



HAL
open science

Réflexions sur quelques nouveaux modèles de communication scientifique

Jessica Schmitz

► **To cite this version:**

Jessica Schmitz. Réflexions sur quelques nouveaux modèles de communication scientifique. domain_shs.info.bibl. 2004. mem_00000180

HAL Id: mem_00000180

https://memsic.ccsd.cnrs.fr/mem_00000180v1

Submitted on 29 Oct 2004

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.

Université Libre de Bruxelles
Faculté de Philosophie et Lettres
Section Infodoc

Diplôme d'études spécialisées (D.E.S.) en sciences et technologies de
l'information

Réflexions sur quelques nouveaux modèles
de communication scientifique

Jessica Schmitz

Sous la direction de Monsieur Christian Brouwer

Année académique 2003-2004

Table des matières

Introduction.....	5
Chapitre 1. La crise des périodiques.....	11
Chapitre 2. Des projets fédérateurs.....	17
1. OAI : Open Archives Initiative.....	19
2. SPARC : The Scholarly Publishing and Academic Resources Coalition....	23
3. NDLTD : Networked Digital Library of Theses and Dissertations.....	28
4. Open Archives Forum.....	32
Chapitre 3. Des nouveaux modèles de communication scientifique.....	36
1. Les archives thématiques.....	36
1.1. ArXiv : e-print archive	36
1.1.1. Historique.....	36
1.1.2. Technique et fonctions.....	37
Fonctionnement général.....	37
Recherche d'information.....	39
1.1.3. Situation actuelle.....	40
1.2. @rchive SIC : Archive Ouverte en Sciences de l'Information et de	
la Communication.....	43
1.2.1. Historique.....	43
1.2.2. Technique et fonctions.....	44
Fonctionnement général.....	44
Recherche d'information.....	46
1.2.3. Situation actuelle.....	48
1.3. RePEc : Research Papers in Economics.....	51

Table des matières

1.3.1. Historique.....	51
1.3.2. Technique et fonctions.....	52
1.3.3. Situation actuelle.....	55
2. Les dépôts institutionnels.....	58
2.1. Caltech CODA.....	58
2.1.1. Historique.....	58
2.1.2. Technique et fonctions.....	60
Fonctionnement général.....	60
La recherche d'information.....	63
2.1.3. Situation actuelle.....	64
2.2. Digital Library of Massachusetts Institute of Technology Theses. .68	
2.2.1. Historique.....	68
2.2.2. Fonctions et technique.....	69
Fonctionnement général.....	69
La recherche d'information.....	71
2.2.3. Situation actuelle.....	72
2.3. e-Prints Soton : University of Southampton e-Prints Service.....	74
2.3.1. Historique.....	74
2.3.2. Technique et fonctions.....	75
Fonctionnement général.....	75
La recherche d'information.....	77
2.3.3. Situation actuelle.....	78
3. Les revues en ligne “Open access”.....	83
3.1. BioMed Central : The Open Access Publisher.....	83
3.1.1. Historique.....	83

Table des matières

3.1.2. Technique et fonction.....	84
Fonctionnement général.....	84
La recherche d'information.....	86
3.1.3. Situation actuelle.....	88
3.2. HighWire.....	91
3.2.1. Historique.....	91
3.2.2. Technique et fonctions.....	92
Fonctionnement général.....	92
La recherche d'information.....	96
3.2.3. Situation actuelle.....	97
3.3. Erudit : Promouvoir et diffuser la recherche universitaire.....	99
3.3.1. Historique.....	99
3.3.2. Technique et fonctions.....	101
Fonctionnement général.....	101
Recherche d'information.....	105
3.2.3. Situation actuelle.....	106
4. Tableau récapitulatif.....	109
Conclusion.....	115
Bibliographie.....	121
Articles et monographies.....	121
Sites Internet.....	128

Introduction

Depuis un certain temps déjà, les prix des périodiques, spécialement en sciences, techniques et médecine, n'ont cessé d'augmenter de manière vertigineuse. On a constaté une augmentation de plus de 470 % entre 1970 et 1995 [Conseil 2002, p.3], et de 186 % entre 1986 et 1997 [Webster 1999]. Le résultat de cette augmentation est l'impossibilité pour les bibliothèques d'assurer pleinement leurs rôles, les budgets ne pouvant pas suivre cette croissance des coûts. Or, les bibliothèques jouent un rôle crucial dans le monde scientifique, principalement au niveau de la diffusion des résultats de la recherche. Il en résulte que tout le processus de communication scientifique est miné par l'attitude et le jeu commercial de quelques grands groupes d'éditeurs internationaux. En Belgique, à l'ULB, le recteur et le directeur des bibliothèques ont tiré la sonnette d'alarme, en décembre 2003, auprès de toute la communauté universitaire par l'envoi d'un mail mettant en lumière une situation que trop de personnes ignoraient jusqu'alors.

Cette situation, couplée à l'avancée des technologies de réseau et à la mobilisation d'une part de la communauté scientifique, a accéléré la mise en place de nouveaux modes de communication. La motivation première des projets nés dans ce cadre n'est pas toujours l'opposition aux éditeurs commerciaux mais plutôt une volonté d'utiliser les technologies de pointe pour améliorer (en temps et en lieu) la diffusion des résultats de la recherche. Ces projets qui rendent l'accès à l'information plus ouvert sont de plus en plus nombreux et, pour une partie d'entre eux, ont acquis une

Introduction

certaine maturité et de la notoriété qui en font des sources d'information très utilisées.

Le propos de ce mémoire est, par l'analyse de trois services issus de trois modèles, de mener une réflexion critique sur ces nouveaux moyens de communication scientifique. Sans pouvoir, ni vouloir prédire l'avenir, la dissection (limitée) des neuf projets nous permettra de faire ressortir des informations sur les origines, le fonctionnement, l'utilisation, les avantages et les inconvénients de ces outils.

Les trois modèles que nous avons choisis d'étudier sont les dépôts thématiques, les dépôts institutionnels et les revues en ligne. Ils proposent tous un accès plus ouvert – c'est-à-dire gratuit ou à un moindre prix – à l'information scientifique. Les archives thématiques ou disciplinaires sont, en général, ouvertes à tous, tant au niveau du dépôt que de la consultation de documents. Les dépôts institutionnels concernent, potentiellement, toutes les disciplines enseignées au sein de l'institution mais n'accueillent que les documents créés en son sein. Enfin, les revues en ligne peuvent concerner tous les domaines. Elles remplissent les mêmes fonctions que les revues sur papier, particulièrement la mise en place et le maintien du processus de *peer review* et se servent des nouvelles technologies pour proposer un accès à moindre prix et enrichi de fonctionnalités propres à la diffusion en ligne.

Chacun de ces trois modèles sera illustré par l'analyse de trois services¹. Ces services ont été choisis parmi de nombreux autres sur base de

1 Dépôt thématique : arXiv, @rchive SIC et RePEc ; dépôt institutionnel : Caltech CODA, MIT These et E-Prints Soton ; revues en ligne : BioMed Central, HighWire

Introduction

leur maturité et de leur succès apparent. Ils servent également à faire la preuve d'une certaine diversité géographique, disciplinaire et fonctionnelle. Il ne s'agit cependant pas d'un échantillon représentatif (les dépôts thématiques ne sont pas répartis en trois tiers concernant chacun la physique, l'économie et les sciences de l'information et de la communication). Ces exemples servent à montrer la diversité possible et la flexibilité des modèles.

Les neuf exemples seront analysés en trois points. L'historique remontera aux origines du projet : depuis quand est-il en place ? qui en sont les initiateurs ? quelles sont les raisons et motivations qui ont poussé à sa réalisation ? Le deuxième point abordera le fonctionnement et la technique du projet. Plus précisément, il sera divisé en deux parties, la première abordera le fonctionnement général du projet (modalité de dépôt, de consultation, navigation dans le site, types de documents hébergés, disciplines abordées, ...) tandis que la seconde s'attachera aux moyens de recherche d'information disponibles. Enfin, l'analyse se clôturera sur la situation actuelle du projet. Cette partie sera rédigée sur base des statistiques disponibles et sur une brève analyse de type SWOT (Strengths, Weaknesses, Opportunities, Threats) .

Les sources pour la réalisation de ce travail se trouvent principalement sur les sites web des différents projets. Mais également dans des périodiques spécialisés qu'ils soient en ligne ou sur support papier, des monographies sur la communication scientifique et de nombreux sites internet directement liés, ou non, aux initiatives envisagées dans le cadre de

et Erudit.

Introduction

ce travail.

L'approche de ces différentes sources ne se fait pas sans précautions. La publication sur internet étant ouverte à un public large, une attention particulière a été portée à la provenance des articles et à leurs auteurs. Les sites universitaires et les articles écrits par des personnes reconnues dans le monde de la bibliothéconomie et de la communication scientifique ont été privilégiés.

Une disparité dans la quantité de documents disponibles concernant les neuf exemple a été constatée. C'est pourquoi certains sont plus développés que d'autres. Pour les mêmes raisons, des éléments différents sont mis en avant selon les exemples.

Enfin, il n'est pas évident, pour un non spécialiste d'aborder un site concernant un domaine pointu et d'en évaluer les différentes caractéristiques. Cette remarque est particulièrement valable pour l'évaluation des possibilités de recherche : comment construire une bonne équation de recherche en astrophysique quand on n'est pas astrophysicien ? Il n'est pas toujours facile dans ces conditions d'évaluer la pertinence des outils disponibles.

La suite de ce travail se présente en trois chapitres. Le premier sert de mise en contexte et aborde la crise des périodiques. Il s'agit d'un problème complexe qui implique de nombreux acteurs et modifie un certain nombre de processus, ceux-ci seront présentés de manière concise et, nous l'espérons, claire.

Le deuxième chapitre présente différentes initiatives fédératrices qui

Introduction

encouragent et guident les projets orientés *Open Access*. Ces fédérateurs sont nombreux mais nous en avons retenus quatre : l'*Open Archives Initiative* (OAI), *The Scholarly Publishing and Academic Resources Coalition* (SPARC), *Networked Digital Library of Theses and Dissertations* (NDLTD) et l'*Open Archives Forum*. Pour chacun de ces fédérateurs, nous envisagerons les objectifs, l'historique et les réalisations.

Enfin, les trois nouveaux modèles seront envisagés successivement dans le troisième chapitre : d'abord les dépôts thématiques, ensuite les dépôts institutionnels et les revues en ligne. Ce troisième chapitre sera clôturé par un tableau récapitulatif accompagné de commentaires dont l'axe principal sera la durabilité des différents projets.

Chapitre 1. La crise des périodiques

Ce premier court chapitre a pour objectif la description de la situation qui a encouragé le développement de nouveaux modes de communication scientifique.

La “crise des périodiques” est causée par l'augmentation impressionnante des prix des périodiques. Ce phénomène est, nous le verrons dans la suite de ce travail, un des facteurs qui est à l'origine de l'apparition de nouveaux modèles de communication scientifique et d'initiatives les encourageant.

A l'origine de cette crise qui affecte l'ensemble de la communication scientifique, se trouve une augmentation sans précédent du prix des abonnements aux périodiques, principalement en sciences, technique et médecine (STM). L'Association of Research Libraries (ARL) indique qu'entre 1986 et 1997, le prix de ces périodiques a augmenté de 169 % pour aboutir à un coût d'abonnement annuel moyen variant de 5 000 \$ à 20 000 \$ [Webster 1999]. En France, entre 1980 et 1999, c'est une augmentation de plus de 500 % qui a été observée pour le coût moyen des abonnements aux revues de médecine et pharmacie, et environ 400 % pour les sciences, 300 % pour le droit, l'économie et la gestion et 200 % pour les lettres et les sciences humaines [Chartron 2004, p. 15]. Ce phénomène est encore accru par le désabonnement, réaction première des bibliothèques face aux prix trop élevés, que les éditeurs compensent en augmentant leurs tarifs [Chartron 2000].

Les causes de l'augmentation première sont diverses. Duane Webster, dans sa participation à la 65^{ème} conférence de l'International Federation of

Chapitre 1. La crise des périodiques

Library Associations and Institutions (IFLA) en août 1999 pointe du doigt quatre causes principales à la crise des périodiques STM : le déséquilibre et l'imperfection² du marché, la place croissante des éditeurs commerciaux, le renforcement des éditeurs STM et la prolifération des titres [Webster 1999].

Si les éditeurs commerciaux se sont trouvés impliqués dans la communication scientifique c'est en partie parce la publication en STM est complexe, parce qu'ils ont su proposer des produits pratiques, une distribution dans les temps et de la crédibilité aux universités [Webster 1999]. Mais aussi parce que la crise financière des années 1980 et l'idéologie de privatisation ont poussé les acteurs tels que universités et sociétés savantes à céder leurs revues aux éditeurs commerciaux privés [Dassetto 2004, p. 14]. Ces éditeurs se sont concentrés, au fil de rachats, en quelques grands groupe financiers internationaux comme Reed Elsevier, Wolters Kluwer, Wiley, soumis aux investissements de leurs actionnaires et aux lois du marché [Chartron 2004, p.15]. Ces grands groupes sont tout à fait à même d'augmenter les prix comme ils l'entendent [Roucoux 2004a, p. 75]. Il appartient ensuite aux bibliothèques de faire ce qu'elles peuvent avec les budgets qu'elles ont. Le désabonnement est en général la solution qui s'impose.

La question vient alors de savoir quelles revues il faut garder. Quelles sont celles dont les chercheurs, les étudiants et les professeurs ne peuvent se passer ? L'*Institute for Scientific Information* (ISI) a mis au point un outil

2 Imparfait parce que caractérisé par deux logiques opposées, d'un côté les éditeurs commerciaux imposent leurs prix, de l'autre, les scientifiques ne vendent pas leurs résultats, ils les mettent à disposition contre une reconnaissance du travail accompli [Webster 1999].

Chapitre 1. La crise des périodiques

pour répondre à cette question. A partir du *Science Citation Index*, on calcule l'*impact factor* ou facteur d'impact des revues c'est-à-dire la fréquence avec laquelle l'article "moyen" d'une revue donnée a été cité durant une année donnée dans les articles dépouillés [Dassetto 2004, pp. 13-14]. Les revues à facteur d'impact élevé sont les revues les plus indispensables, celles dans lesquelles l'information de pointe est contenue et celles, donc, qui sont les plus nécessaires au public des bibliothèques. L'*impact factor* permet la définition d'un groupe de périodiques desquels une bibliothèque ne peut se passer, sous peine de priver son public d'informations vitales : les *core journals*.

Ces revues sont aussi, par la force des choses, les plus prestigieuses, celles dans lesquelles il est bon d'être publié car l'évaluation du travail d'un chercheur est basée sur les répercussions de sa recherche ; un bon chercheur est un chercheur cité. Ici, également, nous nous trouvons face à un cercle vicieux : les articles publiés dans les *core journals* sont les plus lus, donc les plus cités. Ce qui ne fait qu'augmenter l'*impact factor* du périodique qui, dès lors, est plus lu. De plus, l'*impact factor* introduit des changements dans la manière de publier des scientifiques. Les articles vont être mentionnés avec plus d'auteurs qu'il n'en ont réellement et ces auteurs ne vont pas hésiter à s'auto-citer pour faire croire à un impact plus grand de leur article.

Les périodiques ont, comme rôle principal (particulièrement en STM) celui d'assurer la validité des résultats et la qualité de la recherche via le *peer review*, ils permettent également d'établir la propriété intellectuelle d'une découverte [Roucoux 2004a, p. 76].

Le *peer review* ou arbitrage par les pairs est l'épreuve obligatoire par

Chapitre 1. La crise des périodiques

laquelle passe un article avant d'être publié dans une grande revue. Une fois proposé, l'article est évalué par ce comité de scientifiques spécialisés dans le domaine de la revue. Celui-ci va accepter, refuser ou suggérer des modifications à apporter à l'article avant de le publier. Ce processus est le garant de la qualité des articles et donc du périodique. Il va influencer plus d'une chose, parmi lesquelles on peut citer la carrière des auteurs des articles et le financement de la recherche. Pourtant, les quelques études menées sur ce processus, rarement remis en cause par ceux qui y participent, ne prouvent pas son efficacité. De plus, une idée nouvelle exprimée dans un article pourrait très bien ne pas être publiée car remettant en cause les connaissances et idées jusque là considérées comme "justes" et dont les membres du comité scientifique peuvent se considérer comme les gardiens. Dans cette optique, l'arbitrage par les pairs apparaît comme un frein possible à l'innovation et à la diffusion d'idées nouvelles, tout en restant une garantie de qualité des articles [Roucoux 2004a, pp. 76-78].

En réponse à cette situation, l'*Open Access* prône l'accès gratuit pour tous à l'information scientifique, présentée sous format numérique et accessible via le réseau internet [<http://www.earlham.edu/~peters/fos/overview.htm>]. Cette proposition de solution à la crise qui mine l'information scientifique peut prendre différentes formes et poursuivre différents objectifs. Elle est encouragée par des projets fédérateurs qui encourage la diffusion et l'amélioration de la diffusion de l'information via ces nouveaux modèles de communication scientifique. Ces deux points forment la suite et la partie principale de ce travail.

Chapitre 1. La crise des périodiques

Chapitre 2. Des projets fédérateurs

Nous appelons projet fédérateur les initiatives regroupant différents projets d'archives ouvertes sur base de critères définis mais qui peuvent être très englobant. Ces projets peuvent encourager la création de dépôts en ligne, envisager des mesures pour améliorer leur interopérabilité, mener des actions d'information, proposer des répertoires,... bref, autant de bonnes raisons pour se proposer comme point d'accès (ou de recherche) à différents dépôts ou revues en ligne d'accès plus ou moins ouvert.

Un intérêt plus important sera porté sur l'élaboration de méta-données qui rendent interopérables les différents dépôts. Ces méta-données permettent la création de services ajoutés dont un des plus importants est sans conteste la recherche d'information. Ceci permet au chercheur d'effectuer une seule fois sa requête pour obtenir les documents provenant de plusieurs dépôts. Les dépôts institutionnels ont également un grand intérêt pour ce type de services. Etant le plus souvent multidisciplinaires, la recherche transversale leurs permet de faire connaître leur production à un public plus large. En effet, le chercheur, s'il doit soumettre sa requête à tous les dépôts institutionnels préférera se tourner vers un dépôt thématique tel que arXiv s'il est physicien ou RePEc s'il est économiste.

Nous abordons, dans la suite de ce chapitre, trois projets fédérateurs connus et souvent rencontrés dans l'analyse des différents modèles. Il s'agit de l'*Open Archives Initiatives* (OAI), la *Scholarly Publishing and Academic Resources Coalition* (SPARC) et la *Networked Digital Library of Theses and Dissertations* (NDLTD). Il existe des projets plus spécifiques et d'ampleur plus restreinte, l'*Open Archives Forum*, quatrième et dernier fédérateur

Chapitre 2. Des projets fédérateurs

analysé, en est un bon exemple. Chacun des projets présentés sera entrevu sous différents aspects : ses objectifs, son historique et ses réalisations, qu'elles soient techniques ou autres.

Citons également le *Directory of Open Access Journals* (DOAJ) qui concentre en un point, l'accès à 1.219 périodiques d'accès ouvert. Pour 319 d'entre eux, il est possible d'en rechercher les articles [<http://www.doaj.org>]. L'*Initiative de Budapest pour l'Accès Ouvert* (BOAI) est également une preuve de la mobilisation autour de l'accès ouvert. Parrainée par l'*Open Society Institute* (financée par George Soros), l'initiative de Budapest est à la fois une déclaration de principe, de stratégie et de soutien financier ayant pour but d'appuyer l'accès libre à l'information scientifique [<http://www.soros.org/openaccess/fr/index.shtml>]. Enfin, citons encore le *Washington DC Principles for free access to science* qui encourage les éditeurs *not-for-profit* à travailler avec le monde scientifique et universitaire. Cette initiative est liée à HighWire qui est cité comme exemple et qui abrite la plupart des signataires des principes de Washington [<http://www.dcprinciples.org/> ; <http://www.dcprinciples.org/statement.pdf>].

1. OAI : Open Archives Initiative

<<http://www.openarchives.org>>

L'Open Archives Initiative a pour objectif le développement et la promotion de standards d'interopérabilité visant à faciliter la diffusion de contenus. Au départ, l'OAI trouve ses origines dans la volonté d'améliorer l'accès aux dépôts d'e-prints pour augmenter la disponibilité des textes issus de la communication scientifique [<http://www.openarchives.org/documents/FAQ.html>].

L'augmentation des dépôts et bibliothèques numériques a eu pour conséquence que les chercheurs devaient consulter séparément les différentes sources d'information et effectuer plusieurs fois leurs requêtes. Une certaine uniformisation des différentes initiatives et l'augmentation de leur interconnectivité étaient indispensables pour la réalisation de recherches transversales. D'autant plus que ces dépôts utilisent des bases de données qui ne sont pas accessibles aux moteurs de recherche généraux ; elles font partie de ce qui est appelé le *deep web* ou web caché [Suleman 2001, p. 126].

L'OAI et les résultats auxquels elle a abouti sont basés sur les travaux d'Herbert Van de Sompel, des chercheurs et des bibliothécaires du laboratoire national de Los Alamos avec qui il a travaillé. En 1999, lors de la rencontre de Santa Fe, les problèmes générés par l'augmentation des dépôts d'e-prints ont été abordés. Les conclusions tirées de cette rencontre insistent sur la nécessité de développer des services pour permettre la recherche à travers les multiples dépôts, ainsi que sur l'intérêt de la copie de documents d'un dépôt vers un autre. Avec la coexistence des dépôts

Chapitre 2. Des projets fédérateurs

thématiques et institutionnels, un texte peut être disponible dans l'un et l'autre à condition que son auteur le dépose deux fois (ou plus). Si la copie peut se faire automatiquement, un seul dépôt suffit. Le texte intitulé *Santa Fe Convention* reprend le travail effectué. Au coeur de celui-ci se trouve la définition d'une interface permettant aux serveurs d'e-prints de rendre public les métadonnées des documents qu'ils abritent. Ces métadonnées peuvent ensuite être utilisées par les services de recherche fédérateurs ou par d'autres dépôts désirant identifier les documents pour les copier ensuite [Lynch 2001].

