵ portée par l'association Science Europe réunissant un ensemble de financeurs européens, plan appuyé par la Commission européenne. L'initiative soutient la construction d'un système nativement ouvert comme OA2020, mais excluant toute formule d'abonnement, écartant dès lors les revues hybrides des modèles possibles.

Ces reformulations progressives de l'open access ont conduit à brouiller les cartes car si l'objectif n'a pas changé : « accéder rapidement aux résultats de la recherche », les modalités pour y parvenir n'ont pas cessé d'évoluer, fluctuant notamment entre une volonté de renversement d'acteurs en place à des négociations renouvelées pour bâtir un système éditorial nativement ouvert avec des coûts contrôlés. Les paradoxes sont nombreux et les

4 <http://www.berlin12.org/>

5 https://www.scienceurope.org/wp-content/uploads/2018/09/Plan_S.pdf

stratégies menées sont parfois devenues illisibles pour beaucoup, et notamment pour les chercheurs car leur communauté semble avoir perdu le pilotage du système au profit d'une vision politique et technocratique surplombante, animée par une reconquête de la maîtrise de la communication scientifique. Cette vision est prioritairement ciblée pour la « recherche financée sur projets » investissant des budgets considérables en sciences de la nature, médicales et technologiques, de façon moindre pour les sciences humaines et sociales. Ceci a cependant provoqué une généralisation rapide des postures et une injonction à l'accès immédiat sur toutes les publications semble, pour beaucoup, l'horizon à atteindre, perturbant considérablement l'équilibre des revues de nombreux champs de recherche (Chartron, 2018).

3 Questions géopolitiques au centre des débats

La géopolitique est une discipline qui s'interroge sur les rapports entre espace et politique, à l'intersection de la géographie, des sciences politiques et des sciences militaires, qui traite des relations entre pouvoir et territoire. Rosière (2003) la définit comme une discipline spécialisée dans « *l'étude de l'espace considéré comme un enjeu* ». La géopolitique renvoie donc aux questions de stratégies de pouvoir des Etats et des acteurs socio-économiques et aux questions des relations internationales. Nous avons déjà sollicité cette entrée pour rendre compte de l'économie des données, des déséquilibres, et des asymétries à l'œuvre dans la collecte et la valorisation des données personnelles au niveau des grandes plateformes internationales (Chartron, Broudoux, 2015). De la même façon, il semble pertinent de considérer qu'il y a désormais une problématique prioritaire de géopolitique de l'open access, prenant en compte les politiques des différents pays et les déséquilibres potentiels dans l'objectif de déployer un open access large, durable, partagé au sein des communautés scientifiques.

3.1 L'alignement des politiques et des modèles économique (vs) l'alignement au sein des communautés

Les déséquilibres potentiels pourraient se traduire par exemple par le fait que l'accès ouvert ne puisse concerner que les chercheurs de certains pays pour une discipline, que l'accès ouvert ne puisse être mis en œuvre que par des acteurs publics, ou uniquement dans certains périmètres linguistiques. Le système pourrait alors devenir très déséquilibré, discriminant et appauvri. Des initiatives d'harmonisation sont nécessaires, mais plusieurs logiques sont alors envisageables...

- Alignement des politiques et des modèles économiques :

Cet alignement tend à vouloir s'appuyer sur un modèle commun alors que des équilibres différents (types de recherche, type d'éditeurs) caractérisent les champs scientifiques. L'accélération observée ces dernières années pour parvenir à un open access total en 2020 a conduit à des prises de position assez brutales, souvent généralisées sans précaution. Cette vision fut souvent celle des acteurs politiques et celle des financeurs de la recherche relayée également par des postures militantes. Elle a commencé avec des arbitrages peu discutés sur la durée des embargos pour la voie verte (6 mois ou 12 mois exclusivement, pourquoi ?) et elle s'affirme aujourd'hui par le manifeste OA2020 et par le Plan S. Le manifeste OA2020 mise prioritairement sur le modèle APC sans évaluer les risques d'inadéquation à certains champs scientifiques, à certains chercheurs et le risque de verrouillage potentiel qu'il peut installer autour de quelques revues. Le Plan S des financeurs européens, de la même façon, fait l'hypothèse de fonds APC disponibles pour tous, rejetant le modèle des revues hybrides. La recherche sur financement par projet pourrait alors devenir la seule recherche à être reconnue et à pouvoir être diffusée et partagée sans barrière.

- Alignement dans les communautés :

Une autre logique est possible même si elle peut apparaître complexe en termes de négociations, il s'agit de privilégier le raisonnement par communauté scientifique (le périmètre pouvant être élargi au cercle d'échanges significatifs entre chercheurs). La majorité des chercheurs travaillent à l'échelle internationale, un accès ouvert ne peut donc s'arrêter aux frontières des Etats ou des grands blocs politiques, sinon au risque de fractionner l'espace scientifique et l'espace de communication associé. A l'échelle de gros projets tels que les neurosciences, le réchauffement climatique, ces barrières sont déjà levées au sein de périmètres négociés (voir les grandes plateformes de travail de type <https://www.humanbrainproject.eu/> ou <http://www.copernicus.eu/>). De façon plus complexe, le dispositif SCOAP3 en physique opère sur une mutualisation mondiale de la communauté scientifique grâce à la médiation du CERN :« *Contrairement aux autres modèles Open Access où les bibliothèques, les auteurs ou les agences paient des frais pour chaque article, les éditeurs participants au SCOAP3 ont un seul contrat avec le CERN qui centralise les frais de traitement des articles pour les services de publication fournis. Les éditeurs à leur tour éliminent ou réduisent les coûts d'abonnement à tous les clients. Les partenaires SCOAP3 redirigent les fonds précédemment utilisés pour payer les abonnements dans un pool commun qui est ensuite utilisé pour les paiements des APC de façon centrale. Cela garantit des opérations*

efficaces et permet de fixer de manière compétitive un tarif d'APC tout en utilisant principalement des fonds déjà disponibles dans le système de publication »⁶. Mais il est vrai que très peu de communautés sont structurées par un si petit nombre de revues centrales (11 revues) et peuvent s'appuyer sur un acteur médiateur comme le CERN.

Les chercheurs qui travaillent sur les mêmes objets, ne souhaitent pas être mis en tension entre des politiques différentes selon leur pays d'appartenance, particulièrement à l'heure des coopérations de recherche de plus en plus nombreuses, des co-signatures d'articles de plus en plus étendues. De la même façon, les acteurs éditoriaux opérant dans des espaces scientifiques spécifiques, ont besoin de repères stabilisés pour continuer à développer leurs services et continuer à innover. L'instabilité actuelle n'est particulièrement pas propice aux petits acteurs, ni à l'initiative privée.

Les modèles vers l'open access durable et soucieux de maintenir en priorité des critères de qualité et de diversité nécessitent des débats concertés avec les acteurs qui portent les revues, les chercheurs qui y publient et qui les lisent. Les équations pourront certainement ainsi être trouvées pour tendre vers un open access optimum mais il ne peut être que le résultat d'une auto-détermination des communautés. La publication doit rester au service de la recherche et s'adapter aux besoins des champs scientifiques.

3.2 Relations internationales : le poids respectif des pays producteurs d'articles scientifiques

Par ailleurs, quel que soit l'alignement international visé, le poids relatif de chaque pays a un rôle décisionnel pour transformer le système de façon durable. Si, par exemple, les Etats-Unis ou la Chine ne suivent pas les mêmes politiques que l'Europe, de nombreuses communautés se trouveront en tension dans leur mode de communication. A un niveau général, le rapport de l'OST-HCERE (2018) donne une vision globale des rapports mondiaux dès lors qu'il s'agit de considérer les acteurs majeurs, les grands producteurs d'articles dans le système de communication scientifique. La figure 1 compare la croissance globale du nombre de publications entre 2000 et 2015 dans les 20 premiers pays producteurs d'articles scientifiques. Dans cet intervalle de temps, les Etats-Unis, le Royaume-Uni et l'Allemagne ont préservé leur position dans les premiers rangs mondiaux. Certains pays émergents ont connu une forte croissance et ont progressé vers les premiers rangs mondiaux : la Chine est passée du 8. rang mondial au 2., l'Inde du 12. au 6., la Corée du Sud 13. rang au 9. et l'Iran 20. au 16.. Inversement, la

⁶source : <https://scoap3.org/faqs/>

France est passée du 5^e au 7^e rang, le Canada du 6^e au 10^e, l'Espagne du 10^e rang au 11^e, et surtout le Japon du 2^{ème} rang au 5^{ème} rang.

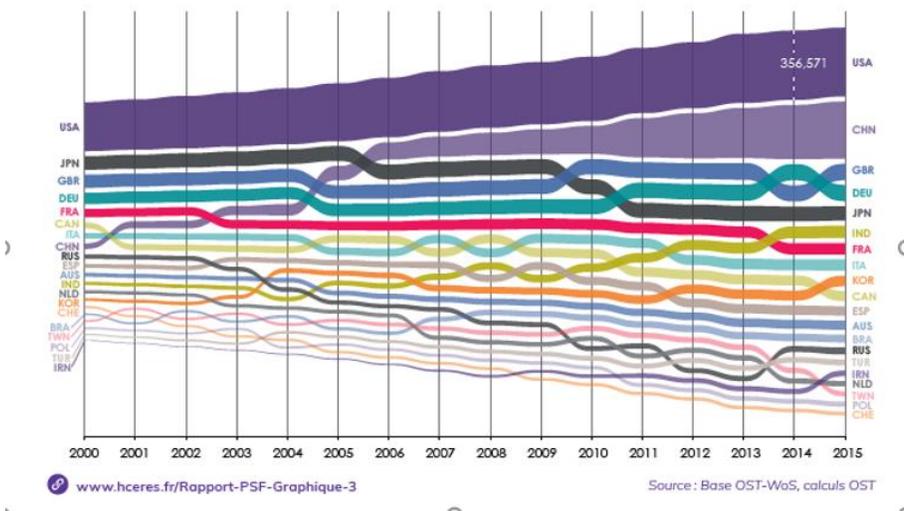


Figure 1 : Nombre de publications des 20 premiers pays, 2000-2015

La transformation des revues et de leur modèle économique nécessite donc un dialogue prioritaire avec les américains, les chinois, les britanniques et les allemands, leaders dans la production d'articles scientifiques. Les autres pays peuvent difficilement revendiquer une politique nationale spécifique qui ne serait pas concertée avec ces puissances, s'ils veulent continuer à œuvrer dans la recherche internationale.

A un autre niveau d'analyse, il conviendrait de regarder ces répartitions pour chaque communauté ou pour chaque périmètre qui constituerait un espace affirmé d'échanges scientifiques (par exemple les SHS francophones). Chaque périmètre afficherait très certainement des poids relatifs des pays qu'il faudrait prendre en considération pour construire l'open access de la science.

3.3 La différence des délais d'embargos dans les politiques publiques

Concernant les questions géopolitiques, la question des différents délais d'embargos autorisés selon les pays est apparue comme un déséquilibre parfois problématique lors des coopérations d'auteurs, ou pour des éditeurs

de revues travaillant majoritairement à l'international. La politique européenne s'est affirmée jusqu'à présent comme la plus exigeante en terme d'ouverture des publications par rapport à celle menée aux Etats-Unis et en Chine.

- *La politique européenne* : par la recommandation de 2012, cette politique soutient la voie verte et la voie dorée, fixe des embargos à 6 mois en STM et 12 mois en SHS au maximum. La nouvelle initiative du Plan S (4 septembre 2018), appuyée par la Commission promeut désormais un OA immédiat, sans embargo, rendant caduque le modèle hybride. Par ailleurs, les Etats membres ont aussi décliné différemment les recommandations de la Commission pour cet alignement 6/12 mois jusqu'à présent, on pourra consulter des analyses plus fines à ce sujet (Gayoso, 2018).
- *La politique américaine* : le Memorandum de l'*Office of Science and technology (OSTP)* du 22 février 2013 cadre la politique américaine jusqu'à présent, demandant à chaque agence fédérale (de + 100 millions de dollars de budget) de finaliser son plan open access. L'embargo ne doit pas dépasser 12 mois. Des mesures budgétaires supplémentaires ont été fournies par les agences en ce qui concerne le paiement des APC. Le système est moins centralisé qu'en Europe, chaque agence de la recherche a établi son plan et le service Chorus⁷ est un outil au service du suivi global. Pour les sciences humaines et sociales, seule l'agence *Department of Education Institute for Educational Studies (DOE/IES)* est concernée par une politique d'open access.
- *La politique chinoise* : en 2014, la Fondation nationale des sciences naturelles de Chine (NSBC) et l'Académie chinoise des sciences (CAS) ont mis en place un mandat obligatoire de dépôt avec un embargo de 12 mois au maximum pour tous les chercheurs. L'alignement se fait sur le modèle américain. La politique européenne d'open access immédiat (OA2020) commence à être suivie avec des négociations de type *offsettings* évoquées précédemment.

3.4 Des stratégies d'influence et identitaires adossées à l'open access

Croiser le mouvement open access et la dimension internationale conduit aussi à s'interroger sur le « dessous des cartes », sur les calculs opérés sous

⁷ <http://www.chorusaccess.org/resources/us-agency-public-access-plans/>

couvert d'ouverture au monde. On pourrait faire l'hypothèse, d'une part, d'une stratégie d'influence propre aux pays développés et, d'autre part, d'une stratégie identitaire des pays en développement.

- *Stratégie d'influence ou d'un « néo-colonialisme » renouvelé* : dès lors que la recherche des chercheurs d'une zone politique circulerait librement dans l'ensemble des Etats, on pourrait y déceler la stratégie d'une nouvelle forme de domination. Le rôle pionnier des grandes universités américaines dans divers mouvements *Open*, notamment le développement des MOOC, serait aussi à analyser de façon critique dans cette mise en perspective. Il en va de même des programmes d'accès subventionnés aux grandes plateformes internationales de revues scientifiques pour les pays africains, captant les attentions sur des problématiques de pays développés.
- *Stratégie d'une reconnaissance identitaire à l'échelle internationale* : inversement, l'open access peut être considéré comme une opportunité inédite de mise en visibilité d'une recherche locale, régionale jusque-là peu considérée, écrasée par les grands canaux de distribution en place et privilégiant des problématiques des pays développés. On peut citer ainsi la mobilisation des pays hispanophones avec l'affirmation de la plateforme Scielo et nous évoquerons plus loin la plateforme africaine AJOL.

3.5 Corollaire : complexité et déséquilibre pour l'activité éditoriale

Dans ce dernier point, c'est sur la complexité de la transformation du modèle économique des éditeurs que nous voudrions insister. Il ne s'agit pas ici de revenir sur la question du juste prix des publications scientifiques, largement débattue par ailleurs.

A l'échelle internationale, la mutation des modèles économiques par des financements amont (APC) installe une équation compliquée pour maintenir les recettes, car cette transformation se fait inégalement selon les pays, selon des calendriers variables, conduisant à maintenir simultanément la modalité de l'abonnement pour les pays défavorables aux APC. La crainte du double paiement des revues est à l'origine du développement des politiques d'*offsettings* mais le paysage est encore très fragmenté car l'engagement n'est pas total, ne serait-ce qu'au niveau des différents pays européens ne partageant pas la même vision⁸. Par ailleurs, l'asymétrie des politiques publiques en terme d'open access peut conduire

⁸<http://esac-initiative.org/offsetting/>

à des paradoxes au niveau national que nous avons déjà soulignés : **acheter des ressources étrangères et simultanément offrir au monde sa production nationale en supportant seul les coûts de la transition de la production locale ; affaiblir certains secteurs éditoriaux nationaux du fait des pressions exercées alors que ces pressions n'existent pas dans d'autres pays ; libérer les articles des auteurs uniquement rattachés à des pays engagés dans des politiques d'off-settings** (Chartron, 2018).

4 Développement durable pour les pays du Sud

Comme souligné précédemment en évoquant les « stratégies identitaires », l'open access est une fabuleuse opportunité pour les pays du Sud et les pays en développement pour afficher, faire reconnaître la spécificité de leurs recherches. Les travaux menés dans la thèse d'Hanae Lrhoul sur le terrain de la recherche médicale au Maroc ont par exemple mis en évidence les enjeux locaux sur des questions de santé publique prioritaires telles que la tuberculose, les hépatites, les leishmanioses, les méningites, les toxoinfections alimentaires collectives et la mortalité infantile (Lrhoul, 2017), mais ces sujets intéressent peu l'actualité des grandes revues internationales.

De nombreux projets numériques liés à la valorisation de cette recherche locale, notamment ceux subventionnés par la Commission européenne, sont restés sans suite, une fois passée la phase d'impulsion financière (Kane, 2018). D'autres semblent vouloir s'intégrer de façon plus pérenne dès lors qu'ils sont pris en charge par des acteurs locaux. C'est notamment le cas de la plateforme AJOL (African Journal Online) qui, aujourd'hui, donne accès à 500 revues scientifiques de 31 pays africains, le projet initié par l'INASP en 98, est porté désormais par une association à but non lucratif basée en Afrique du Sud. Le constat est évident : les chercheurs africains ont besoin de partager les recherches menées sur leur propres continents, ces publications peinent à être intégrées dans les publications des pays développés, elles sont sous-utilisées, sous-valorisées, peu visibles et donc peu citées sur le terrain même de l'Afrique. AJOL travaille à développer cette identité avec des filtres scientifiques requis dans les communautés scientifiques.

Car, effectivement, l'exigence de qualité et de rigueur reste par ailleurs prioritaire. La parole rapportée de chercheurs évoque souvent ce manque de confiance dans les revues nationales (Lrhoul, p.213) « *En ce qui concerne les critères de choix des revues de publication, les entretiens avec les chercheurs de la Faculté de Médecine Pharmacie ont permis d'identifier deux clusters : le premier concerne les jeunes chercheurs qui confirment avoir publié dans des*

revues nationales au début de leur carrière. Cela constitue pour eux un apprentissage à la rédaction scientifique et une préparation à la publication dans des revues indexées. Le deuxième cluster, regroupe les PES (professeur d'enseignement supérieur) qui ne publient que dans les revues indexées et refusent catégoriquement de collaborer dans les revues nationales. Cette dernière catégorie juge que les comités de rédaction acceptent les publications sans validation... ». De nouvelles revues en open access sont aussi initiées fréquemment dans des pays en développement saisissant l'opportunité de publier des articles locaux, mais leur qualité reste problématique dans bien des cas, « Les revues indiennes en libre accès ont soumis depuis 2014, 2578 requêtes d'inclusion dans le DOAJ ; le Brésil en a soumis 2.048, tandis que l'Indonésie se classe première, avec 3.662 demandes. Tom Olijhoek, rédacteur en chef du DOAJ, a déclaré à SciDev.Net que près de la moitié des candidatures sont rejetées, généralement en raison de leur faible qualité »⁹.

Il s'agit donc de consolider ces revues, avec l'exigence d'un contrôle de qualité comparable aux revues établies. Pour le service AJOL, un effort soutenu a été conduit pour certifier un niveau de qualité des revues hébergées, ainsi le cadre de référence « Pratiques et les Normes pour la Publication dans les Revues » (JPPS) fournit des critères d'évaluation détaillés pour la qualité des pratiques d'édition des revues du Sud et est utilisé pour évaluer les revues hébergées sur la plates-forme (INASP & AJOL., 2017). De la même façon, l'intégration de Scielo Citation Index dans le *Web of Science en 2014 a permis d'augmenter la crédibilité et la visibilité des revues et des recherches de l'Amérique latine et des Caraïbes* (Vélez-Cuartas et al., 2016).

Toutefois, le problème de la valorisation des sujets de recherche locaux dépasse aussi le seul constat d'un manque de revues nationales ou régionales. Les enquêtes de terrain mettent en évidence un décalage plus profond : « *le choix des thématiques des chercheurs n'est lié ni aux priorités nationales de la recherche, ni aux problématiques réelles de la santé au Maroc, ni aux recommandations d'organismes internationaux comme l'OMS. Il est plutôt lié aux projets de collaborations qui les incitent à travailler sur des thématiques à intérêt international, pour publier dans des revues indexées dans Scopus et le WoS.* » (Lrhoul, 2017, p.207). Le constat soulève donc deux autres problèmes de fond : le manque de financement de la recherche locale pour être attractive et celui des mesures d'évaluation qui reproduisent aveuglément les standards des pays développés.

⁹ Scidev.net, 27/09/18, Contrôle de qualité du libre accès, Fatima Arkin, <https://www.scidev.net/afrique-sub-saharienne/edition/article-de-fond/libre-acces-qualite.html>

5 Conclusion : l'open access doit rester open

Le constat qui est actuellement fait d'une lente progression vers l'open access, 30 ans après le début du mouvement, devrait donc s'interroger sur les raisons des résistances développées. Le virage politique au plus haut niveau qu'a pris le mouvement depuis ces dernières années semble avoir sous-estimé des valeurs prioritaires pour les chercheurs : la qualité, la rigueur, les marques de confiance construites au fil des décennies avec des équipes de revues. La dimension « open access », bien que majeure et souhaitée par les chercheurs, ne peut balayer les autres valeurs du système de communication scientifique.

Dans cette communication, nous avons insisté sur les risques de déséquilibre que pourrait provoquer un open access brutal, non réfléchi pouvant mettre en tension les chercheurs. Si des alignements sont nécessaires pour faire aboutir un accès ouvert le plus large possible, il ne s'agit très certainement pas de penser que les mêmes modèles seront applicables à tous les champs scientifiques. L'open access doit donc rester open en accord avec les « Cultures épistémiques » (Knorr-Cetina 1999). Mais les équations sont complexes, les négociations internationales d'actualité, et les stratégies d'influence bien présentes.

Les mouvements ne sont pas finis, quelle en sera l'issue ? Une attention particulière est aujourd'hui donnée à des secteurs majeurs (médecine, ingénierie, environnement...) sous l'impulsion des financeurs de la recherche, bien décidés à basculer vers des modèles de diffusion qu'ils souhaitent pleinement contrôler. Des ralliements internationaux au Plan S sont déjà en marche dans le secteur biomédical, par le Wellcome Trust britannique et la fondation américaine Bill & Melinda Gates... à suivre.

Références

- BLONDEAU, Olivier ; LATRIVE Florent (2000). *Libres enfants du savoir numérique. Une anthologie du "Libre"*. Editions de l'Éclat, « Hors collection », 2000, 504 p. ISBN : 9782841620432. URL : <https://www.cairn.info/libres-enfants-du-savoir-numerique--9782841620432.htm>
- CASTELLS, Manuel (1998). *L'Ère de l'information. Vol. 1, La Société en réseaux*. Paris, Fayard, 1998.
- CHARTRON, Ghislaine (2018). *L'Open science au prisme de la Commission européenne*. In *Education et sociétés*, 2018/1 (n° 41), p. 177-193. DOI : 10.3917/es.041.0177. URL : <https://www.cairn.info/revue-education-et-societes-2018-1-page-177.htm>

- CHARTRON, Ghislaine (2016). Stratégie, politique et reformulation de l'open access. In *Revue française des sciences de l'information et de la communication* [En ligne], 8 | 2016. DOI : 10.4000/rfsic.1836.
URL : <http://journals.openedition.org/rfsic/1836>
- CHARTRON, Ghislaine ; BROUDOUX, Évelyne (2015). Enjeux géopolitiques des données, asymétries déterminantes. In *Big Data - Open Data : Quelles valeurs ? Quels enjeux ? Actes du colloque « Document numérique et société », Rabat, 2015*. Louvain-la-Neuve, De Boeck Supérieur, « Information et stratégie », 2015, p. 65-83. DOI : 10.3917/dbu.chron.2015.01.0065. URL : <https://www.cairn.info/big-data-open-data-quelles-valeurs--9782807300316-page-65.htm>
- CURIEN, Nicolas ; MUET, Pierre-Alain (2014). *La société de l'information*. Rapport pour le Conseil d'analyse économique. 310p. La Documentation française. 2014. <https://www.ladocumentationfrancaise.fr/rapports-publics/044000180/index.shtml>
- GAYOSO, Emile (2018). *Analyse comparée des politiques d'open access dans l'Union européenne*, pour le comité de suivi de l'édition scientifique, MESRI, <http://www.dicen-idf.org/wp-content/uploads/2018/09/Diaporama-Etat-des-lieux-d%E2%80%99open-access-en-UE-Gayoso-E..pdf>
- INASP & AJOL. (2017). *Guide pour encadrer les Pratiques et Normes pour la Publication dans les Revues (JPPS)*. [En ligne] <https://www.journalquality.info/en/wp-content/uploads/sites/3/2017/09/INASP-JPPS-Standards-Guide-FR-Digital.pdf>
- KANE, Khardiatou (2018). *Documentation numérique en Afrique : Evaluation de l'offre et des usages en sciences humaines à l'Université Cheikh Anta Diop de Dakar*, Thèse de doctorat (dir G. Chartron), Ecole doctorale Abbé Grégoire, Conservatoire national des arts et métiers, Paris, septembre 2018.
- KNORR-CETINA K. (1999). *Epistemic Cultures : How the Science Makes Knowledge*, Cambridge (UK), Harvard University Press, 1999.
- LRHOUL, Hanae (2017). *La production scientifique des chercheurs de la faculté de médecine et de pharmacie de Casablanca : mesures, cartographie et enjeux du libre accès*. Thèse de doctorat (dir G. Chartron), Ecole doctorale Abbé Grégoire, Conservatoire national des arts et métiers, Paris, décembre 2017, <https://tel.archives-ouvertes.fr/tel-01815122>
- OST-HCERES (2018). *Rapport sur la position scientifique de la France dans le monde, 2000-2015* http://www.obs-ost.fr/fr/indicateur/analyses_et_indicateurs_de_reference
- ROSIERE, Stéphane (2003). *Géographie politique et géopolitique*. Paris, Ellipses, 2003.

SCHIMMER R., GESCHUHN K. K., VOGLERA. (2015). *Disrupting the subscription journals' business model for the necessary large-scale transformation to open access*, doi :10.17617/1.3.

VELEZ-CUARTAS, G. ; LUCIO-ARIAS, D. ; LEYDESDORFF L. (2016). Regional and global science: Publications from Latin America and the Caribbean in the SciELO Citation Index and the Web of Science. In *El Profesional de la Información*, 25(1)

Articles, périodiques savants et plateformes : repenser la publication savante dans un contexte numérique et dans un monde multipolaire

Jean-Claude Guédon
Professeur honoraire, Université de Montréal

Résumé

Après la 2e Guerre mondiale, la publication savante s'est vue graduellement dominée par des entreprises commerciales multinationales. L'évaluation de la recherche, au lieu d'analyser les contributions de recherche, utilise des indicateurs fondés sur les citations de revues par d'autres revues. La concurrence entre individus, institutions et même entre pays a dès lors été arrimée à la compétition entre revues, telle qu'exprimée par des classements.

Dans le contexte numérique, les revues savantes n'ont plus grand sens et la « fonction éditoriale » n'est plus une, mais correspond à plusieurs fonctions qui peuvent se dissocier les unes des autres et se localiser parmi des acteurs nouveaux. La revue savante, pourtant, perdure.

Abstract

After the 2nd WW, scholarly publishing grew to be gradually dominated by multinational businesses. Research evaluation, instead of dealing directly with research contributions, uses indicators based on journal citations in other journals. The competition between individuals, institutions and even countries has been determined by the competition between journals as expressed through rankings.

In the digital world, scholarly journals lose much of their meaning and the publishing function turns out to be made up of at least four functions which can be dissociated from each other, and which can be associated with unfamiliar actors. Yet, the scholarly journal goes on. Platforms are similar to virtual shelves endowed with algorithmic possibilities. These allow the

La plate-forme est une sorte d'étagère virtuelle dotée d'algorithmes gérant les rapports entre documents, les rapports entre individus et documents, et les rapports entre individus. D'où son immense pouvoir. L'article et la revue numériques dépendent des plates-formes.

Pour corriger les problèmes que rencontrent actuellement les publications et communications savantes, il faut repenser la nature et le rôle des revues à l'intérieur des plates-formes. Il faut les dissocier de l'évaluation de la recherche. En fait, les plates-formes peuvent assez facilement correspondre à des territoires et leur inter-connexion permet de penser à une universalité de la connaissance qui se fonde sur la résolution de problèmes liés à ces territoires.

L'Amérique latine a ouvert la voie à cette nouvelle façon d'internationaliser la production des connaissances ; internationalisation qui s'oppose terme à terme à la mondialisation mono-centrée qui domine actuellement. Le monde arabe, par l'existence d'une unité que confère le partage d'une même langue, peut reprendre le modèle de l'Amérique latine.

management of documents with documents, of individuals with documents, and of relationships between individuals. This explains its immense power. Articles and journals depend upon platforms.

To correct the problems that scholarly communications and publications presently encounter, it is necessary to revisit the nature and the roles of journals as they exist within platforms. Journals must be disconnected from research evaluation. Platforms can be compared to territories and their inter-connections allows imagining a form of scholarly universality that also leaves room for the solution of territory-bound problems.

Latin America has pioneered this new approach to the internationalization of knowledge production. This kind of internationalization is directly opposed to the single-centre globalization that presently dominates world research. The Arabic world, thanks to the kind of unity conferred by a shared language, can make good use of the model proposed by Latin America.

Introduction

Depuis environ une génération, la réticulation des ordinateurs du monde par les protocoles de l'Internet est largement accomplie, et les effets de cette transformation technologique se révèlent de façon toujours plus intense par les perturbations qu'ils engendrent. Dans toutes les sociétés du monde, de nouveaux types de documents circulent de façon massive par des canaux inédits. Pensons, par exemple, aux photos prises par des téléphones portables, mais aussi aux commentaires brefs rapidement disséminés au sein de groupes qui ne recourent que partiellement (tout en les dépassant) les anciennes formes de granularité sociale (famille, amis, voisinages, professions, etc.), ou qui en créent de nouvelles : les « amis » de Facebook ne sont pas exactement les amis d'antan ! Et il en va autant pour les « voisinages » et autres repères de sociabilité qui structuraient jusqu'à récemment l'intersection des langues et des territoires. Les documents eux-mêmes diffèrent de ceux du passé ; leur production et dissémination s'appuient sur des processus entièrement nouveaux ; et, effet en retour, les liens entre être humains se tissent également différemment du passé, situation qui soulève des interrogations autant philosophique – la question de l'identité, par exemple – que politique – les recompositions des « partis » ou des « causes ».

Dans le monde de la recherche et de la production des connaissances, les perturbations sont tout aussi nombreuses et profondes. L'irruption des ordinateurs en réseaux bouleverse les chercheurs, leur travail, et leurs rôles de manière aussi radicale : la production des documents issus de la recherche s'est beaucoup transformée ; leur dissémination aussi ; la nature même de ces documents a profondément évolué, par exemple en incorporant données et logiciels ; et les relations entre chercheurs s'organisent sur la base de moyens, et selon des échelles inconnus jusqu'à ces derniers temps. Les chercheurs prennent aussi le contrôle de nouveaux moyens de communication qui s'étendent souvent au-delà des limites traditionnelles du monde la recherche. Ce qu'accomplit un chercheur, ce qu'il est, comment il se situe dans sa société et au-delà, telles sont certaines des questions fondamentales qui accompagnent ses activités. Les mots perdurent, offrant ainsi une certaine illusion de stabilité : la connaissance,

le chercheur, la revue scientifique maintiennent aisément leur présence dans les textes. Pour autant, et plus profondément, ces mêmes mots voient leur signification glisser tout en remettant en jeu l'ensemble des valeurs qui sous-tendent le projet essentiel de créer, ensemble, entre êtres humains, un stock toujours plus riche, plus prédictif, d'énoncés validés sur l'ensemble de la réalité.

Comment tracer une route dans un paysage aussi fluide constitue l'enjeu du petit texte qui suit. Plus précisément et concrètement, la conclusion tentera de suggérer des étapes concrètes pour le monde de langue arabe, perçu ici comme exemple d'un ensemble humain, culturel et linguistique susceptible de jouer un rôle important dans la constitution d'un processus de construction des connaissances impliquant l'ensemble de la planète, sans limiter ce processus à une extension, souvent décrite comme une mondialisation, du mouvement intellectuel né en Europe lors de la « Révolution scientifique » du XVII^e siècle.

Ce qui sera dit ici dans le contexte d'une communication en pays arabophone pourrait être aisément transposé dans le monde hispanophone (qui, d'ailleurs, nous allons le voir, fait figure de pionnier en la matière, soutenu par le monde, linguistiquement proche, des lusophones), les mondes chinois et hindi, etc. L'importance de ces groupes linguistique, une fois reconnue, conduit à imaginer une reconstitution du système de communication et de publications savantes multi-polaire, plus riche en diversité, en inclusivité, plus ouvert à la variété des préoccupations humaines sans pour autant voir sacrifiée l'universalité de l'effort collectif de création de la connaissance. Appelons ce mouvement « internationalisation » de la science, en opposition au mouvement de mondialisation qui, lui, repose sur une structure uni-polaire.

1. Les transformations silencieuses des publications savantes avant l'Internet

Les glissements silencieux du sens que l'on peut attacher au « périodique savant » n'ont pas attendu l'Internet pour se manifester. Dans le monde de la recherche, la période qui a immédiatement suivi la deuxième guerre mondiale a vu s'instaurer de telles transformations dans le monde de la publication savante. Longtemps dominée et contrôlée par les sociétés

savantes et les associations scientifiques, la publication savante, particulièrement dans les sciences de la nature, la médecine et l'ingénierie, a dû brutalement faire face à une croissance extrêmement rapide des activités de la recherche immédiatement après la guerre. Celle-ci avait déjà révélé l'importance de ces connaissances pour l'armement en général. Le démarrage de la « guerre froide » et la course à l'armement qui a suivi ont contribué à intensifier le financement public de la recherche scientifique de manière spectaculaire. Par voie de conséquence, le nombre des publications a également crû exponentiellement, avec le résultat que les sociétés savantes n'ont pas pu, ou su, comment répondre à cette brusque augmentation de la demande de publication de la part des chercheurs.

Cette incapacité des sociétés savantes à répondre aux besoins des chercheurs a ouvert la porte aux éditeurs commerciaux qui, jusque là, n'avaient jamais prêté une attention centrale aux possibilités de profits dans le monde des publications savantes : des textes ésotériques, souvent incompréhensibles pour tout individu non-spécialiste dans la question traitée, souvent diffusés seulement à quelques centaines d'exemplaires, n'offraient guère des perspectives lucratives évidentes. Aider les chercheurs à se repérer dans cette littérature avec des index, des « Abstracts », etc. paraissait une bien meilleure route pour un projet commercial bien conçu. Sauf que, dans les années cinquante, l'accélération de la demande de publication, et l'augmentation du nombre de chercheurs-lecteurs ouvraient de nouvelles possibilités y compris l'augmentation du prix des abonnements. Restait malgré tout un problème épineux : comment assurer la réputation de nouvelles revues.

En allant très, très vite, ce dernier problème ne fut réellement résolu à la satisfaction des maisons d'édition commerciales que grâce à la résolution d'un autre problème de la recherche lié au coup d'accélérateur de l'après la guerre : il s'agit en l'occurrence de la difficulté à s'orienter dans des champs de recherche de plus en plus interdisciplinaires avec les outils bibliographiques traditionnels. Les index et bibliographies classiques, parce que situés essentiellement dans les limites définies par les disciplines (et contribuant aussi à les renforcer), révélaient de plus en plus leurs carences. En suggérant de suivre les citations des articles scientifiques, Eugene Garfield reprenait en fait une idée empruntée au monde des juristes – dans ce dernier cas, les références à des jurisprudences – mais, du même

mouvement, mettait en place un nouveau dispositif de navigation mieux adapté aux besoins de l'interdisciplinarité. Mais les citations se comptent aussi, et cette manière de les aborder allait bientôt donner le moyen de construire des outils d'évaluation. De cette capacité de quantification émergea un tour de passe-passe intellectuel aux conséquences fondamentales : l'évaluation, jusque-là fondée sur la réputation et sur une analyse directe des contenus, se métamorphosa en classements. Ces classements se présentaient sur la base de nombres très précis, comme en témoigne l'usage général de trois décimales dans la publication habituelle du facteur d'impact. L'utilisation de trois décimales renvoie à une prétention de précision très élevée, et cette haute précision est souvent confondue avec la justesse de l'indicateur. Or, il suffit de penser à une balance très précise, mais fondamentalement fautive, par exemple par défaut de construction. Une telle balance peut apparaître très fiable, plus fiable en fait qu'une balance beaucoup plus juste, mais moins précise. La précision accordée au facteur d'impact ne garantit nullement sa justesse quand il s'agit d'estimer la qualité d'une recherche. Et même si l'on considère que le facteur d'impact est un indicateur, et non une mesure, il faut se souvenir qu'un indicateur indique quelque chose, quelque qualité, en général pour aider à prendre quelque décision. Si tel est bien le cas pour le facteur d'impact, nul besoin de trois, deux, ou même une décimale

En fait, par le biais de la quantification, l'évaluation des revues est subrepticement recadrée à l'intérieur d'une problématique générale de concurrence entre revues, elle-même gérée par un classement. Ce classement, on l'a vu, est établi de manière « rigoureuse », parce que quantitatif. Ce faisant, les maisons d'édition disposaient dès lors de moyens rhétoriquement efficaces pour établir la « qualité » de leurs revues. De plus elles savaient exactement ce qu'il fallait rechercher dans la gestion des revues : il fallait les faire progresser dans les classements obtenus par le truchement des facteurs d'impact.

Le système s'est complètement verrouillé quand les bibliothèques ont commencé à faire usage de la même métrique pour construire leurs collections de revues : vendeurs et acheteurs, en somme, s'étaient mis d'accord sur la façon d'évaluer à la fois la valeur commerciale, la valeur d'usage, et la valeur intellectuelle des revues. Mais ces trois valeurs

n'étaient plus abordées que sous l'angle de la capacité d'attirer des citations dans un club de revues choisies par l'organisme (privé) de Garfield.

Cette convergence entre les « valeurs » des revues a permis alors de moduler le prix des abonnements des revues les mieux classées en fonction du facteur d'impact. L'effet général a été la hausse générale des prix d'abonnements aux revues savantes. Finalement, les sociétés savantes ont suivi le mouvement, mais sans bien se rendre compte que leur passage à un régime financier copié sur celui des maisons commerciales les conduisait à se transformer subrepticement en maisons d'édition au comportement quasi-commercial. Du même mouvement, les revues des sociétés, jusque-là porte-voix de communautés de recherche, se sont effectivement transformées silencieusement en marchandise. Avant même l'arrivée de l'Internet « grand public », c'est-à-dire avant 1995 environ, la mutation était achevée et les maisons commerciales d'édition achevaient un double mouvement au terme desquels on les retrouve en position dominante et concentrées au sein d'une structure d'oligopole.

L'un des grands pionniers de cette métamorphose, Robert Maxwell, avait même tenté de prendre le contrôle de ISI, la compagnie fondée par Eugene Garfield, ce qui conduisit à une bataille légale qui dura des années. Cette bataille révèle surtout que les maisons d'édition avaient bien compris comment articuler valeur intellectuelle et valeur pécuniaire et voulaient simplement assurer leur contrôle le plus total possible des instruments nécessaires pour étayer une telle construction. Les évolutions récentes, tel Scopus créé par Elsevier, ou encore le rôle encore ambigu de Digital Science en relation avec Springer Nature confirment cette analyse.

2. Les transformations moins silencieuses des publications savantes en contexte numérique

Dans l'expansion du contexte numérique, l'extension graduelle de l'Internet correspond à la dernière phase du processus : l'ordinateur n'est plus seulement un outil pour calculer ; il devient aussi un élément intégral d'un système de communication dont les limites, potentiellement, sont celles de l'humanité. Toute l'économie de la communication et de la publications s'en est trouvée bousculée, transformée, avec des redistributions importantes de pouvoir entre les différents acteurs du

système, et aussi entre différentes régions du monde. Les communications et publications savantes n'ont évidemment pas échappé à cette révolution technologique, mais les spécificités du domaine se sont traduites par des modes d'appropriation des nouvelles technologies tout aussi spécifiques.

Depuis le XVII^e siècle, la notion de publication savante pouvait exprimer ses caractéristiques particulières en se référant à quatre fonctions fondamentales : l'enregistrement, la certification (souvent assurée par une forme ou une autre d'évaluation par les pairs), la dissémination et la préservation. Dans l'imprimé, ces quatre fonctions se regroupaient assez naturellement autour d'un seul acteur : les maisons d'édition. Seule la préservation, et ce à partir du XIX^e siècle, se divisait entre celles-ci et les bibliothèques. De cette agrégation des fonctions facilitée par l'imprimé est né le mythe de la « fonction éditoriale », fonction traitée comme ensemble insécable et intouchable.

La montée du numérique n'a cessé de montrer à quel point la « fonction éditoriale » était en fait très friable et pouvait être désagrégée pour se répartir autrement parmi les principaux acteurs du milieu. Une façon simple de montrer ceci consiste à pointer dans la direction de toutes les nouvelles maisons d'édition qui sont nées depuis le début de ce siècle. Loin de soutenir la thèse d'une fonction éditoriale indivisible, l'émergence du numérique renvoie en fait à tout un champ d'innovations et d'explorations où chaque phase du processus de publication est soumise à critique, révision, transformation et déplacement. Pour ce qui est des déplacements, certains éléments de la publication apparaissent effectivement en des sites inhabituels, voire déconcertants : des bibliothèques, des cercles de chercheurs, des organismes de financement de la recherche, etc. Les fonctions de publication retravaillées par la numérisation peuvent aussi se reconstruire sur la base de réseaux d'institutions, ou bien en tissant de nouveaux liens entre des parties d'une institution : les alliances entre presses universitaires et bibliothèques constituent d'excellents exemples de cette mutation. Dans d'autres cas encore, certaines fonctions de la publication vont se placer dans des sites tout aussi inédits : pensons aux éléments d'enrichissement des dépôts institutionnels qui sont testés par des réseaux de dépôts tel OpenAIRE en Europe, y compris des procédures d'évaluation. Pensons aussi aux épi-revues qui viennent s'installer au-

dessus d'archives en accès libre. Pensons enfin à la liaison qui s'établir entre ensembles de données et les articles publiés à partir de ces ensembles.

D'autres glissements peuvent être observés en contexte numérique : par exemple, la dissémination – anciennement un problème de messageries et de distribution – se voit remplacée par des questions d'accessibilité : accessibilité technique, légale ou financière : tout est désormais disponible partout à condition de disposer de la bande passante nécessaire, des droits d'accès et, dans le cas des abonnements aux revues, de l'argent pour financer l'accès. De plus, l'accessibilité effective requiert aussi la visibilité et donc la découverte. Cette question place dès lors les moteurs de recherche bien plus au cœur de la dissémination que ne l'ont jamais été les index et bibliographies d'antan. Or, les moteurs de recherche relèvent de compagnies, parmi lesquelles on retrouve évidemment Google, mais aussi les bibliothèques, surtout lorsqu'elles participent à un organisme comme OCLC et son Worldcat. Ces entités se situent loin du monde traditionnel de l'édition.

Les autres fonctions de la publication sont tout aussi sujettes à redistribution et reconfiguration.

- Est-il vraiment nécessaire d'attendre la parution d'un article dans une revue pour que l'article soit enregistré ? Ne serait-il pas plus simple et efficace de gérer cet enregistrement initial dans l'institution qui chapeaute les chercheurs ou leurs équipes ? Cela permettrait en plus aux institutions de recherche de mieux suivre la production effective de leurs chercheurs.
- Est-il nécessaire de lier étroitement la certification d'une contribution scientifique à sa parution dans un organe de publication particulier ? Sans grande imagination, il est concevable d'envisager une évaluation par les pairs organisée autrement que par des revues savantes et leurs éditeurs. Les mêmes chercheurs pourraient appuyer la tâche de la certification sur des bases institutionnelles autres, par exemple des réseaux de dépôts institutionnels.
- Plus fondamentalement encore, la notion de propriété qui relie de manière si constante articles et revues est-elle bien nécessaire dans les domaines savants ? Pourquoi une « revue » ne pourrait-elle pas

rediffuser dans un nouveau secteur de connaissances un travail initialement conçu dans le cadre d'un domaine particulier? Quand, au XIXe siècle, Dmitri Mendeleïev voyait ses articles traduits et publiés en allemand, sa réaction (et celle des revues russes correspondantes) n'était pas de s'offusquer d'une telle initiative, mais, bien au contraire de se réjouir d'une visibilité élargie et diversifiée.

Au sein de la montée du numérique, les transformations des fonctions de publication ne sont pas apparues en premier. Bien avant que l'unité factice de la fonction éditoriale ne fasse l'objet de remarques critiques, d'autres phénomènes, plus étroitement liés à la tradition de la publication héritée de l'imprimerie, se sont manifestés. Le libre accès lui-même est un excellent exemple de cet ensemble de phénomènes que l'on pourrait définir comme la réaction initiale au numérique. En effet, le libre accès, initialement conçu comme un libre échange d'information savante, n'a pas attendu le numérique pour être imaginé et désiré. Mais il fallait le numérique pour lui donner un tour et une réalité pratiques indispensables à son succès. Sans un coût marginal proche de zéro de la copie et de la transmission, le libre accès demeure en effet hautement problématique.

À peu près au moment où le libre accès se formulait de plus en plus clairement, soit vers la fin des années 90, les maisons d'édition commençaient aussi à comprendre comment ce coût marginalement négligeable de la copie et de la transmission pouvait conduire à des plans d'affaires innovants. En particulier, l'idée de créer des bouquets d'abonnements, ou « Big Deal », à toutes les revues d'une maison d'édition remplissait plusieurs objectifs:

- La possibilité d'offrir l'ensemble des revues d'une maison d'édition avec une forte ristourne est parfaitement compatible avec l'objectif d'augmenter les revenus de la maison d'édition : il suffit, acheteur par acheteur, de s'assurer que le coût du bouquet total est supérieur au coût de l'ensemble des abonnements qui précédaient la mise en place du « Big Deal » ;
- Gérer l'accès à un bouquet est plus simple que gérer un accès sur mesure, acheteur par acheteur, et donc moins coûteux ;

- Capter plus de revenus d'institutions dont les budgets comportent des limites très claires signifie que moins d'argent demeure disponible pour les maisons d'édition ne pratiquant pas le « Big Deal » ;
- Le « Big Deal » conduit également à une redéfinition du « marché » des revues : bien que structuré par des classements portant sur des revues individuelles (le facteur d'impact), les « Big Deals » tendent à gommer cette forme de concurrence pour, au contraire, avantager les revues appartenant à un bouquet, quel que soit leur classement, au détriment des petites maisons d'édition et, a fortiori, des revues individuelles : le « Big Deal » constitue en effet un élément fondamental du processus de concentration qui caractérise l'édition savante depuis la fin des années 90 et qui a conduit à une situation que beaucoup perçoivent comme l'instauration d'un oligopole d'éditeurs .
- Enfin, le « Big Deal » permet de contrôler les mouvements de défection qui pourraient se produire parmi les bibliothèques en les mettant en face de choix douloureux : conservez le « Big Deal » ou bien faites face à des coûts d'abonnements encore plus élevés au total si vous décidez de revenir à la constitution de collections de revues visant les véritables usages de vos communautés.

Une autre voie commerciale a vu le jour à peu près au même instant que le « Big Deal ». Elle aussi est liée au numérique, puisqu'elle repose sur le libre accès. Du lecteur, ou de la bibliothèque qui achète les abonnements, elle déplace les coûts de la publication vers l'auteur qui veut publier sa contribution au savoir, ou à tout organisme qui veut bien se substituer à l'auteur (ou les auteurs) : organismes subventionnaires, institutions de recherche (et en particulier les bibliothèques).

En comparaison avec les transformations subies par les revues savantes avant les années 90, ces derniers phénomènes ont été plus visibles, et donc mieux repérés par l'ensemble des acteurs impliqués dans la communication et la publication savantes. Ils ont eu pour conséquence un élargissement du débat avec, en particulier, l'implication nouvelle des chercheurs eux-mêmes, des organismes de subvention, et, dans une certaine mesure, des institutions de recherches. Le proverbe voulant que « à quelque chose,

malheur est bon » trouve peut-être là sa seule confirmation : l'élargissement du débat a également contribué à l'approfondir et c'est peut-être la raison derrière le rôle de plus en plus assuré qu'assument les organismes subventionnaires dans un domaine – la communication et publication savantes – où ils ne s'étaient jamais aventurés. Les débats et controverses qui accompagnent, par exemple, l'énonciation du Plan S en Europe, constituent un excellent indice de ce nouvel état de fait.

3. Repenser le rôle des périodiques savants : un avenir plein de passé

Des parties qui précèdent, quelques thèmes importants s'imposent : non seulement les maisons commerciales sont devenues dominantes dans le domaine des revues savantes, mais, dans le même mouvement, elles assoient une forme d'évaluation de la recherche qui, elle aussi, est devenue dominante : le facteur d'impact. Sans ce rôle dans l'évaluation de la recherche les revues savantes n'auraient pas pris ce nouveau rôle de « king maker » ; elles seraient encore ce qu'elles ont presque toujours été depuis le XVII^e siècle – l'organe, le moyen de diffusion et de projection, la voix en somme de diverses communautés de recherche. Par contraste, les revues indexées dans les grandes bibliographies, tel le Web of Science, ou Scopus, ne correspondent pas (ou plus) à des communautés de recherche, mais plutôt à des clubs conçus pour étayer des hiérarchies et sélectionner des élites.

L'intrication d'objectif commerciaux – quête de profits – et d'objectifs savants n'apparaît que dans la 2^e moitié du XX^e siècle et cet entremêlement du financier et du savant sert plutôt mieux le premier que le second et, de plus, il monte l'objet « revue » en épingle précisément au moment où l'on peut se demander si, dans le contexte de la numérisation, les revues ou périodiques doivent encore jouer un rôle aussi central.

La réponse à cette dernière question doit s'effectuer en deux temps. En toute logique, première perspective adoptée ici, la revue ne devrait plus avoir le rôle central que lui avait conféré l'imprimerie en l'inventant. Le bouquet d'articles initialement conçu pour mécaniser l'échange épistolaire et le rendre plus efficace n'a plus réellement raison d'être dans un monde où l'on navigue d'articles en articles en fonction des besoins de lectures et

des indications laissées dans les citations. Les DOI et autres CrossRef s'appuient sur les articles ; les grands moteurs de recherche, tel Google Scholar, pointent sur des articles et non des revues. L'information savante se disperse dans un nombre élevé d'organes de publication et les chercheurs utilisent tous les moyens possibles pour retrouver l'essentiel de l'information nécessaire. Avant l'Internet, cette quête prenait souvent la forme d'un repérage de revues choisies en fonction de leur pertinence thématique, du moins comme première étape d'une heuristique bien conduite, ou comme manière relativement fiable de se tenir au courant des dernières avancées dans la connaissance. On observe encore, de nos jours, ce mode de comportement dans certaines sciences humaines, précisément parce que l'imprimé joue encore un rôle significatif dans des disciplines comme l'histoire et la littérature. Depuis le déploiement de l'Internet, ces pratiques tendent à s'estomper pour se voir remplacées par le questionnement de bases de données conduisant directement à des articles.

Pour autant, l'objet « périodique savant » ne constitue pas que l'aboutissement nécessaire d'un raisonnement logique à un moment de l'histoire, et il ne disparaîtra pas immédiatement si ce raisonnement s'affaiblit ou s'estompe dans une période de mutation technologique affectant ses fonctions fondamentales. C'est ici que le rôle de premier plan, que les revues ont réussi à prendre pour l'évaluation de la recherche, constitue un élément fondamental pour envisager leur destin. En bref, les chercheurs impliqués dans la direction d'une revue savante cherchent à produire une revue de la meilleure qualité possible. Mais, depuis l'avènement du facteur d'impact, cette quête de qualité est largement assimilée au classement d'une revue dans un groupe de revues perçues comme concurrentes. Gagner un rang dans le classement correspond désormais à ce qui est largement interprété comme une hausse de qualité et les maisons d'édition font la promotion de leur revue en s'appuyant précisément sur ce critère.

Une fois parvenu à jouer le rôle de « King maker », la revue occupe une position au pouvoir redoutable : le facteur d'impact initialement prévu pour comparer des revues d'un même domaine s'applique rapidement aux chercheurs qui parviennent à publier dans ces revues. La hausse du facteur d'impact d'une revue entraîne une croissance du nombre des articles soumis, et donc un taux de rejets plus élevé, critère qui est ensuite avancé

pour justifier la haute qualité de la revue d'une autre manière. Mais pour arriver à se faire accepter dans une telle revue, le ou les auteurs d'un article soumis à cet organe de publication doivent se rappeler que la qualité de leur travail ne constitue pas le seul critère d'évaluation. Même excellent, et même situé dans le domaine où œuvre la revue, ce travail doit aussi pouvoir attirer les citations dont dépend la revue pour maintenir sa réputation. Un travail sur un sujet un peu éloigné des centres d'intérêt dominants des chercheurs du domaine, quelle que soit son importance ou son originalité – trop peut nuire – n'attirera pas les regards des chercheurs bien campés sur les sujets en vogue, et bien logés dans des institutions prestigieuses. Donc, le travail peut-être rejeté par simple estimation d'ésotérisme relatif, et ce parce que le classement de la revue est central, et, en fin de compte, détermine aussi le niveau de revenus qu'une telle revue peut engendrer.

L'arc des préoccupations qui va de la conception d'un problème jusqu'à la vente des revues à des grands ensembles de bibliothèques est continue, sans faille, et manifeste par là-même cette capacité d'intrication qui constitue le cœur de la publication savante telle qu'elle fonctionne en position dominante de nos jours. La position de la revue, en fait, est renforcée par ses effets sur les jugements portés sur des chercheurs, sur leurs laboratoires, leurs institutions de rattachement, et même, en fin de compte, leur pays. L'ensemble forme une structure qui n'est pas sans rappeler les poupées gigognes russes. Embauche, promotion, allocation de recherche, prix, etc. touchent les individus et sont gérés par des méthodes de jugement où le rôle de la revue (c'est-à-dire l'endroit où l'on publie) est plus important que ce que l'on publie. Les classements d'universités s'appuient aussi en partie sur le nombre de publications dans des organes « prestigieux ». Et enfin des pays entiers sont eux aussi classés en s'appuyant en partie sur ce qui est interprété comme l'extrait (« output ») scientifique de tout le pays. Évidemment, cette structure en poupées gigognes confère une solidité accrue à l'ensemble du dispositif « revues » et ceci explique la capacité de cet objet à maintenir sa position dominante dans les systèmes de publication et de communication savantes.

Au terme de ce raisonnement, on pourrait penser que la revue est inamovible et demeurera à jamais avec nous. Pour autant, le contexte de la numérisation la mine de l'intérieur, et cette situation permet d'imaginer comment ruser avec l'objet lui-même. En gros, nous suggérons de conserver

le nom mais d'opérer un glissement de sens de façon à ce que l'entité « revue », au lieu d'interférer avec le processus de communication et de publication savantes, le serve au contraire mieux, et ce en contexte numérique.

Pour imaginer le genre de revue dont la recherche contemporaine a besoin, mieux vaut repartir du passé des périodiques savants, et en particulier de leur ancien enracinement dans les communautés scientifiques. Instrument de communication savante, le périodique savant cherche avant tout à recueillir les meilleures contributions de ces communautés pour les disséminer aussi largement que possible auprès de communautés intéressées par ces travaux. Il suffit ensuite d'examiner comment les revues se sont adaptées au monde de la numérisation pour comprendre que, de nos jours, toutes les revues se situent sur une plate-forme, ou tendent à évoluer dans cette direction. Reste à voir ce que confère une plate-forme à une revue, et comment la revue peut évoluer dans ce contexte si caractéristique du monde numérique.

Dans l'histoire récente des revues électroniques ou numériques, le portail a clairement précédé la plate-forme. L'idée d'offrir un point d'entrée unique à un grand nombre de revues paraissait aussi évident que bâtir une bibliothèque. Rapidement, des regroupements de presses universitaires ont exploré cette voie et le projet MUSE de l'université Johns Hopkins aux États-Unis constitue probablement un des premiers exemples marquants d'un nombre important de revues accessibles à partir d'une seule adresse dans le Web.

Le portail ne constitue cependant que le premier pas dans une évolution qui aboutit à la plate-forme. Là où le portail ne constitue qu'une sorte d'étagère virtuelle unique et se comporte comme elle vis-a-vis du lecteur, la plate-forme se dote d'algorithmes qui permettent de guider la recherche du lecteur. La lecture d'un article particulier, par exemple, peut déclencher la suggestion d'autres articles, provenant d'autres revues de la plate-forme pour à la fois guider le lecteur et, dans certains cas, le maintenir occupé par la plate-forme. Dans ce dernier cas, l'intention correspond aux impératifs de ce que l'on appelle souvent « l'Économie de l'attention ».

La plate-forme se caractérise donc par les algorithmes qu'elle incorpore et qui, au total, permettent de gérer trois types de relations : des documents

aux documents, des lecteurs-utilisateurs aux documents, et, finalement, tout comme les réseaux sociaux, des relations entre chercheurs. La plate-forme, par conséquent, est dotées d'un pouvoir immense sur la structuration de l'information, sur ce que l'on peut faire avec cette information, et sur les êtres humains dans leurs relations réciproques.

Dans l'espace actif et abstrait que constitue la plate-forme, la revue apparaît pour ce qu'elle est vraiment : une manière de créer un regroupement de documents différente de ceux que les algorithmes peuvent générer. Vulnérable dès lors, jusque dans son existence même dans l'espace abstrait et virtuel des plates-formes, la revue, paradoxalement peut-être, redevenir ce porte-voix, cet organe de projection d'une communauté particulière qu'elle fut dans beaucoup de cas jusqu'à la 2e Guerre mondiale. Dans l'espace particulier que constitue une plate-forme, la revue n'a nul besoin de « posséder » des articles ; mieux vaut pour elle de se consacrer à les agréger en ensembles cohérents pour les chercheurs interpellés par les questions sous-jacentes.

En portant le regard au-delà de la plate-forme, on peut distinguer ou au moins envisager l'existence d'autres plates-formes. C'est à ce niveau que le libre accès peut prendre son plein sens : la revue, porte-voix d'une communauté, ne peut effectuer son plein rôle dans la navigation du lecteur que si elle est ouverte et accessible. Une simple attribution de source complétant l'attribution due aux auteurs suffit, ce qui signifie que tout article susceptible d'intéresser une autre communauté de chercheurs peut alors être rediffusé dans une nouvelle « revue » conçue selon le modèle présenté ici. Les revues, dans cette perspective, aident à mailler des communautés de chercheurs à l'intérieur d'une plate-forme, ou entre plates-formes, et non à assurer une propriété quelconque sur un ensemble d'articles.

Au final, une conclusion s'impose : dans le monde numérique, au niveau de la publication, ce qui compte désormais ce sont les plates-formes et la façon dont ces plates-formes se relient les unes aux autres pour former un réseau de systèmes de publications à l'échelle de la planète.

4. Repenser l'internationalisation de la communication et publication savantes en opposition à la mondialisation

Depuis quelques décennies, on parle de plus en plus de la mondialisation, et ce mouvement a, bien évidemment, touché les publications savantes de manière centrale. De fait, en nommant ses nouvelles revues, *The International Journal of...*, Robert Maxwell avait bien compris dans les années 50 qu'il fallait viser le niveau l'international pour mieux surplomber et dépasser le caractère largement national des sociétés qui alors dominaient encore la publication savante un peu partout. C'était aussi, évidemment, une nouvelle manière de concevoir le marché des revues, désormais résolument tourné vers tous les pays du monde, plutôt qu'orienté vers le marché national et quelques marchés voisins.

L'arrivée du Science Citation Index de Eugene Garfield entraîna une intensification de ce mouvement vers l'international tout en lui donnant un tour particulier. Garfield voyait en effet la nécessité d'unifier la langue scientifique et utilisa son instrument de repérage de citations pour faire levier sur les revues savantes de divers pays : celles-ci devaient offrir au moins des titres et résumés en anglais, et bientôt des textes complets dans cette langue. Par ailleurs, l'extension du nombre des titres de revues retenus par son index fonctionnait un peu en boule de neige, partant d'un centre très précis : les revues américaines, britanniques, et certaines revues de l'Europe occidentale. Le bénéfice de se voir inscrit dans l'index de Garfield devint suffisamment évident pour que l'effet de levier linguistique fonctionne : en quelques années, l'anglais devint effectivement la *lingua franca* des sciences. Anglais plus extension graduelle d'un centre, d'un noyau de revues anglophones conduisit à une forme d'internationalisation très particulière qui est largement la nôtre et que nous pouvons appeler mondialisation dans la mesure où ce dernier terme renvoie aussi à une forme d'internationalisation à partir d'un centre dominant.

C'est dans ce contexte qu'il faut tenter de repenser l'internationalisation de la production des connaissances. En effet, si le système actuel unifie bien l'effort de production des connaissances, en particulier dans les domaines scientifiques, il réussit presque trop bien. Que l'universalisme de Robert Merton soit parti de la possibilité pour tout humain de participer à l'effort de recherche, et que l'universalisme, plus généralement, avance l'idée qu'un

résultat scientifique validé est valide partout, voilà qui n'est guère gênant, bien au contraire. Mais la dynamique actuelle de la recherche mondialisée repose aussi sur un système d'évaluation qui crée des biais importants dans le choix des questions posées : si, pour se faire publier dans une revue prestigieuse, il vaut mieux entreprendre une recherche dans un domaine déjà bien balisé et occupé pour pouvoir espérer optimiser le nombre de citations reçues, et il vaut mieux appartenir à une institution bien connue, il en découle plusieurs conséquences importantes.

- Le système d'évaluation actuel, fondé sur le classement des revues par le facteur d'impact, conduit à favoriser la concentration des travaux de recherche sur un ensemble de questions ou problèmetemporairement privilégiés pour une raison ou une autre. Cela signifie également que nombre de questions, dont certaines très importantes, ont tendance à être négligées alors qu'elles peuvent être d'une très haute pertinence pour certaines régions et certains pays ;
- Le bénéfice associé à l'appartenance à une institution de recherche bien connue mondialement constitue un pôle d'attraction pour les chercheurs qui débutent leurs carrières dans des situations plus marginales ou périphériques par rapport a ces centres.

Ces deux axes d'analyse affectent évidemment la dynamique de la recherche dans tous les pays du monde en contribuant à renforcer les inégalités entre individus et institutions. À l'intérieur d'un pays, ce phénomène, du moins en théorie, peut être maîtrisé par une politique scientifique nationale adéquate. Mais entre pays, ces phénomènes tendent à s'intégrer dans les processus de concurrence que se livrent les états-nations. Ils se révèlent aussi comme instrument puissant pour drainer vers les pays les plus puissants et les plus riches des ressources intellectuelles dont le besoin est cruellement ressenti dans les pays les plus pauvres. En même temps, des problèmes très graves – pensons à la malaria, mais aussi au virus Ebola, etc. - qui peuvent dévaster des populations et des économies pendant de très nombreuses années,voire décennies restent à l'abandon sauf à se révéler menaçants pour les pays riches .

Une conclusion importante découle de cette analyse : le biais sur les questions privilégiées, et la concurrence pratiquement sans limite entre

institutions de recherche reposent tous les deux sur les mécanismes d'évaluation de la recherche, et ces mécanismes, on l'a vu, s'insèrent eux-mêmes dans un plan d'affaires commercial. On ne peut donc envisager de solution qu'en apportant de très forts correctifs au système actuel de publication.

Dans ce qui précède, on a également vu le rôle stratégique des revues. Il semble donc raisonnable de penser que toute bonne solution à ces problèmes requiert de revoir le rôle des revues, sinon leur disparition. C'est dans ce contexte que le concept de plate-forme prend toute son importance. Des plates-formes établies sur la base de problèmes importants, mais sous-étudiés, accompagnées de plates-formes structurées pour répondre aux besoins de pays, ou de régions entières pourraient donc constituer une réponse adéquate à la question d'internationaliser la connaissance sans la soumettre aux distorsions d'une mondialisation, processus foncièrement centralisateur et, de surcroît, aux effets aggravés par les objectifs financiers des maisons d'édition.

5. Conclusion : un rôle pour la publication savante dans les pays de langue arabe

La communication et publication savantes dans un ensemble linguistique aussi important que celui des pays arabophones constitue à la fois un problème spécifique, et pour autant comparable à d'autres ensembles linguistiques importants. Le monde hispanophone, vu sous cette perspective, présente des caractéristiques analogues, même en incluant l'Espagne. L'Inde, et la Chine, immenses réservoirs humains, présentent, dans les communications et publications savantes, des caractéristiques que l'on retrouve en fait à peu près partout hors les pays riches.

Dans le cas de l'Inde, un commentaire récent résume assez bien un problème constant qui a déjà été posé plus haut :

Si les scientifiques reconnus dans le pays ne désirent pas publier ne serait-ce qu'une partie de leurs contributions de recherche dans des périodiques indiens, ne désirent pas arbitrer sérieusement des manuscrits pour ces périodiques, et, de façon plus fondamentale, pénalisent, directement ou indirectement, ceux qui publient dans ces périodiques, sans égard pour la qualité de leur travail, il s'ensuivra que ces revues continueront à se

débattre et ne réussiront pas à devenir concurrentielles à l'échelle internationale. C'est de cette façon que nous sommes les ennemis des périodiques de recherche reconnus publiés en Inde.

Ce commentaire est particulièrement intéressant parce qu'il pointe un problème réel tout en offrant une solution insatisfaisante – le volontarisme des chercheurs – et néglige le niveau structurel et systémique qui agit à titre de condition de possibilité du type de comportement que critique l'auteur. Bien sûr, les chercheurs établis en Inde ont tort de se laisser séduire par les sirènes du monde de l'édition des pays occidentaux – ici des revues « prestigieuses » dotées d'un fort facteur d'impact parce. Ce faisant, comme tout bon prisonnier pris dans le dilemme bien connu, la solution repose purement sur l'individu. Mais une réelle solution doit pouvoir pousser plus loin et refuser d'évaluer les contributions de recherche en s'appuyant sur un indicateur de signification incertaine comme l'est le facteur d'impact – s'agit-il de visibilité citationnelle, de visibilité tout court, de prestige de la revue, de qualité de la revue (mais que veut dire cette qualité ?) etc. En d'autres mots, le problème auxquels font face les revues de l'Inde, tout comme celles de l'Amérique latine, de la Chine, de partout en fait, même dans les pays riches, repose sur le fait que l'on a érigé ces revues en indices de qualité, que l'on a quantifié cette qualité pour la traduire en classements, et que les premiers rangs dans ces classements sont assimilés à de l'excellence. Tenter de concurrencer ce système en le renforçant par un jeu de concurrence élargie ne peut que déboucher sur des résultats se situant entre l'échec pur et simple, et les résultats hautement ambigus de certaines revues de SciELO en Amérique latine. En effet, reprenant l'obsession générale pour les facteurs d'impact, SciELO, cherche pour ses revues à atteindre le seuil de 1.000 – il ne faut jamais oublier que le facteur d'impact se décline généralement avec trois décimales, de façon, nous explique Garfield, à réduire le nombre de revues classées *ex æquo*. Cet objectif a conduit SciELO à passer des accords ambigus avec les propriétaires du Web of Science qui, à l'époque, était Thomson Reuters. Ce faisant, SciELO a aussi contribué à révéler aux grandes maisons d'édition quelles revues de SciELO pouvaient offrir des perspectives favorables sur le plan commercial et méritaient donc d'être débauchées de la plate-forme SciELO.

Les problèmes rencontrés par SciELO doivent servir de leçon pour les autres projets visant à établir des sites de publications savantes qui ne

veulent pas simplement succomber aux forces de la mondialisation. Si le but est de former un territoire de recherche autonome, il importe alors de pouvoir définir les problèmes que l'on entend explorer. Cela dit, un territoire de recherche ne renvoie pas à un ghetto intellectuel, bien au contraire. En fait, l'objectif est d'établir un territoire ouvert, lié à d'autres territoires, y compris les territoires actuellement dominants en recherche. Mais ceci veut aussi dire que ces nouveaux types de territoires de recherche doivent s'assurer le contrôle de deux paramètres importants :

- La capacité de pouvoir définir des programmes de recherche autonomes répondant à des demandes émanant du territoire lui-même;
- La capacité de pouvoir évaluer les contributions de recherche répondant aux programmes de recherche émanant du territoire doivent se faire selon des critères qui reposent sur les contributions de recherche elles-mêmes, et non sur le prestige/visibilité des revues où l'on publie.

C'est à ce point que les remarques exprimées plus haut sur les nouveaux rapports possibles entre plates-formes et revues savantes prennent toute leur force. En effet, si les revues correspondent désormais à la voix de communautés de recherche, les articles de ces revues peuvent être pris dans la Grande Conversation de la connaissance parce que, en principe, ces communautés ont établi des seuils de qualité à respecter sauf à mettre en danger la réputation même de ces communautés. À noter que ce concept de réputation ne correspond pas à un classement et n'est pas exprimé dans une forme – le nombre – nécessaire à l'établissement d'un classement. Par dessein, la réputation doit être floue et subjective, précisément pour ne jamais pouvoir prendre la place d'une évaluation menée sérieusement sur la base d'une analyse serrée conduite par des spécialistes de la question traitée dans un article. Si on n'agit pas de cette façon, on retombe dans la concurrence de titres fonctionnant à la manière de logotypes (ou logos) dans le monde commercial. Or, le site de publication ne garantit pas mieux la qualité de son contenu que la marque ne garantit la qualité de l'objet ainsi identifié. De la même façon, à l'intérieur d'une même revue, la distribution des citations se rapportant à chaque article indique qu'une très grande majorité d'entre elles provient d'une toute petite minorité d'articles, laissant peu de citations pour la majorité des articles publiés. Procéder

ensuite en calculant une moyenne sur l'ensemble des articles d'une revue conduit à des aberrations sur le plan mathématique et celui de l'interprétation à accorder au facteur d'impact.

Les thèses fondamentales de cette petite présentation sont donc les suivantes :

- La première consiste à assigner l'essentiel des aberrations de la publication savante contemporaine à la façon dont l'objet « revue savante » a été silencieusement retravaillé lorsqu'il s'est trouvé pris dans le collimateur du facteur d'impact et a vu son évaluation réduite à un classement dont la crédibilité ne repose en bout de ligne que sur l'efficacité du nombre comme instrument rhétorique, comme manière de convaincre par illusion d'objectivité ;
- La deuxième consiste à accepter que l'objet revue est trop bien ancré dans les habitudes de travail des chercheurs pour pouvoir être attaquée de front. Mieux vaut donc ruser avec l'objet en conservant le mot « revue savante » ou « périodique savant », mais en faisant glisser le sens de ce mot, et en le faisant glisser d'autant mieux que l'on peut revenir à un legs émanant du passé : la revue comme organe d'expression d'une communauté ;
- La troisième thèse consiste à dire que ce changement en douceur du sens du mot revue peut s'effectuer assez aisément si l'on insère la revue dans une plate-forme qui la surplombe, l'organise, lui donne de nouvelles capacités ;
- Dans le cas de grands ensembles linguistiques tel le monde arabophone, hispanophone, chinois, etc. la quatrième thèse est de défendre la création de plates-formes au gré des dynamiques de recherche à l'œuvre dans ces territoires, création conçue précisément pour renforcer l'effet territoire dans les domaines de la connaissance. L'idée n'est pas de constituer une seule plate-forme, ou encore de conformer toute plate-forme à un modèle centralisé. Les plates-formes peuvent être multiples (par exemple SciELO, Redalyc-CLACSO et La referencia en Amérique latine), diverses (La Referencia est une plate-forme de dépôts institutionnels plutôt que de revues savantes) et peuvent s'organiser de manière

distribuée à travers tout un continent et même au-delà (penser par exemple à AmeliCA, organisme déridé de Redalyc) ;

- Cinquième thèse, une des fonctions importantes de ces plateformes, véritables charpentes de nouveaux territoires de la connaissance, est de préparer échanges et collaborations avec d'autres territoires dotés de leurs propres plateformes, et néanmoins soucieux de travailler de concert, par exemple en soignant l'interopérabilité des dispositifs impliqués. L'universalité de la science ne se situe pas tant au moment de l'interrogation qu'au moment de la publication, mais tout élargissement de la base des questions possibles, dans toute leur diversité, ne peut que conduire à des tests élargis pour les concepts et théories impliqués dans l'exploration de ces questions, et peut même mener à de nouveaux concepts et de nouvelles théories. En d'autres mots, ouvrir la possibilité d'un questionnement plus large de la réalité ne peut que favoriser l'évolution des connaissances.
- Sixième et dernière thèse, les organismes subventionnaires détiennent les ressources pour impulser les changements profonds que ce programme présente. De plus, ils ne sont pas soumis aux classements et n'ont rien à craindre de l'abandon du facteur d'impact. Ils ont donc un rôle fondamental à jouer comme catalyseurs de ce mouvement vers l'internationalisation de la recherche.

Avec ces cinq thèses, on peut commencer à envisager un plan de développement de la recherche scientifique dans toute région, en particulier dans le monde arabophone, qui contribuera à créer le « World Brain » (cerveau mondial) que faisait miroiter l'auteur de science-fiction H. G. Wells dans les années 30, et qui contribuera aussi à offrir un projet d'internationalisation de la recherche qui peut s'opposer efficacement aux mythes colportés par le programme en cours de mondialisation.

De l'économie du savoir à l'accès libre au savoir : Quelles limites ?

*From the knowledge economy to Open Access to knowledge:
What are the limits?*

Mehdi Lahlou¹

Résumé

Le fondement de l'économie - dans les pays les plus avancés - est constitué par l'accumulation en savoirs de toutes sortes permettant à celle-ci de répondre aux multiples besoins, tant matériels qu'immatériels, de la société. De la sorte, cette économie historiquement basée sur le triptyque "Travail, capital physique, ressources naturelles" a aujourd'hui pour pilier majeur, la production du Savoir. Et donc aussi, bien évidemment, la capacité la capacité de le transmettre en accès libre et de le transformer en biens utiles aux humains et à l'ensemble des vivants.

Mots clés. Economie du savoir, Gestion du savoir, Accès Libre.

Abstract

The foundation of the modern economy - in the most advanced countries - is constituted by the accumulation of knowledge of all kinds, allowing it to address the needs, both material and immaterial, of society. In this way, such an economy historically based on the triptych "Labor, physical capital, natural resources" has, today, for major pillar, the production of Knowledge. And therefore, the capacity to transmit it into Open Access and transform it into useful goods to human beings and all living things.

Keywords. Knowledge Economy, Knowledge Management, Open Access.

¹ Mehdi Lahlou est professeur de sciences économiques, Rabat, Maroc.

L'économie du savoir, une approche historique

L'évolution de l'économie mondiale, qui a conduit à la situation que l'humanité connaît aujourd'hui sur tous les plans, notamment politique, économique et technologique, a commencé par la découverte des routes maritimes de l'Atlantique sud et par celle des Amériques par des navigateurs partis de la Péninsule ibérique entre la fin du XV^e et le début du XVI^e siècle. Elle s'est poursuivie en Grande Bretagne, vers le milieu du XVIII^e, puis dans le reste de l'Europe du Nord-Ouest par un ensemble d'inventions techniques et de nouvelles formes d'organisation sociale qui ont donné lieu à ce que les économistes et les historiens ont appelé la Révolution Industrielle.

Cette dernière s'est faite, sur le plan strictement technique, dans un double sens. Dans celui de la multiplication et de la diversification des activités humaines de production et, simultanément, dans celui du passage de ces activités d'un travail physique exercé principalement sur la terre (la Nature au sens large) et les ressources qui en sont extraites directement, soit essentiellement l'agriculture, à des activités industrielles de plus en plus complexes de transformation et d'élaboration de matières premières naturelles, agricoles, forestières et minières, pour en faire des biens de tous usages (soit de consommation, d'amélioration du cadre de vie ou de production).

Parallèlement à ce processus de croissance (au sens d'accroissement des biens et services disponibles) et de complexification ascendante, l'économie, essentiellement européenne au départ, sera fondée tout le courant des XVIII^e et XIX^e siècles sur l'apparition et le renforcement de l'usage de la motorisation et du machinisme (avec, notamment, l'invention du métier à tisser et du premier moteur à explosion). Ceux-ci vont aider les hommes, c'est-à-dire les travailleurs qui usaient jusque là de leur seule force physique, à produire - avec des moyens historiquement jamais connus - davantage de biens agricoles et industriels, en s'appuyant sur de nouvelles ressources énergétiques, tels que le charbon puis le pétrole, et bien plus tard, le nucléaire.

Toutefois, les processus de production et les modes organisationnels que les Européens de l'Ouest, et à leur suite les Nord-Américains puis les Japonais, vont adopter à partir de la même période historique, vont conduire à de nombreuses autres évolutions qui seront le prélude à la Révolution numérique de la fin du XX^e siècle et du début du XXI^e dont, notamment :

- Une Im-matérialisation de plus en plus marquée de l'économie, éloignant le produit final, quelle que soit sa nature, sa composition ou ses usages, des matières premières physiques (soit la terre, représentée par les ressources naturelles du sol et du sous-sol qu'elle offre). Cette "immatérialisation" est notamment exprimée par la part de plus en plus grande prise dans les systèmes de production évolués des services au détriment des activités agricoles (qui impliquent désormais moins de 2 à 3% de la main-d'œuvre productive) ou des activités matérielles de transformation, ou industries, qui ne concernent dorénavant dans la

plupart des pays industrialisés, qualifiés également de pays développés, que moins du quart de la production globale ;

- Une telle évolution - rendant également les activités économiques de production de plus en plus élaborées et de plus en plus complexes et réalisées dans un cadre convenu (l'Usine ou la manufacture) regroupant un nombre croissant de personnes s'activant autour de machines (capital physique) plus compliquées à manipuler et plus performantes année après année - va conduire au recours à des ressources humaines plus qualifiées et mieux formées à la production. Et donc exigeant pour cela de bénéficier d'une structure sociale nouvelle, assurant l'apprentissage des nouveaux modes d'organisation et de transformation des matières premières en produits finis. Ce sera l'Ecole, et donc l'instruction, dont la généralisation et le caractère obligatoire vont devenir communs à tous les pays industrialisés à partir de la fin du XIXe siècle (et la loi de Jules Ferry rendant l'école publique obligatoire en France, entre autres, en 1882). Et dont les faiblesses ou parfois même la faillite sont, aujourd'hui, pour beaucoup dans la faiblesse du capital humain et donc dans les difficultés de développement économique et social des pays dits du "Tiers-monde", soit ceux qui peinent à satisfaire les besoins, y compris de base, de leurs populations ;
- Avec l'école et la multiplication des inventions techniques, les économies européennes ou d'origine européenne (aux Etats-Unis d'Amérique, notamment, ou en Australie et en Nouvelle-Zélande) vont voir se développer, avec le XXe siècle, un processus de transformations organisationnelles et technologiques, marqué plus particulièrement par la robotisation, dont le support sera l'accumulation d'un ensemble de savoirs de plus en plus évolués, fondés sur un système de recherche scientifique – accolé à l'école – qui fera transformer au milieu du même XXe siècle la Révolution industrielle en Révolution scientifique et technique ;
- Et, avec l'avènement de l'informatique et de ce qui sera qualifié à partir de la fin de ce même siècle, de Nouvelles technologies de l'information et de la communication, les économies mondiales – supposées désormais intégrées dans ce qui va être qualifié d'Economie mondialisée ou Economie globalisée – connaîtront une nouvelle phase de transformation des bases de la production (des modes de production, diront les économistes) dite Révolution numérique qui donnera naissance à l'économie du Savoir. Soit un système de production, plus immatériel que physique, fondé sur les connaissances et les compétences humaines, individuelles et collectives, accumulées grâce à des institutions éducatives et de recherche scientifique et technique, mises en place et financées d'abord grâce à des ressources publiques. Les Etats considérant cela comme relevant de leurs missions naturelles, comme rendre la justice ou assurer la défense et la sécurité de leurs citoyens.

Tout cela ne veut pas dire, toutefois, que les sociétés humaines d'avant la Révolution scientifique et technique (ce qui pourrait être aussi,

aujourd'hui, la situation de certains groupes de population en Afrique, en Amérique Latine ou en Asie), ou encore d'avant la Révolution industrielle, ne disposaient pas de savoirs ou procédaient sans intelligence de leurs milieux et des moyens dont elles disposaient. Mais le fait est que ces savoirs étaient nécessairement limités, conduisant par la même les Hommes à déployer de grands efforts physiques pour parvenir à satisfaire leurs besoins de base (vivre, se protéger, se soigner, etc.) à partir d'importantes quantités de ressources naturelles. Alors que, désormais, il est possible de produire mieux et bien plus de biens et de services, grâce aux savoirs accumulés au cours des siècles, pas seulement pour nourrir une population mondiale près de 15 fois supérieure en nombre au début du XXe à ce qu'elle était 5 siècles plus tôt, mais aussi pour améliorer son cadre de vie, lui faire visiter la planète, la distraire... Et cela, en procédant par une moindre utilisation de ressources naturelles et en mettant en œuvre un minimum d'efforts physiques humains². Ce qui correspond, théoriquement et objectivement, au cœur de la cible des Sciences économiques, soit produire au moindre coût - en termes de temps, de finances et d'efforts humains - le plus d'utilités possibles pour l'ensemble des composantes de la société.

Le fondement de l'économie - dans les pays les plus avancés scientifiquement, matériellement et humainement, en termes de démocratie et de libertés individuelles et collectives - est ainsi devenu non pas les ressources naturelles et physiques à la disposition de la société mais l'accumulation en savoirs de toutes sortes permettant qu'elle réponde à ses multiples besoins, tant matériels qu'immatériels. De la sorte, cette économie historiquement basée sur le triptyque "Travail, capital physique, ressources naturelles" a aujourd'hui pour pilier majeur, le Savoir. Et donc aussi, bien évidemment, la capacité de produire ce savoir, de le faire circuler et de le transformer en biens utiles aux humains et à l'ensemble des vivants.

L'Economie du Savoir, quelles définitions ?

L'Organisation de coopération et de développement économique (OCDE), définit l'ère de l'économie du savoir comme "celle où la productivité et la croissance de l'économie - soit la « performance économique » des pays - dépendent de plus en plus du savoir, de l'éducation, et de la technologie.

² "Une moindre utilisation de ressources naturelles et la mise en œuvre d'un minimum d'efforts physiques humains" ne signifient pas, pour autant, une pression moins forte sur la nature. D'où la problématique lancinante de l'environnement. Cette dernière devenue de plus en plus prégnante avec, notamment, l'urbanisation et le développement de moyens de transport gros consommateurs d'énergies fossiles, est en réalité liée à la forte augmentation des besoins humains de tous ordres et également à l'augmentation de la population mondiale, dont le nombre a été multiplié par 7 entre le début du 20^{ième} siècle et le début du 21^{ième}. Il est toutefois évident d'énoncer que si la recherche scientifique a permis de résoudre les problèmes de rareté et de pénibilité, elle devrait aussi permettre de trouver les solutions aux questions environnementale et climatique.

En vertu de la « nouvelle théorie de la croissance », ou « théorie de la croissance endogène », la production de valeur économique découlera ainsi d'investissements soutenus³ dans : a/ la recherche-développement ou l'innovation techno-scientifique ; b) la formation de la main-d'œuvre- soit des ressources humaines- et l'enseignement ; c) de nouvelles formes plus efficaces d'organisation et de gestion du travail"⁴.

Dans ce sens, et pour saisir la part du savoir dans le développement des économies les plus évoluées aujourd'hui, l'OCDE⁵ a initié en 2011, pendant deux années, un projet portant sur le thème "les nouvelles sources de la croissance : le Capital-basé sur le savoir" – CBS (en anglais, the KBC – Knowledge-Based Capital).

De ce projet, il est notamment ressorti que "les études pertinentes portant sur la croissance, relatives à différentes périodes, montrent une relation positive entre les investissements en CBS, la croissance macro-économique et une plus grande productivité du travail. A titre d'exemple, il a été estimé qu'entre 1995 et 2007 au moins 33,7% de l'augmentation de la productivité du travail aux Etats-Unis d'Amérique était due à l'investissement en CBS. Egalement, au cours de la même période, parmi 14 pays de l'Union européenne, il a été calculé que les investissements en CBS ont été à l'origine d'au moins 19,9% de la croissance de la productivité du travail"⁶.

L'idée générale est que ce sont, désormais, les dépenses qui concernent la connaissance, la science et les nouvelles technologies, spécialement les technologies de l'information, qui stimulent la croissance. De la sorte, les théories économiques (portées par l'école des économistes classiques anglais, notamment)⁷ qui soutenaient à la fin du 18^{ème} et le début du 19^{ème} siècles que toutes les nations pouvaient se développer dans le cadre d'une répartition internationale du travail où chacune se spécialise dans la

³ C'est-à-dire de richesses matérielles et immatérielles.

⁴ Eric, Martin. "Qu'est-ce que l'Economie du Savoir ?". Novembre 2012. [En ligne], consulté le 18/11/2018. URL:<https://iris-recherche.gc.ca/blogue/qu%25e2%2580%2599est-ce-que-1%25e2%2580%2599economie-du-savoir>.

⁵ Créée le 16 décembre 1960, cette organisation de coopération économique, dont le siège se trouve à Paris, regroupe aujourd'hui les 36 pays les plus riches au monde, à l'exception de la Chine. Un seul pays musulman en fait partie, il s'agit de la Turquie.

⁶ OECD (2013), "Supporting Investment in Knowledge Capital, Growth and Innovation", OECD publishing. [En ligne], consulté le 18/11/2018 : <https://dx.doi.org/10.1787/9789264193307>.

⁷ L'avantage comparatif est une théorie développée par l'économiste anglais [David Ricardo](#) (18 avril 1772 - 11 septembre 1823 dont la principale œuvre : "Des principes de l'économie politique et de l'impôt", parue en 1817) en vue de promouvoir le commerce international. Pour ce faire, chaque pays doit, selon Ricardo, se spécialiser dans la production de biens – agricoles ou industriels – pour lesquels il dispose des meilleurs atouts, ou, autrement dit, qu'il peut réaliser au plus bas coût.

production de biens pour lesquels elle dispose des meilleurs avantages comparatifs ne tiennent plus, si jamais elles étaient valables par le passé.

Aujourd'hui, tous les rapports produits par toutes les instances pertinentes à l'échelle internationale (de l'UNESCO, à l'OCDE, à la CNUCED, à la Banque mondiale ou au PNUD, etc.) et où il est question de l'état de l'économie mondiale indiquent que le développement économique et social d'un pays ne dépend ni de sa superficie, ni de son niveau d'ensoleillement et de pluviométrie ou de son potentiel agricole ou minier ou encore du nombre de ses habitants mais de la qualité de ses ressources humaines exprimée en termes de niveau éducatif, de qualifications acquises et de performances scientifiques et créatives.

Ainsi, un pays comme la Corée du Sud, beaucoup moins grand et moins peuplé que l'Inde, par exemple, produit bien plus de richesse par habitant que n'en crée un Indien. De même, les Finlandais, habitant un pays bien plus petit et beaucoup moins pourvu en ressources naturelles que l'Égypte ou le Maroc, génèrent individuellement bien plus de valeurs économiques que les Égyptiens ou les Marocains. Les avantages comparatifs sont désormais dépendants – à quelques exceptions près, actuellement, comme celle de l'émirat gazier du Qatar ou du sultanat pétrolier de Brunei - des systèmes nationaux d'apprentissage et d'innovation mis en place par les différents pays à l'échelle mondiale.

“Comme les besoins des industries et des marchés changent constamment, il devient urgent de former des travailleurs polyvalents et flexibles, dotés de compétences transversales qui les rendent adaptables à toutes les situations nouvelles, et de les renvoyer sur les bancs d'écoles pour assurer le renouvellement continu de ces compétences. Il devient important de diffuser les nouvelles technologies et d'organiser les entreprises plus efficacement en vue de maximiser les gains de productivité. Se met alors en place une véritable « économie de l'apprentissage » où le système d'éducation est vu comme un réservoir à idées permettant de développer de la main-d'œuvre de pointe (capital humain) pour les industries à « haute valeur ajoutée » et des innovations brevetables (propriété intellectuelle) qui permettront de soutenir les entreprises et de dynamiser la croissance”.⁸

Les conditions de l'éclosion de l'économie du savoir

Mais, l'économie du savoir, ou la société du savoir dans laquelle elle se développe, suppose la réunion d'un ensemble de facteurs et d'agents qui opèrent en totale cohérence et interdépendance et qui agissent et interagissent, sur le court et le long terme, de façon coordonnée, pour l'essentiel intelligible pour tous, dans un mouvement cumulatif permettant au savoir d'aujourd'hui de s'adjoindre aux savoirs passés dans une dynamique globale d'autonomisation et d'inclusion. Dans cette perspective, l'économie du savoir est déterminée par l'action conjointe de plusieurs agents et par la réunion de nombre de conditions. Dans ce sens, elle dépend :

⁸ [Eric. Martin. "Qu'est-ce que l'économie du savoir ?". \[En ligne\], consulté le 18/11/2018: https://iris-recherche.qc.ca/blogue/qu%25c2%2580%2599economie-du-savoir.](https://iris-recherche.qc.ca/blogue/qu%25c2%2580%2599economie-du-savoir)

- de la présence d'agents sociaux (ensemble d'individus composant la société) et d'agents économiques bien formés, bien éduqués, conscients de leurs capacités et de leur rôle ainsi que de leur poids au sein de la société ;
- du fait que ces agents disposent de libertés individuelles et collectives qui leur permettent d'agir en toute liberté et sécurité dans tous les secteurs de la vie en société, et surtout lorsqu'il s'agit de l'acquisition de connaissances et de leur production, grâce à l'ensemble des moyens de recherche scientifique disponibles, ou de leur diffusion ;
- d'une large et libre diffusion (Open Access)⁹ de l'information, des connaissances et des résultats de la recherche scientifique et technique auprès de toutes les catégories sociales, des classes d'âge, des disciplines scientifiques et des secteurs économiques; chacun en fonction de ses besoins, de ses usages;
- d'une forte contribution de l'autorité politique, représentée par l'Etat. Celui-ci, dont l'organe d'exécution est le gouvernement – qui est, in fine, une émanation de la société - met en place les conditions nécessaires pour garantir et s'assurer que les agents économiques et sociaux disposent de moyens adéquats pour leur formation, ainsi que de liberté et de sécurité, sans lesquelles les Hommes resteraient brimés, incapables d'exprimer leurs capacités créatives. Et, au mieux, s'en iraient ailleurs, là où ils se sentiraient libres, non persécutés pour leurs opinions, leurs préférences, leurs travaux ou leurs recherches. Dans ce cadre, si le capital privé peut être d'une certaine utilité, celle-ci ne serait que marginale. Parce que l'approche privée est à la fois limitée dans l'espace et dans le temps, et aussi dans son ampleur, sa portée ainsi que dans sa finalité, qui réside d'abord dans la recherche du profit individuel.

Aussi, dans l'économie du savoir, le rôle de l'autorité étatique, soit des gouvernements, est central. Déjà, et dès la fin du XIX^e siècle, la mise en œuvre de la scolarité obligatoire en Europe a souvent été le résultat d'une décision politique majeure, considérée en son temps comme une avancée fondamentale sur la voie du progrès social¹⁰.

Aujourd'hui encore le rôle de l'Etat demeure fondamental partout où le capital privé n'ose pas s'investir. Soit parce que les délais de maturation de certaines dépenses de recherche sont trop longs, soit parce que leurs résultats et donc leurs retombées financières ne sont pas garantis, ou parce que le secteur de recherche intéresse davantage la communauté sociale dans son ensemble plutôt que l'entreprise privée. C'est le cas notamment des

⁹ L'open Access (ou aussi « libre accès », ou encore « accès ouvert ») à la littérature scientifique est un mode de diffusion des articles de recherche sous forme numérique, gratuite et dans le respect du droit d'auteur. [En ligne], consulté le 18/11/2018: <https://openaccess.couperin.org/quest-ce-que-lopen-access/>

¹⁰ Voir, ÉDUCATION, Les systèmes scolaires des pays développés. [En ligne], consulté le 18/11/2018. URL: <https://www.universalis.fr/encyclopedie/education-les-systemes-scolaires-des-pays-developpes/l-la-scolarite-obligatoire/>.

recherches liées au réchauffement climatique ou à l'environnement ou à la gestion des ressources halieutiques ou à l'aménagement du territoire, etc.

La contribution de l'Etat est aussi importante pour mettre en place, organiser et consolider les systèmes nationaux d'innovation et de production. A ce niveau, il s'agit pour les autorités publiques de mobiliser l'ensemble des partenaires sociaux qui sont autant de vecteurs de recherche scientifique et de circulation des produits de cette recherche. Il s'agit aussi bien des personnes physiques (les individus) mises en mesure de contribuer d'une façon ou d'une autre à la création de savoirs, que des universités et autres institutions de recherche et de diffusion de l'information. La référence est ici faite particulièrement aux bibliothèques et systèmes d'information permettant à ces savoirs de prendre une forme tangible, grâce à leur structure organisationnelle et à leur système socio-technique composé de technologies (hardware, software et équipements de télécommunication) et de processus techniques, afin d'être par la suite diffusés en "Open Access" auprès du système productif. Autrement dit auprès des entreprises qui représentent non seulement la plateforme qui traduit les résultats de la recherche scientifique en productions finales, matérielles et immatérielles, mais qui sont aussi, potentiellement, des initiatrices de projets de recherche assurant autant leur rentabilité que leur extension et durabilité.

Le rôle des gouvernements est ici primordial pour faire en sorte que tous ces agents contribuent activement aux évolutions scientifiques et techniques de l'ensemble de la société et profitent, in fine, des progrès de tous ordres, permis grâce à elles. Le réseau de relations ainsi tissé, par l'entremise du rôle fédérateur, régulateur et d'impulsion de l'Etat – y compris dans les pays les plus libéraux sur le plan économique – constitue le garant de la prospérité et du bien-être individuels et collectifs. Ceci est d'autant plus vrai que seul l'Etat dispose de la capacité et des moyens pour avoir une vue d'ensemble sur la société, qu'il est aussi le seul à avoir l'obligation d'agir pour l'intérêt général comme il est le seul "maître des horloges" à bord pour mettre en œuvre des programmes non contraints par le temps court, soit celui des entreprises privées.

Dans le monde moderne, c'est ce type de réseau, structurant le cadre de vie social et de production, qui permet à une société et une économie non seulement de prospérer en interne, mais aussi d'être compétitives à l'international. C'est-à-dire, d'être capables d'étendre et d'accroître leur part de marché dans une économie mondialisée où seul l'avantage comparatif dépendant de la qualité des ressources humaines prime sur tous les avantages classiques (ressources naturelles, conditions climatiques, distance, etc.).

De ce qui vient d'être énoncé il découle, plus particulièrement, l'idée que la liberté, dans ses expressions civiles et politiques constitue un des fondements de la société du savoir. Dans ce sens, l'UNESCO soutient que "l'essor de sociétés du savoir exige que soient tissés de nouveaux liens entre savoir et développement, le savoir étant à la fois un instrument de la satisfaction des besoins économiques et une composante à part entière du développement. Dès lors, la dynamique politique, économique et sociale qui sous-tend l'essor de sociétés du savoir met en lumière la relation

indissoluble qui unit la lutte contre la pauvreté et la promotion des libertés civiles et politiques”¹¹.

Comme il en découle que l'école et l'éducation représentent la pierre angulaire qui détermine les sociétés et économies modernes dans leur quête de plus de savoirs, de connaissance, de bien-être et de progrès social. À ces deux piliers s'ajoutent la nature de l'Etat et l'ordonnancement de ses priorités. Que ces dernières soient d'ordre sécuritaire ou militaire, par exemple, plutôt que d'ordre éducatif et socio-économique, et que les bases mêmes du développement économique et social s'en trouvent affectées.

L'encadré ci-après est édifiant à plus d'un titre. Il indique que les évolutions atteintes aujourd'hui par les pays européens, ou de filiation européenne, ne sont pas dues au hasard ou à l'improvisation, mais bien à une volonté délibérée de l'Etat, exprimée et traduite en lois en pleine expansion de la révolution industrielle et technique.

Encadré sur l'Etat et l'instruction obligatoire en Europe de la fin du XIXe siècle¹²

L'intervention des pouvoirs publics est indispensable pour procurer à tout un peuple les moyens de s'instruire, tel est le point que nous avons essayé d'établir, et qui n'admet pas de sérieuses contradictions. Partout où l'Etat s'est abstenu, l'enseignement primaire a été presque nul et l'ignorance extrême : l'on ne citera pas un seul pays dans lequel les individus, même groupés en puissantes associations, les églises établies ou les corporations, soient parvenus à ouvrir un nombre suffisant d'écoles. — Mais est-ce assez que les communes et l'Etat fondent les établissements nécessaires ? Ne faut-il pas que la loi oblige les parents à y envoyer leurs enfants ? Déjà traitée en France par plusieurs écrivains compétents, notamment par M. Victor Cousin dans son livre souvent cité sur l'enseignement en Prusse, la question de l'instruction obligatoire vient d'être imposée récemment à l'attention de tous par les conclusions hardies d'un ministre qui, sortant des bornes d'une approbation purement théorique, n'a pas craint de réclamer l'adoption immédiate d'une mesure qui ne permettrait plus à personne de laisser ses enfants dans une ignorance absolue. La proposition de M. Duruy a paru soulever de si vives oppositions et des appréhensions si sincères, qu'il est nécessaire d'examiner avec une attention scrupuleuse le fondement de ces résistances et de ces alarmes. Avant d'imposer aux citoyens une obligation nouvelle, il faut démontrer trois choses : d'abord que cette mesure est juste, ensuite qu'elle est utile, enfin qu'elle est applicable, c'est-à-dire que dans l'application les inconvénients ne dépassent pas les avantages. C'est sous ce triple rapport que nous considérerons l'instruction obligatoire.

Ceux-là seuls qui nient la distinction entre le bien et le mal peuvent soutenir que la liberté de l'homme est illimitée. Dès qu'on reconnaît que certaines actions sont mauvaises, il faut admettre aussi que nul n'a le

¹¹ “*Vers les sociétés du savoir*”. Rapport mondial de l'UNESCO pour l'année 2005.

¹² L'Instruction du Peuple au XIXe siècle. — III. — *L'Instruction obligatoire et les moyens d'application*. Revue des Deux Mondes, 2e période, volume 62, 1866 (p. 977-1012).

droit de les commettre. Le droit de faire ce qui est contre le droit ne se peut comprendre. Quand une action ne nuit qu'à son auteur, ou lorsqu'elle ne cause aux autres qu'un tort tel qu'il serait plus nuisible de la punir que de la tolérer, la règle à suivre est la tolérance. Au contraire, quand une action porte préjudice à autrui, que le délit est facile à constater et que la punition est utile, la société a le droit et même le devoir d'intervenir. Celui qui a commis un acte injuste et nuisible tombe sous le coup de la législation répressive. Or tel est le cas du père de famille qui ne donne pas à l'esprit de ses enfants cette culture élémentaire sans laquelle ils ne peuvent devenir des êtres intelligents et moraux. Le père, en agissant ainsi, manque à l'accomplissement d'un devoir naturel. En leur refusant la nourriture spirituelle qui leur est indispensable, il nuit à ses enfants tout autant que s'il ne leur donnait pas les aliments que réclame l'entretien de leurs forces. Il nuit aussi à la société en introduisant dans son sein des hommes ignorants, prédisposés à l'erreur, à l'immoralité, au crime même, et qui par conséquent seront pour elle une cause de désordre, de périls et de dépenses. Il y a donc dans le fait de ce père tous les éléments qui constituent un délit que la loi peut empêcher ou punir.

Et c'est bien cette volonté qui donnera, par la suite, aux économies européennes et nord-américaine les ressources financières et en moyens scientifiques et techniques, non seulement pour durer et se renforcer, mais aussi pour dominer le monde, comme elles l'ont fait au cours du XXe siècle à travers la colonisation, ou le font encore de nos jours, à travers un marché mondial dont elles dominent l'ensemble des compartiments, ou, à tout le moins, les vecteurs à forte valeur ajoutée.

Or, quand l'observateur le plus neutre en vient à la situation des sociétés et économies arabes en ce début du XXIe, il est bien obligé de noter que bien que beaucoup parmi ces économies disposent en abondance de la richesse naturelle la plus stratégique du moment de par le monde, soit le pétrole et le gaz, beaucoup parmi elles – notamment celles les moins pourvues en ressources minières – ne permettent pas encore la satisfaction des besoins de base de leurs populations et pratiquement toutes dépendent des pays développés pour les biens industriels de consommation (du téléphone, à l'ordinateur ou à l'avion) ou d'investissement ou encore pour la satisfaction de leurs besoins en produits alimentaires. Ainsi, d'Arabie Saoudite au Maroc, l'essentiel des biens alimentaires vient de France, de Nouvelle Zélande ou d'Amérique du Nord, et la quasi-totalité des équipements utilisés pour l'extraction de pétrole comme de phosphates est d'origine anglaise, allemande ou américaine.

Et si tel est le cas jusqu'à aujourd'hui, c'est bien parce que les conditions qui étaient à l'origine des révolutions industrielle, puis scientifique et technique et enfin numérique, de l'Europe à l'Amérique du Nord et au Japon, ne sont toujours pas réunies dans les sociétés arabes, loin s'en faut.

Parmi ces conditions, le respect et la garantie de la liberté des personnes composant la société, l'élévation de leur niveau d'instruction et de qualification et un rôle prééminent de l'Etat dans le développement économique et social des communautés dont il a la charge.

Les conditions de la genèse et de l'extension de l'économie du savoir dans le monde arabe

La société du savoir et la liberté dans le monde arabe

L'une des premières conditions pour la constitution d'une économie du savoir réside dans le respect des libertés individuelles et politiques des personnes y contribuer et en profiter. Dans son rapport annuel, publié en 2005, l'UNESCO soutenait, notamment, que "L'approche centrée sur le « développement humain » et l'« autonomisation », qui est au cœur de l'idée de sociétés du savoir, devrait permettre de mieux mettre en œuvre les droits universels et les libertés fondamentales, tout en améliorant l'efficacité de la lutte contre la pauvreté et des politiques de développement. Car l'essor de sociétés du savoir exige que soient tissés de nouveaux liens entre savoir et développement, le savoir étant à la fois un instrument de la satisfaction des besoins économiques et une composante à part entière du développement. Dès lors, la dynamique politique, économique et sociale qui sous-tend l'essor de sociétés du savoir met en lumière la relation indissoluble qui unit la lutte contre la pauvreté et la promotion des libertés civiles et politiques"¹³.

Une telle affirmation, qui donne son assise à l'idée selon laquelle les sociétés du savoir de l'ère de l'information se distinguent des sociétés anciennes de la connaissance par leur caractère intégrateur et participatif, est notamment fondée sur l'article 19 de la Déclaration universelle des droits de l'homme qui stipule le respect de "la liberté d'opinion et d'expression ainsi que le respect de la liberté d'information, du pluralisme des médias et de la liberté académique". Elle s'appuie aussi sur l'article 26 qui énonce, plus particulièrement, que "Toute personne a droit à l'éducation", que cette "Éducation doit être gratuite, au moins en ce qui concerne l'enseignement élémentaire et fondamental" et qu'elle "doit viser au plein épanouissement de la personnalité humaine et au renforcement du respect des droits de l'homme et des libertés fondamentales". Elle est aussi accolée, comme c'est indiqué à l'article 27, alinéa 1 de la même déclaration, au respect du droit de "prendre part librement à la vie culturelle de la communauté, de jouir des arts et de participer au progrès scientifique et aux bienfaits qui en résultent"¹⁴.

Or, le respect des libertés individuelles et collectives ne semble pas encore être à l'ordre du jour dans la quasi-totalité des pays arabes. Ainsi, comme cela est indiqué dans le tableau ci-après qui présente le classement mondial de ces pays s'agissant de la liberté de la presse, il apparaît que seule la Tunisie émerge parmi les cent premières nations dans ce classement (qui synthétise la situation de la presse dans 180 pays) alors que neuf pays sont

¹³ "Vers les sociétés du savoir". Rapport Unesco – 2005. Editions UNESCO.

¹⁴ La Déclaration universelle des droits de l'Homme, Nations Unies. [En ligne], consulté le 18/11/2018 : <https://www.un.org/fr/universal-declaration-human-rights/>.

classés au-delà du 100^{ième} rang. L'Arabie Saoudite apparaissant parmi le groupe des dix dernières nations.

Tableau 1 : liberté de la presse dans le monde arabe

Pays	Classement
Tunisie	72
Liban	101
Qatar	128
Jordanie	130
Emirats Arabes Unis	133
Maroc	135
Algérie	141
Irak	156
Egypte	163
Arabie Saoudite	172

Source : Reporters sans frontières, Classement mondial de la liberté de la presse 2019. <https://rsf.org/fr/classement#>

Et, de fait, la liberté de la presse constitue le centre de gravité de l'ensemble des autres libertés. Son absence marque l'interdiction, ou à tout le moins, la forte limitation effective ou potentielle de toutes les autres libertés, des libertés politique ou religieuse à la liberté de penser, de s'exprimer ou de procéder à de la recherche scientifique ou à la diffusion de ses résultats. Par ailleurs, le printemps arabe ne semble avoir rien arrangé à ce sujet, exception faite de la Tunisie, du Maroc à la Libye, à l'Egypte ou à la Syrie. Comme l'administration américaine, en place depuis 2017, paraît avoir laissé mains-libres au sujet des droits de l'homme aux gouvernants de certaines monarchies alliées du Golfe – par ailleurs gros acheteurs de matériels militaires. A ce propos, l'assassinat d'un journaliste saoudien¹⁵

¹⁵ Il s'agit du journaliste opposant Jamal Khashoggi, tué le 2 octobre 2018 par un groupe de 15 membres des services de sécurité saoudiens, venus directement de Riadh. Le Monde fr. 25/10/2018. J. Khashoggi soutenait, notamment, que "Ce dont le monde arabe a le plus besoin est la liberté de la presse". [En ligne] consulté le 18/11/2018 : https://www.washingtonpost.com/opinions/global-opinions/jamal-khashoggi-what-the-arab-world-needs-most-is-free-expression/2018/10/17adfc8c44-d21d-11e8-8c22-fa2ef74bd6d6story.html?Noredirect=on&utm_term=.06cb026a2785.

dans le consulat de son propre pays à Istanbul, en Turquie, marque un sommet du genre dans la négation des libertés universellement reconnues, dont la liberté politique ou de presse.

La société du savoir et l'éducation

Alors que l'école avait constitué, comme indiqué par ailleurs, une des principales préoccupations des gouvernements européens, de la fin du 19^{ème} siècle – soit voilà près de 140 ans – et alors que dans l'ensemble des pays développés l'analphabétisme a presque totalement disparu et que l'école, rendue partout obligatoire, intègre la quasi-totalité des enfants scolarisables. Et alors que, dans ces mêmes pays, tous les enfants sont poussés, par tous les moyens autant financiers que pédagogiques, à aller jusqu'aux niveaux supérieurs d'enseignement général ou de formation professionnelle, avec ce que cela implique comme allongement de la durée moyenne de scolarité, l'éducation est encore considérée comme un privilège ou, au mieux, comme une dépense sociale qu'il faut comprimer dans beaucoup de pays arabes. Ceci est traduit par le tableau suivant qui indique, par exemple, que la durée moyenne de scolarité – en 2018 – est égale à moins des deux tiers de la même durée dans, notamment, les trois pays les mieux classés par le Programme des Nations Unies pour le Développement.

Tableau 2 : Indicateur de développement humain (IDH) et durée de scolarité dans le monde arabe :

Pays	Classement IDH	Revenu / habitant En milliers de \$ US	Durée de scolarité moyenne En années	Durée de scolarité prévue En années
Norvège	1	68,01	12,6	17,9
Suisse	2	57,62	13,4	16,2
Australie	3	43,56	12,9	22,9
EAU	34	67,8	10,8	13,6
Qatar	37	116,82	9,8	13,4
Arabie Saoudite	39	49,68	9,5	16,9
Liban	80	13,38	8,7	12,5
Algérie	87	13,8	8,0	14,4
Jordanie	95	8,29	10,4	13,1
Tunisie	95	10,27	7,2	15,1
Libye	108	11,1	7,3	13,4
Egypte	115	10,35	7,2	13,1
Irak	120	17,8	6,8	11
Maroc	123	7,34	5,5	12,4

Source : Rapport sur le Développement Humain, 2018.

Revenu en parité de pouvoir d'achat

Le Maroc, un des pays les moins bien classés à ce niveau y compris au sein du monde arabe, enregistre une durée de scolarité inférieure de plus de la moitié à ce qui est le cas dans des pays comme la Norvège, la Suisse ou l'Australie.

Dans le même sens, ce même Maroc – qui peut acheter coup sur coup, entre 2017 et 2018, deux satellites¹⁶ ou qui fait rouler depuis début 2019 le train le plus rapide d'Afrique¹⁷ – enregistre encore un taux d'analphabétisme de près de 30 % parmi sa population active. A ce propos, il fait encore mieux que l'Égypte, pays de 95 millions d'habitants, avec 35 % de taux d'analphabétisme ou la Mauritanie, avec un taux d'analphabétisme de 42 %.

Et pour être plus expressif, en retenant des statistiques marocaines disponibles au sujet de la population active, soit celle qui est directement impliquée dans le système productif, on constate tout particulièrement que près de 6 actifs occupés sur 10 au Maroc (soit 57,8% de la population active) n'ont aucun diplôme, que 28,6% ont un diplôme de niveau moyen et seulement 13,6% font part d'un niveau de formation supérieur.¹⁸

Une telle configuration éducative, au Maroc, comme dans le reste du monde arabe, est incompatible avec l'économie de l'apprentissage, comme le soutient l'économiste québécois Eric Martin. Pour celui-ci, en effet, "Comme les besoins des industries et des marchés changent constamment, il devient urgent de former des travailleurs polyvalents et flexibles, dotés de compétences transversales qui les rendent adaptables à toutes les situations nouvelles, et de les renvoyer sur les bancs d'écoles pour assurer le renouvellement continu de ces compétences. Il devient important de diffuser les nouvelles technologies et d'organiser les entreprises plus efficacement en vue de maximiser les gains de productivité. Se met alors en place une véritable « économie de l'apprentissage » où le système d'éducation est vu comme un réservoir à idées permettant de développer de la main-d'œuvre de pointe (capital humain) pour les industries à « haute valeur ajoutée » et des innovations brevetables (propriété intellectuelle) qui permettront de soutenir les entreprises et de dynamiser la croissance".¹⁹

¹⁶ Il s'agit du système dénommé Mohammed VI, composé de 2 satellites chargés de la reconnaissance et de l'observation de la terre conçus par les français Thales Alenia Space et Airbus. Le satellite A a été lancé le 8 novembre 2017 et le satellite B le 21 novembre 2018.

¹⁷ Livré par la France et inauguré le 15 novembre 2018, ce train est prévu pour atteindre une vitesse commerciale de 320km/heure entre Kénitra et Tanger au nord du Maroc.

¹⁸ Haut commissariat au Plan, "Note du Haut Commissariat au Plan relative aux principales caractéristiques de la population active occupée en 2018".

¹⁹ [Eric, Martin](https://iris-recherche.qc.ca/blogue/qu%25e2%2580%2599est-ce-que-1%25e2%2580%2599-economie-du-savoir), "Qu'est-ce que l'économie du savoir ?" [En ligne], consulté le 18/11/2018, <https://iris-recherche.qc.ca/blogue/qu%25e2%2580%2599est-ce-que-1%25e2%2580%2599-economie-du-savoir>.

Les Etats arabes, clients d'armements plutôt que promoteurs de recherche scientifique, Ou lorsque l'économie du savoir souffre des priorités martiales des autorités politiques

Comme du temps de la société arabe auscultée par le sociologue tunisien Ibn Khaldoun (1332-1406), la faiblesse et la mauvaise répartition de la richesse nationale, jointes à la montée des mécontentements et à la limitation des libertés induit une augmentation des dépenses de sécurité et militaires. Et, aujourd'hui, le monde arabe, à sa tête l'Arabie Saoudite, affecte plus de ressources financières, en valeur relatives, à l'achat d'armes que n'importe quelle autre région au monde.

Ainsi, en 2017, les dépenses militaires (appelées "charges militaires") exprimées en pourcentage du Produit intérieur brut (PIB), étaient plus élevées au Moyen-Orient, avec 5,2% de la richesse nationale, que dans n'importe quelle autre région au monde. En effet, aucune autre région n'a consacré plus de 1,8 % de son PIB aux dépenses en armements. Les dépenses de l'Arabie saoudite, notamment, avaient augmenté alors de 9,2 % par rapport à 2016. Avec 69,4 milliards de dollars, soit plus de 50 % du PIB du Maroc en 2018, le royaume Wahhabite enregistrait la troisième plus forte dépense militaire au monde, derrière les Etats-Unis d'Amérique et la Chine²⁰.

Bien évidemment, dans une telle configuration prédéterminée par des politiques publiques d'ordre plus sécuritaire qu'économique ou social et fondées sur un instinct de défense, beaucoup plus que sur une volonté d'ouverture, d'inclusion et de progrès, et dans cette même posture où beaucoup de pays manquent de ressources au-delà de celles qu'ils consacrent à la préservation des intérêts de leurs gouvernants, les moyens alloués à l'enseignement ou encore à la recherche scientifique sont nécessairement réduits. Nous avons vu succinctement, ci-dessus, ce qu'il en était pour l'éducation et l'école dans le monde arabe. Le tableau 3 ci-dessous indique que seule la Libye affectait plus de 0,8 % de son PIB à la recherche scientifique, en 2013. Mais, depuis cette date ce pays a également chaviré dans la guerre civile, ou dans la guerre tout court, à l'image de la Syrie ou du Yémen.

²⁰ Communiqué de presse, 2 Mai 2018. [En ligne], consulté le 18/11/2018. <https://www.sipri.org/> .

Tableau 3 : Dépenses intérieures brutes de recherche et développement (DIRD) dans certains pays arabes, en pourcentage du PIB, 2009/2013

Année Pays	2009	2013	
Arabie Saoudite	0,07		
Bahreïn	0,04	0,04	
Emirats Arabes Unis		0,49 (2011)	
Egypte	0,43	0,68	
Irak	0,05	0,03	
Jordanie	0,43 (2008)		
Libye		0,86 (2014)	
Koweït	0,11	0,30	
Maroc	0,64 (2006)	0,73 (2010)	
Oman	0,13	0,17	
Qatar		0,47 (2012)	
Tunisie	0,71	0,68 (2012)	

Source : Rapport de l'UNESCO sur " la Science, vers 2030".

<https://unesdoc.unesco.org/ark:/48223/pf0000246417>

A titre de comparaison, l'effort de recherche, mesuré en rapportant les dépenses intérieures en recherche et développement au Produit intérieur brut (ratio DIRD/PIB), a atteint 2,34 % dans les pays de l'OCDE, en 2016²¹. Cette année là, ce même ratio a atteint 2,11 % en Chine, 2,74 % aux Etats-Unis, 3,14 % au Japon et 4,23 % en Corée du Sud. La "palme" revenant alors à Israël – Etat dominant aujourd'hui le Moyen-Orient et au-delà -, avec un ratio DIRD/PIB de 4,25 %, soit – en proportion – 6,25 fois l'effort de R&D égyptien en 2013.

En guise de conclusion

Comme nous venons de le voir dans les paragraphes ci-avant, la constitution de l'économie du savoir et de la société qui en est le réceptacle et le moteur, est un processus cumulatif, historique et social qui a impliqué la mise en œuvre de l'ensemble des paramètres culturel, politique, philosophique, artistique, technique, etc., marquant cette société.

Aujourd'hui, si la société européenne ou d'affiliation européenne, et si des sociétés comme la japonaise ou la coréenne du sud, sont au niveau de développement économique et humain qu'elles ont atteint et si les produits et services de leurs systèmes productifs et de recherche scientifique

²¹ "Panorama de l'effort de R&D dans le monde". Ministère de l'enseignement supérieur, de la recherche et de l'innovation. France. Note flash du SIES n° 19, novembre 2018.

irriguent plus ou moins intensément – sous leur contrôle – l'ensemble des régions de la planète, c'est bien parce qu'elles ont su profiter de tous ces paramètres, en les mettant au service de l'inclusion et du bien-être de tous leurs citoyens.

De ce point de vue, ce n'est absolument pas un hasard de l'histoire, si les pays où les citoyens sont les plus libres, les mieux éduqués et informés, les plus qualifiés, sont les plus développés et les plus riches, abstraction faite de leurs ressources naturelles, de leurs situations géographiques ou de leur organisation politique et institutionnelle.

A l'opposé, si la quasi-totalité des pays en voie de développement, et avec eux l'ensemble des pays du monde arabe, se trouvent dans la situation de dépendance financière, économique, politique et scientifique qui est la leur de nos jours, c'est bien parce, pour la plupart, il font encore très peu cas des droits humains fondamentaux, comme la liberté de pensée, d'expression ou d'organisation, l'accès à l'éducation, à l'information et aux résultats de la recherche scientifique et technique instantanément et librement, ou encore méconnaissent les principes de base de la démocratie, dont le droit de leurs citoyens d'élire librement leurs dirigeants et de contrôler les politiques publiques mises en œuvre par ces derniers, en leur nom. A ce propos, la seule comparaison entre les dépenses en R&D réalisées par la Corée du Sud, le Maroc ou l'Arabie Saoudite, en dit bien long sur les écarts économique, scientifique et, en forme de synthèse, de développement humain entre ces trois pays, qui peuvent être considérés comme représentatifs des situations de tous les autres. L'un, des pays qui ont bâti une économie du savoir en la fondant sur l'épanouissement et la créativité de son capital humain, les deux autres en faisant de leurs populations de simples consommatrices de produits industriels, scientifiques et culturels venus d'ailleurs, notamment – surtout - des Etats-Unis d'Amérique et d'Europe.

Références succinctes

ALBIGÈS, L. M. et Marine VASSEUR, « *Le développement des écoles primaires à la fin du XIX^e siècle* », *Histoire par l'image*. [En ligne], consulté le 21/11/2018. <http://www.histoire-image.org/fr/etudes/developpement-ecoles-primaires-fin-xixe-siecle>.

LAVELEYE, de É. “*L’Instruction du Peuple au XIX^e siècle. — III. — L’Instruction obligatoire et les moyens d’application*”. [*Revue des Deux Mondes*, 2^e période](#), volume 62, 1866 (p. 977-1012).

HARARI, Y.N. “Sapiens – une brève histoire de l’humanité”. Albin Michel, Paris, 2015.

HARARI, Y.N. “Homo deus – une brève histoire de l’avenir”. Albin Michel, Paris, 2017.

HARARI, Y.N. “21 leçons pour le XXI^e siècle”. Albin Michel, Paris, 2018.

Haut commissariat au Plan, “Note du Haut Commissariat au Plan relative aux principales caractéristiques de la population active occupée en 2018”.

MARTIN, E. “*Qu’est-ce que l’Economie du Savoir ?*”. [En ligne], consulté le 21/11/2018. <https://iris-recherche>.

qc.ca/blogue/qu%25e2%2580%2599est-ce-que-1%25e2%2580%2599economie-du-savoir.

OECD. “*Supporting Investment in Knowledge Capital, Growth and Innovation*”, OECD publishing. [En ligne], consulté le 21/11/2018
<https://dx.doi.org/10.1787/9789264193307>

PNUD. “*Rapport sur le développement humain*”. 2001 et 2018

UNESCO. “*Vers les sociétés du savoir*”. Rapport annuel - 2005. Editions de l’UNESCO

UNESCO. “*Rapport sur La science vers 2030*”. [En ligne], (Consulté le 21/11/2018).

<https://unesdoc.unesco.org/ark:/48223/pf0000246417>.

ZWEIG, S. “*Magellan*”. Editions Le Livre de Poche, Paris. Août 2018.

Thème 2

Aspects technologiques de l'Open Access

Mettre en pratique les recommandations sur les archives ouvertes de nouvelle génération de COAR pour l'archive ouverte institutionnelle, Horizon Pleins textes de l'IRD

*Putting into practice the recommendations of
COAR Next Generation Repositories report
into the context of the institutional repository,
Horizon Pleintextes from IRD*

Pascal Aventurier, IRD- France

Résumé

Longtemps les archives ouvertes ont été cantonnées à des fonctions, déjà très importantes, de stockage et de diffusion de l'information. Elles ont connu des évolutions fonctionnelles progressives depuis les années 2000, cependant un récent rapport de COAR (Confederation of Open Access Repositories) sur les archives ouvertes de nouvelle génération explique que celles-ci n'ont pas développé pleinement leur potentiel et ne sont pas assez connectées. Ce travail analyse comment les principales recommandations du rapport COAR pourraient être appliquées à l'archive ouverte institutionnelle Horizon Plein textes de l'IRD.

Mots-clés. Archive ouverte ; Accès ouvert ; Standards ; Système d'information ; COAR ; Fonctionnalités.

Abstract

For a long time, open archives have been confined to important functions of storage and dissemination of information. They have experienced progressive functional evolutions since the 2000s, however a recent report from Confederation of Open Access Repositories (COAR) on open archives of new generation explains that they have not developed their full potential and are not enough connected. This work analyzes how the main recommendations of the COAR report could be applied to the IRD's Institutional repository "Horizon Pleintextes".

Keywords. Repository; open access; standards; information system; COAR; Functionalities

1 Introduction

L'Institut de Recherche pour le Développement (IRD) est un organisme de recherche public français qui propose un modèle original : le partenariat scientifique équitable avec les pays en développement, principalement ceux des régions intertropicales et de l'espace méditerranéen, pour lier réduction des inégalités de développement, changements globaux, atteinte des objectifs du développement durable et innovation fondée sur la recherche. L'IRD a inscrit dans ses textes le recensement et la diffusion de ses connaissances scientifiques et plus récemment, dans son contrat d'objectifs et de performance 2016-2020, la diffusion des documents en libre accès. Ainsi est née il y a plus de 30 ans la base documentaire Horizon qui est l'une des plus anciennes archives ouvertes institutionnelles françaises (Couperin 2017).

Longtemps les archives ouvertes ont été cantonnées à des fonctions, déjà très importantes, de stockage et de diffusion de l'information. Elles ont connu des évolutions fonctionnelles progressives depuis les années 2000 comme l'ajout de fonctions bibliométriques, la gestion des listes de publications des chercheurs, la production de statistiques de consultation, la mise à disposition de flux RSS, ou encore une exposition des données avec le protocole OAI-PMH. Cependant un récent rapport du groupe de travail « Behaviours and Technical Recommendations of the COAR Next Generation Repositories Working Group » de COAR (Confederation of Open Access Repositories) sur les archives ouvertes de nouvelle génération explique que celles-ci n'ont pas développé pleinement leur potentiel (Rodrigues et al. 2017). Elles devraient évoluer pour être mieux intégrées dans l'écosystème de la production scientifique mais aussi dans les systèmes d'information des institutions. Le rapport du MIT sur le futur des bibliothèques (MIT 2016) va aussi dans ce sens et recommande que les bibliothèques génèrent des plateformes de contenu ouvertes et interopérables qui explorent de nouvelles méthodologies innovantes pour la découverte, l'organisation, le partage ou la production d'informations.

L'objectif de ce travail est tout d'abord de présenter et d'analyser comment les recommandations du rapport COAR pourraient être appliquées à Horizon, puis de dégager les fonctionnalités qui semblent les plus importantes, enfin de préciser le niveau d'avancement et de standardisation des développements libres réutilisables.

2 L'archive ouverte institutionnelle Horizon

Horizon¹ recense maintenant plus de 98000 documents dont 65000 PDF téléchargeables en libre accès en son volet Pleins textes. L'accroissement d'Horizon est d'environ 3000 documents par an (Cavet 2014). Les statistiques de consultation montrent qu'environ 7500 fichiers sont téléchargés chaque jour, dont 70% par des publics des pays du Sud.

Horizon utilise pour sa partie gestion documentaire le logiciel Cadic Intégrale de la société CADIC et le moteur SearchServer pour l'indexation des références et des PDF depuis 2006. Des efforts importants sont faits

¹<http://www.documentation.ird.fr/>

pour numériser des documents qui constituent un patrimoine important pour l'Institut (ROSSI 2017).

Horizon est dupliquée quotidiennement en SQL sur un environnement qui permet de fournir des services à valeur ajoutée aux chercheurs tels que des identifiants pérennes pour les références bibliographiques, des *landing pages*² des notices, des listes dynamiques de publications (individuelles, UMR) exportables en différents formats, intégrables (fonction 'embedded') dans des pages web personnelles ou de laboratoires, et des flux RSS. C'est dans cet environnement que sont disponibles au format Dublin Core les 98000 *records* de l'entrepôt OAI. En effet, Horizon est moissonnée en OAI-PMH par, entre autres, le CLACSO (Consejo Latino Americano de Ciencias Sociales)³, Base (Bielefeld Academic Search Engine⁴), OpenAire⁵ pour l'Union Européenne ou Isidore⁶ pour les SHS, le portail de l'information publique environnementale Tout sur l'environnement⁷.

Horizon dispose d'un troisième environnement : Horizon "Bibliométrie". Il est alimenté par les publications de l'IRD repérées dans le Web of Science (WOS) de manière semi-automatique à l'aide d'une base de reconnaissance des segments d'adresse des auteurs. Cette base "Bibliométrie" gère l'attribution des publications à l'unité de recherche, pour les chercheurs de l'IRD, mais également pour ceux des UMR (Unité Mixte de Recherche⁸) ; les métadonnées sont en outre enrichies par la mention des pays du Sud signalés dans l'affiliation d'au moins un des co-auteurs. Les références bibliographiques des chercheurs de l'IRD collectées pour la base "Bibliométrie" sont également importées dans Horizon et reçoivent un identifiant interne. Les résultats de la veille sur le WOS sont diffusés via un bulletin mensuel interactif et des listes de publications dynamiques pour chaque UMR et chaque chercheur de l'IRD. La base "Bibliométrie" est également utilisée pour alimenter des systèmes d'information institutionnels sur les activités de recherche.

Enfin, Horizon interagit avec l'archive nationale française HAL⁹. Une fonctionnalité permet au chercheur de déposer aisément un PDF dans le sous-ensemble HAL-IRD via l'identifiant interne IRD qui déclenche l'autocomplétion des métadonnées bibliographiques. Le chercheur se contente alors d'ajouter sur HAL le texte intégral dans sa version autorisée. Dans l'autre sens, tout dépôt dans HAL d'une (co) publication d'un agent IRD est signalé automatiquement à l'équipe Information Scientifique et Technique et est ajouté dans Horizon, après vérification et demande éventuelle de complétion.

² Les *landing pages* sont les pages sur lesquelles pointent les moteurs de recherche, par exemple.

³ <https://www.clacso.org.ar/>

⁴ <https://www.base-search.net/?l=fr>

⁵ <https://www.openaire.eu/>

⁶ <https://www.rechercheisidore.fr/>

⁷ <https://www.toutsurlenvironnement.fr/>

⁸ Une Unité mixte de recherche (UMR) est une structure administrative constituée par la mise en commun par différents établissements d'enseignement supérieur et de recherche de moyens humains et matériels fondée sur une stratégie et un projet scientifique communs.

⁹ <https://hal.archives-ouvertes.fr/>



Figure 1 : L'interface de la page d'accueil Horizon Plein textes (version en pré-prod au 02/09/2018)

La figure 1 présente l'interface et quelques fonctionnalités, dont les statistiques de consultation.

Les services actuellement offerts par Horizon

Le tableau 1 ci-dessous présente les services actuellement offerts par la base Horizon.

Tableau 1 : Services offerts par la base Horizon

Type de service	Description
Services aux chercheurs - Déposer	<p>Deux modes d'alimentation :</p> <p>Manuel : dépôt par mail du document, puis enregistrement par les documentalistes ou veille du service IST sur les revues, les sites d'éditeurs, les sites d'UMR puis enregistrement par les documentalistes</p> <p>Environ 1500 notices par an</p> <p>Semi-automatique : veille hebdomadaire sur le Web of Science via une équation de recherche sur les affiliations IRD, export dans EndNote : import mensuel des notices dans Horizon</p>

	<p>Environ 1400 notices par an Dans tous les cas, ajout et vérification des droits pour le texte intégral</p>
<p>Services aux chercheurs : Bibliométrie</p>	<p>Bulletin de veille mensuel, plus de 550 abonnés Suivi des publications : rapport bibliométrique Production de listes dynamiques individuelles ou d'UMR</p>
<p>Valoriser : dans les systèmes externes et les moteurs de recherches</p>	<p>Moissonnage OAI-PMH et compatibilité avec les directives européennes (OpenAire) et les infrastructures nationales (Isidore, ...) Indexation Google optimisée (déclaration de sitemap¹⁰ pour le référencement des PDF)</p>
<p>Valoriser: les fonctionnalités de la plateforme</p>	<p>Recherche dans les métadonnées Recherche plein texte dans le PDF Flux RSS Export Dernières saisies</p>
<p>Valoriser : les statistiques de consultation</p>	<p>Mesurer l'impact de la production de l'institution ou de sa production personnelle pour le chercheur via l'analyse des consultations de nos documents sur internet : indicateurs par document, ou sur des ensembles de documents (thématique, auteur, langue...) présentation dynamique des documents via des outils de mapping données librement accessibles par des URL libres</p>

Le fonds documentaire IRD (FDI), collection patrimoniale des publications des scientifiques de l'Institut, a été distingué fin 2017 par l'attribution du label CollEx, collection d'excellence pour la recherche, pour une période de cinq ans (2018-2022) reconductible. Ce fonds, conservé sur le site de la Délégation régionale Île-de-France à Bondy, est intégralement répertorié dans la base Horizon. Ce label permet de répondre à des appels à projets pour la constitution de fonds ou le développement de services pour les chercheurs.

Horizon tient également lieu de catalogue collectif de bibliothèque pour les dix bibliothèques du réseau international de l'IRD.

2.1 Evolutions envisageables pour la base Horizon

Plusieurs fonctionnalités sont à l'étude à différents niveaux. Elles n'ont pas encore été planifiées mais apporteraient de la valeur ajoutée à Horizon. Il s'agit de :

¹⁰ Un sitemap est un fichier XML ou texte standardisé qui indique au moteur de recherche les adresses URL des pages et leur relation pour pouvoir mieux les indexer.

- L'implémentation du module HAL, qui est une fonctionnalité du logiciel Cadic Intégrale, dans Horizon pour permettre d'envoyer directement des métadonnées dans HAL. Cela permettrait aux Unités Mixtes de Recherches pour lesquelles HAL (l'archive ouverte française) est le système d'évaluation de référence, d'avoir l'ensemble de leurs métadonnées dans une seule base ;
- L'intégration de l'identifiant numérique pour chercheurs, Orcid¹¹, afin de permettre une alimentation automatique des comptes chercheurs Orcid à travers cet identifiant unique ;
- L'ajout d'information pour lier la publication au jeu de données et réciproquement ;
- L'implémentation de l'API ISTE¹² qui permet de lier les 20 millions d'articles scientifiques achetés par ce projet et mis à disposition des chercheurs des organisations françaises.
- La mise à disposition des données bibliographiques en données liées ouvertes (linked open data) et alignées avec des vocabulaires de référence permettant ainsi de s'affranchir encore davantage du silo de données que représente une base de données.

3 Les recommandations du rapport COAR

En 2016, COAR (Coalition of Open Access Repositories) a lancé un groupe de travail pour identifier les fonctionnalités centrales pour la nouvelle génération d'archives ouvertes ainsi que les architectures et technologies qui permettraient de les implémenter. L'orientation globale est de mieux connecter les archives ouvertes en réseau pour constituer un patrimoine mondial de connaissances en encourageant la création de services à valeur ajoutée entre les entrepôts de données de documents que sont les archives ouvertes¹³.

Le rapport COAR identifie de nouvelles fonctionnalités principales et présente des recommandations pour l'adoption de nouvelles technologies, standards et protocoles qui peuvent aider les entrepôts à mieux s'intégrer dans l'environnement du web et à jouer un rôle plus large dans l'écosystème de la communication scientifique. Le rapport présente onze nouveaux comportements, technologies, standards ou protocoles qui faciliteraient le développement de services novateurs comme le travail en réseau, la validation par les pairs, la notification ou l'usage d'évaluation.

3.1 Présentation des recommandations du rapport

Les onze comportements "centraux" cités dans le rapport Next Generation of repositories de COAR sont traduits et présentés ci-dessous avec les technologies associées à chacun d'entre eux et une description.

1- Exposer les identifiants.

¹¹Orcid : <https://orcid.org/>

¹²Istex : <https://www.istex.fr/>

¹³ En anglais, le terme "repository" convient pour les entrepôts de données et de documents. En français, le terme "document repository" a conservé la traduction "archive ouverte" alors que les anglo-saxons ont abandonné le terme "open archive" et utilisent "document repository".

Les identifiants doivent être disponibles pour la landing page mais aussi pour le document en PDF par exemple.

Technologie associée : Signposting

2- Déclarer les licences au niveau de la ressource.

La licence et les conditions d'utilisation doivent être appliquées au niveau de chaque objet.

Technologies associées : Signposting, Creative Commons Licence

3- Découvrir à travers la navigation.

Lien direct vers les objets sans passer par la landing page car les outils ne peuvent pas suivre les liens et pour pouvoir accéder aux métadonnées en temps réel via une API.

Technologies associées : Signposting

4- Interaction avec des ressources (Annotation, Commentaire, Review)

Possibilité d'interagir avec les ressources : commenter, mention "j'aime",...

Technologies associées: Activity Streams 2.0, Atom syndication format, Web annotation Model and Web Annotation Protocol, International Image interoperability Network

5- Transfert de ressources

Les ressources (et/ou leurs métadonnées) ne devraient pas être copiées, principalement pour diminuer les temps d'accès.

Technologies associées : IPFS, Resourceync, SWORD

6- Batch Discovery.

Recherche de plusieurs ressources à travers différents entrepôts pour repérer et récupérer des documents pour le textmining ou la préservation par exemple.

Technologies associées : ResourceSync, SignPosting, Sitemaps

7- Collecter et exposer l'activité

L'activité de collecte et d'exposition dynamique et en temps réel des activités (par exemple, des informations sur les changements, ajouts, commentaires, annotations, peer-review, accès, téléchargements, etc.) concernant les objets scientifiques qu'ils hébergent.

Technologies associées : Activity Streams 2.0, Linked Data Notifications, ResourceSync (Change notifications), Signposting, WebMention, WebSub, Other messaging protocols (AMQP, Kafka)

8- Identification des utilisateurs

L'identification des utilisateurs permet d'être reconnu par le système, et de recevoir des recommandations, pour découvrir de nouvelles productions scientifiques potentiellement intéressantes pour lui.

Technologies associées : ORCID ;Social NetWorkIdentities, WebId

9- Authentification des utilisateurs

Demander aux utilisateurs de s'identifier eux mêmes avec des identifiants quand ils interagissent avec des ressources.

Technologies associées : http Signatures, OpenIdConnect 1.0, WebID/TLS

10- Exposer des métriques standardisées

Partage des données d'interaction ou des statistiques de consultation des utilisateurs, soit en mode Push (en utilisant Sushi), soit en mode Pull avec les protocoles de récupérations de statistiques par les agrégateurs (par exemple, OpenAire utilise Piwik).

Technologies associées : Counter, SUSHI, ETAG, Usage Metrics Service Provider for repository IRUS-UK ou Piwik RAMP. Voir la partie 5 transfert de ressources.

11- Préservation des ressources

Permettre la préservation sur le long terme. Le groupe de travail a statué que la préservation est une activité extrêmement complexe, impliquant l'adoption de politiques, normes, pratiques et technologies appropriées déjà mises en oeuvre par des communautés existantes.

Par contre la mise à disposition suivant le comportement 5 "Batch transfert" relève bien de la responsabilité des archives ouvertes et entrepôts.

Technologies associées : les mêmes que le point 5.

3.2 Analyse des recommandations

Parmi les technologies exposées, certaines sont déjà bien connues mais peut-être pas suffisamment implémentées dans les archives ouvertes comme par exemple la Creative Commons Licence, SWORD ou OpenIdConnect pour l'identification.

Les technologies Activity Streams, mais surtout RessourceSync et Signpost, semblent vraiment importantes et sont, d'après COAR en train d'être implémentées dans des plateformes.

Activity Streams : <https://www.w3.org/TR/activitystreams-core/>

Activity Streams 2.0 est une recommandation du W3C depuis le 23 mai 2017. Cette spécification décrit un modèle pour représenter les activités potentielles et terminées en utilisant le format JSON. Elle est destinée à être utilisée avec des vocabulaires qui détaillent la structure des activités et définissent des types spécifiques d'activités d'un nouvel objet pertinent ou d'une modification de métadonnées existantes. Il permettra aussi de proposer du contenu à des utilisateurs.

Signposting : <http://signposting.org>

Signposting est une approche pour rendre le web académique plus convivial pour les machines en exposant les relations en tant que liens typés dans les en-têtes de liens HTTP.

Par exemple dans une "landing page", la page html décrivant la ressource pour les consultations humaines, l'en-tête HTTP contient des informations typant la relation auteur et le reliant avec son identifiant Orcid (figure 2). Cette information est directement interprétable par les machines¹⁴.

