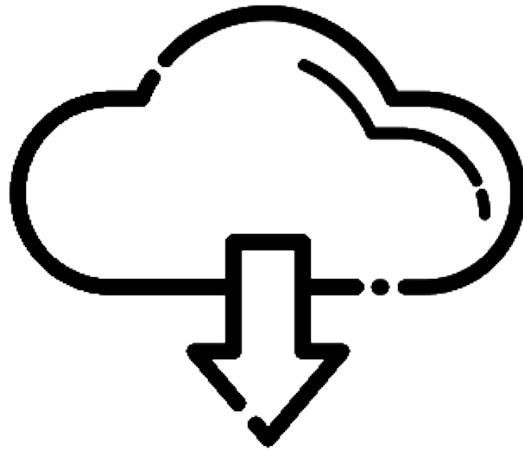


Les clés de la Science Ouverte

OPEN ACCESS, Loi pour une République numérique, Gold Open Access, DONNÉES DE LA RECHERCHE, Dark Data des SHS, Accompagnement des chercheurs, URFIST, Débat autour du MOOC,...

Une revue proposée par le Master 2
Management et Valorisation de
l'Information Numérique (MAVINUM)
Promotion 2017-2018





LOADING DIDAK'TIC



SCIENCE OUVERTE, FAIRE SAUTER LES VERROUS !

Un *Plan national pour la Science Ouverte* a été dévoilé le 4 juillet dernier par la ministre de l'enseignement supérieur, de la recherche et de l'innovation, Frédérique Vidal. Ce plan reprend les préconisations formulées par l'Union européenne dans l'*Amsterdam Call for Action on Open Science* en 2016. **Mais qu'est-ce que la Science Ouverte ?** La Science Ouverte est un mouvement international qui défend l'idée que la science est un bien commun et qui encourage donc « *une approche transversale de l'accès au travail scientifique, des visées du partage des résultats de la science mais aussi une nouvelle façon de FAIRE de la science en ouvrant les processus, les codes, les méthodes* » (DIST-CNRS, *Livre blanc – Une Science ouverte dans une République numérique*, 2016). L'idéal d'une Science Ouverte se concrétise petit à petit grâce à des politiques publiques en faveur de l'*Open Access*. L'expression « *Open Access* » peut être traduite en français par « *libre accès à l'information scientifique et technique* ». L'objectif du mouvement de l'*Open Access* est de « *rendre accessible gratuitement, en ligne, toute production issue d'un travail de recherche* » (ENSSIB, 2012).

Une Science Ouverte, c'est donc faire en sorte que les publications scientifiques et les données de la recherche soient ouvertes « à tous, chercheurs, entreprises et citoyens, sans entrave, sans délai, sans paiement » (*Plan pour une Science Ouverte*). Une Science Ouverte est censée avoir des retombées positives tant pour la recherche et les chercheurs que pour la société toute entière. Mais les obstacles à surmonter sont encore nombreux.

Dans ce second numéro de Didak'TIC, nous commençons par **définir l'information scientifique et technique (IST)** et par présenter des acteurs de l'IST encore trop souvent méconnus : les URFIST. Ensuite, nous vous proposons **un dossier qui revient sur les principaux enjeux de l'Open Access (page 13)**. L'*Open Access*, en ce qui concerne les publications des chercheurs, n'est plus à présenter et est aujourd'hui bien installé dans le paysage de la recherche française, notamment avec HAL – qui compte parmi les principaux réservoirs mondiaux de contenu IST –, et avec l'émergence de nouveaux modèles de publication comme les épi-revues. En 2016, l'article 30 de la *loi pour une République numérique* est venu renforcer la dynamique de l'*Open Access* en France, inscrivant dans la loi les recommandations de l'Union européenne en la matière. Cependant, même si le modèle encore prépondérant de l'abonnement est en déclin, les grands éditeurs commerciaux ont plus d'un tour dans leur sac et ne sont pas prêts de dire adieux à leurs profits colossaux ! En effet, ceux-ci investissent depuis quelques années la voie dorée en proposant un modèle économique auteur-payeur très lucratif que nous détaillerons.

Notre **deuxième dossier est centré sur les données de la recherche (page 39)**. Celles-ci ont pendant longtemps été négligées et ce n'est que récemment que l'on a pris la mesure de leur valeur pour la science. Leur préservation, leur gestion, leur partage et leur réutilisation constituent un nouvel axe de l'*Open Access*. Il est essentiel d'éviter que le phénomène d'accaparement des publications scientifiques par les éditeurs ne se reproduise avec les données de la recherche. En France, la *loi pour une République numérique* a donc cherché à les protéger de l'avidité des grands éditeurs commerciaux. La diffusion et la réutilisation des données de la recherche ouvrent des perspectives nouvelles à la science et viennent modifier en profondeur les pratiques des chercheurs qui ont besoin d'être accompagnés.

Mais pourquoi avons-nous choisi de vous parler cette année du thème de la Science Ouverte ? En tant que futurs professionnels de l'information et de la documentation, le mouvement de la Science Ouverte nous concerne tout particulièrement car :

- Les professionnels de l'information et de la documentation ont eu un rôle politique de premier plan dans le mouvement de l'*Open Access* et ont encore un **rôle politique** déterminant à jouer pour permettre sa généralisation et l'avènement d'une ère de la Science Ouverte.
- Les professionnels de l'information et de la documentation, de par leurs compétences, ont un **rôle de médiation documentaire** à jouer. Les réservoirs en information scientifique et technique en libre accès sur internet se multiplient, ce qui crée un phénomène de fragmentation du savoir. C'est là que les bibliothèques universitaires (BU) jouent un rôle de « passerelle vers les ressources électroniques » (*Barthelemy et al, 2016*). Les bibliothécaires localisent les archives ouvertes et les revues en *Open Access* qui sont en adéquation avec les besoins du public de la BU, les signalent dans le catalogue et les valorisent.
- Enfin, les professionnels de l'information et de la documentation ont **une fonction d'aide et de support technique**. L'*Open Access* rend nécessaires leurs compétences et est venu renouveler leurs missions. Ils accompagnent les chercheurs lorsqu'ils déposent leurs publications dans des archives ouvertes, ils permettent le suivi bibliométrique, ils numérisent et indexent des corpus, ils créent des bases de données et effectuent la maintenance des entrepôts institutionnels.

Comme l'explique le rapport « *Moving towards an Open Access future : the role of academic libraries* » publié en 2012, avec le développement de l'*Open Access*, le professionnel de l'information a un rôle de premier plan. En effet, l'apport des BU reposera de moins en moins sur leur fonds mais dépendra plutôt des services et de l'accompagnement qu'elles fourniront aux étudiants et aux chercheurs : « **Le professionnel de l'information sera la bibliothèque de demain¹** ».

Sonia SALAMI et Constance PICQUE.

BIBLIOGRAPHIE

BARTHELEMY Antoine, BAUDRY Julien, BRAUD Aurélia, CHARAZAC Christelle et GALOT Delphine, « Open access en bibliothèque universitaire : de nouveaux enjeux de médiations », *Revue française des sciences de l'information et de la communication* [En ligne], 8 | 2016.

FABRE Renaud, « L'application de la loi "pour une République numérique, un guide partagé pour le travail de la science, éléments de périmètre », DIST du CNRS le 31/03/2017. URL : http://www.cnrs.fr/dist/z-outils/documents/analyse-systemique_5points-alerte.pdf

DIRECTION DE L'INFORMATION SCIENTIFIQUE ET TECHNIQUE - CNRS. *Livre blanc — Une Science ouverte dans une République numérique*. Nouvelle édition [en ligne]. Marseille : OpenEdition Press, 2016. URL : <http://books.openedition.org/oep/1548>

ENSSIB, *Dictionnaire de l'ENSSIB*, entrée « open access », notice créée le 07 mai 2012 mise à jour le 25 février 2016.

HAMEAU Thérèse, « Aller vers le libre accès : le rôle des bibliothèques universitaires », article publié le 21/09/2012 sur le site de l'INIST (CNRS) : <http://openaccess.inist.fr>

NAEGELEN Pierre, « Plan national pour la science ouverte : beaucoup d'avancées et trois questions », publié le 13/07/2018 sur *hypothese.org*. URL : <https://openarchiv.hypotheses.org/4767>

SIAN Harris, *Moving towards an open access future: the role of academic libraries*, compte rendu d'une table ronde organisée par SAGE, en partenariat avec la British Library, Août 2012.

¹ SIAN, 2012. Traduction citation p.15 : « As one participant put it: the information professional is the library of the future »

Manifestes et déclarations en faveur de la Science Ouverte

- Budapest open access initiative : www.budapestopenaccessinitiative.org/read
- San Francisco Declaration on Research Assessment (DORA), 2013 : sf.dora.org/
- Leiden manifesto for research Metrics, 2015 : www.leidenmanifesto.org/
- Jussieu Call for Open science and bibliodiversity, 2017 : jussieucall.org/

Politiques publiques de la France et de l'Union européenne

- *Plan d'action national de la France pour une action publique transparente et collaborative pour 2018-2020*

Lien : www.etalab.gouv.fr/wp-content/uploads/2018/04/PlanOGP-FR-2018-2020-VFFR.pdf

- *Recommendation on access to and preservation of scientific information*, 25 avril 2018, European commission

Lien : http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51636

- *Amsterdam call for action on open science*

Lien : www.government.nl/documents/reports/2016/04/04/amsterdam-call-for-action-on-open-science

- *Open Science and its role in universities : a roadmap for cultural change*, mai 2018 (LERU)

Lien : www.leru.org/publications/open-science-and-its-role-in-universities-a-roadmap-for-cultural-change

- *Plan national pour la Science Ouverte*, 4 juillet 2018

Lien : http://cache.media.enseignementsup-recherche.gouv.fr/file/Actus/67/2/PLAN_NATIONAL_SCIENCE_OUVERTE_978672.pdf



SOMMAIRE

ÉDITORIAL 3

INFORMATION SCIENTIFIQUE ET TECHNIQUE	
Zoom sur l'Information Scientifique et Technique et les URFIST	7

UN PEU D'HISTOIRE	
Les revues scientifiques	12

DOSSIER : OPEN ACCESS	13
CCSD : moteur de l'Open Access en France	14
Loi pour une République numérique : éclairage sur l'article 30	18
Gold Open Access : quel modèle économique pour demain ?	25

UN PEU D'HISTOIRE	
Le facteur d'impact	32

DÉBAT	
Open education : Le MOOC au cœur des universités	33

DOSSIER : DONNÉES DE LA RECHERCHE	39
Ouverture des données de la recherche : de quoi parle-t-on ?	40

LE SAVIEZ-VOUS ?	
Quels outils pour exploiter les données de la recherche ?	44

DOSSIER : DONNÉES DE LA RECHERCHE	
Gestion et diffusion des données de la recherche : quels enjeux ?	45
Données de la recherche : le Dark Data des SHS	50
Le partage des données de la recherche dans la pratique	55
Open Research Data : le rôle des universités dans l'accompagnement des chercheurs	59

LE SAVIEZ-VOUS ?	
Les financements accordés par l'ANR	64



Le magazine Didak'TIC
ITIC Bâtiment Marc Bloch (Bât. E) Université
Paul-Valéry Montpellier - Route de Mende
34199 Montpellier Cedex 5 - www.didaktik.fr

DÉPÔT LÉGAL de la revue Didak'TIC :
décembre 2017, ISSN 2605-8812,
DATE DE PARUTION de Didak'TIC n°2 :
octobre 2018.

Didak'TIC

DIRECTRICE DE LA PUBLICATION :
Lise VERLAET.

PILOTES DU MAGAZINE : Sonia SALAMI,
Constance PICQUE.

RÉDACTION

Directrice de la rédaction : Sonia SALAMI
Rédacteurs et Rédactrices : Gilane BENFKIH,
Alison DEVILLIERS, Cédric MARION, Carole-
Anne MATEL, Constance PICQUE, Laure
PIERRAT, Sonia SALAMI.

CONCEPTION TECHNIQUE

Pour la partie PRINT : Responsable et
maquettiste : Constance PICQUE.
Iconographes : Sonia SALAMI, Constance
PICQUE.
Pour la partie WEB : Webistes : Sonia SALAMI,
Constance PICQUE, Cédric MARION.

COMITÉ DE VALIDATION : Hans DILLAERTS,
Lise VERLAET, Jean-Michel DUBOIS, Mireille
BACHELOT.

DIFFUSION Imprimé en France

Zoom sur l'Information Scientifique et Technique et les **URFIST**

Sonia SALAMI



Réseau IST (CC0)

Après avoir défini l'information scientifique et technique et dressé un bref historique de ses évolutions et de ses enjeux, cet article présente des institutions spécifiques à la France et encore trop souvent méconnues : les URFIST. Organisées en un réseau national, les URFIST ont pour objectif de favoriser l'accès à l'information scientifique et technique dans l'enseignement supérieur et la recherche.

QU'EST-CE QUE L'INFORMATION SCIENTIFIQUE ET TECHNIQUE (IST)?

La désignation « *information scientifique et technique* » est apparue dans les années 60 pour désigner : « ***l'ensemble des informations produites par la recherche et nécessaires à l'activité scientifique comme à l'industrie*** ». L'IST correspond donc aux informations produites par la communauté scientifique (enseignants-chercheurs, chercheurs, doctorants, étudiants de master) dans toutes les disciplines. L'IST se décline sous différentes formes de publications et inclut :

- **Les publications relevant de la littérature blanche**² c'est-à-dire : les articles scientifiques, les revues scientifiques, les ouvrages scientifiques.

- **Mais aussi les nombreuses publications relevant de la littérature grise** comme : les thèses, les mémoires, les habilitations à diriger des recherches (HDR), les pré-publications, les actes de colloques, les conférences, les rapports de recherche, les carnets de recherche, les jeux de données, les bases de données bibliographiques, les notices de brevets, les spécifications techniques décrivant des processus

de fabrication, la documentation technique accompagnant les produits³.

LES ÉVOLUTIONS DE L'IST

L'information scientifique et technique est née avec les premières revues scientifiques du 17^e siècle⁴. Le développement récent des nouvelles technologies de l'information et de la communication a profondément modifié l'IST, même si les enjeux fondamentaux restent inchangés c'est-à-dire (Schöpfel, 2011) :

- Protéger les droits d'auteur d'une idée ou d'une découverte scientifique (**enregistrement**).
- Diffuser les résultats de la recherche (**diffusion**).
- Soumettre les résultats de la recherche à la critique et au débat (**évaluation par les pairs**).
- Conserver ces résultats de façon pérenne (**archivage**).

On peut résumer les évolutions récentes de l'IST en cinq points :

1. Tout d'abord, **la numérisation de la production scientifique est venue bouleverser les modes de publication**. Avant les années 80 l'édition de l'IST est encore publiée par des sociétés savantes sous forme papier même si des éditeurs commerciaux se développent. Puis, avec le début des bases de données et la généralisation d'internet (qui a

vu la numérisation de la production scientifique)⁵. Les grands éditeurs commerciaux se trouvent rapidement en situation de quasi-monopole – les plus petits éditeurs n'ayant pu suivre la transformation numérique. Ils imposent alors aux BU de façon unilatérale le modèle de l'abonnement à des bouquets de revues aux coûts de plus en plus élevés⁶. La conséquence de cela est le point suivant.

2. **La question de l'accès à l'IST devient un enjeu central** car les institutions n'ont plus les moyens d'assumer les coûts des abonnements. Se développe alors le libre accès qui entraîne une diversification des modèles de publication⁷.

3. L'autre évolution importante est **l'inflation du nombre de publications** qui est corrélée à un **élargissement du public scientifique**⁸.

4. Au niveau de la diffusion, on observe une **accélération de la circulation des travaux** scientifiques en raison du développement du libre accès. Par ailleurs, on observe qu'**internet a facilité l'accès à la littérature grise** autrefois assez confidentielle et difficile d'accès (au point que l'on peut se demander si l'appellation de « *littérature grise* » est toujours pertinente)⁹.

5. Enfin, le numérique a entraîné **le développement d'infrastructures numériques** comme les bibliothèques numériques, les réseaux sociaux de chercheurs, les outils collaboratifs, les agrégateurs de revues électroniques...¹⁰

LA CRÉATION DES URFIST

Les Unités Régionales de Formation à l'Information Scientifique et technique (URFIST) comme leur nom l'indique, « *sont chargées de développer la*





Le logotype du réseau des URFIST © Urfist

maîtrise de l'information scientifique et technique (IST) dans l'enseignement supérieur et la recherche¹³».

Les URFIST existent depuis plus de trente ans (Noël, 2013). Elles ont été créées en 1982 à la suite de plusieurs rapports nationaux ayant pointé du doigt l'importance de l'information scientifique et technique (IST) et la nécessité de sa gestion :

- le Plan Calcul de 1966¹²
- le deuxième Plan Calcul de 1971
- le Rapport Nora-Minc de 1978 intitulé « L'informatisation de la société »

Dans un contexte où les bibliothèques universitaires (BU) étaient assez peu fréquentées et où se développaient des Banques de données en ligne dont l'usage restait assez confidentiel (seuls quelques enseignants-chercheurs et de rares étudiants savaient alors les utiliser), le besoin de formation se faisait sentir. Les URFIST ont alors été conçues comme « *un pont entre bibliothèque et université¹³* » et pour cette raison sont dirigées par un binôme composé d'un conservateur des bibliothèques et d'un universitaire (aujourd'hui, il s'agit souvent d'un enseignant-chercheur).

Chaque binôme est épaulé par un responsable administratif. Ainsi un URFIST est composé de trois personnes, à l'exception de l'URFIST de Paris qui dessert une zone géographique plus importante et qui est donc composée de six personnes.

Les URFIST sont des structures interacadémiques dont le but est de desservir l'ensemble du territoire¹⁴. Au départ, huit URFIST ont été créées. Il y en a sept aujourd'hui :

- URFIST de Bordeaux
- URFIST de Lyon
- URFIST de Nice
- URFIST de Paris
- URFIST de Rennes
- URFIST de Strasbourg
- URFIST de Toulouse

MISSIONS, ACTIVITÉS ET PUBLIC DES URFIST

Les missions des URFIST et le public qu'ils visent n'a fait que s'étendre depuis leur création. Au départ, les URFIST se concentrent sur la formation de formateurs – leur but est principalement la formation à l'interrogation des bases de données en ligne au niveau régional. **Cependant, aujourd'hui les missions des URFIST se sont élargies** et par conséquent leurs activités se sont diversifiées comme l'explique le blog des URFIST¹⁵.

Malgré cette diversification, la formation reste au centre des activités des URFIST - et ce depuis leur création - à travers la « *conception et mise en œuvre d'une offre cohérente de formation* », la « *formation de formateurs* » et la « *conception, réalisation et diffusion d'outils pédagogiques adaptés¹⁶* ». Les URFIST ont par ailleurs un rôle important de veille et de re-

cherche en information scientifique et technique qui leur permet d'avoir une fonction d'expertise et de conseil dans le domaine. Enfin, les URFIST prennent part « *à des activités éditoriales dans le domaine de la maîtrise de l'information* », animent des réseaux professionnels et scientifiques, organisent et interviennent dans des colloques ou lors de journées d'études.

Les domaines de formation, de veille, de recherche et d'expertise des URFIST concernent :

- la culture informationnelle de façon générale
- l'information scientifique et technique
- les ressources en IST pour l'enseignement supérieur et la recherche
- les méthodes et les techniques de recherche de l'IST
- les méthodes et les techniques d'analyse et d'exploitation de l'IST
- la production, l'édition, la communication et la diffusion scientifique¹⁷

Encore trop souvent méconnues, les URFIST s'adressent à un public très large composé : d'enseignants-chercheurs, de formateurs à la maîtrise de l'IST (ingénierie de formation et pédagogie), de professionnels de l'information et de la documentation de l'enseignement supérieur et de la recherche, de doctorants et d'étudiants avancés (à compter du master 2). A titre d'exemple, en 2013, malgré la petite taille de leurs effectifs, les URFIST ont réalisé « *621 actions, près de 3 500 heures de formation dont ont bénéficié plus de 9 300 personnes¹⁸* ».

CRÉATION DU GIS DU RÉSEAU DES URFIST

Le Groupement d'Intérêt Scientifique est une instance consultative récente créée en 2017 à la suite des « *Recommandations du groupe de travail URFIST*¹⁹ » présidé par Pierre Carbone en 2014. Dans l'objectif de structurer les URFIST en un réseau national doté d'implantations territoriales (Serres, 2015), le groupe de travail propose (entre autres) de constituer un groupement d'intérêt scientifique (GIS).

Le GIS du réseau des URFIST a pour objectif « *de faire émerger une*

politique de recherche commune, de mutualiser les innovations en matière de formation et de diffusion et de développer une veille en partenariat »²⁰. Cela passe par un renforcement du « *lien recherche-formation dans l'activité des URFIST* » et du « *lien entre universités et laboratoires autour de la recherche en IST* ».

Les missions du Conseil scientifique du GIS²¹ sont de :

- rendre un avis sur les projets et demandes de financement soumis par les partenaires
- formuler des recommandations sur les orientations scientifiques du GIS et faire des propositions d'actions

- rendre un avis sur le rapport d'activité et le rapport financier présentés annuellement par le directeur du GIS
- organiser l'évaluation scientifique des programmes de recherche sélectionnés par le GIS dans les conditions déterminées par le conseil des partenaires

■ Sonia SALAMI

1 Définition donnée par le gouvernement. Accessible sur le site web : www.enseignementsup-recherche.gouv.fr

2 La littérature blanche correspond aux documents « diffusés par "voie normale", c'est-à-dire soumis à la formalité du dépôt légal et qui sont consultables grâce, entre autres, à des outils de recherche et à une organisation documentaire capable de maîtriser le flux incessant de l'information » (Halima, S. (2007). *La littérature grise : face méconnue de la documentation scientifique (1re partie)*. Documentation et bibliothèques, 53(4), p.205. Accessible sur le site web : <https://www.erudit.org>

3 Source : site du gouvernement

4 Pour avoir plus de détail, voir article de Carole-Anne MATEL sur l'histoire des revues scientifiques

5 Voir résumé de l'intervention de Rémi Barré dans NOËL, 2012.

6 Voir dans ce numéro l'article de Laure PIERRAT sur la loi pour une République Numérique

7 Pour avoir plus de détail sur les épi-revues, voir les articles de Sonia SALAMI, Cédric MARION et Constance PICQUE « *Le CCSD, un acteur majeur de l'Open Access en France* »

8 Voir SCHÖPFEL, 2001 : « *Pourquoi cette augmentation, que certains décrivent comme une explosion, en regrettant de voir aujourd'hui beaucoup trop de titres ? En fait, il n'y a pas trop de titres. Mabe (2003) a démontré le rapport direct entre le nombre des revues, le nombre des chercheurs et le budget global de la recherche, avec une corrélation positive significative entre ces trois chiffres. Autrement dit : le nombre des revues augmente au même rythme que celui des chercheurs et des dépenses pour la recherche.* »

9 HALIMA, 2007, p. 210

10 www.cnrs.fr/inshs/recherche/ist/information-scientifique.htm

11 Blog des URFIST, section « *A propos* »

12 « *Le plan Calcul était un plan gouvernemental français lancé en 1966 par le président Charles de Gaulle sur l'impulsion de Michel Debré et d'un groupe de hauts fonctionnaires et d'industriels, destiné à assurer l'autonomie du pays dans les technologies de l'information, et à développer une informatique européenne* » (définition Wikipedia).

13 « *Les URFIST* » Bulletin des bibliothèques de France (BBF)

14 « *Chaque URFIST a pour mission de couvrir plusieurs académies ; le champ de l'URFIST de Paris est particulièrement vaste : 10 académies dont les DOM-TOM. Seules les académies de Lille et Amiens ne sont pas couvertes, en effet la 8ème URFIST a disparu à la suite de la création de l'UFR IDIST à Lille (devenue depuis département SID)* » (Carbone, 2014)

15 Ibidem.

16 Ibidem.

17 Blog des URFIST, section « *A propos* »

18 Ibid. *Les URFIST touchent aujourd'hui un large public : « Leur bilan n'est pas négligeable : en 2013, 621 actions, près de 3 500 heures de formation dont ont bénéficié plus de 9 300 personnes.* » (Carbone, 2014, p.1)

19 Titre du rapport (CARBONE, 2014)

20 Ibidem p. 10.

21 *Les missions du conseil scientifique du GIS sont tirées de l' « Appel à candidatures pour former le Conseil Scientifique du GIS « Réseau des URFIST ».*

BIBLIOGRAPHIE

CARBONE Pierre, *Recommandations du groupe de travail URFIST*, rapport, 2014.

COMBEROUSSE Martine, « Histoire de l'information scientifique et technique », *Bulletin des bibliothèques de France (BBF)*, 2000, n° 2, p. 144-145.

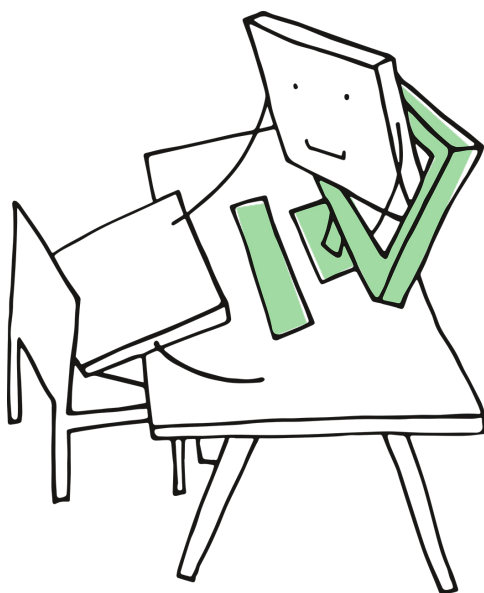
NOËL Elisabeth, « Vers une rénovation des URFIST », article publié sur le site de l'ENSSIB le 19/12/2015. URL : www.enssib.fr/breves/2015/02/19/vers-une-renovation-des-urfist

NOËL Elisabeth, « Trente ans de politiques d'information scientifique et technique (1982–2012) », compte rendu de la journée nationale du réseau des URFIST du 26 novembre 2012.

