



Archives ouvertes – Vers une obligation de dépôt ?

Synthèse sur les réalisations existantes, les pratiques des chercheurs et le rôle des institutions

Dominique L'Hostis ⁽¹⁾ & Pascal Aventurier ⁽²⁾

- (1) INRA Nantes - Services Déconcentrés d'Appui à la Recherche
Equipe Régionale en Information Scientifique et Technique
courriel : Dominique.L'Hostis@nantes.inra.fr
- (2) INRA Avignon - Services Déconcentrés d'Appui à la Recherche
Equipe Régionale en Information Scientifique et Technique
courriel : Pascal.Aventurier@avignon.inra.fr

Avec la collaboration de :

Hélène Bosc

Chargée de mission INRA
Animatrice du groupe de travail Scientific Publishing d' Euroscience
http://www.euroscience.org/WGROUPS/SC_PUBLISHING/index.htm

Avertissement

Etude réalisée dans le cadre du projet d'archive institutionnelle "Prodira" et soumise aux instances de direction de l'INRA.

Navigation dans le document

Dans le but de faciliter la lecture et permettre un survol rapide de ce texte assez long, une version résumée est proposée en tête du document. Des renvois permettent au lecteur de passer alternativement de la version courte à la version complète.

- En version papier, utiliser les renvois de pages et le sommaire,
- En version électronique : à partir de la version résumée, cliquer sur les titres pour consulter le texte complet, et réciproquement sur les liens « lecture rapide » pour revenir du texte complet à la version courte.

Utilisation & diffusion



Paternité - Pas d'Utilisation Commerciale - Pas de Modification
2.0 France <http://creativecommons.org/licenses/by-sa/2.0/fr/deed.fr>

Vous êtes libres de reproduire, distribuer et communiquer cette création au public selon les conditions suivantes :



Paternité. Vous devez citer le nom de l'auteur original



Pas d'Utilisation Commerciale. Vous n'avez pas le droit d'utiliser cette création à des fins commerciales.



Pas de Modification. Vous n'avez pas le droit de modifier, de transformer ou d'adapter cette création.

Sommaire

Résumé du document (lecture rapide).....	5
1- Introduction.....	9
2- Le libre accès : un acquis récent mais désormais largement partagé	10
3- Rappel des principales étapes récentes du mouvement	10
3.1- L’auto-archivage – définition du libre accès et caractéristiques des archives	11
3.2- Interopérabilité et visibilité des archives.....	14
4- Avantages de l’auto-archivage des publications pour les différents acteurs ...	16
4.1- Pour les auteurs.....	16
4.2- Pour les institutions de recherche.....	18
4.3- Pour les lecteurs potentiels : citoyens – société	18
4.4- Pour les organismes de financement de la recherche.....	19
4.5- L’enrichissement des dépôts de documents : une valeur ajoutée potentielle du Libre accès.....	19
5- Des prises de position insuffisamment suivies dans la pratique	20
5.1- Réticence ou attentisme des auteurs : bons ou mauvais arguments ?	20
5.2- Le positionnement des institutions de recherche : une clé de réussite	22
6- De l’incitation à l’obligation de dépôt	25
6.1- Une politique volontariste au sein des organismes.....	25
6.2- Une forte incitation de la part des organismes financeurs et des institutions politiques	27
6.3- Lien entre dépôt et évaluation	30
6.3.1- L'exemple du Royaume Uni : IRRA (Institutional Repositories and Research Assessment).....	30
6.4- Eléments à prendre en compte pour lier évaluation et dépôts des documents	31
6.5- L’archive institutionnelle : le meilleur des modèles ?	31

6.6- Conclusion : pour une obligation de dépôt.....	32
7- Conclusion générale.....	34
Glossaire.....	35
Bibliographie	35
Annexe 1 - De nouveaux modèles éditoriaux	38
Annexe 2 - Adresses de ressources électroniques.....	41

Résumé du document (lecture rapide)

N.B. : Cliquer sur les titres pour consulter la version complète du document

[Chapitre 1: Introduction](#) (page 9)

[& Chapitre 2 : Le libre accès : un acquis récent mais désormais largement partagé](#) (page 10)

Le mouvement du libre accès aux documents scientifiques a pour but l'accès gratuit et immédiat aux publications scientifiques, par deux voies différentes :

- la création de revues alternatives, substituant au classique modèle « abonné-payeur », un nouveau modèle éditorial « auteur/institution-payeur » (l'auteur ou son institution paie pour publier, mais la diffusion est ensuite gratuite et en libre accès) ;
(voir [annexe 1 : « Nouveaux modèles éditoriaux »](#), p.38)
- l'auto-archivage des publications : les auteurs déposent leurs travaux dans des «dépôts institutionnels», hébergés en local par leur institution de rattachement et/ou dans des archives disciplinaires accessibles gratuitement via Internet.

[Chapitre 3 - Rappel des principales étapes récentes du mouvement](#) (page 10)

Il a été initié, de 1991 à 1997, par des communautés de scientifiques (en physique théorique d'abord, puis en mathématiques et en sciences cognitives) qui ont développé des serveurs pour déposer leurs publications avant parution (preprints), puis après publication dans un journal scientifique (postprints). L'enjeu était la libre circulation des connaissances scientifiques et un certain refus de la politique des éditeurs scientifiques (coût exorbitant des abonnements, cession totale des droits sur les publications aux éditeurs, délai de diffusion des connaissances trop long entre la soumission de l'article et la parution).