¹⁴<https://signposting.org/author/springer/>

```

Link:<http://orcid.org/0000-0003-4018-5153>
; rel="author",
<http://orcid.org/0000-0002-2787-7166>
; rel="author",
<http://orcid.org/0000-0003-3749-8116>
; rel="author"

```

Figure 2 : un exemple d'en tête de signposting

Signposting sera implémenté dans Dspace ⁷¹⁵ et est déjà implémenté dans OJS (Open Journal System¹⁶), un logiciel libre pour la gestion complète de journaux scientifiques, dans des entrepôts de données et dans le schéma Datacite pour décrire les jeux de données scientifiques¹⁷.

Resourcesync : <http://www.openarchives.org/rs/toc>

Resourcesync est une spécification s'appuyant sur les sitemaps qui peuvent être utilisés par les administrateurs pour fournir des informations permettant aux systèmes tiers de rester synchronisés avec les ressources de l'entrepôt-source à mesure qu'elles évoluent (création, mise à jour, etc..). ResourceSync a été créé par la SLOAN Foundation¹⁸ et le JISC¹⁹.

L'utilisation de vocabulaires contrôlés en format normalisé SKOS²⁰ pour décrire les objets, permettant ainsi une description avec des référentiels communs et de mieux relier ou agréger les données est très importante. Cette recommandation n'apparaît que très peu dans le rapport.

Enfin, il est nécessaire d'avoir des agrégateurs ou concentrateurs de données ("hub" en anglais) interopérables pour permettre les échanges d'information à travers les archives ouvertes et les entrepôts de données. On peut se demander s'il s'agira de nouveaux systèmes ou si des plateformes comme Openaire ou HAL peuvent jouer ce rôle.

4 L'application de ces recommandations dans le contexte d'Horizon Pleins textes

Le principe d'échanges de contenus est très intéressant pour une archive ouverte "modeste" et offre de nouvelles possibilités bien supérieures au protocole OAI-PMH. Les technologies mentionnées sont plutôt légères à mettre en oeuvre et les comportements listés dans le rapport n'impliquent pas de modification fondamentale de l'architecture d'une archive ouverte.

La COAR préconise de focaliser les services sur les ressources plutôt que sur l'archive. Avec des outils qui ne relèvent certes pas encore des préconisations techniques de type Signposting ou autres, la base Horizon propose déjà des services associés aux ressources, qu'il s'agisse des sitemaps permettant de rendre visibles les PDF dans des moteurs de recherche, ou

¹⁵ <https://wiki.duraspace.org/display/DSPACE/DSPACE+Release+7.0+Status>

¹⁶ <https://pkp.sfu.ca/ojs/>

¹⁷ <https://signposting.org/adopters/>

¹⁸ <https://sloan.org/>

¹⁹ <https://www.jisc.ac.uk/>

²⁰ <https://www.w3.org/2004/02/skos/>

encore via les urls pérennes pour les métadonnées, ou les listes de publications dynamiques : c'est bien la ressource en elle-même, et non pas l'archive Horizon Pleins textes, qui est "poussée" ici sur le web.

Le rapport conforte des développements que nous avons envisagés comme l'attribution d'un identifiant Orcid pour les chercheurs, les échanges de métadonnées ou les liens entre archive ouverte (dépôt de documents) et entrepôt de données.

L'implémentation des fonctionnalités peut à moyen terme remettre en cause l'outil sur lequel est développé notre archive ouverte.

Cependant la base Horizon et son architecture actuelle permettraient de mettre en place facilement des recommandations du rapport. La base Horizon possède des permaliens vers les landing pages qui pourraient être utiles pour implémenter Signposting et elle présente déjà des sitemaps qui pourraient être reliés à la spécification Resourcesync. Le fait de connaître et de maîtriser la structure de la base peut aussi être un atout, bien que les recommandations ne portent pas sur la structure de la base en elle-même.

Parmi les faiblesses, on peut noter que Horizon a deux fonctions : un catalogue pour gérer des fonds papier, le Fonds Documentaire de l'IRD et ceux des bibliothèques du réseau IRD, et une fonction d'archive ouverte institutionnelle. Il faudrait pouvoir continuer à différencier ces deux fonctions dans les échanges de métadonnées et de documents.

Le rapport COAR insiste sur l'importance d'un système d'archives ouvertes distribué, et non pas centralisé. Il faudrait donc que les archives ouvertes connectées suivent l'ensemble des recommandations pour réaliser pleinement le potentiel des technologies projetées dans le rapport.

Enfin, deux besoins ne sont pas abordés dans les recommandations :

le besoin de listes bibliométriques n'est pas évoqué alors que les archives ouvertes participent à l'identification des productions scientifiques des organismes de recherche et dans certains organismes représentent l'une des fonctionnalités principales.

Le web sémantique et les données liées ouvertes (linked open data). Pourtant, le linked open data est une composante de plus en plus importante du web, en particulier dans la recherche d'information scientifique. En exposant les données - données bibliographiques ou autres- converties dans le standard RDF, celles-ci peuvent être publiées et liées avec d'autres données, ce qui permet de naviguer entre les données de la même manière qu'entre les pages web qui sont reliées avec des liens hypertextes. En exposant les données d'une archive ouverte en linked open data on expose les données en dehors des silos de données et on augmente le potentiel de réutilisation, permettant à d'autres projets, utilisateurs finaux, ou à des outils qui savent lire ces formats normalisés d'exploiter ces données. Certaines archives ouvertes exposent leurs données dans ce format via un sparqlendpoint comme dans la plateforme Persée par exemple (<http://data.persee.fr/>). Pour augmenter le potentiel de réutilisation il est important de lier ses données avec des vocabulaires contrôlés, (Gottschling

et Subirats2014) qui sont communément utilisés et centraux dans le nuage du linked data²¹ ou des sources de qualité plus spécifiques.

5 Conclusion

Cette réflexion doit être considérée comme un préalable à l'intégration progressive des fonctionnalités et des standards communs qui permettront de rendre l'archive ouverte plus attractive et de mieux répondre à de nouveaux besoins institutionnels et scientifiques. Elle pourra aussi permettre à d'autres archives ouvertes d'effectuer la même analyse.

Les technologies préconisées sont en cours d'implémentation dans différentes plateformes qui agrègent des données provenant des archives ouvertes et des entrepôts de données, et proposent des services à valeurs ajoutées. Il s'agit de :

- OpenAIRE²² – Europe
- National Institute of Informatics (NII)²³ - Japan
- US Next Generation Repositories Implementers Group
- CARL Open Repositories Working Group²⁴ - Canada

Il faudrait que la grande majorité des entrepôts ou archives ouvertes adoptent ces technologies pour que les échanges puissent réellement s'opérer.

La plupart de ces technologies devraient être disponibles en standard sur les principaux logiciels libres permettant de construire des archives ouvertes ou des entrepôts de données (Dspace, Eprints, Dataverse...) et devraient sans doute aussi être développées dans les plateformes nationales (HAL), ou sur des outils moissonnant différentes sources comme BASE (Bielefeld Academic Search Engine).

Le 6 juin 2018, un workshop sur l'adoption des recommandations a mis en avant les priorités suivantes (SHEARER et al. 2018) :

- Travailler avec les développeurs des logiciels d'archives ouvertes pour faire connaître ces recommandations et développer les cas d'usages pour assurer une plus grande compréhension et un plus grand engagement de la communauté;
- Donner plus d'information à ces logiciels sur les technologies du rapport ;
- Aider à trouver des fonds de la part des organisations intéressées par ces technologies dans les différentes plateformes ;
- Maintenir un annuaire des implémentations dans les différents logiciels ;

²¹<https://lod-cloud.net/>

²²<https://www.openaire.eu/>

²³<https://www.nii.ac.jp/en/>

²⁴<http://www.carl-abrc.ca/advancing-research/institutional-repositories/open-repositories-working-group/>

- Assurer un suivi sur le développement des fonctionnalités de réseaux sociaux.

Ces actions, que le groupe de travail Next Generation of Repositories de COAR devra mettre en place, permettront de suivre les différentes implémentations pour réaliser une infrastructure distribuée et globale, associée à des services de qualité, qui bénéficiera à l'ensemble de la communauté académique.

Remerciements

L'auteur remercie l'équipe IST (Informatique Scientifique et Technique) de l'IRD pour leurs conseils et leurs contributions sur l'évolution technique de l'archive ouverte Horizon Plein textes

Références

- CAVET, Dominique (2014). Horizon l'archive institutionnelle de l'IRD. 45p. In *Archive nationale HAL, archives institutionnelles, quelle complémentarité ? Urfist de Bordeaux*. 2014/06/11, [En ligne]. Disponible à : http://weburfist.univ-bordeaux.fr/wp-content/uploads/2016/03/2014_URFIST-CAVET_IRD.pdf (Page consultée le 30 août 2018).
- COUPERIN. G. TAO – Groupe “ Etude Enquête Archives Ouvertes”
COUPERIN : Résultats de l'enquête (2017). [En ligne]. Disponible à : http://www.couperin.org/images/stories/AO/Depouillement_enquete_AO2017.pdf (Page consultée le 30 août 2018).
- GOTTSCHLING, Maxie ; SUBIRATS, Imma (2014). 7 things you should know about LinkedData. [Enligne]. Disponible à : <https://www.coar-repositories.org/community/events/archive/repository-observatory-third-edition/coar-talks-ir-cris-interoperability/second-edition-linked-open-data/7-things-you-should-know-about-open-data/>. (Page consultée le 30 août 2018).
- MIT. MASSACHUSETTS INSTITUTE OF TECHNOLOGY (2016). Institute-wide Task Force on the Future of Libraries. Preliminary Report. [En ligne]. Disponible à : <https://future-of-libraries.mit.edu/sites/default/files/FutureLibraries-PrelimReport-Final.pdf> (Page consultée le 30 août 2018).
- RODRIGUES, Eloy; SHEARER, Kathleen; COAR Next Generation Repositories WG (2017). Behaviours and Technical Recommendations of the COAR Next Generation Repositories Working Group », [Enligne]. Disponible à : <https://www.coar-repositories.org/files/NGR-Final-Formatted-Report-cc.pdf> (Page consultée le 30 août 2018).

ROSSI, Pier Luigi (2017). Numérisation et bibliothèques électroniques : bilan de 15 années de formations réalisées par l'IRD au bénéfice des professionnels de pays de l'Afrique francophone, 12 p. In *Colloque International du Cinquantenaire de l'Ecole des Bibliothécaires, Archivistes et Documentalistes (EBAD) : Les Sciences de l'Information Documentaire au Service de la Recherche, de la Formation, de l'Intégration et du Développement Durable*, Dakar (SEN), 2017/11/14-16.. [En ligne]. Disponible à : <http://www.documentation.ird.fr/hor/fdi:010071833> (Page consultée le 30 août 2018).

SHEARER Kathleen; RODRIGUES Eloy; BOLLINI , Andrea, CABEZAS, Alberto; CASTELI, Donatella; CARR, Les; CHAN, Leslie; HUMPHREY, Chuck; JOHNSON, Rick; KNOTH, Petr Knoth, MANGHI, Paolo MATIZIROFA, Lazarus; PERAKAKIS, Pandelis; SCHIRRWAGEN, Jochen; SMITH, Tim; VAN DE SOMPEL, Herbert; WALK, Paul; WILCOX, YAMAJI, Kazu. (2018). Next Generation Repositories. Scaling up repositories to a global knowledge commons. »Zenodo. [En ligne]. Disponible à : <https://doi.org/10.5281/zenodo.1306452>. (Page consultée le 30 août 2018).

Text data mining and data quality management for research information systems in the context of open data and open science

Otmane Azeroual¹, Gunter Saake², Mohammad Abuosba³, Joachim Schöpfel⁴

¹German Center for Higher Education Research and Science Studies (DZHW), Schützenstraße 6a, 10117 Berlin, Germany

²Otto-von-Guericke-University Magdeburg, Department of Computer Science, Institute for Technical and Business Information Systems - DatabaseResearch Group, P.O. Box 4120, 39106 Magdeburg, Germany

³University of Applied Sciences HTW Berlin, Department of Computer Science and Engineering, Wilhelminenhofstraße 75A, 12459 Berlin, Germany

⁴GERiiCO Laboratory, University of Lille 3, 59650 Villeneuve-d'Ascq, France

Abstract

In the implementation and use of research information systems (RIS) in scientific institutions, text data mining and semantic technologies are a key technology for the meaningful use of large amounts of data. It is not the collection of data that is difficult, but the further processing and integration of the data in RIS. Data is usually not uniformly formatted and structured, such as texts and tables that cannot be linked. These include various source systems with their different data formats such as project and publication databases, CERIF and RCD data model, etc.

Internal and external data sources continue to develop. On the one hand, they must be constantly synchronized and the results of the data links checked. On the other hand, the texts must be processed in natural language and certain information extracted. Using text data mining, the quality of the metadata is analyzed and this identifies the entities and general keywords. So that the user is supported in the search for interesting research

information. The information age makes it easier to store huge amounts of data and increase the number of documents on the internet, in institutions' intranets, in newswires and blogs is overwhelming. Search engines should help to specifically open up these sources of information and make them usable for administrative and research purposes.

Against this backdrop, the aim of this paper is to provide an overview of text data mining techniques and the management of successful data quality for RIS in the context of open data and open science in scientific institutions and libraries, as well as to provide ideas for their application. In particular, solutions for the RIS will be presented.

Keywords. Current research information systems, Research information systems, Research information, Standardization, Text analysis, Data mining, Knowledge discovery database, Data quality management, Open data, Big data, Open science.

1 Introduction

The flood of research information in higher education and research organizations is steadily growing. Research information systems (RIS) are used to support the collection, integration, processing, storage and presentation of research information. With the help of RIS it is possible to get an up-to-date overview of research activities, to process and manage information on scientific activities, publications, applications, projects as well as research data, press and media reports, etc. For the evaluation and interpretation of the research activities of a scientific organization, data quality and validation are crucial. In order to ensure the quality of RIS in the long term, it is recommended that data quality management (data profiling, data cleansing and data monitoring) be viewed as a holistic and continuous improvement process.

Data quality management becomes even more crucial in the new ecosystem of open data and open science, an “approach to research that is collaborative, transparent and accessible” (European Commission, 2017). The policy of open science has a direct impact of research management and research evaluation insofar henceforth they have to consider and assess new forms of research practice and aspects of openness. Moreover, the process of research evaluation is changing, shifting from a traditional top-down strategy towards more interaction, dialogue and co-production (Tatum, 2017).

The knowledge of research activities and research results is becoming an increasingly important factor for the success of an organization. Yet, this knowledge requires, alongside quantitative data, more and more qualitative information extracted from a large variety of heterogeneous and often unstructured sources on the web, such as data and text files in digital libraries, open repositories, social networks and other platforms. For this reason, the new technology of *text mining* or *text data mining* of the rapid growth of digital, unstructured and semi-structured internal and external data is becoming increasingly important as a scientific method.

Text mining is a technique for analyzing documents or texts and extracting new knowledge unknown to the user. Thus, this developed technology is relevant for all scientific institutions and libraries that produce texts (such as editorial offices and publishers), edit text archives (libraries) or in which large amounts of text are constantly being generated. The aim of text mining techniques is to uncover interesting information or patterns in unstructured textual documents by allowing them to process the vast amount of natural language words and textures, and to handle the treatment of uncertain and fuzzy data. Text mining is a new area of research to solve problems of information overload through the use of data mining techniques, machine learning, computational linguistics (natural language processing), information retrieval, and knowledge management.

The aim of this paper is to present the technology of text data mining and the data quality management and to show how both can improve the performance of RIS, especially in the context of open data and open science.

2 Research information systems

A current research information system (CRIS) or research information system (RIS) is a *database management* or a *specialized federated information system*, that provides a continually updated, comprehensive directory of researchers and research activities (like for example, employees, projects, publications, patents, cooperation partners, awards and conferences etc.) of an organization. It serves to make the portfolio of research activities more visible and thus provide an overview for all interested or users.

The selection of specifications to be included in the RIS is based on the German Research Core Dataset (RCD). The RCD is a standard for the collection, provision and exchange of research information. The RIS in the context of the RCD is the instrument for the uniform definition and structuring of research information. To reach reliable data quality and optimized reuse of research information in and out of the institutions and libraries of distributed and heterogeneous databases, it is important to establish a high degree of connectivity through RCD for the effective exchange of research information in order to avoid duplication and multiple input of research information and the resulting sources of error. In addition to improving data quality, the RCD can greatly simplify work and reporting processes.

The building blocks of a RIS architecture are distinguished in three-stage processes (*data collection and data integration, data storage and data presentation*) (as illustrated in Fig. 1).

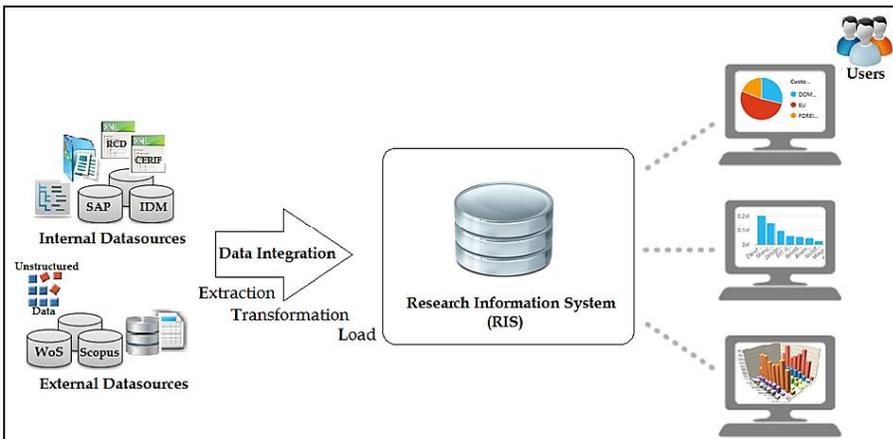


Figure 1. Architecture of RIS (Azeroual et al. 2018c, d)

The *data collection* contains the internal and external data sources. The RIS collects information about the research activities and research results associated with institutions and their scientists by automatically synchronizing existing data volumes with various data sources, so that this qualified research information is on hand to the management as a better basis for decision-making. For an automated data import from existing

systems, a connection of internal as well as external application systems can be realized. These application systems, which ideally can be included to collect research information, are the identity management system and the campus management system as internal systems, as well as public publication databases and project databases. In the case of publications, these include, for example, Web of Science, Scopus or PubMed and the financial system (e.g. SAP) of third-party funds management for the data on third-party funded projects and the personnel management system (e.g. PAISY) for information about the scientific staff.

Offers for the standardized collection, provision and exchange of research information in RIS are the Research Core Dataset (RCD) data model and the Common European Research Information Format (CERIF) data model.

The **data integration** transfers, processes and compresses the required data from source systems (different information systems with different data formats and structures), which is described as the Extraction, Transformation and Load (ETL) process. In the extraction phase, data is extracted and pulled out from various source systems or source documents and provided for further processing steps in the input layer of the RIS. The transformation phase has the task to prepare the data for the loading process and to clean it up in order to convert the data extracts into a uniform internal format. The cleansing of incorrect source data is absolutely necessary and will be detected and corrected by plausibility checks. This data cleansing is done in four steps: filtering, harmonization, aggregation, and enrichment. In the next and last phase, the data is loaded into the RIS, there they are stored in a structured and normalized.

The aim of the ETL process is to ensure that the prepared data can be stored efficiently and persistently in data storage.

The **data storage** includes the RIS and its applications that merge, manage and analyze data held at the underlying level. If the data is present in the RIS, the reports required for evaluation can be set up.

In the **data presentation** is the target group specific preparation and presentation of the analysis results for the end user (decision-makers). These are made available as reports using the business intelligence tools. Beside to various reporting options, portals and websites of the organizations can also be filled here.

3 Standardization of research information

To ensure the aggregation, standardized collection and interchangeability of research information and to be able to integrate as many decentralized stocks of research information as possible, there are (inter)national standards for supporting RIS.

Two measures for the standardization of research information are the well-known international standard CERIF of the European organization

euroCRIS¹ and Research Core Dataset (RCD)², which the German Council of Science and Humanities recommended in 2013. To support and advise the higher education institutions and non-university research institutions of the German science system on the implementation on the RCD, a corresponding helpdesk was established at the German Center for Higher Education Research and Science Studies (DZHW)³ in Berlin, which answers questions about the content and technical of the RCD and provides an information base and cooperate events and workshops on the RCD.

- CERIF is a format recommended by the EU for its members that provides descriptions for research information (such as *persons, projects, organizations, publications, patents, products, funding, equipment*etc.). The CERIF model⁴ can serve as a basis for RIS as an exchange format among various systems of research information.
- The implementation of the RCD at higher education institutions and non-university research institutions in Germany takes place on a voluntary basis. The RCD is a young data standard for research activities and considers it to be more effective and simpler for the complex and inconsistent reporting process for all research stakeholders from the German science system, as well as to reduce the collection of data from research activities and to facilitate the comparability of research data. The RCD defines core data that should be kept in aggregated form by all higher education institutions and non-university research institutions for various documentation and reporting processes. These include the focus areas of (a) *employees*, (b) *support for young researchers*, (c) *third-party funded projects*, (d) *patents and spin-offs*, (e) *publications* and (f) *research infrastructures* (see Biesenbender and Hornbostel, 2016 for details on the structure and logic of the RCD). In order to make the data compatible and interoperable at the international level, the technical RCD data model⁵ offered was developed on the basis of the CERIF format.

The following Fig.2 illustrates the example of the XML schema as a data exchange format for CERIF and RCD.

¹<https://www.eurocris.org/>

²<https://www.kerndatensatz-forschung.de/>

³<https://www.dzhw.eu>

⁴See <https://www.eurocris.org/cerif/main-features-cerif>

⁵See https://kerndatensatz-forschung.de/version1/technisches_datensmodell/

A recent report written by the EC “Working Group on Rewards under Open Science” provides a matrix for the evaluation of research careers fully acknowledging open science practices (European Commission, 2017). This matrix OS-CAM provides a comprehensive framework that can be used to develop evaluation systems for various situations and contexts. It proposes a large number of possible evaluation criteria for the assessment of six domains and 24 open science activities, i.e. datasets and research results, risk management, peer review, knowledge exchange, teaching and even personal qualities.

OS-CAM tries to describe doing open science on the individual level, and it illustrates the multiple dimensions and complexity of what is called “open science”. On the institutional level, the Harvard Open Access Project (HOAP)⁸ from the Berkman Klein Center may provide another framework for the assessment of open science policies and services in universities, research organisations and other scientific structures.

RIS are not designed for a given science policy but are (or at least should be) “policy agnostic”, able to comply with traditional research strategies as well as with the emerging new ecosystem of open science. Of course, insofar this new ecosystem changes the way research is done and modifies the way it is evaluated, it will have an immediate and direct impact on the RIS themselves. For instance, the French National Plan calls for the recognition of open science in assessments of researchers and institutions and requests to prioritise quality over quantity when evaluating research (MESRI, 2018).

Regarding the impact on RIS, we can distinguish three aspects.

1. Information sources: The RIS must be able to process data from new information sources, beyond the traditional data providers and outside academia, in order to monitor and evaluate new configurations of openness, involving other people, institutions and activities like NGOs, social networks, corporate companies etc. The reliability of these sources may be uncertain, such as the quality of the data; probably this means to adopt a strategy similar to the “big data” approach in the corporate sector.
2. Results: Open science changes the way how research is valued, and open science monitoring requires other metrics than the traditional H-index, impact factor article counting etc. “Evaluation outcomes aimed at present and future role of openness in research” (Tatum, 2017). Thus RIS must be able to produce open science related knowledge in a reliable way, representing societal impact (such as altmetrics⁹), openness (such as the SPARC open access spectrum¹⁰) or other qualitative information about participation, networking etc.
3. Process: Finally, RIS must take into account the transformation on the research evaluation process itself. As already mentioned above, research evaluation is shifting from a traditional top-down strategy

⁸<https://cyber.harvard.edu/research/hoap>

⁹<https://altmetrics.org/manifesto/>

¹⁰<https://sparcopen.org/our-work/howopenisit/>

towards more interaction, dialogue and co-production (Tatum, 2017), and this transformation takes place on all levels, on the individual (career assessment, grant awarding), local (institutional), organisational and national level. For the RIS, this means more transparency and more participation from the very beginning; both are critical factors for the success of a RIS, especially in the environment of open science (Schöpfel, 2015). Also, its implementation should be an opportunity to make researchers aware of the new challenges (de Castro, 2018). In other words, a closed RIS in an environment of open science makes no sense.

Open data means data sharing and data reuse, and this applies also to research information. How do RIS produce reusable data? How do they reuse data produced from other systems? And how do they make research information available as open as possible? And do they, really? These questions are all related to standards and policy.

Returning to quality, (Azeroual et al. 2018c) highlight the need for high quality data produced by RIS is an essential consideration for evaluating research activities of all institutions. Because of the increasing heterogeneity of data input, of new evaluation processes and of the requirement of standard, machine-readable output (interoperability with other systems and infrastructures, cf. FAIR guiding principles), the quality especially of metadata becomes a crucial asset of the development and administration of RIS (Simons, 2017).

The impact of open science on the implementation of a RIS has been described by different case studies, with special attention to the role of academic libraries and information professionals (see for instance Clements and Proven, 2015 and Brennan, 2018). In the following chapters, our paper will focus on the metadata problem, especially when related to unstructured and semi-structured data. How can RIS produce and process high quality output based on low quality input?

5 Text data mining

5.1 Definition

For over thirty years, text mining has been a relevant topic in research, although at the beginning it was only used in a few disciplines (Upshall, 2014). The term text mining is a special form of data mining and is referred to as "*text data mining*" or "*text analytics*". The text data mining refers to computer-based methods for the semantic analysis of texts, which support the automatic or semi-automatic structuring of texts, especially very large amounts of texts (Heyer et al. 2008). In particular, text data mining aims to identify and extract knowledge that is implicit in the text that the user of the information system does not know.

5.2 Metadata – handling unstructured and semi-structured data

Higher education institutions and non-university research institutions that manage their research information with larger data sets in RIS face the challenge because their information is not available in structured data but

in semi-structured or unstructured data. The unstructured data comprises, among others, text documents or text publications in Word or PDFs, memos, emails, RSS feeds, wikis, images, video or audio data etc. Semi-structured data include, for example, CERIF or RCD XML or HTML files. Such data can be a time-consuming burden for RIS staff in institutions and libraries to seek and thus make bad decisions, because the unstructured and semi-structured data are stored in different formats and designated by different technologies. To solve this problem, metadata, which is data about data, should be stored with unstructured or semi-structured data. By implementing text data mining techniques, for instance classification and clustering algorithms (see below), metadata about research activities can be generated and organized (Seth et al. 2009 and Bhanuse et al. 2016).

5.3 Integrating text data mining into research information systems

Today, information on scientific activities and outcomes of science organizations is collected, maintained and published in a variety of forms and vast amounts of data through RIS. However, these are mostly unstructured in various forms and media and present a great challenge. A traditional, manual approach is not enough to investigate large amounts of data or document collections from various meta-databases (e.g. PubMed, Scopus etc.). Therefore, text data mining techniques can be helpful to analyze large numbers of text publications, library-produced data or even the content of blogs or tweets. The following Fig. 3 shows a framework for supporting text data mining in RIS.

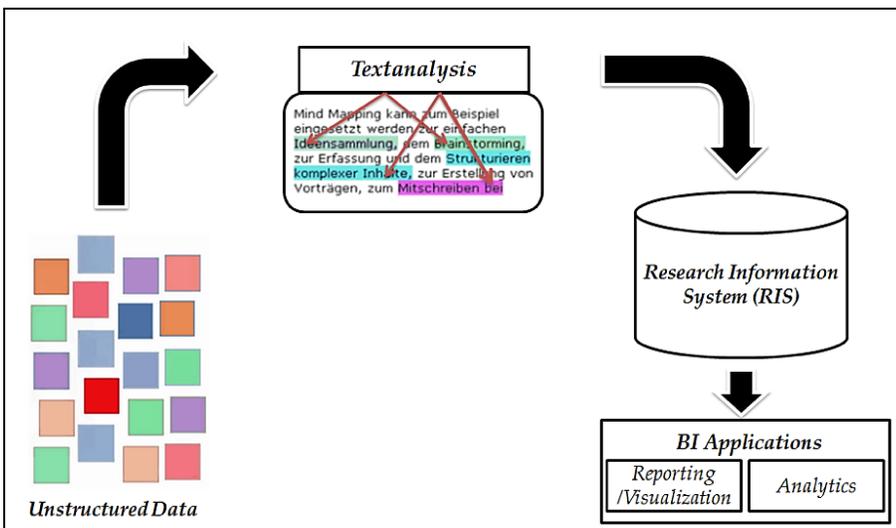


Figure 3. Overview of support text data mining framework to RIS

The peculiarities of text data mining are the structure of the data as well as the origin of the data. The data type to be examined is unstructured text from internal or external sources.

In the following section, the most important techniques of text data mining for supporting RIS in the context of open data and open science are presented in an overview for the institutions and libraries that use them.

6 Techniques in text data mining to support research information systems

Because of the vast amount of textual data available in RIS from various internal and external sources, it is not possible to read them completely and correctly and perform a manual analysis. In this case, text data mining is needed to convert text into data that can then be examined using various analysis techniques. These techniques are classification, clustering, NLP, information extraction and information retrieval.

6.1 Document classification / Categorization

Classification assigns research information in different categories. For example, in this way publication texts can be classified according to their content-related focus (e.g. computer science, economics, politics etc.). Classification procedures identify patterns to make statements about objects based on existing research information. First, the already existing objects are grouped according to their known feature or behavior regarding the problems to be analyzed in different classes. From this set of objects a classification model is developed, with which one can then predict the class affiliation of a new object. In the classification analysis methods of artificial intelligence or decision-tree-oriented methods can be used.

6.2 Document clustering

Clustering consolidates documents into clusters. Clusters of documents can be used to quickly find similar documents and detect duplicates. Unlike classification, clustering does not use a predefined set of terms or taxonomies that are used to group the documents (Liao et al. 2012). Instead, groups are created based on the document features that appear in a set of documents to be clustered.

6.3 Natural Language Processing (NLP)

In order to extract research information from natural language texts, it is necessary to use methods of natural language processing. NLP is an application of computational linguistics, which is responsible for the communication and interaction of humans and computers. The goal is for computers to understand human statements, from single words to complex texts, and to process their content or based on given circumstances, generating natural language as output (Chopra et al. 2013). For the automatic semantic analysis of texts, the use of ontologies is an essential tool. Ontologies are controlled, structured vocabularies that contain entities of a domain and their relationships to each other. They are necessary to make texts interpretable for the computer (Carstensen et al. 2010). Ontologies are used for named entity recognition (NER), which identifies proper names of entities (such as persons, organizations, etc.) in texts. This can be used to automatically keyword texts.

6.4 Information extraction

In information extraction (IE), research information is found in textual documents and translated into structured form (that is, identifying specific text passages and assigning them to structured objects). For example, research information about persons, organizations, publications, projects, etc. can be found and correlated.

6.5 Information retrieval

Information retrieval (IR) is the finding or retrieving of relevant documents that provide answers to specific questions. Here, texts are searched for using keyword requests, and a common example is keyword searches using search engines like Google, Yahoo, or Bing. Thus, the amount of relevant documents should be found by means of keywords and presented to the user in a result set.

Many institutions are still limited in their RIS still on the pure data collection and often go cumbersome ways to make it usable for themselves. With these steps of text data mining techniques, all institutions can better analyze and understand their stored texts in order to make appropriate decisions.

7 A model for research information processing

To assist the scientific institutions and libraries, the following developed process could be used to extract new and interesting information from different sources of unstructured documents using the techniques of text data mining and to derive new conclusions (see Fig. 4). In carrying out these featured text data mining techniques, there are a variety of supported commercial open source software solutions that offer multiple text data mining techniques. Nowadays, a developer, but no longer an expert, has to be in order to harness the potential of texts.

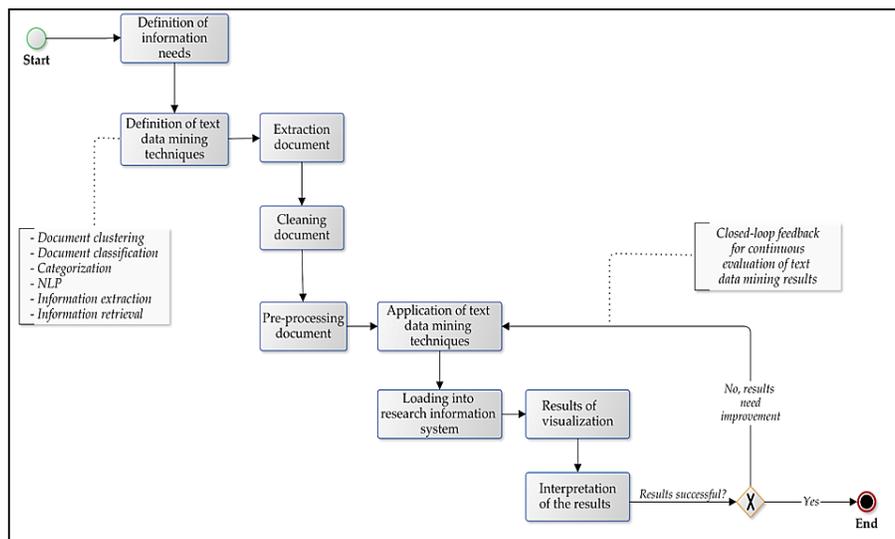


Figure 4. Process of knowledge discovery using text data mining techniques

Text data mining techniques can be used not only in scientific institutions, but also by libraries to improve their services, such as improving search and support in keyword coding.

8 Data quality management as a success factor for research information systems

In the scientific institutions and libraries, the management places the highest demands on the quality of the research information to be processed and evaluated. One of the biggest hurdles in securing data quality is typically the heterogeneity of the various internal and external data sources. Owning the right technology is just one way of ensuring data quality in RIS. Much more important to success is the use of the right methodology. This can be achieved by using the following data quality management as a cycle with its three phases, which must be run continuously to achieve a sustainable improvement and increase in data quality in RIS (as shown in Fig. 5).

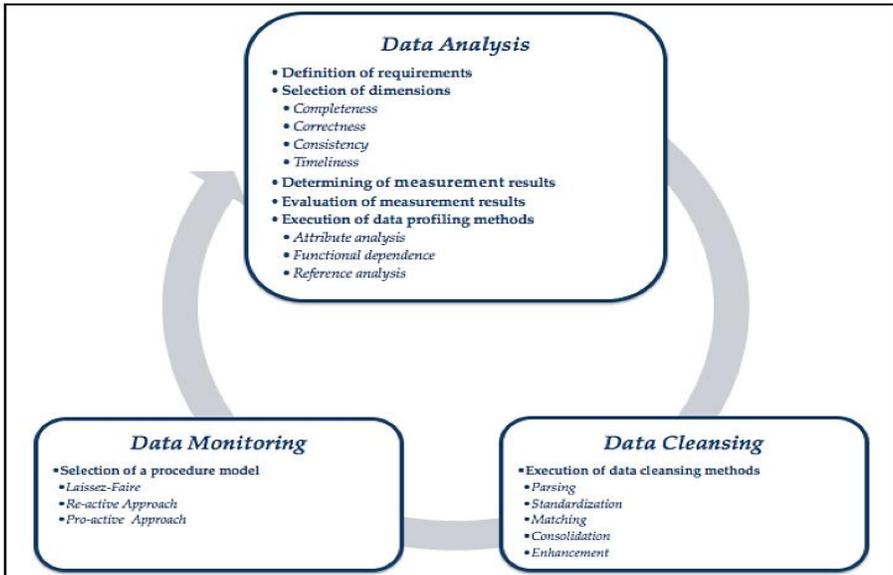


Figure 5. Data quality management cycle in RIS

The three phases capture data analysis, data cleansing and data monitoring, and together form a proven, practical approach to data governance - towards controlled data management in RIS that enables institutions to analyze, cleanse and control their research information. These phases are explained in more detail below.

In the first phase, the data quality criteria and corresponding requirements for them must be defined. Requirements can be set by different parties, e.g. especially by users of a RIS, but also by the RIS administrator. By determining the requirements, the data quality can be measured with appropriate metrics - completeness, correctness, consistency and timeliness (Azeroual, et al. 2018c), because they are easy to measure and represent a particularly representative illustration of the reporting for the users lead to an improved and transparent basis for decision-making (Azeroual, et al. 2018c). After the data quality measurements have been made in RIS, the institutions must review their measurement results. With the help of data profiling, causes of quality problems can be revealed (Azeroual, et al. 2018b). This data analysis and its methods are used to better understand the structure, relationships and content of different data sources in RIS and to detect and automatically correct potential errors (Azeroual, et al. 2018b). Data profiling is considered an important component in analyzing and improving data quality before data can be integrated into RIS (Azeroual, et al. 2018b). The institutions and libraries must use data profiling as early as possible to get an accurate picture of the state of their data. Because data profiling is the foundation for planning the most meaningful approach to correcting and harmonizing information assets.

Data cleansing can correct the identified issues (incomplete, incorrect, inconsistent, etc.) to increase the informational value and usefulness of the

research information (Azeroual, et al. 2018a, d) (Azeroual and Abuosba, 2017). Using data cleansing methods and technologies such as parsing, standardization, enrichment, matching, and merging within the data cleansing process will liquidate data integration issues in RIS (Azeroual and Abuosba, 2017) (Azeroual, et al. 2018a, d). As these methods are essential for achieving and maintaining maximum data quality for more success in RIS (Azeroual, et al. 2018a, d).

If a certain data quality has been achieved, it should be preserved as long as possible. During monitoring, therefore, the data is continuously checked before being stored in the RIS because the data is constantly changing. Such long-term continuous assurance, enhancement and increase of data quality in RIS requires specific measures such as "laissez-faire", "re-active approach" and "pro-active approach" (Azeroual, et al. 2018a, d). At certain intervals, a periodic check of the entire research data should take place again and again. Even if no mistakes are made in the input of research information into RIS, relocations, mergers and other factors repeatedly result in changes to the data. Therefore, only through a permanent data control, the device is able to provide anytime information about their data quality status, as well as to increase confidence in the existing data.

To successfully ensure data quality, institutions and libraries must find a methodology and solution that includes all three key components: data analysis, data cleansing and data control. By combining this sophisticated methodology with their appropriate technology, data quality in RIS is significantly improved and increased. This will lead to sound decisions not only for information purposes and reporting but also for the creation of the information value. Than both are based on the availability of high quality data.

9 Conclusion

Text data mining has become progressively important as a scientific method in recent years and its potential is very high. On the one hand, most of the information is in text form and on the other, knowledge and information about research activities play a very important role in the success of an institution. The automated acquisition of research information and the associated advent of vast amounts of data make automated data analysis necessary in many areas. Likewise, the steady growth of scientific activities (such as publications, projects, etc.) and information is making computer-aided techniques of classification, clustering, NLP, information extraction, and information retrieval increasingly important. Text data mining techniques can be actively used not only by institutions, but also actively by libraries to improve their services or to develop new services. In addition, more and more tools are being developed to make the relatively complicated process of text data mining and the associated advantages usable for users without appropriate text data mining knowledge. However, a minimum level of knowledge in the area of databases is also necessary.

In summary, one can say that text analytic methods or text data mining techniques are today a component of advanced method development in institutions and libraries. The application and further development of such

methods or techniques allows a long-term cost savings in the processing of service projects, addressing the increased expectations of scientific and scientific actors in the context of the establishment of open data, open science and big data, increasing the quality of research projects and the reuse already manually processed text corpora as training data of adaptive algorithms.

When research activities by institutions and libraries are mis-captured and entered into RIS, and changes and updates occur daily in the dataset, then here is the solution, a permanent and controlled process cycle (data analysis, data cleansing and data monitoring) and continuous quality management to detect and correct errors. The subject of data quality should therefore be treated as a high priority business process to guarantee and increase the value proposition of the research information produced, especially in the new environment of open science, and to reduce the efforts of the staff and to save the costs of the institutions.

Acknowledgements

This work has been funded by the German Center for Higher Education Research and Science Studies (DZHW) and by the German Federal Ministry of Education and Research (BMBF) in the context of the project “Helpdesk to facilitate implementation of the Research Core Dataset” (<http://kerndatensatz-forschung.de/>).

References

- Azeroual, O. and Abuosba, M. (2017). Improving the data quality in the research information systems. *International Journal of Computer Science and Information Security*, 15(11): 82-86, November 2017.
- Azeroual, O., Saake G. and Abuosba, M. (2018a). Data quality measures and data cleansing for research information systems. *Journal of Digital Information Management*, 16(1): 12-21, February 2018.
- Azeroual, O., Saake G. and Schallehn, E. (2018b). Analyzing data quality issues in research information systems via data profiling. *International Journal of Information Management*, volume 41, pages 50-56, April 2018. DOI: <https://doi.org/10.1016/j.ijinfomgt.2018.02.007>
- Azeroual, O., Saake G. and Wastl, J. (2018c). Data measurement in research information systems: metrics for the evaluation of data quality. *Scientometrics*, volume 115, pages 1271-1290, April 2018. DOI: <https://doi.org/10.1007/s11192-018-2735-5>
- Azeroual, O., Saake G. and Abuosba, M. (2018d). Investigations of concept development to improve data quality in research information systems (Untersuchungen zur Konzeptentwicklung für eine Verbesserung der Datenqualität in Forschungsinformationssystemen). *Proceedings of the 30th GI-Workshop on Foundations of Databases (Grundlagen von Datenbanken)*, volume 2126, pages 29-34, CEUR-WS, May 22-25, 2018, Wuppertal, Germany.

- Bhanuse, S. S., Kamble, S. D., Kakde, S. M. (2016). Text mining using metadata for generation of side information. *Procedia Computer Science* 78, 807-814.
- Biesenbender, S., Hornbostel, S. (2016). The research core dataset for the German science system: Developing standards for an integrated management of research information. *Scientometrics*, volume 108, pages 401–412, March 2016. DOI: <https://doi.org/10.1007/s11192-016-1909-2>
- Brennan, N. (2018). CRIS systems as key e-infrastructure elements to support open science implementation within the European research area. In *CRIS2018: 14th International Conference on Current Research Information Systems*, Umeå, June 13-16, 2018. <https://dspacecris.eurocris.org/handle/11366/695>
- Carstensen, K.-U., Ebert, C., Jekat, S., Klabunde, R. and Langer, H. (2010). *Computerlinguistik und Sprachtechnologie: Eine Einführung*. Heidelberg: Spektrum Akademischer Verlag, Springer.
- de Castro, P. (2018). The role of current research information systems (CRIS) in supporting open science implementation: the case of strathclyde. *ITLib* (2). <https://dspacecris.eurocris.org/handle/11366/691>
- Chopra, A., Prashar, A. and Sain, C. (2013). Natural Language Processing. *International Journal of Technology Enhancements and Emerging Engineering Research*, Vol 1, Issue 4, pages 131-134, November 2013.
- Clements, A. and J. Proven (2015). The emerging role of institutional CRIS in facilitating open scholarship. In *LIBER Annual Conference 2015*, London, June 25th, 2015. <https://dspacecris.eurocris.org/handle/11366/393>
- European Commission (2017). *Evaluation of research careers fully acknowledging open science practices*. Directorate-General for Research and Innovation Open Science and ERA, Brussels. https://ec.europa.eu/research/openscience/pdf/os_rewards_wgreport_final.pdf
- Heyer, G., Quasthoff, U. and Wittig, T. (2008). *Text Mining: Wissensrohstoff Text*, 1. Auflage, W3L-Verlag, Herdecke.
- Liao, S.-H., Chu, P.-H. and Hsiao, P.-Y. (2012). Data mining techniques and applications - A decade review from 2000 to 2011. *Expert Systems with Application*, 39(12): 11303-11311. DOI: <https://doi.org/10.1016/j.eswa.2012.02.063>
- MESRI (2018). *National plan for open science*. Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, Paris.

- Schöpfel, J. (2015). The SAFARI syndrome. Implementing CRIS and open science. In *euroCRIS Membership Meeting at AMUE*, Paris, 11-12 May 2015. <https://www.slideshare.net/Schopfel/safari-syndrome-48709700>
- Seth, S., Rüping, S., Wrobel, S. (2009). Metadata extraction using text mining. In: Solomonides, T., Hofman-Apitius, M., Freudigmann, M., Semler, S. C., Legré, Y., Kratz, M. (Eds.), *Healthgrid Research, Innovation and Business Case*. IOS Press, Amsterdam, pp. 95-104.
- Simons, E. (2017). Open science: the crucial importance of metadata. In *euroCRIS Strategic Membership Meeting Autumn 2017*, CVTI SR, Bratislava, Slovakia, Nov 20-22, 2017. <https://dspacecris.eurocris.org/handle/11366/627>
- Tatum, C. (2017). What is the evaluative object of open science? In *22nd Nordic Workshop on Bibliometrics and Research Policy*, Helsinki, 09 November 2017 https://figshare.com/articles/What_is_the_evaluative_objectofOpenScience/5624728
- Upshall, M. (2014). Text mining: Using search to provide solutions. In *Business Information Review*, 2014, Vol. 31, Nr. 2: 91–99.

A biographical note on the authors

M.Sc. Otmane Azeroual is a researcher at the German Institute for Higher Education Research and Science Studies (DZHW) in Berlin, Germany. After studying Business Information Systems at the University of Applied Sciences Berlin (HTW), he began his Ph.D. in Computer Science at the Institute for Technical and Business Information Systems (ITI), Database and Software Engineering Group of the Otto-von-Guericke-University Magdeburg, Germany and at the Department of Computer Science and Engineering of the University of Applied Sciences (HTW) Berlin, Germany. His research interest is in the area of database systems, information systems, data quality management, business intelligence, big data, information security, cloud data management, open data, project management and industry 4.0.

Prof. Dr. rer. nat. habil. Gunter Saake is a full professor of Computer Science. He is the head of the Databases and Software Engineering Group at the Otto-von-Guericke-University Magdeburg, Germany. His research interests include database integration, tailor-made data management, database management on new hardware, and feature-oriented software product lines.

Prof. Dr.-Ing. Mohammad Abuosba is a full professor of Department of Computer Science and Engineering at the University of Applied Sciences (HTW) Berlin, Germany. His research areas are engineering, IT systems engineering (focus on database systems, product data management), modeling, quality management and project management.

Dr. Joachim Schöpfel is a senior lecturer in Library and Information Sciences at the University of Lille (France), researcher at the GERiiCO laboratory and consultant at the Ourouk consulting firm. He was Manager of the INIST (CNRS) scientific library from 1999 to 2008, head of the University of Lille department of information sciences from 2009 to 2012 and director of the French Atelier National de Reproduction des Thèses (ANRT) from 2012 to 2017. He teaches library marketing, auditing, valorization and digitization of cultural heritage collections, intellectual property and information science. His research interests are scientific information and communication, especially open access and open science, research data and grey literature. He is member of euroCRIS, of the NDLTD board of directors and of Knowledge Exch.

The role of the objects identification systems in ensuring permanent access to the digital content of the digital repositories: an analytical study

Le rôle des systèmes d'identification des objets pour l'accès permanent au contenu numérique des dépôts numériques : une étude analytique.

Boumarafi Behdja et Mettai Khaled NTIDRN, Constantine - Algérie

Résumé

L'environnement numérique a généré de gigantesques changements dans les aspects de l'accès, l'organisation, l'identification des objets numériques, depuis les années 1990s, et le développement du web sémantique et les improvisations continues sur le web. Ceci entraîne des difficultés concernant la récupération des données par les utilisateurs, à cause de l'instabilité des URLs et les liens, c'est pourquoi, les identifiants électroniques sont devenus une question importante dans l'infrastructure numérique pour les managers des dépôts ainsi que la société de recherche, le besoin d'archiver et préserver ce contenu numérique dans le but de donner l'accès à la société académique et bénéficier de ce contenu riche, et renforcer la possibilité de réutilisation des données pour différentes raisons. Pour réaliser cette étude nous avons analysé les différents systèmes d'identification dont le but commun est de garantir l'accès permanent au contenu des résultats de recherche dans l'environnement numérique à long terme. Et reporter les bonnes pratiques à suivre pour adopter un bon système d'identification durable qui garantira un accès permanent et à long terme aux objets numériques des dépôts institutionnels et du web.

Mots-clés. identifiant numérique, système, liens numériques, accès permanent, dépôts institutionnels

Abstract

The Digital environment has been undergoing big changes with regard to access, organization, identification of digital objects since the 1990s. With the advent of the semantic web and the ongoing improvements of the web, as a result, digital resources can change and possibly cannot be retrieved easily by the users' community due to the instability of URLs and the links. Therefore, the persistent identifiers become a crucial issue in the digital infrastructure for the digital repositories managers as well as for the research community. The need to archive and preserve such digital content to allow the users' community to benefit from this rich content and strengthen the possibility of data reuse for different purposes. This study is undertaken to analyze the different identification systems that have a common goal which is the guarantee of permanent access to the content and research data in the digital environment in the long term. It also reports the best practices that can be followed to adopt the most sustainable system that ensure permanent and long-term access to the digital objects in institutional repositories and the web.

Keywords: digital identifiers, systems, digital repositories, links, permanent access.

1 Introduction

URLs were never intended to be what they've become: an arcane way for a user to identify a site on the Web. Unfortunately, we've never been able to standardize URLs, which would give us a more useful naming system. Arguing that the current URL system is sufficient is like praising the DOS command line, and stating that most people should simply learn to use command line syntax. The reason we have windowing systems is to make computers easier to use, and more widely used. The same thinking should lead us to a superior way of locating specific sites on the Web. (Dougherty, 1996).

The internet has provided many opportunities to link the digital resources and different types of non-textual resources to locate the digital objects in the web environment. The World Wide Web with its bibliographical identifiers has facilitated access to the content of the links over time. It is clear that in the linked data era, the available links on the web have been increasing, as many systems identifiers have been adopted, some of which are funded and supported by international organizations such as the digital identifier object (DOI) foundation. In effect, DOI is used by most of academic journals to archive and sustain the long term access to the scholarship in form of articles and other intellectual production. The increasing number of identification systems and greater diversity of resources we deal with are both contributing to identifier proliferation. On one hand, systems for identifying and managing digital objects are becoming essential for digital content providers as the growth of the digital content market accelerates, on the other hand, the digital repositories have a major role in supporting the publication of research outputs of universities and research centers. This tends to be very important in terms of the sustainability of object identifiers in different forms, where different objects can be linked by different identification systems. The emerging of new identification systems are seen as an alternative to the URLs that appeared to be inconsistent in guaranteeing the long term preservation and long term access to scholarship.

2 Research Methodology

This research is conducted by adopting an analytic theoretical approach aiming to review the previous literature on the topic of systems identifiers to provide a framework for the present study. It attempts to give an overview about the digital objects identification, its different components and its role in ensuring the permanent access to the resources for the long term. In the process, it aims to answer the following questions:

- 1) What are the types of identification systems used to link the digital objects in the web environment?
- 2) What is the most used identification system in the digital repositories to sustain the persistency of the object links?
- 3) What identification systems most used by digital repositories and international organizations of the research communities?
- 4) Is the DOI the most identification system adopted by repositories?
- 5) Is the interoperability an important factor in the identifiers system environment?
- 6) Are there any other identifiers that meet and keep abreast with the current developments of the digital repositories environment ?

3 Literature review

The literature review presented in this section will trace previous studies on objects identification systems to identify the different identifiers that can be used in the digital environment as an essential part of the e-research infrastructure and the digital repositories in order to provide an overview of the development of linking the identification of objects on the web.

Khedmatgozar et al.(2015), conducted an evaluative study with the aim to identify qualified digital identifier system in the digital environment. The authors used seven criteria used a comparative evaluation on common identifier systems. They used a comparative evaluation of common identifier systems to identify which one is most appropriate. They found that six identifiers, including, DOI, Handle, UCI, URN, ARK and PURL, are appropriate digital identifier systems. Results of their study showed that NBN, MARIAM and ISNI identification systems are suitable choices for digital identification in certain specialized fields for their benefits such as digital content chains, networks integration, digital right management, cross referencing, digital libraries and citation analysis. Such results of this study can help digital environment experts to diagnose digital identifier and their effective use in applied fields.

Hakala (2005) described the state of the bibliographic identifiers used in digital environment by using a model of seven layers to define the scope of the identifiers that exist within the digital library field. The study recommended that identifiers must be developed in order to cater for the growing needs of libraries and other organizations involved with creation and dissemination of publications. Also, traditional identifiers must be modified in order to make them more suitable to the needs of electronic publishing.

Park et al. (2011) examined the success factors in adopting the digital object identifier system by extending DeLone and McLean's information system success model. The authors provided an overview of digital object identifier systems by collecting data from 41 digital content providers. The

findings indicated that information quality is the strongest factor that influences organizational benefits through perceived usefulness and user satisfaction. The digital content provider can be a moderator in the successful adoption of digital object identifier systems.

Bellini et al (2010) presented the Italian initiative that aimed to implementing an NBN persistent identifiers (PI) infrastructure based on a novel hardware and software architecture. The researchers described a distributed and hierarchical approach for the management of an NBN namespace, illustrated assignment policies and identifier resolution strategies based on request forwarding mechanisms. Their testing approach raised technical issues concerning the guarantee of the uniqueness identifier. The authors recommended more investigation to establish permanent and reliable connections between NBNs and other persistent identifiers such as DOI, which would favor the implementation of a multi-standard global resolver.

Chandrakar (2005) discussed the digital object identifier system and provided an overview to describe the digital object identifier (DOI) system. The author described the implementation of the handle system by the Corporation for National Research Initiatives, where a handle is designed to provide an efficient, extensible, and secured global name to an intellectual object. He concluded that the management of intellectual objects in a digital environment such as the internet which is flooded with various kinds of objects like research articles, e-books, electronic theses and dissertations etc. requires the existence of persistent, reliable identifiers for each distinguishable piece of content and associated services activated by these identifiers to manage access and other digital rights. The DOI is the essential part of the electronic publishing especially for the management and the access of the concerned resources. The DOI system is the new technology developed for persistent identification and interoperable exchange of intellectual property on digital networks.

McMurry et al (2017) proposed a notion that identifier practitioners should take in the design of identifiers where provisioning, documentation, and referencing can address many of the identifier problems encountered in the life science data cycle—leading to more efficient and effective science. The study concluded that the practices must be both easy and rewarding to the groups that do the implementing. In the broader context of scholarly publishing, DOIs succeeded because they were well aligned with journals' business goals (tracking citations) and because the cost was worth it to them. However, object identifiers face different challenges than do referencing data identifiers. In addition the ecosystem-wide challenges that should be addressed.

4 Different identification systems in use by digital repositories

Persistent identifiers attempt to solve the problems of resource identification and long-term access to online digital materials. A persistent identifier allows the resource to be uniquely identified in a way that will not change if the resource is renamed or relocated, and will persist regardless of the protocol used to access it. This means that a resource can be reliably referenced for future access by humans and software.

4.1 Use identifiers systems

Identifier systems are widely used in the digital repositories to support its main goals in guarantying the long-term preservation and access to the materials as well as easy retrieval of the scholarship, by unique identifiers that remain unchanged over time. According to a paper published in the JISC (2008) using persistent identifiers infrastructure will provide for utilities such as :

4-1-1 Access to resources over time : persistent identifier infrastructure can support access to resources as they are moved from one repository or data store to another.

4-1-2 Discovery: persistent identifiers provide the link between the discovery of resources from federated repositories and access to those resources. The metadata returned from a discovery service can contain an identifier that is resolved to the location of the resource at the time of discovery.

4-1-3 Enhanced resource management : persistent identifiers can be used to model complex relationships between resources and make it possible to manage compound resources and distributed copies of a resource, to distinguish between different versions of a resource, and to help decide which is the most appropriate copy or version for a user.

4-1-4 Rights management : persistent identifiers can be used to associate information rights with resources. They can also be used to track the relationship between a resource and derived works.

With the proliferation of digital resources and the content in digital forms, and the release of some types of persistent identifiers such as ARK, DOI, Handle, ISBN, ISSN, PURL, URI, URL and URN. Which are the most used identifiers in the repositories environment.

4.2 ARK Standard (Archival Resource key)

The Archival Resource Key (ARK) system was developed by Kunze and Rodgers (2008) and is currently maintained by the California Digital Library. The canonical form of the identifier is a URI beginning with ark:/, followed by a five- or nine-digit number representing a Name Assigning Authority, followed by another slash and the object name. The name may be qualified by adding a component path (introduced and subdivided DCC State of the Art Report by slashes) indicating finer granularity and a

variant path (introduced and subdivided by dots) indicating a specific version of the resource. While the identifier itself refers to a specific resource, the identifier followed by '?' refers to metadata about that resource, and the identifier followed by '??' refers to a commitment statement from the current provider about curating the resource. An ARK is (currently) made resolvable by prepending it with `http://` and the host part of a Name Mapping Authority capable of resolving the identifiers of the Name Assigning Authority; this part is expected to vary over time, and mechanisms are proposed for recovering from a faulty Name Mapping Authority specification.

4.3 Handle system

The Handle System was developed by the Corporation for National Research Initiatives with funding from the (US) Defense Advanced Research Projects Agency, and has been running since 1994. The form of the Handle identifier is a code representing a naming authority (usually a pair of numbers separated by a dot) followed by a slash, followed by a locally assigned identifier; the official way of representing this as a URI is to prefix it by 'hdl:'. The system operates its own system of resolvers, meaning it is possible to retrieve information about an identified resource independent of any other Internet application, but an HTTP proxy server is provided which will redirect URLs of the form `http://hdl.handle.net/hhandlei` to a URL in the Handle record for the resource. The proxy supports load balancing between several URLs.¹ Today it is among the most used identification system in the digital repositories.

¹Handle System Web site, URL: <http://handle.net/>

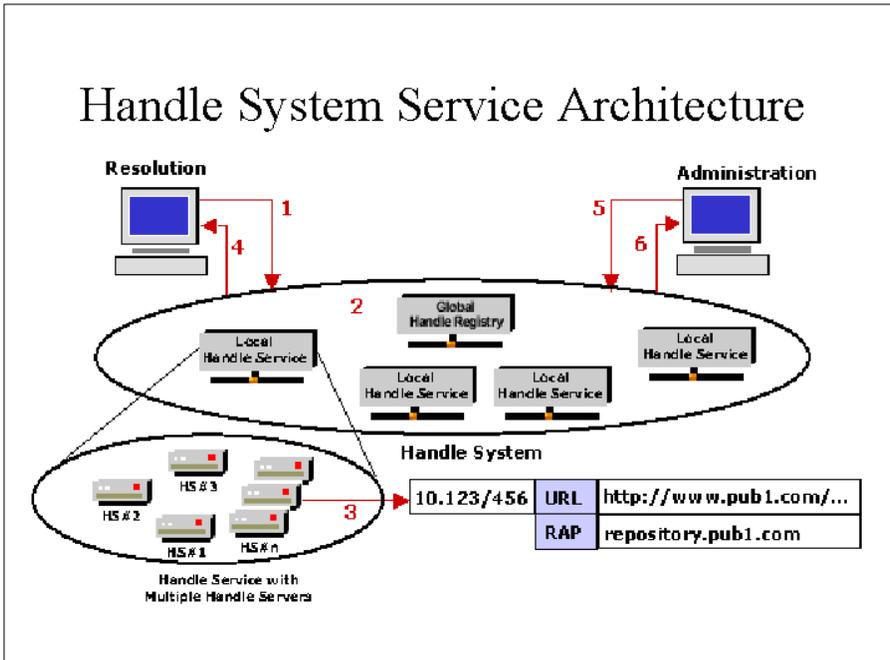


Figure 1: Handle System Service Architecture (<http://handle.net/>)

4.4 PURL System (Persistent Uniform Resource Locator)

PURL System is also among the used identifier systems by digital repositories. It is being adopted by many repositories, due to the diversity of identifier systems. Prasad and Guha (2005) mentioned that PURL is a general naming and resolution service developed and implemented by Online Computer Library center inc, and was developed as an interim service to resolve the problem of broken links or inconsistent links, and PURLs are Unique Identifiers assigned to digital resources and no two resources can be assigned the same PURL to avoid duplication. However, PURLs look like URLs, but instead of directing the client to the Resource, it goes to the PURL Resolver where from the Actual URL is supplied, which ultimately brings out the requested resource.

4.4.1 The components of PURL system

The PURL Looks like URL, it consists of three parts : protocol, resolver address, and user assigned name. The PURL resolver software is freely available, allowing other organizations to set up their own resolver.

- **Protocol:** to access the PURL Resolver.
- **Resolver Address:** the IP address or domain name of the PURL Resolver.
- **Name:** Name of Resource assigned by the user.

Figure 2 shows the position of those parts.

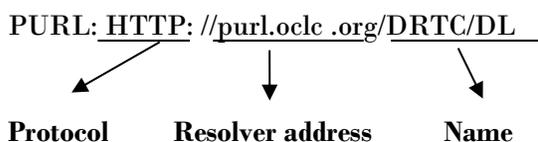


Figure 2: Source: URL: <https://drtc.isibang.ac.in>

Whenever the location of the resource changes; the new location needs to be updated in the PURL Resolver Database.

4.5 DOI System

The Digital Object Identifier (DOI) System supported and funded by DOI Foundation in an initiative that grew out of discussions on electronic publishing within the enabling Technologies Committee of the Association of American Publishers (AAP). The system was first seen at the February 1997 meeting of the Professional /Scholarly Publishing (PSP) division of AAP in Washington DC. Prior to this date, the AAP has identified a technology provider for the system, the Corporation for the National Research Initiatives (CNRI) of Reston, Virginia, and it was the handle system on which the DOI was demonstrated at the PSP Meeting (Simmonds, 1999). The DOI system tends to be a standard way to identify content at the first, and was intended to solve some problems and overcome some challenging including:

- The lack of copyright protection.
- Technologies were developed, but most vendors were not engaged in conversations with book publishers.
- Technologies were designed to facilitate electronic publishing, without understanding the specific and unique needs of book publishers.

4.5.1 *The international efforts to foster DOI Research infrastructure*

Since the system DOI was first adopted in 1997 since then, it received much international support and interest and has got the attention of many organizations and stakeholders to adopt it as new and sustainable identification system. Risher and Rosenblatt (1998) described the factors that led to the adoption of DOI system by a group of prestigious academic and professional organizations in an annual meeting of the American library Association (ALA) and the Association of Research Libraries (ARL) and the special meeting of the National Federation of Abstracting

and Indexing Services (NFAIS). Participants to that meeting recognized the need for the DOI system, to demonstrate the DOI system. Five publishers with existing electronic products were asked to provide access to their files so that the DOIs could be added to their products to facilitate retrieval, persistence overtime. Figure 3 shows the difference between DOIs and URLs. The first DOI registration agency began in 2000; by early 2009 around 40 million DOI names had been assigned through eight registration agencies. The most widely known application of the DOI system is the Crossref cross-publisher citation linking a service which allows a researcher link from a reference citation directly to the cited content on another publisher's platform, subject to the target publisher's access control practices. Other applications include government documentation, books, and data, **Simons(2012)** mentioned that in Australia, There were a number of research data projects in 2012 at the Australian institutions that related to the implementation of DOIs and a collaborative, open discussion between people working on some of these projects has begun. Two distinct approaches and subsequent workflows have emerged from these discussions. One approach is to develop an interface for researchers that allow them to mint or place a mediated request for a DOI name for their data collection. Another approach is to mint DOIs for data collections that have already been created and are maintained by research data administrators. This addressed the international interest of using DOI system that is internally-supported and to refer to the data different forms and the scientific production and their unique location on the long run.

4.5.2 The DOI as persistent identifier

The role of the DOI system as a persistent identifier is to assign an entity to the content like for example (10.1000/123). The DOI system provides a resolution from that name to a current URL. When the content, previously at URL xyz.com, is moved to a new URL newco.com, a single change in the DOI directory is made: all instances of the DOI name identifying that content (even if already recorded in print, as bookmarks, etc.) will resolve to the new URL, without the user having to take any action concerning the change. Therefore, the DOI name is persistent and remains unchanged as shown in figure3.

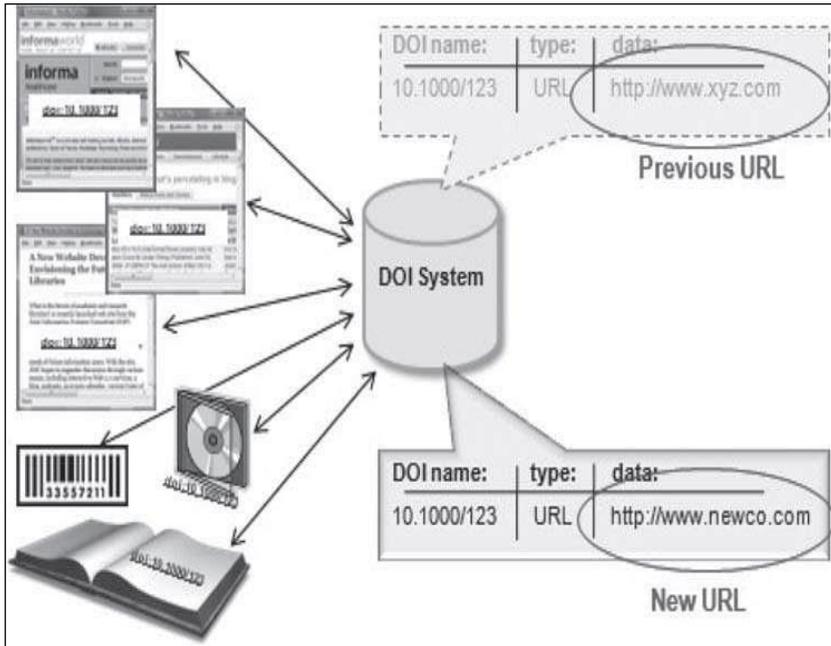


Figure 3: the persistency of DOI. Source (International DOI Foundation).

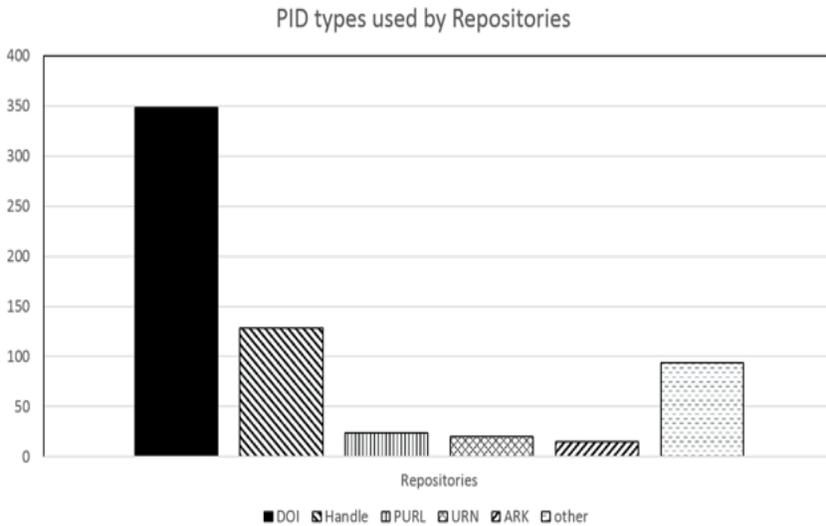


Figure 4: PID Types used by Repositories (Klump and Huber, 2017)

As can be seen from the figure 4, there are many types of persistent identifiers that can be used to give access to digital repositories, but the most persistent identifiers used by repositories are the DOI system, followed by the HANDLE system then URN. The others system are less used by repositories, this can be due to the wide usage of DOI system by

international and scientific organizations since the release of the system in 1997, and its acceptance in the research community as a sustainable identification system and for being adopted by the international academic journals and the important international publishers to ensure the different needs of the publishing industry. It is also due to the different advantages and features of the DOI identification system over the other persistent identifiers as well as its suitability and its many aspects to be the most trusted PI for the international research communities.

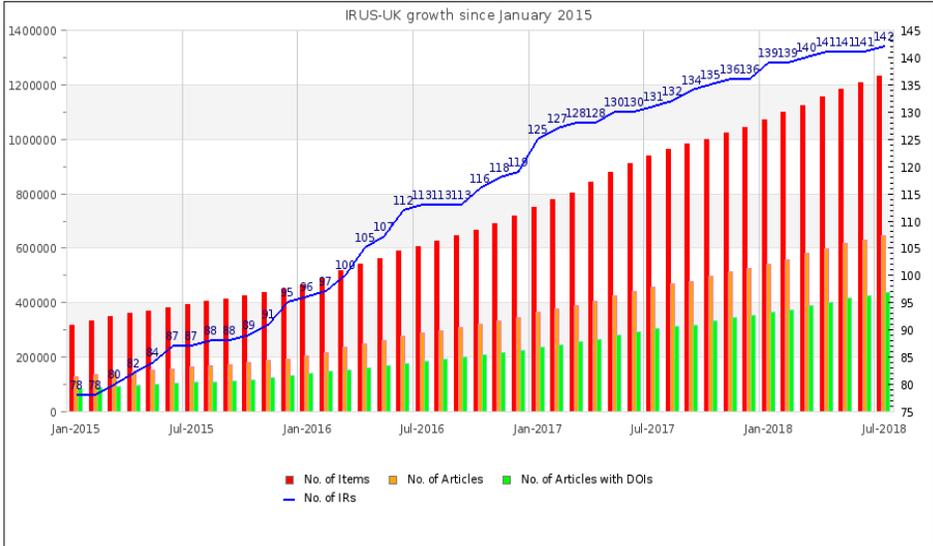


Figure 5: IRUS-UK Growth since January 2015. (Source, IRUS-UK)

As can be seen from the Figure 5, the usage of DOIs in ensuring the permanent and long-term access to the scholarly publications is on the rise since 2015 according to the institutional Repositories Usage Statistics In the UK (IRUS-UK) most of the academic journals have assigned to their articles DOIs. The adoption of DOIs system by the major publishers in the different academic fields has been established and became widely used as an efficient, stable, and persistent identifier system. With the increasing number of the institutional repositories that is the place where the scientific outputs produced by the research communities (faculties, researchers, students.etc) that must be preserved and remained over time to be reused any time. DOI system is more suitable for enhancing access to the digital materials in its different forms and since it has been funded and supported by IDF organization, it became the most used by repositories and the major academic journals in fields such as medicine, computer science,

etc due to its dynamic mechanisms in the process of easy retrieval and access to the materials over time.

4.6 RIDIR Project (Resourcing Identifier Interoperability for Repositories)

RDIR is a project funded under the auspices of the Joint Information Systems Committee 'Repositories and Preservation' program. It is investigating the requirements for, and benefits of, the clear use of PI in order to facilitate interoperability between digital repositories of different types. The RIDIR project's main objectives were to engage with the identifier and repository communities to understand their requirements, to build a fully working demonstrator, and to raise awareness of PI interoperability issues. The project was not about shared PI services themselves as such services would be available and focusing on the use of PI. As a demonstration, they have created a Lost Resource Finder service, which allows a user to be redirected to the new, correct location for a resource when a PI is broken. The new location can either be specified by a repository manager, or it can be based on crowd intelligence – based on other users of the service searching for and discovering the new location of the resource within the service.

Guaranteeing access to data over an indefinite period of time is problematic for research institutions, and the institutional repositories that adopted different persistent systems as their infrastructure to preserve and provide an easy and fast access to their scholarships, might be undergoing further development to integrate the new technologies in achieving fast and interoperable queries between the repositories to meet the search needs in the light of the ever-fast changing of information technologies. They need to make partnerships to help in the collaborative preservation and access by strong and secured persistent identifier systems to help in the Re-use and the citability of the research outputs of the universities through the digital repositories. In this case universities should keep up with the new trends and integrate their IRs with the new project to increase the interoperability between the repositories, the citability as well as the traceability of the resources with the objective of Re-use and long-term access. Nonetheless, there are ecosystem-wide challenges and many efforts should be undertaken to address these issues in the relevant initiatives.

5 Discussion

From the analysis of the different identification systems covered in this study, it can be seen that the lack of reliable links since the advent of the web in 1990 led to the emerging of alternate persistent identifiers that can guarantee the referencing and the locating the digital resources. The development and release of the identifier systems that is used by digital

repositories is allowing universities to preserve and provide access to the resources and the digital scholarship in general. The DOI systems was of great advantages to be adopted for setting up a strong persistent identifier infrastructure due to different advantages in comparison with other persistent identifiers as shown by previous research (Khedmatgozar et al.2015). The success of DOI in the usage by digital repositories over other persistent identifiers(Klump and Huber, 2017) supported the reasons of generating international funding and interest as a strong and sustainable persistent identifier. However, with the current changes and the interoperability issues might change more aspects of how effectively can the repository items interact, be discoverable and findable with other persistent systems and in case of PI broke, all those issues can be addressed by the repository managers to be aware of the changing in the digital environment, so that they can ensure the permanent access to the digital scholarship by the research community.

6 Conclusion

With the rapid growth of digital resources in the web environment, the digital identification systems and the different identifiers systems tend to strongly enhance their role in the visibility of the content preserved in the digital repositories. They can allow the Re-use of this rich content over time without encountering technical issues concerning the unique location and identifier of the item on the web. The institutional Repositories of universities over the world used different identifier systems to built and give permanent access to their intellectual and research outputs. The description of the different identification systems, including URL, URN, PURL, Handle, DOI, showed that each identifier has different mechanisms but all of them provide the digital the possibility of preserving resources for certain time. Furthermore, the URLs have proven according to many studies that they don't meet the requirements of persistency, and cannot be reliable to setting up to a sustainable persistent identifier; as a result, the DOI system was adopted to fix the technical issues of URLs. The international support for DOI made it the most persistent identifier due to its meeting the requirement of persistency and being used by the digital repositories through the implementation of Handle resolver. Many major publishers have adopted the DOI system because of the different advantages of the systems in different aspects regarding the long-term preservation and access, it can also meet the requirement of sustainability of scholarly content in the digital repositories as pillars in the new environment of scholarly communications.

References

- Aridas, T., Sidman, D. (2001). Digital object identifier: What is it and why is it so great?. *PrintMedia*, 15, (10), pp. 22- 36.
- Bellini, E., et al, (2010). Persistent identifier distributed system for digital libraries. *International Cataloguing and Bibliographic Control : Quarterly Bulletin of the IFLA UBCIM* 39, (2), pp. 30-36.
- Bellini, E., et al, (2012). Persistent Identifiers Interoperability Framework. Retrieved from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.474.1967&rep=rep1&type=pdf>
- Chandrakar, R. (2005). Digital object identifier system: an overview. *The Electronic library*. 24, (4), pp.445-452.
- DOI Foundation. Retrieved from: <http://www.doi.org/>
- Hakala, J. (2006). The seven levels of identification: An overview of the current state of identifying objects within digital libraries. *Electronic library and information systems*, 40, (4), pp. 361-371.
- Handle System. Retrieved from: URL: <http://handle.net/>
- JISC.(2008).persistent identifiers. Retrieved from: http://www.rsp.ac.uk/documents/briefing-papers/technical_persistentids.pdf
- Khedmatgozar, H.R. et al. (2015). Digital identifier systems: a comparative evaluation. *Iranian journal of information processing and management*, 30, (2), pp. 523-546.
- Klump, J., Huber, R. (2017).20 years of Persistent Identifiers – which systems are here to stay? *Data Science Journal*. 16 (9), pp. 1–7.
- Kunze, J. , Rodgers, R. P. C. (2008).The ARK identifier scheme. Internet Draft. Retrieved from the Internet Engineering Task Force Website: <http://tools.ietf.org/html/draft-kunze-ark-15>
- Lunghi, M., Cirinnà, C., Bellini, E.. (2013). Trust and persistence for internet resources. *JLIS.it*. 4, (1), pp. 375–390. Retrieved from:DOI: 10.4403/jlis.it-5494
- McMurry J.A., et al. (2017)Identifiers for the 21st century: How to design, provision, and reuse persistent identifiers to maximize utility and impact of life science data. *PLoS Biol*. 15 (6), pp.1-18.
- Park, S., et al(2011).Examining the success factors in the adoption of digital object identifier systems. *Electronic commerce research and applications*. 10, pp.626-636.
- Prasad, A.R.D., Guha, N. (2005). Persistent identifiers for digital resources. DRTC-HP international workshop on building digital libraries using DSpace, 7-11, March, Bangalore.
- PURL Community. Retrieved from: <http://www.purlz.org/>
- Resourcing Identifier Interoperability for Repositories (RIDIR) project. Retrieved from:<http://www.jisc.ac.uk/whatwedo/programmes/reppres/ridir.aspx>
- Simons, N. (2012). Implementing DOIs for research data, D-lib magazine. 18, (5/6).

Open Access dans les bibliothèques universitaires arabes : Penser autrement l'accès à la science dans les pays du Sud

Open Access in Arab University Libraries: Think differently about access to science in the South

Wahid Gdoura, Université de Manouba- Tunis

ISD- Labo SILAB, wahidgdoura@gmail.com

Résumé

Il s'agit de repenser le rôle des professionnels de l'information dans le mouvement du libre accès dans un contexte local et saisir l'impact d'un système d'information documentaire ouvert sur l'avenir des bibliothèques universitaires. Nous soulevons les questions de l'universalisme de la science et de la diversité des cultures, les écarts observés entre les pays du Sud et du Nord quant au processus de production des savoirs, de leur transfert et de leur mise gratuitement en ligne. Les bibliothécaires universitaires arabes sont peu engagés dans ce mouvement d'accès ouvert, et s'ils le font c'est pour perpétuer le déséquilibre entre l'offre éditoriale du Nord et du Sud. On s'est proposé de penser la médiation des connaissances et le libre accès en bibliothèques universitaires arabes autrement.

Mots-clés. Libre accès, Universalisme de la science, Diversité culturelle, Bibliothèque universitaire arabe.

Abstract

The aim of the following paper is to rethink the role of information professionals in the open access movement in a local context, and to understand the impact of an open information system on the future of university libraries. We raise the questions of the universalism of science and cultural diversity as well as the differences observed between the countries of the South and the North in the process of knowledge production, transfer and free online access to it. Arab librarians in university libraries seem not to be very engaged in this open access movement, and if they do so it is to perpetuate the imbalance between the editorial offer of the North and the South. The following paper proposes to think about the mediation of knowledge and open access in Arab university libraries differently.

Keywords. Open Access, Universalism of Science, Cultural Diversity, Arab University Libraries.

1 Introduction

*« Pour voir le monde autrement,
il faut commencer par penser autrement »
Olivier Bowden, Ecrivain et historien britannique*

Les avantages du libre accès sont largement partagés par la communauté scientifique internationale. La littérature scientifique abonde de travaux sur les bénéfices d'Open Access (OA) aussi bien pour les chercheurs que pour les institutions et la société. Suber (2012) et Babini (2014). L'accès ouvert est utile parce qu'il assure un accès rapide et sans barrières de contenus numériques. Aussi, il augmente la visibilité de la recherche scientifique, il améliore la productivité des chercheurs et accroît le facteur d'impact. L'accès ouvert profite également aux institutions de l'état qui supervisent leurs financements dans la recherche, il leur fournit des indicateurs pour évaluer les publications et programmer les futurs projets de recherche. Enfin, il bénéficie à la société, grâce à la médiatisation des connaissances et favorise la transparence et la bonne gouvernance de la recherche.

De son côté, la littérature arabe reprend pratiquement ces avantages et souligne la prédisposition des chercheurs arabes à adopter OA. Ce paradigme leur offre une opportunité pour se libérer de leur isolement et leur permet de mieux confronter leurs idées avec la communauté scientifique internationale et enfin contribuer à la construction du savoir universel. Ce discours « d'intention » exprimé depuis plus de douze ans, date des premières enquêtes sur OA et le chercheur arabe Gdoura (2006) Ibnlkhayat(2007) Jaber(2016) n'a pas été « réellement » concrétisé sur le terrain.

Un autre acteur de la communication scientifique censé jouer un rôle de premier ordre dans cette dynamique : le bibliothécaire arabe semble aussi peu engagé dans ce mouvement de l'accès ouvert, avec une faible contribution à la création de dépôts numériques institutionnels et à la gestion des métadonnées et à l'utilisation des protocoles d'OA. Par contre, il est assez actif dans la sélection et la fourniture de ressources électroniques étrangères. Les professionnels de l'information arabes, en raison de la carence de la production scientifique locale mise en libre accès, ne font que sélectionner et diffuser une abondante offre éditoriale libre en provenance des pays du Nord. Cette offre, en dépit de sa qualité certaine, ne répond pas nécessairement aux besoins de la recherche scientifique arabe.

La problématique de l'Open Access dans les bibliothèques universitaires arabes (BU) ne peut pas être abordée seulement sous l'angle de l'**accessibilité** aux ressources numériques ouvertes, indépendamment du système de recherche dans ces pays et des questions d'usage et d'utilité de la production scientifique internationale. Cette question revient au centre des débats Sud-Nord et les interrogations touchent à des enjeux locaux divergents: dans les pays du Nord l'accès ouvert est perçu comme « *une stratégie politique visant à promouvoir l'innovation et stimuler la croissance* » dans une économie de connaissance Diallerts(2012) alors que dans les pays

du Sud on se préoccupe du flux unilatéral de l'information scientifique qui aggrave la fracture cognitive et accentue la dépendance à l'économie des grandes puissances.

D'où l'objet du présent travail qui étudie le libre accès dans les bibliothèques universitaires du Sud, les obstacles à l'adhésion des bibliothécaires au paradigme OA et se propose de penser la médiation des connaissances et l'accès ouvert en BU arabes autrement.

A partir de l'analyse de la littérature bibliothéconomique et de la synthèse des Journées d'études sur le libre accès en Tunisie (JELA de 2013 à 2018) nous nous proposons de repenser le rôle du bibliothécaire universitaire dans le mouvement du libre accès dans un contexte local et saisir l'impact d'un système d'information documentaire ouvert sur l'avenir des BU arabes. La première partie du travail étudie l'OA sous l'angle des rapports entre l'universalisme de la science et la diversité des contextes de recherche. Quant à la deuxième partie, elle s'intéresse, tout d'abord, à la contribution des BU du Nord au développement d'OA avant de se pencher sur les actions que peuvent mener les BU arabes pour soutenir ce mouvement, redynamiser la médiation de l'information et réintégrer l'espace universitaire.

2 Open Access entre l'universalisme de la science et la diversité des contextes de recherche

La science moderne prétend à l'universalisme puisqu'elle est régie par les jugements de la raison, la quête de la vérité et l'objectivité, on parle de « science idéale ». Toutefois, la science, est aussi une « production sociale des connaissances », elle est produite et renouvelée en rapport avec l'état des connaissances d'une époque et d'une culture donnée. Pathy avance que la science « est aussi prise dans un contexte et axiologique, et repose sur des valeurs admises par la culture dans laquelle elle s'insère » Pathy(1999). Et l'on se demande, à l'ère de la mondialisation envahissante comment la science peut soutenir l'universalisme si on ne tient pas compte de la diversité culturelle et des différents contextes de recherche scientifique Bienaymé(2007). Nous posons cette question, par ce que le développement du nouveau modèle de communication scientifique est tributaire des conditions de création et de diffusion des connaissances.

3 L'enseignement supérieur et la recherche scientifique : universalisme ou uniformisation sous l'effet de la globalisation

Les systèmes de l'enseignement supérieur et de la recherche scientifique (ESRS) ont subi de profonds changements depuis plus de deux décennies sous l'effet de la stratégie de la globalisation répondant à une variété d'intérêts socio-économiques et politiques.

Selon des experts de l'Unesco Altbach(2009) « par mondialisation, nous entendons la réalité déterminée par une économie mondiale de plus en plus intégrée, les nouvelles technologies de l'information et de la communication, l'émergence d'un réseau international du savoir, le rôle de

la langue anglaise ainsi que d'autres forces indépendantes de la volonté des établissements universitaires ».

Il y a une divergence de vues sur les jugements dont la globalisation fait l'objet. Deux positions s'y opposent: pour certains « *elle constitue l'expression de la modernité et du progrès de l'humanité avec l'accélération des flux de communication et l'augmentation de leur volume*, Carpentier(2012), alors que pour d'autres, la globalisation « *n'est rien d'autre que la face invisible du néolibéralisme économique, phase actuelle du développement du capitalisme* » Carpentier(2012).

Sur le plan universitaire, la globalisation tend à « uniformiser » les structures et le fonctionnement des systèmes actuels de l'ESRS à travers la définition des modalités d'accès à l'université, de programmes, des modes de gouvernance, d'autonomie, et surtout par l'adoption de normes de qualité. Dans ce cas, on ne peut plus parler d'une éducation comparée, ni de diversité culturelle et les comparaisons internationales portent sur la capacité des systèmes d'ESRS à s'appliquer à ces normes. Carpentier souligne qu'*'il ne s'agit plus de comparer la diversité mais de classer selon un système unique de normes élaboré par le centre* » Carpentier(2012).

Quant à la recherche, la globalisation est perçue comme un accès à la science moderne produite dans les pays du Nord selon les priorités et les besoins de leurs sociétés Piron(2016). La coopération internationale tend à faire participer les chercheurs du Sud à trouver des solutions aux problèmes propres au Nord. A titre d'exemple, les chercheurs maghrébins impliqués dans des programmes de coopération scientifique avec l'Union européenne ont contribué aux projets de recherche européens en tant qu'associés et non porteurs de projets. De même les doctorants maghrébins ont servi de « main d'œuvre de connaissance » pour les laboratoires de recherche européens.

Il s'agit d'un alignement sur la recherche européenne qui pourrait, s'il persiste, conduire à une « aliénation scientifique ». On peut avancer que l'universalisme de la science dans bien des cas pourrait mener à une uniformisation scientifique beaucoup plus qu'à une science plurielle.

Voyons l'état des systèmes de recherche dans les pays arabes et les obstacles à une communication scientifique « équilibrée » entre les communautés de chercheur du Nord et du Sud.

4 Les systèmes de recherche et d'innovation dans le monde arabe

Les sociétés arabes connaissent des problèmes de développement après une période de décolonisation et d'édification de systèmes d'enseignement supérieur et de recherche scientifique. Consacrés au départ à la formation de cadres pour répondre aux besoins de la société, ces systèmes se lancent dans un travail de recherche universitaire, pour former dans un premier temps des docteurs avant de concevoir des programmes de recherche-développement (R-D) en vue d'un partenariat avec leurs environnements socio-économiques.

Toutefois les enseignants universitaires ne consacrent que 5% de leur temps de travail à la recherche contre 33% aux USA et au Japon (Alecso, 2008). Ils s'engagent dans des activités pédagogiques pour affronter la massification de l'enseignement supérieur. Et lorsqu'ils s'adonnent à la recherche c'est surtout pour obtenir des promotions de grades beaucoup

plus que pour faire avancer les programmes R-D. Cette recherche est demeurée peu visible et mal diffusée.

5 L'édition scientifique arabe en difficulté

5.1 Edition et diffusion de la recherche arabe

Quelques traits caractéristiques du modèle d'édition scientifique :

- Secteur public : l'édition scientifique est prise en charge essentiellement par le secteur public. Les universités publiques ont créé leurs propres structures d'édition pour publier les travaux de leurs chercheurs. Le secteur privé est peu intéressé de la publication académique en raison des coûts élevés et l'absence de rentabilité économique.
- Secteur fragile et en difficulté : on peut énoncer les traits suivants:
 - Rareté des travaux de recherche
 - Nombre réduit d'auteurs productifs
 - Lenteur de l'édition : problèmes liés à l'arbitrage, le montage, le financement, etc.
 - Diffusion restreinte : faibles circuits de distribution et diffusion de publications universitaires à l'échelle nationale et internationale.
 - Revues académiques irrégulières, peu indexées et répertoriées dans les grandes bases de données internationales.

En raison de ces difficultés, beaucoup de chercheurs, surtout en sciences fondamentales et appliquées, préfèrent publier dans des revues internationales.

Les chercheurs arabes préfèrent publier dans des revues internationales à facteur d'impact dans l'objectif est d'obtenir une meilleure visibilité et participer au débat scientifique et aux échanges avec la communauté internationale. Or, ils sont soumis à des contraintes au niveau du choix des sujets à traiter qui ne répondent pas forcément aux priorités locales et du choix de la langue de rédaction scientifique (anglais et français).

Guédon souligne que les revues internationales « *dictent le choix* » de sujets à traiter et parle d'une « *opposition trompeuse* » entre une science internationale et une science locale : « *Poursuivre dans un pays pauvre des questions scientifiques de peu de pertinence pour ce pays ou sa région peut conduire au paradoxe suivant : des laboratoires coûteux ,et donc rares, et des chercheurs, tout aussi rares, peuvent être amenés à contribuer à des questions importantes pour la science émanant des pays riches, mais de peu d'importance pour le pays en question*» Guédon(2014).

En dépit de ces difficultés, les universités arabes ont réalisé quelques performances: Selon le dernier classement des universités de Shanghai (2018) figurent 13 universités arabes d'Egypte, d'Arabie Saoudite, de Tunisie et de Liban. Deux universités tunisiennes apparaissent pour la première fois : Université Tunis Manar et Université de Sfax sont classées dans la position 801-900 (Shanghai Ranking, 2018).

5.2 Le libre accès dans les universités arabes

L'implication des chercheurs et éditeurs arabes dans le mouvement du libre accès est faible. Quelques chiffres illustrent ce peu d'engagement :

Pour les revues académiques arabes en OA :

- d'après DOAJ (2015) on recense 611 revues arabes soit 5.82% de la totalité de périodiques signalés dans ce répertoire. Ces revues sont inégalement réparties entre es pays arabes : Egypte : 533 revues (87.23%) en raison de la présence d'un grand éditeur égyptien : HindawiPublishing Corporation qui publient des revues internationales et non arabes, mais « soupçonné d'être prédateur ». Quant aux autres pays arabes, ils publient 78 revues en OA(12.76%).
- Répartition revues libre par langue : anglais (95%) bilingues (3%)
- Répartition par spécialité : médecine (41%) sciences fondamentales (29%) sciences appliquées(23%) sciences humaines et sociales (6%)Gdoura(2015).

5.3 Pour les Archives ouvertes arabes

- Nombre d'archives ouvertes dans le monde arabe selon DOAR et ROAR : 63 dépôts BenRomdhane(2016)
- Cependant, les pratiques d'auto archivage dans des dépôts numériques étrangers s'améliorent, souvent par obligations internationales. Les chercheurs maghrébins par exemple, associés à des projets de recherche européens, sont tenus de déposer leurs publications en OA (recommandation de l'Union européenne du 17 juillet 2012).

5.4 L'impact Factor, objet de controverse

Ce modèle d'évaluation quantitative des citations de recherche, développé depuis 1961 par l'Institute for Scientific Information de Philadelphie (ISI) est objet de critiques, on lui reproche quelques insuffisances:

- **modalités de calcul du facteur d'impact** : fondées sur une période de deux ans qui ne convient pas à toutes les disciplines. A cela s'ajoutent d'autres éléments : la périodicité, le changement de titre de la revue, l'indice d'actualité, le seuil de déclin, l'indexation défailante des auteurs collectifsBarthez Durand(2016).
- **facteurs humains** : en rapport avec l'évaluation des articles et les dérives des comités de lecture : effet de langue étrangère (anglaise), appartenance géographique, etc. L'expertise en «double aveugle » risque de se transformer en évaluation à « double vitesse » en défaveur des scientifiques du Sud.

Pour contourner ce modèle d'évaluation, des communautés scientifiques ont proposé des alternatives pour faire des calculs : (Barthez Durand,2016)

- Index européen des citations/ European Citation Index in Humanities développé par l'Européen Scientific Foundation (2003)
- Modèle CiteBase du dépôt Arxiv.org pour les sciences fondamentales et appliquées

- Google Scholar présente en 2004 un autre index jugé insuffisant.
- Enfin le dernier né dans les pays du Sud : facteur d'impact arabe (2015), ce nouveau modèle de calcul a été proposé par une équipe d'universitaires arabes. L'égyptien Mahmoud Abdel-Ati en explique les raisons Abdel-Ati(2017) :
 - *Absence d'une instance scientifique arabe unifiée qui prend en charge le classement et l'indexation des revues en langue arabe, selon un facteur d'impact, notamment en sciences humaines et sociales (SHS).*
 - *Renforcer la position de la langue arabe comme outil de communication scientifique. Beaucoup de travaux en SHS relatifs aux sociétés arabes sont destinés beaucoup plus aux lecteurs arabes qu'aux lecteurs étrangers.*
 - *Regrouper des chercheurs autour de thématique de recherche locale et régionale, certains thèmes ne sont pas acceptés par les revues anglaises à haut Impact Factor, ce qui a poussé plusieurs scientifiques arabes, cherchant un rayonnement international, à ne pas se préoccuper de problèmes propres à leurs pays.*
 - *Réponse à la politique d'exclusion exercée par Thomson Reuters sur les revues académiques de langue arabe.*
 - *La plateforme Impact factor arabe applique des outils méthodologiques internationaux pour classer ces périodiques. Son premier rapport date de 2015 et classe 167 revues arabes à Arab Impact Factor.*

Autres solutions proposées pour les pays du Sud mais par des chercheurs du Nord :

- Guédon(2014) suggère la création de territoires autonomes autour de formes d'évaluation autonomes à l'exemple d'une plate-forme de publications scientifiques panarabe.
- Piron (2016) mène déjà une action en collaboration avec plusieurs intervenants: projet SOHA (Science ouverte en Haïti et Afrique francophone) : création de plate-forme de revues africaines et haïtiennes en libre accès, lancement d'une structure d'édition « Editions sciences et biens communs »ESBC, création d'un MOOC d'initiation à la recherche, etc.

5.5 Conclusion

Le blocage du mouvement de libre accès dans les pays du Sud et plus précisément dans les pays arabes, réside dans le fait que la globalisation de la science s'est substituée à l'internationalisation de la recherche, sans se soucier de développer ou mettre à niveau les environnements de recherche du Sud et promouvoir les productions scientifiques locales, et du coup le flux de l'IST a pris une direction unilatérale du Nord vers le Sud, Il nous semble important de repenser l'universalisme de la science dans le sens de la pluralité des cultures et des contextes de recherche dans différentes régions du monde. Guédon parle de territoires autonomes : « *L'internationalisation réussie de la science passera par une mosaïque de territoires scientifiques*

définis par une autonomie relative en ce qui concerne l'orientation des recherches et les méthodes d'évaluation appliquées aux chercheurs. » Guédon(2016).

6 Open Access dans les bibliothèques universitaires arabes: la médiation des connaissances face au déséquilibre de l'offre éditoriale

En plus des défaillances des systèmes de recherche et d'innovation dans les pays arabes, nous comptons aborder les obstacles liés à la mise en œuvre de systèmes d'information documentaires ouverts. Les bibliothécaires universitaires arabes se rendent compte, de plus en plus, qu'ils sont entraînés de perpétuer ce « déséquilibre » entre l'offre éditoriale du Nord et du Sud. Comment, donc, penser la médiation des connaissances et l'accès gratuit au savoir autrement ?

7 Adhésion des BU américaines et européennes au paradigme OA

Les bibliothécaires contribuent, avec les autres acteurs de la communication scientifique aux transformations du modèle de l'édition savante : d'un modèle de publication onéreux et fermé à un modèle libre et ouvert. La problématique de l'OA s'est élargie pour englober les données de la recherche, c'est la science ouverte.

Rappel : Les professionnels de l'information dans les bibliothèques d'universités américaines et européennes se sont impliqués dans ce débat autour du libre accès depuis l'expansion du numérique. Inquiétés par la montée inflationniste des abonnements de revues scientifiques et l'émergence d'oligopoles d'éditeurs commerciaux, les bibliothécaires se sont regroupés pour se concerter sur les mesures à prendre face à la hausse des prix. Ainsi en 1989, on voit paraître un Bulletin d'information: Newsletter on Serials Pricing Issues (NOSPI) Mahe(2011) qui a constitué une tribune aussi bien pour les bibliothèques que pour les éditeurs et les agences d'abonnements pour réfléchir sur les mesures à prendre pour réduire les prix.

De leur côté, les associations de bibliothèques y contribuent dans cette réflexion à l'exemple de ARL (Association of Research Libraries) d'Amérique du Nord qui remet en cause le modèle économique de l'édition scientifique et revendique un accès sans contraintes aux publications académiques. Elle constitue en 1997 le groupe SPARC (Scholarly Publishing Coalition) pour encourager la publication de revues alternatives. SPARC a essaimé en Europe (SPARC Europe) et au Japon (SPARC Japan). La France a lancé le consortium COUPERIN.

7.1 Textes fondateurs

Conscients des enjeux de OA, les bibliothécaires se sont ralliés aux chercheurs pour développer un discours de soutien / Advocacy à ce sujet. Ils signent des déclarations en faveur de libre accès à l'exemple de l'Initiative de Budapest (2002) puis le texte de l'IFLA en 2002 (IFLA *Internet Manifesto pour la liberté d'accès à l'information*) puis en 2004 (IFLA *Statement on Open Access to Scholarly Literature and Research*

Documentation). Cette Fédération a constitué une section OA pour développer les échanges et le retour d'expérience sur l'accès ouvert.

7.2 Outils

En plus du discours de promotion d'OA les bibliothécaires participent à la production d'outils de travail et de signalement et de valorisation de ce paradigme. On cite :

- Open Archive Initiative (OAI, 1991) et le protocole de partage des métadonnées OAI-PMH.
- Moteur de recherche OAIster : créé en 2002 par les bibliothèques de l'Université de Michigan puis géré à partir de 2009 par OCLC.
- Répertoire Directory of Open Access Journal (DOAJ) lancé en 2003 par la bibliothèque de l'Université Lund de Suède.

De même, ils soutiennent les premières initiatives de création d'archives ouvertes ARxiv, de création de logiciels open source : E-Prints (S.Harnad de l'université Southampton). Ils développent la plateforme DSpace (bibliothèque de l'université de Massachussets Institute of Technology MIT, 2002) et la California Digital Library (université de Californie 1997) et la plateforme HAL (Hyper Articles en Ligne) du CNRS de France.

7.3 Rôle des bibliothèques universitaires

Voyons le champ d'intervention des bibliothécaires universitaires dans ce mouvement :

- Création et gestion de dépôts numériques institutionnels.
- Gestion des budgets consacrés au paiement des frais de publication en Gold Open Access à savoir les APC (Articles Processing Charges). Plusieurs universités et laboratoires de recherche confient l'administration des APC aux bibliothèques universitaires qui sont amenées à transformer une partie du budget des abonnements en budget de financement des APC.
- Gestion des métadonnées des dépôts institutionnels en utilisant le protocole OAI-PMH pour assurer l'interopérabilité des fichiers et faciliter la recherche de ressources en OA.
- Evaluation des ressources électroniques libres : Altmetrics
- Recherche de ressources et veille : le bibliothécaire effectue une sélection de ressources libres, en fonction des besoins des usagers, à travers la recherche d'articles pertinents, le signalement de revues en OA et des archives ouvertes.

Le libre accès est devenu le moteur du changement des bibliothèques de recherche. Pour s'engager pleinement dans ce mouvement OA les BU ont mis en place des structures de diffusion et de gestion de contenus numériques libres.

L'université Harvard crée le Office for Scholarly Communication et gère Digital Access to Scholarship at Harvard (DASH).

L'université de Liège lance le portail PoPuPs pour la diffusion des revues libres de l'Université de Liège et le dépôt de thèse BICTEL et d'articles ORBI

A travers cette littérature on peut dire que la bibliothèque se réinvente pour adhérer à l'ère de l'Open Knowledge, de l'Open Science. Son rôle et ses fonctions se renouvellent avec la réduction de la pression économique sur les budgets d'acquisition et son repositionnement dans le système de recherche et d'innovation. Le rôle du bibliothécaire comme médiateur se redéfinit et se renforce dans l'environnement de recherche. Une nouvelle vision de la bibliothèque de recherche se dessine actuellement. «*La gestion des données de la recherche, le mouvement de l'OA puis la promotion active de la science ouverte ont été déterminants pour les activités des bibliothèques* » (Swiatek, 2017). Les actions des organisations professionnelles tendent vers la promotion des bibliothèques en s'engageant dans la voie de la science ouverte.

« Pour LIBER, les années 2018-2022 doivent permettre aux bibliothèques de renforcer leurs récents services et leurs compétences numériques afin de projeter leurs actions sur la durée, de soutenir le développement à venir de la science ouverte, et de pérenniser leur positionnement central dans la politique européenne de l'IST. » (Swiatek, 2017).

8 Le libre accès dans les bibliothèques universitaires arabes

Quand les BU arabes « s'effacent » : une mission menacée :

Les BU ont toujours pour mission fondamentale la transmission des connaissances et la défense des valeurs sociales de l'information : égalité d'accès au savoir pour tous et liberté d'expression et de circulation de tous les courants de pensée. Elles apportent leurs concours pour l'excellence universitaire. Toutefois, les BU arabes sont traversées actuellement par de profondes mutations et voient leur mission menacée. Les facteurs qui influent sur ce changement se résument comme suit :

- Le numérique bouscule le lieu physique de la bibliothèque, et les collections et les ressources d'information transitent vers l'espace virtuel.
- Les usagers abandonnent la bibliothèque au profit des réseaux numériques et des moteurs de recherche.
- Evolution des relations des étudiants et chercheurs à la connaissance. On parle de lecture numérique « fragmentée », de rédaction collective de contenus numériques.
- Les conditions d'apprentissage et les temps académiques sont fortement influencés par le système LMD (licence, master, doctorat) qui ne prévoit beaucoup de temps pour le travail personnel au sein de la bibliothèque.
- Difficultés matérielles, restrictions budgétaires et insuffisances de ressources humaines dans les BU.

Pour affronter ces difficultés les bibliothécaires universitaires arabes sont appelés à repenser la BU dans cette phase de transition numérique, en se penchant sur certaines notions ; on se limite ici à deux notions : le lieu et les services. L'espace physique de la bibliothèque continue à subsister étant sa position symbolique comme lieu de savoir et de mémoire. Il ne se consacre plus à une pure fonction documentaire, puisqu'une partie de ses collections

sont accessibles à distance, mais surtout à une fonction communicationnelle et sociale. On emprunte la notion de troisième lieu du professeur américain de sociologie urbaine Ray Oldenburgh pour réétudier ces fonctions Servet(2010)., *Third Place* est utilisée initialement pour désigner un lieu pour la vie sociale est entendu ici pour désigner la bibliothèque comme lieu de rencontres intellectuelles, d'échanges d'idées et de détente. En d'autres termes, la BU est un lieu polyvalent : lieu de rayonnement de livres, de lectures, de débats, de conférences, d'expositions, de restauration et de loisirs. La notion de services doit être aussi revisitée, pour indiquer la transition numérique : on passe de la conservation à la transmission et du dépôt de livres et revues au flux de ressources numériques.

Enfin, on peut dire que la BU ne sera plus un lieu fermé de méditation mais plutôt un lieu ouvert, de partage et de libre accès à l'IST. Le public a toujours besoin du professionnel de l'information pour l'accompagner dans la recherche, la sélection et l'analyse de ressources électroniques. Il le sollicite pour consolider sa maîtrise de l'information (Information Literacy). On peut soutenir une tendance à la « ré-intermédiation » dans les bibliothèques. Cependant, le rôle du professionnel arabe ne se limite pas à communiquer les connaissances, payantes ou gratuites, produites dans les pays riches, mais aussi à chercher, valoriser et transmettre en OA les connaissances produites dans les pays arabes.

9 Penser autrement le libre accès dans les BU arabes: réajustement de l'offre éditoriale libre

Dans la continuité de notre réflexion sur l'universalisme de la science et la prise en considération de la diversité des cultures, des contextes de recherche, des thématiques scientifiques prioritaires et des méthodes d'évaluation pour les pays du Sud, nous pensons qu'un libre accès « équilibré » ne peut se réaliser sans un engagement fort des BU. Il est question de s'interroger sur le rôle que pourraient jouer les BU arabes dans la transmission libre de la production scientifique locale permettant sa diffusion et sa visibilité et sa valorisation.

- **Conduire des campagnes de sensibilisation en faveur du libre accès:**

Elles s'adressent aussi bien pour les chercheurs que pour les bibliothécaires. Il s'agit de faire la promotion de ce paradigme et d'expliquer son utilité et ses enjeux: remédier aux défaillances de l'édition universitaire locale et rendre les travaux de recherche nationale accessibles librement pour tous.

- Toutefois, l'implication des bibliothécaires et chercheurs dans ce mouvement depuis plus d'une douzaine d'années n'a pas donné de résultats tangibles. La Déclaration de Riyadh sur OA en 2006 est restée « lettre morte », les séminaires et les actions d'universités et d'associations professionnelles sur ce sujet (AFLI) ne sont pas venus à bout de la réticence des acteurs de la communication scientifique. Les universitaires commencent assez modestement, à intégrer le mouvement du libre accès. Leur engagement ne pourrait se confirmer que sur instruction des autorités publiques et des bailleurs de fonds. La dernière Journée d'études sur le libre accès à Tunis (JELA 10avril

2018) a été consacrée aux questions portant sur la mise en œuvre d'une politique nationale sur l'accès ouvert. L'autorité de tutelle, en l'occurrence le ministère de l'ESRS de Tunisie a exprimé sa prédisposition à mettre en place un dispositif juridique pour encourager les chercheurs à adhérer à ce projet.

- **Création et gestion de dépôts numériques institutionnels:**

Cette fonction est indiquée pour les bibliothécaires universitaires arabes comme il est le cas pour beaucoup de leurs collègues étrangers. L'objectif est de rassembler tous les travaux de recherche en arabe et en langues étrangères de tous les chercheurs de l'université ou de l'institution universitaire et de les diffuser dans des archives ouvertes. Ainsi, ces ressources peu ou mal communiquées auparavant seront rendues accessibles librement. Le public universitaire arabe puisera aussi bien dans des dépôts numériques nationaux qu'internationaux rendant la recherche de l'information plus exhaustive. C'est une étape essentielle pour « équilibrer » relativement le volume de connaissances produites dans divers environnements de recherche.

- **Gestion des métadonnées et conservation des ressources :**

Pour élaborer certains outils de travail arabes (OCR arabes, protocoles..) il est important de développer la coopération entre BU arabes, échanger des expériences ou rédiger des manuels de bonnes pratiques « Best Practices ».

- **Accompagnement des chercheurs dans ce mouvement de libre accès :**

Certaines actions sont possibles : sensibilisation aux enjeux de OA, formation à l'archivage des travaux de recherche, assistance juridique pour expliquer les droits d'auteurs (licences Creative Commons). Mais ces actions nécessitent de nouvelles compétences chez les professionnels de l'information.

- **Compétences et formation des bibliothécaires :**

Pour adhérer au paradigme d'OA et remplir de nouvelles fonctions, il est important de former le personnel en exercice et assurer une formation initiale pour les futurs professionnels de l'information. On parle d'un nouveau métier : « bibliothécaire Open Access » et de nouvelles compétences liées à l'accès ouvert et à la science ouverte, Dehaye(2017).

- **Gestion des APC (Articles Processing Charges) : le modèle auteur-payeur en débat :**

Le libre accès ne veut pas dire toujours gratuité. Les APC constituent l'un des modèles économiques de la voie dorée soutenu par les multinationales de l'édition numérique puisqu'on continue à les payer d'une autre manière. On change tout simplement les modalités de paiement des publications : de lecteur payeur à auteur-payeur. Ce modèle perpétue l'inégalité d'accès à l'information scientifique. Hachani juge que cette approche est pénalisante pour les auteurs du Sud.Hachani(2016). Il présente quelques chiffres : 28% des revues libres recensées dans DOAJ (2013) imposent des tarifs surtout les

périodiques en sciences « dures ». Les tarifs varient de 700 USD à 7000USD selon la notoriété de la revue. Ces prix sont excessivement élevés pour les chercheurs du Sud qui se trouvent souvent privés de publier dans des revues à haut facteur d'impact.

Quant aux BU arabes, elles seront confrontées à de nouveaux défis : elles continuent à supporter les prix des revues payantes (papier et électronique), et viennent s'ajouter les APC, sans parler des dépenses de renouvellement des équipements informatiques et des frais de formation du personnel.

Quelles sont les solutions à envisager ?

La voie verte : on ne fait que réitérer l'intérêt des dépôts institutionnels ou thématiques dans l'accès ouvert : centralisation et accessibilité aux publications scientifiques

- Programmes pour un OA en faveur des chercheurs du Sud : ce sont des initiatives lancées par des universités américaines, des éditeurs, et des collectifs d'universitaires pour permettre aux chercheurs des pays en développement d'accéder librement aux revues scientifiques : on cite The Chicago Emerging Nations Initiative, the Researchfor Life, JSTOR, etc.Hachani (2016).

10 Conclusion générale

Le paradigme du libre accès a été conçu dans le cadre de systèmes de recherche innovants et compétitifs dans les pays du Nord pour répondre aux besoins de l'économie du savoir. L'adhésion à ce paradigme dans les pays du Sud a rencontré des obstacles

qui ne sont pas liés seulement à un problème de visibilité et d'accessibilité aux contenus libres en ligne. Ceci nous a conduit à examiner les contraintes qui pèsent sur les systèmes de la recherche scientifique dans les pays en développement et les conditions de médiation des connaissances dans les bibliothèques universitaires en introduisant la notion de l'offre éditoriale.

Open Access ne peut être conçu dans les pays arabes sans une mise à niveau des systèmes de recherche et d'innovation pour surmonter la carence en publications scientifiques. Le rôle des BU arabes consiste à assurer un certain équilibre dans la fourniture de ressources de l'information scientifique produites aussi bien à l'échelle locale qu'internationale. A travers un engagement fort dans la voie verte et la création et gestion de dépôts numériques de travaux de recherche arabe, les bibliothécaires universitaires pourraient contribuer à réduire le flux unilatéral de l'IST et valoriser la production académique locale.

Bibliographie

- Abdel-Ati, M. (2017) *Arab Impact Factor une plateforme pour éditer des travaux de recherche* (en arabe). <http://arabimpactfactor.com/> (consulté le 13/7/2018).
- Accroître la diffusion des connaissances par l'augmentation de l'accès aux productions scientifiques des chercheurs.* Rapport. Université du Québec (2015). http://www.quebec.ca/communications/documents/Rapportlibre_acces2.pdf (consulté le 28/04/2018)
- Alecso, 2008. *La stratégie arabe pour le développement de la créativité dans l'enseignement supérieur.* Tunis : Alecso, 2008, p.38(Livre en arabe).
- Altbach, P (et al.) 2009 *Evolutions de l'enseignement supérieur au niveau mondial.* www.unesco.org
- Babini, D (2014).*Le libre accès à l'information scientifique et académique*<https://core.ac.uk/download/pdf/85142222.pdf> (consulté le 13/7/2018)
- Barthelemy, A (al.) 2018. Open access en bibliothèque universitaire : de nouveaux enjeux de médiations. *Revue française des sciences de l'information et de la communication*
- Ben Romdhane, M (2016). Les archives ouvertes dans le monde arabe : entre stagnation et évolution. *Revue maghrébine de documentation et d'information*, n°25, pp117-135
- Bienaymé A. (2007) Universalisme de la science économique et diversité culturelle.-*Diogene*, n°129. <https://www.cairn.info/revue-diogene-2007-3-page-58.htm> (consulté le 30/5/2018)
- Carpentier, Claude (2012) L'éducation face au défi de la globalisation : entre local et global. *Carrefours de l'éducation*, n°34, pp.7-13
- Chartron, G et Schopfel, J.(2017) . Open access et open science en débat. *Revue française des sciences de l'information et de la communication*, n°11 <https://journals.openedition.org/rfsic/3331> (consulté le 30/5/2018)
- Chartron, G., (2012). La valeur des services documentaires en prise avec le numérique. *Bulletin des Bibliothèques de France* 57 (5), 14-18. <http://bbf.enssib.fr/consulter/bbf-2012-05-0014-003> (consulté le 24/6/2018)
- Delhayé, M (2017). Bibliothécaire Open Access. In : *Bulletin des bibliothèques de France*, decembre, pp.20-21
- Dillaerts, H. (2012) Libre accès à la communication scientifique et contexte français : prospective, développement et enjeux pour la créativité et l'interdisciplinarité. *Thèse de doctorat en Sciences de l'information et de la communication.* Conservatoire national des arts et métiers. Paris
<http://www.theses.fr/2012CNAM0838>(consulté le 30/5/2018)
- Durand-Barthez, M(2016).- *Historique et critique du facteur d'impact*<http://hal.archives-ouvertes.fr/hal-01356185> (consulté le 20/6/2018)

- Gayoso, E (2017).- Synthèse du rapport « Monitoring the transition to Open Access Universities UK » .- in *Bulletin CNAM Comité de suivi de l'édition scientifique*, 2017
- Gdoura, W (2006). *La communication scientifique et l'accès libre à l'information scientifique : les chercheurs et bibliothèques universitaires arabes*. Tunis : Ed. Aleco. (Livre en arabe).
- Gdoura, W. (2015). L'édition scientifique dans le monde arabe : réalité et perspectives ; in *The second Saudi international conference on scientific publishing*, Riyadh, King Saud University 11-13 october 2015
- Guédon, J.C (2014). *Le libre accès et les territoires de la science : mondialisation ou internationalisation des sciences ?* Actes du colloque international Libre accès et recherche scientifique : vers de nouvelles valeurs. Tunis 27-28 novembre 2014/ sous dir. M. Ben Romdhane.- Tunis : Institut supérieur de documentation, 2014
- Hachani, S. Piron, F (2016). La pratique des frais demandés aux auteurs par les revues en libre accès : Une approche pénalisante pour les pays des Suds. *Revue maghrébine de documentation et d'information*, n°25, pp.215-237
- Harris, S. (2012) Moving towards an open access future: the role of academic libraries. A report on a roundtable commissioned by SAGE, in association with the British Library
<https://studysites.uk.sagepub.com/repository/binaries/pdf/Library-OARreport.pdf>(consulté le 24/4/2018).
- Ibn Lkhayat, N. (2007). Les universitaires marocains et le libre accès à l'information scientifique et technique: Où en est-on? In *Revue Maghrébine de documentation et d'Information*, n°17, 65-100.
- Jaber, J. (2016).-Are we now. A review of arabic studies in open access 2005-2015 (en arabe). *Revue maghrébine de documentation et d'information*, n°25, pp21-36
- Keller, A (2015). Library support for Open access journal publishing : a needs analysis.-*Insights*, n°28, november
<https://insights.uksg.org/articles/10.1629/uksg.256/> (consulté le 5/5/2018)
- Mahe, A (2011). Bibliothèques et archives ouvertes. *Bulletin des bibliothèques de France*, T56, n°1
- Morse. L (al.)(2013)*Libraries and their Role in Open Access: Challenges and Opportunities*. <https://fr.slideshare.net/frenaville/libraries-and-their-role-in-open-access-challenges-and-opportunities> (consulté le 24/4/2018)
- Ollendriff, C. (2015) Bibliothèques académiques et archives ouvertes : quels enjeux en France ? in : Cavalier, F., Poulain, M. (dir.) (2015) *Bibliothèques universitaires : nouveaux horizons*. Éditions du Cercle de la librairie, Paris. PP.207-218
- Paty, M. (1999). L'Universalisme de la science : une notion philosophique à l'épreuve de l'histoire. *Revue africaine de philosophie*, n°1, avril. pp.1-26.
<https://halshs.archives-ouvertes.fr/halshs-00167296/document> (consulté le 26/7/2018)

- Piron, F (et al.) (2016) Faire du libre accès un outil de justice cognitive et d'empowerment des universitaires des pays des Suds. *Revue maghrébine de documentation et d'information*, n°25, pp91-105.
- Salaün J.M.(2014) L'Open Access, quels enjeux pour les bibliothèques universitaires ? <https://archinfo24.hypotheses.org/2411> (consulté le 15/7/2018)
- Servet, Mathilde (2010), Les bibliothèques troisième lieu : une nouvelle génération d'établissements culturels, *Bulletin des Bibliothèques de France* n°4, pp.57-63. <http://bbf.enssib.fr/consulter/bbf-2010-04-0057-001> (consulté le 14/3/2018)
- Shanghai Ranking, 2018.- Academic Ranking World Universities. <http://www.shanghairanking.com/Shanghairanking-Subject-Rankings/> (consulté 3/8/2018)
- Suber, P.(2012).- Open access. Massachusetts : MIT Press. <http://www.acces-deschis.ro/documente-descriere/studii/76-open-access-peter-suber-pdf/file> (consulté le 15/4/2018)
- Swiatek, C. (2017) Les orientations Skills and Competencies au sein de Liber. *Bulletin des bibliothèques de France*, decembre, pp.23-31
- Vandegrift, M, Bolick, J. (2014). "Free to All": Library Publishing and the Challenge of Open Access. *Journal of Librarianship and Scholarly Communication* 2(4):eP1181. <http://dx.doi.org/10.7710/2162-3309.1181> (consulté le 30/4/2018).
- Xiang Yang Huang (al.) 2017. *Towards Open Science: China's Scientific Research and Libraries*. IFLA WLIC Wroclaw. www.ifla.org (consulté 25/3/2018)

La naissance du DICAMES, l'archive ouverte du Conseil africain et malgache pour l'enseignement supérieur : récit de la mise en place d'une innovation sociale et technologique en Afrique francophone subsaharienne

The birth of DICAMES, the open archive of the African and Malagasy Council for Higher Education: an account of the implementation of a social and technological innovation in sub-Saharan French-speaking Africa

Florence Piron, Université Laval, Zakari Liré, CAMES et Université Laval, Djossè Roméo Tessy, Université Laval, Thomas Hervé MboaNkoudou, Université Laval, Antonin Benoit Diouf, UNIDEP

Résumé

Le DICAMES est l'archive numérique institutionnelle du Conseil africain et malgache pour l'enseignement supérieur (CAMES). C'est une plateforme ouverte qui vise la conservation et la diffusion en libre accès de toute la production scientifique des pays d'Afrique francophone subsaharienne. Cet article raconte les principales étapes du processus de création du DICAMES en montrant les difficultés, les défis, mais aussi les solutions trouvées.

Mots clés. Archive ouverte, Afrique subsaharienne, libre accès, recherche-action, CAMES

Abstract

DICAMES is the institutional digital archive of the African and Malagasy Council for Higher Education (CAMES). It is an open platform that aims at the conservation and open access dissemination of all scientific production in sub-Saharan French-speaking African countries. This article describes the main steps in the process of creating the DICAMES, focusing on the difficulties, challenges and solutions found.

Keywords. Open Archive, sub-Saharan Africa, Open access, action research, CAME

11 Introduction

Le DICAMES est l'archive numérique institutionnelle du Conseil africain et malgache pour l'enseignement supérieur (CAMES). Située à l'adresse web <https://savoirs.cames.online/>, c'est une plateforme ouverte qui vise la conservation et la diffusion en libre accès de toute la production scientifique des institutions d'enseignement supérieur et de recherche des 19 pays d'Afrique subsaharienne qui sont membres du CAMES, presque tous francophones. L'originalité de cette archive créée avec le logiciel Dspace est qu'elle est à la fois centralisée et décentralisée. En effet, pilotée par le CAMES qui a assuré son installation, sa maintenance et sa gestion globale, cette infrastructure numérique offre à chaque institution de l'espace CAMES la possibilité de se doter d'un portail qu'elle pourra gérer de manière autonome et participative. Autrement dit, le DICAMES est une réalisation panafricaine qui mutualise des services technologiques, juridiques et stratégiques sous la houlette du CAMES pour mieux lutter contre l'invisibilité et la faible accessibilité numérique de la science produite dans ces institutions africaines et la conserver de manière pérenne, dans des conditions optimales.

L'envergure de ce projet a conduit le CAMES à développer une approche méthodologique collaborative innovante, basée sur le « faire ensemble » et mobilisant l'équipe de documentalistes et d'informaticiens du CAMES, une petite équipe de recherche de l'Université Laval, plusieurs experts internationaux bénévoles, des bibliothécaires « pionniers » et des informaticiens québécois et africains, tous invités à échanger sur les différents aspects du projet et à suivre sa progression dans des groupes de discussion sur WhatsApp et Facebook. Ces échanges ont permis de mettre au jour toutes sortes de difficultés et d'y répondre de manière collective, activant l'intelligence collective et la motivation du groupe de travail informel ainsi créé. Cette approche coopérative¹ semble plus féconde que l'approche « client » habituelle pour installer des sites web à l'aide de logiciels complexes comme Dspace.

Dans ce texte, nous retraçons les principales étapes du processus de création du DICAMES en mettant l'accent sur les difficultés, les défis, mais

¹ L'auteur et les auteurs de ce texte tiennent à remercier chaleureusement tous ceux et celles qui ont contribué à la naissance du DICAMES, incluant le Conseil de recherche en sciences humaines du Canada qui a subventionné le partenariat entre le CAMES et Florence Piron (Subvention 892-2017-2087), la Faculté des lettres et des sciences humaines de l'Université Laval qui a co-financé la participation de Florence Piron au colloque ICOA et d'autres partenaires qui ont contribué d'une manière ou d'une autre au rayonnement du projet. Nous dédions ce texte à la jeunesse africaine et à son courage.

aussi les solutions imaginées grâce à la solidarité et à l'agilité créative des membres de l'équipe de travail.

12 Premiers pas

Bien que basée sur plusieurs réflexions antérieures menées au sein de l'équipe du CAMES, en particulier par la chargée de communication Erika Djadjo, l'idée du DICAMES est née de manière très concrète lors d'une rencontre de 30 minutes organisée par Erika à Dakar en juin 2015. Étaient présents d'une part, Florence Piron, coresponsable de la première phase du projet de recherche-action SOHA² et Thomas Hervé MboanKoudou, son assistant de recherche camerounais, et, d'autre part, le professeur Claude Lishou de l'Université Cheikh Anta Diop de Dakar et expert responsable du virage numérique du CAMES, notamment du programme Silhouette. Ce programme lancé en 2013 avait introduit les technologies numériques dans le processus de constitution et de gestion des dossiers de promotion des enseignants-chercheurs et des enseignantes-chercheuses de l'espace CAMES. Dans le cadre de ce programme, le CAMES a reçu des milliers de thèses et d'articles en format numérique.

Bien au courant du manque chronique de documentation scientifique dont se plaignent les universitaires et les étudiants et étudiantes d'Afrique subsaharienne, nous avons eu immédiatement l'idée d'utiliser ces fichiers pour nourrir ce qui allait devenir le dépôt institutionnel du CAMES, le DICAMES. La deuxième idée concernait le logiciel à utiliser. Malgré les mérites du logiciel Invenio que nous présenta le lendemain Antonin Benoit Diouf, bibliothécaire en chef de l'UNIDEP, nous avons rapidement choisi le logiciel Dspace en raison de sa forte communauté francophone et aussi parce que c'était clairement le logiciel le plus utilisé par les dépôts institutionnels d'Afrique anglophone avec lesquels le DICAMES nous paraissait devoir être parfaitement interopérable. Une rapide vérification dans les répertoires internationaux des dépôts institutionnels (DOAR et ROAR) ayant montré le très petit nombre, à l'époque, de dépôts dans les universités d'Afrique francophone subsaharienne (sept), une troisième idée est vite apparue : que le DICAMES soit non seulement l'archive institutionnelle du CAMES, mais offre à toutes les institutions membres actives du CAMES un portail individualisé, une « Communauté » dans le jargon du logiciel Dspace, pour qu'elles y déposent leur production scientifique. Cette mutualisation éviterait que chacune soit obligée

²Ce projet visait à comprendre, pour mieux les dépasser, les obstacles au développement du libre accès et de la science ouverte en Afrique francophone subsaharienne et Haïti (Piron, Regulus et al. 2017). Son site est <http://projetsoha.org>.

d'investir seule dans la création d'un dépôt « individuel », à la mode occidentale, mais sans les ressources des universités du Nord.

La situation des dépôts en Afrique subsaharienne francophone n'a jamais été facile et s'est plutôt dégradée depuis 2015 (Diouf, 2010 ; Ezema, 2013 ; Irifeke et al., 2011 ; Kavulya, 2007 ; Schopfel et Soukouya, 2013 ; Piron, Diouf et al., 2017). En effet, en septembre 2018, il existait 3436 dépôts institutionnels recensés dans le répertoire Open DOAR, dont 155 en Afrique et seulement trois en Afrique francophone subsaharienne. Sur ces trois, deux n'étaient pas fonctionnels. L'existence du DICAMES, construit avec le logiciel libre DSpace comme 43% des dépôts institutionnels répertoriés, pouvait donc offrir d'un seul coup à toutes les universités d'Afrique francophone subsaharienne la possibilité de se doter d'un portail de type dépôt institutionnel sans avoir à en gérer les dimensions techniques, ce qui nous semblait constituer un premier pas vers une appropriation collective possible et la réussite du projet.

Dans l'enthousiasme, le professeur Lishou, avec l'autorisation du Secrétaire Général du CAMES, invita Florence Piron à faire une présentation de ce concept au Gabon le mois suivant, lors d'une réunion du Comité consultatif général (CCG) du CAMES dont sont membres statutairement tous les recteurs, rectrices, présidents et présidentes des 80 universités de l'espace CAMES. Florence a alors pré-enregistré une conférence sur support powerpoint et s'est tenue disponible pendant ses vacances, pendant que Thomas Hervé MboaNkoudou se rendait sur place à Libreville, pour répondre aux questions éventuelles. L'horaire de la réunion du CCG de Libreville n'a toutefois pas permis de diffuser la conférence.

Pendant les deux années suivantes, plusieurs efforts ont été faits de part et d'autre pour faire avancer l'idée. Des experts du Nord et du Sud ont été consultés, notamment Bernard Pochet de l'Université de Liège et Pierre Lasou de l'Université Laval. Thomas MboaNkoudou s'est initié au logiciel Dspace et a pris contact avec plusieurs acteurs de la communauté Dspace. Florence Piron a analysé des dépôts multi-institutionnels, notamment celui de la région de Washington DC, pour bien évaluer la robustesse du concept « multi-institutionnel » du DICAMES. Zakari Liré, responsable du centre d'information et de documentation du CAMES, a été désigné comme responsable du projet au CAMES. Une correspondance s'est installée, un peu timide, entre Thomas et les informaticiens du CAMES à propos du logiciel Dspace. Une des difficultés a été de bien faire comprendre à l'équipe

du CAMES les avantages du DICAMES par rapport à la bibliothèque numérique qu'ils étaient en train de créer.

Puis au printemps 2017, le rythme s'est accéléré. Le CAMES a officiellement invité Florence Piron à présenter le concept lors de la réunion du Comité consultatif général du CAMES le 17 juillet 2017 à Cotonou. En vue de bien préparer cette rencontre cruciale, Zakari Liré est venu faire un stage de deux semaines à l'Université Laval en juin 2017 (il a depuis commencé un doctorat). Ces deux semaines ont été décisives. Nous avons pu rédiger un premier jet des directives du DICAMES et nous avons exploré différents scénarios, notamment avec l'appui du responsable du dépôt institutionnel de l'Université Laval, Pierre Lasou.

La réunion de Cotonou a été un grand succès et a permis la validation officielle et cruciale du projet par le CCG du CAMES. En effet, à la suite de la présentation complète du projet dans ce qui fut la première version d'une longue série d'itérations (Piron 2017a), plusieurs recteurs au leadership incontesté ont tout de suite compris l'importance de cette archive ouverte capable de rendre disponible en libre accès la production scientifique africaine francophone, que ce soit au bénéfice du monde entier ou des étudiants et étudiantes de leur université. Ils ont noté plusieurs avantages, notamment la lutte contre le plagiat. En effet, contrairement aux idées reçues selon lesquelles le libre accès facilite le plagiat en simplifiant le « copier-coller » d'un fichier à un autre, la plupart des recteurs, rectrices, présidents et présidentes présents à Cotonou ont très bien compris que la mise en libre accès des thèses, mémoires et articles pouvaient permettre de les protéger de tout plagiat grâce aux logiciels anti-plagiat qui pourraient alors fonctionner. La discussion a aussi porté sur la décolonisation de la science africaine et sur le rôle qu'une archive comme le DICAMES, bien référencée dans le web scientifique, pourrait jouer pour la désenclaver et la mettre en lumière, ainsi que pour contribuer au renforcement professionnel des enseignants-chercheurs et enseignantes-chercheuses dans le contexte du passage au système LMD (licence, master, doctorat ; Charlier et Croché, 2010 ; Banque mondiale, 2014).

13 Un projet de recherche-action

Un partenariat a alors été noué entre Florence Piron, en tant que chercheuse sur la science ouverte en Afrique francophone subsaharienne, et le CAMES pour monter ensemble un projet de recherche-action visant la mise en œuvre du DICAMES et l'optimisation de son appropriation pendant sa première année d'existence. Ce projet intitulé « Optimisation de

l'appropriation d'une archive scientifique panafricaine ouverte » a été soumis au programme d'engagement partenarial du Conseil de recherches en sciences humaines du Canada (CRSH) en septembre 2017 et a bénéficié en novembre 2017 d'un financement de 24 980 dollars canadiens pour une année, jusqu'en novembre 2018.

Ce projet, qui relève des humanités numériques et des sciences de l'information et de la communication, répond à deux besoins identifiés par le CAMES : l'augmentation de la visibilité et de l'accessibilité sur le web des savoirs produits en Afrique francophone subsaharienne et la stimulation de la productivité scientifique des universités de l'espace CAMES afin de mieux répondre aux grands défis de cette région du monde. En particulier, le DICAMES est apparu comme une réponse concrète au régime postcolonial des savoirs marqué par la faible accessibilité de la science africaine francophone dans le monde, mais aussi d'une université africaine à une autre (Decung et Mukuku, 2016 ; Nkolo, 2016 ; Piron et al., 2017 ; Tessy, 2016). Cette situation non seulement nuit à la qualité des recensions des écrits et donc de la recherche qui se fait en Afrique (qui peut être répétitive ou mal informée), mais elle renforce l'extraversion de la science africaine vers les pays du Nord : exode des cerveaux, absence de recherche locale pertinente, assujettissement à l'agenda d'organismes du Nord (Hountondji, 2001 ; Mvé-Ondo, 2005 ; Piron, 2017c). Le concept d'injustice cognitive (Piron, MboaNkoudou, Pierre et al., 2016) issu du projet SOHA nous a semblé pouvoir rendre compte de cette situation qui empêche les chercheurs et les chercheuses d'Afrique de déployer le plein potentiel de leurs capacités de recherche au service du développement local durable de leur pays (Fredua-Kwarteng, 2015). L'existence d'une archive scientifique panafricaine ouverte pourra réduire l'extraversion de cette science dans une visée de justice cognitive et de stimulation de la recherche africaine, pour et par des Africains. La justice cognitive est un idéal épistémologique, éthique et politique visant l'éclosion de savoirs socialement pertinents partout sur la planète et non pas seulement dans les pays du Nord, au sein d'une science pratiquant un universalisme inclusif, ouvert à tous les savoirs (Piron, Regulus et DibounjeMadiba, 2016 ; Piron, MboaNkoudou, Regulus et al., 2016 ; Piron, Tessy et al. 2016 ; Piron, 2018 ; UNESCO, 2015).

La question de recherche-action de notre projet est la suivante : « Comment optimiser l'appropriation du DICAMES, l'archive scientifique ouverte créée et gérée par le Conseil africain et malgache pour l'enseignement supérieur, par les acteurs et les actrices du monde

universitaire d'Afrique francophone subsaharienne ? ». Nos objectifs initiaux étaient de créer l'archive, puis d'observer et d'analyser le processus d'appropriation du DICAMES par différents acteurs universitaires africains au moyen d'une enquête sur place et à distance. Ces résultats doivent nous permettre d'améliorer de manière continue l'ergonomie et les fonctionnalités de l'archive DICAMES afin d'en faciliter l'utilisation en recherche et en enseignement, par exemple l'intégration de ses fichiers dans les recensions des écrits et les plans de cours. Finalement, nous voulions produire des tutoriels et un manuel de formation sur l'archivage numérique dans le DICAMES afin d'accompagner les utilisateurs et utilisatrices dans leurs efforts d'appropriation de cette technologie. Sur ce point, l'appartenance de l'équipe de l'Université Laval au LabCMO, un laboratoire de recherche sur les usages du numérique, ainsi que nos réflexions critiques sur le développement international (Escobar, 1995) nous ont aidés collectivement à envisager dès le début les risques liés à un transfert technologique du nord vers le sud qui ne se soucierait pas de l'appropriation durable de cette technologie par ses utilisateurs – une expérience regrettable trop souvent vécue...

En effet, d'autres projets antérieurs de dépôt institutionnel ont montré le risque bien réel que ce bel objet technologique ne soit que peu ou pas utilisé par ceux et celles qui pourraient le faire fonctionner ou l'utiliser dans l'enseignement, la recherche ou l'apprentissage, sans compter bien sûr les difficultés reliées à l'accès au web que cette utilisation exige. Rappelons que cette archive est une véritable innovation sociale et technologique dans le monde universitaire d'Afrique francophone subsaharienne. Or, selon Gaglio (2011), le point central du processus d'innovation est l'appropriation. Pour Breton et Proulx (2002), l'appropriation renvoie « aux utilisations particulières qu'un individu ou un groupe peut faire d'un bien, d'un instrument, d'un objet » et ne peut être dite optimale que lorsque trois conditions sont remplies par l'utilisateur : 1) il ou elle démontre un minimum de maîtrise technique et cognitive de l'objet technique; 2) cette maîtrise doit s'intégrer de manière significative et créative aux pratiques de l'utilisateur ; 3) l'appropriation ouvre vers des possibilités de réinvention ou même de participation directe des usagers à la conception de l'innovation, sans pour autant surestimer l'autonomie de l'utilisateur (Jauréguiberry et Proulx, 2012). Le modèle proposé par Bar et al. (2016) ajoute l'idée que le processus d'appropriation n'est pas linéaire, mais cyclique, reposant sur une négociation constante de sens entre usagers et concepteurs. Notre projet a donc continuellement été mené avec le souci de cette appropriation par différents groupes d'acteurs, dans différents contextes et sur différentes versions du dispositif, afin de pouvoir comparer les effets de différentes

conditions d'appropriation. Par exemple, nous ajoutons des tutoriels pour répondre aux questions des utilisateurs et utilisatrices, même s'ils n'étaient pas prévus au départ.

La méthode prévue de collecte des données était mixte, à la fois qualitative et évaluative. Au plan qualitatif, elle proposait de traiter comme des données les commentaires et réactions partagés spontanément par les utilisateurs de l'archive dans des messages courriel ou WhatsApp adressés à l'équipe. À cela devaient s'ajouter des audits d'interface ainsi que des tests utilisateurs mis en place afin de déceler et de corriger d'éventuelles aberrations ergonomiques ou des difficultés dans l'utilisation. Les résultats de cette enquête feront l'objet d'un autre article.

14 Un projet du CAMES

Contrairement à d'autres projets similaires menés ailleurs ou antérieurement, le DICAMES est un projet panafricain porté par le CAMES sans le support d'une institution partenaire du Nord. En effet, Florence Piron, bien qu'employée de l'Université Laval, a contribué au projet au nom de son engagement personnel en faveur de la justice cognitive et dans le cadre de ses recherches, sans obtenir de mandat ou de support particulier de la part de son université (à part son salaire de professeure). Cet ancrage africain du DICAMES est en lui-même un phénomène libérateur source de fierté et d'empowerment collectif et a suscité un enthousiasme proportionnel chez tous ceux et celles qui y ont contribué, notamment les assistants de recherche béninois et camerounais de Florence Piron et les bibliothécaires des universités pilotes. Il se peut aussi que cette situation ait généré quelques craintes au CAMES, bien vite disparues.

Du point de vue du CAMES, la création du DICAMES répond à trois de ses missions statutaires :

- promouvoir et favoriser la compréhension et la solidarité entre les États membres ;
- instaurer une coopération culturelle et scientifique permanente entre les États membres ;
- rassembler et diffuser tous documents universitaires ou de recherche : thèses, statistiques, informations sur les examens, annuaires, annales, palmarès, information sur les offres et demandes d'emploi de toutes origines

Le caractère multi-institutionnel du DICAMES, rendu possible par la technologie utilisée (Dspace), répond parfaitement aux deux premières missions car, pour que l'archive fonctionne durablement, la coopération et la bonne coordination à long terme de toutes les institutions africaines partenaires sera essentielle. Cette archive facilitera bien sûr l'accès de tous à la documentation et à l'information scientifique et technique pertinente sur l'Afrique et contribuera ainsi à améliorer la qualité de l'enseignement et de la recherche dans l'espace CAMES, une priorité du Plan stratégique du CAMES sans cesse défendue par son Secrétaire général, le professeur Bertrand Mbatchi. Selon la vision du CAMES exprimée publiquement lors des cérémonies du Cinquantenaire du CAMES en juin 2018, le DICAMES permettra :

- une meilleure utilisation de la production scientifique de l'espace CAMES dans l'enseignement universitaire, dans les administrations ou dans le système scolaire, d'où une amélioration de la formation générale et à la recherche, basée sur une recherche documentaire plus efficace et puissante
- une valorisation en Afrique de l'activité de recherche grâce à son rayonnement mondial sur le web, alors que la science est un élément important du développement durable
- un avancement plus rapide et efficace des connaissances produites dans les universités africaines grâce à un meilleur partage des résultats obtenus
- une meilleure lutte contre le plagiat
- un appui à l'instauration des programmes LMD en facilitant l'accès à une documentation scientifique de qualité par tous les étudiants et étudiantes de master et de doctorat et par leurs enseignants et enseignantes.