SCHÖPFEL Joachim, « Les mutations du paysage de l'information scientifique » dans *La formation des doctorants à l'information scientifique et technique*, Villeurbanne : Presses de l'enssib, 2011. URL : <http://books.openedition.org/pressesenssib/941>

SERRES Alexandre, « Le Réseau des URFIST à un tournant de son histoire », publié sur le blog des URFIST le 16/02/2015. URL : urfistinfo.hypotheses.org/2725

« Les URFIST », *Bulletin des bibliothèques de France (BBF)*, 1983, n° 3, p. 285-287. URL : <http://bbf.enssib.fr/consulter/bbf-1983-03-0285-002>



Les revues scientifiques

Sonia SALAMI

BIBLIOGRAPHIE

BACH Jean-François, « L'édition scientifique victime de son succès », *discours à l'Académie des Sciences*, 03/02/2015 URL : http://www.academie-sciences.fr/pdf/discours/s170614_bach.pdf

CHARTRON, Ghislaine et SCHÖPFEL Joachim, « Open access et Open science en débat », *Revue française des sciences de l'information et de la communication*, 11, 2017. URL : <http://journals.openedition.org/rfsic/3331>

GUEDON, Jean-Claude, « A l'ombre d'Oldenburg : Bibliothécaires, chercheurs scientifiques, maisons d'édition et le contrôle des publications scientifiques », ARL Meeting, Toronto Université de Montréal, 2001. URL : <https://halshs.archives-ouvertes.fr/halshs-00395366/document>

VERLAET Lise, CHANTE Alain, « La notion d'autorité dans l'édition de la recherche : évolution des formes, changements d'autorité ? », *Communication & langages*, vol. 192, no. 2, 2017, pp. 27-46 URL : <https://www.cairn.info/revue-communication-et-langages1-2017-2-p-27.htm>

La naissance des revues scientifiques au XVII^e siècle

Avant l'apparition des revues scientifiques, le mode de communication des savants était la correspondance. En 1665, les deux premières revues scientifiques sont créées presque en même temps en France et en Angleterre : la revue *Philosophical Transactions of the Royal Society of London* et le *Journal des sçavans* (Bach 2017). **La publication de revues permet aux sociétés savantes de diffuser les découvertes sur un support stable, pérenne et surtout validé par la communauté.**

En effet, les communautés scientifiques sont des « *espaces de rencontre et de débat* » entre érudits (Caune, 2013). Elles sont considérées comme des cercles d'experts auprès de l'Etat et décident si telle ou telle découverte ou avancée relève du domaine scientifique. Ce sont des autorités savantes car elles garantissent la conformité, la valeur scientifique et l'authenticité. Elles détiennent de fait une certaine notoriété. Ce sont ces critères qui permettent à un chercheur d'être reconnu par ses pairs et de gagner en visibilité.

Les revues scientifiques représentent un bouleversement non seulement en ce qui concerne la diffusion et la validation des découvertes scientifiques mais aussi en ce qui concerne la manière de travailler des savants qui peuvent alors « *se référer à des supports écrits pour appuyer et construire leurs recherches* » (Verlaet et Chante, 2017).

Le développement des revues scientifiques 1850-1914

Après la création des premières revues scientifiques, celles-ci **se développent progressivement jusqu'au XX^e siècle**. On note quatre évolutions majeures, comme le relatent Verlaet et Chante (2017) :

- Au milieu du XVIII^e siècle apparaissent les **premières revues spécialisées** comme les *Annales de chimie* (Bach, 2017).
- En 1789, la déclaration des droits de l'Homme, article II annonce la liberté d'imprimer, ce qui **ouvre la voie à d'autres thématiques et augmente le rythme des publications**.
- Durant la III^e République, le savoir se diffuse dans les universités, où de nouvelles disciplines apparaissent et où les **travaux de recherche peuvent se faire librement**.

À la fin du XIX^e siècle, le délai de publication d'un article reste long, ce qui a des conséquences sur la prise en compte des nouvelles découvertes par les chercheurs. Cependant, **en raison du développement de l'Information Scientifique et Technique, on assiste à l'arrivée des professionnels de l'information qui apportent des nouvelles techniques en documentation** qui permettent de faire face à l'augmentation du nombre des publications. **Les éditions scientifiques restent néanmoins toujours gérées par les communautés et des entreprises d'érudits**. Ce qui n'est plus le cas aujourd'hui ! En effet, l'édition scientifique n'est pratiquement plus assurée par des sociétés savantes mais est à présent aux mains de grands éditeurs commerciaux en situation d'oligopole, lesquels réalisent d'énormes profits.

■ Sonia SALAMI

DOSSIER: OPEN ACCESS



Le CCSD, un **acteur majeur** de l' *Open Access* en France

Sonia SALAMI, Cédric MARION
Constance PICQUE

Portant l'esprit du libre accès, le Centre pour la Communication Scientifique Directe (CCSD) est une structure ayant pour vocation d'aider les chercheurs mais également leur environnement institutionnel (universités, organismes de recherches...) en leur fournissant des outils permettant l'archivage, la diffusion ainsi que la valorisation de leurs publications et de leurs données scientifiques. Pour répondre aux besoins des chercheurs, de nombreux services sont proposés par le CCSD. Créé en 2000 par le CNRS, le CCSD est depuis devenu un acteur majeur de l'Open Access en France. **Dans cet article, nous présentons les principaux projets portés par le CCSD.**



Le savoir accessible en ligne (CC0)

ZOOM SUR LA PLATEFORME HAL

Sous cet acronyme qui nous rappelle avec nostalgie cette intelligence artificielle (ndlr : HAL 9000) qui dans *2001, l'Odyssée de l'espace* de Stanley Kubrick était chargée d'enquêter sur un signal émis à l'autre bout du système solaire, nous retrouvons les termes suivants : **Hyper Articles en Ligne**. HAL est une archive ouverte lancée en 2001 par le Centre pour la Communication Scientifique Directe (CCSD) du Centre National pour la Recherche Scientifique (CNRS).

HAL propose une plateforme en ligne destinée au dépôt et à la diffusion de publications scientifiques, déjà publiées ou non, provenant des établissements d'enseignement supérieur et de recherche français ou étrangers, des laboratoires publics ou privés. HAL est une archive multidisciplinaire qui a été « choisie par l'ensemble de la communauté scientifique et universitaire française pour la diffusion des savoirs » (site du CCSD). Elle accueille différents types de publications scientifiques comme les thèses, les articles, les chapitres d'ouvrages, les habilitations à diriger des recherches, les communications de congrès, les brevets, les rapports et même des cours. Plus récemment HAL a permis d'accueillir les données

de la recherche : vidéos, images, sons, cartes et depuis septembre 2018 des logiciels ! Principalement francophone, la plateforme HAL a vocation de permettre un libre accès aux contenus d'information scientifique et technique (IST) qu'elle héberge. Elle contribue à la libre diffusion et à la valorisation du savoir scientifique.

A ce jour, l'archive ouverte HAL est un des principaux réservoirs mondiaux de contenu IST. Elle représente la plus grande base de données pour la recherche francophone avec plus d'un million de publications (notices comprises) provenant de chercheurs et d'institutions francophones du monde entier.

La plateforme HAL entend répondre à de nombreux enjeux. Nous pouvons citer entre autres :

- l'amélioration de la visibilité de la production des laboratoires et des activités de la recherche
- une amélioration de la circulation des publications entre les chercheurs
- une meilleure gestion des bibliographies
- l'assurance d'une pérennité d'accès aux résultats de la recherche avec une sauvegarde des contenus qui est garantie et une adresse URL unique et stable
- le respect des exigences des financeurs publics de la recherche (comme la Commission européenne via le programme Horizon 2020) qui en contrepartie de leur financement demande de plus en plus souvent que les résultats de la recherche (articles) et données (jeux de données et data papers) soient publiés en libre accès de façon pérenne.
- De répondre aux nouvelles possibilités de la loi pour une République numérique de 2016 qui autorise la publication en *Open Access* des ver-

sions *post-print* des articles après le respect d'une période d'embargo, de 6 mois en Science Technique et Médical (STM) et de 12 mois en Sciences Humaines et Sociales (SHS).

LES AUTRES ARCHIVES OUVERTES

Le CCSD propose de nombreuses déclinaisons de sa plateforme d'archive ouverte HAL. Entre autres nous pouvons citer :

- **HAL-SHS, l'archive ouverte thématique entièrement dédiée aux disciplines des sciences humaines et sociales.** Elle regroupe 27 disciplines allant de l'anthropologie biologique à la science politique en passant par le droit.
- **MédiHAL, lancée en 2010 est l'archive ouverte qui a vocation à agréger les données visuelles, audiovisuelles et sonores produites dans le cadre de la recherche scientifique.** De très nombreux formats de fichiers sont acceptés. Cette archive ouverte correspond donc à l'entrepôt de données scientifiques « multimédia » de HAL. Il faut noter que les données visuelles, audiovisuelles et sonores qui sont déposées dans HAL apparaissent également dans MédiHAL.
- **TEL, acronyme de Thèses-En-Ligne a été créée en 2001. Cette plateforme a pour objectif de promouvoir l'auto-archivage en ligne des thèses de doctorat ainsi que des habilitations à diriger des recherches.** TEL est pluridisciplinaire et permet à tout usager, ayant soutenu sa thèse d'auto-archiver son travail. TEL est largement plébiscité : 71% des thèses soutenues en France en 2016-2017 ont été déposées sur le serveur TEL de HAL .

LES PORTAILS INSTITUTIONNELS SUR HAL

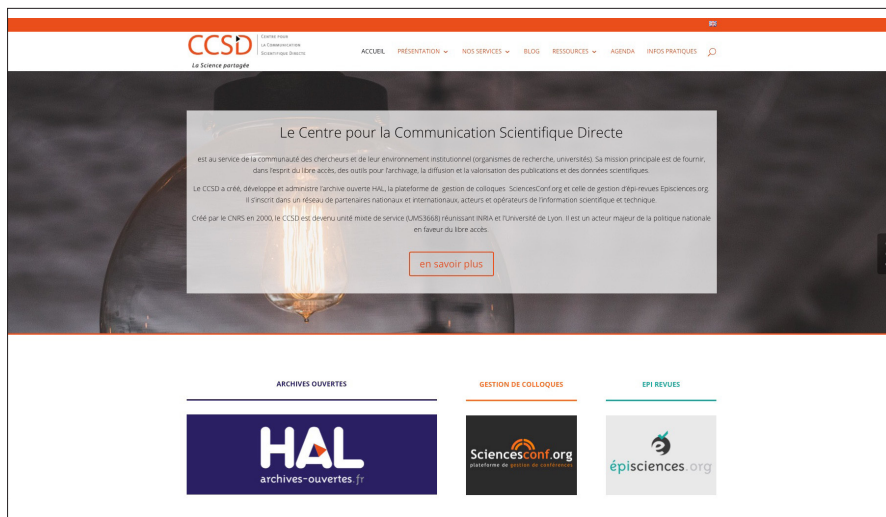
Si HAL a vu le jour en 2001, c'est à partir de 2006, c'est suite au *Protocole d'Accord pour la publication scientifique sur des archives ouvertes* que HAL est choisie comme plateforme commune d'archivage en libre accès par des établissements français favorables au libre accès. Plus tard, en 2013, est signé une convention de partenariat en faveur des archives ouvertes et de la plateforme HAL dont le but est « *le développement coordonné des archives ouvertes sur la base du développement de la plateforme mutualisée HAL et de son interconnexion avec les archives institutionnelles des établissements partenaires* ».

Les archives institutionnelles et HAL s'articulent ensemble : chaque établissement partenaire a ainsi la possibilité de créer un portail dans HAL qui lui est dédié et où il peut verser les publications de ses chercheurs en *Open Access*. Ces portails permettent aux institutions de valoriser les travaux de leurs chercheurs tout en étant autonome dans la gestion.

Autour du projet fédérateur HAL, s'est mis en place une politique très dynamique de développement des archives ouvertes institutionnelles, avec au total 49 nouveaux portails institutionnels ouverts dans HAL depuis 2013. L'objectif à terme étant de réaliser une centralisation de toutes les archives institutionnelles françaises.

LA PLATEFORME DE GESTION DES RÉFÉRENTIELS HAL

Le CCSD propose également une plateforme permettant de gérer l'ensemble des référentiels utilisés par



Page d'accueil du site internet du CCSD

HAL, elle concerne notamment les auteurs, les structures de recherches, les revues et discipline, les projets de l'Agence Nationale de la Recherche ainsi que les projets européens. Le dispositif se nomme AURÉHAL pour *Accès Unifié aux Référentiels HAL. Module de recherche des formes d'auteurs présents dans HAL*.

LA GESTION DE CONFÉRENCES SCIENTIFIQUES

Dans l'objectif de répondre à la demande du CNRS de mettre en place un outil de gestion scientifique de conférences pour la communauté de l'Enseignement Supérieur et de la Recherche, le CCSD a ouvert à la fin de l'année 2010 la **plateforme de gestion de congrès scientifiques Sciencesconf.org**. Cette plateforme recense 5849 conférences et on observe une croissance forte du nombre de conférences organisées.

LE PORTAIL DATA.ARCHIVES-OUVERTES.FR

Le portail data.archives-ouvertes.fr est un espace regroupant une documentation qui présente la base de données appelée Triplestore qui contient « **toutes les données pré-**

sentes dans l'archive ouverte HAL ainsi que dans les différents référentiel d'AURÉHAL » (source : site du CCSD). Le portail contient également les différents schémas de données, une interface d'interrogation de la base de connaissance structurée en RDF par le protocole SPARQL endpoint, ainsi qu'un accès au téléchargement d'une archive zip de l'ensemble des contenus du triplestore, cette archive zip est actualisée tous les mois. Cette base de données contient toutes les données présentes dans l'archive ouverte HAL et également toutes les données présentes dans les différents référentiels d'AURÉHAL.

UNE PLATEFORME POUR ÉPI-REVUES

Lancé en 2013 par l'Institut Fourier en partenariat avec le Centre pour la Communication Scientifique Directe (CCSD), et l'INRIA, **la plateforme episciences.org regroupe des épi-revues en libre accès**. La plateforme episciences.org se veut pluridisciplinaire et a vocation à accueillir des revues de toutes les disciplines scientifiques. On peut déjà y retrouver de nombreuses revues s'appliquant notamment à l'informatique et aux mathématiques appliquées ou aux sciences humaines et sociales.

Mais quelle est sa particularité ? Cette plateforme est une nouveauté dans le monde des modes de publication en libre accès car elle allie les deux voies du libre accès, la voie dorée et la voie verte. En effet, une épi-revue combine les archives ouvertes et l'évaluation par les pairs des articles.

D'une part, la plateforme met à disposition des articles qui proviennent des archives ouvertes compatibles avec Episciences comme HAL, arXiv ou CWI. D'autre part, la plateforme dispose de comités de lecture, tel que Episciences IAM et Episciences-Maths qui examinent les articles proposés et échangent avec l'auteur à partir de l'outil de *workflow* éditorial intégré. Lors de la publication, un label *Episciences IAM* est attribué à l'article pour assurer de sa qualité (Riverieux et al., 2014).

Ainsi, **Episciences.org se sert de la voie verte pour obtenir des articles dans les archives ouvertes, mais les transforme en véritable publications sans pour autant passer par les circuits éditoriaux traditionnels.** Les publications sur cette plateforme donnent lieu à des accès gratuits aux articles scientifiques pour le lecteur (pas d'abonnement) et pour l'auteur de l'article (pas d'APC). Le principe de l'Epi-revue est de réaliser des revues à moindre coûts. Ce système permet également au chercheur de conserver ses droits d'auteurs.

LES AUTRES PROJETS DU CCSD

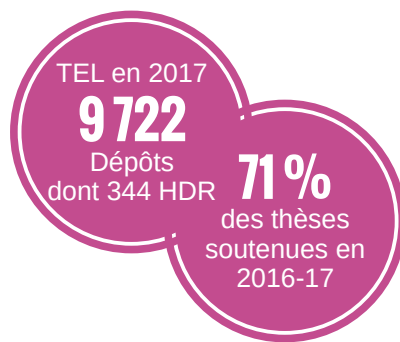
Le CCSD participe également à d'autres projets en lien avec l'*Open Access* comme :

● **Le projet Campus-AAR** : *Le Campus Archives Audiovisuelles de la Recherche* a pour but de constituer et de permettre l'exploitation « *des patrimoines scientifiques des sciences humaines et sociales (SHS) sous forme d'archives audiovisuelles (AV) numériques* » en *Open Access*.

● **Le projet Héloïse** qui est un site d'information pour les chercheurs qui compile les politiques éditoriales des revues scientifiques françaises quant au dépôt des articles en *Open Access*. Les éditeurs peuvent y signaler leur politique.

● **Le projet ISIDORE** qui est une plateforme donnant accès à des ressources numériques en SHS en libre accès. Sont moissonnées et indexées des archives constituées de corpus, de publications scientifiques et d'actualités de la recherche.

- **Sonia SALAMI**
- **Cédric MARION**
- **Constance PICQUE**



BIBLIOGRAPHIE

Définition du Centre pour la communication scientifique directe (CCSD) selon Wikipédia

Définition de l'Hyper articles en ligne (HAL) selon Wikipédia

Présentation générale d'Épisciences, accessible en ligne :
- jao2015.sciencesconf.org/
- conference/jao2015/pages/epiScience_JCouperin_2015_10_13.pdf

Sites internet :

AURÉHAL : aurehal.archives-ouvertes.fr

HAL : hal.archives-ouvertes.fr

HAL-SHS : halshs.archives-ouvertes.fr

Héloïse : heloise.ccsd.cnrs.fr

ISIDORE : www.rechercheisidore.fr

Campus-aar : campusaar.hypotheses.org

Data : data.archives-ouvertes.fr

MédiHAL : medihal.archives-ouvertes.fr

Sciences conf : sciencesconf.org

TEL : tel.archives-ouvertes.fr

Episciences : episciences.org

CCSD : www.ccsd.cnrs.fr

1 Extrait de la Convention de partenariat en faveur des archives ouvertes et de la plateforme mutualisée HAL ". Lien vers le document : http://cache.media.enseignementsup-recherche.gouv.fr/file/HAL/93/3/01_Convention_HAL_246933.pdf

Loi pour une République numérique, Éclairage sur l'article 30

Laure PIERRAT

Une législation dans le domaine du numérique était nécessaire afin d'encadrer et de sécuriser juridiquement des pratiques existantes. La loi pour une République numérique est venue combler ce vide. Elle a, entre autres, pour ambition de favoriser l'Open Access dans le domaine des publications scientifiques et de lutter contre la mainmise des grands éditeurs commerciaux sur les productions scientifiques.



Projet de loi pour une république numérique (source : www.republique-numerique.fr)

La loi n° 2016-1321 pour une République numérique, promulguée le 7 octobre 2016, a trois objectifs :

- Garantir la protection des individus notamment en matière de données personnelles.
- Assurer au plus grand nombre l'accès à internet (droit au maintien d'une connexion, rendre plus facile l'accès à internet aux personnes en situation de handicap, accompagnement des individus vers les services numériques).
- Favoriser la libre circulation des données, du savoir et de la connaissance.

C'est ce dernier aspect qui va nous intéresser, et en particulier la mesure de l'article 30 qui concerne l'ouverture des publications scientifiques. Après avoir fait un bilan de la situation des publications scientifiques en *Open Access* avant la loi, nous verrons quels impacts a eu cette loi.

POURQUOI FAIRE UNE LOI EN FRANCE EN FAVEUR DES PUBLICATIONS SCIENTIFIQUES EN OPEN ACCESS?

Pour comprendre les enjeux d'une loi en faveur des publications scientifiques en *Open Access*, il faut revenir sur les politiques éditoriales des grands groupes avant l'entrée en vigueur de la loi pour une République numérique et se pencher sur les politiques européennes.

La crise du modèle de l'abonnement

Bien souvent, lorsque le chercheur signe un contrat d'édition pour la publication de ses travaux, il est contraint de céder ses droits d'auteur (et ce, même s'il n'est pas rémunéré). **Dans ce type de contrat, les éditeurs cherchent à obtenir une cession la plus large possible des droits d'auteurs** (notamment en ce qui concerne les droits de diffusion, de reproduction), et à obtenir une cession exclusive. Par conséquent, **si le chercheur accepte de telles conditions, il ne peut plus faire usage librement de ses propres publications puisque celles-ci appartiennent désormais à l'éditeur.**

Ce modèle économique aboutit à une **privatisation du savoir** tant il entrave la circulation des publications des chercheurs en érigeant des barrières payantes. En effet, pour faire valoir une connaissance scientifique (souvent produite avec de l'argent public), les chercheurs doivent publier leurs écrits dans une revue scientifique. Ils passent ainsi par des maisons d'édition à qui ils vont devoir, dans la plupart des cas, céder leurs droits d'auteurs. Selon les pays, le droit est différent. Si bien qu'un éditeur anglo-saxons du secteur privé est en droit d'imposer la cessation des droits exclusifs aux chercheurs empêchant toute mise à disposition au public. Face à la pression exercée par la concurrence et à l'obligation d'une reconnaissance par des pairs pour valider leurs travaux de recherches, il est difficile pour les chercheurs de refuser. D'autant plus que l'article, une fois publiée, leur permettra d'acquérir une certaine visibilité, de voir leur carrière avancer et surtout de percevoir des financements publics.

La diffusion des publications des chercheurs ainsi limitée aux revues

scientifiques est donc souvent disponible de manière payante pour le lecteur (via abonnement ou paiement d'une somme non négligeable par article consulté)¹.

Ce système a deux conséquences négatives :

- D'une part, **la main-mise des éditeurs entrave le partage des connaissances.** En effet, les scientifiques, pour mener à bien leurs recherches, doivent avoir accès aux travaux publiés, par d'autres chercheurs. Or, comme le souligne Martin Clavey pour Rue89, « *la plupart des articles scientifiques se trouve encore derrière des paywalls liés à des abonnements que les laboratoires, les bibliothèques universitaires et les centres de recherche ne peuvent pas toujours payer, alors que les articles sont écrits par leurs propres chercheurs* » (Clavey, 2015).

- D'autre part, **les institutions publiques payent deux fois les publications.** Il y a là un dysfonctionnement majeur quant aux investissements publics dans la recherche. Il est important de noter que les auteurs sont des chercheurs et que la majorité d'entre eux sont des fonctionnaires (donc payés par les institutions publiques), les évaluateurs, qui contribuent la plupart du temps bénévolement, le sont eux aussi. Enfin, les bibliothèques universitaires qui sont contraintes d'assumer le coût souvent exorbitant des abonnements (qui ne cessent d'augmenter), sont, elles aussi financées par l'argent public.

N'ayant plus les moyens de se payer l'accès aux revues scientifiques, un certain nombre de bibliothèques ont dû se désabonner.

Les prix des abonnements aux revues scientifiques ont, en effet, augmenté

« de 200 % à 300% entre 1975 et 1995 et de 22 % à 57% entre 2004 et 2007 » (Bernault, 2015).

En raison de cette situation, le contexte s'est tendu autour des années 2010-2012. **La contestation des politiques éditoriales des grandes maisons d'édition scientifique a abouti à un appel au boycott** lancé par certains chercheurs, comme le mathématicien britannique Sir William Timothy Gowers (Moatti, 2013). Cet appel a aussi été relayé par l'université de Harvard (Benjamin, 2012) qui entendait dénoncer les politiques des grands groupes éditoriaux.

Ainsi, pour répondre à cette crise du financement des abonnements et aux recommandations européennes sur le sujet, il s'est avéré nécessaire de légiférer pour limiter le décalage entre le droit et les pratiques IST existantes. **En effet, les scientifiques ont pris très tôt conscience de la possibilité qu'offraient les nouvelles technologies de s'affranchir des éditeurs commerciaux et de construire des dispositifs de Science Ouverte.** A titre d'exemple, le journaliste scientifique Martin Clavey, dans son article, relate l'utilisation d'un hashtag pirate pour faciliter le partage d'articles entre chercheurs. Il cite également le nom des sites Sci-Hub et Libgen considéré par certains comme une solution, et par d'autres comme du piratage (Clavey, 2015).

L'évolution de la législation autour de l'Open Access au niveau européen

L'Union européenne, consciente de l'intérêt de construire une science ouverte, a émis une série de recommandations en ce sens. Ainsi en 2012, la Commission européenne (CE) à formuler un certain nombre

de prescriptions en vue de favoriser le libre accès aux publications issues de la recherche. Elle a fixé notamment des délais après lesquels il est possible de diffuser en libre accès les publications : les délais d'embargo sont de 6 mois après la date de publication, sauf pour les publications dans les domaines des sciences sociales et humaines où le délais est de 12 mois. En parallèle, la CE cible les publications issues de recherche financées au moins pour moitié par des fonds publics. La CE met en application de manière concrète sa politique de soutien au mouvement de l'Open Access via son programme Horizon 2020, en obligeant les chercheurs financés par ce programme à publier en libre accès leurs publications.

Il est important de noter que la CE préconise, à l'attention des Etats membres, d'établir « des politiques claires en matière de diffusion des publications scientifiques issues de la recherche financée par des fonds publics et du libre accès » (DIST – CNRS, 2017). **C'est dans ce contexte que l'Allemagne adopte, en 2013, un texte qui modifie le droit d'auteur en instaurant un droit d'exploitation secondaire.** L'auteur peut, sous certaines conditions, diffuser son travail en Open Access après l'avoir publié chez un éditeur. Ces dispositions suivent les recommandations européennes. Sont concernés par la loi, les écrits nés d'une activité de recherche financée au moins pour moitié par des ressources publiques et la publication est autorisée à la suite d'une période d'embargo. A noter cependant qu'en Allemagne la période d'embargo est fixée à 12 mois sans distinction de domaine scientifique.

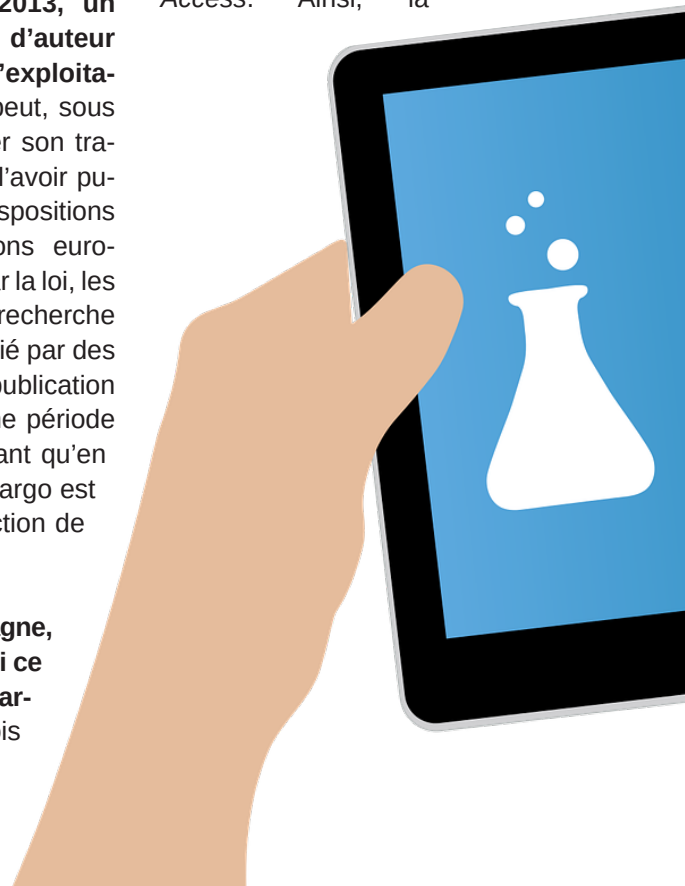
L'Italie, à la suite de l'Allemagne, a créé la même obligation si ce n'est que la période d'embargo est plus longue : 18 mois

pour les textes en science-technique-médical (STM) et 24 mois pour les sciences humaines et sociales (SHS).

On notera même si cela n'est pas notre propos, que **la politique en faveur des publications scientifiques en libre accès est présente également hors de l'Union européenne** comme aux États-Unis où « le Consolidated Appropriation Act oblige les agences fédérales recevant plus de 100 millions de dollars par an pour financer des travaux de recherche à s'assurer que les articles résultant de ces recherches soient accessibles en Open Access au plus tard douze mois après leur publication » (Bernault, 2015).

LES APPORTS DE LA LOI POUR UNE RÉPUBLIQUE NUMÉRIQUE EN MATIÈRE D'ÉDITION SCIENTIFIQUE EN OPEN ACCESS

La loi française a largement suivi les recommandations de l'Union européenne en matière d'Open Access. Ainsi, la



France protège le libre accès aux écrits scientifiques issus « *d'une activité de recherche financée au moins pour moitié²* » par des fonds publics après une période d'embargo.

