

# Accès ouvert, auto-archivage et weblogs

Par Christophe Deschamps

Ci-dessous une chronique publiée dans Veille Mag de décembre 2005. Je me suis un peu éloigné de mes centres d'intérêt habituels et remercie [Olivier Ertszcheid](#) pour son aimable et attentive relecture.

Le 17 novembre dernier les 176 participants du Sommet Mondial sur la Société de l'Information (SMSI) se sont engagés par la Déclaration de Tunis à combler le fossé numérique existant entre pays du Nord et du Sud d'ici 2015.

L'agenda<sup>[1]</sup> qui l'accompagne précise que si la mise en oeuvre de moyens financiers, politiques et juridiques est indispensable elle ne saurait suffire à la réduction de la fracture numérique et invoque sans les nommer précisément les principes de l'accès ouvert et de l'auto-archivage des publications scientifiques, concrétisant ainsi l'idée qui veut que la meilleure façon d'aider l'indigent à se nourrir est de lui apprendre à pêcher.

Ces principes ont été définis au début des années 90 par des scientifiques aux prises avec les évolutions du fonctionnement d'un système éditorial centré depuis la fin du XVII<sup>ème</sup> siècle sur l'obligation de révision par les pairs (*peer review*) et de publication dans des revues savantes.

Le premier facteur d'évolution est venu d'un ressentiment croissant des chercheurs envers les grandes maisons d'édition comme Elsevier, Springer ou Blackwell, qui, tout en leur demandant de céder leurs droits sur les articles qu'ils produisaient, augmentaient (augmentent toujours!) les tarifs d'abonnement aux revues nécessaires à leurs travaux. Sentiment d'exaspération également devant la stratégie commerciale de certains d'entre eux qui, contre un peu de battage publicitaire, n'hésitaient pas à communiquer des résultats de recherches inédits à la presse généraliste avant de les adresser à l'ensemble de la communauté scientifique.

Le second est lié à ce fameux mécanisme d'évaluation des articles par les pairs qui, bien qu'indispensable, rend certains travaux obsolètes avant même leur publication et ralentit ainsi la dynamique de la recherche.

Le troisième facteur est bien sûr l'immixtion dans ce contexte du Web et de sa formidable capacité à diffuser instantanément l'information.

La conjonction de ces éléments va rapidement produire un résultat prévisible puisque dès 1995 le chercheur en sciences cognitives Stevan Harnad lance une « *subversive proposal* » dans laquelle il appelle les chercheurs à publier leurs prépublications en libreaccès sur des serveurs web ouverts dans lesquels la communauté scientifique mondiale pourra piocher à l'envi<sup>[2]</sup>.

Cette idée d'auto-archivage est indissociable de celle d'accès ouvert (open access) qui consiste à proposer gratuitement des revues scientifiques numériques de qualité.

Dix ans plus tard il existe un peu plus de 7000 serveurs de prépublications scientifiques et techniques et environ 1000 revues libres contre 25000 payantes. Le succès n'est pas total mais l'idée que l'information scientifique doit être libre et gratuite fait son chemin, d'autant que certaines études tendent à montrer que le taux de citations des articles en accès ouvert est souvent plus élevé que celui des articles de revues classiques, un argument qui devrait pousser de plus en plus de chercheurs à adopter ce modèle.

Pourtant chaque système génère ses propres blocages et celui-ci n'échappe pas à la règle.

