
Écriteure : la littératie informationnelle à la croisée de l'offre et des services des intermédiaires

Evelyne Broudoux



Édition électronique

URL : <http://journals.openedition.org/rfsic/4738>

DOI : 10.4000/rfsic.4738

ISSN : 2263-0856

Éditeur

Société Française de Sciences de l'Information et de la Communication

Ce document vous est offert par Conservatoire national des arts et métiers (Cnam)

le cnam

Référence électronique

Evelyne Broudoux, « Écriteure : la littératie informationnelle à la croisée de l'offre et des services des intermédiaires », *Revue française des sciences de l'information et de la communication* [En ligne], 15 | 2018, mis en ligne le 01 janvier 2019, consulté le 15 octobre 2019. URL : <http://journals.openedition.org/rfsic/4738> ; DOI : 10.4000/rfsic.4738

Ce document a été généré automatiquement le 15 octobre 2019.



Les contenus de la *Revue française des sciences de l'information et de la communication* sont mis à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.

Écrilecture : la littératie informationnelle à la croisée de l'offre et des services des intermédiaires

Evelyne Broudoux

Introduction

- 1 L'écosystème scientifique à l'heure de la participation sur le web s'est modifié entraînant une diversification des pratiques de diffusion des résultats de la recherche (Chartron, Broudoux, 2009) et la mesure de la participation sociale (Broudoux, 2013). L'ouverture de l'environnement scientifique du web communicationnel a aussi autorisé la promotion du genre éditorial du blog qui s'est imposé comme espace de discussion pour des jeunes chercheur.r.se.s en sciences humaines et sociales (Mayeur, 2017). Enfin, la croissance des données interrogeables et l'interopérabilité des références scientifiques modifient les supports de diffusion eux-mêmes qui ne se cantonnent plus à des articles dans des revues, l'article n'étant déjà plus – pour certaines disciplines – qu'un modèle à remplir servant à divulguer des résultats de calculs sur des données dynamiques.
- 2 Dans ce contexte, nous nous intéressons aux espaces informationnels qui organisent l'écriture et la lecture scientifique exigeant une nouvelle littératie informationnelle (Simonnot, 2009), car ceux-ci enchevêtrent différents secteurs d'activité avec des infomédiaires dont il s'agirait d'identifier les stratégies.
- 3 Nous proposons ici d'observer les usages et non usages de l'offre logicielle visant l'écriture, la lecture, la préservation des données, des références, des citations et leur traitement. Nous établissons le lien entre cette panoplie d'outils, services, plateformes, avec les acteurs de la recherche responsables de cette offre : start-ups émanant de projets de recherche, éditeurs et groupes de médias scientifiques. La méthode employée est celle de la recherche des propriétaires et des acquéreurs des services qui permet de mettre en

évidence le phénomène de concentration chez les acteurs du marché de la communication scientifique.

- 4 Pour cela, nous nous appuyons sur une récente enquête internationale et multidisciplinaire de grande ampleur étudiant les usages de « 101 innovations » de la communication scientifique (Kramer, Bosman, 2016), et qui concerne l'utilisation d'outils dès la recherche d'informations jusqu'à la diffusion et l'évaluation des résultats. Si les résultats de cette enquête ont été analysés avec des focus différents pour l'espace francophone, avec la médecine (Benoist, 2016) et la recherche agronomique (Aventurier, 2017), les « 101 innovations » méritent d'être ré-étudiées du point de vue des résultats concernant les activités d'écriture (Kembellec, Broudoux, 2017). Afin de la compléter, nous proposons une exploration des services et créations logicielles liées au web de données, technologie qui n'apparaissait pas – en tant que telle – dans l'analyse des « 101 innovations » en 2016. En effet, la gestion des données et leur remise à disposition ainsi que l'annotation sémantique des articles représentent des innovations notables où les outils sont encore à l'état de prototypes.

Analyse de l'enquête du point de vue de l'écriture

- 5 L'enquête multidisciplinaire sur l'utilisation des outils de communication scientifique réalisée en 2015-16 par la Bibliothèque universitaire d'Utrecht (Kramer, Bosman, 2016) mérite de s'y intéresser pour son exhaustivité : six communautés de recherche (Sciences physiques, Ingénierie & Technologie, Sciences de la vie, Médecine, Sciences Sociales et Économie, Arts et Sciences humaines, Droit) y sont représentées. Disponible en sept langues, elle intéresse l'ensemble des acteurs liés à la production scientifique : enseignants-chercheurs, chercheurs, bibliothécaires, éditeurs, ingénieurs et autres personnels impliqués dans la recherche. Elle a permis de recueillir 20 663¹ réponses informant sur l'utilisation d'outils² pour dix-sept activités de recherche réparties en six étapes (voir détail en Annexe). Trois questions générales complétaient l'enquête : le développement le plus important à venir dans la communication scientifique, le soutien éventuel au libre accès et celui à la science ouverte.
- 6 Notons une répartition équilibrée entre les répondants issus des STM et des SHS : les Sciences humaines et économiques répondent à 25 %, les Arts et Humanités à 13 % auxquelles il faut ajouter le Droit (2 %) ; de leur côté les Sciences de la vie répondent à 20 %, la Médecine (15 %), les Sciences de l'ingénieur (15 %) et les Sciences physiques (10 %). Les rôles proposés révèlent une majorité d'enseignants-chercheurs composée de Pr, MCF, ATER (42 %), suivis par les Doctorants (19 %), Postdoc (11 %) et Étudiants en Licence et Master (8 %). Les Bibliothécaires et Documentalistes (7 %), acteurs de l'Industrie et du Gouvernement (3 %) et les Éditeurs (2 %) représentent une minorité des répondants³.
- 7 La première remarque est que les classes créées à partir des premières années de publication ne révèlent pas de fossé générationnel dans l'usage ou la recommandation des outils entre les seniors et juniors de la recherche. L'adoption des innovations n'est donc pas imputable à l'âge. 30 % des répondants ont démarré une activité de publication il y a 30 ou 20 ans et 23 % sont des publiants débutants, 16 % des répondants ne publie pas (ce qui correspond aux rôles de la recherche qui distinguent les enseignants-chercheurs des autres professions qui n'ont pas pour vocation à « publier »).