En outre, elle garantit un droit d'exploitation secondaire, mécanisme largement inspiré de la loi allemande.

Présentation des dispositions de l'article 30

L'article 30 traite principalement de la diffusion des publications scientifiques par les auteurs.

Qui est concerné ? Selon la loi sont concernés les chercheurs dont les travaux sont financés pour moitié par des fonds publics.

Qu'est ce qui est concerné par cet article de loi ? La loi parle d'écrits scientifiques publiés dans un périodique paraissant au moins une fois par an.

Dans ce cas, il est possible de déposer la version post-print de l'article : c'est-à-dire la version de l'article qui a été validée par les pairs. L'article est donc validé scientifiquement mais pas encore mis en forme selon le format de la revue.

Quand déposer ? Si la revue publie déjà en *Open Access*, la diffusion immédiate dans une archive ouverte est possible, en fonction du contrat établi avec l'éditeur. Dans le cas contraire, le droit d'exploitation secondaire dû à l'auteur se déclenche 6 mois après la date de la première publication dans le domaine des sciences et techniques. Ce délai passe à 12 mois pour les sciences humaines et sociales.

Où déposer ? La loi indique « *dans un format ouvert par voie numérique* ». Cette formulation est suffisamment large pour recouvrir un grand nombre de possibilités sur lesquelles nous reviendrons.

Que se passe-t-il s'il y a un contrat de cession des droits d'auteur au profit de l'éditeur ?
Le dernier alinéa de l'article 30

stipule que rien ne peut contrevenir à l'application de cette mesure : ni loi étrangère, ni contrat d'édition signé. La loi garantit, ainsi, au chercheur un droit d'exploitation secondaire. Ce droit est un droit inaliénable et donc il passe outre les clauses contraires de tout type de contrat ou de loi étrangère.

Retour sur les choix du législateur

Il faut bien comprendre que le législateur français a été confronté à une forte opposition entre les chercheurs et les éditeurs. Il fallait trouver un juste équilibre pour ne pas court-circuiter les éditeurs, c'est ainsi que le législateur a imposé certaines conditions.

● **La loi souligne la nécessité d'un caractère public des financements.** En indiquant que seuls sont concernés les écrits scientifiques issus « *d'une activité de recherche financée au moins pour moitié* » par l'argent public, le gouvernement souhaite rétablir un juste équilibre quant aux investissements réalisés dans la recherche.

Une loi 2.0 tant sur le fond que sur la forme : le saviez-vous ?

Pour la première fois dans l'histoire juridique française, une proposition de loi a été mise en ligne sur une plateforme afin que l'ensemble des citoyens puissent intervenir dans le processus en postant des commentaires et en soumettant des amendements, avant son adoption.

Cette pratique a été déterminante pour trancher certaines questions concernant l'*Open Access*, notamment grâce à la mobilisation du monde scientifique (chercheurs et professionnel de l'IST).

C'est ainsi que l'ancien député Emeric Brehier, alors rapporteur pour la Commission Culture, a pu constater que l'article 30 a été « *l'une des dispositions qui a eu le plus connexions sur la plateforme, c'est-à-dire le plus de participation. [...] Cela a eu des conséquences sur le travail législatif. Nous avons pu nous appuyer sur les résultats de cette consultation pour valider certaines formulations notamment pour maintenir les délais d'embargos* » (Tambou, 2017).

● **Délai d'embargo : où fixer le curseur ?** Le législateur a fixé un délai d'embargo de 6 mois dans le domaine des sciences et techniques et 12 mois pour les lettres, sciences humaines et sociales. Ce délai a été choisi pour établir un compromis entre les volontés des chercheurs et des éditeurs : il a été décidé de se conformer aux recommandations européennes. Ce délai a l'avantage d'être suffisamment raisonnable pour que la portée de l'information ne soit pas obsolète au moment où le scientifique pourra exercer son droit d'exploitation secondaire. D'autre part, ce laps de temps doit également garantir une marge d'exclusivité pour l'éditeur.

● **Est concerné un écrit scientifique publié dans un périodique paraissant au moins une fois par an.** La loi ne s'applique ainsi qu'aux écrits publiés dans des périodiques. Ce droit d'exploitation secondaire exclut donc toute publication réalisée dans une monographie³. De même pour l'utilisation des images, provenant d'un tiers, le législateur a précisé que ce droit n'était applicable qu'aux écrits scientifiques, les images ne sont donc pas réutilisables. Il faut négocier pour pouvoir diffuser ce type de contenu.

● **Choix du dépôt :** La loi laisse au chercheur une grande liberté quant au choix du lieu de dépôt. Il n'y a pas d'espace plébiscité, il peut donc choisir de diffuser sa publication dans une archive ouverte qu'elle soit internationale ou nationale ou institutionnelle. Le dépôt en archives ouverte est fortement conseillé puisqu'il garantit un archivage pérenne et une visibilité accrue de la publication. Parallèlement, le chercheur n'est pas non plus limité en nombre puisqu'il peut multiplier les points de dépôts. Cependant, bon nombre d'institutions et de personnels de l'IST font le choix



Projet de loi pour une république numérique

de HAL en tant que support de dépôt. Rappelons que la loi donne au chercheur un droit mais pas une obligation de diffusion. Cela a de l'importance sur l'effectivité du texte : la loi donne aux chercheurs le pouvoir de faire progresser ou non l'*Open Access*. C'est ainsi que les Allemands qui avaient opté pour ce même système, à la suite d'une étude sur les impacts de cette loi, ont réalisé qu'un an après la mise en application du texte législatif, les dépôts n'avaient progressé que de 10 %.

UNE LOI ADOPTÉE, ET APRÈS ?

Pour certains partisans de l'*Open Access*, cette loi ne va pas assez loin et devrait contenir une obligation de dépôt en *Open Access*. Pour les éditeurs en revanche, elle est source d'inquiétudes et il subsiste des doutes sur le périmètre de son application.

Vers une obligation de dépôt en *Open Access* ?

L'article 30 de la loi rappelle que « [l']auteur dispose [...] du droit de mettre à disposition gratuitement dans un format ouvert » sa publication. C'est donc un droit que l'auteur peut choisir d'exercer ou non.

Cependant, si le dispositif législatif ne s'impose pas aux chercheurs, il est intéressant de noter que **certaines institutions obligent les chercheurs rattachés à leurs établissements à assurer un libre accès à leurs publications** : c'est le cas pour les chercheurs de l'INRIA, qui doivent déposer leurs publications dans HAL. Ces mêmes politiques existent également à l'échelle des laboratoires de recherche, mais aussi pour les universités : c'est le cas en France pour l'université d'Angers qui utilise le mandat de dépôt obligatoire.

Enfin, **certaines programmes de financement obligent les chercheurs à publier leurs articles en libre accès en contrepartie de l'aide apportée** : c'est le cas des projets financés au niveau européen par H2020. Cette politique est très développée aux Etats-Unis où les agences de financement conditionnent l'attribution de leurs aides à une publication des résultats en *Open Access*. Pourrait-on imaginer un système équivalent au niveau de l'ANR ? En effet, les incitations se font de plus en plus fortes pour assurer un libre accès aux résultats de la recherche, sans qu'il n'y ait pour autant une obligation à proprement parler.

Finalement la loi a levé le blocage légal ce qui a permis à certaines institutions de mettre en place des politiques incitatives voire même d'établir des obligations au dépôt en *Open Access*.

Inquiétudes des éditeurs face à la loi

L'entrée en vigueur de cette loi remet en cause le modèle économique établi par les éditeurs scientifiques. Selon une étude britannique réalisée auprès des bibliothèques, la majorité d'entre elles résilieraient leurs abonnements aux revues payantes si les articles scientifiques étaient disponibles gratuitement moins de 6 mois après leur publication dans des revues.

Face à ces difficultés, un grand nombre de revues, les plus petites, seraient condamnées. Il ne resterait alors que les grandes maisons d'édition déjà très favorisées. Celles-ci se retrouveraient ainsi en position de quasi-monopole.

Il faut bien comprendre que la fronde des chercheurs à l'égard des politiques éditoriales était menée principalement contre les grandes maisons d'éditions telles que Elsevier, Springer ou encore Wiley qui à elles seules, possèdent près de 6400 revues scientifiques et dont les coûts d'abonnement augmentent de façon continue. **Mais les grands éditeurs commerciaux avaient anticipé ce changement et tentent depuis les années 2000 de subvertir la dynamique de l'Open Access à leur avantage en développant le modèle économique auteur-payeur.** Pour une analyse détaillée voir l'article suivant qui passe au crible le modèle auteur-payeur⁴.

Des zones d'ombre au niveau de l'application

Un certain nombre d'aspects restent encore à clarifier. En effet, alors qu'aucun décret n'est encore paru, **la question de la rétroactivité fait couler beaucoup d'encre.** Les chercheurs ayant publiés avant la promulgation de cette loi se demandent s'ils peuvent exercer leur droit d'exploitation secondaire ? Il en est de même quant à la problématique de la **portée de cette loi face aux éditeurs d'autres pays.** La communauté scientifique s'interroge sur la possibilité qu'offre cette loi de publier en *Open Access* lorsqu'ils ont publié des écrits auprès de maisons d'édition étrangères. La réponse à ces questions sera apportée par la jurisprudence. Cependant, ce cas de figure pourrait avoir lieu uniquement si un éditeur décidait d'attaquer un chercheur ayant diffusé un article en *Open Access*. Or, les éditeurs n'iront pas sur ce terrain qui se révélerait être un véritable désastre en terme d'image.

Dans la pratique, les auteurs d'écrits scientifiques prennent le parti de faire usage de leur droit rétroactivement, c'est ainsi que, notamment sur la

plateforme HAL, de nombreux dépôts d'articles anciens ont été faits. Il en va de même pour les articles issus de publications effectuées par des maisons d'édition étrangères. Pour Couperin « *certaines juristes considèrent [que la loi] s'applique aux articles publiés antérieurement [avant sa promulgation]. Le Conseil scientifique du CNRS a publié une recommandation en ce sens*⁵ ».

CONCLUSION

La loi pour une République numérique, à travers l'article 30, cherche à promouvoir un nouvel équilibre entre le chercheur et son éditeur. Même si des incertitudes demeurent, **l'article 30 donne les moyens à la communauté scientifique de se réapproprier sa production et d'en faciliter la diffusion.** Récemment, le 4 juillet dernier, le *Plan national pour une Science Ouverte* propose d'aller plus loin que la *loi pour une République numérique* et de passer d'une incitation à une obligation de publication en accès ouvert des articles et des livres « *issus de recherches financées par appel d'offres sur fonds publics* ».

■ Laure PIERRAT

¹ Pour donner un ordre d'idée, Richard Monvoisin via son blog cortecs.org donne un ordre de prix pour la consultation d'un article dans une revue : « au hasard, l'article *The DRESS syndrom : litterature review*, de Cacoub & al, dans *The American Journal of Medicine Volume 124, Issue 7, juillet 2011, pages 588–597*, il nous en coutera la bagatelle de 31,50 dollars US. »

² Extrait de l'article 30 de la loi pour la République numérique

³ Cette nuance se révèle être d'autant plus importante pour la publication des actes de colloques. En effet, s'ils sont publiés dans une monographie : l'auteur ne bénéficiera pas de ce droit, a contrario, s'ils sont publiés dans une revue le chercheur pourra à l'issue de la période d'embargo diffuser sa publication en *Open Access*.

⁴ Pour une analyse plus en détail, voir dans ce numéro l'article de Sonia SALAMI « *Gold Open Access : quel modèle économique pour demain?* »

⁵ Couperin [en ligne], disponible sur le <https://openaccess.couperin.org/la-loi-numerique/>, consulté le 9 mai 2018.

BIBLIOGRAPHIE

BENJAMIN Anna, « Harvard rejoint les universitaires pour un boycott des éditeurs », *Le Monde* [en ligne], publié le 25 avril 2012. URL : https://www.lemonde.fr/sciences/article/2012/04/25/harvard-rejoint-les-universitaires-pour-un-boycott-des-editeurs_1691125_1650684.html

BERNAULT Carine, « Revues scientifiques et droit d'auteur : la rupture de l'open access », *Hermès, La Revue- Cognition, communication, politique*, CNRS-Editions, 2015, Le XXe siècle saisi par la communication, Vol. 2 : Ruptures et filiations (71), p. 92. URL : <https://halshs.archives-ouvertes.fr/halshs-01160567/document>

CLAVEY Martin, « Un hashtag clandestin pour partager la science inaccessible », *Rue89* [en ligne], publié le

8 septembre 2015, URL : <https://www.nouvelobs.com/rue89/rue89-sciences/20150908.RUE0500/un-hashtag-clandestin-pour-partager-la-science-inaccessible.html>

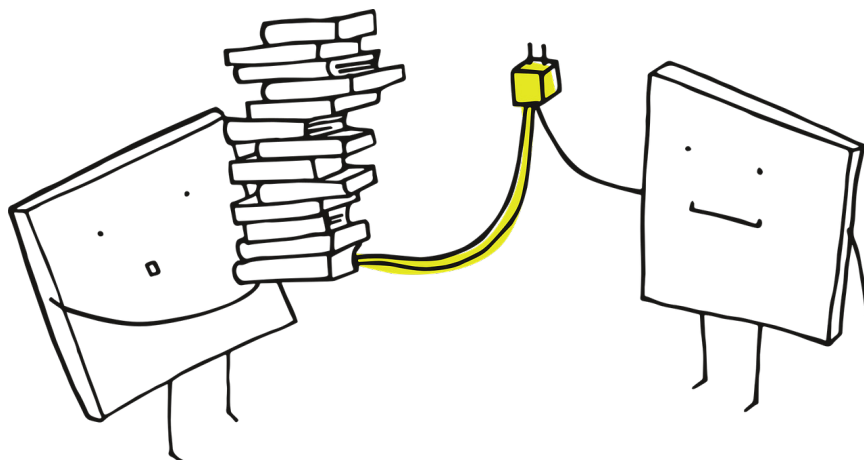
COMMISSION EUROPEENNE, « Recommandation de la Commission du 17.7.2012 relative à l'accès aux informations scientifiques et à leur conservation (C(2012) 4890) », publié le 17 juillet 2012. URL : https://ec.europa.eu/research/science-society/document_library/pdf_06/recommendation-access-and-preservation-scientific-information_fr.pdf

DIRECTION DE L'INFORMATION SCIENTIFIQUE ET TECHNIQUE – CNRS, *Livre blanc — Une Science ouverte dans une République numérique*. Nouvelle édition [en ligne]. Marseille : OpenEdition Press, 2016. URL : <http://books.openedition.org/oep/1548>

MOATTI Alexandre, « Du rifici dans l'édition », publié sur bibnum.hypotheses.org [en ligne] le 15 juillet 2013.

MONVOISIN Richard, « Le coût de la connaissance – Boycott d'Elsevier », publié sur Cortecs.org le 22 juillet 2013. URL : <https://cortecs.org/publications/le-cout-de-la-connaissance-boycott-delsevier/>

TAMBOU Olivia, « L'article 30 de la Loi République numérique vu par le député Emeric Brehier », publié sur blogdroiteuropeen.com [en ligne], le 17 janvier 2017. URL : <https://blogdroiteuropeen.com/2017/01/17/larticle-30-de-la-loi-republique-numerique-vu-par-le-depute-emerich-brehier/>



GOLD OPEN ACCESS: QUEL MODÈLE ÉCONOMIQUE POUR DEMAIN ?

Sonia SALAMI



Gold Open Access (CC0) © C.Picque

La prise en compte de l'aspect économique de l'Open Access dans les politiques des états prend de plus en plus d'importance. Depuis quelques années les rapports, les chiffrages et les initiatives en faveur de la généralisation du modèle auteur-payeur se multiplient car celui-ci serait l'alternative idéale au modèle économique de l'abonnement qui prédomine actuellement. Qu'en est-il réellement, **va-t-on vers une généralisation du modèle auteur-payeur ?** Le modèle auteur-payeur est-il plus vertueux que le modèle de l'abonnement? Sera-t-il le modèle dominant de l'édition scientifique commerciale de demain ?

LA NAISSANCE DU MODÈLE AUTEUR-PAYEUR

Le modèle économique de l'édition scientifique numérique le plus répandu est celui de l'abonnement (modèle lecteur-payeur), son accès n'est donc pas libre. Avec ce modèle les grands éditeurs commerciaux font des profits colossaux qui ne se justifient nullement par les services qu'ils apportent à la communauté scientifique.

En effet, **une fois passée la transition numérique, le prix des abonnements aurait dû diminuer** en raison des économies réalisées par les éditeurs (il n'y a plus d'impression ou de diffusion de revues papier) mais leur coût n'a fait qu'augmenter (depuis les années 90 + 290 % !). Ce prix est d'autant plus injustifié que les revues scientifiques vendues par les éditeurs reposent en grande partie sur le travail d'enseignants chercheurs que les éditeurs ne rémunèrent pas. En effet, les chercheurs qui sont rémunérés par les institutions publiques pour leurs travaux, produisent les articles que les éditeurs vendent, assurent également la qualité scientifique des revues en évaluant les articles (évaluation par les pairs) et vont parfois jusqu'à assurer la mise en page et la mise en ligne, le tout, sans être rémunérés.

En réaction à cette situation absurde où les éditeurs font des profits colossaux et où les institutions paient deux fois (une première fois pour que les chercheurs fassent leurs recherches et une seconde fois pour avoir accès aux résultats) **s'est développé le libre accès ou Open Access**. La déclaration de Budapest en 2002 définit deux modèles pour l'Open Access, la voie verte et la voie dorée. Ces deux modèles sont conçus comme étant complémentaires. Si la **voie verte** correspond au dépôt dans des archives ouvertes institutionnelles des productions scientifiques, la **voie dorée** désigne toutes les revues ou tous les ouvrages publiés nativement en libre accès. **Face au développement de l'Open Access, les éditeurs se sont mis eux-aussi à proposer un modèle économique en libre accès : le modèle auteur-payeur.**

On considère que la voie dorée a une vraie valeur ajoutée par rapport à la voie verte car :

- elle donne un accès libre aux articles sans période d'embargo,
- les articles sont bien référencés,
- la version fournie est la version finale qui a été mise en forme (travail éditorial),
- les articles ont été relus, corrigés et validés par des pairs¹.

LES MODÈLES ÉCONOMIQUES DE LA VOIE DORÉE

Contrairement à ce qu'on pourrait penser, le fait qu'un document soit en libre accès n'est pas synonyme de gratuité et ne signifie pas que le document soit libre de droit. **Le libre accès a pour effet de rendre invisible pour le lecteur les différents modèles économiques (parfois très lucratifs!) qui sont à l'œuvre.** Ainsi le terme de voie dorée ou de Gold Open Access recouvre des modèles économiques

très divers et le modèle auteur-payeur proposé par les éditeurs n'est qu'une de ses modalités.

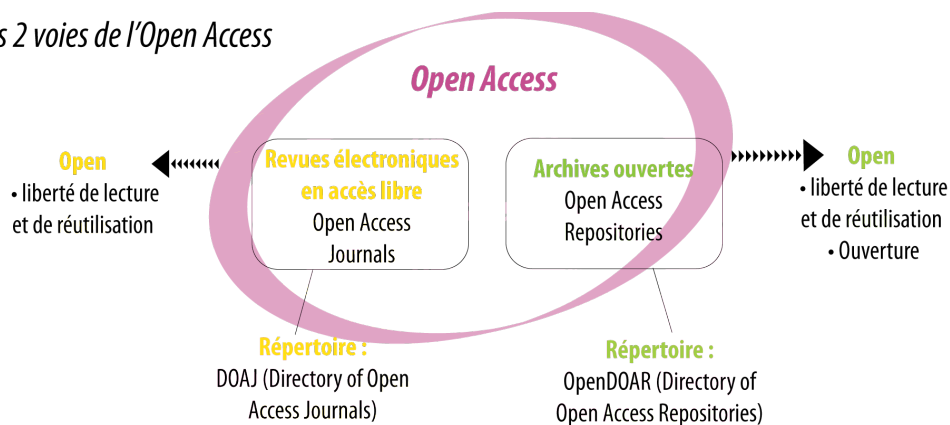
Les modèles économiques de la voie dorée sont principalement² :

- **Le sponsor** où la revue est financée par un organisme à but non lucratif (société savante, institution, organisme de recherche, association ou fondation).
- **Le freemium** où la revue en Open Access est financée par des services payants.
- **Le modèle auteur-payeur** où l'auteur (la plupart du temps l'établissement, le laboratoire ou le bailleur de fonds) paye des frais de publication ou APC (*Articles Processing Charges*) pour que les articles soient publiés en Open Access.
- **Le modèle hybride** qui est un mélange entre l'abonnement et le modèle auteur-payeur. Dans ce modèle, les revues sont accessibles par abonnement mais les auteurs ont la possibilité de payer en plus des APC pour que leur article soit en libre accès³.

POURQUOI BASCULER DANS UN MODÈLE ÉCONOMIQUE AUTEUR-PAYEUR ?

En 2014, 13 % des articles mondiaux, soit environ 260 000 articles, auraient été publiés en Open Access doré selon l'étude commanditée par la Commission européenne (Archambault & all, 2014). Malgré de nombreuses initiatives en faveur de l'Open Access (voie verte et voie dorée) et la progression de celui-ci, le modèle économique de l'abonnement reste prépondérant. Face à l'augmentation régulière du prix des abonnements et face à une

Les 2 voies de l'Open Access



Les 2 voies de l'Open Access, inspiré du schéma de Bibliothèque de l'Université Pierre et Marie Curie» © C.Picque

baisse généralisée des dotations, les bibliothèques ont de plus en plus de mal à financer l'accès aux revues.

Pour répondre à cette situation, en juin 2012 le rapport **Finch Accessibility, sustainability, excellence : how to expand access to research publications**, recommande au gouvernement du Royaume-Uni d'encourager la voie dorée, et plus particulièrement le modèle auteur-payeur et le modèle hybride, afin de développer l'Open Access. Dans ce rapport, la voie verte est limitée à une fonction de conservation et non de diffusion des publications scientifiques, et devrait surtout servir à la publication de la littérature grise et les données de la recherche. Ce rapport estime que lors de la transition entre le système d'abonnements et le modèle auteur-payeur, il y aura des surcoûts. Suite à ce rapport, le Royaume-Uni a donné la priorité à la voie dorée.

Plus tard, le rapport *Disrupting the subscription journals' business model for the necessary large-scale transformation to Open Access* réalisé par la Max Planck Digital Library après avoir réalisé une analyse chiffrée, estime que le budget des institutions aujourd'hui dédié aux abonnements est suffisant pour financer un modèle entièrement en Open Access sur le mo-

dèle voie dorée auteur-payeur. Avec la généralisation du modèle auteur-payeur, les fonds des bibliothèques jusqu'à présent alloués à l'achat des abonnements deviendraient un budget servant à financer les publications. Selon ce rapport, la généralisation d'un tel modèle permettrait d'éradiquer le modèle délétère de l'abonnement et de rendre accessible toutes les publications en Open Access sans durée d'embargo (contrairement à la voie verte). **La généralisation du modèle auteur-payeur n'est donc pas une stratégie visant à combattre et à écarter les éditeurs commerciaux mais au contraire une démarche qui cherche à les inclure pour atteindre l'objectif du 100 % Open Access.** La recherche étant internationale, un tel basculement de modèle économique ne pourrait se faire qu'à la condition que les organismes de recherche se coordonnent à l'échelle mondiale.

LES POLITIQUES EN EUROPE FACE AU DÉVELOPPEMENT DU MODÈLE AUTEUR-PAYEUR

Compte tenu du fait que les APC se généralisent, la nécessité de les prendre en compte s'est faite jour. En effet, selon le DOAJ (*Directory of Open Access Journals*) un tiers des revues en Open Access s'appuient mainte-

nant sur le modèle auteur-payeur via le paiement d'APC et de plus en plus de revues fonctionnant sur le modèle de l'abonnement basculent dans un modèle hybride où il est possible, via le règlement d'APC, de diffuser des articles en Open Access.

Ainsi, **évaluer le montant des APC, quelle que soit la politique choisie par les différents états est devenue un enjeu important.** C'est ainsi que le Royaume-Uni a été développé l'outil **Monitor UK** qui centralise les dépenses en APC des différentes institutions. Au niveau de l'Union Européenne, il y a le **projet OpenAPC** et en France le **projet Conditor** de la *Bibliothèque Scientifique Numérique (BSN)* qui cherche entre autre à évaluer les coûts des APC. **Un rapport de 2014 de l'INRA** cherche à estimer le coût d'un basculement de l'institution dans un modèle de publication Open Access auteur-payeur (Archambault & all, 2014) et plus récemment le **consortium Couperin a lancé une enquête nationale** pour évaluer les coûts d'un tel basculement au niveau national.

La douzième conférence de Berlin qui s'est tenue en décembre 2015 se fonde sur le rapport de la Max Planck Digital Library pour préconiser également la généralisation du modèle auteur-payeur. C'est un virage important qui a des conséquences au niveau des différentes stratégies des pays européens. Aujourd'hui, non seulement le Royaume-Uni, mais aussi l'Allemagne, la Finlande et les Pays Bas ont opté pour une politique en faveur de l'Open Access doré qui inclut les éditeurs commerciaux. Ces États favorisent la négociation de contrats fixant pour chaque établissement un coût global permettant à la fois l'accès au contenu des revues et la diffusion des publications de ses chercheurs, c'est-à-dire incluant le montant des

APC (Delhaye et Lutz, 2017).

En France, au contraire, il y a une opposition entre éditeurs privés et institutions publiques et la politique de l'Open Access n'est pas en faveur du modèle auteur-payeur. Celle-ci favorise la voie verte avec le dépôt dans les archives ouvertes⁴ et pour la voie dorée vise plutôt à soutenir et développer une édition publique dirigée par les bibliothèques et les presses universitaires (Chartron, 2016) et encourage par ailleurs le développement du modèle freemium.

LA GÉNÉRALISATION DU MODÈLE AUTEUR-PAYEUR UNE FAUSSE BONNE IDÉE ?

L'idée de réaffecter les fonds actuellement dédiés aux abonnements pour financer à la place des APC qui permettent de mettre les articles en libre accès paraît séduisante. Cependant la généralisation d'un tel modèle économique au niveau mondial n'irait pas sans produire des effets pervers (Voir « *Déclaration des trois sociétés savantes françaises de mathématiques...* » 2012).

Les principaux risques sont :

- Tout d'abord, **le modèle auteur-payeur n'est pas un modèle généralisable de la même manière dans tous les secteurs scientifiques** et en particulier, il n'est pas adapté pour les SHS (Sciences Humaines et Sociales).
- Ensuite, **la transition d'un modèle où il faut payer pour lire les articles (abonnement) à un système où il faut payer pour publier (auteur-payeur) remplace des inégalités d'accès par des inégalités de publication** entre pays riches et pays pauvres, entre établissements, entre

différents laboratoires d'un même établissement. La généralisation de la concurrence pour publier existera même au sein d'un même laboratoire, car les budgets étant limités, il faudra choisir quel chercheur pourra publier et quel chercheur ne pourra pas. Enfin, il sera presque impossible à un chercheur n'appartenant à aucun laboratoire de publier. Il y existera donc une véritable censure économique à la source.

- Enfin, **cela ne réglerait pas le problème des revues prédatrices renforcés par l'édition scientifique avec le modèle auteur-payeur et pourrait même le renforcer.** En effet, dans ce modèle économique, plus une revue publie d'articles, plus elle gagne de l'argent grâce aux APC, ainsi on peut craindre une diminution du contrôle éditeurs qui chercheront à publier le plus possible pour gagner davantage d'argent.