### **Publication ouverte et blogs personnels**

Le chiffre de 7000 serveurs cité précédemment est approximatif car le premier problème est bien de les identifier afin de les référencer. Une piste d'amélioration pourrait venir des weblogs qui, en assurant la transition d'un système centralisé vers un système décentralisé, ferait de chaque scientifique l'éditeur de ses propres prépublications, justifiant ainsi pleinement le terme d'auto-archivage. Techniquement il me semble que la difficulté est loin d'être insurmontable. Les plateformes de blogs génèrent automatiquement des fils RSS<sup>[3]</sup> qui sont une version simplifiée du langage XML. Il suffirait d'une adaptation des métadonnées de ce format vers celui qui a été défini pour les publications scientifiques, l'Open Access Initiative Protocol for Metadata Harvesting (OAI – PMH), pour que chaque article puisse être indexé par les robots d'un « moissonneur OAI » de type OAIster. Par ailleurs, doubler cette solution d'un système déclaratif tel qu'on peut en trouver sur les moteurs généralistes, ou d'un service comme Ping-o-Matic, permettant d'informer les moteurs de la mise à jour de son blog, consoliderait le système. En effet, un scientifique qui publie actuellement sur ce type de serveur doit finalement faire confiance à ceux qui le gèrent quant à la propagation de son article vers le plus grand nombre de serveurs ouverts possibles. Sachant que son but est évidemment d'être lu par le maximum de confrères il ne serait pas impensable de mettre à sa disposition les outils et services lui permettant d'assurer personnellement sa promotion.

De nombreux thésards et chercheurs utilisent déjà les blogs à cet usage mais sont confrontés à au moins deux obstacles. Le premier concerne le point évoqué plus haut de la non-compatibilité du protocole Rss avec le protocole OAI-PMH qui les oblige à dupliquer leur article sur un serveur ouvert.

Le second est celui du mélange entre le côté intime du blog et la rigueur qui est de mise sur un média professionnel. Certains chercheurs-bloggeurs ont déjà pu Outre-Atlantique faire les frais du ton un peu trop libre de leurs billets et c'est tout leur parcours professionnel qui risque d'en pâtir car, sur le web plus qu'ailleurs, les écrits restent. Pour éviter ce type de contamination dommageable il faut sans doute imaginer des services tournés uniquement vers l'univers professionnel. Nul doute que la communauté Open Source dans laquelle les scientifiques sont fortement représentés pourrait facilement y pourvoir.

### **Revue par les pairs ou gestion de la réputation ?**

Mais n'y a-t-il pas moyen d'aller encore plus loin en donnant à chaque scientifique la possibilité de proposer en accès ouvert des articles dignes de confiance sur son propre weblog. Il ne s'agirait pas de passer outre la nécessaire validation par les pairs mais tout

simplement d'en modifier les modalités, car de quoi s'agit-il ici sinon de gérer la réputation de quelqu'un, et donc la confiance qu'on peut lui accorder ? C'est une problématique que l'on retrouve depuis les débuts du web et qui connaît déjà plusieurs solutions allant de la notation des vendeurs sur eBay aux systèmes plus complexes de recommandations présents sur les services de « réseautage » en ligne[4] professionnels de type Viaduc, OpenBC ou 6nergies. C'est également l'idée par exemple d'un service comme Squidoo (en beta) qui permet à chacun de créer une page personnelle centrée sur ses domaines d'expertise et qui est dotée par défaut d'un mécanisme de notation apte à assurer (ou non !) sa réputation.

Un tel système amélioré et sécurisé ne serait-il pas susceptible de se substituer, ou, à tout le moins, d'épauler le mécanisme traditionnel de revue par les pairs ?

Enfin gageons qu'un dispositif un peu élaboré de labellisation personnelle (tagging) semblable à celui proposé par le service en ligne del.icio.us[5], pourrait parfaitement assurer un classement viable des articles de la communauté scientifique à court et moyen terme, quitte à le doubler plus tard d'un système de classement plus classique. Deux services allant en ce sens existent d'ailleurs déjà, il s'agit de Connotea, créé par la revue Nature et de CiteUlike.

Pour en revenir à la réduction de la fracture numérique il me semble évident que les outils permettant aux scientifiques des pays du Sud de s'informer et de publier leurs travaux existent en partie ou pourraient facilement être déployés en modifiant de manière minimale certains services existants.

Le fond du problème n'est donc pas là mais, d'une part, dans certains choix qui remettraient en cause le fonctionnement actuel du système de l'édition scientifique en modifiant notamment le critère d'évaluation constitué par le facteur d'impact[6] d'une revue, c'est-à-dire sa notoriété et, par effet miroir, celle des scientifiques qui y sont publiés. D'autre part bien sûr, dans le refus des éditeurs scientifiques de renoncer à leurs revenus.