Tout au long de l'année 2000, la réflexion sur l'OAI et le travail sur l'interface de Santa Fe se sont poursuivies en prenant en compte d'autres acteurs confrontés aux mêmes problématiques : les bibliothèques, musées, éditeurs commerciaux de revues, membres de la communauté scientifique. L'OAI est, durant cette période, financée par la *Coalition for Networked Information* et la *Digital Library Federation* et accueillie à l'université de Cornell. En janvier 2001, les spécifications établies lors de la convention de Santa Fe et revues en 2000 lors de différents ateliers sont rendues publiques. Elles forment ce qui est maintenant appelé l'*Open Archives Metadata Harvesting Protocol* ou OAI-PMH pour *Open Archives Initiative Protocol for Metadata Harvesting* [Lynch 2001].

L'OAI-PMH fonctionne comme une interface entre producteurs de données et producteurs de services. Le producteur de données maintient un (ou plusieurs) serveur supportant le protocole comme moyen d'expression des méta-données. Le producteur de service utilise ces méta-données comme base d'un service (par exemple : la recherche transversale

Chapitre 2. Des projets fédérateurs

d'information). Ces deux rôles ne sont pas mutuellement exclusifs, un producteur de données peut aussi offrir des services, et inversement [<http://www.openarchives.org/documents/FAQ.html>].

Le plus petit point commun de ces serveurs est l'utilisation des métadonnées telles que définies par le Dublin Core dans sa version “*unqualified*” ou simple³. Les responsables de dépôt sont libres de fournir des métadonnées plus complètes, elles ne seront cependant peut-être pas utilisées par les services basés sur l'OAI-PMH. La “communauté e-prints” travaille d'ailleurs sur la définition de méta-données plus adaptées à la description de ce type de document.

Le protocole ne gère donc pas les accès aux documents. Si beaucoup d'utilisateurs du OAI-PMH manifestent pour un accès ouvert et gratuit à la documentation qu'ils hébergent, cet engagement ne participe pas à l'OAI. D'ailleurs, l'utilisation du protocole est ouverte à tous et des éditeurs commerciaux peuvent très bien l'utiliser [Lynch 2001].

Il est difficile, voire impossible, de donner une liste complète des fournisseurs de données utilisant le protocole de l'OAI. Ceux-ci n'étant pas obligés de se faire connaître. Jusqu'à présent, 178 dépôts de tout type sont enregistrés sur le site de l'OAI [<http://www.openarchives.org/Register/BrowseSite.pl>]. Un des services

³ Le Dublin Core Metadata Initiative est une organisation dédiée à la promotion et l'adoption de standards de métadonnées interopérables, ainsi qu'au développement d'un vocabulaire de méta-données pour décrire les ressources qui rendent possibles des systèmes plus performants de recherche d'information [<http://dublincore.org/about/>]. La version “*unqualified*” ou simple du Dublin Core est limitée aux 15 éléments principaux du Dublin Core Metadata Element Set [<http://dublincore.org/resources/faq/#whatisthedifference>].

Chapitre 2. Des projets fédérateurs

qui utilise les méta-données issues du protocole, oaister, permet la recherche à travers les dépôts de 327 institutions [<http://oaister.umdl.umich.edu/o/oaister/>]. Les fournisseurs de services sont beaucoup moins nombreux, nous en trouvons 17 sur le site de l'Open Archives Initiative [<http://www.openarchives.org/service/listproviders.html>].

Le succès rencontré par l'OAI-PMH est sans doute dû au fait qu'il peut être utilisé par tous les producteurs d'information qui le désirent. Il est ainsi utilisé aussi bien par des éditeurs commerciaux comme BioMed Central que par des institutions universitaires (California Institute of Technology, Australian National University, University of Southampton,...). Sa simplicité, voulue par ses créateurs, est également facteur de succès ; il peut être d'application pour toutes sortes de documents : e-prints, livres, contribution à une conférence,... Ces deux avantages et le côté pionnier de l'OAI font de celle-ci le fédérateur sans doute le plus connu dans le domaine des fournisseurs de documentation en ligne orienté *Open Access*.

2. SPARC : The Scholarly Publishing and Academic Resources Coalition

<<http://www.arl.org/sparc>>

Deuxième fédérateur présenté dans ce travail, SPARC est une association qui regroupe des universités, des bibliothèques de recherche ainsi que d'autres organisations et qui a été créée pour répondre de manière constructive aux dysfonctionnements du marché de la communication scientifique. La stratégie adoptée pour ce faire est d'augmenter la compétition et supporter les projets *Open Access* pour ainsi contrer les prix élevés et croissants pratiqués par les éditeurs commerciaux de revues scientifiques. Ce problème étant plus aigu en sciences, techniques et médecine, SPARC se concentre jusqu'à présent sur ces domaines [<http://www.arl.org/sparc>].

L'origine de l'association se trouve dans le milieu des bibliothèques qui, à cause de l'augmentation des prix des périodiques ne sont plus aptes à assurer leurs missions. SPARC est un projet qui cherche une réponse pratique à ce problème de la "crise des périodiques" [Case 2002, pp. 2-3].

En octobre 1997, l'*Association of Research Libraries* (ARL) lance SPARC. L'initiative doit servir de catalyseur pour créer un changement qui rendrait le marché de la communication scientifique plus compétitif et dont la conséquence première devrait être la diminution des prix des périodiques. Le premier partenariat est conclu en juin 1998 avec l'*American Chemical Society* (ACS). L'objectif est de publier un nouveau journal tout les ans pendant trois ans. Le premier, *Organic Chemistry* est lancé pour concurrencer *Tetrahedron Letters* édité par Elsevier. La différence de prix

Chapitre 2. Des projets fédérateurs

d'abonnement est flagrante : le périodique de l'ACS est à 2 300 \$ tandis que celui d'Elsevier coûte 8 000 \$ [Case 2002, pp. 8-10].

SPARC se fait peu à peu connaître et de nombreuses institutions s'y joignent. Leurs frais de participation permettent de soutenir financièrement les projets lancés par SPARC.

En 2001, face à l'intérêt de certaines institutions européennes (principalement des universités et des bibliothèques), SPARC Europe est créé. L'objectif est le même que SPARC mais est mené dans une perspective européenne [Case 2002, p. 14 ; <http://www.sparceurope.org>].

Pour atteindre ses objectifs, SPARC a créé trois programmes mettant en oeuvre des moyens d'action différents : SPARC Alternatives, SPARC Leading Edge et SPARC Scientific Communities.

SPARC Alternatives soutient la création de périodiques en compétition directe avec ceux des éditeurs commerciaux classiques. Il a entre autre permis la création d'*Evolutionary Ecology Research* (abonnement : 305 \$) par Michaël Rosenzweig pour concurrencer *Evolutionary Ecology* (800 \$) édité par Wolters Kluwer et également créé par M. Rosenzweig. Ce dernier, accompagné par l'ensemble du bureau éditorial, a quitté Wolters Kluwer pour créer son propre périodique avec le soutien de SPARC. On peut également citer *PhysChemComm* concurrençant *Chemical Physics Letters* d'Elsevier, *Theory and Practice of Logic Programming* concurrençant le *Journal of Logic and Algebraic Programming* d'Elsevier également, et le *Journal of Machine Learning Research* s'opposant à *Machine Learning* édité par Kluwer [Case 2002, pp. 15-16]. Les intitulés des périodiques soutenus par SPARC sont souvent très proches de ceux auxquels

Chapitre 2. Des projets fédérateurs

ils s'opposent", ceci doit certainement entraîner une certaine confusion dans l'esprit des lecteurs et il n'est pas certain que SPARC en bénéficie. Ajoutons un autre bémol, SPARC encourage la création de nouveaux périodiques mais il n'est pas assuré que les bibliothèques vont s'y abonner. Il n'est pas forcément facile pour elles de se désabonner d'un périodique édité par Elsevier, Wolter Kluwer ou un autre. D'une part, elles risquent d'être confrontées à la pression des chercheurs habitués à travailler avec ces coûteux périodiques. D'autre part, l'abonnement peut appartenir à un ensemble négocié et il est dès lors difficile de se séparer d'une partie seulement de cet ensemble. Ajoutons que les restrictions économiques et budgétaires ne permettent vraisemblablement pas aux bibliothèques de s'abonner aux périodiques de SPARC (même s'ils sont moins coûteux) tout en conservant (parce qu'elles y sont obligées) l'abonnement envers les éditeurs commerciaux.

SPARC Leading Edge a pour mission le support de nouveaux modèles d'édition scientifique, il fait la publicité des entreprises qui utilisent les nouvelles technologies pour obtenir des avantages compétitifs ou qui introduisent des nouveaux modèles commerciaux. Parmi ces entreprises, nous pouvons citer le *New Journal of Physics* qui met les articles à disposition gratuitement sur le web et qui, pour se financer, fait payer la publication des articles par leurs auteurs (tout comme BioMed Central). L'*Internet Journal of Chemistry* quant à lui, utilise la 3D et des techniques d'animation pour enrichir son contenu et le rendre plus interactif. Il permet par exemple de montrer la composition d'une molécule en trois dimensions apportant ainsi un véritable plus par rapport à un périodique publié sur

Chapitre 2. Des projets fédérateurs

papier.

Enfin, SPARC Scientific Communities supporte l'aggrégation de contenus autour des besoins spécifiques d'une communauté d'intérêt. Ce programme encourage également la collaboration entre les scientifiques, leurs sociétés et leurs institutions. Il encourage aussi les universités à développer leur propre infrastructure de publication électronique. Les projets soutenus dans ce cadre sont, entre autres, eScholarship de la California Digital Library, Earthscape de Columbia et CogNet du MIT [Case 2002, pp. 17-18]. eScholarship de l'université de Californie a pour objectif le développement d'un modèle de communication scientifique financièrement viable. Et ce, en utilisant les nouvelles technologies. De manière plus générale, il s'agit d'améliorer les différents domaines ayant part à la communication scientifique : création, *peer review*, gestion, diffusion et préservation de ces documents. Sont pris en compte pour cela, le dépôt individuel des résultats de la recherche, la publication sur le web et l'édition électronique de monographies [<http://www.cdlib.org/programs/escholarship.html>]. Le Columbia Earthscape développé au sein de l'université de Columbia est une aggrégation complète de ressources concernant les sciences de la terre et de l'environnement [<http://www.earthscape.org>]. Enfin, CogNet du Massachusetts Institute of Technology se présente comme une communauté électronique de chercheurs en sciences cognitives et cérébrales qui contient des textes importants ainsi que d'autres types d'informations susceptibles d'intéresser les chercheurs (offres d'emplois, annonces de conférences, groupes de discussions, ...) [Case 2002, p. 18 ; <http://cognet.mit.edu>].

Chapitre 2. Des projets fédérateurs

SPARC est une initiative qui provient du monde des bibliothèques scientifiques (de l'Association of Research Libraries) qui, face à l'augmentation des prix des périodiques, se sont vues obligées de réagir. Ses trois programmes s'inscrivent dans une perspective plus générale qui vise à apporter de réels changements dans le monde de la communication scientifique. Pourtant, l'efficacité de SPARC Alternative peut être mise en question car il n'est pas certain que les bibliothèques auront les moyens de s'abonner à ces nouveaux périodiques. Et si, cet aspect de SPARC est sans doute le plus connu et le plus visible, les deux autres ne doivent pas être minimisés.

3. NDLTD : Networked Digital Library of Theses and Dissertations

<<http://www.ndltd.org>>

NDLTD est une initiative entièrement centrée sur le monde universitaire puisqu'elle concerne principalement les travaux de fin d'études. Son objectif est de promouvoir l'adoption, la création, l'utilisation, la diffusion et la préservation de thèses et mémoires électroniques (ETD, pour *Electronic Theses and Dissertations*) [<http://www.ndltd.org>].

Le concept d'ETD est abordé dès 1987 lors d'un meeting orchestré par l'University Microfilms International (UMI). Comme suite à cet événement, Virginia Tech (Virginia Polytechnic Institute and State University) prend en charge le développement d'une DTD (Document Type Definition) SGML (Standard Generalized Markup Language) applicable aux thèses et mémoires. Mais l'action de Virginia Tech ne s'arrête pas là et continue d'envisager la problématique de la publication électronique des travaux de fin d'étude, notamment les aspects de production, d'archivage et d'accès. En 1996, le SURA (Southeastern Universities Research Association) finance la continuation de ce projet basé à Virginia Tech. Ses objectifs concernent aussi bien les étudiants que les chercheurs. Pour les étudiants, il s'agit de leur faire connaître les produits de la publication et des bibliothèques numériques de manière à ce qu'ils puissent appliquer cette connaissance dans leur recherche et dans la création de leur propre thèse ou mémoire électronique. Les universités sont également concernées par l'apprentissage des bibliothèques numériques et de l'intérêt que cela peut leur apporter, notamment dans la diffusion de leur production et de leur potentiel

Chapitre 2. Des projets fédérateurs

intellectuel. A un niveau supérieur, le projet a pour objectif l'amélioration de l'enseignement via un plus grand partage des connaissances. Partage qui pourrait être accéléré par le fait que les résultats de la recherche seront plus rapidement et plus complètement disponibles (si les chercheurs utilisent la publication électronique de manière appropriée) [<http://ndltd.org/info/description.en.html>].

Dès 1996, d'autres universités vont rejoindre le projet NDLTD qui se donne quelques nouveaux objectifs, notamment en matière de préservation et de prix des documents. De nombreux projets de thèses et mémoires numériques ont vu le jour dans les universités du monde entier et les accès à ces dépôts ne font, en général, qu'augmenter [Suleman 2001b].

En plus de mener des actions pour la création de dépôts numériques, NDLTD est actif dans le domaine de la standardisation depuis 2001. Il se concentre sur le développement de métadonnées standardisées pour rendre interopérables les ETD. Ceci a permis la création et l'utilisation d'un catalogue commun à tous les projets membres de NDLTD [<http://www.ndltd.org/standards/index.en.html>]. Pour rendre ce catalogue possible, NDLTD s'est basé sur le travail effectué dans le cadre de l'Open Archives Initiative, en ce compris l'utilisation du "*Protocole for Metadata Harvesting*" et a défini un format de métadonnées spécifique aux thèses et mémoires [Suleman 2001c].

L'utilisation du protocole par les dépôts attachés à NDLTD rend possible la recherche d'information transversale. Quatre moyens de l'effectuer sont disponibles sur le site [<http://www.ndltd.org/browse.en.html>]. Le premier est la recherche sur

Chapitre 2. Des projets fédérateurs

les collections de l'OCLC (Online Computer Library Center). Le second ne concerne qu'une petite partie des collections, seulement 14 sites d'ETD sont concernés ; le fonds documentaire, ici, correspond à un ancien état du projet. Le troisième moyen permet une recherche sur l'ensemble du catalogue via une interface développée par la société VTLS (Visionary Technology in Library Solutions). Enfin, la quatrième technique est rendue possible grâce à un outil de l'OCLC. Tout comme la recherche avec VTLS, elle permet d'agir sur l'ensemble du catalogue du projet mais de manière plus élaborée, elle met en oeuvre des possibilités de recherche plus avancées et plus fines.

Actuellement, NDLTD regroupe 216 membres parmi lesquels 189 sont des universités. Le reste est composé d'institutions variées telles que bibliothèques, centre de recherche ou initiatives opérant dans le domaine de la publication en ligne. Parmi les membres universitaires, nous retrouvons la Bibliothèque Interuniversitaire de la Communauté française de Belgique (BICfB) qui regroupe les universités belges francophones.

Les objectifs proposés par NDLTD sont sans doute très appréciés par les universités détentrices d'un dépôt institutionnel. La création d'un catalogue unique permet aux chercheurs de consulter des documents provenant d'universités dans les dépôts desquelles ils ne seraient pas allés chercher. Les universités moins "réputées" bénéficient donc, à travers NDLTD, d'un moyen de faire connaître leurs travaux à un public large. Sans un moyen de recherche transversale, il y a fort à parier que certains dépôts n'auraient jamais été consultés par des chercheurs n'appartenant pas à la région dans laquelle se trouve l'institution. Certains services liés à l'Open

Chapitre 2. Des projets fédérateurs

Archives Initiative et utilisant son protocole peuvent également offrir cette possibilité de recherche transversale mais NDLTD a la particularité de ne s'appliquer qu'aux thèses et mémoires, offrant ainsi des possibilités nouvelles aux universités.

NDLTD se présente donc comme un fédérateur très spécifique mais qui, grâce à son public, joue un rôle très important, les universités font, en effet, partie des acteurs les plus concernés par la communication scientifique.

4. Open Archives Forum

<<http://www.oaforum.org>>

Dernier fédérateur présenté, l'Open Archives Forum est un exemple d'une action plus régionale, limitée à l'Europe.

L'Open Archives Forum a pour objectif de donner un point de vue européen sur la diffusion de l'information et sur les initiatives européennes d'archives ouvertes. Il regroupe les projets et initiatives nationales ainsi que toutes autres institutions intéressées par les archives ouvertes [<http://www.oaforum.org>].

Le projet a débuté en octobre 2001 pour une durée de 24 mois, il est donc à présent terminé. Ses initiateurs et gestionnaires proviennent d'Allemagne, d'Italie et de Grande-Bretagne ; il s'agit du *Computing Centre* de l'université Humboldt à Berlin, de l'*Istituto per la Scienza e la Tecnologia dell'Informazione – Consiglio Nazionale delle Ricerche* et du *UK Office for Library and Information Networking (UKOLN)* hébergé par l'université de Bath [<http://www.oaforum.org/documents/d54projpres.php>].

Pour la réalisation des objectifs, l'OAF a organisé plusieurs ateliers (*workshops*). Le premier a eu lieu en mai 2002 et les suivants en décembre 2002, mars et septembre 2003 (soit pour clôturer le projet) [<http://www.oaforum.org/documents/>]. Même si l'OAF n'était prévu que pour une durée de 24 mois, tous les documents et apports à la communauté des archives ouvertes sont encore disponibles sur le site.

En plus de la tenue et l'organisation d'ateliers tout au long du projet, l'OAF a réalisé un site internet [<http://www.oaforum.org>] regroupant toutes les informations susceptibles d'intéresser les réalisateurs

Chapitre 2. Des projets fédérateurs

d'une archive en ligne : manuel pour l'implémentation du protocole de l'OAI (OAI-PMH), bibliographie sur l'OAI et autres sujets y afférant. Ajoutons à cela que le site est également un moyen d'accéder à la base de données construite par l'OAF. Celle-ci met à disposition des informations sur les projets et produits en relation avec l'OAI. Ainsi, elle liste les services, dépôts, projets, logiciels, organisations, protocoles et schémas de métadonnées. Il est de plus possible d'effectuer une recherche simple par mots-clés sur le contenu de cette base. Elle n'est pas limitée à ce qui se fait en Europe mais seuls les services, dépôts ou projets européens sont invités à s'y inscrire et ce, bien que le projet soit à présent terminé.

Centré sur l'Europe, l'Open Archives Forum a permis de regrouper des informations utiles, auparavant éparées, pour toutes personnes impliquées dans la gestion d'une archive en ligne. Avec l'organisation d'atelier, il a également rendu possible l'échange d'informations et la création de contacts entre les différents acteurs. Arrivé à son terme en 2003, il faut espérer que la base de données et les informations disponibles seront quand même mises à jour car elles présentent un intérêt certain pour le développement des actuels et futurs dépôts. Moins qu'un fédérateur de projet, l'Open Archives Forum est un fédérateur d'informations variées susceptibles d'intéresser un public large de professionnels.

Ces projets fédérateurs sont d'une ampleur variée et ne concernent pas l'entièreté des initiatives *Open Access*. S'ils offrent un soutien, des lignes de conduite ou des déclarations d'intentions, il n'existe aucune obligation de les suivre. Mais, dans le cas du protocole de l'*Open Archives Initiative*, le mettre en place est certainement un plus. Le chapitre suivant va nous

Chapitre 2. Des projets fédérateurs

amener à considérer de près trois de ces nouveaux modèles. Et pour chacun, nous analyserons trois exemples. Nous verrons quelle place ils réservent aux projets fédérateurs dans l'aperçu que nous donnerons de leurs historiques, leur fonctionnement et la situation dans laquelle ils se trouvent actuellement.

Chapitre 3. Des nouveaux modèles de communication scientifique

1. Les archives thématiques

Les archives thématiques ou disciplinaires sont moins attachées à une institution qu'à une communauté de chercheurs. C'est dans ce modèle que nous retrouvons la première alternative aux éditeurs commerciaux : arXiv qui est également notre premier exemple.

1.1. ArXiv : e-print archive

<http://xxx.arxiv.org>

ArXiv est une archives thématique concernant premièrement et principalement les physiciens, communauté habituée à la circulation des pré-tirages, sur support papier comme sur support électronique.

1.1.1. Historique

Premier projet du genre, arXiv est lancé et développé en 1991 par Paul Ginsparg, physicien membre du laboratoire de Los Alamos [http://en.wikipedia.org/wiki/Paul_Ginsparg]. Aujourd'hui, il est présenté comme suit sur le site officiel :

“ArXiv is an e-print service in the field of physics, mathematics, non-linear science, computer science, and quantitative biology. The content of arXiv is conform to Cornell University academic standards. ArXiv is owned, operated and funded by Cornell University, a private not-for-profit educationnal institution. ArXiv is also

partially funded by the National Sciences Foundation.”

En 2001, Paul Ginsparg a quitté le laboratoire de Los Alamos et est devenu professeur à l'université Cornell, le projet arXiv a également effectué ce déplacement.

Les premiers documents accueillis dans cette archive concernaient principalement la physique. Au fil du temps, le domaine d'application s'est élargi et concerne aujourd'hui les mathématiques, l'informatique et la biologie quantitative [http://en.wikipedia.org/wiki/ArXiv.org_e-print_archive].

Prendre cette initiative était une manière de contourner les délais et les coûts imposés par la publication sur support papier. Il faut ajouter que l'utilisation de prétirages⁴ était déjà courante parmi les physiciens [De La Vega 2000, p.126], la mise en ligne qui permet une accessibilité immédiate est donc dans l'ordre des choses ; arXiv a, d'une certaine manière, permis l'informatisation d'une pratique habituelle.

1.1.2. Technique et fonctions

Fonctionnement général

Les documents disponibles sur arXiv doivent obligatoirement concerner l'un des domaines suivants : physique, mathématiques, sciences non linéaires, informatique et biologie quantitative. Il s'agit de textes scientifiques et techniques. L'archive n'est pas utilisable pour diffuser

⁴ Prétirage, preprint : “article soumis à publication à une revue, non encore publié et destiné à l'être, qui circule de façon informelle dans la communauté. Il fait partie de ce qui est défini plus largement sous le terme de “littérature grise”” [De La Vega 2000, p. 125].

Chapitre 3. Des nouveaux modèles de communication scientifique

d'autres types d'information (telles que des opinions, des nouvelles du monde politique ou académique, ...).