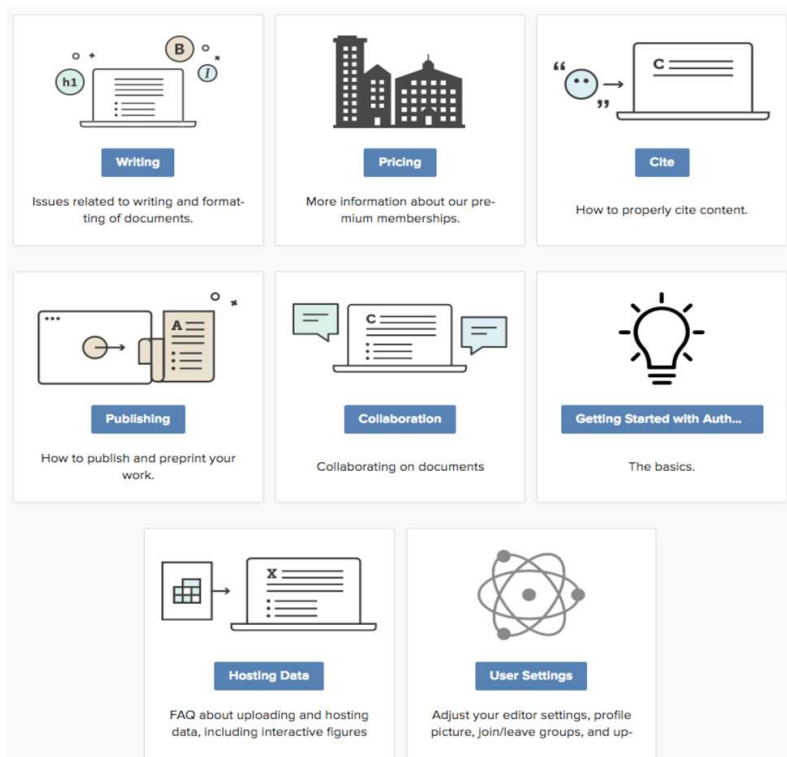
- 8 La deuxième remarque est que la présentation des outils dans ce questionnaire correspond à une vision linéaire de la création d'une publication, dans laquelle l'écriture intervient comme une étape-maillon au sein d'une chaîne d'actions et l'annotation comme un acte de lecture se déroulant pendant la découverte d'information. Or, comme nous l'avons montré dans une vision d'écriteure (Broudoux, Kembellec, 2017), l'annotation est un acte d'écriture qui se déroule pendant la lecture, elle est multiforme et son potentiel technique s'est récemment démultiplié. D'autre part, rappelons que le processus d'écriture n'est pas linéaire (Oriol-Boyer et al., 1998) et prend des formes différentes selon le support-papier ou logiciel (Sharples, 2002). Enfin, l'écriture collaborative est étudiée depuis plus de vingt ans dans l'objectif d'un soutien logiciel à ce processus (Sapsomboon et al., 1997 ; Crozat, 2012)⁴.

1.1 Ecriture, lecture et citation

- 9 L'observation des réponses à cette enquête reposant sur un choix d'outils présélectionnés⁵ pour la lecture-écriture-annotation permet d'affirmer que la majorité des usages est toujours très concentrée sur les outils classiques. La lecture de fichiers pdf avec Acrobat Reader (88 %) et l'écriture avec Word (89 %) sont plébiscitées. On note cependant l'arrivée de Google Docs/Drive (32 %) comme outil d'écriture, dépassant LaTeX (18 %) - ce qui laisse imaginer qu'une écriture à plusieurs s'y déroule - et une lecture en ligne via html (40 %), ce qui indique une lecture-écran sur navigateur habituelle pour près de la moitié des usagers. Si l'on s'intéresse maintenant aux autres logiciels sélectionnés, nous voyons arriver rangés dans les outils de lecture des logiciels permettant des annotations, les LGRB Mendelej (14 %) et Readcube (6 %), et à partir des PDF, iAnnotate (2 %) ou UtopiaDocs (0,5 %), ainsi qu'Hypothes.is (0,9 %) qui est une couche logicielle permettant d'annoter des contenus sur les sites web, de manière individuelle ou collective. Pour ce qui est de l'activité d'écriture, la différence entre dominants comme Word et nouveaux entrants est encore plus notable. Scrivener (2,5 %), Overleaf (0,2 %) et Authorea et Scalar pour moins de 1 % révèlent des utilisations très marginales qui confirment le fait que l'activité d'écriture reste très classique.
- 10 Pourtant, cette offre logicielle intègre des fonctionnalités qui pourraient devenir indispensables à l'écriture scientifique dans les toutes prochaines années. Alors que Scrivener est un outil d'écriture individuel intégrant de la documentation (images, fichiers pdf, etc.) dans un espace de travail et personnalisable selon le genre éditorial choisi (article, fiction, poésie, essai, etc.), Overleaf et Authorea sont des plateformes totalement dédiées à l'écriture-édition d'articles scientifiques intégrant des fonctionnalités d'administration de documents et de données.
- 11 Overleaf (2015) est un outil de publication et de publication collaboratif en ligne éditant des balises LaTeX avec une vue en texte enrichi qui facilite la compréhension et l'utilisation du langage, avec l'édition de formules, l'ajout de commentaires et le versioning de documents. Mais le véritable intérêt de cette plateforme réside en une pluralité de gabarits et de modèles bibliographiques adaptés à différentes revues. La proposition d'administrer la soumission aux revues dont la plateforme gère les modèles excède les fonctionnalités d'écriture. Elle révèle une stratégie d'acteur se positionnant en intermédiaire entre le rédacteur d'article et l'éditeur de revues scientifiques.
- 12 La plateforme Authorea (2012) offre à l'écriture d'articles scientifiques, des possibilités collaboratives sur les documents, le stockage de jeux de données et la génération de

visualisations, ainsi que la gestion des citations automatisées à partir de CrossRef et autres répertoires de métadonnées scientifiques. A cet égard, la capture d'écran des principales fonctionnalités (Fig. 1) révèle une orientation didactique intéressant les doctorants-publants débutants : indications générales sur l'édition d'articles scientifiques, apprentissage de la citation (« how to properly cite the content »), publication et stockage du preprint de son travail sur un répertoire institutionnel (« how to publish and preprint »). Comme le revendique la plateforme, chaque document est aussi un répertoire *Git* qui peut héberger des données. Il est ainsi possible d'y ajouter des notebooks Jupyter et iPython, de copier/coller du script dans des fichiers html (comme Plotly, Bokey, Carto) et de les appeler au sein du document pour générer des infographies et des visualisations de données.

