L'accès libre en mouvement Journées d'étude sur les archives ouvertes du consortium COUPERIN, 2-3 avril 2009

Joachim Schöpfel, Hélène Prost

▶ **To cite this version:**

Joachim Schöpfel, Hélène Prost. L'accès libre en mouvement Journées d'étude sur les archives ouvertes du consortium COUPERIN, 2-3 avril 2009. 2009. sic_00396957

HAL Id: sic_00396957

https://archivesic.ccsd.cnrs.fr/sic_00396957

Preprint submitted on 19 Jun 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'accès libre en mouvement

Journées d'étude sur les archives ouvertes du consortium COUPERIN, 2-3 avril 2009

Joachim SCHÖPFEL, MCF Université Charles de Gaulle Lille 3, GERiiCO

Hélène PROST, INIST-CNRS

Introduction

Dans le cadre de la faculté de médecine de Paris, le consortium COUPERIN a organisé du 2 au 3 avril 2009 ses premières journées d'étude sur les archives ouvertes. Un vrai succès. Plus de 200 professionnels de l'information ont suivi l'invitation pour assister à 19 conférences de haut niveau sur le contexte national et européen et sur les projets et réalisations français. S'il fallait résumer les communications et débats en trois phrases, les voici :

Les communications et discussions techniques sur les applications et initiatives locales témoignent d'un réel dynamisme du mouvement vers l'accès libre à l'information scientifique en France.

Malgré son statut administratif incertain, l'archive HAL du CCSD conserve son rôle central dans le dispositif français et prépare de nouveaux produits et services pour les communautés scientifiques.

Les archives françaises veulent et doivent s'insérer dans le contexte européen et international, avec des contraintes de normalisation et interopérabilité.

Les communications de ces journées ont été publiées sur le web sous forme de blog[1]. Nous nous contenterons de résumer quelques thèmes au cœur de la question des archives ouvertes.

Réussite technique …

Tous les intervenants étaient du même avis : le dispositif français autour du CCSD, basé sur le principe d'échange entre l'archive nationale HAL et les systèmes locaux, est opérationnel, offre une solution fonctionnelle aux différentes communautés scientifiques et place la France parmi les nations avancées en matière d'archives ouvertes.

A. Alastuey (CCSD) a présenté les projets d'évolution de HAL dont l'extraction de collections thématiques sous forme d'épis-journaux, l'intégration de certains types de documents dans la base (multimédia) ou dans des portails institutionnels (rapports, cartes…), l'archivage pérenne avec le CINES, une plate-forme pour la gestion de congrès ou encore l'indexation avec un moteur sémantique.

[1] http://journeesao.wordpress.com/

De même, les présentations du logiciel ORI-OAI, outil d'indexation et de référencement développé pour les archives institutionnelles universitaires (A.C. Bernaudin) et les communications sur les métadonnées (J.F. Lutz, A.C. Bernaudin et E. Romand-Monnier) témoignaient de la maturation du mouvement vers l'accès libre à l'information en France. La prochaine version 1.5 d'ORI-OAI intégrera le schéma TEF (thèses électroniques), et la fonctionnalité d'échange avec STAR (ABES) et HAL suivra.

… mais échec politique

Le bilan politique est plus contrasté. G. Gourdet (COUPERIN) a regretté qu'il n'y ait pas/plus de structure de coordination et pilotage, ni pour l'ensemble du dispositif français, ni pour HAL. F. Thibault et G. Bru du MENSR ont dressé le même constat d'échec : l'avenir administratif du CCSD dans un CNRS en évolution paraît incertain, ce dispositif n'a eu aucun impact sur les négociations avec les éditeurs, et le succès en terme de dépôts reste assez limité à ce jour[2]. En plus, sans connaître les statistiques d'usage des archives ouvertes, il est difficile de mettre en question les modèles économiques des éditeurs commerciaux qui s'appuient sur les chiffres d'accès.

De son côté, J.C. Guédon a constaté un manque de communication entre les différents acteurs d'IST en France. Peut-être est-ce juste une période de transition. Le CCSD annonce la constitution d'une UPS et la mise en place d'un nouveau comité de pilotage opérationnel, la BNF travail sur un schéma numérique pour les bibliothèques, la bibliothèque numérique fait partie du plan Bresson : histoire à suivre, donc.

Promotion et travail en réseau

Plusieurs intervenants dont J.C. Guédon et N. Lossau (DRIVER) ont souligné la nécessité d'une organisation stable et durable, aussi bien au niveau national qu'international, avec un travail en réseau et une fédération d'institutions mais aussi une structure capable de mener une politique de promotion et de communication, notamment auprès des organismes scientifiques.