Le dépôt de texte est ouvert à tous. Chaque contribution est soumise à un spécialiste du domaine et depuis peu, un système d'acceptation (*endorsement*) des nouveaux auteurs a été mis en place. Ceux-ci doivent se trouver un "parrain" reconnu par arXiv pour pouvoir soumettre leurs articles. Le texte doit répondre à certains critères techniques comme le format (L^AT_EX, de préférence) et le "poids" du document. Il doit également être accompagné d'un résumé et être rangé dans une des subdivisions (éventuellement plusieurs mais ce n'est généralement pas utile au vu des domaines extrêmement précis qui sont couverts).

Il est possible pour un auteur de publier une nouvelle version de son travail mais les anciennes resteront accessibles.

L'accès au texte, quant à lui est libre et gratuit. De nombreux sites miroirs⁵ facilitent cet accès en répartissant les requêtes et les accès aux documents sur plusieurs serveurs.

La page d'accueil se veut avant tout pratique, elle permet de rentrer directement dans le vif du sujet. Elle présente les différents domaines et leurs subdivisions. Pour chacun des domaines et certaines subdivisions, il est possible d'accéder directement aux nouveaux *abstracts (new)*, aux *abstracts* et articles les plus récents, c'est-à-dire déposés dans les cinq derniers jours où il y a eu des dépôts (*recent*), de rechercher des *abstracts*

⁵ Afrique du Sud, Allemagne, Australie, Brésil, Chine, Corée du Sud, Espagne, Etats-Unis (deux miroirs en plus du site original), France, Grande Bretagne, Inde, Israël, Italie, Japon, Russie, Taïwan.

(*abs*) et des documents archivés dans les différentes subdivisions (*find*).

La page d'accueil propose également plusieurs liens “à propos d'arXiv” qui sont destinés aux utilisateurs tant du côté de ceux qui cherchent des informations que de ceux qui publient des textes.

Recherche d'information

La recherche de documents peut être effectuée de deux manières différentes : l'utilisateur peut envoyer sa requête via e-mail⁶ ou bien utiliser l'interface web. Cette seconde solution sera envisagée plus en profondeur.

La recherche à partir du site s'effectue obligatoirement au sein d'un domaine ; il n'est pas possible d'effectuer une recherche transversale sur l'ensemble des documents. Dans le cas d'une recherche en physique, il est possible très simplement de se limiter à une ou plusieurs subdivisions.

Les autres moyens de limiter sa recherche sont : la limitation chronologique (par défaut, une recherche s'effectue sur l'année en cours et non sur l'ensemble des années) et la limitation sur les champs de recherche (auteur, titre, contenu de l'*abstract*, contenu du document, commentaires, référence du périodique dans lequel l'article a éventuellement été publié, classification (*Association for Computing Machinery, Mathematic Subject Classification* ou arXiv), numéro de rapport).

Les résultats d'une recherche présentent premièrement la

⁶ Le destinataire du mail définit le domaine concerné, la requête elle-même est à indiquer dans le sujet du mail. Elle est, néanmoins, restreinte à la recherche de mots-clés dans les champs “auteur” et “titre” ; il est également possible de demander un document dont on connaît l'identifiant ; les différentes commandes possibles sont détaillées sur le site.

Chapitre 3. Des nouveaux modèles de communication scientifique

classification par sujet, le numéro de document, le titre, les auteurs (et l'institution à laquelle ils sont rattachés), les éventuels commentaires. Ils donnent également accès à l'*abstract* et au texte entier (au format .ps, .pdf ou autres).

Sur le site d'arXiv, aucune mention n'est faite concernant l'interopérabilité du dépôt ou le respect de l'un ou l'autre protocole pour la recherche d'information. Néanmoins, sur le site de l'Open Archive Initiative [<http://www.openarchives.org/Register/BrowseSites.pl>], il est mentionné comme soutenant le protocole OAI. ArXiv fait également partie des dépôts sur lesquels il est normalement possible d'effectuer une recherche (sur les méta-données uniquement) depuis le site Oaister [<http://oaister.umdl.umich.edu/o/oaister/viewcolls.html>].

1.1.3. Situation actuelle

Les statistiques d'utilisation disponibles sur le site [http://arxiv.org/show_monthly_submissions] montrent que depuis 1991, le nombre de documents soumis tous les mois suit une tendance forte à la hausse ; même si certains mois sont plus creux (en début et milieu d'année, soit les congés académiques habituels). Au 23 juin 2004, le nombre total de documents disponibles s'élevait à 279 628 (disponibles également depuis les 17 sites miroirs).

Actuellement, arXiv en tant que pionnier, bénéficie d'une certaine réputation dans le monde des scientifiques. Cette réputation et le fait qu'il s'agit d'un système bien rodé (mis en place en 1991, nous pouvons supposer qu'aujourd'hui, tout fonctionne correctement) font la force d'arXiv.

Chapitre 3. Des nouveaux modèles de communication scientifique

Ajoutons à cela qu'il a été mis en place au sein d'une communauté habituée à communiquer de manière informelle via les pré-tirages et qui privilégie la rédaction d'articles plutôt que de monographies. Tout laisse à croire qu'arXiv restera fonctionnel pour encore un certain temps.

Néanmoins, nous pouvons nous poser la question de l'avenir d'arXiv une fois que Paul Ginsparg, son concepteur, se retirera. Comme expliqué dans l'historique, arXiv a fait avec lui le trajet de Los Alamos à Cornell, que se passera-t-il plus tard si Paul Ginsparg quitte Cornell ? Il faut espérer que la relève sera assurée, au risque de voir arXiv tomber en désuétude faute d'un animateur dynamique.

Malgré ces avantages, certains inconvénients peuvent être cités. L'interface graphique, si elle est très efficace n'est pas pour autant très conviviale. L'université californienne de Davis donne accès à une interface de recherche pour la partie mathématique d'arXiv [<http://front.math.ucdavis.edu/>] qui n'offre pas de fonctionnalités supplémentaires mais qui permet d'effectuer une recherche dans un environnement plus agréable. Cette interface graphiquement pauvre est sans doute à considérer comme un moyen de faciliter l'accès du site aux personnes qui ne disposent pas de connexion rapide à Internet. Le site reste ainsi accessible à un plus grand nombre de personnes.

Un autre inconvénient est l'absence de *peer review*. Au départ, les soumissions étaient évaluées par une personne experte du domaine. Depuis janvier 2004, un système d'acceptation (*endorsement*) a été mis en place. Pour pouvoir soumettre un texte et être publié sur arXiv, un auteur doit se

Chapitre 3. Des nouveaux modèles de communication scientifique

faire “parrainer” par un scientifique de son domaine qui a déjà publié des articles dans arXiv. Ce système a été mis en place pour garantir la qualité des publications d'arXiv alors que le nombre de soumissions augmente.

ArXiv présente donc certains avantages aux physiciens, mathématiciens et autres qui veulent publier leur production ou rechercher de l'information, notamment par la gratuité de ces deux types de démarche et par la réputation acquise au fil du temps.

1.2. @rchive SIC : Archive Ouverte en Sciences de l'Information et de la Communication

<http://archivesic.ccsd.cnrs.fr/>

Autre dépôt thématique, autre domaine, autre lieu, @rchive SIC propose des articles sur les sciences de l'information et de la communication. Situé en France, la majorité des articles sont en français.

1.2.1. Historique

Le point de départ de la création d'@rchive SIC est une prise de conscience de l'état des revues scientifiques dans ce domaine et de leur faible diffusion envers les chercheurs. En 2000, le Ministère de la Recherche français demandait l'élaboration d'un outil visant à mesurer la production scientifique en sciences humaines et sociales. Pour ce faire, une série d'enquêtes ont été menées, notamment dans le domaine des revues en sciences de l'information et de la communication. Cette enquête montre que sur 263 revues de niveau mondial seules 17 sont connues par 50 % des chercheurs, et, au niveau national, seules 12 sur 40. Et cela ne signifie pas qu'ils les consultent régulièrement [Gallezot 2003a].

Face à ce constat, trois chercheurs (Gabriel Gallezot, Ghislaine Chartron et Jean-Max Noyer) ont proposé la création d'une archive ouverte en SIC [Gallezot 2002]. Celle-ci a été mise à disposition des chercheurs/auteurs en juin 2002. Ses trois objectifs principaux sont d'augmenter la diffusion des articles, de faciliter l'accès au savoir contenu dans ces articles et enfin de lutter contre la “babelisation⁷” de la discipline

⁷ “Phénomène social de non-communication, conséquent à des tentatives avortées d'échanges de message, à cause de différences dans les langages utilisés” [Office

Chapitre 3. Des nouveaux modèles de communication scientifique

tout en conservant une certaine diversité. Plus largement, la création de cette archive ouverte devrait également permettre :

- une plus grande visibilité des articles,
- un accès facilité à des écrits trop dispersés,
- un encouragement à des échanges pluridisciplinaires sur les objets de la recherche,
- une cartographie de la communauté InfoCom,
- le maintien d'une mémoire collective concourant à la structuration de la discipline,
- une réflexion sur les nouvelles fonctionnalités éditoriales liées aux variations des pratiques cognitives en cours [Gallezot 2003b].

L'@rchive SIC telle qu'elle est disponible actuellement se caractérise par rapport à d'autres projets par son approche spécifiquement francophone (voire française), même si la langue française n'est pas obligatoire pour la publication d'articles.

1.2.2. Technique et fonctions

Fonctionnement général

Les documents publiés au sein d'@archive SIC sont soit des articles publiés ou en cours de publication, soit des *working papers*. Les auteurs ne doivent pas spécialement appartenir à une institution reconnue. Le dépôt est ouvert à tous mais l'inscription est obligatoire. Ce dernier peut également être soumis par l'institution à laquelle appartient son auteur.

Aucun *peer review* n'est assuré par @rchive SIC, les articles étant de

québécois de la langue française, Grand dictionnaire terminologique, <http://w3.granddictionnaire.com>]

Chapitre 3. Des nouveaux modèles de communication scientifique

toute façon déjà acceptés pour publication par d'autres revues ou en passe de l'être. Chaque domaine est chapeauté par un modérateur qui accepte ou refuse la publication d'un article. Un modérateur peut parfois être responsable de plusieurs domaines.

Une fois le texte accepté, il est enregistré dans une ou plusieurs catégories ; il y en existe 22 :

- Bibliométrie, scientométrie
- Cinéma, art, esthétique
- Collectivité territoriale
- Communication et information scientifique
- Conflit, stratégie, veille
- Documentation
- Droit de l'information / communication
- Economie, industries culturelles
- Edition électronique
- Education formation
- Espace publique
- Gestion des connaissances
- Géopolitique
- Histoire de l'information / communication
- Hypertexte / hypermédia
- Ingénierie des systèmes d'information
- Muséologie
- Médias de masse
- Organisation et communication
- Sociologie de l'information / communication
- Théorie de l'information / communication
- Autres

Au niveau technique, l'@rchive SIC utilise le logiciel libre GNU EPrints, spécialement développé comme système d'auto-archivage à l'Université de Southampton. La version utilisée est toujours la première mais quelque peu modifiée. Cela pour :

- simplifier le dépôt d'un texte,

Chapitre 3. Des nouveaux modèles de communication scientifique

- rendre le multi-dépôt possible. Tout comme arXiv, il est possible de référencer un article dans plusieurs thématique,
- ouvrir le dépôt à des documents aux formats .rtf, .doc, .zip et vérifier les en-têtes de formats,
- générer automatiquement des versions .html et .pdf,
- permettre de partager la modération,
- permettre la recherche *full-text*,
- offrir une version bilingue,
- offrir des statistiques de dépôt.

Un intérêt particulier du logiciel GNU EPrints est sa compatibilité avec les normes de l'Open Archive Initiative, ce qui rend l'archive interopérable.

L'entièreté de l'archive est hébergée au sein du Centre pour la Communication Scientifique Directe (C.C.S.D.) du Centre National de la Recherche Scientifique (C.N.R.S.) qui gère également d'autres dépôts ouverts (Thèse en Ligne, les archives de l'Institut Jean Nicod, mémSIC : mémoires de 3ème cycle en S.I.C., ...) ainsi que le miroir français d'arXiv. Le fait qu'@rchive SIC soit ainsi hébergé, géré techniquement par le C.S.S.D. et lié au C.N.R.S. peut être vu comme une garantie de la pérennité de l'accès aux documents déposés, même s'il n'est pas exclu qu'un jour, le C.C.S.D. cesse ses activités par manque de moyens financiers.

Recherche d'information

Les possibilités de recherche d'articles au sein de l'archive sont variées et spécifiées très clairement dès la page d'accueil. Trois possibilités sont offertes aux visiteurs : parcourir la liste des auteurs ou des domaines, rechercher dans les mots clés ou dans le texte intégral, ou encore accéder à

Chapitre 3. Des nouveaux modèles de communication scientifique

un document dont l'identifiant est connu.

La recherche par mots-clés est simple ou avancée. La recherche simple est limitée aux champs titre/résumé en français/mots-clés en français, titre de la revue ou du colloque qui a accueilli l'article, auteur, domaine d'application. La recherche avancée s'étend aux champs résumé en anglais, mots-clés en anglais ; elle permet aussi de spécifier le type (article ou communication) et le statut (working paper, soumis à évaluation, publié ou en cours de publication) du document.

La recherche *full text* est effectuée grâce à un module extérieur appelé "ht://Dig" qui n'a pas été mis au point par le C.C.S.D. Celui-ci propose un système d'indexation et de recherche complet pour les intranets ou les domaines internet [<http://sourceforge.net/projects/htdig> ; <http://www.htdig.org>].

Dans un premier lieu, les résultats d'une recherche n'indiquent que le titre de l'article, son auteur, les références complètes de la revue ou du colloque, ainsi que le statut de l'article. L'accès au texte lui-même s'effectue via une page d'information plus complète comprenant les résumés en anglais et en français, les mots-clés attribués et les différents domaines dans lesquels le document est référencé. Les articles sont en général disponibles sous plusieurs types de formats : .rtf, .pdf, .html, éventuellement en .doc.

Il est également possible d'être tenu au courant des nouveaux articles déposés via un système d'abonnement par e-mail. L'utilisateur peut choisir la fréquence à laquelle il veut être alerté de l'arrivée de nouveaux documents (quotidienne, hebdomadaire ou mensuelle) et peut se limiter à certains domaines. L'abonnement aux nouveautés est bien sûr gratuit, il

Chapitre 3. Des nouveaux modèles de communication scientifique

suffit de se créer un compte sur le site, compte qui permet également le dépôt de document.

1.2.3. Situation actuelle

@rchive Sic propose à ses lecteurs une série de statistiques concernant l'évolution de l'archive depuis avril 2002 [http://archivesic.ccsd.cnrs.fr/info/index_fr.html].

La création de comptes est, en moyenne, de 21,26 par mois mais l'écart par rapport à cette moyenne est très varié : on trouve un pic de 85 comptes créés en octobre 2003 et un creux de 3 comptes créés en septembre 2002. L'évolution du nombre d'articles déposés tous les mois suit de près celle de la création de comptes et on retrouve les mêmes pics et creux. En octobre 2003, 77 documents ont été déposés dans l'archive. Ce nombre est dû au dépôt des articles du colloque C.I.F.S.I.C. (Conférence Internationale Francophone en Sciences de l'Information et de la Communication) qui a eu lieu à Bucarest du 28 juin au 2 juillet 2003 [mail de Christian Rossi reçu le 21/07/2004].

Au 27 juillet 2004, le nombre total de documents déposés est de 364. Cela peut paraître peu en comparaison avec un dépôt tel qu'arXiv. Cependant, @rchive SIC existe depuis moins longtemps, concerne des domaines plus restreints et un public principalement francophone, moins nombreux que les anglophones.

La plupart des articles (12 %) concerne la sociologie de l'information et de la communication mais ce domaine est suivi de près par l'édition électronique et la théorie de l'information et de la communication (8 % chacun) ainsi que par les domaines concernant la documentation (7 %), les

Chapitre 3. Des nouveaux modèles de communication scientifique

hypertextes et hypermédia (7 % également). Les chercheurs dans ces domaines que l'on retrouve en tête sont les plus susceptibles de s'intéresser à ce qu'est une archive ouverte en elle-même, à son fonctionnement et à sa possible évolution.

@rchive SIC a l'avantage de s'adresser aux professionnels des sciences de l'information et de la communication, au sens large. Ces personnes sont supposées être intéressées par ce genre de projets et sensibilisées à la problématique de la communication scientifique.

Alors que la plupart des projets “*open access*” sont (exclusivement) anglophones, @rchive SIC est majoritairement francophone mais ne se limite pas au français puisqu'il est possible d'effectuer une recherche sur les mots-clés et résumés en anglais, que certains articles sont écrits en anglais et que l'interface est disponible dans ces deux langues. Néanmoins, il est utile de se poser la question de la place des publications francophones alors que la majorité de la communication scientifique se fait aujourd'hui en anglais. La publication d'articles en français est une force autant qu'une faiblesse. Une force car elle permet la diffusion internationale de textes qui, à cause de leur langue, n'auraient vraisemblablement pas trouvé de place dans des revues internationales en anglais. Une faiblesse car le français reste moins lu et moins connu que l'anglais, et s'adresse donc à un public plus restreint.

@rchive SIC ne s'inscrit pas dans une volonté de “contourner” l'édition scientifique commerciale. Dans cette optique, les articles ou communications ne sont acceptés que s'ils ont déjà été évalués ou sont en passe de l'être (en étant communiqués lors d'un colloque, en étant destinés

Chapitre 3. Des nouveaux modèles de communication scientifique

à une revue ou en étant déjà publiés dans une revue). Ceci permet d'assurer une certaine qualité des textes puisqu'une évaluation a déjà été faite, ou sera faite, en plus du modérateur de domaine propre à @rchive SIC.

Pour terminer, notons que depuis octobre 2003, parallèlement à @rchive SIC, on trouve le projet mémSIC qui permet le dépôt et la consultation de mémoires, rapports de stage concernant également les sciences de l'information et de la communication. Ces deux dépôts se complètent pour donner un aperçu de ce qui se fait dans le domaine des SIC, principalement en France.

1.3. RePEc : Research Papers in Economics

<http://repec.org>

Troisième et dernier dépôt thématique, RePEc propose une quantité importante d'information et de services. Fonctionnant sur un modèle décentralisé, il constitue ainsi une ressource non négligeable pour les économistes.

1.3.1. Historique

RePEc, tel qu'il se présente actuellement est l'évolution du projet NetEc. Le projet NetEc voit ses origines remonter à 1991 et au travail de Thomas Krichel à l'université de Loughborough. Le rôle de NetEc était d'améliorer la communication scientifique dans le domaine des sciences économiques via les médias électroniques (listes de diffusion, gopher, wais, ftp, http, ...). NetEc se présentait sous la forme de différents projets (BibEc, WoPEc, CodEc, WebEc, BizEc et HoPEc) ayant chacun des rôles différents mais travaillant à l'accomplissement des mêmes objectifs. Ces projets ont évolué, certains font encore partie aujourd'hui de RePEc [Krichel 1997].

RePEc se veut la réalisation matérielle du concept d' *Open Library* mis au point par Thomas Krichel. Dans cette optique, il est possible non seulement d'ajouter et de référencer du contenu dans RePEc mais aussi de créer des services à partir des données de RePEc. Par exemple : mettre au point une interface de recherche et de présentation de documents ne concernant que certains sous-domaines des sciences économiques ou ne concernant que certains types de documents. C'est ainsi que plus de 100 personnes provenant de 44 pays différents travaillent bénévolement (en

Chapitre 3. Des nouveaux modèles de communication scientifique

tout cas, sans rémunération de la part de RePEc, il est possible que ces personnes soient payées par l'institution à laquelle ils appartiennent) aux différents services offerts par RePEc [Krichel 2000]. Parmi ces participants, nous en retrouvons quelques uns en Belgique. Ceux-ci sont :

- Energy, Transport and Environnement Workgroup (K.U.L.) ;
- International Economics Working Group (K.U.L.) ;
- LICOS Centre for Transition Economics (K.U.L.) ;
- Centrum voor Economisch Studiën (K.U.L.) ;
- Faculty of Economics and Business Administration (R.U.G.) ;
- Faculté des sciences économiques, sociales et politiques (U.C.L.) ;
- Service of International and Interregional Economics (U.Lg) ;
- Vlerick Leuven Gent Management School.

1.3.2. Technique et fonctions

Plus que simplement donner accès aux documents, RePEc tente d'offrir un maximum d'informations concernant autant les auteurs et les institutions que les ressources documentaires en sciences économiques. Cet objectif ne peut être approché qu'en mettant en place un travail collaboratif et décentralisé (car demandant des ressources matérielles et humaines importantes qu'une institution seule ne peut pas fournir) [Krichel 2000].

Le site web de RePEc joue ainsi un rôle de centralisateur et de guide pour les utilisateurs des services et pour les éventuels créateurs de services.

En terme de documents disponibles via RePEc, ceux-ci peuvent être de plusieurs types : *working papers* (135 000), articles (143 000), logiciels (1 000), listes d'information sur des livres et chapitres de livres (750), des auteurs et leurs publications (5 200) ainsi que sur des institutions (8 050)

Chapitre 3. Des nouveaux modèles de communication scientifique

qu'on limiterait à tort aux universités puisque nous retrouvons certains acteurs étatiques telle que la Réserve fédérale des Etats-Unis ou encore certaines banques. Une grande partie des documents renseignés dans RePEc sont accessibles en ligne mais pas toujours gratuitement étant donné que certains partenaires de RePEc sont des éditeurs commerciaux (par exemple : Kluwer Academic Publisher). Dans ce cas, un système de *pay-per-view* est en général proposé au visiteur, sauf si celui-ci est déjà abonné au périodique qu'il souhaite consulter.

Comme expliqué plus haut, RePEc se compose de différents services. Chacun mériterait une analyse poussée mais ce n'est malheureusement pas possible dans le cadre de ce travail. Un bref aperçu de chacun est néanmoins indispensable.

