



HAL
open science

Usages, pratiques et besoins des chercheurs concernant les serveurs d'archives ouvertes

Angélique Gandossi

► **To cite this version:**

Angélique Gandossi. Usages, pratiques et besoins des chercheurs concernant les serveurs d'archives ouvertes. Sciences de l'information et de la communication. 2011. mem_00655746

HAL Id: mem_00655746

https://memic.ccsd.cnrs.fr/mem_00655746v1

Submitted on 2 Jan 2012

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.



Angélique GANDOSI

MASTER1, MENTION ICD
(Option Sciences de l'Information et du Document)

MÉMOIRE DE STAGE
Mission effectuée du 2 mai au 17 juin 2011

au
SERVICE COMMUN DE DOCUMENTATION
Lille1

**Usages, pratiques et besoins des chercheurs concernant les serveurs
d'archives ouvertes**

Sous la direction de :
M. J. SCHÖPFEL (tuteur universitaire)
Mme I. LE BESCOND (tuteur professionnel)
Mr J. Roche (Directeur de la Bibliothèque Universitaire de Lille1)

Soutenu le 30 juin 2011 à l'UFR IDIST
Université Charles de Gaulle, Lille 3 (Campus Pont de Bois)
BP 60 149, 59 653 Villeneuve d'Ascq Cedex

Année Universitaire 2010/2011

Sommaire

Remerciements.....	4
Introduction	6
Présentation du Service Commun de Documentation de Lille1	8
Les projets à venir.....	10
Description des missions	11
I) Les archives ouvertes, un outil de mise en valeur de la publication scientifique souterraine	12
1) « Archive ouverte » ou « Open Archive »	12
1.1) Définition :	12
1.2) « Open Archive » : Un terme à différencier de « l'Open Access »	13
1.3) Archives ouvertes, Auto-archivage et Archivage	13
2) Objectifs, types de documents et types d'archives	14
2.1) Les Objectifs identifiables des archives ouvertes	14
2.2) Les types de documents contenus dans une archive ouverte	16
2.3) Les différents types d'archives existantes	17
3) L'histoire des archives ouvertes	19
3.1) Début du mouvement pour le libre accès : quelles motivations ?	19
3.2) Les premières archives ouvertes.....	19
3.3) Le projet E-Biomed et la Pétition PLOS	20
3.4) Naissance du Protocole OAI-PMH	21
3.5) Les principales déclarations.....	21
3.6) Lancement du projet RoMEO	22
II) Les pratiques actuelles des chercheurs.....	23
1) Des pratiques de publications distinctes et les principes associés	23
1.1) Les moyens traditionnels, l'auto-publication et les archives ouvertes	23
1.2) Les revues en accès libre	24
1.3) Le principe « d'auteur-payeur » vs le principe de « lecteur-payeur »	24
2) Le dépôt des chercheurs.....	26
2.1) Lieu de dépôt privilégié par les chercheurs.....	26
2.2) Le dépôt des chercheurs sur HAL	27
2.3) Le dépôt des chercheurs de Lille 1 sur HAL	36

2.4) Facilité d'utilisation de HAL corrélé à l'âge des chercheurs.....	42
2.5) Le dépôt des chercheurs corrélé au nombre de dépôt et au temps de dépôt	43
3) La Consultation des chercheurs.....	45
3.1) Lieu de consultation et de recherche d'information privilégié par les chercheurs	45
3.2) Les caractéristiques de la consultation utilisée par les chercheurs	46
4) La Recherche.....	49
III) Les Besoins des chercheurs en matière d'archives ouvertes	50
1) L'exemple de HAL.....	50
2) L'exemple de AUSpace	52
3) L'exemple de Papyrus (Université de Montréal)	54
4) Les besoins demandés par les chercheurs	57
IV) La Perception des serveurs d'archives ouvertes	59
1) Avantages et Bienfaits à l'utilisation des sites d'archives ouvertes ?	59
2) Freins et Réticences à l'utilisation des archives ouvertes?	60
Conclusion	61
Bibliographie	62
Annexe 1 : Grille d'entretien utilisée :.....	64
Annexe 2 : Retranscription de l'entretien 1.....	69
Annexe 3 : Retranscription de l'entretien 2.....	77
Annexe 4 : Retranscription de l'entretien 3.....	86
Annexe 5 : Retranscription de l'entretien 4.....	94

Remerciements

Je tiens à remercier Mme Isabelle Le Bescond, conservatrice du Service Commun de Documentation de Lille1, et Mr Joachim Schöpfel, Enseignant-chercheur de Lille 3, pour l'aide et les conseils qu'ils m'ont apporté tout au long du stage.

Je voudrais aussi remercier toute l'équipe du Service de Documentation de Lille 1 qui m'ont aidé à m'intégrer et qui ont pris sur leur temps libre pour me fournir une présentation dans les moindres détails de la Bibliothèque Universitaire, de leur travail et de leurs missions.

En dernier lieu, je remercie les enseignants-chercheurs ayant bien voulu m'accorder un peu de leur temps pour réaliser des entretiens mais aussi Aline Colbeaux, Frédérique Créton et Leila Djeibien, étudiantes en Master 1 ICD (Lille 3), qui m'ont soutenu et donné des conseils supplémentaires durant ce stage.

Résumé

Le Centre Commun de Documentation de Lille1 désire mettre en place un serveur d'archives ouvertes destiné aux chercheurs dans le but de leur permettre d'archiver de façon pérenne leurs documents scientifiques et techniques et cela en toute sérénité. Dans ce cadre-là, mon stage a consisté à réaliser une étude comparative et une recherche approfondie sur les sites d'archives ouvertes tels que HAL, OATAO, SPIRE et ORBI, et cela sous forme de rapport d'audit. J'ai ainsi essayé de distinguer les différents services (exemple : service de dépôt, service de consultation...) mis en place par les sites d'archives ouvertes, la composition de ces services en terme d'éléments structurants (exemple : divers critères de consultation, ...) mais je me suis aussi placée du côté des chercheurs pour essayer de comprendre leurs pratiques actuelles en matière de dépôt, de consultation et de recherche de documents. Enfin, par l'intermédiaire d'entretiens semi-directifs, j'ai voulu savoir quels étaient leurs réels besoins en terme de services, d'architecture du site, ... mais aussi de connaître leur avis et leur perception du site qu'ils utilisent actuellement.