Figure 1. Aide à la prise en mains des espaces de travail d'Authorea



- 13 Enfin, à la question « Quels outils/sites utilisez-vous pour gérer vos références ? », les réponses ordonnaient EndNote, Mendeley, Zotero, RefWorks, « autres », Papers, Citavi et RefMe. L'examen du nombre de réponses affectées à chaque logiciel fait apparaître aujourd'hui un gagnant à 36 % (EndNote), suivis par Mendeley (19 %) et Zotero (14 %). Comme pour les outils de lecture-annotation et écriture, c'est un gagnant logiciel « propriétaire » qui rassemble les usages. Des quatre logiciels de gestion bibliographiques « libres ou gratuits » qu'avaient présentés G. Kembellec et C. Scopsi en 2012, seuls Mendeley Desktop et Zotero apparaissent tirer leur épingle du jeu alors que JabRef et BibDesk ne sont pas mentionnés.

Tendances vers l'open : oui mais pas trop...

- 14 Les questions sur le libre accès (86 % pour, 4 % contre et 10 % ne se déterminent pas) et l'open science - précisée comme « la création ouverte, le partage ouvert et l'évaluation

ouverte, lorsque c'est possible » - (81 % pour, 4 % contre et 15 % ne se déterminent pas) révèlent le fort intérêt des répondants pour ces thématiques. Cependant, celui-ci est à pondérer par les pratiques réelles.

- 15 À la question : « Quels outils/sites utilisez-vous pour archiver/partager les publications ? », c'est sans surprise que l'on constate que ResearchGate⁶ (43 %) dépasse nettement les dépôts institutionnels (28 %). Mais cet intérêt pour la diffusion des travaux à des pairs en réseau ne se confirme pas lorsqu'il s'agit de connaître les retours des « pairs » sur une plateforme spécialisée. En effet, à la question : « Quels outils/sites utilisez-vous pour l'évaluation par les pairs au-delà de ce qui est mis en place par les revues ? » qui concerne – comme il est précisé dans l'explication de la question, « les commentaires et les évaluations de pairs, pré-publication et post-publication », les répondants ont sélectionné parmi les sept plateformes de peer-reviewing proposées : PubMed Commons, Peerage of Science, Publons, PubPeer, PaperCritic⁷, Academic Karma et Rubriq. Le plus fort taux d'utilisation étant PubMed Commons (4 %) et le plus faible Rubriq (0,5 %). Ce très faible investissement sur les plateformes d'open peer-reviewing peut révéler un désintérêt pour la lecture des évaluations et commentaires mais aussi une méconnaissance de leur existence, ou encore une surcharge cognitive due à la pléthore d'outils et plateformes disponibles. En tout cas, les pratiques n'apparaissent pas confirmer un engouement pour l'open science qui se manifesterait par des formes d'engagement sur des plateformes dédiées.
- 16 Pourtant, les modèles de publication avec un peer reviewing ouvert existent bien mais sont encore au stade des incitations. En témoigne, début février 2018, une lettre ouverte⁸ signée par des Revues principalement en Sciences de la vie appartenant à des éditeurs nord-américains⁹ et publiée par ASaPbio (*Accelerating Science and Publication in biology*), une initiative financée par des fondations dans l'objectif de promouvoir l'innovation et la transparence dans la communication des sciences de la vie. Elle engageait ses signataires à rendre public les rapports d'examen des articles évalués par les pairs, qu'ils soient anonymisés ou non. Les bénéfices attendus sont : 1) une responsabilisation accrue des réviseurs et des rédacteurs, 2) des opportunités de formation pour sensibiliser les étudiants au processus d'évaluation par les pairs, 3) une meilleure compréhension par les lecteurs des articles dans le champ et 4) un moyen de valoriser l'évaluation par les pairs.
- 17 L'examen de cette lettre-manifeste révèle un unique lien hypertexte qui pointe vers l'article « *What is open peer review? A systematic review* » (Ross-Hellauer, 2017), revu par les pairs en mode ouvert et révisé par l'auteur. Publié sur la plateforme F1000Research, il sert à démontrer la faisabilité du processus : l'évaluation des réviseurs est réalisée sous forme de commentaires auxquels l'auteur de l'article peut répondre. A ce workflow, il faut remarquer le soin apporté à la possibilité de citer les versions successives de l'article, mais aussi les commentaires de l'article.

Figure 2. Interface de publication des articles sur la plateforme F1000Research. Sur l'encart de droite (C), le processus de révision est rendu public et mis à jour au fur et à mesure des modifications demandées et réalisées par les auteurs

The screenshot displays the F1000Research article page for 'What is open peer review? A systematic review [version 2; referees: 4 approved]'. The interface is divided into three main sections:

- Section A (Article Details):** Shows the article title, author (Tony Ross-Hellauer), and a note that the article is included in the Science Policy Research gateway.
- Section B (Metrics and Actions):** A sidebar containing metrics (6828 views, 1591 downloads) and action buttons (Get PDF, Get XML, Cite, Export, Track, Email, Share).
- Section C (Open Peer Review):** A detailed view of the review process, including a table of invited referees and their reports.

Version(s)	1	2	3	4
REVISED Version 2 published 31 Aug 2017	read report	read report	read report	read report
Version 1 published 27 Apr 2017	read report	read report	read report	read report

The abstract text is as follows:

Background: "Open peer review" (OPR), despite being a major pillar of Open Science, has neither a standardized definition nor an agreed schema of its features and implementations. The literature reflects this, with numerous overlapping and contradictory definitions. While for some the term refers to peer review where the identities of both author and reviewer are disclosed to each other, for others it signifies systems where reviewer reports are published alongside articles. For others it signifies both of these conditions, and for yet others it describes systems where not only "invited experts" are able to comment. For still others, it includes a variety of combinations of these and other novel methods.

Methods: Recognising the absence of a consensus view on what open peer review is, this article undertakes a systematic review of definitions of "open peer review" or "open review", to create a corpus of 122 definitions. These definitions are systematically analysed to build a coherent typology of the various innovations in peer review signified by the term, and hence provide the precise technical definition currently lacking.