Le CAMES s'est aussi engagé à adopter et publiciser une politique explicite en faveur du libre accès aux travaux scientifiques africains afin de renforcer l'adhésion de ses institutions membres au DICAMES. Cette volonté s'exprime notamment dans le contenu final de la Lettre d'entente entre chaque institution d'enseignement supérieur et le CAMES en vue du déploiement des collections du DICAMES, ainsi que dans les directives du

DICAMES qui mentionnent la possibilité d'utiliser des licences Creative Commons³.

Toutefois, au-delà de cet engagement fort et constant, de petits pépins ont parfois ralenti l'avancée du DICAMES et notamment la rédaction finale du contenu de la Lettre d'entente et son envoi aux institutions partenaires qui n'a eu lieu qu'en novembre 2018. Mentionnons par exemple une mauvaise interprétation de la mission du DICAMES de la part de certains membres du conseil des ministres dont relève le CAMES, des craintes relatives au coût du projet – en fait bien moindre que ce qui était imaginé – et à son acceptabilité dans le monde universitaire africain (où le projet a rencontré au bout du compte plus d'enthousiasme que de réticences), un débat récurrent sur la qualité à exiger des documents scientifiques qui seront archivés dans le DICAMES et surtout une incompréhension de l'idée même d'archive scientifique ouverte par certains acteurs du monde universitaire de l'espace CAMES pour lesquels le web est soit un espace frivole incompatible avec la recherche scientifique, soit un monde inquiétant et menaçant, soit un outil inaccessible en raison de ses coûts et de la fracture numérique. L'appropriation du projet au sein de l'équipe interne du CAMES a nécessité plusieurs rencontres et démonstrations puisqu'il ne s'agit ni d'un pur projet de bibliothèque, ni d'un pur projet informatique, mais d'un projet global alliant informatique, sciences de la communication et sciences de l'information.

15 L'année 2018

Elle a été décisive dans la création du DICAMES.

Commençons par la création du site en janvier 2018, juste à temps pour sa présentation publique au Comité consultatif général du CAMES de janvier 2018 à Niamey, par Florence Piron. Le constat de cette avancée majeure par rapport à la présentation de juillet 2018 à Cotonou a enthousiasmé les dirigeants et dirigeantes d'université présents aux deux rencontres et a permis à certains d'entre eux de mieux comprendre encore le projet.

En parallèle du site du DICAMES, nous avons créé en avril 2018 le site « Centre de ressources du DICAMES » conformément au plan de travail de notre projet de recherche. Outre les directives du DICAMES, ce site propose des outils de découverte et d'appropriation du DICAMES pour les

³Dans le présent texte, nous ne pouvons pas analyser ces Directives qui sont publiques et disponibles sur le Centre de ressources du DICAMES à <http://dicames.scienceafrique.org>. Elles feront l'objet d'un autre article.

bibliothécaires qui seront responsables des dépôts pendant les premières années, ainsi que pour les candidats aux CTS qui devront prendre l'habitude d'y déposer les fichiers numériques de leur dossier. Il comporte aussi une section FAQ et une section consacrée aux dépose-thons, un concept que nous avons imaginé pour rassembler des bénévoles et les faire déposer en groupe des fichiers issus de la bibliothèque numérique du CAMES.

En 2018, nous avons aussi effectué un essentiel travail de terrain, allant à la rencontre des utilisateurs et utilisatrices du DICAMES (bibliothécaires, étudiants et étudiantes, enseignants et enseignantes, dirigeants et dirigeantes d'université) pour leur expliquer le projet sous toutes ses coutures et recueillir leurs réactions. Ces missions ont été cruciales pour l'optimisation du site et le déploiement du projet, ayant permis le recrutement de 10 universités pilotes dès le mois de janvier 2018. Sur le plan de la recherche, ces missions ont bien montré la pertinence de réfléchir aux usages d'une technologie avec ses utilisateurs et de ne pas se limiter à l'implantation clé en main d'un dispositif technologique.

Ainsi, Florence Piron, avec le soutien de Zakari Liré, a eu l'occasion de rencontrer plusieurs recteurs et présidents d'université au Bénin, au Burkina Faso, au Cameroun et en Côte d'Ivoire entre mars et juin 2018, notamment lors des activités lors du Cinquantenaire du CAMES à Ouagadougou, et de noter leurs réactions et inquiétudes face au projet, ce qui a permis d'enrichir et de compléter les directives du DICAMES. Pendant ce travail de terrain, Florence a aussi mobilisé plusieurs bibliothécaires et doctorant-e-s qui se sont engagés à trouver, numériser et déposer dans le DICAMES un maximum de documents scientifiques de leur université : nous les avons surnommés les « bénévoles du DICAMES ». Il reste le problème des moyens de leur engagement, notamment l'accès à un bureau, au web et à des scanners capables de numériser la production imprimée, thèses et revues... Nous avons effectué une mini-enquête sur ces scanners et les sources de financement possibles, mais il est clair que les universités participantes doivent s'engager sur ce plan.

De son côté, Djossè Tessy a effectué deux missions : l'une au Niger en juillet 2018 à l'occasion d'une réunion des CTS du CAMES, ce qui lui a permis de rencontrer plusieurs utilisateurs potentiels du DICAMES dans les jurys mis en place par le CAMES, et l'autre au siège social du CAMES à Ouagadougou en octobre 2018 où il a pu former des employés et des stagiaires à l'utilisation des formulaires de dépôt. Finalement, une mission

de Florence en Guinée et au Sénégal en décembre 2018 termine ce programme.

Pour la communication interne du projet et la mobilisation constante des acteurs et actrices, notamment des responsables locaux du DICAMES dans les universités pilotes, une solution originale a été trouvée après plusieurs tâtonnements : un groupe de discussion sur le réseau social WhatsApp créé vers la fin 2017 et qui s'est graduellement élargi à tous les nouveaux alliés et acteurs ou actrices rencontrés sur le terrain dans différents pays tout au long de l'année. La communication, d'emblée complexe, entre ces différentes personnes résidant dans des lieux géographiquement distants a pu se faire efficacement par cet usage de WhatsApp, le réseau social le plus fréquenté en Afrique. La possibilité d'interactions aussi bien en temps réel qu'en différé par de courts messages a permis de bâtir une base de connaissances utile à l'action immédiate sur certaines étapes du processus. C'est un excellent outil d'idéation collective utilisé dans un processus à caractère scientifique, ce qui de prime à bord ne paraît pas évident. Florence y poste régulièrement des comptes rendus des avancées du projet, ce qui permet à tous et toutes de rester informés, de poser des questions, d'organiser des rencontres, etc.

Sur le plan du rayonnement, outre la présentation du DICAMES par l'équipe du CAMES un peu partout en Afrique francophone subsaharienne, notons un article du magazine de vulgarisation scientifique Sci-Dev paru en juillet 2018 sur le DICAMES (Irwin et Amadou, 2018), une présentation du DICAMES à un groupe de bibliothécaires américains à Chapel Hill aux États-Unis en octobre, une conférence de Zakari Liré sur le DICAMES au Congrès canadien des professionnels et professionnelles de la documentation en novembre et la communication au colloque ICOA de Rabat dont est tiré le présent texte. Le premier dépose-thon a eu lieu à Québec le 5 mai 2018, suivi de celui du 26 novembre qui a eu lieu simultanément à Québec, Paris et à Parakou.

16 Les alliés et alliées du DICAMES

Disposant d'un petit budget pour un projet aussi complexe, l'équipe a dû aller à l'essentiel dans les dépenses et mobiliser de nombreux alliés bénévoles. L'essentiel du budget (24 980 \$ canadiens) est en fait allé à trois postes clés. D'une part, il a servi à payer les informaticiens de la compagnie à but non lucratif koumbit.org, spécialisée dans le web solidaire, qui ont été engagés pour installer le logiciel Dspace en décembre 2018. Bien que vivant à Montréal, le chargé de projet chez Koumbit, Nizar Koubaa est tunisien et a très vite compris l'impact social du DICAMES, ce qui a permis d'établir

une excellente collaboration. Nous avons également eu recours brièvement à la firme québécoise InLibro, spécialisée en infrastructure technologique documentaire, pour nous aider à configurer l'interface du site du DICAMES.

Mais cette interface, assez peu modifiée par rapport à ce que propose par défaut la version 6 de Dspace que nous avons utilisée, a surtout été l'œuvre de l'assistant de recherche béninois de Florence, Djossè Roméo Tessy, doctorant en communication publique, qui a pris la relève de Thomas Hervé MboaNkoudou en septembre 2017. Djossè, grâce à un salaire versé par le projet, a non seulement travaillé à la configuration du site et de ses collections, mais il a élaboré les formulaires de dépôt et a construit le Centre de ressources du DICAMES. Le reste de la subvention a servi à financer en partie les missions de terrain présentées plus haut, ainsi que des dépenses connexes liées à l'informatique (hébergement, Handle, etc.). Le CAMES a déjà pris le relais et financé l'hébergement du DICAMES jusqu'en janvier 2021.

Les alliés et alliées bénévoles ont joué un rôle très important et nous les en remercions. Certains sont des experts de DSpace et nous ont aidés en nous envoyant des fichiers nécessaires à la configuration du site ou en répondant à nos questions concernant la traduction des formulaires ou le référencement du DICAMES. Nous avons aussi recruté des alliés bénévoles précieux parmi les doctorants et doctorantes, les enseignants et enseignantes et les étudiants et étudiantes en sciences de l'information de plusieurs universités, notamment à l'Université de Dschang, à l'Université de Parakou, à l'Université Félix-Houphouët-Boigny de Cocody et à l'Université d'Abomey-Calavi. Ces personnes, confrontées aux difficultés de trouver des ressources documentaires sur l'Afrique, ont bien compris la révolution que constituerait un DICAMES bien rempli. En raison de la surcharge de travail des bibliothécaires universitaires et du peu de moyens et de temps dont ils et elles disposent pour de nouveaux projets exigeants sur le plan technologique tels que le DICAMES, les doctorant-e-s bénévoles, futurs enseignants-chercheurs et enseignantes-chercheuses, sont appelés à jouer un rôle majeur dans le déploiement du DICAMES et le désenclavement de la science de leur pays, de leur continent. Leur mobilisation fait partie du caractère innovateur de notre manière de piloter le projet.

17 Conclusion : les défis du DICAMES

Les défis sont nombreux. Le plus grand défi de l'appropriation du DICAMES dans les universités de l'espace CAMES est bien sûr la fracture numérique (Alzouma 2008 ; Banque mondiale 2016 ; Nyirenda-Jere et Biru 2015 ; van Dijk 2012). L'absence corollaire de virage numérique dans la

plupart des universités de l'espace CAMES se manifeste par un accès précaire à la connexion wifi et aux laboratoires informatiques dans les campus, par une numérisation très rare des travaux imprimés (sauf en Côte d'Ivoire où le processus est bien lancé par l'Université Virtuelle de Côte d'Ivoire), mais aussi par de faibles compétences numériques chez la plupart des universitaires qui, en grande majorité, ne touchent à un ordinateur qu'en entrant à l'université (Piron, MboanKoudou, et al., 2016). Ce « fossé numérique » (Wackermann 2008) et la faible littératie numérique globale (Hoeschsmann et DeWaard 2015) dans cette région du monde constituent un défi à l'appropriation d'une innovation comme le DICAMES. Nous pensons à recommander pour l'avenir des dispositifs spécifiques tels que la possibilité de consulter l'archive sur un serveur local (par exemple, un ordinateur dans une bibliothèque universitaire), sans connexion Internet et avec une mise à jour régulière.

Toutefois, les fondements solides que nous avons établis depuis juin 2015 grâce à l'engagement, au dévouement et à l'agilité créative d'un petit nombre de personnes efficaces, compétentes, persévérantes, capables de travailler en équipe malgré la distance et qui se respectent et s'apprécient nous permettent d'espérer une durabilité réelle pour ce grand projet collectif. La suite de l'histoire dans quelques années !

Références

- ALZOUMA, Gado (2008). Téléphone mobile, Internet et développement : l'Afrique dans la société de l'information ?. In *tic&société*, vol. 2, n° 2. doi:10.4000/ticetsociete.488.
- BANQUE MONDIALE (2014). Améliorer la qualité et la quantité de la recherche scientifique en Afrique, [En ligne]. Disponible à : <http://www.banquemondiale.org/fr/region/afr/publication/improving-the-quality-and-quantity-of-scientific-research-in-africa> (Page consultée le 12 novembre 2018)
- BANQUE MONDIALE (2016). Les dividendes du numérique, [En ligne]. Disponible à : <http://documents.worldbank.org> (Page consultée le 12 novembre 2018)
- BAR, François ; WEBER, Matthew S. ; PISANI, Francis (2016). Mobile Technology Appropriation in a Distant Mirror: Baroquization, Creolization, and Cannibalism. In *New Media & Society*, vol. 18, n°4, pp. 617-36. doi:10.1177/1461444816629474.
- BRETON, Philippe ; PROULX, Serge (2002). Usages des technologies de l'information et de la communication. In *L'explosion de la communication à l'aube de XXIe siècle*. Montréal: Boréal, 2002. 389p.

- CHARLIER, Jean-Emile ; CROCHE, Sarah ; NDOYE, Abdou Karim (2009). Les universités africaines francophones face au LMD : Les effets du processus de Bologne sur l'enseignement supérieur au-delà des frontières de l'Europe. Louvain-La-Neuve: Editions Academia, 2009.
- DECUNG, Laurie ; MUKUKU, Françoise (2016). Overview of Open Knowledge in Francophone Africa. Report. Brighton, Institute of Development Studies, 2016.
- DIJK, Jan A. G. M. van (2012). The Evolution of the Digital Divide. In Digital Enlightenment Yearbook. Amsterdam : IOS Press, 2012. pp.57 7
- DIOUF, Antonin (2010). Quelques bibliothèques numériques sénégalaises [En ligne]. Disponible à : <http://antoninbenoitdiouf.com/2010/06/08/n%c2%b0-51-quelques-bibliotheques-numeriques-senegalaises/> (Page consultée 11 novembre 2018)
- ESCOBAR, Arturo (1995). Encountering Development. The Making and Unmaking of the Third World. Princeton : Princeton University Press. 290 p.
- EZEMA, J. Ifeanyi (2013). Local contents and the development of open access institutional repositories in Nigeria University libraries: Challenges, strategies and scholarly implications. In Library Hi Tech, vol. 31, n°2, pp. 323-340. <https://doi.org/10.1108/07378831311329086>
- FREDUA-KWARTENG, Eric (2015). The case for developmental universities. University World News. [En ligne]. Disponible à : <http://www.universityworldnews.com/article.php?story=20151028020047530> (Page consultée 11 novembre 2018)
- GAGLIO, Gérald (2011). Sociologie de l'innovation. Paris : Presses universitaires de France, 2011. 128p.
- HOECHSMANN, Michael ; DEWAARD, Helen (2015). Définir la politique de littératie numérique et la pratique dans le paysage de l'éducation canadienne, Habilomédias. [En ligne]. Disponible à : [site http://habilomedias.ca](http://habilomedias.ca) (Page consultée 11 novembre 2018)
- HOUNTONJJI, Paulin (2001). Le savoir mondialisé: déséquilibres et enjeux actuels. La mondialisation vue d'Afrique, Université de Nantes/Maison des Sciences de l'Homme Guépin.
- IRWIN, Aisling ; AMADOU, Himadou (2018). Lancement prochain d'une archive numérique du CAMES. [En ligne]. Disponible à : https://www.scidev.net/afrique-sub-saharienne/r-d/actualites/projet-dicames.html?fbclid=IwAR2yWpkkHZ_a6P25c_MiIbXILKaHo4P4bie8r0lzK4_2do9KdNmLGWFQ348 (Page consultée 11 novembre 2018)
- JAUREGUIBERRY, Francis ; PROULX, Serge (2012). Trois approches classiques pour penser les usages. In Poche - Société, juillet, pp. 32-56.

- KAVULYA, M. Joseph (2007). Digital libraries and development in Sub-Saharan Africa: A review of challenges and strategies. In *The Electronic Library*, vol. 25, n°3, pp. 299-315. <https://doi.org/10.1108/02640470710754814>
- MAVHUNGA, C. Chakanetsa (2017). *What Do Science, Technology, and Innovation Mean from Africa ?* Cambridge : MIT Press, 2017.
- MVE-ONDO, Bonaventure (2005). *Afrique : la fracture scientifique / Africa: the Scientific Divide*, Futuribles. Consulté à l'adresse <https://www.futuribles.com/en/base/bibliographie/notice/afrique-la-fracture-scientifique-africa-the-scient/> (Page consultée le 12 novembre 2018)
- NYIRENDA-JERE, Towela ; BIRU, Tesfaye (2015). Développement de l'Internet et gouvernance de l'Internet en Afrique. *Internet society*.
- NKOLO, P. Niclair (2016). Open Access et valorisation des publications scientifiques : les défis de l'Afrique francophone. In *Justice cognitive, libre accès et savoirs locaux. Pour une science ouverte juste, au service du développement local durable*. Québec : Éditions science et bien commun, 2016.
- PIRON, Florence (2018). Justice et injustice cognitives : de l'épistémologie à la matérialité des savoirs humains », In *Les Classiques des sciences sociales. 25 ans de partage des savoirs. Sous la direction d'Émilie Tremblay et RicarsonDorcé*. Québec : Éditions science et bien commun, pp. 259-273. [En ligne]. <https://corpus.ulaval.ca/jspui/handle/20.500.11794/30244> (Page consultée le 12 novembre 2018)
- PIRON, Florence (2016). La science ouverte peut-elle réduire l'aliénation épistémique ?. Colloque Des sciences métissées sont-elles possibles ?. In *Réflexions sur la pluralité des savoirs, des normes et des pratiques*. 84e congrès de l'ACFAS, Montréal, 9-13 mai.
- PIRON, Florence (2017a). Libre accès aux publications scientifiques. In *Options politiques*. [En ligne]. Disponible à : <http://policyoptions.irpp.org/magazines/june-2017/le-libre-acces-aux-publications-scientifiques/> (consultée le 12-11-18)
- PIRON, Florence (2017b). Une solution pour la mise en valeur de la production scientifique des universités de l'espace CAMES : une archive numérique ouverte en libre accès. Assemblée des recteurs du CAMES (Conseil africain et malgache pour l'enseignement supérieur), Cotonou, 17 juillet.
- PIRON, Florence ; NKOUDOU, Thomas Mboa ; REGULUS, Samuel et al. (2016). « Une autre science est possible. Récit d'une utopie concrète, le projet SOHA ». *Possibles*, vol. 40, n°2. [En ligne]. Disponible à : <http://redtac.org/possibles/> (Page consultée le 12 novembre 2018)
- PIRON, Florence, DIOUF, Antonin Benoît ; DIBOUNJE MADIBA, Marie Sophie ; MBOA NKOUDOU, Thomas Hervé ; PIERRE, Anderson ; OUANGRE, Aubierge Zoé ; TESSY, Djossè Roméo ; ACHAFFERT, HamissouRhissa ; LIRE, Zakari (2017). Le libre accès vu d'Afrique francophone subsaharienne. In *Revue française des sciences de l'information*

et de la communication, n°11. [En ligne]. Disponible à : <http://rfsic.revues.org> (Page consultée le 12 novembre 2018)

PIRON, Florence ; MBOA, Nkoudou Thomas Hervé ; ANDERSON, Pierre ; DIBOUNJE MADIBA, Marie Sophie ; ALLADATIN, Judicaël ; FALL Assane ; RHISSA ACHAFFERT, Hamissou ; MICHEL, RencyInson. (2016). Vers des universités africaines et haïtiennes au service du développement local durable : contribution de la science ouverte juste. In *Justice cognitive, libre accès et savoirs locaux. Pour une science ouverte juste, au service du développement local durable*, Québec : Éditions science et bien commun, 2016.

PIRON, Florence ; REGULUS, Samuel ; DIBOUNJE MADIBA, Marie Sophie (dir.) (2016). *Justice cognitive, libre accès et savoirs locaux. Pour une science ouverte juste, au service du développement local durable*. Québec : Éditions science et bien commun, 2016.

SCHÖPFEL, Joachim ; SOUKOUYA, Maebena (2013). Providing Access to Electronic Theses and Dissertations : A Case Study from Togo. In *D-Lib Magazine*, n°19, vol. 11/12. <https://doi.org/10.1045/november2013-schopfel>

TESSY, Djossè Roméo (2016). Les obstacles à l'adoption du libre accès par les étudiants et étudiantes du Bénin. In *Justice cognitive, libre accès et savoirs locaux, au service de la science ouverte juste*. Québec : Éditions science et bien commun, 2016.

UNESCO (2015). *Des clés pour la promotion de sociétés du savoir inclusives : Accès à l'information et au savoir, liberté d'expression, respect de la vie privée et éthique sur un Internet mondial*. UNESCO Publishing.

WACKERMANN, Gabriel, éd. (2008). *Le développement durable*. Paris : Ellipses, 2008.

Ce texte est sous licence Creative Commons CC-BY 4.0 qui encourage sa réutilisation <https://creativecommons.org/licenses/by/4.0/deed.fr>.

Vers une stratégie de développement de la voie dorée marocaine : cas des revues en Sciences Humaines et Sociales

Towards a strategy for the development of the Moroccan gold open access: the case of journals in Human and Social Science

H. Lrhoul ; H. Zerouali ; A. A. Bachr

Laboratoire Mediation-Information-Knowledge-Society (MIKS)
ESI-Maroc

Résumé

Cette communication porte sur la problématique de la visibilité de la production marocaine en sciences humaines et sociales (SHS) et plus spécifiquement sa position par rapport à l'open access. Pour analyser le paysage éditorial des revues marocaines en SHS, nous avons constitué notre corpus en se référant aux données bibliographiques nationales. L'étude de la visibilité internationale s'est faite sur Scopus. Plusieurs résultats problématisables ressortent de cette analyse et concernent les indicateurs en termes de visibilité, d'édition et de langue de publication ainsi qu'en termes d'autres contraintes qui limitent leur référencement dans les BD internationales. Enfin, des recommandations sur le plan de politiques éditoriales sont proposées et pourraient s'inscrire dans une stratégie globale d'accélération de l'OA au Maroc.

Mots-clés. édition numérique ; SHS ; open access ; revues marocaines ; production scientifique ;

Abstract

This paper focuses on the issue of the visibility of Moroccan production in the humanities and social sciences and more specifically its position relative to open access. To analyze the editorial landscape of Moroccan journals in humanities and social sciences, we compiled our corpus with reference to national bibliographic data. The study of international visibility was done on Scopus. Several problematic results emerge from this analysis. They concern the indicators in terms of visibility, edition and language of publication as well as other constraints that limit their referencing in international databases.

Finally, recommendations on the editorial policies are proposed and could be part of a global strategy to accelerate OA in Morocco.

Keywords. digital publishing; Humanities and Social Sciences; Moroccan journals; open access; scientific production.

1 Introduction

Dans cette communication, nous souhaitons approcher la production marocaine en sciences humaines et sociales à l'aune des enjeux du libre accès à la connaissance scientifique lequel constitue aujourd'hui un nouveau mode de représentation et de circulation des savoirs, mode confronté à des barrières pour ne pas dire des frontières et dont les logiques ne sont pas seulement scientifiques comme il sera souligné dans ce travail.

Notre choix porte sur les sciences humaines et sociales car nous estimons que ces dernières sont foncièrement concernées par l'ensemble des défis et problématiques actuelles. En effet les sciences humaines et sociales expliquent les faits qui traversent les sociétés humaines. Elles sont de ce fait nécessaires à la réflexion, à l'action et à la décision. Ce n'est pas pour rien que le rapport mondial sur les sciences sociales de l'UNESCO co-publié avec le Conseil International des Sciences Sociales¹ en 2010 fait le point sur les acquis, les critiques, les problèmes ainsi que les perspectives et les défis majeurs qui renforcent leur place au sein des savoirs.

À ce titre, ce travail propose des perspectives et des recommandations qui pourraient s'inscrire dans une stratégie de développement d'un modèle pérenne pour l'intégration des revues marocaines en sciences humaines et sociales dans les nouveaux modes de représentation et de circulation de la connaissance. Après un état des lieux sur le libre accès, nous explorons la sphère éditoriale en SHS au Maroc et plus spécifiquement celle concernant les revues. Le corpus mobilisé s'appuie sur la base de données de la Fondation du Roi Abdul-Aziz Al Saoud pour les Etudes Islamiques et les Sciences Humaines, sur les bases de données internationales et la liste des revues indexées par l'Institut marocain pour l'information scientifique et technique. L'analyse permettra de connaître cette production et de comprendre sa quasi-absence dans les index internationaux.

2 Problématique

La mondialisation de la science et son internationalisation ouvrent de nouvelles frontières et facilitent le libre partage et l'échange du savoir scientifique. Dans ce schéma d'ouverture, l'Open Access n'est plus un choix mais une nécessité pour adhérer aux transformations de la chaîne de production et de diffusion de la science. Les acteurs responsables de la recherche scientifique au Maroc ne pourraient envisager le développement de l'accès et de la diffusion de l'Information Scientifique et Technique sans intégrer la voie de l'Open Access et contribuer à un meilleur positionnement de leurs chercheurs sur la scène internationale.

En effet, ce nouvel environnement de publication scientifique contribuerait à faire sortir les chercheurs et les universités de l'ombre, en mettant le patrimoine scientifique marocain à la disposition de la communauté nationale, régionale et internationale. Dès lors, mener des initiatives pour limiter les barrières d'accès à l'IST et sensibiliser les politiques à promouvoir le partage libre des travaux de recherche financés par les

¹UNESCO 2010, Rapport mondial sur les sciences sociales. Division dans les savoirs. En Ligne <http://unesdoc.unesco.org/images/0018/001883/188395f.pdf>

deniers publics devient un enjeu stratégique pour la recherche et la production scientifique.

Soulignons que dans les pays du Sud, plusieurs auteurs ont largement relevé les enjeux du mouvement de l'open access qui présentent pour eux un certain nombre d'avantages, étant donné que les chercheurs manquent généralement de moyens financiers pour s'abonner aux revues scientifiques et assurer l'accès à l'information scientifique et technique essentielle à leurs activités. Aussi, le fait de rester à l'écart du mouvement marginalise encore plus les chercheurs du Sud par rapport à la communauté internationale puisqu'ils sont peu nombreux à vouloir publier dans des revues en open access ou à autoarchiver leur production scientifique dans des dépôts institutionnels. Cela va à l'encontre du principe de « communauté scientifique » qui est censée échanger les connaissances et collaborer pour résoudre les problèmes de l'humanité².

Par ailleurs, au moment où les différentes voies de l'Open Access connaissent un développement sans pareil dans les pays développés, les institutions académiques marocaines tardent à intégrer le mouvement de l'open access par manque de connaissance³ des nouveaux paradigmes de la communication scientifique, les limitations technologiques et humaines. Plusieurs études ont montré que l'open access ne connaît pas la même évolution selon les disciplines : alors que les chercheurs en médecine sont les pionniers en ce qui concerne la création d'archives ouvertes et le partage de leurs travaux de recherche sur des plateformes en open access⁴, l'adhésion à ce mouvement tarde à venir dans le champ des SHS, où les chercheurs semblent moins impliqués dans le paysage de l'édition électronique et du libre accès⁵.

Afin d'accompagner les éditeurs vers les transformations numériques des revues scientifiques et l'intégration du mouvement de l'open access, nous envisageons dans cette communication d'explorer les caractéristiques de la publication scientifique dans le champ des sciences humaines et sociales, de recenser les revues marocaines en open access et d'analyser le paysage de l'édition actuel.

² Price, D. ; Beaver, D. (1966). Collaboration in an invisible college. *American psychologist*, 21(11), 1011.

³ Comme préconisé par plusieurs résultats d'enquêtes sur le degré de connaissance de l'open access dans les pays arabes

Ben Romdhane, M. ; Ouerefelli, T. (2013). L'offre des archives ouvertes dans le monde arabe : recensement et évaluation.

Carlson, S. (2015). An Assessment of Institutional Repositories in the Arab World.

Gdoura, W. ; Bouazza, A. ; Ibnkhayat, N. (2009). Attitudes des universitaires arabes à l'égard des revues en ligne et des archives ouvertes : cas du Maroc, d'Oman et de la Tunisie.

Lrhoul, H. (2017). La production scientifique des chercheurs de la faculté de médecine et de pharmacie de Casablanca : mesures, cartographie et enjeux du libre accès (Doctoral dissertation, Conservatoire national des arts et métiers-CNAM)

⁴ Bo-Christer Björk, Mikael Laakso, and Patrik Welling. Anatomy of Green Open Access

⁵ As Martin Paul Eve. Comments, "the humanities still trail behind the sciences in open publishing" [6] (p. 1).

3 Caractéristiques de la publication scientifique en SHS

Au moment où la recherche en Sciences et Techniques et Sciences médicales (STM) est considérée comme internationalisée, structurée autour de communautés homogènes et de champs disciplinaires bien définis, celle en Sciences Humaines et Sociales (SHS) dépend toujours de différentes écoles de pensée ou d'institutions de recherche⁶. Ce qui explique l'hétérogénéité des communautés scientifiques en SHS et leur mode de communication. Il s'agit d'une communauté qui privilégie les thématiques de recherche publiées dans des langues locales, généralement dans des revues en version imprimée. Aussi, les chercheurs en SHS se sentent moins concernés par la publication électronique dans les revues prestigieuses à l'instar des chercheurs en STM.

Une autre différence entre les deux communautés est liée Selon Ferwerda⁷ à la rédaction scientifique. Alors que les articles en STM sont généralement courts avec une présentation concise de la méthodologie et des résultats de recherche, ceux publiés en SHS donnent plus d'importance à la définition des concepts, à l'argumentation et au développement des idées avant de présenter les résultats⁸. Par conséquent, les chercheurs en SHS privilégient les longs articles ou les monographies.

De ce fait, le choix de la monographie comme support de publication limite leur visibilité dans les bases de données internationales qui recensent plus d'articles que d'ouvrages, plus de publications anglophones qu'en d'autres langues et plus de thématiques à intérêt international que celles traitant des problématiques locales.

L'usage des bases de données internationales, pour l'évaluation de la visibilité et l'impact de la recherche en SHS, ne peut donc être exhaustif pour trois principales raisons. *Primo*, il s'agit de la sous indexation des ouvrages, support de base des chercheurs en SHS, lesquels ne pouvant être comparés aux articles aussi bien quant à leur mode de production que celui de leur diffusion. *Secundo*, le champ de recherche axé sur des problématiques locales ne correspondant pas aux thématiques choisies par ces bases de données. *Tertio*, le choix des chercheurs en SHS de publier souvent en langues locales entrave leur visibilité dans les bases de données « Web of Science » et « Scopus » notamment.

Un autre élément qui entrave la visibilité de la production scientifique en SHS à l'échelle internationale, concerne l'usage du facteur d'impact pour le classement des revues. Cet indicateur de mesure ne peut s'appliquer aux revues en SHS dont la durée de vie et l'impact peut atteindre une dizaine d'année ou plus, ce qui ne correspond pas aux critères de qualité du facteur

⁶ Groupement français de l'industrie de l'information (2009). L'édition scientifique française en SHS – Rapport de synthèse. En ligne <https://www.gfii.fr/uploads/docs/l-edition-scientifique-francaise-en-sciences-sociales-et-humaines.pdf>

⁷Ferwerda, E. Open access in humanities and social sciences (2013). In 8th Munin Conference Scholarly Publishing, Tromso, Norway, 25–26 November 2013.

⁸Landscapes of Research: Perceptions of Open Access (OA) Publishing in the Arts and Humanities. Julia Gross and John Charles Ryan.

d'impact qui mesurent les citations des travaux scientifiques sur la base de deux ans.

Malgré ces limites précitées, les gestionnaires de la recherche dans les pays du Sud utilisent le même standard d'évaluation et d'avancement de leurs chercheurs quel que soit leur discipline. Il s'avère donc nécessaire de sensibiliser les universitaires sur les limites des indicateurs bibliométriques classiques et de compléter par d'autres modes d'évaluation. Pour combler les lacunes des bases de données internationales et leur critère de jugement mal adaptés aux SHS et à la littérature non anglophone, les pays du Nord optent pour deux stratégies : d'une part, ils mettent en place des instances d'évaluation de la recherche qui prennent en considération les revues francophones et nationales qui ne sont pas indexées dans ces bases de données. D'autre part, ils adhèrent au mouvement de l'open access en créant des bases de données de la production nationale, en renforçant leurs structures éditoriales et en créant des portails de revues nationales. Il serait intéressant de s'inspirer du modèle des pays du nord qui conjuguent les standards locaux d'évaluation et l'open access, stratégie qui permet de valoriser leurs revues en SHS

Enfin, le rapport⁹ sur l'édition des SHS en France pointe du doigt le manque de structures éditoriales pour augmenter l'impact de leurs chercheurs, alors que les STM disposent de circuit éditorial performant.

Dans ce qui suit, il sera question du contexte marocain et plus spécifiquement des problèmes auxquels les éditeurs marocains en SHS doivent faire face. Nous explorons aussi les opportunités offertes par l'open access pour améliorer le paysage de l'édition académique.

4 Edition scientifique en SHS au Maroc

L'édition académique en SHS concerne, dans cet article, les périodiques de recherche spécialisés en SHS, édités par des éditeurs académiques ou privés, en version imprimée ou électronique. En effet la plupart des éditeurs de revues scientifiques en SHS sont des éditeurs académiques qui relèvent des universités ou des institutions de recherche s'inscrivant dans le champ des SHS. De ce fait, ils dépendent du faible soutien des universités (local, personnel, frais d'impression) et ne peuvent disposer des ressources humaines compétentes pour offrir des services de qualité et créer des revues innovantes. En ce qui concerne les éditeurs privés, les difficultés financières liées à la baisse des ventes et le manque de rentabilité des revues scientifiques les fragilisent et menacent la pérennité et la garantie de la régularité des revues actuelles.

Certes, le fait de dépasser le modèle de revues imprimées et d'intégrer le paysage de l'édition électronique en open access pourrait renforcer le tissu éditorial actuel, en réduisant les frais d'impression et en permettant aux communautés scientifiques d'augmenter leur audience et impact. Mais cela nécessite des investissements dans l'innovation technologique et dans la mise à jour des connaissances des chercheurs pour maîtriser ces nouveaux outils et se détacher du support physique.

⁹Groupement français de l'industrie de l'information (2009). L'édition scientifique française en SHS – Rapport de synthèse.

De plus, une vision claire des instances de pilotage de la recherche scientifique serait nécessaire. Actuellement, chaque partie (des éditeurs militant pour conserver leur gain, des chercheurs intéressés par la large diffusion de leurs publications, des universités qui manquent de ressources pour payer les abonnements aux revues scientifiques) défend ses intérêts ou ses valeurs sans aucune stratégie nationale ou possibilité d'instaurer des politiques d'OA aux résultats de la recherche.

5 Open access : opportunité pour les revues scientifiques en SHS

Il n'est plus à démontrer que l'open access offre de multiples avantages aux chercheurs des pays du Sud pour accéder, produire et diffuser leurs recherches et contribuer ainsi au développement de la science mondiale. Sur le plan de la production et de la diffusion des publications, la gestion électronique du processus éditorial et la mise en ligne en open access pourraient contribuer à la large diffusion du savoir scientifique local et à l'amélioration de sa qualité. Le chercheur devient partie prenante du circuit de l'édition scientifique en tant que producteur de revues scientifiques, grâce au développement de logiciels de création de revues électroniques et pour la diffusion de leur production scientifique. Aussi, plusieurs avantages sont offerts aux éditeurs : un nouvel environnement d'édition scientifique permettant de gérer plus de soumissions d'articles avec moins de ressources financières et moins de temps, publication dans les délais, augmentation de la visibilité des revues.

Cependant, les chercheurs et éditeurs marocains restent à l'ombre en dépit de toutes les opportunités offertes par l'open access pour offrir des revues de qualité et augmenter la visibilité et l'impact de la recherche marocaine. Quelles sont donc les entraves qui les empêchent d'adopter la voie dorée de l'open access ?

Les études menées sur l'OA au Maroc ont principalement visé à comprendre la perception de ce phénomène au niveau de la population des chercheurs et à recommander des actions incitatives dans les politiques de recherche mais ces études n'ont jamais abordé la question sous l'angle des éditeurs de revues scientifiques. Nous pensons que ces derniers sont des acteurs prépondérants pour le développement du libre accès et la sensibilisation des auteurs à ce paradigme de publication.

Soulignons aussi que la plupart des revues scientifiques marocaines font face à plusieurs contraintes qui limitent leur référencement dans les bases de données indexées, à savoir, la non disponibilité de revues en version numérique, la faible maturité des infrastructures d'édition scientifique, le manque de financement et d'éditeurs professionnels. D'où la nécessité d'une restructuration en profondeur pour répondre aux normes de qualité de l'édition scientifique.

Autre élément de la problématique réside dans le fait que les rares études marocaines sur la question de l'OA ont concerné le champ des sciences dures alors que les sciences humaines et sociales sont encore plus importantes en termes d'effectifs de chercheurs et de supports de publications. Il serait donc judicieux de focaliser notre recherche sur les revues en SHS afin de proposer une stratégie d'accélération du modèle de l'OA au Maroc.

6 Objectifs et méthode de collecte des données

L'objectif de notre étude est d'analyser le paysage des éditeurs marocains et de mener une réflexion sur un modèle de publication pérenne en open access. Elle vise plus particulièrement à :

- Analyser la sphère éditoriale marocaine en SHS ;
- Recenser les revues en open access en SHS ;
- Proposer des pistes de renforcement de l'OA comme moyen de diffusion des résultats de recherche.

Pour recenser et analyser la production scientifique dans le domaine des SHS, nous avons choisi dans cette étude de constituer notre corpus en nous basant sur les bases de données internationales et sur certaines bases de données bibliographiques nationales.

Pour répondre à ces questions, nous allons mener une étude au niveau des éditeurs de revues marocaines dans le champ des sciences humaines et sociales. Le recensement des revues se fera sur la base du catalogue de la bibliothèque de la Fondation du Roi Abdul-Aziz Al Saoud pour les Etudes Islamiques et les Sciences Humaines. Nous nous basons aussi sur le répertoire des revues en open access (DOAJ) et le portail de l'Institut marocain de l'Information Scientifique et technique (IMIST) pour inventorier les revues marocaines en open access.

Le modèle qui nous motive ici est celui de Scielo. L'accessibilité et la large diffusion des publications scientifiques de la région Latino-Américaine, ont augmenté le taux de citation et l'impact des chercheurs de la région. L'indexation de Scielo par l'éditeur Clarivate Analytics est la meilleure preuve de l'apport des bases de données locales et de l'open access sur la valorisation de la science des pays du Sud.

7 Résultats de l'étude

7.1 Production scientifique marocaine en SHS

La croissance de l'édition scientifique en SHS est corollaire de la prolifération des publications scientifiques dans ce domaine. A cet effet, nous analysons les articles publiés dans les bases de données internationales, malgré les limites précitées de ces dernières, pour donner une idée générale sur la faiblesse de cette production par rapport au STM et à la production locale.

L'analyse de la production scientifique en SHS dans la base de données Scopus est un indicateur¹⁰ de la visibilité de la recherche et l'impact des chercheurs marocains dans ce domaine à l'échelle internationale. Selon le baromètre¹¹ de l'IMIST, les sciences humaines et sociales représentent le champ disciplinaire le moins productif au Maroc, le nombre de publications scientifiques entre 2007 et 2017 ne dépasse pas 2624 articles. Alors que la

¹⁰ Les bases de données internationales restent les seuls outils utilisés pour l'évaluation de la qualité et l'impact de la recherche à l'échelle nationale et internationale, malgré leurs biais d'indexation de la recherche en SHS et celle des pays en développement.

¹¹ Baromètre de la production scientifique au Maroc. <https://barometre.imist.ma/>

production scientifique en physique dépasse 25 600 articles et celle en médecine atteint 8884 pendant la même période.

Sur le plan des sous disciplines, le domaine des sciences sociales est le plus productif, pendant la période 2007-2017 (voir figures 1 et 2), avec 1080 articles scientifiques, suivi du management (623 articles), de l'art et des sciences humaines (366 articles) et de l'économie (257 articles).

Figure 1 : production scientifique marocaine par champ disciplinaire

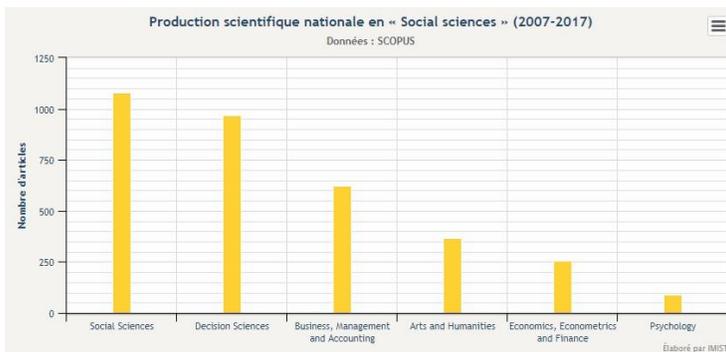
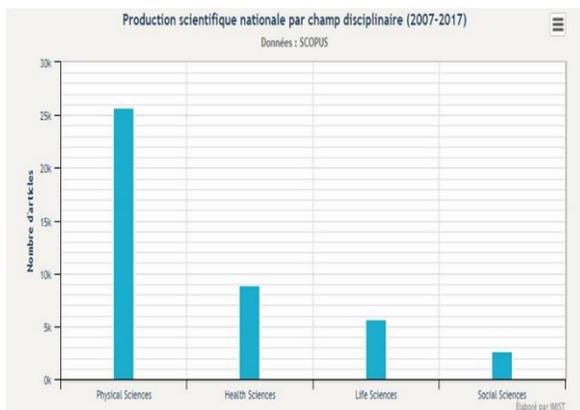


Figure 2 : Production scientifique marocaine en sciences sociales



En ce qui concerne l'indexation des revues marocaines, HesperisTamuda est la première revue spécialisée dans le domaine des sciences humaines et sociales, qui a été indexée par la base de données internationale « Web of Science » depuis 2016.

7.2 Edition scientifique marocaine en SHS

Analyser le paysage éditorial marocain est une étape primordiale pour vérifier les possibilités d'envisager un modèle de publication pérenne en

open access. Notre étude concerne les éditeurs scientifiques¹² marocains dans le domaine des sciences humaines et sociales.

Le rapport de la Fondation Al Saoud pour les Sciences Humaines et Sociales recense 189 titres de revues académiques marocaines, publiés généralement en version papier. Ce nombre paraît important si on le compare avec le nombre de revues françaises (300 revues en SHS), sauf que la qualité ne rejoint pas l'évolution de la quantité ; les périodiques français jouissent d'une grande visibilité à l'échelle internationale, ce qui témoigne de leur qualité¹³. Cela peut être lié aux financements publics attribués aux éditeurs, qui représentent selon le même rapport plus de deux fois leurs recettes¹⁴.

Pour analyser le paysage des éditeurs marocains, nous avons opté pour les critères suivants : langue de publication, ancienneté, type d'éditeur, politique d'open access, indexation dans des portails nationaux ou internationaux.

Dans l'analyse qui va suivre, Nous nous sommes basés, non pas sur le rapport de la Fondation Al Saoud¹⁵, mais sur la liste de 446 revues marocaines recensées par la fondation Al Saoud. Cependant, une précision s'impose. En effet, cette liste comprend des revues scientifiques ainsi que des bulletins, rapports, etc. dont nous ne pouvons juger la valeur scientifique en nous basant uniquement sur le titre. Nous comptons approfondir cette analyse quantitative dans une recherche ultérieure, en consultant les revues en version papier et en menant des entretiens avec leurs éditeurs.

7.3 Langue de publication

L'analyse des langues de publication des revues marocaines, telles qu'elles figurent dans le tableau N°1, fait ressortir le fait que la grande majorité des éditeurs 87% sont arabophones, alors que 13% seulement éditent des revues en langue française.

Tableau 1 : Répartition des revues par langue de publication

Langue	Nombre de revues	Pourcentage
Arabe	390	87
Français	57	13
Total	446	100

L'analyse du tableau N°1 montre que les éditeurs et chercheurs marocains doivent faire face à la contrainte linguistique qui entrave la visibilité des travaux de recherche à l'échelle internationale, en développant des

¹²Selon le rapport sur l'édition française en SHS, le secteur de l'édition SHS peut être défini comme l'édition de travaux de recherche originaux, exigeant du lecteur une connaissance préalable de la discipline dans laquelle l'ouvrage s'inscrit. Ils peuvent être sous forme de livres ou de périodiques spécialisés.

¹³ Groupement français de l'industrie de l'information (2009). L'édition scientifique française en SHS – Rapport de synthèse.

¹⁴estimées à 9 à 10 M€

¹⁵ Rapport sur l'édition et le livre

plateformes nationales de diffusion des revues locales qui s'intéressent aux thématiques du pays.

7.4 Typologie des éditeurs

Aussi, l'analyse de la liste des éditeurs des revues nationales scientifiques illustre que la grande majorité de ces éditeurs agissent sous la responsabilité d'enseignants-chercheurs.

Le manque de financement institutionnel et de rentabilité de l'édition de revues scientifiques en SHS menace la pérennité des revues marocaines universitaires, qui perçoivent rarement un soutien indirect des universités (local ou équipement). Leur continuité dépend ainsi de la volonté et capacité des chercheurs à assurer le financement et la durabilité de leurs revues.

La présentation des acteurs de l'édition en SHS au Maroc révèle d'une part une précarité (consistance, périodicité, pérennité) due principalement au manque de moyens financiers et humains, et d'autre part, l'absence de stratégies nationales de l'édition scientifique.

Tableau 3 : Répartition des revues par année d'édition

Date d'édition	Nombre de revues
1950-1990	63
1990-2000	67
2000-2010	129
2010-2015	280

Le tableau N°3 permet de constater que les revues marocaines sont généralement jeunes, et que la dernière décennie a connu la création d'un grand nombre de revues scientifiques. Cela est loin d'être un indicateur de croissance de l'édition scientifique marocaine, le rapport de la Fondation Al Saoud a soulevé les problèmes de périodicité, d'irrégularité de parution et de non pérennité des revues¹⁶.

7.5 Editeurs marocains et open access

L'open access pourrait être considéré comme alternative au modèle économique non viable des revues marocaines. Dans l'objectif d'accompagner les éditeurs marocains à intégrer ce mouvement et profiter de toutes les opportunités offertes par l'open access, nous souhaitons émettre des recommandations pour améliorer la qualité, la diffusion et la visibilité de ces revues. En premier lieu, nous étudions le degré d'adhésion des revues marocaines à l'open access. En second lieu, nous proposons des pistes de réflexion pour renforcer la visibilité des revues locales.

La première difficulté réside dans l'analyse des politiques éditoriales des éditeurs, l'objectif étant de vérifier leur adhésion au mouvement de l'OA. A l'inverse des revues indexées dans des bases internationales avec des politiques éditoriales claires qu'on peut retrouver sur les sites Sherpa Roméo ou Dulcinea, la majorité des revues marocaines sont publiées en

¹⁶Fondation du Roi Abdul Aziz Al-Saoud pour les Etudes Islamiques et les Sciences Humaines (2017).
Edition et Livre au Maroc : Rapport annuel sur l'état de l'édition et du livre dans les domaines de la littérature, sciences humaines et sociales 2015/2016.

version imprimée et ne disposent pas de politique de diffusion en open access.

Pour cartographier les revues marocaines publiées en open access, nous présentons dans la partie suivante les revues numériques indexées soit dans des portails nationaux comme celui de l'IMIST (tableau 4), soit dans le répertoire international des revues en open access (DOAJ).

Tableau 4 : Liste des revues indexées dans le portail de l'IMIST

Titre de la revue	Date d'édition électronique
(Bahtheyya) البحتية	2013
Revue de Gestion et d'Economie (JBE)	2014
Langues, Cultures et Sociétés (LCS)	2015
Revue de Recherches en Economie et en Management Africain (CREMA)	2015
Revue Marocaine de Recherche en Management et Marketing (REMAREM)	2015
Organisation et Territoires (OT)	2015
Le Pédagogue (LP)	2015
Espace Géographique et Société Marocaine (EGSM)	2016
Revue Marocaine de Gestion et d'Economie (RMGE)	2016
Recherches et Pratiques Marketing (RPM)	2016
Public & Nonprofit Management Review (PNMReview)	2016
African Management Review(RAM)	2016
Revue Marocaine d'Economie (RME)	2016
Recherches et Applications en Finance Islamique (RAFI)	2016
Journal of Information Systems Management and Innovation (ISMI)	2016
Revue Africaine des Migrations Internationales (RAMI)	2016
Journal of Applied Chemistry And Environment Protection (JACEP)	2016
Revue des Etudes et Recherches en Logistique et Développement (RERLED)	2017
Revue du Contrôle de la Comptabilité et de l'Audit (RCCA)	2017

Revue "Tidighin" des Recherches Amazighes et Développement (Tidighin)	2017
(MONAZAAT) مجلة منازعات الأعمال	2017
Langues, Cultures et Communication (L2C)	2018
International Review of Economics, Management and Law Research (IREMLR)	2018
Faits de Langue et Société (FLS)	2018
International Journal of Arabic Linguistics (IJAL)	2018
Revue Marocaine de Management, Logistique et Transport (RMLT)	2018
Revue Marocaine de la Prospective en Sciences de Gestion (RMPSG)	2018
Geopolitics and Geostrategic Intelligence (GGI)	2018
Revue Langues et Littératures (R2L)	2018
Revue de Publicité et de Communication Marketing (RPCoM)	2018
Revue Internationale de la Finance Entrepreneuriale (IREF)	2018
International Arab Journal of English for Specific Purposes (IAJESP)	2018
Journal d'Economie, de Management, d'Environnement et de Droit (JEMED)	2018
Journal of Applied Language and Culture Studies (JALCS)	2018

Le tableau 4 montre que seulement 34 revues marocaines sont publiées sur le portail national des revues scientifiques, géré par l'Institut Marocain de l'Information Scientifique Marocain. Cela peut être justifié par le choix de diffusion des revues en version imprimées et non pas en version électronique. Il est cependant intéressant de remarquer que le nombre de revues en version électronique a connu une évolution importante entre 2013 et 2018 grâce aux efforts de l'IMIST pour éditer les revues locales et aux initiatives de l'Université Hassan II Casablanca pour créer un portail de revues marocaines et de sensibiliser les éditeurs à numériser leurs revues et à utiliser le logiciel Open Journal System (OJS) pour gérer leur processus éditorial¹⁷.

La mise en ligne des revues locales est primordiale pour avoir le feedback de la communauté scientifique internationale et améliorer la qualité des

¹⁷Lrhoul, H. (2017). La production scientifique des chercheurs de la faculté de médecine et de pharmacie de Casablanca : mesures, cartographie et enjeux du libre accès (Doctoral dissertation, Conservatoire national des arts et métiers-CNAM)

publications diffusées en open access. Néanmoins, il nous paraît utile de préciser que l'indexation dans le portail national de l'IMIST n'est pas un gage de qualité, à l'encontre des bases de données internationale où l'évaluation de la qualité des revues se fait par leurs comités éditoriaux. C'est pour cette raison, que nous présentons, dans le tableau N°5 les revues nationales indexées dans le DOAJ, qui exige des éditeurs le respect des standards internationaux de la qualité.

La recherche par affiliation de l'éditeur dans le DOAJ nous permet de recenser 12 revues marocaines publiées dans plusieurs disciplines. Dans le domaine Sciences sociales, une seule revue électronique, sur les 34 diffusées sur le portail de l'IMIST, est indexée dans le DOAJ.

Tableau 5 : Liste des revues indexées dans le DOAJ

Titre de la revue	Editeur	Date d'indexation
Revue Marocaine de Recherche en Management et Marketing	Université Mohammed V de Rabat	2018

Il est indéniable que la numérisation et la mise en ligne des revues est la clé principale de la voie de l'open access, mais cela n'est pas suffisant pour augmenter la visibilité et l'impact des publications locales.

En effet, l'indexation dans le DOAJ nécessite un effort colossal pour s'aligner sur les standards internationaux, déployer les compétences humaines et les infrastructures technologiques (postes informatiques, internet, logiciel de gestion des revues électroniques). Dès lors, les éditeurs académiques ne pourraient assurer ces standards et intégrer la science internationale sans des financements pérennes, le soutien des universités et des responsables chargés de la planification de la recherche scientifique.

Enfin, de véritables stratégies nationales doivent être définies, depuis les plateformes pour recenser les publications et les revues nationales, en passant par les processus d'amélioration de leur qualité, jusqu'à la mise au point des index de citation pour la mesure de leur impact ; car ces revues ne sont pas incluses dans le DOAJ ou dans les bases de données internationales « WoS » et « Scopus ».

8 Perspectives et recommandations

Au terme de ce travail qui constitue une première étape d'un projet scientifique qui soulève la problématique de la représentation et de la circulation de la connaissance en sciences humaines et sociales à l'aune des enjeux du libre accès, des perspectives et des recommandations sont avancées. Trois points forts sont retenus et énoncés ci-dessous :

1. La mise en place d'une politique de mutualisation et d'une stratégie nationale d'appui aux éditeurs académiques pourraient renforcer les universitaires porteurs de projets et les motiver à créer des revues numériques.

2. La création d'un portail qui donne accès aux revues scientifiques nationales est nécessaire, tout en mettant l'accent sur l'évaluation de la qualité des revues et de leur comité éditorial. Aussi, la mise en place de services innovants aux lecteurs permettrait d'augmenter le taux de consultation et la visibilité des publications locales.
3. L'accompagnement des comités éditoriaux et l'exigence du respect des standards de qualité internationale faciliteraient l'indexation des revues locales dans les bases de données internationales.

Par ailleurs, pour envisager l'open access aux publications scientifiques en SHS, nous avons besoin de données sur les contraintes qui entravent le rayonnement des éditeurs marocains, notamment, les ressources humaines et matérielles, la rentabilité des revues scientifiques, ainsi que les financements nécessaires en amont pour basculer vers l'openaccess. Nous envisageons d'explorer ce volet dans une prochaine étude, en mobilisant la méthode qualitative de recueil de données. Cette démarche nous permettra d'analyser les attitudes des éditeurs et des gestionnaires de la recherche scientifique quant à la problématique de l'openaccess.

Bibliographie

- ALIZON, S. (2018). Inexpensive Research in the Golden Open-Access Era. *Trends in ecology & evolution*, 33(5), 301-303.
- ANTELMAN, K. (2006). Self-archiving practice and the influence of publisher policies in the social sciences. *Learned Publishing*, 19(2), 85-95.
- BEAUDRY, G. (2010). La communication scientifique directe : un nouveau champ éditorial. *Hermès, La Revue*, (2), 51-57.
- BEN ROMDHANE, M. ; Ouerefelli, T. (2013). L'offre des archives ouvertes dans le monde arabe : recensement et évaluation. En ligne : http://archivesic.ccsd.cnrs.fr/sic_00829058
- CARLSON, S. (2015). An Assessment of Institutional Repositories in the Arab World. En ligne : <http://www.dlib.org/dlib/may15/carlson/05carlson.html>
- CHARTRON, G. (2014). Open access et SHS: Controverses. *Revue européenne des sciences sociales*, 52(1), 37-63.
- CUSCHIERI, S. (2018). WASP: Is open access publishing the way forward? A review of the different ways in which research papers can be published. *Early human development*. En ligne :
- CUSCHIERI, S. (2018). WASP: Is open access publishing the way forward? A review of the different ways in which research papers can be published. *Early human development*.
- ENNAS, G., & DI GUARDO, M. C. (2015). Features of top-rated gold open access journals: An analysis of the scopus database. *Journal of Informetrics*, 9(1), 79-89.
- FARIDA, I., TJAKRAATMADJA, J. H., FIRMAN. (2015). A conceptual model of Open Access Institutional Repository in Indonesia academic libraries: Viewed from knowledge management perspective. *Library Management*, 36(1/2), 168-181.

- FERWERDA, E (2013). Open access in humanities and social sciences. In 8th Munin Conference n Scholarly Publishing, Tromso, Norway, 25–26 November 2013.
- FONDATION DU ROI ABDUL AZIZ AL-SAOUD POUR LES ETUDES ISLAMIQVES ET LES SCIENCES HUMAINES (2017). Edition et Livre au Maroc : Rapport annuel sur l'état de l'édition et du livre dans les domaines de la littérature, sciences humaines et sociales 2015/2016. En ligne : http://www.fondation.org.ma/publications/rapport_fr_2017.pdf
- GDOURA, W., BOUAZZA, A., IBNLKHAYAT, N. (2009). Attitudes des universitaires arabes à l'égard des revues en ligne et des archives ouvertes: cas du Maroc, d'Oman et de la Tunisie. In *Revue tunisienne de communication*, N°5152, pp.9-44.
- GROSS, J., RYAN, J. C. (2015). Landscapes of research: Perceptions of Open Access (OA) publishing in the arts and humanities. *Publications*, 3(2), 65-88.
- Groupement français de l'industrie de l'information (2009). L'édition scientifique française en SHS – Rapport de synthèse. En ligne : <https://www.gfii.fr/uploads/docs/l-edition-scientifique-francaise-en-sciences-sociales-et-humaines.pdf>
<http://unesdoc.unesco.org/images/0018/001883/188395f.pdf>
<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0020961>
- JACKSON, R., & RICHARDSON, M. (2014). Gold open access: the future of the academic journal?. In *The Future of the Academic Journal* (Second Edition) (pp. 223-248).
- LRHOUL, H. (2017). La production scientifique des chercheurs de la faculté de médecine et de pharmacie de Casablanca : mesures, cartographie et enjeux du libre accès (Doctoral dissertation, Conservatoire national des arts et métiers-CNAM). En ligne : <https://tel.archives-ouvertes.fr/tel-01815122>
- LRHOUL, H., BACHR, A. (2013). Le chercheur universitaire marocain et le libre accès à l'information scientifique et technique. In *Actes du Colloque «L'université à l'ère du numérique»*. Institut Universitaire de la Recherche Scientifique, Rabat le (Vol. 1).
- UNESCO (2010), Rapport mondial sur les sciences sociales. Division dans les savoirs Éditions de L'UNESCO. EN ligne : <http://unesdoc.unesco.org/images/0018/001883/188395f.pdf>

Analyse des articles scientifiques et libre accès à la littérature : Quels Bénéfices pour le Développement de la Science

*Literature Mining and Open Access: advantages for Science
Development*

SEKKAL Houda¹, AMROUS Naila^{1,2}, BENNANI Samir¹

RIME TEAM-Networking, Modeling and e-Learning- LRIE Laboratory- Research in
Computer Science and Education Laboratory

1 Mohammadia School Engineers (EMI) - MOHAMMED V UNIVERSITY IN RABAT
Agdal AV. Ibn Sina

Agdal Rabat BP. 765 Morocco

2 Ecole des Sciences de l'Information (ESI), Rabat, Maroc. Laboratoire "Mediation -
Information - Knowledge - Society" (MIKS).

Résumé

Le libre accès permet à la communauté scientifique d'accéder aux productions scientifiques sans restriction. Plusieurs techniques et méthodes ont été développées par différents chercheurs afin de permettre l'analyse des articles scientifiques pour faciliter le repérage des résultats et des informations pertinentes. S'inscrivant dans une approche qualitative, l'objectif de cet article est de mettre en exergue les bénéfices du libre accès aux articles scientifiques pour le développement de la science. Le présent travail de recherche porte sur l'analyse des travaux qui traitent l'extraction de l'information présente dans les articles scientifiques. Le résultat de cette étude permet de dégager le lien existant entre la promotion du libre accès et le développement de la science.

Mots-clés. Libre accès, revue de littérature systématique, fouille textuelle, traitement de langage naturel, Fouille de la littérature.

Abstract

Open access allows researchers to freely access to scientific productions without barriers. Different tools were developed by researchers to automatically mine scientific articles in order to extract rich scientific information about a specific subject. The goal behind this project is to find the impact and the benefit of the open access to scientific publications. To do so, we analyzed articles that propose methods to mine scientific literature. The results will allow us to extract advantages behind open access to scientific publications from a science development point of view.

Keywords. Open access, systematic literature review, text mining, natural language processing, Literature mining.

Introduction

La recherche scientifique a connu un développement exponentiel durant les dernières décennies. Selon P. Accuosto et al. (2017), le volume de publications scientifiques double chaque neuf an. Ce progrès remarquable de publications constitue un vrai défi devant les chercheurs. En effet, chaque chercheur doit lire un maximum de travaux de ses pairs afin de rester à jour dans son domaine de recherche et de pouvoir participer à son tour au progrès scientifique. La lecture permanente et régulière des travaux scientifiques est une activité qui demande beaucoup de temps ainsi qu'un investissement financier de la part des chercheurs. Ces deux facteurs forment un vrai défi et risquent d'entraver le progrès scientifique (Lutz Bornmann et Rüdiger Mutz, 2015).

Les articles scientifiques sont parmi les documents textuels les plus précieux de la connaissance humaine (P. Accuosto et al., 2017), C'est une mine qui déborde de savoir et qui permet à l'humanité de développer les technologies dans tous les domaines de la vie. Selon M. Funamori (2017), les articles scientifiques encapsulent des données inédites ne se trouvant dans aucun autre document et qui sont très importantes pour les chercheurs. Leurs extractions et leurs analyses sont donc indispensables. Sauf que l'accès au texte intégral de ces articles scientifiques est souvent soumis à des restrictions de la part de la majorité des éditeurs : frais d'abonnement aux revues, frais d'accès au texte intégral, etc.

Avant d'entamer sa recherche, chaque chercheur est dans l'obligation d'étudier les articles en rapport avec son domaine de recherche afin de recenser les dernières avancées et pouvoir développer son domaine. Pour cela, il est indispensable d'offrir à la communauté scientifique des solutions efficaces et fiables permettant un accès plus rapide à la connaissance.

L'Open Access est l'une des solutions pour garantir un accès libre aux données scientifiques. Selon l'UNESCO (2015), l'accès libre aux publications scientifiques est leurs « mise à disposition gratuite sur l'Internet public, permettant à tout un chacun de lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces articles, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale, sans barrière financière, légale ou technique autre que celles indissociables de l'accès et l'utilisation

d'Internet ». Or, malgré les efforts fournis par les professionnels et les communautés scientifiques pour promouvoir le libre accès, le chercheur est toujours confronté à des barrières financières et légales quant à l'exploitation des articles scientifiques.

Il existe différentes formes d'utilisation des articles scientifiques : la forme la plus commune est celle de la lecture des articles par le chercheur pour rédiger une revue de littérature ou pour être informé des dernières explorations en lien avec son domaine. Une autre forme d'utilisation des articles scientifiques existe, celle du traitement automatique du texte scientifique. Cette pratique fait appel à des techniques de fouille et d'analyse du texte (Text Mining) et du traitement du langage naturel (Natural Language Processing). L'objectif de cette analyse est l'extraction de données pertinentes des articles scientifiques ce qui exige la disponibilité du texte intégral de chaque article à analyser. L'analyse du texte scientifique ou Literature Mining est un domaine de recherche en pleine évolution qui est apparu avec l'apparition des articles scientifiques en version électronique. Ce domaine étant très important pour la recherche scientifique, est confronté à une problématique principale, celle de la restriction d'accès aux articles scientifiques. L'accès libre au texte intégral des travaux scientifiques permet donc aux chercheurs de traiter automatiquement les articles scientifiques afin d'en extraire les informations utiles et pertinentes dans certains domaines de recherche. Or, après plusieurs recherches, nous avons conclu qu'aucun travail n'expose les pratiques et les utilisations du texte intégral des articles scientifiques dans le domaine de l'analyse textuelle des articles scientifiques (Literature mining) par les chercheurs. Nous voulons donc, à travers ce travail répondre à la question suivante qui constitue notre problématique de recherche : La restriction aux articles scientifiques peut-elle entraver l'évolution et le développement de la science ? Réciproquement, la réponse à cette question permettra de conclure que le libre accès aux articles scientifiques est un moyen incontournable pour accélérer l'évolution scientifique à travers la mise à disposition des travaux scientifiques aux chercheurs.

Basé sur une approche qualitative, le présent travail a pour objectif de déceler l'importance de la disponibilité des publications scientifiques en libre accès pour le bon développement de la recherche scientifique. La

recherche qualitative désigne selon Taylor et Bogdan (1984) la production et l'analyse des données descriptives, telles que les paroles écrites ou dites ainsi que le comportement observatoire des personnes. Avec l'avènement d'Internet et la publication des articles scientifiques en format électronique, un nouveau domaine de recherche est apparu, celui de l'analyse des articles scientifiques par l'application des techniques de Texte Mining et du traitement du langage naturel (NLP). Nombreux sont les auteurs qui ont contribué à l'étude de ce sujet en proposant des techniques et méthodes d'extraction de l'information utiles et riche des articles scientifiques.

Ce travail s'intéresse plus particulièrement à l'analyse des travaux qui ont proposé des méthodes d'extraction de l'information du texte intégral des articles scientifiques et de dégager l'apport du Libre accès quant à la résolution de leurs problématiques respectives.

Les questions de recherche que nous souhaitons traiter sont les suivantes :

Q1: Quels sont les objectifs derrière l'analyse des articles scientifiques ?

Q2: Quels sont les domaines concernés par l'analyse des articles scientifiques ?

Q3: Quelles sont les différentes techniques d'analyses d'articles scientifiques ?

Q4: Quelles sont les bases de données utilisées pour collecter les articles scientifiques ?

Q5: Quel est l'apport de la Littérature Mining pour la promotion du libre Access ?

1 Travaux similaires au présent travail de recherche

Plusieurs travaux ont été réalisés dans le contexte de la fouille de données scientifiques (ou Literature Mining) mais rares sont ceux qui ont révélé le lien entre l'extraction de l'information des articles scientifiques et le libre accès aux productions scientifiques. Lynette Hirschman (et al, 2002) a présenté un aperçu des travaux sur l'analyse des articles scientifiques dans le domaine biomédical. Dans son travail de recherche, l'auteur a évalué les travaux déjà réalisés dans ce contexte et a essayé de relever les défaillances et les problèmes rencontrés dans le processus d'extractions de données dans le domaine biomédical. Hormis ces travaux, nous estimons qu'aucun travail n'a été présenté sous l'ongle du libre accès aux articles scientifiques.

2 Méthodologie suivie pour répondre aux questions de recherches

Le présent travail a pour but d’apporter des éléments de réponses aux questions de recherche précitées. Nous avons choisi, pour répondre à ces questions, de suivre la méthode proposée par Kitchenham et Charters (2007) afin de donner un aperçu des travaux dirigés sur l’extraction de la connaissance des articles scientifiques. Nous allons pour cela adopter les étapes suivantes : Premièrement nous allons présenter la stratégie de recherche à savoir : les mots clés utilisées pour interroger les bases de données digitales ainsi que les sources et les bases de données interrogées puis détailler le processus de recherche. Deuxièmement nous allons expliquer la procédure suivie pour sélectionner les articles jugés pertinents par rapport à nos questions de recherche. Troisièmement, on va procéder à l’analyser et l’extraction des données. Les étapes adoptées sont détaillées dans la partie qui suit :

2.1 La stratégie de recherche dans les bases de données

La stratégie de recherche vise à détailler le processus de recherche suivi pour collecter l’ensemble d’articles pertinents à la présente problématique de recherche qui traitent l’analyse textuelle des articles scientifiques. Nous avons commencé par choisir les termes qui vont constituer les mots clés de notre équation de recherche, ensuite nous avons choisi les bases de données ayant le plus grand volume d’articles et de journaux indexés pour garantir un maximum de pertinence au niveau de la sélection des articles à analyser.

2.1.1 Mots clés de recherche employés dans les bases de données des articles scientifiques

Afin d’assurer un maximum de pertinence dans notre processus de sélection d’articles, nous avons utilisé plusieurs équations de recherche pour chaque base de données qui sont détaillées au niveau du tableau suivant :

Base de données interrogée	Equation de recherche
ACM Digital Library	scientific AND (article OR publication OR paper) AND (information OR data) AND (extraction OR mining OR analysis)
	(“scientific publication”) AND (“scientific article”) AND (mining OR analysis extraction) AND (“full text”)

Google Scholar	Mining AND scientific AND article
IEEE Xplore	("papers mining") AND ("journals mining") AND "articles mining
	Full AND text AND articles AND mining
	scientific AND (article OR publication) OR paper) AND (information OR data OR text) AND (extraction OR mining OR analysis OR "text mining")
	Scientific AND Article AND information AND extraction

Tableau 1. Les équations de recherches utilisées pour chaque base de données.

Les opérateurs booléens ont été utilisés afin d'apporter plus de précision à nos résultats de recherches au niveau des bases de données. Le choix d'utilisation de chaque opérateur est détaillé comme suit :

- L'utilisation de l'opérateur AND : pour joindre les principaux termes qui constituent nos questions de recherche.
- L'utilisation de l'opérateur OR : pour joindre les synonymes.
- L'utilisation du symbole " " : pour chercher une suite de termes continues.

Cette première étape nous a permis de collecter 143 articles scientifiques.

2.1.2 Le processus de collecte des articles scientifiques

Nous avons adopté deux formes de sélections pour la collecte d'articles scientifiques :

- Dans une première étape, nous avons collecté les articles directement des bases de données mentionnés dans le tableau 1 en utilisant les équations de recherches indiquées.
- Dans une deuxième étape, et après avoir étudié les articles scientifiques jugés pertinents selon les critères détaillés au niveau du tableau 2, nous avons établi une autre sélection, celle des articles candidats à être pertinents et se trouvant au niveau des listes de références desdits articles.

2.2 Procédure de sélection des articles scientifiques collectés

Pour répondre à nos questions de recherche, il est indispensable d'avoir une sélection d'articles scientifiques pertinente. Plusieurs critères ont été

élaborés afin de choisir les articles scientifiques adéquats. Le **tableau 2** présente deux types de critères comme l'ont mentionné Kitchenham et Charters (2007) : des critères d'inclusion et des critères d'exclusion. Pour qu'un article soit accepté pour l'analyse, ce dernier doit satisfaire au moins un critère d'inclusion et ne doit satisfaire aucun critère d'exclusion.

Critères d'inclusion	Critères d'exclusion
<ol style="list-style-type: none"> 1- L'article présente une méthode d'extraction des informations ou des données des articles scientifiques. 2- Articles scientifiques datant de 2014 à 2018 (5 dernières années) 	<ol style="list-style-type: none"> 1- Les articles où il n'y a pas d'expérience faite sur un ensemble d'articles scientifiques ; 2- Les articles qui traitent l'extraction des métadonnées et résumé uniquement ; 3- Les articles avec moins de 4 pages ; 4- Les articles écrits avec une langue autre que l'Anglais et le Français ; 5- Les revues de littérature ; 6- Les articles qui ne présentent pas une analyse des résultats obtenus.

Tableau 2. Les critères d'inclusion et d'exclusion des articles scientifiques.

La démarche de notre sélection d'articles s'est déroulée en respectant les critères indiqués dans le tableau 2, nous avons commencé par sélectionner des 143 articles collectés, les articles scientifiques qui respectent les critères d'inclusions. Ensuite, dans une deuxième étape, nous avons entamé la procédure d'exclusion des articles selon les critères d'exclusion. On a commencé par éliminer les articles avec moins de 4 pages, ceux qui sont rédigés dans une autre langue autre que le français et l'anglais et les revues de littératures. Vient ensuite l'étape qui consiste en la lecture du texte intégral de chaque article afin de vérifier le reste des critères d'exclusion à savoir : les critères (1), (2) et (6). Le tableau 3 est le résultat de notre sélection et constitue l'ensemble des articles scientifiques qui feront l'objet de notre analyse.

Titre de l'Article	Année de	Editeur	Conférence ou journal d'indexation
--------------------	----------	---------	------------------------------------

	public a- tion		
AppTechMiner- Mining Applications and Techniques from scientific articles	2017	ACM	WOSP 2017
Multi-level mining and visualization of scientific text collections	2017	ACM	WOSP 2017
Scientific Monitoring by Mining Scientific Papers	2014	ACM	MEDES'14
Challenges in adapting text mining for full text articles to assist pathway curation	2014	ACM	BCB '14,
Business intelligence in banking: A literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation	2014	Elsevier	Expert Systems with Applications 42
A Method for Thematic and Structural Visualization of Academic Content	2017	IEEE	17th International Conference on Advanced Learning Technologies
Convolutional Neural Networks for Figure Extraction in Historical Technical Documents	2017	IEEE	14th IAPR International Conference on Document Analysis and Recognition
Extracting Code Segments and Their Descriptions from Research Articles	2017	IEEE/ACM	International Conference on Mining Software Repositories (MSR)
Sentiment Analyzer: Analysis of Journal Citations from Citation Databases	2014	IEEE	
Exploring Concept Graphs for Biomedical Literature Mining	2015	IEEE	BigComp 2015
Identification of Protein Interaction Methods from Biomedical Literature	2014	IEEE	

Analyzing the Usage of Standards in Radiation Therapy Clinical Studies	2017	IEEE	
Literature Search Tool for the Extraction of Disease-Associated Genes using Frequent Itemset Mining	2016	IEEE	International Conference on Computing, Analytics and Security Trends (CAST)
A Corpus-based Analysis of the Scientific RA Genre and RA Introduction	2014	IEEE	International Conference on Advanced Mechatronic Systems
Extracting Semantic Relations for Scholarly Knowledge Base Construction	2018	IEEE	12th IEEE International Conference on Semantic Computing
Keyphrase extraction. Abstracts instead of full papers	2014	IEEE	25 th International Workshop on Database and Expert Systems Applications
RLIMS-P 2.0: A Generalizable Rule-Based Information Extraction System for Literature Mining of Protein Phosphorylation Information	2015	IEEE	IEEE/ACM transactions on computational biology and bioinformatics
Towards a Hybrid Human-Computer Scientific Information Extraction Pipeline	2017	IEEE	13th International Conference on eScience

Tableau 3. Les articles scientifiques sélectionnés selon les critères d'inclusion et d'exclusion.

Le résultat de la sélection des articles selon les critères d'inclusion et d'exclusion a généré 18 articles pertinents par rapport à notre problématique de recherche et qui vont faire l'objet de notre analyse.

2.3 Extraction de données du texte intégral des articles scientifiques pour répondre aux questions de recherche

Afin de répondre à nos questions de recherches, les articles collectés figurant dans le **tableau 3** vont être analysé manuellement afin d'extraire les données suivantes : (1) l'objectif de la méthode, (2) le domaine d'application de la méthode présentée, (3) recensement des techniques utilisées, (4) la source utilisée par l'auteur pour collecter les articles à traiter, (5) les résultats escomptés.

Questions de recherche	Données à extraire de l'article
1 Q1 : Quels sont les objectifs derrière l'analyse des articles scientifiques ?	(1) L'objectif de la méthode présentée par l'article.
2 Q2 : Quels sont les domaines concernés par l'analyse des articles scientifiques ?	(2) Le domaine d'application de la méthode présentée.
3 Q3 : Quelles sont les différentes techniques d'analyses d'articles scientifiques ?	(3) Les techniques utilisées pour analyser les articles.
4 Q4 : Quelles sont les bases de données utilisées pour collecter les articles scientifiques ?	(4) La source que l'auteur a utilisé pour collecter le texte intégral des articles scientifiques.
5 Q5 : Quel est l'apport de la littérature mining pour la promotion du libre Access ?	(5) Les principaux résultats escomptés

Tableau 4. Les données à extraire qui vont permettre de répondre aux questions de recherche.

Il est à noter que les bases de données desquelles a eu lieu la récupération du texte intégral des articles scientifiques (4) ainsi que les résultats des méthodes proposées par chaque article (5) vont constituer des éléments de réponse à la question Q5 et qui constitue l'objet du présent travail de recherche.

3 Résultats obtenus et discussions

Nous nous sommes penchés lors de la réalisation du présent travail sur les travaux qui s'intéressent à l'analyse du texte intégral des articles scientifiques. L'objectif est de savoir si la restriction d'accès au texte intégral peut entraver le processus d'extraction automatique de l'information scientifique d'une part, et de recenser les bases de données utilisées par les auteurs offrant un accès libre aux articles d'autre part. Il

est également important de savoir quels sont les domaines qui utilisent le plus la technique de récupération automatique de l'information des articles scientifiques et quelles sont les types de données recherchées dans les travaux scientifiques. Le but est de trouver le lien existant entre la disponibilité de ces productions scientifiques et l'évolution des domaines de recherche. Dans ce contexte, nous avons analysé chaque article en répondant aux cinq questions de recherches.

3.1 Les domaines d'applications et les objectifs derrière l'analyse automatique de la littérature (Q1 et Q2)

L'extraction et l'analyse des informations dans les articles intéressent un grand nombre de chercheurs issus de domaines différents. Il est alors important de présenter les objectifs de chaque méthode d'extraction de l'information présentée par les articles ainsi que les domaines d'application.

La figure 1 présente les domaines d'application de notre sélection d'articles scientifiques. D'après la même figure (figure 1), deux domaines sont prédominants : l'informatique (computer science) et le domaine biomédical. Ensuite viennent les domaines suivants : Les banques, l'analyse des sentiments, le domaine Optical et informatique matérielle. Dans le domaine du computer science, se déclinent plusieurs sous domaines que traite notre sélection d'article. On trouve en premier lieu le domaine de traitement du langage naturel (Natural LanguageProcessing : NLP). D'après l'analyse des articles scientifiques, c'est le domaine où ont été effectués le plus de travaux sur l'extraction automatique des informations des articles scientifiques. C'est un résultat qui est évident puisque l'extraction de l'information des articles scientifiques repose principalement sur les techniques du NLP car les articles scientifiques sont en langage naturel. Dans le même contexte du domaine informatique, on trouve comme sous domaine traité par les articles le Génie logiciel (software Engineering). Parmi les travaux traités également, on trouve un article de Chatterjee (et al, 2017) proposant une méthode d'extraction des codes d'algorithmes pour les réutiliser dans d'autres projets.

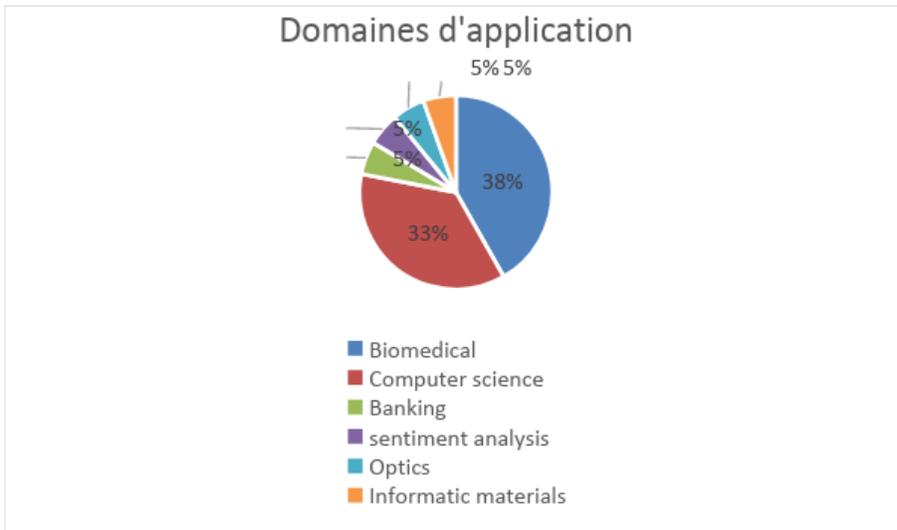


Figure 1. *Les domaines des articles analysés par les auteurs.*

Les tendances générales en termes d'objectifs d'extraction automatique d'informations des articles scientifiques recensés à partir de notre sélection sont :

- La construction d'une base de connaissance.
- L'exploration sémantique et linguistique du contenu ;
- L'aide à l'élaboration des revues de littérature ;
- L'extraction de codes de programmation ;
- Génération d'un graphe de connaissance d'un domaine particulier ;
- Extraction des phrases clés (keyphrases) d'un article scientifique ;
- Extraction d'éléments biomédicaux.
- Création d'un graph de concept ;
- Détermination des tendances d'utilisation de différentes méthodes et techniques dans un domaine particulier ;
- Extraction automatique des figures et des images ;
- Classification des citations des articles scientifiques.

2.1. Les techniques utilisées pour analyser les articles scientifiques (Q3)

Pour extraire les données des articles scientifiques, chaque auteur a proposé une méthode afin de répondre au mieux à sa problématique tout en respectant le contexte des articles à traiter : le domaine, la performance des outils à utiliser, ... etc. Afin de répondre à notre question de recherche (Q3) relative aux techniques utilisées par notre sélection d'articles pour réaliser l'extraction de l'information, nous avons choisi d'établir une cartographie (figure 2) représentant les différentes techniques utilisées par chaque article de la sélection en fonction de la date de publication dudit articles. Les petits nœuds en rose représentent les techniques du NLP les moins cités, les plus gros nœuds en vert représentent les techniques du NLP les plus cités par les articles sélectionnés. Les petits nœuds en bleu représentent les années de publication de notre sélection d'articles. Cette cartographie va nous permettre de déterminer d'une part, l'évolution d'utilisation des techniques de texte mining et du NLP (Natural Language Processing) en fonction du temps, et d'autre part, de recenser les techniques les plus utilisées par les différents auteurs.

Pour réaliser cette cartographie, nous avons utilisé le logiciel de cartographie GEPHI (Bastian M et al., 2009), un logiciel libre qui permet l'analyse et la visualisation des réseaux de données. Nous avons créé une liste de techniques pour chaque article (les techniques utilisées étaient mentionnées dans la partie méthodologie de chaque article scientifique) en fonction des date de publication. Ces données ont été chargées sur Gephi. La cartographie réalisée (figure 2) nous montre les techniques de texte mining et de NLP les plus utilisées par notre sélection d'articles en fonction du nombre d'occurrence (occurrence count) de chaque technique dans l'ensemble des articles. Nous remarquons ainsi que les deux techniques de classifications : NamedEntity Recognition¹ et Regular expression² sont les plus utilisées durant les années : 2014, 2016 et 2017. Nous pouvons donc conclure que ces techniques sont d'actualité et que les chercheurs du domaine du NLP continuent toujours de les appliquer dans leurs travaux de recherche. Cette cartographie nous permet également de détecter les méthodes classiques fréquemment utilisées pour le traitement du langage

¹ Selon Jurafsky (2018), c'est de rechercher chaque mention de reconnaissance d'une entité nommée dans le texte et étiqueter son type.

² Selon Jurafsky (2018), c'est une notation algébrique pour caractériser un ensemble de chaînes.

naturel à savoir : tokenisation³, lemmatisation⁴, stemming⁵, etc. Cela permet aux nouveaux chercheurs désirant apprendre les techniques du NLP de cerner les principales méthodes utilisées. La cartographie nous permet également de voir les techniques et méthodes utilisées selon l'année de publication des articles afin de dégager les tendances de techniques de texte mining et de traitement des articles scientifiques par année. Selon la même figure (figure 2) nous ne pouvons pas constater de grande différences au niveau des tendances car les dates sont très rapprochées (de 2014 à 2018). Par contre les résultats de la cartographie peuvent être intéressants si on choisit une tranche de date assez large (les dix dernières années par exemple).

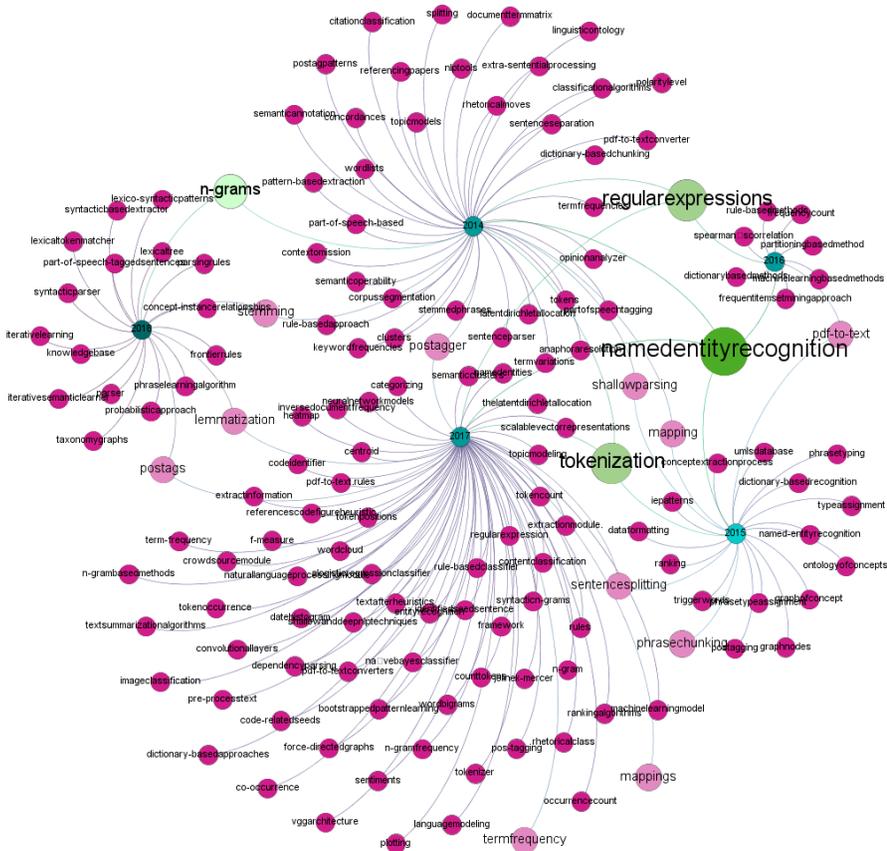
Figure 2. *Cartographie des techniques d'extraction classées par date de publication des articles*

³ Selon Jurafsky (2018), elle consiste à séparer les mots ou les marquer du texte courant.

⁴ Selon Jurafsky (2018), la tâche détermine que deux mots ont la même racine, malgré leurs différences.

⁵ Selon Jurafsky (2018), fait référence à une version plus simple de la lemmatisation dans laquelle nous supprimons principalement les suffixes à la fin du mot.

3.2 Les bases de données utilisées par les auteurs pour récupérer le texte intégral des articles scientifiques analysés (Q4)



Bases de données utilisée	Le volume traité	La période du Corpus utilisé
ACL Anthology Network dataset :	21,213 articles	1965 à 2013

The spanish society for Natural language Processing (SEPLN)	505 PDF	2008 à 2016
Le corpus a été constitué grâce à un projet en commun entre : l'agence universitaire de la francophonie, l'éditeur John LibbeyEurotext et l'Université de Paris-Sorbonne.	674 documents	-
PharmaGKB.	35 articles	-
ACM, Elsevier, INFORMS, Association for IS	219 articles	-
IEEE Xplore et the international review of research in open and distributed learning.	20 articles scientifiques	
ACM et IEEE xplore	8 articles scientifiques	2011 à 2015
Google scholar	10 papiers	-
PubMed.	237 articles	-
Bio Creative dataset	300 articles	-
MEDLINE	531 articles	De 2010 à 2015.
PubMed central	-	-
Journal of Optics	50 articles	De 1997 à 2001
CiteseerX	10 K d'articles	2004- 2014
ACM Digital Libraries (financé par le Gouvernement Russe.)	244 articles	-
Pub Med Open Access Subset, et les résumés de la DB : Medline.	-	-
Poly Info digital data base	6090 Articles	2006-2016

Tableau 5. Les bases de données utilisées par les auteurs d'articles.

Au niveau du tableau 5, les cases mentionnées avec un tiret (-) signifient qu'il n'y a aucune information trouvée.

La réponse à la question Q4 nous permet de déterminer les sources ou les bases de données desquels les chercheurs ont récupéré le texte intégral des articles scientifiques.

Si les auteurs confirment dans leurs articles que le problème de disponibilité des articles scientifiques n'entrave pas leurs processus de recherche et

d'analyse d'articles scientifiques, on peut donc conclure que la question du libre accès est loin d'être une problématique pour l'évolution de la science. Or, d'après notre traitement de la sélection d'articles (18 articles scientifique) qui concernent le traitement du texte intégral des articles scientifiques à travers l'application des techniques détaillées dans la partie relative à la procédure de sélection des articles du présent travail, nous constatons que sur les 18 articles, les auteurs de 10 articles confirment avoir choisi de récupérer le texte intégral dans des bases de données offrant des articles de journaux ou de conférences en libre accès. Ceci ne peut être qu'un indicateur important montrant la nécessité d'avoir un accès libre et gratuit aux publications scientifiques surtout dans le domaine de l'extraction automatique des données scientifiques afin de permettre l'évolution du domaine. Les autres articles ayant récupéré le texte intégral dans les bibliothèques digitales offrant un accès restreint sont principalement des travaux réalisés dans le cadre de partenariat entre des laboratoires de recherche et des entités finançant les projets.

3.3 Les résultats et les défis de l'analyse automatique des articles (Q5)

Pour répondre à la question sur la relation entre la littérature mining (l'analyse des publications scientifiques avec les techniques du texte mining et le NLP) pour la promotion du libre Access, nous pouvons constater d'après les résultats d'analyse de la sélection d'articles traitant l'analyse textuelle des articles scientifiques que la disponibilité du texte intégral des articles à traiter est une nécessité et non pas un choix. En effet, plusieurs auteurs (Manabu Torii et al., 2015), surtout du domaine biomédical, ont confirmé que l'analyse du texte intégral au lieu du résumé est très importante car le texte intégral contient des informations riches ne figurant pas au niveau du résumé. D'autre part, selon Roselyne B. Tchoua et al. (2017) le domaine du Littérature Mining est un domaine émergent mais reste peu mature malgré sa grande importance puisqu'il permet de récupérer des informations précieuses rapidement des publications scientifiques. Pour le promouvoir et le développer, il est nécessaire de réfléchir à des méthodes plus efficaces, rapides et simples permettant l'extraction des données scientifiques pour faciliter l'avancement scientifique. Or de tel progrès ne peuvent être réalisés sans la disponibilité du texte intégral des articles scientifiques en libre accès afin d'effectuer plus

d'expérimentations dans le sujet et encourager les chercheurs à le développer sans réfléchir aux coûts générés.

4 Conclusion et discussions

Ce travail de recherche présente les travaux réalisés dans le domaine de l'analyse textuelle des articles scientifiques. L'objectif est de déceler l'importance de la disponibilité des publications scientifiques en libre accès pour le développement de la recherche scientifique. A travers cette recherche, nous avons présenté des éléments de réponses à cinq questions de recherches à savoir : (Q1) : Quels sont les objectifs derrière l'analyse des articles scientifiques ? ; (Q2): Quels sont les domaines concernés par l'analyse des articles scientifiques ?; (Q3) : Quelles sont les différentes techniques d'analyses d'articles scientifiques ? (Q4): Quels sont les bases de données utilisées pour collecter les articles scientifiques ?; (Q5) : Quel est l'apport de la Littérature Mining pour la promotion du libre Access ?

L'extraction automatique des données scientifiques des articles scientifiques est une technique ayant une importance cruciale pour la recherche. En effet, selon le domaine de recherche, pouvoir récupérer l'information scientifique sur un aspect particulier permet aux auteurs de déceler les problèmes et les défis dudit domaine, les résultats de recherches, les techniques les plus efficaces, etc. A titre d'exemple, parmi les résultats escomptés, nous avons remarqué que dans beaucoup de travaux, la récupération automatique de l'information scientifique permet de faciliter la rédaction des revues de littérature ainsi que de créer des graphes et des bases de connaissances qui est une étape importante dans le domaine de l'Intelligence Artificielle. Dans le domaine médical par exemple, l'analyse des articles scientifiques est primordiale et très utilisée pour l'extraction des éléments biomédicaux.

Nous avons conclu également que sur les 143 articles sélectionnés traitant l'analyse textuelle des articles scientifiques, uniquement 18 articles traitent l'analyse du texte intégral des articles scientifiques. Deux principales causes sont à constater : une première cause est liée à la difficulté du processus d'extraction et du traitement du texte intégral des articles scientifiques suite à leurs complexités comme l'ont annoncé Roselyne B (et al, 2017) dans leur article. D'après ce dernier, le domaine d'analyse des articles scientifiques a besoin de plus d'efforts de la part des chercheurs

pour le développer. Dans ce cas, les chercheurs ont besoin donc d'accès libre aux articles pour pouvoir développer des systèmes d'extraction plus sophistiqués. Une deuxième cause est liée à l'indisponibilité du texte intégral des articles scientifiques. En effet, la partie du présent travail liée aux bases de données utilisées par les auteurs pour récupérer le texte intégral des articles scientifiques met en exergue les bases de données utilisées par les auteurs des 18 articles sélectionnés pour récupérer le texte intégral. On constate donc que 4 articles sur 18 ont utilisés des éditeurs payant contre 9 articles qui ont fait appel à des bases de données en open accès. Nous avons constaté également qu'en plus de ces 4 articles, un article qui est le résultat d'un projet de recherche financé, a utilisé une base de données payante. Il est également constaté que les autres travaux parmi ces 18 articles ayant utilisé des bases de données offrant un libre accès aux articles scientifiques ont pu traiter jusqu'à 20 000 articles à la fois. Ce qui justifie que le libre accès aux articles scientifiques encourage les chercheurs à développer leurs domaines de recherche.

Il est à noter également que certaines communautés scientifiques mettent à la disposition des chercheurs des bases de données d'articles scientifiques en accès libre pour leur permettre de faire des évaluations et tester leurs méthodologies d'extraction surtout dans le domaine du traitement du langage Naturel (NLP) comme les bases de données : ACL Anthology Network dataset et The Spanish Society for Natural Language Processing (SEPLN).

Le résultat de l'étude a permis de dévoiler le lien entre la disponibilité de libre accès aux articles scientifiques et le développement de la science. Ce dernier a montré que la disponibilité du texte intégral des publications scientifiques encourage les chercheurs à développer ce domaine émergent mais pas encore mature qu'est le domaine de l'analyse textuelle de la littérature (Littérature Mining). Ce travail de recherche permet donc de divulguer un aspect stratégique du libre accès celui du développement de la connaissance. Le libre accès est un moyen non seulement favorable pour permettre aux chercheurs d'accéder librement au texte intégral des publications scientifiques mais aussi d'explorer la connaissance pour la développer.

Les résultats ont bien montré la grande importance du traitement du texte intégral des articles scientifiques dans beaucoup de domaines surtout le domaine médical. Il est donc primordial de mettre à la disposition des chercheurs un accès libre aux articles scientifiques pour les encourager à développer le domaine de l'analyse textuelle de la littérature dont les avantages sont importantes. Parmi les axes de recherche futures qui peuvent être développés est celui de la comparaison entre les résultats d'analyse textuelle des résumés avec ceux de l'analyse textuelle du texte intégral des articles scientifiques dans un domaine précis sur une tranche de date assez large (20 dernières années). Cela permettra de confirmer encore plus la nécessité de mettre en libre accès les travaux et les résultats scientifiques.

Références

- ACCUOSTO, Pablo; RONZANO, Francesco; FERRÉS, Daniel; SAGGION, Horacio. (2017). Multi-level mining and visualization of scientific text collections: *Exploring a bi-lingual scientific repository*. In *6th International Workshop on Mining Scientific Publications*. Canada: ACM, 2017. 8p. pp. 9–16.
- Agarwal, S; Singh, M; Dan, S; Goyal, P; Mukherjee, A. (2017). AppTechMiner: *Mining Applications and Techniques from Scientific Articles*. In *6th International Workshop on Mining Scientific Publications*. Canada: ACM, 2017. 8p.
- Al-Zaidy, R. A; Giles, C. L (2018). Extracting Semantic Relations for Scholarly Knowledge Base Construction. In *25th International Workshop on Database and Expert Systems Applications*. IEEE, 2018. 8p. pp. 56–63.
- Amigud, A; Arnedo-Moreno, J; Daradoumis, T; Guerrero-Roldan, A. E (2017). A Method for Thematic and Structural Visualization of Academic Content. In *17th International Conference on Advanced Learning Technologies*. IEEE, 2017. 5p. pp. 230–234.
- Bastian, M; Heymann, S; Jacomy, M (2009). Gephi: *an open source software for exploring and manipulating networks*. In *International AAAI Conference on Weblogs and Social Media*. France: Association

- for the Advancement of Artificial Intelligence, 2009. 2 p. pp. 361-362.
- Bertin, M (2014). Scientific Monitoring by Mining Scientific Papers. In 6th International Conference on Management of Emergent Digital Ecosystems. Saudi Arabia: ACM, 2014. 5p. pp. 196–200.
- BORNMANN, Lutz; MUTZ, Rüdiger. (2015). Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. In *Journal of the Association for Information Science and Technology*, n°66, pp.2215–2222. [en ligne]. Disponible à :<https://arxiv.org/ftp/arxiv/papers/1402/1402.4578.pdf> (Page consultée le: 16/08/2018)
- Chatterjee, P; Gause, B; Hedinger, H; Pollock, L (2017). Extracting code segments and their descriptions from research articles. In 14th International Conference on Mining Software Repositories. IEEE, 2017. 11p. pp. 91–101.
- FLAMERIE, Frédérique (2015). Le libre accès : pourquoi et comment? Université Pierre-et-Marie-Curie, [En ligne]. Disponible à : http://australe.upmc.fr/access/content/group/bupmc_04/midi_biblio_03_oa_hal.pdf (page consultée le 17/08/2018).
- Francis, A (2016). Literature Search Tool for the Extraction of Disease-Associated Genes using Frequent Itemset Mining. In International Conference on Computing, Analytics and Security Trends. India: IEEE, 2016. 6p. pp. 306–311.
- Funamori, M (2017). Open Science and the Academy: *A Theoretical Discussion*. In 6th IIAI International Congress on Advanced Applied Informatics. n°1, 7p. pp. 109–115.
- Hirschman, Lynette; Park, Jong C; Tsujii, Junichi; Wong, Limsoon; Wu, Cathy H. Accomplishments and challenges in literature data mining for biology. In *The MITRE Corporation*. n° 12 .USA,2002. Pages 1553–1561.
- Jhamb, D; et al. (2014). Identification of protein interaction methods from biomedical literature. In 4th International Conference on

- Computer Advancements in Biomedical Science. IEEE, 2014. 6p. pp. 1–6.
- Jurafsky, D. (2018). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Third Edition draft. USA, 2018.
- Kitchenham, B; Charters, S (2007). Guidelines for performing Systematic Literature reviews in Software Engineering. In EBSE Technical Report, n°45, pp. 1-51.
- Louvigne, S; Shi, J; Sharmin; S(2014). A corpus-based analysis of the scientific RA genre and RA introduction. In Proceedings of the 2014 International Conference on Advanced Mechatronic Systems. Japan: IEEE, 2014. 5p. pp. 123–127.
- Moro, S. P; Cortez; Rita P. (2015). Business intelligence in banking: *A literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation*. In Expert Systems with Applications, vol. 42, no. 3, pp. 1314–1324.
- Parthasarathy, G; Tomar, D. C (2014). Sentiment analyzer: *Analysis of journal citations from citation databases*. In International Conference - Confluence The Next Generation Information Technology Summit (Confluence). IEEE, 2014. 6p. pp. 923–928.
- Popova, S; Danilova, V (2014). Keyphrase extraction abstracts instead of full papers. In 25th International Workshop on Database and Expert Systems Applications. DEXA, 2014. 5p. pp. 241–245.
- Ravikumar, K. E; Waghlikar, K. B; Liu, H. (2014). Challenges in adapting text mining for full text articles to assist pathway curation. In 5th International Conference on Bioinformatics and Computational Biology. ACM, 2014. 8p. pp. 551–558.
- Song, M. (2015). Exploring concept graphs for biomedical literature mining. In International Conference on Big Data and Smart Computing. IEEE, 2015. 8p. pp. 103–110.
- Taylor, S.J; Bogdan, R. (1984). *Introduction to Qualitative Research Methods: the Search for Meanings*. Wiley, 1984.

- Tchoua, R. B. et al. (2017). Towards a hybrid human-computer scientific information extraction pipeline. In 13th International Conference on eScience. USA:IEEE, 2017. 10p. pp. 109–118.
- Torii, M; et al. (2015). RLIMS-P 2.0: *A generalizable rule-based information extraction system for literature mining of protein phosphorylation information*. In transactions on computational biology and bioinformatics. n° 1. iee/acm, 2015. 13p. pp. 17–29.
- United Nations Educational, Scientific and Cultural Organization (2015). Introduction to open access, Open access for library schools, Vol.:1; 2015. [En ligne]. Disponible à : <http://unesdoc.unesco.org/images/0023/002319/231920E.pdf> (page consultée le 30/10/2018)
- Yu, C. N; Levy, C. C; Saniee, I. (2017). Convolutional Neural Networks for Figure Extraction in Historical Technical Documents. 14th IAPR International Conference on Document Analysis and Recognition. IEEE, 2017, pp. 789–795, 2018.
- Zhen, Y; et al. (2017). Analyzing the usage of standards in radiation therapy clinical studies. In International Conference on Biomedical Health Informatics. IEEE, 2017. 4p. pp. 349–352.

Le contrôle par les pairs ouvert : Où en sommes-nous ?

Open peer review: where are we at?

S. Hachani, U. Alger 2 – Algérie

Résumé

Le contrôle par les pairs est une étape des plus importantes dans le processus de la publication scientifique. Il est sujet à des multiples critiques qui en ont fait un sujet de controverses et de problèmes. La raison principale est qu'il est entrepris de manière anonyme ce qui a engendré un nombre de préjugés ayant entaché le processus. Les tentatives de l'ouvrir (passant de l'expertise en aveugle à l'expertise en double aveugle) ont eu des résultats mitigés. L'irruption d'Internet a permis de l'ouvrir complètement sur la toile faisant du contrôle une opération faite au vu et su de la communauté avec questions et réponses affichées sur le site en question. Bien qu'ayant éliminé certains des préjugés dont le processus était coutumier, il n'en est encore qu'aux premiers balbutiements et n'est pas accepté unanimement par toute la communauté.

Mots clés. contrôle par les pairs ouvert, F1000, Biology Direct, R.I.O., Internet communication scientifique, Internet

Abstract

Peer review is a very important stage of the scholarly communication. It has been subject to numerous controversies and problems. The main reason is the secrecy by which it is undertaken and that has generated a number of biases that have marred the process. The opening up of the process (such as going from blind to double blind reviewing) has had mixed results. The Internet irruption has allowed a complete opening of the process with questions and answers regarding the submission being seen live on the site. Although it has eliminated some biases the process was known for, it is still in its first stages and is not unanimously accepted by the scientific community.

Key words. peer review, open peer review, F1000, Biology Direct, R.I.O., scholarly communication,

Internet

Introduction

Le monde est de nos jours sujet à des changements et des mutations que l'on aurait difficilement imaginés, encore moins prévus, il ya de cela quelques décennies. En effet, nous sommes rentrés dans une ère d'automatisation, de vitesse et de facilité d'accès à l'information en général des plus ahurissantes. Le dénominateur commun à toute cette révolution est l'utilisation de l'outil informatique et dont la prépondérance est telle qu'il est impossible d'imaginer notre vie sans ce fidèle compagnon. La fin du 20e siècle a vu l'irruption d'une autre révolution qui a emmené cette tendance à l'automatisation encore plus loin : Internet. Cette ouverture, cette immédiateté et surtout instantanéité de l'information a touché toutes les facettes de notre vie. Dans le domaine scientifique, Internet est utilisée par les chercheurs dans toutes les étapes de la recherche scientifique : de la conception, à l'expérimentation, à l'envoi et la publication. Toutes ces étapes, classiques pour le chercheur, sont devenues entièrement automatisées et il n'est pas un seul chercheur qui utilise le papier ou la poste pour rédiger ou envoyer sa soumission. Cette suite d'étapes indispensables à la communication scientifique et qui a été standardisée progressivement depuis les premiers journaux scientifiques du 17e siècle comporte une étape des plus importantes et primordiales de la sphère de la recherche scientifique : le contrôle par les pairs. Cette étape qui sert de tamis aussi bien qualitatif que quantitatif de ce qui est publié a toujours été sujet à des controverses, des abus, des dépassements qui ont été largement documentés surtout dans la littérature anglo-saxonne. Aussi bien véridiques que (dans certains cas) imaginaires, ils n'en représentent pas moins un sujet de discordance qui empoisonne les relations entre les différents protagonistes de la publication scientifique. Les préjugés en constituent la pièce d'achoppement car elles empêchent l'expertise d'être objective et surtout dénuées de toutes arrière-pensées. Qu'ils soient institutionnels, personnels, esthétiques, nationaux, sexuels, etc..., les préjugés plombent cette opération. L'une des raisons les plus prégnantes, pour ne pas la plus prégnante, est le mode d'application du processus : l'anonymat. En effet, il est tout de même inique que dans un tribunal (comme le contrôle par les pairs a été décrit), "l'accusé" ne connaisse pas son juge. Si certaines formes (passant de l'aveugle au double aveugle) ont essayé de remédier à cet état de fait, il n'en demeure pas moins que le processus est toujours en butte à de accusations multiples et variées. L'irruption d'Internet avec sa propension à l'ouverture et à la transparence semble être l'antidote à toutes

ces accusations de copinage, de passe droits, de subjectivité. Le raisonnement serait de dire : vu que l'anonymat protégeant l'expert est l'une des raisons de toute cette situation, l'ouverture partielle (ou complète) du processus rendrait ces agissements difficiles à cacher et donnerait des expertises des plus objectives et dénuées de tous les dépassements que l'anonymat a permis. Cette recherche essaiera d'investiguer quelles sont les expériences les plus récentes, les plus révolutionnaires que le processus a connu et surtout où sont arrivées ces expériences (d'où le titre de la communication).

5 Le contrôle par les pairs : un petit historique

Le contrôle par les pairs est l'action par laquelle les travaux de recherche scientifiques sont jugés par les collègues de la spécialité (d'où l'appellation pairs). La date à laquelle cette opération a commencé à être appliquée est des plus problématiques et des plus imprécises. La raison de cette imprécision est du au fait que le processus en lui-même n'a jamais (jusqu'à récemment) été uniformément accepté, ni unanimement défini. Les raisons sont multiples mais l'une des plus importantes est le fait que le processus n'était pas indispensable car la production scientifique n'était pas des plus foisonnantes surtout à la période où les premiers périodiques scientifiques ont été fondés.

5.1 *Philosophical Transactions* et *Journal des Sçavans* : deux titres temporellement proches mais scientifiquement éloignés

Il est communément accepté que la date du 5 janvier 1665 marque la première fois un périodique scientifique est lancé dans le monde : *Le Journal des Sçavans*. Son but était de rapporter tout ce qui intéressait les savants et hommes de lettres de par le monde. De l'autre côté de la Manche, et exactement trois mois plus tard, en date du 06 mars 1665, Henry Oldenburg fondait *The Philosophical Transactions*. Ces deux titres furent suivis dans les années par une foultitude de titres qui font du périodique scientifique *contrôlé par les pairs* l'outil de prédilection et le medium préféré de publication des chercheurs. Nous avons insisté et mis *par les pairs* en italique car, encore une fois, le processus est à la base et au cœur de la Science. A cet effet, cette opération a été appelé " le pivot autour duquel tout le processus est bâti "(Ziman 1968, p.111) ou alors [les experts sont] " les gardiens du temple de la Science " (Crane 1967) alors que Hargens (1988) compare la publication scientifique à " une Olympiade mentale pour le

prestige et la reconnaissance ". Cette Olympiade, comme toute compétition sportive, a besoin d'arbitres, de juges, de *referees* fonction assurée par l'expert (e).

5.2 Les premières manifestations du contrôle par les pairs

"Officiellement", ce sont *The Philosophical Transactions* qui ont imposé les premières un contrôle en amont de ce qui est publié. En effet, le premier numéro stipulait expressément..." Que les *Philosophical Transactions* soient composées par Mr. Oldenburg, imprimées le premier lundi de chaque mois, si nous avons assez de matériel pour le faire et que la brochure soit autorisée par le même Conseil [des *Philosophical Transactions*] étant *premièrement revue* (c'est nous qui soulignons) par quelques membres de la même institution". De son côté, Dennis De Sallo le premier " éditeur " du *Journal des Sçavans* n'avait nullement la prétention de garantir ni de contrôler l'information mais tout simplement la signaler. Il avait déclaré : " Nous aspirons à rendre comptes les idées des autres sans les garantir " (Rennie 2003, p.2). Nous avons cité ces deux exemples des deux premiers journaux scientifiques pour faire ressortir la différence entre les deux titres qui, bien que temporellement proche, sont de nos jours scientifiquement lointain. Cette idée de la prépondérance, de la "supériorité " scientifique des *Philosophical Transactions* sur *Le Journal des Sçavans* est développée par Francis André qui fait ressortir la supériorité britannique actuelle en matière d'édition scientifique comme preuve de cette différence (André, 2005, p.9) et surtout par Jean Claude Guedon (Guedon, 2001) qui déclare sans ambages : "... régulièrement, la revue est comparée au '*Journal des Sçavans*', comme si les deux publications étaient des sœurs jumelles - j'ai moi-même commis cette erreur il y a quelques années. Alors qu'il est évident que toutes deux sont des périodiques, traitent de la philosophie de la nature et furent lancées à quelques mois d'intervalle seulement, il n'est pas si évident qu'elles poursuivaient exactement les mêmes buts. La publication française reproduisait un modèle d'échanges épistolaires manuscrits, typique de la République des Lettres, quelque peu cancanier et porté sur l'actualité ; de ce fait, elle se rapproche plus de '*Scientific American* ' que d'une revue scientifique moderne et apparaît ainsi fermement ancrée dans le domaine du journalisme scientifique, alors émergeant. Bien que le *Journal* publiât de temps à autre quelques papiers originaux, ceux-ci apparaissaient comme une expression particulière de

l'actualité parmi d'autres types de nouvelles ; par contraste, '*Phil Trans*', bien que diffusant aussi des informations récentes, avait véritablement pour but de créer un registre public des contributions originales à la connaissance. En d'autres termes, la publication parisienne suivait la mode tandis que la revue londonienne servait à valider l'originalité. Là réside la différence significative (et profonde) entre les deux périodiques". Nonobstant ces querelles franco-britanniques, il n'en demeure pas moins qu'une publication (*Philosophical Transactions*) est reconnue comme l'archétype de la publication scientifique par son côté originale et surtout la nature de ses publications alors que la seconde (*Le Journal des Sçavans*) est moins cotée car " quelque peu cancanier et porté sur l'actualité ". D'un autre côté, Ray Spier situe *une forme* de contrôle par les pairs à une date beaucoup plus ancienne exactement au 9e -10e siècle avec un traité de médecine (*Adabatabib–Ethique du médecin*) en vingt chapitres et dont l'auteur était *Ishak Ibn Ali Errahwi* (de la ville de Raha en Syrie).En plus, des différents aspects moraux et pratiques de la profession médicale, cet ouvrage comporte aussi une partie qui s'apparente à un contrôle par les pairs , une sorte d'ordre des médecins qui statuerait sur qualité des soins donnés par le médecin. Ce dernier devait à chaque visite du patient faire des doubles des notes de la condition du patient. Selon la guérison ou le décès du patient, les notes étaient examinées par le conseil local des médecins qui devait décider de la justesse ou des fautes professionnelles éventuelles qui découlaient de la pratique de la médecine par le médecin en question.

L'éventuelle erreur médicale pouvait déboucher sur une interdiction de pratiquer la médecine et aussi à des suites juridiques, les parents du patient décédé pouvant attaquer en justice le docteur pour mauvaise pratique ou faute professionnelle. Ajlouni et El Khalidi (1997) évoquent le système de la *Hisbah* (contrôle de qualité) et du *Muhtasib* (contrôleur de qualité avec les pouvoirs d'un juge) et dont les étapes ressemblent ce qui a été décrit dans le livre de *Ishak Ibn Ali Errahwi* avec la notable différence que le médecin devait remettre aux proches du malade une copie des ordonnances et qui, selon le résultat du traitement, pouvait déboucher soit le paiement d'honoraire ou une demande de compensation. Le *Hakim* (médecin en chef) qui était un docteur connu devait juger de la justesse ou non du traitement. Ces formes de contrôle par les pairs dont s'enorgueillissait la médecine arabe aux 10e - 11e siècles peuvent être considérées comme les premiers embryons

de contrôle par les pairs tels qu'on les connaît dans la publication scientifique moderne.

Après cela, il est difficile de situer exactement une date précise de l'application du contrôle par les pairs pour une raison très simple : sa définition. En effet, si le concept lui-même est connu (surtout de la communauté scientifique), ce qu'il recèle comme "ingrédients" est méconnu même de la part de cette même communauté. En d'autres termes, le contrôle par les pairs n'a pas évolué linéairement se développant crescendo mais a été appliqué selon l'humeur des différents éditeurs en chef et surtout selon les spécialités. S'il n'existe pas de narration temporelle précise qui puisse tenir lieu d'historiographie du processus, Weller (2002, p.3-6) dans un ouvrage de référence du sujet en fait une relation des plus complète et des plus détaillée. On voit dans cette relation que le contrôle par les pairs *moderne* est situé par l'auteur au début du 20^e siècle et qu'il est appliqué et accepté de manière très inégale et surtout qu'il n'est pas unanimement accepté. Cet état de fait, de prime abord contradictoire étant donné la prépondérance du processus dans la communication scientifique, tient au fait que la masse documentaire produite était *quantitativement* gérable. C'est la continuelle et surtout exponentielle augmentation de la masse documentaire, surtout après la seconde guerre mondiale, qui impose le processus comme l'étalon valeur de la science publiée. C'est Derek J. de Solla Price (1963) qui, dans son livre séminal "Little science big science", a exemplifié (entre autres) cette explosion documentaire qui nous a conduit à la situation actuelle. De Solla Price a calculé que le nombre de chercheurs, et de ce fait les publications, doublent chaque 10-15 ans et qu'en conséquence depuis 1600 cette augmentation a atteint une magnitude de 10⁶. D'un autre côté, Price a calculé qu'à l'époque de l'étude, le nombre de périodiques scientifiques se situait aux environs de 50 000 dont 30 000 encore publiés qui publiaient quelques 6 000 000 d'articles avec une augmentation d'au moins 500 000 articles par an (p.8).

6 Les préjugés : un mal dont le contrôle par les pairs ne peut se départir

Le contrôle par les pairs est une opération purement humaine. De ce fait, elle ne peut (et n'est pas) une opération faite objectivement. Pour être plus clair, l'expertise a toujours été faite par des humains (et l'est encore...) et cet état de fait a induit des réactions, des approches ... humaines. Il est

pratiquement impossible de demander à une personne de mettre de côté sa personnalité et juger un travail sur ses valeurs intrinsèques. Il est clair que dans le meilleur des mondes, ceci est le but mais"Ni les éditeurs ni les experts ne sont infaillibles. Les éditeurs peuvent se tromper dans le choix des experts... ou les éditeurs peuvent mal interpréter ou mal appliquer les conseils des experts. Les experts eux-mêmes peuvent ne pas être suffisamment experts, informés, consciencieux ou équitables" (Harnad 2000). Cette tirade de Harnad exprime toute la complexité du travail des différents protagonistes du processus qui se retrouvent à gérer une situation qu'ils (et elles) n'arrivent pas à faire d'une manière correcte et surtout objective. La raison ? La nature humaine qui ne peut se départir (quoiqu'elle fasse) de cette propension à avoir des penchants (conscients ou inconscients, voulus ou pas) contre ou pour une personne, une institution, une religion, un sexe, une esthétique, une langue, un pays etc. Nous parlons là des préjugés¹ et qui sont LE problème qui a taraudé le processus depuis ces premières manifestations modernes. En effet, la littérature sur le sujet abonde d'études qui prouvent que le processus est profondément sujet à des controverses qui ont l'ont entaché et, en même temps dévalorisé, jusqu'à un certain point. Il serait fastidieux de tous les citer car le processus en est inondé mais parmi les plus connus et les plus prépondérants, Shatz (2004, p.37-45) en cite quatre des plus flagrants et qui sont : les préjugés Ad Hominem, les préjugés institutionnels, les préjugés idéologiques et les préjugés esthétiques. Etant donné leur importance, nous avons jugé utile d'en expliquer quelques aspects.

6.1 Les préjugés Ad Hominem

Ce sont des préjugés que l'on peut avoir pour ou contre une personne. Ils peuvent prendre plusieurs formes : on peut éprouver de la haine (ou de la compassion) pour une personne, on peut aussi être en compétition avec une personne. D'un autre côté, on peut avoir ce préjugé (positivement) si on doit juger le travail d'une personne et que l'on sait que le résultat (éventuellement positif) aiderait cette personne à avoir une promotion ; dans ce cas on aurait tendance à être plutôt indulgent.

¹ Nous avons toujours utilisé le terme "*préjugé*" traduction personnelle du mot "*bias*" exclusivement utilisé dans la littérature anglo saxonne. Il faut dire que le sujet est pratiquement traité dans la sphère anglo saxonne seulement et que cette quasi exclusivité a entraîné une normalisation des concepts.

6.2 Les préjugés institutionnels

Ce sont des préjugés qui tendent à juger une soumission non pas intrinsèquement mais selon l'institution du soumissionnaire. En d'autres termes, on aurait tendance à penser qu'une soumission émanant d'un chercheur de Harvard est *automatiquement* meilleure que celle émanant d'un chercheur de l'Université d'Alger 2 (...). S'il est vrai que les institutions académiques ne se valent pas toutes, il n'en demeure pas moins que ce préjugé peut être ...préjudiciable aux chercheurs du Sud qui se voient pratiquement catalogués comme inférieur à leur confrères du Nord. A cet effet, dans une étude devenu classique sur le sujet, Peters et Ceci (1982) ont démontré l'existence de ce préjugé. En effet, les auteurs ont soumis de nouveau douze articles déjà publiés par des chercheurs prestigieux et très productif appartenant aux douze plus prestigieuses universités américaines en psychologie. Ces articles publiés le furent dans des journaux dont le taux de rejet avoisinait les 80% pratiquant le processus d'expertise non aveugle. Des altérations mineures (tel que les noms et les institutions) furent introduites, le contenu (titre, mots clés, méthodologie, etc.....) fut aussi légèrement changé. Ensuite ces soumissions furent envoyées aux mêmes périodiques qui les avaient acceptés dix-huit à trente-deux mois auparavant. Les résultats furent, comme suit : sur trente-huit éditeurs et experts, seul trois détectèrent les resoumissions, neuf des douze resoumissions continuèrent le processus de contrôle par les pairs, des neufs resoumissions, une seule fut acceptée et huit furent rejetées, seize des dix-huit experts recommandèrent le rejet et les éditeurs les suivirent dans leur décision. Les raisons évoquées étaient, dans la plupart des cas, décrits comme "*défauts méthodologiques sérieux*". Bien que cette étude ait été critiquée pour des faiblesses méthodologiques, elle n'en demeure pas moins symptomatique d'une vérité que vit tout chercheur : les recherches ne sont pas toutes traitées sur le même pied d'égalité. Un chercheur dont le nom serait John Clark émergeant au "*Tri-Valley Center for Human Potential*" (nom créé par Peters et Ceci dans leurs resoumissions) est certes moins prestigieux que celui d'un éminent psychologue professeur, par exemple, à l'*Université de Chicago*.

6.3 Les préjugés idéologiques

Ce sont des préjugés qui tendent à influencer positivement ou négativement la position d'un expert. Dans les deux cas, l'expert peut soit accepter et être attentif aux arguments qui confirment les siens ou alors être contre les

arguments qui vont à l'encontre de ses propres points de vue. De plus, le préjugé idéologique a été décrit comme un préjugé " à la mode " en ce sens que les recherches qui vont à l'encontre des théories acceptées et ayant " pignon sur rue " ont de fortes chances de se voir rejeter alors que les recherches qui abondent dans le sens de ces théories ont de fortes chances de se voir accepter. Ceci a engendré un total déséquilibre dans les taux des acceptations en conséquence d'une autre donnée : les experts étant choisis parmi les plus reconnus des scientifiques, il y a de fortes chances à ce que ces derniers abondent dans le sens du statu quo et acceptent les soumissions ne sortant pas des sentiers battus. Enfin, il est (presque) normal pour un scientifique de ne pas prendre de risques et d'adouber une idée qui n'a pas eu le consensus de la communauté.

6.4 Les préjugés esthétiques

Une recherche bien présentée, agencée et propre a-t-elle plus de chances d'être acceptée que celle dont la présentation laisserait à désirer nonobstant le contenu ? Il est connu que la plupart des scientifiques sont comme on dit "tête en l'air " et l'image stéréotypée est celle d'un bureau plein de livres, de papiers et de schémas le tout dans un désordre magistral. Ne dit-on pas qu'" un bureau bien rangé est le signe d'un esprit dérangé" ? Cette image du savant cheveux en l'air, cravate mal nouée, chemise déboutonnée colle bien à l'idée que l'on se fait du savant tout à sa passion. Mais est-ce que cela influe sur l'impression que les experts ont sur leur jugement de la soumission ? Il semblerait bien que oui car un travail bien présenté est tout de même plus facile et plus agréable à lire et analyser.

Au-delà de ces préjugés les plus prégnants et surtout les plus documentés, il existe d'autres dont l'influence sur le travail de l'expert est aussi importante. On citera le préjugé sexuel qui semble agir contre les femmes, le préjugé national qui serait un préjugé qui agirait contre certaines nationalités pour des raisons diverses et variées, le préjugé religieux qui agirait contre une religion donnée ou alors le préjugé linguistique qui agirait contre une langue donnée (et surtout *au profit* de la langue anglaise , lingua franca de la science).

7 L'expertise en double aveugle : est-ce la solution ?

Face à toutes ces dissensions et critiques, il était (et est) impératif de trouver une solution à un problème qui minait le processus du contrôle par

les pairs. La forme la plus usitée étant l'expertise en aveugle (où l'expert connaît l'auteur alors que ce dernier ignore le nom de son expert) et qui s'était avérée, comme vu en supra, source de problèmes, il a été proposé de se débarrasser de la cause principale de ces préjugés en dévoilant (aussi) l'identité de l'expert. Cette action est supposée enlever l'impunité dont jouit l'expert grâce à l'anonymat de l'expertise dans l'expertise en aveugle. C'est l'expertise en double aveugle qui rendait l'expert plus justiciable et responsable de ce qu'il écrit et émet comme avis. Le raisonnement était que l'expert, fort de son anonymat, pouvait agir sans avoir de compte à rendre, le "démasquer" le rendrait plus consciencieux et moins sujet à des dérives (conscientes ou inconscientes) entachant son travail. Malgré la "revolutionarité" du concept dans un processus connu pour sa frilosité à expérimenter, cette expérience a elle aussi montré ses limites malgré certaines avancées. Il est loisible de voir que le premier et plus important avantage est la totale ignorance des deux protagonistes de l'un l'autre ce qui, en principe, donnerait une expertise basée sur le contenu intrinsèque et non pas sur la personne ou l'institution qui a soumis (éliminant de facto le préjugé personnel et aussi institutionnel). Dans ce cas, même une sommité ou même un Prix Nobel pourrait avoir sa soumission refusée. Il est pratiquement impossible et inconcevable pour un expert, aussi chevronné, sur de lui-même et connu soit il, de refuser ou critiquer le travail d'un ténor de la spécialité. Il le ferait s'il est ignorant de son identité et de sa stature. Cet état de fait rendrait l'expertise une expertise des plus objectives possibles et surtout des plus impersonnelle. Quant aux inconvénients, ils sont beaucoup plus prégnants et nombreux et peuvent être résumés en :

- L'expertise en double aveugle ajoute à la charge de travail du journal
- L'expertise en double aveugle n'est pas effective dans l'anonymat (les expert(e)s arrivent à deviner qui est qui surtout dans les spécialités restreintes)

- l'absence d'informations sur l'auteur empêche l'expert(e) de détecter les conflits d'intérêts

- l'absence d'informations sur l'auteur empêche l'expert(e) aussi d'être en mesure de suivre la progression des travaux de l'auteur(e)

- l'absence d'informations sur l'auteur empêche l'expert(e) aussi de voir est ce que le travail est nouveau ou est ce juste la même problématique ressassée plusieurs fois (les Anglo saxons utilisent l'expression de " "

salami slicing " qu'on pourrait traduire, ou du moins en donner un équivalent, par " auto – plagiat ")².

Après toutes ces expérimentations qui traduisant le malaise dans lequel le processus se trouve et qui a mené la communauté scientifique à se départager en deux camps diamétralement opposés chacun se battant pour ses droits, il était impératif de trouver une solution pour rendre au processus sa place dans le landernau de la publication scientifique. La prochaine étape sera celle de l'ouverture complète du processus et qui dépassera les différentes expérimentations. Cette révolution (car il s'agit d'une révolution) sera l'avènement d'Internet. Outil, s'il en est, synonyme d'ouverture, la toile permettra d'ouvrir complètement le processus et d'en alléger les inconvénients dont il souffrait dans l'univers papier.

8 Internet : l'incontournable medium

Il serait présomptueux d'essayer de présenter Internet et son influence sur notre vie de tous les jours et surtout dans la vie des chercheurs et leur maniement quotidien de ce medium des plus indispensables. Mais le fait est que nous vivons par et avec Internet et son influence sur notre vie est des plus présentes et pressantes. Un medium qui couvre plus de la moitié de la population mondiale (avec il est vrai des différences et des fractures des plus évidentes), qui permet une interaction pratiquement en temps réel et raccourcit les distances, ne peut (et est) qu'être un medium des plus bénéfiques .Parmi les nombreux sites qui recensent et donnent des statistiques sur la situation d'Internet de par le monde , *Internetworldstats* donne les chiffres et statistiques les plus sérieuses et les plus mises à jour .En date du 31 Décembre 2017,il y avait 4 156 932 140 personnes ayant accès à Internet ce qui représente 54.4 % de la population mondiale. Il existe aussi de profondes disparités entre les différentes régions et continents. Ainsi, l'Amérique du Nord est la plus connectée avec 95% de taux de pénétration alors que l'Afrique n'est connecte qu'à hauteur de 35.2 %. On remarquera aussi que l'Afrique et l'Asie qui représentent presque les trois quart de la population mondiale (72 %) sont classées dernière et avant

² En fait le *salami slicing* désigne une opération qui consiste au partage en " tranches " d'un même travail afin d'en faire une série de travaux. Si le procédé n'est pas répréhensible en soit, il n'en demeure pas moins symptomatique du " publish or perish " qui guide les chercheurs et qui fait qu'on est jugé sur la longueur de son CV nonobstant la qualité.

dernière en taux de pénétration avec respectivement 35.2 et 48.1%. Cette fracture numérique bien présente tend cependant à se résorber tel que les chiffres de l'évolution du taux de pénétration le montrent : ils sont de l'ordre de 9941 % et de 1670 % pour respectivement l'Afrique et l'Asie. Ces chiffres démontrent aussi que la fracture tend à diminuer ainsi que le fossé entre have et have not. Le tableau suivant résume ce qui a été dit en supra :

INTERNET USAGE STATISTICS
The Internet Big Picture
World Internet Users and 2018 Population Stats

WORLD INTERNET USAGE AND POPULATION STATISTICS DEC 31, 2017 - Update						
World Regions	Population (2018 Est.)	Population % of World	Internet Users 31 Dec 2017	Penetration Rate (% Pop.)	Growth 2000-2018	Internet Users %
Africa	1,287,914,329	16.9 %	453,329,534	35.2 %	9,941 %	10.9 %
Asia	4,207,588,157	55.1 %	2,023,630,194	48.1 %	1,670 %	48.7 %
Europe	827,650,849	10.8 %	704,833,752	85.2 %	570 %	17.0 %
Latin America / Caribbean	652,047,996	8.5 %	437,001,277	67.0 %	2,318 %	10.5 %
Middle East	254,438,981	3.3 %	164,037,259	64.5 %	4,893 %	3.9 %
North America	363,844,662	4.8 %	345,660,847	95.0 %	219 %	8.3 %
Oceania / Australia	41,273,454	0.6 %	28,439,277	68.9 %	273 %	0.7 %
WORLD TOTAL	7,634,758,428	100.0 %	4,156,932,140	54.4 %	1,052 %	100.0 %

NOTES: (1) Internet Usage and World Population Statistics estimates in Dec 31, 2017. (2) CLICK on each world region name for detailed regional usage information. (3) Demographic (Population) numbers are based on data from the [United Nations Population Division](#). (4) Internet usage information comes from data published by [Nielsen Online](#), by the [International Telecommunications Union](#), by [GfK](#), by local ICT Regulators and other reliable sources. (5) For definitions, navigation help and disclaimers, please refer to the [Website Surfing Guide](#). (6) The information from this website may be cited, giving the due credit and placing a link back to www.internetworldstats.com. Copyright © 2018, Miniwatts Marketing Group. All rights reserved worldwide.

Figure 1- Tableau représentant Internet à travers le monde (Internetworldstats-2017)

9 Sol Tax : le premier "Open Peer Reviewer" ...en 1959

Il est de notoriété que parler de contrôle par les ouvert (ou Open Peer Review) équivaut à parler d'Internet qui a permis cette ouverture si convoitée et synonyme de justice et d'équité. Ceci implique automatiquement à ramener la date de cette révolution, au mieux, au milieu des années 90. En effet, Internet tel qu'on le connaît de nos jours n'a pris son envol et est devenu " grand public" que vers cette date. Cependant, il faut reconnaître à l'anthropologue américain Sol Tax l'honneur d'avoir ouvert le contrôle par les pairs sans Internet et ce dès le numéro de lancement de septembre 1959 de *Current Anthropology*, titre qu'il avait fondé. Il explique sa stratégie dans une lettre aux associés (Tax 1959 a) dans laquelle il expose le but et les étapes et appelle *Current Anthropology* une "expérimentation sociale" qui est le fait des associés repart(e)s à travers

toutes les contrées du monde. C'est dans la deuxième contribution (Tax 1959 b) qu'il donne des détails de son expérimentation et dans laquelle il fait preuve d'un esprit pionnier et révolutionnaire étonnant pour l'époque. Après la présentation des objectifs de la revue et ses spécificités, Tax décrit la manière dont les manuscrits seront traités en ces termes : " étant donné que les manuscrits varieront en approche, portée et complexité, les éditeurs les traiteraient de manières différentes 1- certains seront lus par quelques experts et seront acceptés et publiés

2-un manuscrit peut être un important noyau pour l'intercommunication entre les spécialistes dans le domaine couvert par la soumission. Dans le cas, de tels manuscrits et après qu'une recherche ait été lue et provisoirement acceptée, elle sera dupliquée et envoyée à une liste de lecteurs. Cette liste comprendra des noms suggérés par l'auteur et aura deux catégories générales de personnes :

a) des lecteurs experts dans le domaine en considération. Ils peuvent ajouter des informations, argumenter l'interprétation ou alors ne rien dire. Dans toutes les figures de cas, l'auteur verra les commentaires faits par les lecteurs et nous conseillera sur la meilleure manière de traiter chaque réponse soit par incorporation dans l'original (avec remerciements), par inclusion (avec la réplique appropriée) ou ce qui paraît la meilleure option. Ainsi, dans un seul numéro, nous aurons la problématique centrale, les informations additionnelles pertinentes, les arguments principaux et la réfutation.

b) les lecteurs dont les centres d'intérêt sont à la marge du sujet traité par la recherche mais qui n'y est pas central. A titre d'exemple, les personnes qui abordent le sujet soit en tant que partie d'un plus grand tout ou comme le tout duquel ils sont intéressés par les parties. Ainsi, nous aurons un cadre inclusif et étendu et une occasion d'apprendre d'autres sciences et de partager avec elles nos découvertes.

Cette manière d'agir semble être un moyen prometteur pour faciliter l'intercommunication. En envoyant chaque recherche à des personnes aux approches différentes, nous aurons un niveau différent de précision et d'exhaustivité et une plus grande opportunité de répondre aux questions soulevées aux frontières de notre champs d'investigation " (Tax 1959 b, p.8).

Nous avons reproduit *in extenso* cette contribution de Tax pour montrer à quel point ses idées étaient révolutionnaires. En effet, il propose, et ce en 1959 et alors que l'ouverture dont nous nous prévalons de nos jours n'était

même pas envisageable, d'ouvrir le contrôle par les pairs et de laisser à l'auteur (e) le choix de choisir des personnes qui valideraient sa recherche, innovation qui est appliqué par certains sites de nos jours (Biology Direct). De plus (et sans Internet et son interactivité), *CurrentAnthropology* publie le cycle de l'article avec *problématique centrale, informations additionnelles pertinentes, arguments principaux et réfutation*. Cette manière d'agir est maintenant pratiquée dans certains sites (*R.I.O.- Research Ideas and Outcomes*) où la recherche en gestation est suivie en direct et l'article est exposé dans ses différentes étapes sur Internet. Tax a réussi la gageure de lancer ce qui est de nos jours innovateur avec une différence notoire : il n'avait pas la puissance, l'ubiquité, la rapidité et l'étendue qu'a Internet de nos jours. C'est pour cela que Sol Tax mérite, sans aucun doute, le titre de père de *L'Open Peer Review* et ce à une époque où il ne bénéficiait pas des fonctionnalités extraordinaires qu'Internet permet de nos jours.

10 Les premières manifestations de contrôle par les pairs ouvert

Si le contrôle par les pairs traditionnel et dont la date d'application et d'utilisation est pour le moins sujette à controverse à cause de raisons multiples (modalités d'application, différence entre les spécialités, absence de raisons quant à l'application), on peut décemment situer les deux premières formes de contrôle par les pairs ouvert aux environs de la fin des années 70 et 90 avec respectivement *Brain and Behavioural Sciences* et *Psychology*. S'il est vrai que Sol Tax a innové et ce dès la fin des années 50 en pratiquant un contrôle ouvert sans Internet, le contrôle par les pairs ne s'est ouvert et atteint son apogée qu'avec Internet dont le côté grand public et universalité n'a été atteint que vers la moitié des années 90 – début des années 2000. C'est à partir de cette date qu'on peut parler d'Internet comme le phénomène universel qu'il est actuellement. Deux titres aux destins différents et opposés et sous l'impulsion d'un pionnier du libre accès (Stevan Harnad) peuvent se prévaloir du titre de premiers journaux pratiquant le contrôle par les pairs ouvert à l'ère d'Internet (nous mettons à l'ère d'Internet pour différencier leurs travaux de ceux de Sol Tax qui avait fait du contrôle par les pairs ouvert *sans Internet*). Ce sont respectivement *Brain and Behavioural Sciences* (1978-) et *Psychology* (1990 - 2002). Si *Psychology* n'est plus vivant (pour employer un terme catalographique) à cause de raisons multiples (dont la viabilité financière) , *Brain and Behavioural Sciences* quant à lui pratique une nouvelle forme de contrôle par les pairs appelée " *Open Peer Commentary* "

(Commentaires Ouverts par les Pairs). Dans l'éditorial inaugural , Harnad (1978) reconnaît d'ailleurs que B.B.S. est calqué sur le modèle de *CurrentAnthropology*. Mais quel est la particularité de B.B.S et de son " Open Peer Commentary " ? Le site la présente de la manière suivante: " Les travaux de recherche particulièrement controversés et significatifs dans les domaines de la psychologie, la biologie comportementale ou les sciences cognitives sont publiés avec 10 à 25 commentaires pour chaque article émanant de ces disciplines et des disciplines connexes en plus des réponses de l'auteur à ces commentaires. Le résultat est un forum unique et fascinant pour la communication, la critique et la stimulation et particulièrement l'unification de la recherche dans les sciences comportementales et des neurosciences qui vont de la neurobiologie moléculaire à l'intelligence artificielle et la philosophie de l'esprit ". De son côté, Harnad (2011) l'explique de la manière suivante : " un journal traditionnel commence par conduire le processus du contrôle par les pairs et publie alors les articles ayant répondu à ses critères de certification. BBS (et *CurrentAnthropology* sur lequel BBS est modélisé) fait circuler premièrement les articles (acceptés et certifiés) à environ 100 chercheurs dans toutes les spécialités et dans le monde entier les invitant à soumettre un commentaire de 1000 mots qui critique ou complète l'article cible. L'auteur répond alors aux commentaires (10-20 ou plus) et tout est coédité dans le journal ".

Ces deux exemples (surtout *Brain and Behavioural Sciences* qui est toujours vivant contrairement à *Psychology* qui est mort pour rester dans le jargon catalographique) qui ont bénéficié de la personnalité de leur fondateur et de l'expérience de *CurrentAnthropology* ont pratiqué une forme de contrôle par les pairs innovante : le commentaire par les pairs ouvert. Non seulement ont-ils ouvert le processus mais aussi ils lui ont ajouté une autre couche d'ouverture en permettant aux expertises et aux commentaires d'être partie prenante de l'article et de permettre de voir son évolution et sa maturation. Ceci était impossible dans le monde papier où le circuit de la soumission était balisé : auteur (e)-éditeur (trice)-expert (e)-éditeur (trice)- auteur (e) et surtout sous le sceau du secret ce qui avait induit toutes les critiques et problèmes dont souffre le processus .Il faut remarquer, à toutes fins utiles, la révolutionnarité de *CurrentAnthropology* qui dans un monde ne connaissant la connectivité qui est la notre de nos jours a réussi la gageure d'ouvrir un processus connu pour son côté secret et inique.