Mots clés

Archives ouvertes, Accès Libre, Bibliothèque universitaire de Lille1, HAL, Entrepôt de documents, Archive Institutionnelle, Prépublication, Post-publication, pratiques des chercheurs, besoins des chercheurs, service de dépôt, service de consultation.

Abstract

The Lille1 Library wishes to implement a open archives website for researchers in order to enable them to archive their scientific and technical documents in a lasting way. In that context, my internship was to conduct a comparative study and thorough search on Open Archives sites as HAL, OATAO, SPIRE and ORBI in a form of audit report. I tried to distinguish the different services set up by theses websites, the composition of theses services in terms of structural elements but I also place on the side of researchers in order to understand their current practices regarding filing, consulting and search documents. Finally, I wanted to know what their real needs in terms of services, website architecture,... but also their views and their perception of the site they use now.

Keywords

Open Archive, Open Access, Lille1 Library, HAL, Digital Repository, Institutional Repositories, Preprint, Postprint, Researchers practices, the needs of Researchers, Deposit Service, Consultation Service

Introduction

Depuis quelques années, les nouvelles technologies et les services proposés par celles-ci envahissent le terrain. Elles modifient notre manière de communiquer, de travailler et d'accéder à la connaissance. Ces nouvelles technologies remettent en cause des équilibres établis antérieurement, des habitudes bien ancrées en nous, et il est parfois bien difficile de les changer. C'est le cas notamment pour le passage du livre papier au livre numérique, de la revue papier à la revue numérique : dans les deux cas, nous éprouvons des difficultés à changer de support visuel. Ces mêmes technologies bouleversent le champ de l'édition et l'on vient à se poser la question : « allons nous assister à la fin de l'édition scientifique traditionnelle ? ». Cependant, tel n'est pas le cas, l'édition traditionnelle scientifique devient une édition numérique et « multi-supports » capable de s'adapter aux nouveaux besoins des chercheurs.

Dans ce contexte, les archives ouvertes et le mouvement pour le libre accès apparaissent et prennent de l'ampleur. Dans son rapport, M. J. Schöpfel [1]¹ démontre que le nombre de sites d'archives ouvertes est passé de 56 en 2008 à 150 en 2009, soit une augmentation de 174%. A cela, il faut ajouter les archives ouvertes créées et développées en 2010 et en 2011 et les projets éventuels à venir.

Le Service Commun de Documentation de Lille¹ a très vite compris l'ampleur de ce phénomène et tous les enjeux qui se trouvent derrière. Il a ainsi mis en place un serveur de thèses, HDR et documents numérisés en histoire des sciences nommé *IRIS*². Aujourd'hui, il désire mettre à disposition de sa communauté académique une plateforme d'archives ouvertes contenant diverses publications produites par l'Université de Lille¹.

Dans cette perspective, le projet de mon stage a consisté à faire une recherche et une analyse en amont de l'existant et des besoins des chercheurs en matière d'archives ouvertes, pour ensuite fournir des éléments fonctionnels et techniques pour la mise en place possible d'un tel projet.

Dans ce mémoire, je me suis positionnée du côté du chercheur pour comprendre leur vision et leur utilisation des archives ouvertes et j'ai alors cherché à répondre à la question suivante : « En quoi les serveurs d'archives ouvertes

¹ Afin d'assurer une visibilité optimale de ce mémoire, les citations et les références bibliographiques seront suivies d'un numéro que vous retrouverez dans la bibliographie à la fin de ce mémoire.

² Pour accéder au site IRIS (archive institutionnelle de l'Université de Lille¹ pour la diffusion des thèses, HDR et pour les documents numérisés en Histoire des Sciences), voir le lien : <https://iris.univ-lille1.fr/dspace>

répondent mieux ou non aux besoins des chercheurs que les éditeurs traditionnels ? Quelle est alors la perception des chercheurs par rapport à ce nouveau mode de diffusion ? ».

Pour pouvoir répondre à cette question, j'ai commencé par réaliser une analyse approfondie de HAL³. Cette analyse approfondie m'a permis de voir la structuration globale d'un site avec les services souvent mis en place. J'ai ensuite effectué une analyse statistique du dépôt avec les chiffres trouvés sur HAL pour comprendre qui déposait le plus, les types de documents déposés... Enfin, pour comprendre l'utilisation réelle de ces sites par les chercheurs, j'ai réalisé des entretiens semi-directifs avec des chercheurs ou chercheurs enseignants en informatique et mathématique.

Dans ce cadre-là, je m'attacherai dans une première partie à définir les archives ouvertes et je rattacherai cette définition à un bref historique qui retracera les moments les plus importants de la mise en place de ces archives. Ensuite, dans une seconde partie, j'analyserai les pratiques actuelles des chercheurs en matière de dépôt, de consultation et de recherche au sein de ces archives ouvertes. Dans une troisième partie, j'aurai à cœur de comprendre quels sont les besoins des chercheurs en termes de services, d'architecture du site,... Enfin, l'objectif de la dernière partie sera d'expliquer et comprendre la perception qu'ont les chercheurs des archives ouvertes.

³ Hyper Article en Ligne

Présentation du Service Commun de Documentation de Lille1

Le Service Commun de Documentation de Lille1, aussi appelé « Bibliothèque universitaire », fut construit en 1965 par l'architecte Noël LeMaresquier. Il se trouve au centre du campus de Lille1, et est conçu comme un « point d'ancrage et d'amarrage », où chaque discipline s'y arrête et s'y rencontre pour se documenter, lire...