Results: This quantifiable data yields rich information on the range and extent of differing definitions over time and by broad subject area. Quantifying definitions in this way allows us to accurately portray exactly how ambiguously the phrase "open peer review" has been used thus far, for the literature offers 22 distinct configurations of seven traits, effectively meaning that there are 22 different definitions of OPR in the literature reviewed.

Conclusions: I propose a pragmatic definition of open peer review as an umbrella term for a number of overlapping ways that peer review models can be adapted in line with the aims of Open Science, including making reviewer and author identities open, publishing review reports and enabling greater participation in the peer review process.

- 18 Trois zones sont visibles sur la capture d'écran (voir Fig. 2) du portail de publication F1000Research : les métadonnées identifiant l'article (A), ses statistiques de consultation (B), partage, etc., et le suivi effectué par ses réviseurs (C). À noter que le système incite à la citation en procurant une aide afin de réaliser un processus normalisé (voir Fig. 3).

Figure 3. Interface de citation d'un commentaire de réviseur sur la plateforme F1000Research. La citation de la version évaluée est normalisée par une mise entre crochets

The screenshot shows a 'Referee Report' from Richard Walker, dated 08 May 2017, which is 'Approved'. A 'Views' popup shows 86 views. A 'HOW TO CITE THIS REPORT' popup provides the following information:

HOW TO CITE THIS REPORT:
Walker R. Referee Report For: What is open peer review? A systematic review [version 2; referees: 4 approved]. F1000Research 2017, 6:588 (doi: 10.5256/f1000research.12273.122299)
The direct URL for this report is: <https://f1000research.com/articles/6-588/v1/referee-response-22299>

NOTE: It is important to ensure the information in square brackets after the title is included in this citation.

Buttons for 'CLOSE' and 'COPY CITATION DETAILS' are visible.

Analyse des acteurs de l'offre logicielle de la communication scientifique

- 19 Les acteurs majeurs de l'offre logicielle dans la communication scientifique sont :
- Les chercheurs eux-mêmes impliqués dans l'Information scientifique et technique, l'organisation de la recherche et le développement logiciel,
 - Les institutions publiques et fondations privées, financeurs de la recherche,
 - Les groupes de publication et de médias, les agrégateurs de revues, les bibliothèques numériques,
 - Les entreprises au service des éditeurs et des bibliothèques,
 - Les nouveaux entrants et entreprises privées extérieures au champ de la publication scientifique.
- 20 La dynamicité du secteur de la communication scientifique se laisse observer avec le lancement de nouveaux services dédiés à l'écriture et la consultation d'articles, à la gestion des données et des documents. Nous avons mis en perspective sur le Tableau 1 les acquisitions récentes de services web et de plateformes scientifiques par les principaux acteurs de la communication scientifique : les groupes de publications et médias (ex : Elsevier, Holtzbrinck [propriétaire à 53 % de Springer], Wiley), les nouveaux entrants (ex : Onex). Apparaissent aussi dans le tableau, les créateurs de services comme les fondations privées (ex : Mellon) ou les projets-startup de chercheurs (ex : Academia.edu). Pour réaliser ce tableau, nous avons interrogé différents portails (Crunchbase, Acquiredby, Scholarly Kitchen¹⁰) dans l'objectif de retracer les différentes acquisitions des groupes spécialisés dans les technologies de publication de l'information scientifique.
- 21 Le rachat en 2017 de Bepress par Elsevier après celui de SSRN en 2016 signale l'entrée de l'acteur dans le domaine de l'accès ouvert (Schonfeld, 2017) puisqu'il gagne avec Bepress Digital Commons, un dépôt institutionnel utilisé par plus de 500 lycées, universités, facultés de droit, écoles de médecine, centres de soins de santé, bibliothèques publiques et centres de recherche dont l'objectif était de préserver et valoriser leurs publications et collections.

Tableau 1. Propriétaires et financeurs des services web et plateformes scientifiques avec l'année de rachat de services et entreprises

	Dépôts institutionnels, Archives ouvertes, Réseaux, Données	LGRB, métriques, Ecriture, Annotations	Gestion infrastructures de recherche, équipes, CRIS, ID	Gestion de workflow sur plateformes, blockchain
Elsevier	Bepress ¹¹ (2017) SSRN (2016) Preprints + Réseau social SHS	Mendeley (2015) Plum Analytics (2017)	Atira (2012) : Pure gestion de projets et métriques de la recherche	Aries Systems Corporation (2018) Gestion du peer-review

Holtzbrinck (Digital Science)	<i>Figshare</i>	<i>ReadCube</i> <i>Altmetric</i> <i>Overleaf</i> (2015) <i>Symplectic</i>	<i>Tetrascience</i> <i>Peerwith</i> <i>IFI Claims</i>	<i>Katalysis</i> <i>Dimensions</i> <i>GRID</i>
Wiley		<i>Athorea</i> (2018) (The Winnower)	<i>Profiles International</i> (2012) <i>Ranku</i> (2016)	<i>Atypon</i> (2016)
Academia.edu		<i>Plasmyd</i> (2013) Moteur de recherche		
Mellon Foundation	<i>Manifold</i> (UMN) <i>Humanities</i> <i>Commons</i> (MLA)	<i>Hypothes.is</i>		<i>Vega</i> (WVU) <i>Editoria</i> <i>Fulcrum</i> (UM)
Onex		<i>Clarivate analytics</i> (2016) (Kopernio, Publons, EndNotes)		

- 22 Les logiciels de gestion de références bibliographiques transformés progressivement en bibliothèques numériques annotables et partagées entre lecteurs ont été rachetés par les éditeurs. Le cas typique est celui de Mendeley fondé en 2008 et racheté par Elsevier en 2013 qui dispose ainsi d'un dispositif d'observation du lectorat par discipline scientifique. Ce tableau relativise l'indépendance de services qui ayant démontré leur utilité sont voués à être rachetés par les acteurs dominants qui concentrent les savoir-faire. Bien qu'incomplet car il ne couvre pas tous les acteurs, il indique que les infrastructures de gestion de la recherche sont aussi sous contrôle des groupes de médias scientifiques tout comme les recherches innovantes impliquant la gestion des projets de recherche, les identifiants de chercheurs, la technologie blockchain pour le peer-reviewing, etc.