Ce mouvement prit de l'ampleur et fut relayé par les organismes financeurs de la recherche, donnant lieu à plusieurs déclarations de scientifiques et d'organismes de différents pays réunis à Budapest (2002), puis à Berlin (2003, 2004 et 2005). Les chercheurs peuvent déposer leurs documents sur des sites institutionnels (repository(ies) en anglais) qui partagent une base technologique commune (le protocole OAI-PMH) permettant de collecter dans des moissonneurs (OAIster) une description des informations bibliographiques (métadonnées).

Ils peuvent aussi soumettre leurs articles dans des revues en libre accès : le dépôt du document peut être payant (l'auteur ou l'institution paye pour publier) mais l'accès du document est gratuit, c'est le modèle « auteur-payeur ».

Les éditeurs scientifiques, dans un premier temps hostiles à ce mouvement du libre accès, acceptent maintenant, pour plus de 90 % d'entre eux (voir le site [Sherpa/Roméo](#)) des dérogations à leurs contrats de cession des droits d'auteur et autorisent les chercheurs à déposer leurs articles sur un serveur institutionnel. Ils adaptent progressivement leurs modèles économiques en proposant soit une option de libre accès, payante au choix, mais aussi depuis très peu de temps en donnant la possibilité aux auteurs de diffuser sur le site des éditeurs la version manuscrite (= preprint) de leur article soumis à la revue, avec possibilité pour la communauté scientifique de commenter en ligne le manuscrit soumis (exemple : « *Atmospheric Chemistry and Physics* » et « *Nature* »).

Chapitre 4 – Avantages de l’auto-archivage pour les différents acteurs (page 16)

Pour les chercheurs

- Une plus grande visibilité des publications ;
- Une accessibilité accrue, et donc un taux de citation bien plus important selon toutes les études mentionnées (gain de 25 à 250 %) ;
- Une dissémination immédiate, et un facteur d’immédiateté plus grand (les articles sont cités plus tôt après leur parution) ;
- L’augmentation du facteur d’impact pour la plupart des titres de revues en libre accès.
- Une pérennité des informations ;
- On note également l’émergence de nouvelles mesures d’impact applicables aux articles eux-mêmes, indépendamment des revues.

Pour les institutions

- Un entrepôt de données institutionnel constitue une vitrine externe de l’organisation scientifique (voire une vitrine pour l’ensemble de la recherche française quand toutes les archives ouvertes des organismes français seront visibles dans la future archive nationale française) ;
- Il permet aussi de suivre la production scientifique interne (production potentielle d’indicateurs) et peut éventuellement être utilisé pour l’évaluation scientifique ;
- Les avantages cités précédemment pour les chercheurs le sont aussi pour les institutions ;
- Le Libre Accès représente enfin un outil essentiel de diffusion universel et direct des connaissances vers l’enseignement et le grand public ainsi que vers les pays en voie de développement ;
- Les potentialités du libre accès sont encore à développer. Ainsi les chercheurs pourraient déposer d’autres types de documents : données expérimentales, corpus de document... et créer ainsi de la valeur ajoutée aux sites institutionnels.

Chapitre 5 – Des prises de position insuffisamment suivies dans la pratique (page 20)

En dépit du positionnement de nombreux organismes en faveur de l’auto-archivage au plan national et international, on constate encore des taux de dépôt assez faibles : environ 460 archives institutionnelles sont recensées, mais elles ne représentent en moyenne que seulement 15 % de la production scientifique de leurs institutions avec des différences importantes selon les disciplines.

Les raisons sont les suivantes :

- Les chercheurs ne connaissent pas ou mal le principe du Libre Accès ;
- Ils sont attachés au modèle Peer-Review des éditeurs, peinent à s’émanciper des canaux traditionnels et évoquent des craintes par rapport au modèle auteur/institution-payeur ;
- Le dépôt dans les archives institutionnelles est trop souvent perçu comme une contrainte administrative et une perte de temps ;
- les chercheurs se soucient peu de l’encadrement juridique du dépôt.

Pour être pleinement opérationnelles, les archives ouvertes institutionnelles doivent s’approcher des 100 % de taux de dépôt.

Différentes enquêtes montrent que les chercheurs sont prêts à auto-archiver si leur organisme les y contraint.

Le positionnement des institutions de recherche est donc un facteur de réussite pour les archives ouvertes et une politique d’obligation de dépôt doit être mise en place dans les institutions, soutenue par un dispositif d’accompagnement des chercheurs (sensibilisation aux enjeux, encadrement juridique...).

Tous les organismes de recherche français disposent ou vont disposer à court terme d’une archive ouverte institutionnelle, avec des approches très différentes.

Certains organismes s'appuient sur des bases de données déjà existantes : c'est le cas du Cemagref qui dispose d'un recensement exhaustif de ses publications (sans le texte intégral) et va s'appuyer sur cet acquis pour démarrer son archive ouverte.

D'autres organismes, comme l'Inserm ou l'INRIA ont choisi d'utiliser HAL, l'outil développé par le CNRS et utilisé pour la plateforme nationale.