A partir de ces expériences pionnières, le monde de la publication scientifique a commencé à expérimenter différentes formes de contrôle par les pairs ouverts. Nous disons différentes car ouvrir le processus étant acquis avec l'irruption d'Internet, chaque site, chaque spécialité a essayé des variantes de contrôle que nous allons voir. Nous avons choisi trois qui semblent parmi les plus innovantes et les plus pionnières. Ce sont: *Research Ideas and Outcomes (RIO)*, *Faculty of 1000* et *Biology Direct*. Ces trois sites ne sont pas les seuls à pratiquer le contrôle par les ouverts 3 mais ils ont poussé l'ouverture d'un processus, connu pour sa propension à être fermé, à un niveau qu'il est intéressant de voir et d'analyser.

11 F 1000, BiologyDirect et R.I.O.: trois sites à l'avant-garde du contrôle par les pairs ouverts

L'ouverture du processus du contrôle par les pairs étant relativement récente, il est normal que les expériences soient peu nombreuses. Après *Psychology* et *BBS* surtout mus par la personnalité de leur fondateur, Stevan Harnad, les autres expériences ont soit disparus (Naboj, Philica) ou alors ont été suspendus (tel que *ETAI*). Les expérimentations présentées sont parmi les plus innovatrices et surtout permettent à toute personne munie d'Internet de voir les différentes étapes du processus pratiquement en temps réel et surtout avec une transparence et une absence d'anonymat qui seraient propices à des expertises des plus objectives car les experts sont connus, leurs expertises aussi et leur décision finale quant à la publication est publique.

11.1 F 1000 où le contrôle se fait post publication



Figure 2-Page d'accueil de F 1000

On pourra entre autres citer *Journal of Medical Internet Research (JMIR)*, *Philica*, *BMJ Rapid Responses*, *Atmospheric Chemistry and Physics (ACP)*, *Journal of Interactive Media in Education*, *Electronic Transaction on Artificial Intelligence (ETAI)*

Faculty of 1000 est un site se spécialisant en médecine et en biologie. Il comporte un nombre de rubriques dont les plus importantes et relatives à notre travail sont celles sous la rubrique " F 1000 Research " et plus particulièrement la sous rubrique "Article".

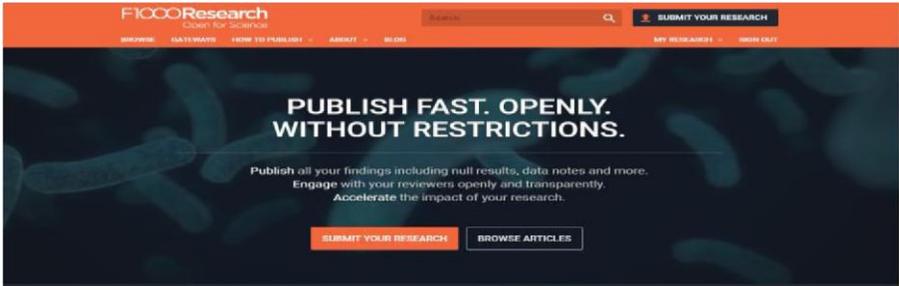


Figure 3- F 1000 Research et section Articles

En effet, cette sous rubrique détaille la manière dont les différentes soumissions sont expertisées de la manière la plus ouverte possible. Cette sous rubrique est au cœur de l'ouverture prônée par la communauté dans les expertises. En effet, elle ouvre le processus à un point qui aurait été non simplement impossible mais aussi inimaginable il y a de cela quelques années. Qu'on en juge.

11.1.1 Articles

Ce sont des articles soumis de la manière la plus classique possible par le biais d'un système où les différentes parties de la soumission sont incluses (titre, auteur(s), mots clés, etc.). Après avoir soumis, l'auteur doit suggérer des experts pour son travail. Il est même aidé dans sa tâche par des conseils donnés par le site pour l'orienter. Il est à signaler que l'auteur suggère et l'équipe éditorial contacte les experts.

Le schéma suivant explique le processus :



Figure 3- Processus de publication F1000

Une fois cette étape franchie, les expertises sont reçues par le comité éditorial qui les publie avec l'article, les réponses des auteurs et la version finale. Par exemple, l'article "*First draft genome assembly of the Argane tree*" n'a pas eu encore d'expertises et se présente de la manière suivante :



Figure 4-Présentation d'un article à expertiser

On remarquera que la mention "Awaiting peer review" (en attente d'expertise) veut dire les experts n'ont pas soit répondu à l'appel à expertise ou alors n'ont pas encore expertisés le travail.

De son côté, et là est le côté révolutionnaire du processus, l'article "*Haematology of N'Dama and West African Shorthorn cattle herds under natural Trypanosoma vivax challenge in Ghana*" a eu trois réponses contradictoires : qui veut dire expertise négative. L'article apparaît sur le site sous cette forme :

✓ qui veut dire expertise
? qui veut dire expertise positive avec
✗ enfin



Figure 5-Article expertisé et décisions des experts

De plus, ces expertises sont intégralement sur le site avec le nom de l'expert sous "read report". Ce qui veut dire les auteurs de l'article en question savent qui a accepté, qui a émis des réserves et refusé la publication de leur article. Nous prenons pour exemple l'expert ayant refusé l'article, son expertise se lit comme suit :

Open Peer Review
Current Referee Status: ✔ ✘ ✘

Version 1

Referee Report 01 Jun 2018
Sophie Thévenon, UMRI 12135/12135, French Agricultural Research Centre for International Development (CIAT) - Montpellier, France

✘ Not Approved

The manuscript aims at assessing the effect of infections by *Trypanosoma vivax* on the hematological parameters of N'Dama and WASH (West African Short-horn) cattle, raised in two natural environments in Ghana. The purpose is to highlight the trypanotolerant character of these breeds.

Major comments:

The article suffers from several major problems and is not suited for indexing.

The bibliography is quite incomplete: major papers written by Traill et al.¹⁻² and Mattoli et al.³ are not cited. These authors worked on N'Dama cattle raised in Congo and Dambila respectively and on the relationships between productivity, anemia and infection. Mattoli et al. 1998 (Acta Tropica) showed that N'Dama cattle suffered from high febrile challenge. Traill et al. 1994 showed that N'Dama infected by trypanosomes had lower PCV values and lower weight gain than non-infected N'Dama. In addition, an experimental infection published by Berthier et al. (2015) presented anemia evolution in 5 cattle breeds of West Africa under *T. congolense* infection and show of N'Dama and WASH were less anemiated than Zebu Fulani and Boran.

The experimental design presented in the article does not bring robust elements on anemia control during *T. vivax* infection and on the comparison between N'Dama and WASH. There is not any susceptible breed that could be compared to N'Dama and WASH. It is thus not possible to know if the *T. vivax* strains are highly pathogenic or not. Since N'Dama and WASH are not raised in the same area under the same agro-ecological context, it is not possible to compare these two breeds.

Because only 4 and 7 animals were positive to *T. vivax* PCR, an Anova cannot be used. Only a non-parametric test can be used.

Figure 6- Expertise ouverte sur le site avec nom et texte

De plus, on peut même lire la réponse des auteurs aux critiques et autres remarques des experts. Cela prend la forme suivante :

Version 1

Author Response 10 août 2018
Paa Kobina

Responses to Second Reviewer's comments

1. Actual concentration of ethidium bromide and volumes of the PCR components have been provided in the methods section of the revised text, under the **Trypanosome detection** sub-section.
2. A revised **Dataset 1** has been provided in which the differential white blood cell counts (Neutrophils, Lymphocytes, Eosinophils, Monocytes and Basophils) are presented as both absolute numbers and %. **Table 1** in the revised text also has the differential white blood cell counts presented as both absolute numbers and %.

Response to Third Reviewer's comments

1. The comment on incomplete bibliography has been fully addressed in the revised text.
2. The comment on absence of susceptible breed that could be compared to N'Dama and WASH has been addressed in the revised text by providing data on trypanosusceptible Sanga and Zebu breeds that we had sampled together with the N'Dama and WASH but did not report in the initial text.
3. We agree with the comment that it is not possible to compare N'Dama and WASH since the two breeds were not raised in the same area under the same agro-ecological context. This aspect has consequently been removed from the revised text.
4. In the revised text, the non-parametric Kruskal-Wallis test, rather than ANOVA, was used.
5. The reviewer made some suggestions which have been incorporated in the revised text.

Normal haematological values in Table 1:
 For ease of reference, we have incorporated the normal values of haematological parameters for cattle (Jain, 1993) into **Table 1**.

Competing interests: None Close

Figure 7-Réponse de l'auteur à l'expertise

De plus, même un lecteur potentiel peut intervenir et émettre des remarques sur le travail s'il juge qu'il (elle) est assez spécialisé pour intervenir et sous une rubrique "respond or comment "(répondre ou commenter) Cela prend la forme suivante :

Respond or Comment

Your "User Comment" will be automatically labelled Author Response, Referee Response or Reader Comment. User comments must be in English, comprehensible and relevant to the article under discussion. We reserve the right to remove any comments that we consider to be inappropriate, offensive or otherwise in breach of the User Comment Terms and Conditions. Commenters must not use a comment for personal attacks. When criticisms of the article are based on unpublished data, the data should be made available.

Competing Interests
 Please disclose any competing interests that might be construed to influence your judgment of the article's or referee responses' validity or importance.

Post

Figure 8- Exemple de remarques potentielles d'un lecteur

Enfin tout ce processus peut être résumé dans "le tableau de bord" de la soumission (l'expression est de nous) et qui résume toutes les étapes, versions, expertises et réponses dans la figure suivante :



Figure 9- " Tableau de bord" complet d'une soumission

11.2 Biology Direct : quand l'auteur choisit ses propres experts



Figure 10- Page d'accueil de Biology Direct

Biology Direct se présente comme [un site] " servant la communauté de la recherche en science de la vie étant un journal pratiquant le contrôle par les pairs ouvert fournissant aux auteurs et lecteurs une alternative au modèle traditionnel du contrôle par les pairs "(Page d'accueil du site).Ce préambule dans la page d'accueil démontre la politique de *Biology Direct* qui se veut à l'avant-garde des nouvelles fonctionnalités offertes par l'Internet et appliquées au contrôle par les pairs. Il faut dire que le processus est des plus importants dans des spécialités telles que la biologie et la médecine. En effet, une formule, un médicament, un protocole médical proposé ne peuvent être appliqués à la légère et le contrôle est autrement plus important que dans les sciences sociales, par exemple, où une théorie, un avis ou une orientation idéologique n'ont pas d'effets *immédiats* sur la santé publique ni le bien être et devenir de la population.

11.2.1 Le processus en détail

Le processus de contrôle par les pairs à *Biology Direct* est complètement ouvert et comporte les étapes indispensables suivantes :

-L'auteur soumet son article à travers le système de soumission en ligne .Les responsables du processus font une vérification de base quant à l'acceptabilité scientifique de la soumission.

-L'auteur *choisit lui-même* ses experts (au nombre de huit) à partir d'une liste du comité éditorial. Les responsables du processus contactent alors les experts choisis au nom de l'auteur et coordonnent le processus.

-La soumission est acceptée pour publication si trois experts désignés retournent leurs expertises. -Le nom des experts et leurs expertises sont publiés avec tout article publié dans *Biology Direct*.

Ces étapes de base sont cependant entrecoupées par des étapes intermédiaires donnant au processus toute son originalité. Elles sont de deux sortes : pour les experts désignés et pour les auteurs.

Pour les experts désignés :

Une fois l'expert accepte d'expertiser, il doit aussi accepter que son nom et le contenu de son expertise figure avec l'article. Les expertises peuvent être très critiques et même très négatives mais cela n'empêche pas la publication qui reste du ressort de l'auteur 4.L'auteur est le seul responsable de sa soumission : contrairement au schéma traditionnel où un éditeur décide du sort de la soumission, dans *BiologyDirect* c'est l'auteur qui décide s'il publie, révisé ou retire sa publication. Les experts désignés doivent déclarer ne pas avoir de conflit d'intérêts avec les auteurs.

Une expertise en deux étapes

La première est purement technique et est une invitation à expertiser. Accepter d'expertiser veut dire que l'expert approuve une acceptation et publication éventuelles. Il peut aussi (s'il juge qu'il ne peut expertiser le travail) désigner un autre expert. Dans le cas ce dernier accepte, son nom figurera avec le nom de cet expert. *Dans la deuxième*, l'expert fournit des commentaires relatifs au travail à expertiser. Il peut aussi ne pas commenter dans quel cas il sera indiqué *sous son nom* la mention suivante: " Cet expert n'a pas fait de commentaires à publier ".Le rapport d'expertise est essentiellement composé de questions à travers le système d'expertise en ligne. Il y est demandé à l'expert soit :

-d'*endosser* la publication (tout en sachant que leur expertise aussi négative soit elle sera publiée avec l'éventuelle publication)

-*rejeter* la soumission comme étant non scientifique. Dans ce cas, les éditeurs de sections prendront une décision de concert avec les éditeurs en chef de *Biology Direct*.

Après cela, experts rédigent leur expertise qui doit comporter ;

-résumé de l'expertise : c'est un bref aperçu de la soumission qui donne un avis sur l'originalité, la signification et l'originalité de la soumission

-recommandations de l'expert : elles doivent préciser les faiblesses et être constructives pour l'auteur afin de corriger d'éventuelles failles. Elles sont départagées en lacunes " mineures " et " majeures ". **-lacunes mineures** : cette section est utilisée afin de signaler, par exemple, des erreurs d'orthographe ou typographiques. Elle n'apparaît pas avec l'expertise.

-commentaires confidentiels expert-éditeur : les experts peuvent aussi faire des commentaires confidentiels à l'éditeur concernant des problèmes d'éthiques ou politiques qui ne sont pas publiés avec l'article soumis à expertise

Les experts sont priés de remettre leur expertise dans un délai ne dépassant pas les 14 jours.

Les auteurs ayant reçu un minimum 03 expertises, peuvent décider de :

-réviser leur soumission et répondre aux commentaires des experts

-publier sans prendre en ligne de compte les commentaires des experts. Dans ce cas, les expertises seront publiées avec l'article afin que toutes les recommandations et/ou problèmes éventuels soient signalés aux lecteurs.

-retirer leur soumission plutôt que répondre aux recommandations ou publier sans répondre aux recommandations.

En dernier lieu, une fois les commentaires adressés par les auteurs, les experts ont deux options :

-se déclarer satisfait des réponses et ne font pas de commentaires additionnels

Les experts ont aussi la possibilité de refuser le travail s'il n'est pas de substance scientifique ou ne répond pas aux critères de scientificité.

-faire des recommandations supplémentaires auxquelles les auteurs ne sont pas obligés de répondre.

Cependant, toutes ces recommandations seront publiées si l'auteur décide de publier comme il en a le droit.

Cette dernière étape du processus doit être complétée par les experts dans un délai de 07 jours.

-Pour les auteurs

"*Biology Direct* opère un modèle dont l'auteur est l'acteur principal ainsi qu'un contrôle par les pairs complètement transparent". C'est en ces termes que le site présente le modèle de contrôle par les pairs prôné par *Biology Direct*. Il donne ensuite les étapes à suivre et qui sont résumées dans ce qui suit :

-sujets couverts : l'article doit s'insérer dans les 08 sections couvertes par *Biology Direct*.

-vérification technique : avant d'être envoyé en expertise, le comité éditorial fait une vérification de base afin de voir est ce que le travail adhère aux lignes directrices de soumission.

-Sélection d'experts : le soumissionnaire doit choisir huit experts parmi les membres du comité éditorial. Ces experts désignés ne doivent pas être des co-auteurs ou des collègues proches. A noter que le comité éditorial et les éditeurs de sections peuvent désigner d'autres experts.

Le contrôle par les pairs :

Le processus étant guidé par l'auteur, il peut, comme souligné auparavant, décider soit de publier en prenant en ligne de compte les commentaires, publier sans les prendre en ligne de compte ou alors retirer sa soumission. S'il décide de soumettre en incluant les commentaires des experts, il procède comme suit : Il doit insérer une section " experts "comme paragraphe juste après le résumé qui aurait la forme suivante : "Experts : cet article a été expertisé par XX, YY, ZZ (NB : si un expert a été désigné par un membre du comité éditorial, la forme serait : XX désigné par AA)

Il doit aussi insérer une section "contrôle par les pairs ouvert" dans une nouvelle page entre le résumé et l'introduction qui aurait la forme suivante :

" Contrôle par les pairs ouvert : cet article a été expertisé par XX, YY, ZZ (NB : si un expert a été désigné par un membre du comité éditorial, la forme serait : XX désigné par AA). Pour la totalité des expertises, voir la section des commentaires d'experts ".

Insérer ensuite les commentaires des experts avec leurs noms sous l'en-tête " commentaires d'experts " à la fin du manuscrit avant les références bibliographiques. Elle aurait la forme suivante :

"Commentaires d'experts. Rapport d'expertise 1 : Nom, affiliation (désigné par Nom, affiliation le cas échéant). Commentaires d'experts : ----- ". NB : si l'expert n'a pas fourni d'expertise, indiquer : " Cet expert n'a fourni aucun commentaire à publier ".

Dans le cas l'auteur veut répondre aux commentaires d'un expert ; il les insère précédés de la mention réponse d'auteur.

Ci-dessous une vue d'un échange entre auteurs et experts :

Reviewer comments: R1_2 (3), "For each query sequence, a genome in the reference database with the minimum score is selected as its source genome" what score? You never mentioned any score. How is it being calculated?"

Author's response: *Sorry for the misunderstanding. To avoid this misunderstanding, we have revised the manuscript to use "similarity score" instead of "distance", "similarity", or "score". The similarity score represents the similarity between a query sequence and a genome in the reference database. It can be calculated by formula (2).*

Figure 11- Exemple d'échange entre expert et auteur

11.3 *ResearchIdeas and Outcomes* (R.I.O.) ; une nouvelle manière de faire la science



Figure 12- Page d'accueil de RIO

ResearchIdeas and Outcomes (R.I.O.) est un journal multidisciplinaire qui pratique le contrôle par les pairs ouvert et de plus introduit des innovations qui en font un modèle de transparence dans toutes les étapes du cycle de publication et en particulier dans le contrôle par les pairs qui, on le verra, prend avec R.I.O. une toute autre dimension. Le processus est divisé en deux parties et qui sont :

11.3.1 Contrôle par les pairs pré-soumission :

Alors que généralement, la soumission est contrôlée *après la soumission*, dans RIO il est possible de faire faire le processus avant même la soumission. Ceci est possible grâce à ARPHA (AuthoringReviewingPublishingHostingArchiving - Créer Contrôler Publier Héberger Archiver). La nouveauté et l'innovation dans ARPHA réside dans ce contrôle par les pairs *pré-soumission* et qui applique une action faite habituellement entre collègues et la rend publique .En effet, une fois l'auteur prêt a soumettre, il se connecte à son compte ARPHA et

demande d'une manière transparente et formelle à un ou plusieurs collègues de donner leur avis sur la soumission en l'état actuel. ARPHA fournit des fonctionnalités du genre Google Doc afin de permettre cette opération. Le (s) collègue (s) (appelé *expert pré soumission*) se connecte à ce stade à son compte ARPHA, lit le travail et remplit un formulaire déclarant le travail publiable. L'expertise finale/commentaires et le nom de l'expert apparaissent avec le travail après publication. Malgré cela, RIO se réserve le droit de rejeter un travail pseudo-scientifique, présentant des problèmes éthiques même si l'*expert pré soumission* a émis une expertise positive. Tout ce cycle est résumé dans le schéma suivant :

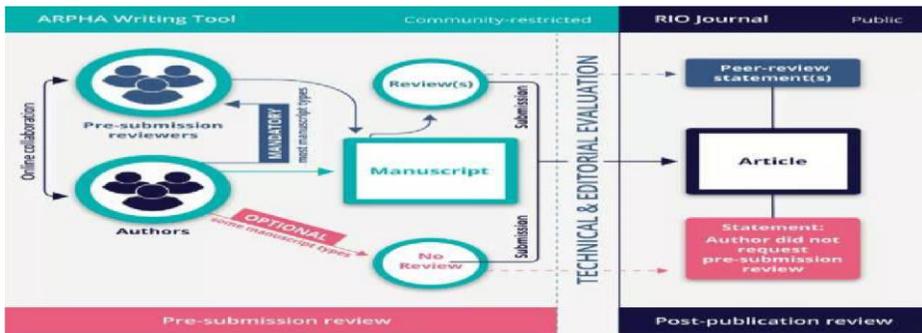


Figure 13- Etape "expertise pré soumission »

7.3.2-Contrôle par les pairs post publication

Après avoir passé le contrôle par les pairs *pré soumission* ainsi qu'obtenu l'aval de RIO, la soumission est sujette à un contrôle par les pairs *post publication*. Ceci permettra, entre autres, aux soumissionnaires d'établir la priorité sur les extrants (d'où *output* dans l'appellation du journal). Il est à signaler qu'il n'existe pas de choix quant à la publication des expertises qu'elles soient pré ou post publication : elles font partie de la publication. En fait, cette étape est la troisième du processus étant donné que le contrôle *pre soumission* et le contrôle interne de RIO en sont les deux premières. Le processus se divise en deux étapes et qui sont :

1-les auteurs peuvent demander la validation interne post publication par le journal ou alors garder la publication en sa version originale soumise après contrôle *pré soumission* et *contrôle interne de RIO*

2- les auteurs peuvent aussi demander la publication des versions révisées. Ceci leur permettra d'améliorer leur travail si les expertises post publication sont négatives. Seul RIO peut donner l'approbation mais seulement après

qu'un nombre suffisant et acceptable d'expertises post publication aient été reçus. Cette étape prend la forme suivante :

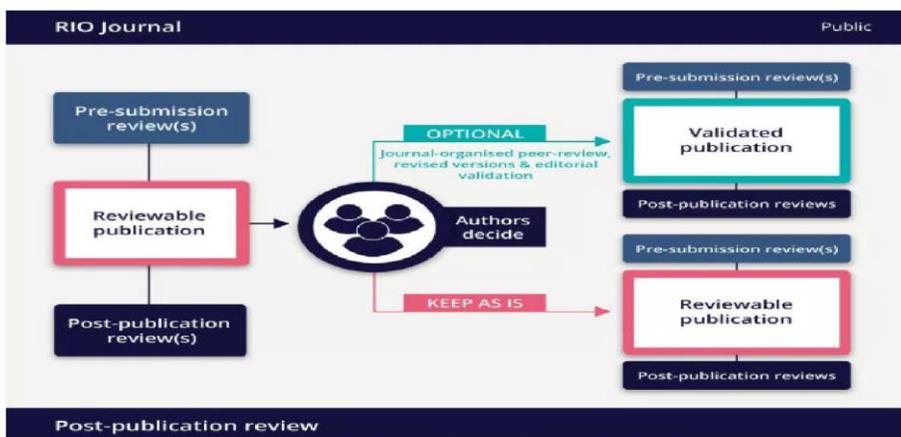


Figure 14- Expertise post publication

D'un autre coté, on peut voir un exemple de contrôle ouvert et ceci en cliquant sur un article donné. Prenons l'exemple de : "*Effectiveness of peer-mediated learning for English language learners: A meta-analysis*". Cet article a eu deux 👍 (expertises positives) pour contrôle par les pairs pré soumission. Les deux experts ont leur expertise en entier à la suite de leur nom. Elle peut être lue en cliquant sur "seereview"

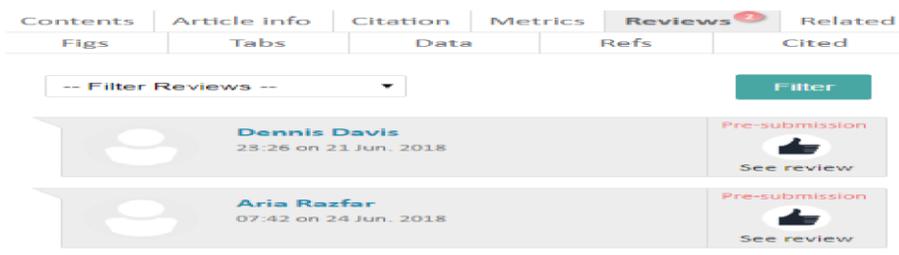


Figure 15- Exemple d'expertise ouverte post publication

Ils ont tous les deux donné leur accord pour la publication tel qu'indiqué dans la figure suivante :



Figure 16- Exemple de soumission acceptée

D'un autre coté, l'article "*Gravitationalwavespeed:undefined. Experimentsproposed*" a eu deux  (expertises négatives) pour contrôle par les pairs pré soumission. Les deux experts ont leur expertise en entier à la suite de leur nom. Elle peut être lue en cliquant sur "seereview".



Figure 17- Exemple de soumission refusée

Les deux experts ont tous les deux refusé la soumission tel qu'indiqué dans la figure suivante :



Figure 18- Exemple de soumission refusée

De même que l'auteur à la suite de l'expertise a répondu à certains questionnements émis par l'expert :

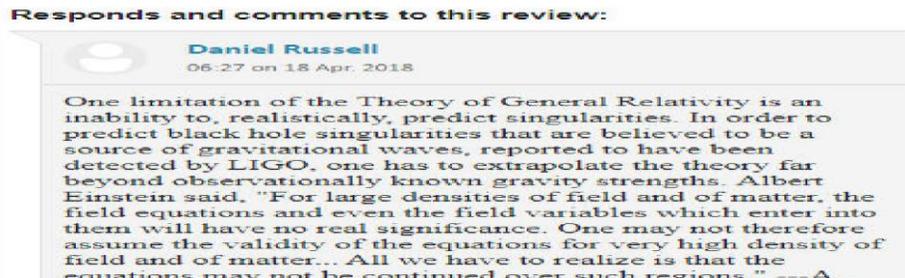


Figure 19- réponse de l'auteur à l'expert

De même que l'auteur a répondu au deuxième expert d'une manière plutôt brusque où on sent l'agacement de l'auteur (sentiment que tout chercheur a éprouvé mais qu'il ou elle ne peut exprimer dans le schéma traditionnel. Qu'on en juge !

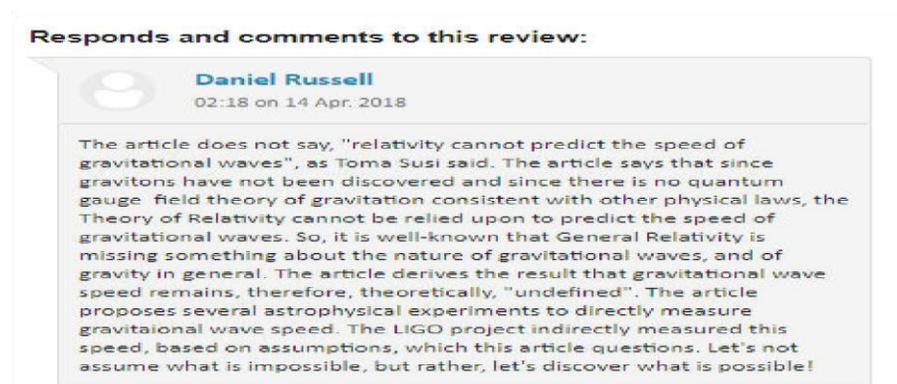


Figure 20- Réponse de l'auteur à l'expert

Le travail à expertiser peut être aussi accepté après révisions . C'est le cas de "*Origin and diversification of hoverflies: a revision of the genera Asarkina and Allobaccha – A BIG4 Consortium PhD project*". Dans ce cas, après l'expertise, l'expert accepte la soumission mais émet des réserves. Cela prend la forme suivante :

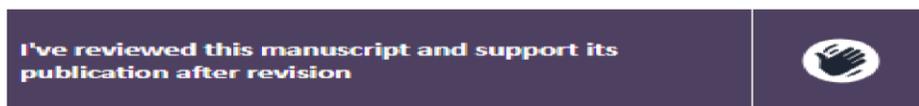


Figure 21-Exemple de soumission acceptée avec réserves

D'un autre côté, certains articles sont sous le label "reviewable" (expertisable). Ils n'ont pas encore eu d'expertises mais les experts pré soumission sont cités nommément. Pour l'article "*Impaired cognitive processes influence expressive language skills in attention deficits*", cela prend la forme suivante :



Figure 22- Exemple d'expertise pré soumission avec nom de l'expert

Conclusion

Le contrôle par les pairs a été avec l'Homme depuis qu'il a commencé à essayer de communiquer. Cette action aussi anodine soit elle implique un *feed back* que nous faisons tout le temps sans en être conscient. Ce processus est, dans la sphère scientifique et académique, une opération des plus importantes et indispensables. En effet, la science se bâtit par incrémentations et chaque " pierre " doit être solide pour supporter la suite de la construction. Cette analogie à la construction est des plus pertinentes car depuis la nuit des temps, la science a progressé en s'appuyant sur ce qui a été accepté, normalisé et intériorisé par la communauté pour avancer. Les théories, innovations et autres créations ont du être acceptées pour devenir la norme. Certaines de ces innovations ont été plus tard supplantées car la science découvre chaque jour et se régénère. La publication scientifique périodique contrôlée par les pairs, moteur indispensable de la science publiée, n'a pas accepté ni appliqué le contrôle par les pairs que lorsque la masse scientifique produite est devenue exponentielle et ingérable et que le volume produit devait être freiné. Le contrôle par les pairs a toujours eu deux fonctions principales : une fonction quantitative et une autre qualitative. La première avait pour but de freiner l'explosion documentaire tant décrite alors que la deuxième sert de tamis entre ce qui est publiable et ce qui est non publiable (avec des variations entre les deux pôles). Ce tamis qualitatif a toujours été sujet à des controverses et des critiques qui ont en fait un sujet ardemment discuté et disputé. La raison est que le contrôle par les pairs est considéré comme la porte menant à des avantages pécuniaires, des financements, des promotions et autres postes auxquels chaque chercheur veut accéder. Cette compétition entre chercheurs dont certains (les experts) sont investis du rôle (quasi mystique) de Gardiens du Temple a donné lieu à des problèmes. Cette problématique a engendré de nombreuses recherches ayant toutes conclut à l'inéluctabilité d'un changement d'un processus qui, dans sa forme originelle, ne donnait pas satisfaction quoique ces mêmes recherches aient aussi conclut que le processus est indispensable pour la bonne marche de la science. Alors que le débat battait son plein, l'irruption d'Internet et sa propension à l'ubiquité, l'omniprésence et quasi instantanéité apparut. Cette révolution qui de nos jours couvre plus de la moitié de la population mondiale a été utilisé pour ouvrir un processus connu pour sa tendance à être secret. D'une opération quasi mystérieuse et méconnue, propice à tous les dépassements,

elle est devenue une opération qui se passe en direct avec tous les protagonistes connus par nom, expertise et aussi dans un climat d'ouverture des plus totale. Les sites présentés représentent l'avant-garde car le processus est totalement transparent et surtout interactif laissant voir, dans certains cas, une sorte de conversation entre experts, auteurs et éditeurs qui n'aurait pas été imaginable quelques deux à trois décennies en arrière. L'expert dans ces schémas n'est plus l'omnipotent Gardien du Temple qui n'aurait de comptes à rendre qu'à l'éditeur (et accessoirement sa probité et sa conscience) mais est redevable devant toute la communauté de ce qu'il émet comme imprimatur. Bien que les expériences, aussi bien celle présentées que d'autres, soient encore récentes et même sujet elles-mêmes de nouveaux problèmes non connus dans l'univers papier traditionnel comme la création par des chercheurs véreux d'adresses mails fictives utilisées pour expertiser les soumissions de leurs propres collègues et même leur propres soumissions ! (The Guardian 2014), il n'n demeure pas moins qu'elles ont introduit une dose d'équité et de justice qui manquaient dans le système traditionnel fermé. De même que cette ouverture bien que réclamée à cors et a cris par la communauté n'est pas sans créer de nouvelles donnes comme celle concernant les experts non confirmés qui hésiteraient à critiquer un travail présenté par un ténor de la spécialité, problème qui ne se posait pas dans le système traditionnel. Malgré ces hésitations, l'avenir est à ce genre d'expérimentations qui, dans un monde de plus en plus ouvert, ne peut accepter que l'on soit jugé et connaitre son verdict sans connaitre son juge. Ce nouvel *Weltanschauung* est appelé à devenir la norme dans un monde académique aux ressources de plus en plus disponibles et ouvertes.

Bibliographie

AJLOUNI, Kamel M., AL-KHALIDI, Usama (1997). Medical records, patient outcome, and peer review in eleventh century Arab medicine, *Annals of Saudi Medicine*, 17 (03), p.326-327

ANDRE, Francis (2005). *Libre accès aux savoirs = open access to knowledge*. Paris: Futuribles. 73 p.

CRANE, Diana (1967). The Gatekeepers of science: some factors affecting the selection of articles for scientific journals,

The American sociologist, 2 (4), 195-201

DE SOLLA PRICE, Derek J. (1963). *Little Science, Big Science*. New York: Columbia University Press. 301p.

GUÉDON, Jean Claude, *In Oldenburg's long shadow: librarians, research scientists, publishers, and the control of scientific publishing*, In: *Creating the digital future*, Proceedings of 138th ARL Membership Meeting, Toronto [Ontario], May 23 - 25 2001, Baker, Shirley K. Ed.

HARGENS, Lowell L. (1988). Scholarly consensus and journal rejection rates, *American Sociological Review*, 53, (1), p. 139-151

HARNAD, Stevan (2000). The Invisible hand of peer review, *Exploit Interactive*, 5 [Enligne]. Disponible à : <https://archipel.uqam.ca/188/1/nature2.html> (Page consultée le 22 Aout 2018)

RENNIE, Drummond. *Editorial peer review: its development and rationale*, p.2 In: *Peer review in health sciences*, Goodlee, Fiona and Jefferson, Tom Eds., BMJ Books, London, 2003, VIII-392 p.

PETERS, Douglas P., CECI, Stephen J. (1982). Peer-review practices of psychological journals: the fate of published articles, *Behavioral and Brain Sciences*, 5 (2) , p.187-195

SHATZ, David (2004). *Peer review: a critical inquiry*, Lanham [Md., USA]: Rowman & Littlefield, 249 p.

SPIER, Ray (2002). The History of the peer review process in science, *Trends in biotechnology*, Vol.20, n°8, p.357-358

SWAINE, Jon (2014). Academic journal retracts articles over 'peer review ring' with bogus scholar, *The Guardian*, July 10th 2014 [enligne] Disponible à : <https://www.theguardian.com/media/2014/jul/10/academic-journal-retractsarticles-peer-review-ring> (Page consultée le 28 Aout 2018)

TAX, Sol (1959 a). Letter to associates. *Current Anthropology*, 1 (Pre issue), p.8-9

TAX, Sol (1959 b). Review articles. *Current Anthropology*, 1(Pre issue), p. 3

What science has done ? (2011).The Humble beginnings of peer review alternatives: Stevan Hanard and Open Peer Commentary [enligne] Disponible à :
<https://richardfmasters.wordpress.com/2011/06/02/the-humblebeginnings-of-peer-review-alternatives-stevan-hanard-and-open-peer-commentary/> (Page consultée le 22 Aout 2018)

WELLER, Ann C. (2002). *Editorial peer review: its strengths and weaknesses*. Medford [N.J., USA]: Information Today. 344 p.

ZIMAN, J. (1968). *Public knowledge: an essay concerning the social dimension of science*, Cambridge: Cambridge University Press.168p.

Webographie

<https://www.internetworldstats.com/stats.htm> (Page consultée le 22 Aout 2018) <https://f1000.com/> (Page consultée le 22 Aout 2018)

<https://f1000research.com/browse/articles?&selectedDomain=articles> (Page consultée le 26 Aout 2018)

<https://f1000research.com/browse/f1000-faculty-reviews?&selectedDomain=f1000-faculty-reviews> (Page consultée le 26 Aout 2018)

<https://f1000research.com/author/version/17615> (Page consultée le 26 Aout 2018) <https://f1000research.com/for-authors/tips-for-finding-referees> (Page consultée le 26 Aout 2018)

<https://f1000research.com/articles/7-314/v2> (Page consultée le 26 Aout 2018)

<https://biologydirect.biomedcentral.com/> (Page consultée le 27 Aout 2018) <https://biologydirect.biomedcentral.com/about/how-it-works> (Page consultée le 28 Aout 2018)

<https://biologydirect.biomedcentral.com/articles/sections> (Page

consultée le 28 Aout 2018)

<https://biologydirect.biomedcentral.com/submission-guidelines> (Page consultée le 28 Aout 2018)

<https://biologydirect.biomedcentral.com/about/editorial-board> (Page consultée le 28 Aout 2018)

<https://biologydirect.biomedcentral.com/articles/10.1186/s13062-018-0220-y#Sec20> (Page consultée le 28 Aout 2018) <https://riojournal.com>

(Page consultée le 28 Aout 2018) <http://arphahub.com/> (Page consultée le 28 Aout 2018) <http://blog.riojournal.com/2015/09/17/peer-review-at-rio-part-1/> (Page consultée le 28 Aout 2018)

<http://blog.riojournal.com/2015/09/23/peer-review-at-rio-part-2/> (Page consultée le 28 Aout 2018) <https://riojournal.com/articles> (Page

consultée le 28 Aout 2018) <https://riojournal.com/article/29375/list/17/> (Page consultée le 28 Aout 2018)

<https://riojournal.com/article/29375/element/17/24074/> (Page consultée le 28 Aout 2018)

<https://riojournal.com/article/25606/list/17/è^p-t> (Page consultée le 28 Aout 2018) <https://riojournal.com/article/25606/list/17/> (Page

consultée le 28 Aout 2018)

Publication scientifique marocaine en computer science entre l'Open Access et l'accès payant : état des lieux et perspectives

Moroccan scientific publication in computer science between Open Access and paid access

Rhassate, Imane, Mohammadia School Engineers (EMI), Mohammed Vth
University, Rabat, Morocco – imane.rhassate@gmail.com, ORCID ID:
<https://orcid.org/0000-0003-0905-3777>

Amrous, Naila, LaboratoireMediation-Information-Knowledge-Society (MIKS)
ESI, Rabat, amrousnaila@yahoo.fr

Bennani, Samir, Mohammadia School Engineers (EMI), Mohammed Vth University,
Rabat, Morocco – sbennani@emi.ac.ma

Résumé

L'édition électronique a changé la façon dont la plupart des chercheurs accèdent à la littérature scientifique dans leur domaine. Après plus de 17 ans de son émergence, l'avènement de l'Open Access retient de plus en plus l'intérêt des chercheurs des différentes disciplines. Basé sur les recherches antérieures, les résultats de l'étude de cas, et l'analyse de données de la base « Scopus », l'article présente un état des lieux des publications en libre accès des chercheurs marocains en « Computer Science », mesure l'écart qu'elles présentent par rapport à celles en édition payante en tenant compte de leur typologie, distribution, etc. L'approche du sujet est étayée par l'analyse du comportement des chercheurs marocains face à l'Open Access, leur degré d'implication, l'usage qu'ils en font, et les causes de leurs réticences présumées.

L'article conclut à la nécessité de sensibiliser les principaux acteurs de la publication scientifique à œuvrer dans le contexte du Libre Accès.

Mot clés. publication scientifique, computer science, libre accès, Maroc.

Abstract

Electronic publishing has changed the way most researchers access scientific literature in their field. After more than 17 years of its emergence, the advent of Open Access is increasingly attracting the interest of researchers from different disciplines. Based on previous research, the results of the case study, and the analysis of data from the "Scopus" database, the article presents an inventory of open access publications of Moroccan researchers in "Computer Science", measures the gap they present compared to those in paid publishing taking into account their typology, distribution, etc. The subject's approach is supported by an analysis of the Moroccan researchers' behavior with regard to Open Access, their degree of involvement, the use they make of it, and the causes of their alleged reluctance.

The article concludes that it is necessary to raise awareness among the main actors in scientific publishing to work in the context of Open Access.

Keywords. Scientific publishing, computer science, Open access, Morocco.

1 Introduction

Les publications scientifiques semblent être le canal le plus utilisé par les chercheurs pour partager leurs connaissances et communiquer leurs résultats de recherche. Néanmoins, les prix d'abonnement aux journaux scientifiques ont explosé dans les 30 dernières années. Ce qui a mené les universités et les établissements de recherche à abandonner leurs abonnements à plusieurs périodiques au profit des sources en libre accès.

C'est dans un contexte caractérisé par une quête continue de l'excellence scientifique et par la volonté d'édifier une société de la connaissance dans laquelle chacun a la possibilité de créer, d'obtenir, d'utiliser et de partager l'information et le savoir, que le mouvement de l'Open Access a vu le jour. C'est en effet au début du XXIème siècle qu'il a été baptisé comme étant «La pratique qui consiste à fournir un accès en ligne à des informations scientifiques gratuites et réutilisables pour l'utilisateur final» [1].

Au sein de la communauté scientifique, la production a, depuis toujours, constitué le reflet réel de l'activité de la recherche, d'où la nécessité de sa communication. *Agostini* considère que "sans publication la science est morte"[2] et la science n'est autre chose que le résultat de la recherche des scientifiques. De plus, la publication des résultats de la recherche constitue la forme de transmission la plus répandue et de ce fait, elle constitue l'objectif même de la recherche scientifique. Un chercheur est généralement évalué par ses publications ; sa promotion professionnelle, sa carrière et sa valeur scientifique en dépendent. Cette évaluation dépend de la quantité et qualité de ses publications scientifiques, de leur apparition dans des revues de notoriété reconnue et, depuis près de deux décennies, de leur présence dans des bases de données en Libre Accès. Pour le chercheur la publication constitue un des principaux modes de validation et de diffusion pour ses travaux et leur présence en ligne en augmente la visibilité et l'impact et en assure une économie d'échelle.

Dans la perspective de profiter des avantages de l'OA et de dépasser les barrières financières de la publication traditionnelle, le Maroc ne devrait-il pas s'orienter vers le Libre Accès à l'information et aux résultats de la recherche ?

C'est dans ce cadre que s'inscrit notre recherche qui a pour objectifs :

- de dresser, dans un premier temps, l'état des lieux des publications scientifiques marocaines en computer science en accès libre, et, de préciser leur volume, leurs langues de publication ainsi que leurs auteurs et universités pionniers ;
- de présenter, dans un deuxième temps, les résultats d'une enquête sur le terrain que nous avons menée afin d'approcher le rapport

qu'entretient le chercheur marocain en "Computer Science" avec l'OA et ce, à travers la prospection de son niveau de sensibilisation, de connaissance, d'usage et de promotion de l'Information Scientifique et technique (IST).

Notre hypothèse étant : si les articles en Open Access sont facilement accessibles, ils seraient plus utilisés par les chercheurs aussi bien dans la réalisation de leurs travaux que dans la diffusion de résultats qui en découlent.

Les paragraphes qui suivent exposent les méthodes utilisées pour l'atteinte de ces objectifs, suivies d'une brève revue de la littérature et d'une description des données recueillies. Les résultats sont ensuite présentés, pour chaque objectif de recherche, et sont suivis d'une discussion et conclusion générales.

2 Méthodologie

La présente recherche à propos de la production scientifique marocaine dans le domaine du computer science en accès libre se donne deux objectifs. En premier lieu, quantifier la part de la production scientifique marocaine en accès libre dans ce domaine. En second lieu, aborder la relation entre le chercheur marocain en Computer Science "CS" et l'accès libre à l'information en termes de recherche, d'usage et de communication.

Ainsi, des objectifs tracés découlent un ensemble de questions de recherches : d'abord, est ce que le chercheur marocain en "CS" publie en Open Access ? Si oui, avec quel écart par rapport à la publication traditionnelle ? Sinon, quelles en sont les causes ?

Dans le même ordre d'idées, nos chercheurs sont-ils encouragés par leurs tutelles, bailleurs de fonds (bourses) et/ou universités à donner accès à leurs travaux librement ? Et quels sont les mécanismes mis en place à cette fin ?

Pour répondre aux objectifs et questions de recherche, nous avons eu recours à deux méthodes : l'analyse de données secondaires et la collecte de données primaires par le biais de l'enquête sur le terrain.

A. L'analyse des données secondaires a porté sur les informations puisées dans la base de données «Scopus», en utilisant la stratégie de recherche suivante :

- l'identification du pays : Maroc / Morocco dans le champ « Affiliation country » ;
- la spécification des publications scientifiques dans le domaine du « computer science »

- la précision du type de document : Open Access ;
- la couverture longitudinale qui va de 1990¹ (année de la première publication en CS en accès libre à l'international) à juin 2018 (date de la génération des derniers résultats) ;
- l'analyse des données recueillies : par l'outil Google Sheets.

L'usage de la base de données « Scopus » est motivé par sa large couverture de plus de 21 500 titres de revues issues de sources validées par des pairs. C'est une base qui prend en compte la localisation géographique de la production scientifique. Elle est considérée par de nombreux spécialistes, comme une référence pour les activités de recherche scientifique et technique.

B. L'enquête sur le terrain a été réalisée par le biais d'un questionnaire électronique envoyé aux laboratoires et aux équipes de recherche doctorales de Rabat dont la spécialité de recherche porte sur le « Computer Science ». La couverture de l'ensemble de la population a été motivée par le souci d'élargir la perspective et couvrir l'ensemble des sources en Open Access.

Notons toutefois que cette recherche comporte quelques limites. Notre recueil d'informations s'est axé sur la base de données bibliographique « Scopus ». Combiner son usage avec d'autres bases de données -telles : le WoS, ou DOAJ- aurait permis d'évaluer, à grande échelle, la production des chercheurs marocains en OA dans le domaine du CS.

Une autre limite relève de la population enquêtée, dont le champ se limite à Rabat. Un échantillon plus étendu couvrant l'ensemble des universités et des villes aurait permis de mieux comprendre le phénomène étudié et de répondre aux questions de décalage accusé par la communauté des chercheurs marocains dans le domaine de CS par rapport à ses homologues aux deux niveaux régional et international.

3 Etat de l'art

La technologie entraîne des changements profonds dans la façon dont la recherche est menée et communiquée. De plus, la publication du contenu scientifique est un processus qui bénéficie au mieux du développement des TIC notamment de l'Internet. Ce dernier reste un canal vital de diffusion de l'information en général et de la production scientifique en particulier.

Le paragraphe suivant contient quelques précisions terminologiques, présente les raisons du choix du domaine «Computer Science» comme

¹Nos recherches sur Scopus donnent en effet l'année 1990 comme année de la première publication en OA dans le domaine de CS indexée dans cette base de données.

champ d'études, des éclaircissements sur la relation entre l'OA et la visibilité scientifique. Il contient enfin, une rétrospective des nouvelles tendances d'accès à l'information, notamment les réseaux sociaux académiques et les sites de piratage (sciHub). A noter que cette revue de la littérature est réalisée exclusivement sur la base des ressources accessibles en OA.

3.1 Terminologie

- Publication scientifique : un canal de communication et de diffusion de l'information scientifique valide. Elle couvre les articles, comptes-rendus de congrès, ouvrages collectifs, thèses, rapports de recherche, monographies... ;
- Open Access journals : journaux dont le modèle de financement ne nécessite pas de faire payer le lecteur ou l'institution pour y accéder. L'utilisateur final peut lire, télécharger, transmettre et/ou imprimer le texte intégral d'une publication [3] ;
- Pour cette recherche, nous utilisons le terme "accès payant" (également appelé "accès fermé") pour désigner la littérature payante [4].

3.2 Le domaine du « Computer Science », pourquoi ?

En matière de recherches, le « Computer Science » se situe comme l'un des domaines les plus actifs, bien qu'il soit considéré comme une activité récente qui trouve ses racines, fort anciennes, dans les travaux des mathématiciens et des électroniciens. Les spécialistes de ce domaine sont considérés parmi les pionniers de la communication scientifique numérique avant l'avènement même du World Wide Web [3]. En outre, depuis le lancement de l'Open Access, les chercheurs en mathématiques, physique, informatique et autres « sciences dures » [5] étaient les premiers à déposer leurs prépublications dans des réservoirs d'Open Access, tels ArXiv et HAL. Ceci a incité Steve Lawrence à se focaliser sur le domaine de Computer Science, afin de démontrer le lien positif entre le Libre Accès aux articles et le nombre de lectures et de citations qui en sont faites [6].

En sus, le CS figure parmi les dix (10) premiers domaines scientifiques les plus répandus pour les revues en OA [7] ; la plupart de ses articles publiés -sous abonnement- peuvent être accessibles également grâce à l'auto-archivage par l'auteur.

Pour ce domaine, les chercheurs soit auto-archivent leurs publications surtout dans leur site web, dont les références sont repérées par Citeseer, ou bien, ils font recours aux archives institutionnelles. L'avantage de ces

dernières est leur conformité au protocole de marquage de métadonnées de l'OAI (Open Archive Initiative) [8].

3.3 L'Open Access et la visibilité des publications scientifiques

L'usage de l'OA plus diversifié et plus facilité par les TIC se trouve être confronté actuellement à des barrières légales, technologiques et marketings [7]. Une autre barrière est liée à la citation et à l'impact de l'auteur. Les chercheurs se posent en effet, souvent, la question de : « si je publie dans une revue en Open Access, mon travail sera-t-il cité [9], autrement dit aura-t-il un impact ?

Les mécènes de l'OA confirment que les journaux à accès ouvert ont un impact semblable à celui des journaux payants. Les auteurs ne devraient aucunement pas craindre la publication dans ces journaux simplement à cause de leur modèle d'accès [9].

Concernant la question de la qualité des journaux en OA, une étude menée par l'éditeur *Thomson ISI* a montré qu'une bonne partie d'entre ces journaux se conforme à des critères rigoureux de sélection. Ils ont tendance à se comporter comme d'autres journaux de la collection de *Thomson ISI*. Certains d'entre eux sont au sommet même de leurs catégories disciplinaires [10].

En effet, plusieurs études prouvent que l'OA ne diminue pas de la qualité des journaux, bien au contraire il augmente le niveau de lectorat, renforce les citations et le taux d'impact. Il permet d'accroître la visibilité, facilite le repérage et l'accessibilité des publications scientifiques en général et ceux en Computer Science en particulier [11] [12]. L'OA a ainsi un effet significatif sur le comportement de citation [13] [14] [15].

Plus précisément, pour le domaine du Computer Science, *Harnad & Brody* (2004) avancent que l'Open Access a un effet avantageux sur les publications scientifiques surtout en termes d'amélioration d'usage et d'impact [16]. *Lawrence* (2001) a trouvé, aussi, que pour le même domaine, les citations sont 3 fois supérieures pour les articles en Open Access que pour ceux accessibles via accès payant [17].

En termes de citation, la majorité des auteurs cités précisent qu'en plus de la qualité, l'originalité et la pertinence de la recherche, le mode d'accès reste un facteur influent dans la mesure de l'usage d'une publication scientifique.

Par ailleurs, une étude de *Gretchen (2011)* confirme que les scientifiques aiment bien les articles en libre accès en tant que lecteurs, mais ils sont moins enthousiastes à l'idée de soumettre leurs articles à des revues à accès libre [18]. C'est parmi les points que nous avons essayé de confirmer ou d'infirmer pour les chercheurs marocains en CS par notre enquête sur le terrain.

3.4 Réseaux sociaux académiques, un nouvel élan pour l'accès libre aux publications

Au-delà des journaux en OA et des archives ouvertes, un intérêt pour la « Science 2.0 » a évolué de manière vertigineuse, au cours des dernières décennies, avec le but d'ouvrir la science à la société globale à des fins de développement économique et technologique. La science devient de plus en plus sensible aux besoins de la société [19].

En outre, la majorité des articles disponibles librement ne sont pas trouvables dans un réservoir ou un journal en Open Access, mais le sont plutôt dans les sites personnels des auteurs [20], ainsi que dans leurs pages sur les réseaux sociaux académiques comme ResearchGate[21] et Academia.eu [22].

Il s'agit de sites Web où les chercheurs peuvent publier leurs articles, chercher et lire les articles publiés par les autres. Ils combinent le rôle de dépôts d'archives avec des fonctions d'un site du réseau social qui permet aux universitaires de créer leur propre profil, de dresser la liste de leurs publications et d'interagir avec leurs pairs. L'inscription est généralement gratuite et les utilisateurs peuvent télécharger tous les documents affichés sur le site. Ces sites offrent une nouvelle façon pour les chercheurs de diffuser leurs travaux et, par conséquent, de changer la dynamique de la communication savante informelle.

Plusieurs études [23] ont démontré l'impact positif en matière de citations entre les articles publiés dans ces réseaux sociaux universitaires (ASN) et d'autres articles de revues similaires. Un article reçoit entre 16 % à 58 % plus de citations que les articles non disponibles en ligne ou uniquement sur d'autres sites comme les pages d'accueil personnelles ou départementales.

3.5 Les articles piratés : une autre facette de l'Open Access !

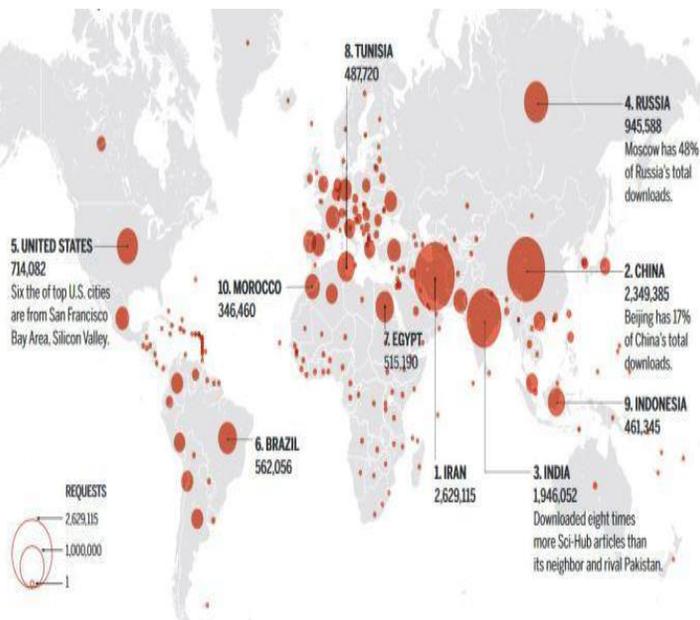
Afin de contourner la domination des trois quarts de la littérature savante par les éditeurs commerciaux [24] et en plus de l'usage de l'Open Access, nombreux sont les chercheurs qui font appel au service de sites web donnant accès au texte intégral des articles payants d'une façon illicite (tel : LibGen[25] et Sci-Hub).

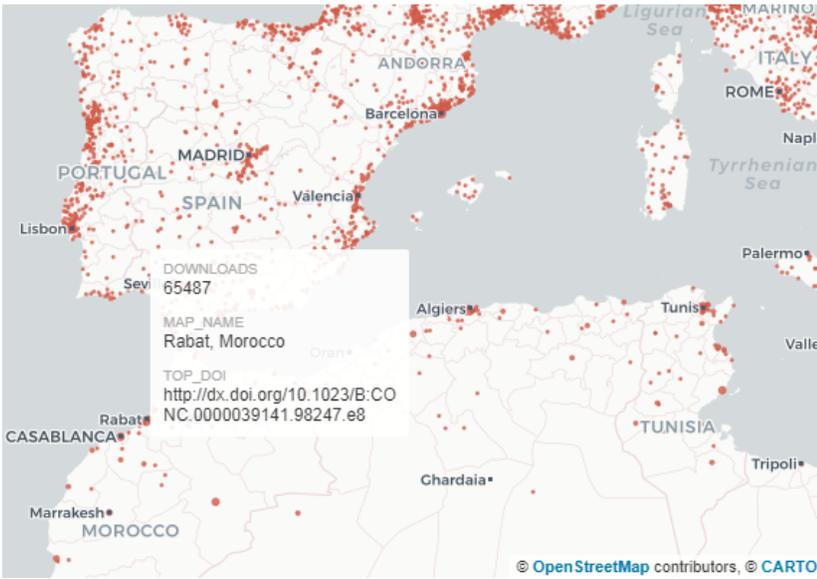
La base de données Sci-Hub, lancée en 2011, se présente comme "le premier site web pirate au monde à fournir un accès de masse et public à des dizaines de millions de documents de recherche" [26]. Cet outil héberge 62 millions d'articles (Mars 2018), c'est-à-dire à 85,1% des articles publiés dans les journaux payants [27].

La recherche que nous avons menée sur l'usage de la base de données Sci-Hub à l'échelle mondiale et à Rabat (Maroc) a montré que les utilisateurs de Sci-Hub proviennent de tous les continents, ce qui montre que le besoin en Open Access ne se limite pas au monde en développement, mais intéresse tous les pays indépendamment de leur rang économique (tels les États-Unis et la Russie). Aussi, la lecture de la mappe dessus montre-t-elle que la publication la plus téléchargée par les chercheurs marocains appartient au domaine du Computer Science. Il s'agit d'un acte de conférence publié dans « IEEE 9th Workshop on Multimedia Signal Processing ».

A la lumière de ce qui précède, notre recherche vise d'abord à dresser un état des lieux de la production scientifique marocaine dans le domaine du CS en OA et ensuite à analyser l'attitude du chercheur marocain à son égard que ce soit en matière de recherche d'informations, d'usage ou de communication scientifique.

Figure 1 : l'usage de la base de données Sci-Hub à l'échelle mondiale et à Rabat (Maroc)





Source : <http://www.sciencemag.org/news/2016/04/whos-downloading-pirated-papers-everyone>

4 Résultats et discussion

La partie suivante présente, dans un premier lieu, les résultats et la discussion à propos de l'analyse de la production scientifique marocaine dans le domaine du computer science en Open Access dans la base de données bibliographique « Scopus ». Dans un deuxième lieu, cette partie expose et discute les résultats de l'enquête sur le terrain concernant la relation entre le chercheur marocain en « CS » et le libre accès à la science.

4.1 La production scientifique marocaine dans le domaine du computer science en Open Access

Les données de notre recherche, puisées dans la base de données Scopus, couvrent pour le CS la période allant de 1990 – 2018. Est considérée ici comme publication, tout texte intégral librement disponible. Cela inclut : draft, preprint et postprint.

4.1.1 *Quantifier la part de la production scientifique marocaine en CS en Open Access*

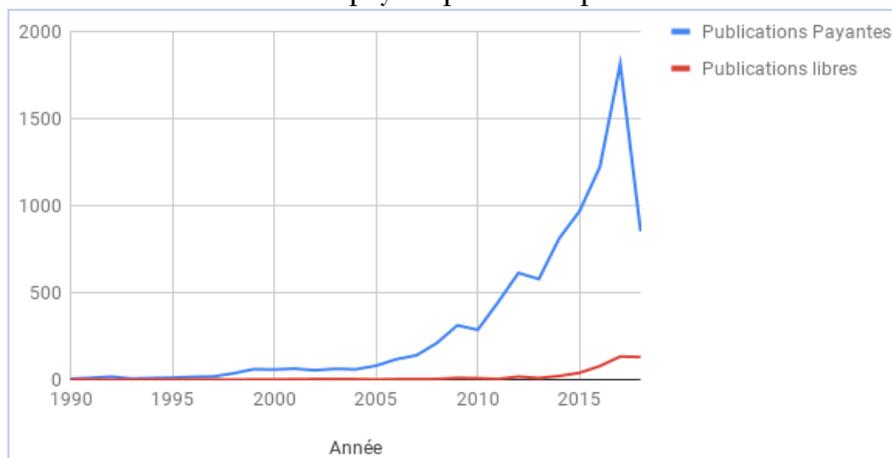
En prospectant toutes les publications scientifiques marocaines en computer science indexées dans la base de données bibliographiques Scopus, les résultats ont montré que la première publication marocaine parue dans une revue indexée par Scopus remonte à 1974, cependant la première publication dans le même domaine accessible en Open Access, indexée par Scopus, n'a été répertoriée qu'en 1990.

4.1.2 La production scientifique marocaine en CS en accès libre VS accès payant

A ce propos, sur un total de 8695 publications scientifiques marocaines, en Computer Science indexées par Scopus, seules 311 sont en OA, à raison de 3,58% (Avril 2018).

La lecture du graphe 1, que nous avons élaboré sur la base des données de scopus, montre l'existence d'un écart entre les publications payantes et celles en OA. Ces écarts deviennent de plus en plus distants au fil des années. A l'inverse de la tendance mondiale de la publication dans le domaine du CS, celle marocaine se voit plus orientée vers les journaux commerciaux que vers ceux libres.

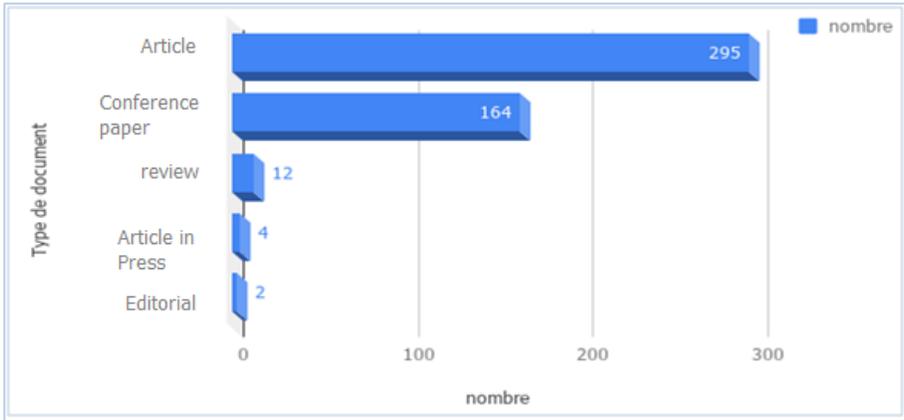
Graphe 1 : Comparaison entre les publications des chercheurs marocains en CS en accès libre et en accès payant pendant la période 1990-2018



4.1.3 Forme et langue de publication marocaines en CS en accès libre

Le graphe 2 stipule que l'article de revue est le canal le plus utilisé par les chercheurs marocains en Computer Science pour diffuser leurs résultats de recherche scientifique. Il est suivi par les communications de conférences.

Graphe 2 : typologie des publications marocaines en OA dans le domaine de computer science



Relativement à la question de la langue, la quasi-totalité des publications marocaines en OA dans le domaine du computer science sont en langue anglaise. En effet, sur les 477 publications marocaines en OA, seuls deux sont en français, le reste est en anglais. A noter que pour l'indexation de ses revues Scopus privilégie la langue anglaise au dépend des autres langues.

4.1.4 Où les chercheurs marocains en CS publient-ils ?

Le tableau 1 fait référence aux dix premières revues spécialisées en CS, disponible en libre accès, dans lesquelles les chercheurs marocains publient le plus. La mesure de citation desdites revues varient entre 3.86 et 0.25.

Tableau 1 : Top 10 des revues dans le domaine du CS en OA

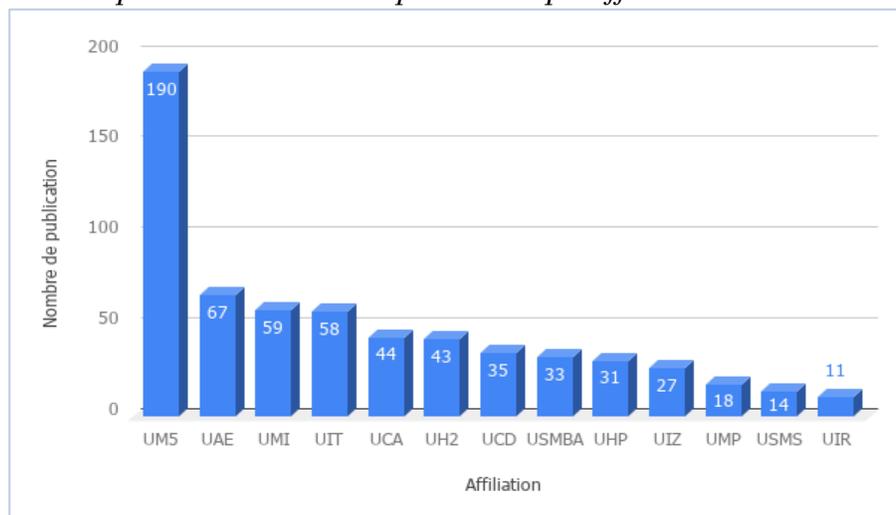
	Titre de la revue	Nombre d'articles	CiteScore [28]
1	Procedia Computer Science	140	1.03
2	International Journal Of Electrical And Computer Engineering	72	1.22
3	ProcediaManufacturing	22	0.70
4	Indonesian Journal Of Electrical Engineering And Computer Science	18	1.22
5	Mathematical And Computer Modelling	18	1.68
6	Eurasip Journal On Wireless Communications And Networking	15	1.73
7	Computers And Mathematics With Applications	14	1.89
8	Journal Of King Saud University Computer And Information Sciences	14	2.05

9	Journal Of Computer Networks And Communications	10	1.21
10	Modelling And Simulation In Engineering	10	0.25

4.1.5 Quel classement pour les universités marocaines ?

L'université Mohammed V de Rabat devance les autres universités par 190 publications en OA dans le domaine du CS. Elle est suivie, avec un grand écart, par l'université Abdelmalek Essadi de Tétouan par 67 publications, puis par l'université Moulay Ismail de Meknès avec 59 publications.

Graphe 3 : Distribution des publications par affiliation marocaine



4.1.6 Etat de la collaboration

L'analyse des résultats sur la collaboration, en termes de Co-publication entre les chercheurs dans le domaine du CS, montre que si les marocains publient en collaboration avec leurs homologues d'autres pays arabes et étrangers, ils préfèrent cependant coopérer avec leurs pairs marocains.

Tableau 2 : Top 10 des co-publications par pays

	Pays	Nombre de publications
1	Morocco	472
2	France	76
3	Spain	21

4	United States	12
5	Canada	10
6	Italy	10
7	United Kingdom	8
8	China	7
9	India	6
10	Saudi Arabia	6

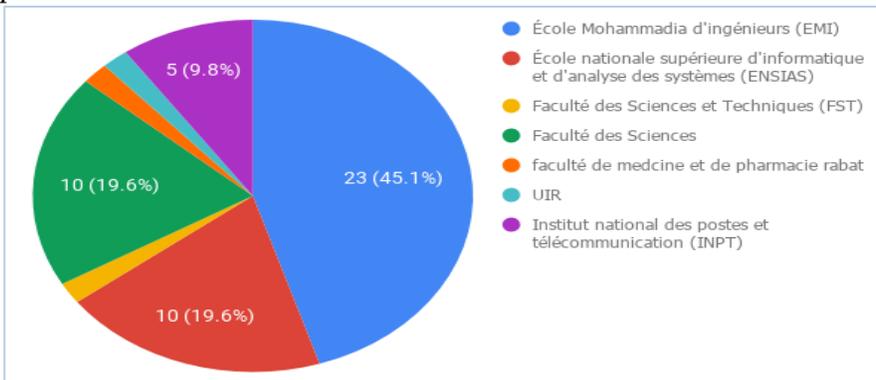
4.2 Le chercheur marocain en « Computer Science » et le libre accès à la science

En réponse à notre objectif relatif aux attitudes des chercheurs marocains à l'égard de l'OA, un questionnaire a été distribué auprès des chercheurs des laboratoires et équipes de recherche doctorales dans la ville de Rabat dont la spécialité de recherche est le « Computer Science ». Le nombre de réponse reçu s'élève à 59, comme cela ressort du graphe 4.

4.2.1 Caractéristiques de la population cible

La majorité des répondants sont des doctorants (88.7%), puis des professeurs assistants (5.7 %), des professeurs habilités (1.9 %) et des professeurs d'enseignement supérieur (1.9%).

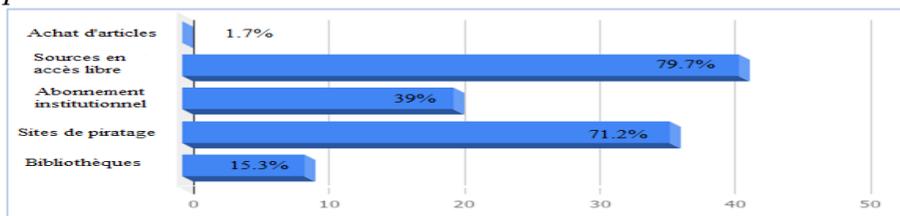
Graphe 4 : Institution d'appartenance des répondants à l'enquête sur les publications des marocains en OA



4.2.2 Accès aux sources d'information

Pour accéder à l'information scientifique, 79.3% de la population, objet de notre enquête, utilise des sources en accès libre, 67.9% d'entre eux font recours aux sites de piratage, 37.7% accèdent à l'information payante à travers l'abonnement de leurs institutions aux bases de données, et, en dernier lieu 16.9% utilisent le fonds documentaire des bibliothèques universitaires ou de recherche.

Graphe 5 : Sources d'information des répondants à l'enquête sur les publications des marocains en OA



4.2.3 Appréhension de l'OA par le chercheur marocain en CS

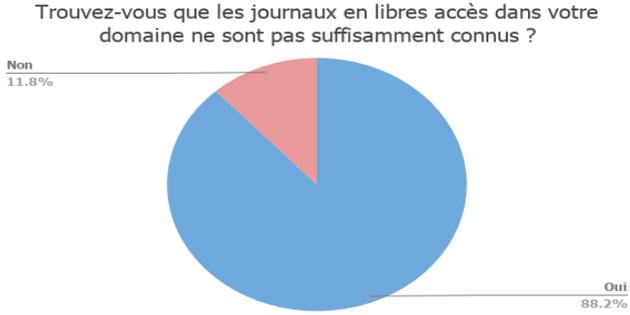
Comme indiqué par le graphe 7, un peu plus de la moitié de la population enquêtée connaît l'existence des journaux scientifiques en OA dans le domaine du « Computer Science ». Toutefois, connaissance ne signifie ni repérage ni usage. Aussi, l'autre moitié non au fait de l'OA devrait normalement être initiée à ce nouveau paradigme de communication combien précieux pour la recherche et l'innovation scientifiques.

Par surcroît, la majorité des chercheurs de la population, objet de notre recherche, n'a jamais publié un article dans une revue en accès libre (82,7%) et n'a aucunement déposé un travail de recherche, quel qu'en soit le type, dans un réservoir institutionnel d'archives ouvertes (92%).

Graphique 6 : Appréhension de l'OA par le chercheur marocain en CS

Connaissances des chercheurs marocains en CS des journaux scientifiques en libres accès





En considération de ce qui précède, parmi les causes de ce résultat, nous pouvons faire référence à la difficulté du repérage des revues scientifiques en CS ayant un accès libre, car 88,2% des chercheurs affirment que les tenants et aboutissants de ces dernières ne sont pas suffisamment maîtrisés.

4.2.4 Usage de l'OA par le chercheur marocain en CS

Le graphe 8 atteste que le lectorat (61%), la visibilité (51%) et l'accessibilité (38%) sont les raisons les plus motivantes des chercheurs marocains pour la publication en OA. De surcroît, la plupart d'entre eux

sont conscients de l'influence positive de l'OA sur le nombre citations des publications.

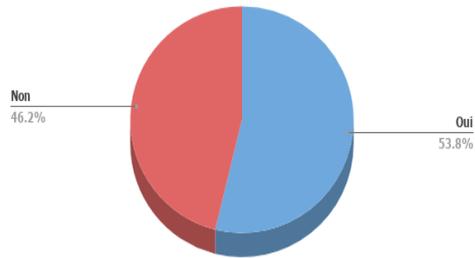
Même si plus de 53,8% des auteurs ont déclaré qu'ils trouvent que les journaux libres sont aussi prestigieux que ceux payants, la majorité d'entre eux révèle que l'établissement de tutelle n'oriente pas le chercheur vers la publication dans un journal en OA.

En outre, un répondant déclare que son encadrant lui interdit de publier dans un journal en OA ; la crédibilité de ces derniers, notamment ceux publiés au Maroc, n'étant pas encore un fait accompli !

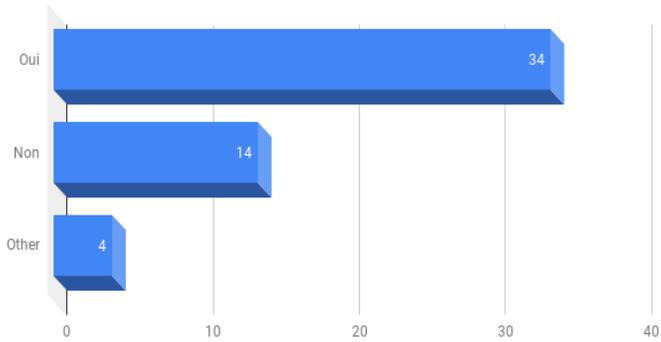
Graph 7 : Usage de l'OA par le chercheur marocain en CS



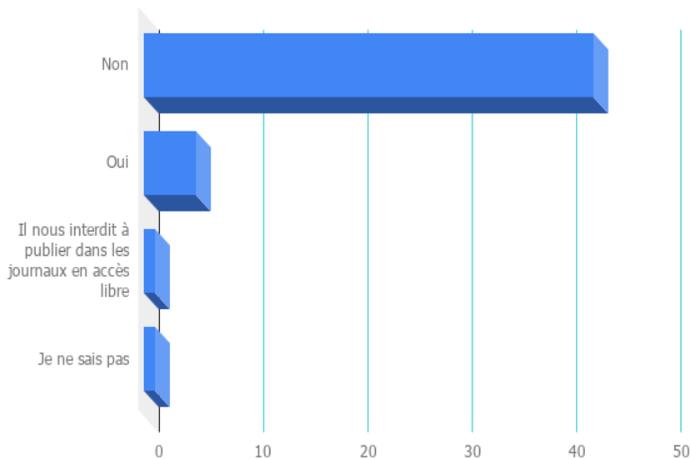
Trouvez-vous que les journaux en libre accès sont aussi prestigieux que ceux payants ?



Pensez-vous que votre publication sera-t-elle plus lue si elle est publiée en accès libre ? Sera-t-elle plus citée ?



Votre université vous encourage-t-elle à publier dans des journaux en particulier ?

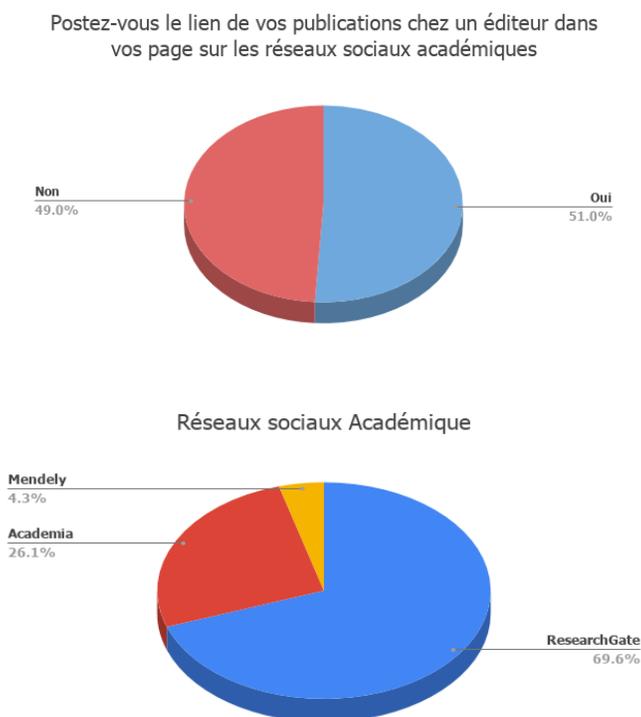


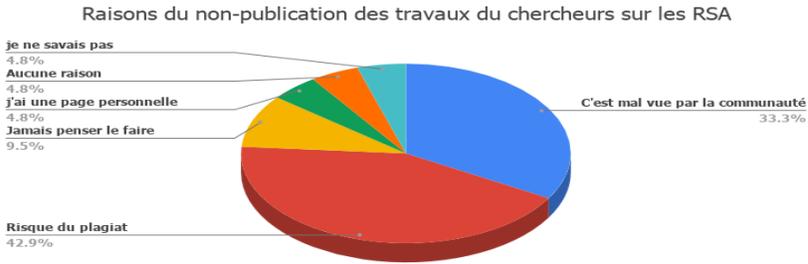
4.2.5 Les réseaux sociaux académiques, l'Open Access et le chercheur marocain

Pour la moitié des chercheurs enquêtés, en plus des sites Web personnels et les dépôts d'archives, les réseaux sociaux académiques (RSA) constituent pour eux un moyen de diffusion de leurs résultats de recherche. C'est ainsi qu'ils utilisent « ResearchGate » à 69,6%, suivi d'« Academia.eu » à 26,1% et de « Mendeley » à 4,3% (Graphe 9).

L'autre moitié des chercheurs hésite à utiliser les RSA comme moyen de partage des publications pour une multitude de raisons. Il s'agit principalement du souci de plagiat (42,9%) et de l'exposition au jugement négatif de la communauté (33,3%). D'autres enquêtés n'ont jamais pensé à publier via les RSA (9,5) ou méconnaissent totalement la possibilité qu'ils offrent (4,8).

Grappe 8 : Usage des réseaux sociaux académiques par le chercheur marocain en CS





4.2.6 Connaissance du Contexte marocain en matière d'Open Access

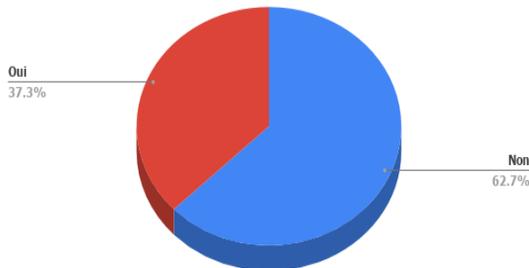
Plus de 62% du public cible de notre enquête affirment leur méconnaissance de l'existence du Portail des Revues Scientifiques Marocaines « PRSM ». Ce faible pourcentage (7,7%) des chercheurs ayant publié dans une revue marocaine semble être justifié (Graph 10).

En effet, le contexte national en matière de l'OA est quasiment ignoré par les chercheurs marocains. Pourtant celui-ci est constitué par plusieurs projets importants initiés par l'IMIST-CNRST. Il s'agit en particulier du réservoir d'archives ouvertes « Toubkal » [29] qui contient plus de 12000 notices de Thèses (1000 en texte intégral), et, du Portail des Revues Scientifiques Marocaines « PRSM » [30] qui héberge plus de 121 revues avec plus de 4000 articles (Juillet 2018) des chercheurs appartenant à des sociétés savantes marocaines.

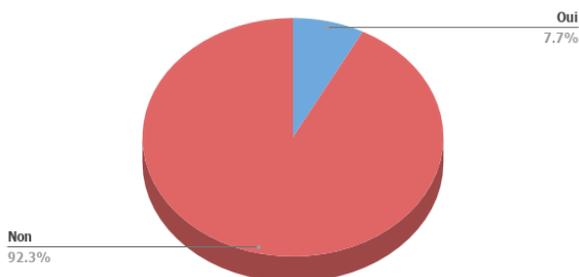
A cet effet, les revues marocaines en OA les plus mentionnées par les répondants sont : e-TI (Revue internationale en Technologies de l'Information) [31] et MATEC WEB of conferences [32].

Graph 9 : Connaissance du contexte marocain de l'Open Access

Connaissez-vous le Portail des Revues Scientifiques Marocaines ?



Avez-vous déjà publié dans une revue marocaine ?



5 Conclusion et recommandations

Il résulte de l'analyse des données de Scopus que l'offre marocaine en matière de l'Open Access dans le domaine du CS reste modeste et que la publication marocaine tend plus vers les revues à accès payant. L'autre résultat atteint par l'enquête sur le terrain a montré que les revues existantes sont méconnues par les chercheurs et que ces derniers encouragent l'OA, plus, en tant que consommateurs. S'ils sont conscients de son importance pour les processus de recherche et de communication scientifique (notamment en termes de citation, d'impact et de visibilité des publications), ils sont loin de l'adopter. En l'absence de mécanismes d'encouragement et de mobilisation pour publier en OA (bourses, prix...), nos chercheurs hésitent encore à faire part de la communauté du libre accès quand il s'agit de publier leurs propres résultats de recherche.

Il s'avère ainsi qu'au-delà de la sensibilisation à l'offre internationale, il faille encourager les initiatives locales et inciter les chercheurs à publier et à partager leur production scientifique sur des plateformes libres nationales. Ces initiatives doivent être accompagnées d'un travail de formations afin d'aider les acteurs de l'OA du pays à dépasser les barrières techniques, linguistiques, légales et économiques. Il s'agira aussi de positionner les revues marocaines et de les faire connaître à l'échelle nationale, régionale et pourquoi pas internationale. Une combinaison de mesures leur permettrait de surmonter les barrières susmentionnées et d'être concurrentielles. Les revues marocaines en OA doivent garantir les mêmes services offerts par un éditeur commercial (peerreviewing, label qualité, marketing, etc.).

Cependant, être conscient des avantages qu'offre l'OA et avoir des revues libres en bonne et due forme ne suffisent pas pour adopter le mouvement et changer les habitudes des chercheurs envers la publication scientifique. Son

adoption nécessite une véritable politique et une stratégie pour son application. Celles -ci devraient être soutenues par une assise juridique menant le chercheur marocain à déposer gracieusement ses publications dans les archives ouvertes de leur tutelle, notamment pour les travaux émanant des recherches financées par des fonds publics.

Dans le même ordre d'idée, la récompense académique et l'appui des établissements de recherches restent la façon la plus prometteuse pour atteindre les objectifs du libre accès.

Pour l'avenir de la communication savante dans notre pays, nous incitons d'autres chercheurs à creuser davantage les données puisées dans Scopus et à profiter des résultats collectés par l'enquête que nous avons menée.

Références

- [1] European Commission (2017). Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. [En ligne]. Disponible à : http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_a_pilot/h2020-hi-oa-pilot-guide_en.pdf (page consultée le 20 avril 2018)
- [2] BEN ROMDHANE, Mohamed (1996). Analyse des publications scientifiques : caractéristiques, structures et langages. [En ligne]. Disponible à : <http://www.enssib.fr/bibliotheque-numerique/documents/1652-analyse-des-publications-scientifiques-caracteristiques-structures-et-langages.pdf> (page consultée le 30 avril 2018)
- [3] Budapest Open Access Initiative : initiative internationale, lancée en 2002, pour rendre les articles de recherche dans tous les domaines universitaires librement accessibles sur l'Internet.
- [4] Fang Hua et al (2017). Open access: concepts, findings and recommendations for stakeholders in dentistry. [En ligne]. Disponible à : <https://doi.org/10.1016/j.jdent.2017.06.012> (page consultée le 10 avril 2018)
- [5] WOJCIECHOWSKA, Anna (2006). Usage des archives ouvertes dans les domaines des mathématiques et de l'informatique. [En ligne]. Disponible à : <https://www.cairn.info/revue-documentaliste-sciences-de-l-information-2006-5-page-294.htm> (page consultée le 25 avril 2018)
- [6] LAWRENCE, Steve (2001). "Online or Invisible." preprint. [En ligne]. Disponible à : <http://www.neci.nec.com/~lawrence/papers/online-nature01/>. "Free Online Availability Substantially Increases a Paper's Impact," Nature 411 (page consultée le 10 mai 2018)
- [7] BJORK, B-C. (2004) Open Access to scientific publication : an analysis of the barriers to change. Information research, Vol.9, No.2, paper 170. [En ligne].

Disponible à : <https://helda.helsinki.fi/handle/10227/647> (page consultée le 20 mai 2018)

- [8] HARNAD, Stevan et al. (2018). The green and the gold roads to Open Access. [En ligne]. Disponible à : https://www.researchgate.net/publication/37537731_The_green_and_the_gold_roads_to_Open_Access (page consultée le 25 mai 2018)
- [9] DAVIS, Philip M.; WALTERS, William H. (2011). The impact of free access to the scientific literature: a review of recent research. [En ligne]. Disponible à : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3133904/>
- [10] PRINGLE, James. Do Open Access journals have impact? (2018). [En ligne]. Disponible à : <https://www.nature.com/nature/focus/accessdebate/19.html> (page consultée le 10 juin 2018)
- [11] SWAN, Alma. The Open Access citation advantage: studies and results to date. [En ligne]. Disponible à : <https://eprints.soton.ac.uk/268516/> (page consultée le 10 juin 2018)
- [12] ANTELMAN, Kristin (2004). Do Open Access articles have a greater research impact? [En ligne]. Disponible à : http://eprints.rclis.org/5463/1/do_open_access_CRL.pdf (page consultée le 10 juin 2018)
- [13] EVANS, JA; REIMER, J. (2009). Open access and global participation in science. [En ligne]. Disponible à : DOI:10.1126/science.1154562
- [14] ANTELMAN, K. (2004). Do open access articles have greater research impact? [En ligne]. Disponible à : http://eprints.rclis.org/5463/1/do_open_access_CRL.pdf (page consultée le 10 juin 2018)
- [15] GARGOURI, Y ; HAJJEM, C ; LARIVIERE, V. et al. Self-selected or mandated, open access increases citation impact for higher quality research. [En ligne]. Disponible à : <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0013636> (page consultée le 20 juin 2018)
- [16] HARNAD, Stevan; BRODY, Tim (2004). Comparing the impact of Open Access vs. non-Oa Articles in the Same Journals. D-Lib Magazine. [En ligne]. Disponible à : <http://www.dlib.org/dlib/june04/harnad/06harnad.html> (page consultée le 10 mai 2018)
- [17] LAWRENCE, Steve (2001). Free online availability substantially increases a paper's impact Nature volume 411, page 521. [En ligne]. Disponible à : <https://www.nature.com/articles/35079151> (page consultée le 10 juin 2018)
- [18] VOGEL, Gretchen (2011). Open Access Gains Support; Fees and Journal Quality Deter Submissions. *Science* : Vol. 331, Issue 6015, pp. 273. [En ligne].

Disponible à : <http://science.sciencemag.org/content/331/6015/273.1> (page consultée le 10 juin 2018)

- [19] 2015 STM: International Association of Scientific, Technical and Medical Publishers. [En ligne]. Disponible à : https://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf (page consultée le 10 mai 2018)
- [21] THELWALL, Mike; KOUSHA, Kayvan (2014). ResearchGate: Disseminating, communicating, and measuring Scholarship? [En ligne]. Disponible à : <https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.23236> (page consultée le 20 juin 2018)
- [22] THELWALL, Mike ; KOUSHA, Kayvan (2013). Academia.edu: Social network or Academic Network? [En ligne]. Disponible à : <https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.23038> (page consultée le 20 juin 2018)
- [23] NIYAZOV, Y ; VOGEL, C ; PRICE, R et al. (2016) Open Access Meets Discoverability: Citations to Articles Posted to Academia.edu. PLOS ONE 11(2): e0148257. [En ligne]. Disponible à : <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0148257> (page consultée le 10 juin 2018)
- [24] PIWOWAR, H ; PRIEM, J ; LARIVIÈRE, V et al. (2017) The State of OA: A large-scale analysis of the prevalence and impact of Open Access articles. PeerJ. [En ligne]. Disponible à : <https://peerj.com/preprints/3119/> (page consultée le 19 avril 2018)
- [25] CABANAC, Guillaume (2015) « Bibliogifts in LibGen? A study of a text-sharing platform driven by biblioleaks and crowdsourcing », *Journal of the Association for Information Science and Technology*. [En ligne]. Disponible à : <https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.23445> (page consultée le 29 avril 2018)
- [26] BOHANNON, John (2016). Who's downloading pirated papers? Everyone. [En ligne]. Disponible à : <http://www.sciencemag.org/news/2016/04/whos-downloading-pirated-papers-everyone> (page consultée le 12 juillet 2018)
- [27] HIMMELSTEIN et al. Research: Sci-Hub provides access to nearly all scholarly literature. [En ligne]. Disponible à : <https://elifesciences.org/articles/32822> (page consultée le 20 juin 2018)
- [28] CiteScore (SCOPUS) mesure l'impact de milliers de revues en calculant le nombre moyen de citations par article sur une période de trois ans.
- [29] Catalogue national des thèses et mémoires : www.toubkal.imist.ma
- [30] Portail des Revues Scientifiques Marocaines : www.prsm.imist.ma
- [31] e-TI, La Revue internationale en Technologie de l'Information : <https://revues.imist.ma/index.php?journal=e-TI>

- [32] Open Access proceedings in Materials Science, Engineering and Chemistry:
<https://www.matec-conferences.org/>

Open Access au Maroc : Vers un dispositif national d'Open Access collaboratif au profit de l'innovation

Open Access in Morocco: Towards a National collaborative Open Access dedicated to Innovation

NABIGH Majda

Laboratoire DICEN-IDF, CNAM, Paris, Île-de-France
nabigh.majda@yahoo.fr

DAOUDI Najima

Laboratoire Lyrica, Ecole des Sciences de l'Information, Rabat (Maroc)
Laboratoire SSL, ENSIAS - Rabat (Maroc)
daoudinajima@yahoo.fr

Résumé

Le Libre Accès aux résultats de la recherche s'est rapidement répandu au niveau international mais de façon déséquilibrée, notamment en Afrique. L'enjeu du libre accès n'est pas seulement scientifique mais aussi économique. Dans ce contexte, nous dresserons, à travers cet article, un panorama de l'état actuel des archives ouvertes dans le monde avec un focus sur les initiatives nationales. Ceci nous mènera à la fin de notre article vers la proposition d'un projet de plateforme nationale d'Open Access collaboratif pour conclure sur un modèle conceptuel illustrant la contribution et l'intérêt de l'Open Access dans le processus d'innovation/ intelligence économique d'une organisation.

Mots-clés. Libre accès, archives ouvertes, innovation, intelligence économique, Maroc.

Abstract

Open Access to research results has rapidly spread internationally but unbalanced, particularly in Africa. The issue of Open Access is not only scientific but also economic. In this context, we will provide an overview of the current state of open archives around the world with a focus on national initiatives. This will take us to the end of our article towards the proposal of a national collaborative Open Access platform project to conclude on a conceptual model on the contribution and interest of Open Access in the innovation / intelligence process of an organization.

Keywords. Open Access, open archive, innovation, competitive intelligence, Morocco.

1 Introduction

L'Open Access constitue une des facettes du mouvement international du « libre » : open data, open source, openness, open innovation,.. Un mouvement qui a contribué à la créativité et à l'innovation scientifique et technique. L'ouverture au public des résultats d'une recherche, permet d'avoir un regard externe qui est souvent source d'une nouvelle réflexion. Un constat confirmé par une étude qui a été conduite en 2010 par le Research Information Network (RIN) et le National Endowment for Science, Technology and the Arts (NESTA). Cette étude a montré que le mouvement du libre a permis d'une part, une grande efficacité de la recherche à travers la promotion de la rigueur scientifique et le traitement de nouvelles problématiques de recherche. D'autre part, il joue un rôle central dans la collaboration entre chercheurs (RIN et NESTA, 2010).

Conscients de l'impact de la recherche sur l'économie et la société, plusieurs pays ont mis en place des initiatives de démocratisation de l'accès aux résultats de la recherche scientifique comme les Etats-Unis, la France, la Suisse, le Canada, l'Allemagne... Un impact qui est visible non seulement dans le milieu académique (conduite de nouvelles études, amélioration d'études antérieures, conception de nouveaux modèles...) mais également dans le domaine industriel (invention de nouveaux produits, amélioration de produits existants)... Ceci nous mène à réfléchir sur la situation du Maroc par rapport à ces pays, notamment les initiatives qu'a connue la scène nationale et à nous interroger sur l'intérêt d'une politique nationale d'Open Access collaboratif appuyée par les pouvoirs publics ainsi que par le monde académique, dont la finalité est d'une part, stimuler la recherche et la collaboration académique, d'autre part, favoriser l'innovation et le développement durable du territoire en appuyant le processus d'intelligence économique des organisations. C'est dans ce cadre que s'inscrit notre problématique : quel modèle d'Open Access mettre au profit du monde académique ? Comment le développement de l'Open Access peut influencer le processus de l'intelligence économique/territoriale ?». Avant de répondre à ces questions, nous présenterons l'état actuel des archives ouvertes dans le monde ainsi que l'état des lieux des initiatives nationales en la matière. Nous présenterons également une analyse fonctionnelle d'une sélection d'expériences internationales d'archives ouvertes ainsi qu'une réflexion autour d'une plateforme nationale collaborative au profit de l'Open Access. Nous terminerons notre article sur un modèle conceptuel illustrant la contribution de cette plateforme dans le processus d'intelligence économique-innovation des organisations à travers lequel nous démontrerons l'impact des résultats d'une recherche scientifique dans l'amélioration de l'édition des outils de gestion des bibliothèques virtuelles, comme exemple.

1.1 Méthodologie de l'étude

Sur le plan méthodologique, nous avons combiné la méthode descriptive, le benchmark ainsi que la méthode d'enquête sur le terrain. La méthode descriptive, consistant généralement à déterminer les pratiques courantes, le statut ou les caractéristiques au niveau de la situation étudiée, nous a permis de dresser une vue panoramique des principales expériences qu'a connu le Maroc. Quant au benchmark, il nous a aidé à examiner fonctionnellement quelques expériences internationales d'archives

ouvertes afin d'en extraire les fonctionnalités indispensables pour la plateforme nationale que nous proposons. Concernant la méthode d'enquête sur le terrain, nous avons utilisé principalement le guide d'interview afin d'avoir des réponses plus élaborées et complètes. L'enquête a été conduite auprès d'un échantillon de 60 étudiants universitaires (cycle master et doctorat) dont 20 étudiants des Facultés des Sciences Juridiques Economiques et Sociales de Salé et Rabat ; 15 étudiants de l'Institut supérieur de commerce et d'administration des entreprises ; 15 étudiants de l'école Mohammadia des Ingénieurs et 10 étudiants de l'Institut Supérieur de l'Information et de la Communication.

Par ailleurs, nous avons exploité la revue de littérature sur l'Open Access afin de donner une vue globale des principaux concepts liés à ce mouvement. Bien que le l'Open Access comprend aussi bien les archives ouvertes que les revues ouvertes, le présent article traitera principalement les archives ouvertes.

1.1 Repères théoriques

Le concept d'Open Access est venu dans le cadre d'un mouvement global en faveur d'un accès gratuit et pérenne aux résultats de la recherche par l'ensemble de la communauté scientifique internationale (FILY, 2005). Un mouvement qui est né de la combinaison de plusieurs facteurs à savoir : l'avènement des technologies de l'information et de la communication (TIC) notamment l'Internet, l'augmentation des tarifs appliqués par les revues scientifiques, la politique des éditeurs à l'égard des auteurs qui exige des délais de publication trop longs et oblige les auteurs à céder leurs droits. C'est dans ce contexte que la première archive ouverte a vu le jour (HARNAD, 1995) avec la création du premier serveur de publications en physique par Paul Ginsparg à l'université de Los Alamos en 1991 (ArXiv). Plus tard, les archives se sont multipliées partout dans le monde, traduisant une prise de conscience progressive dans les milieux universitaires et politiques de la nécessité de mettre les résultats de la recherche en accès libre. Ceci a donné lieu à plusieurs appels au partage des résultats de la recherche, dont le premier est la déclaration de Budapest¹ (BOAI), lancée en 2002, selon laquelle l'Open Access est défini comme « *la mise à disposition gratuite sur l'Internet public, permettant à tout un chacun de lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces articles, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale, sans barrière financière, légale ou technique autre que celles indissociables de l'accès et l'utilisation d'Internet. La seule contrainte sur la reproduction et la distribution, et le seul rôle du copyright dans ce domaine devrait être de garantir aux auteurs un contrôle sur l'intégrité de leurs travaux et le droit à être correctement reconnus et cités.* ».

Un autre appel au libre accès a vu le jour en 2003. Il s'agit de la déclaration de Bethesda² selon laquelle « *l'auteur et les titulaires du droit d'auteur accordent à tous les utilisateurs un droit d'accès gratuit, irrévocable, mondial et perpétuel et leur concèdent une licence leur permettant de copier, utiliser, distribuer, transmettre et visualiser publiquement l'œuvre... à condition d'en*

¹<http://www.soros.org/openaccess/fr/read.shtml>(version française)

²<http://biomedcentral.com/openaccess/bethesda/>

indiquer correctement l'auteur...Ils accordent également aux utilisateurs le droit de faire un petit nombre de copies papier de l'œuvre pour leur usage personnel...La version complète de l'œuvre est déposée dès sa publication dans au moins un réservoir en ligne subventionné par un organisme reconnu œuvrant pour le libre accès, la diffusion sans restriction, l'interopérabilité et l'archivage à long terme.».

La troisième déclaration qui a marqué le mouvement du libre, est celle de Berlin³ qui a été lancée également en 2003 et qui a repris la même définition du libre accès donnée par la déclaration de Bethesda en étendant ce modèle à l'ensemble des résultats de la recherche et au patrimoine culturel. La particularité de cette déclaration réside dans le fait qu'elle a donné au libre accès une validité institutionnelle au vu de la nature des signataires qui sont principalement des institutions, contrairement aux déclarations précédentes dont les signataires sont des individus.

L'Open Access signifie donc l'accès immédiat, permanent, libre, gratuit et électronique aux résultats de la recherche. On en distingue principalement deux voies⁴ :

- **La voie verte (Green Open Access)** : La voie verte correspond à la diffusion des résultats de la recherche via une archive ouverte. Concrètement, il s'agit pour un chercheur de déposer son travail, de l'archiver, dans un réservoir en ligne. Ce mode de diffusion ne remplace pas l'édition, mais permet une diffusion plus large. En effet, un auteur pourra y déposer un article déjà publié dans une revue, et ainsi le rendre accessible gratuitement à tous.
- **La voie dorée (Gold Open Access)** : La voie dorée, c'est un mode d'édition qui vient en parallèle du circuit classique. Une revue Open Access c'est une revue qui ne sera pas soumise à abonnement payant et, par conséquent, son accès sera gratuit pour le lecteur. Il faut toutefois être vigilant, car accès gratuit ne signifie pas forcément publication gratuite. Des éditeurs demanderont souvent des APC (Article Processing Charges), des frais à la charge de l'auteur, de son institution, au moment de la publication, pour « libérer l'article ». Ce mode de publication auteur-payeur revient à déplacer les coûts du budget documentaire vers le budget de la recherche.

Il existe également une voie dite « hybride » proposée par les éditeurs en réaction à la montée des archives ouvertes, qui suppose que les auteurs, leurs institutions ou les agences de recherche financent en amont la mise en ligne d'articles mis en libre accès. En d'autres termes, il s'agit de publier des articles en accès gratuit et d'autres en accès payant (abonnement). Le choix étant celui de l'auteur plus que celui de l'éditeur. La situation hybride réside en ce que des articles d'une revue sont en Open Access pendant que d'autres de la même revue ne le sont pas. Lors de la soumission de son

³http://www.zim.mpg.de/openaccess-berlin/BerlinDeclaration_fr.pdf

⁴HAL. Guide pratiques : mieux connaître l'Open Access et HAL.

article, l'auteur se voit donc proposer une option de paiement pour passer en Open Access (REYMONET, 2012).

2 Open Access dans le monde : Etat actuelle des archives ouvertes

Aujourd'hui on assiste à une prolifération des plateformes d'archives ouvertes à l'échelle internationale, avec une forte concentration au niveau de l'Europe, suivie de l'Asie et l'Amérique du nord. L'interrogation durépertoire internationalOpenDOARfait ressortir ce constat comme illustré dans la figure 1 :

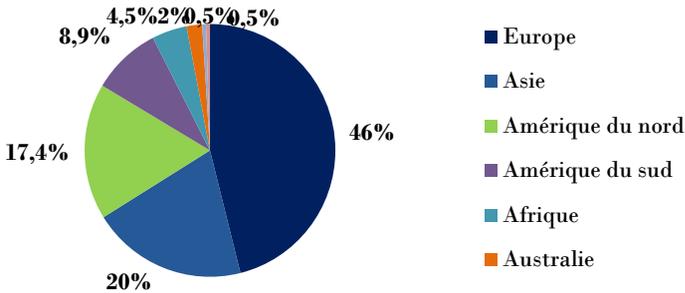


Figure 3. Répartition des répertoires d'archives ouvertes dans le monde par continent⁵

Néanmoins un déséquilibre est clairement constaté en observant la répartition du nombre d'entrepôts en libre accès dans le monde. Les trois pays leaders à savoir : les Etats Unis, le Royaume Uni et le Japon ont presque le tiers du nombre total des entrepôts recensés par l'OpenDOAR comme le montre la figure 2.

⁵Les données du graphique ont été puisées du OpenDOAR à la date du 10 septembre 2018 accessible sur :<http://www.opendoar.org/onechart.php?cID=&ctID=&rtID=&clID=&IID=&potID=&rSoftWareName=&search=&groupby=c.cContinent&orderby=Tally%20DESC&charttype=pie&width=600&height=300&caption=Proportion%20of%20Repositories%20by%20Continent%20-%20Worldwide>

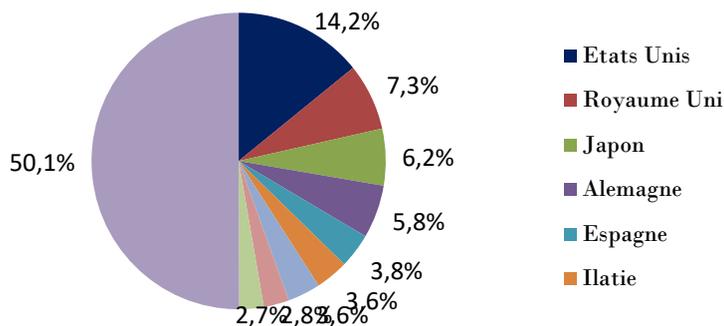


Figure 4. Répartition des répertoires d'archives ouvertes dans le monde par pays⁶

Les pays de l'Afrique, ont un faible pourcentage, soit 4,5%, par rapport au nombre total de dépôts recensés par l'OpenDOAR, soit 3777 dépôts en septembre 2018. L'Algérie fait partie des pays qui ont un nombre important d'archives ouverte soit 8,2% du nombre total enregistré en Afrique qui est de 158 dépôts.

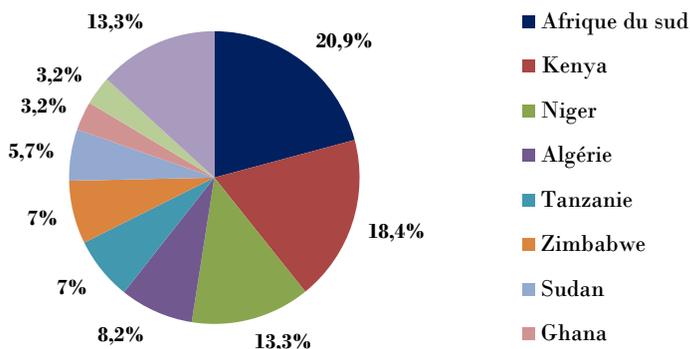


Figure 5. Répartition des répertoires d'archives ouvertes dans les pays de l'Afrique⁷

Par ailleurs, le nombre limité des initiatives d'Open Access en Afrique peut-être expliqué, entre autre, par le faible taux de pénétration d'internet. Le

⁶Les données du graphique ont été puisées du OpenDOAR à la date du 10 septembre 2018 accessible sur : <http://www.opendoar.org/onechart.php?cID=&ctID=&rtID=&clID=&IID=&potID=&rSoftWareName=&search=&groupby=c.cCountry&orderby=Tally%20DESC&charttype=pie&width=600&height=300&caption=Proportion%20of%20Repositories%20by%20Country%20-%20Worldwide>

⁷Les données du graphique ont été puisées du OpenDOAR à la date du 10 septembre 2018 accessible sur : <http://www.opendoar.org/onechart.php?cID=Africa&ctID=&rtID=&clID=&IID=&potID=&rSoftWareName=&search=&groupby=c.cCountry&orderby=Tally%20DESC&charttype=pie&width=600&height=300&caption=Proportion%20of%20Repositories%20by%20Country%20-%20Africa>

continent vient au dernier rang à l'échelle mondiale, comme le montre la figure 4 :

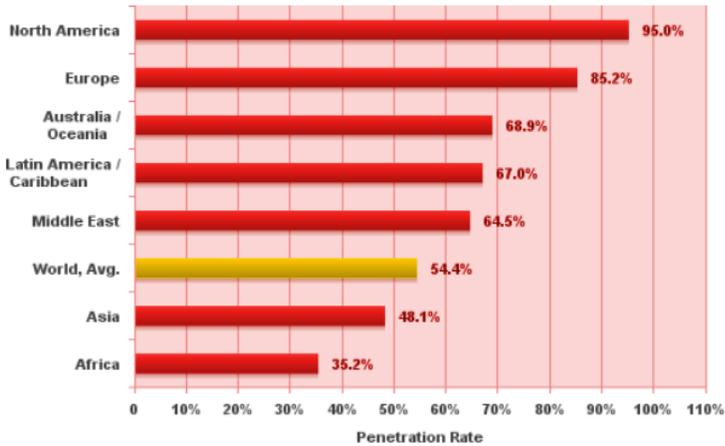


Figure 6. Taux de pénétration d'Internet par continent⁸

D'autres facteurs explicatifs de la situation du libre accès en Afrique (NWAGWU et ALLAM, 2009) sont liés à la méconnaissance du libre accès et son intérêt pour la recherche et les chercheurs, la faiblesse des moyens accordés à la recherche scientifique, la fragilité des infrastructures technologiques. Ce qui constitue autant de chantiers sur lesquels les instances nationales des pays d'Afrique doivent apporter un regard attentif.

3 L'Open Access au Maroc

3.1 L'open access dans les pratiques info-documentaires des universitaires Marocains

Dans le but d'avoir une vue globale sur la connaissance et l'intérêt de l'Open Access du point de vue de l'étudiant marocain. Nous avons conduit une étude auprès d'un échantillon de 60 étudiants universitaires appartenant au cycle de master et doctorat sur l'axe Rabat-Salé. L'analyse des résultats a montré une connaissance confuse de la signification du concept, seule 10% des enquêtés a pu donner une définition exacte du terme et 60% l'on automatiquement confondu avec les outils open source.

⁸ Internet WorldStat. [En ligne]. Disponible sur : <https://www.internetworldstats.com/stats.htm> (Consulté le 15 janvier 2017)

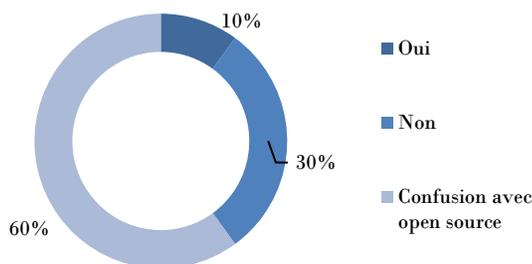


Figure 7. Connaissance de la signification de l'Open Access

Néanmoins, les étudiants ont affirmé avoir recours à des plateformes offrant le texte intégral des travaux de recherche de type thèses, articles et mémoires. Les principales plateformes citées sont : Hal, theses.fr, memoireOnline et cairn.info. Aucun des enquêtés n'a fait mention des initiatives nationales d'Open Access comme Risalaty, Toubkal, SIJIL...

Concernant le processus d'identification de la thématique de recherche pour un projet de master ou de thèse, les étudiants se limitent à consulter les projets soutenus dans leurs établissements de recherche, disponibles dans les bibliothèques ou des thématiques inspirées des travaux identifiés dans des plateformes internationales d'archives ouvertes. Aucun des étudiants n'a indiqué avoir recours à Otrohati⁹ pour l'identification des travaux traités dans d'autres établissements de recherche. Ce qui entrainera sur le long terme une redondance des thématiques et une stagnation de la recherche.

Par ailleurs, les enquêtés ont souligné le rôle important que jouent les réseaux sociaux comme Facebook et LinkedIn dans la diffusion des informations sur les manifestations scientifiques nationales et internationales grâce aux fonctionnalités qu'ils offrent (création de groupes, discussion, commentaires...). Ils ont jugé important la mise en place d'initiatives nationales permettant non seulement de collecter les résultats de la recherche scientifique mais qui permettront plutôt de connecter les chercheurs des différentes structures de recherche autour d'une plateforme unique favorisant l'échange et le partage.

3.2 Expériences nationales

3.2.1 Projet Tempus IStEMag

Le projet Tempus IStEMag¹⁰ a été lancé en 2009 par la Commission Européenne dans le but d'optimiser l'accès à l'information scientifique et technique (IST) dans les pays du Maghreb. Il marque la coopération entre 14 universités européennes (Belgique et Roumanie) et maghrébines (Maroc, Algérie et Tunisie). Les objectifs spécifiques du projet consistent à établir un bilan détaillé des conditions d'accès à l'IST pour les enseignants-

⁹Otrohati est une plateforme de signalement des thèses en cours de préparation dans les centres des études doctorales (CEDoc) marocains

¹⁰Projet d'optimisation de l'accès à l'information scientifique et technique dans les universités du Maghreb. [En ligne] :<https://ec.europa.eu/programmes/erasmus-plus/projects/eplu-project-details/#project/510986-TEMPUS-1-2010-1-BE-TEMPUS-JPCR>(Consulté le 5 août 2018)

chercheurs et les étudiants des universités maghrébines ; créer des portails documentaires des bibliothèques universitaires ; créer des consortia documentaires afin d'appuyer la négociation collective de l'accès en ligne aux principales sources d'information ; définir et mettre en place une politique institutionnelle d'archives ouvertes ; lancer effectivement les activités de ces structures et politiques et diffuser cette expérience en vue de sa généralisation à l'ensemble des institutions d'enseignement supérieur.

En 2014, un séminaire de clôture du projet Tempus ISTE¹¹Mag sous le thème «Bilan final du projet ISTE¹¹Mag et la signature de la charte de pérennité du projet» a été organisé au siège du Centre de Recherche sur l'Information Scientifique et Technique à Ben Aknoun en Algérie. À l'issue de cet événement, il a été décidé, dans le cadre de ladite charte, d'assurer la généralisation du projet et sa diffusion auprès d'autres acteurs locaux ; son appropriation et sa prise en charge par les bénéficiaires locaux ; l'inscription des politiques adoptées dans une dynamique locale et notamment la politique d'archives ouvertes ; la prise en compte de la politique d'archives ouvertes maghrébine arrêtée tout en respectant les spécificités locales ; la durabilité et la visibilité du projet en garantissant la survie du portail fédérateur et des plateformes locales ; le soutien technique par le maintien du comité technique créé à cet effet et la fourniture à ses membres des conditions adéquates pour la continuation des actions entamées et le transfert des compétences et du savoir-faire acquis.

3.2.2 Portail des revues scientifiques Marocaines

L'Institut Marocain de l'Information Scientifique et Technique (IMIST) a mis en place le Portail des revues scientifiques Marocaines¹¹. Une plateforme qui a pour objectif la valorisation et la diffusion à l'échelle nationale et internationale de la production scientifique, toutes disciplines confondues, éditée par les laboratoires publics et les sociétés savantes à but non lucratif. Dans le cadre de ce projet, l'IMIST assure la mise en ligne et l'archivage des nouvelles revues scientifiques ou celles déjà existantes. Les revues intéressées par ce service prennent contact avec l'IMIST à travers un formulaire qui fixe les conditions générales que doit satisfaire le rédacteur en chef de la revue ainsi que les informations concernant cette dernière, la rédaction et le format de publication. L'IMIST oblige les revues à donner libre accès au texte intégral de tous les articles publiés sur la plateforme. Le Portail prend uniquement en charge les articles disponibles sous format électronique. Pour les revues papiers déjà existantes, un embargo maximum d'une année est permis sur le dernier numéro publié.

3.2.3 Toubkal

Toubkal¹² est une plateforme nationale qui a été mise en place par l'IMIST en collaboration avec les établissements d'enseignement supérieur nationaux et les chercheurs marocains. Elle s'inscrit dans le cadre d'une politique gouvernementale visant la promotion et le développement de la recherche scientifique et la réalisation ainsi du e-Maroc à travers la

¹¹Portail des revues scientifiques Marocaines. [En ligne] : <https://revues.imist.ma/> (Consulté le 5 août 2018)

¹²Toubkal-IMIST. [En ligne] : <https://toubkal.imist.ma/> (Consulté le 5 août 2018)

fourniture de services en ligne (documents électroniques, traitement des demandes d'information...). Cette plateforme se veut un catalogue des thèses et mémoires et projets de fin d'étude soutenus dans les différents établissements de formation et universités marocaines; les thèses de médecine ainsi que les thèses soutenues par les chercheurs marocains dans des établissements étrangers. Toubkal donne accès à au contenu des thèses, en totalité ou en partie, sous réserve de remplir le formulaire d'autorisation de diffusion par le l'auteur de la thèse.

3.2.4 Rissalaty

Rissalaty¹³ est une archive ouverte initiée par l'Université Hassan II-Casablanca. Elle qui permet d'accéder aux thèses de doctorat et aux thèses d'exercice, notamment en médecine et pharmacie et en médecine dentaire. Les thèses dont les auteurs ont signé une autorisation de diffusion en ligne sont accessibles sur internet sans restriction. D'autres sont accessibles uniquement aux membres de l'Université Hassan II Casablanca via leur Espace Numérique de Travail (ENT). La plateforme permet d'accéder aux thèses grâce à un moteur de recherche. Il est également possible d'affiner les résultats de sa recherche à travers des filtres tels que le titre, l'auteur, l'année de publication, la discipline, etc. Afin d'identifier les contenus les plus consultés le site permet de générer des statistiques sur le nombre de consultations et de téléchargements

3.2.5 SIJIL

SIJIL¹⁴ est considéré comme le dépôt institutionnel où sont publiés les thèses et travaux de recherche du personnel, enseignants-chercheurs, administratifs et étudiants des établissements de l'Université Mohammed V Rabat. La plateforme permet d'effectuer une recherche sur l'ensemble des thèses et articles mais permet aussi une navigation par auteur, par sujet en encore par année.

3.2.6 Dépôt institutionnel de l'Université Hassan II Casablanca

Il s'agit d'une expérience semblable à celle qui a été conduite par l'Université Mohammed V Rabat. Le dépôt institutionnel de l'Université Hassan II Casablanca¹⁵ est une plateforme créée grâce à au logiciel Dspace et qui offre à la communauté scientifique un accès libre et gratuit aux publications des enseignants et des chercheurs de l'Université. L'interface est en est accessible aussi bien français qu'en anglais et contient des flux RSS pour alerter les utilisateurs sur le nouveau contenu.

3.3 Synthèse des expériences nationales

Malgré ces quelques initiatives qu'a connue la scène nationale, la comparaison du nombre d'initiatives d'archives ouvertes qu'a connu le Maroc, par rapport à d'autres pays arabes, de l'Afrique et à l'échelle

¹³Rissalaty. [En ligne] : <http://rissalaty.univcasa.ma:8080/xmlui/> (Consulté le 5 août 2018)

¹⁴SIJIL. [En ligne] : <http://ens.um5.ac.ma/page/d%C3%A9p%C3%B4t-institutionnel-sijil> (Consulté le 5 août 2018)

¹⁵Dspace@UH2C. [En ligne] : <http://dspace.univcasa.ma:8080/jspui/> (Consulté le 5 août 2018)

mondiale, montre que l'Open Access est encore à un stade embryonnaire malgré les efforts déployés par certaines institutions nationales comme l'IMIST qui a lancé plusieurs chantiers depuis 2014 à savoir :

- L'Etude et l'évaluation des états des lieux de l'usage des ressources électroniques par les chercheurs marocains
- L'accompagnement de la semaine internationale de l'Open Access à travers le monde par l'organisation de la première édition de la journée sur l'Open Access au Maroc

Il s'avère ainsi urgent d'entreprendre un certain nombre d'actions dans le but d'incarner le libre accès comme pratique inhérente au processus de la production scientifiques des chercheurs marocains. Il s'agit de :

- Remonter aux instances de décision nationales l'importance de l'engagement des pouvoirs publics dans la mise en place de projets d'Open Access notamment en allouant les ressources financières adéquates ;
- Créer un comité national permanent qui travaillera en coordination avec les différentes universités et établissements de recherche sur la mise en place d'une politique nationale d'Open Access ;
- Elaborer une stratégie nationale de communication appropriée sur l'Open Access, pour d'abord bien assimiler le concept et pour in fine incarner la culture d'ouverture des résultats de la recherche scientifique au public ;
- Multiplier les formations sur le Libre Accès au sein des établissements de formation pour en assimiler le fonctionnement et l'intérêt ;
- Faire des visites d'échange et de partage d'expériences avec les pays ayant acquis une expertise dans le domaine. Afin de capitaliser sur leurs expériences et en puiser les bonnes critiques.

4 Vers un dispositif national d'Open Access collaboratif

Le Maroc a certes connu peu d'initiatives d'Open Access. Néanmoins, elles attestent d'un intérêt grandissant à la mise en public des résultats de la recherche scientifique. Il convient ainsi de capitaliser sur ces initiatives dans une optique de mise en place d'un dispositif national de démocratisation de l'accès à la production scientifique sous forme d'une plateforme collaborative d'archives ouverte.

4.1 Objectifs

A travers la plateforme proposée nous visons à atteindre les objectifs suivants :

- Mettre à la disposition du monde académique une plateforme unique permettant à l'ensemble de la communauté scientifique d'avoir une vue globale et complète sur l'état de la recherche, dans le but d'améliorer la qualité et les sujets traités par les chercheurs ;

- Favoriser la connexion et la collaboration scientifique entre les chercheurs appartenant aux différentes structures de recherches ;
- Offrir à tout un chacun la possibilité de valoriser travaux de recherche qui n'ont pas eu l'occasion d'emprunter le long circuit de publication (délais strictes, complication du processus...) ;
- Mesurer qualitativement et quantitativement à travers des indicateurs précis l'évolution de la recherche au sein des structures de recherche (attribution de score de recherche par CEDoc, taux de production scientifique par labo...)

4.2 Architecture et fonctionnement du dispositif

Nous proposons la mise en place d'une plateforme nationale collaborative destinée à l'ensemble des chercheurs Marocains. Cette plateforme, comme le montre la figure 5, sera fédératrice de l'ensemble des contenus produits par les chercheurs appartenant aux différentes structures de recherche.

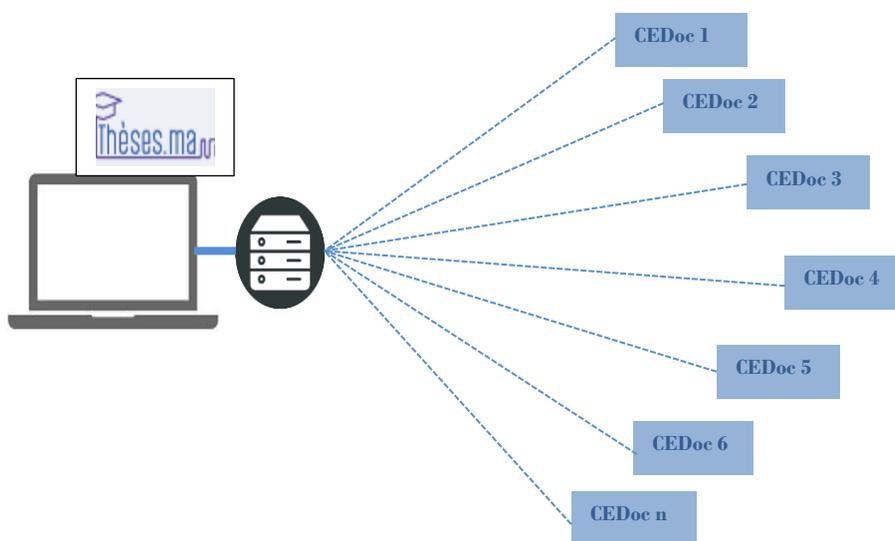


Figure 8. Architecture du dispositif nationale d'Open Access

Il s'agit notamment de signaler différents types de contenus : les thèses en cours, les thèses soutenues, les articles, les actes de conférences, les posters...L'utilisateur final pourra ainsi consulter à partir d'une interface unique toute la production scientifique des chercheurs marocains.

Concrètement, il est proposé d'organiser l'interface, dénommée thèses.ma, autour de trois modules collaboratifs :

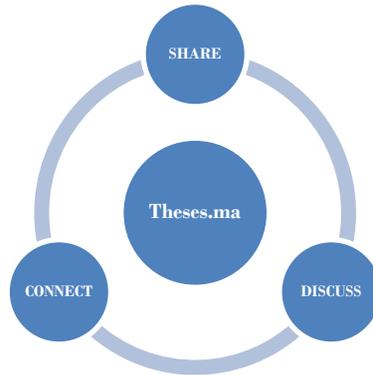


Figure 9. Modules de la plateforme theses.ma

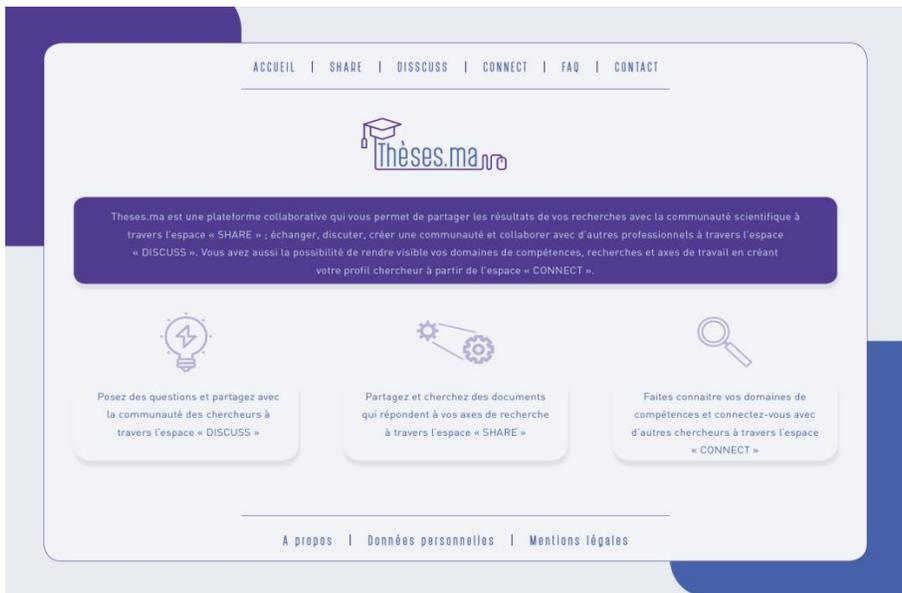


Figure 10. Plateforme Thèses.ma

- **Module de partage de contenu « SHARE »** : cet espace permettra de partager des thèses et tout document de recherche produit par le chercheur marocain

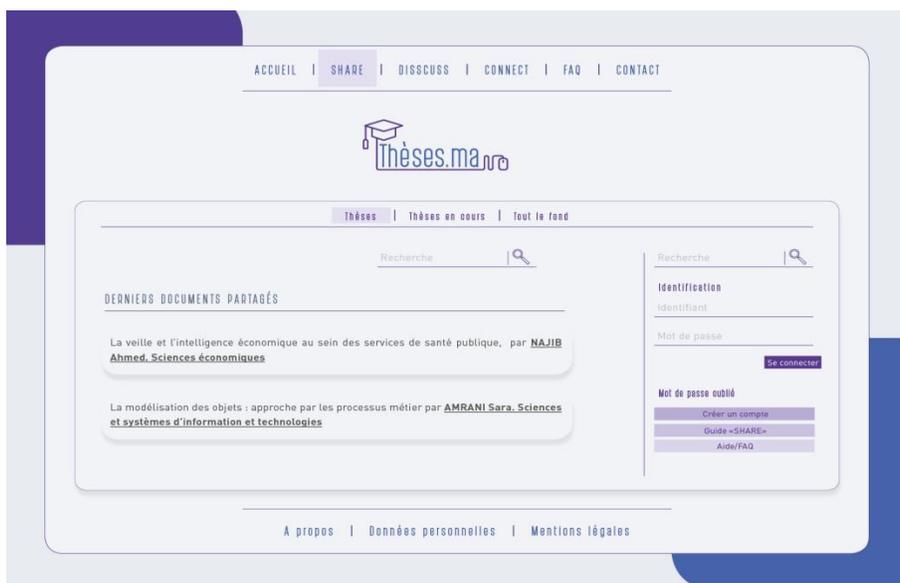


Figure 11. Module « SHARE » de la plateforme Thèses.ma

- **Module de type forum de discussion « DISCUSS »** : un chercheur peut à travers cet espace poser une question, faire un sondage, discuter avec la communauté des chercheurs, partager une information (appel à communication, demande de collaboration...)

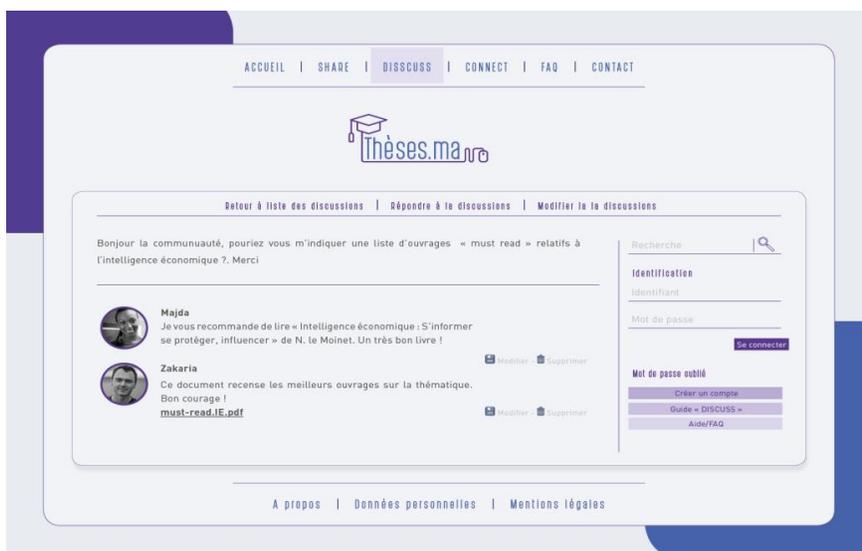


Figure 12. Module « DISCUSS » de la plateforme Thèses.ma

- **Module de type annuaire « CONNECT »** : chaque chercheur peut rendre visible ses domaines de compétences, valoriser son parcours

scientifique et se connecter à d'autres professionnels en créant son profil chercheur à partir de cet espace.



Figure 13. Module « CONNECT » de la plateforme Thèses.ma

Sur le plan fonctionnel, une fois inscrit au CEDoc, le doctorant recevra au moment de l'inscription un lien l'invitant à se connecter à la plateforme theses.ma ce qui lui permettra de créer son propre profil de chercheur et compléter toutes les informations concernant sa thèse à signaler comme en cours de préparation. Le CEDoc validera les informations indiquées, afin d'en garantir l'exactitude, avant leur intégration dans la plateforme globale. Le chercheur pourra ainsi, à l'issue de cette phase, utiliser les différents modules de la plateforme.

Par ailleurs, il est important de noter que dans le but de garantir une alimentation continue de la plateforme proposée. Il est essentiel de mettre en place les conditions qui assureront sa continuité dans le temps. C'est dans ce cadre que nous préconisons la mise en place d'une politique globale d'Open Access, à décliner au niveau de chaque Centre des Études Doctorales (CEDoc). Celle-ci consistera en l'élaboration d'un document à soumettre aux doctorants au moment de l'inscription, selon lequel ces derniers autorisent le personnel chargé de la gestion des études doctorales au sein des CEDoc le signalement la thèse au moment de sa préparation et sa mise en ligne au niveau de la plateforme après la soutenance.

4.3 Fonctionnalités et ressources

La plateforme, notamment le module « SHARE », doit offrir un certain nombre de fonctionnalités. Pour ce faire, nous avons effectué un benchmark sur la base d'une sélection de cinq initiatives d'archives ouvertes dans plusieurs pays. Ces initiatives couvrent plusieurs disciplines et donnent accès à des thèses, articles, e-book, actes de conférences... Il s'agit respectivement de :

- **Theses.fr (France):** est la vitrine des thèses de doctorat françaises en cours de préparation ou soutenues depuis 2006 dans les établissements ayant choisi d'abandonner le dépôt de la thèse papier au profit du support électronique (plus de 6 000 thèses). L'accès au texte intégral des thèses n'est pas systématique car conditionné par l'autorisation de mise en ligne accordée par le docteur et/ou les ayant-droits de la thèse. (443100 documents).
- **UNBC's Digital Institutional Repository (Canada):** il s'agit du dépôt institutionnel de l'University of Northern British Columbia. Il comprend des projets de troisième cycle, des thèses et des mémoires, des publications du corps professoral et des rapports. (plus de 29800 documents).
- **BCU Open Access (Royaume-Uni):** c'est le répertoire du Birmingham City University (BCU). Il donne accès aux résultats de recherche et d'entreprise produits par le personnel de la BCU. (plus de 3000 documents).
- **HKUST Repository (Chine):** Ce répertoire collecte, conserve et diffuse les résultats scientifiques créés par la communauté de The Hong Kong University of Science and Technology. Il répertorie également les publications scientifiques des chercheurs de l'université et fournit une interface propre au corps professoral dans lequel ils peuvent partager leurs publications, études bibliométriques, domaines d'intérêt et projets de recherche (plus de 11280 documents).
- **Dépôt institutionnel de Bouira (Algérie):** L'Université de Bouira propose aux visiteurs de son dépôt institutionnel des fichiers téléchargeables gratuitement, notamment des thèses de doctorat, des mémoires de master dans différentes disciplines technologiques et littéraires, des actes de conférences et des articles scientifiques de haut niveau (plus de 3000 documents).

Le tableau 1 ci-dessous présente les principales fonctionnalités qu'offrent les interfaces de navigation dans les différents dépôts faisant objet du benchmark :

	Theses.fr	UNBC's Digital Institutional Repository	BCU Open Access	HKUST Repository	Dépôt institutionnel de Bouira
Recherche simple	☒	☒	☒	☒	☒
Recherche avancée :					
Titre	☒	☒	☒	☒	☒

Date	<input checked="" type="checkbox"/>				
Auteur	<input checked="" type="checkbox"/>				
Résumé	-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	-	-
Langue	-	-	-	<input checked="" type="checkbox"/>	-
Mots clés	-	-	<input checked="" type="checkbox"/>	-	-
Discipline	-	-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	-
Collection	-	-	-	<input checked="" type="checkbox"/>	-
Région/Ville	-	<input checked="" type="checkbox"/>	-	-	-
Etablissement	<input checked="" type="checkbox"/>	-	-	-	-
Type de contenu	-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	-
Table des matières	-			<input checked="" type="checkbox"/>	-
Navigation directe par :					
Date	<input checked="" type="checkbox"/>	-	<input checked="" type="checkbox"/>	-	<input checked="" type="checkbox"/>
Auteur	<input checked="" type="checkbox"/>	-	<input checked="" type="checkbox"/>	-	<input checked="" type="checkbox"/>
Discipline	<input checked="" type="checkbox"/>	-	<input checked="" type="checkbox"/>	-	<input checked="" type="checkbox"/>
Etablissement	<input checked="" type="checkbox"/>	-	<input checked="" type="checkbox"/>	-	-
Ecole doctorale	<input checked="" type="checkbox"/>	-	-	-	-
Directeur de thèse	<input checked="" type="checkbox"/>	-	-	-	-
Collection	-	<input checked="" type="checkbox"/>	-	-	-
Choix de la langue de navigation	<input checked="" type="checkbox"/>	-	-	-	-
Abonnement au fil RSS	<input checked="" type="checkbox"/>	-	<input checked="" type="checkbox"/>	-	-
Export des résultats de la recherche	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	-
Tri des résultats de la recherche	<input checked="" type="checkbox"/>	-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Tableau 2. Comparatif des interfaces de navigation des répertoires d'OA

Ce tableau dresse les fonctionnalités que doit offrir l'interface de navigation de la plateforme proposée. Un effort d'adaptation est à faire concernant certaines fonctionnalités, comme la recherche : recherche par CEDoc,

faculté, université ; intégration des travaux en arabe avec possibilité de recherche et de navigation en arabe.

Par ailleurs, la réussite de ce projet dépend également de la combinaison de plusieurs ressources à la fois humaines et matérielles. En termes de ressources humaines, il est important d'impliquer plusieurs profils dans dont chacun assurera un volet du projet, à savoir:

Profil	Fonction
Documentalistes	Traitement de la partie métadonnées : création de fiches signalétiques des thèses, indexation...
Informaticiens	Partie technique du projet : mise en place de la plateforme, paramétrage, maintenance, demande d'évolution...
Représentants de chaque CEDoc	Coordination et gestion des relations aussi bien avec les chercheurs et les instances de décision de l'université que l'équipe de travaillant sur la plateforme nationale
Equipe projet « Otrohati »	Equipe de travail permanente assurant la gestion du projet et la coordination du travail avec les représentants de chaque CEDoc

Tableau 3. *Ressources humaines requise pour le projet*

En plus des ressources logicielles, un budget est à consacrer pour les ressources matérielles. Qui dit plateforme nationale dit un nombre considérable d'internautes qui peuvent interroger simultanément cet entrepôt de données. Ce qui exige la mise en place d'un serveur d'une grande capacité, capable de répondre à plusieurs requêtes en même temps.

Par ailleurs, la mise en place de cette plateforme collaborative n'est pas la finalité ultime. Son impact, comme indiqué au début de notre article, s'étend du milieu académique vers celui industriel. En effet, l'ouverture des résultats de la recherche scientifique constitue un enjeu majeur pour la compétitivité des entreprises et des territoires. Les plateformes offrant du contenu en Open Access quel que soit leurs types, représentent des mines d'informations qui peuvent servir une entreprise (de produits ou de services) / un territoire, notamment les résultats des études, des recherches, des recommandations...etc. Ces outputs scientifiques auront un impact certes dans l'appui de l'innovation des entreprises s'ils sont bien exploités dans le processus de collecte/traitement de l'information stratégique dont toute entreprise a besoin pour sa compétitivité.

5 Open Access et intelligence économique : Intérêt de l'Open Access pour la compétitivité des territoires

A l'échelle internationale, l'intelligence économique s'est imposée comme pratique indispensable au service de la compétitivité et du développement

durable. Il devient ainsi primordial pour toute organisation d'identifier et exploiter l'information permettant l'alignement de ses stratégies aux exigences de son environnement spécifique.

L'intelligence économique en tant que pratique est considérée (BESSON et POSSIN, 1996) comme *« la capacité d'obtenir des réponses à des questions en découvrant des intelligences entre deux ou plusieurs informations préalablement mémorisées. L'entreprise mettra au service de cette capacité tous les moyens dont elle dispose pour saisir des opportunités ou détecter des menaces »*.

Le groupe de travail présidé par Henri MARTRE, a défini en 1994 l'intelligence économique comme *« l'ensemble des actions de recherche, de traitements et de diffusion (en vue de son exploitation) de l'information utile aux acteurs économiques... »*.

D'autres définitions mettent l'accent sur la nature de l'information traitée dans le processus d'intelligence économique (LEVET Jean-Louis et PATUREL, 1996) notamment l'information économique dont se sert une organisation en vue d'une action immédiate ou ultérieure, individuelle ou collective. Cette action est atteinte à travers trois types de fonctions qu'assure l'intelligence économique (CLERC Philippe, 1995) à savoir : la maîtrise du patrimoine scientifique et technique, la détection des menaces et des opportunités, l'élaboration de stratégies d'influence. L'intelligence économique constitue ainsi un outil d'interprétation permanente de la réalité des marchés, des techniques et des modes de pensée des concurrents et partenaires, de leur culture, de leurs intentions et de leurs capacités à les mettre en œuvre.