- **EconWPA** (University of Washington) dont l'objectif est la “*free distribution of working papers in economics*”. Pour ce faire, EconWPA propose aux personnes dont l'institution ne serait pas partenaire de RePEc de déposer leurs *working papers*. Ceux-ci seront repris par après dans la base RePEc.
- **IDEAS - Internet Documents in Economics Access Service** (University of Connecticut, department of Economics) se présente comme “[*the largest bibliographic database dedicated to Economics*]” et donne accès à toutes les informations de RePEc.
- **EconPapers – Economics at your fingertips** (Economic Research Institute at the Stockholm School of Economics) propose un accès à la base de données RePEc.
- **NEP – News Economics Papers** (Bernardo Bátiz-Lazo & Marco

Chapitre 3. Des nouveaux modèles de communication scientifique

Novarese) est un service d'alerte pour les nouvelles contributions, disponible sur inscription. Il est possible de n'être informé que sur certains sous-domaines, ceux-ci sont chapeautés par des spécialistes qui sélectionnent les nouveaux documents intéressants.

- **EDIRC - Economics Departments, Institutes and Research Centers in the World** (University of Connecticut, Department of Economics) liste toutes les institutions (8 078) et leurs coordonnées dans 217 pays.
- **LogEc** (Stockholm School of Economics) présente les statistiques de différents services de RePEc (NetEc, IDEAS, Socionet, EconPapers et NEP).
- **Citations in Economics** (Institute of Economic Research at Hitotsubashi University, Tokyo) “*provides citation analysis for documents distributed on the RePEc digital library*”. Ce service n'est pas encore très développé et ne prend en compte que les références aux documents électroniques.
- **WebEc – World Wide Web Resources in Economics** (Center for Innovative Education, Helsinki School of Economics & Faculty of Social Sciences, University of Helsinki) catégorise les informations libres d'accès en sciences économiques disponibles sur le web.
- **WoPEc – Working Papers in Economics** (Jose Manuel Barrueco & Thomas Krichel) utilise RePEc mais n'en retire que les *working papers*.
- **BibEc** (Jose Manuel Barrueco & Thomas Krichel) complète WoPEc en référençant des documents sur support papier.
- **inomics – The Internet Site for Economists** fournit un accès à RePEc ainsi que d'autres services pour les économistes (offres d'emploi, annonces de conférences,...).

- **SPZ – Socionet Personal Zone** se présente comme “[a] Workplace for Researchers, Tutor and Students in the Social Sciences Information Space”.
- **Socionet** met à disposition du public russophone un accès à RePEc ainsi que des services complémentaires.

Tous ces services fonctionnent de manière indépendante (sauf bien sûr, quand ils utilisent les données provenant d'un ou de plusieurs autres services) ; leur principal point commun est leur rattachement à RePEc. Il est important pour l'utilisateur, le visiteur de RePEc de s'informer sur le contenu de chacun de ces services de manière à utiliser celui qui sera le plus approprié à ses besoins.

1.3.3. Situation actuelle

Les statistiques disponibles sur LogEc (ne concernant donc que les services NetEc, IDEAS, Socionet, EconPapers et NEP) montrent que, de manière générale, l'utilisation des différents services disponibles est en augmentation, même si certaines périodes sont moins fastes que d'autres. Nous pouvons également noter une augmentation du nombre de documents disponibles. Depuis janvier 1998 jusque juin 2004, ce ne sont pas moins de 8 340 892 fichiers de tous types qui ont été téléchargés par les utilisateurs de RePEc.

En ce qui concerne le type de document, les statistiques montrent que ce sont les *working papers* qui sont le plus souvent téléchargés et dont les abstracts sont le plus souvent lus. Viennent ensuite les articles de périodiques. La consultation de livres, chapitres de livres et le téléchargement de logiciels semblent anecdotiques par rapport au succès des deux types de documents précédents (articles et *working papers*).

Chapitre 3. Des nouveaux modèles de communication scientifique

La décentralisation et la possibilité pour qui veut de créer de nouveaux services permet d'augmenter l'offre de documentation ainsi que les moyens d'y accéder. Ceci permet une offre plus variée de documentation et donc plus susceptible de rencontrer les besoins des utilisateurs. Dans l'état actuel des choses, nous pouvons déjà constater que RePEc ne se limite pas à un seul type de document ; l'offre de service semble vouloir être exhaustive.

La division en de nombreux services n'est pas toujours claire et pourrait décourager l'utilisateur. Il est nécessaire que l'utilisateur de RePEc prenne le temps de comprendre comment fonctionne RePEc et ce que lui offre les différents services. S'il trouve cet apprentissage trop long ou trop complexe, pourquoi ne se contenterait-il pas d'une recherche sur Google ?

Il serait intéressant pour RePEc d'uniformiser un maximum les interfaces des différents services. Cela renforcerait la cohérence du projet et permettrait aux utilisateurs d'utiliser plus facilement les différents services.

Le nombre d'institutions et de personnes impliquées dans RePEc peut être le garant d'une survivance à long terme. Il y a peu de chance que tous se retirent de RePEc en même temps. D'un autre côté, l'importance du nombre de participants n'est pas forcément un gage de qualité.

Si l'offre de documentation de RePEc est large, diversifiée et pour la plupart d'accès gratuit, il est nécessaire pour l'utilisateur de s'habituer et de comprendre son fonctionnement. Cette démarche favorisera les utilisateurs du monde des sciences économiques qui ont souvent besoin d'accéder au type de documentation contenue dans les bases de données de RePEc. Par contre, l'utilisateur ponctuel ne fera sans doute pas l'effort d'apprendre à

Chapitre 3. Des nouveaux modèles de communication scientifique

utiliser RePEc et préférera peut-être effectuer sa recherche via des moteurs de recherche plus généralistes.

2. Les dépôts institutionnels

Ce deuxième modèle se différencie du premier principalement par le critère de regroupement des documents qui, ici, n'est pas le thème mais la provenance du document. Celui-ci doit avoir été créé au sein de l'institution pour pouvoir apparaître dans le dépôt. Nous verrons que cela ne se fait pas toujours selon les mêmes modalités

2.1. Caltech CODA

<http://library.caltech.edu/digital>

Le dépôt du California Institute of Technology est notre premier exemple. Si un dépôt institutionnel est, en général, multidisciplinaire, nous verrons que cet aspect est parfois limité.

2.1.1. Historique

Caltech CODA (pour *Collection of Open Digital Archives*) est le nom donné au dépôt institutionnel lancé en 1999 par le “California Institute of Technology” [<http://www.caltech.edu>] situé à Pasadena. Une réflexion sur la situation de la publication scientifique était déjà entamée bien avant 1999 et le projet CODA. En effet, une conférence avait été organisée en 1997 sur ce thème. Il en est ressorti une certaine vision de ce qui pourrait être un système alternatif de communication scientifique. L'idée principale sur laquelle reposerait ce système est la séparation du processus de certification de celui de dissémination du document [Schaffner 1997].

L'idée fait son chemin et en 1999, le “Scholar's Forum”, dans un article, propose un nouveau modèle pour la communication scientifique. Il y définit les différentes fonctions que devrait assurer le système, les

Chapitre 3. Des nouveaux modèles de communication scientifique

principaux acteurs à impliquer, la base de données documentaire (inspirée de celle développée par Paul Ginsparg pour arXiv), les différents dispositifs à mettre en place ainsi que les standards et protocoles à utiliser [Buck 1999].

C'est également en 1999 qu'est mise en place la phase de lancement de CODA. Le projet va prendre son envol en 2000 avec l'élargissement à différents types de documents (en réalité, à tous les documents considérés comme intéressants), l'amélioration du matériel informatique qui sera mieux adapté à l'ampleur du projet et l'augmentation de l'implication des bibliothécaires. Ceux-ci vont apprendre comment fonctionne le système et établir un standards de métadonnées flexibles. En 2000, CODA se compose principalement de l'archive du CSTR - *Computer Science Technical Reports* - à laquelle s'ajoute, en 2001, les archives CAV 2001 (pour *Cavitation 2001 Proceeding*) et Caltech EERL (pour *Caltech Earthquake Engineering Research Laboratory Reports*). En 2002, ce sera au tour de Caltech PARADISE (*Parallel and Distributed Systems Group E-Tech Reports*) d'enrichir l'archive de Caltech [Douglas 2002].

L'archive s'est également enrichie de :

- Caltech BOOK, *Books by Caltech Authors* en 2002 ;
- Caltech CACR, *Center for Advanced Computing Research* en 2004 ;
- Caltech CDSTR, *Control and Dynamical Systems Technical Reports* en 2003 ;
- Caltech ETD, *Electronic Theses and Dissertations* en 2001 ;
- Caltech GALCITFM, *Graduate aeronautical Laboratories (Fluid Mechanics) Technical Reports* en 2001 ;

Chapitre 3. Des nouveaux modèles de communication scientifique

- Caltech JSNC, *11th Joint Symposium on Neural Computation* en 2004 ;
- Caltech KHR, *W.M. Keck Laboratory of Hydraulics and Water Resources Technical Reports* ;
- Caltech LESSGS, *Large-Eddy Simulation and Subgrid-Scale Modeling for Turbulent Mixing and Reactive Flows* en 2003 ;
- Caltech LIB, *Library System Papers and Publication* en 2001 ;
- Caltech OH, *Oral History* en 2002⁸.

On le voit, de nouvelles archives sont venues périodiquement enrichir le dépôt institutionnel de Caltech ; et actuellement, d'autres sont encore en développement.

2.1.2. Technique et fonctions

Fonctionnement général

Comme expliqué brièvement dans l'historique, le dépôt institutionnel du California Institute of Technology est composé de 19 archives indépendantes dont cinq sont en développement. La création d'une archive au sein de Caltech CODA doit répondre à différents critères bien définis :

“A. Each CODA repository must consist of a body of works sponsored by Caltech professional faculty.

B. Works in CODA repositories :

1. must be produced, submitted, or sponsored by

⁸ Ces informations proviennent des pages spécifiques des différentes archives ou de leurs statistiques ; la date à laquelle commencent les séries statistiques sont alors considérées comme la date de début de projet. En ce qui concerne W.M. Keck Laboratory of Hydraulics and Water Resources Technical Reports, l'information n'est pas disponible.

Chapitre 3. Des nouveaux modèles de communication scientifique

Caltech professional faculty,

2. must be complete and in final form,
3. must be scholarly, research, or educational material of enduring value.

C. The author/owner of each work must grant to Caltech the non-exclusive right to preserve and distribute the work in perpetuity.”

[<http://library.caltech.edu/digital/criteria.htm>]

Il n'y a pas, dans ces critères, d'obligation concernant les types de documents que l'on retrouve dans Caltech CODA. Cela ne semble pas important tant que le contenu du document présente un intérêt. Néanmoins, ce critère est pris en compte dans deux archives : Caltech BOOK et Caltech ETD (*Electronic Theses and Dissertations*), cela tient bien évidemment à leur contenu.

Les domaines abordés dans le cadre de Caltech CODA dépendent des archives créées. Actuellement, toutes les disciplines enseignées à Caltech⁹ n'ont pas une archive dédiée ; leur création dépend entièrement de la motivation du personnel.

Bien que chaque archive soit indépendante, les interfaces d'accès à la documentation sont identiques ; excepté pour ETD (*Electronic Theses and Dissertations*), OH (*Oral Histories*) et Cav2001 (*Fourth International Symposium on Cavitation, 2001*). Les mêmes menus sont proposés aux

⁹ “Biology, chemistry and chemical engineering, engineering and applied science, geological and planetary sciences, humanity and social sciences, physics, mathematics & astronomy” [<http://www.caltech.edu>]

Chapitre 3. Des nouveaux modèles de communication scientifique

visiteurs et la même simplicité que l'on retrouve chez arXiv est de mise : pas d'image, texte noir (ou bleu pour les liens) sur fond blanc permettant ainsi un chargement rapide de la page et une compatibilité avec la plupart des navigateurs web. Caltech ETD, OH et Cav 2001 bénéficient, eux, d'une mise en page plus riche.

Graphiquement, ETD semble s'inscrire dans le site général de la bibliothèque. Contrairement aux autres archives de CODA, ETD est directement accessible depuis la page principale du “*Caltech Library System*” [<http://library.caltech.edu>] via les liens “*Electronic These*” puis “*Electronic Theses Only*” qui mènent directement à la page d'accueil de Caltech ETD. En plus des possibilités également offertes par les autres archives, ETD consacre une section spéciale aux étudiants ; ceux-ci sont, en effet, les premiers concernés par la publication électronique dans le cadre d'ETD. Il leur est d'ailleurs précisé dans cette section que Caltech participe au projet NDLTD (Networked Digital Library of Theses and Dissertations).

La page d'accueil de l'archive Caltech *Oral Histories* [<http://oralhistories.library.caltech.edu>] bénéficie aussi d'une mise en page plus recherchée. Cela va sans doute de pair avec les objectifs du projet. Le principal est de “[*purpose of*] recording the personal memories of the distinguished scientists, teachers and administrators of the Institute”. Ce qui, sans négliger l'intérêt scientifique de cette démarche, peut aussi bien servir le prestige de l'institution. Le projet méritait donc une interface graphique plus élaborée. Ajoutons que l'archive Caltech OH peut être téléchargée dans son intégralité en un fichier unique.

Enfin, le cas de Cav 2001 est différent. Cette archive a été mise en

Chapitre 3. Des nouveaux modèles de communication scientifique

place dans le cadre d'un symposium : les participations à celui-ci ont pu être mises en ligne via Cav 2001 avant le début des conférences. Le parcours de l'archive s'effectue par "session" suivant ainsi la structure de la manifestation [Buck 2002].

Au niveau technique, le logiciel utilisé pour le support des différentes archives est GNU EPrints de l'université de Southampton et ETD-db développé par Virginia Tech (Virginia Polytechnic Institute and State University) pour Caltech ETD.

Toutes les archives du dépôt respectent le *Protocole for Metadata Harvesting* proposé par l'Open Archives Initiative. Celui-ci rend l'archive interopérable, l'utilisation d'un protocole commun aux différents dépôts ouverts permet la création de services transversaux. Chaque dépôt étant répertorié séparément comme fournisseurs de données, ils participent comme entités séparées aux différents services basés sur les standards définis par l'OAI.

La recherche d'information

Chaque archive étant indépendante, il n'est pas possible, sur le site de Caltech CODA, d'effectuer une recherche transversale. L'uniformité des moyens de recherche d'information est néanmoins de mise. Trois moyens d'accès à l'information sont proposés : le menu "*main*", parcourir et recherche (simple ou avancée). La page principale ("*main*") de chaque archive permet une recherche très simple soit sur l'enregistrement du document (auteur, titre, ...) soit une recherche *full text*. Il est également possible de consulter directement les derniers documents qui ont été déposés dans l'archive.

Chapitre 3. Des nouveaux modèles de communication scientifique

Le parcours des archives peut se faire sur deux et parfois trois critères. Nous retrouvons dans tous les cas les parcours par sujets et par années et il est parfois possible de parcourir par personnes ; c'est le cas pour Caltech BOOK.

La recherche simple peut s'effectuer sur les champs suivants : *Title/Abstract/Authors/Creators/Date* ; *Full Text/Title/Abstract/Authors/Creators/Date* ; *Authors/Creators/Date* ou uniquement sur le champ *Date* qui permet une recherche sur un intervalle précis. La recherche avancée, quant à elle, offre un nombre de champs de recherche plus étendu et plus précis : *Full Text* ; *Title* ; *Authors/Creators* ; *Abstract* ; *Uncontrolled Keywords* ; *Subjects* (subdivision du domaine de l'archive) ; *EPrint Type* ; *Department* ; *Editors* ; *Status* ; *Refereed* ; *Journal/Publication Title* ; *Date*.

Une fois que l'utilisateur sait dans quelle archive il doit situer sa recherche, il n'est plus difficile pour lui de retrouver l'information qu'il désire.

2.1.3. Situation actuelle

Il n'existe pas de statistiques disponibles pour l'ensemble du dépôt mais pour chaque archive en particulier. Ces statistiques concernent principalement l'accès aux documents ; celui-ci est facilement déduit du nombre de fois qu'un fichier est téléchargé. Certaines archives mettent en ligne deux types de format : .ps et .pdf. Dans tous les cas, c'est le format .pdf qui remporte le plus de succès auprès des utilisateurs.

En général, le téléchargement de documents est en hausse, cette hausse est plus ou moins marquée selon les archives. Deux exceptions à ce

Chapitre 3. Des nouveaux modèles de communication scientifique

mouvement : Caltech PARADISE (*Parallel and Distributed System Group*) dont le nombre de téléchargement diminue et Cav 2001 (*Fourth International Symposium on Cavitation, June 2001*) dont le nombre de téléchargements semble plus ou moins stable depuis la fin du symposium. Par son statut et son utilisation spécifique, cette situation n'est pas étonnante, cette archive ne s'enrichissant normalement pas.

Une approximation faite depuis le site d'oaister, qui permet une recherche à travers les sites respectant le protocole de l'Open Archives Initiative, donne pour résultat le chiffre de 2 331 documents disponibles pour l'ensemble du dépôt (au 12 septembre 2004) [<http://oaister.umdl.umich.edu/o/oaister/browse.html>].

En plus des 14 archives déjà “en service”, cinq supplémentaires sont en développement : ASCI (*Accelerated Strategic Computing Initiative Technical Reports*), BIC001 (*Biological Imaging Center's Frog (Xenopus) Gastrulation Images*), EQL (*Environnement Quality Laboratory Technical Reports*), ES (*Engineering & Science (journal)*), GALCITSM (*Graduate Aeronautical Laboratories (Solid Mechanics) Technical Reports*). Avec ces cinq dépôts supplémentaires, il est probable que l'intérêt des chercheurs pour Caltech CODA augmente encore.

Le fait que la mise sur pied d'une archive se fasse sur base volontaire devrait être la garantie de l'utilisation de l'archive par les personnes concernées, au moins en ce qui concerne les dépôts de documents.

Malgré la disparité des dépôts, le fait qu'ils ne forment pas ensemble une seule grosse archive, l'interface de consultation reste la même pour chacun (sauf ETD, OH et Cav 2001 comme expliqué plus haut). Le visiteur

Chapitre 3. Des nouveaux modèles de communication scientifique

ne se sentira vraisemblablement pas perdu s'il visite les différents dépôts les uns après les autres. Par contre, il ne pourra pas effectuer de recherche sur l'ensemble de Caltech CODA, à moins de passer par un service OAI de type OAIster [<http://oaister.umdl.umich.edu/o/oaister/>] qui ne permet pas d'effectuer des recherches *full text*.

Autre désavantage de CODA : la création d'une archive se fait sur base d'initiative plus ou moins personnelle, ce qui a pour conséquence que toutes les disciplines enseignées au California Institute of Technology ne sont pas reprises dans le dépôt institutionnel sauf éventuellement, via l'archive des thèses (Caltech ETD) et des livres (Caltech BOOK).

Le dépôt institutionnel de Caltech est le fruit d'une réflexion sur la communication scientifique commencée au plus tard en 1997. L'alternative au circuit d'édition de la documentation scientifique dite "classique" semble avoir été mûrement réfléchi. Dans un article de 1999 qui propose un nouveau modèle pour la communication scientifique [Buck 1999], le groupement d'universités est considéré comme un des acteurs principaux dans la mise en place de ce nouveau modèle. Pourtant, Caltech n'est associé à aucune autre université (exceptées celles dont il utilise les logiciels comme EPrints de Southampton et ETD de Virginia Tech), pas même au projet eScholarship [<http://www.cdlib.org/programs/escholarship.html>] de l'université de Californie. Il faut préciser que le projet eScholarship a été mis en place quelques années après Caltech CODA et, pour cette raison, il n'était peut-être pas possible de joindre les deux archives.

Par la centralisation des différentes archives au sein d'un seul dépôt institutionnel, Caltech CODA offre un bon outil à ses membres, tant pour la

Chapitre 3. Des nouveaux modèles de communication scientifique
recherche que pour la publication d'information.

2.2. Digital Library of Massachusetts Institute of Technology Theses

<http://theses.mit.edu>

Si Caltech CODA acceptait tous les types de documents dans son archive, cet exemple du MIT montre que ce n'est pas toujours le cas. Ce genre de limitation permet la diffusion de documents peu et difficilement consultables : les thèses.

2.2.1. Historique

La “Digital Library of Massachusetts Institute of Technology Theses” est un dépôt institutionnel un peu particulier par rapport à ce que nous avons vu précédemment avec la “Collection of Open Digital Archives” de Caltech, puisqu’il concerne exclusivement les thèses. De plus, si la majorité des thèses disponibles sont datées de ces dernières années, un processus de numérisation a permis à ce dépôt de proposer des travaux datant de la fin du XIX^{ème} siècle. Celles-ci sont néanmoins peu nombreuses car l’objectif n’a jamais été d’offrir au public la totalité des thèses du MIT. A vrai dire, si le public a aujourd’hui accès à ces thèses, c’est un peu le fruit du hasard : la sélection de ces thèses ne se fait sur aucun critère précis, seul entre en compte la demande des utilisateurs pour une version papier.

Au départ, en 1998, ces thèses ont été numérisées pour servir uniquement le “Document Services department” de la bibliothèque du MIT. Ces versions numérisées étaient (et sont toujours) très utiles lorsqu’une demande d’achat d’un exemplaire de thèse était adressée au département. Une fois la thèse numérisée, il ne restait plus qu’à l’imprimer et l’envoyer à

son destinataire.

Avant de rendre l'accès public, il était obligatoire d'enrichir cette bibliothèque numérique de métadonnées qui permettraient aux utilisateurs d'effectuer un minimum de recherche sur le fonds documentaire. Ces métadonnées ont été extraites de "Barton", le catalogue en ligne de la bibliothèque du MIT qui, évidemment, renseigne toutes les thèses.

Petit à petit, et bien qu'il n'y ait pas eu de promotion, ni de publicité pour ce nouveau service, des scientifiques du monde entier se sont intéressés à cette importante source d'information et l'ont utilisée [Stone 2000].

Aujourd'hui, le dépôt de thèses du MIT n'est plus le seul à pouvoir accueillir une part de la production intellectuelle de l'Institut. Depuis 2000, le MIT associé à la société Hewlett-Packard a travaillé à la mise au point d'un logiciel *open source*, disponible gratuitement : Dspace. Celui-ci, un peu comme GNU Eprints permet la mise en place facile de dépôts de documents numériques. Le MIT en a bien sûr profité et propose à ses membres la création d'archives numériques avec Dspace. Dix "communautés" en ont déjà profité [<http://dspace.mit.edu>].

2.2.2. Fonctions et technique

Fonctionnement général

"Digital Library of theses" ne donne accès qu'à un nombre limité de thèses. Même pour les plus récentes, l'accès en ligne n'est pas assuré puisque cela dépend du bon vouloir de l'auteur.