Le problème du multi-dépôt institutionnel (cas où un document doit être présent dans plusieurs archives institutionnelles du fait de l'affiliation de ses auteurs), n'apparaît pas très clairement solutionné pour le moment.

Chapitre 6 – De l'incitation à l'obligation de dépôt (page 25)

Au niveau national et international, l'obligation de dépôt est soutenue à deux niveaux :

- Au sein de quelques organismes de recherche qui l'exigent de la part de leurs chercheurs.
Le site ROARMAP (Registry of Open Access Repository Material Archiving Policies) recense 29 déclarations de politique d'auto-archivage mais les initiatives de dépôts obligatoires dans les universités sont actuellement très isolées et peu nombreuses.
Parmi les 6 organismes "pilotes" recensés, on peut citer :
 - l'Université de Technologie du Queensland (Australie) qui atteint 100% de l'auto-archivage de sa production scientifique ;
 - l'Université de Minho (Portugal) qui attribue une incitation financière aux laboratoires déposant leurs publications dans l'archive ouverte institutionnelle ;
 - Electronics & Computer Science Department de Southampton qui a mis en place depuis 2001 une politique d'auto-archivage atteint 50% de taux de dépôt et souhaite rendre le dépôt obligatoire.
- Par les organismes financeurs de la recherche publique
Quatre des huit Research Councils du Royaume Uni (BBSRC, ESRC et MRC) ont réaffirmé leur politique en faveur de l'auto-archivage.

Le Wellcome Trust (1^{er} organisme de financement pour la recherche biomédicale) exige que tout travail publié avec son soutien financier soit déposé dans le site PubMed Central (projet du NIH permettant le libre accès aux publications de la recherche médicale).

La politique volontariste du NIH en faveur de l'auto-archivage a échoué. Seulement 4% des articles sont auto archivés.

La communauté européenne a publié en janvier 2006 un rapport comprenant une dizaine de recommandations dont le libre accès aux travaux financés par la CE. L'Euroscience Workgroup on Science publishing propose d'obliger d'auto-archiver toutes les publications mais de ne mettre en ligne que celles qui ont les droits de diffusion et en respectant les périodes d'embargo des éditeurs.

Au niveau français, un cadre de coopération national se met progressivement en place (protocole d'accord EPST et Conférence des Présidents d'Université signé en juillet 2006).

Les déclarations d'incitations d'organismes de pilotage de la recherche ont toutefois une portée limitée car ils n'ont pas réellement les moyens d'imposer leur volonté aux organismes de recherche. Les organismes financeurs de la recherche doivent également subventionner le libre accès dans les revues scientifiques (en libre accès complète ou au choix) afin de soutenir le système éditorial scientifique.

Les documents déposés dans les archives institutionnelles pourraient être utilisés pour l'évaluation scientifique collective ou individuelle. Dans ce cadre, l'auto-archivage obligatoire des documents, sans exiger leur mise en libre accès, permettrait d'atteindre 100% de la production scientifique et faciliterait le travail de l'évaluation.

Au Royaume Uni, le projet IRRA (Institutional Repositories and Research Assessment) concrétise ce lien entre archive institutionnelle et évaluation. Le dispositif envisagé facilite le travail des scientifiques

qui déposent des dossiers aux instances d'évaluation ainsi que celui des commissions évaluatrices. Il sera utilisé pour l'évaluation scientifique à partir de 2008.

Conclusion générale (page 34)

Les archives institutionnelles permettent de stocker (archiver) l'ensemble de la production scientifique sous une forme numérique pour des fonctions de conservation et de préservation. De nombreux travaux prouvent qu'elles permettent surtout de mieux **valoriser, et d'augmenter la visibilité des** travaux scientifiques.

Mais le taux de dépôt demeure très peu important (de l'ordre de 15%) à moins d'une politique très volontariste des institutions.

L'obligation de dépôt apparaît comme la seule méthode permettant d'atteindre l'exhaustivité dans une archive ouverte institutionnelle, seule garantie d'une bonne représentativité de la production d'un organisme et condition incontournable pour tirer tous les bénéfices attendus du Libre Accès. Cette obligation devrait être liée à l'évaluation : elle garantirait un taux de dépôt important et faciliterait le travail des instances d'évaluation.

Lier l'obligation de dépôt au processus d'évaluation présente deux avantages :

- augmenter le taux de documents déposés dans l'archive ouverte institutionnelle et garantir une mise à jour régulière des données,
- réduire considérablement le travail des évaluateurs et des chercheurs pour l'évaluation.

Annexes

Annexe 1 – De nouveaux modèles éditoriaux (page 38)

Annexe 2 – Adresses de ressources électroniques (page 41)

1- Introduction

[[Lecture rapide](#)]

L'avènement de l'internet a permis très tôt (dès 1990, avant la création du web) la diffusion directe des connaissances scientifiques. Depuis une dizaine d'années, il révolutionne l'accès à la littérature scientifique.