Deux secteurs d'innovation liés au web de données

La gestion et la mise à disposition des données

- 23 En France, des démarches en faveur de la gestion et du partage des données de la recherche ont été lancées dès le début des années 2010 ; des enquêtes sur les attentes des chercheurs en matière de données de la recherche à l'Université de Lille 3 (Prost, Schöpfel, 2015) et à l'Université de Rennes 2 (Serres et al., 2017) ont montré leur intérêt mais aussi la nécessité de la formation des doctorants et des chercheur.se.s confirmé.e.s.
- 24 La gestion des données est sans doute le secteur le plus prometteur dans la découvrabilité scientifique. La vision de Jim Gray (2007) est souvent invoquée par ceux qui pensent que les données massives représentent une révolution dans la façon dont les recherches sont conduites et affectent les manières de penser. Elle repose sur l'idée que nous serions entrés dans l'ère du 4^e paradigme de la découverte scientifique, avec la transformation

des méthodes scientifiques due à l'exploitation intensive des données. Le premier paradigme serait constitué par la recherche empirique, le deuxième par la théorisation, le troisième serait celui de la simulation de phénomènes complexes autorisée par l'informatisation, le quatrième étage étant constitué par les données massives produites et générées par les simulateurs, capteurs et autres instruments. Seule une petite partie de ces données étant utilisée, le reste n'étant pas pertinent pour la recherche initialement conduite reste à explorer pour d'autres objectifs. C'est l'ouverture de la découvrabilité qui caractériserait l'eScience, car ce type de données produites de manière automatisée représente de la donnée brute exploitable. Dans cette optique, le big data et le web de données apportent des opportunités de mise à jour de liaisons inédites.

- 25 Zenodo (2013) est une archive ouverte de données de recherche pour la préservation et la mise à disposition de contenus de recherche, éducatifs et informatifs, issu d'un projet Cern/OpenAIRE. Ouverte aux SHS, elle fonctionne par communauté de recherche. L'accès au contenu de Zenodo est cependant ouvert à tous, « à des fins non militaires uniquement ». Un récent plantage de la base de données indique que la sollicitation intense de ce genre de services exige un dimensionnement qui s'adapte aux usages.
- 26 La plateforme française ISTEEX (Initiative d'excellence de l'Information Scientifique et Technique) récemment lancée par l'Inist a pour objectif principal de proposer, à l'ensemble de la communauté de l'enseignement supérieur et de la recherche, un accès en ligne aux collections rétrospectives de la littérature scientifique dans toutes les disciplines. Le projet initial de 2012, porté par quatre partenaires¹², vise l'acquisition de ressources documentaires et la création d'une plateforme d'hébergement et de mise à disposition des données. Les acquisitions se composent de collections rétrospectives de revues, de livres électroniques, de grands corpus de documents patrimoniaux numérisés et de bases de données, laissant augurer une grande richesse documentaire.
- 27 La plateforme se compose actuellement :
- 28 – d'un espace de recherche et téléchargement de corpus,
- 29 – de fonctionnalités de gestion de recherche bibliographique : intégration d'Istex à Google Scholar et Zotero et sous forme d'extension pour les navigateurs, un bouton d'accès aux plein textes s'affiche lorsqu'ils sont disponibles, à partir des identifiants documentaires,
- 30 – de l'ouverture au requêtage par une API qui permet d'attaquer directement les données :
- un module d'apprentissage,
 - un espace des données ISTEEX interrogeable en sparql,
 - la liste des jeux de données interrogeables.
- 31 Les services intéressants notre article sont un moteur de recherche aux fonctionnalités pointues de requêtes et l'accès systématique au texte intégral des documents ainsi qu'aux services de traitement des données : fouille de textes, extraction de données, production de synthèses documentaires et de corpus terminologiques.

Les services reposant sur l'annotation

- 32 L'annotation qui est le processus qui consiste à associer à un document des informations complémentaires comme des remarques (explication, commentaire critique, etc.) ou des notes (référence bibliographique, url, etc.) qui peuvent être de différente nature (textuelle, imagière, documentaire) a été progressivement incluse dans les outils

d'écrilecture et de publication. Les différents scénarios d'utilisation de l'annotation, éventuellement combinables, vont d'une utilisation individuelle et collaborative à une déclinaison manuelle et automatisée.

- 33 Si l'annotation individuelle consiste souvent à poser des commentaires à la fin d'un billet de blog, l'annotation collective se manifeste par la possibilité de commenter à plusieurs des portions de textes et à s'entre-commenter. On peut parler d'annotation collaborative lorsqu'un projet d'annotation est lancé visant un objectif partagé explicitement de manière collective par plusieurs participants (Merzeau, 2012). Alors que l'annotation manuelle est posée par un acteur-auteur ou éditeur, l'annotation automatisée est générée par un programme suivant un certain nombre de critères. Par exemple, les liens hypertextes extradocumentaires qui pointent vers des ressources externes sont interopérables. Dans la communication scientifique, l'annotation automatisée peut être « sémantisée » selon la définition donnée par les acteurs informaticiens du web sémantique : « l'accrochage d'un élément de l'ontologie à un fragment de document, qui implique de reconnaître (ou de déduire) la présence de l'élément ontologique dans la forme de surface du document » (Prié, Garlatti, 2004). Dans cette optique, l'annotation fait partie de l'éditorialisation sémantique des publications scientifiques (Verlaet, Dillaerts, 2016).
- 34 *Remarq* est un exemple de nouveau service éditorial (lancé au printemps 2017) – à visée annotative manuelle - conçu pour l'enseignement et la recherche et qui offre des fonctionnalités de recommandation, de collaboration, de commentaire, de partage d'article, le tout sur le site de l'éditeur et en utilisant la version enregistrée de l'article scientifique. Par exemple, un éditeur qui inclurait l'outil dans son offre miserait sur la création d'un réseau social doté de fonctionnalités de partage et d'annotation d'articles. Les remarques sur les sites d'éditeurs peuvent être lues par n'importe qui, mais les annotateurs sont validés par le personnel de *Remarq* avant que leurs annotations publiques puissent apparaître (Staines, Martone, 2018). En adoptant ce service pour six de ses revues, l'*American Mathematical Society* (AMS) expérimente un système destiné à impliquer son lectorat. *RedLink* qui offre ce service est représentatif des nouveaux acteurs faisant le lien entre bibliothèques, éditeurs et utilisateurs (rédacteurs, auteurs, lecteurs). La conception de ce service transversal s'appuie sur le lieu d'études (l'école, la bibliothèque), l'article consulté (la revue, l'éditeur), et les ajouts, partages et apports réalisés par des lecteurs, organisés en groupes de travail. L'offre professionnelle consiste en la tenue de tableaux de bord spécifiques aux bibliothèques et éditeurs, des moteurs de prospection de ventes pour les éditeurs aux outils éditoriaux. *RedLink* propose aussi *RedLink Network*, un service gratuit de synchronisation des informations d'identification et des besoins de service entre bibliothèques et éditeurs. Ces services misent sur la transversalité permise par l'interopérabilité entre acteurs et plateformes.
- 35 L'annotation lorsqu'elle est sémantisée peut aller beaucoup plus loin que celles de l'imprimé qui se situent dans ses marges, au bas des pages, à la fin des chapitres ou à la fin du document lui-même. Le W3C¹³ définit les annotations Web comme une tentative de recréer et d'étendre cette fonctionnalité en tant que nouvelle couche d'interactivité et de liaison sur le Web. L'objectif étant de permettre à n'importe qui d'annoter n'importe quelle ressource, que ce soit une page Web, un livre électronique, une vidéo, une image, un flux audio ou des données sous forme brute ou visualisée. Les annotations Web peuvent être liées, partagées entre des services, retracées dès leurs origines, recherchées et découvertes et stockées à la demande de leur auteur. Pour servir cette vision

