

L'édition électronique : petite mise au point (mars 2008)

Emmanuelle Picard, SHE, INRP-ENS (epicard@club-internet.fr)

Claire Lemercier, IHMC, CNRS (Claire.Lemercier@ens.fr)

L'édition en ligne, i.e. la publication sur Internet d'articles, thèses, communications à des colloques, livres... est actuellement en plein développement. Elle permet d'accroître la visibilité de l'activité scientifique, d'améliorer sa diffusion, voire de la mesurer ou d'en saisir la nature. De ce fait, elle se trouve au cœur de la triple problématique de la diffusion, de la réception et des usages de l'activité éditoriale en sciences humaines.

Elle est susceptible par ailleurs de devenir rapidement le support essentiel des opérations de bibliométrie dans les années à venir, comme c'est déjà le cas dans d'autres champs du savoir. Une première base de données réalisée à partir des publications électroniques a ainsi vu le jour sur le modèle du *Web of Science* (base de données sur les publications scientifiques qui existe depuis une cinquantaine d'année et qui est à l'origine de la mise en place de la bibliométrie couramment utilisée en sciences exactes) : Scopus, qui appartient à la compagnie Elsevier (<http://info.scopus.com/overview/what/>).

Deux modalités différentes de la publication électronique des textes scientifiques doivent être distinguées.

Ils peuvent, d'une part, être mis en ligne directement sous un format html ou, encore mieux, xml (page Internet) ou sous forme d'un fichier à télécharger (format doc, rtf, pdf...) à partir d'une page web (personnelle, d'un laboratoire, dédiée à un événement...). On peut retrouver ces documents par le biais d'une recherche sur un moteur généraliste, de type Google ou Yahoo ; mais les résultats seront alors noyés parmi un ensemble de pages comportant des données proches, quelle que soit leur nature (scientifique mais aussi commerciale) : c'est ce que l'on appelle le « bruit ».

On peut, d'autre part, déposer les textes sur des portails, ou dépôts, répondant aux normes d'une mise en ligne raisonnée et standardisée, adaptée aux productions scientifiques. Il s'agit de sites qui acceptent les dépôts de textes accompagnés de « méta-données ». Ces méta-données, présentes pour chaque texte (auteur, titre, lieu de publication, date, résumé, mots-clés...) et répondant à un protocole internationalement reconnu (OAI-PMH), permettront ensuite un référencement beaucoup plus efficace dans des moteurs de recherche spécifiquement dédiés aux productions scientifiques¹. Enfin, ces sites génèrent pour chaque texte déposé une adresse URL stable, ce qui facilite ensuite la citation du texte en ligne.

Les portails ou sites de dépôt sont de différentes sortes : les uns sont consacrés exclusivement aux revues (articles, compte rendus, notes de lecture...) ; les autres accueillent toutes sortes de productions (articles, mais aussi thèses, contribution à des ouvrages collectifs, communications...).

1 On peut trouver un premier aperçu de la richesse des ressources disponibles sur l'Album : <http://album.revues.org/index31.html> (rubrique « l'édition électronique ») et <http://album.revues.org/index28.html> (« moteurs de recherche spécialisés »).

Les portails de revues

Ils prennent en charge la mise en ligne des versions électroniques des sommaires, articles, résumés, compte rendus, notes de lecture... sous un double format, permettant la lecture en ligne et le téléchargement, et autorisant des recherches en « plein texte ». L'accès aux revues peut être totalement gratuit (accès à tous les documents sans restriction aucune), mais il est le plus souvent payant pour les numéros récents (on parle de barrière mobile : à 4 ans, elle signifie que les 4 dernières années sont payantes, les années précédentes étant gratuites ; certains ont des barrières mobiles de 2 ans, comme Erudit pour les revues québécoises). La plupart du temps, une partie du contenu éditorial est accessible gratuitement, en particulier les sommaires, résumés, informations diverses et compte rendus.

C'est dans le domaine anglophone que les premiers portails ont vu le jour, essentiellement pour les « sciences dures ». En sciences humaines, le principal est à l'heure actuelle Jstor (<http://www.jstor.org>) : portail avant tout patrimonial, il contient les articles publiés jusqu'à il y a cinq ans de plus de 700 titres de périodiques en anglais, à quelques exceptions près. On trouve également Muse (<http://muse.jhu.edu>), qui compte 350 revues, elles aussi majoritairement anglophones, mais, à la différence de Jstor, accessibles jusqu'au dernier numéro. L'accès à ces deux portails est payant et ne se fait que par le biais d'un abonnement institutionnel (bibliothèque, grand établissement...).