Gold Open Access (CC0) © C.Picque

Les auteurs Christian Fuchs et Marisol Sandoval, ainsi que le directeur scientifique du Réseau National des Bibliothèques de Mathématique, *Frédéric Hélein*, questionnent la convergence de différents rapports en faveur de la généralisation du modèle auteur-payeur. En effet, **ce modèle avantage beaucoup les grands éditeurs commerciaux (au détriment des autres)**. Cette convergence semble être le résultat du travail de lobbying mené par les grands éditeurs et un lien entre la Max Planck Gessellschaft et Spinger-Nature questionne l'objectivité du rapport qu'elle a produit et qui a

été suivi notamment à la conférence Berlin 12 : « *En fin de compte on peut se demander dans quelle mesure les orientations de la Max Planck Gesellschaft sont indépendantes des intérêts de groupes comme Springer-Nature, sachant qu'elle est associée à la Max Planck Förderstiftung, fondation dotée d'un capital de 500 millions d'euros, fondée et présidée par Stefan van Holtzbrinck, propriétaire de 53 % du capital de Springer-Macmillan (qui a racheté Nature) et co-dirigée par le président de la Max Planck Gesellschaft* ».

Lise Verlaet, enseignante-chercheuse au LERASS CERIC, voit dans le développement du modèle auteur-payeur **une véritable revanche des éditeurs face au mouvement du libre accès qui se développait et parle à ce sujet de « détournement du modèle du Free Gold Open Access ou libre accès institutionnel »**. Le modèle auteur-payeur ne cesse de se développer et, selon le rapport de de la DIST / CNRS 2015, d'ici une dizaine d'années, il pourrait supplanter le modèle de l'abonnement.

QUELLES ALTERNATIVES ?

Le modèle auteur-payeur n'est donc pas un modèle idéal et il semble même urgent **d'arrêter de soutenir indirectement l'hégémonie des grands éditeurs commerciaux, notamment en modifiant le mode d'évaluation des chercheurs.**

En effet, l'évaluation des chercheurs repose en partie sur le nombre d'articles publiés et leur facteur d'impact or **le facteur d'impact** n'est en rien un critère pour évaluer la pertinence scientifique d'une publication (Guédon, 2013), et celui-ci encourage les chercheurs à publier dans les revues avec un facteur d'impact élevé, or les

classements des revues en fonction de leur impact favorisent les grands éditeurs commerciaux qui appuient leur modèle économique sur les abonnements, les APC et le modèle hybride. De plus, le facteur d'impact ne prend pas en compte la qualité des revues (Fuchs et Sandoval, 2013) et ne protège donc pas des revues prédatrices. L'Université de Liège a trouvé une alternative : l'évaluation de ses chercheurs ne se fait que sur les publications déposées dans son archive ouverte institutionnelle.

Il ne faut pas oublier qu'il **existe d'autres choix pour la voie dorée** en dehors du modèle auteur-payeur ou hybride notamment avec le **freemium** (modèle encouragé par la France avec *OpenEditions*) qui permet à la fois un accès gratuit aux publications qui sont publiées nativement en libre accès et une gratuité pour les auteurs. Au-delà du clivage voie verte/voie dorée, on voit également se développer les **épi-revues** comme *episcience.org*. Les épi-revues sont des revues électroniques en libre accès réalisées à moindre coût à partir d'articles publiés dans des archives ouvertes (appartenant à la voie verte) en ajoutant une évaluation par les pairs et une éditorialisation (propres à la voie dorée)⁶.

Tout cela montre que nous sommes face à une vraie **diversification des modèles d'Open Access**, et que le modèle dit de la voie dorée tel qu'il a été défini au départ recouvre à présent des modèles très différents et n'apparaît plus très pertinent. Pour cette raison, Christian Fuchs et Marisol Sandoval dans leur article « *The Diamond Model of Open Access Publishing* » proposent de redéfinir la voie dorée comme étant limitée au modèle auteur-payeur et d'exclure le modèle hybride qui pour eux ne relève pas de la voie dorée.

La voie dorée serait alors à différencier d'une troisième voie plus vertueuse : **la voie diamantée, ou voie platine** en fonction des auteurs. La « *voie platinum* » ou platine, « *vise une gratuité au lecteur et à l'auteur. Elle a été introduite aux États-Unis en 2011 par Crawford pour se distinguer de la voie dorée fondée sur les APC* » (Chartron 2016). Les termes de voie diamantée ou de voie platine sont parfois assimilés ou distingués. Par exemple pour certains, les épi-revues relèveraient de la voie diamantée et les freemium de la voie platine.

La voie diamantée désignerait les revues en libre accès publiées par des organisations à but non lucratif, avec une gratuité pour les auteurs et pour les lecteurs et de meilleures conditions pour les salariés⁷.

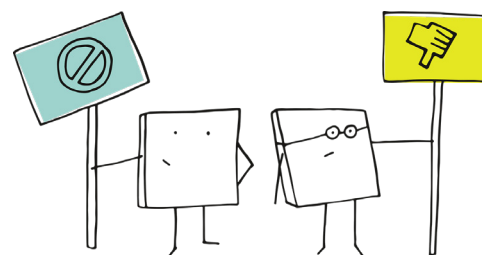
Jean-Claude Guédon va dans le même sens. L'édition scientifique est déjà financée par des fonds publics et fait des profits très importants. Il propose donc de **se réapproprier la diffusion en généralisant la voie dorée nationale** c'est à dire financée par les états ce qui aboutirait à une disparition des éditeurs privés actuels et permettrait une gratuité pour les auteurs et les lecteurs.

CONCLUSION

En conclusion, nous avons montré qu'une généralisation du modèle auteur-payeur pour la voie dorée n'est en rien une solution miracle pour lutter contre le coût des abonnements, ni pour remédier aux problèmes rencontrés actuellement par la voie dorée par rapport à la fiabilité des évaluations et aux revues prédatrices.

Le modèle-auteur payeur semble plutôt être une stratégie des éditeurs pour reprendre la main face au mouvement de l'*Open Access* qui les menace. **Le partage et l'opposition entre voie verte et voie dorée semble avoir perdu en pertinence depuis la déclaration de Budapest de 2002 et le libre accès se cristallise plutôt en trois pôles : voie verte, voie dorée et voie diamantée ou platine.** En effet, il semble que nous soyons dans une **période de transition où les modèles économiques de l'Open Access se diversifient**, de nouveaux modèles font leur apparition. Les stratégies en faveur de l'*Open Access* varient selon les États et il est encore trop tôt pour savoir quel sera ou quels seront les modèles économiques qui domineront demain. Cependant, un changement majeur est peut-être sur le point de s'opérer, le *Plan National pour une Science Ouverte* du 4 juillet 2018, s'inspire des préconisations européennes et propose de rendre obligatoire la publication en *Open Access* « *des articles et livres issus de recherches financées par appel d'offres sur fonds publics* », ce qui accélérerait le déclin du modèle de l'abonnement – surtout si l'initiative est suivie par d'autres pays. Cependant, on note qu'**une telle mesure ne menace en rien le modèle auteur-payeur qui a encore de beaux jours devant lui !** Même si ce plan préconise l'exploration de nouveaux modèles économiques de publication en *Open Access*...

■ Sonia SALAMI



1 La validation par les pairs permet dans une revue scientifique de sélectionner le contenu. Le but est de vérifier la rigueur de la méthodologie et d'évaluer la validité des résultats présentés. Dans les archives ouvertes (voie verte) il n'y a pas de contrôle. Définition du CNRS : « l'évaluation par les pairs désigne la validation d'un article par un comité de lecture composé de scientifiques, experts dans le même champ disciplinaire que le contenu de l'article. Ce processus est destiné à lui assurer une qualité scientifique. » (voir <https://lab.hypotheses.org/1453>)

2 Voir l'article « Modèles économiques du Gold » disponible sur le site dédié à l'Open Access du Le consortium Couperin (Consortium unifié des établissements universitaires et de recherche pour l'accès aux publications numériques) : <http://openaccess.couperin.org/modeles-eco-du-gold/>

3 Avec ce modèle les institutions paient deux fois : une fois via un abonnement pour avoir accès aux articles et une seconde fois pour diffuser les articles en Open Access.

4 Voir le développement de HAL comme archive nationale liée aux archives institutionnelles (voir Cédric Marion et Carole-Anne Matel "le CCSD : moteur de l'Open Access") et plus récemment la Loi pour une République numérique en 2016 (voir dans ce numéro l'article de Laure PIERRAT « Loi pour une République numérique éclairage sur l'article 30 »).

5 Voir en particulier : « Publications, Open Access, etc. : une liste de liens utiles avec quelques commentaires critiques » partie « Une fausse bonne idée qui pourrait s'avérer catastrophique »

6 Pour en savoir plus sur episcience.org, voir dans ce numéro l'article de Sonia SALAMI, Cédric MARION et Constance PIQUE « Le CCSD : moteur de l'Open Access en France ».

7 « In theDiamond Open Access Model, not-for-profit, non-commercial organizations, associations or networks publish material that is made available online in digital format, is free of charge for readers and authors and does not allow commercial and for-profit re-use. » (Caraco, 2014)

BIBLIOGRAPHIE

Articles sur les modèles économiques de la voie dorée :

CHARTRON Ghislaine, « Stratégies, politiques et reformulations de l'Open Access », *Revue française des sciences de l'information et de la communication* [En ligne], 8 | 2016, mis en ligne le 24 mars 2016. URL : <http://journals.openedition.org/rfsic/1836>

CHARTRON Ghislaine, « Open Access : Origine, enjeux, actualités : présentation au CNAM », 11 février 2013, URL : <http://www.openaccess-shs.info/wpcontent/uploads/2012/12/Pr%C3%A9sentation-de-Ghislaine-Chartron-11-f%C3%A9vrier.pdf>

CONSORTIUM COUPERIN, « Modèles économiques du Gold », site COUPERIN, URL : <http://openaccess.couperin.org/modeles-eco-du-gold/>

CONTAT Odile « Publications scientifiques et libre accès : une mise au point sur le Gold Open Access » dans Ressources pour les correspondants IST de l'InSHS : http://corist-shs.cnrs.fr/gold_open_access

DIRECTION DE L'INFORMATION SCIENTIFIQUE ET TECHNIQUE – CNRS, *La publication scientifique aujourd'hui : financement, usages numériques*, étude, 04/06/2015 URL : http://www.cnrs.fr/dist/z-outils/documents/Distinfo2/DISTetude_4.pdf

DELHAYE Marlène, LUTZ Jean-François, « Mesurer les dépenses d'APC : méthodologie et étude de cas », *Revue française des sciences de l'information et de la communication*, 11 | 2017. URL : <https://journals.openedition.org/rfsic/3238>

VERLAET Lise, « Modèles d'affaires de l'Open Access. Réflexions autour du projet Numerev », *Les Cahiers de la SFSIC, Société française des sciences de l'information et de la communication*, 2017, p.39-50. URL : <https://hal.archives-ouvertes.fr/hal-01676350/document>

Les dangers du modèle auteur-payeur et de sa généralisation :

BÊCHE Vincent, « Les revues "prédatrices" un danger pour les

chercheurs ! », *openarchiv.hypotheses.org* 09/03/2014. URL : <https://openarchiv.hypotheses.org/2044>

CHILLIARD Yves, COCHARD Hervé, MOULIA Bruno, BADEL Eric, DURAND Jean Louis, FONTAINE Sébastien, GIROUSSE Christine, FORTERRE Yoël, POULIQUEN Olivier, FOURNIER Meriem, « La science menacée par une bulle spéculative de l'édition », dans *lemonde.fr*, 11/11/2013. URL : https://www.lemonde.fr/sciences/article/2013/11/11/la-science-menacee-par-une-bulle-speculative-de-l-edition_3511861_1650684.html

CNRS, « L'avis du Comité d'éthique du CNRS du 29 juin 2012 sur Le libre accès aux publications scientifiques ("Open Access") », « Lettre du COMETS n° 3 », 2013. URL : http://www.cnrs.fr/comets/IMG/pdf/lettre_comets_no3.pdf

DELHAYE Marlène, « Le rapport finch », *marlenescorner.net*, 31/08/2012. URL : <https://marlenescorner.net/2012/08/31/le-rapport-finch/>

GUEDON Jean Claude, « Avenir de l'Open Access avec le Gold », conférence prononcée à la journée *Open Access* Couperin, 2015. URL : https://webcast.in2p3.fr/video/avenir_de_l_open_access_avec_le_gold

HENRI Agnès, VAN TIGGELEN Bart, LERICHE Marie-Anne « réflexion sur la voie dorée des publications scientifiques » *Reflets de la physique* n°30, 2012. URL : http://lpm2c.grenoble.cnrs.fr/UserFiles/File/Reflets%2030_publications%20scientifiques.pdf

HUET Sylvestre « *Open access* du rêve au cauchemar », *sciences.blogs.liberation.fr*, 04/10/2013 URL : <http://sciences.blogs.liberation.fr/2013/10/04/open-access-du-reve-au-cauchemar/>

HUET Sylvestre « *Open access*, du rêve au cauchemar : l'avis de J.C. Guédon » sur 04/10/ 2013 URL : <http://sciences.blogs.liberation.fr/2013/10/04/open-access-du-reve-au-cauchemar-bis/>

SOCIETE FRANÇAISE DE STATISTIQUE (SFDS), SOCIETE DE MATHEMATIQUES APPLIQUEES ET INDUSTRIELLES (SMAI), SOCIETE MATHEMATIQUE DE FRANCE (SMF), *Déclaration des trois sociétés savantes françaises de mathématiques, Open Access : mise en garde et effets pervers du système auteur-payeur, Seconde version*, septembre 2012. URL : http://smf.emath.fr/files/open_access_trois_societes_sept.pdf

UNIVERSITE RENNE 1, « Les revues hybrides », site de l'université de Renne 1. URL : <https://openaccess.univ-rennes1.fr/les-revues-hybrides>

Les rapports en faveur d'une généralisation du modèle économique auteur payeur :

ANDRO Mathieu et al., Estimation des dépenses de publication de l'Inra dans un modèle théorique « *Gold Open*

Access », *Documentaliste-Sciences de l'Information* 2014/4 (Vol. 51), p. 70-79. URL : <https://hal.archives-ouvertes.fr/hal-01097171/document>

FINCH Janet et al., *Accessibility, sustainability, excellence: how to expand access to research publications* [en ligne], juin 2012. <<http://www.researchinfonet.org/wp-content/uploads/2012/06/Finch-Group-report-FINAL-VERSION.pdf> [Traduction et résumé du rapport Finch, URL : <http://openaccess.inist.fr/?Accessibilite-durabilite>]

HOUGHTON Jonh et SWAN Alma « Going for Gold? The costs and benefits of Gold Open Access for UK research institutions: further economic modelling Report to the UK *Open Access* Implementation Group June 2012 » URL : http://repository.jisc.ac.uk/610/2/Modelling_Gold_Open_Access_for_institutions_-_final_draft3.pdf

MAX PLANCK DIGITAL LIBRARY, « Area-wide transition to open access is possible, A new study calculates a redeployment of funds in Open Access » article publié le 27 avril 2015. URL : <https://www.mpg.de/9202262/area-wide-transition-open-access>

SCHIMMER Ralf, GESCHUHN Kai Karin, VOGLER Andreas, « Disrupting the subscription journals' business model for the necessary large-scale transformation to Open Access A Max Planck Digital Library Open Access Policy White Paper ». URL : https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_2148961

Autres alternatives

ALBERGANTI, M, 2013. « Economie de la publication scientifique et libre accès: un débat relancé par la mort d'Aaron Swartz », *Slate.fr* 21/01/2013. URL : <http://www.slate.fr/story/67263/suicide-aaron-swartz-economie-publication-scientifique-libre-acces>

CARACO Benjamin « La voie diamanté du libre access », *hypothese.org*, 27/01/2014 URL : <https://oadesk.hypotheses.org/298>

CHAFAI Djalil « Coût des publications : un exemple instructif », *Gazette des maths*. 2012 (version modifié du 4 novembre 2013) <http://smf.emath.fr/files/cout-des-publications-propositions-concretes.pdf>

DACOS Marin, conférence « Autres modèles de Libre Accès » à la journée *Open Access* Couperin 2015 URL : https://webcast.in2p3.fr/video/autres_modeles_de_libre_acces

FUCHS Christian et SANDOVAL Marisol « The Diamond Model of Open Access Publishing: Why Policy Makers, Scholars, Universities, Libraries, Labour Unions and the Publishing World Need to Take Non-Commercial, Non-Profit Open Access Serious », *Triple C : Communication, Capitalism & Critique*, vol 11, n°12, 2013 URL : <https://www.triple-c.at/index.php/tripleC/article/view/502>

« Quels sont les différents modèles de l'open access » : <https://sharelex.org/t/quels-sont-les-differents-modeles-de-lopen-access/170>

SILVY Christine, « De l'Open Access à l'Open Data : enjeux et perspectives. » Séminaire CBGP, 6 janvier 2015. URL : <https://www.fosteropenscience.eu/sites/default/files/pdf/5219.pdf>

Le facteur d'impact

Sonia SALAMI

BIBLIOGRAPHIE

CoopIST-CIRAD « Le facteur d'impact et ses indicateurs associés pour évaluer une revue » URL : <https://coop-ist.cirad.fr>

LARIVIERE Vincent, « Facteur d'impact: définition du concept » vidéo de la chaîne youtube du CIRST (Centre interuniversitaire de recherche sur la science et la technologie) appartenant à l'Université de Montréal. URL : <https://www.youtube.com/watch?v=7GeBlkwICzs>

En 1960, le chercheur en science de l'information Eugène Garfield crée le facteur d'impact qui est un indicateur bibliométrique censé évaluer la notoriété d'une revue. Pour une **année N**, le **facteur d'impact (FI)** se calcule en faisant le ratio entre le **nombre de fois où des articles de la revue ont été publiés au cours des deux années précédentes (N-1 et N-2) sont cités au cours d'une année N et le nombre d'articles que cette revue a publié au cours des deux années précédentes (N-1 et N-2)**. Le nombre de citations est calculé selon la base de données bibliographique *Web of Science Core Collection* (WoS) qui compte environ 12 000 revues :

$$\text{FI (année N)} = \frac{\text{nombre de citations des articles de la revue publiés en N - 1 et N - 2 faites par les articles des revues du WoS en année N}}{\text{nombre d'articles de la revue publiés en N - 1 et N - 2}}$$

Comme l'explique Vincent Larivière, chercheur au Centre interuniversitaire de recherche sur la science et la technologie, **cet indicateur ne mesure l'impact qu'à court terme** et, quand il a été conçu, son seul but était d'aider les bibliothèques à choisir les revues auxquelles elles devaient s'abonner, en indiquant aux bibliothécaires les revues les plus citées. Cependant, un glissement s'est opéré : il a été utilisé pour évaluer un article, puis pour évaluer les chercheurs.

Autrement dit, **cet indicateur qui servait à l'origine à mesurer l'impact d'une revue est devenu un critère pour évaluer la qualité de la recherche et donc pour évaluer les chercheurs**. Or, l'obtention d'une bonne évaluation conditionne l'avancement du chercheur et le

montant des subventions reçues par le laboratoire. **Ce critère d'évaluation a donc incité les chercheurs à privilégier certaines revues qui jouissaient d'une plus grande visibilité** et a donc été un des facteurs ayant favorisé l'émergence et la domination des grands éditeurs commerciaux. L'usage détourné qui est fait de cet indicateur est très critiqué aujourd'hui tant pour les effets pervers qu'il entraîne que pour ses nombreux biais, notamment car :

- il prend en compte l'autocitation et il peut de cette manière là être gonflé artificiellement
- il ne mesure pas la qualité d'un article

- il favorise les chercheurs qui publient beaucoup
- il est très variable selon les disciplines qui ne peuvent être comparées

Aujourd'hui, les initiatives en faveur de la Science Ouverte remettent en question le facteur d'impact et proposent de prendre en compte de nouveaux critères pour l'évaluation des chercheurs. Par exemple, le *Plan national pour la Science Ouverte* (4 juillet 2018) souhaite « *réduire l'emprise de l'évaluation quantitative au profit de l'évaluation qualitative* » et donc « *reconnaître la science ouverte dans les évaluations des chercheurs et des établissements* ».

■ Sonia SALAMI

OPEN EDUCATION : LE MOOC AU CŒUR DES UNIVERSITÉS

Constance PICQUE

A l'heure du numérique, les universités se transforment. Telle est la volonté politique menée à travers l'Hexagone. Mais pourquoi tant d'engouement pour le MOOC ?

Née aux États-Unis, cette nouvelle forme pédagogique en faveur de l'Open Access n'a-t-elle que des vertus ?



Le MOOC au cœur des universités (CC0)

Au cours de ces dix dernières années, nos universités françaises ont connu de profonds bouleversements pédagogiques au regard de la révolution Internet et du développement rapide des nouvelles technologies numériques. Ces changements se traduisent par le développement de nouvelles pratiques informationnelles telles que les ENT, appelés environnements numériques de travail, et les plateformes d'apprentissage en ligne pour l'enseignement à distance. Mais pas seulement... Aujourd'hui, de plus en plus d'universités se mettent à la création de MOOC (*Massive Open Online Courses*), nouveau dispositif de formation ouvert à distance.

DÉFINITION DU MOOC

Apparu pour la première fois en 2011, **le MOOC est un nouvel outil pédagogique de l'Open Education qui permet de suivre une formation en ligne, gratuitement et sans exigence de diplômes ou de pré-requis sur une période allant en moyenne de 5 à 6 semaines.** Accessible tous et pouvant être suivi par un nombre illimité de personnes, ces nouveaux dispositifs de formation permettent de toucher des participants de tous les pays et de profils variés : étudiants classiques, professionnels en activité, autodidactes ou encore retraités (Bourcieu, 2013).

Le MOOC s'applique aussi bien à la formation continue qu'à la formation initiale. De manière individuelle ou collective, les participants ont donc accès à des modules de cours sous forme de vidéos d'une dizaine de minutes, à des activités régulières et autocorrectives, ainsi qu'à des forums destinés aux échanges entre apprenants et enseignants. Assurés par des outils du Web 2.0, le MOOC favorise également l'échange et l'interaction.

Nous pouvons distinguer deux formes de MOOC : les **xMOOC**, issus des cours traditionnels, qui visent à valider les compétences acquises en délivrant un certificat de réussite ; et les **cMOOC**, issu de l'approche « *connectiviste* » (théorie de l'apprentissage développée par George Siemens et Stephen Downes fondée sur les apports des nouvelles technologies). En France, les MOOC sont principalement diffusés sur les plateformes suivantes : *coursera.org*, *fun-mooc.fr* et *OpenClassrooms.com*. A l'international, nous pouvons également citer les plateformes *udacity.com* et *edx.org*. En somme, le MOOC est sensé permettre un libre accès à l'éducation.

CONTEXTE

Depuis les années 90, les gouvernements successifs pensent que l'université française doit être adaptée au monde de l'entreprise. De nombreuses réformes ont été adoptées en vue de cet objectif, comme la loi relative aux « *libertés et responsabilités des universités* » (dite « loi Pécresse ») et le plan dit « *Réussite en licence* » de 2007, puis la loi relative à l'Enseignement supérieur et la Recherche de 2013 (dite « loi Fioraso ») (Pinto 2014). L'objectif de ces différentes lois est de développer le numérique dans le système éducatif entre autres, à travers

la création de plateformes ENT et à travers une « *professionnalisation* » des études censées permettre aux diplômés d'acquérir une « *expérience professionnelle* » favorable à l'insertion.

Par ailleurs, parallèlement au développement de l'*Open Access*, une compétition féroce entre universités du monde entier a vu le jour, sans qu'il y ait de lien de causalité. Aujourd'hui, tous les établissements de l'enseignement supérieur se

Définition des 2 formes de MOOC par Matthieu Cisel, doctorant à l'ENS Cachan :

xMOOC : Cours à vocation de transmission et duplication du savoir (cf. Coursera, Udacity et edX). Il n'y a pas de révolution majeure dans les modes d'apprentissages qui restent très frontaux et sans aucune valeur ajoutée. Les activités sont individuelles, les apprenants sont peu ou pas encadrés ni accompagnés, les connaissances sont transmises au travers de cours magistraux plus que construites, l'évaluation cible avant tout les savoirs et non les compétences.

cMOOC : (c pour connectivisme), apparus un peu plus tard, ces MOOC ont pour ambition de changer la façon d'enseigner en profondeur. Il n'y a pas vraiment d'ordre séquentiel, les contenus libres d'accès sont développés et mis en réseau aussi bien par les apprenants que par les professeurs. Le rôle de l'enseignant est décentré et reporté sur la facilitation des échanges entre participants, notamment via les réseaux sociaux.

¹ CONCYTEC : Consejo nacional de ciencia, tecnología e innovación tecnológica (en français : Conseil national de la science, de la technologie et de l'innovation technologique du Pérou)

retrouvent ainsi confrontés aux dérivés du classement de Shanghai fondé sur un modèle d'université à l'américaine qui encourage à la fois la concurrence, la sélection à l'entrée et la fusion des universités en grands pôles. En effet, ce classement consiste à lister les meilleures universités mondiales selon la qualité de l'enseignement et de l'institution, la taille de cette dernière et le nombre de publications. A titre d'exemple, l'Université Pierre et Marie CURIE, classée au 36^{ème} rang mondial, est actuellement la première université française reconnue à l'international. Ce classement peut jouer dans certains financements, et sur l'attractivité des bons étudiants étrangers. Par exemple, au Pérou, le CONCYTEC¹ donne des bourses au mérite pour des étudiants en thèse, mais il faut que l'université d'accueil soit dans les 100 premiers du classement de Shanghai. Cette ouverture à la concurrence impose aux universités la recherche d'une plus grande visibilité.

Alors, pour faire face à cette course au gigantisme (Mercier, 2013) – mais aussi parce que les universités sont touchées de plein fouet par la baisse des subventions –, le gouvernement a décidé de mettre en place un nouveau dispositif de rapprochement des universités, appelé Pôles de Recherche et d'Enseignement Supérieur (PRES). L'objectif du PRES est d'augmenter

la visibilité des universités françaises. Nous pouvons citer, par exemple, la COMUE (Communauté d'Universités et d'Établissements) du Languedoc-Roussillon.

UNIVERSITÉS CONFRONTÉES AUX DÉRIVÉS DU « CLASSEMENT DE SHANGHAI »

Entre temps, ont émergées de nouvelles méthodes pédagogiques issues des technologies de l'information et de la communication dont le

MOOC, apparu aux États-Unis dans les années 2010 en réponse à la crise des universités privées américaines.

AVÈNEMENT DU MOOC

Devant l'engouement et le nombre grandissant d'inscrits (Lubnau, 2017), notamment auprès de la génération *Digital Native*, l'État a décidé de développer les MOOC gratuits avec des certifications payantes car ceux-ci participeraient à un meilleur équilibre économique et social.

Du côté des universités et des enseignants-chercheurs, soumis à la pression des accréditations et des rankings internationaux (Bourcieu, 2013), le MOOC donne une meilleure visibilité.