Si les éditeurs ont adapté leur offre d'abonnement aux contenus électroniques, de nouveaux moyens de fournir le libre accès aux publications arrivent à maturité :

- Les archives ouvertes : sites Internet où l'auteur dépose directement ses publications, après l'évaluation par les pairs (post-publications) ou avant (pré-publications). La diffusion est cependant soumise à la modération des responsables de l'archive ouverte ;
- Les revues en libre accès : les éditeurs diffusent gratuitement sur Internet des revues à comité de lecture. Par contre la soumission des articles est en général payante.

Dans le même temps, les organismes de recherche français (CNRS, INSERM,...) s'engagent dans cette voie en signant la déclaration de Berlin sur le libre accès (octobre 2003): http://www.inist.fr/openaccess/article.php3?id_article=38.

De 2005 à 2006 on remarque une évolution importante du nombre d'archives institutionnelles (dépôt de la production scientifique des chercheurs affiliés à un organisme) ou d'archives thématiques (PubMed Central, projet du NIH) et un nombre croissant de prises de position par des instances de pilotage de la recherche, des institutions parlementaires etc...

Mais force est de constater que malgré ces incitations pour le libre accès, les résultats ne sont pas à la hauteur des espérances : le taux de dépôt des documents en libre accès représente souvent moins de 15% du total des publications des institutions (Hajjem et al. 2005b).

Les enjeux sont très importants tant pour les institutions (gestion de leur production scientifique, production d'indicateurs, visibilité et impact scientifique des recherches) que pour les scientifiques (consultation accrue des publications qui entraîne une augmentation du nombre de citations, délais de publication inexistantes alors qu'une revue diffuse un article accepté quelquefois plusieurs mois après son acceptation..).

Quels sont les obstacles ? : Le positionnement des éditeurs commerciaux, des sociétés savantes ? L'inertie des chercheurs et de leur institutions ? Comment favoriser le dépôt des connaissances scientifiques en libre accès ? Ne faut-il pas alors s'orienter vers une obligation de dépôt des documents ?

Ce document en s'appuyant sur les études récentes ou en cours, tente de faire le point sur les initiatives françaises et étrangères pour inciter ou "obliger" au dépôt de documents dans les archives ouvertes institutionnelles.

Ce document ne traite que de la voie d'auto-archivage des publications dans les archives institutionnelles.

En complément, l'annexe 1 "[Nouveaux modèles éditoriaux](#)" présente la mouvance actuelle des modèles éditoriaux : modèle de revue « classique » proposant le libre accès au choix, modèle de revue en accès entièrement libre fonctionnant selon le modèle "auteur-payeur".

2- Le libre accès : un acquis récent mais désormais largement partagé

[[Lecture rapide](#)]

Depuis les années 1990, une campagne internationale menée initialement par quelques chercheurs « activistes » comme Paul Ginsparg (principalement dans le domaine de la physique et des mathématiques) et Stevan Harnad (pour toutes les disciplines) plaide en faveur du libre accès à l'information scientifique et technique (Harnad 2003).

Une première voie est la création de revues alternatives, substituant au classique modèle « abonné-payeur », un nouveau modèle éditorial « auteur/institution-payeur », (l'auteur ou son institution paie pour publier, mais la diffusion est ensuite gratuite et en libre accès).

Une seconde voie est d'inciter les auteurs à déposer leurs travaux dans des « dépôts institutionnels » (repositories) accessibles gratuitement via Internet, hébergés en local par leur institution de rattachement et/ou dans des archives disciplinaires.

Bien que leur mise en œuvre suscite de nombreuses interrogations sur les plans économique, juridique, technique et organisationnel : ces dernières années on note via différentes étapes et des textes fondateurs importants, la montée en charge vers des solutions du libre accès (Open Access) et l'intérêt croissant des universités, du monde de la recherche et des organismes de subventionnement pour ce mouvement.

3- Rappel des principales étapes récentes du mouvement

[[lecture rapide](#)]

Voir sur le site de l'INIST « Le libre accès à l'information scientifique et technique », la page consacrée aux textes de références : http://www.inist.fr/openaccess/rubrique.php3?id_rubrique=11

- 2002 : Budapest Open Access Initiative (BOAI) : les signataires s'engagent à promouvoir le libre accès par l'une ou l'autre des deux démarches complémentaires préconisées
 - BOAI-1 : l'auto-archivage des articles (pré et post-publications) ;
 - BOAI-2 : publication dans les revues en libre accès.
- 2003 :
 - Déclaration de Bethesda pour l'édition en libre accès : définition précise de l'édition en libre accès (BOAI-2), mais engagements individuels plus qu'institutionnels. (participants essentiellement anglosaxons) ;
 - Version française :
http://www.inist.fr/openaccess/imprimersans.php3?id_article=58
 - 1^{ère} Déclaration de Berlin, pour le libre accès à la connaissance en sciences exactes, sciences de la vie, sciences humaines et sociales
 - signataires : CNRS, INSERM (octobre 2003), INRA, INRIA (juillet 2004),
 - « *Les auteurs accordent aux utilisateurs un accès libre, gratuit, irrévocable, et la possibilité de le réutiliser sous réserve de reconnaissance de paternité,*
 - *Le dépôt se fait sur un serveur de type public offrant les garanties : accès ouvert, distribution sans limitation, inter-opérabilité (protocole OAI-PMH), et archivage à long terme.* » ;
- 2004 : Berlin 2 ou comment concrétiser le libre accès à la littérature scientifique ... ;
- 2005 : Berlin 3 ou comment accélérer cette concrétisation : notamment en recommandant aux institutions d'inciter leurs communautés de chercheurs à auto-archiver leurs publications.