d'infrastructure d'annotation décentralisée et ouverte, un groupe de travail *Open Annotation Collaboration*¹⁴ a diffusé des avis jusqu'en 2013 pour l'*Open Annotation Community Group* du W3C qui a publié en 2016 trois recommandations¹⁵, définissant un cadre de modélisation, le vocabulaire utilisé et le protocole décrivant le mécanisme de transport pour créer, gérer et récupérer les annotations.

- 36 L'exemple de Dokie.li prouve que cet ensemble de possibilités est exploitable dans le cadre de la communication scientifique. Il s'agit d'un éditeur client pour l'écriture d'articles scientifiques, avec des vues générables du document (html, pdf), les annotations et les interactions sociales étant gérées directement depuis l'outil d'écriture. Dokie.li implémente les spécifications W3C d'annotation sémantique décrites plus haut à des fins de peer-reviewing. Quel que soit le lieu de publication de l'article, les lecteurs peuvent laisser des annotations et des réponses s'ils possèdent leur propre espace d'enregistrement. Si l'article est publié sur un espace de stockage personnel, il est possible de proposer de stocker les annotations et les réponses en leur nom et activer également les réponses anonymes. Toutes les annotations et les réponses peuvent être enregistrées avec une licence Creative Commons.
- 37 Des identifiants uniques (URI) sont générés automatiquement pour chaque section de l'article afin de faciliter la liaison entre utilisateurs et s'y référer. Il est possible d'ajouter des identifiants à tous les concepts jugés importants à tous les niveaux de granularité. Il est aussi possible d'ajouter un balisage descriptif à tout concept ou extrait de prose qui a un identifiant. Dokie.li génère un balisage RDFa afin que les idées soient exposées en tant que données liées pour que d'autres puissent les interroger, les réutiliser et les visualiser. L'exemple de la figure 4 illustre un processus de liaison pour spécifier une référence.

Figure 4. Référence explicative pointant vers une portion de page web pour l'annotation avec Dokie.li

Resources and tools

Please make [pull requests](#) or [join the chat](#) to make recommendations and updates to this document.

Authoring Tools

The screenshot shows a web page with a dropdown menu open. The menu is titled 'agrees with' and contains the following options: agrees with, cites, cites as authority, cites as data source, cites as evidence, cites as metadata document, cites as potential solution, cites as potential reading, cites as related, cites as source document, **cites for information** (highlighted), compiles, confirms, contains assertion from, corrects, credits, critiques, and derides. The background text includes 'Footnote', 'Reference', 'agrees with', 'ations and', 'social inter', 'http://example.org/article#results', 'What's up?', 'Review', 'dokieli: dec', 'bookmarking); allows annotations to be', 'assigning Creative Commons license to', 'identifiers. Uses [W3C Web Annotation](#)', '[Notifications](#).', 'Self-publishing and archi', 'aning.', 'al storage,', 'nanitics and', 'ed Data'.

Conclusion

- 38 D'une part, il s'agissait ici d'effectuer un bilan de l'offre logicielle visant l'écriture, la lecture, la préservation des données, des références, des citations et leur traitement. Si l'aide logicielle à l'édition de textes s'est diversifiée, en termes d'usages, les outils traditionnels sont loin de céder leur place à des services innovants et n'apparaissent pas vouloir les intégrer. La composition générale de cette offre logicielle révèle pourtant son attraction pour les groupes de publications en place mais aussi pour de nouveaux entrants qui pratiquent des stratégies d'agrégation de services afin de capturer du lectorat.
- 39 Le travail sur les données, leur gestion, leur liaison apporte un certain nombre de promesses. Mais l'accessibilité à l'information qui passe par la médiation des données suppose aujourd'hui des capacités élargies pour les chercheurs en matière de communication scientifique. En particulier, l'aptitude à (ré)utiliser les données apparaît un enjeu incontournable pour les sciences humaines et sociales sommées d'acquérir de nouvelles compétences informationnelles (Serres, 2012).
- 40 Alors que le travail collaboratif – consistant à partager en ligne des tâches et des documents – est devenu la norme pour certaines professions (ex : les ingénieurs), il apparaît que dans l'autorialité scientifique, ce travail collaboratif peine à être reconnu. On peut imaginer qu'un travail sur les liaisons sémantiques impliquerait d'abord de reconnaître la nature collaborative de la recherche scientifique. Par exemple, ici, le peer-reviewing qui représente la trace des choix qualitatifs des communautés scientifiques, est soutenu par un workflow logiciel en tant qu'activité autonome... et les usages ne sont pas au rendez-vous. En parallèle, on constate que différents essais d'open peer-reviewing sont périodiquement lancés impliquant des revues, des éditeurs et des auteurs – réviseurs et commentateurs en nombre limité et sur une courte période, mais ces projets « ne font pas école » et ne sont pas généralisables. Le fait que ces essais aient été peu couronnés de succès indique les communautés scientifiques ne sont pas prêtes à l'adopter. Il faudrait donc s'interroger plus avant sur leurs motivations.

