L'archive ouverte HAL-SHS : Comment ça marche, pourquoi s'en servir ?

Caroline Abela

▶ **To cite this version:**

Caroline Abela. L'archive ouverte HAL-SHS : Comment ça marche, pourquoi s'en servir ?. 2009. sic_00407275

HAL Id: sic_00407275

https://archivesic.ccsd.cnrs.fr/sic_00407275

Submitted on 24 Jul 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'archive ouverte HAL-SHS : Comment ça marche, pourquoi s'en servir ?

Caroline Abela

Tired of having your work locked in the ivory tower ? Break out [1]**. C'est ainsi que le mouvement du libre accès à l'information scientifique incite les chercheurs à sortir du modèle économique des publications payantes pour rendre leur production accessible au plus grand nombre. Cette incitation obéit à deux logiques qui s'entrecroisent : d'une part la documentation scientifique est coûteuse, d'autre part le libre accès à l'information scientifique est devenu un enjeu de société, un appel humaniste au partage de la connaissance considérée comme « bien commun ». Dans ce contexte, les archives ouvertes ont pour vocation de faciliter la communication des résultats de la recherche.**

Ce qu'est une archive ouverte :

Une archive ouverte (*Open Archive*) est un site Web sur lequel sont déposées des ressources en ligne issues de la recherche scientifique et de l'enseignement et dont l'accès se veut ouvert, c'est-à-dire ici : en accès libre et gratuit. Il s'agit en fait d'une bibliothèque numérique qui rassemble deux types d'information : les références bibliographiques des publications scientifiques d'une part et, lorsque les droits d'auteur le permettent, une copie électronique de la publication d'autre part. Ainsi, HAL-SHS (Hyper article en ligne - Sciences de l'homme et de la société) est un portail du CNRS qui propose en accès libre sur Internet soit des références bibliographiques seules, soit des références accompagnées du texte intégral du document.

- *En savoir plus*

Pour une histoire concise et claire du concept d'archive ouverte voir : FILY, Marie-Françoise. Introduction au concept d'archive ouverte. *Archivsic*. Juin 2005 [ref. du 09 novembre 2008]. Disponible sur : http://archivesic.ccsd.cnrs.fr/sic_00001523/en/

Pour un glossaire de termes relatifs aux archives ouvertes voir le lexique mis en ligne par l'INRIA : http://www.inria.fr/publications/archiveouverte/lexique.fr.html

Un peu d'histoire :

La première archive ouverte a vu le jour en 1991. Il s'agit d'ArXiv, une archive administrée par la bibliothèque de l'Université de Cornell, initialement dans le domaine de la physique, et qui propose aujourd'hui un réservoir de plus de 500 000 publications (octobre 2008). En France, plusieurs événements marquent l'histoire du

[1] Slogan de l'affiche de promotion du *Self Archive project*, 2004. Cette affiche est disponible sur le site consacré aux archives ouvertes de Couperin.org : Les archives ouvertes pour les établissements d'enseignement supérieur et de recherche

mouvement en faveur du libre accès, qui vise à une meilleure circulation et diffusion de l'information scientifique à travers, d'une part, les revues en ligne et d'autre part, les archives ouvertes. Différents organismes, dont le CNRS, signent en 2003 la Déclaration de Berlin, qui encourage au niveau institutionnel le dépôt des publications scientifiques en libre accès. En 2006, un accord inter établissement est signé entre le CNRS et différents centres de recherche comme le CEMAGREF, le CIRAD, l'INRA, l'INRIA, l'INSERM, l'Institut Pasteur et l'IRD, mais aussi avec la Conférence des grandes écoles, la Conférence des présidents d'universités et enfin l'Académie des sciences, qui s'engagent à la mise en place d'une plateforme commune. Dès 2000, le CCSD (Centre pour la communication scientifique directe) est créé par le CNRS. Il lance en 2002 le serveur multidisciplinaire HAL, qui se déclinera en différents portails dont, en 2005, HAL-SHS pour les sciences humaines et sociales et TEL pour les thèses. Depuis sa création cette archive thématique est en progression constante en nombre de dépôts (plus de 10 000 en janvier 2008, plus de 16 000 en novembre 2008).