Il ressort de ces définitions l'intérêt et le rôle central que joue l'information au sein du processus d'intelligence économique. Plus l'information collectée est d'une valeur ajoutée, plus l'action d'une organisation sera efficace. C'est dans ce cadre que nous proposons le modèle illustré dans la figure 7 qui montre la place des plateformes d'Open Access dans le processus d'innovation / intelligence économique d'une organisation :

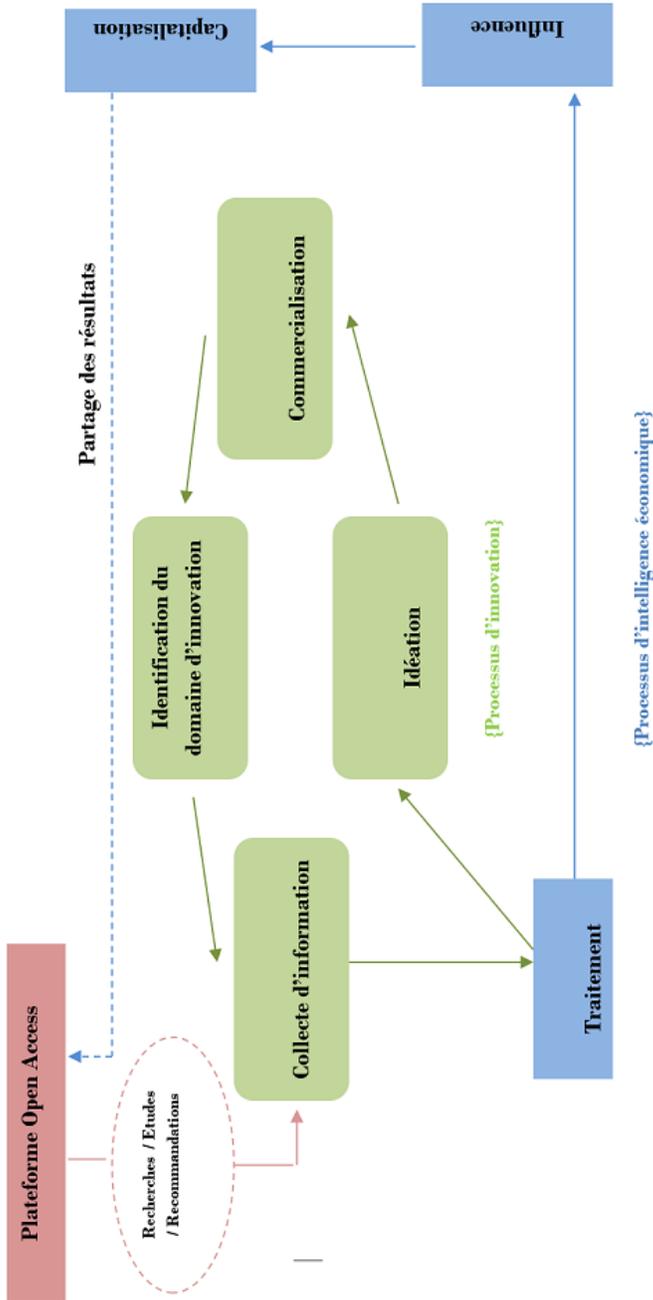


Figure 14. *Modèle conceptuel proposé : Articulation Open Access – Innovation- Intelligence économique*

Pour expliquer la valeur ajoutée des plateformes d'archives ouverte dans l'appui du processus d'innovation des organisations, nous proposons l'exemple suivant :

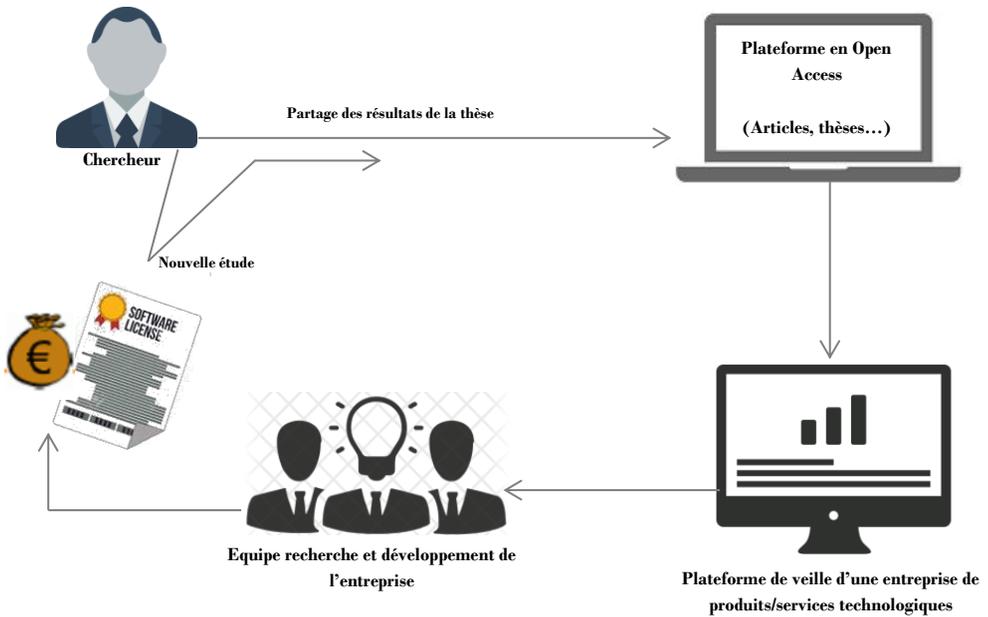


Figure 15. Exemple de la contribution de la recherche scientifique en Open Access au processus d'innovation /intelligence économique

1. Un chercheur lambda travaillant dans le domaine des sciences de l'information a élaboré une thèse sur les logiciels de gestion des bibliothèques virtuelles. Dans sa thèse il a effectué deux études. La première est une enquête auprès des bibliothèques pour étudier leurs besoins en termes de fonctionnalités que doit couvrir le logiciel à utiliser pour créer et gérer une bibliothèque virtuelle. La deuxième étude est sous forme d'un benchmark de l'ensemble des outils offerts sur le marché international. A la fin de sa thèse le chercheur relève des écarts entre les besoins des bibliothèques et l'offre du marché.
2. Le chercheur a partagé les résultats de cette étude sur une plateforme nationale qui offre en Libre Accès les résultats de la recherche scientifique.
3. Une entreprise qui travaille dans le domaine des solutions informatiques de gestion de contenu a intégré dans son processus de veille comme source d'information, les plateformes d'Open Access. Le chargé de veille de

cette entreprise a identifié lors de la collecte des informations intéressant le domaine d'activité de l'entreprise l'étude publiée par le chercheur.

4. Cherchant à améliorer sa compétitivité, l'entreprise a travaillé sur les recommandations et les écarts relevés par le chercheur pour développer un nouveau logiciel de gestion des bibliothèques virtuelles plus complets que les outils offerts sur le marché. Cet outil sera encore utilisé dans d'autres études que d'autres entreprises peuvent exploiter pour élaborer de nouveaux outils plus pointus et répondants au maximum aux exigences des utilisateurs.

Ainsi il convient de conclure que l'ouverture des résultats de la recherche au public ne peut être que bénéfique au développement de la recherche au sein d'un Etat, à l'appui de l'innovation au sein des entreprises et au rayonnement scientifique et économique d'un pays. La plateforme d'archive ouverte jouera le rôle d'une source de veille informationnelle pour une organisation sur les thématiques touchant son domaine d'activité.

6 Conclusion

Le présent travail constitue une proposition visant la contribution au déploiement des Archives Ouvertes dans les différents CEDocs, à travers une plateforme unique favorisant la collaboration entre chercheurs. Notre approche s'est axée sur l'analyse de l'état des lieux des archives ouvertes dans le monde et au niveau national. Nous avons également démontré la contribution de la plateforme des archives ouvertes dans le processus d'intelligence économique-innovation des organisations, notamment le rôle de source de veille informationnelle que jouera la plateforme pour ces organisations.

Notre proposition est motivée par l'essor international qu'a connu le mouvement du Libre Accès où plusieurs pays se sont mis à développer politiques adaptées à leur contexte, obligeant les chercheurs à se mettre dans la voie de l'Open Access comme la commission européenne qui a élaboré le programme Horizon 2020 qui comporte l'obligation d'assurer le Libre Accès aux publications issues des recherches qu'il aura contribué à financer (COMMISSION EUROPEENNE, 2014). Il en va de même pour les projets financés par le Wellcome Trust¹¹⁰ au Royaume-Uni ou ceux financés par les National Institutes of Health¹¹¹. La France a également élaboré son propre programme. Il s'agit de « Persée¹¹² » qui a pour objet la numérisation des revues en SHS et le développement des archives ouvertes à travers le dispositif HAL (Hyper Article en Ligne).

¹¹⁰Projets Wellcome Trust. [En ligne] : <https://wellcome.ac.uk/funding/scheme-finder> (Consulté le 11 septembre 2018)

¹¹¹NIH Public Access Policy. [En ligne] : <https://publicaccess.nih.gov/index.htm> (Consulté le 11 septembre 2018)

¹¹²Portail Persée. [En ligne] : <https://www.persee.fr/a-propos> (Consulté le 5 août 2018)

Ces pays ont développé une expertise qui mérite d'être analysée pour être décliné aux niveaux de plusieurs pays, comme le Maroc, qui manquent d'initiatives plus offensives en faveur du Libre Accès.

L'intégration de la « société du Libre Accès », si on peut l'appeler ainsi, ne peut se faire qu'à travers une politique nationale favorisant le libre accès à la production du chercheur marocain.

C'est dans ce cadre que nous envisageons, comme perspective à ce travail, la mise place d'une proposition de politique nationale d'open accès en travaillant avec une équipe multidisciplinaire afin d'aborder aussi bien les aspects conceptuels que techniques et juridiques du projet.

Bibliographie

BEN ROMDHANE, Mohamed, OUEREFELLI, Taek. (2013). *L'offre des archives ouvertes dans le monde arabe : recensement et évaluation*. [En ligne]. Disponible sur : http://archivesic.ccsd.cnrs.fr/sic_00829058 (Page consultée le 12 aout 2018).

BESSON, Bernard, POSSIN, Jean-Claud (1996). *Du renseignement à l'intelligence économique : Détecter les menaces et les opportunités pour l'entreprise*. Dunod, 1996. p. 20

CHANIER, Thierry. (2004). Archives ouvertes et publication scientifique : comment mettre en place l'accès libre et aux résultats de la recherche ? Paris : L'Harmattan, 2004. 186 p.

CHARTRON, Ghislaine. (2016). Stratégie, politique et reformulation de l'Open Access. *Revue française des sciences de l'information et de la communication*. [En ligne]. Disponible sur : <http://journals.openedition.org/rfsic/1836> (Page consultée le 17septembre 2018).

CLERC, Philippe (1995). Encyclopaedia Universalis. p. 195.

Commission européenne. (2014). *Le programme-cadre de l'UE pour la recherche et l'innovation*. [En ligne]. Disponible sur : https://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_FR_KI0213413FRN.pdf (Page consultée le 10 aout 2018).

Déclaration de Budapest. [En ligne]. Disponible sur : <http://www.budapestopenaccessinitiative.org/> (Page consultée le 10 aout 2018)

Erasmus+ Project Results. Optimisation de l'accès à l'information scientifique et technique dans les universités du Maghreb. [En ligne]. Disponible sur : <https://ec.europa.eu/programmes/erasmus-plus/projects/eplu-project-details/#project/510986-TEMPUS-1-2010-1-BE-TEMPUS-JPGR> (Page consultée le 12 août 2018)

HAL. Guide pratiques : mieux connaître l'Open Access et HAL. [En ligne]. Disponible sur : <https://hal-clermont-univ.archives-ouvertes.fr/public/GuidePratiqueOpenAccess.pdf> (Page consulté le 5 juillet 2018)

- HARNAD Stevan. (1995). Universal FTP archives for esoteric science and scholarship: a subversive proposal. Association of Research Libraries, Washington DC. [En ligne]. Disponible sur : <http://www.arl.org/scomm/subversive/toc.html> (Page consultée le 12 aout 2018).
- JONCHERE, Laurent. (2013). *Synthèse sur les politiques institutionnelles de libre accès à la recherche*. Archive SIC.[En ligne]. Disponible sur : <https://archivesic.ccsd.cnrs.fr/docs/.../Synthese-politiques-LA-Jonchere-fev-2013.pdf> (Consulté le 10 août 2018)
- LEVET, Jean-Louis ; PATUREL, Robert. (1996). *L'intégration de la démarche d'intelligence économique dans le management stratégique*. 5ème Conférences de l'Association Internationale de Management Stratégique, les 13-14-15 Mai, Lille. [En ligne]. Disponible sur : www.drdsi.cerist.dz/SNIE/david.pdf (Consulté le 10 août 2018)
- NWAGWU, Williams; ALLAM, Ahmed. (2009). *Building Open Access in Africa*, International Journal of Technology Management, 2009, 45 (1/2): Enligne]. Disponible sur:http://www.inderscience.com/www/pdf/ijtmv45n12_oainafrika.pdf (Consulté le 10 août 2018)
- 6.1** OpenAccess France. Politiques nationales et européennes. [En ligne]. Disponible sur :<https://openaccess.couperin.org/politiques-nationales-et-europeennes/>(Consulté le 10 août 2018)
- Research Information Network.*Open science case studies*. [En ligne]. Disponible sur:http://www.rin.ac.uk/system/files/attachments/RINews_Issue_11_0.pdf (Consulté le 9 août 2018)
- REYMONET, Nathalie (2012). *Open Access : voies et modèles économiques*. [En ligne]. Disponible sur:<https://appui-recherche.univ-paris-diderot.fr/file/369/download?token=buVgq7Pw>(Page consultée le 12 aout 2018).
- The RIN / NESTA report Open to All? Case Studies of Openness in Research* (2010)[Enligne].Disponible sur :<http://www.rin.ac.uk/our-work/data-management-and-curation/openscience-case-studies>(Page consultée le 12 aout 2018).
- VALLUY, Jérôme. (2017). *Libre accès aux savoirs et accès ouvert aux publications*. *Revue française des sciences de l'information et de la communication*. [En ligne]. Disponible sur : <http://journals.openedition.org/rfsic>(Consulté le 10 août 2018)

Annexes

Liste des dépôts d'archives ouvertes présentées dans le benchmark

BCU Open Access. [En ligne]. Disponible sur : <https://www.bcu.ac.uk/research/open-access>(Page consultée le 12 aout 2018).

Dépôt institutionnel de Bouira.[En ligne]. Disponible sur : <http://dspace.univ-bouira.dz:8080/jspui/>(Page consultée le 12 aout 2018).

HKUST Repository. [En ligne]. Disponible sur : repository.ust.hk/(Page consultée le 12 aout 2018).

Theses.Fr.[En ligne]. Disponible sur : <http://www.theses.fr/>(Page consultée le 12 aout 2018).

UNBC's Digital InstitutionalRepository.[En ligne]. Disponible sur : <https://unbc.arcabc.ca/>(Page consultée le 12 aout 2018).

Indicateurs de consultation, indicateurs stratégiques : leur production à partir de l'analyse des consultations des documents d'une archive en libre accès

Accesses indicators, strategic indicators: their production from log files analysis of an open access archive

ROSSI Pier Luigi

Institut de recherche pour le développement (IRD)

32, avenue Henri Varagnat

93140 Bondy - France

Mail:rossi@ird.fr

Résumé

L'archive documentaire de l'IRD (Institut de recherche pour le développement) est constituée par l'ensemble des productions scientifiques de l'Institut. Actuellement plus de 66.500 documents sont disponibles gratuitement sur Internet. En associant, en mode relationnel, la base de données bibliographiques et celle concernant les consultations de ces documents (« fichiers log Apache »), nous pouvons produire des tableaux de bord dynamiques en fonction de plusieurs critères et de plusieurs indicateurs. Nous allons présenter plusieurs indicateurs originaux ainsi que nos méthodes d'analyse et de production.

Mots-clés. Indicateurs, accès internet, libre accès, fichiers journaux, analyses statistiques.

Abstract

The IRD (French National Research Institute for Sustainable Development) archive contains all the scientific outputs of the Institute. Currently more than 66,500 documents are freely available on the Internet. By associating, in relational mode, the bibliographic database and accesses database ("Apache log files"), we can produce dynamic dashboards according to several criteria and several indicators. We will present several original indicators, our analysis methods and indicators production.

Keywords. Indicators, internet access, free access, log files, statistical analysis.

1 Introduction

L'IRD (Institut de recherche pour le développement)¹ est un institut de recherche français au service des pays méditerranéens et intertropicaux qui font de la science et de l'innovation l'un des premiers leviers de leur développement. Il s'agit d'un établissement public français à caractère scientifique et technologique (EPST) placé sous la double tutelle du ministère de l'Enseignement supérieur et de la Recherche et du ministère des Affaires étrangères et du Développement international².

Le décret du 1er décembre 1955 (ORSTOM, 1955) organisant les services de l'Office de la recherche scientifique et technique outre-mer, publié au Journal officiel de la République française du 21 décembre 1955, définit l'existence d'un centre de documentation rattaché à la Direction de l'Institut. L'article 6 stipule que « le Centre de documentation est chargé d'assurer le dépouillement, la conservation et la diffusion de la documentation scientifique et technique se rapportant aux activités de l'O.R.S.T.O.M. ».

Depuis 1955, l'Institut a donc créé une archive institutionnelle avec la volonté de préserver et de diffuser ses productions scientifiques. Ce fonds documentaire patrimonial se compose actuellement (août 2018) de 98.400 documents et constitue le fonds documentaire de l'IRD (FDI).

Cette collection a été informatisée depuis 1986 avec la création de la base de données bibliographiques Horizon (Roux-Fouillet, 1988). Dès 1996, nous avons initié sa numérisation (Rossi, 1997) et 66.500 documents sont actuellement (août 2018) disponibles gratuitement sur Internet (Rossi et Ngoma-Mouaya, 2000).

Compte tenu des implantations géographiques des centres de recherche et des missions des chercheurs, tout au long de l'histoire de l'Institut, une partie significative de travaux scientifiques réalisés (plus de 39.000 documents en accès libre) concerne les pays du continent africain.

2 Fichiers journaux³ et filtrage de données

Le site hébergeant l'archive institutionnelle de l'IRD fonctionne avec Apache HTTP Server⁴. Ce logiciel stocke de manière détaillée l'historique des consultations concernant les échanges de fichiers qui se réalisent entre le serveur et le client (la machine qui a consulté les fichiers disponibles sur le serveur) (Agostiet *al.*, 2012).

¹ Voir : <http://www.ird.fr/l-ird/presentation> (Page consultée le 6 août 2018).

² L'IRD a été créé en 1943 sous le nom d'Office de la recherche scientifique coloniale (ORSC). En 1953, il est devenu Orstom (Office de la recherche scientifique et technique d'outre-mer). En 1984, il prend le nom d'Institut français de recherche scientifique et technique pour le développement en coopération, tout en conservant son acronyme et devient IRD en 1998. Voir : <http://www.ird.fr/l-ird/historique> (Page consultée le 6 août 2018).

³ Voir : [https://fr.wikipedia.org/wiki/Historique_\(informatique\)](https://fr.wikipedia.org/wiki/Historique_(informatique)) (Page consultée le 7 août 2018).

⁴ Voir : <https://httpd.apache.org/> (Page consultée le 7 août 2018).

```
154.126.12.202 - - [12/Feb/2017:17:50:33 +0100] "GET /exl-
doc/pleins_textes/pleins_textes_5/b_fdi_08-09/10230.pdf HTTP/1.1" 200
2300612
"http://www.google.fr/search?ei=bZGgWJHPiIzhvATnxpywDQ&q=trachysp
haera+fructigena+pdf&btnG=" "Mozilla/5.0 (Linux; U; Android 4.2.2; en-US;
HTC Butterfly Build/JDQ39) AppleWebKit/534.30 (KHTML, like Gecko)
Version/4.0 UCBrowser/11.1.5.890 U3/0.8.0 MobileSafari/534.30"
```

Figure 1. *Format du fichier d'historique, en mode « combiné »*

Le format du fichier d'historique, en mode « combiné »⁵, contient plusieurs informations (figure 1) incluant: l'adresse IP du client (154.126.12.202), l'heure à laquelle la demande a été reçue (12/Feb/2017:17:50:33 +0100), l'URL du fichier consulté (/exl-doc/.../10230.pdf), le code d'état renvoyé par le serveur (200), le site depuis lequel le client a lancé sa requête⁶ (http://www.google.fr ...), la question éventuelle posée par le client (trachysphaerafructigenapdf), le type d'appareil utilisé (... Mobile ...) et le navigateur utilisé par le client (...Safari ...).

Pour mener à bien notre étude, nous avons géolocalisé les adresses IP en utilisant la base de données GeoLite Country de MaxMind⁷. Dans l'exemple de la figure 1, l'adresse IP est située à Madagascar.

Le fichier journal est filtré (KrishnagandhietDhas, 2016) en conservant uniquement les lignes ayant accès aux fichiers PDF avec le statut « 200 »⁸. Le fichier résultant est ensuite décomposé en plusieurs sous-fichiers. Le champ contenant le type de navigateur inclut également la signature des robots de recherche. Cela permet de générer des fichiers journaux relatifs à l'indexation effectuée par les moteurs de recherche, avec notamment le moteur de recherche Google (figure 2).

```
66.249.64.221 - - [07/Feb/2017:12:01:04 +0100] "GET /exl-
doc/pleins_textes/divers17-01/010068365.pdf HTTP/1.1" 200 631342 "-"
"Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
```

Figure 2. *Indexation d'un fichier effectuée par « Googlebot »*

Nous rejetons et générons un fichier "spam" avec toutes les lignes du fichier journal qui masquent par un tiret ("-") le champ référent (le champ du site à partir duquel le client a lancé sa requête) ou le champ navigateur : nous présumons que les requêtes qui masquent ces deux champs peuvent être assimilées soit à une attaque par déni de service⁹ soit à un « referer spam »¹⁰. Selon la même logique, on rejette également les accès provenant de la même adresse IP et consultant le même fichier qui sont réalisés dans un intervalle inférieur à 5 minutes.

⁵ Voir : <https://httpd.apache.org/docs/2.4/fr/logs.html> (Page consultée le 7 août 2018).

⁶ Voir : [https://fr.wikipedia.org/wiki/Référent_\(informatique\)](https://fr.wikipedia.org/wiki/Référent_(informatique)) (Page consultée le 7 août 2018).

⁷ Voir : <https://dev.maxmind.com/geoip/legacy/geolite/> (Page consultée le 7 août 2018).

⁸ Voir : <https://tools.ietf.org/html/rfc2616#page-58> (Page consultée le 7 août 2018). Le statut « 200 » du code http indique que la requête a été traitée avec succès.

⁹ Voir https://fr.wikipedia.org/wiki/Attaque_par_déni_de_service (Page consultée le 7 août 2018).

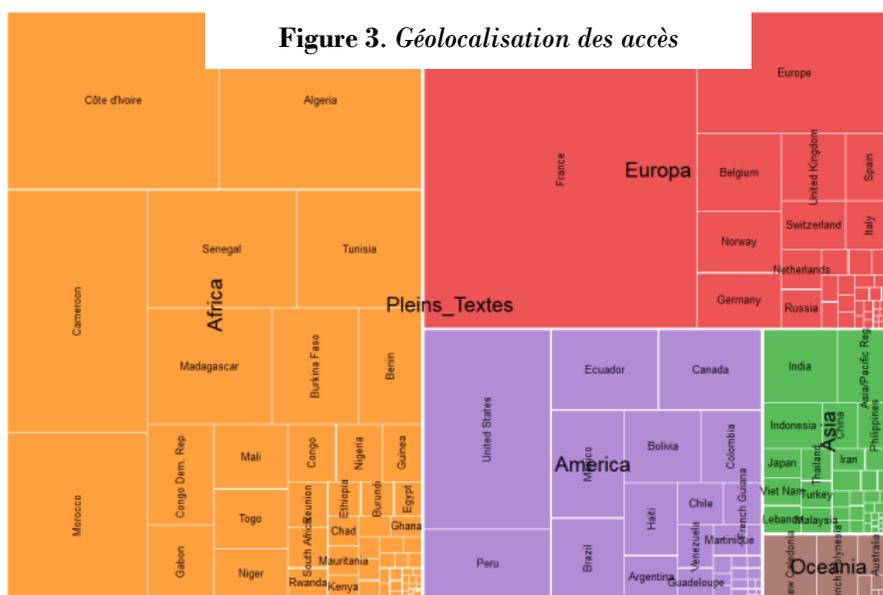
¹⁰ Voir https://fr.wikipedia.org/wiki/Referer_spam (Page consultée le 7 août 2018).

Le fichier de travail filtré (Jansen, 2006) qui se compose de tous les accès « validés » (« humanaccess ») contient 2.748.000 lignes, soit 39% du fichier brut d'origine. Le fichier des « accès robots » et le fichier des « accès spam » représentent respectivement 42% et 19% du fichier brut d'origine.

3 Répartition géographique des accès

Après ces différentes étapes de filtrage (Kaur et Aggarwal, 2017), nous avons un fichier constitué (pour l'essentiel) de « humanaccess ». Il est prêt à produire des indicateurs de statistiques et d'accès (Martinez-Comeche, 2017 ; Rossi et Thiaw, 2012 ; Rossi *et al.*, 2018).

Comme tous les accès ont été géolocalisés, il est possible de produire des statistiques en fonction de la répartition géographique par pays et par continents. Nous présentons ces résultats à l'aide de graphiques « treemap »¹¹ qui nous ont parus bien adaptés à la présentation de ce type de données (figure 3).



¹¹ Voir : <https://fr.wikipedia.org/wiki/Treemap> (Page consultée le 7 août 2018). Voir également : Treemaps for Space-Constrained Visualization of Hierarchies: The History of Treemap Research at the University of Maryland. <http://www.cs.umd.edu/hcil/treemap-history/index.shtml> (Page consultée le 7 août 2018).

Si l'on considère les données « validées » de l'année 2017¹², les accès sont situés dans 234 pays dont 56 pays africains. La France est le premier pays où les consultations sont localisées avec 459.021 accès, soit 17% du total. Pour les pays du continent africain, on compte 1.290.723 accès, 47% du total (Europe: 29%, Amériques: 18%, Asie: 5,1%, Océanie: 1,6%). La Côte d'Ivoire est le premier pays d'Afrique avec 199.424 accès, soit 7,3% du total. Les pays du Maghreb (Algérie, Maroc et Tunisie) génèrent 392.516 accès, soit 14% du total.

Le Nigéria est le premier pays anglophone d'Afrique avec 14.201 accès, soit 0,5% du total. Les États-Unis sont le premier pays des Amériques avec 134.998 accès, soit 4,9% du total. L'Inde est le premier pays d'Asie avec 28.800 accès, soit 1,1% du total. La Nouvelle-Calédonie est le premier pays d'Océanie avec 18.406 accès, soit 0,7% du total.

Ces statistiques peuvent témoigner des contextes géopolitiques tendus caractérisant certains pays africains. Notre Institut a, par exemple, produit un grand nombre de publications sur le Tchad (1.344 documents) et la République centrafricaine (787 documents) qui sont librement accessibles. Le nombre de consultations dans ces pays est relativement faible: pour le Tchad, on compte 5.110 accès en 2017, soit 0,19% du total et pour la République centrafricaine 3.560 accès, soit 0,13% du total.

Ces résultats s'expliquent par la prise en compte du nombre d'internautes par rapport à la population du pays (pénétration internet), qui est respectivement de 5% pour le Tchad et de 5,4% pour la République centrafricaine (Internet World Stats 2018). Ces éléments factuels sont amplifiés par la forte instabilité politique et sociale qui caractérise ces pays depuis plusieurs années. Le prix de l'accès à Internet est également un facteur limitant dans plusieurs pays africains (Cable.co.uk 2017).

L'analyse de géolocalisation des accès peut être croisée avec la langue des documents. Le graphique des consultations des documents en français¹³ est visuellement semblable au graphique de la figure 3, mais les pays francophones sont caractérisés par des pourcentages plus élevés. On compte 2.217.559 accès à des documents en français (ce qui représente 80,7% des de l'ensemble des consultations), situés dans 224 pays dont 55 pays africains.

La France est le premier pays où les consultations sont localisées avec 447.349 accès, soit 20,2% du total. Pour les pays du continent africain, on compte 1.243.482 accès, 56% du total (Europe: 32%, Amériques : 8,2%, Asie: 1,9%, Océanie: 1,6%).

La Côte d'Ivoire est le premier pays d'Afrique avec 198.283 accès, soit 8,9% du total. Les pays du Maghreb (Algérie, Maroc et Tunisie) génèrent 388.633 accès, soit 18% du total. Ce filtrage de données (langue du

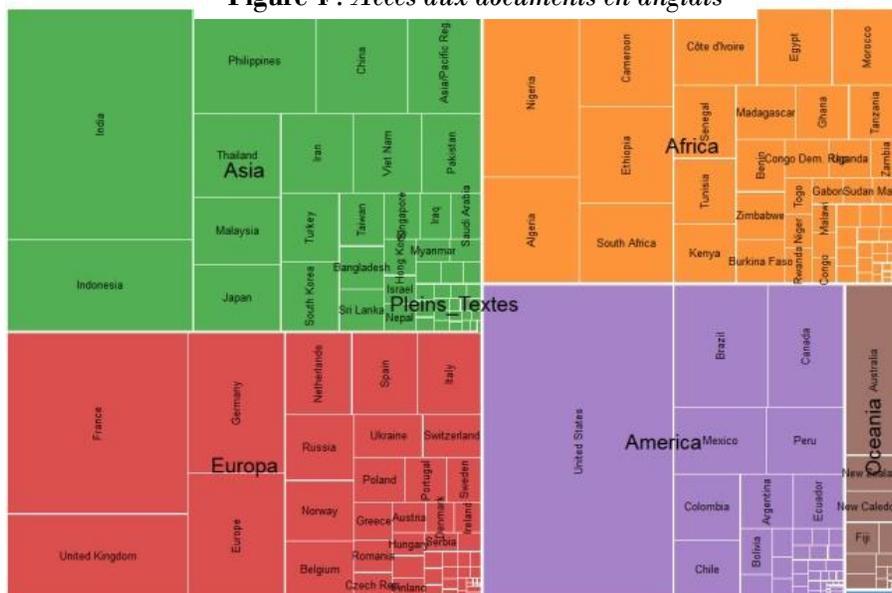
¹² Le tableau de bord complet est accessible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/2017.htm (Page consultée le 9 août 2018). Le tableau de bord contient, entre autre, des informations sur les répartitions et les fréquences des langues des documents consultés. On consultant le tableau on constate que les documents en espagnol sont deux fois plus consultés que les documents en français et que ceux en anglais sont deux fois moins consultés que ceux en français.

¹³ Le tableau de bord complet est accessible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/FR-2017.html (Page consultée le 9 août 2018).

document :français), fait apparaître Haïti (0,9%, 20866 accès) dans les Amériques et le Liban (0,2%, 5302 accès) en Asie.

En ce qui concerne les documents en anglais (figure 4)¹⁴, on compte 326.351 accès, situés dans 228 pays, dont 55 pays africains. L'Asie est le premier continent avec 94.952 accès, soit 29% du total. L'Inde représente 8,2% du total avec 26.667 accès.

Figure 4 : Accès aux documents en anglais



Les États-Unis sont le premier pays où les accès sont situés avec 37.244 accès, soit 11,4% du total.

Pour les pays du continent africain, on dénombre 71.435 accès, 22% du total (Europe: 24%, Amériques: 22%, Océanie: 3,3%). Le Nigeria est le premier pays d'Afrique avec 10.104 accès, soit 3,1% du total.

En ce qui concerne les documents en espagnol (figure 5)¹⁵, on compte 234.059 accès, situés dans 161 pays. Les Amériques sont le premier continent avec 221.135 accès, soit 94% du total. L'Equateur est le premier pays avec 43.295 accès, soit 18% du total. Les pays andins les moins avancés (Equateur, Pérou et Bolivie) génèrent 113.344 accès, soit 48% du total. L'Espagne est le premier pays d'Europe avec 6.368 accès, soit 2,7% du total.

¹⁴ Le tableau de bord complet est accessible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/EN-2017.html (Page consultée le 9 août 2018).

¹⁵ Le tableau de bord complet est accessible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/ES-2017.htm (Page consultée le 9 août 2018).



Figure 5 : *Accès aux documents en espagnol*

Ces résultats sont révélateurs de l'impact des productions de l'IRD, publiées en espagnol, dans les pays d'Amérique latine : en 2017 2.241 fichiers en libre accès ont été consultés au moins une fois et la moyenne des consultations est de 104, une moyenne 4 fois plus importante que celle des consultations des documents en langue anglaise.

4 Tableaux de bord¹⁶, construction et indicateurs

En croisant de façon relationnelle la base de données bibliographiques et celles des consultations, il est possible de générer de façon dynamique des tableaux de bord en fonction de plusieurs critères : auteur, année de publications, langue du document, thématique des documents, pays à l'origine des consultations.

Ces tableaux de bord sont constitués de plusieurs éléments. Si l'on considère le tableau de bord du « thème auteur »¹⁷ (analyse des consultations des documents en libre accès d'un auteur), le premier élément est un tableau qui résume les principales données.

¹⁶ Les tableaux de bord sont réalisés en mode dynamique avec des formulaires au format html intégrant des scripts en php et des requêtes vers la base de données sql contenant les tables des statistiques de consultations et des métadonnées bibliographiques des documents du fonds documentaire de l'IRD. Les graphes sont réalisés avec des bibliothèques javascript. Pour les graphes treemap nous utilisons « protovis » (voir : <https://github.com/mbostock/protovis>, page consultée le 9 août 2018) et pour les autres graphes nous utilisons « highcharts » (voir : <https://www.highcharts.com/>, page consultée le 9 août 2018).

¹⁷ Le tableau de bord complet est accessible à http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/Roose-2017.htm (Page consultée le 7 août 2018).

Indicateur	Valeur
Nombre de fichiers « Pleins_Textes » consultés	350
Nombre de consultations « Pleins_Textes »	27989
Nombre moyen de consultations par fichier « Pleins_Textes »	79,97
Facteur D (nombre de documents consultés au moins "n" fois)	75
Somme des consultations des Top 75	24120
Consultations Top 75 sur le total des consultations des documents de Roose, Eric	86,2%

Figure 6 : Résumé des principaux indicateurs d'un tableau de bord

Pour cet exemple (figure 6), au cours de l'année 2017, 350 différents fichiers du même auteur (Roose, Eric) ont été consultés au moins une fois. Pour ces documents, 75 ont été consultés au moins 75 fois au cours de l'année 2017. Cette valeur, indiquant le nombre de documents consultés au moins n fois (calculée avec un script en php), détermine le « Facteur D »¹⁸ de l'auteur pour la période annuelle des consultations.

Dans le tableau de bord de l'auteur on affiche ensuite le tableau des « n » documents les plus consultés (la valeur de n est égale au Facteur D). Pour cet exemple (figure 7), le document « Top 1 » à été consulté 2.319 fois en 2017, le pays ayant le plus consulté ce document (premier pays consultant) est le Cameroun (291 accès) et l'indice d'appropriation du Cameroun est de 0,13 (soit 13% de l'ensemble des consultations). Si l'on parcourt ce tableau, on remarque que le deuxième document porte sur la Côte d'Ivoire (voir le titre de la référence bibliographique) : dans ce cas l'indice d'appropriation du premier pays consultant (Côte d'Ivoire) est 0,70.

Figure 7 : Extrait du tableau des documents les plus consultés

Tableau des Top 75 (nombre de fichiers = Facteur D) (2017)

Classement	Nombre de consultations	Premier pays consultant	Indice d'appropriation par le premier pays consultant	Référence bibliographique
Top 1	2319	Cameroun (291)	0.13	Roose, Eric. Causes et facteurs de l'érosion hydrique sous climat tropical : conséquences sur les méthodes antiérosives. 1984 (Sc. Terre)
Top 2	1617	Côte d'Ivoire (1128)	0.70	Roose, Eric; Chéroux, Michel; Humbel, François-Xavier (collab.); Perraud, Alain (collab.). Les sols du bassin sédimentaire de Côte d'Ivoire. 1966 (Sc. Terre)
Top 3	1323	Cameroun (174)	0.13	Roose, Eric (ed.); Duchaufour, H. (ed.); De Noni, Georges (ed.). Lutte antiérosive : réhabilitation des sols tropicaux et protection contre les pluies exceptionnelles. 2012 (Sc. Terre)
Top 4	1051	Côte d'Ivoire (179)	0.17	Roose, Eric. Impact du défrichement sur la dégradation des sols tropicaux. 1984 (Sc. Terre)
Top 5	955	Morocco (571)	0.60	Roose, Eric (ed.); Sabir, M. (ed.); Laouina, A. (ed.); Benchakroun, F. (collab.); Al Karkouri, J. (collab.); Lauri, P. (collab.); Qarro, M. (collab.). Gestion durable des eaux et des sols au Maroc : valorisation des techniques traditionnelles méditerranéennes. 2010 (Sc. Terre)
Top 6	762	Benin (141)	0.19	Roose, Eric. Dégradation des terres et développement en Afrique de l'Ouest. 1985 (Sc. Terre)
Top 7	734	France (401)	0.55	Roose, Eric; Blancaneaux, Philippe; Freitas, P.L. de. Un simple test de terrain pour évaluer la capacité d'infiltration et le comportement hydrodynamique des horizons pédologiques superficiels : méthode et exemples. 1993 (Sc. Terre)

¹⁸ Cet indicateur, Facteur D, ou Downloading factor, s'inspire directement de l'indice H. Voir : https://fr.wikipedia.org/wiki/Indice_h (Page consultée le 7 août 2018).

L'observation complète du tableau des documents les plus consultés, montre que si un document porte sur un pays spécifique (dont le nom apparaît généralement dans le titre du document), « le premier pays consultant » est le pays indiqué dans le titre et son indice d'appropriation est élevé.

Par ailleurs, ce tableau permet d'observer que le nombre des consultations est indépendant de l'année de publication des documents : le déterminant des consultations, et, par conséquent, de leur nombre, est le contenu du document ainsi que son titre. Les pages de résultats des moteurs de recherche¹⁹ qui s'affichent après une requête comportent le titre des documents trouvés et ceux-ci déterminent, très probablement, les choix de consultation des internautes.

Le tableau de bord « auteur », comporte ensuite le graphe « treemap » avec la répartition géographique totale des consultations. Dans le cas spécifique, avec les outils intégrés d'analyse dynamique du graphe on constate que les consultations venant de l'Afrique constituent 17.509 accès (63% du total), réparties sur 52 pays.

Le nombre total des fichiers consultés et le nombre total des consultations sont ensuite repartis en fonction de la langue des documents et représentés sur des graphes. La comparaison de ces valeurs permet de générer un indice de spécialisation en fonction de la langue. Pour l'exemple choisi (figure 8),

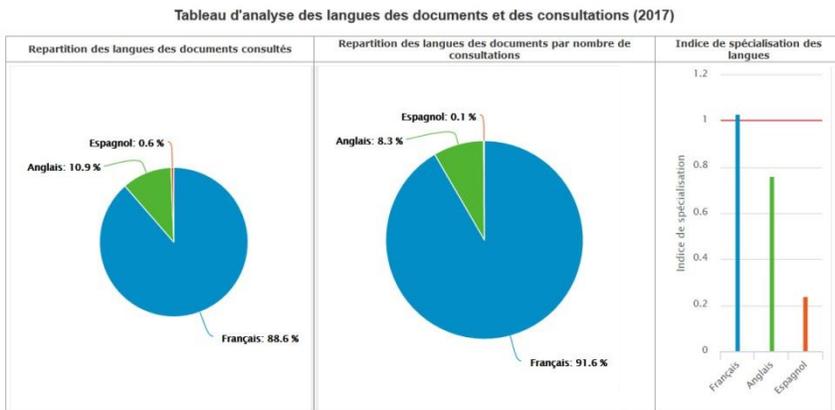


Figure 8 : Répartition des accès en fonction de la langue des documents

on constate ainsi que les documents en français sont légèrement sur-consultés (1,03) et que les documents en anglais (0,76) et en espagnol (0,24) sont sous-consultés.

Le tableau de bord intègre un graphe dynamique qui permet de calculer et visualiser l'indice de Pareto²⁰. Pour cet exemple on constate qu'il suffit de considérer 16% des fichiers le plus consultés pour atteindre 80% des consultations.

¹⁹ Il est à noter que Google est à l'origine de plus de 85% des requêtes faites par les internautes vis-à-vis de notre serveur. Ce résultat est obtenu par le comptage de champ « referer » dans la table des accès « validés » de l'année 2017.

²⁰ Voir : https://fr.wikipedia.org/wiki/Indice_de_Pareto (Page consultée le 10 août 2018).

5 Répartitions thématiques des consultations des pays

Dans la base de données bibliographique du fonds documentaire de l'IRD, chaque document est associé à un thème (indexation des traitements effectuée par les documentalistes de l'IRD). Les thèmes de la base de données bibliographiques sont les suivants : Sciences de l'ingénieur (Sc. Ing.), Océanographie-Hydrobiologie (Oceano + Hydrobio.), Santé (Santé), Sciences de la Terre (Sc. Terre), Sciences végétales et animales (Sc. Veg + Ani), Sciences humaines et sociales (Sc. Hum. + Soc.).

Les données d'accès peuvent être filtrées par rapport à un pays et associées aux thèmes des documents. Ces données "pays" peuvent être comparées à la distribution globale des accès par thème. Ceci montre un index d'intérêt (indice de spécialisation thématique) pour les thèmes par rapport à chaque pays (figure 9).

Pour l'Afrique du Sud²¹, les documents concernant l'Océanographie-Hydrobiologie et les Sciences végétales et animales sont sur-consultés. Les documents concernant les Sciences de l'ingénieur et les Sciences de la Terre sont sous-consultés.

Pour l'Algérie²², les documents concernant les Sciences de l'ingénieur, les Sciences de la Terre et les Sciences végétales et animales sont sur-consultés. Les documents concernant les Sciences humaines et sociales sont sous-consultés.

Pour la Côte d'Ivoire²³, les documents concernant les Sciences humaines et sociales sont sur-consultés. Les documents concernant tous les autres sujets sont sous-consultés.

Pour l'Égypte²⁴, les documents concernant l'Océanographie-Hydrobiologie et les Sciences végétales et animales sont sur-consultés. Les documents concernant les Sciences de l'ingénieur et les Sciences humaines et sociales sont sous-consultés.

²¹ Le tableau de bord complet est accessible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/ZAF-2017.htm (Page consultée le 9 août 2018).

²² Le tableau de bord complet est accessible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/DZA-2017.htm (Page consultée le 9 août 2018).

²³ Le tableau de bord complet est accessible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/CIY-2017.htm (Page consultée le 9 août 2018).

²⁴ Le tableau de bord complet est accessible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/EGY-2017.htm (Page consultée le 9 août 2018).

Pour le Maroc²⁵, les documents concernant les Sciences de l'ingénieur et les

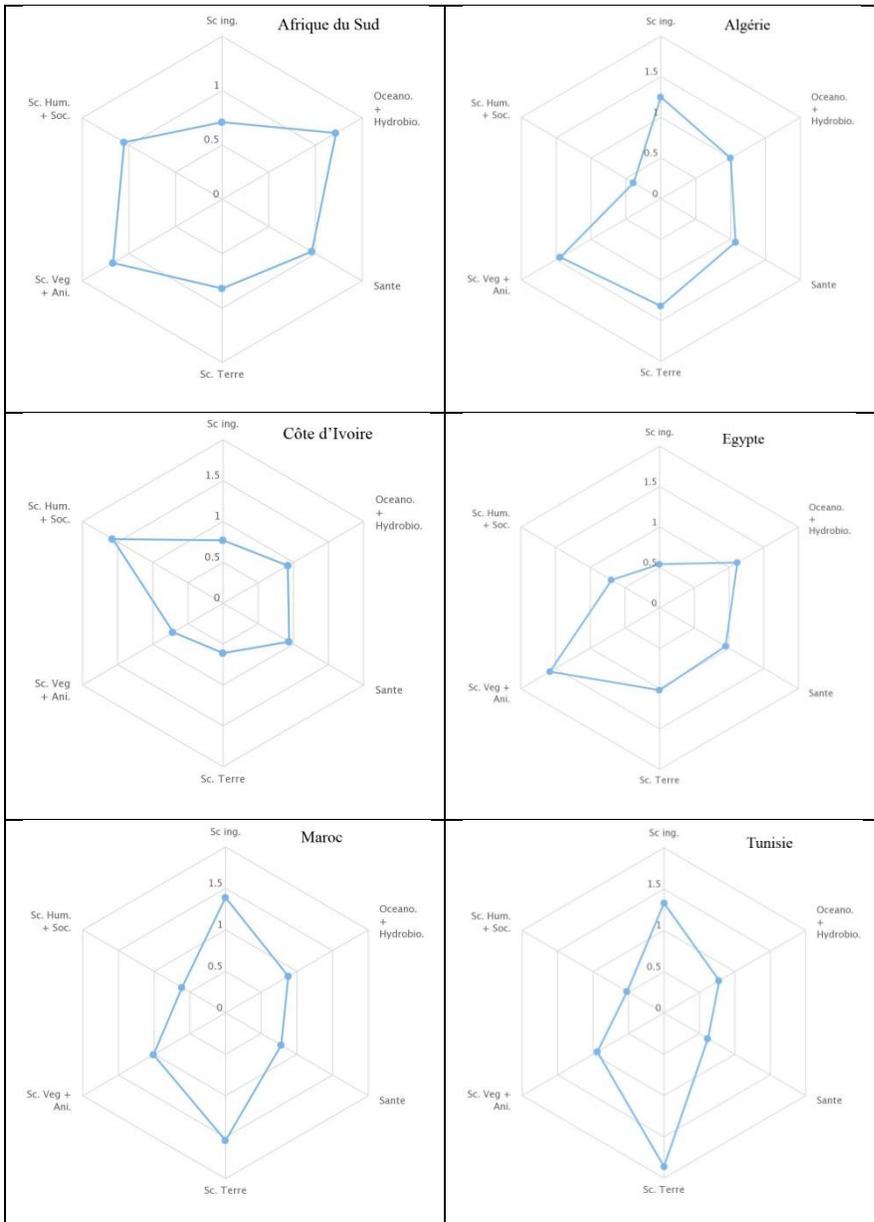


Figure 9. Répartitions thématiques des consultations de 6 pays d'Afrique Sciences de la Terre sont sur-consultés. Les documents concernant les

²⁵ Le tableau de bord complet est accessible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/MOR-2017.htm (Page consultée le 9 août 2018).

Sciences humaines et sociales, l'Océanographie-Hydrobiologie et la Santé sont sous-consultés.

Pour la Tunisie²⁶ la répartition des thématiques des consultations est comparable à celle du Maroc. On remarque néanmoins un très fort intérêt pour les documents concernant les Sciences de la Terre.

En analysant les profils disciplinaires de ces pays, on remarque que les graphes de l'Afrique du Sud et de l'Égypte (pays essentiellement anglophones) sont semblables. Il en est de même pour le Maroc et la Tunisie avec des intérêts très forts pour les Sciences de l'ingénieur et les Sciences de la Terre.

6 Pédologie et Hydrologie

Au cours de 2016 et 2017, nous avons mené une campagne de numérisation pour traiter toutes nos productions scientifiques concernant la Pédologie (Fargier 2015) et l'Hydrologie. Ce sont deux disciplines qui ont caractérisé l'histoire scientifique de notre institut et qui ont permis de produire un nombre important de publications particulièrement significatives pour les pays dans lesquels elles ont été réalisées : aussi bien en Afrique qu'en Amérique latine.

En 2017, 6.696 documents concernant la Pédologie ont été consultés au moins une fois et le nombre total des accès est de 331.600²⁷. Les accès se répartissent dans 213 pays dont 56 pays africains. Les documents ont donc été consultés en moyenne 50 fois, pour toute l'année 2017. La France est le premier pays où les consultations sont localisées avec 55.813 accès, soit 17% du total. Pour les pays du continent africain, on compte 179.310 accès, 54% du total (Europe : 28%, Amériques : 13%, Asie : 4,2%, Océanie : 0,7%). L'Algérie est le premier pays d'Afrique avec 39.048 accès, soit 12% du total. Les pays du Maghreb (Algérie, Maroc et Tunisie) génèrent 83.392 accès, soit 25% du total. Six documents, dont 3 portant sur les techniques d'analyse des sols, totalisent en 2017 plus de 3.000 accès chacun.

En 2017, 6.242 documents concernant l'Hydrologie ont été consultés au moins une fois et le nombre total des accès est de 212.895²⁸. Les accès se répartissent dans 200 pays, dont 55 pays africains. Les documents ont donc été consultés en moyenne 34 fois pour toute l'année 2017. La France est le premier pays où les consultations sont localisées avec 34.274 accès, soit 16% du total. Pour les pays du continent africain, on comptabilise 105.605 accès, 50% du total (Europe : 26%, Amériques : 20%, Asie : 3,2%, Océanie : 0,8%). L'Algérie est le premier pays d'Afrique avec 20.409 accès, soit 9,6% du total. Les pays du Maghreb (Algérie, Maroc et Tunisie) génèrent 49.976 accès, soit 23% du total. Le Brésil est le premier pays des Amériques avec 7.970 accès, soit 3,7% du total. Trois documents

²⁶ Le tableau de bord complet est accessible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/TUN-2017.htm (Page consultée le 9 août 2018).

²⁷ Le tableau de bord complet est accessible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/Pedologie-2017.htm (Page consultée le 9 août 2018).

²⁸ Le tableau de bord complet est accessible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/acces/Hydrologie-2017.htm (Page consultée le 9 août 2018).

concernant les techniques d'analyse hydrologique et la construction de barrages, totalisent en 2017 plus de 3.000 accès chacun.

L'indice de Pareto est de 11.9% pour la Pédologie et de 12.1% pour l'Hydrologie, ce qui signifie qu'il faut prendre en considération environ 12% des documents les plus consultés pour atteindre 80% des consultations. Ces deux valeurs, tout à fait comparables, sont largement en dessous de 20%. Ainsi les consultations des documents se concentrent sur une part « relativement » faible des documents disponibles.

L'indice de spécialisation des consultations par rapport aux langues que nous avons mis au point compare les pourcentages du nombre des documents consultés par rapport au nombre total des consultations des documents en fonction de la langue. Les tableaux de bord de ces deux disciplines fournissent des informations intéressantes par rapport à l'analyse des langues des documents en fonction des fréquences des consultations.

La figure 10 ci-dessous, illustre les résultats obtenus par rapport à ces deux disciplines. Pour la Pédologie on observe que les documents en anglais (IS = 0,69) et en portugais (IS = 0,41) sont sous-consultés et que les documents en espagnol sont sur-consultés (IS = 1,63). Pour l'Hydrologie on observe que les documents en anglais (IS = 0,54) sont sous-consultés et que les documents en espagnol (IS = 1,86) et en portugais (IS = 2.66) sont sur-consultés.

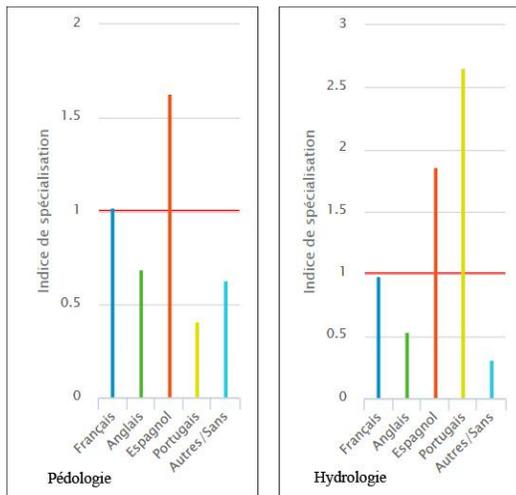


Figure 10 : *Indice de spécialisation de langues des documents de Pédologie et d'Hydrologie en fonction des accès*

7 Conclusions

Le projet de numérisation que nous avons lancé en 1996 a permis la numérisation et la mise en libre accès d'environ 66% de la production scientifique de l'IRD (plus de 66.500 documents en format pdf). Le projet se poursuit et environ 4.000 fichiers pdf par an s'ajoutent à l'archive ouverte.

L'analyse des statistiques de consultation de nos documents en libre accès montre l'impact des documents produits par l'IRD vis-à-vis des utilisateurs de l'internet, notamment ceux des pays en développement et des pays émergents. L'association, en mode relationnel, des données de consultation et des métadonnées décrivant les documents disponibles permettent de réaliser des tableaux de bord relativement élaborés et diversifiés : auteur, année de publication, langue du document, thématique des documents, pays à l'origine des consultations.

Dans ces tableaux de bord nous pouvons introduire des indicateurs originaux et spécifiques aux types de données dont nous disposons : « downloading factor », indice d'appropriation, indice de spécialisation en fonction de la langue, indice de spécialisation relatif à la thématique du document.

La combinaison de ces données avec les thèmes et les langues des documents montre des spécificités pour chaque pays (thématique) ou pour les aires linguistiques (pays francophones versus pays anglophones) du continent africain.

L'analyse montre que les documents les plus consultés sont souvent des documents publiés dans les années 90 sur les méthodologies, les techniques d'analyse et la construction de structures de génie civil (barrages).

La langue des documents apparait comme un facteur déterminant des consultations en fonction des pays : dans les pays francophone d'Afrique l'essentiel des consultations concernent des documents en français. Une vision stratégique des publications scientifiques vis-à-vis de ces pays devrait éventuellement prendre en compte ce facteur de la langue plutôt que de privilégier « systématiquement » des choix de publication en langue anglaise dans des revues à fort facteur d'impact.

Puisque l'on peut constater que les répartitions thématiques des consultations de chaque pays ne sont pas homogènes, le choix et les orientations des recherches pourraient en tenir compte pour privilégier des thématiques sur-consultées ou, à contrario, pour produire des publications scientifiques pouvant créer ou développer des intérêts plus originaux.

En tout état de cause, nous n'envisageons pas l'utilisation de ces résultats à des fins d'évaluation (des individus, des laboratoires, des institutions)²⁹ mais plutôt comme des éléments permettant d'élaborer des

²⁹ Une littérature très copieuse existe sur les questions de bibliométrie (voir : <https://fr.wikipedia.org/wiki/Bibliométrie>, page consultée le 9 août 2018), de scientométrie (voir : <https://fr.wikipedia.org/wiki/Scientométrie>, page consultée le 9 août 2018) et d'altmetrics (voir : <https://en.wikipedia.org/wiki/Altmetrics>, page consultée le 9 août 2018) et les usages qui sont faits en matière d'évaluation.

réflexions afin de comprendre l'utilisation de nos productions scientifiques par les internautes ainsi que l'impact qu'elles peuvent avoir, notamment dans les pays des Suds.

Bibliographie

- AGOSTI, Maristella ;CRIVELLARI, Franco, DI NUNZIO, Giorgio Maria (2012). Web log analysis: a review of a decade of studies about information acquisition, inspection and interpretation of user interaction. In *Data Mining and Knowledge Discovery*, vol. 24, n° 3, pp. 663-696. <https://doi.org/10.1007/s10618-011-0228-8>
- Cable.co.uk (2017). *Study of broadband pricing in 196 countries reveals vast global disparities in the cost of getting online*. [En ligne]. Disponible à : <https://www.cable.co.uk/about/media-centre/releases/new-worldwide-broadband-price-league-unveiled/> (Page consultée le 10 août 2018).
- FARGIER, Nathalie (2015) Numériser la littérature grise scientifique. In *I2D Information, données et documents*, vol. 52, n° 1, pp. 61-62. [En ligne]. Disponible à : <https://www.cairn.info/revue-i2d-information-donnees-et-documents-2015-1-page-61.htm> (Page consultée le 10 août 2018).
- Internet World Stats (2018). *Internet Users Statistics for Africa*. [En ligne]. Disponible à : <https://www.internetworldstats.com/stats1.htm> (Page consultée le 10 août 2018).
- ITU/UNESCO Broadband Commission for Sustainable Development (2017). *The state of broadband 2017: broadband catalyzing sustainable development*. ITU, Unesco, 2017, 104 p. [En ligne]. Disponible à : https://www.itu.int/dms_pub/itu-s/opb/pol/S-POL-BROADBAND.18-2017-PDF-E.pdf (Page consultée le 10 août 2018).
- JANSEN, Bernard J. (2006). Search log analysis: what it is, what's been done, how to do it. In *Library & Information Science Research*, vol. 28, n° 3, pp. 407-432. <http://dx.doi.org/10.1016/j.lisr.2006.06.005>
- KAUR, Navjot; AGGARWAL, Himanshu (2017) A Novel Semantically-Time-Referrer based Approach of Web Usage Mining for Improved Sessionization in Pre-Processing of Web Log. In *International journal of advanced computer science and applications*, vol. 8, n° 1, pp. 158-168. [En ligne]. Disponible à : http://thesai.org/Downloads/Volume8No1/Paper_22-A_Novel_Semantically_Time_Referrer_based_Approach.pdf (Page consultée le 10 août 2018).
- KRISHNAGANDHI, Geetha ;DHAS, Suresh Gnana (2016). Web log mining: a study. In *IIOAB Journal*, vol. 7, n° 9, pp. 6-15. [En ligne]. Disponible à : https://www.iioab.org/articles/IIOABJ_7.9_6-15.pdf (Page consultée le 13 août 2018).
- MARTINEZ-COMECHÉ, Juan-Antonio (2017). Determinación de grupos de usuarios de bibliotecas digitales mediante el análisis de ficheros log. In *Revista Española de Documentación Científica*, vol. 40, n° 3, pp. 1-19. [En ligne]. Disponible à : <http://dx.doi.org/10.3989/redc.2017.3.1420> (Page consultée le 13 août 2018).

- ORSTOM (1955) *Office de la Recherche Scientifique et Technique Outre-Mer : organisation - activités : 1944-1955*. Paris : ORSTOM, 1955, 182 p. [En ligne]. Disponible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers12-03/010027861.pdf (Page consultée le 10 août 2018).
- ROSSI, Pier Luigi (1992). Servers and online bibliographic databases in developing countries: the African reality. In: *Raitt D.I. (ed.). Online information 92*. Oxford : Learned Information, pp. 431-435. ISBN 0-904933-83-0. [En ligne]. Disponible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/pleins_textes_6/b_fdi_35-36/41308.pdf (Page consultée le 10 août 2018).
- ROSSI, Pier Luigi (1997) Economie et portabilité : une chaîne d'édition électronique destinée à la dissémination de l'information primaire. In : *Forum initiatives 97*. Hanoi : AUF, 6 p. multigr. [En ligne]. Disponible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/pleins_textes_6/divers1/010022348.pdf (Page consultée le 10 août 2018).
- ROSSI, Pier Luigi ; Ngoma-Mouaya Marcel (2000). "Pleins_Textes" : IRD (Institut de Recherche pour le Développement) electroniclibrary. In: *Online information 2000 proceedings*. Oxford : Learned Information, pp. 201-206. [En ligne]. Disponible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/pleins_textes_5/TAP/010024168.pdf (Page consultée le 10 août 2018).
- ROSSI, Pier Luigi ; THIAW, Anastasie (2012) Log analysis and textmining on internet access to dissertations of the INSEPS (Institut National Supérieur de l'Education Populaire et du Sport) Dakar, Sénégal. In *African Research and Documentation*, vol 118, pp. 79-90. [En ligne]. Disponible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers13-05/010058664.pdf (Page consultée le 10 août 2018).
- ROSSI, Pier Luigi ; TRAORE Minata ; MAÏGA DIALLO, Fatoumata (2018) Publications en libre accès des universités du Burkina Faso : analyse d'impact et visibilité internationale. In *027.7 Zeitschrift für Bibliothekskultur*, vol. 5, n° 1, pp. 52-64. [En ligne]. Disponible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers18-02/010072183.pdf (Page consultée le 10 août 2018).
- ROUX-FOUILLET, Jean-Paul (1988) Horizon : base bibliographique ORSTOM : présentation. In : *Séchet Patrick (ed.). Séminfor 1, premier séminaire informatique de l'ORSTOM : bases de données et systèmes d'information : quelles méthodes ?* Paris: ORSTOM, pp. 285-296. [En ligne]. Disponible à : http://horizon.documentation.ird.fr/exl-doc/pleins_textes/pleins_textes_4/colloques/26249.pdf (Page consultée le 10 août 2018).
- SHNEIDERMAN, Ben (1992). Tree visualization with tree-maps: 2-d space-filling approach. In *ACM Transactions on Graphics*, vol. 11, n° 1, pp. 92-99. doi.org/10.1145/102377.115768.

L'impact de la recherche en Tunisie entre webométrie et métriques alternatives

The impact of research in Tunisia between webometrics and altmetrics

Hanan Gmara

Doctorante chercheuse en science de l'information

Email : gmara.hanan@gmail.com

Tel : 00216 55355722

Résumé

Il s'agit dans la présente étude d'examiner l'impact de la recherche à travers les sites web des structures de recherche tunisiennes.

Nous explicitons dans un premier temps à travers les méthodes webométriques par le biais du collecteur de données SocSciBot des représentations cartographiques soulignant l'interaction hypertextuelle entre les sites web de notre corpus (universités et structures de recherche).

L'objectif étant de savoir ce que nous pouvons inférer de ces représentations hypertextuelles sur la logique des relations entre les sites des structures de recherche. Le but est de dresser un état descriptif des caractéristiques du réseau constitué. Nous poursuivons les analyses par le biais des métriques alternatives « altmetrics » pour examiner la présence de ces structures dans les réseaux sociaux ainsi que de mesurer de nouvelles formes d'attention et d'impact scientifique et social. Le recours complémentaire aux méthodes des métriques alternatives reliant en

particulier la visibilité dans les réseaux sociaux est justifié par l'usage de la communauté scientifique (chercheurs et institutions) de ces outils pour la collecte et la diffusion de l'information. Cette recherche tourne autour du libre accès, étant donné que l'information existante sur ces sites web et les outils utilisés (logiciels) le sont également en libre accès; les licences libres viennent compléter les règles pour faciliter ces accès à l'information sur internet.

Mots-clés. webométrie, métriques alternatives, analyse des liens, évaluation de la recherche, web social, recherche scientifique, libre accès.

Abstract

The objective of this review is to study the impact of web based search system of the Tunisian research structures.

In the first instance, we will make explicit through the webometrics techniques (or methods) in particular via the datalogger SocScibot cartographic representations highlighting the hypertextual interactions between the different websites of the Tunisian corpus (universities, researches centers/institutes). The objective is to determine the relational logic between the websites of the different organizations in order to produce a descriptive chart of the network characteristics. The analysis will be extended through the metric alternatives « altmetrics » to verify the presence of the said structures in the social networks and thus

measuring all new forms of attention, social and scientific impact. The complementary use of metric alternatives methods linking in particular visibility in social networks is justified by the scientific community (researches, institutes) as a major tool for collection and dissemination of information.

This study revolves around the free access since the available information and used tools (software) are also in free access. Free licensing come to supplement the rules of intellectual property to facilitate the access to information on line.

Keywords. webometrics, altmetrics, linkanalysis, evaluation of research, social web, scientific research, open access.

1 Introduction

Le web représente depuis une quinzaine d'années un média privilégié de communication et de diffusion de l'information. Les sites web se révèlent indispensables en tant qu'outils d'information et de valorisation de l'information. Etant donné que le web est considéré comme une base de données non contrôlée (Pinède et Vieira, 2008)¹, nous sommes intéressée à étudier cet objet de recherche à l'aide d'instruments de mesures qu'il est nécessaire de calibrer.

Le web est en train de transformer les modalités de la communication savante, de la dissémination des informations, de la collaboration et de l'échange scientifique. Auparavant, l'information scientifique était réservée à des érudits. De nos jours, le web diffuse mondialement ces informations sous des formes variées. Le web en tant que véhicule d'information est au cœur de la communication scientifique.

Le libre accès (open access) est la disponibilité en ligne de contenus numériques mis à disposition directement et gratuitement, sans contrainte juridique, financière ou autre. Au départ, il s'agissait simplement du partage de logiciels, puis cela s'est étendu à toutes sortes de contenus : scientifiques, philosophiques, économiques, administratifs, etc.

¹David Reymond, Nathalie Pinède-Wojciechowski, Lise Vieira. Intérêt stratégique d'un Observatoire des Sites web en sciences de l'information et de la communication. *16e congrès de la SFSIC*, 2008, Compiègne, France. 2008.

Cette recherche tourne autour du libre accès, étant donné que l'information existante sur ces sites web en libre accès et ce mouvement (le libre accès) souligne l'importance de l'extension de l'accès à des savoirs nombreux, disponibles pour tous ; les licences libres viennent compléter les règles de la propriété intellectuelle pour faciliter ces accès à l'information sur internet ainsi que la mesure et l'évaluation de cette dernière.

D'un autre côté, le monde de la recherche scientifique est un terrain pour expérimenter des méthodes d'évaluation par la webométrie². Cette dernière offre un ensemble de méthodes et de mesures permettant d'étudier la structure et le processus de communication académique sur le web. L'analyse des liens hypertextes (analysis links), la plus connue des approches webométriques, est devenue très sophistiquée, et l'arrivée des moteurs de recherche indexant ces liens a provoqué des avancées quantitatives et qualitatives dans d'autres méthodes webométriques.

Nous souhaitons présenter dans le cadre de cet article, les résultats préliminaires de notre recherche doctorale qui vise à étudier les sites web des structures de recherche tunisiennes. Ladite finalité s'accompagne d'une nécessité croissante pour ces institutions d'évaluer la quantité et la qualité des écrits scientifiques qu'elles diffusent sur le web, ainsi que l'évaluation de l'hypertextualité interne et externe supportée par le système de communication internet, en relation avec la très active littérature scientifique internationale (Thelwall, 2003) (Thelwall et Vaughan, 2004). Le recours complémentaire aux méthodes des métriques alternatives reliant en particulier les réseaux sociaux (Facebook) est justifié par l'usage de la communauté scientifique (chercheurs et institutions de recherche) de ces outils pour la collecte et la diffusion de l'information qui est présente en libre accès ainsi que les outils d'analyse webométrique et altmétrique en libre accès aussi.

2 Visibilité des sites web scientifiques en Tunisie : une recherche en cours

Avec l'informatique et internet sont apparues de nouvelles possibilités de créer, diffuser des savoirs ; conséquemment il est possible de mesurer et d'évaluer l'information scientifique sur le web par le recours à des méthodes adaptées à ce médium.

Le web, depuis sa création, est un web qui diffuse des informations académiques et scientifiques et cette tendance se maintient jusqu'à nos jours. Certes les usages ont évolué avec l'événement du web sémantique, mais chose certaine le web reste un des outils essentiel, voire même un médium essentiel pour communiquer sur le fonctionnement, les activités et publications des structures de recherche comme le soulignent les structures de recherche tel que le centre national de recherche scientifique français (CRNS) : « [...] Si la mission essentielle d'un laboratoire est de produire de la connaissance, il lui incombe également de la transmettre, et pas

²« ...l'étude des aspects quantitatifs de la construction et l'utilisation des ressources d'information, structures et technologies sur le Web en utilisant des méthodes bibliométriques et infométriques. »

Source : Boutin, Eric. La recherche d'information sur Internet au prisme de la théorie des facettes. Diss. Université du Sud Toulon Var, 2008.

uniquement selon les voies académiques en direction des experts de la discipline ou par le biais de l'enseignement universitaire. En effet, dans le monde d'aujourd'hui, la communication occupe une place de plus en plus importante. Elle est au centre de la vie de la société et la communauté scientifique ne peut pas y échapper »³.

Nous nous demandons si les structures de recherche en Tunisie ont recours à cet outil pour la visibilité du fruit de leurs réflexions et recherches ?

Le point de départ de notre problématique se base sur le constat des résultats des différents classements des universités à l'échelle internationale qui montrent que les universités tunisiennes sont hors classement, une image des universités plutôt négative. « *Le dernier classement de l'International colleges and universities (2016), spécialisé dans le classement des collèges et universités du monde, nous classe 70^e à l'échelle du continent africain, loin derrière l'Égypte (7^e), le Maroc (20^e), l'Algérie (29^e)* »⁴.

Le rapport annuel de la société « Times Higher Education's Emerging Economies », publié en 2016, classe entre autres les meilleures universités du monde arabe. Ce classement se base sur le niveau de l'enseignement, la recherche scientifique et les mentions des recherches, les universités tunisiennes y sont absentes. Une situation qui reflète, selon cette source, la réalité de l'enseignement supérieur et la recherche scientifique en Tunisie⁵.

La Tunisie n'est pas non plus présente dans le classement de Shanghai 2014-2015-2016-2017)⁶. En fonction de ce dernier classement, nous pouvons noter que l'université tunisienne est en perte de vitesse ces dernières années.

Ce résultat est lié peut être au problème de visibilité des écrits des chercheurs tunisiens qui sont souvent affiliés à d'autres laboratoires à l'étranger et qui, quand ils publient, préfèrent le faire sous le nom du laboratoire (par exemple le laboratoire de biologie et pathologie végétales) et non leur affiliation à leurs universités. Le classement des universités se base sur un ensemble de données telles que : les indicateurs de qualité des sites web des universités, qui tiennent compte de leur dimension, de leur visibilité dans les moteurs de recherche et de la richesse de contenus ainsi que des citations reçues par rapport aux publications des structures de recherche. Le classement des universités à l'échelle mondiale nous amène en tant que chercheuse dans le domaine des sciences de l'information à nous interroger sur la visibilité de la recherche scientifique. Cette étude permettra d'étudier, de connaître et d'évaluer le degré de rayonnement de la recherche scientifique tunisienne à travers le web⁷.

³Site de l'université Clermont Auvergne. [En ligne] : <http://www.lpc-clermont.in2p3.fr/spip.php?rubrique31> (visité le : 08/07/2018)

⁴Journal électronique « leaders », [En ligne] : <http://www.leaders.com.tn/article/13280-l-enseignement-superieur-les-technologies-de-l-information-desormais-a-l-honneur> (visité le : 21-09-2017)

⁵Journal électronique « webdo », [En ligne] : <http://www.webdo.tn/2016/06/24/aucune-universite-tunisienne-dans-le-classement-des-meilleures-universites-arabes/> (visité le : 20-09-2017)

⁶ Site de classement de Shanghai [En ligne] : <http://www.shanghairanking.com/fr/> (visité le : 20-09-2017)

⁷D'après le « webometrics ranking of world universities » est une initiative du laboratoire cybermetrics, un groupe de recherche appartenant au conseil supérieur de investigations científicas (CSIC), la plus grande institution publique de recherche en Espagne, avec 126 centres et instituts dans tout le pays, rattaché au ministère de l'éducation. Le laboratoire cybermetrics a pour objectif

L'intensification de la compétition internationale pour l'excellence scientifique s'accompagne d'un foisonnement des instruments de mesure destinés à évaluer systématiquement les universités et les organismes de recherche pour tenter d'améliorer leurs performances. Si ce mouvement fait l'objet d'un véritable engouement de la part des gouvernements et des agences publiques dans plusieurs pays (Auranen et Nieminen, 2010) comme la Tunisie, il suscite également de nombreuses initiatives pour étudier la situation de la production scientifique ainsi que la mesure de la visibilité de cette dernière.

Un observatoire des sites web universitaires tunisiens à été mené en 2009 indique qu' :*« il existe bel et bien une fracture scientifique traduite par les études webométriques effectuées mais qui ne signifie pas pour autant une absence, sur le plan réel, de collaboration effective (vu que des relations existent réellement) mais dénoterait plutôt d'un grave problème de visibilité et de conscience collective quant au rôle d'un conséquent de cette visibilité »*⁸. L'étude que nous souhaitons réaliser essaye d'examiner si cette fracture existe encore ou bien elle est réduite ou encore elle s'est développé après presque 7 ans vu que la culture numérique est devenue dans nos jours dans le monde entier aussi bien à l'échelle nationale qu'internationale ancrée voir même cruciale dans le domaine scientifique. Ainsi que l'infrastructure technologique en Tunisie est en constante amélioration.

L'objectif principal de notre étude est d'examiner le degré de visibilité de l'information scientifique tunisienne à travers les sites des structures de recherche. Plus précisément d'examiner la proportion de l'information scientifique tunisienne présentée sur le web à travers l'étude des sites proprement dits et la présence de ces structures dans les réseaux sociaux à travers des outils en libre accès.

3 La méthodologie

Notre démarche méthodologique s'appuie sur une collecte de données issues de diverses sources du web qui nous permettront de décrire l'état actuel du réseau, et son inscription dans le web en général. Nous procédons par une analyse webométrique de la zone DNS⁹ des sites web qui a pris la forme du

d'analyse quantitative de l'internet, il a conçu des indicateurs qui visent à mesurer l'activité scientifique sur le web. Le but original du classement était de favoriser la publication sur le web, pas de classer les établissements. Mais ces indicateurs donnent des renseignements utiles, car ils ne sont pas fondés sur le nombre de visites ou la conception des pages, mais sur la visibilité globale des universités sur le réseau. » Source : Hervé Théry. *Palmarès des Universités mondiales, "Shanghai" et les autres*. M@ppemonde, Maison de la géographie, 2009. [En ligne] : <https://halshs.archives-ouvertes.fr/halshs-00490220/document> (visité le : 13-01-2018)

8 S. MAHMOUD. « Observatoire des sites web universitaire en Tunisie : fracture scientifique ou collaboration », 2009.

⁹ « Le serveur DNS (Domain Name System) est un service dont la principale fonction est de traduire un nom de domaine en adresse IP. ... le serveur DNS agit comme un annuaire que consulte un ordinateur au moment d'accéder à un autre ordinateur via un réseau. Autrement dit, le serveur DNS est ce service qui permet d'associer à site web (ou un ordinateur connecté ou un serveur) une adresse IP ... Conçu en 1983 par Jon Postel et Paul Mockapetris. Chaque fournisseur d'accès à Internet dispose notamment de ses propres serveurs DNS, avec des adresses IP qui prennent souvent la forme d'une succession de nombres de chiffres (194.158.122.10 par exemple). » Source : Journal du net [en ligne] : <https://www.journaldunet.fr/web-tech/dictionnaire-du-webmastering/1203373-dns-serveur-dns-domain-name-system-definition-traduction/> (Visité le : 27/07/2018)

dénombrément de ses sous domaines, des pages web et de son réseau d'hyperliens. Le réseau des relations hypertextes internes entre les différents espaces du site laisse apparaître le caractère mécanique de la liaison entre les sous domaines.

Enfin, la dispersion des liens vers les zones DNS partenaires en montrera la diversité.

Cette approche permettra de révéler le mode éditorial spécifique de chacun des sites.

Nous explicitons à travers les méthodes webométriques par le biais du collecteur de données « SocScibot¹⁰ » qui est un logiciel d'indexation web en libre accès pour l'analyse de liens sur un ou plusieurs sites pour produire des statistiques sur le(s) réseau(x) de l'interconnexion.

Il est à noter que « SocScibot » est notre outil collecteur pour ce travail, ce choix est principalement motivé par la stabilité de cet outil. L'usage de « SocScibot » est associé à un autre logiciel « Pajek » qui permet les représentations cartographiques des liens hypertextuels, cet outil est en libre accès pour la communauté scientifique.

Pajek (l'équivalent slovène du mot araignée) est un programme, pour l'analyse de larges réseaux dans divers domaines : généalogie, réseaux informatiques, réseaux de transport, réseaux sociaux, réseaux inter/intra organisationnels... il convient de préciser à ce stade que Pajek est un logiciel libre, ce qui est en fait un atout non négligeable. Il offre fondamentalement à l'utilisateur¹¹ :

- la liberté d'étudier comment le logiciel fonctionne et de l'adapter à ses besoins ;
- la liberté de diffuser les améliorations apportées au logiciel.

Il est encore pertinent d'attirer l'attention sur l'« utilité participative »¹² des logiciels libres d'autant qu'il est presque impossible d'y introduire des fonctionnalités cachées ; ce qui garantit la libre circulation des informations et la liberté d'expression de chacun quel que soit son équipement¹³. L'objectif est de savoir ce que nous pouvons inférer de ces représentations hypertextuelles sur la logique des relations entre les sites des structures de recherche. Le but est de dresser un état descriptif des caractéristiques du réseau constitué.

¹⁰SocSibot est un collecteur de données conçu pour la recherche. A l'aide de programmes de soutien tels que les outils SocScibot et Cyclist, il peut être utilisé pour conduire des analyses de liens sur un site ou un ensemble de sites ou encore de diriger un moteur de recherche sur un ensemble de sites. Ces programmes peuvent aussi être utilisés dans l'enseignement et la recherche pour illustrer le fonctionnement des analyses des liens et des moteurs de recherches. Le résultat fourni par SocSibot n'est pas le nombre total des pages à partir du site, c'est plutôt le nombre de pages que ce dernier peut trouver en suivant les liens dans le site. Certains liens peuvent ne pas être liés au site ou sont liés à des pages externes au site, ou encore revêtent des formes que SocSibot ne peut pas repérer (Javascript par exemple).

¹¹ G. Datondji, Le logiciel libre : un modèle économique viable pour un développement durable des TIC en Afrique, in *Revue Maghrébine de Documentation et d'information*, n°13-14-15.

¹²C'est la contribution d'un utilisateur ou d'un développeur à l'amélioration de la qualité et de la performance d'un logiciel.

¹³ G. Datondji, Op. cit.

Nous poursuivons les analyses par le biais des métriques alternatives « altmetrics » qui a été utilisé pour la première fois dans un tweet de Priem de septembre 2010 : « *I like the term #articlelevelmetrics* »¹⁴ (il faut préciser que les « article levelmetrics » et ce que nous appelons aujourd'hui les « altmetrics » ne réfèrent pas tout à fait à la même chose, les « altmetrics » englobant les « article levelmetrics », mais aussi une panoplie d'« objets » de recherche qui peuvent être des agents ou des documents, ainsi que nous le verrons plus loin.), « *but it fails to imply *diversity* of measures. Lately, I'm liking #altmetrics.* » (Priem, 2010). Le mois suivant¹⁵ est paru un manifeste en faveur de l'utilisation de ces métriques prétendument « alternatives » (Priem et al., 2011), dont les auteurs expliquent que, dans le contexte de l'augmentation incessante du nombre de publications scientifiques, les chercheurs ont besoin de « filtres » pour les aider à sélectionner les travaux les plus pertinents. Selon eux, les « filtres » existants, à savoir l'évaluation par les pairs, le nombre de citations et le facteur d'impact, par exemple, sont imparfaits et ne montrent qu'une infime partie de l'impact réel de la recherche : « *Unlike citation metrics, altmetrics will track impact outside the academy, impact of influential but uncited work, and impact from sources that aren't peer-reviewed* » (Priem et al., 2011).

Le programme utilisé « nibbler » (qui est un outil en libre accès pour étudier l'accessibilité, le référencement ainsi que la présence des sites web dans les médias sociaux) pour examiner la présence de ces sites dans les réseaux sociaux ainsi que de mesurer de nouvelles formes d'attention et d'impact scientifique et social. Le recours complémentaire aux méthodes des métriques alternatives reliant en particulier la visibilité dans les réseaux sociaux est justifié par l'usage de la communauté scientifique (chercheurs et institutions) de ces outils pour la collecte et la diffusion de l'information, d'autant plus que l'étude webométrique ne montre pas l'utilisation de l'information scientifique par les non-auteurs. Les statistiques d'utilisation de l'information scientifique issues des métriques alternatives peuvent donc aider à mieux comprendre les modes de ce type d'information et sont complémentaires aux indicateurs bibliométriques classiques.

4 Analyse webométrique : résultats de SocSibot

Notre démarche s'appuie sur une collecte de données issues de diverses sources du web pour la constitution du corpus (sites web : Ministère de l'enseignement supérieur/ centre de calcul EL-Khawarizmi), les sites des universités et les sites des structures de recherche. L'ensemble des sites collectés après traitement, s'étalant sur la période allant du 22 avril 2018 au 23 Juin 2018.

Le logiciel utilisé nous permet de recueillir les graphes des éditions et de les disperser selon un format compatible avec le logiciel Pajek spécialisé dans la visualisation et le traitement exploratoire des graphes. L'outil de

¹⁴Haustein, Bowman et Costas, 2016, p. 376)

¹⁵Nous renvoyons à la seconde version du manifeste parue en septembre 2011, dans laquelle le tiret a été retiré de l'orthographe initiale *alt-metrics*.

¹⁶Priem, J., Taraborelli, D., Groth, P. et Neylon, C. « Altmetrics: A manifesto ». 2011 [En ligne] <http://altmetrics.org/manifesto/> (visité le : 02/04/2018)

visualisation et de traitement des graphes illustrant le distinguo relationnel des différents sites. Cette approche permettra de révéler le mode éditorial spécifique de ce site. Plus précisément le collecteur SocSibot, développé par Mike Thelwall, extrait les contenus de sites ou de groupes de sites ainsi que les différents réseaux à divers niveaux d'agrégation du Advanced Model Document du même auteur puis d'en produire un format utilisable par un outil de traitement des graphes. De nombreux travaux de webométrie utilisent des données collectées via des index du web (Aguillo, 2006) et l'usage de l'un de ces index en particulier est biaisé (Aguillo, 2008). Néanmoins l'outil SocSciBot permet la production de réseaux d'hyperliens selon quatre échelles (Thelwall, 2002 et 2004) pris par la distinction des sites au sens DNS (nom de domaine, sites du sous-domaine, répertoires et pages) et l'agrégation des liens de chacun des niveaux permettant de discuter des éléments différenciés du réseau hypertextuel. Cet échelonnage des réseaux est particulièrement adapté aux sites des instances de recherche tunisiennes que nous allons le découvrir dans la partie empirique.

4.1 Réseau universitaire

Le logiciel webométrique SocSibot nous a produit un ensemble de graphiques que nous avons pu analyser et traiter par l'outil Pajek.

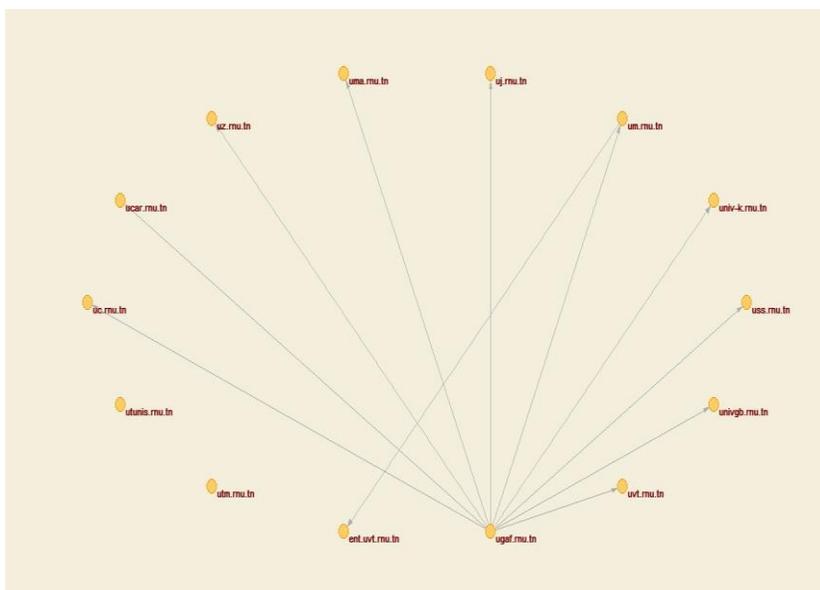


Figure 1 : Réseau hypertextuel entre sites d'universités

Le réseau hypertextuel des sites des universités, selon la théorie des graphes (un graphe est une représentation mathématique de réseaux formés de nœuds inter liés par des chemins¹⁷). Les nœuds dans notre cas c'est les

¹⁷ Bruno Courcelle. Introduction à la théorie des graphes : définitions, applications et techniques de preuves.

DNS des sites de notre corpus. Dans un graphe dirigé, les chemins représentent des relations directionnelles entre les nœuds¹⁸.

L'observation de la figure 1 reflète un paysage plutôt désertique. Une absence totale de liaison entre les sites des universités est constatée sauf pour l'université de Gafsa, qui semble être le facteur commun à toutes ces institutions universitaires, ce qui est assez intrigant et nous laisse perplexes quant au rôle joué par cette université.

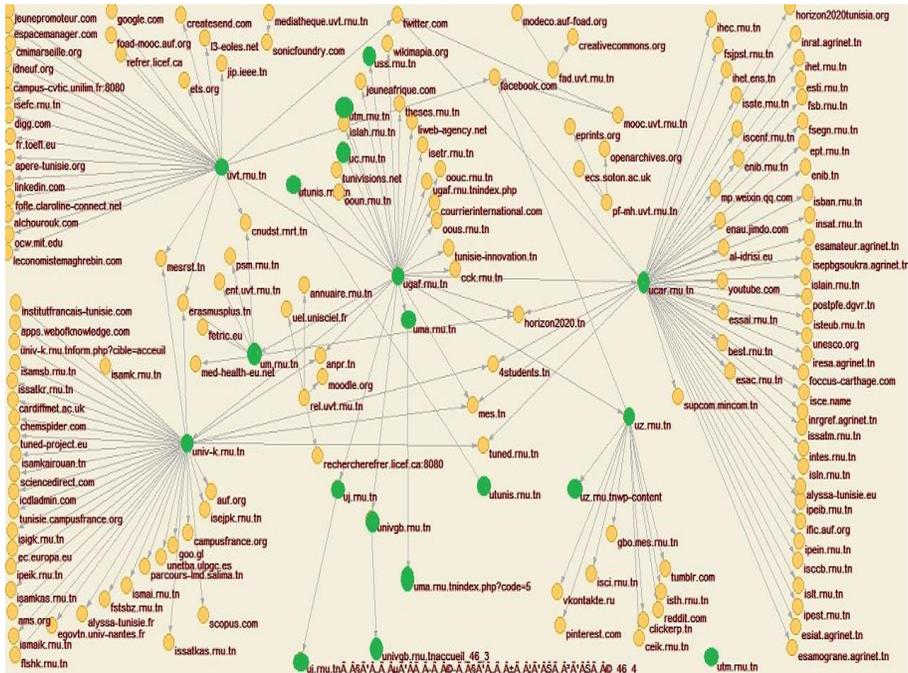


Figure 2 :Réseau relationnel extérieur des universités

Légende :

● : Sites des universités ● : liens externes

La figure ci-dessus présente les universités mentionnées auparavant et expose les liens vers lesquels elles pointent. Il est à noter que seuls les sites des universités de Carthage, de Gafsa, de Kairouan, de Zitouna ou encore l'université virtuelle (UVT) renvoient soit vers des liens qui semblent à priori, n'intéresser qu'elles ; soit des liens qui pointent avec d'autres

¹⁸ Bruno Courcelle, Op. Cit.

universités en même temps sur par exemple le site www.horizon2020.tn¹⁹ qui est lié aux universités de Carthage, de Gafsa et de Monastir. Par ailleurs le site www.erasmusplus.tn²⁰ est lié aux universités de Monastir, de Kairouan et l'UVT. Les deux sites précédemment cités sont relatifs aux programmes de recherche à l'échelle nationale, qui financent des actions et des projets liés à la recherche scientifique

Pour rendre l'analyse plus significative, nous avons procédé à l'intégration du site de Ministère de tutelle dans les réseaux relationnels ce qui a permis de construire des graphiques assez significatifs.

4.2 Le positionnement du site de Ministère

La première figure positionne le site du Ministère au centre du réseau et pointe vers de nombreux sites tels que les universités, les facultés et autres. L'observation du réseau mettant en relation les différentes institutions avec le Ministère (mes.tn). Il nous semble que le Ministère est l'élément fondamental de l'ensemble du réseau. Quelques liens concentrés dans le centre mettent en relation quelques facultés mais d'une manière peu significative.

¹⁹ « Horizon 2020 est le plus grand programme de recherche et d'innovation jamais réalisé par l'Union européenne (UE). En favorisant la transition des grandes idées, du laboratoire au marché, il conduira à des avancées révolutionnaires, des découvertes et des premières mondiales. Outre l'intérêt que lui portent les investisseurs des secteurs public et privé, il bénéficie d'un financement de près de 80 milliards d'euros sur 7 ans (de 2014 à 2020). Désormais pays associé, La Tunisie, adhérente à ce programme cadre, bénéficie des mêmes avantages que les pays d'Europe ». Source : Portail Tunisien Horizon 2020 [En ligne] <http://horizon2020.tn/> (visité le : 16-06-2018)

²⁰ « Le nouveau programme Erasmus+ vise à soutenir des actions dans les domaines de l'enseignement, de la formation, de la jeunesse et du sport pour la période 2014-2020. Erasmus+ regroupe et remplace sept programmes: le programme d'action dans le domaine de l'éducation et de la formation tout au long de la vie (Erasmus, Leonardo da Vinci, Comenius, Grundtvig et Jean Monnet), le programme Jeunesse en action, cinq programmes de coopération internationale (Erasmus Mundus, Tempus, Alfa, Edulink, le programme pour la coopération avec les pays industrialisés) et la nouvelle action pour le sport. » Source : Bureau National Erasmus+ Tunis [En ligne] <https://www.erasmusplus.tn/fr/article/0060-a-propos> (visité le : 17-06-2018)

Sfax (CBS) et celle de l'institut des régions arides de Médenine (IRA) et l'institut national de la recherche agronome de Tunis (INRAT).

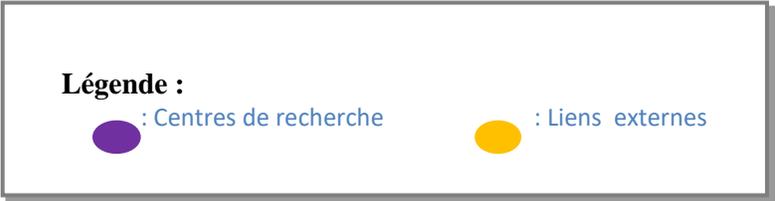
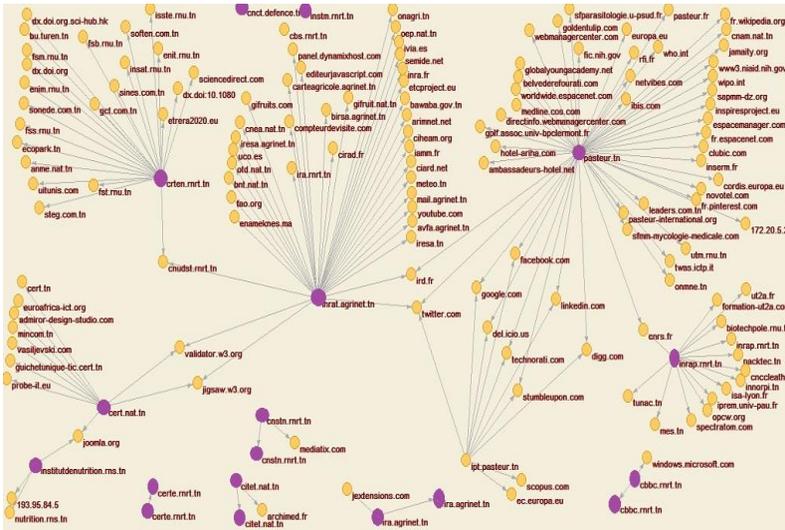


Figure 6 : Maillage relationnel des centres de recherche

Le maillage relationnel des sites des centres de recherche, représentés dans la figure no 6 par les cercles mauves (voire en annexe les noms des institutions ainsi que leurs URLs), laisse apparaître des liens éparpillés autour de chaque site des centres de recherche en question. L'institut pasteur et (www.pasteur.tn) et l'institut national de la recherche agronomique de Tunis (www.inrat.agrinet.tn) possèdent le plus de liens sortants vers des sites de tous genres (nationaux et internationaux) dont des laboratoires et des centres de recherche, des éditeurs et autres comme la poste tunisienne, le centre nationale universitaire de documentation scientifique et technique (CNUDST), Yahoo, Wikipédia, Facebook...

Un constat remarquable ressort de la figure no 6 : il n'y a pas de liens directs entre les centres représentés.

4.3.2 Réseau des technopoles

Nous constatons dans la figure no 7 une absence totale de relations entre leurs sites. C'est un paysage vide en effet, les technopoles semblent n'avoir aucun lien entre eux.



Figure 7 : Entre technopoles

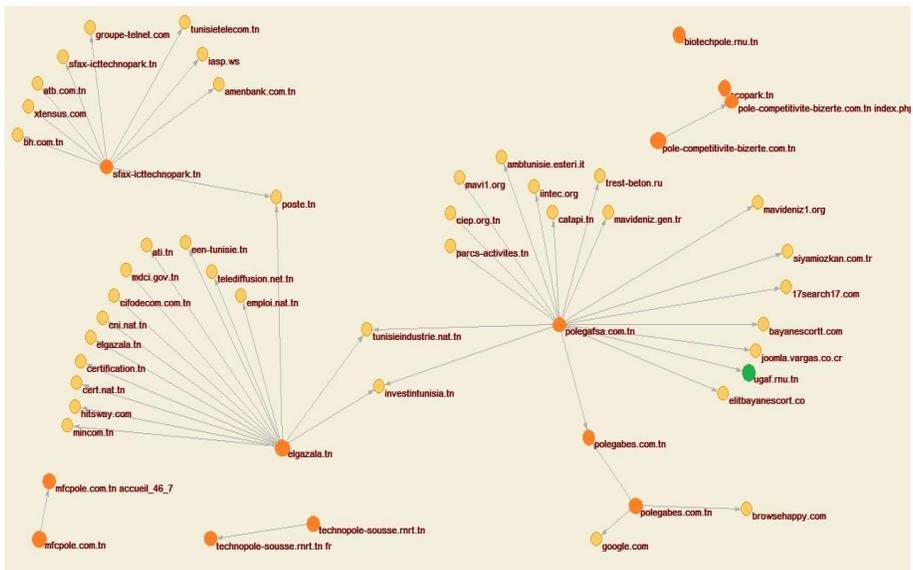


Figure 8 : Technopoles (Outlinks)

Nous constatons dans cette figure 8 que, parmi les neuf technopoles, trois possèdent un réseau assez important. La figure laisse apparaître une structure en faisceaux. Les trois technopoles, avec les cercles oranges font référence à différents sites, le pôle de compétitivité de Gafsa « PCG » est en relation avec l’université de Gafsa. Plus encore les technopoles, les centres et les laboratoires/unités de recherche relevant d’un même axe de recherche n’ont pas automatiquement de liens directs.

4.3.3 Réseau des laboratoires et unités

Nous nous intéressons dans ce qui suit aux liaisons éventuelles entre les divers laboratoires et unités de recherche relevant de différentes universités.

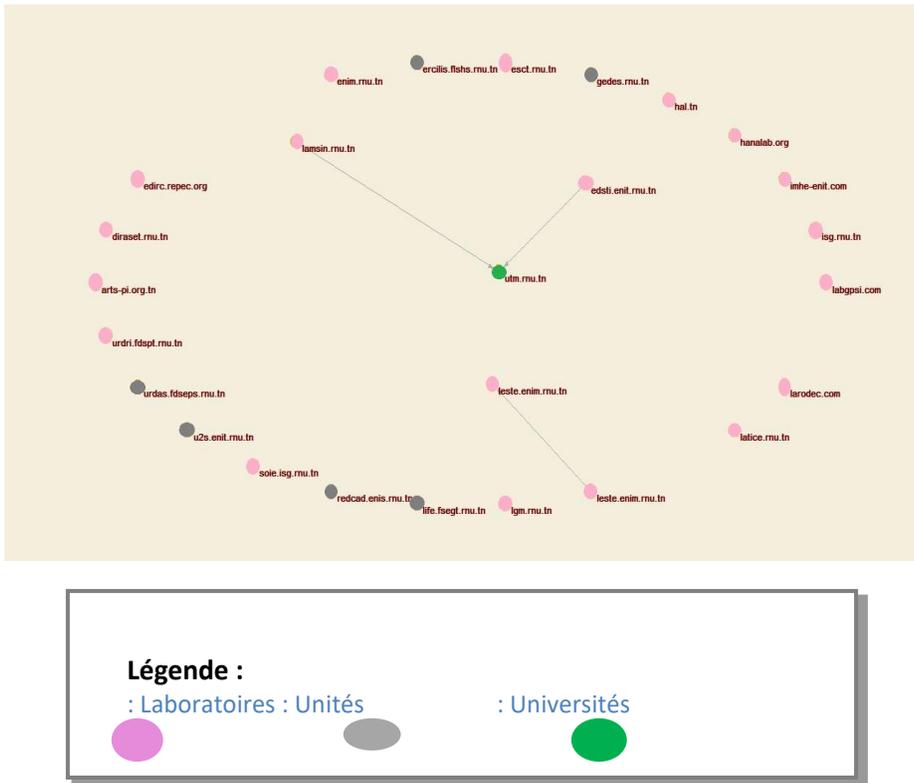


Figure 9 : Laboratoires et unités de recherche

La figure no 9 dénote une faible liaison entre les laboratoires et les unités de recherche sauf pour le laboratoire de modélisation mathématique et numérique dans la science de l’ingénieur (LAMSIN) le laboratoire de modélisation en hydraulique et environnement (LMHE) qui lui-même en relation avec l’université de Monastir et non pas entre eux.

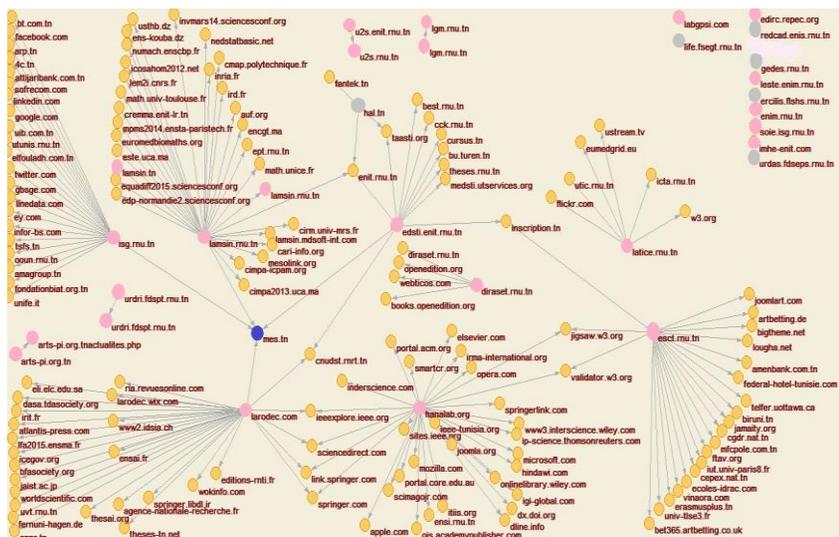


Figure 10 : Réseau des laboratoires et unités de recherche (Outlinks)

Le maillage relationnel des sites d'un ensemble de laboratoires représentés dans la figure no 10 par des cercles roses et les unités représentés par des cercles gris, laisse apparaître des structures éclatées autour des sites des laboratoires en question. La figure laisse apparaître une structure en faisceaux émanant des laboratoires vers divers sites

Les laboratoires qui possèdent le plus de liens sortants vers des sites de tous genres sont : Lamsin, Hanalab, Larodec, Edsti, Esct et Lattice (voir annexe 1).

Nous constatons aussi qu'une partie des laboratoires n'ont pas de liens sortants vers d'autres structures comme : Leste, Ercilis, Soie, Imhe, Edirc, Labgpsi, Img, U2s (voir annexe 1).

Quant aux unités de recherche, la figure no 10 démontre que l'unité Hal a trois liens sortants et que les autres unités représentées dans la figure n'ont aucun lien sortant.

5 Données altmetrics des structures de recherche

L'analyse croissante des réseaux sociaux scientifiques montre l'existence d'une demande de visibilité et intérêt accru pour le dépôt et la publication des travaux en libre accès. Les chercheurs tunisiens trouvent-ils l'information présentée par les structures de recherche sur les réseaux sociaux intéressant ? Certaines raisons pourraient être facilement levées.

L'analyse des réseaux sociaux est décrite comme une nouvelle méthodologie d'étude des communautés et des institutions en ligne. Afin d'analyser en profondeur la nature de leur présence sur le Web et de mieux comprendre leurs stratégies communicationnelles, nous avons constitué notre corpus comme suit, en tableau 1.

L'observation du tableau nous montre qu'une partie des structures sont présentes sur les réseaux sociaux particulièrement sur Facebook avec un d'intérêt modéré. Nous avons constaté qu'un nombre restreint de structures associe leurs sites à des pages sur le réseau social Facebook à l'instar de : l'institut national de recherche et d'analyse physico-chimique (INRAP), le centre de recherche et des technologies des eaux (CERTE), le technopole de Sousse : mécanique, électronique et informatique, pôle Elgazala des technologies de la communication, etc. D'autres structures n'ont pas associé leurs sites web aux réseaux sociaux, c'est-à-dire elles se sont contentées de faire des pages sur Facebook et/ou indépendantes (ne sont pas en relation avec les sites web qui en dérivent) telles que l'Institut Pasteur, le pôle de compétitivité de Bizerte, le pôle industriel et technologique de Gabès, etc. Il y a aussi d'autres structures ne sont présente sous aucune forme dans les réseaux sociaux comme : le centre de biotechnologie de Sfax (CBS), l'institut des régions arides de Médenine (IRA), l'équipe de recherche en civilisation et littérature de Sfax (ERCILIS), etc.

Nous estimons que ces pratiques ont une influence négative sur la visibilité et le référencement des sites web étudiés.

En examinant les pages et les groupes relatives aux structures de recherche sur le réseau social Facebook, nous constatons que peu d'entre elles sont présentes d'une manière dynamique, c'est-à-dire un nombre restreints de structures se servent de leurs pages/groupes dans les réseaux sociaux pour communiquer sur leurs productions scientifiques, pour la mesure de d'attention, l'audience et l'usage.

6 Conclusion

La disponibilité de l'information sur le web d'une manière générale et l'usage des outils webométrique et des métriques alternatives en libre accès ont permis de réaliser ce travail et qui montre que le maillage relationnel entre les différents sites des structures de recherche est relativement pauvre.

L'analyse des liens dénote une liaison dans la majorité des cas assez faible entre les structures étudiées. Il est à noter aussi que pour une partie des sites étudiés ont plus de liens sortants (outlinks) vers d'autres liens de tous genres (nationaux et internationaux) en relation avec le domaine de la recherche scientifique ou autres.

L'analyse à travers les métriques alternatives nous révèle qu'une partie des structures sont présentes sur les réseaux sociaux (particulièrement Facebook) avec un niveau d'intérêt social modéré. Seules quelques structures associent directement leurs sites à des pages sur Facebook.

D'autres structures se contentent de faire des pages ou des groupes indépendamment de leurs sites web, ce qui reflète quelque part l'influence négative de leur visibilité sur le web d'une manière générale.

En application concrète de la webométrie existe, entre autres, le classement international des sites web des universités ou encore le classement international des sites web des hôpitaux. L'idée sous-jacente est de motiver et les structures de recherche et les chercheurs à être présents sur le Web de manière à refléter leurs activités. Si la performance sur le Web d'une institution est en dessous des espérances présumées du fait de sa performance académique, les autorités concernées devraient revoir leur politique et promouvoir l'amélioration des publications électroniques quantitativement et qualitativement.

Nous pensons, au terme de ce travail, avoir donné une esquisse significative de la situation qui prévaut au sein du réseau des structures de recherche tunisiennes.

Nous pouvons présenter un ensemble de recommandations qui aident les structures de recherche à améliorer leur visibilité à travers web.

Etant donné que les signaux provenant des réseaux sociaux ont de l'impact sur l'amélioration du classement dans les moteurs de recherche les structures de recherche peuvent associer les sites web aux pages dans les réseaux sociaux.

Avoir des sites web pour chaque structures (laboratoires, unités, centres de recherche) ou bien au moins avoir une rubrique à part entière.

- De préférence aussi avoir une sorte de « marketing » de contenu ou de publication pour la mise à jour continue de site web.
- Développer une stratégie à long terme pour le référencement et l'évolution de la visibilité de site web.

Parmi les perspectives de ce travail est de faire une étude exploratoire à travers des entrevues auprès des responsables des structures de recherche pour compléter et interpréter les données recueillies à travers les analyses webométriques et les métriques alternatives utilisés dans cet essai. L'objectif est d'avoir une idée sur :

- l'usage du numérique et du web par les spécialistes de la recherche scientifique (institutions, chercheurs) ;
- l'existence d'une politique de diffusion des écrits scientifiques sur le web ;
- Le budget alloué à la recherche scientifique répond-t-il aux coûts des projets de recherche ? ;
- L'idée que les chercheurs préfèrent ou pas publier dans des revues prestigieuses pour avoir un facteur d'impact élevé.

Tableau 1 : Structures de recherche dans les réseaux sociaux

Sites web structures de recherche	Nombre d'abonnés dans les réseaux sociaux	Nbre de j'aime	Niv. d'intérêt social	Liens entrants	Liens sortants
Centres de recherche					
http://www.certe.rnrt.tn	Page associées au site web sur Facebook (2512 abonnés)	53 likes (Facebook)	Modéré	11104 liens + 59 domaines	4 liens
http://www.cbtc.rnrt.tn	Pas d'association aux réseaux sociaux. Page indépendante sur Facebook CBBCTechnopoleBorjCedria : https://www.facebook.com/CBBCTechnopoleBorjCedria/ (1259 abonnés)	--	--	--	--
http://www.crtcn.rnrt.tn (site web ne fonctionne pas)	Pas d'association aux réseaux sociaux.	--	--	--	--
http://www.cnstn.rnrt.tn	Pas d'association aux réseaux sociaux. Page indépendante sur Facebook : https://www.facebook.com/Cnstn-المركز-الوطني-للعلم-والتكنولوجيا-والتكنولوجيا-693234947372481/?ref=br (3776 abonnées)	--	--	--	--
http://www.inrap.rnrt.tn	Page associées au site web sur Facebook. Page indépendante : https://www.facebook.com/INRAPtn-115556305152816/ (1917 abonnées)	1876 likes sur Facebook	Modéré	--	--
http://www.instm.rnrt.tn	Pas d'association aux réseaux sociaux (un groupe sur Facebook avec 711 membres : https://www.facebook.com/groups/30578057519/about/)	--	--	--	--
http://www.cbs.rnrt.tn	Pas d'association du site web aux réseaux sociaux.	--	--	--	--
http://www.ira.rnrt.tn	Pas d'association du site web aux réseaux sociaux.	--	--	--	--
http://www.institutdenutrition.rns.tn	Pas d'association du site web aux réseaux sociaux. Page indépendante sur Facebook : https://www.facebook.com/Institut-national-de-	--	--	--	--

	nutrition-394749010727080/ (1544 abonnées)				
http://www.pasteur.tn	Pas d'association du site web aux réseaux sociaux. Page indépendante sur Facebook : https://www.facebook.com/InstitutPasteurDeTunis/ (8545 abonnées)	--	--	--	--
http://www.citet.nat.tn	Pas d'association du site web aux réseaux sociaux. Page indépendante sur Facebook : https://www.facebook.com/InstitutPasteurDeTunis/ (5451 abonnées)	--	--	--	--
http://www.cert.nat.tn	Pas d'association du site web aux réseaux sociaux. Page indépendante sur Facebook : https://www.facebook.com/CERT.PageOfficielle/ (3537 abonnées)	--	--	--	--
http://www.cnet.nat.tn (site web ne fonctionne pas)	Pas d'association du site web aux réseaux sociaux.	--	--	--	--
Technopoles					
http://www.ecopark.tn	Pas d'association du site web aux réseaux sociaux. Page indépendante sur Facebook : https://www.facebook.com/ecopark.tn/?ref=br_rs (211 abonnées)	--	Modéré	1088 pages sur 14 domaines accèdent	--
http://www.polegafsa.com.tn	Page associée au site web + une page indépendante sur Facebook : https://www.facebook.com/polegafsa/ (2005 abonnées)	266 likes	Modéré	185 liens	11 liens
http://www.sfax-icttechnopark.tn	Page associée au site web + une page indépendante sur Facebook : https://www.facebook.com/Technopole.de.Sfax/?hcr ef=ARR7qhoktc0B_a317_AHQMUqoS5Tdsnt_53HNk9zTfhk1JY_BjHWT89oyPHJ1Koo5AY&fref=nf (6075 abonnées)	159 likes	Modéré	10258 pages +32 liens	39 liens
http://www.technopole-sousse.rnrt.tn	Page associée au site web : https://www.facebook.com/CRMNTunisie/	8 likes	Faible	61301 pages sur 141 domaines	141 liens

http://www.elgazala.tn	Page associée au site web : https://www.facebook.com/Elgazala-Technopark-391558404369877/ (2936 abonnées)	2826 likes	Modéré	12206 pages sur 90 domaines	--
http://www.pole-competitivite-bizerte.com.tn	Pas d'association du site web aux réseaux sociaux. Page indépendante sur Facebook : https://www.facebook.com/polebizerte/ (1795 abonnées)	--	--	--	--
http://www.biotechpole.rmu.tn (site web en cours de construction)	Pas d'association du site web aux réseaux sociaux. Page indépendante sur Facebook : www.facebook.com/Biotechpole-Sidi-Thabet-officiel-151886198213669/ (2791 abonnées)	--	--	--	--
http://www.polegabes.com.tn	Pas d'association du site web aux réseaux sociaux.	--	--	--	--
http://www.mfcpole.com.tn	Pas d'association du site web aux réseaux sociaux. Page indépendante sur Facebook : https://www.facebook.com/mfcpole/ (2277 abonnées)	--	Modéré	--	--
Laboratoire					
http://www.leste.enim.rmu.tn	Pas d'association du site web aux réseaux sociaux.	--	--	--	--
http://www.arts-pi.org.tn/cristal.php	Pas d'association du site web aux réseaux sociaux. (un groupe sur Facebook avec 25 membres : https://www.facebook.com/groups/191878707500389/about/)	--	--	--	--
http://www.hanalab.org	Pas d'association du site web aux réseaux sociaux. Page indépendante sur Facebook : https://www.facebook.com/hanalabaudio/ (333 abonnées)	--	--	--	--
http://www.lamsin.rmu.tn	Pas d'association du site web aux réseaux sociaux. (un groupe sur Facebook avec 494 membres : https://www.facebook.com/groups/lamsin/about/)	--	--	--	--
http://www.edsti.enit.rnu.tn/fr/structures_02.php?id=17&code=LR-99-ES19&cod	Pas d'association du site web aux réseaux sociaux.	--	--	--	--

http://www.esct.rnu.tn/site/index.php/recherche/labo-thema	<p>Pas d'association du site web aux réseaux sociaux. - 1^{ère} page indépendante sur Facebook : https://www.facebook.com/THEMA-1834189560201833/ (344 abonnées)</p> <p>- 2^{ème} page indépendante sur Facebook : https://www.facebook.com/M%C3%A9thodes-Marketing-1872897092929451/ (49 abonnées)</p> <p>- 3^{ème} page indépendante sur Facebook : https://www.facebook.com/Urquarg-258610341226727/ (13 abonnées)</p> <p>- 4^{ème} page indépendante sur Facebook : https://www.facebook.com/Larequad-1771339266467447/ (74 abonnées)</p>	--	--	--	--
http://www.imhenit.com (Site web ne fonctionne pas)	--	--	--	--	--
http://www.urdiri.fdspt.rnu.tn/	Pas d'association aux réseaux sociaux.	--	--	--	--
http://www.diraset.rnu.tn	<p>Pas d'association aux réseaux sociaux. Page indépendante sur Facebook : https://www.facebook.com/Diraset-Etudes-Maghr%C3%A9bines-191477120872553/timeline (2866 abonnées)</p>	--	--	99 liens et 9 domaines	24 liens
http://www.lattice.rnu.tn	Pas d'association aux réseaux sociaux	--	--	--	--
http://www.soie.isg.rnu.tn (site web en cours de maintenance)	Pas d'association aux réseaux sociaux	--	--	--	--
http://www.larodec.com	<p>Pas d'association du site web aux réseaux sociaux. (un groupe sur Facebook avec 72 membres : https://www.facebook.com/groups/larodec/)</p>	--	--	74 liens et 21 domaines	2 liens
http://www.urdas.fdsps.rnu.tn/ (site web ne fonctionne plus)	--	--	--	--	--

http://www.hal.tn	Pas d'association aux réseaux sociaux.	--	--	362 liens et 20 domaines	14 liens
http://www.u2s.enit.rnu.tn	Pas d'association aux réseaux sociaux. (un groupe sur Facebook avec 30 membres : https://www.facebook.com/groups/150278435049288/)	--	--	4 liens et 2 domaines	0 liens
http://www.redcad.ens.rnu.tn (site web en cours de maintenance)	--	--	--	--	--
http://www.gedes.rnu.tn	Pas d'association aux réseaux sociaux.	--	--	1 page et 1 site	--
http://www.ercilis.flsh.s.rnu.tn	Pas d'association aux réseaux sociaux.	--	--	49 liens et 6 domaines	5 liens

Bibliographie

- LAFOUGE, Thierry ; LE COADIC, Yves-François ; MICHEL, Christine (2002). *Éléments de statistique et de mathématique de l'information : infométrie, bibliométrie, médiométrie, scientométrie, muséométrie, webométrie*. Paris : Presses de l'ENSSIB, Ecole nationale supérieure des sciences de l'information et des bibliothèques, 2002. 306 p.
- GOASDOUE, Rémi (2016). *La bibliométrie évaluative : une redéfinition des valeurs scientifiques* In journal international de recherche en éducation et formation. Neuchâtel: Association pour le développement des méthodes d'évaluation en éducation, n° 1, pp. 57-72
- GINGRAS, Yves (2014). *Les dérives de l'évaluation de la recherche : du bon usage de la bibliométrie*. Paris : Raison d'agir éditions, 2014. 122 p.
- MAHMOUD, Saloua (2009). *Observatoire des sites web universitaires en Tunisie : fracture scientifique ou collaboratoire*. Tunis :Fenzi, 2009. 254 p.
- Mathieu Andro, Gaëtan Tröger (2013). *Statistiques et visibilité des bibliothèques numériques : quelles stratégies de diffusion ?*. 7 p. 2013.
- MATHIEU F. (2004), *Graphes du web, Mesures d'importance à la PageRank*, Thèse de doctorat informatique, Université de Montpellier II.[en ligne]
<http://papyrus.lirmm.fr/Document.htm&numrec=031984291916600>
- PONTILLE, David ;TORNY, Didier (2013). *La manufacture de l'évaluation scientifique : algorithmes, jeux de données, outils*

- bibliométriques. In Réseaux : communication, technologie, société, n° 210, pp.25-61.
- REYMOND, David (2006). Intérêt informationnel et opérationnel de la webométrie pour la modélisation conceptuelle et la production de sites internet. In Actes du colloque international "Information dans les organisations : dynamique et complexité " 6 et 7 avril 2006 : Tours : Université Bordeaux Montaigne, 2006. 344 p. pp. 113-129.
- ROSTAING, Hervé ; (2001). Le Web et ses outils d'orientation. Comment mieux appréhender l'information disponible sur l'Internet par l'analyse des citations ? Bulletin des bibliothèques de France, 2001, n° 1, pp. 68-77. [En ligne] Disponible à : <http://bbf.enssib.fr/consulter/bbf-2001-01-0068-007> (Page consulté le : 16/08/2018).
- THELWALL, Mike (2017). Avoiding obscure topics and generalising findings produces higher impact research. Scientometrics, n° 110, pp. 307-320. [En ligne] Disponible à : http://wlv.openrepository.com/wlv/bitstream/2436/620209/1/Avoiding_gObscureTopicsPreprint.pdf (consulté le : 28/04/2018)
- THELWALL, Mike (2009). Introduction to webometrics: Quantitative web research for the social sciences. New York : Morgan & Claypool, 2009. 116 p.
- THELWALL, Mike; KOUSHA, Kayvan (2017). Goodreads: A social network site for book readers. In Journal of the Association for Information Science and Technology, n° 68, pp. 972-983. [En ligne] Disponible à : http://wlv.openrepository.com/wlv/bitstream/2436/609243/1/GoodReadsASocialNetworkSiteForBookReaders_preprint.pdf (consulté le : 22/03/2018)
- THELWALL, Mike; KOUSHA, Kayvan (2017). ResearchGate articles: Age, discipline, audience size and impact. Journal of the Association for Information Science and Technology, 68(2), 468-479. [En ligne] Disponible à : http://wlv.openrepository.com/wlv/bitstream/2436/609261/1/ResearchGateArticles_preprint.pdf (consulté le : 20/03/2018)
- THELWALL, Mike; KOUSHA, Kayvan (2017). SlideShare presentations, citations, users and trends: A professional site with academic and educational uses. In Journal of the association for information science and technology, n° 68, pp. 1989-2003. [En ligne] Disponible à : <http://wlv.openrepository.com/wlv/bitstream/2436/619242/1/SlideSharePaperPreprint.pdf>(consulté le : 22/03/2018)
- THELWALL, Mike ; SUD, Perdeep ; VIS, Farida (2012). Commenting on YouTube videos: From Guatemalan rock to El Big Bang. In Journal of the American Society for Information Science and Technology, n° 63, pp. 616-629. [En ligne] Disponible à : <http://eresearch-ethics.org/position/researching-the-public-web/>(consulté le 25/04/2018)

THELWALL, Mike; SUD, Perdeep; Wilkinson, Dadid (2012). Link and co-link network diagrams with URL citations or title mentions. In *Journal of the American society for information science and technology*, n° 63, pp. 805-816. [En ligne] Disponible à :http://www.scit.wlv.ac.uk/~cm1993/papers/URCitationsTitleMentionNetworks_preprint.pdf(consulté le: 25/04/2018)

ANNEXES

Annexe I : liste des structures de recherches et leurs URLS

Centres de recherche

Le Centre De Recherche Et Des Technologies Des Eaux (CERTE)	http://www.certe.rnrt.tn
Le Centre de Recherche et des Technologies de l'Energie (CRTEN)	http://www.crten.rnrt.tn
Le Centre de Biotechnologie De Borj-Cédria (CBBC)	http://www.cbcc.rnrt.tn
Le Centre National Des Sciences Et Technologies Nucléaires (CNSTN)	http://www.cnstn.rnrt.tn/
L'institut National De Recherche Et D'analyse Physico-chimique (INRAP)	http://www.inrap.rnrt.tn
L'institut national des sciences et technologies de la mer (INSTM)	http://www.instm.rnrt.tn
Le Centre de Biotechnologie De Sfax (CBS)	http://www.cbs.rnrt.tn
L'institut des Régions Arides De Médenine (IRA)	http://www.ira.rnrt.tn
L'Institut National de la Recherche Agronomique de Tunisie (INRAT)	http://www.inrat.agrinet.tn
L'institut national de nutrition et de technologie Alimentaire (INNTA)	http://www.institutdenutrition.rns.tn
L'institut Pasteur De Tunis	http://www.pasteur.tn
Le Centre Int. des Technologies de L'environnement De Tunis (CITET)	http://www.citet.nat.tn
Le Centre d'Etudes Et de Recherche des Télécommunications CERT	http://www.cert.nat.tn
Le Centre National de la Cartographie et de la Télédétection	http://www.cnct.nat.tn

Technopoles

Pôle de Compétitivité Monastir/El Fejja (Manouba) : Textile et Habillement	www.mfcpole.com.tn
Pôle Industriel et Technologique de Gabès (Pol.i.tech-Gabès)	www.polegabes.com.tn
Biotechnologie appliquée à la santé et Industries Pharmaceutiques (Sidi Thabet)	www.biotechpole.rnu.tn
Pôle de compétitivité de Bizerte : Industrie Agro-alimentaire	www.pole-competitivite-bizerte.com.tn
Pôle Elgazala des Technologies de la Communication	www.elgazalacom.nat.tn

Technopole de Sousse : Mécanique, Electronique et Informatique	www.technopole-sousse.rnrt.tn
Technopole de Sfax : TIC et Multimédia	www.sfax-icttechnopark.tn/
Pôle de Compétitivité de Gafsa "PCG"	www.polegafsa.com.tn

Laboratoires de recherche

Le laboratoire d'Études des Systemes Thermiques et Énergétiques (LESTE)	http://www.leste.enim.rnu.tn
Laboratoire génie mécanique (monastir)	http://www.lgm.rnu.tn
Laboratoire de Recherche en génie logiciel, applications distribuées, systèmes décisionnels et imagerie intelligente [ENSI]	http://www.riadi.rnu.tn
laboratoire CRISTAL (Association de la recherche tunisienne des sciences pour l'image)	http://www.arts-pi.org.tn/cristal.php
HANA HeterogeneousAdvenaced Networking & applications	http://www.hanalab.org
Laboratoire de modélisation mathématique et numérique dans la science de l'ingénieur	http://www.lamsin.rnu.tn
	http://www.edsti.enit.rnu.tn/fr/structures_02.php?id=17&code=LR-99-ES19&code1=LR-99-ES19&code2=&code3
Laboratoire THEMA "Théories Economiques, Modélisations et Applications"	http://www.esct.rnu.tn/site/index.php/recherche/lab-o-thema
Génie des procédés et systèmes industriels	http://www.labgpsci.com
Laboratoire de recherche en économie quantitative du développement	http://www.edirc.repec.org
Laboratoire de recherche opérationnelle, de décision et de contrôle de processus (LARODEC)	http://www.larodec.com
Le Laboratoire de Droit des relations Internationales, des Marchés et des Négociations DRIMAN	http://www.urdri.fdspt.rnu.tn

Etude bibliométrique des publications sénégalaises sur les maladies non transmissibles : libre accès vs accès payant

Bibliometric study of Senegalese publications on non communicable diseases: open access vs paid access

Kouakou Kouassi Sylvestre ; Diakhate Djibril. Laboratoire de recherche en sciences de l'information et de la communication (LARSIC). Université Cheikh Anta Diop (Sénégal).

Résumé

Cette étude vise à présenter un panorama de la production scientifique sénégalaise sur les MNT de 1960 à 2018. Via une approche bibliométrique, elle montre que cette production, bien qu'étant la plus importante en Afrique de l'Ouest francophone, demeure très faible avec seulement 808 publications. En outre, elle est d'un impact scientifique très limité avec 5957 citations. Aussi, la collaboration nationale entre institutions de recherche est moins dense alors que celle internationale, elle est beaucoup plus importante avec comme partenaire principal la France.

Par ailleurs, la proportion de publications dans des revues en Open Accès pour ces dix dernières années est très insignifiante avec seulement environ 10% des articles (soit 34/353).

Mots clés. Bibliométrie, maladie non transmissible, libre accès, Sénégal.

Abstract

This study aims to present a panorama of Senegalese scientific production on NCD from 1960 to 2018. Through a bibliometric approach, it shows that this production, although the most important in francophone West Africa, remains very weak with only 808 publications. In addition, it has a very limited scientific impact with 5957 citations. Also, the national collaboration between research institutions is less dense whereas the international one, it is strong with as main partner France. Moreover, the proportion of publications in Open Access journals for the last ten years is very insignificant with only about 10% of articles (ie 34/353).

Keywords. Bibliometrics, noncommunicable disease, open access, Senegal.

1 Introduction (problème et questions de recherche)

Les maladies non transmissibles (MNT) représentent aujourd'hui un vrai problème de santé publique dans le monde et plus particulièrement dans les pays à revenu faible comme le Sénégal. Ces maladies notamment celles cardiovasculaires, les cancers, les affections respiratoires et le diabète sont les premières causes de mortalité dans le monde avec en 2012, 38 millions (68%) des 56 millions de décès (OMS, 2016). Les pays à revenu faible ou intermédiaire à elles seules ont enregistré 28 millions (près de 75%) des décès. Au Sénégal, en (absence de chiffres plus récents) 2014, 34% des 98 000 décès constatés étaient dus aux MNT (OMS, 2014b). De plus, il est à remarquer que le taux de morbidité et de mortalité des MNT est en constante croissance au Sénégal depuis 2011. Selon les derniers chiffres de l'OMS datant de 2014, il est passé de 30% à 34%, soit une croissance de 4% entre 2011 et 2014 (OMS, 2014 ; 2011). Cette situation alarmante n'est pas singulière au Sénégal. Ainsi, la directrice générale de l'OMS (2007-2017) avait lancé l'appel suivant : « (...) *Les pays doivent progresser sur tous ces fronts pour atteindre la cible globale qui est de réduire de 25 % la mortalité imputable aux 4 principales MNT d'ici 2025...* » (OMS, op.cit.). Il apparaît clairement que les états sont invités à développer des politiques afin d'enrayer l'épidémie des MNT.

Disposer d'une cartographie des publications scientifiques sur les MNT au Sénégal, permettrait d'élaborer des politiques publiques efficaces. L'approche bibliométrique est la mieux indiquée pour effectuer cet état des lieux. Cette approche permet de quantifier la production scientifique des institutions de recherche, des chercheurs ou des pays, et de connaître les tendances et évolutions de la recherche. L'analyse bibliométrique en donnant un aperçu des activités de recherche dans un domaine spécifique, pourrait être utilisée comme outil d'aide à la décision.

Très peu d'études bibliométriques sur les publications en sciences médicales ont été réalisées en Afrique. Ainsi, Diakhaté (2014) et Mègnibêto (2014) ont respectivement dressé une cartographie des connaissances au Sénégal et au Bénin en mettant l'accent sur la visibilité scientifique des institutions de recherche pour le premier et la collaboration scientifique entre les chercheurs pour le second. Seules quelques recherches en Afrique du Nord se sont intéressées aux productions médicales. Salem et coll. (2011) ont dressé le profil bibliométrique des thèses soutenues à la faculté de médecine de Monastir (Tunisie) en soulignant les disciplines les plus pourvoyeuses de thèses. Quand, Lrhoul et coll. (2015), eux ont établi une cartographie des publications de la faculté de Médecine et de Pharmacie de Casablanca en mettant l'accent sur l'évolution de la production et les réseaux de collaboration nationale et internationale.

La présente étude a pour objectif de décrire et d'analyser la production scientifique sénégalaise sur les MNT afin de guider la mise au point de mesures politiques de recherche qui permettront d'instaurer un environnement favorable à la lutte contre ces maladies.

Elle ambitionne de répondre aux questions suivantes : comment évolue la production scientifique sur les MNT au Sénégal? Quel est l'impact scientifique de cette production? Quel est son rayonnement international? Quel est le niveau de collaboration des chercheurs sénégalais? Quelle est la proportion de la production issue des périodiques en Access ? Y a-t-il une la corrélation entre la nature du périodique avec l'impact scientifique de la publication ?

2 Méthodologie

2.1 Constitution du corpus

Malgré l'existence d'une pluralité de solutions pour le recueil de données bibliographiques, notre choix s'est porté sur Scopus pour un certain nombre de raisons. En matière d'études bibliométriques, *Scopus* est une référence derrière Web of Science de Thomson et il est assez exhaustif dans le signalement des publications en Sciences Techniques et Médecine (STM). A la différence de Pubmed, il est possible avec Scopus de faire un filtrage des publications selon le modèle d'accès. Un critère essentiel dans cette étude où il est question de comparaison entre libre accès et accès payant. Enfin, l'absence d'indice de citation dans Pubmed, un indicateur important dans la mesure de la notoriété des auteurs et de la qualité des publications, nous a résolument orientés vers Scopus.

La stratégie de recherche adoptée a été d'abord de réaliser un tableau de concepts en anglais et en français en partant des mots clés les plus pertinents dans les thématiques des MNT. Ensuite, à l'aide du thésaurusMeSH¹ (MedicalSubjectHeading) France des descripteurs associés aux concepts ont été identifiés (tableau 1).

¹ Le MeSH (MedicalSubjectHeadings) est le thésaurus de référence dans le domaine biomédical. NLM (U.S. National Library of Medicine)- <http://mesh.inserm.fr/FrenchMesh/>

Diabète	Asthme	Hypertension	Cancer	Sénégal
Diabetes Mellitus	asthma	Hypertension artérielle		Sénégalais
Troubles du métabolisme du glucose	Respiratory Tract Diseases	Cardiovasculaire		
Glucose Metabolism Disorders	Maladies de l'appareil respiratoire	Cardiovascular		
Maladies métaboliques				
Metabolic Diseases				

Tableau 1. Plan de concepts et des termes associés

Les tests opérés dans Scopus nous ont convaincus de réaliser une seule interrogation par thème en connectant les différents descripteurs par les opérateurs booléens. L'objectif était de pouvoir disposer de données permettant de juger spécifiquement l'évolution dans le temps de la production scientifique dans chaque maladie.

Cependant pour les besoins l'analyse globale qui porte sur les MNT, il a été plus judicieux de recourir à une seule équation de recherche.

Un éclatement (une équation de recherche par thème) certes aboutirait au même résultat, mais créerait des redondances qui altéreront l'interprétation des données.

[TITLE-ABS KEY ("hypertension artérielle" OR hypertension OR cardiovasculaire OR cardiovascul ar)OR TITLE-ABS-KEY (asthma OR asthma OR "Respiratory Tract Diseases" OR "Maladies de l'appareil respiratoire") OR TITLE-ABS-KEY (diabetes OR "Diabetes Mellitus" OR "Troubles du

*metabolisme du glucose" OR "Glucose Metabolism Disorders"OR "Maladies metaboliques" OR "Metabolic Diseases")OR TITLE-ABS
KEY (cancer) AND AFFILCOUNTRY (senegal)]*

Le filtrage géographique réalisé à partir du champ « Affiliation institutionnelle » a permis de restreindre les résultats de recherche aux seules publications des chercheurs des institutions sénégalaises. Un second filtre a été appliqué « document type » afin de sélectionner les articles scientifiques de recherche en excluant les monographies, chapitres de livre, les notes de synthèses, les références bibliographiques, les notes de synthèses, les revues de littérature, les résumés, etc. Un dernier filtre a été posé selon le modèle d'accès (Libre accès et Accès payant) en perspective de la comparaison entre les deux systèmes.

2.2 Préparation des données

Les résultats issus de la recherche sur Scopus ont été exportés sous format RIS² afin de les rendre exploitables par les outils de traitement bibliographique parmi lesquels Bibexcel³ dont l'utilisation dans les études bibliométriques est de plus en plus répandue (Persson, Danell, et Schneider 2009). Avec ce dernier il a été d'abord question d'extraire du corpus de données et de représenter sous forme de tableau les champs bibliographiques sur lesquels se fonde notre étude. Une préparation qui a rendu possible :

- Le classement des revues en fonction des citations et du nombre d'articles reçus
- L'extraction et le classement des publications par pays et par institution de recherche
- L'exportation des données vers Excel pour une analyse statistique.

Ensuite, Bibexcel du fait de sa compatibilité avec le logiciel Vosviewer a été utile dans la mise en forme des données bibliométriques pour une présentation cartographique des réseaux de collaboration.

² RIS (Systèmes d'Information Recherche) principalement utilisés pour la mise en œuvre rapidement et facilement des fonctionnalités de citations et bibliographies dans des documents texte en utilisant une grande variété d'applications.

³Bibexcel est une application développée par OllePersson qui permet plusieurs types d'analyse bibliométrique. Le programme offre des passerelles avec Excel et SPSS et les logiciels de cartographie des réseaux

1.1 Choix des indicateurs bibliométriques

La bibliométrie, application des méthodes statistiques aux livres, articles et autres moyens de communication (Rostaing 1996) offre un certain nombre d'indicateurs pour mesurer l'activité scientifique d'un chercheur ou d'une institution sur une période donnée. En ce qui concerne notre étude nous avons choisi les indicateurs en fonction de leur capacité à apporter des réponses pertinentes aux différentes questions posées :

- Le nombre de publications dont l'intérêt est exclusivement quantitatif. Nous considérons les articles dans cette étude, car ils représentent le mode d'expression privilégié des connaissances nouvelles (Okubo 1997). La citation des articles, un indicateur de mesure de la qualité. Elle permet de mesurer l'impact d'un article ou d'un auteur sur la communauté scientifique. Cependant dans la perspective de Seglen (1992), elle n'est ni équivalente, ni corrélée de manière non équivoque à la qualité scientifique.
- Le H-index ou l'index de Hirsh du nom de son inventeur qui permet de calculer le facteur d'impact d'un auteur ou d'un journal à partir du nombre de citations reçues dans une base de données (Hirsch 2005). Il soutient que le facteur h d'un auteur est le nombre d'articles de cet auteur, qui ont reçu au moins h citations chacun. Du fait que la base Scopus ne propose pas directement les H-index des revues, nous avons été obligés de recourir au SCImago Journal Rank (SJR) qui utilise la base de données Scopus pour classer annuellement les revues en fonction de plusieurs critères parmi lesquels le H-index. Selon Schöpfel et Prost (2009) le SJR est l'équivalent du Facteur d'Impact (IF) de l'Institute of Scientific Information et bénéficie d'une couverture plus large que celle de l'IF et meilleure surtout dans les sciences médicales. D'où son intérêt pour le classement des titres de revues de notre corpus.

3 Présentation des résultats

3.1 La production scientifique sénégalaise sur les MNT

Entre le 04 avril 1960 et le 30 juin 2018, seulement 808 articles ont été publiés dans des revues scientifiques à comité de lecture. Comme le montre la figure 1, de 1960 à 1970, le Sénégal n'a enregistré qu'une seule publication par an. Ce n'est qu'à partir des années 1970 que le Sénégal va enregistrer plus d'un article sans en atteindre 20 par an. Ce nombre va chuter et stagner autour de 10 articles par an jusque dans les années 2010, où l'on observe une croissance nette et continue des publications allant

jusqu'à une soixantaine par an. Cette tendance va se poursuivre et probablement doubler, car au moment de la rédaction de cet article (au mois de juillet), l'on notait pour l'année 2018, déjà 58 publications.

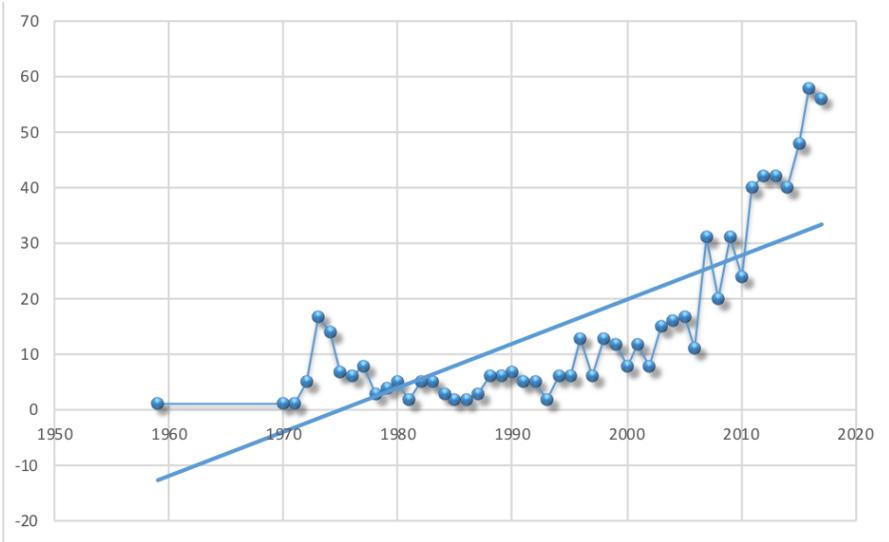


Figure 1 : Evolution des publications sur les MNT

Suivant le tableau 1, les publications sénégalaises sur les MNT concernent en majorité le cancer (n= 314) soit 39% suivi des maladies cardiovasculaires (n= 284) avec 35%, puis des articles sur le diabète (n=168) et seulement n=42 pour l'asthme.

Maladies	Total
Maladies cardiovasculaires	284
Asthme	42
Diabète	168
Cancer	314
Total	808

Tableau 2. Distribution des publications suivant lesMNT

L'évolution des publications (figure 2) sur chacune de ces maladies montre que les premières publications sur les MNT au Sénégal ont concerné le cancer (de 1960 à 1970), ensuite les publications sur les maladies cardiovasculaires et le diabète dès les années 70. Ce n'est qu'au milieu des années 70, que le premier article sur l'asthme a été publié. Les courbes d'évolutions montrent que si pour l'ensemble des maladies, les publications ont progressé considérablement à partir de 2010, il n'en est pas de même pour l'Asthme qui n'a connu d'évolution de publications que très récemment (2015) pour chuter aussitôt.

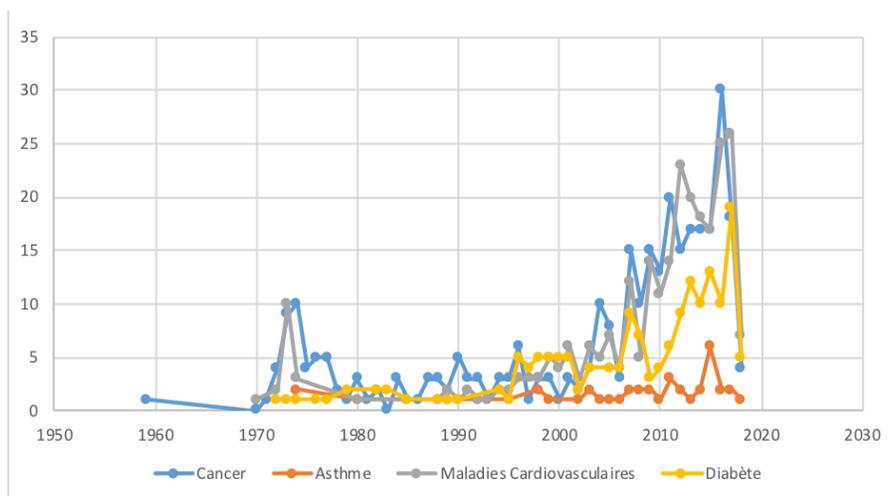


Figure 2. Évolution des publications par maladie

3.2 La répartition suivant le titre de revue

La production scientifique dans les MNT est riche de 808 articles répartis dans 305 revues scientifiques. Autant il est important de souligner la faiblesse de la production en 60 ans de recherche autant il est intéressant de faire remarquer la diversité des revues. L'ensemble des publications est distribué entre 305 titres de revues, soit une moyenne de 2, 33 articles par revue.

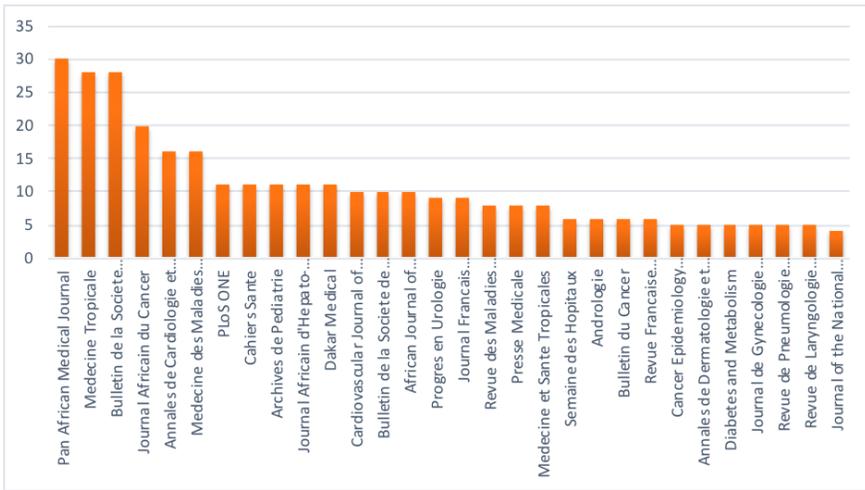


Figure 3. Répartition des publications dans les revues spécialisées

Le classement des revues par nombre d'articles publiés (figure 3) montre une prépondérance de celles généralistes et géographiquement connotées.