On trouve également des revues de sciences humaines sur :

- Sciencedirect : <http://www.sciencedirect.com/>
- Springer Link : <http://www.springerlink.com/home/main.mpx>
- I-revues : <http://irevues.inist.fr/>
- Periodicals Archive Online : <http://pao.chadwyck.co.uk/marketing/index.jsp>
- La Criée : <http://periodiques.wordpress.com/>

Cette liste n'est bien sûr pas exhaustive et demande à être complétée.

Dans le domaine francophone européen, trois portails principaux se partagent actuellement le terrain dans le domaine des SHS :

- le plus ancien est Revues.org : <http://www.revues.org/>

Né en 1999, il compte aujourd'hui 90 revues en ligne, 120 revues adhérentes et 30 adhésions en cours d'examen. L'équipe du Centre pour l'édition électronique ouverte (CLEO : EHESS, CNRS, Universités d'Avignon et de Provence) en est responsable. Il fonctionne avec un logiciel libre, Lodel. Les revues adhérentes à Revues.org sont chargées de la gestion de la mise en ligne des articles et du contenu du site de la revue, l'adhésion étant gratuite.

- vient ensuite Cairn : <http://www.cairn.info/accueil.php>

Né en 2005, issu d'une collaboration entre l'Université de Liège et des éditeurs privés, il compte aujourd'hui 35 062 articles parus depuis 2001 dans 150 revues de recherche et de débat. Le portail s'inscrit dans une logique commerciale, rétrocédant aux revues une partie des produits de ventes d'articles et abonnements. Le travail de mise en ligne est financé par le reversement par les revues à Cairn des subventions reçues pour mise en ligne.

- enfin, Persée : <http://www.persee.fr/>

Ce portail, à la différence des deux précédents, a une vocation uniquement patrimoniale : il n'accueille que les collections anciennes (antérieures à 2000) des revues, qui sont numérisées à cette occasion. Le site est mis en œuvre dans le cadre d'une collaboration entre la Direction de l'enseignement supérieur et des bibliothèques et l'université Lumière Lyon 2. Son plan de développement pour 2008 prévoit par ailleurs l'accueil de monographies de différents types, en plus des revues.

La consultation de Persée est libre et gratuite. La consultation de Revues.org dépend du choix fait par la revue concernée : l'accès peut être totalement libre ou payant pour une période déterminée. La consultation sur Cairn est toujours payante dans les limites de la barrière mobile fixée par les revues, à savoir en général quatre années, et permet aux particuliers l'achat par article. Les bibliothèques et les établissements scientifiques s'abonnent généralement aux portails ou aux revues : les utilisateurs ont alors l'accès complet et gratuit à l'ensemble des données.

Le CNRS a acheté, non seulement pour ses personnels, mais pour toutes les personnes rattachées à une UMR (doctorants inclus), un accès global à un grand nombre de portails (et bases de données). L'intérêt est que cet accès se fait par mot de passe, donc depuis tout ordinateur, y compris privé. Ces mots de passe doivent être demandés au directeur du laboratoire. Le site <http://biblioshs.inist.fr> permet d'activer ces accès et par là même de repérer où sont mises en ligne un grand nombre de revues.

Globalement, retrouver une revue en ligne n'est pas très compliqué : il suffit généralement de taper son titre dans Google. Cependant, certains sites ont comme objet de recenser les revues en ligne et d'indiquer les adresses Internet où les consulter :

- le Directory of Open Access Journals : <http://www.doaj.org> recense les lieux de dépôts des revues en accès libre.

- la bibliothèque universitaire de Ratisbonne a créé un site extrêmement précieux en la matière, puisqu'il recense, pour chaque revue en ligne, le lieu où elle se trouve et les conditions d'accès aux articles : l'*Elektronische Zeitschriftenbibliothek* (EZB, Electronic Journals Library ou Bibliothèque des revues électroniques) accessible à l'adresse suivante :

<http://rzblx1.uni-regensburg.de/ezeit/index.phtml?bibid=AAAAA&colors=7&lang=en>

Les portails multi-supports

Ces portails contiennent des documents de différentes natures et de différents statuts, qu'il s'agisse d'articles, de collections de documents ou de livres numérisés.