Du côté des apprenants, qui peuvent être étudiants ou non, ces derniers ont des motivations très diverses lorsqu'ils suivent des MOOC. Pour les étudiants, les MOOC peuvent représenter une opportunité pour découvrir d'autres disciplines, pour approfondir

des connaissances et les MOOC peuvent s'inscrire dans leur parcours pédagogique en tant que complément d'un cours (parfois recommandés par les enseignants). Pour les apprenants non étudiants, les MOOC sont principalement une opportunité d'ouverture sur le monde, un hobby, une manière d'acquérir de nouvelles compétences professionnelles ou d'actualiser ses connaissances.

C'est ainsi que la Fonction Publique Territoriale (FPT) diffuse régulièrement des MOOC à destination de ses agents ou à destination d'un public intéressé afin d'expliquer le fonctionnement des institutions, de présenter de nouvelles dispositions juridiques ou d'apporter des réponses à des débats de sociétés en lien avec la FPT. Selon le site du CNFPT (Centre National de la Fonction Publique Territoriale) « *Les buts de ces formations sont de couvrir des besoins spécifiques, prolonger ou accentuer l'efficacité des formations présentes et de s'adapter aux contraintes temporelles et géographiques des agents. L'objectif du CNFPT est de rendre accessible au plus grand nombre ses formations numériques et d'inciter les agents territoriaux à placer le numérique au cœur de leur parcours de formation* ». Cet exemple montre que les MOOC sont envisagés comme une réponse à un besoin spécifique d'un agent dans son cadre professionnel, comme un complément des formations dispensées en présentiel par le CNFPT ou comme un moyen de répondre aux questionnements d'un public intéressé. **Cet exemple illustre bien par ailleurs la principale limite actuelle des MOOC qui sont souvent assez simples et superficiels et qui ne constituent qu'une porte d'entrée, une initiation à un sujet, un complément et ne remplacent pas une vraie formation ou un cursus universitaire. Ils ne sont d'ailleurs pas conçus pour**



cela, bien souvent l'enjeu premier est d'améliorer la visibilité de l'institution réalisatrice du MOOC.

En France, la plateforme « FUN MOOC » (France Université Numérique) a vu le jour en 2015. Encadrée par le ministère de l'Enseignement Supérieur et de la Recherche, FUN est dédiée exclusivement aux MOOC. Pour renforcer son action, le gouvernement a également lancé le programme « Investissements d'Avenirs » en partenariat avec l'Agence National de la Recherche (ANR). Ce programme a pour vocation d'accélérer la création de MOOC, de cursus en ligne et de dispositifs de formations numériques de qualité, afin de développer l'attractivité internationale des établissements d'enseignement supérieur et de recherche, de promouvoir des dispositifs pédagogiques innovants par le numérique et de conforter une dynamique de formations universitaires tout au long de la vie (Extrait du programme). En parallèle, un site web officiel a été développé pour référencer toutes les formations, y compris les MOOC, proposées par les établissements de l'enseignement supérieur (Lien vers le site web : <http://www.sup-numerique.gouv.fr>).

*FUN EST DEVENUE
LA RÉFÉRENCE
DES FORMATIONS
PROFESSIONNELLES ET
CONTINUES*

Pour autant, les certifications délivrées par les MOOC ne sont pas reconnues immédiatement dans la vie professionnelle. Si bien qu'un an après son lancement, FUN s'attaque à la certification des connaissances et à la formation continue qui permettent



Page d'accueil de la plateforme FUN MOOC - <https://www.fun-mooc.fr>

aux participants d'obtenir notamment des crédits européens « ECTS » (European credits transfer system), dans le cadre de leur cursus universitaire. En l'espace de trois ans, la plateforme FUN est donc devenue la référence des formations professionnelles et continues des établissements publics d'enseignement supérieur en France, après le CNED (Centre National d'Enseignement à Distance), actuellement détenteur du monopole de la formation à distance (Lubnau, 2017). **Mais les Moocs remplissent-ils vraiment leurs promesses ? Quelles sont leurs limites ?**

UTOPIE LIBÉRALE

Fondée sur une idéologie mêlant transparence, participation de tous (l'interactivité) et connaissance pour tous, le MOOC tiendrait plus du marketing académique que de la pédagogie numérique selon le sociologue, Jérôme Valluy (2014). Le MOOC ne serait qu'une utopie libérale, un fantasme occidental de lutte contre la fracture culturelle au nom de l'idée d'une culture universelle accessible et profitable pour tous. Cet accès à la culture reste pourtant limité et ne facilite pas l'accès aux études longues.

PÉDAGOGIE EN QUESTION

Un des principaux problèmes est le **taux d'échec très élevé**. Il y a un taux d'abandon qui atteint 90 % (Valluy, 2014). Cela vient du fait qu'il n'y a pas d'interaction avec les professeurs ou les autres étudiants, que le cours n'est pas forcément adapté au niveau des participants. De plus, il faut être dans un environnement propice à l'étude pour se concentrer car avec les MOOC, il n'y a pas de salles de cours... Il faut également une discipline personnelle pour travailler tous les jours.

Par ailleurs, le MOOC est présenté comme favorable à l'échange et à l'interaction mais il reste fondé sur une approche pédagogique descendante. Il est difficile de créer de l'interactivité et d'apporter des réponses spécifiques à chaque étudiant lorsque le nombre de participants est très important. Comme le souligne Patrick Zylberman, professeur d'histoire de la santé qui a participé à un MOOC intitulé « SRAS, une révolution de la gouvernance mondiale des épidémies ? » en 2014, **le point faible du MOOC reste la pédagogie, ou plus précisément la relation pédagogique** (article paru sur le blog du Centre Virchow-Villermé en 2014).

ET LES ENSEIGNANTS ?

Quant aux enseignants, certains craignent une surcharge de travail au niveau de la préparation des supports et du suivi des étudiants. Outre l'aspect technique, les MOOC mobilisent des ressources humaines importantes. En effet, **la construction d'un MOOC demande à l'enseignant d'accomplir un travail de préparation conséquent et de grande qualité.** D'après l'expérience du professeur Patrick Zylberman, la charge de travail est en effet plus importante depuis la production du MOOC jusqu'à la délivrance des attestations : « *Il m'a fallu un mois pour écrire les cent pages qui constituent le texte de ce MOOC. Pour un cours normal, on prend des notes, on a un diaporama. Ici, l'improvisation est impossible, en raison même de la nécessité de minuter chaque section du cours lors de sa réalisation. [...] Pendant le MOOC, je consacrai un jour par semaine pour répondre à quelques questions du forum de discussion et réaliser la synthèse hebdomadaire* ». Pourtant, les universités peinent encore à le reconnaître. Selon Jean-Marie Guillot, maître de conférences en Bretagne, **il est nécessaire de réajuster de la rémunération des enseignants et tout particulièrement les grilles de valorisation fondées sur les activités de cours/TD/TP.** En effet, un enseignement en ligne et un enseignement assuré en présentiel ne peuvent pas être évalués de la même façon.

Au contraire, le gouvernement voit une possible extension de la logique des MOOC à l'enseignement supérieur comme un moyen de faire des économies. Par exemple, dans le cadre du troisième Programme d'investissements d'avenir, le ministère de l'Enseignement supérieur, de la

Recherche et de l'Innovation a lancé en 2018 un appel à projet « Nouveaux cursus à l'université » qui a, entre autre, pour objectif **d'expérimenter de nouveaux modèles de formation et de nouvelles pédagogies en intégrant le numérique.** C'est ainsi que l'Université Paul Valéry de Montpellier (entre autres) a reçu des financements pour son projet *Nexus* qui propose d'intégrer progressivement des modules d'apprentissage numérique en licence. Dans le contexte actuel de stagnation des dotations, de sur-effectif, ce genre de projet n'est pas annodin et semble avoir comme objectif de diminuer des coûts de la formation. Certains opposants tentent de démontrer que ce n'est qu'une illusion et que l'enseignant en présentiel coûterait plus cher. Si selon Anne Lubnau-Wimez (2017), généraliser l'apprentissage par le numérique à distance devient rentable dès lors que le nombre d'apprenants est plus nombreux à s'inscrire qu'il n'aurait été en présentiel, il semblerait pourtant que plus le nombre de participants est important, plus il est difficile de les accompagner : « **Soit le numérique se fait à bas prix, [...], et l'offre de formation confine à l'escroquerie. Soit les MOOC sont techniquement et pédagogiquement sophistiqués et leurs coûts de production deviennent exorbitants** » (Jérôme Valluy, 2014).

CONCLUSION

Finalement, l'accès à la formation n'a jamais été aussi facile, est-on pour autant dans l'ère de l'*Open Education*, c'est-à-dire l'éducation pour tous ? En effet, les MOOC ont beaucoup de limites. Il semblerait qu'ils ne soient pas un véritable outil pédagogique mais plus un instrument de « marketing académiques ». Pour autant,

le MOOC dans les universités a de belles perspectives devant lui.

Si le modèle du MOOC inspire les formations universitaires de demain, ne remettrait-il pas en cause le système de l'enseignement supérieur ? N'est-il pas un cache misère, un concept à la mode qui cache une situation où les différentes réformes de l'université depuis les années 80 rognent sur les moyens et plus récemment ont pour but d'en restreindre l'accès ? Le numérique seul, sans politique publique et sans investissements, peut-il inverser cette tendance ? Quelles seraient les conséquences sur l'université, sur les enseignants, sur les infrastructures (i.e. il faut des serveurs...), sur les valeurs de ces formations ? Affaire à suivre...

■ Constance PICQUE

BIBLIOGRAPHIE

ALBÓ Laia, HERNÁNDEZ-LEO Davinia, OLIVER Miquel, "Blended MOOCs: University Teachers' Perspective". CHANGEE/WAPLA/HybridEd@ EC-TEL. 2015. p. 11-15.

BOURCIEU S., LEON O., « Les MOOC, alliés ou concurrents des business schools ? » L'Expansion Management Review 2013/2 (n°149), p. 14-24. URL : http://ceur-ws.org/Vol-1599/2HybridEd_2015.pdf

DANIEL Jonh, « Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility ». *Journal of Interactive Media in Education*. 2012(3), p.Art. 18. URL : <https://jime.open.ac.uk/articles/10.5334/2012-18/>

DELPEYROUX Stéphanie, ROUVEIX Florence, GUYON Magali, BACHELET Rémi, TRUCHE Corinne, « Hybridation d'un MOOC: quels moyens déployer? », QPES 2017 : Questions de Pédagogie pour l'Enseignement Supérieur, Jun 2017, Grenoble, France. QPES, Actes QPES 2017 : Questions de Pédagogie pour l'Enseignement Supérieur, pp.405 - 415, 2017. URL : <https://hal.archives-ouvertes.fr/hal-01686343/document>

GILLIOT Jean-Marie, « La place de l'enseignant dans un monde de MOOCs ? », Blog Technique innovantes pour l'enseignement supérieur, 26 novembre, 2012.

URL : <https://tipes.wordpress.com/2012/11/26/la-place-de-lenseignant-dans-un-monde-de-moocs/>

MERCIER Arnaud. « Classer, marchandiser et manager : quel idéal de l'Université opposer aux dérives en cours ? », *Questions de communication*, vol. 24, no. 2, 2013, pp. 125-144. URL : <https://www.cairn.info/revue-questions-de-communication-2013-2-p-125.htm>

LUBNAU-WIMEZ Anne « Nicolas Oliveri, Apprendre en ligne, quel avenir pour le phénomène MOOC ? », *Communication et organisation* [En ligne], 52 | 2017. URL : <http://journals.openedition.org/communication-organisation/5789>

PAPADOUDI Hélène, « Le statut didactique des TIC, les enseignants et la médiation technique. » Frantice.net, Université de Limoges, 2014, *Les TICE dans l'enseignement et la formation. Quels usages pour quelles pédagogies ?* p. 87-99. URL : <https://hal.univ-lorraine.fr/hal-01438229/document>

PINTO Vanessa, « La « réussite pour tous » passe-t-elle par la « professionnalisation de l'enseignement supérieur ? », *Savoir/Agir*, 2014/3 (n°29), p 67-73. URL : <https://www.cairn.info/revue-savoir-agir-2014-3-page-67.htm>

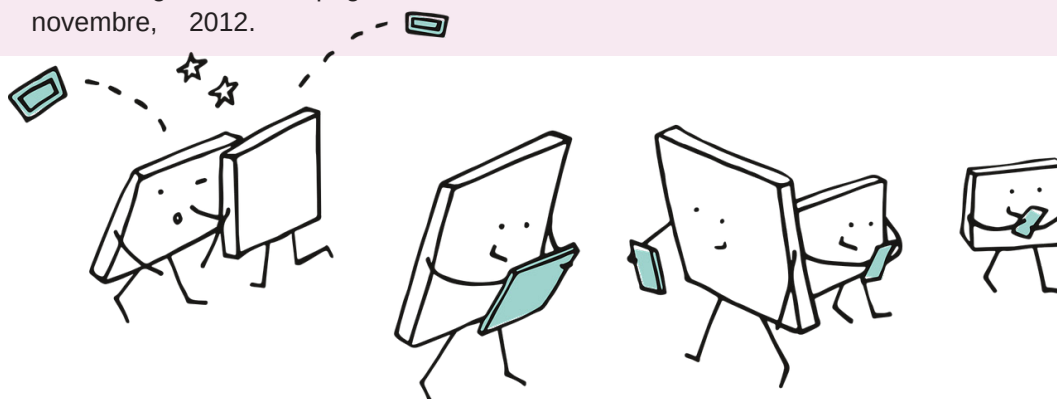
RAPP Lucien, « Les MOOCs, révolution ou désillusion ? Le savoir à l'heure du numérique », L'Institut de l'entreprise, 2014. URL : https://www.letudiant.fr/static/uploads/mediatheque/EDU_EDU/4/1/256141-les-mooc-revolution-ou-desillusion-rapport-de-lucien-rapp-institut-de-l-entreprise-septembre-2014-original.pdf

VARDI Moshe Ya'akov, « Will MOOCs destroy academia? », *Communications of the ACM*, 2012, vol. 55, no 11, p. 5-5. URL : <https://cacm.acm.org/magazines/2012/11/156587-will-moocs-destroy-academia/fulltext>

VALLUY Jérôme, « Ce sont des publicités, pas des cours », *France Soir*, 31/01/2014.

WILHELM Carsten, « Dans le halo des MOOC, la rationalité communicationnelle de la formation 2.0 », *Communication et organisation*, 49/2016. URL : <https://www.cairn.info/revue-communication-et-organisation-2016-1-p-87.htm>

Article paru en 2014 sur le blog du Centre Virchow-Villermé en Allemagne, « Enseigner avec un MOOC retour d'expérience du professeur Patrick ZYLBERMAN » URL : <https://virchowvillermé.eu/fr/education-fr/enseigner-avec-un-mooc-retour-dexperience-du-professeur-patrick-zylberman/>



DOSSIER: DONNÉES de la RECHERCHE



Ouverture des données de la recherche : de quoi parle-t-on ?

Alisson DEVILLIERS, Laure PIERRAT, Sonia SALAMI

COMMENT DÉFINIR LES DONNÉES DE LA RECHERCHE ?

Les données de la recherche, aussi appelées *research data* en anglais, sont définies par l'Organisation de Coopération et de Développement Économiques (OCDE) dans leur rapport de 2007 Principes et lignes directrices pour l'accès aux données de la recherche financées sur fonds publics comme « *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 les résultats de la recherche* ».

La définition données par l'OCDE est très restrictive car les données de la recherche ne se limitent pas aux « *enregistrements factuels* ». En effet, l'INIST (l'Institut de l'Information Scientifique et Technique) du CNRS définit 5 types de données :

- **Les données d'observation** : données capturées en temps réel, habituellement uniques et donc impossibles à reproduire.
- **Les données expérimentales** : données obtenues à partir d'équipements de laboratoire, qui sont souvent reproduction mais parfois coûteuses.
- **Les données computationnelles ou de simulation** : données générées par des modèles informatiques ou de simulation, souvent reproductibles si le modèle est correctement documenté.
- **Les données dérivées ou compilées** : données issues du traitement ou de la combinaison de données « *brutes* », elles sont souvent reproductibles mais coûteuses.
- **Les données de référence** : collection ou accumulation de petits jeux de données qui ont été revus par les pairs, annotés et mis à disposition.

Mais les données de la recherche sont très diverses et hétérogènes et même la typologie de l'INIST n'épuise pas toutes les possibilités : un article du carnet des URFIST sur *hypothese.org* donne des exemples qui montrent toute l'étendue que peut recouvrir le terme de « données de la recherche » : « les "données" d'un linguiste peuvent des écrits ou des discours, des enregistrements de locuteurs ; les "données" d'un médiéviste sont des sources archivistiques, archéologiques, épigraphiques, iconographiques, littéraires ; les "données" d'un géologue rassemblent des coupes et observations de terrain consignées sur un carnet, des résultats de carottage, des analyses d'échantillons, des données sismographiques » (*urfistinfo.hypothèses.org*).

On peut tout de même définir deux tendances :

- En sciences humaines et sociales, l'exploitation des documents (textes ou images) est, de loin, la source la plus répandue. Néanmoins les enquêtes, les entretiens, les observations, ainsi que les archives se révelent également être des sources importantes.
- Dans le domaine des sciences exactes, la typologie des données peut varier mais il y a majoritairement des données expérimentales (exemple : chromatographe) ou de données de simulations (exemple : simulations sismiques).

La définition donnée par l'OCDE est également trop restrictive sur un deuxième point, **les données de la recherche ne sont pas seulement des données « nécessaires à la validation des résultats »**. En effet, le « chercheur produit plus de données que celles qui sont stricto sensu nécessaires pour valider le résultat de la recherche » (Rosemberg, 2015) or

ces données n'étant pas directement impliquées dans les résultats « *pourraient être utilisées par d'autres chercheurs dans le cadre d'un projet de recherche inédit* » (Rosemberg, 2015). C'est même un des principaux intérêts du mouvement d'ouverture des données.

A QUI S'ADRESSENT LES DONNÉES DE LA RECHERCHE ?

Il existe différents types de données collectées en fonction des phases du projet de recherche. Les données ne sont pas figées. Ainsi, il faut distinguer les données brutes, des données traitées ou dérivées et de celles qui ont été analysées. **Ces différents types de données ne sont pas destinées aux mêmes acteurs.**

- **Les données brutes (ou données primaires)** sont des données non interprétées et qui n'ont été soumises à aucun traitement. A destination des experts, elles sont difficilement interprétables et s'adressent donc à un public pointu très restreint.

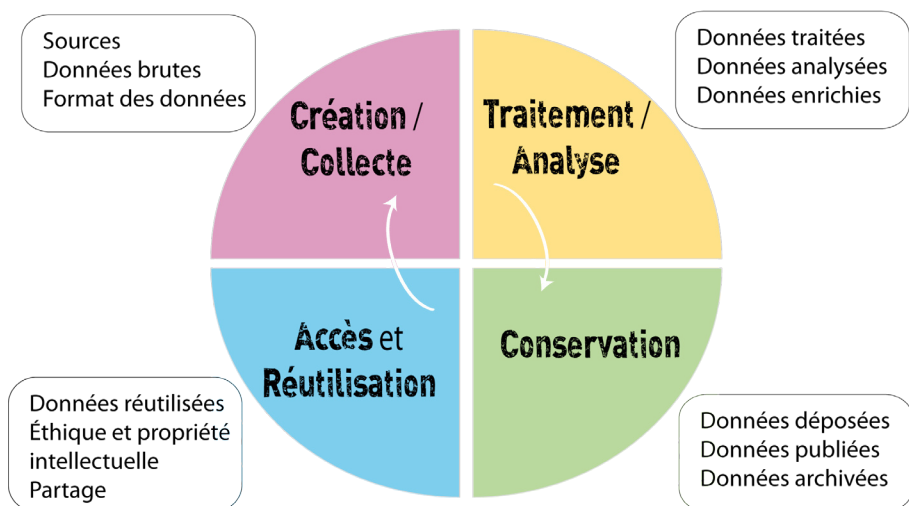
- **Les données traitées** (les données « *produites après calibration/éta-lonnage ou correction des données brutes* ») et les données dérivées (celles qui « *présentent un résumé ou une représentation/vue spécifique des données (agrégation, compilation, calcul, réorganisation* » : Jacquemot-Perba, 2015), s'adressent quant à elles à un public composé de scientifiques et/ou de citoyens avertis.

- Enfin, **les données analysées** s'adressent à tous : il s'agit des données publiques.

POURQUOI PARTAGER LES DONNÉES DE LA RECHERCHE ?

Il est souvent question du partage et du libre accès aux résultats de la recherche, pourtant les données issues de la recherche ont autant d'importance que les résultats publiés. Le chercheur produit généralement des données seulement en vue de publier des articles (ou autres travaux). **Cette culture de l'article a longtemps entraîné un désintérêt pour les données issues de la recherche mais les choses sont en train de changer.**

Le cycle de vie de la donnée



Les 2 voies de l'Open Access, inspiré du diaporama "Fiche pratique IST Agropolis : L'Open Access et les données de la recherche" sur *slideshare.net* © C. Picque

● L'intérêt de partager les données est de permettre leur réutilisation et donc l'exploration de nouvelles pistes qui n'avaient pas été envisagées au départ par les chercheurs à l'origine de la collecte des données, ce qui permet de rentabiliser la recherche et d'accélérer l'innovation car la collecte des données peut être coûteuse. De plus, la réutilisation des données permet de favoriser de nouvelles collaborations, il n'y a plus de frontières, et d'éviter la duplication des travaux.

● L'intérêt de mettre à disposition les données de la recherche est aussi d'**aboutir à une science plus intègre en facilitant l'évaluation des données**, ainsi que leur critique : ce qui réduit le risque de fraude et favorise le débat sur la méthode avec une possibilité de remettre en cause des résultats.

● Enfin, le partage des données permet de **rendre plus visible la communauté scientifique**. En effet, les travaux peuvent être visibles non seulement à partir des articles publiés mais aussi à partir des données elles-mêmes. Le partage des données multiplie donc les chances pour les chercheurs d'être cités. En effet, l'auteur des données est identifiable¹ et a aussi la possibilité de publier des *data papers* (une publication scientifique qui a vocation de décrire un ou plusieurs jeux de données, plutôt que de résultats d'analyse).

Cependant, il est important de rappeler que tout n'est pas partageable. Par exemple, les données engageant la sécurité nationale sont exclues. Les données personnelles sont protégées et leurs publications doivent faire l'objet d'anonymisation. D'autre part, les données relevant du droit à la propriété intellectuelle ou liées à une exploitation commerciale sont difficilement partageables.

UNE VOLONTÉ POLITIQUE POUR INCITER LES CHERCHEURS À DÉPOSER EN LIBRE ACCÈS LES DONNÉES DE LEUR RECHERCHE

En France et dans de nombreux pays de l'Union européenne, l'évolution de **la législation vise à permettre aux chercheurs de publier en libre accès** à la fois les résultats de leurs recherches et leurs données. Mais, au-delà de cette possibilité (qui n'est pas une obligation), **de plus en plus d'organismes publics finançant la recherche exigent en contrepartie de leur contribution que la publication des résultats de la recherche, ainsi que les données soient publiés en Open Access.**

Par exemple, en 2014, l'Union Européenne lance « Horizon 2020 », reconnu pour être le plus grand programme de recherche et d'innovation d'Europe. Il est ouvert à la participation des équipes des 28 Etats membres de l'UE et des pays associés. Ce programme a pour objectif de s'assurer que l'Europe atteint un niveau scientifique et technologique de classe mondiale qui stimule la croissance économique. Avec près de 80 milliards d'euros de financement sur 7 ans (de 2014 à 2020), Horizon 2020 se concentre sur trois domaines clés : l'excellence scientifique, la primauté industrielle et les défis sociétaux (Union européenne, Horizon 2020 en bref, 2014). La participation à ce programme est ouverte aux chercheurs du monde entier à condition que ceux-ci diffusent les résultats qu'ils produisent

aussi rapidement que possible². Les chercheurs sont ainsi poussés à partager leurs données de recherche, sous peine de sanctions financières. Ces dernières sont à déposer sur le site internet du projet et/ou dans un entrepôt en *Open Access*.

COMMENT PARTAGER LES DONNÉES DE LA RECHERCHE SUR LE WEB³?

Les chercheurs qui souhaitent donner un accès libre à leurs données et ainsi les valoriser ont le choix entre deux voies principales et complémentaires :

● **Publier un *Data paper* en libre accès**, c'est-à-dire une publication dont l'objet est de décrire « *un ou plusieurs jeux de données plutôt que des résultats d'analyse* » dans une forme « *structurée et lisible pour un humain* ». Il existe des revues spécialisées dans la publication de *Data papers* appelées *Data journals*.

● **Publier ses données dans un entrepôt de données (*Data repository*) ou réservoir de données** contenant « *des données de recherche brutes ou élaborées qui sont décrites par des métadonnées de façon à pouvoir être retrouvées* ». Il existe des annuaires répertoriant les entrepôts existants pour aider le chercheur dans son choix.

Ces deux dispositifs, le *Data paper* et l'entrepôt, permettent à la fois de valoriser les données de la recherche mais aussi de valoriser le travail du chercheur qui pourra être cité si celles-ci sont réutilisées. En effet, un *Data paper* est une publication citable au même titre que n'importe quelle publication et un identifiant pérenne attribué au jeu de donnée permet de relier les deux.



TOUTES LES DONNÉES PEUVENT-ELLES ÊTRE PARTAGÉES ?

En 2016, la loi pour une République numérique veut permettre la réutilisation des données dès lors que celles-ci sont issues de recherches financées au moins pour moitié par des fonds publics et qu'elles ne sont « pas protégées par un droit spécifique ou une réglementation particulière et qu'elles ont été rendues publiques par le chercheur, l'établissement ou l'organisme de recherche, leur réutilisation est libre⁴ ». En effet, **certaines restrictions existent pour protéger notamment les données personnelles, les données sensibles menaçant la sécurité nationale ou les données ayant une valeur commerciale.**

On a donc d'une part, une volonté politique en faveur du partage et de la réutilisation des données et d'autre part « le secret médical, le secret des affaires, le droit d'auteur ou le règlement européen sur la protection des données personnelles (RGDP) s'imposent à tous et pondèrent, réduisent voire annulent toute possibilité d'ouverture des données⁵ ». Le gouvernement a donc publié en décembre 2017 un guide juridique pour aider les chercheurs à y voir plus clair et à déterminer s'ils doivent partager leurs données et le cas échéant sous quelles conditions. Le document *Ouverture des données de la recherche, guide d'analyse du cadre juridique en France*⁶ différencie les données dont la communication est obligatoire et celles dont la communication est possible sous certaines conditions, de celles dont la publication est interdite par principe :

● Communication obligatoire :

- Données géographiques et relatives à des émissions de substances dans l'environnement.

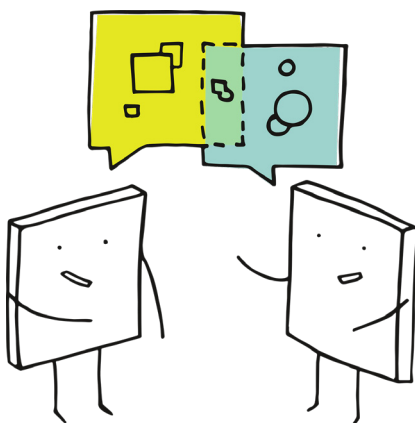
● Communication interdite par principe :

- Données présentant des risques pour la protection du secret de la défense nationale.
- Données présentant des risques pour la sûreté de l'État, la sécurité publique, la sécurité de l'établissement,
- Secrets professionnels.