Il s'agit de la revue Pan African Medical Journal qui domine le classement avec 30 articles, suivi de Médecine Tropicale (28 articles) et du Bulletin de la Société Médicale d'Afrique Noire de Langue Française (28 articles). Les revues spécialisées dans les MNT sont classées 4^e (Journal Africain du Cancer), 12^e (Cardiovascular Journal of Africa) et 18^e (Revue des Maladies Respiratoires) avec respectivement 20, 10 et 8 articles. La première et unique revue locale (Dakar Medical) vient en 11^e position avec seulement 11 articles.

3.3 La collaboration scientifique nationale sur les MNT

La production sénégalaise sur les MNT au niveau local est dominée par l'Université Cheikh Anta Diop (271 publications), les Centres Hospitaliers Universitaires en l'occurrence le CHU de Fann et le CHU Aristide le Dantec avec respectivement 174 et 118 articles signés (figure 4). Les centres médicaux non universitaires notamment l'Hôpital Général de Grand Yoff, le Centre Marc Sankalé, les centres hospitaliers régionaux, malgré une présence relativement faible témoignent d'une certaine vivacité de la recherche médicale au Sénégal.

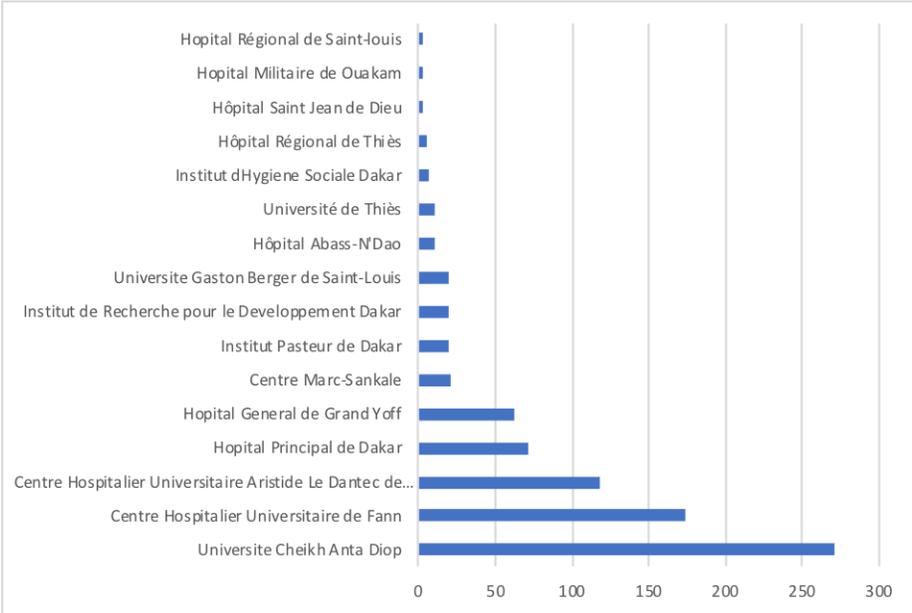


Figure 4. Distribution des publications suivant les acteurs locaux

Au niveau national, les acteurs de la recherche sur les MNT sont très peu nombreux et gravitent autour des universités, Centres Hospitaliers Universitaires (CHU), et hôpitaux régionaux. Ils ont établi à travers les co-publications un système de collaboration, mais qui reste faible comme en atteste la transparence de la carte (figure 5) et le peu de nœuds représentés dans les groupes. Au sein de ces derniers, les différentes entités de recherche s'agglomèrent autour de leaders tels que le CHU Aristide le Dantec (vert), l'université Cheikh Anta Diop (jaune), l'hôpital général de Grand Yoff (rouge) et le CHU de Fann (bleu).

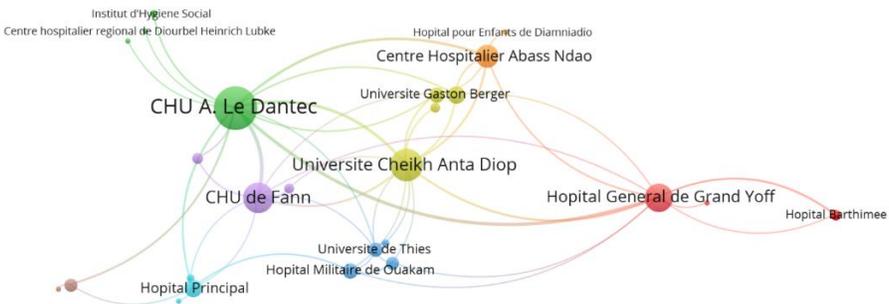


Figure 5. Cartographie des co-publications nationales

3.4 La collaboration scientifique internationale sur les MNT

En bibliométrie, pour mesurer la collaboration internationale qui se calcule sur la base des co-signatures des publications, il existe deux méthodes :

- Le comptage entier qui attribue un crédit entier à chaque pays d'origine d'un auteur cosignataire d'une publication internationale,
- Le comptage fractionnaire qui repose sur le principe de partager le crédit aux auteurs cosignataires.

Chacune des méthodes présente des avantages et des inconvénients. Pour notre étude, nous avons choisi la deuxième méthode qui nous semble la plus équilibrée du fait parfois du nombre souvent excessif de coauteurs dans les publications en médecine. Ainsi, remarquons-nous pour des raisons historiques et de proximité scientifique que la France (80 co-signatures) représente le pays qui collabore le plus avec le Sénégal suivi des États Unis (30) et de l'Italie (7) (Figure 6). Sur cette liste, le premier pays africain arrive 4e. Il s'agit du Cameroun, suivi de la Cote d'ivoire, du Nigéria et du Burkina, respectivement 6e, 8e et 9e avec à peine 5 collaborations.

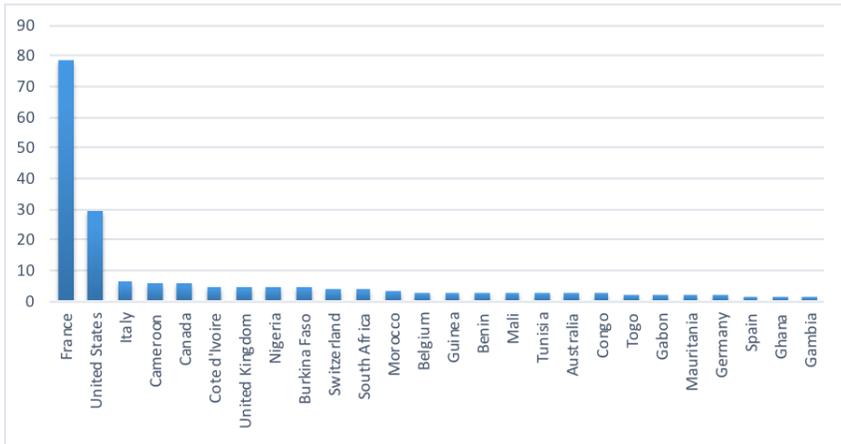


Figure 6. Classement des pays qui collaborent le plus avec le Sénégal

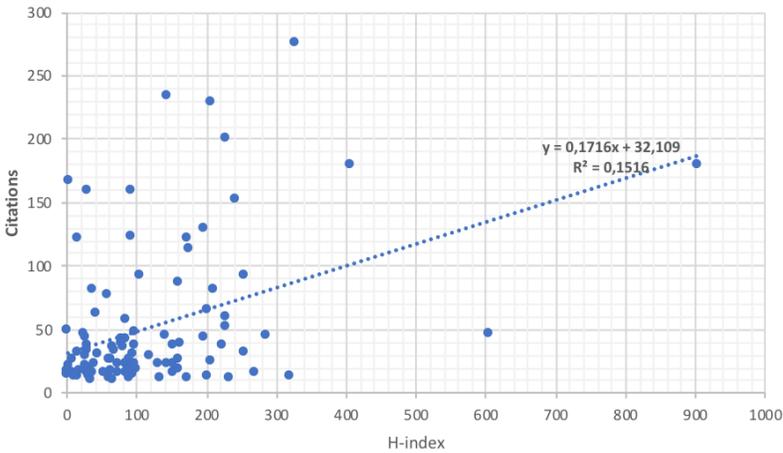
3.5 Impact scientifique des publications sénégalaises sur les MNT

A la lumière du tableau 3 ci-dessous, il apparaît clairement que le nombre de citations n'est pas toujours corrélé au nombre d'articles publiés. Autrement dit ce n'est pas la quantité de la production qui fait notoriété.

En effet, sur le corpus de 808 articles, moins d'une centaine (n=95) mobilisent plus de la moitié des citations (56%).

Cependant, en comparant le nombre de citations avec l'impact des revues, on constate que la quasi-totalité des articles les plus cités a été publiée dans des revues à H-index élevé. Sur les 18 titres ayant reçu au moins 2% des citations, seules 2 revues (Archives of Internal Medicine et Nature Genetics) ont un impact facteur faible (inférieur à 35).

Figure 7 : Évolution des citations en rapport avec le H-index des revues



Cette tendance se confirme dans la figure 7 ci-dessus, qui montre que la visibilité des articles et leur impact éventuel dépendent des revues dans lesquelles ils sont publiés. Il existerait donc une corrélation positive ($R^2 = 0,1516$) entre le nombre de citations et la réputation de la revue mesurée.

H-index	Revue	Citation	Article	% Citation
326	Journal of the National Cancer Institute	276	4	5%
144	Bulletin of the World Health Organization	234	2	4%
206	International Journal of Cancer	229	3	4%
227	Journal of Infectious Diseases	201	3	3%
404	Cancer Research	180	3	3%
902	New England Journal of Medicine	180	1	3%
1	Archives of Internal Medicine	167	1	3%
27	Medecine Tropicale	160	28	3%
92	Pediatric Nephrology	160	1	3%

241	PLoS ONE	153	11	3%
195	CardiovascularResearch	129	1	2%
91	Pediatric Blood and Cancer	124	4	2%
172	Cancer Epidemiology Biomarkers and Prevention	122	5	2%
14	Nature Genetics	122	1	2%
174	International Journal of Epidemiology	113	2	2%
105	Journal of MedicalVirology	94	3	2%
252	The Lancet Oncology	93	3	2%
159	Journal of Hypertension	87	4	1%
35	Annales de Dermatologie et de Venerologie	82	5	1%
209	EuropeanRespiratory Journal	81	1	1%
58	HumanHereditiy	77	1	1%
200	AIDS	65	1	1%
41	Archives of CardiovascularDiseases	63	2	1%
227	American Journal of Epidemiology	60	2	1%
83	Current Opinion in Nephrology and Hypertension	58	1	1%
226	Journal of ClinicalMicrobiology	52	2	1%
Total		3412	95	56%

Tableau 3 : Classement des revues selon le nombre de citations

3.6 Libre accès Vs Accès payant : quel impact scientifique ?

Depuis l’initiative de Budapest, BOAI (Budapest Open Access Initiative) en 2002 appelant les acteurs de l’information scientifique et technique notamment les chercheurs et les professionnels de l’information à mettre en place des stratégies concourant à rendre facilement accessible les publications, le mouvement du libre accès est arrivé en phase de maturation. Il n’a pas encore réussi à renverser le monopole des grands éditeurs scientifiques, mais contribue fortement à faire coexister avec leur modèle payant un autre proposant un accès ouvert. Une quinzaine d’années plus tard, les différents domaines de la science se sont enrichis d’archives ouvertes, de revues en accès libre (Suber 2012). Les éditeurs classiques ne voulant pas se mettre en marge de ce mouvement proposent désormais certains de leur contenu en accès libre (Farchy et Froissart 2010). Fort de constat, il nous a paru nécessaire de greffer dans cet article une étude comparative entre les deux modèles. Nous avons choisi de démarrer l’étude comparative à partir de 2008, date de publication du

premier article en Open accès dans notre corpus. Les données extraites se présentent comme suit :

	Accès payant		Open Access	
	Quantité	Pourcentage	Quantité	Pourcentage
Revue	182	91%	18	9%
Publication	319	90%	34	10%
Citation	2462	89%	289	11%

Tableau 4. Distribution des publications selon le modèle d'accès

Les données analysées confirment la nette domination du système payant contrôlé par les majors de l'édition scientifique. La part du libre accès dans les publications sénégalaises en MNT reste alors faible (10%) (Tableau 4). Une tendance suivie par les revues en libre accès (9%) et les citations (11%). Concernant les revues en accès payant, l'analyse d'impact (tableau 5) conforte la perception évoquée précédemment selon laquelle l'importance du nombre d'articles publiée influe moins sur le nombre de citations que la notoriété des revues dans lesquelles ils sont publiés. En effet, 6,5% des articles (n=21) mobilisent plus de la moitié (52%) des citations du corpus. Et la quasi-totalité des revues dans lesquelles ces publications ont été faites dispose d'un fort H-index. Pourtant, certaines d'entre elles, pour ne pas dire la plupart n'ont publié qu'un seul article du corpus. C'est le cas du *New England Journal of Medicine* (H-index 902), du *Pediatric Nephrology* (H-index 92) et *Pediatric Blood and Cancer* (H-index 91), et respectivement 1^{er}, 3^e et 4^e dans l'indice de citation. L'analyse de l'impact des revues en libre accès offre une lecture différente. L'évolution des citations suit à une exception près celle des publications (tableau 6). Les revues qui présentent le plus de publications sur les MNT comptabilisent les indices de citation les plus élevés. *Plos One*, par exemple compte plus de la moitié des citations (53%) avec ses 11 articles, suivie d'*Archives of Cardiovascular Diseases* (2 articles, 22% des citations) et de *The Lancet Global Health* (2 articles, 11% des citations). Par ailleurs, faisons remarquer que l'évolution des citations en fonction de l'impact des revues, reste identique, quel qu'en soit le mode d'accès (payants ou libre). Cependant, les citations restent toujours plus importantes dans les revues en libre accès à fort H-index.

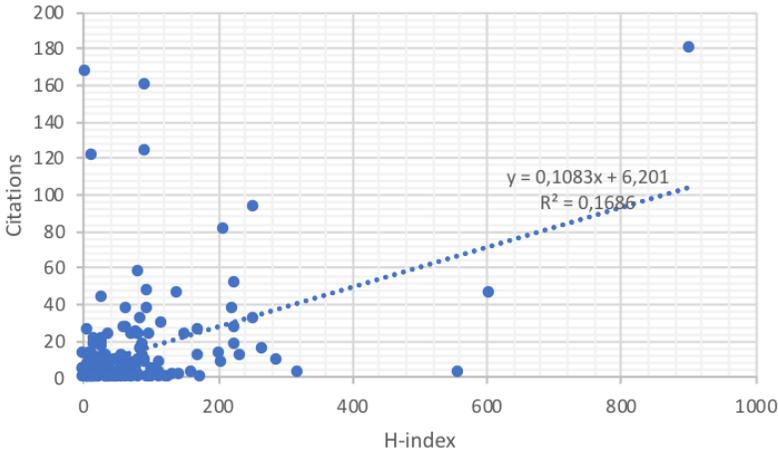


Figure 8. Citations en rapport avec le H-index des revues en accès payant

S'agissant de l'interdépendance entre l'évolution des citations et l'impact des revues, le calcul des régressions (Figure 9 et 10) démontre une corrélation positive et ceci indépendamment des modes d'accès. Il serait toutefois intéressant de faire remarquer que cette association positive est plus prononcée dans le libre accès que dans le système payant ($R^2 = 0,3197$ contre $R^2 = 0,1686$).

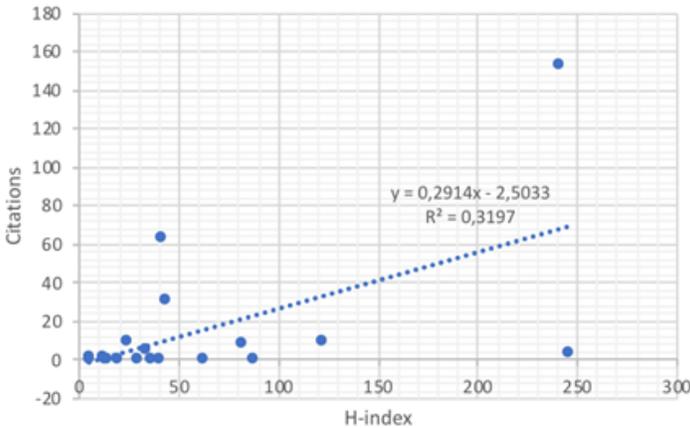


Figure 9 : Citations en rapport avec le H-index des revues en OA

Évolution des citations selon le modèle d'accès : Accès payant

Rang	H-Index	Revue	Citation	Article	% Citation
1	902	New England Journal of Medicine	180	1	7%
2	1	Archives of Internal Medicine	167	1	7%
3	92	Pediatric Nephrology	160	1	6%
4	91	Pediatric Blood and Cancer	124	4	5%
5	14	Nature Genetics	122	1	5%
6	252	The Lancet Oncology	93	3	4%
7	209	European Respiratory Journal	81	1	3%
8	83	Current Opinion in Nephrology and Hypertension	58	1	2%
9	226	Journal of Clinical Microbiology	52	2	2%
10	95	Materials Science and Engineering C	48	1	2%
11	603	JAMA - Journal of the American Medical Association	47	1	2%
12	141	Journal of Acquired Immune Deficiency Syndromes	46	1	2%
13	221	International Journal of Radiation Oncology Biology Physics	38	3	2%
14	95	Diabetes Research and Clinical Practice	38	1	2%
15	65	Fundamental and Clinical Pharmacology	37	1	1%
Total			1291	21	52%

Tableau 5. Classement des revues en accès payant selon le nombre de citations

Open Access

Rang	H-index	Revue	Citations	Articles	% Citations
1	241	PLoS ONE	153	11	53%
2	41	Archives of Cardiovascular Diseases	63	2	22%
3	43	The Lancet Global Health	31	2	11%
4	24	International Journal of General Medicine	10	1	3%
5	122	Clinical Microbiology and Infection	10	1	3%
6	82	BioMed Research International	8	1	3%

7	34	Archives de Pediatrie	5	1	2%
8	246	Kidney International	4	1	1%
9	12	International Journal of Surgery Case Reports	2	1	1%
10	5	African Journal of Urology	2	3	1%
11	87	International Journal of Nanomedicine	1	1	0%
12	62	The Scientific World Journal	0	1	0%
13	14	New Microbes and New Infections	0	1	0%
14	36	Journal of Food and Drug Analysis	0	1	0%
15	40	International Journal of Endocrinology	0	1	0%
Total			289	29	100%

Tableau 6. Classement des revues en Open Access selon le nombre de citations

4 Discussion

4.1 Discussion relative à la productivité scientifique sénégalaise sur les MNT

Notre étude révèle que la recherche sur les MNT est en constante progression au Sénégal. Ainsi, sur les dix dernières années (2007-2017), les publications annuelles ont doublé, passant de 30 à 60. Une recherche rapide sur Scopus avec la même équation de recherche, et pour la même période, montre que le nombre de publications du Sénégal sur les MNT (n=808) est de loin supérieur à celui des autres pays francophones de l'Afrique de l'Ouest francophone (à titre d'illustration : Cote d'Ivoire n=539, Bénin n=249). Toutefois, il reste largement en deçà de celui des pays anglophones du même espace géographique tels que le Nigéria n=6491 et le Ghana n=1104. Ces résultats sont identiques à ceux obtenus par Menigbeto à l'issue de son étude comparative des publications scientifiques au Ghana, au Bénin et au Sénégal. Il affirmait ceci :

« Senegal is ranked first, Ghana second and Benin third regarding the number of citations per cited documents or citable documents. The total number of citations Senegalese papers received is the highest. As a result, eventhough the Senegalese annual number of papers is lower than the Ghanaian, Senegal's h-index is higher ». (Menigbeto, 2013)

Deux raisons essentielles peuvent justifier la lente progression des publications sénégalaises sur les MNT.

D'abord, les MNT ont été pendant longtemps négligées dans le financement de la recherche en science médicale au Sénégal. En effet, la priorité a été accordée depuis les indépendances aux maladies infectieuses (pandémiques et épidémiques) notamment le paludisme, le SIDA, la tuberculose, le choléra, etc. Ce qui a poussé plusieurs laboratoires et chercheurs à consacrer leurs recherches à ces maladies. Pour les quelques chercheurs sur les MNT, leur capacité de production est extrêmement limitée, car ne disposant pas de moyens financiers suffisants.

Ensuite, le manque d'équipe de chercheurs et de spécialistes effectuant une recherche fondamentale et appliquée sur les MNT, combiné à la jeunesse des programmes de formation en médecine dans les universités de province pourrait expliquer cette faible production sur les MNT. À titre d'illustration, on dénombre au Sénégal pour un total de 15 millions d'habitants seulement 15 pneumologues (quotidien, 2018) et 80 cardiologues selon le professeur Serigne Abdou Bâ, chef du service de cardiologie de l'hôpital Aristide Le Dantec de Dakar.

Par ailleurs, l'on note un regain d'intérêt pour les MNT depuis environ deux décennies. Cette situation s'explique par le fait que les maladies « anciennes » comme le paludisme, la tuberculose, etc. sont en passe d'être vaincues et que parallèlement l'on a assisté à l'émergence de ces « nouvelles » pathologies (les MNT) en Afrique et principalement au Sénégal. L'Union africaine (2013) indique que les risques de développer 1 MNT se sont considérablement accrus du fait des changements démographiques et de modes de vie. Les gouvernements africains ont pris de conscience de la gravité de la situation et ont adopté en 2001 à Durban en Afrique du Sud des résolutions sur les maladies cardiovasculaires (Yaya et Kengne, 2014). Ces concours de circonstances vont favoriser le financement et la création de centre de recherche et de société savante dans

le domaine. Ainsi, en 2016, la multinationale pharmaceutique britannique Glaxo Smith Kline (GSK) a lancé en 2014 un projet dénommé *AfricaNcd Open Lab*. Le but étant de créer un réseau de chercheurs et de groupe d'universitaires africains travaillant sur les MNT auquel GSK va fournir un soutien scientifique, technique et financier (depuis, 4 ans, GSK octroie des bourses bisannuelles de recherche à hauteur de 100 000 £, soit 73 millions de F CFA).

4.2 Discussion relative à la collaboration scientifique nationale et internationale

Au niveau national, la collaboration entre différentes équipes de recherche est très insignifiante. Et cela en dépit de la création par les autorités universitaires, dans le cadre du contrat de performance des universités sénégalaises en 2010 de laboratoires et d'écoles doctorales avec parfois une transdisciplinarité qui incite à la collaboration. Notre étude a montré que les publications entre les institutions de recherche est beaucoup plus marquée par celles l'université Cheikh Anta Diop de Dakar et les CHU de la capitale et beaucoup moins avec les autres universités de province. Ce constat pourrait se comprendre sous deux ordres.

Premièrement, les professeurs de l'université Cheikh Anta Diop de Dakar cumulent pour la plupart des fonctions de chefs de service dans ces centres hospitaliers. Ainsi, leurs équipes de recherche sont à la fois constituées de chercheurs de l'université et des spécialistes officiants dans ces CHU. Il existe donc une proximité géographique et parfois hiérarchique qui influence considérablement la collaboration.

Deuxièmement, il est difficile de trouver des équipes de chercheurs sur les MNT dans les universités de province (de par la création relativement récente de leur programme de formation en sciences médicales) ce qui logiquement influence sur la possibilité de réaliser des co-publications. En outre, une « saine » concurrence existe entre les universités, notamment entre l'UCAD et l'UGB, ce qui fait que les équipes existantes travaillent de façon isolée dans leur « chapelle » et espère produire beaucoup plus que l'autre afin d'avoir à « juste titre » une reconnaissance étatique et attirer des financements.

Toutefois, la collaboration prise sous l'angle de la co-publication entre chercheurs c'est-à-dire ne tenant pas compte de la variable institution de

rattachement aurait probablement été forte. C'est d'ailleurs, ce que l'étude de Diakhaté (2014) sur la visibilité de la recherche a montré. Elle a révélé que les équipes de recherche à l'UCAD sont plus tournées vers les collaborations locales avant de s'ouvrir à l'international : « *La collaboration est très active au niveau local, entre collègues d'une même faculté, d'un même établissement. Parfois très dense entre deux collègues (...)* » Diakhaté (2014). Également, la faible collaboration nationale intra-institutions est à nuancer à cause de la durée tenue en compte (1960 à aujourd'hui). En effet, créée depuis 1957, l'université de Dakar devenue université Cheikh Anta Diop en 1986 a été la seule université du Sénégal jusque très récemment c'est-à-dire en 1990 avec la création de l'UGB et 2007 pour les universités de Ziguinchor et de Bambey.

Pour la collaboration internationale, nous constatons sans surprise que le premier pays avec lequel les institutions sénégalaises sur les MNT publient le plus est de loin la France qui concentre à elle seule 60 co-publications. Cela peut se comprendre pour des raisons historiques, de proximité scientifique et de langue. En plus d'avoir créé la première université sénégalaise, la France a formé la quasi-totalité des professeurs et chercheurs sénégalais tant en sciences médicales que dans d'autres disciplines. Ces derniers publient pour la plupart du temps avec leurs professeurs ou collègues des laboratoires et universités en France, dans lesquels ils ont réalisé leur thèse ou diplôme de spécialisation. En outre, la deuxième position occupée par les Etats-Unis (30 co-publications) et les positions relativement acceptables du Canada (5^{ème}) et du Royaume-Uni (7^{ème}) avec un peu plus de 5 publications dans un pays, dont la langue officielle, est le français sont conformes à l'étude bibliométrie des publications béninoises sur Scopus menée par Mègnigbèto en 2012. Elle est révélatrice de l'importance du monde anglo-saxon, notamment de la langue anglaise dans la recherche scientifique. Par contre, ce qui nous semble très surprenant, c'est la position très reculée (4^{ème}, 6^{ème}, 8^{ème} et 9^{ème}) des premiers pays africains dans les statistiques de collaboration. Nous avons supposé au vu des revues dans lesquels les chercheurs sénégalais ont le plus publié que les pays africains arriveraient en tête de la collaboration internationale. Nous émettons donc l'hypothèse que cette collaboration prédominante avec les pays occidentaux est guidée non par des raisons

scientifiques, mais bien plus par des opportunités de visibilité et d'accroissement d'audience (Mêgnigbêto, 2013).

Par ailleurs, une cartographie du réseau des co-publications donne une lisibilité plus intéressante des relations entre partenaires, des sous-groupes et de l'influence de chacun dans le groupe auquel il appartient. Chaque nœud représente un partenaire dont la taille dépend du nombre de liens qu'il entretient avec les autres membres du groupe reconnaissables sous la même couleur. En d'autres termes plus le nœud est gros plus il est influent dans le groupe.

Ainsi pour ce qui est du réseau de collaborations internationales (figure 10), la carte fait ressortir cinq groupes majeurs de partenariat :

- Les Anglo-saxons (bleu) dominés par les États-Unis, le Royaume-Uni, l'Afrique du Sud et le Nigéria.
- Les francophones (vert) avec une forte influence de la France et du Sénégal
- Un groupe de partenaires plus hétéroclites (rouge) constitués par une majorité de pays européens, sud-américains et de quelques pays asiatiques avec une distribution des forces entre l'Allemagne, l'Espagne, l'Italie et le Brésil
- Le groupe des pays de l'Asie du Sud-est et l'Australie (cyan) caractérisé par une faible densité, mais avec une très forte interaction impulsée par la Chine, l'Australie et l'Inde
- Le groupe des pays arabes (jaune), dans la périphérie du réseau avec une relative domination de l'Arabie Saoudite et de l'Egypte.

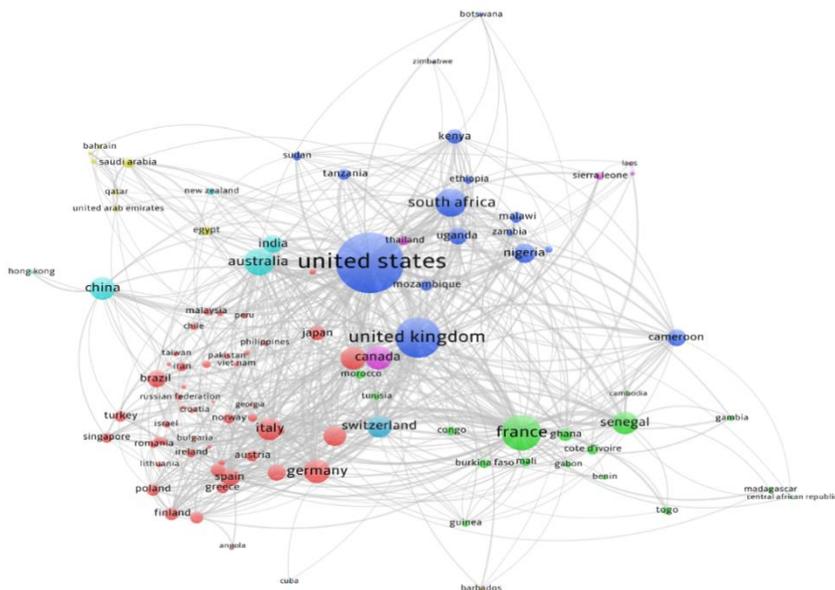


Figure 10 : Cartographie des co-publications internationales

4.3 Discussion relative à l'impact scientifique des publications

L'impact scientifique des publications sénégalaises sur les MNT a été mesuré à partir des indices de citation (h-index) des revues dans lesquelles les chercheurs ont publié. A l'observation des résultats, nous notons un impact scientifique faible avec 5957 citations, soit un ratio d'environ 8 citations par articles. Cela semble très peu par rapport à la période considérée (10 ans) d'autant plus que nous notons une croissance considérable des publications sur les MNT depuis 2006 ; et que selon des experts en bibliométrie : « *il est établi que les citations les plus nombreuses se font pendant les trois à quatre premières années* » (documentation française, 2003). Cette situation s'expliquerait par une abondante publication dans des revues à très faible facteur d'impact. En clair, sur les trente premières revues qui comptent le plus de publications des chercheurs sénégalais (au moins 4 articles), seul 6 font partie des revues ayant un h-index supérieur à 35. De plus, seulement 56% des publications ont été cités une fois.

4.4 Discussion relative à la faible proportion de publications enLibre accès

La quantité insignifiante de publication dans des revues en accès libre peut se comprendre par la quête de revues à fort facteur d'impact. Selon l'assertion générale « gratuité ne rime pas avec qualité », autrement dit, une revue libre d'accès ne serait pas « fiable » et donc n'offrirait pas une grande valorisation ni un impact scientifique significatif. Même si ces revues favorisent la diffusion et l'accessibilité aux publications, elles n'ont pas un fort facteur d'impact, contrairement aux revues payantes qui elles sont éditées et indexées par les bases de données des éditeurs et par conséquent pris en compte dans le calcul des facteurs d'impact. Les chercheurs sénégalais pour leur carrière sont évalués sur l'impact scientifique de leurs publications, mesuré à partir du facteur d'impact des revues dans lesquelles ils ont publié. Ainsi, les revues indexées, qui se trouvent être payantes dans leur large majorité sont tout naturellement leur premier choix.

5 Conclusion

Cette étude est pionnière de l'analyse bibliométrique de la production scientifique sur les MNT au Sénégal. La base de données choisie est Scopus. Les indicateurs recherchés étaient la production annuelle depuis les indépendances (1960), l'indice de collaboration nationale et internationale mesurée à partir des co-publications institutionnelles, l'indicateur d'impact scientifique obtenu à partir du nombre de citations et du H-Index (SJR) des revues. Aussi, la proportion des publications en libre accès a été évaluée et analysée comparativement à celle en accès payant sous l'angle de l'impact scientifique.

Les résultats obtenus montrent que sur la période indiquée, la production sénégalaise sur les MNT est très faible avec seulement 808 publications parues dans 305 revues. Deux raisons essentielles à savoir, l'insuffisance de financement et la jeunesse des programmes de formation dans le domaine justifient cet état des faits. Toutefois, au cours de ces dix dernières années avec une nette progression d'environ 60 publications annuelles contre moins de 20 avant, le Sénégal demeure le leader scientifique dans le domaine des MNT en Afrique de l'Ouest francophone.

En outre, au niveau de la collaboration, l'étude a fait ressortir qu'elle est plus orientée internationale que nationale. Pour la collaboration nationale, elle est plus forte entre l'université Cheikh Anta Diop de Dakar et les CHU de la capitale. La collaboration internationale est quant à elle pour des raisons historiques dominée par la France.

Par ailleurs, la production sénégalaise est d'une insignifiante influence scientifique avec 5957 citations, soit un ratio d'environ 8 citations par articles. L'explication tient à l'index-h relativement faible des revues dans lesquelles les chercheurs sénégalais ont publié.

Enfin, la proportion de publications dans des revues en libre accès pour ces dix dernières années est très faible avec seulement environ 10% des articles (soit 34/353). Ce choix clairement orienté vers les revues payantes est mû par la quête de revues à fort facteur d'impact, garantie d'une valorisation professionnelle.

Cette étude malgré sa robustesse comporte quelques limites du point de vue essentiellement méthodologique. En effet, certains articles publiés en 2018 peuvent ne pas avoir encore été indexés par Scopus au moment de notre recherche. Un autre biais possible fut la limitation de notre corpus aux articles scientifiques, excluant de fait une probable riche production originale sur les MNT, notamment les chapitres d'ouvrages et les monographies. De plus, sachant que les corpus sont différents d'une base de données à une autre, nous ne pouvons prétendre à une recension et une analyse exhaustive de la production sénégalaise sur les MNT. Aussi, un référencement erroné de « l'affiliation institutionnelle » des auteurs, ou du « descripteur géographique » pourrait influencer négativement sur la constitution du corpus.

Au total, cette étude a dressé un panorama de la production scientifique sénégalaise sur les MNT. Elle a montré l'évolution de la recherche dans ce domaine, les institutions les plus productives et l'impact scientifique des publications. Les résultats, nous l'espérons permettront aux décideurs politiques, aux autorités universitaires et aux gestionnaires de programmes de recherche de mieux comprendre la progression de la recherche dans le domaine et les acteurs principaux, afin de développer un programme de recherche en adéquation avec les enjeux actuels de la recherche sur les MNT. Toutefois, même si cette étude permet de dégager une tendance générale sur les MNT, elle ne propose pas une analyse plus fine qui révélerait les thématiques les plus abordées pour chacune des MNT.

Bibliographie

- BOUDRY, Christophe ; DURAND-BARTHEZ, Manuel. (2017). Publications en libre accès en biologie-médecine: historique et état des lieux en 2016. *Ethics, Medicine and Public Health*, vol. 3, no 1, p. 169-181.

- DIAKHATE, Djibril. (2014). Cartographie des connaissances pour une visibilité scientifique numérique des universités africaines: le cas de l'Université Cheikh Anta Diop de Dakar (Ucad). In: *Colloque international sur le Libre Accès (CILA14)-International Colloque On Open Access (ICOA14)*.
- FARCHY, Joëlle, et FROISSART Pascal. 2010. « Le marché de l'édition scientifique, entre accès « propriétaire » et accès « libre », The scientific publishing market, from "proprietary" to "open" accès ». *Hermès, La Revue*, n° 57: 137-50.
- HIRSCH, Jorge E. 2005. « An index to quantify an individual's scientific research output ». *Proceedings of the National academy of Sciences of the United States of America* 102 (46): 16569–16572.
- LRHOUL, Hanae ; CHARTRON, Ghislaine ; BACHR, Ahmed & Benammar, Othman. (2015). La datavisualisation comme outil de pilotage de la recherche scientifique médicale au sein de la Faculté de Médecine et de Pharmacie de Casablanca. Dans *Big Data - Open Data : Quelles valeurs ? Quels enjeux : Actes du colloque « Document numérique et société »*, Rabat, 2015 (pp. 165-181). Louvain-la-Neuve, Belgique : De Boeck Supérieur.
- MÊGNIGBÊTO, Eustache. (2013). Scientific publishing in West Africa: Comparing Benin with Ghana and Senegal. *Scientometrics*, vol. 95, no 3, p. 1113-1139.
- NEES VAN, Eck, et WALTMAN Ludo. 2009. « Software survey: VOSviewer, a computer program for bibliometric mapping ». *Scientometrics* 84 (2): 523–538.
- OKUBO, Yoshiko. 1997. « Indicateurs bibliométriques et analyse des systèmes de recherche: méthodes et exemples ». *OECD Science, Technology and Industry Working Papers*.
- OLLE, Persson; DANELL Rickard; et SCHNEIDER J. Wiborg. 2009. « How to use Bibexcel for various types of bibliometric analysis ». *Celebrating scholarly communication studies: A Festschrift for Olle Persson at his 60th Birthday*, 9–24.
- Organisation Mondiale de la Santé (2014). Rapport sur la situation mondiale des maladies non transmissibles 2014. [En ligne]. Disponible à : www.who.int/nmh/publications/ncd-status-report-2014/fr (Page consultée le 30 avril 2018)

Organisation Mondiale de la Santé (2016). Ce que les ministères doivent savoir à propos des maladies non transmissibles : notes sectorielles.

POTTIER, Pierre ; BUFFENOIR, Kevin ; CASTILLO, Jean-Marie ; ETIENNE Jérôme ; RICHARD Isabelle (2015). La recherche en éducation médicale dans les facultés de médecine en France. Etat des lieux à partir d'une revue systématique des travaux publiés. *Pédagogie Médicale*, vol. 16, no 2, p. 91-103.

ROSTAING, Hervé. (1996). *La bibliométrie et ses techniques*. Sciences de la Société. [En ligne]. Disponible à:http://crrm.u3mrs.fr/web/IMG/pdf/La_bibliometrie_et_ses_techniques.pdf. (Page consulté le 17 juillet 2018)

SCHÖPFEL, Joachim, et PROST Hélène. 2009. « Le JCR facteur d'impact (IF) et le SCImago Journal Rank Indicator (SJR) des revues françaises: une étude comparative ». *Psychologie française* 54 (4): 287–305.

SEGLÉN, Per O. 1992. « The skewness of science ». *Journal of the American society for information science* 43 (9): 628–638.

SUBER, Peter. 2012. « Open access overview ».

THONON, Frédérique ; SAGHATCHIAN, Mahasti ; NERFIE, Alexia ; DELALOGÉ, Suzette. (2015). Tendances et évolutions de la recherche française sur le cancer du sein : étude bibliométrique. *Bulletin du Cancer*, 102(5), 417-427.

Union Africaine (2013). État de l'hypertension en Afrique. *Conférence des ministres de la santé de l'UA*

YAYA, Sanni et KENGNE, André. (2014). L'essor des maladies non-transmissibles dans les pays en développement: nouveau destin, nouvelles priorités.

Publication scientifique en accès libre : Opinions et habitudes des chercheurs

Scientific publication in open access: Opinions and habits of researchers

Noureddine Elmqaddem

Laboratoire Mediation-Information-Knowledge-Society (MIKS)
ESI, Rabat – Morocco - nelmqaddem@esi.ac.ma

Résumé

En nous basant sur une enquête par questionnaire, nous avons étudié les opinions et les habitudes des chercheurs quant à la publication dans des revues en accès libre ou le dépôt dans des archives ouvertes. Il s'agissait plus particulièrement de démystifier le libre accès à l'information scientifique et mettre au clair ses forces et ses faiblesses. Contrairement à nos attentes, la majorité des enseignants enquêtés se sont montrés familiers avec ce concept, ses principes et ses implications pour l'auteur et le lecteur, mais aussi pour la communauté scientifique toute entière. Ceci est vrai malgré une faible contribution à la publication en accès libre et à l'auto-archivage due principalement à quelques incertitudes et fausses idées à l'égard de ces deux modes de publication. Cette recherche nous a permis de confirmer que l'implication de plusieurs personnes, incluant les leaders politiques, les institutions, les bibliothécaires ainsi que les chercheurs et les utilisateurs finaux, est nécessaire pour promouvoir le libre accès de façon éclairée tout en répondant aux principaux points de résistance des auteurs potentiels.

Mots-clés. Libre accès ; Accès ouvert ; Publication scientifique ; Archives ouvertes ; auto-archivage.

Abstract

With the help of a questionnaire survey, we studied the opinions and habits of researchers about publishing in open access journals or self-archiving. More specifically, we tried to demystify open access to scientific information and to clarify its strengths and weaknesses. Contrary to our expectations, the majority of teachers surveyed were familiar with this concept, its principles and its implications for the author and the reader, but also for the entire scientific community. This is true despite a weak contribution to open access publishing and self-archiving due mainly to some uncertainties and misconceptions about these two modes of publication.

This research allowed us to confirm that the involvement of many people, including political leaders, institutions, librarians as well as researchers and end-users, is necessary to promote open access in an informed way while responding to key resistance factors of potential authors.

Keywords. Open access; OA; Scientific publishing; Open archives; Self-archiving.

1 Introduction

Selon l'Initiative de Budapest pour l'Accès Ouvert (2004), « par "accès libre" à cette littérature, nous entendons sa mise à disposition gratuite sur l'Internet public, permettant à tout un chacun de lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces articles, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale, sans barrière financière, légale ou technique autre que celles indissociables de l'accès et l'utilisation d'Internet. »

Par ailleurs, et comme le note Eggis(2018), outre l'intérêt de partage (de l'information et la connaissance), le libre accès permet de répondre à une situation économique inédite qui a vu le jour à partir des années 1990 : *l'accroissement très important du coût des abonnements* (augmentation de 350 % entre 1986 et 2008), l'augmentation du nombre de revues et *la concentration des éditeurs*, lesquels se trouvent en position de force pour négocier avec les bibliothèques.

La prolifération des technologies de l'information et de la communication, notamment Internet (et ses différents services : FTP, Group news, messagerie électronique, Web...) a alors aidé à l'émergence du mouvement du libre accès à l'information scientifique.

Si la publication dans des revues traditionnelles (payantes) limite l'accessibilité aux lecteurs qui ont un abonnement direct ou à travers une bibliothèque qui paie pour un abonnement institutionnel auquel les membres ont accès, la publication dans des revues à accès libre ou l'auto-archivage permettent l'accessibilité gratuite à tout le monde.

La gratuité d'accessibilité est souvent considérée comme étant la différence principale entre les deux modes de publication, libre ou payante. Néanmoins, les choses ne s'arrêtent pas là et plusieurs facteurs doivent être pris en compte pour déterminer les frontières entre les deux types de publication et d'accès aux résultats de la recherche scientifique.

Pour les partisans du mouvement de l'accès ouvert, rendre les articles scientifiques accessibles au public et s'éloigner des modèles d'édition traditionnels à but lucratif permettront de faire progresser les découvertes médicales et scientifiques et bénéficieront ainsi à la science et à la société en général. Mais les critiques contestent si l'accès ouvert est économiquement durable et si la publication en libre accès sans examen par les pairs, ou avec un examen limité, fait réellement avancer la science. En outre, dans les cas où les revues ne facturent pas l'accès à leurs articles évalués par un comité éditorial, comment les éditeurs couvrent-ils les coûts de l'examen par les pairs et de la publication, et cela affecte-il la qualité de la recherche publiée ? Et puis, quels sont les droits et obligations de l'auteur dans tout ceci ?

Nous allons alors mener une enquête par questionnaire auprès des enseignants de l'école des sciences de l'information pour vérifier quelles sont leurs opinions et leurs habitudes à l'égard de la publication en accès libre. Les résultats de cette enquête devront nous aider à mieux situer l'accès libre par rapport à l'édition classique, déterminer ses avantages et ses défauts et corriger les fausses-idées qu'on peut avoir à son égard.

2 Problématique

Malgré une croissance significative du nombre de documents de recherche disponibles en accès libre, principalement par l'auto-archivage des auteurs dans les archives institutionnelles, le pourcentage du nombre d'articles publiés chaque année en libre accès est encore bas et ceci est lié directement à l'implication des personnes concernées par la recherche scientifique. Plusieurs enseignants chercheurs ne connaissent pas encore bien les principes de l'accès libre, ses raisons d'être, les droits et obligations qu'il donne à l'auteur qui publie un article en accès libre, et aux lecteurs de cet article et puis les avantages à court et à long terme.

Ainsi, nous avons mené une recherche à travers laquelle nous essayons de démystifier ce concept et corriger les fausses idées qu'on peut avoir à l'égard de ce mode de publication scientifique.

3 Objectifs

A travers l'étude des opinions et des habitudes des enseignants chercheurs à l'égard du libre accès à l'information scientifique, nous cherchons à montrer les avantages et les inconvénients de la publication dans des revues à accès libre ou l'auto-archivage et à corriger quelques fausses idées à propos de ces deux modes de publication. Des clichés à la réalité, ce travail de recherche donnera quelques clarifications que nous espérons aideront les personnes concernées à y voir plus clairement.

4 Méthode de recherche

Pour notre recherche, nous avons mené une enquête de terrain basée sur un questionnaire distribué en ligne. Nous avons pris comme population de l'enquête les enseignants permanents de l'école des sciences de l'information (ESI).

5 Analyse des résultats de l'enquête

Bien que nous ayons fait en sorte à ce que le questionnaire soit facile à remplir et en avons informé les personnes enquêtées, seulement 14 enseignants ont participé à cette enquête. En tout cas, la réticence à répondre au questionnaire est en elle-même un indicateur intéressant pour notre recherche. Nous pouvons interpréter ceci par le fait que les gens sont en général mal à l'aise quand il s'agit de parler de l'Open Access. Même si tout le monde en a entendu parler, il y en a qui n'ont pas une connaissance approfondie de ce mode de publication.

Les professeurs qui ont répondu au questionnaire se composent majoritairement de femmes (78,6%) et appartiennent principalement à la tranche d'âge 36-55 (70,14%) suivie de la tranche 46-55. 64,3% ont entre 5 et 15 années d'ancienneté dans l'enseignement et la recherche.

Nous remarquons que les enseignants qui ont moins de 5 années d'ancienneté n'ont pas répondu au questionnaire. Ceci peut être dû au fait que cette catégorie appartient à une nouvelle génération qui s'intéresse moins à l'accès libre ou tout simplement ne se sent pas à l'aise quand il s'agit de remplir un questionnaire surtout lorsqu'il porte sur un sujet non

familier pour eux/elles. Et puis, les membres de cette catégorie appartiennent principalement au domaine de l'informatique et des technologies de l'information qui est représenté par seulement 28,6% des répondants. En fait, ce sont les enseignants qui appartiennent au domaine des sciences de l'information qui ont répondu le plus au questionnaire (64,3%) et à peine 7,1% pour les autres domaines. Le domaine des langues et communication marque 0%. Il paraît alors que le libre accès intéresse plus les spécialistes de l'information que les autres.

Comme le montre le graphique ci-dessous, presque tous les professeurs enquêtés connaissent la forme de publication dite « voie verte », plus connue par l'appellation auto-archivage ou dépôt par l'auteur dans des « archives ouvertes », ainsi que la voie dorée, connue par « publication dans des revues en accès libre » quel que soit leur mode de financement.

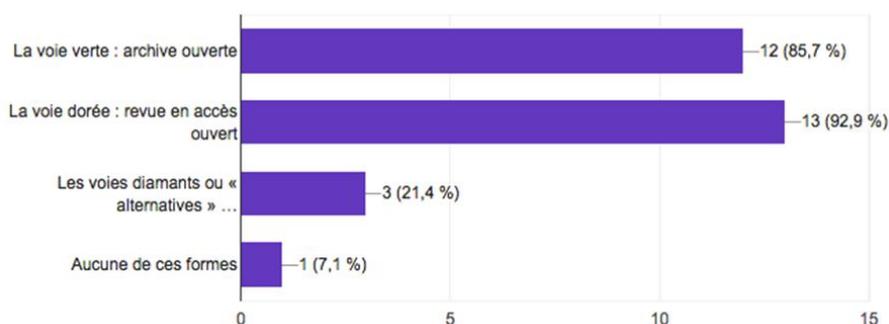


Figure 1 : Cochez la case ou les cases de la forme d'accès libre que vous connaissez.

Mais seulement 28,6% ont déjà publié dans une revue en accès libre :

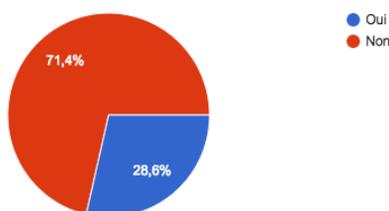


Figure 2 : Avez-vous déjà publié dans une revue en libre accès ?

Et seulement 15,4% ont fait un dépôt dans une archive ouverte :

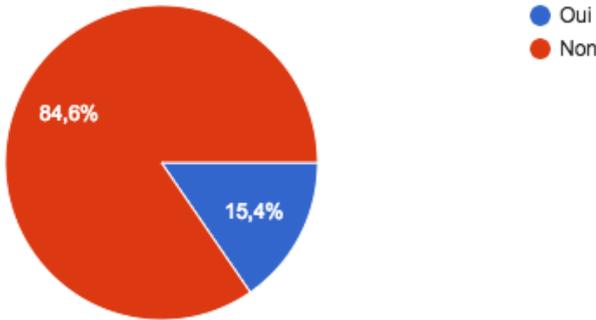


Figure 3 : Avez-vous déjà fait un dépôt dans des archives ouvertes ?

Pour la question concernant les modes connus de l'accès libre, presque tous les répondants sont familiers avec les archives ouvertes (dites voie verte) et les revues en accès libre (dites voie dorée). Alors que pour la question concernant la raison de la faible publication dans des revues en libre accès ou de l'auto-archivage, la réponse est principalement la non-connaissance de ces modes de publication. Ceci paraît contradictoire, mais il ne l'est pas. En fait, il y a plusieurs niveaux de connaissance d'un phénomène quelconque, qui varient de l'expérimentation personnelle et l'implication directe dans ce phénomène au simple fait d'en avoir entendu parler. Ceci dit, ces répondants considèrent comme étant une sorte de familiarité avec ces deux modes de publication libre le simple fait de savoir qu'ils existent ou d'en avoir entendu parler.

Cette explication est appuyée par le fait que 78,6% n'ont aucune idée de ce que sont les APC (Article Processing Charges) qui constituent l'un des moyens de financement les plus importants de la voie dorée. Et 64,3% ne connaissent pas les journaux prédateurs ou les revues prédatrices ; d'ailleurs, 92,9% n'ont jamais eu affaire à un journal prédateur.

Les réponses concernant les droits et frais de l'accès libre montrent que la majorité des répondants distinguent bien les implications de ce mode de publication. En fait, 71,4% savent que l'accès libre n'implique pas la gratuité à tous les niveaux. Pour 69,2%, la publication n'est pas gratuite. Et pour 85,7%, l'accès est gratuit. Cependant, 64,3% pensent que l'auteur ne paye pas s'il publie dans une revue en accès libre. Ceci est en partie vrai, car la voie dorée comprend différents modèles de financement, comme le modèle auteur-payeur (APC), le modèle sponsor et le modèle freemium. Donc, l'auteur ne paye pas toujours et automatiquement et même dans le cas du modèle auteur-payeur, certains éditeurs proposent des adhésions institutionnelles pour pouvoir bénéficier de remises sur les coûts des APC, d'autres des adhésions individuelles à vie permettant de publier un ou

plusieurs articles par an, qui se substituent alors aux APC.¹ L'université ou l'institution à laquelle appartient l'auteur peut prendre en charge les APC, et des facilités pouvant aller jusqu'à l'exonération complète des APC sont proposées parfois par quelques revues si l'auteur ne peut pas les payer, surtout s'il appartient à un pays en voie de développement.

Mais, bien que le modèle auteur-payeur ne soit pas le seul moyen qui permet à l'éditeur en accès libre découvrir les coûts liés à l'évaluation par les pairs et le comité éditorial et que la majorité des revues en libre accès n'exigent pas le paiement des APC, comme nous allons le montrer plus loin, le principe de l'auteur-payeur constitue l'un des facteurs qui découragent les auteurs à publier dans des revues en libre accès.

Seulement 28,6% des répondants savent que l'accès libre est un modèle économique qui permet la valorisation commerciale. Comme nous venons de l'expliquer, il y a des coûts liés à l'évaluation par les pairs et le personnel éditorial que l'éditeur en libre accès doit couvrir. Pour cela, rien n'empêche de chercher des moyens de financement en s'appuyant sur diverses méthodes de marketing et de valorisation commerciale. Mais ceci ne doit pas se contrarier avec les principes de l'accès libre.

Globalement, les répondants ont montré une bonne connaissance des droits et obligations relevant de la publication en libre accès. D'ailleurs, 78,6% savent que dans le cas d'un travail de recherche publié en accès libre, l'utilisateur doit citer la source.

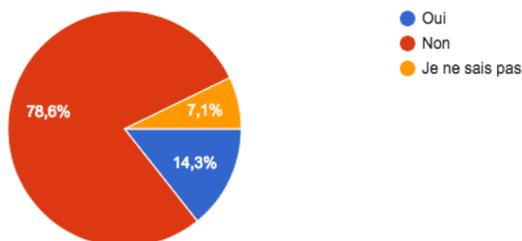


Figure 4: Dans le cas d'un travail de recherche publié en libre accès, l'utilisateur n'a pas besoin de citer la source.

Concernant la qualité, pour 64,3%, la publication en libre accès n'implique pas une mauvaise qualité.

¹Voir : <https://openaccess.couperin.org/modeles-eco-du-gold/>

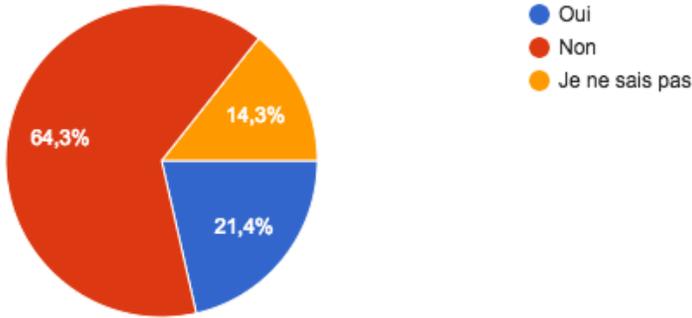


Figure 5 : Les articles publiés en accès libre ne sont pas de bonne qualité.

Et pour 85,7%, l'accès libre ne favorise pas le plagiat.

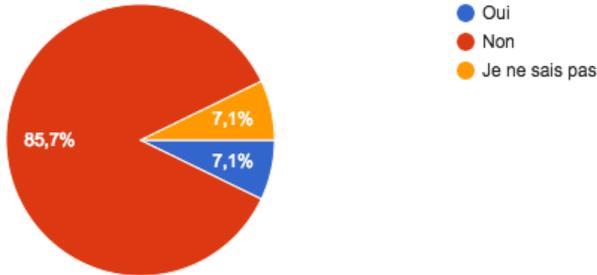


Figure 6 : L'accès libre favorise le plagiat.

Et puis, selon 78,6%, le libre accès ne nous dirige pas vers la décadence de la recherche scientifique.

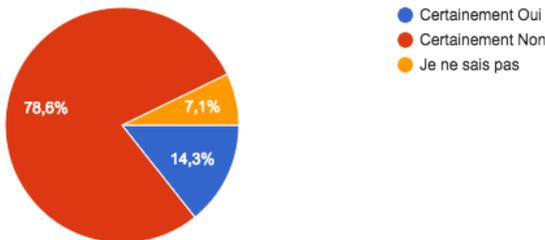


Figure 7 : Le libre accès nous dirige vers la décadence de la recherche scientifique

Pour 71,4%, le libre accès n'entraînera pas la fin de l'édition et des comités de lecture.

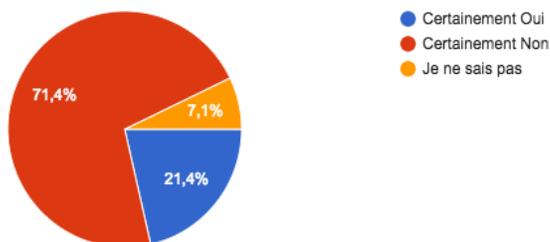


Figure 8 : Le libre accès entraînera la fin de l'édition et des comités de lecture

Pour les questions qui concernent l'impact et la visibilité, 64,3% savent que les revues en accès libre ont bien un facteur d'impact tout comme les revues classiques.

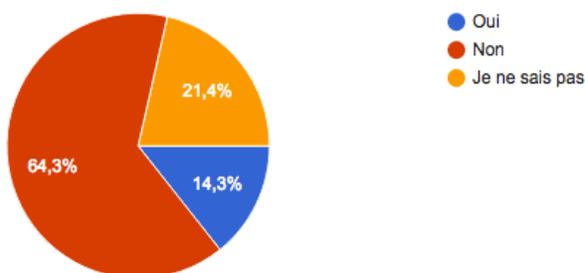


Figure 9 : Les revue en accès libre n'ont pas de facteur d'impact

Et pour 78,6%, publier en accès libre augmente le nombre de citations.

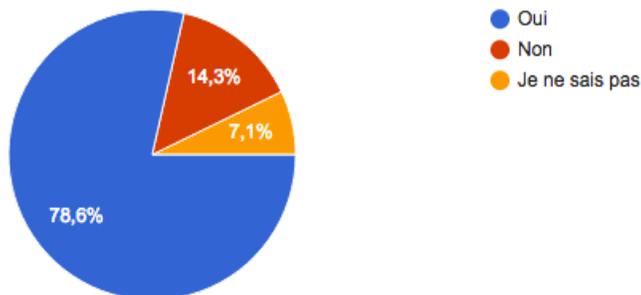


Figure 10 : Publier en accès libre augmente le nombre de citations.

Seulement 21,4% considèrent que l'accès libre a un intérêt bas pour eux/elles

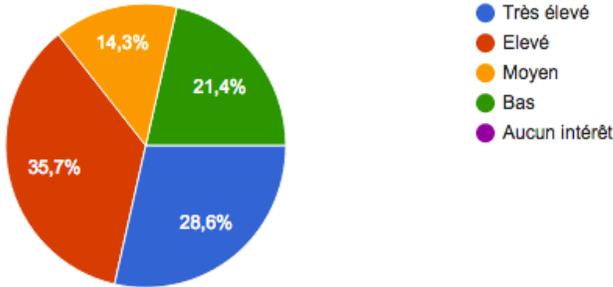


Figure 11 : Le degré d'intérêt du libre accès pour moi

Et à peine 28,6% considèrent que les résultats de la recherche scientifique ne sont pas intéressants pour le grand public

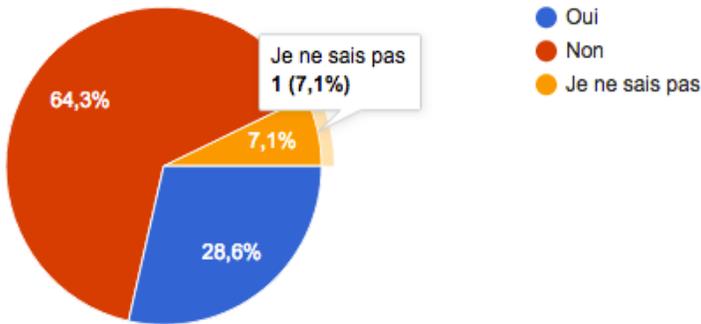


Figure 12 : Les résultats de la recherche scientifique n'ont pas d'intérêt pour le grand public, car c'est trop spécialisé

Pour 92,9%, l'accès libre est un très bon moyen pour faciliter aux lecteurs l'accès gratuit aux résultats de la recherche scientifique.

Et pour 71,4%, il est profitable à tout le monde, aussi bien aux chercheurs des pays développés, qu'aux chercheurs des pays en voie de développement.

Ceci est vrai, quoique le manque de moyens qui permettraient aux chercheurs dans les pays pauvres d'accéder aux publications des revues payantes fait que ces derniers profitent plus de l'accès libre. Alors que pour les pays riches, bien que l'abonnement dans des revues payantes est parfois très élevé, les institutions, universités et centres de recherche de ces pays font toujours le nécessaire pour garantir à leurs chercheurs l'accessibilité aux articles scientifiques publiés dans ces revues.

6 Discussion et synthèse

D'après les résultats de notre enquête, les enseignants de l'ESI s'intéressent bien au libre accès à l'information scientifique et se sont montrés familiers avec ce concept, malgré la faiblesse de leur participation à ce mode de publication. Les raisons de cette réticence ne sont pas claires, mais peuvent être liées au manque de clarté quant aux droits et obligations de l'auteur qui publie en libre accès. En tout cas, il paraît logique que le concept de l'accès libre soit familier aux spécialistes de l'information. Selon Chartron&Schöpfel, les Sciences de l'Information et de la Communication « sont doublement concernées par l'Open access, en tant que communauté scientifique et en tant que discipline intéressée par la production et la circulation de l'information au sein des structures et organisations scientifiques »(Chartron&Schöpfel, 2017)

Pour encourager tout le monde à publier encore plus en libre accès, l'implication de plusieurs personnes est nécessaire. Bibliothécaires, administration, professeurs et tout autre spécialiste de l'information, tous doivent participer pour rendre accessible à tout le monde une information sur mesure répondant directement à leurs préoccupations. Ceci devrait alors avoir une grande influence sur le futur, à court et à long terme, du libre accès à l'information scientifique au Maroc et dans le monde.

Dans ce qui suit, nous allons discuter quelques points qui ont fait objet de notre enquête et qui portent sur les frais de publication, le contrôle et la qualité du contenu publié, les droits d'auteurs, la rapidité de publication, l'accessibilité et puis la lisibilité des articles publiés en libre accès.

6.1 Frais de publication

Comme nous l'avons noté plus haut, les frais requis pour publier dans des revues en libre accès sont parfois très élevés et souvent légitimés par la couverture des frais d'édition qui aident l'éditeur à couvrir les coûts liés à l'évaluation par les pairs et le personnel éditorial.

Il est à noter que lorsqu'il s'agit de parler du modèle économique basé sur le financement par APC, il y a deux extrêmes : d'un côté, les principaux éditeurs commerciaux qui réclament des dépenses d'environ 3 000 à 5 000 USD par article publié (Morris, 2005). De l'autre côté, des universitaires engagés dans le mouvement de l'OA (Odlyzko, 1997) qui proposent que les revues puissent fonctionner avec des budgets très bas et être « gratuits » aux deux extrémités, étant donné que les académiciens peuvent remplir presque toutes les fonctions nécessaires comme étant une partie de leurs devoirs universitaires de toute façon, sans compensation monétaire supplémentaire. (Björk, Shen, & Laakso, 2016)

Le modèle économique d'APC ou auteur-payeur a grandi rapidement depuis que BioMedCentral (BMC) et le PLoS² l'ont mis en place dans les années 2000. (Claudio, 2017)

² Public Library of Science

Plusieurs auteurs payent les APC même si la somme d'argent et parfois très élevée (allant jusqu'à 30000Dh), car ils sont conscients des avantages de ce mode de publication pour eux, pour les lecteurs et pour la communauté scientifique en général. Toutefois, il faut savoir que la majorité des revues en libre accès n'exigent pas les APC. Le DOAJ³ recense aujourd'hui près de 11928 revues parmi lesquelles 8710 (73%) n'exigent pas le paiement des APC. En fin de compte, il s'agit plus d'une fausse idée que d'une réalité.

The screenshot shows the DOAJ website interface. At the top, there is a navigation bar with 'Home', 'Search', 'Browse Subjects', 'Apply', 'News', 'About', 'For Publishers', and 'API'. A 'SUPPORT DOAJ' button is on the right. Below the navigation bar, there are search filters: 'share | embed', '10', 'order by ... relevance', and 'search all'. A sidebar on the left has sections for 'Journals vs Articles', 'Subject', and 'Article processing charges (APCs)'. The main content area displays a search result for 'Global Justice: Theory, Practice, Rhetoric' with details like ISSN, URL, peer review status, subject, date added, and last updated. A 'License: Publisher's own license' badge is visible on the right.

Figure 13 : 8710 revues recensées par DOAJ n'exigent pas le paiement des APC

Et puis, comme le précise l'Initiative de Budapest pour l'Accès Ouvert (2004): « puisque le prix constitue un obstacle à l'accès, ces nouvelles revues ne factureront pas l'abonnement ou l'accès, et se tourneront vers d'autres méthodes pour couvrir leurs frais. Il existe, pour cette fin, de nombreuses sources de financement alternatives, parmi lesquelles les institutions et les gouvernements qui financent la recherche, les universités et laboratoires qui emploient les chercheurs, les dotations allouées par discipline ou par institution, les amis de la cause du libre accès, les profits générés par la vente d'enrichissements apportés aux textes de base, les fonds libérés par la transformation ou la disparition des périodiques facturant un abonnement traditionnel ou un prix d'accès, voire les contributions des chercheurs eux-mêmes. Nul besoin de favoriser une solution plutôt qu'une autre pour toutes les disciplines et toutes les nations, ni d'arrêter de chercher de nouvelles alternatives originales. »

6.2 Qualité du contenu publié et droits d'auteurs

On associe parfois l'accès libre à l'accès non-contrôlé et donc de mauvaise qualité, ce qui n'est pas toujours vrai. Tout comme dans le cas des revues

³Le Directory of Open Access Journals répertorie toutes les revues en libre accès

traditionnelles, la publication dans des revues à accès libre nécessite la validation d'un comité éditorial constitué généralement de spécialistes du domaine, notamment des universitaires.

L'accès libre est associé parfois aussi au plagiat. On pense que publier un contenu en libre accès veut dire donner aux autres le droit de réutiliser ce contenu sans citer l'auteur propriétaire. En fait, on pense que l'accès libre ne permet pas à l'auteur d'un contenu de préserver ses droits sur ce contenu. Ceci n'est pas vrai non plus. Même quand on publie en libre accès, on garde toujours nos droits, notamment le droit d'être cité et reconnu comme auteur de ce contenu.

Par ailleurs, le fait de rendre un contenu facilement accessible aux autres est en soi un moyen de protéger ce contenu contre le plagiat, qui devient facile à détecter.

Ceci dit, même dans les publications en libre accès, on doit citer la source. C'est une sorte de reconnaissance des droits d'auteur.

Selon l'étude de Bourrion et al.(2006), les deux réticences les plus significatives vis-à-vis du dépôt en accès ouvert sont la crainte d'une moindre garantie des droits d'auteur et celle d'un plagiat. La première crainte trahit d'abord la méconnaissance du fait que l'accès ouvert conserve tous ses droits à l'auteur. Mais c'est aussi une étrange négation de la spoliation de ces droits opérée par les éditeurs traditionnels ; il est vrai que ces éditeurs ont coutume de défendre leurs intérêts financiers au nom du droit d'auteur ! La crainte d'un plagiat n'est pas davantage fondée. A la différence des éditeurs traditionnels, les revues en accès ouvert et les archives ouvertes enregistrent automatiquement, à la seconde près, la date de réception du document, avant de le mettre en ligne. C'est donc dès la réception de son premier envoi (la prépublication) que l'auteur peut faire valoir l'antériorité de son travail.

6.3 Lisibilité des articles publiés

Une autre fausse idée est liée à la lisibilité des articles publiés dans des revues en libre accès. En fait, on pense souvent que le facteur d'impact des revues à accès libre est bas par rapport à des revues à accès payant. Ceci n'est pas vrai. Même les revues à accès libre peuvent avoir un facteur d'impact élevé. D'ailleurs, ceci est consolidé par le principe même de liberté d'accès. Si le facteur d'impact est lié à la renommée d'une revue, alors les revues en accès libre ont plus de chance d'avoir un facteur plus élevé car leur liberté d'accès leur permet une popularité plus facile. L'accessibilité des articles publiés dans des revues à accès libre encourage les auteurs à lire et réutiliser ces articles dans leurs travaux de recherche ce qui augmente la citation de ces articles et par conséquent augmente la visibilité et le facteur d'impact de la revue. Le croisement des références est parmi divers moyens qui consolident ce principe de citation et de visibilité des articles publiés dans des revues scientifiques.

Comme l'explique Claudio (2017), il existe des preuves montrant que la grande visibilité offerte par le libre accès peut conduire à un nombre plus élevé de citations. Le nombre de citations est devenu de plus en plus important en tant que mesure de la production scientifique d'un

scientifique. Des mesures telles que l'indice h (h-index)⁴ sont supposées être des indicateurs de la productivité d'un scientifique et de l'impact de ses recherches. Le fait d'avoir des publications citées dans d'autres publications augmente le h-index, ce qui peut être important pour certains professionnels de la recherche et leurs institutions universitaires. Selon Hirsch(2005), l'indice h est le nombre de papiers avec un numéro de citation supérieur ou égal à h. Cet indice a pour but de quantifier la productivité scientifique et l'impact d'un scientifique en fonction du niveau de citation de ses publications.

6.4 Rapidité de publication et accessibilité de l'information scientifique

Parmi les raisons principales qui encouragent les chercheurs à publier dans des revues en libre accès, il y a la rapidité de publication et l'accessibilité de ces articles aux lecteurs. Pour attirer plus d'auteurs à publier dans leur revue ou journal à accès libre, les éditeurs font plus d'effort notamment au niveau du temps alloué à la révision des articles par le comité de lecture et à la publication des articles acceptés. Ceci est surtout le cas lorsque les APC sont pris en charge par l'auteur ou son institution. Cependant, dans le cas des revues classiques, le temps qu'il faut entre la soumission et la publication est parfois très long. Cette longue attente du papier est souvent légitimée par le fait que les examinateurs sont des gens occupés et qu'il leur faut beaucoup de temps pour soumettre leurs remarques. Ceci fait que le mode de publication en libre accès est très attrayant pour les auteurs qui cherchent à réduire le temps d'attente de la publication de leurs articles.

Bien qu'elle ne fournisse pas de délais exacts, une étude dans laquelle ont été examinées 135 revues répertoriées dans le Scopus Citation Index a montré que les journaux à accès ouvert avaient un délai plus court entre l'acceptation et la publication par rapport à l'édition traditionnelle. (Claudio, 2017)

Quant à l'accessibilité, c'est un objectif citoyen et noble. L'auteur qui publie dans une revue à accès libre est conscient de l'importance de ce geste qui rendra l'information scientifique accessible à tout le monde, quoique soient leurs moyens, sans être obligés d'avoir un abonnement dans une revue ou une bibliothèque, etc. A court terme, le bénéfice revient au lecteur, qui se voit capable d'accéder gratuitement à l'information scientifique qui l'intéresse, et à l'auteur, qui voit augmenter la visibilité de ses publications et aussi la reconnaissance morale et matérielle de sa contribution dans la société. Mais, et encore mieux, ceci est bénéfique pour la société scientifique entière, surtout à long terme. En fait, nous sommes dans une période de transition vers un monde où le système éditorial de la recherche sera transformé en un système ouvert sans barrières d'accès (Conseil européen, 2016). Tout le savoir humain sera accessible à tout le monde, partout, sans aucune contrainte.

⁴L'indice h (ou indice de Hirsch) est un indice suggéré en 2005 par Jorge Hirsch.

6.5 Inconvénients

Malgré les avantages incontestables du libre accès à l'information scientifique, nous admettons que ce mode de publication a quelques inconvénients, parmi lesquels on note le manque de contrôle de la qualité des articles publiés dans quelques revues en libre accès. Certes, ceci ne concerne pas les revues de renommée, mais d'autres qui, ayant besoin de couvrir les coûts d'édition, sont parfois poussées à publier un plus grand nombre d'articles au détriment de la qualité globale. Et puis, l'auto-archivage ne concerne pas toujours des articles validés par les pairs. Par ailleurs, la présence de plusieurs journaux prédateurs ou revues prédatrices décourage parfois les auteurs à faire appel à l'accès libre. Au lieu de passer son temps à vérifier si la revue est fiable ou non, l'auteur préfère publier dans une revue classique qu'il connaît bien.

L'autre inconvénient est lié aux frais de publication. Bien que le consommateur final n'ait pas à payer pour lire un article publié en libre accès, des frais de publication sont parfois exigés par l'éditeur pour couvrir les frais d'examen et d'édition des articles à publier. Et c'est souvent l'auteur - peut-être par son employeur ou une subvention de recherche - qui a la responsabilité de couvrir ces coûts (APC). Pour plusieurs auteurs, payer les APC est un frein de taille qui les décourage d'avoir recours au libre accès. Cependant, et comme nous l'avons précisé plus haut, la majorité des revues à accès libre n'exigent pas le paiement des APC. Prendre ce facteur comme raison pour éviter de publier en libre accès n'est pas bien fondé. Ceci relève plus de l'ignorance de tous les droits et toutes les obligations de ce mode de publication, mais reste une des raisons principales de la réticence à l'égard de la publication dans des revue à accès libre.

En fin, la soumission d'une contribution à un périodique scientifique à accès libre s'effectue selon les procédures habituelles. Mais elle présuppose bien sûr l'existence de telles revues dans le champ disciplinaire concerné. Dans le cas contraire, actuellement le plus fréquent, seul l'auto-archivage est possible. Néanmoins, Il est essentiel que l'auto-archivage soit effectué sur un site institutionnel garantissant la pérennité de l'archive.

7 Conclusion

L'enquête de terrain nous a permis d'avoir une idée sur les habitudes et les opinions des enseignants-chercheurs de l'école des sciences de l'information. Et ceci nous a aidés à élaborer nos interprétations et conclusions. Contrairement à nos attentes avant l'enquête, les résultats démontrent que le concept d'accès libre à l'information scientifique est familier pour la majorité des chercheurs. Malgré une sorte de réticence à l'égard de la publication dans des revues en libre accès ou l'auto-archivage, les répondants sont bien conscients des avantages de ce mode de publication et de ses retombées positives aussi bien pour eux en tant que chercheurs que pour les lecteurs en tant qu'utilisateurs finaux et pour la société en général. Nous avons essayé de corriger quelques fausses idées et mettre encore plus au clair les avantages et les inconvénients de la publication en libre accès. Nous avons aussi argumenté en faveur de la nécessité de contribution de tout le monde, leaders politiques, institutions, bibliothécaires et chercheurs, à la promotion de ce mode de publication dont les avantages à

long terme sont encore plus intéressants et palpables. En fait, il faut adopter une approche basée sur la persuasion et l'encouragement des auteurs. Il est à noter que nous sommes encore dans une période de transition pendant laquelle les retombées sont difficiles à mesurer et les frontières entre les forces et les faiblesses du libre accès (en comparaison avec la publication classique) encore floues. Toutefois, il est certain que nous allons dans la bonne direction, une direction qui peut cependant bifurquer vers de nouvelles voies donnant naissance à de nouvelles possibilités. C'est ce qui a eu lieu il y a déjà quelques années avec ce qu'on appelle « open peerreview ». En fait, comme le souligne Jacobs (2006), certaines revues ont commencé à expérimenter un examen par les pairs qui est ouvert. Les articles sont dans ce cas affichés sur le Web pour une discussion ouverte entre les groupes de pairs et l'auteur, ils sont modifiés à la lumière de cela et sont ensuite examinés officiellement par des évaluateurs sélectionnés de manière traditionnelle. Aux modes de publication actuels, de nouveaux s'ajouteront avec le temps, mais il est certain que la liberté d'accès à l'information aura toujours une place importante dans nos préoccupations.

Références

- Björk, B.-C., Shen, C., & Laakso, M. (2016). A longitudinal study of independent scholar-published open access journals. *PeerJ*, 4. <https://doi.org/10.7717/peerj.1990>
- Bourrion, D., Boutroy, J. L., Giordanengo, C., & Krajewski, P. (2006, juillet 18). Les chercheurs en Lettres et Sciences Humaines et les Archives Ouvertes. Consulté à l'adresse <https://halshs.archives-ouvertes.fr/halshs-00086084/document>
- Chartron, G., & Schöpfel, J. (2017). Open access et Open science en débat. *Revue française des sciences de l'information et de la communication*, (11). Consulté à l'adresse <http://journals.openedition.org/rfsic/3331>
- Claudio, L. (2017, mars 30). Pros and Cons of Open Access vs Traditional Publishing in Scientific Journals. Consulté 20 mai 2018, à l'adresse <https://www.linkedin.com/pulse/pros-cons-open-access-vs-traditional-publishing-journals-luz-claudio>
- Conseil européen. (2016). Conclusions du Conseil européen - 28 juin 2016 - Consilium. Consulté 20 juillet 2018, à l'adresse <http://www.consilium.europa.eu/fr/press/press-releases/2016/06/28/euco-conclusions/>
- Eggis, G. (2018). BiblioGuides: Libre accès et licences libres: Le Libre Accès à l'information scientifique. Consulté 21 avril 2018, à l'adresse [//paris-sorbonne.libguides.com/open-access/libre-acces-IST](http://paris-sorbonne.libguides.com/open-access/libre-acces-IST)
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569–16572. <https://doi.org/10.1073/pnas.0507655102>

Initiative de Budapest pour l'Accès Ouvert - Libre accès à l'information scientifique et technique. (2004). Consulté 27 mai 2018, à l'adresse <http://openaccess.inist.fr/?Initiative-de-Budapest-pour-l>

Jacobs, N. (2006). *Open Access: Key Strategic, Technical and Economic Aspects*. Elsevier.

Morris, S. (2005). The true costs of scholarly journal publishing. *Learned Publishing*, 18(2), 115-126. <https://doi.org/10.1087/0953151053584975>

Odlyzko, A. (1997). The Economics of Electronic Journals. *First Monday*, 2(8). <https://doi.org/10.5210/fm.v2i8.542>

Valoriser les publications d'un laboratoire universitaire dans l'environnement de la science ouverte. Retour d'expérience de la collection GERiiCO sur HAL

*Valuing the publications of a university laboratory
in the environment of open science
Feedback from the GERiiCO collection on HAL*

Joachim Schöpfel est enseignant-chercheur (MCF-HC) en Sciences de l'information et de la communication à l'Université de Lille et membre du laboratoire GERiiCO. Après avoir dirigé l'UFR IDIST de 2009 à 2012 et l'Atelier National de Reproduction des Thèses de 2012 à 2017, il travaille aujourd'hui comme consultant indépendant dans le domaine de l'information scientifique. Ses intérêts scientifiques : le libre accès à l'information, les données de la recherche, les thèses et mémoires, la science ouverte.

ORCID 0000-0002-4000-807X joachim.schopfel@univ-lille.fr

Hélène Prost est ingénieur d'études à l'Institut de l'Information Scientifique et Technique (INIST-CNRS) et membre associé au laboratoire GERiiCO de l'Université de Lille. Elle participe à différents projets de recherche relatifs à l'évaluation des collections, la fourniture de documents et l'analyse des usages, ainsi qu'à la littérature grise, les données de la recherche et le libre accès; elle est référente publications de la collection GERiiCO sur HAL. Elle est l'auteur de plusieurs publications.

ORCID 0000-0002-7982-2765 helene.prost@inist.fr

Amel Fraisse est Maître de conférences en Sciences de l'information et de la communication à l'Université de Lille et membre du laboratoire GERiiCO. Ses travaux de recherche s'inscrivent dans le domaine du Traitement Automatique de la Langue et plus particulièrement la collecte, le traitement et la diffusion d'information multilingues.

ORCID 0000-0002-8693-8862 amel.fraisse@univ-lille.fr

Stéphane Chaudiron est professeur en Sciences de l'information et de la communication à l'Université de Lille et directeur du laboratoire GERiiCO. Il est

co-directeur de la revue *Etudes de Communication*, président du comité scientifique de la revue *I2D - Information, Données, Document*, responsable pédagogique du parcours de Master VeCIS (Veille et Communication de l'Information Stratégique), membre du conseil d'UFR, membre de l'Association for information science and technology (ASIST), membre du chapitre français de l'International society for knowledgeorganization (ISKO) dont il a été président de 2005 à 2009, co-animateur du groupe de travail « Veille et analytique » du Groupement français de l'industrie de l'information » (GFII) et membre de l'Association des professionnels de l'information et de la documentation (ADBS). Il a été chargé de mission "politique éditoriale" auprès de la Présidence de l'Université de Lille 3 et membre du conseil d'administration de cette université.

stephane.chaudiron@univ-lille.fr

Résumé

La question de la diffusion des résultats de la recherche et, en particulier, le libre accès aux publications des chercheurs est au cœur de la politique pour la science ouverte. Comment peut se positionner un laboratoire de recherche universitaire ? Comment peut se traduire la politique pour la science ouverte sur le terrain d'un campus universitaire ? Sous forme d'un retour d'expérience, notre étude analyse la mise en place de la collection du laboratoire GERiCO de l'Université de Lille sur l'archive ouverte nationale HAL. L'objectif de l'initiative est double : d'une part, assurer une visibilité maximale et un impact au-delà de la communauté disciplinaire, à travers des médias sociaux et le référencement des moteurs de recherche ; d'autre part, contribuer à l'évaluation de la production scientifique du laboratoire.

Nous présentons les ressources mobilisées et les actions mises en oeuvre, analysons les résultats en termes de dépôts, d'usage et de services, et évoquons les facteurs de succès, les problèmes rencontrés et quelques perspectives pour le futur développement. En particulier, nous comparons le contenu de la collection HAL avec les résultats de la base de données scientométrique d'Elsevier (Scopus) et du moteur de recherche Google Scholar, et nous montrons le potentiel de la collection pour visualiser les relations au sein du laboratoire (analyse de réseaux) et son rayonnement international.

Mots clés. laboratoire universitaire, production scientifique, publication, archive ouverte, évaluation, valorisation, HAL, science ouverte

1 Introduction

Le 4 juillet 2018, lors de la conférence annuelle de la Ligue des Bibliothèques Européennes de Recherche à Lille, la Ministre de l'Enseignement Supérieur, de la Recherche et de l'Innovation, Frédérique Vidal, a annoncé un Plan national pour la science ouverte. L'ambition est de rendre "les résultats de la recherche scientifique ouverts à tous, sans entrave, sans délai, sans paiement." Le premier axe du Plan est de généraliser l'accès ouvert aux publications, de rendre obligatoire la publication en accès ouvert des articles et livres issus de recherches financées par appel d'offres sur fonds publics et de soutenir l'archive ouverte nationale HAL.

Lancée en 2002 par le Centre pour la Communication Scientifique Directe (CCSD) du CNRS, l'archive pluridisciplinaire HAL (= Hyper articles en ligne) est devenue au fil des ans l'une des plus importantes plateformes de la "voie verte" du libre accès à l'information scientifique. "Voie verte" veut dire: l'auto-archivage des publications scientifiques par les auteurs eux-mêmes, sur une plateforme dédiée (Harnad et al., 2004). C'est, avec les revues en libre accès ("voie dorée"), l'une des deux principales stratégies pour atteindre, selon les mots de la Ministre, "à terme 100% de publications scientifiques françaises en accès ouvert".

Cette stratégie de la "voie verte" s'appuie avant tout sur deux acteurs : l'auteur, dans la mesure où il détient les droits intellectuels pour effectuer le dépôt de ses propres publications (article, chapitre, communication, mémoire etc.), et son institution (organisme de recherche, université, école etc.), dans la mesure où elle a la possibilité d'inciter voire d'imposer l'auto-archivage (Thirion & Rentier, 2014) et où elle dispose également des ressources et de la légitimité pour une archive institutionnelle (Lynch, 2003). En revanche, un troisième acteur sur l'échiquier de l'Enseignement Supérieur et de la Recherche (ESR) est plus ou moins absent et peu visible dans cette stratégie : le laboratoire de recherche.

Or, les chercheurs de l'ESR français sont tous rattachés à des laboratoires de recherche, des unités constituées affiliées à une université et/ou à un ou plusieurs organisme(s) de recherche. Le laboratoire structure le cadre de travail des chercheurs ; les projets de recherche sont organisés autour des laboratoires, et c'est dans le cadre de leur laboratoire que les chercheurs sont évalués par le Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur (HCERES).

A partir de la collection du laboratoire GERiCO de l'Université de Lille, nous interrogeons le rôle potentiel d'un laboratoire universitaire pour une

stratégie de la “voie verte”, dans l’environnement d’une politique de science ouverte. Quel est l’intérêt pour un laboratoire, en termes d’impact et de visibilité ? Quels sont les facteurs-clés pour réussir une telle initiative, et quels sont les verrous ? Nous présentons un premier bilan, suivi de quelques perspectives pour un futur développement d’une telle collection.

2 Un nouvel écosystème

Comme un rapport récent de l’OCDE le souligne, ce terme de “science ouverte” a beaucoup de significations différentes ; dans un sens très large, il désigne tous les efforts pour rendre le processus scientifique plus ouvert et inclusif pour l’ensemble des acteurs concernés, tant au sein des communautés scientifiques, qu’à l’extérieur du milieu de la recherche (Dai et al. 2018). Malgré les différences d’approches et d’interprétations disciplinaires, il y a consensus qu’une telle politique de science ouverte impacte l’ensemble des acteurs et des dimensions de la recherche, tels que les mécanismes de financement, les modes d’évaluation (*peerreview*, indicateurs etc.), les infrastructures scientifiques, le transfert des connaissances, la propriété intellectuelle et la gestion des données de la recherche, mais aussi les nouveaux modèles de “crowdsourcing” et de la science citoyenne.

La question de la diffusion des résultats de la recherche et, en particulier, le libre accès aux publications des chercheurs est au coeur de la Science ouverte. Les enjeux sont connus : rendre les résultats scientifiques accessibles sans délais et sans restrictions, aux autres chercheurs, aux acteurs économiques et aux citoyens, augmenter l’impact et la visibilité des structures scientifiques, maîtriser les coûts de l’information scientifique. Les différentes options du libre accès sont également connues, dont notamment la publication dans des revues en libre accès (“voie dorée”) et la diffusion via des archives ouvertes (“voie verte”). Cette dernière option est facilitée par la Loi pour une République numérique de 2016 et la création d’un droit d’exploitation secondaire pour les chercheurs français (article 30, cf. CNRS-DIST 2016).

Comme son modèle américain, le service d’e-prints arXiv¹, l’archive ouverte HAL² a été conçue sur le principe de la communication directe entre chercheurs, pour faciliter et accélérer l’échange d’articles, avant même leur publication dans une revue. Sa mission est le dépôt et la diffusion non

¹<https://arxiv.org/>

²<https://hal.archives-ouvertes.fr/>

seulement d'articles scientifiques de niveau recherche, publiés ou non, mais aussi de thèses, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Avec l'évolution du mouvement du libre accès (cf. Suber, 2012) et des archives institutionnelles, puis, notamment après la signature à l'Académie des Sciences le 2 avril 2013, de la "Convention de partenariat en faveur des archives ouvertes et de la plateforme mutualisée HAL" entre les établissements et organismes de l'ESR français, HAL est devenue une sorte d'archive institutionnelle nationale, une "infrastructure nationale mutualisée hébergeant des archives institutionnelles ou vers laquelle les autres archives institutionnelles sont fermement invitées à déverser leur contenu" (Bauin, 2014, p.3).

Cette position singulière au cœur de l'écosystème des archives ouvertes et de la stratégie "voie verte" en France a été confirmée récemment par le Plan national pour la science ouverte; ce Plan rend obligatoire la publication en accès ouvert des articles et livres issus de recherches financées par appel d'offres sur fonds publics, il annonce la création d'un fond pour la science ouverte et s'engage à soutenir l'archive ouverte nationale HAL; il déclare également vouloir simplifier le dépôt par les chercheurs qui publient en accès ouvert sur d'autres plateformes dans le monde (MESRI, 2018).

Autour de l'archive ouverte HAL s'est construite une offre de services importante et variée, faite de portails et de collections, de référentiels, d'applications, d'un serveur de thèses de doctorat et d'habilitations (TEL), d'une archive ouverte de photographies et d'images scientifiques (MédiHAL), d'une plateforme pour organiser des événements scientifiques (sciencesconf.org) et d'une plateforme pour réaliser des "épirevues" à moindre coût afin de mettre en œuvre le libre accès aux versions électroniques des articles (Episciences).

Une partie des universités et organismes scientifiques ont fait le choix de créer leur archive institutionnelle sur HAL, en agrégeant les dépôts de leurs communautés scientifiques à partir des affiliations. Il s'agit par exemple de Sorbonne Université, de l'Ecole Normale Supérieure, de l'Ecole Centrale de Paris, d'Aix-Marseille Université, de l'Université de Toulouse Paul Sabatier, de l'Université de Lorraine, de l'Ecole Polytechnique et de l'Université de Nantes. Parmi les organismes, on trouve l'INRIA, l'INSERM, l'IRD et le CIRAD, mais aussi le CEA, l'Institut Pasteur ou l'Observatoire de Paris.

GERiiCO fait partie des laboratoires, instituts et autres structures de recherche qui ont décidé la création d'une collection sur HAL, proposant un accès unique aux publications de leurs chercheurs et quelques services annexes. A titre d'exemple, des 66 laboratoires de l'Université de Lille, pour 59 d'entre eux, on trouve des dépôts sur HAL (en tout, 7 582 documents, plus 29 363 notices) mais seulement 20 laboratoires ont créé leur propre collection, pour rendre visible leur production; leurs dépôts de documents représentent un peu plus de 60% des documents de tous les laboratoires.

3 Méthodologie

Sous forme d'un retour d'expérience, nous présentons la mise en place de la collection du laboratoire GERiiCO de l'Université de Lille sur l'archive ouverte nationale HAL. Les ressources mobilisées et les actions mises en œuvre font l'objet d'une description succincte, suivie d'une analyse des résultats en termes de dépôts, d'usage et de services et d'une synthèse des facteurs de succès et des problèmes rencontrés. L'étude terminera par quelques perspectives pour le futur développement d'une telle collection de laboratoire.

Les analyses statistiques du contenu et de l'usage de la collection ont été menées en avril et mai 2018, grâce aux outils de la plateforme HAL et en exportant les résultats sous forme de tableaux Excel pour d'autres analyses plus poussées. L'analyse scientométrique des auteurs et co-auteurs a été menée en juillet 2018 avec Gephi, à partir des données extraites en avril et mai. La comparaison scientométrique des publications de GERiiCO dans Scopus, le Web of Science et Google Scholar a été faite en juillet 2018.

4 La collection GERiiCO sur HAL

Le laboratoire GERiiCO³ est un pôle de recherche à vocation internationale en sciences de l'information et de la communication de la région Hauts-de-France. Créé en 2006 comme une équipe d'accueil de l'Université de Lille 3, GERiiCO compte au 1er janvier 2018 35 enseignants-chercheurs titulaires, trois professeurs émérites et 38 doctorants. L'orientation majeure de GERiiCO est l'analyse des pratiques, des processus et des dispositifs info-communicationnels saisis dans leurs dimensions langagières et discursives, technologiques et symboliques.

Les premiers dépôts de la part d'auteurs de GERiiCO remontent à 2003. Pendant plusieurs années, il n'y a pas eu de position officielle ou

³ Groupe d'Études et de Recherche Interdisciplinaire en Information et COmmunication cf. <https://geriico-recherche.univ-lille3.fr/>

d'incitation de GERiCO concernant le libre accès ou HAL ; tous les ans ont été déposés entre 20 et 40 documents ou références sur HAL, par les auteurs eux-mêmes. La situation évolue après la dernière campagne d'évaluation (AERES) du laboratoire en 2014 lorsqu'il est décidé de créer des notices sur HAL pour la production de GERiCO, à partir des listes bibliographiques 2008-2012 des membres titulaires. En parallèle, une collection GERiCO est officiellement créée sur HAL, sous la responsabilité d'un administrateur, avec l'aide du SCD. L'objectif de la mise en place de cette collection est double : d'une part, assurer une visibilité maximale et un impact au-delà de la communauté disciplinaire, à travers des médias sociaux et le référencement des moteurs de recherche; d'autre part, pallier l'absence d'un système d'information recherche pour la gestion des publications au sein du laboratoire (cf. figure 1).

gériico Groupe d'Études et de Recherche Interdisciplinaire en Information et Communication

Accueil Consultation Recherche Ajouter le texte intégral

Accueil

Notices	Dépôts
829	436

Ai-je le droit de déposer ma publication en texte intégral dans HAL?
 En savoir plus sur la loi du 7/10/2016

Est exactement par Contient ISSN
 Commenter ISSN

Connaître la politique des éditeurs également sur

Chercher une publication

Bienvenue dans la collection HAL du laboratoire GERiCO

Equipe d'accueil de l'Université de Lille (EA 4073), elle fédère 39 enseignants-chercheurs de plusieurs départements. L'orientation majeure de GERiCO est l'analyse des pratiques, des processus et des dispositifs info-communicationnels saisis dans leurs dimensions langagières et discursives, technologiques et symboliques.

Derniers dépôts

- Joachim Schöpfel. Vers une culture de la donnée en SHS : Une étude à l'Université de Lille. [Rapport de recherche] Université de Lille. 2018. (hal-01846849)
- Quoc-Tan Tran. Knowledge Organization Systems and cultural interoperability in open humanitarian settings. Fifteenth International ISKO Conference: "Challenges and Opportunities for Knowledge Organization in the Digital Age", Jul 2018, Porto, Portugal. (http://www.iskolberko.org/about-isko-porto-2018/ (hal-01833678))
- Abdou Beukes Sow, Stéphane Chaudiron. Analyse des pratiques info-communicationnelles des chercheurs sénégalais selon l'approche sense-making. Colloque international "Les sciences de l'information documentaire au service de la recherche, de la formation, de l'intégration et du développement durable", Nov 2017, Dakar, Sénégal. (http://senegal.ird.fr/toute-l-actualite/evnements-scientifiques/colloque-international-les-sciences-de-l-information-documentaire-au-service-de-la-

Contacts

Référent publications :
 Hélène Prost
 Lab.Gerico@gmail.com
 Archive créée et administrée sur la plateforme HAL du CCSD, et réalisée avec le soutien de la MSH de Dijon

Authentification

Se connecter ? Se créer un compte ? Retrouver son identifiant ou son mot de passe ?... **CLIQUEZ, C'EST ICI !**

Dernières actualités du Libre Accès

LE PRINCIPE DE « SCIENCE OUVERTE » DEVIENDRA LE MODE OPÉRATEUR D'HORIZON EUROPE (29/06/2018)
 Dans un communiqué en date du 7 juin 2018, la Commission européenne annonce qu'elle propose une enveloppe de 100 milliards d'euros pour la recherche et l'innovation. Pour elle, le nouveau programme « Horizon Europe est le programme de recherche et d'innovation le plus ambitieux à ce jour. »
 La Commission demande plus d'ouverture : « Le principe de « science ouverte » deviendra le mode opératoire d'Horizon Europe, exigera un accès ouvert aux publications et aux données, ce qui favorisera la (...) »
 - Actualités

LIRE LE BILLET

LA LIBRE CIRCULATION DE LA CONNAISSANCE DANS L'UNION EUROPÉENNE (14/06/2018)

Figure 1. Page d'accueil de la collection GERiCO sur HAL (capture d'écran, 26 juillet 2018)

A l'occasion du bilan de l'activité de GERiCO pour la nouvelle campagne d'évaluation 2019 (HCERES), les références des listes bibliographiques 2013-2017 des chercheurs de GERiCO ont servi pour la création des notices sur HAL. En même temps, un compte HAL a été créé pour chaque chercheur n'ayant pas un tel compte au préalable. La saisie sur HAL a été réalisée par deux membres de GERiCO et une vacataire du SCD, après dédoublement des listes, et uniquement pour les publications pas encore

signalées sur HAL. Une partie des références ont été importées en format BibTex, d'autres ont été générées à partir de leur DOI, le reste a été saisi manuellement.

La saisie de la production 2013-2017 est terminée ; elle a été accompagnée d'un débat au sein du laboratoire, sur l'intérêt d'une telle action et sur la qualité du résultat.

4.1 Documents et notices

A l'issue de la saisie des références du bilan 2013-2017, la collection GERiCO sur HAL comptait 1 253 notices, dont 413 avec le texte intégral (33%). Elle est composée de treize catégories de documents, avant tout de communications dans un congrès (36%), d'articles de revues (32%) et de chapitres d'ouvrages (15%). Mais on trouve également des rapports, thèses, HDR et posters (tableau 1).

Type document	Nombre
Communications	450
Articles	403
Chapitres d'ouvrages	184
Directions d'ouvrages	60
Autres	41
Rapports	35
Indéfinis	28
Thèses de doctorat	19
Posters	14
Ouvrages	12
Mémoires	3
HDR	2
Brevets	2
TOTAL	1 253

Tableau 1 : Typologie des documents de la collection GERiCO sur HAL (N=1 253)

Les publications les plus anciennes de la collection datent de 1993 à 1999. En fait, 4% correspondent à des publications avant la création de GERiiCO en 2006. Mais il n'y a pas eu de projet pour signaler d'une manière rétrospective la production issue des deux unités de recherche dont la fusion est à l'origine de GERiiCO.

Trois quarts des publications de GERiiCO sont écrites en français, le reste est en anglais, plus quelques textes en italien, polonais, espagnol et allemand (figure 2).

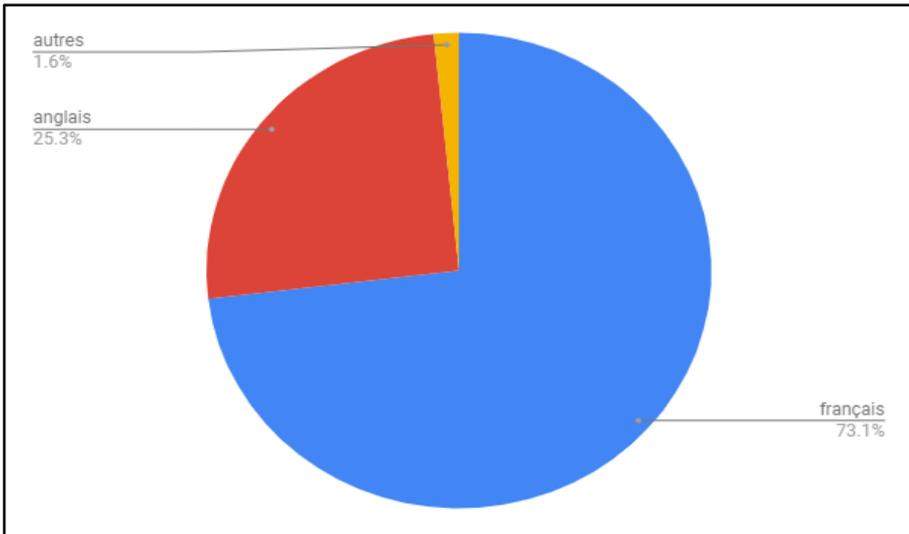


Figure 2. La langue des publications de la collection GERiiCO sur HAL (n=1 253)

Nous reviendrons plus loin sur l'aspect international de la collection.

4.2 Exhaustivité et visibilité

L'objectif de la saisie des références à partir des listes bibliographiques du bilan pour le HCERES, en complémentarité de l'auto-archivage par les auteurs est d'obtenir une vitrine exhaustive et fiable de la production des chercheurs de GERiiCO. Pour deux raisons : les grandes bases de données scientométriques (Web of Sciences, Scopus) ne contiennent qu'une petite partie des publications françaises, notamment en SHS (Schöpfel & Prost 2009) ; et le référencement par les outils de recherche comme Google Scholar, Microsoft Academic, Bielefeld Academic Search Engine ou Dimensions reste opaque et manque de fiabilité.

La comparaison directe des publications de GERiiCO sur HAL, dans Google Scholar et dans Scopus confirme ce constat. En effet, l'interrogation de Scopus sur le champ affiliation fournit 74 notices liées à GERiiCO, soit 5,9% de la collection. Les années de publication de ces notices s'étendent de 2006 à 2018, mais seule l'année 2016 atteint 10 notices. Quant à l'interrogation de Google Scholar, on obtient 1 020 réponses dont une partie est du bruit - soit des pages web sans caractère de publications (appel à communication, annonce de conférence...), soit des publications qui ne font pas partie de la production de GERiiCO (communications d'une conférence organisée par GERiiCO...).

Figure 3 illustre les effets des deux montées en charge de saisie dans HAL en 2016 et 2018 sur l'exhaustivité de la collection. Le nombre de notices présentes dans Google Scholar est plus important que celui détenu dans HAL seulement pour l'année de publication 2012, année où le laboratoire GERiiCO a accueilli le colloque *Communiquer dans un monde de normes*, mais aussi pour 2017 et 2018, car la visibilité dans Google Scholar est effective sitôt la saisie sur Internet. Mais encore une fois, une partie des références de Google Scholar sont des erreurs, dans la mesure où elles ne correspondent pas aux publications des chercheurs de GERiiCO.

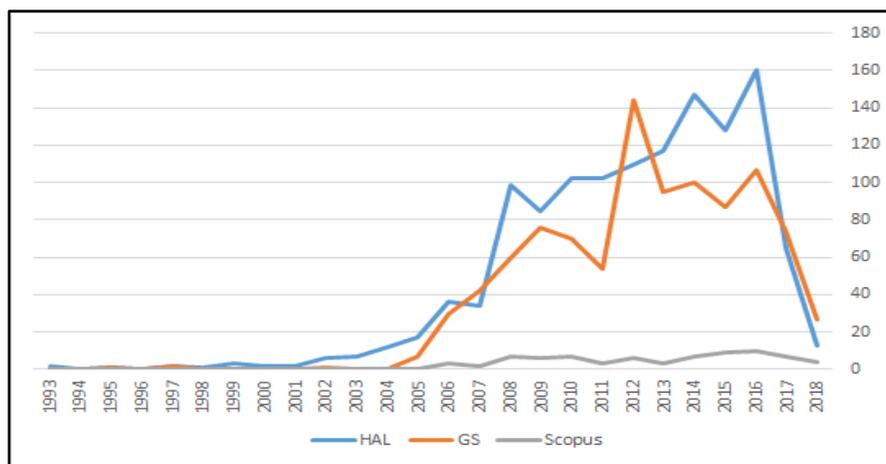


Figure 3. Comparaison de la collection HAL avec Google Scholar et Scopus (29 juillet 2018)

Quant à l'exhaustivité de la collection sur HAL, le seul bémol est que par rapport aux listes exhaustives du bilan pour le HCERES, une petite partie des publications ne correspond pas aux critères initiaux de HAL. En effet, au départ, il n'était pas possible d'intégrer certains types de documents

dans HAL, tels que les articles de blog, les notes de lecture ou les traductions; c'est désormais possible, mais la saisie rétrospective de ces notices reste à faire. Autrement dit, certaines de ces références ne figurent pas dans HAL mais sont potentiellement visibles sur Internet, à condition d'être signées avec l'affiliation GERiCO.

4.3 Consultations et téléchargements

Dans le cadre de ce travail de recherche, nous nous sommes basés uniquement sur le nombre de consultations des notices et de téléchargements des fichiers comme indicateur d'influence scientifique.

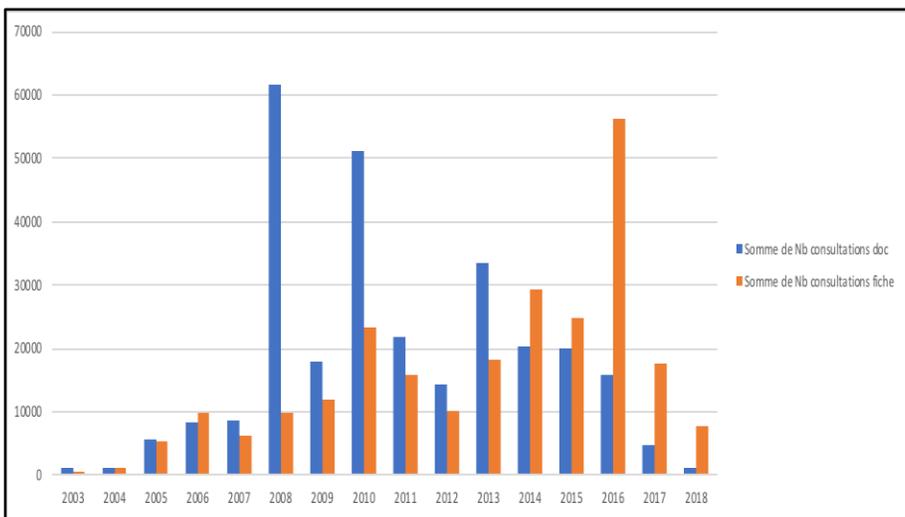


Figure 4. Consultations et téléchargements de la collection GERiCOpar année de dépôt (2003-2018)

La figure 4 présente le total des consultations et de téléchargements par document et par référence déposés entre 2003 et 2018. Nous pouvons constater que des documents déposés en 2008 et 2010 ont été marqués par un nombre de consultations très importants. On remarque également un nombre de consultation de références important pour 2016, qui coïncide avec la date de dépôt sur HAL d'une grande partie de la collection GERiCO.

La figure 5 montre l'évolution du nombre total de documents déposés dans HAL. Ce nombre a évolué de façon constante entre les années 2003 et 2015. Le nombre de documents déposés sur cette période est passé de 0 à 50 documents avec une moyenne de 4-5 dépôts par année. Un nombre de dépôt

important (450 documents) a été réalisé en 2016. Ceci s'explique par les actions internes menées au sein du laboratoire pour inciter les chercheurs à déposer leurs documents sur la plateforme HAL.

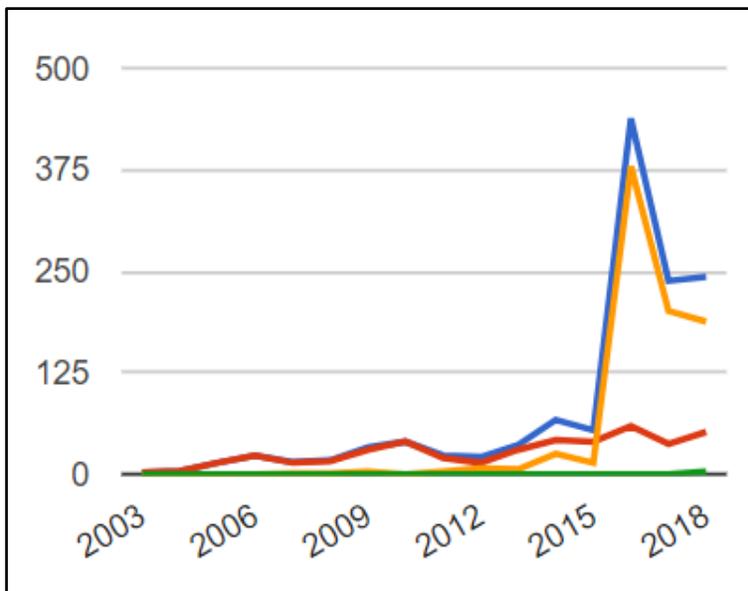


Figure 5. Dépôts sur HAL (n=1 265, juillet 2018). Bleu : total des dépôts. Jaune : références. Rouge : documents. Verts : annexes.

Regardons les statistiques d'usage de plus près. Depuis la création de la collection, les 413 documents déposés ont été téléchargés 286 658 fois, avec une médiane de 276 téléchargements. Quant aux 840 notices sans documents, elles ont été consultées 248 401 fois, avec une médiane de 83 (chiffres de mai 2018). Premier constat : tous les documents ont été téléchargés au moins trois fois, et toutes les notices ont été consultées au moins dix fois. On est donc face à un usage du genre "longue traîne", certes avec quelques dépôts très fortement demandés mais surtout avec un grand nombre de documents et de notices occasionnellement consultés et téléchargés. A titre d'exemple, les chiffres de téléchargements pour les documents (figure 6).

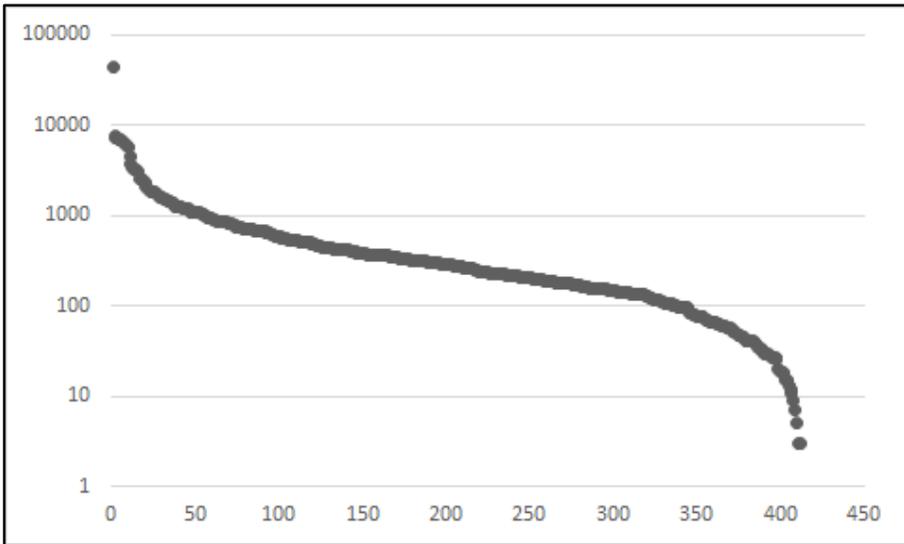


Figure 6. Téléchargements des documents déposés ($n=413$, mai 2018, échelle logarithmique)

A priori, il s'agit d'usage réel, "humain", dans la mesure où l'outil web analytics de HAL filtre l'essentiel de l'activité machine (robots etc.). On peut donc partir du principe que la mise en ligne des documents et notices HAL rend la totalité de la production scientifique du laboratoire visible, et pas seulement une partie.

La date de dépôt joue un rôle, mais l'impact sur l'usage est relatif. A titre d'exemple, les statistiques pour le téléchargement des documents déposés (figure 7).

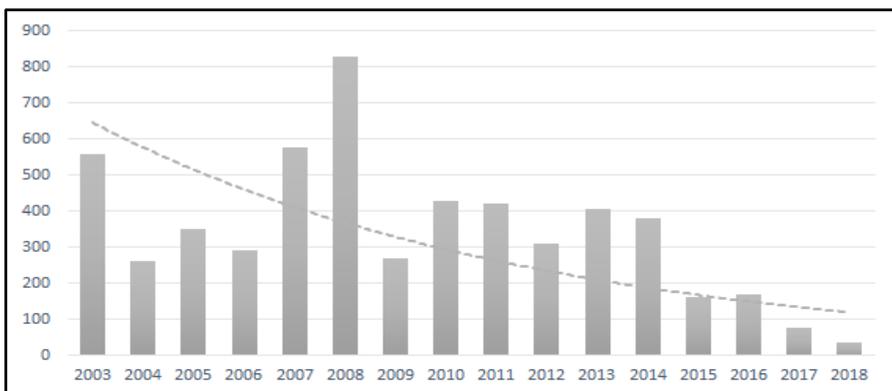


Figure 7. Le téléchargement des documents en fonction de l'année du dépôt ($N=413$, mai 2018, médiane)

Comme la courbe de tendance l'indique, la médiane des téléchargements augmente avec la durée des dépôts. La figure 7 montre aussi que cela s'applique avant tout aux 5-10 dernières années et qu'au-delà, l'impact semble faiblir ; mais ceci est peut-être aussi un effet du petit nombre de dépôts durant les premières années de la collection.

Une dernière observation. Quand on compare les types de documents, on constate quelques différences - là aussi, davantage de tendances que de différences significatives (tableau 2).

	Nombre	Téléchargements
Articles	156	273
Communications	136	261
Chapitres	34	283
Thèses	19	288
Rapports	14	300

Tableau 2. Le téléchargement par type de document (mai 2018, médiane)

On voit ainsi que l'usage moyen des chapitres d'ouvrage, thèses et rapports est en moyenne plus élevé que pour les articles et communications, un phénomène déjà constaté dans le contexte des études sur la littérature grise (Schöpfel & Prost 2016). Mais là encore, il ne faut pas sur-interpréter les chiffres, à cause du petit nombre des rapports, thèses et chapitres.

4.4 Le rayonnement international

Plus haut, nous avons évoqué la part des langues étrangères et notamment de l'anglais qui correspond à un quart des publications. La collection sur HAL fournit deux autres informations sur le rayonnement international du laboratoire.

D'une part, la collection contient un nombre non négligeable de communications dans des conférences nationales et internationales. Ce mode de publication (en conférences) a un impact important sur la dissémination et la valorisation scientifique des travaux de recherche au sein de la communauté scientifique, étant donné le nombre important de chercheurs présents simultanément au moment de la publication. Pour 435 des 450 communications de la collection, la notice contient le pays de la

conférence, et il est possible de générer des statistiques d'internationalité (figure 8).

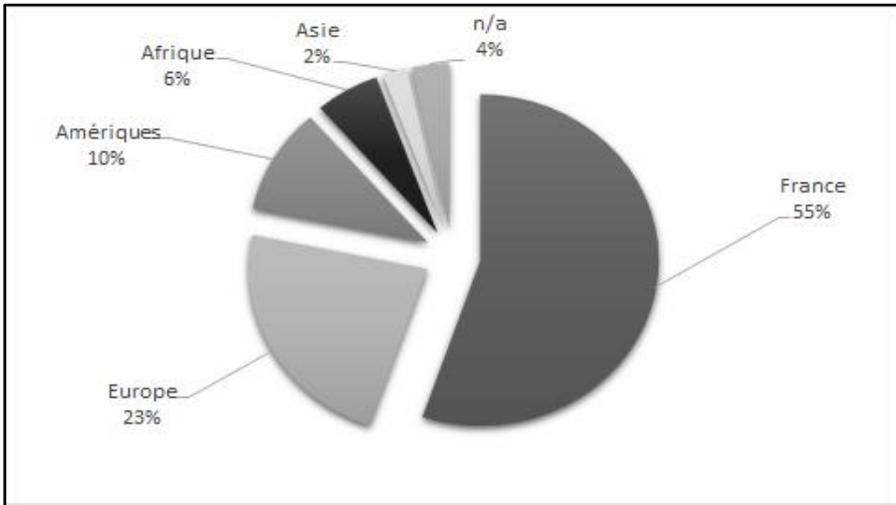


Figure 8. Pays et continents des conférences de la collection HAL (N=450)

Plus de la moitié des communications ont été présentées en France ; un quart environ dans d'autres pays européens. Il n'y a pas de tendance claire vers davantage ou moins de communications à l'internationale, et le rapport international/national reste relativement stable dans le temps, en oscillant entre 0,6 et 1,0 par an.

L'information géographique dans les affiliations des auteurs et coauteurs fournit une autre information sur le rayonnement international et les partenariats avec d'autres pays. Vu son intérêt pour l'image du laboratoire, cette information est visualisée sur la page d'accueil de la collection, à l'aide d'une application HAL (figure 9).



Figure 9. Carte des collaborations, collection GERiCOsur HAL (N=1 268, 28 juillet 2018)

La carte est interactive, et on peut modifier les années de publication des références pour visualiser les anciens partenariats ou au contraire, les plus récents, ou pour étudier l'évolution géographique des collaborations.

En revanche, la qualité des résultats de ce “widget” est conditionnée par la qualité des métadonnées des auteurs, en particulier de l'adresse de son affiliation. Nous y reviendrons plus loin.

4.5 Les réseaux des auteurs et coauteurs

Grâce aux informations sur les auteurs et les coauteurs, la collection sur HAL permet aussi une analyse scientométrique des réseaux scientifiques au sein et autour du laboratoire. Ainsi, nous avons mené une analyse des relations entre les différents chercheurs présents dans la collection, à partir des publications communes. Cette analyse a pour objectif de quantifier la communication d'un individu ou d'un groupe, non seulement en termes de volume, mais également de visibilité, d'influence, de partenariats, d'insertion dans les réseaux. Concrètement, chaque chercheur qui est auteur ou coauteur d'au moins un document dans la collection GERiCOsur HAL est représenté par un nœud dans le réseau. Les relations entre les nœuds représentent une publication commune entre deux chercheurs.

Deux exemples: dans la figure 10, le document est représenté par deux nœuds : [*Joachim Schöpfel, Hélène Prost*] et une relation entre les deux nœuds.



Figure 10. Exemple d'une publication à 2 auteurs avec 2 affiliations différentes

Le deuxième exemple (figure 11) est représenté par cinq nœuds [*Romarc Besançon, Stéphane Chaudiron, Djamel Mostefa, Ismaïl Timimi, Khalid Chokri*] et dix relations [*Romarc Besançon__StéphaneChaudiron, Romarc Besançon__Djamel Mostefa, Romarc Besançon__IsmaïlTimimi, Romarc Besançon__Khalid Chokri, Stéphane Chaudiron__Djamel Mostefa, Stéphane Chaudiron__Ismaïl Timimi, Stéphane Chaudiron__Khalid Chokri, Djamel Mostefa__IsmaïlTimimi, Djamel Mostefa__Khalid Chokri, Ismaïl Timimi__Khalid Chokri*].



Figure 11. Exemple d'une publication multi-auteur avec trois affiliations différentes

L'application de cette analyse à l'ensemble des publications de la collection GERiCOsur HAL aboutit à un réseau avec 592 nœuds et 3 852 relations. Ce réseau global est visualisé par figure 12.

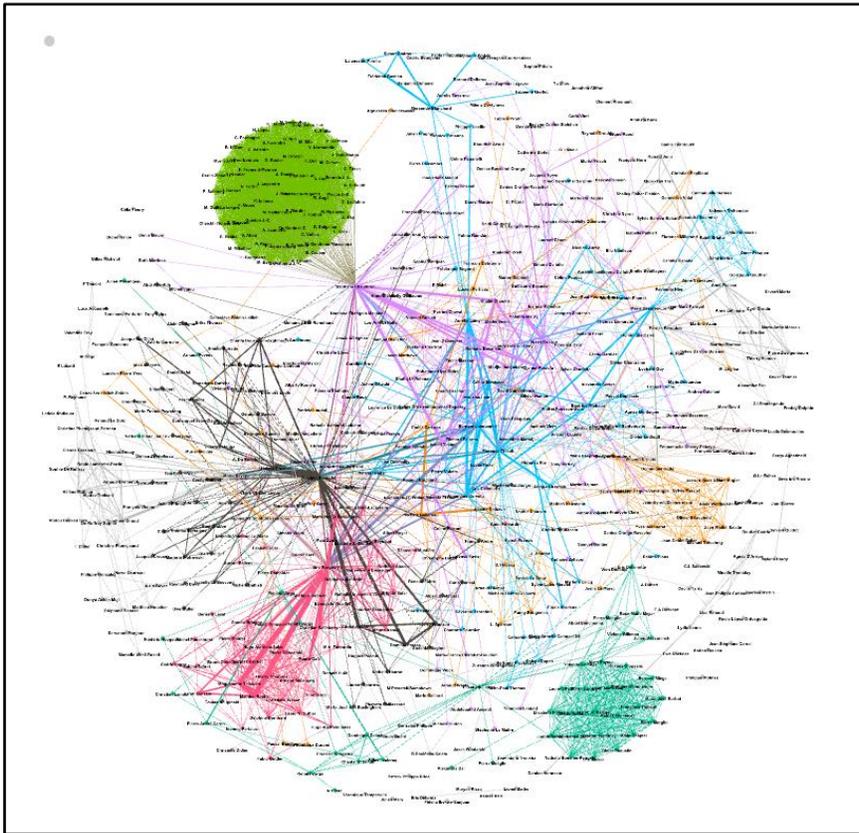


Figure 12. Réseaux internes et externes entre les chercheurs GERiCO

Comme le montre la figure 12, nous remarquons l'existence d'espaces de communautés très variées. Nous divisons cet espace selon deux catégories:

1. *Les communautés fermées* : nous qualifions de *fermées* les communautés ayant un nombre de relations intra-communauté (internes) plus important que le nombre de relations inter-communautés (externes).
2. *Les communautés ouvertes* : nous qualifions de *ouvertes* les communautés ayant un nombre de relations inter-communautés plus important que celui d'intra-communauté.

Nous remarquons aussi que certains nœuds jouent un rôle pivot entre sous-communautés, ce qui permet un meilleur impact et ouverture des travaux de recherche.

Nous n'allons pas plus loin ici dans l'analyse, puisqu'elle sert uniquement d'illustration du potentiel d'une collection de laboratoire sur HAL. Une telle analyse est utile pour la compréhension des liens et dynamiques au

sein d'un laboratoire de recherche. Elle est également intéressante pour l'état des lieux dans le cadre du bilan à destination des instances d'évaluation.

5 Facteurs de succès et problèmes rencontrés

5.1 Un projet collaboratif

La collection GERiiCOsur HAL a été lancée au printemps 2015, sous une interface customisée et un guichet unique pour l'accès à l'ensemble des dépôts que les membres de GERiiCO avaient effectués antérieurement sur HAL ou l'un des autres portails de HAL (TEL, ArchiveSIC, HAL-SHS, HAL-Lille3). La préparation et mise en œuvre de cette collection a été faite en étroite concertation avec le SCD de Lille 3 et le CCSD, dont les supports de formation à destination des gestionnaires de collection, au moment du déploiement de la version 3 de HAL fin 2014, nous ont été fort utiles. De nouvelles fonctionnalités ont facilité le dépôt de notices : citons par exemple, l'import automatique des principales métadonnées d'un article via le DOI ou la détection immédiate de doublon au moment de la saisie, si la notice existe déjà dans la collection. En 2017, dans le cadre d'une politique de la science ouverte, la Direction Recherche de Lille 3 a recruté une vacataire afin d'accélérer l'alimentation de HAL ; plusieurs jours de son travail ont été consacrés à la collection GERiiCO.

La customisation du site de la collection a été réalisée en partenariat avec la MSH de Dijon qui propose des modèles de mise en page et feuille de style CSS pour les portail et collections sur HAL. Cette coopération avec les instances locales et nationales de l'écosystème HAL a été (et reste) indispensable pour la mise en œuvre et le développement d'une telle initiative.

Les compétences mises en œuvre dans cette phase du projet relèvent du métier de l'information-documentation (conception d'un dispositif documentaire, paramétrage, gestion documentaire, métadonnées etc.). En termes comptables, la mise en place et l'alimentation de la collection n'ont pas eu besoin d'un budget d'investissement ou de fonctionnement particulier. Les seules dépenses sont celles des ressources humaines, en termes de charge de travail, dont une vacation en 2017 (SCD, Direction Recherche, GERiiCO) ; mais il n'y a pas eu d'estimation de coûts, et une partie du travail a été effectuée sur temps personnel.

Après le lancement de la collection, plusieurs membres de GERiiCO, dont la référente publications HAL, ont saisi les références du bilan de 2014 pour obtenir une couverture exhaustive de la production entre 2008 et 2012. La saisie de ces références a pris plusieurs mois et a été terminée fin 2016. Les

résultats ont été présentés le 9 février 2017 par Joachim Schöpfel et Hélène Prost dans le séminaire doctoral GERiiCO 2016-2017 “De l’influence en SIC” coordonné par Patrice de la Broise, sous l’aspect “Mesurer l’influence scientifique en termes d’impact : potentiel et limites des nouveaux indicateurs scientométriques”. C’était l’occasion notamment de discuter les statistiques d’usage de la collection et de comparer sa visibilité avec Google Scholar et Scopus.

Le débat entre collègues et le soutien de la part du bureau et de la direction du laboratoire sont essentiels pour la réussite d’une telle initiative. Le débat interne, avec la coopération au sein de l’écosystème HAL, est une sorte de garant pour la durabilité de l’initiative à moyen terme.

Les critiques ont porté sur la forme, non pas sur le fond. L’intérêt d’une collection de laboratoire sur HAL n’a pas été remis en question lors des échanges au sein du laboratoire. Les interrogations concernaient d’une part la mise en œuvre (pilotage, gouvernance, interlocuteurs) et d’autre part le résultat, les références mises en ligne (erreurs de saisie, doublons etc.). Dans certains cas, il s’agissait simplement d’une incompréhension de la différence entre dépôt de références (notices) et auto-archivage des publications, et la discussion n’a pas seulement clarifié l’aspect juridique, mais elle a également contribué à une meilleure compréhension des enjeux de la science ouverte.

5.2 Accès vs. signalement

Nous avons indiqué plus haut qu’à l’issue de la saisie des références du bilan 2013-2017, la collection GERiiCOsur HAL comptait 1 253 notices, dont 413 avec le texte intégral (33%). Pour un catalogue de références dont la finalité est de signaler la production d’un laboratoire, c’est beaucoup. Pour une archive ouverte dont la fonction est la communication directe entre chercheurs, ce n’est pas assez.

L’explication de ce taux de 33% est effectivement la saisie exhaustive des références issues des bilans du laboratoire à partir de 2016. Des 350 notices de la collection créées entre 2003 et 2015 par les auteurs, 83% contiennent un lien avec le document déposé sur HAL. Cette situation s’inverse complètement à partir de la saisie des références ; des 903 notices créées entre 2016 et 2018, plus que 22% contiennent ce lien. Pour la plupart, il s’agit de notices créées sous forme d’auto-archivage, par les auteurs. Seulement dans quelques cas isolés, le document a été ajouté à la notice avec l’autorisation des auteurs.

Cet effet d’une stratégie de signalement exhaustif de la production scientifique d’une structure fait débat. Les uns regrettent l’abandon du

principe de la communication directe entre chercheurs, d'autres saluent le potentiel pour une évaluation scientométrique en dehors des outils commerciaux (Scopus, Web of Science) ; certains préconisent une distinction nette des dispositifs de l'écosystème de la science ouverte, par une séparation explicite des archives ouvertes (100% accès libre aux documents) et des outils d'évaluation (référentiels).

Pour évaluer notre approche, nous avons comparé la collection GERiiCO avec les autres collections de laboratoire de l'Université de Lille sur HAL.

Pour l'ensemble des 66 laboratoires, le taux de dépôts sur HAL avec documents (texte intégral) est de 20% (19% pour les laboratoires avec une collection HAL, 23% pour les autres), avec une médiane de 32%. La figure 13 montre le rapport entre ce taux et le nombre total des dépôts sur HAL ; la collection de GERiiCO est entourée d'un cercle rouge.

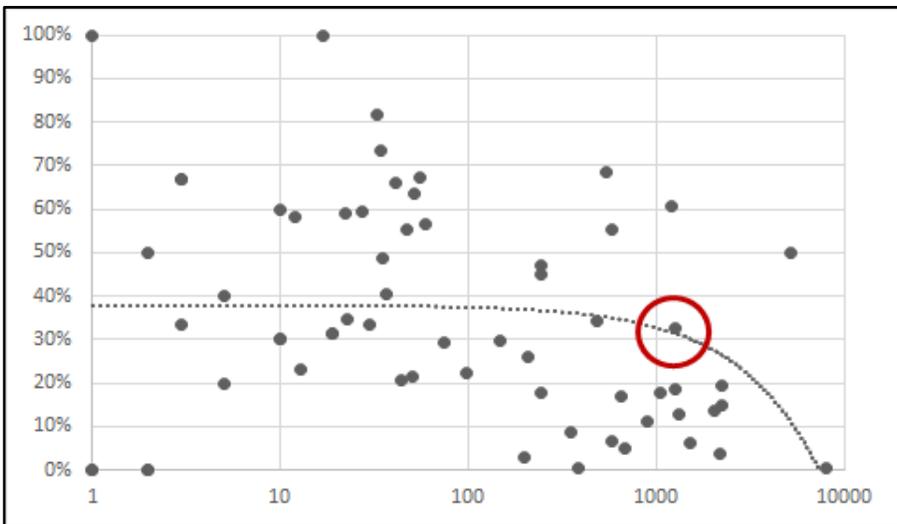


Figure 13. Taux de dépôts avec document et nombre de dépôts pour les laboratoires de l'Université de Lille (mai 2018, axe du nombre de dépôts en échelle logarithmique)

Par rapport au nombre de dépôts, la collection de GERiiCO est bien au-dessus de la moyenne des laboratoires de l'Université de Lille (médiane = 52 dépôts). Quant au taux de dépôts avec texte intégral, la collection se trouve juste au-dessus de la moyenne. Par ailleurs, la courbe de tendance visualise un phénomène analysé ailleurs : les grandes collections ont souvent un taux de texte intégral moins élevé (Prost & Schöpfel 2014), avec des exceptions bien sûr (ici il s'agit de l'UMR 9189 CRISTAL (Centre de

Recherche en Informatique, Signal et Automatique de Lille), avec plus de 5000 dépôts et un taux de texte intégral de 50%.

Le même schéma avec uniquement les laboratoires en SHS, tous avec une collection HAL, montre deux groupes avec des approches différentes (figure 14).

La médiane des dépôts est 533, celui du taux de texte intégral de 22%. GERiiCO se trouve donc bien dans le quart supérieur des laboratoires SHS, dans un groupe de neuf laboratoires qui semblent viser comme GERiiCO l'exhaustivité de la collection sur HAL, au détriment du taux de texte intégral. Cependant, un autre groupe de trois laboratoires semble être sur une autre trajectoire, avec un taux élevé de dépôts avec texte intégral mais peu de dépôts. Probablement, ces laboratoires ne font pas (encore) de saisie de références.

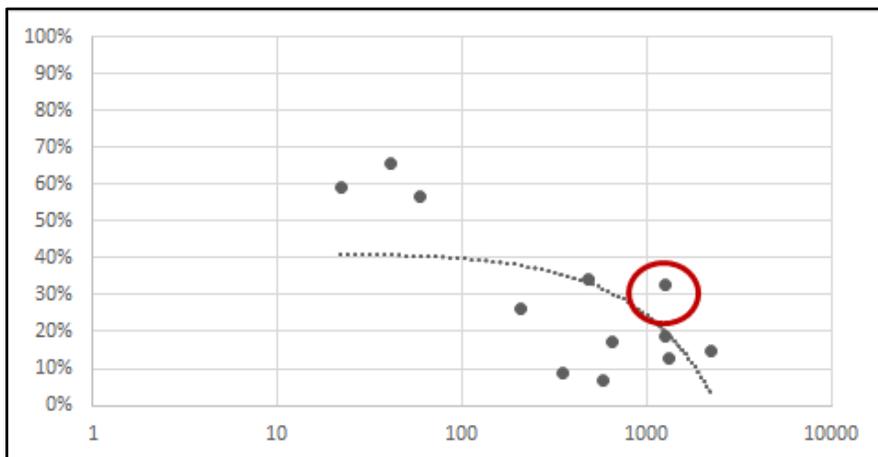


Figure 14. Taux de dépôts avec document et nombre de dépôts pour les laboratoires SHS de l'Université de Lille (mai 2018, axe du nombre de dépôts en échelle logarithmique)

5.3 Métadonnées et identifiants

Le problème majeur rencontré et critiqué est la qualité des métadonnées (cf. liste en annexe). Il y a deux problèmes : des métadonnées erronées, et des métadonnées incomplètes ou absentes.

- Métadonnées erronées : souvent des fautes d'orthographe ou d'écriture du nom et prénom d'auteur, de l'affiliation, du titre de la publication, du nom de la revue, du nom de la conférence, des fautes dans le résumé etc. L'ordre des auteurs est un autre point sensible.

- **Métadonnées incomplètes ou absentes** : absence d'une affiliation ou affiliation incomplète (absence de pays), oubli d'un auteur, absence de pagination, absence d'un identifiant (DOI ou autre).

D'une manière générale, il peut y avoir deux raisons : soit il s'agit d'une erreur de saisie lors du dépôt et de la création des métadonnées, par l'auteur ou par une autre personne du SCD ou du laboratoire ; soit l'erreur provient de la liste bibliographique du bilan HCERES.

Le contrôle lors de la création des métadonnées est relativement léger, et la procédure de dépôt a été simplifiée pour faciliter et accélérer l'auto-archivage. Il y a donc un risque réel de créer des doublons de notices ou de variantes d'écriture pour les auteurs, les institutions, les revues et conférences.

Lors de la deuxième vague de saisie des références, l'index des noms d'auteurs a fait l'objet d'un contrôle systématique, tout comme les doublons. Aussi, un stagiaire a relevé plusieurs dizaines d'anomalies de métadonnées dont une partie a déjà été corrigée. De même, le personnel du SCD est intervenu pour encourager les chercheurs à créer leur propre IDHal, l'identifiant auteur unique dans HAL, de façon homogénéiser les métadonnées auteur, tout en permettant au chercheurs de générer leurs curriculum-vitae et bibliographies.

5.4 Indexation

Une autre question se pose pour l'indexation des dépôts. Le principe de l'auto-archivage veut que l'indexation (mots-clés, domaine) soit réalisée par le déposant, c'est-à-dire par l'auteur mais l'ajout de mots-clés n'est pas obligatoire. La plateforme HAL n'a pas de fonctionnalité ou d'outils pour une gestion documentaire de l'indexation. Certes, il est possible de corriger ou compléter l'indexation d'une notice mais il n'y a pas de terminologie contrôlée et/ou structurée, et il n'est pas possible de corriger l'indexation de plusieurs notices à la fois, en remplaçant systématiquement un mot-clé par un autre etc.

Le principe d'une indexation en mode "crowdsourcing" facilite le dépôt et ne devrait pas avoir d'impact sensible sur le référencement par les moteurs de recherche puisqu'ils préfèrent indexer le texte intégral (y compris titre et résumé) et sont capables à faire face à une indexation en langage naturel. En revanche, l'effet n'est pas satisfaisant pour la navigation dans la collection, sur le site de la collection, et pour une analyse scientométrique à partir de l'indexation.

Un problème particulier se pose pour la saisie des références à partir des listes bibliographiques qui ne contiennent ni de mots-clés ni de domaines

thématiques. L'option retenue a été une indexation *a minima* du domaine SIC, sans mots-clés, d'autant plus que, depuis le déploiement de la dernière version de HAL, la saisie de mots-clés n'est plus obligatoire pour la validation de la notice.

6 Perspectives

La mise en place d'une collection de laboratoire sur HAL contribue à la transformation de l'utilisation individuelle de HAL (l'auto-archivage par les chercheurs) en démarche institutionnelle, avec la création d'un site web, la désignation d'un responsable de collection (référént publications), le suivi des dépôts et la correction des métadonnées. Quelles sont les perspectives d'un tel projet ? Voici quelques éléments de réflexion, par rapport à la nouvelle politique nationale pour la science ouverte, par rapport à la stratégie institutionnelle de l'Université de Lille, et par rapport à l'intérêt d'une telle collection pour l'évaluation des chercheurs et du laboratoire, y compris par la mise en place d'outils à valeur ajoutée.

6.1 Alimentation et mises à jour

Le choix d'une collection HAL vitrine de la production scientifique d'un laboratoire implique une alimentation continue, avec des mises à jour effectuées par les auteurs ou d'autres.

Maintenir le statu quo et continuer comme avant voudrait dire, une alimentation continue par les dépôts des chercheurs avec des saisies de références au moment des bilans, tous les cinq ans. L'inconvénient est que l'auto-archivage par les chercheurs correspond actuellement à environ 30-40% de la production de GERüCO et que le reste, plus de la moitié, restera donc invisible sur HAL d'un bilan à l'autre. Un autre inconvénient est la charge de travail que cela implique - d'abord la rédaction des listes pour le bilan par les chercheurs, puis la compilation pour le rapport d'activité du laboratoire et en parallèle, la saisie dans HAL. Et comme indiqué, cette procédure est également une source d'erreurs.

On pourrait chercher des alternatives locales, essayer d'inciter les chercheurs à utiliser Zotero pour leur bibliographie ou saisir leur références directement dans HAL, puis les exporter pour le bilan d'activité. La solution viendra probablement de la nouvelle politique pour une science ouverte, au niveau national aussi bien qu'au niveau institutionnel.

- National : dans le cadre du Plan national pour la science ouverte, il y aura sans doute une incitation forte voire, dans certains cas (financement ANR) une obligation de publier les résultats en open access, dont une partie se retrouvera sur HAL ; en d'autres termes,

la part des dépôts devrait augmenter d'ici un ou deux ans. En parallèle, la plateforme HAL devrait se doter de nouvelles fonctionnalités et de nouveaux outils pour faciliter les échanges avec d'autres dispositifs, simplifier l'auto-archivage et rendre l'utilisation de HAL plus attractive.

- Institutionnel : l'Université de Lille a lancé sa propre archive ouverte institutionnelle, LilloA⁴, et il n'est pas exclu qu'elle s'aligne sur le modèle de Liège, avec l'obligation d'utiliser ce dispositif pour toute demande d'avancement ou de subvention.

Dans les deux cas, la part de l'auto-archivage devrait augmenter ; si LilloA devient réellement un outil de suivi et d'évaluation de la production des chercheurs de l'Université de Lille, il faudra s'assurer de l'alimentation de la collection HAL à partir de LilloA.

A suivre, donc, et à revoir d'ici un ou deux ans afin d'ajuster la procédure au sein de GERiiCO, le cas échéant. Il faudrait voir également comment récupérer le cas échéant les métadonnées et les documents déposés par un chercheur sur une autre plateforme de son choix. Par un moissonnage systématique ? On ponctuellement, au moment de la saisie des références ? Pour l'instant, il n'y a pas de solution.

6.2 Dépôt du texte intégral

La loi numérique a donné un cadre légal à "l'exploitation secondaire" des publications, y compris de l'auto-archivage sur HAL. Aussi, certains éditeurs autorisent les dépôts des preprints de leurs articles. Les outils de HAL permettent de filtrer ces publications.

Actuellement, la collection GERiiCO sur HAL contient 240 références d'articles sans texte intégral, publiés entre 1993 et 2017. La loi numérique autorise la diffusion en libre accès de ces documents. On pourrait donc imaginer une démarche ciblée pour inciter les chercheurs d'ajouter leur publication à la référence sur HAL, ce qui est une affaire de quelques moments.

Pour 34 références d'articles, dont trois publiés en 2018, l'éditeur autorise l'auto-archivage. De nouveau, on pourrait imaginer une démarche d'incitation pour déposer ces articles - soit par les auteurs, soit par d'autres collègues, avec l'autorisation des auteurs.