➢ *En savoir plus*

Le site du libre accès à l'information scientifique et technique du CNRS propose, entre autres, une rubrique « textes de référence » sur l'histoire du mouvement du libre accès :
http://openaccess.inist.fr/

Fonctionnement et enjeux de l'archive HAL-SHS :

Le principe d'un serveur comme HAL-SHS repose sur celui de l'auto-archivage : les chercheurs sont invités à déposer eux-mêmes leurs publications sur le serveur, s'appropriant ainsi les moyens de diffusion de leur production. Les différents types de publications qui peuvent être déposés sur HAL-SHS avec le texte intégral sont les suivants : des articles dans des revues avec ou sans comité de lecture ; des conférences invitées ; des communications avec ou sans actes ; des ouvrages scientifiques ou des chapitres d'ouvrages. Une rubrique « autres publications » laisse aussi la possibilité de déposer par exemple des *working papers*, des posters ou encore un compte rendu de lecture d'ouvrage. Il est aussi possible de présenter un document en texte intégral sans référence de publication. Si le texte est finalement publié il sera possible de le signaler. Comment déposer dans HAL-SHS ? On ne peut déposer sur le serveur qu'après s'être identifié (il suffit d'ouvrir un compte, il s'agit d'une inscription libre). Différents contributeurs ou déposants sont cependant possibles : l'auteur de la publication mais aussi, par exemple, des documentalistes. Chaque document déposé est accompagné d'une notice descriptive du document que le contributeur doit renseigner. Celle-ci peut prendre en compte l'affiliation de chaque auteur. Quel que soit le format utilisé pour le dépôt, HAL-SHS l'enregistre en format PDF. Le dépôt, qui comprend la rédaction de la notice et la mise en ligne du document, est rapide (compter 10 mn). Chaque dépôt fait l'objet d'une validation technique et juridique sous 24h : l'équipe d'HAL-SHS vérifie qu'il s'agit bien d'une publication de niveau scientifique (correspondant à ce qui serait publié dans une revue ou un ouvrage scientifique). Elle considère également la lisibilité du document, son intégralité, le respect des droits d'auteur et ceux des éditeurs. Selon une règle imposée par HAL, qui veut que les auteurs sont responsables de leurs dépôts, le retrait des documents déposés n'est pas possible en archives ouvertes (*cf.* le mode d'emploi de HAL). Il est par contre possible de déposer des nouvelles versions d'un même texte.

Le fonctionnement d'une archive est caractérisé essentiellement par le stockage et la conservation des données (les notices et les documents en texte intégral), la stabilité des URL (qui assure la pérennité de l'archivage), l'interopérabilité et le moissonnage des métadonnées. L'interopérabilité des différents systèmes d'archives est rendue possible par l'utilisation d'un protocole commun : l'OAI-PMH (the *Open Archives*

Initiative Protocol for Metadata Harvesting). Les métadonnées qui composent chaque référence bibliographique (titre, mots-clés, etc.) sont ainsi moissonnées par les moteurs de recherche, comme Google bien sûr, mais aussi Google Scholar (qui repère les notices avec le texte intégral du document) et aussi des moteurs de recherche spécialisés comme OAIster : qui permet de trouver des ressources électroniques dans un catalogue commun de plus 1 000 archives.

De ce fait, alors même qu'une archive comme HAL-SHS n'a pas pour objectif premier l'évaluation des publications déposées et celle des auteurs, mais plutôt la diffusion de l'information scientifique, elle sert finalement bel et bien à cette évaluation : les données moissonnées par Google Scholar peuvent être utilisées *via* l'application *Harzing publish or perish* pour la création d'indicateurs de notoriété. Ainsi, l'enjeu d'HAL-SHS et des archives ouvertes en général est la visibilité des publications scientifiques, sur le plan national et international. HAL-SHS permet une communication scientifique directe à l'ensemble de la communauté scientifique ainsi qu'à un public plus large. Cette archive permet donc d'accroître sa lisibilité et son lectorat vis-à-vis de ses pairs, mais aussi des organismes de financement comme l'ANR (une rubrique est d'ailleurs prévue dans HAL-SHS pour signaler les documents faisant l'objet d'un projet ANR). Un article déposé sur des archives ouvertes serait cité deux à trois fois plus que s'il était seulement publié dans une revue papier (voir plus loin *En savoir plus*). Dans tous les cas, le déposant ou contributeur peut suivre les statistiques de consultation de ses propres dépôts directement sur HAL-SHS.

Par ailleurs, les références des publications déposées dans HAL-SHS sont automatiquement versées dans les outils administratifs que sont les fiches CRAC (fiches d'activités des chercheurs CNRS) ou le système d'information Labintel. HAL-SHS permet aussi d'exporter très facilement une liste de ses publications (en les classant par exemple par type de publication : articles dans des revues à comité de lecture, etc.), de se constituer sa propre page Web de publications (rubrique « services » sur la page d'accueil de HAL-SHS). L'archive permet également d'utiliser des fils RSS ou bien d'enregistrer un profil personnel afin de recevoir par mail des alertes sur les nouveaux dépôts effectués, et ce en fonction de ses centres d'intérêt.