La Bibliothèque Universitaire est un bâtiment classé comme monument historique et est original de part sa forme ronde et son aspect panoramique (disposant de multiples fenêtres). Sa superficie globale atteint les 8000 m² et elle est conçue sur trois niveaux : un sous-sol, un rez-de-chaussée et un premier étage.

Actuellement, la Bibliothèque Universitaire se prépare à être rénovée en « Learning Center ». C'est la raison pour laquelle le premier étage est fermé au public.

Ainsi, avant cette préparation à la rénovation, les collections destinées aux publics se trouvaient sur deux niveaux :

- au Rez-de-chaussée, on pouvait retrouver les collections des licences-masters. Ces collections étaient réparties dans trois salles distinctes. Il y avait ainsi une salle des Sciences et Technologies (Informatique, Mathématiques, Astronomie, Physique...), une salle de culture générale (Langues, Art, Urbanisme, Littérature, Histoire, Géographie) et une salle des Sciences Humaines (Sociologie, Économie, Droit, Gestion).

- au premier étage, on pouvait accéder à des collections concernant la recherche réparties dans six salles différentes nommées par des lettres⁴ et à une salle informatique.

A l'heure actuelle, le rez-de-chaussée reste ouvert au public et les collections des licences-masters sont toujours accessibles via ces trois mêmes salles. Cependant, les collections du premier étage ne sont plus en accès-libre : une partie a été placée en magasin et une autre partie dans la salle de culture générale. Aujourd'hui pour pouvoir les consulter ou les emprunter, cela fonctionne comme pour les livres du magasin : il faut remplir une fiche de demande de consultation (couleur verte).

Le rez-de-chaussée dispose également d'une banque de prêt, d'un bureau de renseignement permettant aux lecteurs d'obtenir une aide à la recherche documentaire, ou une information pratique concernant la Bibliothèque Universitaire ou le Campus. Il comprend aussi des photocopieurs ainsi que 11 postes

⁴ Exemple : La salle C comprenait des documents de Chimie, la salle G-H des documents de Sociologie, Economie, Gestion et Géographie,...

informatiques dédiés à la recherche sur le catalogue de la Bibliothèque Universitaire et 2 postes connectés au site web de la bibliothèque.

Les projets à venir

Afin de s'adapter aux besoins évolutifs des lecteurs (population étudiante, population des enseignants-chercheurs,...), la Bibliothèque universitaire désire mettre en place plusieurs projets innovants.

Le « Learning Center »⁵ est un des premiers projets qui verra le jour d'ici 2014. Il consiste d'une part en une rénovation et en un agrandissement du bâtiment afin de l'adapter aux objectifs HQE (Haute Qualité Environnementale) et aux besoins pédagogiques des étudiants. D'autre part, il a pour objectif de devenir un « lieu vivant », une « place to explore and to stay » [2], autrement dit un « lieu de rencontre entre diverses pratiques ». Outre cela, le Learning Center proposera une plus grande offre en terme d'Information et de Documentation, un « Internet Café », un « expérimentarium », un « auditorium », des salles de conférence, quelques salles de cours et de formation et des salles d'exposition. Face aux nouvelles exigences des lecteurs, cette nouvelle bibliothèque augmentera son volume horaire (60 heures par semaine) et le nombre de places assises (1000 au lieu de 650 places). Par là même, c'est aussi tout un équipement électronique qu'elle mettra à disposition du public et une technologie RFID⁶ permettant la lecture et l'enregistrement d'étiquettes radiofréquence⁷.

Si le premier projet a déjà été approuvé, le deuxième n'est qu'à la phase de réflexion. La Bibliothèque Universitaire de Lille 1 désire en effet créer une plateforme d'archives ouvertes permettant de répondre aux besoins des chercheurs de Lille 1. Le but revient alors à proposer aux chercheurs des services fiables et utiles et un moyen de diffusion « rapide » de leurs documents scientifiques sans forcément devoir passer par une plateforme centralisée (comme HAL⁸). Ce serveur devra aussi favoriser le libre accès⁹ à ces documents. S'il est approuvé, ce projet devra fournir les éléments nécessaires à sa mise en place tels que des services de dépôts, de consultation,...

⁵ Voir le site informationnel de la Bibliothèque Universitaire de Lille1 que vous trouverez sur cette page : <http://bloc.univ-lille1.fr>

⁶ En anglais « Radio Frequency Identification » ou « Identification par Radio Fréquence ».

⁷ Il s'agit d'étiquettes émettant une onde radio, souvent apposées sur les livres des nouvelles bibliothèques afin de rendre l'utilisateur plus autonome. Ces étiquettes sont lues par un lecteur qui capte et transmet l'information.

⁸ Pour le bien de l'étude, la plateforme HAL sera décrite et analysée dans la première partie sur les archives ouvertes.

⁹ Le mouvement pour le libre accès sera aussi expliqué dans la première partie de ce document.

Description des missions

Durant ce stage, ma mission principale a consisté à réaliser un travail de recherche et d'analyse sous forme d'audit pouvant permettre la mise en place future d'une plateforme d'archives ouvertes. Cette mission s'est déroulée en plusieurs étapes :

- phase de lecture de publications scientifiques sur le thème des archives ouvertes et phase d'observation des principales archives ouvertes existantes.
- phase d'analyse des plateformes existantes (HAL, OATAO, ORBI, SPIRE) avec une analyse des pratiques en matière de dépôt des chercheurs de Lille1.
- phase d'analyse des différents utilisateurs (qui sont-ils ? de quels laboratoires viennent-ils ?...)
- description du service attendu. Cette dernière phase s'appuie sur l'analyse de la plateforme d'archives ouvertes HAL, sur son analyse critique (points forts et points faibles) et sur des entretiens semi-directifs.
- réalisation d'un rapport d'audit.