- Gallica : <http://gallica.bnf.fr>

il s'agit de la bibliothèque numérique de la BNF, qui met gratuitement à disposition des utilisateurs un ensemble de documents numérisés issus de ses fonds. L'accès aux textes est cependant conditionné par la notion de « domaine public », date à partir de laquelle un document édité est considéré comme libre de droits, en l'occurrence 70 ans après la mort de l'auteur. Il n'y a encore que peu de revues de sciences humaines qui aient fait l'objet d'une telle numérisation, mais le principe est de développer des collaborations concernant les collections anciennes.

- Erudit : <http://www.erudit.org>

D'abord été mis en place en 1998 à l'initiative des Presses de l'Université de Montréal, comme site de diffusion numérique des revues, il s'est élargi à la publication de livres électroniques, des thèses et autres documents. Il contient actuellement environ 50 revues avec une barrière mobile de 2 ans, aux côtés d'un grand nombre d'autres documents.

- Presses universitaires de Lyon :

http://presses.univ-lyon2.fr/rubrique.php3?id_rubrique=2

Elles développent un programme de publication électronique destiné à rendre librement accessible un ensemble de publications scientifiques, ainsi que certains ouvrages épuisés de leur fonds ; l'ensemble de ces publications est en libre accès.

Les archives ouvertes

Ces sites sont destinés au dépôt de productions scientifiques, depuis les versions préliminaires d'un texte jusqu'aux versions publiées, que ce soit dans un ouvrage collectif, des actes de colloque ou une revue, voire des communications jamais éditées. Ils sont fondés sur le dépôt volontaire par les chercheurs eux-mêmes (auxquels ils sont réservés). Très nombreux dans les pays anglo-saxons, ils se développent plus lentement en Europe. En France, les deux principaux sites d'archives ouvertes en SHS sont :

- HAL-SHS : principal site d'archives ouvertes français, géré par le CNRS. Cette plateforme fait l'objet d'un consensus de la part des institutions scientifiques françaises comme devant être le lieu unique de dépôt des travaux scientifiques en sciences humaines et sociales : <http://hal.archives-ouvertes.fr/>. Une partie du dépôt HAL est spécifiquement dédiée aux thèses et aux habilitations (TEL) : <http://tel.archives-ouvertes.fr/>

- @rchiveSIC : archives ouvertes en sciences de l'information
<http://archivesic.ccsd.cnrs.fr/>

The Directory of Open Access Repositories – OpenDOAR, <http://www.opendoar.org>, donne accès à une liste de sites d'archives ouvertes sur toute la planète.

Le rôle des méta-données et les moteurs de recherche spécifiques

Cet ensemble représente des millions de pages et de références éparpillées sur des sites très divers, aux modes d'accès et d'interrogation variés. L'enjeu est de développer les outils qui permettent de les interroger ensemble en même temps.

Les méta-données constituent la clé d'entrée dans la mise en place d'une édition électronique de qualité. Il s'agit de l'ensemble des informations accompagnant les documents (de son auteur à ses mots-clés) et qui permettent d'effectuer une recherche par le biais de moteurs de recherche spécifiques. L'enjeu est de constituer, sur Internet, un espace dédié à l'activité scientifique, au sein duquel puissent se mener des recherches qui ne seront pas

polluées par tous les autres domaines d'activité. Ce mode de fonctionnement se fonde sur le principe de l'interopérabilité (tous les sites donnant accès en ligne à ce type de documents souscrivent à un protocole, appelé OAI-PMH, qui assure la reconnaissance des documents qu'ils contiennent).

Ainsi « encadrés », les documents peuvent être repérés et proposés comme réponses à la requête d'un internaute. Par ce biais, on peut retrouver tous les documents déposés par tel auteur, ou dans telle revue, ou sur tel sujet. Parmi les moteurs de recherche spécifiques au web scientifique, signalons-en trois en particulier :

- OAIster : <http://www.oaister.org/>

OAIster est un projet de l'*University of Michigan Digital Library Production Service*. Il constitue le plus grand moteur de recherche actuel, « moissonnant » parmi la documentation scientifique disponible en accès libre dans 800 dépôts OAI dans le monde, soit actuellement plus de 15 millions de documents.