● Communication sous conditions :

- Données présentant des risques pour la protection du potentiel scientifique et technique de la nation.
- Le cas des zones à régime restrictif (ZRR) Données protégées par le droit d'auteur et autres droits de propriété intellectuelle.
- Données relatives aux personnes, à la vie privée.
- Données statistiques.
- Données liées à un contrat avec un tiers non soumis à une obligation de service public.

■ Alisson DEVILLIERS
■ Laure PIERRAT
■ Sonia SALAMI



¹ Doranum nous rappelle que « les données sont citables et que la paternité du chercheur est reconnue » notamment par le biais du PID « identifiant pérenne objet »

² En France, ce n'est pas moins de 3 milliards d'euros de subventions qui ont été obtenues, ce qui nous classe en première position ! Une somme qui va s'accroître dans le temps, puisque qu'en effet de 2018 à 2020 30 milliards d'euros seront encore en jeu dans les appels à propositions.

³ Cette partie se fonde sur le module de formation en ligne de l'INIST « Une introduction à la gestion et au partage des données de la recherche » [En ligne : <http://www.inist.fr>]

⁴ Voir dans ce numéro l'article de Laure PIERRAT "Loi pour une République numérique éclairage sur l'article 30.

⁵ Préface de : "Ouverture des données de la recherche, guide d'analyse du cadre juridique en France", décembre 2017.

⁶ Disponible sur le site web : <http://www.bibliothequescientifiquenumerique.fr>

BIBLIOGRAPHIE

JACQUEMOT-PERBAL Marie-Christine, COSSERAT Françoise, CIOLEK-FIGIEL Anne, « Gestion et diffusion des données de la recherche », support de stage du 12 mai 2016. URL : <https://drive.google.com/file/d/0BxKZLWq08xXTW5VOEUtd2FSRE0/view>

COMMISSION EUROPEENNE, *Horizon 2020 en bref. Le programme-cadre de l'UE pour la recherche et l'innovation*, 2014. URL : https://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/H2020_FR_KI0213413FRN.pdf

ROSEMBERG Noémie, « De la définition des données de la recherche », *hypothese.org*, 30/09/2015. URL : <https://donneesshs.hypotheses.org/39>

Quels outils pour exploiter les données de la recherche ?

Sonia SALAMI

Les données prennent une importance croissante dans la recherche avec le mouvement de l'ouverture des données qui permet leur réutilisation mais aussi car les outils de mesure sont aujourd'hui très précis et récoltent énormément de données qu'il faut ensuite analyser. Les chercheurs ont donc besoin d'outils adaptés pour les exploiter et pour travailler de façon collaborative sur les données comme :

- **Les outils de calculs puissants et de traitement statistique.** Par exemple le logiciel R.

- **Les outils de *text and data mining***

La notion de *text and data mining* (TDM) ou encore, en français, la fouille de textes et de données, est apparue au début des années 1990, dans le domaine du Marketing. Aujourd'hui le TDM s'est beaucoup répandu dans le domaine scientifique avec le développement des archives ouvertes. Cependant, le TDM subit, un frein juridique, car la technique de la fouille de contenus implique la création d'une copie de l'œuvre à traiter ainsi que la sécurisation de l'accès à ces données. On retrouve notamment plusieurs lois sur le droit du TDM, comme par exemple l'article 38 de la loi (Petite loi) qui confère un droit au TDM en insérant une exception au droit d'auteur et au droit du producteur de base de données. On note qu'aujourd'hui beaucoup d'entrepôts de données proposent des fonctionnalités de *Text and Data Mining* (TDM).

- **Les outils de visualisation**

Par exemple, le logiciel *Gephi* ou *The Vistorian*. Les outils de visualisation des données

ont pour objet de faciliter. Les principales représentations des données sont : la cartographie, les représentations temporelles, l'analyse de réseaux et l'analyse de données textuelles. Avec l'évolution des sources de données, les outils de visualisation doivent également s'adapter, que ce soit pour prendre en considération des données hybrides ou pour correspondre au monde de la recherche actuel qui est de plus en plus pluridisciplinaire et international et qui rend nécessaire des outils collaboratifs.

- **Outils de collaboration autour des données**

De plus en plus de disciplines s'intéressent aux données et pour les exploiter, il est nécessaire de pouvoir réaliser des analyses décloisonnées et transdisciplinaires. Le travail des chercheurs est de plus en plus collaboratif, en particulier au niveau de la recherche européenne qui fédère des chercheurs et des laboratoires situés dans différents pays. La France et l'Union européenne développent donc :

- **Des outils de stockage sécurisés et de partage de jeux de données** : plateformes d'hébergement équipées de dispositifs d'authentification et d'autorisation.

- **Des outils de travail collaboratif en réseau pour les données** : cloud, environnement virtuel de recherche (EVR) dédiés aux données et à la gestion de workflow.

▪ Sonia SALAMI

Gestion et diffusion des données de la recherche : quels enjeux ?

Sonia SALAMI

L'ouverture des données de la recherche s'inscrit dans le mouvement de la Science Ouverte. La gestion des données et leur diffusion sont les deux étapes clés de l'ouverture des données de la recherche. C'est en amont, avant même la collecte des données qu'il faut envisager leur gestion pour qu'elles puissent être effectivement conservées et réutilisées sur le long terme. En France, de nouvelles dispositions juridiques visent à assurer une diffusion ouverte des données et à éviter leur captation par les éditeurs.



Diffusion d'une donnée de la recherche (CC0) © C.Picque

L'OUVERTURE DES DONNÉES DE LA RECHERCHE ET LA SCIENCE OUVERTE

L'ouverture des données de la recherche, tout comme le libre accès pour les publications scientifiques, s'inscrit dans **le mouvement de la Science Ouverte (Open Science)**. La Science Ouverte est un mouvement mondial qui veut développer : « *une approche transversale de l'accès au travail scientifique, des visées du partage des résultats de la science mais aussi une nouvelle façon de FAIRE de la science en ouvrant les processus, les codes, les méthodes* »¹.

L'ouverture des données de la recherche a des retombées positives importantes pour les chercheurs et la Science :

- En permettant un plus grand contrôle et une plus grande transparence dans les publications scientifiques qui étaient la plupart du temps publiées sans les données ayant permis leur existence.
- En augmentant la visibilité et l'impact d'une étude.
- En faisant progresser la science plus rapidement notamment grâce à la réutilisation des données et à la possibilité de réaliser des méta-analyses.

L'ouverture des données de la recherche aura aussi un **impact au niveau économique** (en évitant de recréer des données qui sont déjà disponibles et en permettant leur réutilisation par des entreprises), **au niveau de la société** (en incluant les citoyens et en leur donnant davantage confiance dans la science). L'ouverture des données de la recherche représente également un **enjeu patrimonial** pour l'histoire des sciences. Cependant, en pratique, leur gestion et leur diffusion ne va pas sans poser de problèmes.

L'IMPORTANCE D'UNE BONNE GESTION DES DONNÉES OU DATA MANAGEMENT

La quantité des données produites par la recherche augmente de façon exponentielle mais la majorité des données se perdent à cause de négligences humaines et faute d'avoir fait l'objet d'un traitement adapté. En effet, selon une étude de 2013, 17 % des données sont inaccessibles seulement deux ans après leur publication et 80 % des données de la recherche sont perdues après 20 ans (Vines et al, 2014) ! A l'heure de l'ouverture des données de la recherche, le *Data Management* est donc devenu un enjeu majeur car il faut plus que jamais préserver ces données.

Mais comment définir le Data Management ?

Le *Data Management*, c'est la prise en charge des données durant tout leur **cycle de vie**, depuis le début du projet scientifique jusqu'à sa finalisation, puis au-delà, pour permettre la réutilisation des données sur le long terme (voir schéma page suivante du cycle de vie des données).



Data management (CC0) © C.Picque

Le Data Management permet d'assurer :

- La collecte des données de la recherche.
- La fiabilité, la traçabilité et la sécurité des données dans le temps en empêchant leur altération, leur falsification, leur suppression.
- Le stockage pérenne des données et leur lisibilité dans le futur pour permettre leur consultation.
- L'intelligibilité des données pour des personnes extérieures afin qu'elles puissent être utilisées à nouveau via la création de métadonnées.
- La définition des droits d'accès et de réutilisation des données.

Tous ces éléments sont censés être définis en amont du projet scientifique dans le **Plan de Gestion des Données (PGD)**, ou *Data Management Plan* (DMP). Il s'agit d'un document qui fixe les processus à mettre en place tout au long de la recherche pour garantir l'accès et l'exploitation ultérieure des données. Ce document définit le projet, les acteurs du projet, les financements, décrit et explicite les données, définit les modalités de stockage et de sauvegarde, ainsi que

¹ DIST - CNRS. Livre blanc — *Une Science ouverte dans une République numérique. Nouvelle édition [en ligne].* Marseille : OpenEdition Press, 2016 (généré le 29 avril 2018 : « L'application de la loi "pour une République numérique, un guide partagé pour le travail de la science, éléments de périmètre ».

les droits d'accès et de réutilisation. Pour que le PGD soit fiable, il faut le mettre à jour tout au long du projet. L'ouverture des données de la recherche étant devenu une priorité de la recherche publique au niveau mondial, **de nombreux organismes de financements et universités exigent pour les recherches qu'ils financent la tenue d'un PGD.**

IMPORTANCE DE LA GESTION DES DONNÉES POUR LEUR DIFFUSION ET LEUR RÉUTILISATION

La gestion et la diffusion des données de la recherche sont interreliées car sans une bonne gestion, la diffusion et la réutilisation des données sont impossibles. Tout d'abord, les principales causes de perte de données de la recherche sont des adresses e-mail obsolètes et des modalités de stockage non pérennes.

Ces causes peuvent facilement être évitées lorsqu'on définit un PGD. Ensuite, **les données doivent pouvoir être exploitées par d'autres équipes de recherche d'où la nécessité de les rendre intelligibles. Il faut élaborer une documentation (Data Paper)** venant accompagner ces données. Elle devra préciser entre autres les conditions de recueil des données (la méthodologie suivie, l'échantillon, les instruments...) et comporter un dictionnaire des données. Documentation et données doivent être liées réciproquement par des identifiants pérennes.

Rendre intelligible les données passe aussi par la création de métadonnées selon des standards définis (metadata standards). Les métadonnées doivent répondre aux questions : Qui ? Que ? Où ? Quand ? Comment ? Pourquoi ? Elles sont à la

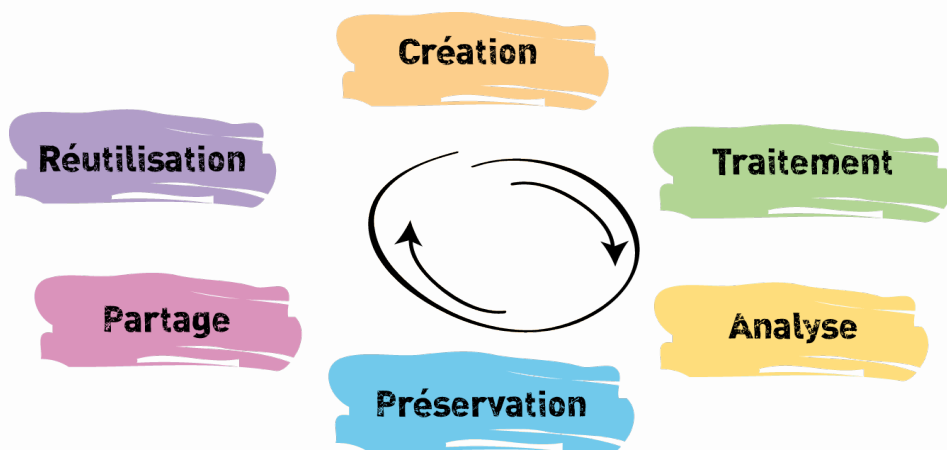


Figure 1 : Cycle de vie des données, inspiré du schéma de Nathalie Clairoux, « Introduction à la gestion des données de la recherche » (CC0) © C.Picque

fois lisibles pour un humain et pour un ordinateur ce qui permet, entre autre, un accès aux données via une requête dans un moteur de recherche.

Enfin, pour permettre une réutilisation directe des données, il faut, avant même la collecte, penser à la lisibilité de celles-ci dans le futur, et **veiller à l'interopérabilité des données en privilégiant les formats ouverts.** Par exemple, diffuser des données sous forme PDF permet leur consultation mais difficilement leur réutilisation.

ASSURER UNE DIFFUSION OUVERTE DES DONNÉES DE LA RECHERCHE

En dehors du management des données, afin d'assurer une diffusion ouverte des données de la recherche, il est nécessaire de prendre des mesures pour que celles-ci restent publiques et ne soient pas captées par des intérêts privés (comme cela a été le cas pour les publications scientifiques par les éditeurs).

En effet, Le Livre blanc *Une Science ouverte dans une République numérique* établi par la DIST du CNRS qualifie de « *captation abusive* » l'accapa-

rement des données scientifiques de la recherche publique par les éditeurs. En cause, le droit de la propriété intellectuelle qui, dans le domaine scientifique gêne l'accès aux publications et aux données, entrave les progrès de la science et, surtout, entraîne une privatisation de la production scientifique par les éditeurs.

La *loi pour une République numérique* de 2016 essaye de pallier ce phénomène en donnant la possibilité aux chercheurs (mais non l'obligation), après une période d'embargo, de déposer en libre accès leurs publications lorsqu'elles sont issues de recherches financées au moins pour moitié par des fonds publics. **De même, la loi pour une République numérique entend anticiper une potentielle captation par les éditeurs des données de la recherche qui empêcherait leur réutilisation.** Dès lors que les données sont issues de recherches financées au moins pour moitié par des fonds publics et qu'elles ne sont « *pas protégées par un droit spécifique ou une réglementation particulière et qu'elles ont été rendues publiques par le chercheur, l'établissement ou l'organisme de recherche, leur réutilisation est libre* ».

Certains grands éditeurs (comme Elsevier et Springer) conscients de la valeur de ces données ont encouragé les chercheurs à publier leurs données de recherche avec leurs articles. Les données accompagnant les articles sont alors soumises aux mêmes conditions de publication que ces derniers : leur accès est payant (abonnement) et des clauses d'exclusivités limitent la réutilisation de ces données.

Cette tendance représentait le risque de voir se produire une appropriation progressive des données de la recherche par quelques grands éditeurs. En effet, les données de la recherche représentent une manne financière pour les grands éditeurs et ceux-ci, de par leur position de quasi monopole, pouvaient progressivement accaparer la plupart des données de la recherche de la même manière qu'ils avaient accaparé les articles et obligé les universités à payer pour avoir accès aux données et pour pouvoir les réutiliser.

De plus, cette tendance était contraire aux **exigences d'une Science Ouverte** où les données de la recherche doivent être accessibles non seulement aux chercheurs mais à toute la société, et doivent pouvoir être réutilisées - y compris

par des entreprises, dans un but lucratif. Le législateur a pris en compte ce risque de captation des données de la recherche et la nouvelle loi stipule que l'éditeur d'un écrit scientifique ne peut plus « limiter la réutilisation des données de la recherche rendues publiques dans le cadre de sa publication. »

Cependant, comme l'explique le conservateur des bibliothèques et juriste Lionel Maurel, la nouvelle loi protège seulement les données de la recherche de la convoitise des éditeurs et non de celle des réseaux sociaux comme *Academia*, *Research Gate* ou *Figshare* (qui appartient d'ailleurs à *MacMillan Publishers*).

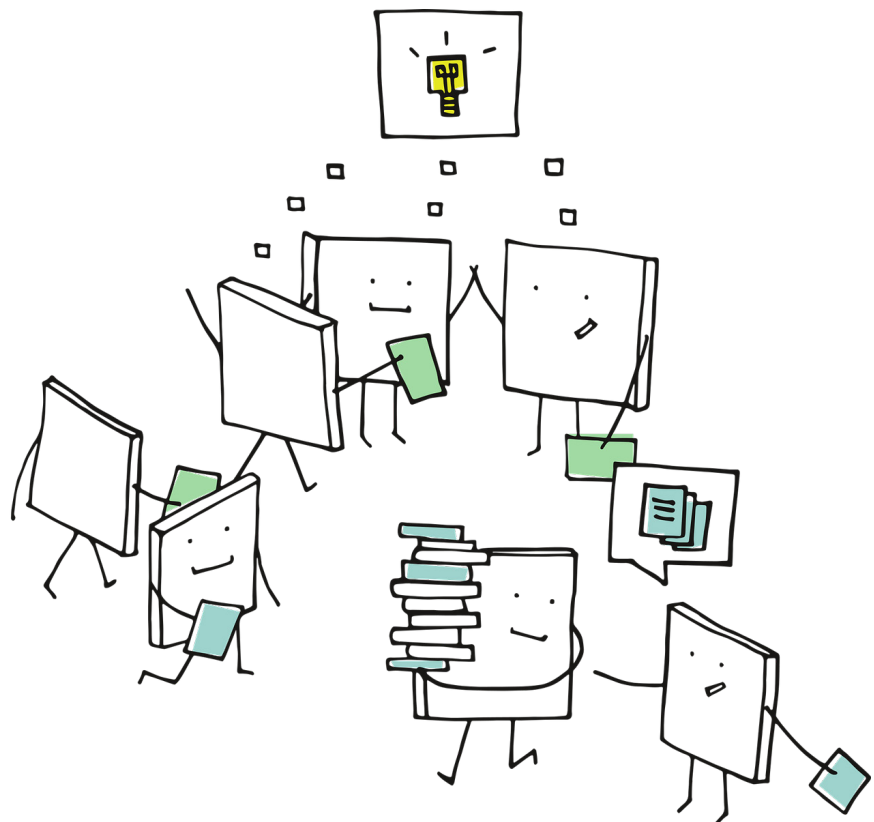
CONCLUSION

La gestion et la diffusion des données de la recherche sont devenus des enjeux centraux pour la recherche au niveau mondial et la France, avec la *loi pour une République numérique* a pris en compte les freins liés à la diffusion.

On peut s'attendre prochainement

à une inscription dans la loi d'une obligation de diffusion des « données de recherche issues de programmes financés par appels à projets sur fonds public » (et non plus d'une possibilité) comme le préconise le *Plan National pour la Science Ouverte*. Cependant, **reste à savoir si cette politique d'ouverture des données sera accompagnée de réels moyens humains et financiers** sans lesquels des infrastructures de stockage et de diffusion performantes ne pourront pas être mises en place et sans lesquels les chercheurs ne pourront être véritablement accompagnés pour atteindre l'objectif d'une Science Ouverte. Le *Plan National pour la Science Ouverte*, conscient de cette difficulté, préconise la création de « la fonction d'administrateur des données et le réseau associé au sein des établissements » ainsi que la création d'un fond pour la Science Ouverte. Affaire à suivre...

■ Sonia Salami



BIBLIOGRAPHIE

Les données de la recherche :

Le Blog des BU de l'université de Nice Sophia Antipolis, « Les données, quel enjeu pour la recherche et les bibliothèques ? ». URL : <http://bibliotheque-blogs.unice.fr/httbu/2017/01/06/les-donnees-quel-enjeu-pour-la-recherche-et-les-bibliotheques/>

GAILLARD Rémi, « De l'Open Data à l'Open Research Data : quelle(s) politique(s) pour les données de recherche ? », Mémoire pour l'obtention du Diplôme de Conservateur des Bibliothèques, 2014. URL : www.enssib.fr/bibliotheque-numerique/documents/64131-de-l-open-data-a-l-open-research-data-quelles-politiques-pour-les-donnees-de-recherche.pdf

GARVANÈSE Guillaume, « Préserver les données de la recherche à l'ère du Big Data », publié le 09.09.2016 sur *Le journal du CNRS*, URL : <https://lejournal.cnrs.fr/articles/preserver-les-donnees-de-la-recherche-a-leredubig-data>.

MÜLLER Catherine (ENSSIB), « Bibliothèque, Open Science, Open Data et données de la recherche au Canada : quels enjeux ? ». URL : <http://www.enssib.fr/recherche/enssiblab/les-billets-denssiblab/bibliotheque-de-recherche-open-access-open-data-gestion>

ROBERGE Alexandre, « Les données scientifiques en voie d'extinction. Une étude fait un portrait alarmant de la recherche scientifique dont les données se perdent annuellement », *cursus.edu*, 08/02/2015, Dernière MAJ 11/03/2015, URL : <https://cursus.edu/articles/33591/les-donnees-scientifiques-en-voie-dextinction#.W8P6p2gzblU>

FAYET Sylvie, « Données » de la recherche, les mal-nommées. URFIST, 15 novembre 2013 URL : <https://urfistinfo.hypotheses.org/2581>

VINES Timothy et al., « The Availability of Research Data Declines Rapidly with Article Age ». *Current Biology*. *Current Biology*, 24, 94–97, January 6, 2014 URL : <http://www.cell.com/current-biology/abstract/S0960-9822%2813%2901400-0#Summary>

Gestion et diffusion des données de la recherche :

ABERKANE Ourida et S. YAHIA Mohamed (Inist – CNRS), support d'intervention « Gestion et diffusion des données de la recherche », 25 novembre 2016, URL : http://www.inist.fr/IMG/pdf/support_formation_urfist_strasbourg_20161125.pdf

ANDRÉ Francis, « Déluge des données de la recherche ? Petit manuel d'immersion. Curation, infrastructures et partage », dans *Big Data. Nouvelles partitions de l'information. Actes du séminaire IST Inria, octobre 2014*, sous la direction de Calderan Lissette, Laurent Pascale, Lowinger Hélène et al. Louvain-la-Neuve, De Boeck Supérieur, Information et stratégie, 2014, Chapitre 5, p. 77-95.

CLAIROUX Nathalie, « Introduction à la gestion des données de la recherche » 2016, URL : <https://fr.slideshare.net/nclairoux/gestion-donnees-recherche-plan-de-gestion-des-donnees-archivage-prservation>

INIST, « Une introduction à la gestion et au partage de données », formation en ligne, 16 septembre 2014, URL : http://www.inist.fr/donnees/co/Donnees_recherche_web.html

INRA, « Enjeux sur la réutilisation des données ». URL : <https://www6.inra.fr/datapartage/Reutiliser/Enjeux-sur-la-reutilisation-de-donnees>

Le Blog des BU de l'université de Nice Sophia Antipolis, « Gérer et diffuser les données de la recherche: quels enjeux ? ». URL : <http://bibliotheque-blogs.unice.fr/httbu/2017/01/26/gerer-et-diffuser-les-donnees-de-la-recherche-enjeux/>

MORLOCK Emmanuelle, ROMIER Geneviève « Les enjeux du "Data Management Plan" pour les chercheurs », *scienceconf.org*, Conférence du 25/05/2016. URL :

MEDIADIX, URFIST, « Les données numériques de la recherche journée d'étude », 10 octobre 2014.

Loi pour une République numérique :

DIRECTION DE L'INFORMATION SCIENTIFIQUE ET TECHNIQUE - CNRS. *Livre blanc — Une Science ouverte dans une République numérique*. Nouvelle édition [en ligne]. Marseille : OpenEdition Press, 2016. Voir en particulier la partie « Incidence prévisibles de la loi pour une République numérique sur l'Enseignement Supérieur et le Recherche ». URL : <http://books.openedition.org/oep/1548>

Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique article 3.

MAUREL Lionel, « L'ébauche d'un statut pour les données de la recherche » sur *scinfolex.com* le 3 novembre 2016. URL : <https://scinfolex.com/2016/11/03/quel-statut-pour-les-donnees-de-la-recherche-apres-la-loi-numerique/>

Les données de la recherche, le *Dark Data* des SHS

Constance PICQUE

Le résultat est sans appel, les Sciences Humaines et Sociales (SHS) sont en retard par rapport aux sciences techniques et médicales (STM) en ce qui concerne l'ouverture des données de la recherche. Pour mieux comprendre les chercheurs en SHS, une enquête a été menée sur leurs pratiques et leurs attentes à l'heure de l'Open Research Data.



SHS face aux données de la recherche (CC0)

Face aux exigences européennes et celles d'autres agences de financement, **les établissements de Recherche et les bibliothèques françaises se mobilisent pour participer à l'élaboration de politiques en faveur du partage des données de la recherche.**

Entre 2015 et 2016, Hélène Prost (chercheuse à l'INIST¹ du CNRS), Joachim Schöpfel (maître de conférences en SIC² à l'Université de Lille 3), Alexandre Serres (maître de conférences en SIC et co-responsable de l'URFIST³ de Rennes) Marie-Laure Malingre (conservatrice des bibliothèques et co-responsable de l'URFIST de Rennes), Morgane Mignon (ingénieur d'étude), Cécile Pierre (conservatrice des bibliothèques, responsable du département recherche, SCD⁴ de Rennes 2) et Didier Collet

(Technicien de bibliothèque, Département Recherche, SCD de l'Université Rennes 2) ont réalisé **deux projets de recherche au sein des universités de Rennes 2 et Lille 3** pour faire un point sur la gestion de leurs données de recherche et leurs attentes dans le domaine des Sciences Humaines et Sociales (SHS).

AVÈNEMENT DE L'OPEN RESEARCH DATA

Depuis le début des années 2000, l'ouverture des résultats de la recherche est devenu l'objet d'une préoccupation à l'échelle européenne et internationale. A travers la déclaration de l'OCDE⁵ en 2003 sur l'accès aux données de la recherche financée par des fonds publics, la Commission Européenne a pour ambition de mettre en place, à l'horizon 2030, une infrastructure collaborative des données de la recherche, garantissant leur accessibilité, leur utilisation, leur réutilisation et leur fiabilité (Gaillard, 2014). Ainsi, pour assurer l'intégrité des jeux de données, l'ensemble des agences de financement public imposent aujourd'hui une politique de gestion des données par le biais notamment du DMP⁶ (*Data Management Plan*).

Bien que l'objectif de départ était d'analyser les pratiques des chercheurs en SHS, les deux enquêtes ont permis de mettre en exergue l'insuffisance des infrastructures et les services auprès des universités de Rennes 2 et Lille 3 pour répondre à l'ambition des agences de financements françaises.

NATURE DE LA DONNÉE

Qu'elle soit brute (*raw data*), d'observation, traitée (*processed*) ou dérivée (*derived*), la donnée de recherche, « *Datasets* » en anglais, est un enregistrement factuel qui peut contenir des informations qualitatives ou quantitatives nécessaires à la validation des résultats d'un projet de recherche.

Tout comme la publication scientifique, les données appartiennent à des produits de la recherche. Pour illustrer cette définition,

les deux enquêtes nous donnent un aperçu des données manipulées dans le domaine des SHS.