I) Les archives ouvertes, un outil de mise en valeur de la publication scientifique souterraine

1) « Archive ouverte » ou « Open Archive »

1.1) Définition :

La notion d'archive ouverte, traduction de l'anglais « Open Archive » se réfère à ce que l'on appelle un « Digital Repository ».

Ce « Digital Repository » est défini par Annaig Mahé [3] comme étant un « réservoir de documents scientifiques et techniques auto archivés, en accès libre en ligne, sans barrière économique, ni juridique, intégrant le protocole OAI-PMH »¹⁰.

Yves Lemay, dans son article paru dans un ouvrage du traité des sciences et techniques de l'information [4], va quant à lui définir les archives ouvertes comme étant des « entrepôts d'informations contenant des documents numériques quelle que soit leur antériorité accessible par une interface de type moteur de recherche ».

Au vue de ces deux définitions, une archive ouverte se caractérise par les cinq éléments suivants :

- un réservoir ou un entrepôt de documents ou d'informations intégrant le protocole OAI-PMH
- son contenu doit être des documents scientifiques numériques
- son contenu est obligatoirement disponible en ligne via un serveur ou un moteur de recherche
- son contenu est disponible gratuitement
- ses documents ne sont pas forcément des documents récents.

Ce réservoir de documents doit donc permettre aux chercheurs de déposer leurs publications scientifiques assez rapidement et aux internautes de les consulter librement. Les documents déposés peuvent être des textes intégraux ou seulement des notices bibliographiques, au choix du déposant.

¹⁰ Définition provenant de l'ouvrage d'Annaig Mahé.

1.2) « Open Archive » : Un terme à différencier de « l'Open Access »

Les « Open Archives » sont les résultats du mouvement pour « l'Open Access »¹¹.

Ce mouvement est né dans les années 1990. L'enjeu était de mettre à disposition en ligne, gratuitement et sans restriction les résultats de la recherche. Ce mouvement a mobilisé de nombreux acteurs (chercheurs, instituts de recherche, bibliothèques...). Il serait vu comme étant la conséquence de nombreux facteurs : « augmentations démesurées des prix des maisons d'éditions, le retard entre la remise et la publication d'un article, la cession forcée des droits d'auteurs aux maisons d'édition et la restriction de la diffusion des résultats de recherche » [5].

L'Open Access peut-être alors traduit par le terme « Libre Accès ». Ainsi selon K. Guma [6], l'Open Access signifie « l'accès gratuit et immédiat aux publications scientifiques ».

A contrario, l'Open Archive est le « **stockage** des publications scientifiques sur des serveurs fonctionnant avec un logiciel Open Source, permettant une interopérabilité avec d'autres serveurs et offrant un accès totalement gratuit » [6].

1.3) Archives ouvertes, Auto-archivage et Archivage

On peut dire que les archives ouvertes dépendent largement de l'auto-archivage comme le montre Yvon Lemay dans son article [4].

En effet, auto-archiver revient à donner la possibilité au chercheur de déposer lui-même ses articles (post-publications et prépublications) dans des archives ouvertes.

Si l'on veut comprendre un peu plus cette définition d'auto-archivage, il faut commencer par éclaircir celle d'archivage.

On peut penser qu'archiver un document consiste simplement à le stocker dans un entrepôt disponible en ligne. Cependant l'opération de « stockage » reste une opération minimale qui ne perdure pas : on stocke l'information en attendant de la réutiliser et de la supprimer.

Ici, l'archivage a bien évidemment pour but, comme le souligne Yves Lemay [4], de préserver un document sur un support quelconque, qui peut avoir une valeur de preuve et/ou une valeur de mémoire d'informations. Ce document contiendra des signes de validation permettant de le rendre authentique.

¹¹ Ce mouvement sera décrit plus bas (partie suivante).

2) Objectifs, types de documents et types d'archives

2.1) Les Objectifs identifiables des archives ouvertes

La publication au sein d'une archive ouverte répond à un certain nombre d'objectifs. Ces objectifs peuvent être distingués selon le statut de l'internaute (Chercheur ou Lecteur).

2.1.1) Objectifs pour le chercheur

Pour le chercheur, ces objectifs consistent principalement à lui donner un **espace de dépôt** lui permettant d'archiver ses documents en toute sérénité. Cependant d'autres objectifs peuvent compléter ce premier objectif :

- Répertorier et sauvegarder toutes les publications d'un chercheur ou d'un groupe de chercheurs en un lieu unique
- Permettre une **diffusion rapide** de ses publications scientifiques.
- Accroître la **Visibilité** de ses travaux de recherches aussi bien au niveau national qu'international, et cela grâce au protocole OAI-PMH¹² permettant à un site et à ses ressources d'être bien référencées sur Internet.
- Assurer la **pérennité** d'un document scientifique.
- Faciliter la **communication entre pairs** grâce à la Visibilité des documents.
- Obtenir la **reconnaissance de ses pairs**.
- Augmenter la **notoriété** du chercheur, afin qu'il soit reconnu et cité par des lecteurs, ou des partenaires potentiels.
- Offrir aux scientifiques des **outils et solutions techniques conviviales** permettant de mettre en place les nouveaux circuits des publications scientifiques.
- **Objectif économique** : proposer une alternative moins onéreuse à la publication scientifique.

¹² Protocole qui sera expliqué plus bas dans la partie « Bref Historique ».

2.1.2) Objectifs pour le lecteur

Pour le lecteur, ces objectifs se résument principalement à la facilité d'accès, à l'ergonomie de la recherche et à la qualité des outils de consultation mis en place. On peut relever d'autres objectifs dont voici une liste non-exhaustive :

- Favoriser un **accès libre, gratuit et immédiat** aux ressources scientifiques et à la connaissance.
- Pouvoir localiser **et utiliser la littérature d'importance** pour lui.
- Disposer de **ressources de qualité**.
- Disposer d'une **interface de recherche unique**.
- Avoir la possibilité d'**être informé en temps réel des nouveaux dépôts**.