La mise en place du groupe de travail sur les archives ouvertes de COUPERIN (GTAO) avec une communication via le nouveau site web[3] est un pas dans cette direction (F. Rivière) ; les partenariats et projets européens du CCSD vont dans le même sens (Europository, PEER, partenariat INRIA-Max Planck Gesellschaft).

Mais l'essentiel, a souligné J.C. Guédon, est de ne pas perdre de vue le chercheur. C'est lui, avec ses objectifs, habitudes et valeurs, qui doit être et rester au centre des réflexions et projets.

Droit

[2] Le chiffre de 15% des publications françaises déposées dans HAL paraît largement surestimé, au vu du dernier rapport de l'OST.
[3] http://www.couperin.org/archivesouvertes/

A plusieurs reprises, les journées ont abordé des aspects juridiques liés aux archives ouvertes :

En 2007, CPU, COUPERIN et CNRS ont travaillé sur les aspects juridiques liés à HAL ; les résultats n'ont été rendus publics.

Le site web du GTAO de COUPERIN contient une liste FAQ avec plusieurs questions juridiques.[4] F. Rivière pense que le conseil juridique constitue une des perspectives d'évolution du site du GTAO.

A plusieurs reprises, la mise en place d'un site SHERPA/ROMEO[5] français a été annoncé afin de recenser les positions et conditions des éditeurs français en matière d'archives ouvertes et d'auto-archivage, un projet pour lequel le CCSD se positionne comme opérateur technique.

Deux questions concernaient directement HAL : Comment le CCSD garantit-il l'exercice du droit moral au retrait du dépôt d'une œuvre ? Réponse : le CCSD devra revoir sa position (= pratiquement pas de retrait de dépôt). Est-ce que le CCSD accepte une licence Creative Commons[6] ? Réponse : non, tous les dépôts sont gérés avec la licence HAL.

Trois cas

OATAO

Sandrine Malotaux a présenté l'Open Archive Toulouse Archive Ouverte[7] (OATAO) qui regroupe les dépôts de quatre établissements d'enseignement supérieur de Toulouse, à savoir l'Ecole Nationale Vétérinaire (ENVT), l'Institut National Polytechnique (INPT), l'Institut Supérieur de l'Aéronautique et de l'Espace (ISAE) et l'Ecole Nationale de Formation Agronomique (ENFA). Soutenu par le Service Commun de Documentation et la Direction de la Recherche de l'INPT, le projet a pour objectif de garder la maîtrise de la diffusion et de l'archivage de la recherche, de fournir des indicateurs pour l'évaluation, de renforcer les équipes autour d'un projet transversal et de proposer des services aux chercheurs. L'interface d'utilisation est souple, pour permettre au chercheur d'effectuer ses dépôts en toute autonomie.

La politique de dépôt diffère selon les établissements ; volontaire à l'INPT, 15% des résultats de la recherche de l'institut sont publiés ; inversement, l'obligation de dépôt à l'ISAE rend visible 70% de la production scientifique.

Regroupant 1652 documents au 1er mars 2009, le contenu global de l'archive présente les caractéristiques suivantes : 32% des documents datent de moins de trois ans ; 57% des documents sont en langue anglaise (87,5% si l'on exclut les thèses) ; 38% sont des articles, 40% sont des thèses, 20% des conférences.

[4] http://www.couperin.org/archivesouvertes/spip.php?rubrique33
[5] http://www.sherpa.ac.uk/romeo/
[6] http://fr.creativecommons.org/
[7] http://oatao.univ-toulouse.fr/

Elaborées à l'aide de Google Analytics, les statistiques d'usage montrent que 57% des visites viennent de France. L'archive est bien référencée car 86% des visites s'effectue à partir de moteurs de recherche ; 79% arrivent via Google. En moyenne, pour chaque visite active, 5 documents sont consultés en texte intégral. Si l'ensemble des articles et des thèses est largement consulté, on enregistre en moyenne 49 et 109 consultations respectivement par article et par thèse.

Les chercheurs approuvent largement le principe de dépôt et sont satisfaits de l'usage; alors qu'ils utilisent très peu les services web 2.0 proposés par l'interface, 60% d'entre eux recommandent la consultation de l'archive, si l'accès à leurs publications est demandé. Le temps moyen nécessaire pour un dépôt est de 13 minutes pour une thèse et de 33 minutes pour un autre document ; le dépôt représente une charge de travail équivalent à 0,75 ETP (équivalent temps plein).