➢ *En savoir plus*

Sur le lien entre citation d'une publication et dépôt dans une archive ouverte, voir la bibliographie proposée par l'Université de Southampton et l'Université de Cornell dans le cadre du projet *Open Citation* : « The effect of open access and downloads ('hits') on citation impact: a bibliography of studies". [ref. du 23 novembre 2008]. Disponible sur : http://opcit.eprints.org/oacitation-biblio.html

Voir également l'article en ligne : HARNAD, Stevan ; BRODY, Tim. Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals. *D-Lib Magazine*. June 2004 [ref. du 09 novembre 2008]. Disponible sur : http://www.dlib.org/dlib/june04/harnad/06harnad.html)

La foire aux questions de l'Université Lumière Lyon 2 autour de l'idée d'auto-archivage : http://archives.univ-lyon2.fr/faq-boai/

Droits d'auteur - *Creative Commons* :

Quelles sont les règles juridiques à respecter pour le dépôt des documents électroniques ? On distingue l'avant et l'après publication du document. Avant la publication l'auteur est détenteur des droits sur son texte. Lorsque le document est publié il faut considérer s'il s'agit d'une publication sous contrat avec un éditeur ou non. S'il y a contrat, il est possible de déposer la version PDF de l'éditeur à condition que celui-ci donne pour cela son accord explicite. Si l'auteur a signé une clause d'exploitation électronique du document par

l'éditeur, il doit demander l'autorisation de dépôt à l'éditeur et respecter les règles d'embargo. S'il n'y a pas de contrat, l'auteur a le droit de diffuser son document. Dans tous les cas, la version « auteur » du document est toujours déposable, c'est-à-dire sans la charte graphique et les corrections de l'éditeur, à condition de respecter bien sûr par ailleurs le droit des images, et d'obtenir l'accord des coauteurs éventuels.

Du point de vue du droit d'auteur le texte non remanié par l'éditeur est en effet la propriété de l'auteur. Cela n'empêche pas, dans la pratique, d'interroger l'éditeur afin de connaître son point de vue. Pour un certain nombre d'éditeurs ce point de vue est clairement exprimé sur un outil commun à usage des chercheurs : SHERPA RoMEO *(Rights Metadata for Open archiving)*. Il s'agit d'un consortium d'universités et d'instituts de recherche anglo-saxons qui a établi une classification des éditeurs et des titres de revues en fonction des autorisations de publication en ligne qui leur sont liées. SHERPA RoMEO distingue quatre groupes d'éditeurs. Les « *Green publishers* » autorisent le dépôt sur des archives ouvertes des *pre-print* (ou prépublications) : le document qui n'a pas encore été relu et corrigé par un éditeur, mais qui est soumis à un éditeur (en fait le fameux document « auteur »). Ils autorisent aussi les *post-print* (ou postpublications) : le document relu et corrigé par l'éditeur (ou comité de lecture, ou équivalent) mais qui ne comporte pas la mise en page et la charte graphique de l'éditeur. C'est la dernière version avant publication du document sous sa forme éditeur. Les « *Blue publishers* » autorisent seulement les postpublications. Les « *Yellow publishers* » autorisent seulement les prépublications. Les « *White publishers* » n'autorisent aucun dépôt. Plus exceptionnellement, certains permettent le dépôt du PDF de l'éditeur.

Par ailleurs, un auteur peut tout à fait préciser les droits qu'il accorde à l'utilisateur de ses publications en libre accès, en se servant pour cela des licences Creative Commons. Ces licences permettent d'autoriser l'usage de la publication à condition que la paternité de l'oeuvre soit reconnue et le nom de l'auteur cité. Elles permettent aussi d'autoriser ou non une modification, d'autoriser ou non un usage commercial de l'oeuvre et enfin de permettre ou pas sa réutilisation.

➢ *En savoir plus*

Indications de la Direction des affaires juridiques du CNRS sur la propriété intellectuelle et les archives ouvertes : http://www.sg.cnrs.fr/daj/propriete/droits/droits7.htm

Caroline ABELA
Ingénieure documentaliste au CNRS
Centre de documentation REGARDS
UMR ADES
CNRS-Université Bordeaux 3 -Bordeaux 2
Mél. : c.abela@ades.cnrs.fr