2.1.3) Autres objectifs

Il existe à l'heure actuelle d'autres objectifs tels que :

- **Objectif administratif** : Faciliter le contrôle de la publication scientifique des différents organismes de recherche (EPST, Universités) par le recensement et l'évaluation des publications des chercheurs via une base centrale.
- Favoriser l'**interopérabilité des archives**, autrement dit « la recherche croisée » (Possibilité de dialogue entre plusieurs systèmes sans aucune perte de données et cela grâce au protocole OAI-PMH).

2.2) Les types de documents contenus dans une archive ouverte

D'une manière générale, une archive ouverte permet de déposer des publications. Il peut s'agir de preprints (prépublications) et/ou de Postprints (Post-publications).

Ainsi, un document « preprint » est selon la définition de Gabriel Gallezot [7] un « texte finalisé par son (ou ses) auteurs, mais non encore publié (évalué, corrigé) par un éditeur. » Les preprints sont donc les articles qui n'ont pas encore été soumis au contrôle des pairs ou experts.

Par contre un document « postprint » est la version publiée dans une revue d'un texte produit par un (des) auteur(s) après acceptation par un comité de rédaction et éventuellement par un comité de lecture (évaluation par les pairs). [8] Ces post-publications sont donc des articles qui ont été certifiés par les pairs et publiés dans des revues scientifiques.

Parmi ces publications (postprints ou preprints), on peut trouver des articles scientifiques, des documents numérisés (exemple : Gallica, Persée), des thèses et HDR¹³, des cours, des ouvrages, des documents multimédias, des communications, des actes de congrès ou de conférence...

Il arrive parfois aussi de trouver des rapports techniques et des études, mais le dépôt de ce type de document s'effectue le plus souvent dans des archives internes c'est-à-dire des archives d'instituts, de laboratoires, d'unités de recherches,...

De plus, ces publications, disponibles dans un format lisible par tous (PDF), sont accompagnées de leur « fiche catalographique concise ou détaillée »¹⁴. Cette fiche contient des métadonnées permettant de renseigner le lecteur sur le contenu du document (résumé, titre, auteur...).

¹³ Habilitation à diriger une recherche

¹⁴ Par fiche catalographique, j'entends une fiche composée de métadonnées. Cette fiche est souvent appelée « notice bibliographique ».

2.3) Les différents types d'archives existantes

Il existe principalement 4 types d'archives ouvertes. Ces 4 types d'archives peuvent se distinguer selon les genres (disciplines), le type de documents et le type d'acteurs.

On compte en premier lieu les « archives institutionnelles » ou « institutional repositories » contiennent les réalisations des chercheurs d'universités, de grandes écoles, de centres de recherches, laboratoires. Ces archives ont pour but principal de valoriser la production scientifique de ces chercheurs.

En deuxième lieu, les archives disciplinaires ont pour objectif de répondre aux besoins spécifiques d'une communauté de chercheurs identifiée et d'améliorer la communication entre ces chercheurs.

En troisième lieu, les archives thématiques sont des types d'archives se focalisant sur un sujet en particulier ou un type de document spécifique.

En dernier lieu, les archives centrales sont des serveurs généralement mis en place au niveau national par un pays en particulier qui sont destinées à recueillir la production scientifique de ce pays afin d'améliorer la visibilité au niveau international.

Archives ouvertes institutionnelles	SPIRE : http://spire.sciences-po.fr/web/
	OATAO: http://oatao.univ-toulouse.fr/
	Archive institutionnelle du CERN: http://weplib.cern.ch/
	Archive institutionnelle de l'IFREMER : http://archimer.ifremer.fr/
	Archive ouverte du CNRS : http://hal.archives-ouvertes.fr/CNRS/fr/
	Laboratoire d'ethnologie et de sociologie comparative de Paris : http://hal.archives-ouvertes.fr/LESC/fr/
	Archive ouverte de l'ENSSIB : http://hal.archives-ouvertes.fr/ENSSIB/fr/
Archives ouvertes disciplinaires	ArXiv : http://arxiv.org/
	E-Lis : http://eprints.rclis.org/
	PubMedCentral : http://www.ncbi.nlm.nih.gov/pmc/
	REPEC : http://www.repec.org/
Archives thématiques	Archives ouvertes de la Didactique des mathématiques : http://hal.archives-ouvertes.fr/DIDACTIQUE/fr/
	Archives ouvertes de la Méditerranée : http://hal.archives-ouvertes.fr/AOM/fr/
	Archives ouvertes des études urbaines : http://hal.archives-ouvertes.fr/ETUDES-URBAINES/fr/
	ArchiveSic : http://archivesic.ccsd.cnrs.fr/
	HAL SHS : http://halshs.archives-ouvertes.fr/
Archives ouvertes centrales	Plateforme HAL : http://hal.archives-ouvertes.fr/
	DARE (archive centrale néerlandaise : Digital Academic REpositories) : http://dare.uva.nl/en
	DEPOT (archive centrale britannique) http://opendepot.org/
	LARA : http://lara.inist.fr/
	DRIVER (Digital Repository Infrastructure Vision for European Research) : http://www.driver-community.eu/

Tableau 1 : Quelques exemples d'archives ouvertes selon le type d'archive

3) L'histoire des archives ouvertes

3.1) Début du mouvement pour le libre accès : quelles motivations ?

Le mouvement pour le libre accès a débuté dans les années 1990 et s'est conforté avec l'essor du World Wide Web.