- Google Scholar : <http://scholar.google.fr/>

Reprenant l'ambition de Google, Google Scholar se présente comme le principal moteur permettant d'interroger les sites d'archives ouvertes et de dépôts de publications scientifiques (bibliothèques, portails...). Il permet de multiples interrogations et grossit par le seul fait de sa notoriété : en effet, les portails qui se développent négocient avec lui leur référencement. C'est le cas de HAL-SHS, ainsi que des portails des revues mentionnés plus haut. Google Scholar non seulement permet de retrouver les textes dont on a besoin, mais propose également de repérer les citations dont ils ont fait l'objet. Dans certaines disciplines de sciences humaines et sociales, il commence à être utilisé pour la bibliométrie.

- In-extenso : <http://www.in-extenso.org/>

Ce moteur de recherche développé par Revues.org est dédié aux sciences humaines et sociales. Il « moissonne » lui aussi les dépôts d'archives ouvertes, ainsi que les pages web spécifiques, en se concentrant sur le web francophone.

Conclusion

- **Consulter** : c'est la partie la plus facile. Chaque mois, le volume des productions scientifiques accessibles en ligne s'accroît. On a donc de plus en plus intérêt à utiliser les moteurs de recherche spécifiques pour trouver l'information et télécharger les textes qui nous intéressent. La seule difficulté réside dans l'accès payant à un certain nombre de ces documents. On peut y remédier en les consultant depuis les postes informatiques des bibliothèques ou sur des ordinateurs branchés sur des réseaux institutionnels (c'est le cas du réseau de l'ENS, par exemple, qui permet l'accès direct à presque tout).
- **Mais surtout déposer** : pratique qu'il s'agit de développer au plan individuel si l'on veut que la consultation ait un sens. Les avantages sont de plusieurs ordres : on assure la pérennité de la diffusion de ses travaux (les sites en question, HAL-SHS par exemple, sont des créations institutionnelles pérennes, à la différence des pages personnelles des chercheurs). Ensuite, on contribue à leur diffusion, particulièrement à l'extérieur de nos frontières, et donc à leur usage. Enfin, et cela découle de la raison

précédente, on augmente les citations dont ils font l'objet... ce qui ne pourra manquer de compter dans les années à venir². La procédure de dépôt sur HAL-SHS n'est pas compliquée en soi, elle prend seulement un peu de temps (quelques minutes par document). Elle est particulièrement indiquée pour les textes difficiles d'accès (communications non publiées, textes anciens...), mais a pour vocation à terme de concerner l'ensemble de la production scientifique sous toutes ses formes. Reste à retrouver ses anciens fichiers... et à scanner ceux dont on n'a plus la version informatique (à déposer en format pdf)³.

- **La question du lieu de dépôt** sera certainement centrale dans les années à venir : dans le modèle anglo-saxon, les institutions mettent en place des dépôts propres (que l'on pourrait qualifier de dépôts locaux) et incitent fortement les chercheurs à les utiliser : on peut donc déposer sur le portail de son université ou de son établissement (c'est la cas à Lyon 2 actuellement). Mais le choix français a été plutôt de développer de gros portails nationaux, type HAL-SHS, par grands domaines du savoirs. Cela ne change pas grand chose en pratique, même si l'on peut penser que la proximité des portails est une incitation bien plus grande au dépôt effectif des travaux.

Il semble aujourd'hui difficile d'ignorer les ressources prodigieuses déjà offertes par l'édition électronique. Mais la France souffre encore d'un fort déficit de visibilité en la matière. Il est donc souhaitable que se développe aujourd'hui largement la pratique du dépôt de nos travaux.

2 De nombreux travaux ont montré que les citations des documents en accès libre sont plus élevées que pour ceux qui ne sont accessibles que par abonnement. Voir par exemple : Hajjem, C., S. Harnad et Y. Gingras, « Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation », *IEEE Data Engineering Bulletin*, vol. 28, n° 4, 2005, pp. 39-47. Harnad, Stevan, Tim Brody, François Vallières, Les Carr, Steve Hitchcock, Yves Gingras, Charles Oppenheim, Heinrich Stamerjohanns et Eberhard R. Hilf, « The Green and the Gold Roads to Open Access ». *Nature Web Focus*, 2004. On trouve ces textes à l'adresse suivante :

<http://www.chss.uqam.ca/PUBLICATIONS/Scientom%C3%A9trie/tabid/64/Default.aspx>

3 L'usage veut que l'auteur demande à l'éditeur, ou à la revue, son accord pour le dépôt dès lors qu'il veut mettre en ligne un tiré à part. Rien n'empêche cependant la mise en ligne d'une version doc ou rtf correspondant au texte non mis en page.