BIBLIOGRAPHIE

- ATTWOOD, T. K., DKELL, D.B., MCDERMOTT, P., MARSH, J., PETTIFER, S.R., THORNE, D. (2010). "Utopia documents : linking scholarly literature with research data", *Bioinformatics*, 26, (18), 15 September 2010, i568–i574, <https://doi.org/10.1093/bioinformatics/btq383>
- AVENTURIER, P. (2016). « 101 Innovations - Présentation et résultats de l'enquête internationale sur l'utilisation des outils de communication scientifique. Focus sur les résultats de l'Inra ». Séminaire des revues scientifiques, Jun 2016, Ivry-sur-Seine, France. 27 p.

- BENOIST, D. (2016). « Outils de communication scientifique : les résultats pour la France ». Billet de blog du 22 juillet 2016. URL : <http://www2.biusante.parisdescartes.fr/wordpress/index.php/enquete-outils-de-communication-scientifique-medecine-france/>
- BROUDOUX, E., & CHARTRON, G. (2009). La communication scientifique face au Web2. 0 : Premiers constats et analyse. In *H2PTM*. Hermès Science-Lavoisier. pp. 323-336.
- BROUDOUX, E. (2013). « L'écosystème scientifique à l'heure de la participation sur le web » in Actes du colloque H2PTM 13, 239-256. <sic_00998364>
- CAPADISLI, S., GUY, A., VERBORGH, R., LANGE, C., AUER, S., & BERNERS-LEE, T. (2017). « Decentralised authoring, annotations and notifications for a read-write web with dokieli », in International Conference on Web Engineering. Springer, 469-481. <http://csarven.ca/dokieli>
- CROZAT S. (2012). « C2M : Chaînes éditoriales collaboratives multimédia ». Actes du colloque CIDE 15 « Métiers de l'information, des bibliothèques et des archives à l'ère de la différenciation numérique », Tunis, Tunisie. URL : <https://stph.crzt.fr/res/crozat2012cide1.pdf>
- GRAY, J. (2009). "A transformed scientific method", In : Hey T., Tansley, S., Tolle, K. (eds). *The Fourth Paradigm : Data Intensive Scientific Discovery*. Microsoft, Redmond (2009).
- KEMBELLEC, G. BROUDOUX, E. (dir.) (2017). *Ecriture augmentée dans les communautés scientifiques. Humanités numériques et construction des savoirs*. Iste Editions. p.
- KEMBELLEC, G., SCOPSI, C. (2012). "Étude comparée de quatre logiciels de gestion de références bibliographiques libres ou gratuits ». *Documentation et Bibliothèques*, Asted Inc., 58 (4), 187-197. <hal-00768168>
- KRAMER, B., and BOSMAN, J. (2016). « Innovations in scholarly communication - global survey on research tool usage" [version 1 ; referees : 2 approved]. *F1000Research*, 5 :692. doi : 10.12688/f1000research.8414.1
- MAYEUR, I. (2017). « La communication scientifique directe vers un public élargi », Revue française des sciences de l'information et de la communication [En ligne], 11 | 2017. URL : <http://journals.openedition.org/rfsic/3224>; DOI : 10.4000/rfsic.3224
- MERZEAU, L. (2013). « Éditorialisation collaborative d'un événement. L'exemple des Entretiens du nouveau monde industriel 2012 ». *Communication et organisation*, (43), 105-122.
- NICHOLAS, D., RODRÍGUEZ-BRAVO, B., WATKINSON, A., BOUKACEM-ZEGHMOURI, C., HERMAN, E., XU, J., ABRIZAH, A. and SWIGÓN, M. (2017), "Early career researchers and their publishing and authorship practices". *Learned Publishing*, 30 : 205-217. doi :10.1002/leap.1102
- ORIOLE-BOYER, C., DRIOL, M., ORIOLE, J. C., & SAMY, T. (1998). « Hypertexte et didactique de la production textuelle en français ». *Revue de l'EPI (Enseignement Public et Informatique)*, (89), 75-94.
- PRIE, Y. GARLATTI, S. (2004). « Annotations et métadonnées dans le Web sémantique », in *Revue I3 Information-Interaction-Intelligence*, Numéro Hors-série Web sémantique, 24 pp.
- PRIEM, J., PIWOWAR, H. A., & HEMMINGER, B. M. (2012). "Altmetrics in the wild : Using social media to explore scholarly impact". arXiv preprint arXiv :1203.4745.
- PROST, H., SCHÖPFEL, J. (2015). « Les données de la recherche en SHS. Une enquête à l'Université de Lille 3 : Rapport final ». [Rapport de recherche] Lille 3. <hal-01198379>
- ROSS-HELLAUER, T. (2017). "What is open peer review ? A systematic review [version 2 ; referees : 4 approved]". *F1000Research*, 6 :588. (doi : 10.12688/f1000research.11369.2)
- SAPSOMBOON, B., ANDRIATI, R., ROBERTS, L., SPRING, M. (1997). "Software to Aid Collaboration : Focus on Collaborative Authoring". Draft of a report to the National Institute of Standards and