De même, on pourrait cibler certains types de documents, dont notamment les rapports (actuellement, 21 notices sans texte intégral) et les

⁴<https://lilliad.univ-lille.fr/chercheur/open-access/lilloa-lille-open-archive>

communications (actuellement, 307 notices sans texte intégral), et sensibiliser les auteurs de leurs droits et de l'intérêt d'une diffusion des documents en texte intégral, si possible.

Par rapport au dépôt du texte intégral, il faut insister sur la finalité du projet (= vitrine exhaustive de la production scientifique du laboratoire) et sur la différence d'une telle collection avec par exemple une bibliothèque numérique. Ici, il n'y aura pas de filtre ou de contrôle qualité, et la modération de la part du CCSD concerne essentiellement la cohérence des métadonnées, sans évaluer le contenu du document déposé. C'est le principe même du libre accès basé sur l'auto-archivage. En revanche, on pourrait imaginer un dispositif d'annotation et de commentaire a posteriori, une sorte de « post peerreview » des documents déposés, à l'instar des épreuves.

6.3 Contrôle des métadonnées

Pour un laboratoire universitaire sans poste de documentaliste, le contrôle systématique des métadonnées de la collection est impossible. Ceci étant, pour améliorer la qualité de la collection et la visualisation des partenariats et collaborations, on pourrait imaginer une sorte d'audit ponctuel pour quelques métadonnées "sensibles" pour identifier et corriger des anomalies et erreurs, en particulier pour les noms d'auteurs, les affiliations et les conférences.

Pour l'instant, le laboratoire invite les chercheurs à vérifier leurs notices et, le cas échéant, à signaler les erreurs à la référente publications, voire les corriger sur HAL.

Le Plan d'action pour la science ouverte annonce plusieurs mesures qui pourraient s'avérer utiles pour améliorer la qualité de ces métadonnées, dont notamment la généralisation d'un identifiant normalisé (ORCID) et l'interconnexion entre HAL et un référentiel national des publications françaises (projet CONDITOR). A voir aussi dans quelle mesure le projet institutionnel autour de LilloA pourra contribuer à la qualité des métadonnées, par exemple par la curation des dépôts.

Faut-il contrôler la qualité des mots-clés, pour obtenir une terminologie homogène et structurée ? Notre réponse est non, pour deux raisons.

- Faisabilité : d'une manière pragmatique, sans documentaliste et sans outils de gestion, vouloir corriger "à la main" la terminologie d'une collection de plus de mille publications est impossible.
- Intérêt : l'intérêt d'une terminologie contrôlée est limitée, aussi bien vis-à-vis du référencement par les moteurs de recherche et

autres outils de découverte que vis-à-vis du potentiel d'applications de la fouille de textes et de données (TDM).

Il serait donc plus judicieux d'utiliser un outil TDM pour faire des recherches et pour naviguer dans les données et métadonnées de la plateforme HAL (ou simplement dans la collection GERiiCO), plutôt que d'essayer d'obtenir une qualité de métadonnées sans les moyens et sans véritable intérêt, juste dans la continuité des techniques documentaires des années 80 et 90.

La génération des métadonnées à partir d'un DOI facilite le dépôt et devrait à terme réduire le taux d'erreur, même si cette procédure n'obtient pas non plus une qualité de 100%. De même, on pourrait imaginer un import direct de métadonnées à partir des plateformes d'éditeurs, mais cela impliquerait des partenariats et licences qui n'existent pas au niveau d'un laboratoire.

6.4 Evaluation

La collection HAL d'un laboratoire universitaire est d'un intérêt multiple pour l'évaluation par le HCERES :

- **Visibilité** : la collection rend la production scientifique du laboratoire visible, bien au-delà de l'indexation par les bases de données scientométrique et d'une façon plus fiable qu'un moteur de recherche.
- **Impact** : diffuser les publications par la "voie verte", via une archive ouverte, augmente leur impact (consultation en ligne, téléchargement, référencement etc.) et leur procure un "avantage de citation".
- **Valorisation** : mettre en place un tel outil demande un investissement en temps et énergie, cela prouve que le laboratoire mène une politique réfléchie pour valoriser les résultats de la recherche de ses chercheurs.
- **Science ouverte** : l'initiative d'une collection sur HAL est le reflet d'une politique locale pour la science ouverte, en conformité avec la politique nationale et la stratégie institutionnelle.
- **Scientométrie** : la collection permet de faire des analyses scientométriques pointues pour affiner l'auto-évaluation du bilan.
- **Suivi et bilan** : la collection a le potentiel d'un outil de gestion pour faciliter le suivi et le bilan de la production du laboratoire.

A l'avenir trois initiatives pourraient renforcer l'intérêt pour l'évaluation: rapprocher la typologie documentaire de HAL de la nomenclature du HCERES, améliorer les outils de gestion de HAL pour mieux répondre aux besoins de l'évaluation, et faire le lien de HAL avec les référentiels du

gouvernement et avec les systèmes d'information recherche, en particulier avec le futur CapLab. Ces trois initiatives ne relèvent pas du périmètre d'un laboratoire. Ajoutons que la politique nationale actuelle va dans le sens d'une utilisation de HAL pour des fins d'évaluation et devrait favoriser ce genre d'initiatives.

6.5 Lien avec les données de la recherche

A ce jour, la collection ne contient pas de données de la recherche. Il y a quelques liens vers des dépôts de données et un tableau d'enquête en complément du dépôt d'un rapport, mais tout cela reste anecdotique, précurseur peut-être, mais sans réelle importance.

Dans les prochaines années il faut s'attendre à ce que la question des données de la recherche se pose à chaque chercheur, dans le contexte des programmes de recherche européens et français, mais aussi comme bonnes pratiques pour la conservation et la communication des résultats. Le Plan d'action pour la science ouverte de la France prévoit de renforcer les liens entre les publications et les données, par le biais des citations et des liens dans les métadonnées, mais aussi par le biais des revues et articles de données. D'une manière ou d'une autre, la collection des publications de GERiiCO sur HAL sera impactée. Et la question se posera sans doute aussi un jour si GERiiCO se dotera d'une collection de données dans un entrepôt de données, qu'il soit local, national ou international.

7 Conclusion

A partir de la collection HAL du laboratoire GERiiCO de l'Université de Lille, nous avons montré comme un laboratoire peut contribuer à la politique de la science ouverte, en mettant en œuvre une stratégie de "voie verte". Nous avons décrit l'intérêt d'une telle approche pour la production scientifique du laboratoire, en termes d'impact et de visibilité, mais aussi pour l'auto-évaluation, pour le bilan d'activité et pour son rayonnement international. Nous avons présenté les facteurs-clés de la réussite d'une telle initiative, dont notamment le soutien par le laboratoire et le partenariat avec le SCD et le CCSD, nous avons évoqué quelques problèmes relatifs aux métadonnées et aux mots-clés, puis nous avons ouvert plusieurs perspectives sur le futur développement d'une telle collection.

Nous n'avons pas mentionné deux points pourtant essentiels pour le positionnement et le développement de HAL - la garantie de conservation à long terme des données (documents) et métadonnées, grâce au partenariat avec le Centre informatique national de l'enseignement supérieur (CINES), et l'interopérabilité de la plateforme avec d'autres dispositifs et infrastructures, grâce à l'application de standards reconnus. Un troisième

aspect fait actuellement l'objet d'une évaluation au niveau du Ministère : l'impact d'une telle démarche sur les modèles économiques et perspectives des éditeurs scientifiques en France. Comment concilier l'auto-archivage avec l'avenir de l'édition commerciale ? Peut-on imaginer des partenariats entre HAL et certains éditeurs ?

En guise de conclusion, nous insisterons sur deux autres aspects. Le premier aspect est le potentiel d'une telle collection, dont ses données et métadonnées sont objet et matière primaire de la recherche en sciences de l'information et de la communication. Dans un travail récent (Mariani et al. 2018), les auteurs ont proposé des mesures permettant d'évaluer l'innovation d'un chercheur dans un domaine scientifique. Pour un auteur ou une publication donnée, il s'agit de repérer les termes novateurs (employés pour la première fois dans un domaine scientifique restreint) et de calculer le nombre d'occurrence de ce terme dans des travaux de recherche ultérieurs dans le même domaine.

D'autres analyses sont possibles, sur les relations et interactions au sein d'une communauté et structure de recherche, sur les partenariats avec d'autres structures et organismes, sur les modes de communication. Les outils de l'analyse des réseaux et de l'exploration des textes et des données ouvrent de nouvelles perspectives à la recherche en SIC; une collection HAL, en tant que corpus enrichi de métadonnées, s'adapte parfaitement à ces outils, à condition d'être exhaustive (ou du moins représentative), d'avoir un nombre significatif de documents en plein texte et des métadonnées de qualité. L'intérêt est double : non seulement alimenter l'auto-évaluation et la discussion stratégique du laboratoire, mais aussi contribuer à la compréhension du travail scientifique et de la production de la connaissance.

Le deuxième aspect est le fait qu'une telle initiative fait partie d'un système (écosystème) plus large. Elle n'a pas de sens comme démarche isolée, mais elle est liée aussi bien à la politique institutionnelle dans les domaines de la science ouverte, du libre accès aux publications et de l'édition scientifique, à la politique nationale de la science ouverte, ses projets, ses financements et priorités, et aux débats et choix au sein des SIC. La conformité du projet local avec les stratégies institutionnelles et politiques lui confère une légitimité indispensable pour un développement dans la durée et permettra aussi, au moins ponctuellement, d'obtenir des moyens supplémentaires pour la saisie et le dépôt, pour la curation et/ou pour certaines analyses.

Faire partie d'un écosystème plus large implique aussi d'anticiper les développements d'autres acteurs et de saisir les opportunités qui se

présentent dans l'intérêt du laboratoire, pour augmenter son rayonnement et son impact. Dans ce sens, la contribution au partage des résultats de la recherche devrait figurer parmi les objectifs prioritaires du prochain projet scientifique du laboratoire, tandis que la science ouverte devrait trouver une place privilégiée dans la définition de ses valeurs et de sa vision d'avenir.

Remerciements

Nous tenons à remercier les collègues et étudiants qui ont contribué à la réussite du projet, en particulier Bernard Jacquemin, Eric Kergosien, Cécile Malleret, Hélène Martin, Clément Plouquet (Université de Lille) et Armelle Thomas (MSH de Dijon).

Références

Bauin, S., 2014. *L'open access à moyen terme : une feuille de route pour HAL*. CNRS Direction de l'Information Scientifique et Technique, Paris. http://corist-shs.cnrs.fr/Rapport_HAL_DIST_2014

CNRS-DIST, 2016. *Livre blanc : une Science ouverte dans une République numérique*. CNRS Direction de l'Information Scientifique et Technique, Paris. <http://books.openedition.org/oep/1548>

Dai, Q., Shin, E., Smith, C., 2018. *Open and inclusive collaboration in science: A framework*. OECD science, technology and industry working papers. 2018/07, OECD, Paris. <https://doi.org/10.1787/2dbff737-en>

Harnad, S., Brody, T., Vallières, F., Carr, L., Hitchcock, S., Gingras, Y., Oppenheim, C., Stamerjohanns, H., Hilf, E. R., 2004. The Access/Impact problem and the green and gold roads to open access. *Serials Review* 30 (4), 310-314. <http://eprints.soton.ac.uk/259939/1/impact.html>

Lynch, C. A., 2003. *Institutional repositories: Essential infrastructure for scholarship in the digital age*. 226, ARL Association of Research Libraries. <http://www.arl.org/resources/pubs/br/br226/br226ir.shtm>

Mariani, J., Paroubek P., Francopoulo, G. (2018). Measuring Innovation in Speech and Language Processing Publications. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. Miyazaki, Japan, May 7-12, 2018.

MESRI, 2018. *Plan national pour la science ouverte*. Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, Paris. <http://www.enseignementsup-recherche.gouv.fr/cid132529/le-plan->

national-pour-la-science-ouverte-les-resultats-de-la-recherche-scientifique-ouverts-a-tous-sans-entrave-sans-delai-sans-paiement.html

Prost, H., Schöpfel, J., 2014. Degrees of openness. Access restrictions in institutional repositories. *D-Lib Magazine* 20 (7/8). <http://www.dlib.org/dlib/july14/prost/07prost.html>

Schöpfel, J., Prost, H., 2009. Le JCR facteur d'impact (IF) et le SCImago journal rankindicator (SJR) des revues françaises : une étude comparative. *La Psychologie Française* 54 (4), 287-305. <http://dx.doi.org/10.1016/j.psfr.2009.07.002>

Schöpfel, J., Prost, H., 2016. Altmetrics and grey literature: Perspectives and challenges. In: *GL18 Eighteenth International Conference on Grey Literature*. November 28-29, 2016, New York Academy of Medicine, New York NY, USA. <https://archivesic.ccsd.cnrs.fr/hal-01405443v1>

Suber, P., 2012. *Open access*. MIT Press, Cambridge Mass. <http://mitpress.mit.edu/books/open-access>

Thirion, P., Rentier, B., 2014. La communication scientifique à la croisée des chemins : Enjeux et stratégies institutionnels. In: *Open Access Week*, 22 octobre 2014, Université de Strasbourg. <https://orbi.uliege.be/handle/2268/173664>

Annexe

Métadonnées sur HAL

Chaque document est décrit par un ensemble de 31 variables :

- *halId_s*: l'identifiant du document sur la plateforme HAL.
- *contributorid_i*, *contributorFullName_s* : l'identifiant et le nom de la personne qui a fait le dépôt sur HAL.
- *Version_I* : la version du document
- *Uri_S* : l'URL du document
- *docType_s* : cette variable peut prendre une valeur parmi les 13 valeurs suivantes: COMM, ART, COUV, DOUV, OTHER, REPORT, UNDEFINED, THESE, POSTER, OUV, MEM, HDR, PATENT.
- *doiId_s*, *nntId_s* : l'identifiant doi
- *Title_S*, *subTitle_s* : le titre du document et le sous-titre s'il existe
- *authId_i*, *authFullName_s*, *authLastName_s* : les identifiants, les noms et les prénoms de l'ensemble des auteurs.

- *producedDate_s*: la date de publication
- *Domain_S*: le domaine qui peut prendre une valeur parmi les 4 valeurs suivantes : 0.shs (Sciences Humaines et Sociales), 1.shs.info (Sciences de l'Information), 2.shs.info.doc (Sciences de l'Information et de la documentation) ou 2.shs.info.com (Sciences de l'information et de la Communication).
- *journalTitle_s, journalPublisher_s, Volume_s, Number_s, Page_s* : lorsqu'il s'agit d'un article de revue ces champs contiennent respectivement, le nom du journal, le nom de l'éditeur, le numéro et le nombre du volume, les numéros de pages.
- *conference_Title_s, conferenceStartDate_s, Country_s* : pour les communications dans des congrès ces champs contiennent le nom de la conférence, la date de début et le pays.
- *Language_s* : la langue du document
- *inPress_book*: 2 valeurs possibles : *oui* si le document est apparu dans un livre de presse sinon *non*.
- *fileType_s* : le type du fichier
- *Nb_consultation_fiche, Nb_consultation_doc* : le nombre de consultations de la fiche et du document
- *submittedDate_s, submittedDateY_i* : la date du dépôt du document sur HAL.

Accès aux données ouvertes de la recherche : Vers l'application d'un système de recommandation optimisant le repérage de jeux de données libres

Access to open research Data: Towards a recommendation model optimizing the recovery of open data sets

Amine Sennouni, Laboratoire Dicen-IDF, CNAM- Paris- France

asenouni@esi.ac.ma

Ahmed Abdelilah Bachr, Laboratoire Mediation-Information-Knowledge-Society (MIKS) ESI-Rabat-Maroc

abachr@esi.ac.ma

Résumé

L'open data représente un gisement de jeux de données numériques essentiels dans la démarche de recherche scientifique et technique, du moment où l'accès et l'usage sont laissés libres pour les usagers. Ces jeux de données sont disponibles dans des entrepôts de données ouvertes. Cependant, leur indexation à l'aide de métadonnées descriptives est forcément réductrice, ce qui pose le problème de leur repérage par la communauté de chercheurs. L'optimisation de l'accès à ces données ouvertes de la recherche peut se penser dans une optique de recommandation, mettant en œuvre les corrélations éventuelles entre plusieurs jeux de données d'une part, et les similarités des profils des chercheurs d'autre part, en vue de leur proposer une offre plus exhaustive d'ensembles de données en relation.

Mots clés. Données ouvertes, système de recommandation, recherche scientifique, jeux de données.

Abstract

Open data represents an essential digital dataset in the scientific and technical research process, as long as access and use are left free for users. These datasets are available in open data warehouses as well as in specialized search engines. However, their indexing with descriptive metadata is generally reductive, which raises the problem of their identification by the research community. Optimizing access to this open data can be thought in a recommender system perspective, implementing correlations between several datasets on the one hand, and the similarities of researchers' profiles on the other hand, in order to bring them a more exhaustive data sets in relation with their scientific field of interest.

Keywords. Recommender system; supervised classification; Modeling; Library, ALS, algorithms, Spark, usage data.

1 Introduction

Selon la définition de l'OCDE¹, « la donnée ouverte de recherche est l'enregistrement factuel couramment considéré dans la communauté scientifique comme nécessaire à la validation des résultats de la recherche ». La disponibilité croissante de jeux de données ouvertes représente une source importante de matériaux de base essentiels à la production scientifique et technique. Toutefois, l'accès à ces données demeure encore délicat et fastidieux, dans la mesure où les chercheurs doivent formuler leurs requêtes sous forme de mots-clés qui ne peuvent être assez précis pour garantir un taux de récupération satisfaisant. D'un autre côté, l'indexation en amont de ces jeux de données, que ce soit par les producteurs ou par les professionnels de l'information, ne peut prétendre à une description exhaustive et précise vu la complexité et la technicité des données de la recherche.

A cet égard, nous pensons qu'un meilleur accès à ces données ouvertes peut se penser selon une démarche de recommandation, mettant en œuvre les corrélations entre jeux de données d'une part, et les similarités des profils des chercheurs d'autre part, en vue de leur suggérer d'autres jeux de données, en plus de ceux récupérés via une recherche multicritère classique.

2 Objectif

Notre objectif est de concevoir un modèle de système de recommandation (SR) adapté à la récupération de jeux de données ouvertes indexées dans un entrepôt de données. Ce SR devra permettre aux chercheurs de repérer des ensembles de données pertinents avec un taux de rappel supérieur aux systèmes traditionnels basés sur la logique booléenne.

3 Méthodologie

Pour atteindre notre objectif, nous commencerons par un benchmark des SR qui permettra de connaître les caractéristiques fonctionnelles des systèmes de recommandation, ensuite, nous évaluerons les systèmes de récupération actuels des principaux entrepôts de données ouvertes et enfin, nous proposerons une solution SR adaptée aux moteurs de recherche de l'open data. Ainsi, les actions suivantes seront menées :

1. Relever les bonnes pratiques et les fonctionnalités pertinentes en matière de SR ;
2. Dresser la liste des lacunes des systèmes actuels de récupération utilisés par les entrepôts de données ouvertes ;

¹Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics. OCDE, 2007 (accessible en ligne sur : <http://www.oecd.org/fr/science/sci-tech/38500823.pdf>)

3. Concevoir un modèle de recommandation orienté contenu, visant à établir les corrélations entre les jeux de données existants et entre les profils des chercheurs ;
4. Le prototypage du modèle.

Dès l'apparition du domaine des systèmes de recommandation, plusieurs approches et méthodes ont émergé, certaines ont été étudiées, expérimentées et comparées. Parfois, plusieurs terminologies sont utilisées pour désigner une même méthode ou approche. C'est pour cela que certains chercheurs se sont intéressés à la classification de ces méthodes et proposent une terminologie unifiée.

4 Etat de l'art

Un aperçu sur la littérature des systèmes de recommandation prouve qu'il y a un consensus en ce qui concerne la ventilation des systèmes de recommandation en trois principales approches: les méthodes basées sur le filtrage collaboratif, les méthodes basées sur le contenu et les méthodes hybrides.

4.1 Approche personnalisée

C'est une approche qui consiste à faire des recommandations à l'utilisateur sur la base de son comportement et usage passés, commençant par les produits achetés ou consultés, en passant par ses mouvements sur la toile (clics, navigation, etc.).

4.2 L'approche objet ou basée sur le contenu (Content-Based Filtering)

Il s'agit de proposer des contenus à partir des caractéristiques de ce contenu, tout en le corrélant avec les intérêts et les préférences des utilisateurs. En outre, cette approche utilise les attributs (les caractéristiques d'un objet), en vue de recommander des contenus similaires. A titre d'illustration, les thèmes généraux, et les mots clés d'un contenu / objet donné, sont comparés aux préférences de l'utilisateur à ce niveau. Pour ce faire, la recommandation objet est faite à la base des algorithmes qui vérifient si un contenu non encore exploré par l'utilisateur, pourrait éventuellement l'intéresser, comparativement à ce qu'il a évalué positivement dans le passé. Un système de recommandation axé sur l'objet ou le contenu peut fonctionner moyennant trois techniques distinctes:

1. Le système se contente de s'ancrer dans les généralités, par vérification si l'objet existe parmi les genres préférés par l'utilisateur, par le biais d'une similarité binaire de 0 et 1.

2. Le second algorithme de Dice se focalise plutôt sur les mots clés caractérisant le contenu en question avec la mesure du degré de la similarité et les croisements possibles entre les mots clés préférés par l'utilisateur et ceux relevés et extraits des différents contenus. Toutefois, cette technique n'est pas tellement efficace, puisque les mots clés ne sont pas toujours révélateurs du fond du contenu, d'autant plus que ce type pousse le système de recommandation à proposer des documents longs chargés par les mots qui apparaissent le plus souvent, peu importe s'ils sont représentatifs du contenu ou non.
3. Méthode de Rochchio (le retour de pertinence) privilégie l'idée de rester à l'écoute du feedback de l'utilisateur, en lui permettant de juger les recommandations proposées par le système, une fois que ce dernier reçoit les évaluations des utilisateurs, il les ventile en deux catégories : la première relative au jugement positif, et la seconde est liée à celui négatif. Par la suite, le centre de chaque groupe est calculé de telle sorte que la requête de l'utilisateur serait plus affinée et pertinente, du moment où il y a des contenus à retenir et d'autres à écarter.

Les avantages et les inconvénients représentés par cette dernière approche sont récapitulés dans le tableau ci-après.

Tableau 1: Les avantages et les inconvénients de la recommandation basée sur le contenu

Avantages	Inconvénients
<ul style="list-style-type: none">- Pouvoir fonctionner, quoiqu'il soit le seul utilisateur dans le système ;- Elle évolue au fur et à mesure avec le système lui-même.	<ul style="list-style-type: none">- Certains attributs peuvent être extraits automatiquement ;- La dose de la subjectivité qu'expriment certains éditeurs, quand ils relèvent les caractéristiques d'un objet donné ;- L'analyse des items par extraction des propriétés qui les marquent, et qui n'est pas toujours facile pour tout type d'items (musique, multimédia, image ou même d'autres items textuels) ;- L'overspecialisation, et qui renvoie à la nature de la recommandation se basant sur son profil, et tous les éléments hautement évalués dans ce dernier ne sont rien d'autre que ce qu'il avait connu avant.

Il est à signaler que cette approche de recommandation est moins utilisée comparativement à celle de voisinage proche (filtrage collaboratif), et que les imperfections de cette approche représentent inversement les atouts de la seconde.

4.3 L'approche de la recommandation sociale ou de filtrage collaboratif (Collaborative Filtering)

Cette approche définit l'utilisateur dans le contexte des autres utilisateurs du système. Partant de la logique de recherche des similarités, le système prédit ce que l'utilisateur apprécie pour chercher des utilisateurs similaires, tout en liant leur comportement d'usage du passé sur celui du futur.

En effet, le filtrage collaboratif s'opère de deux manières différentes :

- Une recommandation orientée utilisateur (voisinage proche) : Il s'agit d'un type d'utilisateurs regroupés en fonction de leurs similarités avec d'autres, en termes des préférences et des intérêts des objets à recommander, à partir du calcul du rating attribué par chacun des utilisateurs à des objets/contenus sous forme d'une matrice de deux axes, le premier correspond aux utilisateurs et le second renvoie au rating (notation) attribué aux objets par ces utilisateurs. Alors que le cas où des utilisateurs jugent de la même manière un contenu donné, ou dont la moyenne des ratings est pondérée, ils auront donc à recevoir des recommandations, sur ce qu'ils aiment et apprécient les uns et les autres.

Pour procéder ainsi, des algorithmes sont déployés, dont nous citons notamment celui de Pearson. Ce coefficient se calcule sur la base des valeurs figurant sur la matrice avec des valeurs extrêmes, le groupe représenté par «+1» signifie une forte corrélation positive, et celui représenté par «-1» correspond à une faible corrélation négative, c'est grâce à cette technique que les meilleurs voisins pertinents sont identifiés, et qu'un seuil de corrélation est défini pour sélectionner les utilisateurs voisins. Toutefois, il est à souligner que les utilisateurs peuvent être suivis du simple (like, dislike) sur la toile jusqu'à la surveillance du comportement de navigation sur un site web.

Pour finir, il faut dire que cette approche rend difficile l'action du système de recommandation en temps réel, quand il s'agit de gros sites, où il existe des millions d'utilisateurs. De ce fait, certains d'entre eux se sont vu implémenter des techniques de traitement des données hors ligne.

- Une recommandation orientée contenu : Contrairement à l'approche « usersnearestneighborhood », celle « d'Item-to-Item » s'axe sur le contenu et non pas l'utilisateur, partant du coefficient de Pearson toujours, cette méthode cherche la corrélation éventuelle entre des contenus similaires pour en proposer d'autres pareils à l'utilisateur. De même, les ratings donnés aux items (objets) sont calculés sur la base du critère de similarité. Cette approche présente des atouts, comme elle porte son lot d'imperfections²(voir le tableau suivant) :

Avantages	Inconvénients
<ul style="list-style-type: none">- Le privilège de ne pas avoir besoin des connaissances sur les contenus ;- La qualité des recommandations croit avec le temps en rapport avec l'évolution du nombre d'utilisateurs, avec la proposition des items de tout genre, sans que l'usager n'ait exprimé quelques items dans ses favoris.	<ul style="list-style-type: none">- Scalability : Le problème d'avoir sur des gros sites des millions d'utilisateurs, ce qui rend la recommandation en temps réel très difficile.- Cold start : Il s'agit du démarrage à froid, quand le système n'a pas au début assez de données sur l'utilisateur et les objets.- Sparsity : Le nombre des contenus est important, au point que parfois les utilisateurs se contentent de noter qu'une partie infime des contenus sur l'ensemble.- Le shilling : Il s'agit de l'action malveillante d'influencer les recommandations, en créant des faux profils pour voter dans le sens de favoriser ou défavoriser certains Items au détriment des autres.

² KEMBELLEC, Gérard et al. Systèmes de recommandation. In : Techniques de l'ingénieur, Novembre 2015, P.1-22.

Tableau 2: Les avantages et les inconvénients de la recommandation basée sur le filtrage collaboratif

Il est à signaler que cette approche de recommandation est plus commune que celle de voisinage proche, et que les imperfections de cette approche représentent les avantages de la première.

4.4 L'approche de la recommandation hybride (HybridFiltering)

La recommandation hybride est une sorte de combinaison des deux approches (objet et sociale), dans la perspective de combler les lacunes de chacune d'elles avec une collecte de caractéristiques des contenus, en allusion à l'approche orientée objet, et par la suite calculer le rating des utilisateurs dans le contexte d'une recommandation sociale, et éliminer ainsi les manquements des deux approches, comme en témoigne le cas de la plateforme d'e-commerce Amazon.

4.5 L'approche basée sur les connaissances (KnowledgeBased)

La base des connaissances acquise sur un domaine précis représente la matière première de la recommandation dans ce type de systèmes, en la confrontant aux connaissances et aux attentes des utilisateurs en termes de caractéristiques pour un item donné. C'est l'exemple des produits recommandés avec des spécifications précises liées au besoin des utilisateurs.

4.6 L'approche démographique (DemographicBased)

Cette famille de systèmes adapte la recommandation à la zone géographique de l'utilisateur (le pays, la ville ou le quartier). La personnalisation dans cette méthode concerne la langue d'affichage ou les items recommandés en fonction de leur particularité et disponibilité dans une zone donnée³.

5 Travaux liés

Il existe de nombreux systèmes de recommandations sous plusieurs sphères (e-commerce, musique, vidéos, presse, jeux, etc.). Quant aux réservoirs de données ouvertes de la recherche, nous soulignons qu'il existe un certain nombre de systèmes liés notamment aux bibliothèques numériques, dont nous citons :

- **BookPsychic**

BookPsychic est un système de recommandation simple pour les usagers de la bibliothèque, à l'image de Netflix ou Amazon, il permet d'évaluer les

³ Gaillard, Julien. Systèmes de recommandation: adaptation Dynamique et Argumentation. Diss. Université d'Avignon, 2014

livres et les DVD, et enregistre les préférences des utilisateurs, pour déboucher sur des listes de recommandation.

En effet, il s'agit de choisir une bibliothèque qui est inscrite dans book psychic, et noter les livres par genre, en vue d'obtenir des recommandations.



Figure 1 : Interface de BookPsychic, système de recommandation d'une bibliothèque numérique

• Huddersfield Book RecommenderSystem :

Il s'agit d'un système qui permet aux utilisateurs de contribuer explicitement dans la génération des recommandations, en ajoutant manuellement des titres similaires à des livres, ou de permettre à d'autres utilisateurs de suivre ce qu'ils prêtent à évaluent.



Figure 2 : Interface de Huddersfield Book RecommenderSystem (système de recommandation pour bibliothèques numériques)

- Ex LibrisBxRecommender

Les services d’information des bibliothèques universitaires à l’instar des bibliothèques publiques ont fait appel aux systèmes de recommandations, en utilisant le ScholarRank, qui établit un classement des ressources disponibles, et les résultats sont ordonnés selon la pertinence dépendant aussi bien de la spécialité de l’étudiant, que du cycle dans lequel il suit ses études⁴.

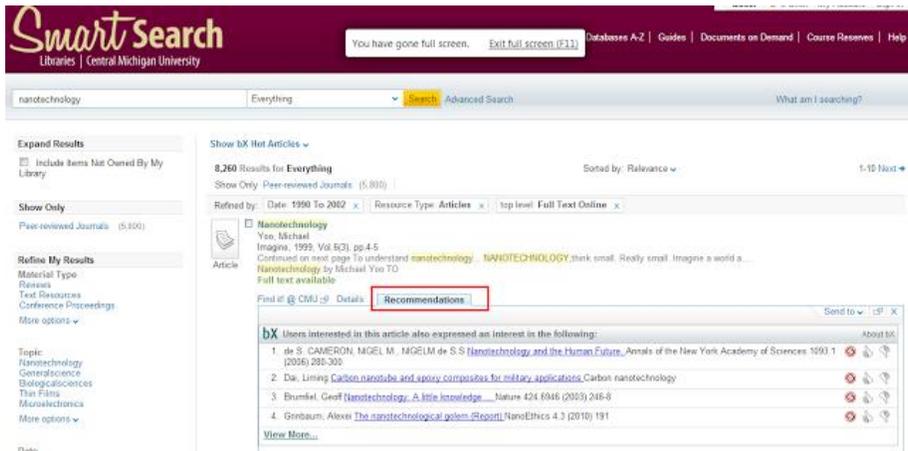


Figure 3 : Interface de Ex LibrisBxRecommender (système de recommandation pour bibliothèques numériques)

Critères / Exemples de solutions de SR	Bibliothèque physique	Bibliothèque numérique	Données explicites	Données implicites	Recommandation sur l'OPAC	Usager authentifié	Usager anonyme	Filtrage collaboratif à base de modèle	Filtrage collaboratif à base de mémoire
	BookPsychic	✗	✓	✗	✓	✗	✗	✓	✗
Huddersfield Book Recommender system	✗	✓	✗	✓	✗	✗	✓	✗	✓
Ex LibrisBxRecommender	✗	✓	✗	✓	✗	✗	✗	✗	✓

Tableau 3 : Comparatif des fonctionnalités des solutions de recommandations dans les services documentaires

⁴Isinkaye F. O., Folajimi Y. O., Ojokoh B. A. “Recommendationsystems: Principles, methods and evaluation.” EgyptianInformatics Journal, Volume 16, Issue 3, November 2015, Pages 261-273

Ainsi et après avoir dressé le comparatif ci-dessus, nous avons opté pour les choix suivants :

- Notre modèle est développé à la base des logs des usagers, représentant des données implicites, plutôt que leurs retours d'appréciation explicites constituant l'option de l'écrasante majorité des systèmes de recommandation déployés dans la sphère des industries culturelles et créatives en général, et les services d'information documentaire plus précisément, tout en combinant l'implicite à l'explicite en back-office de notre modèle et solution à travers une conversion ;
- Notre modèle tient compte de deux cas de figure à savoir, un usager authentifié disposant déjà d'un compte et celui d'un usager anonyme ;
- Notre modèle est implémenté sur le réservoir des données de recherche ouvertes et non pas sur une plateforme dédiée ;
- Notre modèle repose sur l'approche de filtrage collaboratif à base de modèle (faisant appel aux données de collecte, pour construire un modèle de prédiction qui devrait être opérationnel sans avoir besoin de consulter les données de départ).

6 Conception du modèle de recommandation

Afin d'orienter notre travail nous avons fixé quatre objectifs principaux pour notre solution :

1. Exploiter les données implicites des usagers et les convertir en données explicites
2. Profiling des usagers
3. Créer un modèle de recommandation et l'utiliser pour pouvoir générer des recommandations
4. Interfaçage des résultats de recommandations

Ces objectifs constituent des modules qui composent notre application.

Il nous a été demandé par le service informatique de développer une interface pour pouvoir visualiser ces recommandations et tester le système. Une fois testé, le service pourra s'occuper de son intégration dans le réservoir des données ouvertes de la recherche⁵.

⁵Kembellec, *Gérald; Chartron, Ghislaine; Saleh, Imad. Les moteurs et systèmes de recommandation*, ISTE Editions, 2014

Notre solution se base sur les fichiers logs issus des traces d'interaction des chercheurs avec le système, pour leur proposer des contenus qui peuvent leur être pertinents. Etant basé sur le filtrage collaboratif comme approche de recommandation, notre système va permettre :

- D'échanger les idées entre usagers en faisant profiter les uns des préférences des autres.
- De favoriser le transfert des connaissances
- De faciliter la sélection de l'information.

6.1 Scénario de recommandation

Le scénario de recommandation que l'application doit garantir fait distinguer deux cas de traitement :

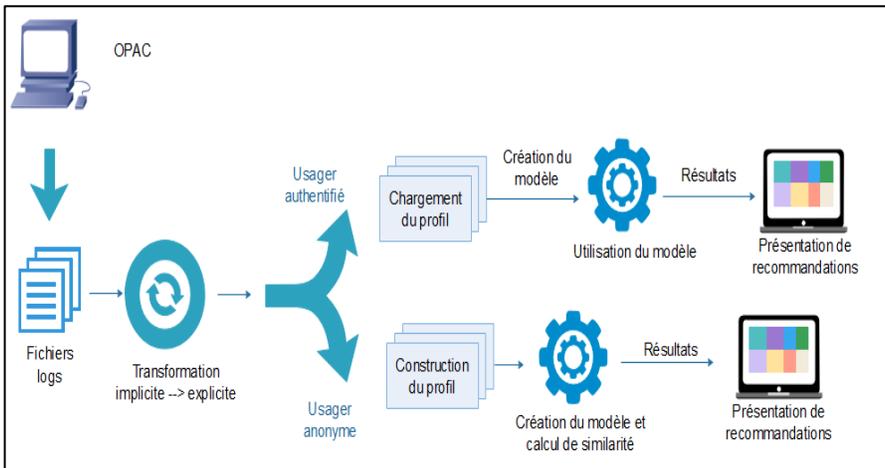


Figure 4 : Schéma du processus de recommandation à travers l'application.

Un utilisateur abonné peut s'authentifier via son login et mot de passe que l'administrateur lui a déjà défini au moment de son enregistrement. Le système charge son profil et lui présente des recommandations en se basant sur ses besoins antérieurs. A travers une analyse de similarité à base de profil, le système lui affiche des recommandations personnalisées⁶.

⁶Jinjiao Lin, Haitao Pu, Yibin Li, JianLian. "Intelligent Recommendation System for Course Selection in Smart Education." *Procedia Computer Science*, Volume 129, 2018, Pages 449-453 ? Smith, éditions FMSH, à paraître. Accessible en ligne : <http://halshs.archives-ouvertes.fr/halshs-00990771>

6.2 Architecture conceptuelle de la solution

Notre système exploite les données contenues dans les logs, à travers un processus de plusieurs étapes. Pour pouvoir générer des recommandations, le processus se fait en plusieurs étapes :

6.2.1 L'extraction de fichiers logs

Dans notre cas, et pour faire des simulations de recommandation, on utilisera un jeu de données que le service a généré depuis les fichiers logs, qui contient des informations de navigation d'usagers structurées comme suit :

- Identifiant de l'utilisateur : identifiant donné par le système à un usager connecté, qui permet de l'identifier.
- Dernier document consulté : c'est le document qui met fin à la recherche, c'est le besoin que l'utilisateur cherche à satisfaire.
- Nombre de rebond de la requête : nombre de requêtes effectuées avant de trouver le jeu de données voulu, ce nombre nous renseigne sur l'intérêt que cet usager actif a vis-à-vis du document trouvé.
- Affichage du document dans les résultats : nombre de fois que le jeu de données est retourné dans les résultats.
- Le temps passé sur l'entrepôt de données ouvertes : c'est le temps investi dans la recherche.

6.2.2 La Transformation des données implicites en données explicites

Comme déjà mentionné dans l'état de l'art, le système de recommandations peut se baser sur deux types d'informations :

- *Des informations explicites* qui représentent les notations que les utilisateurs ont attribuées aux jeux de données ouvertes, sur une échelle allant de 1 à 10 points, celles-ci sont directement exploitables par le système de recommandation.
- *Des informations implicites* qui caractérisent le comportement de navigation de l'utilisateur, que le réservoir collecte durant la navigation. Il s'agit du temps de connexion sur le réservoir, le nombre de rebonds de la requête et le nombre d'affichages du jeu de données correspondant dans les résultats.

Les informations implicites collectées ne sont pas directement exploitables puisqu'elles ne sont pas homogènes, le temps par exemple, ne peut pas être rapproché au nombre de rebonds de la requête, ni au nombre d'affichages du jeu de données dans les résultats.

La méthode répandue dans le traitement des données implicites est de les transformer en notations semblables à celles collectées explicitement⁷.

Le schéma ci-dessous présente les étapes du processus de transformation de données implicites en notations :

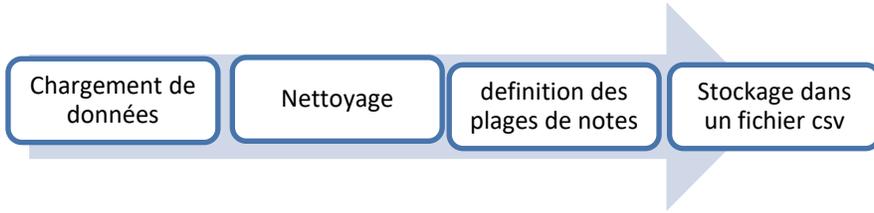


Figure 5: Processus de transformation de données implicites en notations

Pour convertir les données implicites en explicites, nous commençons par nettoyer les données, définir les plages de notes à appliquer, puis stocker le fichier dans un format standard qu'on pourra exploiter par la suite.

Les données employées sont souvent sous format texte (fichier log), les informations contenues dans ce fichier log ont été exportées sous forme de base de connaissance.

6.3 Le nettoyage de données

C'est l'opération de détection et correction des erreurs présentes dans le fichier, selon plusieurs étapes :

- **Traitement des valeurs manquantes :** à travers l'analyse du fichier, on a trouvé que 3 enregistrements contenant des valeurs manquantes, puisqu'il s'agit d'une faible portion, qui représente moins de 5% de la taille du fichier, nous avons enlevé ces enregistrements et nous n'avons conservé que les enregistrements complets.
- Nous avons limité les sessions de connexion à 30 minutes, ce qui est logique, vu qu'un utilisateur ne pourra pas se connecter au-delà de cette durée, et nous avons éliminé toutes les valeurs qui contiennent des durées dépassant 30 minutes. Le fichier résultant contient 500 enregistrements, après avoir été de l'ordre de 600.

⁷OARD, Douglas W., KIM, Jinmook, et al. Implicit feedback for recommender systems. In Proceedings of the AAAI workshop on recommender systems. 1998. p. 81-83.

6.4 Définition des plages de notes

Pour définir les plages de notes, nous devons diviser les données en intervalles de même amplitude. Le nombre d'intervalles est conditionné par l'échelle de notation à adopter. Ainsi, deux échelles de notations se présentent : l'une sur 5 points et une autre sur 10 points. Cette dernière a été adoptée pour garantir plus de précisions en ayant des intervalles de 10 au lieu de 5. Chaque intervalle est associé à une note, ce qui rend plus fiable l'estimation de la notation qu'on se propose d'établir.

Après la définition des plages de notes, on remplace chaque valeur contenue dans la plage par la notation correspondante, comme illustré dans le schéma ci-dessous :

	Nombre de rebond de la requête		Affichage du document dans les résultats		Le temps passé sur l'OPAC	
	Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure
Note 1	[0;	0,8[[0;	2,3[[5;	7,4[
.	[0,8;	1,6[[2,3;	4,6[[7,4;	9,8[
.	[1,6;	2,4[[4,6;	6,9[[9,8;	12,2[
.	[2,4;	3,2[[6,9;	9,2[[12,2;	14,6[
.	[3,2;	4[[9,2;	11,5[[14,6;	17[
.	[4;	4,8[[11,5;	13,8[[17;	19,4[
.	[4,8;	5,6[[13,8;	16,1[[19,4;	21,8[
Note 8	[5,6;	6,4[[16,1;	18,4[[21,8;	24,2[
.	[6,4;	7,2[[18,4;	20,7[[24,2;	26,6[
Note 10	[7,2;	8]	[20,7;	23]	[26,6;	29]

Tableau 4: Définition des notations correspondant à chaque plage de données

Cette opération garantit l'homogénéisation des données, puisqu'on aura des notations sur une échelle unifiée.

6.5 Stockage sous format final

L'étape suivante étant de calculer une note générale qui prend en compte les notations produites. Après des discussions avec l'administrateur de base de données, on s'est mis d'accord sur le fait que « Affichage du document dans les résultats » aura une pondération élevée, puisqu'il exprime au mieux l'intérêt de l'utilisateur vis-à-vis du document.

La formule adoptée pour calculer la note générale :

$$\text{Note générale} = \text{Note (Nombre de rebonds de la requête)} * 30\% + \text{Note (Affichage du document dans les résultats)} * 40\% + \text{Note (Le temps passé)} * 30\%.$$

Le fichier résultant sera ainsi sous cette forme :

Id de l'utilisateur	Document consulté	Note globale estimée
1	Doc 1	Note 1
2	Doc 2	Note 2
...

Tableau 5: Format final du fichier à exploiter

Ce fichier va être transformé par la suite en fichier CSV que l'application exploitera.

7 La validation du modèle

7.1 Authentification de l'utilisateur

L'utilisateur s'authentifie à travers l'application à l'aide de son identifiant et son mot de passe pour bénéficier des recommandations liées à son profil. Une fois authentifié, l'identifiant de cet utilisateur sera communiqué à Spark ou se poursuivront les traitements.

Cette interface permet à l'utilisateur d'effectuer des recherches sur la base de données. Une fois une recherche lancée, il reçoit des recommandations qui lui sont appropriées. La figure ci-dessous montre un aperçu sur les résultats de recherche accompagnés de recommandations.

7.2 Calcul des fonctions d'erreur (RMSE, MSE, K MAPK)

Les mesures les plus utilisées dans la littérature sont : « les fonctions d'erreurs » leur rôle est de mesurer la « distance » moyenne entre les prévisions estimées et les observations correspondantes. Ainsi, on peut dire qu'une prédiction est parfaite tant que la valeur de la fonction est proche de 0. Les fonctions les plus populaires dans l'évaluation des systèmes de recommandation, et qu'on peut exécuter sur Spark, sont :

- *La fonction d'erreur moyenne quadratique (MSE : « Mean Squared Error »)* c'est une fonction qui calcule le carré de la distance entre la prévision et l'observation. Elle est connue par sa grande sensibilité à la précision.
- *La fonction RMSE : « Root Mean Squared Error »* : c'est la racine de l'erreur moyenne quadratique (MSE), elle est largement utilisée à la place du MSE pour évaluer les systèmes de recommandation.
- *La précision moyenne en K MAPK « Mean average precision at K »* : mesure les scores de pertinence moyenne d'un ensemble de

documents top-K présentés en réponse à une requête. Cette mesure est la plus adaptée aux données implicites.

Après plusieurs tests, en jouant sur les paramètres du modèle, on s'est arrêté sur des valeurs optimales qui minimisent le RMSE.

La fonction RMSE n'est pas une bonne mesure pour prédire la préférence des utilisateurs, notamment lorsqu'il s'agit des données implicites. Cette mesure doit être combinée et, dans certains cas, remplacée par d'autres mesures issues de la récupération d'informations telles que la précision⁸. On tiendra compte plutôt dans notre cas des valeurs de la précision moyenne en k, et d'après les résultats de cette mesure qui tendent vers 0, la marge d'erreur de notre système est minime.

7.3 Calcul du temps d'exécution du programme

Un autre critère qu'on peut utiliser pour évaluer la performance de notre système est de connaître le temps d'exécution du script de recommandation. On a procédé par plusieurs tests de calcul du temps d'exécution, et on a pris les mesures ci-dessous :

Test	Durée d'exécution(s)
1	6.64
2	6.40
3	6.23
4	6.12
5	6.37
6	5.83
7	5.48
8	5.59
9	5.71
10	5.20

Tableau 7: valeurs des tests de durées d'exécution du script de recommandation

⁸PARRA, Denis, KARATZOGLOU, Alexandros, AMATRIAIN, Xavier, et al. Implicit feedback recommendation via implicit-to-explicit ordinal logistic regression mapping. Proceedings of the CARS-2011, 2011.

La moyenne des valeurs générées par le test est égale à 5.95 secondes, donc notre système offre les recommandations dans plus au moins 6 secondes.

8 Conclusion

Notre objectif principal était de concevoir et réaliser un prototype d'un moteur de recommandation basé sur les données implicites des usagers des entrepôts de données ouvertes, pour leur recommander des jeux de données susceptibles de les intéresser au diapason avec leur profil et leurs centres d'intérêt.

Pour ce faire, il était question d'explorer quelques champs d'application du machine learning dans les services documentaires dédiés à la recherche, en particulier les systèmes de recommandations et les algorithmes sous-jacents.

Notre étude a abouti au prototypage et l'implémentation d'un système de recommandations adapté au contexte des entrepôts de données ouvertes. De fait, le système a été implémenté sur Apache Spark en langage Scala. Notre système procède à une analyse des données de navigation des usagers à travers plusieurs algorithmes, notamment l'algorithme ALS, pour ensuite présenter les résultats de l'analyse sous forme de recommandations. De plus, le système distingue entre un usager abonné et un usager anonyme en termes de recommandation, car si le premier bénéficie des suggestions de jeux de données, selon son besoin identifié par le système, le second usager (anonyme) bénéficie également des recommandations, mais sur la base des notices qu'il consulte.

Sur un autre registre, des améliorations de notre travail peuvent être envisagées ; il s'agit d'automatiser le processus de transformation de données implicites en données explicites, et à intégrer une composante qui permettra de tirer parti des nombreuses données accumulées par le système de recommandation.

Bibliographie

GAILLARD Rémi. « De l'Open data à l'Open research data : quelle(s) politique(s) pour les données de recherche ? » Mémoire d'étude / janvier 2014, sous la direction de Monique Joly. Université de Lyon

INRA Institut national de la recherche agronomique, Rapport du groupe de travail sur la gestion et le partage des données, Paris, juin 2012, 62 p.

ISINKAYE F. O., FOLAJIMI Y. O., OJOKOH B. A. “Recommendation systems: Principles, methods and evaluation.” *Egyptian Informatics Journal*, Volume 16, Issue 3, November 2015, Pages 261-273

JEROME Denis et Samuel GOËTA. « La fabrique des données brutes. Le travail en coulisses de l’open data », in *Penser l’écosystème des données. Les Enjeux Scientifiques et Politiques des données numériques*, dir. Mabi Plantin Monnoyer

JINJIAO Lin, HAITAO Pu, YIBIN Li, JIAN Lian. “Intelligent Recommendation System for Course Selection in Smart Education.” *Procedia Computer Science*, Volume 129, 2018, Pages 449-453 ? Smith, éditions FMSH, à paraître. Accessible en ligne : <http://halshs.archives-ouvertes.fr/halshs-00990771>

KEMBELLEC, Gérard; CHARTRON, Ghislaine; SALEH, Imad. *Les moteurs et systèmes de recommandation*, ISTE Editions, 2014

KOP, R. The Challenges to Connectivist Learning on Open Online Networks : Learning Experiences during a Massive Open Online Course. *European Distance and E-learning Network annual Conference. 2010*, p.4- 32.

LECHANI, Tamine Lynda et Mohand, BOUGHANEM. *Accès personnalisé à l’information : Approches et techniques. Rapport interne*. RIT, 2005. Disponible sur :

<ftp://ftp.irit.fr/IRIT/SIG/rapport_Perso_0904_VF.pdf>
[consulté le 6 Octobre 2016].

NAAK, Amine. *Papyrus : Un système de gestion et de recommandation d’articles de recherche. Mémoire d’obtention de maîtrise en sciences informatiques*. Université de Montréal, 2009

STEINBROOK, R and KASSIRER, JK. Free the data. *BMJ*. 2010; 341: 811–812

Le libre accès aux données de recherche : quelle politique de données pour les universités marocaines ?

Open Research Data: which Data policy for Moroccan universities?

Nadia BENAÏSSA, Université Abdelmalek Essaâdi. Maroc

Saïda BELOUALI, Université Mohamed Ier. Maroc

Résumé

Cette communication s'inscrit dans la perspective des études sur le libre accès aux résultats de recherche dans le contexte marocain en se focalisant sur l'effervescence autour de l'Open Research Data. L'objectif étant de jeter les bases pour une expérience marocaine en matière de l'ouverture des données de recherche dans les établissements d'enseignement supérieur.

Dans le cadre épistémologique de la recherche, nous nous sommes basé sur le modèle de la traduction ou la théorie de l'acteur-réseau.

Les résultats d'une enquête sur le terrain menée auprès des enseignants chercheurs nous ont permis de dresser un état des lieux des pratiques de gestion des données de recherche (données sources et données produites) et la prédisposition de ces acteurs principaux à s'intégrer dans cette tendance nouvelle.

Mots-clés : Science ouverte, Accès ouvert aux publications, ouverture des données de recherche, Entrepôt de données, Plan de gestion de données.

Abstract

This paper is part of the studies on open access in the Moroccan context, focusing on the growing interest on Open Research Data.

The aim is to lay the foundations for a Moroccan experience in opening up research data in higher education institutions.

In the epistemological framework of research, we rely on the actor-network theory.

The results of a field survey conducted among research professors enabled us to draw up an inventory of research data management practices (source data and data produced) and the predisposition of these main actors to integrate into this new trend.

Keywords: Open Science, Open Access, Open Research Data, Data Repository, Data Management Plan.

1 Introduction

Le tournant numérique pris par la sphère scientifique a transformé les modalités de production et de communication savante.

Assurément, le mouvement du libre accès aux résultats de recherche, venu contre l'hégémonie des éditeurs scientifiques commerciaux, a révolutionné le mode de publication des chercheurs universitaires.

Dépassant la mise en libre accès de la production scientifique sous forme d'articles, de thèses, ou de rapports, la culture du partage entre chercheurs s'est étendue aux données même issues de la recherche ; les données que le chercheur a utilisées pour aboutir à ses résultats.

Au Maroc, malgré que le mouvement du libre accès soit toujours à un stade embryonnaire, plusieurs initiatives éparses ont vu le jour dans certaines universités et centres de recherche et ce, à travers la mise en place d'archives institutionnelles et de revues à accès libre. Plusieurs expériences en témoignent, à savoir : l'expérience de l'Université Mohammed V, l'expérience de l'Université Abdelmalek Essaâdi, l'expérience de l'Université Hassan II, l'expérience du Centre National de la Recherche Scientifique et Technique, pour ne citer que celles-là. La réussite, même partielle, de ces projets au sein des universités, met en évidence la prédisposition des différents acteurs de la recherche scientifique marocains (Enseignants chercheurs, Présidents, Responsables des établissements universitaires, responsables des structures de recherche) à s'inscrire dans le mouvement et à s'aligner aux nouvelles tendances éditoriales qui s'imposent avec acuité dans le monde de la recherche.

Néanmoins, et au sortir de la nouvelle loi marocaine sur le droit d'accès à l'information (B.O. N° 6655, 2018), nous posons-nous certaines interrogations sur la faisabilité et l'applicabilité de l'Open Research Data dans le contexte scientifique et universitaire marocain ? Les mêmes acteurs qui ont exprimé leur position en faveur du libre accès aux résultats de la recherche, sont-ils prédisposés à aller encore de l'avant et mettre en libre circulation les données produites au cours de leurs processus de recherche ? Quels sont les types de données de recherche concernés par cette ouverture et ceux présentant des exceptions ? Quelles sont les contraintes qui s'imposent aux établissements d'enseignement supérieur et de recherche marocains ? Comment peut-on convaincre les responsables des établissements universitaires à établir une politique de données et l'intégrer dans leur politique institutionnelle de recherche ? et enfin, quels sont les besoins des enseignants chercheurs en termes de formation et d'initiation à ce nouveau procédé ?

Dans une première partie nous présenterons une revue de littérature succincte embrassant les différentes facettes de l'Open Research Data, ensuite nous dresserons le contexte dans lequel se positionne notre étude,

par la suite nous présenterons le cadre méthodologique de la recherche et l'analyse des résultats de l'enquête menée auprès des enseignants chercheurs, enfin, nous traiterons l'aspect juridique de l'ouverture des données de recherche dans la loi marocaine.

2 De l'Open Science à l'Open Research Data, une suite de pensées.

A l'aune de la révolution numérique, l'ensemble des secteurs de la société moderne ont connu de profondes mutations ou en passe de l'être. S'il est un secteur dont les fondements même sont remis en question c'est bien celui de l'enseignement supérieur et de la recherche scientifique. Cette mutation a pris son essor avec l'avènement des initiatives de l'open access qui opère sous deux stratégies différentes et complémentaires ; les archives ouvertes (voie verte) et les revues alternatives (voie dorée). Ce courant avant-gardiste fait partie d'une discipline plus générale qu'est l'Open Science.

Dans cette perspective, Blanchard A. conçoit la science comme un bien commun où les modes de communication scientifique sont ouverts à la discussion de tout un chacun. Il met dans cette discipline les publications en accès libre, les archives ouvertes, les carnets de recherche, la science participative, la science citoyenne et l'Open Data (Blanchard A., 2016).

L'Open Science, selon Bartling S. (2014), se réfère à une culture scientifique caractérisée par son ouverture. Les scientifiques partagent immédiatement les résultats de leurs recherches avec leurs pairs favorisant ainsi une communication ouverte et large à travers des outils de diffusion de la connaissance, nouveaux et crédibles.

Les études menées sur l'Open Science ont fait naître plusieurs écoles de pensée. B. Fecher et S. Friesike distinguent cinq écoles de pensée : l'école démocratique, l'école pragmatique, l'école basée sur l'infrastructure, l'école publique et enfin, l'école basée sur la mesure et l'évaluation des chercheurs et de la recherche. Le tableau N°1 nous explique l'objectif principal de chaque école, les acteurs impliqués et les outils et méthodes par lesquels les objectifs seront atteints (B. Fecher et S. Friesike, 2014).

Ecole de pensée	Postulat central	Groupes impliqués	Objectif	Outils et méthodes
Démocratique	L'accès au savoir est inégalement réparti	Scientifiques, politiciens, citoyens	Rendre les connaissances disponibles gratuitement pour tous	Open access, droits de la propriété intellectuelle, Open Data, Open code
Pragmatique	La création de connaissances pourrait être plus efficace si les scientifiques travaillaient ensemble	Scientifiques	Ouvrir le processus de création de connaissances	Sagesse des foules, effets de réseau, Open Data, Open Code
Infrastructure	Une recherche efficace dépendrait des outils et des applications disponibles	Scientifiques et Fournisseurs de plates-formes.	Créer des plates-formes, des outils et des services accessibles aux scientifiques	Plates-formes et outils collaboratifs
Publique	la science doit être accessible au public	Scientifiques et citoyens	Rendre la science accessible aux citoyens	Science citoyenne, Science PR, Science Blogging
Measurement	Les contributions scientifiques d'aujourd'hui nécessitent des mesures d'impact alternatives	Scientifiques et politiciens	Développer un système métrique alternatif pour le calcul de l'impact scientifique	Altmetrics, peer review, citation, impact factors

Tableau N°1 : Les écoles de pensée selon la vision de B. Fecher et S. Friesike sur l'Open Science

L'accès au savoir commence à se démocratiser. Il est rendu ouvert et disponible gratuitement à tout chercheur et citoyen à travers de nouveaux procédés et modes de publication qui ne se limitent pas seulement aux résultats de recherche mais s'étendent, ces dernières années, aux données de recherche.

En effet, les débats sur l'ouverture des données de recherche prennent de plus en plus de l'ampleur. Une seconde révolution touche le domaine de l'édition scientifique et de la communication savante après la quasi-généralisation de l'Open Access dans les institutions de recherche des pays occidentaux. L'enjeu étant de rendre les données produites dans un processus de recherche réutilisables pour une meilleure garantie de la scientificité d'une étude, d'un postulat, d'une recherche, etc. Les arguments de poids avancés par les adeptes de ce courant consistent aussi à réduire les cas de fraudes scientifiques à travers la vérification et la reproductibilité des résultats publiés.

2.1 Données de recherche, essai de définition

Selon l'OCDE (2007), les données de recherche sont « des enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider des résultats de recherche ».

Gaillard R. (2014) fait la distinction entre plusieurs types de données. Il cite la définition de l'ANDS (2013) qui stipule que « Les données [de recherche] peuvent être des données brutes, des données non traitées d'observations de phénomènes particuliers. D'autres sont des données traitées, données produites après formatage ou correction de données brutes. D'autres, des données dérivées, qui présentent un résumé ou une présentation spécifique des données brutes ».

On distingue les données d'observation, les données expérimentales, les données computationnelles ou de simulation, les données statistiques sous format numérique, etc.

Puren M. (2017) énumère quelques types de données de la recherche comme suit :

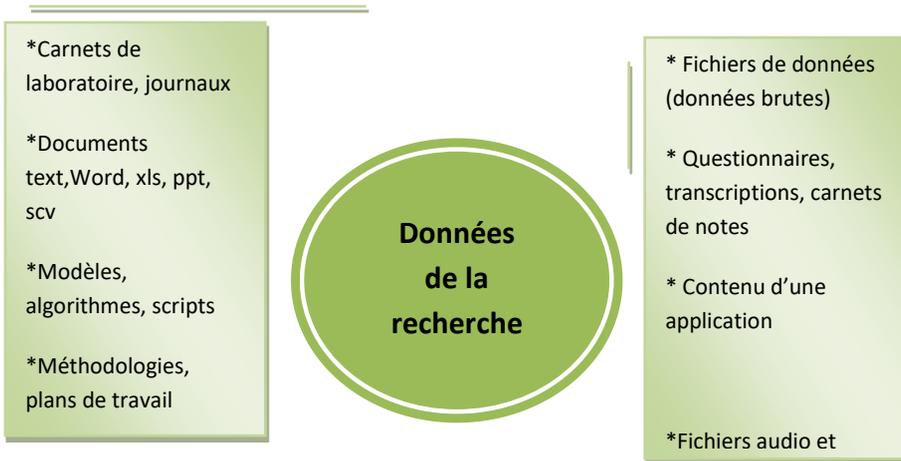


Figure N°1 : Exemples de données de recherche

Source : M. PUREN, INRIA, 2017

Ce sont, de ce fait, des données susceptibles d'être collectées, créées, analysées et stockées sous différents formats numériques.

L'enregistrement de ces données sous la forme d'un ou de plusieurs fichiers électroniques, téléchargeables, citables (DOI) et intelligibles est appelé « Jeu de données » (Balanchard R., 2014) ou « Dataset ».

Gaillard R., définit le jeu de données comme étant « l'agrégation, sous une forme lisible, de données brutes ou dérivées présentant une certaine « unité », rassemblées pour former un ensemble cohérent ».

2.2 La gestion des données de recherche

La gestion des données de recherche englobe toutes les actions de collecte, de traitement, d'analyse, de stockage et d'ouverture pour des fins de réutilisation et de vérification. Le schéma du cycle de vie des données nous résume cette chaîne de production (Figure N°2) :

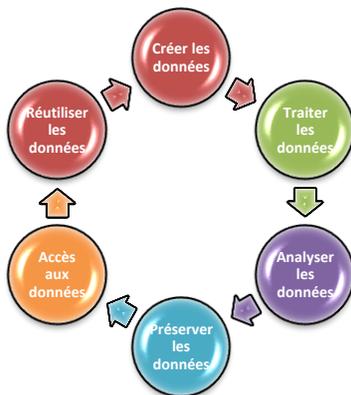


Figure N° 2 : le cycle de vie des données

Source : UK Data Archive, Research Data Life Cycle

Il est donc primordial que les universités et les structures de recherche marocaines élaborent un dispositif de gestion des données liées à la recherche, du moins pour celle financée sur fonds public et de garder une traçabilité des travaux de recherche pour une ultime vérification et réutilisation.

Un travail étroit devrait être établi entre les chercheurs et les services informatique de chaque université pour garantir le stockage, l'échange et la pérennisation des données et des résultats de recherche.

2.3 Les acteurs liés à l'ouverture des données de recherche

Plusieurs acteurs sont étroitement concernés par l'ouverture des données de recherche et ce, à chaque étape du cycle de vie de la donnée. On peut citer :

- Les chercheurs eux même puisqu'ils sont la principale source et les principaux utilisateurs de ces données. Peuvent intégrer le clan des

chercheurs : les enseignants chercheurs universitaires, les chercheurs des centres de recherche et les doctorants ;

- Les établissements d'enseignement supérieur ;
- Les structures de recherche (laboratoires et équipes de recherche) ;
- Les organismes financeurs de recherche
- Les éditeurs ;
- Les bibliothèques universitaires et de recherche ;
- Les entreprises bénéficiaires des avancées de la recherche ;
- Les fournisseurs d'infrastructures de recherche (OpenAIRE et OpenAIRplus) ;
- Les services informatiques.

2.4 Le plan de gestion de données ou Data Management Plan (DMP)

Pour la mise en œuvre d'une politique d'ouverture des données de recherche, il est important de dresser les normes et les modalités que tout acteur impliqué doit respecter. Un plan de gestion de données est alors dressé pour déterminer Qui ?, fait Quoi ?, Comment ? et par Quel instrument ? Le Data Management Plan (DMP) dressera le rôle de chaque acteur, les données susceptibles d'être ouvertes et celle connaissant des restrictions, le format des fichiers, les modes d'archivage, la sécurité, les coûts et les métadonnées descriptives.

L'INIST définit le DMP comme étant un « document rédigé au commencement d'un projet de recherche et qui définit ce que les chercheurs feront de leur données pendant et après le projet, explicitant notamment la mise à disposition des données ». Ce document formel répond souvent aux directives des organismes financeurs de la recherche. Il permet ainsi de planifier les différentes étapes de la gestion des données et les dates butoirs.

Bien gérer ses données faciliterait leur recherche, augmenterait leur visibilité et accessibilité et par conséquent, accroîtrait leur taux de citation par les pairs. On parle ainsi des FAIR Data, diminutif de : Findable (données trouvables grâce aux métadonnées descriptives et aux identifiants), Accessible (données accessibles après autorisation), Interoperable (données interopérables sous formats ouverts) et Reusable (données réutilisables en respectant les closes du droit d'auteur, du droit d'accès aux données et les licences Creative Common's).

Par la suite vient l'étape de l'archivage, de la conservation et du partage des données et ce, à travers un entrepôt de données.

2.5 L'entrepôt de données

Appelé aussi « réservoir de données », c'est une plateforme permettant la collecte, la description, la conservation, la recherche et la diffusion des jeux de données. Les données brutes ou élaborées sont décrites par des métadonnées pour leur garantir une meilleure accessibilité (Aventurier P., 2013).

On distingue les entrepôts de données thématiques, disciplinaires, multidisciplinaires ou institutionnels. Ces différents types d'entrepôts de données sont classés dans des annuaires pour aider les chercheurs et les institutions à trouver le réservoir de données qui répond à leurs besoins tels que Re3data et Databib qui énumère les entrepôts de données multidisciplinaires.

2.6 La valorisation des données de recherche

Outre le dépôt dans des entrepôts de données, les données de recherche peuvent être valorisées en les publiant comme matériel supplémentaire en lien avec l'article ou en les publiant dans des Data papers dont la description des données se fait d'une manière structurée et lisible. Ce type de publication est citable tout comme l'article scientifique (Inist, 2018).

3 Contexte de recherche

Le Maroc n'est pas en reste des orientations internationales en matière du libre accès aux résultats de recherche. Plusieurs initiatives ont été menées il y a plus d'une décennie par des chercheurs universitaires et militants du nouveau mode de publication afin de sensibiliser les universités et les tutelles à suivre les tendances internationales en matière de communication savante. Certaines universités, leader en matière d'introduction des nouvelles technologies de l'information dans leur circuit de gouvernance et de pilotage de la recherche scientifique, ont mis en place des archives institutionnelles et des revues en libre accès pour garantir un meilleur accès et une meilleure visibilité de leur patrimoine scientifique.

Les actions dans le contexte marocain sont, certes, timides et restreintes mais n'est-il pas opportun d'améliorer les expériences réussies en matière de l'open access et l'étendre à l'ouverture des données de recherche ?

Pour se faire, nous avons pris le cas de l'Université Abdelmalek Essaâdi. Une université qui mène une politique de recherche axée sur les systèmes d'information d'aide au pilotage de la recherche scientifique et où la gouvernance numérique de la recherche scientifique est la devise.

Signalons dans ce sens qu'une enquête préalable a été menée dans deux institutions à vocation scientifique et technique de cette université qui avait pour objectif d'étudier la prédisposition des différents acteurs de la

recherche scientifique à s'intégrer dans un projet d'archive institutionnelle et de sensibiliser les décideurs à l'introduire dans leur politique de recherche.

Il nous semble judicieux de mener notre enquête sur l'ouverture des données de recherche auprès de cette même population déjà initiée.

Rappelons dans ce contexte que l'Université Abdelmalek Essaâdi, créée en 1989, réunit 960 enseignants chercheurs, 4790 doctorants et 122 structures de recherche.

Dans l'objectif de soutenir ses enseignants chercheurs et de valoriser leur recherche scientifique, l'UAE a mis en place un outil de pilotage de la recherche scientifique. Il s'agit d'un système d'information qui permet d'obtenir une description de l'existant et une étude des besoins, pour aboutir à une optimisation des ressources et un meilleur engagement dans le processus de l'autoévaluation et de l'évaluation par les pairs. Le système d'information permet de bénéficier, en continu, des données relatives aux activités de recherche des différents acteurs de l'université (enseignants chercheurs, structures de recherche, etc) et de bâtir une conception partagée de son potentiel de recherche en le positionnant dans son environnement régional, national et international. Il dispose d'une application qui intègre la nomenclature et la base de données des indexes de citation : Science Citation Index, Social Science Citations Index et Art & Humanities Citations Index, dans l'objectif ultime d'enregistrer et de valoriser les publications des enseignants chercheurs de l'UAE parues dans les revues internationales indexées. Le système permet aussi de favoriser une meilleure visibilité des activités de recherche de l'université par le biais d'une passerelle automatique des données sur l'espace numérique de l'UAE comme indiqué dans la figure N° 3.

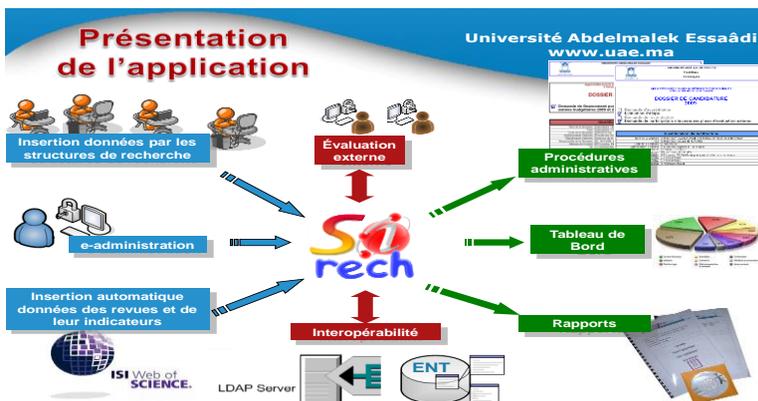


Figure N° 3 : Le système de pilotage de la recherche scientifique à l'UAE

Dans le même sillage, et suite à l'enquête entreprise sur l'open access auprès d'une population d'enseignants chercheurs, de décideurs, d'informaticiens et de bibliothécaires des deux institutions de l'Université Abdelmalek Essaâdi, une archive institutionnelle a été mise en place pour permettre aux chercheurs de déposer leurs contributions. Les décideurs ont opté pour une politique incitative, en premier lieu, suivie d'une solution s'appuyant sur l'évaluation : seules les productions scientifiques contenues dans l'archive institutionnelle seront évaluées. L'archive institutionnelle a été programmée sous Invenio V1.0.0-rc0, un logiciel qui permet à la fois de gérer une archive ouverte de grande taille et aussi de gérer les opérations bibliothéconomiques. L'installation a été effectuée sous Linux, sous la distribution « Ubuntu » version 10.x. comme le montre la figure N°4 :



Figure N° 4 : L'archive institutionnelle de l'UAE

4 Méthodologie de recherche

Une enquête sur le terrain a été menée auprès d'une population constituée de 156 enseignants chercheurs. L'outil de collecte de données que nous avons jugé adéquat est le questionnaire. Ainsi, sur 156 questionnaires distribués, 92 enseignants chercheurs ont répondu, soit un taux de réponse de 59 %. Notre enquête sur l'ouverture des données de recherche vise à répondre à trois objectifs de recherche :

- Dresser une typologie des données collectées et produites par les chercheurs et les structures de recherche
- Déterminer les pratiques de stockage, d'archivage et de partage des données de recherche par les chercheurs et les structures de recherche
- Etudier la prédisposition des décideurs et des chercheurs à adopter et suivre une politique des données de recherche

L'étude quantitative nous a permis de dresser une typologie des données collectées et des données produites par les enseignants chercheurs lors de leurs travaux de recherche et les supports matériels sur lesquels sont stockées ces données. Nous avons pu aussi étudier les pratiques d'archivage des données de recherche par ces chercheurs en présence ou non d'une politique de données suivie par la structure de recherche. L'enquête nous a permis, de ce fait, d'étudier la prédisposition de ces acteurs à rendre leurs données de recherche ouvertes, réutilisables, exploitables et vérifiables par d'autres chercheurs

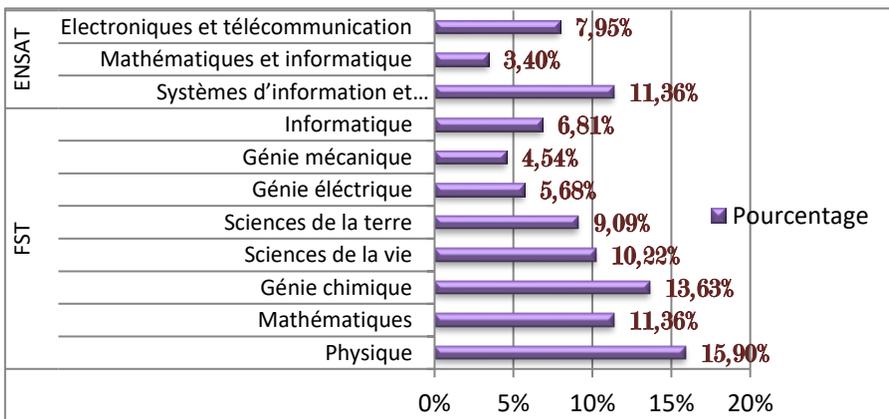
L'étude qualitative nous a éclairé les orientations des responsables universitaires à adopter un plan de gestion de données, à offrir des séances de sensibilisation et de formation sur l'aspect technique et juridique du partage de données et à mettre en place une politique des données de recherche à l'université Abdelmalek Essaâdi. Les résultats de l'étude qualitative ne seront pas présentés dans cette communication vue qu'ils sont en cours de traitement. La présente étude suit les orientations des enquêtes élaborées ou en cours d'élaboration par certaines universités européennes, notamment françaises, qui ont pour objectif de dresser un état des lieux des pratiques de gestion des données de recherche par les chercheurs universitaires.

5 Analyse des résultats

5.1 Affiliation

Les réponses sur l'affiliation des répondants sont dressées dans le graphique N°1 :

Graphique N° 1 : La répartition des répondants par établissement et par département



Les enseignants chercheurs de la FST et de l'ENSA exercent dans des disciplines scientifiques et techniques. Les enseignants de langue et

communication font partie du département des systèmes d'information et de communication.

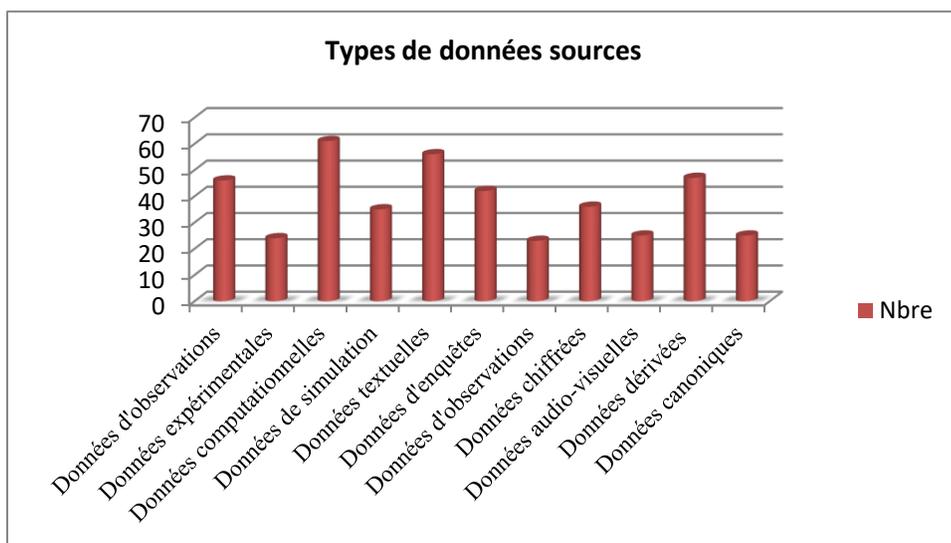
Les répondants appartiennent essentiellement au département de la physique, du génie chimique et des systèmes d'information et de communication. Néanmoins, certains départements sont peu représentés en raison du nombre faible des répondants comme c'est le cas du génie mécanique et génie électrique.

Pour déterminer la typologie des données de recherche collectées et produites par les enseignants chercheurs, plusieurs questions ont été posées sur le type de données sources des chercheurs (données collectées), leur support matériel, sont-elles numérisées, ou en cours de numérisation.

5.2 Typologie des données de recherche (Données sources et données produites)

Les données source

Les réponses sur les données sources collectées par les répondants sont comme suit :



Les réponses montrent que les enseignants chercheurs collectent une variété de données liées essentiellement à leur domaine de recherche et d'enseignement. Viennent en tête les données computationnelles qui répondent à une population exerçant à l'École Nationale des Sciences Appliquées qui forment essentiellement des Ingénieurs informaticiens, suivies par les données textuelles et d'enquêtes. Les données canoniques

telles que les banques de données de structures chimiques sont collectées surtout par les enseignants chercheurs des départements de génie chimique. Les données sources de type «textuelles» intéressent l'ensemble des répondants.

Support matériel

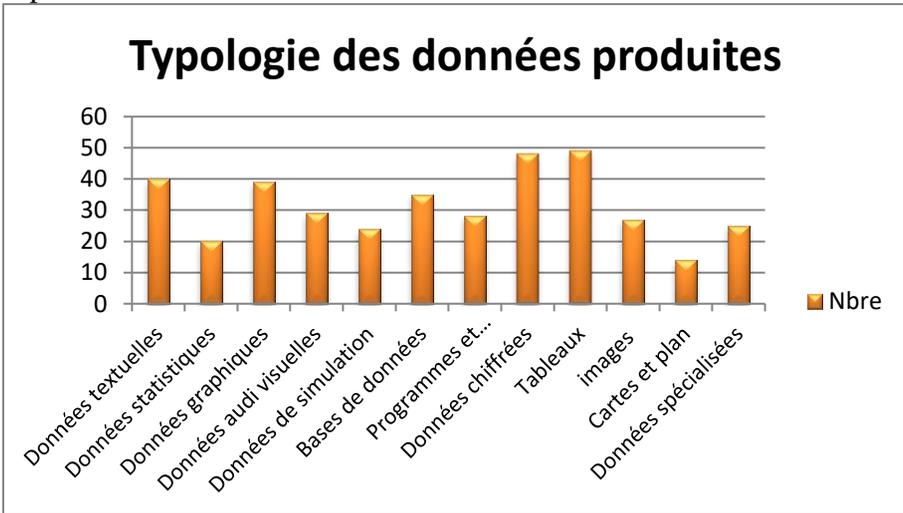
L'enquête a montré que les données sources collectées sont, pour la plupart, numériques 52%.

Pour les données sources initialement non numériques, 57% des répondants les numérisent après collecte.

L'on déduit que les domaines de recherche en sciences et techniques favoriseraient les données source initialement numériques ou numérisées a posteriori ce qui faciliterait leur stockage, leur partage et leur réutilisation.

Données produites

En ce qui concerne la typologie des données produites, les enquêtés ont répondu comme suit :



Graphique N° 3 : Typologie des données produites par les enseignants chercheurs

Comme pour les données sources, les données textuelles, chiffrées et graphiques viennent en tête puisqu'elles sont produites par tout chercheur quelle que soit sa discipline. Pour les données spécifiques, chaque spécialiste choisi celle qui lui convient (cartes et plan, données de simulation, bases de données, etc.).

Pour les données spécialisées, l'on compte par exemple les données de télédétection spatiales, les images satellites, etc.

Stockage de données

Pour notre question sur l'espace nécessaire au stockage de données, les enseignants en sciences et techniques préfèrent un espace de stockage entre 100 GO et 1 TO, soit l'équivalent d'un disque dur externe.

Concernant la fréquence de sauvegarde, la majorité des répondants font des copies de sauvegarde d'une manière irrégulière, en fonction des besoins.

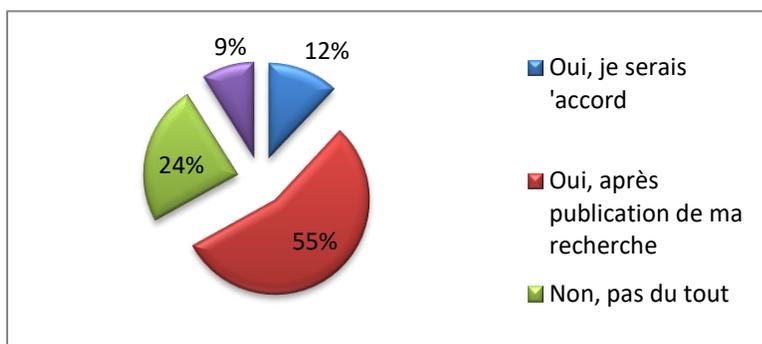
5.3 Les politiques d'archivage

En questionnant les enquêtés sur la présence ou non d'une politique d'archivage au sein de leur structure de recherche, à l'unanimité, les réponses étaient négatives. De même pour l'attribution des métadonnées aux données de recherche.

Ces résultats montrent qu'il serait très difficile d'instaurer la culture d'ouverture et de partage des données de recherche dans les universités marocaines. Chaque chercheur travaille, conserve et archive ses travaux selon ses besoins et ses priorités sans pour autant se soucier d'une politique de gestion et de sauvegarde des données au sein de sa structure de recherche ou de son département.

5.4 Pratiques de partage des données de recherche

Sur notre question accepteriez-vous de partager vos données de recherche avec vos pairs, les résultats étaient comme suit :

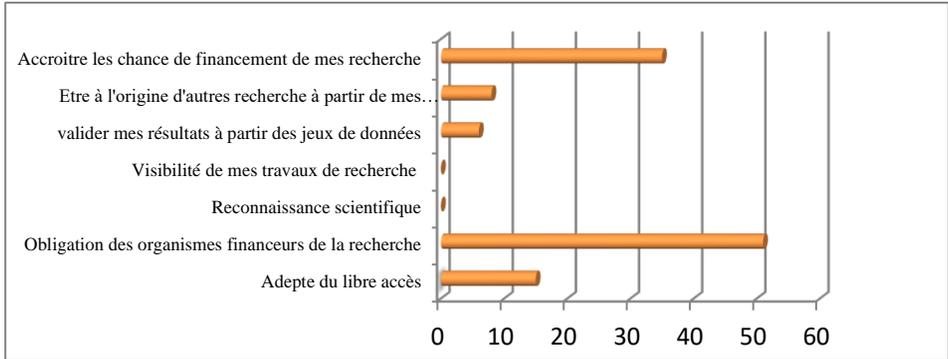


Graphique N° 4 : Opinions sur la diffusion des données de recherche

55% des répondants accepteraient de partager leurs données de recherche après publication de leurs travaux évitant ainsi toute utilisation frauduleuse de leurs données.

Pour les enquêtés ayant répondu positivement ou avec la condition que la recherche soit publiée, une autre question a porté sur les raisons qui les

inciteraient à rendre leurs données de recherche ouvertes a été posée. Les résultats étaient comme suit :



Graphique N° 5 : Raisons qui encourageraient les chercheurs à rendre leurs données de recherche ouvertes

Le financement de la recherche reste la raison principale qui inciterait les chercheurs à mettre en libre accès leurs données de recherche.

En effet, les organismes financeurs de la recherche imposent, de plus en plus, la mise en libre accès des résultats des projets qu'ils ont financés ainsi que les données sous-jacentes.

Dans le contexte marocain, le ministère de tutelle et le Centre National de la Recherche Scientifique et Technique CNRST peuvent imposer la mise en libre accès des publications et des données de recherche respectivement dans des archives institutionnelles et des entrepôts de données pour les projets de recherche qu'ils ont validé et financé.

5.5 Plan de gestion de données

Pour notre question sur le degré de connaissance d'un Plan de gestion de données, seulement 2% ont répondu positivement. Ce qui impliquerait tout un travail de sensibilisation et d'organisation de stages et d'ateliers de formation au profit des différents acteurs de la recherche scientifique, notamment les enseignants chercheurs et les décideurs.

6 L'Open Research Data dans le context juridique marocain

La loi marocaine sur le droit d'auteur et droits voisins publiée dans le Dahir N° 1-05-192 du 15 moharrem 1427 (14 février 2006) portant promulgation de la loi N° 34-05 modifiant et complétant la loi N° 2-00 relative aux droits d'auteur et droits voisins (Bulletin Officiel N° 5400 du 02 mars 2006) stipule dans l'article 3 que « *la loi s'applique aux œuvres littéraires et artistiques (ci-après dénommées «œuvres») qui sont des créations intellectuelles originales dans le domaine littéraire et artistique.* Sont concernés aussi «les programmes

d'ordinateur, les conférences, allocutions, sermons et autres œuvres faites de mots ou exprimées oralement ».

Par ailleurs, l'Article 10 précise que « l'auteur d'une œuvre a le droit exclusif de faire, d'interdire ou d'autoriser [...] (à) rééditer et reproduire son œuvre de quelque manière et sous quelque forme que ce soit, permanente ou temporaire, y compris l'archivage temporaire sous forme électronique ».

Le premier alinéa de l'article 25 précise que l'œuvre est protégée pendant « la vie de l'auteur et 70 ans après sa mort ». Pour ce qui est des limitations, l'article 12 stipule qu'« il est permis, sans l'autorisation de l'auteur et sans le paiement d'une rémunération, de reproduire une œuvre licitement publiée exclusivement pour l'usage privé de l'utilisateur » mais ne s'applique pas à « à la reproduction de la totalité ou de parties de bases de données sous forme numérique ».

L'article 65 précise que « les entités sans but lucratif suivantes : bibliothèques, services d'archives, institutions d'éducation, ...ne sont pas soumises aux dispositions de l'article 64 pour des violations mentionnées à l'article 65 alinéa a, d, e, f, g, ou h. »

Aussi, cette loi est venue principalement protéger les œuvres contre le piratage, comme c'est le cas des logiciels qui représente une création intellectuelle.

Enfin, le contrat d'éditeur dans la loi marocaine est mentionné dans le chapitre VIII. Selon la loi : « le contrat d'édition est le contrat par lequel l'auteur d'une œuvre ou ses ayants droit cèdent à des conditions déterminées à une personne appelée « éditeur », le droit de fabriquer ou de faire fabriquer en nombre des exemplaires de l'œuvre, à charge pour elle d'en assurer la publication et la diffusion». Ce contrat doit être rédigé par écrit et prévoit pour l'auteur ou ses ayants droit une rémunération proportionnelle ou forfaitaire.

Des obligations de l'auteur (Article 46), on pourrait retenir : « faire respecter ce droit et de le défendre contre toute atteinte qui lui serait portée » et que « l'objet de l'édition fournie par l'auteur reste la propriété de celui-ci. L'éditeur en sera responsable pendant le délai d'un an après l'achèvement de la fabrication »

La loi prévoit dans l'article 54 chapitre III la libre utilisation des ressources en limitant et précisant les cas d'utilisation. En ce qui nous concerne on pourrait retenir surtout la libre utilisation à des fins de recherche scientifique ou d'activités d'enseignement.

Cependant, et suite à la lecture de ces closes, peut-on intégrer les données de recherche parmi les œuvres traitées par l'article 3 de ladite loi ? La

réponse reste toutefois indécise vu que l'expression « Données de recherche » n'a pas été clairement mentionnée.

En mars 2018, une autre loi a été publiée en lien avec le droit d'accès à l'information. Il s'agit de la loi marocaine N° 31-13 du 12 mars 2018 (version en langue arabe) et du 03 mai 2018 (version en langue française) sur le droit d'accès à l'information.

Cette loi concerne essentiellement les informations que possèdent les administrations publiques marocaines et les organismes investis de mission de service public. Les universités marocaines et les centres de recherche en sont étroitement concernés.

Selon l'article 2 de ladite loi, on entend par « Information » : « Les données et statistiques exprimées sous forme de chiffres, de lettres, de dessins, d'images d'enregistrement audiovisuel, ou toute autre forme contenues dans des documents, pièces, rapports, études, décisions, périodiques, circulaires, notes, bases de données et autres documents à caractère général, produits ou reçus par les institutions ou les organismes concernés dans le cadre des missions de service public, quel que soit le support, papier, électronique ou autre. »

Plusieurs questions peuvent se poser si l'on souhaiterait impliquer les différents acteurs marocains de la recherche scientifique à s'engager dans ce mouvement d'ouverture des données de recherche :

- Peut-on qualifier de « Bien public » les données produites par les enseignants chercheurs et par les structures de recherche lors de la préparation des travaux de recherche, des rapports d'activités, des expériences scientifiques, etc. ?
- Est-ce que les données de recherche, tout comme les publications financées sur fonds public, sont la propriété du chercheur seul, de son institution d'appartenance (université, centre de recherche) ou des deux ?
- Les institutions d'enseignement supérieur et de recherche marocaines peuvent-elles imposer une politique de données à leurs chercheurs ? Cette obligation de dépôt est-elle protégée par la loi marocaine du droit d'auteur N° 34-05 et par la loi marocaine N° 31-13 ?

En l'absence d'un code de la recherche et d'une désignation précise de ce que c'est qu'une donnée de recherche et où peut-on l'intégrer dans les deux lois étudiées, ces quelques questions resteront sans réponse sauf si l'on considère que les données produites par les établissements d'enseignement supérieur et de recherche dans le cadre de leur mission de service public sont considérées comme un bien public et doivent être communiquées à toute

personne qui en fait la demande sauf dans les cas exceptionnels mentionnés dans l'article 7 relatifs, entre autres, à la défense nationale, à la sécurité de l'Etat, à la vie privée des personnes, les données personnelles, etc.

Enfin, Nous pouvons constater que malgré les modifications apportées à la loi marocaine de droits d'auteur, le vide juridique persiste encore, notamment à l'ère du libre accès à l'information scientifique et technique et aux données de recherche. Des clauses et des règles peuvent être ajoutées pour donner à l'auteur plus de liberté de diffuser ses recherches dans des archives ouvertes ou des entrepôts de données.

Conclusion

En guise de conclusion, cette étude pilote se veut être le premier jet pour une expérience marocaine en matière de l'Open Research Data à travers laquelle nous inciterons les différents acteurs de la recherche scientifique et les tutelles à s'inscrire dans cette nouvelle dynamique prometteuse.

Cependant, malgré que le terrain semble propice en étudiant la prédisposition des acteurs principaux à s'intégrer dans cette tendance, l'aspect juridique reste la principale entrave. La loi N° 31-13 du 12 mars 2018 sur le droit d'accès à l'information reste ambiguë en ce qui concerne l'ouverture des données de recherche.

Il est primordial d'adapter l'arsenal juridique marocain, axé sur la protection des œuvres littéraires et artistiques, aux exigences des nouveaux modes de publication scientifique basés sur les principes de partage et de réutilisation des œuvres et des données sous-jacentes.

Bibliographie

- BARTLING, S., FRIESIKE, S. (2014). Towards Another Scientific Revolution. In. « *Opening Science* ». London : Springer International Publishing, 325 p., Pp. 3-16.
- BLANCHARD, A. (2016). Les humanités numériques, une science « plug and play » ?.Acquitaine : Maison des sciences de l'homme, 130 p.
- COMMISSION EUROPEENNE (2012). Communication de la commission au parlement européen, au conseil, au comité économique et social européen et au comité des régions : pour un meilleur accès aux informations scientifiques
- EUROPEAN COMMISSION (2016). H2020 Programme Guidelines on FAIR Data Management in Horizon 2020, Directorate-General for Research& Innovation [en ligne] disponible sur :

http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

- FECHER, B., FRIESIKE, S. (2014). Open science : one terme, five Schools of Thought. In. « *Opening Science* ». London : Springer International Publishing, 325 p., Pp. 17- 47
- GAILLARD, R. (2014). De l'Open data à l'Open Research Data : quelle (s) politique (s) pour les données de recherche ? [en ligne] disponible sur : <https://www.enssib.fr/bibliotheque-numerique/documents/64131-de-l-open-data-a-l-open-research-data-quelles-politiques-pour-les-donnees-de-recherche.pdf>
- MESRI (2017). Ouverture des données de recherche : guide d'analyse du cadre juridique en France, 45 p.
- OCDE (2007). Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds public [en ligne] disponible sur : <http://www.oecd.org>
- PUREN M. (2017). Créer un plan de gestion des données (DMP). Vandoeuvre-Les-Nancy, INIST-INRIA
- Royaume du Maroc (2018). Loi N° 31-13 sur le droit d'accès à l'information. In. Bulletin officiel, N°6655, 12 mars 2018, Pp. 1438-1443
- Royaume du Maroc (2006). Dahir N° 1-05-192 du 15 moharrem 1427 (14 février 2006) portant promulgation de la loi N° 34-05 modifiant et complétant la loi N° 2-00 relative aux droits d'auteur et droits voisins. In. Bulletin officiel, N° 5400, 2 mars 2006, p. 325
- UK Data Archive. Research Data Lifecycle. [en ligne] disponible sur : <http://data-archive.ac.uk/create-manage/life-cycle>
- Université Abdelmalek Essaâdi. Système d'Information Marocain de la Recherche Scientifique Simarech. Disponible sur : <http://simarech.uae.ac.ma/Simarech/>

Le libre accès en Afrique de l'ouest francophone : état des lieux

Open Access in Francophone West Africa: State of Play

Diéyi Diouf U. Cheikh Anta Diop Sénégal

Résumé

Les TIC offrent d'importantes opportunités d'interactions entre les communautés des chercheurs. Les nombreux obstacles liés à l'instauration du Libre Accès paralysent la circulation de l'information. Si ce mouvement a vu le jour dans les pays développés, il semble important d'analyser son évolution en Afrique de l'ouest francophone. Cette étude abordera ces points avant de proposer quelques pistes visant l'adoption du Libre Accès en Afrique de l'ouest francophone.

Mots-clés.

Libre accès, TIC, Information, Afrique francophone.

Abstract

ICT offer important opportunities of interaction between researchers communities. The numerous obstacles related to the establishment of Open Access paralyze the circulation of information. If this movement has emerged in developed countries, it seems important to analyze its evolution in Francophone West Africa. This study will address these points before proposing some avenues for the adoption of Open Access in Francophone West Africa.

Keywords.

Open Access, Information, Francophone Africa.

1 Introduction

Les bibliothèques ont toujours favorisé le développement et l'innovation dans la mise à disposition de l'information grâce à l'efficacité, la créativité et la proactivité des professionnels qui les gèrent. Ainsi, nous affirmons avec N. Bakhoum¹ que « l'information et le savoir, en tant que vecteurs de pouvoir et de richesse, éliminent la marginalisation économique, les inégalités, le chômage et d'autres pénuries dont souffre le continent africain ».

Les professionnels africains de l'information documentaire ont très tôt pris conscience que de telles insuffisances, au lieu d'être perçues comme une frustration, devraient plutôt être transformées en source supplémentaire de motivation et d'impulsion pour des changements innovants. C'est dans ce sillage que nous avons prédit la mise en place d'un dépôt institutionnel unique sous régional, gratuitement accessible dans le cadre de notre thèse de doctorat (Diouf, 2009)². Ainsi, la promotion du Libre Accès dans le contexte ouest africain francophone s'est avérée efficace pour fournir une infrastructure technique, des solutions de qualité et à valeur ajoutée à la gestion des collections, particulièrement au sein de bibliothèques académiques.

Dans un chapitre démonstratif, nous analyserons l'appropriation du Libre Accès par certaines institutions et/ou bibliothécaires académiques africaines qui s'adaptent aux changements innovants et développent des capacités pour construire des dépôts institutionnels et des bibliothèques numériques. L'objectif de cet article est de démontrer le processus d'installation du Libre Accès en Afrique de l'ouest francophone, les actions et stratégies adoptées par les acteurs, les chercheurs et les institutions elles-mêmes. Enfin, nous décrypterons quelques pistes visant l'adoption générale du Libre Accès dans le processus de diffusion, de partage des savoirs et des enseignements-apprentissages en Afrique de l'ouest francophone.

2 Quelques définitions du concept « Libre Accès »

1) L'initiative de Budapest pour l'accès ouvert soutient : « le libre accès à la littérature scientifique est sa mise à disposition gratuite sur l'Internet public, permettant à tout un chacun de lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces

¹Bakhoum, Nafissatou(2016). Challenges of building digital repositories in Africa. Semaine du Libre accès à l'UCAD. En ligne : [http://bca.ucad.sn/jspui/\(Page consultée le 10 août 2018\)](http://bca.ucad.sn/jspui/(Page consultée le 10 août 2018))

² Diouf, Diéyi (2009). Quelle méthodologie pour l'archivage et la diffusion de la documentation scientifique et technique dans le cadre d'une université africaine : le cas de l'UCAD. Th. Doctorat. Troyes : UTT, 2009, 250 p. En ligne : <http://www.diffusiontheses.fr/> (Page consultée le 02 juillet 2018)

articles, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale, sans barrières financières, légales ou techniques autres que celles indissociables de l'accès et l'utilisation d'Internet. La seule contrainte sur la reproduction et la distribution et le seul rôle du copyright dans ce domaine, devrait être de garantir aux auteurs, un contrôle sur l'intégrité de leurs travaux et le droit à être correctement reconnus et cités (BOAI, 2002).

2) Selon l'INIST, « le libre accès (Open Access en anglais), désigne l'accès permanent et gratuit sur Internet aux données issues de la recherche scientifique et de l'enseignement ».

3) Le libre accès est la mise à disposition en ligne de contenus numériques qui peuvent eux-mêmes être soit libres, soit sous un des régimes de propriété intellectuelle. Le libre accès est principalement utilisé pour les articles de revues de recherche universitaires, sélectionnés par des pairs. (wikipedia.org). Nous tenons à signaler d'entrée de jeu, que le Libre Accès à l'information scientifique recouvre deux réalités distinctes : les revues en libre accès et les archives ouvertes. Il repose sur des standards, des outils et plateformes fonctionnant grâce à des logiciels libres.³

3 Historique

Le serveur de prépublications initié en 1991 par Paul Ginsparg dans le domaine de la haute Physique demeure la pierre angulaire du vaste mouvement du Libre Accès. Puis, ce mouvement a été jalonné de déclarations de principes et de prises de conscience aux difficultés d'accès aux résultats de la recherche et de propositions de solutions pour y remédier, à savoir :

- l'appropriation par la société d'un processus initié par des individus ;
- l'élargissement du mouvement à tous les domaines du savoir.

3.1 Contributions en faveur du Libre accès

Le mouvement du Libre Accès vise entre autres, à permettre une diffusion plus large de l'information scientifique sans aucune barrière économique, de manière à faciliter l'échange, le développement et le partage des savoirs. Ce mouvement s'est considérablement amplifié grâce à de nombreux

³Initiative de Budapest pour l'accès ouvert ((2002). En ligne : (http://guides-formadoct.ueb.eu/libre_acces/) (Page consultée le 09 juin 2018)

soutiens et prises de positions officielles émanant d'un grand nombre d'institutions scientifiques, de chercheurs, de responsables d'universités et de gouvernements, à travers des déclarations internationales (Budapest Open Access Initiative en 2002, Déclaration de Berlin en 2003...). Plusieurs organismes de financement de la recherche ont également joué un rôle moteur en définissant des politiques contraignant les chercheurs qu'ils financent, à diffuser librement et gratuitement les articles qu'ils publient (Wellcome Trust en 2003, le National Institute of Health - NIH en 2005⁴). Nous résumons ci-dessous, certaines étapes qui ont jalonné le processus d'instauration du libre accès :

3.1.1 Entre 1991 et 1997

Des chercheurs ont conceptualisé et mis en place des serveurs pour stocker leurs publications avant parution (prépublications), puis après publication (postpublications) dans une revue scientifique.

3.1.2 En 1999

L'Open Archive Initiative (OAI) a été créée pour promouvoir et développer un ensemble de protocoles communs et des standards d'interopérabilité facilitant la diffusion du contenu des documents.

3.1.3 Entre 2001 et 2003

L'Initiative de Budapest pour l'Accès Ouvert « Budapest Open Access Initiative » (BOAI), entreprise en décembre 2001 a été un grand pas vers l'instauration de ce mouvement. Puis, la Déclaration de Berlin d'octobre 2003 sur le libre accès à la connaissance, sont toutes des initiatives internationales fondatrices du Libre Accès entre autres.

4 Modèles du Libre accès

La philosophie du " Libre Accès" se manifeste partout dans le monde à travers des eBooks, des blogs, des wikis, des forums de discussion et bien sûr, des réseaux sociaux de scientifiques. Nous pouvons noter deux principaux modèles du Libre Accès à la science : la voie dorée (Gold Open Access) et la voie verte (Green Open Access).

4.1 La voie dorée (Gold OA)

La première préoccupation de l'accès ouvert était la bataille avec les éditeurs classiques qui a abouti à la mise en place de plateformes de dépôts d'articles de revues accessibles gratuitement. A ce titre, nous pouvons citer l'Open Access Journals (Gold OA) à l'image d'ArXiv.org, précurseur dans l'accès gratuit à des publications scientifiques en physique, mathématiques,

⁴ Venet, Jean (2014) Libre accès : dossier. 12 p. En ligne : <http://www.inshea.fr/sites/default/files/Open%20access.pdf> (Page consultée le 20 octobre 2018)

informatique et sciences quantitatives. Dans ce modèle, c'est l'auteur qui paye l'accès. Le principe de libérer les connaissances de barrières économiques, juridiques, techniques et autres n'a pas négligé la qualité des contenus en libre accès. L'une des raisons est de lever l'équivoque entre les termes anglais "free" et "poor quality", voulant sous-entendre que si l'article est accessible gratuitement, il est de facto médiocre. Au contraire, des réalisations ont été faites dans ce sens. La Public Library of Science (PloS) et Biomed Central (BMC) sont deux grands serveurs qui fournissent des articles validés (Peer Reviewed) en Physique, Chimie, Médecine, Biologie, etc. Les coûts induits de ces types de publication dont la chaîne est essentiellement numérique, sont supportés par l'auteur à travers des fonds alloués par son université, centre de recherche ou une fondation.

4.2 La voie verte (Green OA)

Suite à la réaction irréversible de grandes institutions scientifiques, des professionnels de l'information documentaire, des techniciens et chercheurs ont également senti l'urgente nécessité de sauvegarder et de valoriser par l'auto-archivage, la production scientifique des universités et centres de recherche pour un accès plus ouvert. C'est la naissance des dépôts institutionnels, faisant référence à la voie dite verte.

5 Le contexte en Afrique de l'ouest francophone

Les chercheurs africains ont souvent été confrontés à des problèmes d'accès à l'information scientifique et technique (IST) dans leurs recherches. Les raisons sont multiples. Il s'agit notamment, des difficultés d'accès permanent à l'Internet selon les localités, des prix exorbitants de la documentation, des délestages intempestifs, la liste est loin d'être exhaustive. A cela, s'ajoute le déficit de contenus africains sur la toile.

À l'instar de la situation qui a conduit à la naissance du Libre Accès dans les pays développés, nous pouvons noter dans la plupart des universités d'Afrique de l'ouest francophone, des phénomènes communs. Pour preuve, le coût d'abonnement aux bases de données demeure hors de la portée des enseignants-chercheurs et étudiants dans ces universités.

Paradoxalement, malgré le gap informationnel signalé dans les suds et le déficit de contenus africains sur la toile, la production scientifique a grandement augmenté. Cependant, le manque de cadre de publication et les revues à haut facteur d'impact demeurent des obstacles non négligeables. A cela, s'ajoutent plusieurs facteurs parmi les plus récurrents, le monopole des grandes maisons d'édition classique qui détiennent toujours avec subtilité, la clé éditoriale en Afrique francophone.

Quant aux chercheurs, ils se trouvent encore dans « l'obligation » de subir ce phénomène pour un souci d'avancement dans leur carrière qui reste sujet à plusieurs contraintes. De plus, certains auteurs développent encore une

forte crainte vis-à-vis du plagiat devenu un phénomène cyber planétaire, ce qui les empêche de s'engager dans la publication en libre accès. Ces derniers ignorent cependant, que la publication en libre accès est un moyen privilégié de protection des résultats de la recherche. Pour les mêmes causes donc, les chercheurs africains ont éprouvé le besoin de trouver des mécanismes de production, de publication, de stockage, de circulation, d'acquisition et de partage de leurs résultats de recherche. C'est dans ce contexte que le Libre Accès est né en Afrique de l'ouest francophone.

Par ailleurs, même si la production scientifique est foisonnante, en Afrique de l'ouest francophone, il est facile de constater qu'information ne circule pas comme le souhaitent les chercheurs. Pour ces raisons et tant d'autres, plusieurs initiatives ont été entreprises dans le but de mieux partager les résultats de recherche à travers un canevas jusque-là considéré comme un phénomène utopique : le Libre Accès.

5.1 Objectifs du libre accès pour l'Afrique

Nous pouvons citer plusieurs objectifs dans l'instauration du libre accès parmi lesquels :

- la diffusion maximale de la science pour les pairs et pour le grand public;
- la réduction du coût de la documentation et de l'abonnement aux périodiques, ce qui laisserait plus de marges de manœuvre aux structures documentaires dans leurs acquisitions ;
- la redistribution des fonds publics dans des activités prioritaires de recherche et d'enseignement ;
- l'appropriation de la communication scientifique par la communauté des chercheurs elle-même.

5.2 Avantages

En Afrique de l'ouest francophone, il est indéniable que le Libre Accès à l'information recouvre plusieurs avantages pour les différents acteurs.

• Les auteurs

Ils bénéficient d'une meilleure visibilité, celle de leurs productions scientifiques et de leurs institutions d'origine ;

- de possibilités de collaboration et de partenariat avec d'autres chercheurs ;

- une mise en ligne plus importante de contenus facilement référencables et citables (meilleur facteur d'impact), etc.

• Les utilisateurs et les bibliothèques

Quant aux utilisateurs, ils pourront désormais accéder gratuitement sans barrières quelconques, à des résultats scientifiques pluriels (bases de données gratuites, sites web thématiques, Webrings, ...). Les budgets consacrés aux ressources payantes pourront être reconvertis et investis dans d'autres priorités des structures documentaires.

6 Etat des lieux en Afrique de l'ouest francophone

6.1 Actions, activités et stratégies

Malgré un constat similaire à celui d'autres contrées à une certaine époque, la réalité est tout autre sur le plan ouest africain francophone. En effet, bien que des avancées aient été notées, le Libre Accès bute encore sur quelques freins. L'état des lieux en Afrique de l'ouest francophone, jusqu'à une date récente (août 2018), fait état de timides expériences qui peinent à gagner l'opinion quelque peu frileuse de nombre d'universitaires, tiraillés entre le désir d'avancer dans leur carrière en publiant dans des revues à haut facteur d'impact et le partage gratuit des résultats de recherche dont la publication en libre accès ne semble pas garantir contre le plagiat.

C'est ainsi que dans des pays comme le Sénégal, le Burkina Faso ou la Côte d'Ivoire, des enseignants, chercheurs, associations de professionnels de l'information et autorités académiques engagés, mènent quotidiennement d'importantes actions/activités visant l'instauration du Libre Accès. Dans ces pays, de fortes actions de communication (journées de sensibilisation ou de lancement du Libre Accès, des colloques, des conférences, la semaine du Libre Accès (Open Access Week), sont définitivement ancrées dans l'agenda des Recteurs, des enseignants-chercheurs, des doctorants et des bibliothécaires. En témoignent les importantes mesures juridiques prises par les autorités de certaines universités ou les décisions d'organismes d'accréditation comme le Conseil africain et malgache pour l'enseignement supérieur (CAMES).

C'est dans cette lancée que le Burkina Faso avait mis en place en 2014, un important dispositif dénommé « Burkina Open Data Initiative (BODI), qui est un Cloud gouvernemental pour la collecte des données administratives des institutions gouvernementales.

Au Sénégal, le Consortium des bibliothèques des établissements d'enseignement supérieur du Sénégal (COBESS) et l'Association sénégalaise des Bibliothécaires, Archivistes et Documentalistes (ASBAD), ont entrepris depuis plusieurs années déjà, d'importantes initiatives visant l'adoption du Libre Accès dans le quotidien des enseignants-chercheurs et étudiants du pays. En 2010, un atelier national de lancement du Libre Accès a été organisé en vue d'accroître la visibilité et l'accessibilité des résultats de la recherche scientifique dans les universités du Sénégal. Cet atelier visait également, le renforcement des capacités des institutions de recherche dans le partage des connaissances à travers la création et la mise en place de dépôts institutionnels au niveau du pays.

L'usage des TIC dans les bibliothèques a également apporté beaucoup de mutations dans le domaine de la recherche scientifique, de la communication du savoir et de l'usage des ressources numériques. Dans l'appropriation et la généralisation de cette nouvelle donne, l'Université

Cheikh Anta Diop de Dakar a initié depuis 2002, un projet d'archivage et de diffusion des travaux académiques qu'elle génère par le canal de ses chercheurs, enseignants et étudiants. Ce projet très ambitieux est confié à la bibliothèque centrale qui en assure la phase d'exécution⁵.

Ce projet a donné naissance à une grande bibliothèque numérique avec l'acquisition de la plateforme libre et gratuite appelée Cyberthèses, devenue Cyberdocs-UCAD. Il comportait déjà en 2008, plus de six mille sept cent quarante-six (6 746) documents scientifiques numérisés, accessibles gratuitement en texte intégral sur le site de la bibliothèque centrale de l'UCAD en intranet d'abord, puis sur l'Internet public. De plus, le Service de numérisation et de valorisation des publications académiques fraîchement créé, avait bénéficié de l'adoption et de l'application d'un arrêté rectoral concernant le dépôt systématique d'une version électronique de tout travail académique avant sa soutenance à l'UCAD. Plus tard, le 04 février 2013, le Recteur de l'UCAD, sortait la circulaire n° 721 adressée aux doyens, directeurs d'écoles nationales supérieures et d'instituts d'université, invitant les candidats sur les listes d'aptitude du Conseil africain et malgache de l'enseignement supérieur (CAMES), à effectuer leur inscription en ligne à partir de la bibliothèque centrale. Cette décision a permis d'enregistrer le dépôt d'articles et de copies de thèses jusque là non disponibles à la bibliothèque universitaire, en sus de ceux déposés volontairement par certains chercheurs engagés.

La bibliothèque universitaire de l'UCAD, après avoir développé plusieurs expériences (DATAD, Cyberthèses-Cyberdocs), a définitivement opté pour la mise en place systématique de tous les travaux scientifiques effectués au sein de cette prestigieuse institution. Ces différentes initiatives ont donné naissance principalement à la « bibliothèque SIST Sénégal ».

Ce dépôt regroupe les travaux académiques (antérieurs à l'année 2000) des enseignants, chercheurs et étudiants des universités du Sénégal. Malgré des difficultés de toutes sortes il convient de noter qu'en plus de ces actions et projets, certaines universités africaines francophones sont parvenues à mettre en place leur propre dépôt institutionnel ou bibliothèque numérique.

Nous dressons ci-dessous, un aperçu (probablement non exhaustif) des dépôts institutionnels déjà référencés en Afrique de l'ouest francophone par le Répertoire des dépôts institutionnels en libre accès ou « Directory of Open Access Repositories (DOAR) ».

⁵ Diouf, Diéyi (2009).L'expérience d'archivage et de diffusion électronique des thèses et mémoires à l'Université Cheikh Anta Diop de Dakar : un pas vers le désenclavement scientifique de l'Afrique. In : *Bulletin des bibliothèques de France (BBF)*, T. 54, n° 2. ISSN 1292-8399 [En ligne] Disponible sur : <http://bbf.enssib.fr/consulter/bbf-2009-02-0084-002.pdf>(Page Consultée le 02 juillet 2018)

6.2 Dépôts institutionnels référencés dans le Directory of Open Access Repositories (DOAR)

Le Sénégal compte quatre dépôts institutionnels.

6.2.1 La bibliothèque numérique de l'UCAD

Ce dépôt institutionnel est le plus important du Sénégal. Il comprend plusieurs collections de documents progressivement mises en ligne parmi celles-ci, nous avons dénombré des thèses et des mémoires déposés à la bibliothèque depuis la création de l'université le 24 février 1957.

6.2.2 Biens Culturels Africains

Le dépôt institutionnel de l'Institut fondamental d'Afrique noire (IFAN) offre un accès gratuit au patrimoine documentaire audiovisuel, iconographique, sonore et textuel de l'IFAN.

6.2.3 Bibliothèque SIST-EISMV

La bibliothèque numérique de l'Ecole Inter-états des Sciences et Médecine Vétérinaires de Dakar (EISMV) donne accès aux thèses et mémoires soutenus dans cet établissement.

6.2.4 IDEP Document Server

Le dépôt de l'Institut pour le développement économique et de la planification (IDEP) recense l'ensemble de la production scientifique des chercheurs de l'IDEP.

En plus des quatre (04) dépôts précités pour le Sénégal, il existe un grand nombre de bibliothèques numériques, fruit de l'appropriation par les professionnels de l'information, des chercheurs engagés et défenseurs du Libre Accès.

Nous dressons ci-dessous, un tableau récapitulatif des dépôts institutionnels existants en Afrique de l'ouest francophone, du moins, ceux référencés dans le DOAR.

Pays	Dépôts institutionnels référencés	Commentaires
Bénin	00	
Burkina Faso	00	
Côte d'Ivoire	01 Projet signalé dans MemOnline	ATTA, Jean Constant Adou. Mém. Master. Dakar : EBAD, 2012
Guinée	00	
Mali	00	
Mauritanie	00	
Niger	00	

Sénégal	04	En plus de ces 4 dépôts référencés, on note une floraison de projets et de bibliothèques virtuelles institutionnelles
Togo	00	

Tableau des dépôts institutionnels d’Afrique de l’ouest francophone référencés

Source :Directory of Open Access Repositories

Il convient de signaler qu’en plus de ces rares dépôts institutionnels, plusieurs bibliothèques numériques ont vu le jour dans la quasi-totalité des universités et centres de recherche d’Afrique de l’ouest francophone. Cependant, ces dépôts ne sont pas signalés dans le DOAR. Cet engouement résulte du désir de partager les résultats de recherche avec les pairs, mais également d’être présent et visible sur la toile. Afin d’atteindre cet objectif, les universités, les bibliothèques et centres de recherches s’organisent et conduisent des projets de numérisation et de gestion de documents numériques. Il s’agit entre autres, de la mise en place de projets et de programmes qui permettent de faciliter l’accès à l’information scientifique, surtout celle produite par les chercheurs africains. Ces réalisations essayent de résorber le fossé informationnel en se tournant définitivement vers le numérique.

6.3 Lesacquis et stratégies

Les différentes décisions prises par les Ministres de l’enseignementsupérieur des pays membres du CAMES visant la dématérialisation des candidatures ontpermis de constituer une importante collection numérique, regroupant la production scientifique des enseignants et chercheurs de l’espace CEDEAO.

Certains présidents d’universités ont même sorti des arrêtés demandant aux enseignants et chercheurs de déposer obligatoirement leurs résultats de recherche dans les bibliothèques universitaires (BU) pour obtenir un quitus leur permettant de postuler au CAMES. C’est ainsi qu’en 2015, la BU de l’UCAD a pu collecter1 057 articles en moins de trois (3) mois.

Quant aux étudiants, l’obligation de dépôt d’une version électronique avant toute soutenance d’un travail scientifique à l’UCAD a été la pièce maîtresse de l’actuel dépôt institutionnel dénommé “Bibliothèque numérique”. Aussi, un important conseil tenu en août 2013 a pris dix (10) décisions présidentielles relatives à l’enseignement supérieur et à la recherche.

Du côté des politiques et de l'Etat, la concertation nationale sur l'avenir de l'Enseignement supérieur au Sénégal (CNAES) en 2014 a significativement contribué à l'innovation et au développement des contenus numériques dans les universités publiques du pays.

D'autres actions comme :

- le Plan de développement de l'Enseignement supérieur et de la recherche (PDESR) 2013-2017/MESR ;
- le Contrat de performance (CDP) 2011-2016 ;
- l'instauration définitive du système LMD avec des formations « Transfer » ;
- la création de campus partenaires au sein des BU ;
- la déclaration de Dakar sur le Libre Accès à l'issue du colloque du WACREN en mars 2016) ;
- la déclaration de Dakar sur le Libre Accès lors du colloque du CODESRIA en avril 2016 ;
- la mise en ligne de revues scientifiques en libre accès grâce au soutien de l'AUF ;
- la création de bibliothèques électroniques issues de la numérisation de collections grâce à la coopération avec des partenaires étrangers ;
- l'implémentation de plateformes collaboratives appelées « Boutiques des sciences » qui sont des interfaces science/société, sont autant d'initiatives visant l'adoption et l'instauration du Libre Accès en Afrique de l'ouest francophone.

7 Résultats et synthèse de l'étude

Il ressort de notre étude :

- une lente instauration du Libre Accès en Afrique de l'ouest francophone pour plusieurs raisons (réticence, ignorance, illettrisme numérique, manque d'infrastructures et/ou de cadres de publication) ;
 - certains pays en sont encore au stade de projets de lois ou de programmes ;
 - d'autres tardent et peinent à appliquer les textes déjà élaborés ;
- Pourtant, au regard des budgets dérisoires des institutions documentaires ;
- de la cherté de la documentation scientifique et technique encore pesante ;
 - du prix d'abonnement aux bases de données toujours élevé ;
 - du gap informationnel signalé dans les Suds ;
 - du problème de littératie numérique de certains étudiants, enseignants et chercheurs ;
 - de la politique déguisée des grandes maisons d'édition, l'adoption du Libre Accès demeure une grande opportunité à ne pas rater par les chercheurs.

Par ailleurs, les dépôts institutionnels que nous avons signalés dans cette étude regroupent une multitude de collections pluridisciplinaires (publications scientifiques, mémoires, thèses, ouvrages scientifiques et

articles déjà publiés dans des revues éditées soit par les établissements ou à l'étranger dans les différentes langues d'enseignement (anglais, français, arabe, italien...). Les paramètres de recherche dans les dépôts référencés sont les suivants :

- Collection (il s'agit du type de document : article, thèse, ...)
- Date de parution
- Auteur
- Sujet.

Aussi, les documents sont en format PDF et l'accès se fait soit à travers la page web de l'établissement, soit par la bibliothèque ou directement sur la page d'accueil du site du dépôt.

Nous tenons à signaler que le site du dépôt institutionnel « Biens culturels africains » rencontre quelques difficultés depuis un bout de temps.

7.1 Difficultés liées à la mise en place de dépôts institutionnels en Afrique francophone

A l'image du reste du continent et à quelques exceptions près, nous avons remarqué que les dépôts institutionnels sont encore rares dans les universités et établissements scientifiques d'Afrique de l'ouest francophone, car sauf référencements récents, nous en avons dénombré seulement 04 dans DOAR au sein de l'Université Cheikh Anta Diop de Dakar au Sénégal. Dans les autres pays, il est certain que les universités et centres de recherche ont développé des initiatives ou des projets visant la mise en place de bibliothèques numériques ou de dépôts institutionnels, mais ils ne sont pas encore référencés.

Ces dernières années, les associations nationales des professionnels de l'information telles que l'ASBAD, l'ASPID, etc.), sont très engagées aux côtés de l'IFLA et d'EiFl, pour faire instaurer définitivement le Libre Accès dans les universités africaines.

C'est ainsi que plusieurs ateliers de formation, des colloques, des conférences et journées scientifiques sont souvent organisés pour une meilleure sensibilisation, du lobbying et des plaidoyers envers les académiques, mais également les décideurs et autorités politiques.

Malgré ces multiples acquis, nous avons noté la persistance de quelques réticences de la part de nombreux scientifiques africains à partager leurs travaux par crainte du plagiat.

A ces freins, s'ajoutent :

- un accès parfois conditionné par une disponibilité d'énergie et d'internet non permanente ;
- un déficit de contenus africains gratuitement accessibles sur certaines plateformes ;

- l'imperfection de certains logiciels (fonctionnalités nécessaires manquantes ou problème de la langue) ;
- un déficit de revues à comité de lecture dans certaines disciplines pour permettre aux chercheurs du domaine, de publier en libre accès leurs résultats de recherche ;
- un déficit de formations continues dédiées aux professionnels de l'information pour la gestion de certaines plateformes et infrastructures de publication ;
- un manque de collaboration des tutelles/autorités ou organes de décision locaux, etc.

Bref, si les chercheurs, autorités académiques, politiques et bibliothécaires africains sont convaincus que les ressources numériques constituent une nouvelle force de publication, d'accès et de partage du savoir pour les Suds où la fracture numérique n'est pas encore totalement vaincue, les universités d'Afrique francophone ne suivent pas le même rythme de création de dépôts institutionnels que leurs pairs anglophones.

Pourtant, les instances de validation et d'accréditation comme :

- le Conseil africain et malgache pour l'enseignement supérieur (CAMES) ;
- la Conférence des Recteurs et Présidents d'établissements membres de l'Agence universitaire de la francophonie (AUF) ;
- l'Agence universitaire de la Francophonie (Afrique de l'ouest) ;
- Le Consortium des bibliothèques de l'Enseignement supérieur du Sénégal (COBESS) ;
- La Fédération internationale des associations de bibliothèques et de bibliothécaires (IFLA) ;
- L'UNESCO, ont posé plusieurs actes et luttent en permanence pour que les universités et autres structures de recherche, puissent initier des dépôts institutionnels, des bibliothèques numériques, des revues en libre accès ou des archives ouvertes (projets DATAD, SIST, Cyberdocs, etc.) ;

Malgré les difficultés et obstacles précités, le libre accès est lentement mais sûrement installé en Afrique de l'ouest francophone et de façon irréversible à l'instar du reste du monde.

Cet élan est fortement soutenu dans les universités, instituts et centres de recherche par des académiques avertis et très engagés, mais également par les tutelles et des partenaires internationaux. Il faut justement noter le cadre législatif étréglémentaire favorable (textes, arrêtes, décisions, déclarations, etc.) à l'instauration du Libre Accès dans les universités francophones africaines de l'ouest.

7.2 Recommandations

7.2.1 Rôle des professionnels de l'information documentaire

Convaincue que l'instauration du libre accès est irréversible malgré les lenteurs notées dans son appropriation, nous estimons qu'il faudra utiliser tous les moyens légaux pour impliquer de façon responsable, les producteurs de contenus scientifiques et les communautés des chercheurs à partager leurs résultats de recherche en libre accès et de les considérer surtout comme un « bien commun ». Pour ce faire, les professionnels de l'information documentaire ont un rôle non négligeable à y jouer. C'est ainsi qu'ils sont appelés à favoriser et à démultiplier la mise en place de dépôts institutionnels dans chaque université. Pour y arriver, il faudra utiliser tous les événements liés à la recherche au sein des campus, pour vulgariser et promouvoir le LibreAccès. Ils profiteront également de ces manifestations, pour encourager la production, l'édition et la diffusion de corpus scientifiques sur la toile en vue d'un accès pour tous et sans barrières quelconques à l'IST. La réussite d'un tel challenge passera certainement par :

- de grandes campagnes nationales de plaidoyers (série d'activités annuelles) en faveur du Libre accès ;
- le renforcement des capacités des producteurs et gestionnaires de contenus (professionnels de l'information, personnel informatique, chercheurs, étudiants) des Services communs de la documentation (SCD) des universités africaines ;
- l'organisation de sessions spéciales de sensibilisation au profit des décideurs (recteurs, doyens, élus locaux, parlementaires et enseignants-chercheurs dits « OPEN », pour démultiplier les acteurs engagés ;
- la pérennisation de sessions de partage avec le public, autour de panels à la fin desquelles, des brochures, tee-shirts, dépliants, pins et matériaux promotionnels à l'effigie « *J'aime le Libre accès* » seront distribués aux différents participants.
- la désignation des recteurs des universités comme présidents des comités de direction des campagnes et l'inscription de ces activités en tête de leur agenda annuel ;
- l'identification et l'organisation des académiques engagés en groupes « d'amis du Libre Accès » parmi les enseignants, les chercheurs, les responsables de laboratoires de recherche et étudiants bénévoles qui acceptent le téléchargement et/ou le dépôt de leurs productions dans les bibliothèques numériques initiées par les BU ;
- le lancement de jeux avec des quiz sur le Libre Accès ;

- l'organisation à la fin du jeu, d'une exposition et récompenser les gagnants avec des outils promotionnels à l'effigie du Libre Accès (disque flash, brochures, casquettes, porte-clés, stylos) ;
- des formations sous-régionales sur les logiciels de gestion de contenus (Invenio, greenstone, koha, PMB, etc.). Profitant de l'avènement de la science ouverte, les bibliothécaires devraient également, démultiplier les cadres de discussion tels que les conférences, les colloques, les journées de vulgarisation et de promotion du Libre Accès. Ils doivent cultiver la proactivité et la polyvalence pour aider les auteurs dans l'édition et à la mise en ligne des résultats de recherche. Toutes ces actions requièrent une solide formation/autoformation permanente (tout au long de la vie), afin de former et surtout de rassurer le public, généralement déconcerté par l'immensité des ressources et la complexité des outils de recherche disponibles à travers l'Internet.

7.2.2 Les états et politiques

Pour faire la promotion du Libre Accès à l'information, les pays africains de l'ouest devraient tous mettre en place, une politique nationale prenant appui sur plusieurs facteurs et acteurs. La première décision importante à prendre par les états est d'amener les établissements scientifiques à suivre le mouvement général du Libre Accès en mettant en place, des dépôts institutionnels avec des règles incitatives ponctuées d'une forte sensibilisation des communautés de chercheurs à y contribuer et à en faire un usage quotidien. Quant aux tutelles et décideurs, la pérennisation de soutiens techniques pour la mise en place de dispositifs en libre accès dans les universités et centres de recherche leur est dévolue prioritairement. Ils doivent élaborer, adopter et surtout faire appliquer des textes de lois favorisant le Libre Accès. Une autre priorité demeure le soutien des acteurs scientifiques à bien mener leurs recherches et les amener à mettre leurs productions en libre accès, dès lors qu'ils utilisent les fonds publics. L'ensemble des acteurs devrait cultiver une synergie d'actions pour la mise en place d'un dépôt institutionnel unique, regroupant toute la production scientifique des chercheurs d'Afrique de l'ouest francophone⁶.

7.2.3 Les chercheurs

Les chercheurs devraient prioritairement s'appropriier des TIC et des opportunités/enjeux du Libre Accès ;

- adhérer, faire adhérer et contribuer à la généralisation du Libre Accès ;

⁶ Diouf, Diéyi (2009). Quelle méthodologie pour l'archivage et la diffusion électronique de la documentation scientifique et technique dans le contexte d'une université africaine ? : le cas de l'UCAD). Lille, ANRT Diffusion, 2013. 259 p. [En ligne] : Accessible sur : <http://www.diffusiontheses.fr/> (Page consultée le 02 juillet 2018)

- ouvrir la science aux citoyens pour mieux réussir le développement durable ;
- Publier systématiquement dans des revues en libre accès en vue d'un partage équitable des résultats de la recherche, mais surtout pour leur visibilité et celle de leurs institutions.

8 Conclusion

La publication en libre accès atteste d'une richesse et d'une vitalité incontestables pour les chercheurs. C'est ainsi que le libre accès gratuit sans barrières quelconques aux résultats scientifiques pluriels sous tous les formats et de toute nature, offre d'importantes opportunités aux communautés des chercheurs et aux populations en général (possibilités de découverte, de collaboration et de partenariat). Le Libre Accès accroît de façon significative, la valeur des ressources libres, mais aussi les travaux des chercheurs et ceux des étudiants. C'est ainsi que les chercheurs parviennent à contourner les insuffisances locales en accédant à l'information mondiale et y contribuent eux-mêmes en publiant sur le réseau des réseaux. Pour que les pays d'Afrique de l'ouest francophone réussissent l'adoption responsable du Libre Accès, il faut nécessairement un engagement de toutes les communautés, à partir de concertations/sensibilisations larges et régulières, car le Libre Accès constitue une réponse à des problèmes récurrents auxquels, les chercheurs des Suds font face au quotidien. En tout état de cause, les universités d'Afrique de l'ouest francophone doivent impérativement créer leur dépôt institutionnel ouvert et gratuit, afin de partager largement la science et de faire circuler les résultats de recherche du Sud au Nord de la Francophonie. Sans partage d'informations, pas de développement local durable.

Références

1. Bakhoun, Nafissatou (2016). Challenges of building digital repositories in Africa. Semaine du Libre accès à l'UCAD. En ligne. Disponible à : http://www.bu.ucad.sn/cyberpac/Telechargement/UCAD_OA_Presentation_N.Bakhoun.pdf (Page consultée le 10 août 2018).
2. DIENG, Moctar (2010). Pour une meilleure visibilité de la recherche universitaires : Le COBESS mise sur le libre accès à l'information scientifique. En ligne. Disponible à : <http://www.osiris.sn/Pour-une-meilleure-visibilite-de.html> (Page consultée le 26 août 2018).
3. DIOUF, Diéyi (2013). Enjeux de la visibilité des résultats de recherche à l'UCAD. Forum mondial des sciences sociales. Montréal : Palais des Congrès, 12-15 mai

4. DIOUF, Diéyi (2009). L'expérience d'archivage et de diffusion électronique des thèses et mémoires à l'UCAD. In : *Bulletin des bibliothèques de France (BBF)*, T. 54, n° 2. ISSN 1292-8399. En ligne. Disponible à : <http://bbf.enssib.fr/consulter/bbf-2009-02-0084-002.pdf> (Page consultée le 20 mars 2018).
5. DIOUF, Diéyi (2016). Le processus d'adoption du libre accès au Sénégal. Communication au Colloque international sur « *La Justice cognitive, science ouverte et valorisation des savoirs locaux* ». Yaoundé, ESSTIC, 25-27 mai 2016 ;
6. DIOUF, Diéyi (2014). La question du libre accès dans la politique nationale de recherche et d'innovation (PNRI) du Sénégal. Montréal, 82^e congrès de l'ACFAS
7. EGGIS, Gilles (2013). Les archives ouvertes pour valoriser la recherche africaine. In : *HAL-Francophonie*
8. Initiative de Budapest pour l'accès ouvert ((2002. En ligne. Disponible à : ([http://guides-formadoct.ueb.eu/libre acces.](http://guides-formadoct.ueb.eu/libre%20acces/)))(Page consultée le 09 juin 2018)
9. KALY, Eugène (2012). Afrique : Publication des résultats de la recherche. La part du continent reste très faible. In : *Le Soleil*, 17 janvier
10. Venet, Jean (2014) Libre accès : dossier. 12 p. En ligne. Disponible à : <http://www.inshea.fr/sites/default/files/Open%20access.pdf> (Page consultée le 20 octobre 2018)

Thème 5

**Aspects didactiques, ressources pédagogiques
ouvertes**

Corrélation entre formation sur les ressources éducatives libres et culture du libre accès : retour d'expérience avec des enseignants universitaires en Tunisie

Correlation between open education resources training and open access culture: experience feedback with university teachers in Tunisia

Sihem ZGHIDI

Institut supérieur de documentation, Université La Manouba, Tunisie.

Haifa BELHADJ ANANE

Faculté des sciences de Bizerte, Université de Carthage, Tunisie.

Résumé

Dans cet article nous analysons l'impact d'une formation à distance, sur les Ressources Éducatives Libres, sur les pratiques des enseignants universitaires tunisiens qui en avaient bénéficié.

Cette étude a montré un changement des pratiques dans l'utilisation des REL. Toutefois, une réticence a été également observée par rapport à la publication en libre accès. Nous interprétons les résultats obtenus et nous présentons des constats.

Mots-clés. Ressource éducatives libres (REL), Libre accès, Licence Créative Commons, Éducation ouverte, Information scientifique et technique.

Abstract

In this paper we analyze the impact of a distance learning experience about Open Educational Resources in the practices of Tunisian university teachers who benefited from it.

This study revealed a change of practice related with the use of OER. However, a reluctant behavior was also noted related with the open access publishing. We present major findings in this paper and make interpretation of the results obtained.

Keywords. Open education resources (OER), Open access, Creative commons, Open education, Scientific and technical Information.

1 Introduction

Il est difficile selon Chartron (2016) de préciser exactement la date d'émergence du mouvement du libre accès. La majorité des chercheurs s'accorde néanmoins à croire que ce mouvement a démarré depuis les années 90 (Nétange, 2009), période qui coïncide avec la création d'ArXiv par Paul Ginspar en 1991. Grâce à l'essor d'internet et le développement des technologies des réseaux et du partage des ressources et le processus d'édition électronique, ce mouvement a connu une remarquable évolution attestée notamment par le nombre évolutif des revues et des dépôts en libre accès. Le nombre total de soumissions d'article affiché dans le compteur automatique d'ArXiv depuis sa création en aout 1991 jusqu'au 6 septembre 2018 (après 27,1 ans) avoisine les 1 433.50 articles¹.

En effet, ayant pour objectif principal d'offrir un accès en ligne et gratuit aux publications scientifiques, de nombreuses actions ont été entreprises sous formes de déclarations (déclaration de Budapest 2002, déclaration de Berlin 2003), de politiques éditoriales, voire de lois stimulant et parfois obligeant les chercheurs à publier sous cette nouvelle forme de publication.

Malgré les différentes actions (déclarations de politiques, colloques, journées, etc.), l'adhésion au mouvement du libre accès reste relativement limitée au point qu'il est toujours nécessaire de renforcer la culture du libre accès auprès des chercheurs et des communautés de pratiques. C'est une réalité qui n'est pas propre à des disciplines déterminées, ni à des structures académiques et de recherche précise et encore moins à des pays particuliers. Ce constat s'observe autant dans les pays nantis que dans les pays en voie de développement.

En Tunisie, pays émergent où ces questions d'accès libre se posent avec acuité dans les domaines de l'éducation et de la recherche, plusieurs questions sont déjà en étude pour essayer de comprendre comment développer la culture du libre accès. Parmi ces questions, il s'agit notamment de savoir comment encourager les enseignants universitaires tunisiens à l'usage des Ressources Éducatives Libres (REL) et à la culture du libre-échange sur les réseaux. Autrement, il est question d'identifier les facteurs qui les motiveraient pour adhérer à ce mouvement universel des REL et du libre accès non comme simples utilisateurs mais aussi comme producteurs et diffuseurs de savoir et de connaissances partagées.

¹ArXiv monthly submission rates. https://arxiv.org/stats/monthly_submissions

Dans une tentative d'apporter des éléments de réponses à ces questions, nous avons procédé à une étude à partir d'un retour d'expérience d'une formation sur les REL organisée par l'université virtuelle de Tunis² (UVT) depuis 2017 au profit d'enseignants universitaires tunisiens dans le but de les encourager à utiliser, concevoir et partager des REL. Cette formation s'inscrit dans un ensemble de cycles de formation en ingénierie de la formation en ligne (IFEL) offerts par l'UVT, qui depuis sa création en 2002 veille à développer et à faciliter la mise en place des formations hybrides et à distance auprès des universités tunisiennes.

Considérant que les REL font partie intégrante des publications concernées par le mouvement du libre accès, nous partons de l'hypothèse que cette formation a eu un impact positif auprès des enseignants pour leur faire découvrir et comprendre plusieurs aspects liés à la culture du libre accès et des données ouvertes dans les domaines de l'enseignement et de la recherche.

À l'issue de six sessions de formation, nous avons jugé utile de mener une enquête d'impact et de satisfaction auprès des enseignants formés pour estimer le degré de leur adoption des principes fondamentaux abordés.

Cette formation est basée sur le référentiel des compétences REL de l'Organisation Internationale de la Francophonie (OIF) auquel l'UVT a participé conjointement avec l'UNESCO, l'ALECSO et l'Open Education Consortium (OIF, 2016). En d'autres termes, notre objectif est de vérifier si cette formation a réussi à encourager les enseignants pour créer et mettre des ressources éducatives en libre accès. À cet égard, notre enquête a tenté de répondre aux interrogations suivantes :

- Quelles compétences, les enseignants ont-ils acquis en suivant cette formation ?
- À quel niveau, les enseignants ont-ils adopté et mis en pratique les principes du mouvement des REL ?
- Les enseignants ont-ils publié des REL dans des entrepôts en accès libre ?

Dans notre démarche, nous clarifierons dans un premier temps le lien entre les REL et le libre accès. Nous présenterons ensuite les détails de la formation de l'UVT sur les REL ainsi que les résultats recueillis auprès de 68 enseignants universitaires tunisiens qui ont répondu à l'enquête. Ces résultats portent particulièrement sur la sensibilisation des enseignants aux questions des droits d'auteurs, de l'adoption des licences Creative Commons

² Université virtuelle de Tunis : www.uvt.rnu.tn

ainsi que de la réutilisation des REL. Nous discuterons aussi bien les points forts que les contraintes qui stimulent ou qui freinent l'adoption de cette philosophie de partage parmi la population cible.

2 Quels liens entre REL et mouvement du libre accès ?

Il est admis que le mouvement mondial des REL œuvre en faveur de la création et de la distribution de ressources d'enseignement libres et gratuites, favorisé en cela par l'essor d'Internet et des technologies éducatives du Web.

Grâce à leur potentiel transformateur, leur facilité d'accès et les valeurs qu'elles véhiculent, les REL sont aussi inscrites au cœur du quatrième objectif du Programme des nations-unies pour le développement durable à l'horizon 2030 qui vise à « assurer l'accès de tous à une éducation de qualité, sur un pied d'égalité, et de promouvoir les possibilités d'apprentissage tout au long de la vie » (ONU, 2015).

Plus concrètement, le mouvement des REL s'inscrit dans la continuité des mouvements de l'Open source et du logiciel libre, des données ouvertes et du libre accès ainsi que des licences ouvertes notamment les Creative Commons.

Adoptée depuis 2002 par l'Unesco lors d'un forum sur l'impact des didacticiels libres, les REL occupent de plus en plus d'importance dans les systèmes éducatifs où elles offrent une opportunité stratégique pour améliorer la qualité de l'éducation et partager librement les connaissances en vue d'un renforcement continu des compétences. Leur principe d'ouverture et de gratuité est reconnu dans la définition de l'Unesco qui les considère comme « des matériels d'enseignement, d'apprentissage et de recherche sur tout support, numérique ou autre, existant dans le domaine public ou publiés sous une licence ouverte permettant l'accès, l'utilisation, l'adaptation et la redistribution gratuits par d'autres, sans restrictions ou avec des restrictions limitées... » (Unesco, 2012).