Ce mouvement prend place avec la remise en cause par les chercheurs du monopole exercé par les grands éditeurs scientifiques sur les revues. Cette remise en cause est liée à deux phénomènes : d'une part, les chercheurs se voient confisquer leurs droits patrimoniaux sur leurs textes et d'autre part, ils voient les prix des abonnements aux revues augmenter sans cesse.

D'autre part, il y a un écart considérable entre la remise du document à l'éditeur et la publication de ce document. Il peut parfois s'écouler plusieurs années avant que le document soit publié, ce qui peut rendre le document obsolète.

Quelques chercheurs vont ainsi vouloir mettre en place un système capable de palier à tous ces problèmes, et permettant de fournir un accès immédiat et sans limite aux documents.

Cette volonté de mettre en place un nouveau système innovant, est aussi marquée par l'ambition de poursuivre des anciennes pratiques. En effet, au départ, l'échange de documents entre chercheurs et la communication entre eux se faisait principalement par courrier. La technologie du Web n'a fait que renforcer ces pratiques, et les rendre accessibles à tous.

3.2) Les premières archives ouvertes

3.2.1) ArXiv

En 1991, Paul Ginsparg, physicien du laboratoire national de Los Alamos, décide de créer la première archive ouverte permettant aux chercheurs en Sciences et Techniques (Physique, astrophysique, Mathématiques,...) de déposer leurs publications. Très rapidement, cette archive prend de l'ampleur et aujourd'hui, plus de 315 000 documents y sont archivés. Ce phénomène qui séduit nombre de chercheurs à l'heure actuelle, serait lié à la vitesse de communication et à la vitesse de « citation » dans d'autres articles.

3.2.2) Cogprints

Devant l'ampleur de ce phénomène, Stevan Harnad, chercheur en Sciences Cognitives, comprend très vite qu'il faut ouvrir ce moyen de diffusion à un plus grand nombre de chercheurs et à d'autres disciplines que la Physique, ou que l'Astrophysique. Pour lui, il ne faut pas que les archives restent « l'exclusivité d'une minorité de chercheurs », mais il faut étendre cette pratique aux autres chercheurs. Selon son expression, citée dans le chapitre 1 de l'ouvrage de Christine Aubry et

Joanna Janik, il faut « toucher les yeux et l'esprit de ses pairs » [9] en utilisant les technologies de l'Internet.

Il crée en 1997 l'archive électronique Cogprints. Elle donne la possibilité aux chercheurs en Psychologie, Neurosciences, Linguistique et autres branches des Sciences de l'Informatique de déposer leurs publications (Biologie, Philosophie, Informatique,...).

3.2.3) SciELO

En 1998, le Chili et le Brésil s'associent pour mettre en place l'archive « SciELO » favorisant le libre accès aux publications de leurs chercheurs. En 2003, Cuba et l'Espagne rejoignent ce projet. Le but était aussi de montrer que la littérature espagnole et portugaise n'était pas « de la science perdue » [9].

3.3) Le projet E-Biomed et la Pétition PLOS¹⁵

En mai 1999, deux chercheurs du National Center for Biotechnology Information lancent le projet E-Biomed, archive destinée à recevoir les préprints et les postprints des chercheurs en biomédecine. Derrière cela, il y a l'idée d'un contrôle, d'une expertise de qualité par les pairs.

Peu de temps après, H.Varmus reprend le projet à son compte et le nomme « PubMed Central » où il compte mettre seulement des postprints.

Cependant, il est confronté à la réticence des éditeurs qui ne désirent pas mettre leurs revues sur ce serveur. Il décide alors de diffuser une pétition sur le Internet nommée « PLOS ». Le but derrière cette pétition consiste pour les signataires à rendre accessibles leurs publications sur ce site six mois après qu'elles soient publiées dans une revue. Cette pétition fut signée par 34 000 chercheurs mais malgré cela, la pétition n'a pas eu l'effet escompté au départ et les éditeurs n'ont pas suivi le mouvement. En effet, peu de chercheurs sont allés au bout de leur démarche initiale consistant à dire « si vous n'obtempérez pas, je ne publie plus dans votre revue ». Le succès et la carrière d'un chercheur dépendaient en majeure partie de la publication en revue.

¹⁵ Terme signifiant « Public Library of Science ».

3.4) Naissance du Protocole OAI-PMH

Face au nombre d'archives ouvertes qui augmente, la recherche d'information devient de plus en plus difficile à trouver. En 1999, on parle d'un « véritable éclatement de l'information ». C'est la raison pour laquelle il y a une réelle volonté de la part des chercheurs de mettre au point un protocole rendant toutes les archives ouvertes « interoperables ».

Suite à cela, les principaux responsables des archives ouvertes décident de se réunir à Santa Fé pour convenir d'un plan d'organisation et cela afin de parvenir à une « interoperabilité minimale » entre les différents sites. Cette réunion donne lieu à l'élaboration d'une convention lancée en Février 2000 qui devient « The Open Archive Initiative Protocole Metadata Harvesting (OAI-PMH) ».

Aujourd'hui cette norme est pour nous un standard informatique international permettant « la recherche et la collecte dans toutes les archives distribuées simultanément comme si elles n'en faisaient qu'une » [9].

3.5) Les principales déclarations

3.5.1) *L'initiative de Budapest*

Pour faire avancer le mouvement du libre accès, quelques chercheurs se réunissent en 2001 à Budapest. Ils lancent en 2002 un appel dans le monde entier afin que d'autres chercheurs rejoignent ce mouvement pour rendre public et accessible à tous la communication scientifique sans restriction et rétribution.

Afin de toucher le plus de monde, cet appel a non seulement été lancé dans les périodiques scientifiques, mais il a aussi tourné sur Internet. Cette initiative fut nommée « la Budapest Open Access Initiative (BOAI) ».

Les signataires devaient s'engager à promouvoir le libre accès par une stratégie au choix :

- La BOAI1 permettait au chercheur d'auto-archiver ses publications dans des archives électroniques ouvertes
- La BOAI2, par la création de nouvelles revues alternatives aux revues commerciales, devait permettre au chercheur de publier dans une revue en libre accès.