« LE MAÎTRE MOT [...] C'EST LA DIVERSITÉ »

Tout d'abord, les chercheurs collectent ou créent des données très diverses telles que les données sources avec un trio de tête : les données personnelles, très largement cités (72 % des répondants à Rennes et 64 % à Lille), les données d'enquêtes (53 % des répondants à Rennes et 47 % à Lille) puis les données observations (43 % contre 41 % à Lille) (figure 1). Pour les données dites de résultats, les données textuelles occupent le premier rang avec 71 % pour Rennes et 75 % pour Lille. Le rapport sur l'enquête de l'Université de Lille révèle également que **les chercheurs ont parfois des difficultés à délimiter la frontière entre les sources utilisées et les données produites.** Ainsi, quelles que soient les disciplines concernées, de multiples types

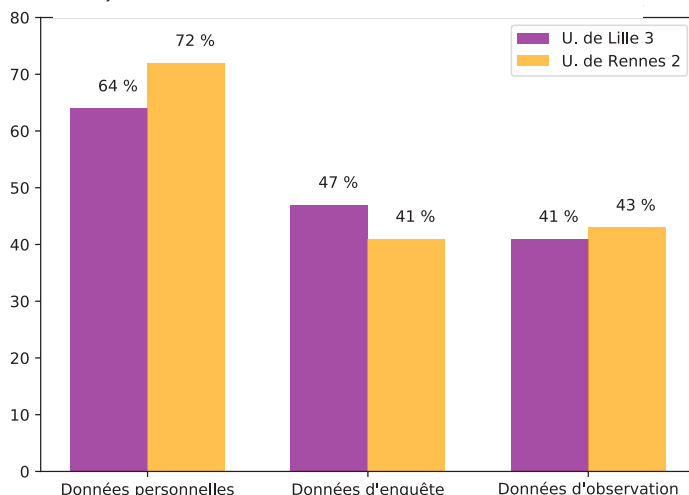


Figure 1. Principales données «Sources» © C. PICQUE

de données sont manipulées au quotidien dans les laboratoires. Pour autant la définition donnée par l'OCDE reste encore ambiguë aux yeux des chercheurs, qui sont pourtant les principaux producteurs de la donnée de la recherche.

PRATIQUES DE STOCKAGE ET D'ARCHIVAGE

La diversité des données pose un certain nombre de problématiques. Les deux enquêtes ont mis en évidence des difficultés sur le plan de la conservation des données :

- **Nécessité d'avoir des espaces de stockage plus importants.** L'émergence des humanités numériques accroît les besoins de stockage. De plus, le besoin d'espace de stockage entre les enseignants-chercheurs est différent. Puisque certains réclament plus d'espace, cela suppose la mise en place d'un service personnalisé de stockage adapté aux besoins de chacun. Or, ce n'est pas la politique mise en œuvre actuellement au sein des deux universités.

- Le papier n'a pas disparu. Outre le contraste volumétrique entre les disciplines, toutes les données ne sont pas numériques parce que le support papier n'a pas totalement disparu dans les pratiques. **Cela suppose que l'accès libre à l'ensemble des données n'est pas possible pour l'intégralité des données.**

- L'utilisation de nombreux logiciels et de formats de données propriétaires soulèvent des **problèmes liés à l'archivage à long terme.**

- Un stockage principalement local et individuel. Les sauvegardes donnent l'impression d'une pratique personnelle, privée plus souvent que professionnelle avec les moyens du

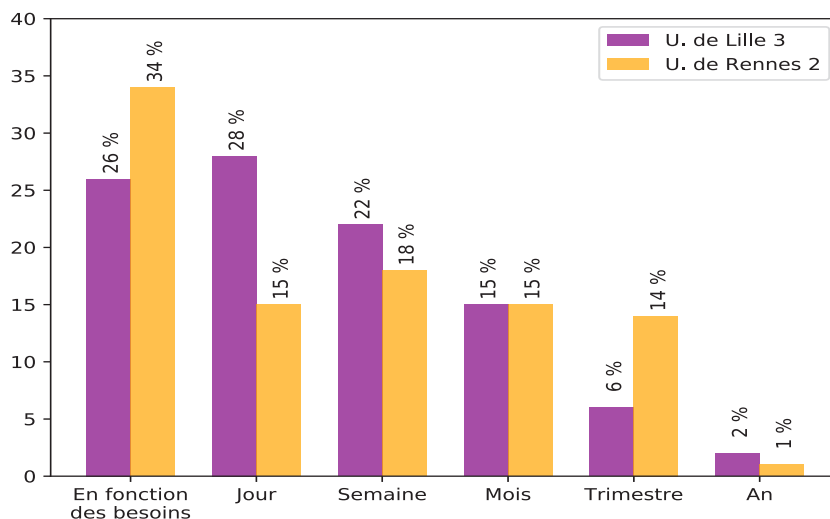


Figure 2. Fréquence de sauvegarde des données © C. PICQUE

bord (disque dur, clé USB) mais avec un certain souci de sécurité (stockage à plusieurs endroits et sur plusieurs supports, sauvegardes plus ou moins régulières) [extrait d'enquête de Lille 3]. En effet, le stockage en local sur ordinateur privée ou professionnel est le mode de sauvegarde le plus fréquent. Parmi les participants, rares sont ceux qui mentionnent un stockage sur le web (dans le Cloud). De même que les personnes interrogées mentionnent très majoritairement l'utilisation de disques durs externe et de clés USB. Le problème soulevé ici est le **risque de perte et le faible niveau de sécurité des données** contrairement au stockage institutionnel et en réseau. Pour autant, le recours au stockage sur le Cloud institutionnel (27 % pour Rennes 2 et 40 % pour Lille 3) semble progressivement se développer d'après l'enquête sur l'Université de Rennes 2.

- **Les fréquences de sauvegardes sont chronophages.** Comme nous montre la figure 2, les sauvegardes sont faites de manière régulière mais variable selon les personnes interrogées. Globalement, l'enquête révèle une absence de politique en terme

de conservation des documents sauf dans le cas de données considérées comme lourdes.

PRATIQUES DE PARTAGE ET DE DIFFUSION

Selon Jane H. Smith (2013), **les chercheurs ne savent tout simplement pas comment partager leurs données.** En SHS, les chercheurs ont tendance à se regrouper en équipe, par incitation ou par choix, que ce soit au sein de leur unité

de recherche ou dans le cadre de projets de recherche, pour autant les deux enquêtes nous montrent que **leurs données sont encore largement inaccessibles et réservées.** Comme le montre la figure 3, la majorité ne partage pas ses données avec d'autres (54 % des répondants à Rennes, 64 % à Lille). Et ceux qui le font, partagent d'abord et surtout avec les collègues du groupe de travail (équipe scientifique) (42 % des répondants à Rennes et 34 % à Lille). Les enquêtes relèvent que les chercheurs ayant déjà déposé leurs données de recherche en libre accès sur le web sont une minorité (16% à Lille), que seul la moitié d'entre eux ont déjà téléchargé les données ou donné l'accès

« L'ARCHIVAGE DES DONNÉES SEMBLE ENCORE INEXISTANT »

d'autres chercheurs.

Par ailleurs, il y a des différences constatées au niveau des pratiques entre les deux universités. **Les enseignants-chercheurs de l'Université de Lille 3 sont nombreux à déclarer ne pas se sentir prêts pour déposer l'ensemble de leurs résultats, ce qui n'est pas le cas pour les chercheurs de l'Université de Rennes 2 qui seraient plutôt d'accord pour diffuser leurs données de recherche en libre accès.** Ce positionnement est nuancé par différentes interrogations (l'idée d'un mouvement jugé irréversible, des interrogations quant aux modalités du partage et des réserves face à l'idée de « tout-partagé »).

Pour autant, **l'ensemble des chercheurs s'accorde à dire que les aspects juridiques liés à la nature de leurs données, le manque de temps et la crainte de plagiat les freineraient.** L'enquête menée à l'Université de Lille 3 précise les deux premières raisons évoquées : l'illisibilité des données brutes (autrement dit, la nécessité de passer du temps pour les rendre lisibles) ou tout simplement la protection de la propriété intellectuelle : « ces données m'appartiennent »... En vue de ces résultats, le partage des don-

nées en Libre Accès est encore loin de faire partie de la pratique ou de la préoccupation courante des chercheurs en SHS. Ce phénomène s'explique également par les spécificités des données scientifiques en SHS : les enjeux et les problématiques ne sont pas les mêmes qu'en STM (Sciences Techniques et Médicale).

BESOINS ET ATTENTES

Bien qu'il soit de la responsabilité des chercheurs de conserver et de diffuser leurs données, ces derniers ne semblent pas prêts à s'inscrire dans une démarche scientifique plus ouverte et transparente comme l'explique l'enquête menée à l'Université de Rennes 2. **Les répondants seraient prêts à rendre leurs données de recherche accessibles en Libre Accès si cela donnait une meilleure visibilité à leurs travaux de recherche et que par conséquent cela leur permettait de créer de nouveaux contacts au sein de la communauté scientifique.** L'adhésion aux valeurs du Libre Accès aux résultats de la recherche scientifique

« TRÈS PEU
OUVRENT LEURS
DONNÉES [...] DANS
UNE DÉMARCHE
OPEN DATA »

n'arrive qu'en troisième position. En termes de services, d'infrastructures et d'outils, les conclusions tirées par les deux études donnent **des pistes de réflexion aux décideurs** :

- Tout d'abord émerge **un besoin technique avec la nécessité d'un archivage et d'un espace de travail collaboratif sécurisé.** Selon les deux enquêtes, les chercheurs souhaiteraient obtenir avant tout un espace d'archivage sécurisé et fiable (54 % des répondants à Lille et 57 % à Rennes) pour le dépôt de leurs données, tel que les Archives Ouvertes comme HAL ou des plateformes spécialisées, plutôt que le serveur de leur université de rattachement ou encore les réseaux sociaux et les blogs. Sur le plan de la sauvegarde des données, ils souhaite-

raient également pouvoir stocker leurs données de recherche au-delà de la durée de vie d'un projet. D'autre part, les chercheurs semblent d'avis qu'il faut mettre en place un plan de gestion de données (DMP). Bien que les chercheurs interrogés déclarent ne pas connaître ce dispositif, l'enquête de Rennes 2 nous montre qu'ils ne sont pas contre cette démarche institutionnelle, sans pour autant accorder une importance à une politique d'établissement pour les métadonnées, ni de système de nommage des fichiers.

- Ensuite, la seconde piste pour les décideurs est **le besoin de services en collaboration avec des techniciens, des professionnels de l'IST ou encore des informaticiens.** Ce besoin se caractérise par une nécessité de sensibilisation au paysage numérique, de formation et de conseils spécialisés. En effet, comme nous le montre la figure 4, les services de conseil sont très plébiscités, particuliè-

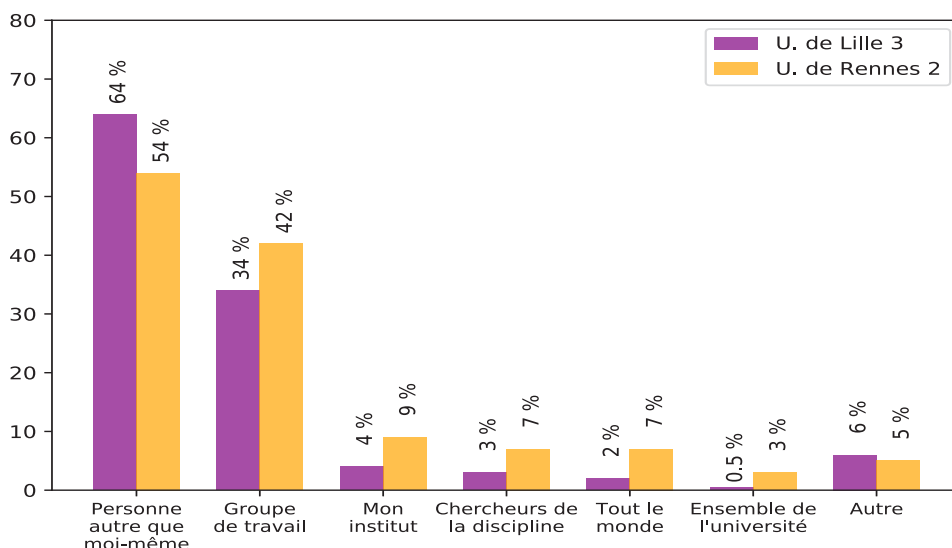


Figure 3. Partage des données © C. PICQUE

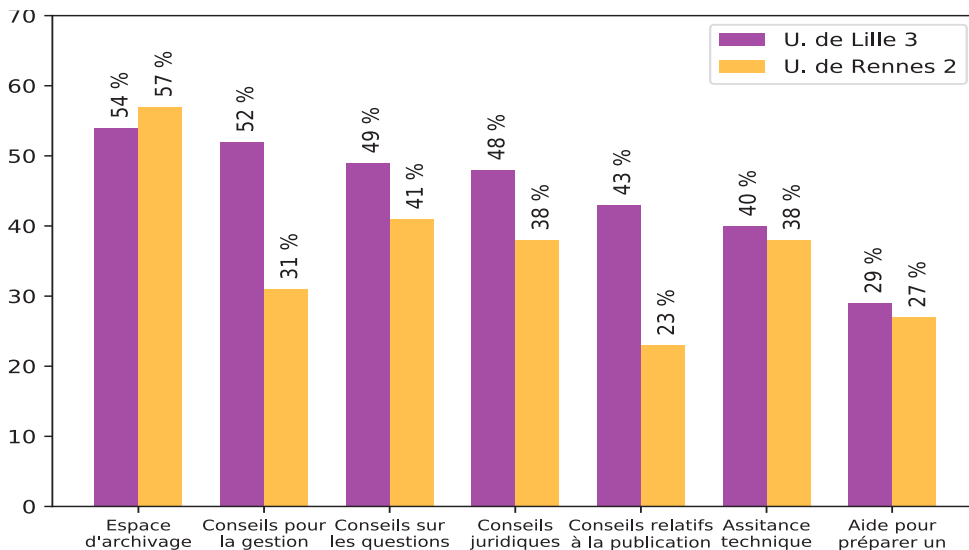


Figure 4. Attentes et besoins des chercheurs © C. PICQUE

rement ceux portant sur des questions techniques et juridiques. Les conseils relatifs à la publication des données ne semblent pas être une priorité.

BILAN

A travers cette synthèse, nous comprenons mieux pourquoi les SHS sont en retard par rapport aux STM en ce qui concerne l'ouverture des données de la recherche. Entre « Culture du partage », « Culture du libre accès » et « Culture de l'IST* », le chemin est encore long mais n'est pas impossible. Il est du ressort des universités et des professionnels de l'information d'accompagner les chercheurs dans une

démarche adaptée et durable.

Comme nous l'avons montré, la tendance est la même dans les deux universités où s'est déroulée l'enquête.

Les enjeux de l'Open Research Data en SHS concernent principalement les dispositifs organisationnels et techniques qui sont à mettre en place au sein des organismes de recherche ainsi que l'accompagnement des chercheurs. En sciences de la terre par exemple, il existe déjà des serveurs de stockage et de partage de données comme Georoc (<http://georoc.mpch-mainz.gwdg.de/georoc/>).

■ Constance PICQUE

Bibliographie

GAILLARD Rémi, « De l'Open Data à l'Open Research Data : quelle(s) politique(s) pour les données de recherche ? », Mémoire pour l'obtention du Diplôme de Conservateur des Bibliothèques, 2014. URL : www.enssib.fr/bibliotheque-numerique/documents/64131-de-l-open-data-a-l-open-research-data-quelles-politiques-pour-les-donnees-de-recherche.pdf

PROST Hélène, SCHÖPFEL Joachim, *Les données de la recherche en SHS. Une enquête à l'Université de Lille 3. : Rapport final* [Rapport de recherche], Lille 3. 2015. Lien : <https://tel.archives-ouvertes.fr/UNIV-LILLE1/hal-01198379v1>

SERRES Alexandre, MALINGRE Marie-Laure, MIGNON Morgane, PIERRE Cécile, COLLET Didier, *Données de la recherche en SHS. Pratiques, représentations et attentes des chercheurs : une enquête à l'Université Rennes 2* [Rapport de recherche], Université Rennes 2. 2017. URL : <https://hal.archives-ouvertes.fr/hal-01635186>

1 INIST : Institut de l'Information Scientifique et Technique

2 SIC : Science de l'information et de la Communication

3 URFIST : Unité Régionale de Formation à l'Information Scientifique et Technique

4 SCD : Service commun de documentation

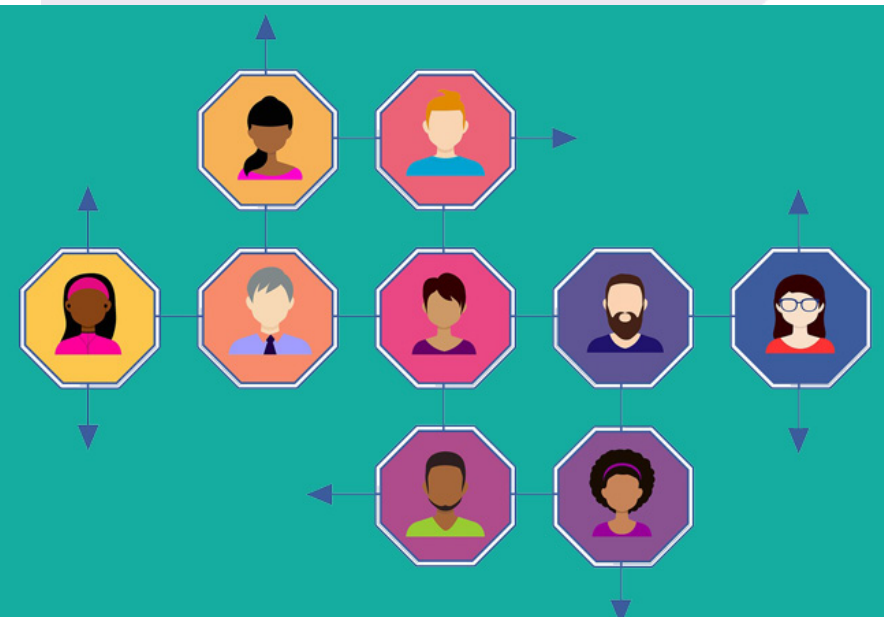
5 OCDE : Organisation de coopération et de développement économiques

6 Qu'est-ce que le DMP ? Retrouvez la définition dans l'article intitulé "Big Data et données de la recherche" paru le 5 juin 2017 dans le magazine Didaktik

Le partage des données de la recherche dans la Pratique

Laure PIERRAT

Le chercheur n'en a pas toujours conscience mais, au cours de ses projets de recherche, il va produire une très grande quantité de données. Si ces données se révèlent être d'une importance capitale tant pour la validation des résultats, que pour leur réutilisation dans de futurs projets de recherche, il est intéressant de se pencher sur la question de la pratique du partage des données par les chercheurs. Y'a-t-il un décalage entre les bonnes pratiques et la réalité du terrain ? Quels sont les leviers pour encourager le partage ?



Le partage des données de la recherche (CC0)

Le partage des résultats de la recherche en *Open Access* semble à présent bien ancré dans le quotidien des chercheurs. Pourtant, il n'en est pas de même pour le partage des données de la recherche qui fait l'objet de politiques publiques plus récentes. A travers un exemple concret, nous allons vous faire découvrir ce que peut être un projet de partage de données en *Open Access*. Ensuite, nous aborderons le partage des données de la recherche dans la pratique et enfin nous nous intéressons aux leviers disponibles pour inciter les chercheurs à davantage partager leurs données.

EXEMPLE CONCRET D'UN PROJET DE PARTAGE DE DONNÉES EN OPEN ACCESS

L'Université Lyon 1 dispose de l'un des plus importants herbiers au monde après celui d'Harvard. Cette collection fait actuellement l'objet d'un travail de numérisation¹.

Quelles données ?

Ce projet illustre bien les formes variées que peuvent revêtir les données de la recherche. Dans le cas présent, il s'agit de la mise en ligne de la collection du prince Roland Bonaparte (1858-1924), qui compte trois millions de plantes séchées provenant du monde entier et qui sont âgées de plus

de 250 ans pour les plus anciennes. Les données partagées ici se révèlent être bien différentes de données chiffrées, de textes, ou d'autres données généralement disponibles.

Partager, pourquoi ?

- **Pour valoriser un patrimoine scientifique unique** : comme l'explique très simplement la directrice de l'herbier de l'Université Claude Bernard Lyon 1 : « *Nous avons l'impression d'héberger des trésors (...) Mais les gens ne connaissent pas cet herbier, ne savent pas où il est, à quoi il sert* ».
- **Pour ouvrir au plus grand nombre l'accès à ces données** et ainsi favoriser un travail collaboratif, multidisciplinaire, sans barrière entre scientifiques et amateurs éclairés, ni même frontière.

Des données pour qui ?

- **Pour le grand public.**
- **Pour des chercheurs** dans d'autres domaines de recherche : cet herbier n'est pas uniquement destiné à la botanique. En effet, Pierre Joly, professeur des universités à Lyon 1, a réutilisé ces données au cours de ses recherches sur le climat. Dans un spécimen végétal très ancien, l'analyse de l'oxygène encapsulé permet de déterminer le climat des siècles passés.

LE PARTAGE DES DONNÉES DANS LA PRATIQUE

Aujourd'hui on estime que 90 % des données de la recherche ne sont pas réutilisables³. Ce même constat est également établi par Hélène Prost et Joachim Schöpfel au cours de leur enquête sur les données de la recherche⁴. Il apparaît que les deux tiers des répondants ne partagent pas leurs données. Les auteurs insistent sur le fait que ceux qui partagent le font prin-

cipalement « *avec les collègues du groupe de travail (équipe scientifique) (34 %). Très peu (< 5 %) ouvrent leurs données davantage et partagent avec l'institution, d'autres chercheurs ou « tout le monde », dans une démarche Open Data au sens strict du terme.* »

Refus de partage ou méconnaissance ?

Dans certains cas il y a une réelle réticence au partage, ainsi lors de l'enquête sur les « *données de la recherche en SHS* » à l'Université Rennes 2⁴, à la question « *avez-vous déjà stocké vos données en libre accès sur le web* » 30 % des chercheurs interrogés répondent « *non et je ne le ferai pas* ».

Il faut dire qu'une majorité des chercheurs se trouvent démunis face à cette pratique. La plupart des chercheurs expriment un désintérêt pour des tâches perçues comme étant administratives et chronophages. C'est aux professionnels de l'information scientifique et technique de former les chercheurs à ces tâches qui sont généralement réalisées par ces derniers sans qu'ils s'en rendent réellement compte. En effet, pour exploiter leurs données ils doivent forcément les documenter, les renommer, les stocker pour les retrouver.

D'ailleurs, une majorité de chercheurs reconnaissent un manque de connaissances par rapport à la gestion de leurs données. Il s'agit alors de proposer une montée en compétences des chercheurs en leur présentant les méthodes, les standards, mais aussi les outils qui sont disponibles.

Ce qui permettrait de mettre fin à de mauvaises pratiques pouvant entraîner jusqu'à la disparition totale des données de la recherche.

De mauvaises pratiques de gestion

La bonne gestion des données demande une bonne méthodologie et cela dès le début du projet de recherche. Il est conseillé, voire même imposé dans certains cas – notamment lors de la participation à des projets financés –, de procéder à la rédaction du **plan de gestion des données** (*Data Management Plan* : DMP)⁵. Le DMP est un document qui décrit la collecte, le traitement, la documentation, le stockage, l'archivage, la conservation et le partage des données pendant et après un projet de recherche. C'est donc un outil qui va permettre au chercheur de planifier le dépôt de ses données et faciliter son travail. **Afin de simplifier la rédaction d'un tel document, un outil est disponible depuis novembre 2016 : il s'agit de DMP OPIDoR⁶**. Il permet aux chercheurs et à ses collaborateurs de rédiger simplement un DMP, un service d'accompagnement donne des conseils et met à disposition des exemples et propose des modèles adaptés au format institutionnel.

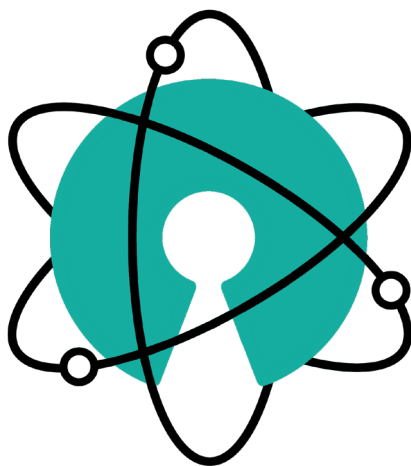
Ensuite, **les jeux de données peuvent être déposés dans des entrepôts**. Le choix de l'entrepôt se fait en fonction des recommandations des parties prenantes du projet s'il y en a, sinon le chercheur doit faire un choix en fonction :

- **De la discipline** : il existe des entrepôts spécialisés qui sont propres à une institution ou à une discipline, des entrepôts multidisciplinaires ou thématiques,

- **Des besoins du chercheur** : a-t-il besoin d'un entrepôt certifié ou seulement d'un entrepôt de confiance reconnu par la communauté scientifique ? Y'a-t-il une nécessité d'avoir un entrepôt hébergé dans tel ou tel pays ? L'entrepôt permet-il de stocker tous types de données ? En effet, certains n'acceptent que des données

liées à des publications (Ex. DRYAD), d'autres acceptent tous les jeux de données (Ex. ZENODO).

Les mauvaises pratiques s'illustrent particulièrement dans le domaine de la conservation des données : Le disque local fait encore aujourd'hui figure de grand favori pour le stockage des données (Prost, 2015).



En ce qui concerne le partage des données en lui-même : plus de 50% des chercheurs déclarent ne pas partager leurs données (Serres, 2017). Ceux qui franchissent le pas privilégient les e-mails avec des pièces jointes, les échanges via serveurs locaux, les échanges via Dropbox ou Google Drive et enfin via clé USB. Ces supports ne garantissent pas la pérennité des données.

Manque de formation et non recours aux professionnels de l'IST

Les chercheurs sont très demandeurs de formation par rapport à la gestion et au partage des données. Cependant, le principal obstacle réside souvent dans la **mauvaise compréhension du périmètre de la gestion de l'information et dans la méconnaissance du rôle des professionnels de l'IST**.

En effet, selon Cherifa Boukacem⁷, dans le domaine des sciences exactes, seuls 9 % des chercheurs se disent formés à la gestion des données. De plus, il apparaît que c'est généralement grâce à leur mobilité qu'ils ont gagné en compétences (par exemple certains se sont formés à la NASA, ou encore par l'intermédiaire de la Bibliothèque du MIT).

D'autre part, toujours selon Cherifa Boukacem, lorsqu'on les interroge, 52 % des chercheurs se disent insuffisamment informés mais **seulement une moitié d'entre eux se disent favorables à une aide de la part des professionnels de la gestion de l'information** : notamment sur les questions d'archivage. Pour ceux qui se disent non favorable à une aide, le manque d'expertise technique des documentalistes vient généralement au premier plan. Rappelons que le professionnel de l'IST n'intervient pas au moment de la production des données par le scientifique mais intervient seulement après, en partenariat, afin d'en faciliter la conservation. Pourtant il s'agit précisément du rôle des professionnels de l'IST qui ont donc **un travail de pédagogie conséquent à mener auprès des chercheurs**.

MOYENS D'AGIR : COMMENT INCITER LES CHERCHEURS À PARTAGER ?

Mise en place d'obligations de dépôt :

Les politiques publiques actuelles, tant au niveau national qu'au niveau européen et mondial encouragent l'ouverture des données de la recherche. Elles souhaitent par ces politiques favoriser l'innovation : les institutions y voient la possibilité d'un retour sur investissement notamment pour les TPE ou pour les PME qui disposeront d'un accès libre à certaines données et pourraient les réutiliser. Il

est important également de rappeler que l'innovation est un vecteur de croissance et d'emploi. Ainsi, à titre d'exemple la cartographie du génome humain va à la fois permettre de progresser dans le domaine médical, mais surtout a permis de créer par le libre accès à l'information 310 000 emplois⁸.