Cette définition des REL croise, en plusieurs points, celle du mouvement du libre accès, inscrite dans les objectifs du congrès de Budapest tenu les 1 et 2 décembre 2001 sur l'autoarchivage et l'accès ouvert. La notion du libre accès a été reprise dans la déclaration de Bethesda (USA) le 11 avril 2003 dont l'une des directives est que « Le/les auteur(s) ainsi que les titulaires du droit d'auteur accordent à tous les utilisateurs un droit d'accès gratuit, irrévocable, mondial et perpétuel et leur concèdent une licence leur permettant de copier, utiliser, distribuer, transmettre et visualiser publiquement l'œuvre et d'utiliser l'œuvre pour la réalisation et la

distribution d'œuvres dérivées, sous quelque format électronique que ce soit et dans un but raisonnable, et ce à condition d'en indiquer correctement l'auteur » (Bernault, 2016, p.22). Ces principes légèrement modifiés ont été ensuite entérinés par la déclaration de Berlin à l'issue d'un congrès tenu par la Société Max-Planck le 22 octobre 2003 (Max-Planck Inst., 2003).

De façon plus concrète, le principe original des REL était spécifiquement destiné à fournir aux étudiants et aux enseignants un accès gratuit aux matériels pédagogiques de l'enseignement supérieur. Le mouvement de l'accès libre a été lancé quant à lui pour permettre le partage en ligne de documents de recherche et de revues universitaires.

En fin de compte, qu'elles soient définies comme REL ou comme ressources en accès libre (archives ouvertes), les matériaux sont gratuits ou normalement bon marché, ou du moins beaucoup moins dispendieux que les matériaux produits dans les circuits commerciaux. Mais comme le terme « accès libre » est un terme générique, toute ressource offerte sous ce label n'est pas nécessairement une REL.

Le droit d'auteur est aussi une question qui doit être prise en compte lors de la publication des REL ou de la publication de documents en libre accès. En effet, avec l'évolution du mouvement du libre, la gestion du droit d'auteur des deux types de ressources, est actuellement, souvent aligné sur les licences Creative Commons. Ainsi, les deux fournissent des options communes pour les éditeurs et les auteurs.

Au vu de ces similitudes, nous considérons que les ressources éducatives font partie intégrante des types de documents pouvant être concernés par l'accès libre à l'information.

3 Contexte de l'étude

C'est précisément pour étudier le degré d'introduction de ces concepts dans le contexte universitaire tunisien et leur appropriation par les enseignants chercheurs, que nous avons choisi de conduire notre étude dans le périmètre d'une formation entièrement en ligne livrée par l'Université virtuelle tunisienne sur les questions des REL et du libre accès. Jusqu'au mois d'avril 2018, nous avons assuré, comme tutrices de formation, six cycles de cette formation programmées dans les orientations majeures de l'UVT et de son plan stratégique 2016-2019.

En effet, se basant sur le référentiel de compétences REL de l'OIF (2016), l'UVT a mis en place, depuis 2017, cette formation à distance au profit des enseignants universitaires. Cette formation couvre la recherche,

la création et la publication des REL. L'objectif en est de les sensibiliser à l'utilisation des REL, mais surtout aux questions des droits d'auteur et l'usage des licences libres.

La formation est proposée entièrement à distance. Elle se déroule sur quatre semaines et porte sur les axes suivants :

- Familiarisation avec les REL
- Recherche des REL
- Évaluation des REL
- Utilisation, création et partage des REL

Au cours de cette formation, les participants sont appelés à consulter des ressources et à réaliser des activités en relation avec chaque aspect de la formation. Ainsi, ils sont tenus de maîtriser les concepts des REL et de l'accès libre, pour être capable de les identifier lors de leur recherche et notamment de maîtriser les questions des droits d'auteurs et des licences Creative Commons qui leurs sont associées.

L'UVT, estime qu'à travers cette formation, les enseignants universitaires seront sensibilisés à ce mouvement et seront enclins à publier leurs ressources éducatives en libre accès aussi bien dans les archives ouvertes que dans les entrepôts de ressources éducatives libres.

4 Objectifs et déroulement de l'étude

L'objectif général de notre étude est de mesurer le degré d'adoption des principes fondamentaux du référentiel REL par les enseignants formés pendant les six cycles de formation auxquels nous avons participé jusqu'en juin 2018.

Étant donné que la formation sur les REL aborde des questions en étroite relation avec le libre accès, notre étude avait comme hypothèse principale de vérifier si cette formation a réussi à sensibiliser les enseignants pour qu'ils créent et partagent des ressources éducatives en libre accès.

Pour réaliser cette étude, nous avons procédé par une enquête sous forme de questionnaire distribué en ligne via Google forms.

L'idée principale de l'enquête était de savoir s'il y a eu une évolution des connaissances et un changement d'attitudes de la part des enseignants par rapport aux notions des REL, de l'accès libre et des droits d'auteur avant et après le suivi de la formation. À cette fin, les questions de l'enquête ont été élaborées dans un esprit de comparaison entre les connaissances, les

motivations et les pratiques des enseignants avant et après la formation sur les REL.

L'enquête a été adressée à tous des participants dans les différents cycles de formation entre 2017 et 2018. L'étude a permis ainsi de couvrir une population cible variée constituée de 141 enseignants dont 68 seulement ont fourni des réponses suite à deux envois successives (juin et juillet 2018). Nous considérons que cet échantillon constitue une masse représentative de l'effectif général, non pas d'un point de vue statistique uniquement mais surtout d'un point de vue des caractéristiques générales des profils des répondants.

Les données de profilage démontrent en effet qu'il ne s'agit pas uniquement de débutants mais plutôt d'enseignants expérimentés chargés d'assurer des cours dont ils ont préparé par eux-mêmes les contenus. En outre, beaucoup d'entre eux ont des publications scientifiques, ce qui constitue un critère fondamental aussi bien pour accéder au poste d'enseignant dans l'enseignement supérieur que pour valider les concours de passage de grade.

L'expérience des répondants est observée aussi à travers leurs tranches d'âge majoritairement situées entre 31-40 ans (52,2%) et 41-50 ans (32,8%) et de leurs anciennetés dans le métier d'enseignement qui dépasse en majorité les dix ans (36,4%). Toutefois, si une contre analyse peut relativiser la qualité de cette expérience en partant du fait que les plus jeunes enseignants seraient plus aptes à maîtriser les technologies numériques pour s'appropriier les concepts des REL, les statistiques montrent aussi que la grande majorité des répondants, y compris les plus âgés, proviennent des disciplines de l'informatique.

Critère	Valeurs
Genre	74,6% femmes 25,4% hommes
Tranche d'âge	52,2% : entre 31- 40 ans 32,8% : entre 41 et 50 10,4% : plus de 50 04,6% : moins de 30 ans
Grades	Tous les grades
Ancienneté dans l'enseignement	36,4% : 11-15 ans 27,3% : 6-10 ans 18,2% : 16-20 ans 9,1% : Plus de 20 ans

Disciplines	informatique, sciences de gestion, biologie, pharmacie, génie mécanique, génie civile, littérature française, mathématiques
-------------	---

Tableau 1 : caractéristiques des profils de la population interrogée

Dans leur globalité, ces indicateurs de profilage donnent une bonne marge de fiabilité aux résultats de l'enquête.

5 Résultats de l'étude

A partir des données recueillies, nous avons procédé à une analyse des résultats en mettant l'accent sur le degré d'appropriation par les enseignants universitaires du concept des REL et des modes de leur utilisation.

Notre analyse a également porté sur l'impact de la formation sur la production et la publication des REL et l'utilisation des licences libres.

5.1 Maitrise des concepts

Dans les détails relatifs aux REL et au libre accès, les résultats de l'enquête font état d'une nette évolution dans l'appropriation de ces concepts par les enseignants concernés. Une comparaison entre l'état des connaissances de ces thèmes avant et après la formation dénote d'une évidente prise de conscience de l'intérêt des REL et de l'accès libre. Les résultats présentés dans la figure 1 montrent, en effet, qu'avant la formation, la majorité des participants n'avaient aucune connaissance ou disposaient des connaissances plutôt faibles dans le domaine du libre accès et des REL.

[Q : Avant de suivre le cycle 2 sur les REL, comment estimez-vous votre connaissance des notions suivantes ?]

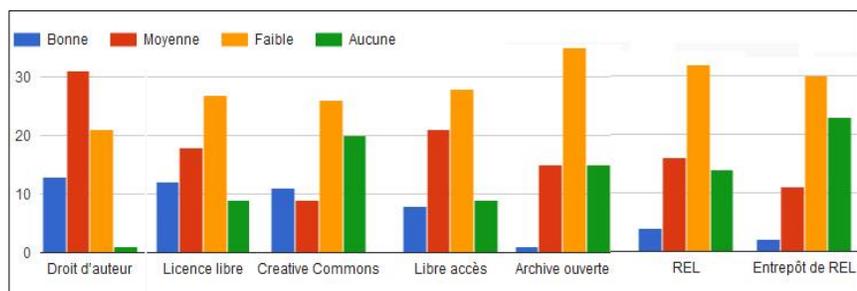


Figure 1 : Compétences avant la formation

Après la formation (fig. 2), la majorité des enseignants interrogés affirment que la formation leur a permis de mieux maîtriser les différentes notions du libre accès, licence libre, REL, entrepôts de REL, droit d'auteurs et Créative Commons.

[Q : La participation au cycle 2 vous a-t-elle permis de comprendre et maîtriser les notions suivantes ?]

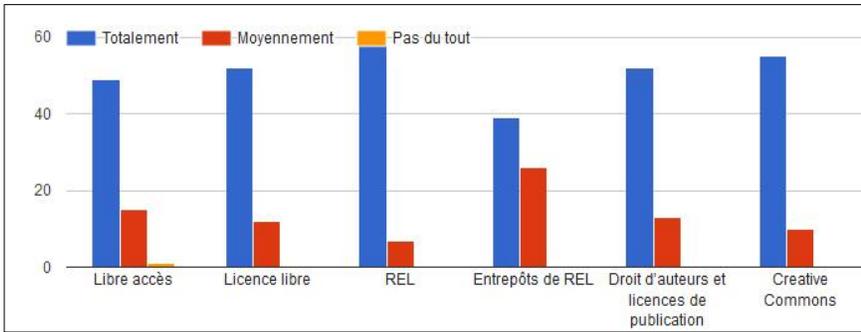


Figure2 : Compétences après la formation

Autour de 85% des répondants (58 sur 68) affirment à titre d'exemple avoir « totalement » compris et maîtrisé le concept des REL après la formation alors qu'ils n'étaient que 4 (5,8%) avant la formation. Environ 80% (55 sur 68) d'entre-deux affirment aussi « totalement » comprendre et maîtriser les *Creative Commons* alors qu'ils ne représentaient que 16% (11 sur 68) avant la formation.

Il est important de signaler qu'à l'issue de la formation, tous les répondants ont affirmé connaître ou comprendre quasiment tous les concepts comme illustré dans la figure 2 alors qu'avant la formation, ils ont déclaré ne disposer d'aucune connaissance à propos d'un ou de plusieurs concepts (fig.1). Les entrepôts de REL, les archives ouvertes et les REL étaient les moins connues alors que les Droits d'auteurs, comme concept courant dans les communautés culturelles et scientifiques, est celui que connaît le plus la population cible.

Notons aussi l'attitude positive des répondants par rapport à l'impact positif de la formation sur leur motivation à s'engager dans l'usage et l'application des différents concepts abordés (fig.3).

À ce propos, notons la seule réserve exprimée au niveau de la motivation mitigée des répondants pour la production et la publication de leurs propres REL, actions pour lesquelles les répondants sont « moyennement » plutôt que « totalement » motivés. Cette constatation traduit le caractère

transitoire du passage de toute appropriation de nouveaux concepts à leur mise en application. Elle impactera, comme il sera démontré sous le point 5.4, la mise en pratique des concepts abordés, notamment la publication des REL et l'utilisation des licences libres qui affichent des taux moins élevés que ceux de la sensibilisation et la motivation. Ce constat devrait d'ailleurs faire l'objet d'une analyse approfondie et de recommandations appropriées.

[Q : La formation vous a-t-elle encouragé/incité à :]

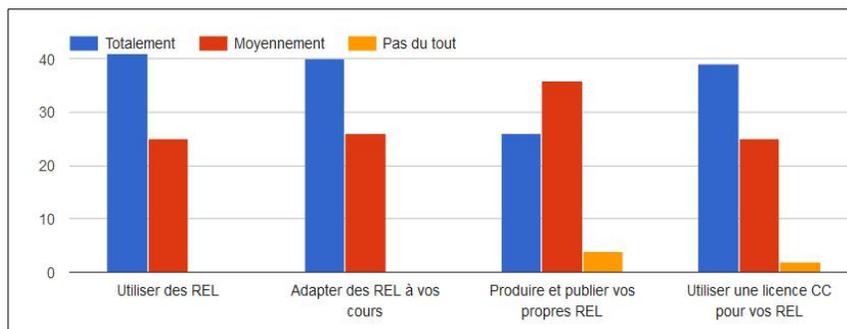


Figure 3 : Niveau d'incitation de la formation sur les concepts

Ce sont là grosso modo les premiers indicateurs prouvant que les résultats de l'enquête ont confirmé chez les enseignants une progression à deux vitesses : d'une part une sensibilisation et une motivation générale pour comprendre et s'approprier des concepts comme les REL et les licences libres, mais d'autre part une lenteur, voire une réticence, dans leur usage et mise en application pour la conception et la production de nouvelles ressources. La formation aurait ainsi atteint son objectif de sensibilisation des enseignants sur les enjeux des concepts clés abordés, mais il lui reste de confirmer, au-delà de la sensibilisation et de la motivation, le niveau d'usage de ces concepts et leur mise en application par les enseignants dans leurs activités pédagogiques et de recherche.

5.2 Utilisation des REL

L'utilisation des ressources est une phase intermédiaire entre la sensibilisation et la production. Elle devrait normalement constituer un premier ancrage des enseignants autour des nouveaux concepts acquis en formation, d'autant plus qu'avant la formation, 61,2% d'entre eux affirment ne jamais avoir utilisé des REL puisqu'ils n'en connaissent pas l'existence, contre seulement 28,4% qui affirment en avoir rarement fait usage (fig. 4).

[Q : Avant de suivre le cycle 2, avez-vous utilisé des REL ?]

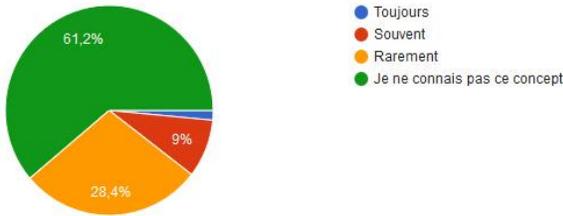


Figure 4 : taux d'usage des REL avant la formation

Après la formation et la découverte des concepts par les enseignants, on comprendrait qu'il n'y avait pas eu une évolution significative dans l'usage des REL. Cette déduction se fonde sur les taux d'usage des licences libres (cf. point 5.5), eux-aussi moins élevés que les taux de sensibilisation et de motivation signalés précédemment (fig. 2, 3).

Plus concrètement, si les REL tirent leur nom de leur publication sous une licence libre, et que seulement 35,4% des enseignants affirment n'utiliser des ressources qu'après avoir vérifié leurs types licences (fig. 5), nous y voyons un indicateur signifiant que l'usage des REL après la formation n'a pas été suffisant. Mais face à cette insuffisance, la formation avait au moins permis d'attirer l'attention des enseignants sur l'existence des REL et leur valeur ajoutée dans l'activité pédagogique. C'est ce que confirment l'analogie entre les deux figures 1 et 2 par rapport à la bonne connaissance et compréhension du concept des REL.

Après les taux élevés de compréhension et la grande motivation pour les concepts, puis les taux moins élevés de leur utilisation, l'étude a permis de démontrer que le niveau de mise en application des concepts notamment dans la production et la publication des REL a été beaucoup moindre.

5.3 Production et publication des REL

En effet, l'engagement pratique des enseignants dans des activités liées aux concepts étudiés reste très timide aussi bien avant qu'après la formation (fig.5). Par rapport à la question relative au dépôt de documents dans une archive ouverte ou dans un dépôt en libre accès, les taux avant et après la formation sont quasi identiques, affichant de part et d'autre des taux très modestes.

[Q1 : Avant de suivre le cycle 2 avez-vous déposé un document dans une archive ouverte ou un dépôt en libre accès ?]

[Q2 : Après la formation, avez-vous publié un de vos documents pédagogiques ou de recherche (cours, exercice, article scientifique, communication, mémoire, etc.) en libre accès]

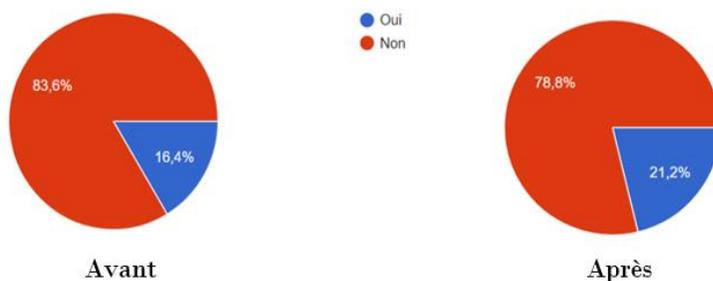


Figure 5 : taux de publication de REL en libre accès avant et après la formation

De façon plus détaillée, avant la formation, seulement deux des onze répondants à cette question (18% mais 2,9% de tous les répondants) affirment avoir déposé des « ressources éducatives » contre 9 sur 11 (82% mais 13,2% de tous les répondants) qui affirment avoir déposé des « articles scientifiques ». Cet écart s'explique par la culture dominante de la recherche scientifique comme seule activité pouvant donner lieu à des publications. Cet écart explique surtout la tendance générale vers la valorisation et la visibilité universitaire par les seuls produits de la recherche scientifique plutôt que par les ressources pédagogiques et les documents académiques comme les mémoires et les thèses (aucun mémoire et une seule thèse déposée dans les archives ouvertes).

Les répondants expliquent leurs faibles taux de publication de REL par trois raisons principales : ne pas connaître comment publier en libre accès (31 sur 55 répondants = 56,4%), ne pas connaître où publier (24 sur 55 répondants = 43,6%) et ne pas connaître le principe du libre accès (21 sur 55 répondants = 38,2%).

Ceci confirme la particularité évoquée sous le point 5.2, relative à la motivation des enseignants qui auraient plutôt tendance à se focaliser plus sur l'usage de l'existant que de produire et publier de nouvelles ressources libres.

Après la formation, seulement 5% des enseignants ont publié en libre accès. Pour expliquer les raisons de cette attitude passive, 72% affirment qu'ils comptent le faire prochainement, 25% déclarent ne pas en être motivés car la publication en libre accès n'est pas reconnue pour la

promotion professionnelle. Environ 15% craignent encore la violation de leurs droits et pensent que la publication en libre accès augmente le risque de plagiat. Autour de 8% ont répondu qu'ils ne savent pas encore où publier.

5.4 Utilisation des licences libres

Le taux de publication limité de REL impacte automatiquement celui de l'usage des licences libres. Toute licence n'est appliquée que si ressource est produite puis conditionnée pour être diffusée. Or, étant donné le taux réduit de la publication des REL par les enseignants, le taux de sensibilisation et de vérification des licences libres est lui-aussi systématiquement réduit avant et après la formation (fig.6).

Il s'avère en effet, qu'environ 43,9% des répondants affirment ne « jamais » avoir vérifié les licences et les droits d'auteur des œuvres qu'ils utilisaient sur Internet avant de suivre la formation. Autant de répondants ne le faisaient que « rarement ». Après la formation, 35,4% déclarent « toujours » vérifier les licences et 60% s'assurent « souvent » de la licence attribuée aux œuvres utilisées et consultées sur le net (fig. 6). La culture des droits d'auteurs et des licences libres en est ainsi impactée.



Figure 6 : taux de vérification des droits d'auteurs avant et après la formation

Venant à la question de l'usage des licences libres pour tous types de documents pédagogiques ou de recherche, les taux ne sont pas meilleurs puisque des 66 répondants à cette question (97%), 51 (78,8%) répondent par la négative et seulement 14 (21,2%) confirment avoir utilisé une licence libre au moins une fois (fig.7).

[Q : *Après la formation, avez-vous publié un de vos documents pédagogiques ou de recherche (cours, exercice, article scientifique, communication, mémoire, etc.) en libre accès ?*]

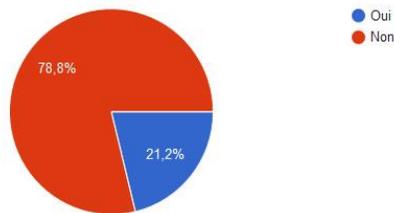


Figure 7 : taux d'usage des licences libres

Dans l'ensemble des 51 répondants qui déclarent ne pas utiliser les licences libres, l'argument majeur pour 37 d'entre eux (72,5%) est l'intention de le faire ultérieurement. Bien que non fondée, une promesse est toujours mieux qu'un refus catégorique, ce qu'un seul répondant d'ailleurs a confirmé sous prétexte qu'il n'a pas la compétence pour le faire.

Du côté des 14 répondants qui affirment avoir utilisé une licence libre, 7 d'entre eux (50%) déclarent le faire sporadiquement pour quelques-unes de leurs publications aussi bien pédagogiques que scientifiques. Les autres se partagent à parts égales les choix entre ressources pédagogiques et ressources scientifiques, archives ouvertes ou entrepôts de ressources pédagogiques.

6 Recommandations

Il est évident au vu des constats élaborés tout au long de notre analyse que la formation de l'UVT a permis aux enseignants de bien comprendre le concept des REL et du libre accès, qu'elle les a sensibilisés vis-à-vis du respect des droits d'auteurs et les a motivés à utiliser les RELs. La formation n'a pas réussi toutefois à engendrer une dynamique productive et de publication de REL sous des licences de libre accès. Bien qu'une majorité compte le faire ultérieurement, nous pensons qu'un soutien supplémentaire est à accorder au corps enseignant, que ce soit de la part du ministère de tutelle ou bien des structures d'attache. Nous proposons à cet égard les recommandations suivantes :

- Élargir les formations de ce type à un public plus large provenant de toutes les universités tunisiennes ;
- Promouvoir la culture des REL et des licences ouvertes dans les universités et les communautés de la recherche scientifique par l'organisation des journées de sensibilisation ;
- Développer une politique nationale des REL et de l'accès libre qui serait transdisciplinaire et transversale aux universités publiques et privées ;

- Créer des référentiels nationaux pour la production des REL, l'accès libre et les licences ouvertes ;
- Renforcer l'équilibre entre ressources pédagogiques et ressources scientifiques dans l'échelle des valeurs de la qualité scientifique des enseignants-chercheurs.

7 Conclusion

En définitive, la décroissance dans les statistiques observées entre la motivation pour les REL, leur usage et production et enfin leur publication sous une licence libre est tout à fait logique et compréhensible. Si l'on procède par une analyse de cause à effet, tout nouveau concept passe impérativement par une première phase d'appropriation et d'adhésion de la personne concernée au bien-fondé de ce concept. Cette même personne doit ensuite passer par une phase de découverte et de validation de ce concept dans son milieu naturel avant de le soumettre à un exercice de production et de démarche qualité. Par conséquent, si les enseignants ont été sensibilisés et motivés aux principes des REL, des licences libres, de l'accès ouvert, des entrepôts de REL et des archives ouvertes, il n'en demeure pas moins qu'il nécessite une longue période de « digestion » de ces concepts avant de passer en phase de production qualitative de ressources éducatives avec des licences libres.

Or, comme annoncé précédemment, notre étude a été conduite quelques mois après la fin de la dernière formation clôturée en avril 2018. Les quelques mois qui séparent la dernière formation et l'enquête réalisée ne sont pas suffisants pour engendrer un effet immédiat sur des sujets qui sortent tout juste d'une formation de sensibilisation. Il faudrait à notre sens revoir les mêmes critères auprès du même public dans un délai d'au moins une année pour voir si les notions acquises en formation auront permis aux enseignants de passer au stade de la production qualitative.

Entre-temps, et au vu des résultats obtenus, nous pensons que l'objectif de sensibilisation de la formation a été atteint sur les différents concepts des REL, des licences libres et des concepts associés. Si les recommandations proposées précédemment seraient tenus en compte, il y a des fortes chances d'avoir dans un avenir proche une catégorie d'enseignants d'un niveau professionnel dans ces différents domaines.

Références bibliographiques

- BERNAULT, Carine (2016). *Open access et droit d'auteur*. Éditions Larcier, 232 p.
- CHARTRON, Ghislaine (2016). Stratégie, politique et reformulation de l'open access. In *Revue française des sciences de l'information et de la communication*, n° 8, janvier 2016. journals.openedition.org, doi:10.4000/rfsic.1836.
- Max-Planck Inst. (2003). *Déclaration de Berlin sur le Libre Accès à la Connaissance en sciences exactes, sciences de la vie, sciences humaines et sociales - Libre accès à l'information scientifique et technique*. 22 octobre 2003, 2 p.
- NÉTANGE, Héloïse (2009). L'accès libre et gratuit à l'information scientifique. In *Revue internationale d'éducation de Sèvres*, n° 51, septembre 2009, p. 12-15. journals.openedition.org, doi:10.4000/ries.612.
- OIF (2016). *Référentiel de compétences sur les Ressources éducatives libres (REL)*. Organisation Internationale de la Francophonie, 2016. 105 p.
- ONU (2015). Objectifs de développement durable - objectif 4 : éducation de qualité. [En ligne]. Disponible à : <https://www.un.org/sustainabledevelopment/fr/education/> (Page consultée le 30 août 2018)
- UNESCO (2012). *Déclaration de Paris sur les REL 2012, Congrès mondial sur les Ressources Éducatives Libres (REL) 2012*, Unesco, Paris, 20-22 juin 2012, 3 p.

Analyse de l'offre des objets pédagogiques numériques dans les archives ouvertes

Analysis of the supply of digital learning objects in open repositories

Mohamed Ben Romdhane ; Rachid Zghibi
mbromdhane@gamil.com, rachid_zghibi@yahoo.fr
Univ. Manouba, ISD, BNP UR11ES43, Tunisie

Résumé

Ce travail de recherche s'inscrit dans le cadre d'une étude exploratoire relative à l'analyse de l'offre des objets pédagogiques dans les archives ouvertes à l'échelle internationale en nous basant sur le répertoire OpenDOAR comme source d'information principale. Parmi les 545 dépôts recensés, 177 ont été sélectionnés formant, ainsi, le corpus de l'étude.

Pour l'analyse du contenu des dépôts retenus, nous nous sommes fondés sur un ensemble des critères tels que l'origine géographique, le logiciel utilisé, la ou les langue(s) des ressources et de l'interface, le schéma des métadonnées adopté, le nombre d'objets pédagogiques, les licences et droits d'auteurs, etc. Les résultats de cette étude montrent que cette offre reste.

Mots-clés. Objet pédagogique, archives ouvertes, libre accès, REL, Ressources Educatives Libres, enseignement supérieur.

Abstract

This work is an exploratory study on the analysis of the supply of learning objects in internationally open archives based on the OpenDOAR directory as the main source of information. Among 545 deposits identified in OpenDOAR, 177 were selected forming the corpus of the study.

To analyzing the content of the selected repositories, we have used a set of criteria such as the geographical origin, the software used, the language (s) of the resources and the interface, the metadata schema adopted, the number of learning objects, licenses and copyrights, etc. The results of this study show that this offer remains modest.

Keywords. Learning object, repositories, Open Access, OER, Open Educational Resources, Higher education.

1 Introduction

Le mouvement du libre accès à l'information scientifique et technique (IST) a commencé depuis environ deux décennies dans l'objectif d'ouvrir l'accès aux résultats de la recherche librement sans barrières techniques, financières ou autres. L'initiative de Budapest pour l'accès ouvert de 2002 définit le libre accès comme suit : «Par "accès libre" à cette littérature, nous entendons sa mise à disposition gratuite sur l'Internet public, permettant à tout un chacun de lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces articles, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale, sans barrière financière, légale ou technique autre que celles indissociables de l'accès et l'utilisation d'Internet. La seule contrainte sur la reproduction et la distribution, et le seul rôle du copyright dans ce domaine devrait être de garantir aux auteurs un contrôle sur l'intégrité de leurs travaux et le droit à être correctement reconnus et cités ». Le principe d'ouverture décrit ici pour les articles scientifiques s'applique à tous les autres types de publications.

C'est ainsi que durant les premières années du mouvement, on s'intéressait à la mise en libre accès, principalement, des publications scientifiques, y compris les pré-publications et les post-publications, et des travaux académiques. Pour cela, les deux voies du libre accès se sont développées à savoir la voie verte qui est les archives ouvertes ou dépôts numériques et qui concerne l'archivage de tous types de travaux académiques et/ou scientifiques et la voie dorée qui est les revues scientifiques en libre accès.

Au fil des années, les politiques envers le libre accès à l'IST se sont développées au niveau des institutions, des pays et même des régions. Le mouvement s'est développé pour toucher non seulement les publications issues de la recherche scientifiques mais aussi les données de la recherche, on parle de plus en plus de l'ouverture de la science ou encore de l'Open Science ou Open Research. D'autres domaines se sont alignés sur cette logique de partage et d'ouverture comme en informatique avec les logiciels libres, en journalisme avec le libre accès à l'information, en administration le droit d'accès aux documents administratifs, l'Open Data, etc.

Le domaine de l'enseignement supérieur, quant à lui, n'a pas tardé à adhérer au mouvement du libre accès. En effet, plusieurs entrepôts d'archives ouvertes appartenant à des établissements universitaires ou à des universités publient en libre accès bon nombre d'objets pédagogiques parallèlement aux autres publications scientifiques et académiques. Certains de ces dépôts sont réservés aux objets pédagogiques qui sont destinées, en particulier, à la communauté universitaire ainsi qu'à quiconque voulant bénéficier de ces mines de ressources d'information en général. Néanmoins, ces ressources, malgré leur importance, restent parfois méconnues par les utilisateurs et manquent de visibilité et d'accessibilité. Partant de ce constat, nous avançons les interrogations suivantes : que représente cette offre d'objets pédagogiques dans ces dépôts d'archives ouvertes ? quelles sont ses caractéristiques ? et dans quelle mesure est-elle qualifiée du libre ?

C'est ainsi que le présent travail s'inscrit dans le cadre d'une étude exploratoire relative à l'analyse de l'offre de ce type des ressources dans les archives ouvertes dans l'objectif de comprendre la répartition de cette offre et ses principales caractéristiques. Les résultats obtenus, pourraient être utilisés par les établissements universitaires pour améliorer leurs offres tout en respectant les caractéristiques essentielles d'un objet pédagogique et les exigences du libre accès.

2 Libre accès à l'IST et objets pédagogiques numériques

2.1 Ressources couvertes par le mouvement du libre accès

Le mouvement du libre accès a débuté durant les années 1990 avec le lancement de la première archive ouverte arXiv. Il vise essentiellement à contourner le prix en hausse des publications scientifiques en particulier celui des revues scientifiques monopolisées par les éditeurs commerciaux. Au fil des années, les déclarations en faveur de l'Open Access venant des différents acteurs de la communication scientifique se sont multipliées donnant, ainsi, naissance aux deux voies du libre accès à savoir la voie verte, représentée par les archives ouvertes et la voie dorée représentée par les revues en libre accès.

Si la publication dans les revues en libre accès obéit à des règles d'évaluation et de qualité bien connues par la communauté des chercheurs, il n'en est pas de même pour les archives ouvertes qui sont beaucoup moins exigeants puisque les règles de dépôt sont généralement déterminées par les institutions en question.

Dans le cadre de ce travail de recherche, notre intérêt s'oriente aux archives ouvertes qui étaient les premiers à se développer dans ce mouvement. Parallèlement aux publications scientifiques (pré-publication et post-publication) et les travaux académiques (principalement les thèses et les mémoires), plusieurs dépôts donnent accès à une grande variété d'objets pédagogiques pour servir la communauté universitaire.

Étant donné son caractère foncièrement éducatif, cette catégorie des ressources devient une composante essentielle de plusieurs dépôts appartenant à des universités et à des établissements d'enseignement supérieur voire la seule composante dans beaucoup d'autres.

Dans le même ordre d'idées, le glossaire sur l'Open Access de l'INIST¹, fait référence aux données de l'enseignement en plus de celles de la recherche dans sa définition d'une archive ouverte en stipulant que "le terme archive ouverte désigne un réservoir où sont déposées des données issues de la recherche scientifique et de l'enseignement et dont l'accès se veut ouvert c'est-à-dire sans barrière. Cette ouverture est rendue possible par l'utilisation de protocoles communs qui facilitent l'accessibilité de contenus provenant de plusieurs entrepôts maintenus par différents fournisseurs de données".

¹<http://openaccess.inist.fr/spip.php?page=glossaire>

Dans la typologie des documents concernés par les archives ouvertes dressée par Gabriel Gallezot on trouve le cours qui n'est pas diffusé sous sa forme classique mais plutôt sous forme de synopsis (ppt) ou contextualisé sur une plateforme pédagogique (GALLEZOT, 2008).

1.2 Objet pédagogique

2.2 Historique et définitions

Selon Olivier Catteau, le terme objet pédagogique (Learning Object en anglais) est devenu d'usage courant à partir de 1994 lorsque Wayne Hodgins a baptisé le groupe de travail de l'association CEDMA (Computer Education Managers Association) Learning Architectures, APIs and Learning Object (CATTEAU, 2008). A l'heure actuelle, il existe plusieurs appellations qui sont utilisées comme des synonymes tels que ressource d'enseignement et d'apprentissage (REA), ressource d'information pour l'enseignement et l'apprentissage (RIEA), ressource éducative, objet d'apprentissage, objet éducatif, matériel éducatif, etc.

De même, il n'existe pas une seule définition de ce qu'on entend par objet pédagogique. Selon le groupe de travail P1484.12 «Learning Object Metadata (WG12)» de l'IEEE-LTSC², un objet pédagogique est «toute entité, numérique ou non, qui peut être utilisée, réutilisée ou référencée lors d'une formation dispensée à partir d'un support technologique»³. Cette définition restreint l'objet pédagogique aux entités, numériques ou non, qui sont utilisées uniquement dans le cadre des systèmes d'apprentissage et de formation assistés par ordinateur et néglige, en revanche, le cas des formations traditionnelles ou classiques.

D'autres définitions plus restrictives sont proposées, dont celle d'Allard Strijker qui considère un objet pédagogique comme des entités fondamentalement numériques utilisables et réutilisables dans des différentes situations d'apprentissage et de formation. Du point de vue technique, ces entités peuvent être simples comme un extrait du texte ou une image, ou complexes tels qu'un fichier pdf ou une ressource hypermédia interactive (STRIJKER, 2004).

Plus restrictive encore, un consortium regroupant des universités de l'Etat de Wisconsin aux Etats-Unis d'Amérique dont le but est de développer des objets pédagogiques réutilisables et accessibles gratuitement via son dépôt Wisc-Online définit l'objet pédagogique comme des petites unités d'apprentissage autonomes et d'une durée comprise entre 2 et 5 minutes. De ce fait, ils sont suffisamment petits pour être intégrés à une activité pédagogique, une leçon, un module ou un cours⁴.

²L'IEEE (Institute of Electrical and Electronics Engineers) est un organisme international de standardisation des domaines du génie informatique, de la technologie biomédicale, de la télécommunication et de l'énergie électrique, etc. Le LTSC (Learning Technology Standards Committee) est un consortium international au sein de l'IEEE qui développe des standards techniques et des recommandations relatifs au domaine des technologies de l'enseignement et de la formation à distance parmi lesquels le standard des métadonnées pédagogiques (LOM).

³<http://ltsc.ieee.org/>

⁴Wisc-Online <https://www.wisc-online.com/>

2.3 Enseignement supérieur : apports des objets pédagogiques en libre accès

Les systèmes d'enseignement supérieur sont, de plus en plus, confrontés à d'énormes défis liés principalement à l'augmentation des effectifs d'étudiants dans le monde entier. D'après l'UNESCO, les établissements d'enseignement supérieur accueilleraient plus de 262 millions d'étudiants d'ici 2025, contre 97 millions en 2000⁵. Cette augmentation risque de ne pas être accompagnée d'un accroissement équivalent des ressources humaines et financières notamment dans beaucoup des pays à faibles et moyens revenus.

En s'inscrivant dans la logique du libre accès, les objets pédagogiques numériques contribuent à faire progresser l'enseignement supérieur partout dans le monde en apportant des solutions aux divers défis qu'affrontent les établissements universitaires ainsi que les enseignants universitaires. En effet, comparé aux médias d'apprentissage traditionnels, les objets pédagogiques se caractérisent notamment par (CATTEAU, 2008, BOURDA, 2002):

- **La réutilisabilité** : du fait qu'ils sont créés sous forme de petites composantes, les objets pédagogiques peuvent être réutilisés dans différents contextes pédagogiques et pour différentes fins (enseignement supérieur ou autres). Ils sont également autonomes et peuvent, par conséquent, être utilisés sous leurs formes initiales indépendamment des autres ;
- **L'agrégation** : les objets pédagogiques de granularité fine peuvent être assemblés à d'autres objets pédagogiques existants pour créer des nouveaux objets répondant aux besoins spécifiques des enseignants universitaires et des étudiants ;
- **L'accessibilité** : les objets pédagogiques sont des entités numériques livrables sur Internet ce qui permet à plusieurs personnes d'y accéder et les utiliser simultanément. Ils sont indexés à l'aide des schémas de métadonnées standardisés permettant de le retrouver facilement ;
- **La collaboration** : les objets pédagogiques favorisent la collaboration des enseignants universitaires qui les utilisent en permettant la création des nouvelles versions ;
- **La rentabilité économique** : les objets pédagogiques en accès libre permettent aux établissements universitaires de faire de substantielles économies en offrant une solution alternative aux coûts élevés de la production de certains matériels pédagogiques tels que les simulations, les animations informatiques interactives, les vidéos, les manuels, etc.

3 Méthodologie

Les dépôts numériques ou archives ouvertes, surtout de type institutionnel, renferment généralement une grande variété de ressources numériques, tels que les articles, les livres, les thèses, les mémoires et bien d'autres.

⁵<http://www.noirsurlblanc.com/population-etudiante-dans-le-monde/>

En plus de ces types de publications scientifiques et académiques, plusieurs dépôts relevant du domaine de l'enseignement supérieur et la recherche scientifique (universités, établissements d'enseignement supérieur, bibliothèque universitaires...), donne accès à une catégorie particulière des ressources, à savoir les objets pédagogiques ou ressources éducatives. Il s'agit de tout type de matériel, quel qu'en soit la forme et le contenu, qui peut être utilisé dans une activité d'enseignement, de formation et d'apprentissage.

Notre étude consiste à analyser les archives ouvertes qui donnent accès à ce type des ressources en nous fondant sur le répertoire OpenDOAR⁶ qui recense, à l'heure actuelle, 545 dépôts renfermant des contenus de type «Learning Object». Notre méthodologie du travail se résume, ainsi, aux deux étapes suivantes:

- Exploration et sélection des dépôts en question selon un certain nombre des critères préalablement définis,
- Analyse du contenu du corpus.

3.1 Exploration et sélection du corpus

Cette étape consiste à explorer les 545 dépôts référencés par le répertoire OpenDOAR afin de sélectionner les dépôts qui répondent aux critères suivants :

- **Accessibilité** : c'est-à-dire vérifier si le dépôt s'ouvre correctement en activant son adresse URL fourni par le répertoire OpenDOAR. L'application de ce critère a permis d'écarter 79 dépôts du corpus (sites introuvables).
- **Objet pédagogique** : il s'agit de vérifier si le dépôt contient réellement des ressources éducatives. A cette fin, nous avons cherché cette confirmation à travers la page d'accueil du dépôt (présentation, annonces, collections et sous collections...) et l'interface de recherche notamment en mode avancé qui permet de rechercher souvent par type des ressources. Il est à noter que plusieurs appellations sont utilisées pour désigner la notion de « Learning Object » tels que cours, objets d'enseignement, objets pédagogiques, matériel didactique, matériel d'apprentissage, Ressources Éducatives Libres, ressources d'enseignement, etc. Cette vérification a permis d'écarter 254 dépôts de la liste retenue de l'étape précédente.
- **Enseignement supérieur** : pour les 212 dépôts retenus, nous avons vérifié si leurs collections d'objets pédagogiques sont destinées à un usage au niveau de l'enseignement supérieur. Ce critère est déterminant car le libre accès concerne uniquement ce niveau d'enseignement. C'est ainsi que nous avons exclu 12 dépôts qui s'adressent à d'autres niveaux d'enseignement et d'apprentissage

⁶ OpenDOAR : The Directory of Open Access Repositories est un répertoire des Archives ouvertes universitaires. Chaque dépôt référencé est visité par le personnel du projet pour vérifier les informations fournies. Ce répertoire qui recense 3520 dépôts au 1^{er} août 2018 est développé par l'Université de Nottingham au Royaume Uni. URL :<http://opendoar.org/>

comme l'enseignement secondaire, la formation continue, la formation professionnelle, l'enseignement primaire, etc.

- **Accessibilité aux ressources** : Enfin, pour les 200 dépôts retenus, nous avons vérifié si les objets pédagogiques sont accessibles au public ou non. Ce critère nous a permis d'écarter 23 dépôts qui exigent une authentification préalable pour pouvoir accéder au texte intégral de toutes les ressources du dépôt.

Cette première étape, nous a permis de retenir 177 dépôts seulement qui constitueront le corpus de notre analyse.

3.2 Analyse du corpus retenu

Pour analyser le contenu des dépôts retenus, nous nous sommes fondés sur les critères suivants : origine géographique (pays et continent), logiciel utilisé, langue(s) des ressources et de l'interface, les REL (Ressources Éducatives Libres), les schémas des métadonnées adoptés, le nombre d'objets pédagogiques et enfin les licences et les droits d'auteurs. Il est à noter que notre étude se limite à une analyse de contenu global et que nous ne prétendons en aucun cas faire une évaluation de ces dépôts (interface, qualité et diversité du contenu, ...) ni des ressources qu'ils renferment (qualité, usage, ...). Ces types d'évaluation nécessitent d'autres démarches et d'autres outils de travail.

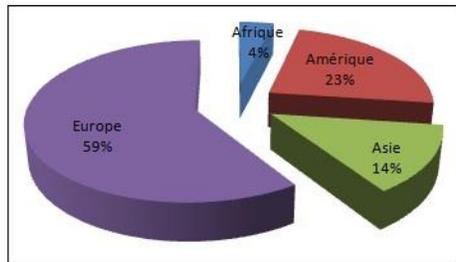


Figure 1 : Répartition des dépôts par continent

4 Résultats et discussions

4.1 Origine géographique des objets pédagogiques

La figure 1 montre que la plupart des dépôts renfermant des objets pédagogiques sont européens 59%. Le continent américain vient en deuxième position avec 23%. L'Asie renferme 14% des dépôts alors que le continent Africain ne renferme que 7 dépôts soit 4% seulement du corpus.

Tableau 1 : Répartition des dépôts par pays

<i>Pays</i>	<i>Nbre</i>	<i>%</i>
<i>United States</i>	18	10,2
<i>France</i>	16	9,0
<i>Germany</i>	15	8,5
<i>Spain</i>	14	7,9
<i>Poland</i>	11	6,2
<i>United Kingdom</i>	10	5,6
<i>Canada</i>	8	4,5
<i>Japan</i>	8	4,5
<i>Indonesia</i>	7	4,0
<i>Ukraine</i>	6	3,4
<i>Colombia</i>	5	2,8
<i>Norway</i>	4	2,3
<i>Mexico</i>	4	2,3
<i>India</i>	4	2,3
<i>Russian Federation</i>	4	2,3
<i>Italy</i>	4	2,3
<i>Brazil</i>	3	1,7
<i>Belarus</i>	3	1,7
<i>Greece</i>	3	1,7
<i>Kenya</i>	3	1,7
<i>Argentina</i>	2	1,1
<i>Czech Republic</i>	2	1,1
<i>Estonia</i>	2	1,1
<i>Taiwan</i>	1	0,6
<i>Botswana</i>	1	0,6
<i>Turkey</i>	1	0,6
<i>Hungary</i>	1	0,6
<i>Slovenia</i>	1	0,6
<i>Portugal</i>	1	0,6
<i>Denmark</i>	1	0,6
<i>China</i>	1	0,6
<i>Sudan</i>	1	0,6
<i>Sri Lanka</i>	1	0,6
<i>Luxembourg</i>	1	0,6
<i>Switzerland</i>	1	0,6
<i>Costa Rica</i>	1	0,6
<i>Thailand</i>	1	0,6
<i>Moldova</i>	1	0,6
<i>Uganda</i>	1	0,6
<i>Nigeria</i>	1	0,6
<i>Finland</i>	1	0,6
<i>Croatia</i>	1	0,6
<i>Bulgaria</i>	1	0,6
<i>Cyprus</i>	1	0,6
44	177	100,0

Quant à la répartition par pays, le tableau 1 montre que 44 pays sont représentés dans le total des dépôts du corpus. Les Etats Unis vient en premier lieu avec 18 dépôts suivi de 5 pays européens à savoir la France, l'Allemagne, l'Espagne, la Pologne et le Royaume Unis avec respectivement 16, 15, 14, 11 et 10 dépôts. Il est à signaler que la France vient en deuxième position vue la politique nationale incitant les universités et les institutions d'enseignement et de la recherche françaises à créer leurs propres dépôts dans le cadre de l'archive ouverte multidisciplinaire choisie par l'ensemble de la communauté scientifique française pour la diffusion des savoirs HAL (Hyper Articles en Ligne) qui est à la fois un entrepôt unique et une application développée et maintenue par le CCSD (Centre pour la Communication Scientifique Directe). Il s'agit

en fait de la seule archive ouverte multidisciplinaire HAL qui fédère les dépôts de chaque institution.

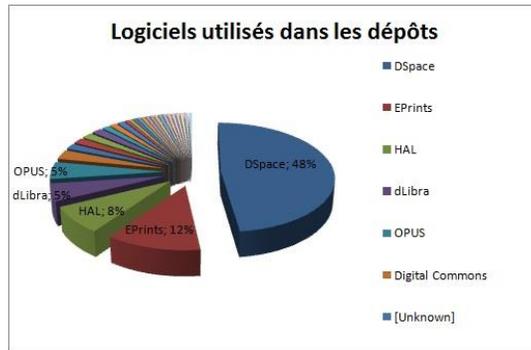


Figure 2 : logiciels utilisés par les dépôts

4.2 Logiciels utilisés dans les dépôts

La figure 2 montre que DSpace est le logiciel le plus utilisé par les dépôts de notre corpus puisque 48% des dépôts gèrent leurs ressources avec ce logiciel. Ce logiciel libre, développé en 2002 à l'initiative de MLT et les laboratoires HP à Cambridge et destiné à la construction d'archives ouvertes et de bibliothèques numériques reste le logiciel le plus utilisé dans les projets Open Access à travers le monde puisqu'il assure l'interopérabilité à travers le protocole OAI-PMH et se base sur des métadonnées standardisées principalement le Dublin Core. Le deuxième logiciel utilisé par notre corpus est Eprints avec 12%. Ce logiciel a été créé en 2000 par l'Université de Southampton. Il est aussi l'un des logiciels libres les plus utilisés dans les archives ouvertes puisqu'il se base sur le protocole OAI-PM. Le troisième logiciel plutôt français HAL est utilisé par 8% des dépôts exclusivement français puisqu'il s'agit d'un produit développé en France par le CCSD et destiné aux institutions françaises pour créer leurs propres archives ouvertes et alimenter le dépôt national français HAL.

1.3 Langue des ressources et des interfaces

Pour chaque dépôt du corpus il s'agit d'identifier la ou les langues des objets pédagogiques offertes en se basant sur les métadonnées ou sur le classement des ressources par langue dans certains dépôts. Les résultats montrent que l'anglais reste la langue la plus utilisée dans les objets pédagogiques offerts par les dépôts de notre corpus 66.1% comme dans leurs interfaces 82.5% (Tableau 2 et 3). N'oublions pas que ces dépôts sont à la base des archives ouvertes universitaires destinés principalement à l'archivage et à la diffusion en Open Access de la production scientifique des universités. Cette production est produite généralement en langue anglaise, langue de la science, pour assurer une meilleure visibilité de ces documents. Ce qui explique le nombre élevé des interfaces en anglais par rapport aux objets pédagogiques.

Tableau 2 : des ressources

Langue	Nbre dépôts	%
English	117	66,1
Spanish	31	17,5
French	17	9,6
German	13	7,3
Polish	11	6,2
Japanese	8	4,5
Indonesian	7	4,0
Ukranian	6	3,4
Russian	6	3,4
Norwegian		
Italian	4	2,3
Deutch		
Portuguese		
Greek		
Pvccknn	3	1,7
Chenese		
Finnish	2	1,1
Czech		
Estonian		
Danish		
Turkey		
Hungarian		
Croatian		
Bulgarian		
Hindi	1	0,6
Thai		
Arabic		
Swahili		
Catalan		

Tableau 3 : langue de l'interface

Langue interface	Nbre	%
English	146	82,5
Spanish	27	15,3
French	24	13,6
Deutch	18	10,2
Polish	11	6,2
Japanese	7	4,0
Russian	6	3,4
Ukranian		
Indonesian	5	2,8
Norwegian		
Italian	4	2,3
Pvccknn		
Greek		
Portuguese	3	1,7
Chenese		
Estonian		
Czech	2	1,1
Catalan		
German		
Hungarian		
Slovenian		
Bulgarian		
Swedish	1	0,6
Turkev		
Croatian		
Thai		

L'espagnol est en deuxième position que ce soit pour la langue des objets pédagogiques, 17,5 % soit 31 dépôts (tableau 2), ou des interfaces 15,3% soit 27 dépôts (tableau 3), malgré que l'Espagne est en quatrième position dans le nombre de dépôts avec 14 dépôts seulement (voir tableau 1). Ceci s'explique par l'usage de cette langue par pas mal d'autres pays comme ceux de l'Amérique Latine.

Il est à noter ici que la langue arabe est représentée par le seul dépôt soudanais (tableau1) et ne figure dans l'interface d'aucun dépôt. En examinant les ressources en arabe de ce dépôt, il s'avère qu'il ne s'agit pas de vrais objets pédagogiques mais plutôt de résumés en arabe de thèses. Dans le même sens, la langue chinoise, malgré qu'elle est la langue la plus parlé sur la planète, elle ne figure que dans les objets pédagogiques et les interfaces de 2 dépôts seulement.

4.3 Nombre des ressources

L'offre des objets pédagogiques dans les archives ouvertes est déterminée en quantifiant ces objets dans ce corpus. Notre analyse montre qu'une bonne partie des dépôts, 18.1% (voir tableau 4), ne permettent pas de quantifier ces objets puisqu'ils ne donnent pas le volume de chaque type de ressource géré par le dépôt. Plus que la moitié des dépôts renferment moins de 100 objets pédagogiques. Il est à noter ici que l'offre des objets pédagogiques dans les archives ouvertes reste très limitée puisque plus que 40% des dépôts renferment moins de 50 objets dont la plupart renferment 1 ou 2 objets seulement. Cette situation pourrait s'expliquer par l'absence de politique claire des institutions incitant les enseignants à déposer leurs objets pédagogiques dans ces archives ouvertes et parla réticence des enseignants à déposer leurs cours dans ces dépôts (Gallezot, 2008). D'ailleurs, les deux dépôts qui renferment 0 objets pédagogiques sont des dépôts qui renferment une rubrique ou une collection Ressources Éducatives Libres qui restent vides.

Par ailleurs, 10 dépôts seulement renferment plus de 5000 objets (tableau 4) dont la plupart sont des dépôts dédiés à des objets pédagogiques à couverture nationale voire internationale.

<i>Nbre ressources</i>	<i>Nbre dépôts</i>	<i>%</i>
001-050	72	40,7
non défini	32	18,1
101-500	24	13,6
051-100	16	9,0
1001-5000	11	6,2
501-1000	10	5,6
>5000	10	5,6
0	2	1,1
<i>Total</i>	177	100%

Tableau 4 : nombre d'objets pédagogiques dans les dépôts

4.4 REL dans les archives ouvertes

L'expression Ressources éducatives libres (REL) a été créée lors du Forum international de l'UNESCO sur les didacticiels libres en 2002. Elle désigne des matériels d'enseignement, d'apprentissage et de recherche sur tout support, numérique ou autre, existant dans le domaine public ou publiés sous une licence ouverte permettant l'accès, l'utilisation, l'adaptation et la redistribution gratuits par d'autres, sans restrictions ou avec des restrictions limitées. Les REL peuvent comprendre des cours/programmes complets, cours de formation, modules, guides de l'étudiant, notes pédagogiques, manuels, articles de recherche, vidéos, outils et instruments d'évaluation et tout autre matériel utile à des fins éducatives.

En 2012, la Déclaration de Paris sur les REL a été officiellement adoptée lors du Congrès mondiale des Ressources éducatives libres (REL) qui s'est tenu au siège de l'UNESCO du 20 au 22 juin à Paris. Elle enjoint les gouvernements du monde entier d'accorder des licences ouvertes aux ressources pédagogiques destinées au grand public et bénéficiant de financements d'État.

Malgré l'importance de ce mouvement pour le domaine de l'enseignement supérieur et les enjeux économiques et éducatifs qu'il représente, il ressort de notre étude que seulement 15 dépôts contiennent ce type d'objets pédagogiques, soit 8.5% (tableau 5).

REL	Total	%
non	161	91,0
oui	15	8,5
(vide)	1	0,6
Total	177	100,0

Tableau 5 : dépôts renfermant des Ressources Educatives Libres (REL)

Nous pensons, qu'il est primordial que les autres universités et institutions (91%) alignent leurs politiques éducatives sur les principes fondamentaux de ce mouvement afin d'améliorer les systèmes d'enseignement supérieur à travers le monde en permettant aux enseignants et aux étudiants un accès libre et légal à un nombre assez important des ressources pédagogiques et aux établissements d'enseignement ainsi qu'aux Etats d'effectuer d'énormes économies.

4.3 Métadonnées

Le terme métadonnées est utilisé pour définir l'ensemble des informations ajoutées à une ressource d'information quelconque pour mieux la qualifier. Il est défini par le NISO (National Information Standards Organization) comme « des informations structurées qui décrivent, expliquent, localisent ou rendent, autrement, plus facile le repérage, l'utilisation ou la gestion des ressources d'information »⁷. En général, elles décrivent les caractéristiques qui concernent le contenu, le contexte et la structure d'une ressource d'information indépendamment de sa forme physique et de son contenu informationnel⁸.

Dans le domaine de l'enseignement, l'apprentissage et la formation, il existe une multitude des schémas des métadonnées dont la plupart sont des profils d'application⁹ issus de la norme Dublin Core et du standard Learning Object Metadata (LOM).

⁷ NISO, Understanding Metadata. www.niso.org/.../press/UnderstandingMetadata.pdf

⁸ Gilland-Swetland, Anne J. Introduction to metadata

http://www.getty.edu/research/conducting_research/standards/intrometadata/.

Voir aussi « What is metadata », ICT Advice for Teachers. URL : <http://www.ictadvice.org.uk>

⁹ Un profil d'application est un ensemble d'éléments de données choisis parmi un ou plusieurs schémas de métadonnées formant ainsi un sous-ensemble adapté aux besoins spécifiques d'une communauté d'utilisateurs voire d'un pays. Le nombre d'éléments sélectionné varie d'un profil d'application à un autre en fonction de l'approche choisie : minimaliste à l'instar de la norme Dublin Core ou structuraliste à l'instar du standard LOM.

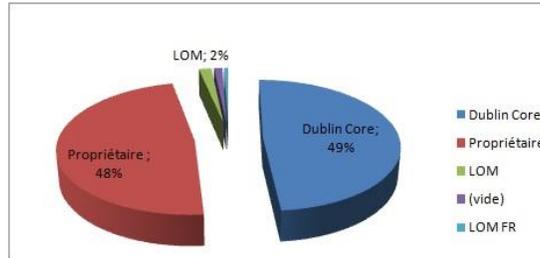


Figure 3 : métadonnées utilisées par les dépôts

Le Dublin Core¹⁰, norme ISO 15836, est un schéma des métadonnées génériques qui a été créé en 1995 par le DCMI (Dublin Core Metadata Initiative). Dans sa version de base 1.1, son schéma élémentaire est composé de 15 éléments de données simples permettant de décrire le contenu, la propriété intellectuelle et l'instanciation particulière des ressources numériques ou non numériques. Tous ces éléments sont répétitifs, optionnels et peuvent apparaître dans n'importe quel ordre¹¹.

Le LOM est un schéma de description et d'indexation de ressources éducatives tant numériques que non numériques. Il a été développé, en 2002, par le groupe de travail IEEE-LTSC-LOM¹² avec pour objectif de faciliter la recherche, l'évaluation, l'acquisition des objets pédagogiques par les enseignants, les apprenants et les outils logiciels.

Il propose un schéma de métadonnées composé de 68 éléments de données optionnels structurés en 9 catégories de données accomplissant chacune une fonction différente : General, Lifecycle, MetaMetadata, Technical, Educational, Rights, Relation, Annotation et Classification.

Ayant le statut d'un standard international, le LOM a servi comme modèle de référence pour le développement de plusieurs profils d'application dans le monde tels que : NORLOM (Norvège), NIME GLAD (Japon), LOMAP (Taiwan), LOMFR (France), CanCore et Normetic (Canada), UK LOM Core (Royaume-Uni), etc.

Bien qu'il ne soit pas réellement adapté à l'indexation des ressources pédagogiques, nous avons constaté, à travers cette recherche, que le Dublin Core est le schéma des métadonnées le plus utilisé puisque 49% des dépôts l'adoptent (figure 3). A notre sens, cette primauté pourrait être expliquée par deux raisons essentielles :

- La simplicité de création des éléments de données du Dublin Core que même des non-spécialistes peuvent élaborer facilement des notices descriptives à leurs objets pédagogiques,

¹⁰Dublin Core Metadata Initiative. <http://dublincore.org/>

¹¹L'ensemble des éléments du Dublin Core est accessible à l'adresse suivante : <http://purl.oclc.org/dc/documents/rec-dces-19990702.htm>

¹² Institute of Electrical and Electronics Engineers, Inc. - Learning Technology Standards Committee - Learning Objects Metadata Working Group

- Un grand nombre des dépôts (48%) utilise le logiciel DSpace (figure 2) qui, par sa conception, génère des métadonnées Dublin Core par défaut.

En revanche, seulement deux dépôts utilisent le LOM et un seul dépôt utilise le profil français LOM-FR pour indexer leurs objets pédagogiques.

Par ailleurs, 48% des dépôts utilisent des schémas propriétaires (figure 3) qui ne correspondent à aucun modèle des métadonnées pédagogiques standardisé et reconnu. Nous pensons que le recours à ce type de schémas, présenterait, sans doute, un obstacle technique assez important interdisant l'interopérabilité et la portabilité des objets pédagogiques et de leurs métadonnées à l'échelle mondiale

4.4 Licences et droits d'auteurs

L'Open Access consiste à un accès libre à l'information scientifique en général et pédagogique dans notre cas, accès qui se veut en ligne, immédiat, gratuit et permanent. Il s'agit d'un accès sans barrières financières juridiques ou techniques. Les archives ouvertes permettent alors une diffusion des différents types de document en libre accès sans pour autant perdre les droits d'auteurs initiales de ces documents. C'est ainsi que dans la définition de l'Open Access de l'initiative de Budapest (voir introduction) on garantit à l'auteur le contrôle de l'intégrité de ses travaux et leurs reconnaissances par la citation correcte : « La seule contrainte sur la reproduction et la distribution, et le seul rôle du copyright dans ce domaine devrait être de garantir aux auteurs un contrôle sur l'intégrité de leurs travaux et le droit à être correctement reconnus et cités ».

<i>Droits</i>	<i>nbre sites</i>
<i>non défini</i>	79
<i>Creative Commons</i>	40
<i>copyright</i>	28
<i>spécifique à chaque ressource</i>	13
<i>Institution</i>	9
<i>Droits d'auteur</i>	3
<i>openaccess</i>	2
<i>(vide)</i>	2
<i>domaine public</i>	1

Tableau 6 : Licences et droits d'auteur

Parallèlement au développement de l'Open Access, les licences Creative Commons (CC) se sont lancés en 2002 et ont connu un succès remarquable de telle façon qu'elles sont souvent associées à l'accès libre. Carol Couture annonce : "Les licences Creative Commons (CC) sont le moyen privilégié pour définir et communiquer clairement les actions permises aux utilisateurs ... Il en existe six principales, correspondant à diverses conditions ou combinaisons de conditions d'utilisation. Toutes comportent l'obligation d'attribution, qui consiste à mentionner l'auteur du document et à fournir un lien vers sa version originale, et permettent à quiconque de le rediffuser et d'en distribuer des copies à des fins non commerciales" (COUTURE, 2018).

Les résultats de notre analyse (tableau 6) montrent que la plupart des dépôts et les auteurs des objets pédagogiques ne déterminent pas les licences d'utilisation de ces ressources 44.6%. Les licences CC sont par contre utilisées par 22.6% des dépôts auxquels s'ajoutent 7.3% des sites dont la licence est spécifique et différente d'une ressource à une autre et dont la plupart sont des licences CC. Il est à noter ici que plusieurs types de licences CC ont été identifiés dans ces dépôts qui méritent une analyse approfondie. Néanmoins, 15.8% des dépôts adoptent le Copyright à toutes les ressources auxquels s'ajoutent 5.1% qui expriment ce copyright comme étant le droit de l'institution. 3 dépôts seulement adoptent les droits d'auteurs du pays d'origine du dépôt (droit d'auteur canadien à titre d'exemple). Le seul dépôt proposant des documents qui sont dans le domaine public est le dépôt polonais Kujawsko-Pomorska Digital Library qui est plutôt une bibliothèque numérique offrant des documents anciens tombés dans le domaine public.

5 Conclusion

Si le mouvement de l'Open Access s'est beaucoup développé du côté des publications scientifiques en faisant pression sur les éditeurs commerciaux et sur les décideurs pour changer la législation en vigueur et adopter des politiques en faveur de l'Open Access, il n'en est pas de même pour les objets pédagogiques. Malgré l'adoption officielle de la Déclaration de Paris sur les REL en 2012 qui appelle les gouvernements du monde entier à placer sous licence libre les ressources pédagogiques destinées au grand public et bénéficiant de financements d'État.

Selon les résultats de cette étude, nous pouvons avancer que l'offre des objets pédagogiques dans les archives ouvertes demeure modeste. En effet, leur nombre est assez limité en comparaison avec d'autres types de ressources tels que les articles des revues et les livres ainsi que le nombre de pays offrant ce type des ressources. De même, les licences ouvertes, principalement les CC, ne sont pas généralisées et très peu des dépôts offrent des Ressources Educatives Libres malgré l'adoption officielle de la Déclaration de Paris sur les REL en 2012 qui appelle les gouvernements du monde entier à placer sous licence libre les ressources pédagogiques destinées au grand public et bénéficiant de financements d'État. D'autre part, la plupart des schémas des métadonnées utilisés pour l'indexation des objets pédagogiques, quand elles existent, ne sont pas appropriés au domaine de l'enseignement et de l'apprentissage.

Reste à vérifier si les universités mettent en place, en plus des archives ouvertes, des plateformes dédiées aux ressources pédagogiques et destinées à l'enseignement classique ou à distance (e-learning).

Références bibliographiques

ALLEN, N., BROWNE, D., FORWARD. M. L., GREEN, C., & TARKOWSKI, A. (2015). Foundation for OER strategy development. [Enligne]. Disponible à: <http://www.oerstrategy.org/home/read-the-doc/>

ANDERSON Terry (2013). Open Access Scholarly Publications as OER. In *The international review of research in open and distance learning*. Vol. 14, n° 2, 81-

95. Open Educational Resources: Opening Access to Knowledge.[Enligne]. Disponible à : <http://www.irrodl.org/index.php/irrodl/article/view/1531/2494> (Consulté le 12/06/2018)
- BOURDA, Yolaine (2002). Des objets pédagogiques aux dossiers pédagogiques (via l'indexation). In Document numérique. Vol. 6, n°1, 2001. [En ligne]. Disponible à : <https://www.cairn.info/revue-document-numerique-2002-1-page-115.htm> (Consulté le 28/08/2018)
- CATTEAU, Olivier (2008). Le cycle de vie de l'objet pédagogique et de ses métadonnées. Thèse de doctorat, Université de Toulouse. [En ligne]. Disponible à : [https://tel.archives-ouvertes.fr/tel-00344701/file/Le cycle de vie de l objet pedagogique et de ses metadonnees These CATTEAU Olivier.pdf](https://tel.archives-ouvertes.fr/tel-00344701/file/Le_cycle_de_vie_de_l_objet_pedagogique_et_de_ses_metadonnees_These_CATTEAU_Olivier.pdf)(Consulté le 18/08/2018)
- COUTURE, Marc (2018). L'accès libre: évolution et enjeux actuels. Dans É. Tremblay et R.Dorcé (dir.), Les Classiques des sciences sociales: 25 ans de partage des savoirs dans la francophonie (p. 145-168). Québec, Canada:Éditions Science et bien commun.Manuscrit récupéré du répertoire R-libre [En ligne]. Disponible à : <http://r-libre.telug.ca/119> (Consulté le 15/08/2018)
- GALLEZOT Gabriel (2008). Archives ouvertes : définition et constats français. In Schedae, 2008, prépublication n°3, (fascicule n°1, p. 23-34. [En ligne]. Accessible à : <https://www.unicaen.fr/puc/images/preprint0032008.pdf> (consulté le 12/08/2018)
- MCGREAL, RORY; et al. (2013). Open Educational Resources: Innovation, Research and Practice. (Commonwealth of Learning (COL);Athabasca University, 2013). 268p. ISBN 978-1-894975-62-9. [Enligne]. Disponible à : <http://oasis.col.org/handle/11599/486>(Consulté le 10/06/2018)
- PIOTKIN Hal (2010). Free to learn. An Open Educational Resources Policy Development Guidebook for Community College Governance Officials. En ligne URL : <https://wiki.creativecommons.org/images/6/67/FreetoLearnGuide.pdf> (Consulté le 16 05/2018)
- PONTIKA Nancy & al. (2015). Fostering Open Science to Research using a Taxonomy and an eLearning Portal. In :iKnow : 15 th International Conference on Knowledge Technologies and Data Driven Business. 21-22 Octobre 2015, Graz, Austria. [En ligne]. Disponible à : http://oro.open.ac.uk/44719/2/kmi_foster_iknow.pdf (Consulté le 15/05/2018)
- SANTOS-HERMOSA Gema, FERRAN-FERRER Nuria, ABADAL Ernest (2017). In *The international review of research in open and distance learning*. Vol. 18, n° 5.[En ligne]. Disponible à : <http://www.irrodl.org/index.php/irrodl/article/view/3063/4300>(Consulté le 12/06/2018)
- STRIJKER, Allard (2004). Reuse of learning objects in context human and technical aspects. Thèse de doctorat de l'université Twente, Enschede,

- Allemagne. [En ligne] Disponible à :
http://doc.utwente.nl/41728/1/thesis_Strijker.pdf (Consulté le 15/08/2018)
- UNESCO (2012). Déclaration de Paris sur les ressources éducatives libres 2012. En ligne URL :
http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/Events/French_Paris_OER_Declaration.pdf (Consulté le 16/05/2018)
- UNESCO (2011). Lignes directrices pour les ressources éducatives libres dans l'enseignement supérieur. En ligne URL :
<http://unesdoc.unesco.org/images/0021/002136/213605f.pdf> (Dernière visite 7/05/2018)

Les MOOC dans le paysage universitaire tunisien : étude du cas de l'Institut supérieur de documentation

MOOCs in the Tunisian university landscape: case study of the Higher Institute of Documentation

Rabii Djebbi

Institut supérieur de documentation, Tunis

El-Khansa Mkada-Zghidi

Institut supérieur de documentation, Tunis

Résumé

Les MOOC représentent une innovation tant au plan de leur mode de diffusion que de leur modèle pédagogique. Leur définition, origine et principes sont d'abord exposés en regard de la diversité de leur typologie et des spécificités.

Puis, sur la base d'une enquête auprès du public universitaire tunisien, nous tentons de fournir des éléments de réponse aux problématiques issues de la gratuité et du libre accès aux MOOC pour savoir dans quelle mesure les apprenants maghrébins, particulièrement les tunisiens, profitent de ces deux avantages.

Mots-clés. MOOC, libre accès, enseignement supérieur.

Abstract

MOOC represent an innovation both in terms of their mode of diffusion and their pedagogical model. First Their definition, origin and principles are exposed according to the diversity of their typology and specificities.

Then, based on a survey of the Tunisian university public, we try to provide elements of answer to the problematic issues from their free and open access to know to what extent the Maghrebian learners, especially Tunisians, take advantage of these two advantages.

Keywords. MOOC, open access, higher Education

1 Introduction

L'origine du MOOC remonte à moins d'une vingtaine d'années lorsque le Massachusetts Institute of Technology crée en 2001 le MIT OpenCourseWare (MIT OCW) et propose des ressources variées sous licence Creative Commons. Il est suivi par plusieurs d'établissements et se trouve rapidement à la tête d'un mouvement qui prendra le nom d'OpenCourseWare Consortium.¹

L'année 2008 est la deuxième date déterminante lorsque l'Université de Stanford offre son programme en informatique et en Sciences de l'ingénieur connu sous Stanford Engineering Everywhere (SEE). Ce programme diffusé sur la plate-forme Coursera deviendra en 2012 leader du mouvement MOOC avec près de dix mille nouveaux inscrits par jour.² En France, le ministère de l'Enseignement supérieur et de la recherche lance en 2013 la Fondation France Université Numérique (FUN) qui fédère les projets des universités et des écoles françaises pour leur donner une visibilité internationale. Au plan maghrébin, dès 2013, des initiatives timides et isolées voient également vu le jour notamment en Tunisie et au Maroc.

A la lumière de l'expansion des MOOC dans la plupart des pays du monde, la problématique de notre étude émane de leurs deux caractéristiques principales, à savoir la gratuité et le libre accès. Celles-ci, contextualisées dans notre environnement maghrébin soulèvent au moins deux questions centrales auxquelles nous essayerons de répondre :

Si les MOOCs marquent « *l'aube de la démocratisation de l'enseignement universitaire* » (Thierry Karsenti, 2013), quel est le facteur qui empêche les pays maghrébins de s'inscrire dans ce mouvement, aussi bien en tant que consommateurs mais aussi en tant que producteurs ? Si, comme l'affirme Mathieu Cisel « *Démocratiser l'enseignement via Internet, ce n'est pas imposer son usage à tous, mais offrir au plus grand nombre l'opportunité de se former. Il appartient ensuite à chacun de se saisir ou non de cette opportunité.* »

³ Quels seraient donc les obstacles qui se dressent face à cette opportunité dans le contexte en question ?

De plus, les populations démunies ne sont-elles pas les plus concernées par les MOOC afin de reculer les limites de l'accès à l'éducation, notamment à l'enseignement supérieur ? Une fois appliqué à notre contexte, pouvons-nous vérifier l'affirmation de Patrick Juignet assurant que « l'accès

1. Voir <http://www.oeconsortium.org/>

2. Voir <https://see.stanford.edu/>

3. Dans <http://blog.educpro.fr/matthieu-cisel/2016/06/13/la-democratisation-de-lenseignement-superieur-par-les-mooc-demagogie-ou-realite/>

potentiellement illimité est en pratique limité par la *langue utilisée, par la censure de quelques États, mais aussi et surtout par la culture et la disponibilité individuelle.*»⁴ Si oui, quelles sont alors les contraintes politiques, culturelles, techniques, financières ou autres qui entravent la réalisation concrète du libre accès dans notre environnement ?

Pour répondre au premier volet de la problématique, nous présentons l'origine des MOOC en les situant dans le cadre empirique de l'apprentissage de type connectiviste (George Siemens et Stephen Downes, 2005), puis nous faisons un état des MOOCs produits par diverses institutions aussi bien à l'étranger qu'au Maghreb arabe.

Quant au second volet, il présente les résultats d'une enquête par questionnaire menée auprès des enseignants et des étudiants du niveau mastère de l'ISD. Les réponses fermées ont fait l'objet d'une approche quantitative et qualitative, quant aux questions ouvertes, elles ont été regroupées par voisinage de sens afin de refléter les justifications ou les propositions du public-cible. L'analyse de ces résultats pourrait servir de base pour mettre au point des stratégies en fonction de perspectives d'avenir bien configurées que la discussion des participants à notre Colloque enrichira, et parmi ces dernières, les possibilités de collaboration institutionnelle entre les pays maghrébins ne sont pas à négliger.

2 Définition du MOOC

MOOC est un acronyme anglais pour Massive Open Online Courses. La traduction française littérale serait «Cours en ligne ouverts et massifs », mais c'est souvent l'acronyme anglais qui est utilisé. MOOC désigne donc un moyen de formation ou d'apprentissage à distance, via internet, en libre accès.

Mais contrairement à l'apprentissage ouvert ou à distance, le MOOC est un phénomène relativement nouveau qui offre une combinaison d'opportunités d'apprentissage en ligne et d'éducation ouverte. Ce domaine d'apprentissage émergent contribue à faciliter la diffusion de la connaissance à travers de multiples réseaux de connexion afin que les expériences éducatives soient réparties entre les structures d'apprentissage, les réseaux et les applications pédagogiques (DalitLévy, 2011).

De plus, de vastes cours en ligne ouverts améliorent les connaissances et l'apprentissage au niveau collégial et deviennent accessibles à des milliers

4 JUIGNET, Patrick. Web et libre accès. In : *Philosophie, science et société* [en ligne]. 2018. Disponible à l'adresse : <https://philosciences.com/philosophie-et-societe/economie-politique-societe/34-web-et-libre-acces>.

de personnes qui peuvent avoir un accès limité aux cours collégiaux traditionnels (Marshall, 2013). Le MOOC représente donc un outil de diffusion de la connaissance en constante évolution. Selon Lévy, avec l'évolution des NTIC, le phénomène des cours en ligne massifs (MOOC) peut même potentiellement avoir « *un impact apocalyptique sur l'apprentissage personnel à travers le monde entier* »⁵.

Ces cours sont proposés par des grandes écoles, des universités, des plateformes de formation, ainsi que par des grandes entreprises. Les MOOC sont des cours souvent gratuits, toutefois, la délivrance d'un certificat après un examen final peut être payante. Signalons aussi que MOOC est parfois employé pour désigner les plates-formes qui dispensent ce type de cours qui sont « *limités dans le temps portant sur un thème spécifique. Ils incluent un ensemble cohérent de ressources pédagogiques, de modalités d'interactions, d'exercices et d'examens conduisant éventuellement à une certification* »⁶. Ils impliquent une équipe pédagogique chargée de l'encadrement des étudiants et du bon déroulement des cours. Il s'agit :

- de cours diffusés exclusivement sur internet
- de cours généralement d'un niveau universitaire et non pas de visioconférences
- de cours entièrement gratuits et en libre d'accès. La délivrance des certificats est payante mais la demande de certification est facultative.

Précisons aussi qu'aucun pré-requis n'est exigé de la part des apprenants et que le contenu de ce type de cours n'est pas libre de droit : il ne peut être ni réutilisé, ni rediffusé sans le consentement de son auteur.

La différence du concept MOOC par rapport à l'e-learning et aux formations à distance se trouve justement au niveau de cette « grande ouverture ». Tout le monde peut se former dans le domaine qui l'inspire. Aujourd'hui des centaines d'universités ont développé leur MOOC, avec des plateformes pour les héberger dans les différents pays. Parmi les plateformes de MOOC les plus connues, citons Coursera, OpenClassrooms, France université numérique (FUN), Udacity et edX.

5. Cité par Federica Minichiello dans *Centre de ressources et d'ingénierie documentaires du CIEP* - <http://www.ciep.fr/revue-internationale-deducation-sevres/lecole-diversite-cultures/phenomene-moocs-massive-open-online-courses>

6. Cisel Matthieu, Bruillard. Éric. Chronique des MOOC. http://sticef.univ-lemans.fr/num/vol2012/13r-cisel/sticef_2012_cisel_13rp.pdf

3 Origine et développement du MOOC

Le MOOC est un phénomène d'apprentissage relativement récent qui est né aux États-Unis. En effet, son origine remonte à moins d'une vingtaine d'années lorsque le Massachusetts Institute of Technology crée en 2001 le MIT OpenCourseWare (MIT OCW) et propose des ressources variées sous licence Creative Commons. Ce prestigieux institut est suivi par des centaines d'établissements et le MIT OCW se trouve rapidement à la tête d'un mouvement qui prendra le nom d'OpenCourseWare Consortium.⁷

L'année 2008 est la deuxième date déterminante lorsque l'Université de Stanford offre son programme en informatique et en Sciences de l'ingénieur connu sous Stanford Engineering Everywhere (SEE). Ce programme est diffusé sur la plate-forme Coursera qui deviendra en 2012 leader du mouvement MOOC avec près de dix mille nouveaux inscrits par jour⁸. Un autre MOOC pionnier a été aussi lancé en 2008, il est nommé « Connectivism and Connective Knowledge », créé par les éducateurs Georges Siemens de l'Université d'Athabasca et Stephen Downes du National Research Council, ce cours a été dispensé en premier lieu à « 25 étudiants de l'Université de Manitoba et 2300 participants y ont pris part en ligne, gratuitement »⁹ et ces derniers étaient libres de participer et d'enrichir le cours avec les outils de leur choix. C'est dans la lignée de cette expérience, de nombreux MOOC vont se mettre en place dans différents établissements.

Le véritable essor des MOOC commence à partir d'un cours d'intelligence artificielle dispensé également sur le site de Stanford en novembre 2011 sous la charge de SebastienThrun, enseignant de robotique. Le cours d'intelligence artificielle « attire environ 160.000 étudiants alors que seulement 10.000 étaient attendus, plus de 15% vont jusqu'au bout »¹⁰. C'est suite à ce succès que Sebastian Thrun fonde la plate-forme Udacity en février 2012.

L'Europe, et en particulier la France, ne sera pas en reste et le Ministère de l'Enseignement Supérieur et de la Recherche lance en octobre 2013, une initiative visant à fédérer les projets des universités et écoles françaises pour leur donner une visibilité internationale. C'est ainsi que la fondation France Université numérique (FUN) est née, elle est l'opérateur de la plateforme du même nom qui offre des MOOC enregistrant ainsi un franc succès avec,

7. Voir <http://www.oeconsortium.org/>

8. Voir <https://see.stanford.edu/>

9. Matthieu Cisel, Éric Bruillard .Chronique des MOOC. <https://hal.archives-ouvertes.fr/hal-01670296/document>

10. idem

à l'heure actuelle, 1 million d'inscriptions à 150 cours disponibles. Ces chiffres illustrent latendance exponentielle des MOOC où en 2013 on estimait les inscriptions à « près de 20 millions d'étudiants de plus de 200 pays. »¹¹

Dans les pays arabes, le phénomène MOOC est plus récent et moins étendu. La Tunisie à lance son premier MOOC FOFLE en 2014, ce MOOC est notamment consacré à l'apprentissage du français¹². La même année, l'Université Virtuelle de Tunis lance sa plateforme TU-MOOC, c'est la première plateforme universitaire tunisienne des MOOC. Un premier MOOC est ouvert au grand public sous le nom « MOOC Radicalisation et Terrorisme »¹³. Il est élaboré dans le cadre d'un partenariat entre l'Université de la Manouba, l'Institut arabe des droits de l'Homme et l'UVT. Le MOOC Radicalisation et Terrorisme vise à analyser ces phénomènes et à capitaliser la réflexion autour des moyens pour y remédier tant à l'échelle individuelle que collective.

En 2017, l'Ecole Supérieure des Communications de Tunis, SUP'COM, spécialisée dans la formation en TIC, a annoncé aussi la création du son premier MOOC/CLOM maghrébin. « Il s'agit d'un cours en ligne ouvert et massif (CLOM/MOOC), porté par Sup'Com, en collaboration avec la Faculté des Sciences et de Technologie de l'Université Hassan 1er de Settat (Maroc) et soutenu par le Bureau Maghreb de l'Agence universitaire de la Francophonie et développé dans le cadre de l'appel à candidatures lancé par l'IFIC. Ce MOOC permet à un public professionnel ou celui à la recherche de l'emploi de se préparer à la certification internationale « LPI Linux Essentials » favorisant ainsi leur insertion sur le marché de travail», précise SUP'COM¹⁴. Toujours en Tunisie, signalons enfin le MOOC sur l'écotourisme proposé en collaboration par l'Université de Jendouba en Tunisie et l'Université Toulouse - Jean Jaurès en France.¹⁵

Quant au Maroc, il lance aussi en 2017 à partir de «campus Casablanca, emlyon businessschool »¹⁶ sa propre plateforme MOOC. Elle a pour objectif de répondre aux enjeux de la formation supérieure sur le continent et de contribuer ainsi à combler l'énorme déficit qu'accuse l'Afrique en matière d'accès à l'enseignement supérieur etaux formations de qualité qui

11. Voir http://MOOC.crifpe.ca/files/RITPU_VOL10_NO2_MOOCvf.pdf

12. Voir https://www.huffpostmaghreb.com/2015/02/26/tunisie-enseignement-MOOC_n_6760588.html

13. Voir <http://www.uvt.rnu.tn/news/554-tu-MOOC-plateforme-des-MOOC-de-1%E2%80%99uvt>

14. Voir <https://www.leconomistemaghrebin.com/2017/02/07/tic-supcom-lance-premier-MOOC-maghrebin/>

15. Voir <https://www.fun-mooc.fr/courses/ujendouba/36001/session01/about>

16. Voir <https://casablanca.em-lyon.com/lancement-de-premiere-plateforme-MOOC-maroc-lafrique-emlyonx/>

permettent de répondre aux enjeux du développement de la région. En 2018, le e-learning center et l'université de Mohamed V de Rabat ont créé un MOOC qui offre des formations en ligne gratuites et certifiantes en comptabilité générale pour une durée de 6 semaines¹⁷.

4 Caractéristiques des MOOC

L'identification des caractéristiques des MOOC sert à dresser une typologie de ce type d'enseignement qui ne se distingue pas seulement par son modèle pédagogique particulier. Dans ce qui suit, nous essayons de passer en revue la typologie issue de ces caractéristiques avant de nous appesantir sur l'aspect pédagogique qui, selon nous, représente l'un de leur caractère innovant le plus déterminant quant au degré d'adhésion des apprenants potentiels.

4.1 Typologie

La première typologie des MOOC a été établie par Daniel (2012) et se limite à la distinction entre le **lexMOOC**, fondé sur le modèle transmissif et axé sur le contenu proposé aux apprenants, et le **cMOOC** centré sur le modèle connectiviste et sur les échanges entre pairs. Elle a été enrichie par Lane (2012) qui a aussi distingué le **tMOOC** centré sur les tâches à réaliser par les apprenants et par Gilliot & al (2013) qui ont suggéré le **iMOOC** établi sur le degré de liberté laissé aux apprenants dans le choix de leurs ressources et de leur mode d'organisation pour réaliser leurs activités d'apprentissage.

Selon un tout autre point de vue (Clark Donald, 2013) est parvenu à distinguer huit types de MOOC selon des critères assez hétérogènes mais tous issus de la pédagogie : *'itis important to define a taxonomy of MOOC not from the institutional but the pedagogic perspective, by their learning functionality, not by their origins'*, il s'agit des :

- **transferMOOC**, ce sont des cours déjà existants et intégrés à une plateforme de MOOC. La renommée de l'enseignant et celle de l'université peuvent leur garantir une large audience auprès des apprenants.
- **madeMOOC**, ici les activités d'apprentissage sont ouvertes, et les échanges et évaluations entre pairs sont encouragés.
- **synchMOOC** qui est en mode synchrone, avec des dates fixes pour les séances des cours et les remises des travaux.

17. Voir <https://www.youtube.com/watch?v=3BPgIKJdvVs>

- **asynchMOOC** en mode asynchrone, sans dates fixes et des périodes plus souples pour remettre les travaux. .
- **adaptiveMOOC** qui utilise des algorithmes adaptatifs pour proposer des expériences d'apprentissage personnalisées.
- **groupMOOC** qui rassemble de petits groupes d'apprenants qui communiquent et réalisent les tâches demandées de manière collaborative.
- **connectivistMOOC** qui se fonde essentiellement sur les interactions entre apprenants qui collectent et partagent leurs ressources et construisent en commun leurs propres parcours d'apprentissage.
- en enfin **miniMOOC** qui est un MOOC très courts de quelques heures à deux ou trois semaines.

Une autre proposition de typologie (Mathieu Cisel, 2016) se fonde sur une approche issue d'une corrélation d'axes pédagogiques tels que :

- **L'objectif du cours** : celui-ci peut se limiter à la diffusion des connaissances via Internet, Cette transmission permet aussi la formation de communautés de pratique à travers les forums, les réseaux sociaux et autres. Enfin, des cours peuvent avoir pour objectif la détection de talents et la co-crédation de savoir par les participants.

- **Le niveau de pré-requis** : même si les MOOC sont ouverts et très souvent gratuits, c'est leur contenu en lui-même qui détermine les pré-requis. Ainsi, les cours introductifs attirent de nombreux participants (par exemple sur Coursera, entre 50 000 à 200 000 personnes), il faut signaler ici que ces types de cours sont souvent pris en charge par des professeurs de renommée internationale. Ceci dit, il existe aussi des MOOC très spécifiques tels les premiers MOOC de Stanford sur l'intelligence artificielle (novembre 2011) qui ont intéressé plus de 150 000 participants.

- **Le type de ressources** : dans les MOOC, les ressources textuelles sont rarement utilisées seules. Les ressources les plus fréquentes sont audiovisuelles, vidéos et présentations powerpoint commentées. Les cours sont généralement divisés en séquences de 5 à 20 minutes. Dans le cas des cMOOC, les ressources sont mises à la disposition des apprenants qui les sélectionnent et les agrègent.

- **Le type d'activités** : il existe une variété d'activités pouvant être soit individuelles, soit collectives. Les quizz sont automatiquement corrigés. Lorsque des activités collectives de type rédactionnel sont proposées, ce

n'est pas l'enseignant qui les évalue mais les pairs, et ce sur la base de critères et d'un barème préétablis.

- **Le degré de contrainte** : cette caractéristique est étroitement liée au MOOC certifiant dans la mesure où, dans ces conditions, l'apprenant est tenu d'effectuer obligatoirement les devoirs demandés par l'équipe pédagogique.

Signalons enfin une dernière typologie d'Eléonore Vrillon (2017) à partir d'une analyse des correspondances multiples (ACM) sur 195 MOOC de la plateforme FUN (France Université numérique) qui a permis de dégager les quatre paires de formes typiques de MOOC suivantes:

- **MOOC spécialiste versus MOOC profane** : le premier type s'adresse aux diplômés du supérieur et est généralement produit par les grandes écoles. Il est généralement certifiant et dure de 6 à 9 semaines. A l'inverse, le MOOC profane ne nécessite aucun prérequis, il dure moins de 6 semaines et n'est généralement pas certifiant.

- **MOOC généraliste d'approfondissement versus MOOC spécialisé introductif** : le premier type nécessite moins de semaines mais il est généralement certifiant. Il est aussi proposé par les grandes écoles. Le second dure plus longtemps, mais il n'est pas certifiant. Il est surtout proposé par des universités.

- **MOOC spécialisé exploratoire versus MOOC spécialisé intermédiaire** : le premier MOOC n'intègre pas la variable du public cible mais celles de l'effort hebdomadaire et du domaine disciplinaire qui correspondent respectivement à 4 heures hebdomadaires et se trouvent surtout dans les domaines du droit, de l'économie et de la gestion. Il est généralement certifiant. A l'inverse, le second MOOC ne nécessite que deux heures hebdomadaires, il se retrouve surtout dans le domaine de la santé et n'est généralement pas certifiant.

- **MOOC spécialisé intensif versus MOOC intermédiaire spécialiste** : le premier type s'adresse à tous les publics sans délivrer de certificat. Il nécessite plus de 4 heures hebdomadaires pendant une durée moyenne de 3 semaines. Les domaines couverts sont souvent l'informatique et les sciences humaines. Le second type est répandu dans le domaine de la santé. Il nécessite entre 2 et 4 heures hebdomadaires pendant 6 à 8 semaines de cours et mène à la délivrance d'un certificat.

La typologie des MOOC montre une grande diversité de leurs caractéristiques qui peut se répercuter sur leur modèle pédagogique.

4.2 Modèle pédagogique des MOOC

Le modèle pédagogique des MOOC se revendique généralement du connectivisme. Ce modèle fait l'objet de nombreuses controverses dues au fait qu'il semble privilégier les technologies éducatives au dépend de la pédagogie. En l'occurrence, Manguenot (2014) remet même en question l'idée selon laquelle « *il y aurait les « bons MOOC » constructivistes (ou connexionnistes, si on adhère aux thèses de Downes et Siemens) et les « mauvais MOOC » transmissifs, voire behavioristes.* »

Pour essayer de clarifier davantage cet aspect, rappelons que, selon certains chercheurs qui seront cités plus bas dans cet article, le connectivisme a succédé aux principaux modèles pédagogiques qui se sont développés sur la base d'une meilleure connaissance des aptitudes psychiques humaines ainsi que d'un ensemble de facteurs socio-économiques dans lesquels la démocratisation de l'accès au savoir n'est pas étrangère ni l'évolution intrinsèque et inhérente de la fonction de l'enseignement.

Historiquement, le modèle transmissif est le premier identifié. La relation établie entre l'élève et le maître se résume en ce que ce dernier, détenteur du savoir, doit remplir la tête de l'élève qui se doit d'accumuler, de mémoriser et de restituer les leçons enseignées. En 1912, Skinner pose les bases de la prise en compte du comportement de l'élève, c'est la naissance du behaviorisme où l'enseignement se fonde désormais sur une évolution progressive des acquis et où les erreurs des élèves et les remédiations des enseignants permettent d'atteindre des objectifs prédéterminés (Watson, 1912).

Le behaviorisme est amélioré par Piaget dans les années 1930 qui, en équilibrant les processus de l'assimilation et de l'accommodation, fonde le modèle du constructivisme où l'élève est en quelque sorte amené à remédier à ses propres erreurs. L'apprentissage par erreur s'effectue sous la houlette de l'enseignant qui tend alors à personnaliser les procédures de l'acquisition des connaissances par son élève. Dans la lignée de ce modèle, la prise de conscience de l'élève de ses propres erreurs favorise le modèle la mise en place des fondements du cognitivisme, essentiellement les stratégies de résolution des problèmes (Freinet, 1947).

Les recherches de Vogotski (1985) suivi par Bruner (1996) sur l'impact des rapports sociaux dans l'acquisition des connaissances, en particulier le rôle de la médiation entre enseignant et apprenant donnent naissance au socioconstructivisme. L'apprenant est désormais placé au centre du processus de l'apprentissage, les interactions avec son tuteur et les échanges

entre pairs le motivent et favorisent la prise en charge de son propre parcours. Ce modèle est largement appliqué dans l'enseignement à distance.

Le socioconstructivisme va perdurer jusqu'en 2008, lorsque Siemens et Downes donnent un cours, gratuit et ouvert à tous, dans lequel ils enseignent le connectivisme tout en l'utilisant comme une méthode d'enseignement. Cette méthode est profondément impactée par les TIC, et se présente actuellement tantôt comme un modèle pédagogique et tantôt comme une théorie de l'apprentissage c'est pourquoi il est aussi appelé apprentissage 2.0. Pour ses fondateurs, *«le savoir étant distribué à travers un réseau de connexions, l'apprentissage, qui est un état constant, consiste en l'habileté de construire et de naviguer les réseaux»*¹⁸. Afin de faciliter l'apprentissage, il s'agit donc essentiellement de maîtriser les outils technologiques, d'entretenir les connexions, de percevoir les liens entre les idées et les domaines, de mettre à jour ses connaissances de manière régulière et d'interpréter correctement l'information sélectionnée. Dans le cas précis des cMOOC, Stephen Downes pense que *«l'impact des cours connectivistes se mesure à l'engagement de chacun et à l'amélioration des connaissances générales de la communauté»*.¹⁹

Le connectivisme émane donc essentiellement du fait que l'apprentissage ne se fait plus de façon individualiste, il serait le processus de connexions au sens large, englobant les connexions neuronales, les connexions entre les hommes, les ordinateurs, mais aussi l'interconnexion entre les différents champs de savoirs. Les cinq principes suivants représentent les fondements de base du connectivisme :

- **La notion sociétale** : elle se concrétise par la présence sociale à travers la création et le maintien des réseaux d'apprenants qui apprennent les uns des autres autant que d'autres membres de l'espace. Les activités des apprenants se reflètent dans leurs contributions sur les outils du web 2.0. (Anderson et Dron, 2011).
- **La diversité des opinions** : l'apprentissage et la connaissance résident dans la diversité des opinions. La valeur de la diversité est une fonction de l'apprentissage: plus les opinions varient, plus le réseau est diversifié et plus cet espace sera efficace (Siemens, 2006).
- **Le contrôle de l'apprentissage** : l'un des éléments dominants dans le connectivisme est l'intégration contrôlée par l'utilisateur sur les contenus au moyen des logiciels et sur la manière de les appréhender (Siemens, 2006).

¹⁸ <http://tecfa.unige.ch/tecfa/maltt/memoire/Laporte2015.pdf>

¹⁹ Idem.

- **L'apprentissage peut résider dans des systèmes non humains** : le savoir peut se trouver tout simplement sous forme de base de données (Siemens, 2004).
- **La prise de décision ainsi que l'obtention des connaissances précises et mises à jour** : c'est ce vers quoi tendent toutes les activités d'apprentissage connectivistes (Siemens, 2006).

Pour résumer les avantages du modèle des MOOC, « savoir-faire » et « savoir-quoi » sont complétés par un « savoir-où », autrement dit, le méta-apprentissage devient aussi important que le l'apprentissage lui-même (Siemens, 2006).

Les principales critiques du connectivisme émanent de Plon Verhagen (2005) et de Bill Kerr (2007). Le premier lui reproche de se préoccuper davantage du contenu de l'apprentissage que de ses objectifs, quant au second, il ne voit pas du tout la nécessité de cette nouvelle théorie par rapport à celles déjà existantes même s'il reconnaît l'influence de la technologie dans un tel environnement d'apprentissage. Pour sa part, Claire David (2015) remet en question le fait que le MOOC soit une innovation pédagogique car il se conforme au modèle behavioriste.

5 Le MOOC vu par le public universitaire tunisien

La connaissance du MOOC par la population universitaire en Tunisie, plus précisément à l'Institut Supérieur de Documentation (ISD) a fait l'objet d'une enquête par questionnaire adressée aux enseignants et aux étudiants des niveaux mastères professionnel et de recherche. Ce questionnaire a été diffusé sur Google forms durant le mois de juin 2018 auprès de 25 enseignants permanents ou vacataires de tous les niveaux et 75 étudiants des mastères professionnels et de recherche, soit une population totale de 100 personnes. Le nombre total des retours s'élève à 52, ce qui donne un taux de réponse de 52 %.

5.1 Profil de la population cible

Les répondants au questionnaire se répartissent à égalité entre hommes et femmes. Cela est assez représentatif de la population tunisienne et en particulier dans l'enseignement supérieur où la proportion des femmes est légèrement supérieure à celle des hommes.

Genre	Nombre	Pourcentage
Femme	26	50 %

Homme	26	50 %
Total	52	100 %

Tableau 1. *Genre des répondants*

Sur l'ensemble des répondants, les étudiants en mastèresont plus légèrement plus nombreux que les enseignants. Cette faible différence peut s'expliquer par le fait que les 75 étudiants à qui a été adressé le questionnaire n'étaient pas très disponibles au mois de juin.

Statut	Nombre	Pourcentage
Etudiant	27	55,1 %
Enseignant	22	44,9 %
Total	52	100 %

Tableau 2. *Statut des répondants*

5.2 Connaissance des MOOC

La majorité des enquêtés a déjà entendu parler des MOOC, soit 82,7% et tous les enseignants figurent dans cette proportion.

Connaissance du MOOC	Nombre	Pourcentage
Oui	43	82,7 %
Non	9	17,3 %
Total	52	100 %

Tableau 3. *Connaissance des MOOC*

Les deux questions ouvertes demandant de préciser la date et les circonstances de cette connaissance des MOOC montrent que cela remonte à 2 ans en arrière (2016) pour plus des deux tiers de notre public(76 %) et que cela s'est produit lors d'un cours pour 53,7% d'entre eux ou par le biais d'un collègue pour 46,3%. Cette proportion non négligeable laisse à penser que, parmi notre population, nombreux auront donc suivi au moins un MOOC.

5.3 Suivi des MOOC

Sur le plan du suivi d'un MOOC, la tendance précédente tend étonnamment à s'inverser, en effet, près de 70% des répondants n'ont jamais suivi un MOOC. Les 16 répondants, soit 30% qui l'ont fait ont surtout cité les 2 MOOC tunisiens (Lutte contre le terrorisme et Développement durable) et rares sont les répondants ayant cité des MOOC produits par des universités étrangères tels que la comptabilité, les finances et l'apprentissage des langues et ce, sans autres précisions. Ces dernières réponses nous interpellent en ce sens que les MOOC se content par centaines, sont disponibles sur Internet et qu'il suffit de les chercher pour

les trouver. C'est ce à quoi la rubrique suivante pourrait apporter des éléments de réponse.

Suivi d'un MOOC	Nombre	Pourcentage
Oui	36	69,2 %
Non	16	30,8 %
Total	52	100 %

Tableau 4. *Suivi d'un MOOC*

5.4 Intérêt pour les MOOC

L'écrasante majorité des répondants souhaite la création des MOOC dans l'enseignement supérieur. Les raisons sont évoquées par la suite.

Souhait d'un MOOC	Nombre	Pourcentage
Oui	49	94,2 %
Non	3	5,8 %
Total	52	100 %

Tableau 5. *Suivi d'un MOOC*

La question ouverte des raisons évoquées pour ce vœu a recueilli 45 retours assez variés, certains ayant cité plus d'une seule raison, qui peuvent être regroupés comme suit en fonction de leur fréquence :

a-la gratuité et la démocratisation de l'accès au savoir : ce sont les deux arguments les plus évoqués par 14 personnes parmi lesquels figure un seul enseignant. Nous trouvons des réponses telles que « *pour démocratiser l'enseignement supérieur* », « *former les gens qui n'ont pas les moyens pour étudier* ». Le fait que cet argument soit le premier cité confirme l'assertion suivante de Mathieu Cisel « *Démocratiser l'enseignement via Internet, ce n'est pas imposer son usage à tous, mais offrir au plus grand nombre l'opportunité de se former. Il appartient ensuite à chacun de se saisir ou non de cette opportunité.* »²⁰. Mais il faut aussi remarquer que sur 52 enquêtés, 14 est nombre très restreint parmi notre public qui a compris que l'offre gratuite des MOOC représente une opportunité à saisir.

b- l'amélioration des connaissances acquises et l'acquisition de nouvelles : ces deux raisons complémentaires ont été citées par 11 personnes, parmi lesquelles figure un enseignant. Elles sont formulées par exemple comme suit : « *acquérir de nouvelles compétences* », « *enrichir les acquis* », « *pouvoir suivre ses études à distance et profiter des connaissances qui peuvent élargir les horizons* », « *le MOOC permet de combler des lacunes* ». Elles s'inscrivent

20 . <http://blog.educpros.fr/matthieu-cisel/2016/06/13/la-democratisation-de-lenseignement-superieur-par-les-mooc-demagogie-ou-realite/>

dans la lignée des MOOC généralistes d'approfondissement versus MOOC spécialisés introductifs cités dans la typologie de Vrillon (2017) et peuvent fournir des pistes lors de la conception de MOOC pour un public spécifique. Deux personnes parmi ces répondants semblent considérer l'apprentissage au moyen des MOOC comme une chance pour progresser dans leur carrière professionnelle « *répondre à des besoins professionnels* », « *pour aboutir à un succès professionnel* ».

c-la facilité de l'accès à l'enseignement à distance : cette raison, et surtout sa position va à l'encontre des réfractaires au connectivisme qui lui reprochent de privilégier les technologies au dépend de la pédagogie. Elle relève du « *contrôle de l'apprentissage* » et elle a été évoquée par 9 personnes parmi lesquelles ne figure aucun enseignant. Rappelons que pour Siemens (2006) l'un des éléments dominants dans le connectivisme est l'intégration contrôlée par l'utilisateur sur les contenus au moyen des logiciels et sur la manière de les appréhender. Pour notre public, le contrôle de l'apprentissage est « *facile et gratuit* », « *on bénéficie de formations par des moyens faciles* », bref « *c'est pratique* ». La manière d'appréhender les contenus au moyen des logiciels est même assimilée à un jeu « *c'est un moyen ludique d'apprentissage* », « *on apprend à jouer avec les techniques d'apprentissage à distance* ».

d- le partage et la collaboration entre apprenants : ce principe des MOOC figure parmi les moins cités, seules 5 personnes ont mis l'accent sur ces aspects, parmi lesquelles figure un seul enseignant. Nous attribuons la faiblesse de ce nombre à l'impact du système éducatif qui n'encourage pas les échanges entre apprenants et qui est dominé par la récompense de l'effort individuel. C'est pourquoi, selon nous, le développement des MOOC pourrait contribuer à induire des changements dans les pratiques et amenuiser les effets de ce lourd héritage. « *se former et interagir avec plusieurs intervenants* », « *travail collaboratif* », « *partager plus de savoirs à distance* ».

e- l'évolution de l'EAD et de l'enseignement supérieur : argument évoqué par 5 enseignants de la manière suivante : « *promouvoir l'enseignement à distance* », « *promouvoir l'enseignement supérieur* ». Le fait que les MOOC soient perçus comme une évolution est en soi un aspect positif. Nous reviendrons sur cette appréciation dans les chapitres suivants.

f- l'autonomie dans l'apprentissage : cette raison a été citée par 3 enquêtés parmi lesquels aucun enseignant. Le terme « *autonomie* » a été cité à deux reprises et celui d' « *auto-formation* » à une seule et dans ce dernier cadre, l'enseignant a été perçu comme n'étant plus l'unique source d'information.

Notons que cette raison est corollaire du travail collaboratif et du partage qui n'a été citée que par cinq personnes. Cela nous interpelle en ce sens que les concepts d'autonomie et de collaboration nous semblent mériter plus d'approfondissement conceptuel afin d'élucider leur contradiction apparente. Nous reviendrons aussi sur cet aspect dans les chapitres suivants.

g- la promotion des MOOC dans les pays arabes et dans le continent africain : 3 enseignants ont été sensibles à la quasi – absence de MOOC dans les espaces géographiques dont ils relèvent : « *encourager l'EAD dans les pays africains* », « *encourager l'enseignement dans les pays arabes* ». Nous ne pouvons qu'adhérer à la nécessité de promouvoir les MOOC dans ces deux environnements et cela est d'ailleurs l'un des objectifs de notre présente contribution.

Pour cerner davantage les opinions des enquêtés nous avons voulu les cibler en fonction des avantages et des inconvénients perçus à travers deux questions à choix fermés qui nous permettront sans doute d'analyser de manière plus fine les divers centres d'intérêt ci-dessus évoqués. Pour clore cette enquête, à ces questions fermées s'ajoutera une rubrique de libre expression au sujet du MOOC afin de recueillir plus d'opinions.

5.5 Evaluation du MOOC

Deux séries de questions fermées à choix multiples ont été soumises au public de l'enquête. La première se rapporte aux avantages du MOOC et la seconde à ses inconvénients. Les propositions des choix ont été élaborées sur la base des principes des MOOC.

5.5.1 Avantages du MOOC

Trois propositions ont été faites au sujet des avantages perçus du MOOC : la gratuité, l'autonomie de l'apprenant et l'accessibilité. Les taux de réponses montrent que l'accessibilité est perçue comme le principal avantage (84,6 %). L'autonomie vient au deuxième rang avec 75% des réponses et la gratuité au dernier avec 69,2 %. Par rapport aux réponses relatives à l'intérêt du MOOC (voir ci-dessus), celles-ci corroborent le fait que l'autonomie de l'apprentissage n'ait pas suscité un grand intérêt parmi les répondants. De plus, toujours dans intérêts manifestés pour les MOOC, elle infirme un tant soit peu le fait que la gratuité jouit d'un rang privilégié, enfin elle rehausse l'attrait représenté par l'accessibilité de ce type de cours

Avantages du MOOC	Nombre	Pourcentage
Gratuité	36	69,2 %
Autonomie	39	75 %
Accessibilité	44	84,6 %

Tableau 6. *Avantages du MOOC*

Les inconvénients du MOOC peuvent donner un autre éclairage sur la perception globale de ce type de cours.

5.5.2 *Inconvénients du MOOC*

Le questionnaire a proposé quatre inconvénients du MOOC, à savoir, la disponibilité de l'apprenant, l'absence de la langue arabe, l'absence de suivi par un enseignant et la formation non diplômante. L'absence de suivi par un enseignant a été l'inconvénient le plus cité, y compris par les enseignants (32, soit 65,3%), cela remet en question le principe de base l'autonomie de l'apprenant qui suit un MOOC et montre qu'il faut soit sensibiliser davantage les futurs adeptes de ce type de cours dans notre pays, soit initier ce que Clark (2013) appelle les synchMOOC qui sont des cours synchrones, avec des dates fixes pour les séances des cours et les remises des travaux.

L'inconvénient suivant relatif à non diplomation des MOOC a été retenu 24 fois (49 %). Or, depuis leur création les MOOC sont des cours de courte durée et la certification n'est pas systématique et est facultative. Cette réponse montre aussi que le public cible pense sans doute à valoriser le MOOC suivi, et nous retrouvons ici la typologie de Clark, notamment les transferMOOC, qui sont des cours intégrés à une plateforme de MOOC et ou la renommée de l'enseignant et celle de l'université sont à même de garantir une large audience. L'avant dernier inconvénient retenu est la disponibilité de l'apprenant, seules 13 personnes (26,5 %) ont retenu cet aspect et cela est un signe encourageant dans la mesure où les MOOC ne durent que quelques semaines en laissant une grande souplesse à l'apprenant pour y accéder à sa guise. Le dernier inconvénient cité est la barrière linguistique signalée par seulement 8 personnes (16,3 %), la faiblesse de ce taux est aussi un facteur encourageant dans la mesure où nous n'avons pas rencontré de MOOC en langue arabe en effectuant la présente recherche. Mais cela ne signifie pas qu'un

pour la création de MOOC en langue arabe ne doit pas voir le jour le plus rapidement possible afin que cette langue trouve sa place parmi les centaines d'offres de MOOC à travers le monde.

Inconvénients du MOOC	Nombre	Pourcentage
Disponibilité de l'apprenant	13	26,5 %
Absence de la langue arabe	8	16,3 %
Absence de suivi par un enseignant	32	65,3 %
Formation non diplômante	24	49 %

Tableau 7. *Inconvénients des MOOC*

5.6 Libre parole sur les MOOC

Pour terminer notre enquête, nous avons donné libre cours aux suggestions du public cible. Sur les 21 retours obtenus, seuls 11 (un enseignant et dix étudiants parfois en émettant plus d'un avis à la fois) sont significatifs dans la mesure où les autres sont demeurés au plan des généralités, par exemple « *encourager les MOOC* », etc. Bien que relativement faible, le contenu de ces retours peut s'avérer révélateur dans la mesure où cette étude est la première du genre et peut servir de repère pour les études à venir.

Les réactions retenues sont examinées à la lumière des cinq principes des MOOC suivants : la notion sociétale ; la diversité des opinions ; le contrôle de l'apprentissage ; l'apprentissage peut résider dans des systèmes non humains ; la prise de décision ainsi que l'obtention des connaissances précises et mises à jour. Ces principes sont classés selon la fréquence de leur évocation par notre population car cela peut aussi être un indicateur à exploiter à l'avenir.

- La notion sociétale : sept commentaires peuvent se ramener à ce principe. Les répondants ont en effet signalé cette notion qui met l'accent sur les interactions entre apprenants « *c'est l'usage des MOOC par les apprenants qui contribue à leur succès* », « *l'interactivité permet une meilleure assimilation du contenu des formations* ».

- La prise de décision ainsi que l'obtention des connaissances précises et mises à jour : sept commentaires peuvent aussi se ramener à ce dernier principe à dominante pédagogique. En fait c'est ce vers quoi tendent toutes les activités d'apprentissage connectivistes, « *les MOOC responsabilisent les apprenants* », « *les MOOC sont une nouvelle pédagogie ouverte sur l'innovation* », « *les MOOC sont des cours qui répondent à un besoin noble, ils sont au service de la société et favorisent l'accès au savoir* ». Citons enfin un dernier commentaire, émis par un enseignant, relatif à l'acquisition des connaissances qui a mis l'accent sur un aspect très important à nos yeux et

que nous devons prendre en considération dans le cas de la création de MOOC à l'échelle maghrébine « *l'apprentissage par les MOOC va toujours dépendre de la qualité des contenus, qui dépendent à leur tour de qui va les soumettre.* » Cette remarque nous rappelle que les universités ont un rôle immense à jouer, et à notre sens, la création d'un consortium et une collaboration entre les compétences de nos pays ne peut que contribuer au rayonnement et au succès des MOOC qui répondent mieux aux attentes des futurs apprenants de notre région.

- **La diversité des opinions** : selon Siemens, cette notion a un effet positif sur l'efficacité des MOOC, or sur ce plan, il convient de souligner la dimension culturelle, et plus spécifiquement maghrébine et africaine qui pourrait aussi être un facteur déterminant en matière d'interactions. Nous relevons en effet les deux commentaires suivants : « *il faut créer un MOOC entre les pays arabes* », « *promouvoir la culture de la création des MOOC dans les pays du Maghreb arabe* ». Nous voyons dans ces opinions que la notion sociétale se concrétise aussi bien par la présence sociale à travers la création et le maintien des réseaux d'apprenants autant dans un espace virtuel que dans un espace géographique communs.

- **l'apprentissage peut résider dans des systèmes non humains** qui est un autre principe du modèle connectiviste et nous le retrouvons à travers les deux commentaires suivants « *maîtriser les e-compétences* », « *la technique ne suffit pas à elle seule, il faut une culture à la technique* ».

- **Le contrôle de l'apprentissage** : un seul commentaire peut se ramener à ce principe à travers l'idée que « *les MOOC aident à découvrir comment apprendre dans un environnement numérique* ». Rappelons que l'un des facteurs dominants dans le connectivisme est en effet l'intégration contrôlée par l'utilisateur sur les contenus au moyen des logiciels et sur la manière de les appréhender (Siemens, 2006).

6 Conclusion

La synthèse des commentaires libres sur les MOOC et leurs recoupements avec les diverses réponses au questionnaire montrent que, si ce type de cours est souhaité par la majorité du public universitaire, ses contraintes et ses principes ne sont pas tous assez bien assimilés en regard du nombre très limité des personnes qui les ont suivis. En effet, certains principes sur lesquels repose ce type de cours ouvert et gratuit n'ont pas été assez soulignés telle la nécessité des interactions, et même des évaluations entre pairs.

Pour en revenir aux deux volets de notre problématique, consistant à essayer d'abord de savoir pourquoi les pays maghrébins tardent à s'inscrire dans le mouvement des MOOC qui marquent pourtant «*l'aube de la démocratisation de l'enseignement universitaire*» (Karsenti, 2013) nous estimons qu'il ne suffit pas, comme le prône Mathieu Cisel, «*qu'il appartient à chacun de se saisir ou non de cette opportunité* » mais que les établissements universitaires ont une lourde responsabilité dans la sensibilisation, la conception et l'offre de MOOC proposant des cours de qualité pour des apprenants potentiels.

En cela nous rejoignons en partie l'affirmation de Patrick Juignetcitée dans le second volet de notre problématique et qui consiste à dire que «*l'accès potentiellement illimité [...] est en pratique limité aussi et surtout par la culture [...]*».

Seule la création des MOOC dans un environnement social bien déterminé et l'incitation à leur usage par un large public pourraient mettre en valeur l'apport immense représenté aussi bien par les technologies éducatives que par des contenus adaptés aux besoins réels et spécifiques des apprenants des pays maghrébins.

Bibliographie

Ouvrage

POMELO, Jean-Charles, et al (2014). *Les MOOC : conception, usages et modèles pédagogiques*. Paris : Dunod, 2014. 145 p.

Document disponible sur le web

CISEL, Matthieu (2016). Utilisations des MOOC : éléments de typologie. Education. Université Paris-Saclay, 2015. [En ligne]. Disponible à : <https://tel.archives-ouvertes.fr/tel-01444125/document> . (Page consultée le 2 juin 2018).

CLARK, Donald (2013). MOOC: taxonomy of 8 types of MOOC, [Enligne]. Disponible à : <https://donaldclarkplanb.blogspot.com/2013/04/moocs-taxonomy-of-8-types-of-mooc.html>(Page consultée le 6 juin 2018).

DOWNES, Stephen (2005). An introduction to connective knowledge.[En ligne]. Disponible à : <http://www.downes.ca/cgi-bin/page.cgi?post=33034> (Page consultée le 2 mai 2018).

DOWNES, Stephen (2006). Learning networks and connective knowledge. Retrieved November 12, 2006, [En ligne]. Disponible à : <http://it.coe.uga.edu/itforum/paper92/paper92.html> (Page consultée le 2 mai 2018).

- FERNANDEZ, Sandra Meza (2013). Enseigner et apprendre en ligne : vers un modèle de la navigation sur des sites Web de formation universitaire. *Education. Université de Strasbourg*. [En ligne]. Disponible à - <https://tel.archives-ouvertes.fr/tel-00974481/document>. (Page consultée le 20 avril 2018).
- KARSENTI, Thierry (2013). MOOC : révolution ou simple effet de mode ? - *International Journal of Technologies in Higher Education*. [En ligne]. Disponible à : http://ritpu.ca:81/img/pdf/RITPU_v10_n02_06.pdf(Page consultée le 15 avril 2018).
- LEVY, Dalit (2011). Lessons Learned from Participating in a Connectivist Massive Online Open Course[Enligne]. Disponible à : http://chais.openu.ac.il/chais2011/download/f-levy-d-94_eng.pdf(Page consultée 5 mai 2018).
- SIEMENS, George (2004). Connectivism: A Learning Theory for the Digital Age December 12, 2004 [Enligne]. Disponible à : <http://devrijerumte.org/content/artikelen/Connectivism.pdf>. (Page consultée 3 avril 2018).
- SIEMENS, George (2006). Connectivism: Learning Theory or Pastime of the Self-Amused?[En ligne]. Disponible à : https://scholar.google.com/scholar?hl=fr&as_sdt=0%2C5&as_vis=1&q=sie mens+2006+connectivism&btnG= (Page consultée 3 avril 2018).

ANNEXES

Déclaration finale

3rd International Colloquium on Open Access



"Accès libre à la science: Fondements, enjeux et dynamiques"
28- 30 novembre 2018

Le 3e colloque international sur "Le libre accès à la science : Fondements, enjeux et dynamiques" s'est tenu à Rabat - Royaume du Maroc du 28 au 30 novembre 2018.

Organisé par l'Ecole des Sciences de l'Information (Maroc), The Maghrebien Information Processing Society (MIPS-Maroc) et l'Institut Supérieur de Documentation (Université de la Manouba-Tunisie), icoa2018 a bénéficié de la participation effective du Haut Commissariat au Plan, via le Centre National de Documentation (Maroc).

Outre le soutien de l'Université Mohamed V, à travers le Centre National de la Recherche Scientifique et Technique (CNRST) qui a hébergé le colloque, celui-ci a profité du partenariat de la Fédération des Universités du Monde Islamique (FUMI) affiliée à l'ISESCO, de l'Université Hassan II de Casablanca, et de la Bibliothèque de l'Université Laval du Québec.

Huit (08) pays ont participé à icoa2018: l'Allemagne, l'Algérie, le Canada, les États-Unis, la France, le Maroc, le Sultanat d'Oman et la Tunisie. Pour des raisons financières, deux autres pays n'ont pas pu être présents: la Côte d'Ivoire et le Sénégal. Toutefois, les communications de leurs chercheurs, retenues par le comité scientifique, sont publiées dans le E-Book.

Plus de 50 conférences et communications y ont été présentées. Ceci en plus des travaux de trois ateliers relatifs l'un à la " Stratégie de publication en Open Access d'un chercheur " ; le deuxième à la "Création de périodiques et critères d'indexation dans le Répertoire des Revues en Open Access (DOAJ)"; et le troisième aux "Spécificités des problématiques et des processus de l'Open Access dans les pays **des Suds**".

Au cours des 11 sessions du colloque (5 en arabe et 6 en français et anglais) et des ateliers, les discussions ont porté sur les questions et problématiques liées aux différentes thématiques proposées par le programme. Les participants ont ainsi mis en évidence les fondements épistémologiques et les évolutions de l'accès libre et ouvert à la science, ainsi que les obstacles qui empêchent leur mise en œuvre, notamment dans les pays du Sud. Les débats ont porté également sur les règles, approches et méthodologies susceptibles d'y conduire. Une attention particulière a été accordée aux moyens de le renforcer auprès de ses décideurs et acteurs ainsi qu'aux défis technologiques, économiques, juridiques et pédagogiques.

Les travaux du colloque se sont concrétisés par le lancement de Nidaear-Ribat (l'appel de Rabat) et la publication des actes d'icoa2018 ».

Le lancement de Nidaear-Ribat (l'appel de Rabat) pour l'accès ouvert à la science n'est au fait que le prolongement de Nidaear-Riyad, le premier appel lancé en faveur de l'accès libre aux publications en 2006.

Nidaear-Ribat réfère dans son préambule à la tradition arabo-islamique qui fait de la connaissance et de son partage un devoir non seulement pour les savants et les chercheurs, mais également pour le grand public. Il exhorte ainsi les universités, les centres académiques et de recherche et les bibliothèques universitaires des pays arabes et islamiques à s'inscrire dans le mouvement du libre accès à la science et les incite à :

- la généralisation de l'accès aux publications par la création et/ou le renforcement d'archives numériques et de périodiques ouverts et par l'incitation à leur utilisation ;
- la structuration et l'ouverture sur les données de la recherche scientifique, que ce soit celles produites par les chercheurs locaux ou par les élites intellectuelles immigrées, et l'encouragement de leur accès et utilisation. L'objectif étant d'impliquer l'ensemble du potentiel scientifique de ces pays dans le processus du développement durable, l'éradication de la fracture numérique et l'intégration dans la société de la connaissance ;
- l'inscription du mouvement de l'Open Access dans une dynamique globale et durable par la couverture des pays du Maghreb, du monde arabe et islamique et par la réalisation de son suivi et accomplissement.

La publication des actes d'icoa2018 en deux tomes, sous format E-Book. Le tome 1 consacré aux articles en langues française et anglaise (~500 p.) et, le tome 2 à ceux en langue arabe (+ de 350 p.). Cf. Actes d'icoa'18 in : <https://cnd.hcp.ma/3eme-Colloque-International-sur-le-Libre-Access-ICOA-18a1030.html>)

Dans le même ordre d'idées, le colloque n'a pas hésité à souligner la nécessité pour le monde arabe et islamique :

- d'élaborer des politiques nationales pour un accès ouvert aux publications scientifiques et aux résultats de la recherche locale et celle réalisée par ses cerveaux à l'étranger ;
- de prévoir et/ou de renforcer les textes législatifs permettant de préserver le droit moral des chercheurs et des innovateurs, notamment la paternité de leur œuvre et le respect de leur intégrité, et d'inscrire leurs actions, autant que possible, dans le cadre des législations et licences existantes, notamment les licences Creative Commons ;
- d'inciter les universités et les instituts de recherche scientifique des pays arabes et islamiques à soutenir et / ou à :
 - créer des revues libres et gratuites dans divers domaines, en donnant l'importance qui se doit aux humanités numériques et en veillant à les doter de comités d'évaluation et de mesure de qualité permettant de reconnaître les travaux qui s'y publient dans la promotion professionnelle et scientifique ;
 - adopter la voie verte de l'accès libre par la construction d'archives numériques institutionnelles pour héberger les thèses de doctorat, les rapports de recherche, les ressources pédagogiques...
- d'activer le rôle des bibliothèques universitaires et des centres d'information et de documentation des pays arabes et islamiques, dans le renfort de l'accès des chercheurs et des lecteurs aux sources de la connaissance et aux résultats de la recherche scientifique, librement et gratuitement ;
- d'exhorter la coopération inter-organisations arabes, islamiques et internationales y compris avec la fédération des Universités du Monde Islamique (FUMI), l'International Federation of Library Associations (IFLA), la Fédération arabe des bibliothèques et de l'information (AFLI)...

- de profiter des expériences des bibliothèques universitaires et des centres de recherche académiques pionniers dans l'utilisation des nouveaux paradigmes de la communication scientifique ;
- de suivre les traces du "Directory of Open Access Journals (DOAJ)", "The Scholarly Publishing and Academic Resources Coalition (SPARC)", et du "Directory of Free Arab Journals (DFAJ)"... et plaider pour un accès libre et ouvert à la science aux niveaux local, régional et international.

**L'appel de Rabat en faveur de l'Open Access
Nidae ar-Ribat**



**De l'accès libre aux publications à l'accès ouvert aux données
de la recherche scientifique**

Prologue

En 2006, Nidaear-Ryadh (l'appel de Ryadh), en faveur du Libre Accès, a été lancé à partir de la capitale Saoudite ar-Ryadh, à l'occasion du 2^e colloque scientifique Golfe-Magreb. L'idée était d'appeler à la création d'une bibliothèque numérique à l'échelle régionale permettant d'intégrer le nouveau paradigme de communication « l'Open Access » dans les pratiques d'accès à l'information scientifique et technique.

Douze années plutard, munis d'initiatives renouvelées et d'expériences documentées sur le plan technique, légal et économique, nous pouvons affirmer que les pays arabes se trouvent encore au tout début de l'effort mondial. Un chemin a été parcouru certes par ces pays, mais les actions restent individuelles, et sont semées d'embuches.

Aussi, les chercheurs de la région continuent-ils à réaliser leurs travaux, évaluer ceux de leurs pairs, siéger dans des jurys, diriger des publications et publier dans des revues à l'ancienne ! Le Libre Accès s'il leur a fait miroiter la possibilité d'une plus grande ouverture sur les publications et sur les résultats et données de la recherche scientifique, continue à être équivoque à leurs yeux.

La non-prise en considération des travaux publiés en libres d'accès pour la promotion professionnelle n'encourage pas les chercheurs à y recourir. Aussi, le non-respect par certaines revues libres des processus d'écriture et d'évaluation, affecte-t-il leur crédibilité et va à l'encontre des intérêts des chercheurs et de leurs bailleurs de fonds. Il en est de même des nombreux échecs essayés dans la mise en œuvre des archives libres.

Un nouvel appel en faveur du Libre Accès à l'échelle régionale s'avère être crucial. Il émane d'une autre capitale cette fois-ci, Rabat, à l'occasion du « 3rd International Colloquium on Open Access » (icoa2018).

L'Appel de Rabat

À Rabat, les participants au 3rd International Colloquium on Open Access (icia2018) ont examiné les projets et politiques du libre accès des différents pays, notamment ceux arabes, et ont exploré comment chercheurs, professionnels de l'information et de l'édition, bailleurs de fonds et décideurs des politiques de l'enseignement et de la recherche pourraient utiliser leurs ressources en commun de manière productive, comment s'assurer la transition vers l'accès libre et ouvert, et comment rendre la publication et les données de la recherche techniquement et économiquement aisées.

Nidaear-Ribat (l'appel de Rabat) réaffirme ainsi l'énoncé de principes dont les origines se prolongent dans la tradition de la civilisation arabomusulmane et dans ses pratiques qui considèrent la science et sa libre circulation un devoir de tout un chacun, et, pour le financement de laquelle al-Wakf jouait un rôle fondateur¹.

Aussi, l'appel de Rabat, exhorte-t-il les pays arabes à s'inscrire dans le mouvement du libre accès aux publications et aux données de la recherche. Il s'agit pour eux de mettre à jour, corriger et capitaliser les initiatives lancées, ici et là, et d'explorer les possibilités de faire converger l'ensemble des projets existant vers un projet fédérateur, tel que SciELO (Scientific Electronic Library Online) latino-américain ; un projet en mesure de tenir compte des avancées opérées à l'international et des spécificités locales.

¹ Al-waḳk, réfère, dans le droit islamique, à une donation faite à perpétuité par un particulier à une œuvre d'utilité publique. De nombreuses donations ont ainsi été faites en faveur des écoles (Medersas), des hôpitaux (Moristane), des bibliothèques (Maktabat–Khizanah) et continuent de l'être. Aussi, de très nombreux manuscrits et œuvres savantes portent des inscriptions référant à leur Waḳk par leurs auteurs (Ibnkhaldoun, Ibn Nafis, Ibn Sīnā (Avicenne). Les Waḳfs permirent aussi de financer, entre autres, les travaux scientifiques et médicaux d'Ibn-Rushd (Averroès), d'Al-Qazi, d'Ibn Sīnā, d'Ali Ibn Isa... et de leur traduction et diffusion.

L'objectif étant de forger un espace numérique dans lequel chercheurs, et données et résultats de la recherche circuleraient librement et qui serait en mesure de renforcer leur potentiel d'innovation, d'assurer le développement d'une science endogène et de contrecarrer l'unilatéralité des savoirs ; la trajectoire de la science devant désormais être croisée (Nord Sud/Sud-Nord).

A cette fin, les participants à icoa2018 incitent le monde arabe à :

- 1. Expliciter et mettre en avant les fondements épistémologiques, les processus et les pratiques du libre accès aux publications et de l'accès ouvert aux données de la science :** Les concepts de base et les évolutions opérées aux trois niveaux étant mal maîtrisées et mal reflétées dans les pratiques des acteurs de la science qu'ils soient auteurs, innovateurs, éditeurs, professionnels de l'information ou décideurs des politiques de la formation et de la recherche.

Les paradoxes de la liberté d'accès aux publications et aux données de la recherche sont en effet mal élucidés. L'un d'entre eux, concerne les aspects juridiques ; produire et transmettre un savoir implique des droits d'auteur, alors qu'en bénéficier requiert liberté et "gratuité" ! Le deuxième, relève de son économie ; les modes de financement du "Libre" ne faisant pas l'unanimité de ses décideurs, éditeurs, chercheurs et utilisateurs.

L'action militante en faveur du Libre requiert l'identification et la maîtrise de ses concepts et fondements, ainsi que des droits et devoirs qui en découlent. Elle requiert également la maîtrise de ses processus et pratiques et des financements qui le conditionnent ;

- 2. Créer la bibliothèque numérique arabe :** qui réunirait l'ensemble des bibliothèques numériques (et celles à numériser) éparpillées à travers les espaces géographiques du monde arabe, y compris la bibliothèque numérique Maghrébine, dont les prémisses se profilent à l'horizon et qui en serait une partie constituante ;
- 3. Structurer et s'ouvrir sur les données de la recherche scientifique :** il s'agit à la fois des données produites par les chercheurs locaux et de celles produites par les élites intellectuelles immigrées, et

l'encouragement de leur accès et utilisation. L'objectif étant d'impliquer l'ensemble du potentiel scientifique de ces pays dans le processus du développement durable, l'éradication de la fracture numérique et l'intégration dans la société de la connaissance.

Cette structuration requiert :

- **de consolider les revues en Open Access** existantes et d'en créer de nouvelles afin de couvrir l'ensemble des disciplines scientifiques et des humanités ; des revues soumises à la rigueur des règles de la méthode scientifique ;
- ***de participer au développement d'outils open-source collaboratifs pour accompagner les revues en Open Access ;***
- **d'accélérer l'effort régional visant à créer et à faire évoluer les archives ouvertes** et inciter les chercheurs à en déposer des copies de leurs articles, rapports et données de la recherche ;
- **de constituer un consortium du monde Arabe d'acteurs** pour établir un cadre opérationnel de financement de l'innovation en matière de publication en Open Access. L'idée étant de favoriser les modèles économiques de publication qui n'impliquent le paiement ni par les auteurs pour publier, ni par les lecteurs pour accéder aux textes.

Revue libre et archives ouvertes requièrent le respect des droits d'auteurs tels qu'ils émanent des licences, notamment « Creative Commons », et, la garantie de leur accessibilité ;

4. **Inscrire le mouvement de l'Open Access dans une dynamique globale et durable** : Et ce, par la couverture des pays du Maghreb, du monde arabe voir même ceux du monde islamique et par l'exécution de son suivi et accomplissement.

**The Rabat call for Open Access
Nidae-ar-Ribat**



**From open access to publications to open access to data of scientific
research**

Prologue

In 2006, Nidaear-Riyadh (the Riyadh call), in favour of Open Access, was launched from the Saudi capital Riyadh, on the occasion of the 2nd Gulf-Maghreb scientific symposium. The idea was to call for the creation of a regional digital library to integrate the new communication paradigm "Open Access" in the practices of access to scientific and technical information.

Twelve years later, with renewed initiatives and documented technical, legal and economic experiences, we can say that the Arab countries are still at the very beginning of the global effort. A great deal of progress has been made by these countries, but the actions remain individual and are fraught with difficulties.

As a result, researchers in the region continue to carry out their work, evaluate those of their peers, sit on committees, direct publications and publish in old-fashioned journals! Open Access, if it has given them the possibility of a greater openness to publications and the results and data of scientific research, continues to be equivocal in their eyes.

Failure to consider open-access work for professional promotion does not encourage researchers to use it. Also, the non-respect by some free journals of writing and evaluation processes, affects their credibility and goes against the interests of researchers and their funders. The same is true of the many failures in the implementation of free archives.

A new call for Regional Open Access has proved to be crucial. It comes from another capital this time, Rabat, on the occasion of the 3rd International Colloquium on Open Access (icoa2018).

The Rabat Call

In Rabat, the participants of the 3rd International Colloquium on Open Access (ICIA2018) examined open access projects and policies of various countries, particularly the Arab ones, and explored how researchers, information and publishing professionals, donors, funders and education and research policy makers could use their resources productively, ensure the transition to free and open access, and make publication and research data technically and economically undemanding.

Nidaear-Ribat (the Rabat call) thus reaffirms the statement of principles whose origins extend into the tradition of the Arab-Muslim civilization and its practices which consider science and its free circulation a duty of everyone, and which has primarily been financed and sustained by the Islamic Al-Waqf¹.

Moreover, Rabat's call urges Arab countries to join the movement of open access to publications and research data. The aim for them is to update, correct and capitalize initiatives launched here and there, and to explore the possibilities of converging all existing projects into a unifying project, such as Latin American SciELO (Scientific Electronic Library Online); a project able to take into account both local specificities and international developments.

The aim is to establish a digital space in which researchers, research data and research results would flow freely and would be able to reinforce potential for innovation, to ensure the development of an endogenous science and to thwart unilaterality of knowledge; the trajectory of science must now cross countries (North South / South-North).

To this end, the participants in icoa2018 encourage the Arab world to work for:

¹Al-Waqf refers, in Islamic law, to a gift made in perpetuity by an individual to a public utility work. From the past till now, many donations have been made to schools (Medersas), hospitals (Moristane), libraries (Maktabat-Khizanah). Also, numerous manuscripts and scholarly works bear inscriptions referring to their Waqf by their authors (Ibnkaldoun, Ibn Nafis, Ibn Sina (Avicenna). The Waqfs also made it possible to finance, among others, the scientific and medical works of Ibn- Rushd (Averroes), Al-Qazi, Ibn Sina, Ali Ibn Isa ... and their translation and diffusion.

1. Explaining and highlighting the epistemological foundations, processes and practices of open access to publications and open access to science data: The basic concepts and developments at the three levels are poorly mastered and poorly reflected in the practices of the actors of science; i.e. authors, innovators, publishers, professionals of information or decision-makers of education and research policies.

The paradoxes of freedom of access to publications and research data are indeed unclear. One of them concerns the legal aspects; producing and transmitting knowledge implies copyright, while enjoying it requires freedom and "free"! The second is its economy; the mode of the "free" financing is not unanimously agreed on by decision-makers, publishers, researchers and users.

Militant action for free requires the identification and mastery of its concepts and foundations, as well as the subsequent rights and duties. It also requires the control of its processes and practices and the financing that conditions it;

2. Creating the Arabic digital library: which would bring together all digital libraries (and those to be digitized) scattered across the geographical areas of the Arab world, including the Maghreb digital library, whose premises are on the horizon.
3. Structuring and opening up on the data of scientific research: Of concern here is the data produced by local researchers and by immigrant intellectual elites, and the encouragement of their access and use. The objective is to involve all the scientific potential within the Arab countries in the process of sustainable development, the eradication of the digital divide and the integration into the knowledge society.

This structuring requires:

- Consolidating existing Open Access journals and creating new ones to cover all scientific disciplines and humanities. All journals should comply with the rigour of scientific methods;

- Participating in the development of collaborative open-source tools to support Open Access journals;
- Accelerating the regional effort to create and develop open archives and encouraging researchers to deposit copies of their articles, reports and research data;
- Establishing an Arab world consortium of actors to establish an operational framework for funding innovation in Open Access publishing. The idea is to favour the economic publishing models which involve no payment by the authors to publish, nor by the readers to access the texts.

Free journals and open archives require the respect of copyrights in conformity with the licenses, including "Creative Commons", and the guarantee of their accessibility.

4. Integrating the movement of Open Access within global and sustainable dynamics through its expansion, execution of its follow-up and accomplishment in the countries of the Maghreb, the Arab world and even those of the Islamic world.

Sommaire des communications en langue arabe

(Tome 1)

المجلد رقم I : عناوين المقالات

المحتوى الرقمي العربي: من الملكية الفردية إلى المشترك الإبداعي داخل الفضاء الأكاديمي
الجزائري
فؤاد بن ضيف الله

الوصول الحر من خلال العموميات الخلاقة: دراسة تطبيقية على مواقع مؤسسات ومرافق
المعلومات العربية
بوعناققة سعاد، سدوس روميصة

تقييم استخدام الباحثين لتراخيص المشاع الإبداعي في عملية النشر العلمي: دراسة ميدانية
بالمعهد الوطني الجزائري للبحث الزراعي INRAA
نضيرة عاشوري، كريمة خلف، فطيمة الشيخ

المعرفة العلمية في الفضاء الافتراضي عبر المواقع الإلكترونية لمؤسسات التعليم العالي: من
ضروريات الإتاحة إلى سبل الإغناء. دراسة وصفية تحليلية لإتاحة الدوريات العلمية عبر
المواقع الإلكترونية للجامعات الجزائرية
عبد الرزاق غزال، وفاء بورحلي

نحو منصة جزائرية مفتوحة لنشر العلوم: دراسة تقييمية لمنصة الدوريات العلمية
منير الحمزة، نور الدين صدار

خدمات المعلومات الإلكترونية بمستودع الأرشيف المفتوح للتظاهرات العملية لجامعة قاصدي
مرباح ورقلة الجزائر
أولادحسيني يوسف، كمال بوكرزازة

المنصة الجزائرية للمجلات العلمية ASJP بين تطوير البحث العلمي وتجسيد التوجه نحو النفاذ
الحر للمعلومات
عبد المالك بن السبتي، رميصة سدوس

خصائص النتائج الفكرية المنشور في دليل دوريات الوصول الحر في مجال علم المكتبات
والمعلومات: دراسة بيبليومترية
سيف الجابري، أصيلة الهنائي

تقييم المستودعات الرقمية بالجامعات الجزائرية باستخدام تقنيات الويبمتركس
شاشة فارس، عبد الغني بن دريدي

أثر الإتاحة الحرة للإنتاج العلمي للجامعات الجزائرية في تصنيف ويب متركس
بوزيد هجيرة

طبيعة العوائق التي تواجه حركة النفاذ المفتوح بالمكتبات الجامعية: دراسة ميدانية بالمكتبات
الجامعية للشرق الجزائري
شورار عفيفة

مقاييس تقييم موثوقية المستودعات الرقمية وفقا لمعايير ISO/DIS 16363 الفرعية: دراسة
تحليلية للمستودعات المؤسساتية بالجامعات البريطانية
محمد الصالح نابتي، عز الدين بودربان، نضيرة عاشوري

إتاحة المنشورات العلمية للباحثين عبر الشبكات الاجتماعية الأكاديمية وتقييمها من طرف النظراء:
دراسة ميدانية مع أساتذة علم المكتبات بجامعات قسنطينة2، وهران1،
الجزائر2
مصبيح وردة، بهجة بومعرافي، سارة بن السبتي

التدريب على الوصول الحر للمعلومات بجامعة قاصدي مرياح ورقلة الجزائر: نحو منظمات متعلمة
صبرينةمقناني، محبوب فيروز

تصنيف ويب متركس للجامعات: هل يشجع فعلا النفاذ المفتوح للبحث العلمي؟
فاطمة شتاب، مهني أقبال

الوصول الحر في البلدان العربية والهيمنة المعرفية الغربية: دراسة فينومينولوجية
جميلة جابر