Seules les images extraites des thèses sont d'accès ouvert et consultables par tous. Le visiteur peut avoir une vue générale de la thèse

Chapitre 3. Des nouveaux modèles de communication scientifique

qu'il désire consulter via les vignettes. Cet aperçu global est très utile pour identifier la table des matières et ainsi définir la partie de la thèse la plus intéressante. Il lui est également possible de naviguer à travers le texte en indiquant le numéro de l'image qu'il veut visionner.

Une fois l'image sélectionnée, l'utilisateur peut la visionner en format .gif en résolution basse (100 dpi) et en nuances de gris. Grâce à ce choix de résolution, le chargement de la page se fait relativement rapidement, selon le type de connexion. Dès que l'image est affichée, il est possible d'agrandir l'une ou l'autre partie, de visionner les pages précédentes ou suivantes et de revenir aux vignettes. Notons que, si visionner une image avec un navigateur web a certains avantages (entre autres, celui de ne pas devoir sauver l'image sur disque dur), cela présente certains inconvénients ergonomiques lorsque le texte est mis en forme en "paysage" alors que le visionnement se fait en mode "portrait". Les navigateurs n'offrent pas la possibilité de tourner l'image¹⁰.

Si la consultation de la thèse via les images s'avère trop fastidieuse, il est possible de commander le document soit en format .pdf, soit sur support papier. Les deux nécessitant une participation aux frais qui peut atteindre plus de \$65 pour un .pdf et plus de \$80 pour une version papier. Dans les deux cas, le nombre de pages influence le prix de la thèse. L'argent ainsi récolté permet à la bibliothèque de continuer d'assurer ses services et notamment, la mise en ligne des thèses.

L'accès aux thèses du MIT n'est donc que partiellement ouvert car si l'utilisateur veut bénéficier d'un peu de confort (nécessaire à une lecture

¹⁰ Ici, c'est le navigateur Mozilla Firefox 0.9.1 qui a été utilisé.

Chapitre 3. Des nouveaux modèles de communication scientifique

entière, accessoire si on ne souhaite la consulter que partiellement), il est obligé de payer. Il est curieux de constater que même les thèses récentes sont disponibles gratuitement uniquement sous forme d'image alors qu'une version électronique – dont il est aisé de tirer un .pdf – existe certainement. L'avantage de ces images pour le MIT et l'auteur de la thèse est qu'il n'est pas facile de “copier-coller” le texte. Ceci dit, si l'éventuel “pillard de thèse” le désire, il est toujours possible d'appliquer aux images une reconnaissance optique des caractères et ainsi, en recréer le texte. Le pendant de cette mesure de protection est qu'il n'est pas possible d'effectuer une recherche *full text* sur ces thèses.

La recherche d'information

Les textes des thèses étant des images, la recherche *full text* – comme nous l'avons déjà dit – est impossible. La recherche simple ne peut s'effectuer que sur deux champs : auteur et titre. La recherche avancée propose en plus le champ résumé et peut porter sur les thèses électroniques ou sur toutes les thèses du MIT. Il est également possible de rechercher un document dont on connaît l'identifiant.

La recherche proposée sur les thèses en ligne n'est donc pas très complexe ni complète. Il est d'ailleurs suggéré sur le site, d'effectuer d'abord une recherche dans Barton, le catalogue complet des bibliothèques du MIT. Il est possible, de là, de spécifier qu'on ne cherche que les thèses du MIT. En utilisant Barton, le visiteur peut limiter sa recherche sur des champs tels que le département dans lequel la thèse a été défendue, son superviseur, le diplôme pour lequel elle a été présentée... Une fois que la requête dans Barton a fourni ses résultats, il ne reste à l'utilisateur qu'à

vérifier (et espérer) que la thèse existe sous format numérique. Notons qu'il est parfois plus simple et plus rapide de parcourir les thèses numériques par années de publication ou par auteur.

2.2.3. Situation actuelle

Aucune statistique concernant le dépôt de thèses du MIT n'est disponible. Nous savons qu'en 2000, le nombre de thèses hébergées s'approchait des 4.000 [Stone 2000, p. 2]. Un bref comptage¹¹ des thèses datées de 2000 à 2004 nous indique qu'à présent, le dépôt de thèses comprend plus de 5.600 documents ; notons qu'il est possible que des thèses plus anciennes aient été ajoutées.

Le dépôt de thèse du MIT est caractéristique par rapport à d'autres projets, en ce sens qu'il était au départ basé sur une démarche rétrospective et propose donc à ses utilisateurs de consulter des thèses anciennes (1879 pour la plus ancienne).

Malheureusement, l'accès aux thèses en ligne, s'il est assuré, n'est pas très ergonomique. Nous l'avons évoqué plus haut, il n'est pas toujours pratique de consulter ces documents sous forme d'image et ce format empêche l'offre de certains services comme la recherche *full text*. Il est bien sûr possible d'obtenir une version sur support papier ou en format .pdf mais cela entraîne des délais et des coûts qu'un chercheur ne peut pas toujours se permettre. Ce système de consultation par image a son intérêt pour le MIT puisque, couplé au prix des versions papier et .pdf, il peut servir de barrière technique au piratage et à la copie de thèse.

¹¹ Comptage grossier réalisé sur base de la recherche de thèses par année, le résultat n'est qu'approximatif et ne sert qu'à donner un ordre de grandeur.

Chapitre 3. Des nouveaux modèles de communication scientifique

Il serait intéressant pour ce service d'augmenter encore le nombre de thèses disponibles, qu'elles soient anciennes en lançant une politique de numérisation rétrospective plus systématique ou qu'elles soient récentes en obligeant le dépôt numérique des thèses. Mais cette obligation n'est pas du ressort du service des bibliothèques, une intervention des autorités supérieures serait indispensable.

Il serait intéressant de voir, dans le futur, de quelle manière le dépôt de thèses du MIT va évoluer alors que face à lui, nous retrouvons le dépôt DSpace qui, pourquoi pas, pourrait également héberger des thèses. DSpace mettant les documents entièrement à disposition des visiteurs et permettant, lui, la recherche *full text*. Nous pouvons nous demander quelle sera la place réservée au dépôt de thèses si ces dernières peuvent être mises à disposition sur DSpace. Peut-être n'offrira-t-il qu'un accès aux thèses plus anciennes qui nécessitent d'être scannées et mises sous format image pour être ensuite diffusées sur le net ce qui pourrait justifier les coûts d'accès aux formats plus ergonomiques. DSpace pourrait donc influencer l'évolution de l'accès aux thèses du MIT.

2.3. e-Prints Soton : University of Southampton e-Prints Service

<http://eprints.soton.ac.uk>

Dernier exemple de dépôt institutionnel, e-Prints Soton nous donne un aperçu service européen, britannique, plus précisément. Le Royaume Uni est sans doute un des pays européen les plus actifs en matière d'initiative *Open Access*.

2.3.1. Historique

Le dépôt institutionnel de l'université de Southampton, e-Prints Soton a été développé dans le cadre du projet TARDis (*Targeting Academic Research for Dissemination and Disclosure*) qui s'inscrit lui-même dans le cadre du programme FAIR¹² (*Focus on Access to Institutional Resources*) financé par le JISC¹³ (*Joint Information Systems Committee*) [<http://eprints.soton.ac.uk/information.html>].

Le projet TARDis a été lancé au mois d'août 2002 [<http://www.jisc.ac.uk/index.cfm?name=projectbrowse>] avec pour objectif la mise sur pied d'un dépôt institutionnel multidisciplinaire qui donnerait accès à la recherche scientifique menée à l'université de Southampton. Ce dépôt serait bâti grâce à un système d'auto-archivage par les chercheurs et

12 Le programme FAIR a pour objectif l'évaluation et l'exploration des différents mécanismes permettant la diffusion et le partage de contenu sur le web [http://www.jisc.ac.uk/index.cfm?name=programme_fair].

13 JISC offre, dans le cadre de l'enseignement supérieur, une guidance stratégique, des conseils et des opportunités à l'utilisation des technologies de l'information et de la communication pour aider l'enseignement, l'apprentissage, la recherche et l'administration [<http://www.jisc.ac.uk>].

Chapitre 3. Des nouveaux modèles de communication scientifique

d'archivage assisté. Le projet TARDIS s'est appuyé sur les expériences pilotes menées par les *School of Ocean and Earth Sciences* et *School of Electronic and Computer Science* dont les bases de données seront finalement intégrées dans le dépôt [<http://tardis.eprints.org>].

Alors que le projet est lancé depuis un an, en août 2003, l'université de Southampton revoit et modifie sa structure générale en 3 “*faculties*” (*Faculty of Law, Arts and Social Sciences, Faculty of Medicine, Health and Life Sciences, Faculty of Engineering, Science and Mathematics*) et 20 écoles. Ces changements ne sont pas sans incidences sur le projet qui se voit obligé de s'adapter pour représenter au mieux la nouvelle structure de l'université [Hey 2004, pp. 2-3].

En plus des deux expériences pilotes, le développement d'un dépôt institutionnel à l'université de Southampton pouvait se baser sur deux éléments déjà disponibles : le dépôt thématique CogPrints [<http://cogprints.ecs.soton.ac.uk/>] et le logiciel d'archivage GNU EPrints [<http://software.eprints.org>]. Ce dernier a d'ailleurs servi au développement d'e-Prints Soton [<http://tardis.eprints.org>]. Il n'est donc pas étonnant que l'université de Southampton ait décidé de se doter d'un dépôt institutionnel.

2.3.2. Technique et fonctions

Fonctionnement général

La consultation des documents hébergés au sein d'e-Prints est ouverte à tous sans restriction, il n'est pas nécessaire de s'inscrire, ni de payer pour y accéder. En revanche, en tant que dépôt institutionnel, le dépôt de textes est uniquement autorisé aux membres de l'université de Southampton qui se

Chapitre 3. Des nouveaux modèles de communication scientifique

sont inscrits sur le site.

Ces chercheurs peuvent déposer tout type de documents : articles de périodique publiés ou non, chapitre de livre, texte de conférence,... ainsi que des documents multimedia. Pour le dépôt lui-même, deux solutions sont offertes. Le chercheur peut auto-archiver ses documents et prendre complètement en main leur dépôt via l'interface web, la rubrique “*help*” lui fournit toutes les informations pour ce faire. Il peut également choisir de transférer ses documents à un service d'e-Prints qui se chargera de les déposer dans l'archive.

Les avantages que le dépôt propose aux chercheurs sont de rendre leur travail de recherche plus visible et plus accessible ; de promouvoir leur travail mais aussi celui de leur communauté à l'intérieur de l'université ; d'offrir un espace de stockage sécurisé pour leur publication qui permet également de répondre aux requêtes *full text* et à la publication des données ; de contribuer aux initiatives globales et régionales qui ont pour but d'assurer une audience internationale aux dernières recherche.

E-Prints Soton est ouvert à toutes les disciplines scientifiques enseignées à l'université de Southampton. Dans la pratique, néanmoins, on se rend facilement compte, via l'option “*browse by subjects*”, qu'elles ne sont pas toutes représentées. Sur les 1 239 documents répertoriés dans les différentes matières selon la classification de la *Library of Congress*, on en retrouve 766 en océanographie (soit 62 %) et 364 en géologie (soit 30 %). Les 8 % restant se retrouvent principalement parmi les autres disciplines scientifiques (mathématique, biologie, chimie, ...). Des disciplines telles que la philosophie, l'histoire, le droit, les langues et littérature ne se retrouvent

Chapitre 3. Des nouveaux modèles de communication scientifique

pas dans le dépôt institutionnel. On peut ainsi se rendre compte aisément des différences qui existent dans la manière de diffuser le savoir scientifique d'une discipline à l'autre.

D'un point de vue technique, e-Prints Soton utilise le logiciel libre GNU EPrints [<http://software.eprints.org> ; <http://www.eprints.org>] pour la réalisation et le maintien de son dépôt institutionnel. L'utilisation de GNU EPrints rend le dépôt compatible avec les normes développées par l'*Open Archive Initiative*. Ce logiciel étant développé au département d'informatique et d'électronique de l'université de Southampton, on imaginait mal e-Prints Soton utiliser un autre produit. D'autant plus que la qualité de celui-ci est avérée par les nombreuses utilisations qui en sont faites de par le monde.

La recherche d'information

Le visiteur peut débiter sa recherche d'information via le formulaire "Quick Search" qui se trouve sur la page d'accueil. Cela lui permettra de rechercher de l'information parmi les auteurs, titres, mots-clés, date et résumé. Depuis cette même page d'accueil, il est possible de parcourir le dépôt institutionnel sur base de différents critères : sujets (présentés selon la classification de la bibliothèque du Congrès, les subdivisions qui ne concernent aucun documents du dépôt ne sont pas présentes) ; facultés, écoles ou autres groupements (sur base de l'organisation de l'université) ; années ; derniers ajouts au dépôt (sans distinction de sujet, type de document,...).

Enfin, deux autres façons de rechercher de l'information : la recherche simple ou avancée. La recherche simple est assez proche de la

Chapitre 3. Des nouveaux modèles de communication scientifique

recherche rapide, elle permet d'interroger en une fois les champs *full text*, titre, résumé, mots-clés, auteur, créateur, éditeur, date en précisant que les résultats doivent satisfaire à toutes ou n'importe quelles conditions. Il est également possible de les trier par titre, années (du plus récent au plus vieux ou inversement), noms d'auteur.

La recherche avancée offre, outre les champs “classiques”, “habituels” (*full text*, titre, auteur, créateur, éditeur, date, résumé, mots-clés, titre de publication, de revue) une recherche sur l'école, le centre d'où émane les documents, le sujet suivant la classification de la *Library of Congress* (toutes les subdivisions de cette classification sont présentées alors que certaines ne concernent aucun documents présents dans le fonds), le type d'e-print (article, livre, section de livre, textes de conférence ou de rencontre, monographie, brevet, thèse, autre) et son statut (publié, en impression, soumis pour publication, non publié).

On retrouve, en général, les mêmes moyens d'accéder à l'information dans les dépôts fonctionnant avec GNU EPrints, ceux-ci sont d'ailleurs indispensables. Les différences que l'on retrouve de l'un à l'autre montre aussi la modularité de ce logiciel et l'adaptation dont il peut faire l'objet par les gérants des différents dépôts. E-Prints Soton est un bon exemple des différentes possibilités de recherche disponibles avec GNU EPrints.

2.3.3. Situation actuelle

Les premiers dépôts de documents dans E-Prints Soton ont eu lieu aux alentours du mois de juillet 2003 et depuis, la quantité de documents disponibles n'a cessé d'augmenter. On peut notamment constater une importante accélération des dépôts durant la période qui va du mois de

Chapitre 3. Des nouveaux modèles de communication scientifique

décembre 2003 au mois de mars 2004, c'est à cette époque que le dépôt semble prendre son envol [<http://eprints.soton.ac.uk/information.html>]. Il aura fallu quelques mois avant que le dépôt ne soit régulièrement utilisé.

La période couverte par les documents s'étend de 1980 à nos jours. Comme cela est presque toujours le cas, les documents les plus anciens sont aussi les moins nombreux, sans doute à cause de la démarche rétrospective que leur dépôt dans l'archive implique. En moyenne, ce sont 50 documents qui ont été déposés par an mais les écarts par rapport à cette moyenne sont forts. En effet, pour la période de 1980 à 1995, ce sont en moyenne, trois documents qui ont été déposés pour chaque année, tandis que de 2001 à 2004, ce sont 264 documents¹⁴ [http://eprints.soton.ac.uk/view/date_effective].

Au niveau des domaines représentés au sein du dépôt, rappelons une absence presque totale des sciences sociales et humaines. Il serait d'ailleurs peut-être intéressant pour l'université de Southampton d'encourager le dépôt de documents dans ces domaines afin que l'archive institutionnelle soit un meilleur reflet de la recherche menée à Southampton. Il est vrai que les habitudes de communication scientifique sont différentes d'un domaine à l'autre mais il est peu probable qu'il n'y ait rien à diffuser électroniquement du côté des sciences sociales et humaines et que les membres de cette communauté n'aient aucun avantages à tirer de l'enrichissement d'un dépôt institutionnel.

D'autant que le dépôt individuel est plus facile à réaliser que, par

¹⁴ Notons que les chiffres pour l'année 2004 ne sont pas complets puisque l'année n'est pas terminée.

Chapitre 3. Des nouveaux modèles de communication scientifique

exemple, dans le cas de Caltech CODA (*Collection of Open Digital Archives*), il n'est pas nécessaire ici de créer un niveau supplémentaire (un dépôt ne concernant qu'un département, une école, une unité de recherche), le dépôt d'un document peut se faire immédiatement par le chercheur. E-Prints Soton est donc directement accessible à tous les chercheurs de l'université.

Dans la mise sur pied de son dépôt institutionnel, l'université de Southampton a pu bénéficier de deux avantages que d'autres institutions n'ont pas forcément eu. D'un côté, citons l'archive thématique CogPrints [<http://cogprints.ecs.soton.ac.uk>] concernant la psychologie et les neurosciences créée sur impulsion de Steven Harnad en 1997 [Gallezot 2002] considérée comme projet pionnier (avec arXiv et RePEc) dans le domaine des archives ouvertes. Du côté technique, c'est au sein de l'université de Southampton qu'est développé le logiciel GNU EPrints qui sert à la création d'archives en ligne, tant institutionnelles que thématiques. E-Prints Soton peut ainsi bénéficier des derniers développements de ce logiciel. Il est aussi presque obligatoire pour e-Prints Soton d'utiliser ce logiciel ; quelle serait l'image de GNU EPrints si l'université dans laquelle il est développé se permettait d'utiliser un autre outil ? Que se passerait-il si GNU EPrints n'était plus développé en continu et dépassé par un autre logiciel (par exemple : DSpace du MIT) ? Le lien qui unit e-Prints Soton à son logiciel comporte certes des avantages (rapide passage aux nouvelles versions, possibilité de retour, de commentaires sur le produit,...) mais aussi des inconvénients principalement celui de ne pas pouvoir aller voir ailleurs en cas de non-satisfaction.

Le lancement du dépôt institutionnel a été réalisé dans le cadre du

Chapitre 3. Des nouveaux modèles de communication scientifique

projet TARDis (*Targeting Academic Research for Deposit and Disclosure*) financé par le JISC (*Joint Information Systems Committee*). Par ces moyens de financement et l'inclusion dans le programme FAIR (*Focus on Access to Institutional Resources*), l'université s'inscrit dans le cadre d'un projet plus général concernant l'entièreté du Royaume-Uni et assurant une certaine cohérence entre les différents projets. Le financement d'initiatives de ce type est sans aucun doute un point fort pour la communication scientifique au Royaume-Uni mais que se passera-t-il si, à un moment donné, ces financements étaient stoppés ? Les projets en cours pourraient-ils continuer normalement ? La marche pour une amélioration de la diffusion des textes scientifiques ne serait-elle pas stoppée ? Et de là, ne serions-nous pas face à un retour en arrière ? Autant de questions et de situations qu'il est impératif d'envisager dès maintenant pour pouvoir y répondre au mieux si elles viennent à se poser un jour.

E-Prints Soton présente de nombreux avantages dont l'un des plus marquant est la simplicité avec laquelle il semble possible de déposer un document et de retrouver de l'information. Espérons qu'il puisse s'enrichir encore et profiter de ses liens avec les développeurs du logiciel GNU EPrints de manière à “coller” au mieux avec les attentes et besoins de ses utilisateurs.

Les dépôts institutionnels sont basés sur un principe assez simple : la collecte de textes provenant d'un même milieu, quel que soit son auteur, son domaine d'application et parfois, son type. A partir de ce principe, les trois exemples examinés nous ont montré que des variantes étaient possibles. Il va sans dire qu'il en existe bien d'autres, chaque université étant libre de

Chapitre 3. Des nouveaux modèles de communication scientifique

faire ce qu'elle veut de son dépôt. Ce type de projet est d'ailleurs de plus en plus développés dans les universités. Citons ici le projet BICTEL/e mené par les universités francophones de Belgique.

3. Les revues en ligne “Open access”

Troisième et dernier modèle envisagé, les revues en ligne regroupent une mine d'informations dont la qualité est souvent assurée par le processus de *peer review*. Grosse différence par rapport aux deux autres types de dépôt : l'aspect financier est bien plus souvent mis en avant.

3.1. BioMed Central : The Open Access Publisher

<http://www.biomedcentral.com>

BioMed Central, comme son nom l'indique est centré sur les sciences bio-médicale, secteur malmené par les éditeurs commerciaux classiques. Seul projet envisagé dans ce travail ayant un but lucratif affirmé, BioMed Central n'en est pas moins *Open Access* dans sa manière de donner accès à l'information. Par contre, la publication d'article est, elle, payante.

3.1.1. Historique

BioMed Central est “*une maison d'édition indépendante engagée à permettre un accès immédiat et gratuit aux travaux de recherche biomédicale revus par pairs*” [http://www.biomedcentral.com/info/about/whatis_fr]. Elle a été lancée au mois de mai 2000 et fait partie du Current Science Group spécialisé dans l'information et les services pour les professionnels du monde bio-médical. Avec une offre de plus de 100 périodiques au moment de son lancement (147 aujourd'hui), BioMed Central peut se poser comme leader de l'édition “Open Access” de son domaine [<http://current-science-group.com>]. Ces 147 périodiques sont répartis en 58 catégories [<http://www.biomedcentral.com/browse/bysubject/>] concernant le domaine des sciences bio-médicales au sens large.

Chapitre 3. Des nouveaux modèles de communication scientifique

Si la consultation des articles est gratuite et accessible à tous, la gratuité n'est plus de mise depuis 2002 du côté des auteurs. En effet, 525 \$ sont demandés aux auteurs pour la publication de leur article une fois qu'il a été accepté. Cette somme sert à payer les coûts de la publication scientifique. BioMed Central n'est donc *Open Access* que d'un "côté", celui de la consultation. Les auteurs, eux, doivent avoir les moyens de payer la publication de leur article. Heureusement, ceux-ci sont en général inclus dans les budgets de recherche.

3.1.2. Technique et fonction

Fonctionnement général

BioMed Central donne accès gratuitement à 147 revues en ligne concernant les différents domaines des sciences biologiques et médicales. L'intérêt de ces revues par rapport à d'autres produits de documentation scientifique disponibles en ligne est qu'elles bénéficient toutes d'un système de *peer review*. Les articles qu'elles proposent sont tous examinés par des professionnels.