3.5.2) *La déclaration de Berlin*

La mise en place de cette déclaration s'est déroulée en 3 étapes. Ces trois étapes correspondent à 3 réunions différentes n'ayant pas eu le même impact.

La première réunion « Berlin 1 » s'est tenue à Berlin en octobre 2003. Elle a donné lieu à une déclaration. Celle-ci visait à montrer l'adhésion des responsables universitaires pour le libre accès. Cependant, même s'il y avait des signataires, cette déclaration n'a eu aucune conséquence au niveau des actions concrètes (actions visant à publier ou auto-archiver).

La deuxième réunion « Berlin 2 » s'est déroulée au CERN à Genève en mai 2004, mais il n'y a eu aucune amélioration notable par rapport à la première réunion.

Enfin, la troisième réunion « Berlin 3 » s'est quant à elle tenue à Southampton en 2005. Le but recherché était que les signataires appliquent réellement une politique institutionnelle favorisant le libre accès. Il s'agissait alors pour les institutions de donner à leurs chercheurs les moyens nécessaires pour déposer dans des revues en libre accès ou dans des archives ouvertes institutionnelles.

Cette déclaration (Berlin 3) a été soutenue par le CERN, l'INRIA et le CNRS.

3.6) Lancement du projet RoMEO

Le projet RoMEO (Rights METadata for Open archiving) a été fondé par la JISC (Join Information Systems Committee) en août 2002. Il visait à répertorier les politiques d'auto-archivage de chaque éditeur. Ces politiques sont consultables sur le site E-Prints RoMEO¹⁶ et sur le site RoMEO Sherpa¹⁷.

Le principe de ces deux sites est de classer les politiques éditoriales d'auto-archivage en 4 couleurs distinctes :

- Les revues blanches caractérisent les revues n'autorisant pas l'auto-archivage en archives ouvertes
- Les revues jaunes caractérisent les revues autorisant l'archivage des prépublications
- Les revues bleues caractérisent les revues qui autorisent l'archivage des post-publications
- Les revues vertes caractérisent les revues autorisant la prépublication et la post-publication

¹⁶ <http://romeo.eprints.org/>

¹⁷ <http://www.sherpa.ac.uk/romeo/>

II) Les pratiques actuelles des chercheurs

En 2009, selon J.Schöpfel [1], on dénombrait 150 archives ouvertes. Aujourd'hui, leur nombre augmente encore. Il paraît indéniable que nous assisterons à une véritable généralisation des archives ouvertes. Selon Daniel Bourrion [10], elles deviendront un outil incontournable pour les chercheurs et le dépôt sera alors soutenu par l'évidence des pratiques.

La question qu'on peut se poser ici est de savoir où les chercheurs déposent-ils leurs articles ? Est-ce qu'ils déposent leurs articles dans des revues, ou est-ce qu'ils les déposent dans des archives ouvertes ? Ou y a-t-il une double pratique ?

A l'issue de cette première question, on se posera la question de la raison expliquant le choix réalisé par le chercheur par un type de publication en particulier (moyen de diffusion de leurs textes).

Enfin, je m'attacherai à analyser et comparer les pratiques des chercheurs dans leur globalité et les pratiques des chercheurs de Lille1 sur le site HAL¹⁸ afin de comprendre s'il existe une pratique différentielle du site.

1) Des pratiques de publications distinctes et les principes associés

1.1) Les moyens traditionnels, l'auto-publication et les archives ouvertes

Les chercheurs voulant publier leurs textes ont le choix entre trois modes de publication distincts.

Tout d'abord, ces chercheurs peuvent choisir le mode de publication traditionnel (les revues traditionnelles) car c'est un moyen assurant une caution scientifique. Il donne une validation scientifique certaine, ce qui rassure les chercheurs.

Cependant, avec le mode de publication traditionnel, l'auteur est obligé de céder ses droits et il doit payer pour avoir accès aux articles publiés dans une revue.

Ensuite, les chercheurs peuvent choisir « l'auto-publication ». Cette auto-publication consiste pour un chercheur à publier leurs travaux sur un site personnel dont il assure la plupart du temps seul la maintenance. L'avantage de ce mode de diffusion est qu'il est relativement simple à mettre en place et à utiliser.

L'inconvénient majeur est le manque de visibilité du document déposé sur ce site-là. En effet, le site n'est pas forcément bien indexé, et pas forcément facile à trouver.

¹⁸ Site Multidisciplinaire d'archives ouvertes « Hyper Article en Ligne ».

Le deuxième inconvénient est l'absence de pérennité : il n'y a aucune garantie que le site restera longtemps en ligne et que le document sera consultable et disponible sur le long terme.

Enfin, les chercheurs peuvent privilégier un dépôt en archive ouverte, soit dans une archive institutionnelle, soit dans une archive disciplinaire. Ces serveurs d'archives leur fournissent une garantie que le document sera consultable sur le long terme (stockage pérenne), et leur assurent une visibilité certaine.

1.2) Les revues en accès libre

Rappelons d'abord la définition d'une revue en accès libre : c'est une « revue répondant aux exigences de qualité des articles par la présence d'un comité de lecture mais qui a un mode de financement permettant une diffusion aussi large que possible, sans restriction d'accès et d'utilisation » [11].

D'après l'étude de Daniel Bourrion [10], ces revues en libre accès ont l'avantage de structurer les publications. Elles organisent les textes d'une façon particulière afin d'en assurer une visibilité maximale. En plus de cela, elles ajoutent un cadre introductif et des critiques qui enrichissent la réception et la compréhension du texte en question.

Ces revues en libre accès répondent parfaitement à l'inquiétude des chercheurs concernant la validation et l'évaluation scientifique.