### **Ressources complémentaires :**

- Site de Inist entièrement consacré à la problématique de l'Open Access : <http://www.inist.fr/openaccess>
- Communication scientifique et Internet : site de l'INRA consacré à l'Open Access : [http://www.tours.inra.fr/prc/internet/documentation/communication\\_scientifique/comsci.htm](http://www.tours.inra.fr/prc/internet/documentation/communication_scientifique/comsci.htm)
- Bimpe (Base d'Information Mutualiste sur les Périodiques Electroniques), sélection d'outils et d'infos sur l'accès ouvert : [http://bimpe.free.fr/outils\\_et\\_infos\\_ao.htm](http://bimpe.free.fr/outils_et_infos_ao.htm)
- Base de données listant l'ensemble des revues en Open Access : <http://archives.eprints.org/>

Articles intéressants à télécharger sur @rchiveSIC, serveur ouvert de publication en sciences de l'information :

- Christian Gerini : L' « Open Access », pour une réelle liberté de la communication scientifique : état des lieux et problématiques, avril 2005, [http://archivesic.ccsd.cnrs.fr/sic\\_00001425.html](http://archivesic.ccsd.cnrs.fr/sic_00001425.html)
- Michèle Battisti : Libre accès à l'information scientifique et technique : état de l'art et perspectives, septembre 2004, [http://archivesic.ccsd.cnrs.fr/sic\\_00001068.html](http://archivesic.ccsd.cnrs.fr/sic_00001068.html)
- Nathalie Pignard : La publication scientifique sur Internet, juillet 2003, [http://archivesic.ccsd.cnrs.fr/sic\\_00001414.html](http://archivesic.ccsd.cnrs.fr/sic_00001414.html)

### Sur le protocole OAI-PMH :

Site de l'Open Archives Initiative : <http://www.openarchives.org/>

### Services utiles pour les veilleurs :

- Chercher dans les archives ouvertes :
  - OAster : <http://oaister.umdl.umich.edu/o/oaister>
  - Citeseer : <http://citeseer.ist.psu.edu/>
  - Hal (Hyper Article en Ligne) (fr) : <http://hal.ccsd.cnrs.fr/>
  - Thèses en ligne (fr) : <http://tel.ccsd.cnrs.fr/>
- Lister les citations d'un auteur ou d'un article
  - Citebase : <http://www.citebase.org>
  - Citeseer : <http://citeseer.ist.psu.edu/>
- Déclarer la publication de nouveaux billets sur un blog
  - Ping-O-Matic : <http://pingomatic.com/>
  - Pingoat : <http://www.pingoat.com/>
- Services de bookmark sociaux pour chercheurs :
  - CiteUlike : <http://www.citeulike.org/>
  - Connotea : <http://www.connotea.org/>
- Service de recherche d'expert :
  - Squidoo (actuellement en version beta) : <http://www.squidoo.com/>

---

[1] La section 90-k invite à appuyer « *les institutions à but éducatif, scientifique et culturel*, notamment les bibliothèques, les archives et les musées, dans leur mission, qui consiste à élaborer et préserver des contenus variés et à offrir un accès équitable, ouvert et peu coûteux à ces contenus, y compris les contenus numériques, pour faciliter l'enseignement formel et informel, la recherche et l'innovation».

[2] Pour être complet il faut noter que ces serveurs existaient depuis 1991 avec la mise en ligne du site arXiv (arxiv.org) par l'Université de Los Alamos. La proposition de Stevan Harnad a accéléré le phénomène.

[3] *Veilleurs syndiquez-vous !*, Christophe Deschamps, Veille Magazine n° 67, septembre 2003

[4] *Se faire un carnet d'adresses avec les services de réseaux sociaux*, Christophe Deschamps, Veille Magazine n° 69, novembre 2003

[5] Del.icio.us (<http://del.icio.us>) permet par exemple de hiérarchiser ses tags au fil de leau. Il propose également de catégories déjà existantes en sappuyant sur les tags d'autres utilisateurs pour chaque nouvelle information à enregistrer

[6] Calcul effectué sur la base du nombre de fois où un article est cité sur une période de deux ans divisé par le nombre d'articles publiés durant cette même période dans cette même revue.