- 2006 :
A noter, en France, la signature le 6 juillet 2006 d'un protocole d'accord définissant une approche coordonnée, au niveau national, pour l'archivage ouvert de la production scientifique française. Outre le CNRS, les principaux organismes de recherche EPIC et EPST (INSERM, INRIA, INERIS, IRD, ADEME ...), la conférence des présidents d'université (CPU), la conférence des grandes écoles et l'institut Pasteur ont signé également cet accord (texte du protocole : <http://www.cge.asso.fr/presse/Protocole-d-accord-archivage-ouvert.pdf>).

3.1- L'auto-archivage – définition du libre accès et caractéristiques des archives

♦ **Définition du « Libre Accès »** tirée de l'initiative de Budapest (BOAI) du 14 février 2002
<http://www.soros.org/openaccess/fr/read.shtml>

« La littérature qui devrait être accessible en ligne gratuitement est celle que les savants donnent au monde sans en attendre de rétribution » [...]

« Par "accès libre" à cette littérature, nous entendons sa mise à disposition gratuite sur l'Internet public, permettant à tout un chacun de lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces articles, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale, sans barrière financière, légale ou technique autre que celles indissociables de l'accès et l'utilisation d'Internet. »

Cette définition couvre aussi bien les documents (articles) publiés dans des **revues en libre accès** que les **documents déposés dans une archive ouverte**.

Pour ceux qui souhaiteraient connaître l'évolution du mouvement et des définitions du Libre accès au cours des années : voir la synthèse de Peter Suber « Open Access Overview »
<http://www.earlham.edu/~peters/fos/overview.htm>

Dans la suite de ce document, seule la voie de l'**auto-archivage** est abordée.

Au cours du temps se sont développées parallèlement deux types d'archives ouvertes : disciplinaires et institutionnelles. Le site "Registry of Open Access Repositories" (ROAR) <http://archives.eprints.org/> répertorie les différentes archives ouvertes au plan mondial et fournit les statistiques d'alimentation.

Ce site propose également un filtre qui permet d'identifier les archives existantes sur différents critères (par type, par pays par logiciel utilisé ...).

<http://archives.eprints.org/?country=fr&version=&type=institutional&order=recordcount&submit=Filter>

L'enregistrement n'étant pas obligatoire, la liste n'est pas forcément exhaustive, mais elle compte néanmoins à ce jour plus de 700 archives ou site de revues enregistrées dans OAISTER (<http://www.oaister.org/o/oaister>) et environ 460 si l'on prend en compte uniquement les archives déclarées «archives institutionnelles ou multi-institutionnelles» (Registry of Open Access Repositories (ROAR) : <http://archives.eprints.org/> - chiffres novembre 2006)

♦ Quelques exemples d'archives

Archive thématique internationale

- ArXiv : <http://fr.arxiv.org/>

Créée en 1991 à l'initiative de Paul Ginsparg, elle comprend des pré- et post-publications, dans les domaines : physique, mathématiques, informatique, biologie quantitative et contient plus de 350 000 articles.

Archive inter-institutionnelle française

- HAL (Hyper Article en Ligne) : <http://www.ccsd.cnrs.fr/accueil.php3?lang=fr>

Archive internationale multidisciplinaire orientée « texte intégral » produite par le CNRS depuis 2001, elle devient la plate-forme fédératrice pour d'autres archives institutionnelles : INRIA, INSERM, bientôt Institut Pasteur ...

◆ **Les caractéristiques des archives :**

- **Dépôt par les auteurs**

Le dispositif de dépôt centré sur le chercheur est fortement recommandé. Il engage sa responsabilité sur la qualité du document et reste maître du mode de diffusion de ses résultats.

On ne dépose que des documents qui ont déjà été validés par les comités de lecture de revues ou des colloques scientifiques (les postpublications, « post-prints ») ou des documents qui pourraient être proposés à des comités de lecture de revues ou des colloques scientifiques (les prépublications, « préprints ») .

- **Modération de base**

Une vérification de base non scientifique est généralement assurée : adéquation entre le contenu et la description, vérification et/ou complément des **métadonnées** (voir le paragraphe 3.2- Interopérabilité et visibilité des archives). Une modération scientifique est appliquée dans certaines archives (travail des vérificateurs par disciplines au CNRS).

- **Permanence des dépôts**

Une fois accepté sur le serveur, le retrait du document est fortement découragé (sauf en cas de circonstances exceptionnelles), mais de nouvelles versions peuvent être proposées et les métadonnées accompagnant le document peuvent être corrigées ou complétées.

Le découragement du retrait favorise la responsabilisation des dépôts et l'intérêt scientifique, et garantit la pérennité de l'accès au document.

- **Appartenance institutionnelle du déposant**

Lors d'un dépôt, le déposant, considéré par défaut comme l'auteur, doit indiquer son rattachement à un laboratoire et un organisme.

Pour certaines archives non-institutionnelles (ArXiv par exemple) un dispositif de parrainage entre auteurs (même d'organismes différents) a été mis en place.