Ces revues sont le fruit d'un travail collaboratif entre des groupes de scientifiques désirant mettre leur revue à disposition du public et BioMed Central qui s'occupe de la gestion technique et commerciale de la revue, et qui se pose comme son éditeur et propriétaire exclusif. Il est du ressort des créateurs de la revue de définir ses objectifs, son domaine d'application, ainsi que de composer un comité de rédaction et un comité de pairs qui assureront la qualité scientifique du contenu de la revue [http://www.biomedcentral.com/info/authors/startajournal_fr]. La création d'une revue n'est donc pas ouverte à n'importe qui, il est

Chapitre 3. Des nouveaux modèles de communication scientifique

indispensable de faire preuve de gages de qualité. Il est donc également possible pour une revue déjà existante et éditée par ailleurs, de migrer entièrement vers BioMed Central.

Le chercheur qui désire publier un article via BioMed Central doit le soumettre auprès d'une revue précise, et non, auprès de BioMed Central. La soumission d'un texte est gratuite ; celui-ci sera soumis au comité de rédaction et aux *reviewers* gratuitement. S'il est accepté pour publication, l'auteur devra en payer les frais. Ceux-ci s'élèvent à 525 \$ pour la plupart des revues mais, depuis juillet 2004 certaines revues¹⁵ ont augmenté leur prix et passent à 1 000 \$ voire 1 500 \$. Notons que l'auteur conserve, sous certaines conditions, ses droits sur le texte publié dans une revue de BioMed Central ; il ne les cède ni à la revue, ni à BioMed Central¹⁶.

Ces frais de publication peuvent ne pas incomber au chercheur si l'institution à laquelle il appartient adhère au programme de BioMed Central (ce programme est également ouvert aux entreprises, les conditions sont néanmoins quelque peu différentes). Il est possible pour une institution scientifique¹⁷ de s'inscrire moyennant une cotisation qui est fonction du nombre de chercheurs en sciences médicales et biologiques de l'institution. Une fois inscrite, ses chercheurs ne doivent plus payer pour

¹⁵ Genome Biology, BMC Biology, BMC Medicine, Journal of Biology
[<http://www.biomedcentral.com/info/authors/apcfaq>]

¹⁶ "Les auteurs restent libres de publier leurs articles sur leur page web, d'en imprimer autant de copies qu'ils le désirent et de les distribuer à leurs collègues, à condition toute fois que l'article soit correctement cité et BioMed Central mentionné" [http://www.biomedcentral.com/info/authors/reasons_fr]

¹⁷ Institutions belges inscrites à BioMed Central : "University of Ghent and University Hospital, KULeuven – Biomedical Library, Université Catholique de Louvain, Université de Liège" [<http://www.biomedcentral.com/inst/cou/56>].

Chapitre 3. Des nouveaux modèles de communication scientifique

publier un article et bénéficient de réduction pour l'accès aux services payant de BioMed Central. L'adhésion est renouvelable tous les ans et son prix dépend du nombre d'articles publiés durant l'année précédente (avec, néanmoins, un minimum de trois articles à payer, qu'ils aient été publiés ou non) [http://www.biomedcentral.com/info/about/instmembership_fr].

Ce système d'inscription des institutions ou entreprises ainsi que le paiement des publications par les auteurs sont à la base du financement de BioMed Central. Cela ne semble pourtant pas être suffisant puisque nous retrouvons, sur le site, plusieurs publicités, notamment une série de “*ads by google*” qui ont comme particularité de “se fondre” dans le site puisqu'elles en suivent la charte graphique et concernent des produits en relation avec le thème du site sur lequel elles se trouvent.

La recherche d'information

Plusieurs moyens sont accessibles aux utilisateurs qui cherchent une information. Il leur est possible de parcourir la liste des revues disponibles par domaines ou par ordre alphabétique. Cela leur donne directement accès aux articles publiés en format .pdf et aux métadonnées de l'article (titre, auteur(s), résumé, ...).

Il est également possible d'effectuer une recherche rapide sur l'ensemble des articles de BioMed Central, PubMed Central et PubMed. Une condition est cependant à remplir pour accéder aux résultats : il est indispensable d'être inscrit sur le site de BioMed Central. Cette inscription est gratuite et permet à BioMed Central de mieux connaître ses utilisateurs puisqu'un formulaire concernant les centres d'intérêt est à remplir. Une fois inscrit, l'utilisateur a également accès à la recherche avancée ainsi qu'à la

Chapitre 3. Des nouveaux modèles de communication scientifique

mailing list signalant la publication de nouveaux articles dans différents domaines au choix.

Cette recherche avancée offre deux possibilités : effectuer une requête sur quatre champs et l'affiner sur un intervalle de temps et/ou sur un périodique ; ou l'utilisateur peut effectuer une recherche booléenne et donc soumettre lui-même une équation de recherche qui peut aussi être limitée dans le temps et/ou sur un périodique.

La recherche avancée peut être définie sur un nombre important de champs différents ; on y retrouve les habituels et indispensables auteur, titre, abstract, texte, périodique, année de publication auxquels on peut ajouter le DOI (Digital Object Identifier), l'institution de l'auteur et, plus exceptionnellement, les différentes parties du texte c'est-à-dire : introduction, méthode, résultats, discussion/conclusion, légendes des illustrations et références.

La recherche d'information dans BioMed Central offre donc de nombreuses possibilités à l'utilisateur. L'inscription étant obligatoire pour effectuer une recherche, une fonctionnalité intéressante a été ajoutée : il est possible de stocker les équations de recherche. Au moment de l'enregistrement (et même par après), l'équation peut servir d'alerte mail à une fréquence choisie par l'utilisateur. Ces requêtes restent disponibles pour modification jusqu'à ce que l'utilisateur les supprime. Ajoutons qu'un historique de recherche est disponible dans l'interface de recherche avancée tant que la fenêtre du navigateur n'a pas été fermée, et ce, même si l'utilisateur a clôturé sa session dans BioMed Central.

3.1.3. Situation actuelle

Malheureusement, nous n'avons pas trouvé de statistiques concernant l'utilisation de BioMed Central par ses auteurs ou par ses visiteurs, ni même de statistiques individuelles pour les périodiques publiés. Pourtant, ces statistiques doivent vraisemblablement être réalisées puisqu'il est possible de consulter les articles les plus lus pour l'ensemble de BioMed Central et pour chaque périodique. De plus, il s'agit d'un genre de statistiques réellement utile pour la gestion de BioMed Central. Il semble donc simplement qu'elles ne soient pas accessibles au public (et il serait vain d'en imaginer les raisons).

Malgré l'absence de chiffres, il n'est pas difficile de constater, sur base de différentes lectures sur le site de BioMed Central et ailleurs, que l'offre, tant au niveau des services que du nombre de revues disponibles, ne cesse d'augmenter. De plus, un bref parcours de certains périodiques montre que les textes les plus récents sont réellement disponibles depuis peu de temps. Ces deux indices sont un indicateur du dynamisme de BioMed Central.

Ce dynamisme est peut-être à mettre en lien avec le fait que BioMed Central est un produit commercial d'un genre nouveau. La viabilité de ce modèle d'édition n'est pas encore assurée (mais, l'économie n'étant pas une science exacte, peut-on jamais se considérer comme assuré de quoi que ce soit ?). Contrairement à d'autres services étudiés précédemment, de l'argent est directement en jeu et il faut supposer que ses propriétaires agiront un maximum pour le garder et le faire fructifier. Cela n'exclut tout de même pas le risque de faillite.

Ajoutons que le public se montrera peut-être moins indulgent et

Chapitre 3. Des nouveaux modèles de communication scientifique

attendra plus de qualité, plus de services de la part de BioMed Central, éditeur commercial, qui, véritablement, vend un produit : la publication de périodiques. Et, comme les chercheurs qui publient sont aussi ceux qui accèdent à l'information, la présentation des documents ne peut pas être baclée sous prétexte qu'elle est gratuite et accessible à tous.

Cette obligation de qualité fait que BioMed Central offre un site internet riche et clair mais aussi des revues qui doivent toutes soumettre leurs articles à un véritable processus de *peer review*.

Le système de financement de BioMed Central et son aspect commercial n'est pas sans faille, en particulier pour les utilisateurs. La publication étant payante, comment doivent réagir les petits groupes de recherche et les petits projets scientifiques faisant partie d'une petite université non membre de BioMed Central ? D'autant que les tarifs de publication dans certaines revues ont augmenté drastiquement (de 525 \$ à 1 000 \$ ou plus). Si l'augmentation des prix continue, BioMed Central ne va-t-il pas se rapprocher des éditeurs commerciaux "classiques" tant critiqués, avec pour unique différence un placement des frais sur la publication plutôt que sur la consultation ? De plus, il faudra, à tout prix, que BioMed Central garantisse l'honnêteté de son principe de paiement ; les détracteurs du système pourraient très facilement le critiquer en faisant remarquer que, si le système dérappe, les articles publiés seront ceux dont les auteurs auront payé le montant le plus important.

L'utilisateur-consultant se voit offrir certaines garanties d'accès à la documentation. BioMed Central a mis au point une charte qu'il entend respecter et faire respecter en cas de changement de propriétaire, assurant

Chapitre 3. Des nouveaux modèles de communication scientifique

la pérennité de l'accès aux textes. De plus, BioMed Central est associé à deux initiatives orientées *Open Access* : SPARC (Scholarly Publishing and Academic Resources Coalition) et PubMed Central. Preuve d'un véritable intérêt pour l'*Open Access* ou manière de rassurer ses utilisateurs ? Dans les deux cas, ces partenariats impliquent certaines règles en matière d'accès à l'information et le respect d'une certaine "philosophie".

Entre *Open Access* et édition commerciale, BioMed Central associe deux logiques qui, de prime abord, peuvent sembler antagonistes, pour offrir un produit qui semble viable (jusqu'à présent) et satisfaisant pour un certain nombre d'utilisateurs du monde des sciences biologiques et médicales.

3.2. HighWire

<http://www.highwire.org>

Le deuxième exemple diffère du premier par deux points : il est issu du milieu universitaire et il n'a pas de but lucratif même si l'accès à l'information est majoritairement payant.

3.2.1. Historique

HighWire, tel qu'il se présente actuellement est le fruit d'un projet débuté en 1995 au sein de l'université de Stanford (Californie). Les missions à atteindre étaient de raccourcir le lien entre l'auteur de textes scientifiques et son lectorat spécialement dans le domaine scientifique, technique et médical, en utilisant les nouvelles technologies de réseau, en partenariat avec d'autres acteurs impliqués dans la même problématique. Il était également question d'assurer que le marché de la communication scientifique électronique ne se développe pas uniquement avec le concours des éditeurs commerciaux "classiques".

En 1995, HighWire débute avec la publication en ligne d'un seul périodique : le *Journal of Biological Chemistry* publié par la *American Society for Biochemistry and Molecular Biology* et considéré comme l'un des plus importants de son domaine. Il sera rejoint par après par les revues *Sciences* et *Proceeding of the National Academy (of Sciences of the United States of America)*. La publication de ces trois revues a servi de banc d'essai à la suite du projet et à la publication d'un nombre de revues bien plus important. L'intérêt de débiter le projet avec le *Journal of Biological Chemistry* résidait dans l'importance des volumes publiés (de 80 à 100 articles par semaine) et

Chapitre 3. Des nouveaux modèles de communication scientifique

dans la complexité du matériel (graphiques, formules mathématiques et notations chimiques, photographies,...). Ces deux “difficultés” ont rendu obligatoire la réalisation d'un système prêt à gérer de l'information complexe et en grande quantité.

En créant HighWire, Stanford se pose en partenaire et conseiller des sociétés scientifiques désirant publier leurs revues en ligne mais n'ayant pas les ressources nécessaires pour effectuer ce changement technologique. De plus, HighWire apporte beaucoup par rapport à l'édition sur support papier car il met en évidence les liens entre les auteurs, les articles, les citations. Il est possible de naviguer d'un article à l'autre via les citations et les auteurs qu'ils ont en commun. En cas d'accès payant à un article, seule la référence est accessible. Il rend possible la recherche avancée ainsi que la consultation d'images en haute résolution, interactives et de matériel multimédia. Aujourd'hui, HighWire produit 687 sites pour différentes revues dont la plupart offrent leur contenu gratuitement ou à moindre prix [<http://www.highwire.org/about/intro.dtl> ; http://www.highwire.org/about/original_info.dtl].

3.2.2. Technique et fonctions

Fonctionnement général

HighWire propose à ses visiteurs l'accès à plus de 650 périodiques et à leurs articles dans les domaines des sciences biologiques, médicales, physiques et sociales. Chacun de ces domaines est réparti en une série de subdivisions qui donnent un aperçu de l'étendue des sujets abordés. On ne trouvera pas dans HighWire, contrairement à Erudit, d'autres types de documents que des articles de périodiques.

Chapitre 3. Des nouveaux modèles de communication scientifique

Si HighWire est une porte d'accès à une grande quantité d'information, cet accès ne se fait pas toujours selon les mêmes modalités ; tout n'est pas gratuit et ce qui est payant ne l'est pas toujours de la même manière (bien qu'il s'agisse d'un éditeur sans but lucratif). Tout d'abord, HighWire propose deux types d'abonnement : institutionnel ou personnel (la demande d'abonnement se fait directement auprès du périodique concerné). Ensuite, sont gratuitement accessibles des périodiques entiers (23), des périodiques entiers mais pour un temps seulement (3 le 19 août 2004 mais il va sans dire que ce nombre varie) enfin, des périodiques (180) entiers mais pour lesquels seuls les anciens numéros sont gratuits. Ces derniers fonctionnent avec une barrière mobile qui rend inaccessibles sans souscription les x derniers numéros. Pour terminer, 298 périodiques sont accessibles via un système de *pay-per-view*. A l'utilisateur de choisir, parmi les offres, l'accès à un article pendant un certain laps de temps ou l'accès à la revue entière, pendant un certain temps également. Les prix de ces différents services varient d'une revue à l'autre. De 5 \$ à 28 \$ par jour pour l'accès à un article, de 30 \$ par jour à 30 \$ pour 30 jours pour l'accès à l'entièreté du périodique.

En plus d'accès aux articles, HighWire propose d'autres services à ses utilisateurs. Si un article cite une référence disponible au sein de HighWire, un hyperlien sera disponible pour atteindre directement le résumé de la référence. Ce service (*Inter-Journal Links*) offre ainsi un réel plus par rapport à une édition sur papier.

Autre plus intéressant, la visualisation d'une image peut se faire de trois manières différentes. Dans le texte, par défaut, une image est placée

Chapitre 3. Des nouveaux modèles de communication scientifique

en petit format (*thumbnail*), cela permet une lecture fluide du texte et un chargement plus rapide de la page. Il est ensuite possible, en cliquant sur l'image, de l'agrandir pour obtenir la taille normale de l'image ou pour visualiser l'image en haute résolution. La possibilité est parfois offerte d'ouvrir les images dans une nouvelle page du navigateur. Ces possibilités de visualisation sont particulièrement intéressantes, voire même indispensable, pour l'imagerie médicale.

Une fois l'utilisateur inscrit gratuitement, d'autres services lui sont offerts, notamment des alertes par email ou sur son PDA (*Personal Digital Assistant*). Deux systèmes d'alertes sont disponibles : eTOCs prévient de l'arrivée de nouveautés dans certains périodiques choisis (ou dans l'ensemble d'HighWire) et CiteTracks annonce les nouveautés correspondant à certains critères (de sujet, d'auteur,...) définis par l'utilisateur.

L'inscription permet également d'établir une sélection de périodiques (*My Favorite Journals*) raccourcissant ainsi le chemin d'accès à ceux-ci, d'avoir une vue générale des périodiques (*My Access*) auxquels on a accès soit gratuitement, soit via les abonnements personnels ou institutionnels. Il est également possible d'ouvrir une fenêtre (*My Site Bar*) d'accès direct aux périodiques favoris et aux outils de recherche de HighWire. Trois nouveaux services sont prévus : *My Topics* permettra de sélectionner les articles sur base de sujets définis par l'utilisateur, celui-ci pourra, en quelque sorte, se créer son propre journal ; *My Folders* donnera accès aux articles et requêtes utilisées auparavant et conservées par l'utilisateur ; et enfin, *My Searches* conservera les préférences de recherche de l'utilisateur.

Seules les revues assurant le *peer review* de ses articles peuvent être

Chapitre 3. Des nouveaux modèles de communication scientifique

publiées via HighWire. Ceci, ainsi que les différentes fonctions d'édition, sont de la responsabilité de l'éditeur de la revue et non de celle de HighWire. Ce dernier ne s'occupe que de la technique de la diffusion en ligne et bien sûr, d'une part de la publicité. En faisant sa propre promotion, HighWire fait aussi celle des revues qu'il héberge. Il aide les organisations qui le désirent et qui paient, à développer la présence de leur périodique sur le web. Il doit également assurer la transition des abonnements à l'édition papier (si elle existe) et donner accès aux personnes qui y ont droit. Pour effectuer ce travail, HighWire propose aux éditeurs de périodiques différents outils plus ou moins avancés. C'est à l'éditeur de faire le choix des options qu'il va offrir ou faire payer à ses lecteurs de manière à correspondre au mieux à la demande de ce public. Aucun prix de ces services n'est disponible directement sur le site mais, HighWire n'étant pas une entreprise commerciale mais une *“non-for-profit division of a highly ranked research university”* nous pouvons supposer que les tarifs pratiqués ont d'abord pour objectif le maintien du site et des services de HighWire. L'offre de service que fait HighWire a ses futurs partenaires est modulaire et semble pouvoir être adaptée aux besoins et aux moyens financiers des éditeurs de périodiques. A un service de base (hébergement, outil de recherche, gestion des méta-données,...) il est possible d'ajouter des fonctionnalités plus avancées (gestion entièrement électronique des manuscrits, diffusion des résumés sur les PDA's des lecteurs, conversion rétrospective d'articles hébergés sur d'autres sites web,...).

L'utilisateur final, le visiteur du site se voit également proposer un grand nombre de services qui sont la preuve que HighWire a su tirer profit

Chapitre 3. Des nouveaux modèles de communication scientifique

des technologies de l'information pour donner à son produit un véritable avantage par rapport à la publication sur support papier.

La recherche d'information

Dès la page d'accueil, le visiteur se voit offrir deux moyens d'accéder à l'information : par parcours et par recherche. Le parcours de l'information peut se faire sur les articles ou les revues par sujet, sur les revues en ordre alphabétique, sur les éditeurs ainsi que sur d'autres listes (revues avec *pay-per-view*, avec abonnement, *back issues*, revues accessibles gratuitement depuis certains pays "défavorisés", ...). L'outil *TopicMap* est un applet java s'ouvrant dans une nouvelle fenêtre de navigateur qui offre une manière visuelle de parcourir les articles par sujets. Il se présente sous forme d'arbre reprenant les différentes matières, de la plus générale à la plus particulière. L'utilisateur peut très simplement faire apparaître les sous-divisions qui l'intéressent en double-cliquant sur une rubrique, la liste des articles regroupés sous celle-ci s'affichera dans le navigateur.

La recherche d'article proprement dite peut être rapide ou avancée. La recherche rapide est accessible sur la page d'accueil. La requête peut être effectuée sur les champs auteur, mots-clés, années, volumes, pages et limitée aux périodiques favoris de l'utilisateur identifié, à l'ensemble d'HighWire ou encore à l'ensemble d'HighWire et de Medline¹⁸.

La recherche avancée peut porter sur les mêmes sources d'informations que la recherche rapide mais il est en plus possible de définir

¹⁸ Medline est la base de données de la bibliothèque nationale de médecine aux Etats-Unis, elle couvre les domaines de la médecine, le nursing, la dentisterie, la médecine vétérinaire, le système des soins de santé et les sciences "pré-cliniques" [http://www.nlm.nih.gov/databases/databases_medline.html].

Chapitre 3. Des nouveaux modèles de communication scientifique

une sélection de périodiques sur lesquels la requête sera effectuée et spécifier si on s'intéresse à tous les articles ou uniquement aux articles de fonds. Les champs disponibles pour effectuer cette requête sont les classiques auteur (deux fois), titre, résumé/titre, *full text*/résumé/titre. Comme c'est en général possible ailleurs, l'utilisateur peut définir une période sur lequel sera effectuée sa recherche. Un historique de la recherche est affiché jusqu'à la fermeture de la fenêtre du navigateur.

Les possibilités de recherche de HighWire sont larges. Ajoutons à cela qu'une rubrique d'aide à la recherche est disponible, elle présente et explique les différentes techniques disponibles pour retrouver un article et donne quelques conseils généraux d'utilisation de ces techniques. Le moins que l'on puisse dire c'est que l'utilisateur d'HighWire a tout les atouts pour mener à bien sa recherche.

3.2.3. Situation actuelle

Encore une fois, les statistiques qui permettraient de juger l'utilisation de HighWire ne sont pas disponibles. Il est possible pour certaines revues de consulter les articles les plus lus et/ou les plus cités mais cela ne donne aucune vue générale et ne permet même pas d'établir une comparaison entre revues. Nous pouvons néanmoins constater que le nombre de revues hébergées par HighWire augmente, de nouvelles seront d'ailleurs bientôt disponibles. HighWire ne semble pas stagner mais au contraire, ne fait que s'agrandir, tant au niveau de la documentation proposée que des services disponibles.

Les outils mis à disposition des visiteurs sont en effet très variés et clairs (tant pour le site en lui-même que pour les services proposés).

Chapitre 3. Des nouveaux modèles de communication scientifique

HighWire propose ainsi un vrai plus par rapport à l'édition papier mais aussi par rapport à d'autres produits d'archives ou de revues en ligne, le *TopicMap*, qui donne, graphiquement, un excellent aperçu des matières disponibles, en est un bon exemple.

HighWire est un produit de l'Université de Stanford, aujourd'hui encore, leur collaboration est étroite. Un projet de recherche a d'ailleurs été mené sur les comportements des utilisateurs des revues hébergées par HighWire [<http://ejust.stanford.edu>]. Le projet eJust va permettre à HighWire d'adapter son offre aux demandes et attentes des utilisateurs. Par exemple, pour le moment, différents moyens de paiement sont disponibles selon les revues. Il n'est pas exclu qu'un jour l'un d'eux soit privilégié et appliqué à toutes les revues payantes.

HighWire ne s'inscrit pas parmi les initiatives *Open Access* les plus connues, il ne respecte pas le protocole proposé par l'*Open Archive Initiative* et n'est donc pas interopérable. Sur base de l'expérience menée par HighWire, un manifeste pour un meilleur accès à l'information scientifique a été écrit. Le *Washington DC principles for free access to science* propose un modèle qui se situe entre l'accès totalement libre et gratuit à la documentation scientifique et l'édition commerciale "classique". La plupart des éditeurs des revues d'HighWire ont signé ce texte qui les engage à continuer de favoriser un large accès à la documentation scientifique, et plus particulièrement dans le domaine médical [<http://www.dcprinciples.org/statement.htm>].