- Technology on the features to be expected in Collaborative Authoring Software. <http://www.sis.pitt.edu/spring/cas/>
- SCHONFELD, R. (2017). "Elsevier acquies bepress". 2 août 2017. Billet de blog. <https://scholarlykitchen.sspnet.org/2017/08/02/elsevier-acquies-bepress/>
- SERRES, A. (2012). *Dans le labyrinthe. Evaluer l'information sur internet*. C&F Editions. 224 p.
- SERRES, A., MALINGRE, M.-L., MIGNON, M., PIERRE, C., COLLET, D. (2017). « Données de la recherche en SHS. Pratiques, représentations et attentes des chercheurs : une enquête à l'Université Rennes 2 : Rapport et Annexes ». [Rapport de recherche] Université Rennes 2. 2017, 159 p.. <hal-01635186v1>
- SHARPLES, M. (2002). *How we write : Writing as creative design*. Routledge.
- SIMONNOT, B. (2009). Culture informationnelle, culture numérique : au-delà de l'utilitaire. *Les Cahiers du numérique*, vol. 5,(3), 25-37. <https://www-cairn-info/revue-les-cahiers-du-numerique-2009-3.htm-page-25.htm>.
- STAINES, H., & MARTONE, M. E. (2018). Community feedback on scholarly content : why it is important and why it should be preserved. *Insights*, 31, 13. DOI : <http://doi.org/10.1629/uksg.418>
- VERLAET, L., DILLAERTS, H. (2016) « L'enjeu du web de données pour l'édition scientifique », *I2D – Information, données & documents 2/2016 (Vol. 53)*, p. 49-49. URL : www.cairn.info/revue-i2d-information-donnees-et-documents-2016-2-page-49.htm

ANNEXES

« 101 innovations »

Le questionnaire correspondant au processus de recherche-publication tel qu'il a été modélisé se décomposait en six phases auxquelles correspondaient pour chacune des entrées de ces phases, une liste présélectionnée de sept outils à cocher, plus une entrée libre.

1. Découverte

- a. Recherche de littérature, données, etc,
- b. Accès à la littérature, données, etc.
- c. Alertes, recommandations,
- d. Lire, visionner, annoter

2. Analyse

- a. Analyse des données, textes, etc.
- b. Partage de Notebooks, protocoles, processus de travail,

3. Écriture

- a. Ecriture, préparation du manuscrit
- b. Gestion des références

4. Publication

- a. Archivage/Partage des publications
- b. Archivage/Partage des données et du code source
- c. Choix de la revue pour soumettre un article
- d. Utilisation des Plateformes, Revues et Sites d'éditeur

5. Portée

- a. Archivage/Partage des posters et présentations
- b. Outils utilisés pour diffuser ses recherches en dehors du milieu académique
- c. Outils de Profils pour chercheurs

6. Évaluation

- a. Outils utilisés pour l'évaluation par les pairs (en dehors du processus mis en place par les revues) : commentaires et évaluations de pairs, pré-publication et post-publication
- b. Mesure d'impact
- c. Outils spécifiques au pays ou à la langue (Sudoc, Francis, Hal, Revues.org, Héloïse, Hypothèses, IdRef, et (autres).

NOTES

1. À l'origine 25 820 réponses desquelles on a retiré les spams. Les données des répondants sont toujours disponibles et comprennent le pays d'affiliation, la discipline de recherche, le rôle dans la recherche et l'année de la première publication.
2. Une limite de cette étude résidait dans le choix non explicité des outils retenus : par exemple, l'absence de Zotero, alors que cet outil est devenu commun dans la gestion en ligne des références bibliographiques ou le choix de retenir pour ResearchGate retenu au détriment d'Academia.edu.
3. Une catégorie « autre » non identifiée récolte cependant 8 %.
4. Il ne nous sera donc pas possible de ré-utiliser cette grille pour une analyse des usages d'outils d'écriture.
5. Précision que le choix des outils n'était pas exclusif, il était donc possible d'en sélectionner plusieurs.
6. On peut se demander pourquoi Academia.edu (spécialisée SHS) n'a pas été proposé. ResearchGate qui est orienté STM et SHS apparaissait sans doute aux auteurs de l'enquête plus à même de refléter la diversité des chercheurs.
7. PaperCritic est un service qui a fermé le 26 mars 2015.
8. URL : <http://asapbio.org/letter>
9. Les éditeurs (nombre de revues) sont : ASCB (3), BioMed Central (67), BMJ (4), Company of Biologists (1), Copernicus Publications (19), EDP Sciences (1), eLife Sciences Publications (1), Elsevier (2), EMBO Press (4), F1000 (1), MDPI (3), Nature Publishing Group (1), PeerJ (1), PLOS (7), Rockefeller University Press (1), Royal Society (4), SfN (1), Wellcome Trust (1).
10. URL : <https://www.crunchbase.com/>, URL : <https://acquiredby.co/> , URL : <https://scholarlykitchen.sspnet.org/>
11. <https://scholarlykitchen.sspnet.org/2017/08/02/elsevier-acquires-bepress/>
12. Le Centre National de la Recherche Scientifique (CNRS), l'Agence Bibliographique de l'Enseignement Supérieur (ABES), le Consortium Universitaire de Publications Numériques

(Couperin1), et l'Université de Lorraine agissant pour le compte de la Conférence des Présidents d'Université (CPU).

13. <https://www.w3.org/annotation/>

14. <http://www.openannotation.org/>

15. Web Annotation Data Model : <https://www.w3.org/TR/annotation-model/>

Web Annotation Vocabulary : <https://www.w3.org/TR/annotation-vocab/>

Web Annotation Protocol : <https://www.w3.org/TR/annotation-protocol/>

RÉSUMÉS

Nous proposons ici d'observer les usages et non usages de l'offre logicielle visant l'écriture, la lecture, la préservation des données, des références, des citations et leur traitement. Nous établissons le lien entre cette panoplie d'outils, services, plateformes, avec les acteurs de la recherche responsables de cette offre : start-ups émanant de projets de recherche, éditeurs et groupes de médias scientifiques. Nous complétons par une exploration des services et créations logicielles liées au web de données et à l'éditorialisation sémantique des articles qui représentent des innovations notables encore à l'état de prototypes.

We propose here to observe the uses and non-use of the software offer for writing, reading, preservation of data, references, quotes and their processing. We establish the link between this range of tools, services, platforms, with the actors responsible for this offer : start-ups from research projects, publishers and media groups. We complete with an exploration of services and software creations related to the web of data and the semantic editorialisation of the articles.

INDEX

Mots-clés : littératie informationnelle, comportement informationnel, outils d'écrilecture, éditorialisation sémantique

Keywords : information literacy, information behavior, reading and writing tools, semantic publishing

AUTEUR

EVELYNE BROUDOUX

Evelyne Broudoux est Maître de conférences HDR au Cnam, à l'INTD (Institut national des sciences et techniques de la documentation). Elle est membre permanent de l'équipe DICEN-IDF (Dispositifs d'information et communication à l'ère numérique). Courriel : evelyne.broudoux@cnam.fr