- **Respect des droits d'auteurs et droits d'éditeurs**

Le rôle des revues est de publier, mais aussi et surtout de sélectionner, contrôler, corriger et valider les travaux, notamment grâce aux procédures d'évaluation paritaire (peer review). A condition de respecter les droits respectifs des auteurs et des éditeurs, tout chercheur peut confier aux archives ouvertes de son organisme la version acceptée par le comité de lecture d'une revue (dernière version « auteur » validée = post-print).

Archives ouvertes et revues traditionnelles peuvent donc cohabiter et se compléter (voir [annexe 2](#), p. 38).

Le site <http://romeo.eprints.org> (site recommandé pour sa clarté) ainsi que le site Sherpa/Romeo <http://www.sherpa.ac.uk/romeo.php> sont devenus des outils de référence pour suivre la politique des éditeurs scientifiques en ce qui concerne l'autoarchivage.

Extraction au 31/07/06 : Statistiques pour les 155 éditeurs enregistrés à ce jour sur le site

Code Couleur RoMEO	Politique d'archivage	Nombre de revues	%	Exemples d'éditeurs
vert	Autorise l'archivage de post-print (ie dernière version auteur après le peer-review)	6483	69%	<ul style="list-style-type: none"> <input type="checkbox"/> Elsevier (1882 journaux) <input type="checkbox"/> CAB International (17 journaux) <input type="checkbox"/> Cambridge University Press (186 journaux)
vert-pâle	Autorise l'archivage de pre-print (ie version auteur avant le peer-review)	2248	24%	<ul style="list-style-type: none"> <input type="checkbox"/> Blackwell (698 journaux) <input type="checkbox"/> Marcel Dekker (83 journaux) <input type="checkbox"/> Nature Publishing Group (47 journaux) <input type="checkbox"/> Oxford University Press (180 journaux)
gris	Archivage non autorisé	604	7%	<ul style="list-style-type: none"> <input type="checkbox"/> American Chemical Society (35 journaux) <input type="checkbox"/> American Physiological Society (16 journaux) <input type="checkbox"/> Lippincott (287 journaux)

Au total : environ **94%** des revues de cette liste autorisent une forme d'auto-archivage par les auteurs.

Journal Policies - Summary Statistics So Far



[Search List](#) | [List of publishers](#) | [More information](#) | [Summary Statistics](#) | [Corrections](#)

Journal Policies - Summary Statistics So Far

[Click here for more detailed statistics](#)

Current Journal Tally: 94% Green!

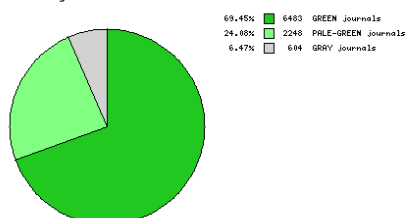
FULL-GREEN = Postprint, PALE-GREEN = Preprint, GRAY = neither yet

Total number of publishers registered at RoMEO to date: 154

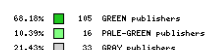
Number of publishers processed so far: 155

Number of journals processed so far: 9339

Journal Policy Chart



Publisher Policy Chart



Très réticents au départ, les éditeurs ont pour beaucoup assoupli leur politique. Le basculement en 2004 de l'éditeur Elsevier dans le camp des éditeurs permettant l'archivage des preprints et postprints est un tournant notable.

- **Règles éthiques et déontologiques**

Les mêmes règles que celles du circuit éditorial classique doivent être appliquées aux documents déposés dans une archive. Franck Laloë, dans une récente intervention (Laloë 2006) rappelle ces « bonnes pratiques »:

- accord de l'ensemble des auteurs et co-auteurs ;
- fourniture de l'ensemble des éléments scientifiques ;
- citation des tiers ;
- pas de plagiat ;
- loyauté vis-à-vis de l'employeur et des entités subventionnant la recherche décrite (pas de divulgation de résultats confidentiels tels que les brevets et mention des soutiens reçus).

- **Garanties techniques / Serveur**

Les archives ouvertes doivent assurer la pérennité des informations que ce soit au niveau de l'accès aux ressources électroniques (archivage) ou à la lecture des formats de documents déposés, dans la majorité des cas, en PDF ou dans un format MSWORD (ce qui implique de conserver dans le temps les logiciels capables de lire ces formats dans la bonne version).