HighWire propose donc une grande quantité de services et d'information dans des domaines qui sont très proches de ceux disponibles

Chapitre 3. Des nouveaux modèles de communication scientifique

au sein de BioMed Central. Ces deux initiatives se posent donc sur un même marché (l'information de pointe en sciences biomédicales) avec un modèle de financement quelque peu différent même si dans les deux cas, les utilisateurs paient. Et si sur le site de BioMed Central, nous retrouvons des informations, des liens vers d'autres services *Open Access* dont HighWire, sur le site de ce dernier, dans la section "*Selected web portal for scientists*" qui propose des liens susceptibles d'intéresser les chercheurs, BioMed Central n'est pas cité. Nous pouvons difficilement envisager un tel oubli de la part de HighWire ce qui nous amène à penser que les deux projets sont bien en concurrence.

HighWire, avec toutes ses revues et services, ses différents moyens d'accès à la documentation payante et ses liens très étroits avec l'université de Stanford a beaucoup d'avantages à proposer à la communauté scientifique des sciences bio-médicales au sens large.

3.3. Erudit : Promouvoir et diffuser la recherche universitaire

<http://www.erudit.org>

Erudit est le dernier exemple analysé de ce travail. Considérant principalement les sciences humaines et sociales, il se différencie de BioMed Central et HighWire par la quantité moindre d'information disponible. Il est le fruit d'une collaboration étroite entre trois universités et le secteur public québécois.

3.3.1. Historique

Erudit propose un accès à des revues en ligne se situant

Chapitre 3. Des nouveaux modèles de communication scientifique

principalement dans le domaine des sciences humaines et sociales (économie, anthropologie, histoire,...) ainsi que, dans une moindre mesure, dans celui des sciences naturelles (géographie, entre autres). À côté de ces revues, Erudit donne également accès à quelques livres, actes, thèses et autres types de documents comprenant des articles et des rapports.

Le projet Erudit a été envisagé dès 1997 au sein des Presses de l'Université de Montréal. Il a débuté en 1998 avec pour objectif la diffusion numérique de revues. Cet objectif envisageait trois dimensions : l'expérimentation de la publication de 1998 à 1999 de cinq revues pour mettre à l'épreuve la chaîne de traitement numérique (tout en conservant la publication sur papier) ; la démonstration de l'intérêt, de l'utilité et du bien fondé de la mise en place d'un centre pour la publication de revues scientifiques en ligne ; la prise en compte et la découverte des enjeux de la publication scientifique numérique [<http://www.erudit.org/documentation/rapport/chap2/chap2.htm>].

En 2000, le projet aboutit à la création d'un groupe interuniversitaire pour l'édition numérique ou "consortium Erudit", composé de trois universités québécoises : l'Université de Montréal, l'Université du Québec à Montréal et l'Université Laval [Iacovella 2004, p.2]. En 2001 et 2002, le projet met sur pied un processus pour la publication, non seulement de revues mais aussi de livres numériques. Dans le même temps, le consortium se voit confier la mission de créer une plateforme pour la diffusion de ces produits numériques. Cette plateforme sera mise en service en octobre 2002 sous le nom d'Erudit.

Si le projet est techniquement réalisé par les universités québécoises,

Chapitre 3. Des nouveaux modèles de communication scientifique

celles-ci ne sont pas les seuls intervenants dans le projet puisqu'il est financièrement soutenu par le fonds FCAR¹⁹ de 1998 à 1999 et par le Fonds de l'autoroute de l'information de 2001 à 2002. Actuellement, Erudit bénéficie du soutien du Fonds québécois de recherche sur la société et la culture, du Fonds de l'autoroute de l'information, du Ministère du Développement économique et régional, du Fonds québécois de recherche sur la nature et les technologies et du Fonds de recherche en santé du Québec [<http://www.erudit.org/info.html>].

Nous le voyons, Erudit est le fruit d'une collaboration étroite entre le monde universitaire québécois et différentes institutions de la région qui encouragent la recherche par leurs subventions.

3.3.2. Technique et fonctions

Fonctionnement général

Erudit propose d'emblée plusieurs sections présentant les types de document que l'utilisateur désire consulter : revues, livres et actes, thèses, “autres documents et données”. Il est également possible d'accéder directement à l'interface de recherche.

Les revues disponibles sont actuellement au nombre de 37 et cinq de plus seront bientôt diffusées sur le site d'Erudit. Les domaines concernés par ces revues sont variés et touchent principalement aux sciences humaines

¹⁹ Fonds pour la Formation de Chercheurs et l'Aide à la Recherche qui stoppe ses activités en 2001 suite à la nouvelle “politique québécoise de la science et de l'innovation” et dont la relève est assurée par le Fonds de recherche sur la nature et les technologies, le Fonds québécois de la recherche sur la société et la culture et le Fonds de recherche en santé du Québec [<http://www.fcar.qc.ca/nateq/fonds.htm>].

Chapitre 3. Des nouveaux modèles de communication scientifique

et sociales. Quelques revues concernent, elles, les sciences naturelles : *Cahiers de géographie du Québec*, *Géographie physique et Quaternaire*, *Médecine/Sciences*, *Phytoprotection*.

Pour assurer la survie d'Erudit, tous les numéros des revues ne sont pas accessibles gratuitement. Un système de barrière mobile a été mis au point. Les abonnés payant ont accès à la totalité des numéros tandis qu'on ne trouve en accès direct et gratuit que les anciens numéros (ceux-ci ne sont pas pour autant dépourvus d'intérêt). Quand un nouveau numéro de la revue paraît, le nombre de numéro en accès libre augmente de manière à barrer l'accès aux x derniers numéros parus. Cette barrière mobile est spécifique à chaque revue et varie de l'une à l'autre. Les revues *Etudes françaises* et *Etudes internationales*, par exemple, autorisent la consultation de certains numéros de l'année en cours. Par contre, les numéros de la revue *Géographie physique et Quaternaire* ne sont disponibles que jusqu'en 2001.

Si, pour certaines revues, la présence de la barrière ne laisse un accès qu'à peu de numéros, d'autres revues ont fait l'objet d'une numérisation rétrospective et peuvent, comme les *Cahiers québécois de Démographie*, proposer des numéros datant de 1989 ou comme *Meta : Journal des traducteurs* des numéros de 1966.

Les articles sont parfois disponibles en différents formats (.pdf et .html), parfois en un seul format (.pdf). C'est le cas pour les anciens articles qui ont été numérisés et pour lesquels une version .html n'a pas été réalisée.

Comme BioMed Central, Erudit ne fait que proposer une plateforme technique pour l'hébergement de revues et n'intervient pas dans le

Chapitre 3. Des nouveaux modèles de communication scientifique

fonctionnement interne de celles-ci. Le mode d'acceptation d'articles, la soumission à un *peer review* reviennent donc aux différentes rédactions.

Jusqu'à présent, Erudit propose treize livres téléchargeables gratuitement par les internautes. Tous sont issus d'une manière ou d'une autre des trois universités initiatrices du projet. Le seul ouvrage non édité directement par une des maisons d'éditions des universités, l'est par un de leur professeur²⁰.

La publication de livres par Erudit n'a pas comme seul objectif la simple mise à disposition gratuite d'ouvrages. Erudit permet aux éditeurs d'expérimenter la publication en ligne et aux chercheurs universitaires, l'étude des usages et nouveaux services pour le lectorat.

Certains livres sont disponibles aux formats Open eBook et Microsoft Reader permettant le téléchargement et la lecture au moyen d'un livre électronique. D'autres ouvrages (par exemple : *Comprendre la famille, actes des symposium québécois de recherche sur la famille*) ne sont disponibles qu'au format .pdf.

Erudit propose également une structure d'accueil pour les thèses. N'importe quelle université peut demander à utiliser ce service mais jusqu'à présent, aucune thèse n'est disponible. Il est vrai que de plus en plus d'universités mettent sur pied leurs propres dépôts de thèses.

²⁰ Internet et la Démocratie est édité par Denis Monière, professeur à l'Université de Montréal.

Chapitre 3. Des nouveaux modèles de communication scientifique

La page des thèses d'Erudit, se pose dès lors en point de départ de la recherche en fournissant quatre liens : les thèses en ligne des universités de Montréal et Laval, NDLTD et Cyberthèses²¹, . Il ne s'agit donc que d'un point d'accès.

La section d'Erudit concernant les “autres types de documents” fonctionne sur le même principe que celle des thèses : Erudit propose un espace et une structure d'accueil aux institutions qui souhaiteraient diffuser leurs documents en ligne. Le service n'est pas uniquement réservé aux universités québécoises qui sont à l'origine d'Erudit, il est ouvert à tous et gratuitement. Pour assurer ce service, Erudit utilise le logiciel DSpace développé au MIT.

Jusqu'à présent, douze unités de recherche utilisent Erudit et diffusent ainsi des articles, des pre-prints, des rapports de recherche, ... Elles totalisent ainsi 844 documents²².

Les différents services rendus par Erudit ne bénéficient pas tous du même succès, l'utilisation des uns et des autres semble variée. Si les revues et les “autres documents” fournissent aux utilisateurs une quantité de données relativement importante, il n'en va pas de même pour les thèses et les livres. Mais ce dernier média se prête sans doute moins à la numérisation et à la diffusion en ligne, d'autant plus que les livres électroniques n'ont pas rencontré un grand succès.

21 Projet mené par l'Université de Montréal et l'Université Lumière Lyon 2 référençant et reprenant les liens hypertextes vers les thèses des institutions participantes [<http://www.cybertheses.org>]

22 Données au 18 août 2004

Recherche d'information

La page d'accueil d'Erudit donne un moyen d'accéder directement à la recherche d'information. Cet outil ne permet d'effectuer des requêtes que sur le contenu des revues, cela n'est pas précisé de manière très claire à l'utilisateur. Il est possible d'effectuer une recherche simple basée uniquement sur des mots-clés. La recherche plus avancée permet d'utiliser les champs titre, auteurs, tous champs sauf texte et de limiter les résultats à un intervalle de temps et/ou à la publication dans une ou plusieurs revues.

Une fois la recherche effectuée, il suffit de cliquer sur le titre de l'article pour le consulter, les autres résultats éventuels restent disponibles.

Si aucune recherche n'est possible sur les livres, ce n'est pas le cas en ce qui concerne les “autres documents et données”. Pour chaque archive d'unité, il est possible d'effectuer une recherche spécifique simple sur l'auteur, le titre ou la date. la recherche avancée permet aussi de choisir un ou plusieurs dépôts comme source d'information et peut s'effectuer sur les index des auteurs, titres, résumés, collections, identificateurs ou tous ces champs à la fois mais seulement trois mots-clés sont autorisés. Ils peuvent être liés par des opérateurs booléens (et, ou, sauf).

Les possibilités de rechercher de l'information dans Erudit ne sont pas très riches. Dans aucun cas il n'est possible d'effectuer une recherche *full text* ni une recherche sur l'ensemble du fonds documentaire. Il est dommage que l'utilisateur doive répéter plusieurs fois sa requête pour être sûr d'obtenir toute l'information disponible. Néanmoins, même si ces possibilités de recherche ne sont pas larges, cela ne signifie pas pour autant qu'il est impossible de retrouver une information.

3.2.3. Situation actuelle

Comme c'est également le cas pour d'autres services présentés ici, Erudit n'offre pas de statistiques. Et, s'il est aisé d'évaluer la quantité de documents disponibles, il est impossible de savoir dans quelle mesure ceux-ci sont exploités. Les treize unités de recherche proposent 844 documents et données diverses, treize livres sont téléchargeables et 37 revues sont consultables (le nombre de numéros disponibles gratuitement ou non varie de l'une à l'autre).

Un des intérêts d'Erudit est que, à côté des revues et de leurs articles, on trouve également d'autres types de documents (monographies, actes de colloques, rapport de recherche, ...) concernant des domaines parfois très variés. Malheureusement, le visiteur pourrait être déçu. Sur la page d'accueil, la section thèse se situe au même niveau que les autres et on s'attend à trouver un fonds de thèse aussi riche que celui des revues. Or Erudit n'héberge aucune thèse, ce manque est en partie compensé par la présence des différents liens.

Il est également précisé que l'objectif d'Erudit n'est pas la simple mise en ligne de documents. Si, comme cela est dit sur le site, il y a effectivement des chercheurs qui s'intéressent à l'utilisation d'Erudit, il faut espérer que ceux-ci utiliseront leurs observations pour améliorer les services d'Erudit ou ceux des trois universités participant pleinement au projet.

Le site du projet manque parfois de clareté. L'outil de recherche est limité aux revues mais cela n'est pas indiqué et la confusion augmente encore puisqu'il est accessible depuis toutes les pages du site. Les derniers numéros des revues ne sont accessibles que sur abonnement mais, là aussi,

Chapitre 3. Des nouveaux modèles de communication scientifique

ce n'est pas précisé de manière évidente ; l'utilisateur ne peut voir que les numéros auxquels il a accès mais pas ceux pour lesquels il doit payer. Le visiteur a donc intérêt à se renseigner d'abord sur le fonctionnement d'Erudit, à faire un bref tour du site pour prendre conscience des opportunités qui lui sont offertes.

Erudit est ouvert à tous dans le monde mais une visite du site montre que la plupart des documents proposés proviennent des universités québécoises. Les universités de Montréal et Laval ont toutes deux leurs propres dépôts de thèses et développent pour elles-mêmes leurs projets de dépôts institutionnels²³. Les universités étrangères préféreront peut-être, elles aussi, développer leurs propres outils car si elles en ont les moyens, avoir son propre dépôt institutionnel est plus prestigieux et est la preuve d'un certain savoir faire. Erudit pourrait, dès lors, se passer de sa section “thèses” (ou éventuellement, l'enrichir en ajoutant des liens vers d'autres dépôts de thèses en ligne), continuer d'enrichir son fonds de revues ou, pourquoi pas, se centrer uniquement sur le monde québécois de manière à conserver une certaine uniformité, un point commun entre les documents proposés. Une autre idée serait de réellement enrichir la section des thèses et proposer en dépôt institutionnel commun aux trois universités participantes mais, comme celles-ci (l'UQAM exceptée) ont déjà toutes un projet en route, il y a donc peu de chance qu'Erudit puisse jouer ce rôle. Si les universités initiatrices développent chacune leur propre dépôt institutionnel avec succès, il y a peu de chance que les services “thèses” et

23 Archimède à l'Université Laval [<http://archimede.bibl.ulaval.ca>] et Papyrus à l'Université de Montréal [<http://papyrus.bib.umontreal.ca>]

Chapitre 3. Des nouveaux modèles de communication scientifique

“autres documents et données” d'Erudit subsistent ou du moins, s'agrandissent encore.

4. Tableau récapitulatif

	Initiateur	Situation géographique	Type de documents	Quantité de documents ²⁴	Domaines abordés
Dépôts thématiques					
arXiv	P. Ginsparg, Laboratoire de Los Alamos	Etats-Unis	Articles (pré- et post-prints)	288 352 soumissions	Physique, mathématique, sciences non-linéaires, informatique, biologie quantitative
@rchive SIC	G. Gallezot (Université de Nice), G. Chartron (Institut National de Recherche Pédagogique), J.-M. Noyer (Université Paris 7)	France	Articles publiés ou en cours de publication, working papers	370 documents	Sciences de l'information et de la communication
RePEc	T. Krichel (université de Loughborough)	Projet décentralisé	Articles, working papers, logiciel, information sur des livres, chapitres de livres, auteurs, publications, institutions	284 000 éléments	Sciences économiques
Dépôts institutionnels					
Caltech CODA	California Institute of Technology	Etats-Unis	Livres, thèses, articles, rapports de recherche,...	2 331 documents ²⁵	Toutes les disciplines enseignées à Caltech
MIT Theses	Massachusetts Institute of Technology	Etats-Unis	Thèses et mémoires	5 600 documents (en août 2004)	Toutes les disciplines enseignées au MIT

²⁴ Chiffres au 8 septembre 2004

²⁵ Approximation depuis le site d'oaister

[<http://oaister.umdl.umich.edu/o/oaister/browse.html>]

Chapitre 3. Des nouveaux modèles de communication scientifique

	Initiateur	Situation géographique	Type de documents	Quantité de documents	Domaines abordés
e-Prints Soton	University of Southampton	Royaume-Uni	Articles, chapitre de livres, textes de conférence, multimédia,...	1 328 documents	Toutes les disciplines enseignées à l'université de Southampton
Revues en ligne					
BioMed Central	Current Science Group	Royaume-Uni	Revues complètes et leurs articles	151 revues	Biologie et médecine
HighWire	Université de Stanford	Etats-Unis	Revues complètes et leurs articles	684 revues	Biologie, médecine, physique et sciences sociales
Erudit	Université de Montréal, Université du Québec à Montréal, Université Laval	Canada	Revues, (livres, thèses, "autres documents")	37 revues, 13 livres	Sciences humaines, sociales et naturelles

	Financement	But lucratif	Gratuité	Lien avec un fédérateur (autre que OAI)	OAI - PM H
Dépôts thématiques					
arXiv	Public (National Science Foundation) et universitaire	Non	Oui	Aucun	Oui
@rchive SIC	Public (CCSD/CNRS)	Non	Oui	CCSD (Centre pour la communication scientifique directe)	Oui
RePEc	Décentralisé, public (JISC) à l'origine	Non	Dans la plupart des cas	Aucun	Oui
Dépôts institutionnels					
Caltech CODA	Universitaire	Non	Oui	Caltech ETD participe à NDLTD	Oui
MIT Theses	Universitaire	Non	Accès aux images : oui, version .pdf ou papier : non	NDLTD	Non

Chapitre 3. Des nouveaux modèles de communication scientifique

	Financement	But lucratif	Gratuité	Lien avec un fédérateur (autre que OAI)	OAI - PM H
e-Prints Soton	Public (JISC)	Non	Oui	Aucun	Oui
Revue en ligne					
BioMed Central	Privé (Current Science Group et publication d'article)	Oui	Accès : oui, publication : non	Supporte SPARC et NDLTD	Oui
HighWire	Universitaire et abonnements	Non	Dans certains cas	Washington DC Principles for Free Access to Science	Non
Erudit	Public, universitaire et abonnements	Non	Dans certains cas	Aucun	Non

Ces deux tableaux se complètent pour donner une vue d'ensemble des différents projets analysés. Le premier tableau met en avant des données concernant l'état des exemples considérés, le second présente des données en rapport avec l'*Open Access* et ses aspects financiers. Les conclusions, comparaisons et remarques que nous pouvons établir sont propres aux exemples choisis, il est inutile de vouloir les appliquer à l'ensemble des projets du même type existant par ailleurs.

Chacun des modèles envisagés présente des avantages et des inconvénients, tous n'ont pas les mêmes fonctionnalités, ni les mêmes objectifs. Les dépôts thématiques centralisent une large quantité de documents de tout type sur un même sujet mais n'assurent pas de *peer review*. Ce qu'ils compensent avec un système de parrainage comme dans arXiv ou une évaluation par un spécialiste comme dans @rchiveSIC.

Les dépôts institutionnels ont l'avantage de donner une certaine visibilité à des documents difficilement consultables : les travaux de fins d'études. Ceux-ci, s'ils ne sont pas soumis au *peer review* sont évalués par

Chapitre 3. Des nouveaux modèles de communication scientifique

les professeurs. Malheureusement, aucun des dépôts considérés n'a mis en place une politique d'archivage des thèses et mémoires obligatoire. Et, si ces dépôts ont une fonction de prestige, ils ne peuvent néanmoins pas être considérés comme entièrement représentatifs de l'activité scientifique de l'université à laquelle ils appartiennent.

Enfin, les revues en ligne assurent, à peu près, les mêmes fonctions que leurs homologues sur papier et mettent en oeuvre l'arbitrage par les pairs. Leur gratuité est néanmoins réduite, soit au niveau de la lecture des articles (HighWire et Erudit), soit au niveau de leur publication (BioMed Central). Nous pouvons le constater, les moyens de paiement sont variés et chacun des trois exemples utilise un système de financement différent.

La question que nous pouvons nous poser à présent est celle de la durabilité de ces différents projets. Nous ne pouvons pas, bien évidemment, prédire l'avenir mais certains éléments peuvent se poser en indice d'une certaine pérennité. Tout d'abord, la quantité d'information recueillie dans certain projet, il est peu probable que les 288 352 articles d'arXiv, les 284 000 éléments de RePEc, les 684 revues de HighWire, ... ne soient un jour, plus disponibles nulle part. Et, en cas de difficulté, il est du devoir de ces projets d'assurer l'accès à l'information qu'ils contiennent. Le soutien de projets fédérateurs, l'adhésion aux textes promotionnant l'ouverture de l'information scientifique peuvent se poser en garants de la pérennité de l'accès à ces documents

La maturité de certains projets tels qu'arXiv est un autre indice. De nombreux physiciens, mathématiciens, biologistes et informaticiens utilisent régulièrement ce dépôt, tant pour y déposer que pour y chercher de

Chapitre 3. Des nouveaux modèles de communication scientifique

l'information.

Nous pouvons également noter l'investissement du secteur public, notamment en France avec le CNRS et le CCSD, en Grande-Bretagne avec le JISC et au Québec avec divers fonds supportant la recherche scientifique. Bien sûr, les universités sont également une forme de financement public indirect.

Celles-ci apportent beaucoup aux différents projets. En matière d'expertise et en tant qu'acteur principal de la communication scientifique, il est impensable pour les différents projets de se passer de leur implication. Ajoutons qu'un dépôt institutionnel fonctionnel et dynamique est un atout important pour une université. La diffusion des résultats de la recherche, en plus d'améliorer la communication scientifique, peut augmenter le prestige d'une université.

Tous ces éléments ne sont que des indices, il est impossible de prévoir le sort de ces nouveaux modes de communication scientifique. Leur avenir dépend certainement de l'implication des chercheurs, de l'utilisation qu'ils feront de ces produits. Plus ils seront utilisés tant pour le dépôt de documents que pour la recherche d'information, plus il y a de chance qu'ils subsistent. C'est au monde scientifique de leurs créer une place au sein de son système de communication et d'évaluation.

Conclusion

Après avoir, en quelques mots, dressé le portrait de la situation catastrophique dans laquelle se trouvent les bibliothèques de recherche, nous avons présenté quelques uns des projets qui encadrent et font la promotion des initiatives pour une meilleure diffusion de l'information scientifique.