Cet engouement des politiques pour l'*Open Science* s'explique également par la volonté de réduire les coûts de la recherche en les mutualisant : il est moins coûteux de réutiliser des données existantes plutôt que de les recréer. N'oublions pas qu'un même jeu de données peut servir à plusieurs disciples. **C'est pourquoi des obligations de dépôt ont été mises en place dans le cadre de projets financés.** C'est le cas pour les projets financés par l'EU dans le cadre du programme H2020 : les chercheurs doivent obligatoirement déposer leurs données dans un entrepôt.

D'autre part, dans le cadre d'une publication, le dépôt peut être une condition obligatoire afin de garantir l'accès aux données sources. C'est ainsi que de plus en plus de revues demandent un accès aux données pour améliorer les conditions dans lesquelles les comités de lecture valident les articles, c'est le cas par exemple de l'éditeur PLOS ONE.

EN RÉSUMÉ

Le partage en *Open Access* des productions scientifiques dans le but de valoriser les résultats de la recherche est à présent bien assimilé par les chercheurs qui l'ont largement adopté. Cependant, il n'en est pas de même pour le partage des données de la recherche. En effet, **l'ouverture des données de la recherche est un mouvement récent qui oblige les chercheurs à repenser en profondeur leurs manières de travailler.** Malgré le fait que différentes institutions encouragent l'ouverture des données, force est de constater qu'il y a une réelle nécessité de clarifier la situation pour ne pas laisser le chercheur seul face à des problématiques de gestion documentaire. C'est pour les professionnels de l'information scientifique l'occasion de se positionner en tant qu'intermédiaire en proposant une assistance aux chercheurs.

■ Laure PIERRAT

- 1 Catherine Lagrange. « Lyon numérise l'un des plus grands herbiers universitaires au monde ». *Le Parisien.fr*, [en ligne] 13/03/2018.
- 2 Nord Littoral. « À Lyon, un herbier numérique pour partager la connaissance des plantes ». 13/03/2018 [En ligne]
- 3 DORANUM. « Zoom sur » [en ligne]. Disponible sur : < <http://doranum.fr/> >
- 4 Pour une analyse plus en détail des causes du non partage des données de la recherche en SHS voir dans ce numéro l'article de Constance PIQUE « Les données de la recherche, le dark data des SHS »
- 5 Pour plus de détail sur le DMP voir dans ce numéro l'article de Sonia SALAMI « La gestion des données de la recherche et leur diffusion, quels enjeux ? »
- 6 DMP Opidor signifie : Data management Plan pour une optimisation du partage et de l'interopérabilité des données de la recherche.
- 7 Cherifa Boukacem. Données de la recherche et contexte des pratiques des chercheurs : injonctions, risques, opportunités. 14/11/2017. Journée d'étude intitulée : « Partager les données de la recherche : pour qui, pourquoi, comment ? » [en ligne].
- 8 Commission Européenne « Communication de la commission au parlement européen, au conseil, au comité économique et social européen » [en ligne]



Open Research Data : le rôle des universités dans l'accompagnement des chercheurs

Carole-Anne MATEL

Les chercheurs de l'enseignement supérieur sont à l'origine du cycle de vie des données car ce sont eux qui les collectent, les produisent, les manipulent et en connaissent la valeur. Face à la dynamique mondiale d'ouverture des données de la recherche (Open Research Data) qui bouscule leurs pratiques, comment les universités peuvent-elles les accompagner dans la gestion et la diffusion de leurs données en libre accès ?



Accompagnement des chercheurs dans les universités (CC0)

LA POLITIQUE DE GESTION DES DONNÉES DANS LES UNIVERSITÉS

Qu'est-ce qu'une politique de gestion des données ?

Au niveau des universités, il a été recommandé par l'OCDE¹, la Commission Européenne et le Ministère de l'enseignement supérieur, **la mise en place de politiques concernant la gestion des données, appelées également *Data Management policy*.**

Ce document est un guide contenant les étapes du projet et un outil de mise en place de la gestion des données de la recherche dans l'institution. **Il doit définir de façon claire des objectifs à atteindre et « des actions à mettre en œuvre »** en prenant en compte toutes les disciplines de l'établissement, les différents types de données produites, les normes actuelles de diffusion, les contraintes, etc.

Pour le rédiger et assurer son utilisation, la participation de la gouvernance², des professionnels de l'Information Scientifique et Technique (IST) et des chercheurs est à mobiliser (Gaillard, 2014).

Partager les responsabilités, l'exemple de Monash

A l'étranger et depuis de nombreuses années, les universités se sont penchées sur la réalisation d'une politique institutionnelle en matière de gestion des données, comme l'Université de Monash, en Australie.

A l'Université de Monash, l'élaboration de la politique de gestion des données a duré 10 ans. **La stratégie adoptée a été celle du partage des responsabilités entre les chercheurs et l'université.** Ainsi le *Data Management policy* contient une série de recommandations à destination des chercheurs qui doivent mettre à disposition leurs données. En contrepartie, l'université s'engage à les accompagner, à les former et à sauvegarder leurs données dans une infrastructure ou du moins à répertorier leurs données dans un catalogue de métadonnées (Bellamy, 2014, Gaillard, 2014).

Cet exemple met en lumière la nécessité de partager les responsabilités entre les différents acteurs, afin d'assurer un déploiement efficace de cette politique.

L'importance de l'audit

Pour pouvoir rédiger une politique de gestion des données, **il faut tout d'abord évaluer l'existant, ce qui passe par la réalisation d'un audit.** L'audit consiste à analyser les pratiques et les besoins en matière de gestion des données dans l'université, et permet donc d'en apprendre davantage sur les habitudes des chercheurs. A la suite des résultats obtenus, des actions pourront être mises en place afin d'améliorer les éventuels dysfonctionnements repérés.

Il existe au Royaume-Uni un outil, le *Data Asset Framework* (www.data-audit.eu) qui « fournit aux organisations les moyens d'identifier, de localiser, de décrire et d'évaluer la façon dont elles gèrent leurs actifs de données de recherche ». Cette solution offre également des exemples de méthodes d'enquêtes et d'entretiens réalisés par d'autres universités comme l'Université de Glasgow et Southampton (Gaillard, 2014).

Du côté français, deux enquêtes sont actuellement disponibles sur Hal (pour plus de détail, voir article précédent de Constance Picque « Les données de la recherche, le dark data des SHS ») :

- « Données de la recherche en SHS. Pratiques, représentations et attentes des chercheurs : une enquête à l'Université Rennes 2 »
- « Les données de la recherche en SHS. Une enquête à l'Université de Lille 3 »

Le public concerné correspond à la communauté des chercheurs en Sciences Humaines et Sociales. Ces enquêtes ont permis d'établir un état des lieux de la situation, de dégager d'éventuels services que les universités pourraient proposer aux chercheurs³. Elles mettent aussi en avant un problème récurrent : le non partage de données. Les raisons évoquées sont diverses, éthiques, juridiques, voire personnelles... Mais certaines disciplines partagent davantage leurs données (logique du *bottom-up*) notamment « les communautés œuvrant dans la génomique, la cristallographie, l'astronomie... » (Bellamy, 2014, p.4).

L'explication est ici historique puisque ce sont ces dernières qui ont commencé à partager leurs données sur le Web. A l'opposé, les Sciences Humaines et Sociales sont les moins enclines à cette notion de partage : certains acteurs ne



partagent pas leurs informations avec d'autres et, s'ils le font, c'est seulement avec leur groupe de recherche. Et seulement une minorité a déjà auto-archivé ses articles sur les archives ouvertes, comme le montre l'enquête de Rennes 2 et de Lille 3.

On voit donc que le partage des données n'est pas encore rentré dans les pratiques de tous les chercheurs et pour les évaluateurs, la phase d'audit est également une occasion pour les sensibiliser. Cependant cela n'est pas suffisant et il est capital d'accompagner la communauté des chercheurs dans ses nouvelles pratiques.

LA FORMATION DES CHERCHEURS

Le bibliothécaire, un formateur au plus près des chercheurs

Parmi tous les professionnels et toutes les institutions de l'IST (URFIST, INIST-CNRS, BSN, etc.), il y a un acteur à prendre en compte : le **bibliothécaire universitaire**. En effet, l'*Open Access* lui donne de nouvelles missions.

Les chercheurs et les enseignants-chercheurs fréquentent de moins en moins les Bibliothèques Universitaires et les Services de documentation. Ils préfèrent utiliser leur poste de travail pour consulter les ressources en ligne que proposent les universités. **L'Open Access est l'occasion pour le bibliothécaire de recréer du lien**

avec les chercheurs, car il possède notamment des compétences sur la gestion des données (Cheval, 2015). Il est donc l'un des acteurs majeurs de la politique des données de la recherche.

Rôle du bibliothécaire universitaire :

- Sensibiliser les chercheurs au libre-accès.
- Aider à l'évaluation de l'existant en identifiant les pratiques et les besoins des chercheurs.
- Apporter un soutien aux chercheurs pour la rédaction du Plan de Gestion des Données (PGD) ou *Data Management Plan* (DMP)⁴.
- Indexer les données.
- Conseiller les chercheurs sur le choix des lieux de dépôts.
- Conseiller les chercheurs sur les questions de droit d'auteur.
- Former les chercheurs aux bonnes pratiques de gestion des données⁵ et/ou leur proposer des stages à l'INIST-CNRS (Institut de l'information scientifique et technique du Centre national de la recherche scientifique) ou aux URFIST (Unités Régionales de Formation à l'Information Scientifique et Technologique).

Certaines de ces activités sont assez familières au bibliothécaire universitaire, d'autres sont nouvelles. Si celui-ci veut parfaire ses compétences et se former, il peut comme les chercheurs s'adresser à l'INIST-CNRS ou aux URFIST qui proposent des ateliers sur la gestion des données.

DoRANum : un outil de formation à distance

Le projet *doranum.fr* (DoRANum : « Données de la Recherche : Apprentissage NUMérique à la gestion et au partage ») a vu le jour pour **répondre au besoin de formation des cher-**

cheurs en gestion et en partage de données. C'est un outil qui peut être utilisé par les bibliothécaires universitaires afin de se former mais aussi afin de former les chercheurs.

DoRANum est une plate-forme de formation à distance ouverte, où l'utilisateur n'a pas besoin de s'identifier au préalable. Financée par le Ministère de l'enseignement supérieur de la recherche et de l'innovation au titre de la BSN (Bibliothèque Scientifique Numérique) segment « 9 » (Formation, compétences et usages) et « 10 » (Données de la recherche), il est copiloté par l'URFIST de Nice et l'INIST-CNRS. La plateforme regroupe différentes ressources de formation et d'auto-formation. Les documents sont trouvés ou créés pour elle, par les partenaires du projet comme l'Université de Nice-Sophia-Antipolis, l'Inra... Ils prennent la forme de quiz, MOOC...



Logotype de DoRANum

Les publics visés pour ce programme sont les « *enseignants-chercheurs, les chercheurs, les doctorants, les professionnels de l'IST et de soutien (documentalistes, archivistes, bibliothécaires, ingénieurs)* » (DoRANum-GRICAD, 2017, p.4). Avec son format, DoRANum peut « *s'inclure dans un programme de formation en présentiel et compléter des formations proposées* » dans les établissements d'enseignement supérieur. La boîte à outils « *formateur* » permet aux professionnels de l'information d'organiser des travaux pratiques à destination de la communauté des chercheurs. Ainsi **le chercheur peut soit travailler en autonomie en se formant seul sur**

la plateforme, soit suivre une formation mêlant du présentiel et distanciel.

Les ressources proposées sont classées par thématiques et outils. Les thématiques correspondent aux grandes questions que l'on peut se poser sur le libre accès des données de la recherche. Toutes sont traitées de manière ludique et synthétique au vu du peu de temps que certains publics peuvent y consacrer. Il y a pour chacune d'entre elles, une fiche synthétique avec une vidéo qui reprend le sujet de manière concise, et un court questionnaire. Les vidéos et les fiches des différentes thématiques peuvent être téléchargées, imprimées ou réutilisées (Licence CC).

Ce projet DoRANum étant toujours en cours de développement, il fournira à l'avenir encore bien d'autres supports à ses (futurs) utilisateurs, dont des distanciels avec les classes virtuelles, des webinaires, des tutoriels multimédias... (DoRANUM, 2017).

Conclusion

Afin d'accompagner au mieux les chercheurs pour qu'ils puissent rendre accessibles et réutilisables leurs données, **l'université doit organiser et systématiser leur mise à disposition** (Prost, Schöpfel, 2015). La mise en place d'une **politique de gestion des données** est une solution qui a déjà fait ses preuves dans de nombreux pays notamment au Royaume-Uni, aux États-Unis et en Australie, ainsi que dans de nombreux organismes de recherche français comme à l'INRA, au CNRS, à l'INRIA... (Gaillard, 2014 ; Serres, 2017).

Effectivement, **il faut avant tout mobiliser plusieurs acteurs** dont la gouvernance et les professionnels de l'IST lors de sa mise en œuvre pour qu'elle



Formation à la gestion des données (CC0)

puisse fonctionner. Il est nécessaire de créer une dynamique et responsabiliser les deux parties (université-chercheur). Il est également important de ne pas négliger **la phase de l'audit** qui est un moyen de communiquer et de commencer à sensibiliser les chercheurs sur l'*Open Research Data*. **La création d'une politique de gestion des données est complexe**, car il faut prendre en compte les attentes, les freins et mesurer la faisabilité, afin d'avoir un outil opérationnel. Une fois la politique déployée, l'action principale consistera à **former et à sensibiliser le public**. Les universités pourront s'appuyer sur différents outils internes et externes. **Quant aux bibliothécaires, ils apporteront certaines de leurs compétences mais devront également envisager de se former pour assurer leurs nouvelles missions**. Les bibliothécaires peuvent mobiliser un outil comme DoRANUM qui a l'avantage d'être ludique et de permettre l'auto-formation à la gestion des données. Cet outil pourrait apporter une solution adaptée aux différents profils de chercheurs en complément, ou pas, d'une formation en présentiel et permettre également aux bibliothécaires qui souhaiteraient de se perfectionner.

■ Carole-Anne MATEL

- 1 Organisation de coopération et de développement économique.
- 2 La gouvernance représente ici les gestionnaires d'un établissement (président, directeur...) qui doivent mettre en place la politique de gestion des données.
- 3 Voir l'article de Constance PICQUE « Les données de la recherche, le Dark Data des SHS » qui revient longuement sur ces deux enquêtes françaises.
- 4 Les articles de Sonia SALAMI « La gestion des données de la recherche et leur partage, quels enjeux ? » et de Laure PIERRAT « Le partage des données de la recherche dans la pratique » reviennent sur cette notion de DMP et son importance.
- 5 (Naegelen, 2015, paragr.1) Pour avoir plus de détail, voir article de Laure PIERRAT « Le partage des données de la recherche dans la pratique ».

BIBLIOGRAPHIE

BELLAMY Dorothée, *Dossier d'information. L'ouverture des données de la recherche (INS-HEA)*, 2014. URL : <http://www.inshea.fr/sites/default/files/Les%20donn%C3%A9es%20de%20la%20recherche.pdf>

CHEVAL Christelle, « Services aux chercheurs, quelle valeur pour les universitaires ? » *I2D – Information, données & documents*, 4 (Volume 53), 2015, p. 58. URL : <https://www.cairn.info/revue-i2d-information-donnees-et-documents-2015-4-page-58.htm>

DATA ASSET FRAMEWORK, (s.d.). Four steps to effective data management. URL : <https://www.data-audit.eu/>

GAGNON Lise, PERETZ Isabelle, FULOP Tamàs, « Musical structural determinants of emotional judgments in dementia of the Alzheimer type ». *Neuropsychology*, 23(1), p. 90-97, 2009. URL : http://www.academia.edu/16893405/Musical_structural_determinants_of_emotional_judgments_in_dementia_of_the_Alzheimer_type QUEL EST LE RAPPORT ? QUE FAIT CETTE REFERENCE ?

GAILLARD Rémi, « De l'Open Data à l'Open Research Data : quelle(s)

politique(s) pour les données de recherche?», Mémoire pour l'obtention du Diplôme de Conservateur des Bibliothèques, 2014. URL : www.enssib.fr/bibliotheque-numerique/documents/64131-de-l-open-data-a-l-open-research-data-quelles-politiques-pour-les-donnees-de-recherche.pdf

LAÏ Paolo, JANIK Joanna et ROLAND Michel, « Stratégie & Enjeux : La gestion et le partage des données de recherche » [interview], *sup-numerique.gouv.fr*, 20/02/2018. URL : <http://www.sup-numerique.gouv.fr/pid33129-cid126420/la-gestion-et-le-partage-des-donnees-de-recherche.html>

NAEGELEN Pierre, « Données de la recherche : quel positionnement et quels rôles pour les bibliothèques ? » [conférence], journée d'étude organisée par Médiad'Oc, l'Observatoire Midi-Pyrénées et l'URFIST de Toulouse à la Maison de la Recherche et de la Valorisation de Toulouse le 15 juin 2015. *canal-u.tv*, 2015. URL : https://www.canal-u.tv/video/universite_de_toulouse/pierre_naegelen_scd_ups_donnees_de_la_recherche_quel_positionnement_et_quels_roles_pour_les_bibliotheques.18892

OCDE, « Principes et lignes

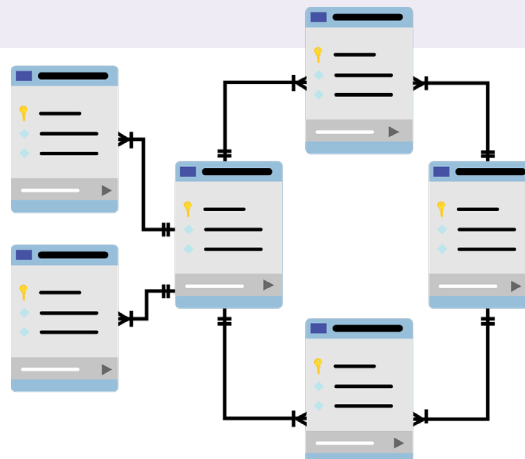
directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics », 2007. URL : <http://www.oecd.org/fr/science/sci-tech/38500823.pdf>

PROST Hélène, SCHÖPFEL Joachim, *Les données de la recherche en SHS. Une enquête à l'Université de Lille 3. : Rapport final* [Rapport de recherche], Lille 3. 2015. Lien : <https://tel.archives-ouvertes.fr/UNIV-LILLE1/hal-01198379v1>

ROLAND Michel, « Présentation de DoRANum » [conférence journée GRICAD], *sciencesconf.org*, 28/03/2017. URL : <https://datalabo2017.sciencesconf.org/142235/document>

SERRES Alexandre, MALINGRE Marie-Laure, MIGNON Morgane, PIERRE Cécile, COLLET Didier, *Données de la recherche en SHS. Pratiques, représentations et attentes des chercheurs : une enquête à l'Université Rennes 2* [Rapport de recherche], Université Rennes 2. 2017. URL : <https://hal.archives-ouvertes.fr/hal-01635186>

URFIST Paris, (s.d.), Missions et statuts. URL : <http://urfist.chartes.psl.eu/qui-sommes-nous/missions-et-statuts>



Financements accordés par l'ANR, décryptage...

Gilane BENFKIH

L'ANR (Agence Nationale de la Recherche) attribue des fonds en fonction d'appels d'offre qu'elle formule. Sur quels critères l'ANR s'appuie-t-elle pour formuler ses appels d'offre et quelles conditions doivent remplir les projets pour être éligibles ?

QU'EST CE QUE L'ANR ?

L'ANR c'est l'Agence National de la Recherche. C'est une agence qui finance des projets publics et des projets partenariaux de recherche scientifiques ou technologiques sélectionnés selon des critères internationaux d'excellence en France.

L'ANR finance des équipes de projets, qu'elles soient publiques ou publiques avec un partenariat privés, sous forme de contrat de recherche à durée indéterminée d'environ 3 ou 4 ans.

D'OÙ VIENNENT LES FONDS DE L'ANR ?

Les fonds de l'ANR proviennent principalement de la subvention d'intervention que lui attribue le Ministère de la Recherche. Une partie de ces fonds (3,3 % en 2015 soit 17,4 M€) sont apportés par des structures avec lesquelles l'ANR a construit des partena-

riats. Ils lui offrent des opportunités de financement ou de cofinancement de projets qui complètent le budget d'intervention de l'ANR.



COMMENT FONCTIONNE CETTE AGENCE ?

Cette agence fonctionne principalement avec des appels à projets (AAP). Ces appels à projets s'inscrivent dans les priorités nationales et concernent toutes les grandes disciplines de la recherche. Depuis 2010, 50 % des financements sont consacrés à la programmation **Non-Thématique** dites « blanches ». La seconde partie est consacrée à la partie **Thématique** qui est divisée en plusieurs disciplines

phases de la recherche scientifique et technologique (Sciences humaines et sociales, Sciences de la vie, Sciences des matières, Biodiversité, Politique publique, Innovation, etc.).

QUEL EST L'OBJECTIF DE L'ANR ?

L'objectif de cet établissement est multiple. Il s'agit tout d'abord d'équiper la France d'une structure « réactive » destinée au financement de la recherche fondamentale et de la recherche appliquée. L'ANR soutient les projets ambitieux et les équipes performantes pour contribuer au développement des sciences et des technologies.

L'agence souhaite favoriser l'émergence de nouveaux concepts en mettant en avant la créativité et le décloisonnement des disciplines malgré la réticence des chercheurs français à l'égard de la pluridisciplinarité. L'ANR

souhaite également cibler et accroître les efforts de recherche sur les priorités (AAP à thématique), encourager et intensifier la collaboration entre projets publics et industriels (AAP partenariaux), et enfin développer des partenariats internationaux (AAP transnationaux).



Financement de la recherche (CC0) © C.Picque

POURQUOI FINANCER DES PROJETS ?

L'ANR finance des projets car ils sont un enjeu majeur de croissance pour un pays, ils permettent des avancées technologiques, économiques et sociales et sont un accélérateur de croissance.

QUELLES PLACES OCCUPENT CES PROJETS EN FRANCE ?

L'appel à projet est nouveau en France et la création de l'ANR a été vécue comme une véritable révolution. Les projets de recherche sont importants cependant, en France ils restent tout de même modestes et il y a de moins en moins de chercheurs.

Un chercheur français est recruté à BAC+8 et touche deux ou trois fois moins qu'un chercheur américain. Un chercheur français peut mieux gagner sa vie et toucher entre 3 500 et 4 000 € s'il devient professeur d'université ou directeur de recherche. Le métier de chercheur est devenu précaire et peine à susciter de nouvelles vocations. Par rapport aux autres pays, la

France finance proportionnellement beaucoup moins de projets. L'ANR, par exemple, finance seulement entre 10 et 15 % des projets soumis contre 25 % chez les Allemands. Les chercheurs se découragent ou tentent d'obtenir des bourses européennes.

L'appel à projet correspond à 22 % des dépenses des financements de recherche. Selon le ministère, le système d'appel à projet favorise les laboratoires les plus riches. Cependant en fonction des secteurs, l'appel à projet ne permet pas aux chercheurs d'être libres de leurs recherches et aux laboratoires de développer leur propre politique scientifique.



European Research Council

« Des pans entiers de la recherche ne sont pas étudiés, car ils ne font pas l'objet d'appels à projets. C'est particulièrement vrai en matière d'environnement, où nous pouvons nous trouver face à des questions polémiques. Lorsque nous voulons travailler sur les effets de certains produits sur la santé, ou sur un autre modèle d'agriculture, plus respectueux de l'environnement, nous nous retrouvons vite à court de financements ¹. »



En laboratoire (CC0)

À QUI SONT DONC REVERSÉS LES FONDS ATTRIBUÉS PAR L'ANR ?

Sur le site de l'ANR nous avons pu constater que les financements des projets étaient extrêmement variés. Entre quelques milliers et plusieurs millions d'euros sont attribués à des projets. Nous nous sommes alors demandé : Comment l'ANR attribue son financement ? Quels sont les domaines scientifiques privilégiés par l'agence ? Comment l'ANR attribue son financement ?

- 1 Pour commencer l'ANR lance des appels à projets (AAP) pour donner lieu à une sélection compétitive.
- 2 Par la suite certains projets sont retenus, sélectionnés et organisés. (1071/ 10032 ont été retenus en 2014).
- 3 Ensuite l'ANR s'occupe de la contractualisation (création du contrat)
- 4 Ainsi l'ANR s'engage à financer le projet, à le suivre mais aussi à évaluer son impact.

+ 15 600 projets ont été financés par l'ANR depuis sa création.

Ainsi pour avoir des chances d'être sélectionné pour être financé :

- Le projet doit tout d'abord avoir une expertise scientifique française ou étrangère externe à l'ANR et le projet et les demandeurs doivent avoir une qualité scientifique appréciable.
- Le projet doit être créatif, collaboratif et interdisciplinaire (pour répondre aux standards internationaux et à la norme ISO-9001).

¹ Source Les échos.fr

QUELS SONT LES DOMAINES SCIENTIFIQUES PRIVILÉGIÉS DE L'AGENCE ?

Tout d'abord le domaine qui était le favori à la création de l'agence était « la recherche environnementale ». Le second fut « les défis sociétaux » c'est-à-dire le changement climatique, la crise de la biodiversité, la politique publique et l'innovation. Par la suite 6 domaines se sont démarqués :

- 1 - L'éco-ingénierie (les villes durables).
- 2 - La biodiversité (les contaminants et leurs effets sur la santé environnementale).
- 3 - L'agriculture.
- 4 - La pêche.
- 5 - Les risques naturels.
- 6 - Le changement climatique.

Et les liens entre la pollution de l'eau et de l'air et les réseaux trophiques comptent aussi parmi les priorités de l'ANR.

Si vous voulez que votre projet fasse partie des 10 % à 15 % des projets qui obtiennent l'aide financière de l'ANR votre projet doit être en rapport avec une ou plusieurs de ces disciplines car elles passeront toujours avant les autres.

Car l'ANR finance en priorité des projets ayant une utilité immédiate plutôt que des projets de recherche fondamentale.

▪ **Gilane BENFKIH**



Co-financement de la recherche (CCO)

ACTIVITÉ



283 collaborateurs au 31/12/16, soit 260 ETPT (Équivalents Temps Plein Travaillé)



595,9 M€
Budget d'intervention 2016



457,6 M€
Alloués sur appels à projets

INTERDISCIPLINARITÉ



74 % des projets financés sont des projets collaboratifs



19,5 %
Des projets internationaux cofinancés avec des agents étrangères



16 %
Des projets financés associent des équipes de recherche du monde académique et de l'entreprise

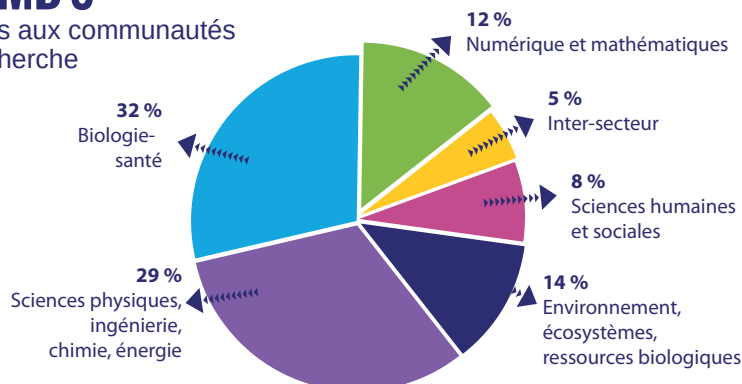
PROJETS FINANCÉS



15 600
Projets financés



6,5 MD€
Alloués aux communautés de recherche



www.didaktic.fr