3.2- Interopérabilité et visibilité des archives

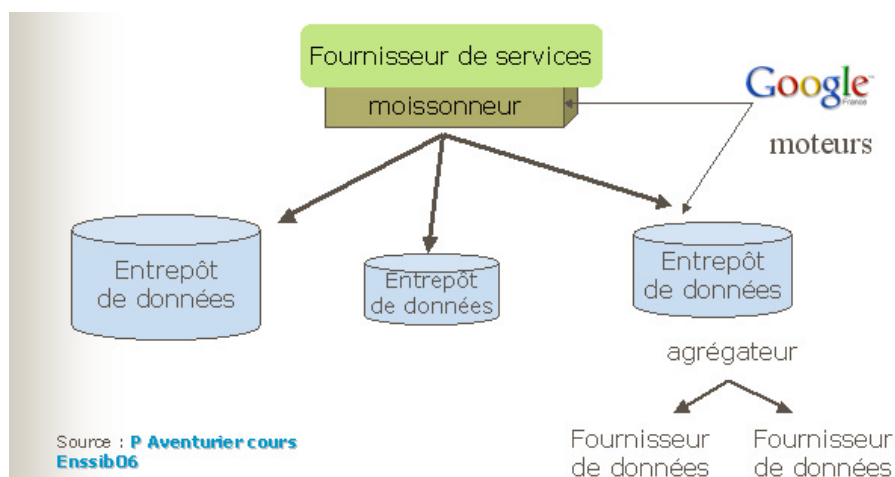
Les entrepôts de données peuvent être constitués à partir de logiciels libres (les plus répandus sont EPrints et DSpace) ou d'outils développés spécifiquement. Leurs caractéristiques communes sont de disposer d'un module de mise à disposition, en utilisant le protocole OAI-PMH (Protocole de Moissonnage des Métadonnées de l'Open Archives Initiative, OAI-PMH) et des métadonnées descriptives des ressources électroniques (auteur, titre...).

Les métadonnées, « ensemble de données structurées décrivant des ressources physiques ou numériques » sont un maillon essentiel pour le partage de l'information et l'interopérabilité des ressources électroniques. Le protocole OAI-PMH permet en effet aux métadonnées des sites d'archives ouvertes d'être récoltées par des moteurs de recherche spécialisés, dits « moissonneurs ». Le jeu de métadonnées géré par ce protocole est le Dublin Core » (Morel-Pair 2005).

Décrire les ressources électroniques avec des métadonnées communes pour chaque institution française par exemple est un fort enjeu pour garantir l'interopérabilité et disposer d'un système commun fiable et performant.

Les métadonnées des sites d'archives ouvertes sont "récoltées" par des moteurs de recherche spécialisés, dits « moissonneurs » qui permettent d'interroger de façon simultanée plusieurs archives comme un seul et unique réservoir de données.

(à noter : à la différence de certains moteurs, les moissonneurs, ou OAI harvesters, traitent uniquement les métadonnées). La figure ci-dessous représente les processus de visibilité des contenus des archives sur Internet.



- Les moissonneurs/agrégateurs :
Le plus connu est OAIster (<http://www.oaister.org/o/oaister>) qui référence 660 entrepôts de données pour plus de 8 millions de documents. L'interface de recherche est cependant limitée aux informations titre, auteur et institutions (créateur), année, langue et type de document.
- Des initiatives commerciales permettent aussi indirectement d'accéder aux contenus des archives ouvertes :
 - SCIRUS : <http://www.scirus.com/srapp/>
Moteur de recherche spécialisé, développé par Elsevier Science, il intègre outre les ressources publiées par le groupe, des documents issus de bases de données, de serveurs de prépublications et d'Open Archives ;
 - SCOPUS : <http://www.scopus.com/>
Egalement développé par Elsevier, c'est une base de données orientée sciences et techniques, qui intègre outre des références issues de bases de données produites par cet éditeur, des références de Pubmed, des références d'articles publiés par Elsevier, et des sites web recensés par le moteur SCIRUS. Seules des ressources validées (« peer reviewed ») sont intégrées ;
 - GOOGLE SCHOLAR : <http://scholar.google.com/> moteur de recherche spécialisé dans la littérature universitaire. La couverture réelle n'est pas bien connue, mais il indexe des articles issus de périodiques avec évaluation par les pairs, des thèses, des livres, des articles en pré-publication, des rapports qui proviennent d'éditeurs scientifiques, de sociétés savantes, de répertoires de préprints, de serveurs universitaires.
- Des technologies associées :
 - OPENURL
« C'est un protocole qui sert à créer des liens contextuels. Concrètement, il s'agit de lier des métadonnées (par exemple, les références bibliographiques d'un article) à la ressource elle-même (l'article en plein-texte).
Pour une ressource (toujours notre article) on connaît le contexte dans lequel elle est citée (la bibliographie qui la cite), décrite (les métadonnées de l'article et qui les a rédigées), utilisée (le lecteur qui la recherche et les droits dont il dispose), et la façon dont on utilise le protocole (quel résolveur, pour obtenir quel service) pour la lier à ce qu'elle décrit (l'article lui-même) » (définition reprise du site : <http://www.figoblog.org/document207.php>).

Ces outils permettent également :

- Des recherches multiformes : sur les métadonnées des articles et/ou en recherche plein texte ;
- La pérennité de l'accès au document : L'URI (Uniform Resource Identifier) vise à attribuer à chaque ressource (et version ou format de ressource) un identifiant unique pérenne (Morel-Pair 2005) ;
- Pour certains outils, l'affichage du nombre de citations et du nombre de téléchargement ;
- Le lien entre les références citées et le texte intégral ;
- L'accès à des documents anciens (avant passage en ligne des revues) dans certains cas de rétronumérisation.

4- Avantages de l'auto-archivage des publications pour les différents acteurs

[[lecture rapide](#)]

4.1- Pour les auteurs

Les chercheurs publient avant tout pour communiquer leurs résultats pour « être lu par les autres, et pour être cité ».