L'interface permet d'évaluer la production scientifique des établissements ; elle propose la répartition des publications par éditeurs et met en relief les différentes affiliations et collaborations entre laboratoires. Cependant, pour obtenir une évaluation au plus juste, il faudrait améliorer la couverture, renforcer la visibilité du service et parvenir à une interconnexion avec HAL. D'autres perspectives : migration vers ORI-OAI, moissonnage au niveau PRES, utilisation du web sémantique, ouverture vers d'autres établissements, obligation de dépôt.

NEEO

François Cavalier a exposé les objectifs du projet NEEO[8] (Network of European Economists Online), soutenu par le consortium NEREUS (Networked Economics Ressources for European Scholars) qui regroupe 23 bibliothèques. Par le biais du portail Economists Online, financé par la Communauté Européenne l'ambition est de proposer l'accès libre aux publications des 500 meilleurs chercheurs, soit environ 50 000 documents (articles de revues, working papers, chapitres de livres, datasets) à l'attention des 30 000 chercheurs en économie et des 250 000 étudiants européens.

Alors que les deux principaux portails concurrents (REPEC, SSRN) détiennent essentiellement des documents en langue anglaise avec accès partiellement payant, Economists Online est un portail européen, multilingue et ouvert car l'archive ouverte permet une meilleure visibilité de la recherche. Appuyé sur une solide structure administrative européenne et différents groupes de travail, le portail propose le moissonnage dans les dépôts propres aux six membres fondateurs. Début 2009, il comptabilise 35 000 références dont 13 000 documents en texte intégral, issus de 900 chercheurs.

Le contenu de la base : 57% working papers, 36% articles. Beaucoup de travaux hollandais et belges.

L'objectif du portail est d'intégrer les datasets, données à partir desquelles sont construits les articles. Malgré les problèmes de formats, de logiciel et d'identifiant et la réticence de certains chercheurs (problème de confidentialité, demande d'un

[8] http://www.neeoproject.eu/

embargo), la création d'un répertoire Dataverse Network[9] permet de déposer et consulter les différents champs de données et de faire le lien avec les publications.

Le portail sera pleinement opérationnel en 2010.

SPIRE

Présentée par Julien Rault, SPIRE[10] est l'acronyme de Science Po Institutional Repository, nom donné à l'archive ouverte de Science Po. Elle fut initiée en 2005, par la bibliothèque, suite à son adhésion au réseau NEREUS. Elle regroupe 2 200 références bibliographiques, essentiellement en économie, dont 50% sont en texte intégral. Utilisée sous D-Space, la bibliothèque et la direction des services informatiques de Sciences Po développent un nouveau logiciel SPIRE, déployé en Java, qui propose une alternative à HAL, ORI-OAI, Eprint ou Castor (Ecole des Mines) et remplacera D-Space fin 2009.

SPIRE rassemble plusieurs briques technologiques existantes et offre un espace de stockage de 500 To. L'archive regroupe, entre autres, différentes métadonnées bibliographiques, des données relatives aux auteurs, aux droits d'auteur et aux statistiques d'usage. Il est possible d'importer des informations relatives à la politique d'archives ouvertes de chaque éditeur ou bien des données LDAP qui permettent l'identification d'une personne physique ou morale.

SPIRE contient les datasets et fait le lien entre ces données et les publications. Connectée à l'archive nationale HAL, SPIRE est bien classée dans le Web ranking des archives institutionnelles.

Questions ouvertes

Plusieurs questions sont restées ouvertes à l'issue de ces journées riches d'informations et de débats. Citons deux qui seront certainement de nouveau à l'ordre du jour en décembre, lors de la conférence internationale sur l'accès libre à l'information à Paris[11] :

Comment organiser l'évaluation des dépôts ? Comment développer ce « capital symbolique » pour le chercheur dont parlait J.C. Guédon ? Par une indexation des publications avec peer review et/ou une gestion du versionnage ? Par une validation scientifique *a minima*, un contrôle de la qualité du contenu et de la forme des dépôts ? Par le calcul d'un facteur d'impact ou d'usage *a posteriori*, après le dépôt, à partir des statistiques d'utilisation voire des annotations par les lecteurs ?

Et comment rendre l'utilisation des archives ouvertes plus simple, plus conviviale et plus en cohérence avec les habitudes des chercheurs ? Par une indexation rigoureuse des différents types de documents, souvent négligée par les archives (J.C. Guédon parle de « fouillis ») ?

En guise de conclusion et pour finir avec une image : les journées d'étude COUPERIN ont montré que la maison des archives ouvertes était solidement

[9] http://thedata.org/ et http://dvn.iq.harvard.edu/dvn/
[10] http://spire.sciences-po.fr/
[11] http://www.inria.fr/actualites/colloques/2009/berlin7/index.en.html

construite. Aujourd'hui il s'agit de rendre les lieux attrayants et habitables, et de s'y installer.