Ces fédérateurs se sont développés peu après l'apparition des premiers services *Open Access* et continuent de jouer un rôle important, notamment dans la sensibilisation des acteurs.

Nous sommes ensuite passés à l'analyse de neuf exemples représentant trois modèles : les dépôts thématiques, les dépôts institutionnels et les revues en ligne. Nous avons vu que chacun, sur des modes différents, proposent aux scientifiques de nouvelles méthodes pour diffuser leurs résultats. Nous avons également pu constater qu'au sein d'un même modèle, des variantes sont présentes. Chaque projet met en avant et développe ce qui lui semble le mieux convenir à la communauté qu'il sert. Tous présentent donc des avantages et des inconvénients.

Malgré toutes ces différences, il existe néanmoins un point commun à tous ces projets : ils n'existent et ne pourront continuer d'exister qu'à la condition que les chercheurs et l'ensemble du monde scientifique et académique les utilisent pour leurs recherches et les alimentent de leurs écrits.

Ces projets ont modifié le comportement de certains chercheurs mais

Conclusion

qu'en est-il de celui des éditeurs commerciaux ? Le projet SPARC est d'augmenter la concurrence de manière à faire baisser les prix, a-t-il eu des effets ? Il est sans doute encore tôt pour mesurer les effets de ces projets sur les éditeurs commerciaux mais certains ont déjà réagit. Depuis peu, Elsevier autorise ses auteurs à archiver leurs articles dans le dépôt de leur institution et d'ainsi le rendre disponible en libre accès. Cet auto-archivage n'était pas réalisable auparavant car l'auteur, en général, cédait l'entièreté de ses droits sur l'article à son éditeur.

L'*Open Access* semble être, pour le moment, un sujet dont il est beaucoup question. Cette année, en Belgique, la revue *Louvain* et les *Cahiers de la Documentation* en ont fait leur thème principal. Sur le web, plusieurs bibliothécaires et documentalistes profitent de la technologie des weblogs pour faire circuler l'information (Open Acces News, BiblioAcid, In Between, ...). Tout cela n'est certainement pas suffisant, le public sensibilisé reste en très petit nombre, mais ces initiatives assurent une certaine visibilité au sujet et il faut espérer que celle-ci ira croissant.

L'information, aussi grande soit son importance, doit s'accompagner d'une action émanant des autorités universitaires, des pouvoirs publics, des chercheurs. Un bon exemple en est certainement le *Joint Information Systems Committee* qui en Grande Bretagne encourage la création de dépôts et de projets favorisant l'accès à l'information. En communauté française de Belgique, le projet BICTEL/e, lorsqu'il sera finalisé, permettra aux chercheurs et étudiants des universités francophones belges, d'auto-archiver leurs thèses et mémoires.

Conclusion

Si l'*Open Access* apparaît comme l'une des principales solutions à la crise des périodiques, elle ne présente pas que des avantages. Un des inconvénients majeurs est, à notre avis, le passage implicite à un "tout électronique". L'objectif principal des projets que nous avons rencontrés est d'améliorer la diffusion des textes, la rendre plus rapide et accessible à un plus grand nombre de personnes. La diffusion via internet rend cela possible : pas de temps "perdu" pour l'impression et l'envoi du courrier, articles consultables par plusieurs personnes en même temps, ... Tous ces avantages ont aussi leur revers : on ne tombe plus par hasard sur un article intéressant dans une revue abandonnée sur un coin de table et on ne rencontre plus personne en allant chercher une revue dans la bibliothèque, la lecture à l'écran n'est pas aussi confortable que sur papier, ...

A ce problème d'ordre social, on peut ajouter le problème de l'archivage et des coûts cachés. Alors que les bibliothèques gèrent l'archivage des revues sur papier, à qui appartient ce rôle dans le cas de documents électroniques ? L'accès aux anciens numéros peut être garanti dans le contrat signé entre la bibliothèque et l'éditeur commercial mais dans le cas de dépôts disciplinaires, institutionnels et des revues dont la consultation est gratuite, il n'y a pas de contrat signé. Signalons toutefois, même avec un éditeur commercial, il arrive qu'un article ne soit plus accessible. Elsevier a ainsi supprimé de son offre plusieurs articles considérés comme "polémiques" ou "mauvais" après avoir passé la barrière du *peer review* et avoir été publié dans l'une de leurs revues[Lapèlerie 2003]. Mais cela est un autre aspect de la publication en ligne sur lequel nous ne nous étendrons pas.

Conclusion

La consultation gratuite d'une ressource en ligne, coûte toujours une certaine somme à son lecteur, même si cela n'apparaît pas directement. Il doit payer la connexion au serveur, doit s'être équipé du matériel informatique adéquat. S'il souhaite conserver l'article, c'est à lui de payer le stockage sur son disque dur ou le papier et l'encre qui lui serviront à se procurer une copie sur papier. Si ces différents éléments sont facilement disponibles pour les chercheurs des pays favorisés, il n'en va pas de même pour les pays en voie de développement.

La publication par les éditeurs commerciaux, qu'il s'agisse de revues ou de monographies, a encore sa place sur nos étagères et nos ordinateurs mais si la croissance des prix continue, les bibliothèques seront obligées de s'en passer ou, au minimum, de diminuer leur offre de documentation à leurs utilisateurs. Augmenter l'utilisation des dépôts institutionnels et disciplinaires, favoriser l'auto-archivage et considérer les revues en ligne comme les égales des revues papiers sont des moyens d'augmenter la concurrence avec les éditeurs commerciaux. Ces derniers, limités à quelques grands groupes, sont en situation d'oligopole, ce qui leur permet d'imposer leurs prix. L'augmentation de la concurrence pourrait les inciter à baisser les coûts des abonnements.

La situation idéale serait peut-être un retour à la normale des prix des périodiques et une co-existence des initiatives *Open Access* et des éditeurs commerciaux. Pour mieux comprendre comment cela serait possible et quelle situation serait, en réalité, la plus équilibrée, la plus réalisable, de nombreux points devraient encore être abordés. Citons simplement celui de

Conclusion

la technique, elle évoluera encore certainement et permettra l'apparition de nouveaux outils. La question des droits d'auteur, même si nous l'avons peu abordée est également omniprésente dans ce débat, et, le droit n'étant pas figé, des évolutions pourraient encore largement être envisagées dans ce domaine.

La communication scientifique est un sujet vaste, les changements induits par la crise des périodiques et les nouvelles technologies n'ont fait que l'étendre et il est impossible d'en aborder toutes les facettes dans le cadre d'un seul mémoire.

Bibliographie

Articles et monographies

[Agostinelli 2003] AGOSTINELLI, S., *Les nouveaux outils de communication des savoirs*, Paris : L'Harmattan, 2003.

[Buck 1999] BUCK A.M., COLES B., FLAGAN R. C., *Scholar's Forum : A New Model For Scholarly Communication*, Pasadena : California Institute of Technology, 1999.

<<http://resolver.caltech.edu/CaltechLIB:1999.001>>

[Buck 2002] BUCK A.M., FLAGAN R.C., *The Scholar's Forum Debut*, Pasadena : California Institute of Technology, 2002.

<<http://resolver.caltech.edu/CaltechLIB:2002.003>>.

[Case 2002] CASE M. M., *Igniting Change in Scholarly Communication : SPARC, Its Past, Present, and Future* (comme publié dans *Advances in Librarianship*, vol. 26), 2002.

<http://www.epublications.de/SPARC_Advances.pdf>

[Chartron 2000] CHARTRON G., SALAÜN J.-M., *La reconstruction de l'économie politique des publications scientifiques*, in, *Bulletin des Bibliothèques de France*, Paris, 2000, t. 45 n°2, pp. 32-42.

[Chartron 2003] CHARTRON G., *Les archives ouvertes dans la communication scientifique*, 2003.

<<http://web.ccr.jussieu.fr/urfist/archives-ouvertes.htm>>.

Bibliographie

[Chartron 2003b] CHARTRON G., *Eléments pour une approche comparée de la publication scientifique*, Forum Universitaire : La communication scientifique en quatre dimensions, 4-6 juin, Archives nationales du Québec, Montréal, 2003, 14 pp.

<http://archivesic.ccsd.cnrs.fr/sic_00000435.html>

[Chartron 2004] CHARTRON G., *Marchands de savoir*, in, *Louvain*, Louvain-la-Neuve, avril 2004, n° 146, pp. 15-17.

[Conseil 2002] CONSEIL DE BIBLIOTHÈQUES, *La publication scientifique : problèmes et perspectives (Rapport du Conseil des bibliothèques)*, Louvain-la-Neuve, 2002.

<http://www.ucl.ac.be/cbib/pub_sc_nv91.pdf>.

[Crawford 2002] CRAWFORD W., *Scholarly Journals and Grand Solutions*, in, *Cites and Insights : Crawford at Large*, juillet 2002, vol. 2, n°9, pp. 1-3.

<<http://cites.boisestate.edu>>

[Dassetto 2004] DASSETTO F., *Il était une fois... des savants et des revues*, in, *Louvain*, Louvain-la-Neuve, avril 2004, n° 146, pp. 12-14.

[Dekeyser 2002] DEKEYSER R., *Bibliotheeksamenwerking voor een alternatieve methode voor wetenschappelijke communicatie*, in, *Entre réel et virtuel : la coopération entre les bibliothèques de recherche en Belgique (Actes du colloque organisé à Bruxelles, le 26 novembre 2001)*, Bruxelles, 2002, pp. 55-71.

[De la Vega 2000] DE LA VEGA J. F., *La communication scientifique à*

Bibliographie

l'épreuve de l'Internet, Villeurbanne : Presses de l'ENSSIB, 2000, 240 pp.

[Dorban 2004] DORBAN M., *La bibliométrie au banc des accusés ?*, in, Louvain, Louvain-la-Neuve, avril 2004, n° 146, pp. 21-22.

[Douglas 2002] DOUGLAS, K., *Cost of Creating Digital Collections at Caltech-CODA*, SPARC Presentation, 2002.

<<http://www.arl.org/IR/Douglas>>

[Doury 2003] DOURY-BONNET J., *Libre accès à l'information scientifique et technique*, in, *Bulletin des Bibliothèques de France*, Paris, 2003, t. 48 n°3, pp. 91-93.

<http://bbf.enssib.fr/bbf/html/2003_48_3/2003-3-p91-doury.xml.asp>

[Esposito 2004] ESPOSITO J., *The devil you don't know : The unexpected future of Open Access publishing*, in, *First Monday*, août 2004, vol. 9 n°8.

<http://www.firstmonday.org/issues/issue9_8/esposito/>

[Gallezot 2002] GALLEZOT G. , CHARTRON G., NOYER J.-M., *Une archive ouverte des publications en InfoCom*, 2002.

<http://archivesic.ccsd.cnrs.fr/sic_00000029.html>

[Gallezot 2003a] GALLEZOT G., *Archive SIC : rôle, fonctionnement et usage*, 2003.

<http://archivesic.ccsd.cnrs.fr/sic_00000600.html>

[Gallezot 2003b] GALLEZOT G., ROSSI C., CHARTRON G., NOYER J.-M., *Conception d'une archive ouverte en SIC : le sens de la technique*, 2003.

Bibliographie

<http://archivesic.ccsd.cnrs.fr/sic_00000943.html>

[Gobin 2004] GOBIN M., *Un panorama de la recherche universitaire belge : le répertoire BICTEL/e des thèses électroniques et e-prints*, in, *Cahiers de la documentation*, Bruxelles : Association belge de documentation, juin 2004, n°2, pp. 89-94.

[Harnad 1997] HARNAD S., *Comment accélérer l'inéluctable évolution des revues érudites vers la solution optimale pour les étudiants et les chercheurs ?*, Southampton, 1997, (version française de Harnad S., *How to Fast-Forward Serials to the Inevitable and the Optimal for Scholars and Scientists*, in, *Serials Librarian*, 1997, n°30, pp. 73-81, traduction par M.-N. Frachon et J.-M. Salaün).

<<http://www.ecs.soton.ac.uk/~harnad/Papers/Harnad/harnad97.revues.français.html>>

[Harnad 1998] HARNAD S., *On-Line Journals and Financial Fire-Walls*, Southampton, 1998.

<<http://www.ecs.soton.ac.uk/~harnad/nature.html>>

[Harnad 2004] HARNAD S., BRODY T., *Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals*, in, *D-Lib Magazine*, juin 2004, vol. 10 n°6.

<<http://www.dlib.org/dlib/june04/harnad/06harnad.html>>

[Harnad 2001] HARNAD S., *Lecture et écriture scientifique "dans le ciel" : Une anomalie post-gutenbergienne et comment la résoudre* (traduction par N. Fresco), Southampton, 2001.

Bibliographie

<<http://www.ecs.soton.ac.uk/~harnad/Tp/text-e.htm>>

[Hey 2004] HEY J.M.N., *An environmental assessment of research publication activity and related factors impacting the development of an Institutional e-Print Repository at the University of Southampton*, Southampton : University of Southampton, 2004, 19 pp.

<<http://eprints.soton.ac.uk/archive/00006218/>>

[Hirtle 2001] HIRTLE P., *OAI and OAIS : what's in a name ?*, in, *D-Lib Magazine*, numéro 7 volume 4, avril 2001.

<<http://www.dlib.org/dlib/april01/04editorial.html>>

[Iacovella 2004] IACOVELLA A., *Les portails de revues en sciences humaines et sociales*, 2004.

<http://archivesic.ccsd.cnrs.fr/sic_00000980.html>

[Institute 2002] INSTITUTE FOR THE FUTURE, *Final Synthesis Report of the e-Journal User Study*, Menlo Park, 2002, 42 pp.

<<http://ejust.stanford.edu/SR-786.ejustfinal.pdf>>

[Jeannin 2000] JEANNIN P. , BEAU V., *Enquête sur les revues d'information et de communication*, Paris : Ministère de la Recherche en France, 2000.

<http://www.iut.tarbes.fr/enquete/info_comm/>

[Jérôme 2001] JÉRÔME S., *Rapport sur le "Workshop on the OAI and Peer Review Journals in Europe" Genève (CERN) 22 au 24 mars 2001*, in, *Cahiers de la documentation*, Bruxelles : Association belge de documentation, décembre 2001, n°4, pp. 59-63.

Bibliographie

[Krichel 1997] KRICHEL T., *About NetEc, with special Reference to WoPEc*, 1997.

<<http://netec.mcc.ac.uk/doc/hisn.html>>

[Krichel 2000] KRICHEL T., *RePEc, an Open Library for Economics*, 2000.

<<http://openlib.org/krichel/papers/salisbury.html>>

[Lapèlerie 2003] LAPÈLERIE F., *Un autodafé électronique*, in, *Bulletin des Bibliothèques de France*, Paris, 2003, t. 48 n°2, pp. 130-132.

<http://bbf.enssib.fr/bbf/html/2003_48_2/2003-2-p130-lapelerie.xml.asp>

[Lognay 2004] LOGNAY C., *La publication scientifique : l'avis d'un chercheur*, in, *Cahiers de la documentation*, Bruxelles : Association belge de documentation, juin 2004, n°2, pp. 86-88.

[Lynch 2001] LYNCH C., *Metadata harvesting and the Open Archives Initiative*, in, *ARL Bimonthly Report*, août 2001, n° 217.

<<http://www.arl.org/newsltr/217/mhp.html>>

[Needleman 2002] NEEDLEMAN M., *The Open Archives Initiative*, in, *Serials Review*, 2002, n°28, pp. 156-158.

[Rohe 1998] ROHE T. A., *How Does Electronic Publishing Affect the Scholarly Communication Process ?*, in, *The Journal of Electronic Publishing*, mars 1998, vol. 3 n° 3.

<<http://www.press.umich.edu/jep/03-03/rohe.html>>

Bibliographie

[Roucoux 2004a] ROUCOUX A., *La science et sa communication*, in, *Cahiers de la documentation*, Bruxelles : Association belge de documentation, juin 2004, n°2, pp. 74-86.

[Roucoux 2004b] ROUCOUX A., *Le peer review : un mal nécessaire ?*, in, *Louvain*, Louvain-la-Neuve, avril 2004, n° 146, pp. 18-20.

[Schaffner 1997] SHAFFNER A. C., *Caltech Conference on Scholarly Communication : March 25-27, 1997*, in, *Serials Review*, Summer 1997, vol. 23 n°2, pp. 85-89.

[Stone 2000] STONE L., COMSTOCK B., GLAVASH K., *Harvesting the Low-hanging Fruit : World Wide Web Access to a Collection of MIT Theses*, Cambridge : Massachusetts Institute of technology, 2000.

<<http://mit.edu/theses/www/fruit.pdf>>

[Suleman 2001a] SULEMAN H., FOX A., *The Open Archives Initiative : Realizing Simple and Effective Digital Library Interoperability*, in, *Journal of Library Administration*, 2001, 35(1/2), pp. 125-145.

[Suleman 2001b] SULEMAN H., ATKINS A., GONÇALVES M.A., *Networked Digital Library of Theses and Dissertations. Bridging the gaps for global access - Part 1 : Mission and Progress*, in, *D-Lib Magazine*, septembre 2001, vol. 7 n°9.

<<http://www.dlib.org/dlib/september01/suleman/09suleman-pt1.html>>

[Suleman 2001c] SULEMAN H., ATKINS A., GONÇALVES M.A., *Networked Digital Library of Theses and Dissertations. Bridging the gaps for global access - Part 2 : Mission and Progress*, in, *D-Lib Magazine*, septembre 2001, vol. 7 n°9.

Bibliographie

<<http://www.dlib.org/dlib/september01/suleman/09suleman-pt2.html>>

[Van de Sompel 2000] VAN DE SOMPEL H., LAGOZE C., *The Santa Fe convention of the Open Archives Initiative*, in, *D-Lib Magazine*, numéro 6, volume 2, février 2000.

<<http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html>>

[Vandooren 2002] VANDOOREN F., *Commercial publishers : friends of foes*, Ethical forum "Scientific publication in internet era", University Foundation Brussels, 21 octobre 2002.

<<http://universitairstichting.be/fr/forum1.php>>.

[Van Parijs 2004] VAN PARIJS P., *Les enjeux éthiques de la publication électronique*, in, *Louvain*, Louvain-la-Neuve, avril 2004, n° 146, pp. 23-25.

[Webster 1999] WEBSTER D., *Début de réponses à la crise des périodiques scientifiques*, 65th IFLA Council and General Conference, Bangkok, Thaïlande, août 20-28, 1999.

<<http://www.ifla.org/IV/ifla65/papers/062-122f.htm>>

Sites Internet

Ces sites ont été visités pour la dernière fois le 15 septembre 2004.

@rchive SIC : *Archive Ouverte en Sciences de l'Information et de la Communication*

<<http://archivesic.ccsd.cnrs.fr>>

Archimède : Système de dépôt institutionnel de la Bibliothèque de

Bibliographie

l'Université Laval

<<http://archimede.bibl.ulaval.ca>>

ArXiv : e-print archive

<<http://xxx.arxiv.org>>

BiblioAcid : Informatique, Documentation numérique et autres questions bibliothéconomiques

<<http://www.biblioacid.org/>>

BioMed Central : The Open Access Publisher

<<http://www.biomedcentral.com>>

Caltech Collection of Open Digital Archives

<<http://library.caltech.edu/digital>>

Cognet : The Brain Sciences Connection

<<http://cognet.mit.edu>>

CogPrints : Cognitive Sciences Eprint Archive

<<http://cogprints.ecs.soton.ac.uk/>>

Columbia Earthscape : An Online Resource on the Global Environment

<<http://www.earthscape.org>>

Current Science Group : Services for the Biomedical Community

<<http://current-science-group.com>>

Cyberthèses : Publication et diffusion en ligne des thèses

Bibliographie

<<http://www.cybertheses.org>>

Digital Library of Massachusetts Institute of Technology Theses

<<http://theses.mit.edu>>

Diglet : A blog devoted to digital library issues from the UCSD Digital Library Planning Working Group (DLPWG)

<<http://gort.ucsd.edu/mtdocs/diglet/>>

DOAJ : Directory of Open Access Journals

<<http://www.doaj.org>>

Dublin Core Metadata Initiative : Making it easier to find information

<<http://dublincore.org>>

eJust : Electronic Journal User Study

<<http://ejust.stanford.edu>>

e-Prints Soton : University of Southampton e-Prints Service

<<http://eprints.soton.ac.uk>>

Erudit : Promouvoir et diffuser la recherche universitaire

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

eScholarship

<<http://www.cdlib.org/programs/escholarship.html>>

Fonds québécois de recherche sur la nature et les technologies

<<http://www.fcar.qc.ca/>>

Bibliographie

GNU EPrints

<<http://software.eprints.org>>

<<http://www.eprints.org>>

Grand dictionnaire terminologique (Office québécois de la langue française)

<<http://w3.granddictionnaire.com>>

HighWire

<<http://www.highwire.org>>

ht://Dig : WWW Search Engine Software

<<http://www.htdig.org>>

In Between : A weblog on scholarly online publishing, open access, and library related technology

<<http://eepi.ubib.eur.nl/iliit/>>

Initiative de Budapest pour l'accès ouvert

<<http://www.soros.org/openaccess/fr/index.shtml>>

JISC : The Joint Information Systems Committee

<<http://www.jisc.ac.uk>>

Medline

<http://www.nlm.nih.gov/databases/databases_medline.html>

NDLTD : Networked Digital Library of Theses and Dissertations

<<http://www.ndltd.org>>

Bibliographie

OAIster... find the pearls

<<http://oaister.umdl.umich.edu/o/oaister>>

Open Acces News : News from the Open Access Movement

<<http://www.earlham.edu/~peters/fos/fosblog.html>>

Open Archives Forum

<<http://www.oaforum.org>>

Open Archives Initiative

<<http://www.openarchives.org>>

Papyrus : dépôt institutionnel de l'Université de Montréal

<<http://papyrus.bib.umontreal.ca>> (non disponible au 15 septembre 2004)

PubMed

<<http://www.pubmed.org>>

RePEc : Research Papers in Economics

<<http://repec.org>>

SourceForge.net : The world's largest Open Source software development site

<<http://www.sourceforge.net>>

SPARC : The Scholarly Publishing and Academic Resources Coalition

<<http://www.arl.org/sparc>>

<<http://www.sparceurope.org>>

Bibliographie

TARDis : Targeting Academic Research for Dissemination and Disclosure

<<http://tardis.eprints.org>>

Washington DC Principles for free access to science

<<http://www.dcprinciples.org>>

Wikipedia, The free Encyclopedia

<<http://www.wikipedia.org>>