« Les auteurs ne sont jamais payés pour les articles qu'ils produisent. Leur rémunération est la reconnaissance de leurs travaux par leurs pairs et leurs employeurs (pour obtenir des financements, de l'avancement ...) d'où l'importance de la diffusion la plus large possible de ces publications, décisive pour l'impact de leurs recherches» (Schmitt 2006).

Ils ont également besoin pour leur activité scientifique de suivre et collecter en temps réel, la production scientifique de leur domaine, au plan international.

L'auto-archivage apparaît comme une **voie de diffusion maximale des résultats** de la recherche, dont les effets en terme d'**impact** sont désormais mesurables.

Le site <http://opcit.eprints.org/oacitation-biblio.html> fournit une liste bibliographique des différentes études réalisées pour mesurer l'effet de l'auto-archivage sur l'impact des articles.

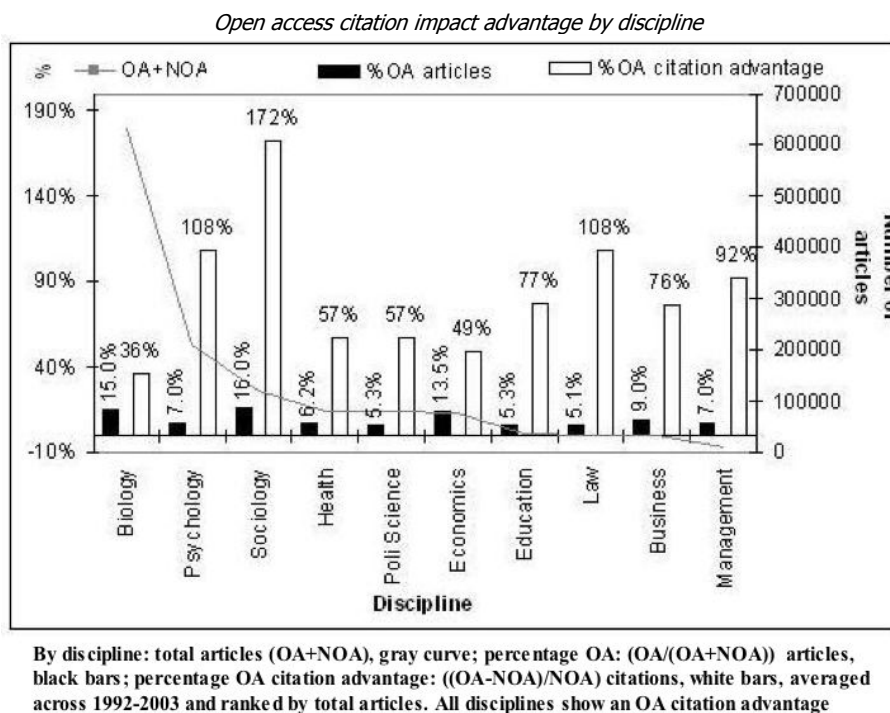
◆ Mesures d'impact - A partir des **taux de citations**

- Depuis l'article de Lawrence (Lawrence 2001) sur la corrélation entre le taux de citation d'un article et sa disponibilité en Libre Accès, de nombreuses études montrent un gain substantiel de citations lorsque les documents sont déposés dans une archive ouverte, avec cependant des différences notables selon les disciplines considérées.

La figure ci dessous présente une récente étude (Harnad and Brody 2004) comparant l'impact (citation) des articles en Open Access (OA) auto-archivés, avec ceux dans le même journal qui ne le sont pas. <http://www.dlib.org/dlib/june04/harnad/06harnad.html>.

- Une étude récente de 2005 (Hajjem et al. 2005a) réalisée sur un échantillon d'environ 1000 journaux (indexés par l'ISI) représentant 4 disciplines (Biologie, Business, Psychologie, et Sociologie) montre des avantages notables en terme de taux de citations si les documents ont été parallèlement déposés dans une archive ouverte. **Le gain serait de 25 à 250 % selon les disciplines.** (<http://eprints.ecs.soton.ac.uk/11687/>)

- Une autre étude des mêmes auteurs (Hajjem et al. 2005b) rapportée par T. Chanier (Chanier 2006) montre des résultats convergents dans six autres disciplines. (<http://sites.computer.org/debull/A05dec/hajjem.pdf>)



◆ Mesures d'impact - Autres types d'évaluation

Des études montrent un lien entre le taux de consultation (téléchargement) et le taux de citation. Dans certains domaines, on peut aussi noter un lien entre la rapidité de la consultation et le taux de citation (Brody et al. 2006; Perneger 2004).

Par ailleurs, de nouveaux modèles émergent pour évaluer l'impact des documents eux-mêmes (et non plus des revues). Adaptés au libre accès, ils permettent le calcul automatique des citations dans l'environnement du libre accès.

- Citebase <http://www.citebase.org>
Une analyse graphique met en évidence le nombre de téléchargements, le temps de latence entre la publication et les premiers téléchargements, et les citations de ses propres citations.
- CiteSeer : <http://citeseer.ist.psu.edu>
Permet de rechercher des documents et des citations dans les domaines scientifique et technique. Sa base de données contient plus de 700 000 documents.
- L'ISI développe depuis peu une nouvelle base de données, le Web Citation Index, qui prend en compte les contenus de certains dépôts d'archives. Intégré au Web of