Certaines revues sélectionnent par exemple leurs articles dans les archives ouvertes et les soumettent à un comité de lecture pour validation. C'est notamment le cas de « L'overlay Journal ».

A ce jour, 3 grands portails permettent de consulter des revues en libre accès :

- Revues.org
- PERSEE
- Erudit.org

1.3) Le principe « d'auteur-payeur » vs le principe de « lecteur-payeur »

Le principe d'auteur-payeur est le principe utilisé pour les serveurs d'archives ouvertes, ou pour les revues en accès-libre. Avec ce principe, l'auteur ou son institution d'appartenance donne une contribution à l'éditeur pour rendre l'article accessible librement à tout lecteur.

Par contre, le principe de lecteur-payeur correspond au modèle traditionnel de l'édition. Le lecteur ne peut avoir accès qu'aux revues auxquelles il a soumis une souscription payante, ou auxquelles son institution a soumis un abonnement auprès d'un ou plusieurs éditeurs.

2) Le dépôt des chercheurs

2.1) Lieu de dépôt privilégié par les chercheurs

D'après B. Bégault [12], il y a une véritable méconnaissance de la part des chercheurs en Sciences de l'Ingénieur qui ne savent pas qu'ils peuvent auto-archiver leurs pré-print sur des sites d'archives ouvertes. Ainsi, ils n'auraient pas de connaissances précises quant aux moyens de communication et de diffusion mis à leur disposition sur Internet.

De même, on constate la peur du changement émanant des chercheurs qui ne désirent pas changer leurs habitudes. B. Bégault exprime bien cette idée dans son article :

« Tant que la validation de la publication électronique n'est pas bien définie, 18 chercheurs, parmi les 23 interrogés, ne souhaitent pas publier ce type de document. Mais si la procédure devenait identique à celle proposée dans les revues traditionnelles, ils disent qu'ils le feraient. »

Derrière cette citation, on voit bien ici que les chercheurs désirent garder des pratiques anciennes sans rien changer parce qu'elles sont bien ancrées en eux mais aussi parce que ce sont des pratiques qu'ils connaissent très bien. En changeant ces pratiques, ils peuvent craindre de ne pas être à la hauteur des nouvelles exigences.

Au regard de cette étude, il faut tout de même relativiser. Même s'il existe une méconnaissance de la part de certains chercheurs, on remarque que certaines archives comptabilisent le plus de dépôts. D'après l'étude de J. Schöpfel [1], il existe en 2009 cinq sites d'archives ouvertes les plus importants en termes de dépôts. Je rappellerai ici les trois premiers sites pour une meilleure compréhension :

- PERSEE¹⁹ contient 259 816 documents
- Gallica dispose de 215 422 documents
- HAL a 143 341 documents

Au regard de ces trois chiffres, j'ai voulu recenser de nouveaux chiffres pour l'année 2011. Cependant, je n'ai pas pu avoir accès au nombre de documents contenu dans PERSEE. Aujourd'hui, Gallica dispose de 1 489 617 documents en ligne (livres, cartes, manuscrits, Images, Presses et Revues, Partitions, Paroles de musiques).

¹⁹ Revues Scientifiques en Sciences Humaines et Sociales

HAL dispose de 169 923 documents avec texte intégral. Ce chiffre ne prend pas en compte les « notices bibliographiques » reliées à aucun document. Avec les notices bibliographiques, HAL disposerait de plus de 200 000 documents²⁰.

Ayant moi-même réalisé des entretiens avec des chercheurs et des enseignants-chercheurs, j'ai voulu en apprendre plus sur leurs pratiques de dépôt. J'ai donc interviewé 3 chercheurs en informatique et un chercheur en Mathématique. Les 4 chercheurs m'affirment déposer leurs documents sur HAL et plus exactement sur HAL-INRIA. Certains affirment avoir plusieurs lieux de dépôts. C'est le cas notamment du chercheur « tout juste sorti du doctorat » qui m'affirme avoir déposé sa thèse sur TEL, mais aussi de la deuxième personne enquêtée qui m'affirme utiliser « GForge », outil de dépôt des codes :

« Enquêtée 2 : C'est GForge pour ce qui est du dépôt des différentes versions, par exemple un code, euh(...), après je vais plutôt utiliser des archives(...), des archivages de preprints, donc ArXiv, HAL, les moyens de(...), en fait de dépôt, de publications de l'INRIA, du CNRS, c'est surtout ça. »

Dans tous les cas, il s'agit surtout d'archives institutionnelles et bien souvent multidisciplinaires qui sont utilisés pour déposer des publications.

2.2) Le dépôt des chercheurs sur HAL

Dans cette partie, j'analyserai les pratiques des chercheurs en matière de dépôt à partir des chiffres trouvés sur HAL. Ces chiffres datent du 16 mai 2011. J'ai choisi de représenter les données chiffrées par deux styles de représentations différentes pour plus de clarté : d'une part des tableaux et d'autre part des graphiques sous forme de secteurs ou de courbes.

Je me suis intéressée dans un premier temps à l'évolution du nombre de dépôts c'est-à-dire aux dépôts par année. Dans un second temps, il m'a semblé très judicieux d'analyser les dépôts par catégorie de documents (type de publication) et par discipline de dépôt.

²⁰ D'après une comptabilisation des documents contenus dans chaque discipline (voir analyse des dépôts de HAL II, 3)

2.2.1) Le dépôt par année de publication

Année de publication	Nombre de documents déposés dans HAL
1930	41
1933	68
1936	91
1939	85
1942	30
1945	51
1948	61
1951	199
1954	274
1957	159
1960	257
1963	330
1966	371
1969	443
1972	510
1975	627
1978	1290
1981	1417
1984	1567
1987	1928
1990	1539
1993	2248
1996	2257
2000	3135
2003	7271
2006	15642
2009	16179
2010	16349

Tableau1 : Nombre de dépôts par année de publication

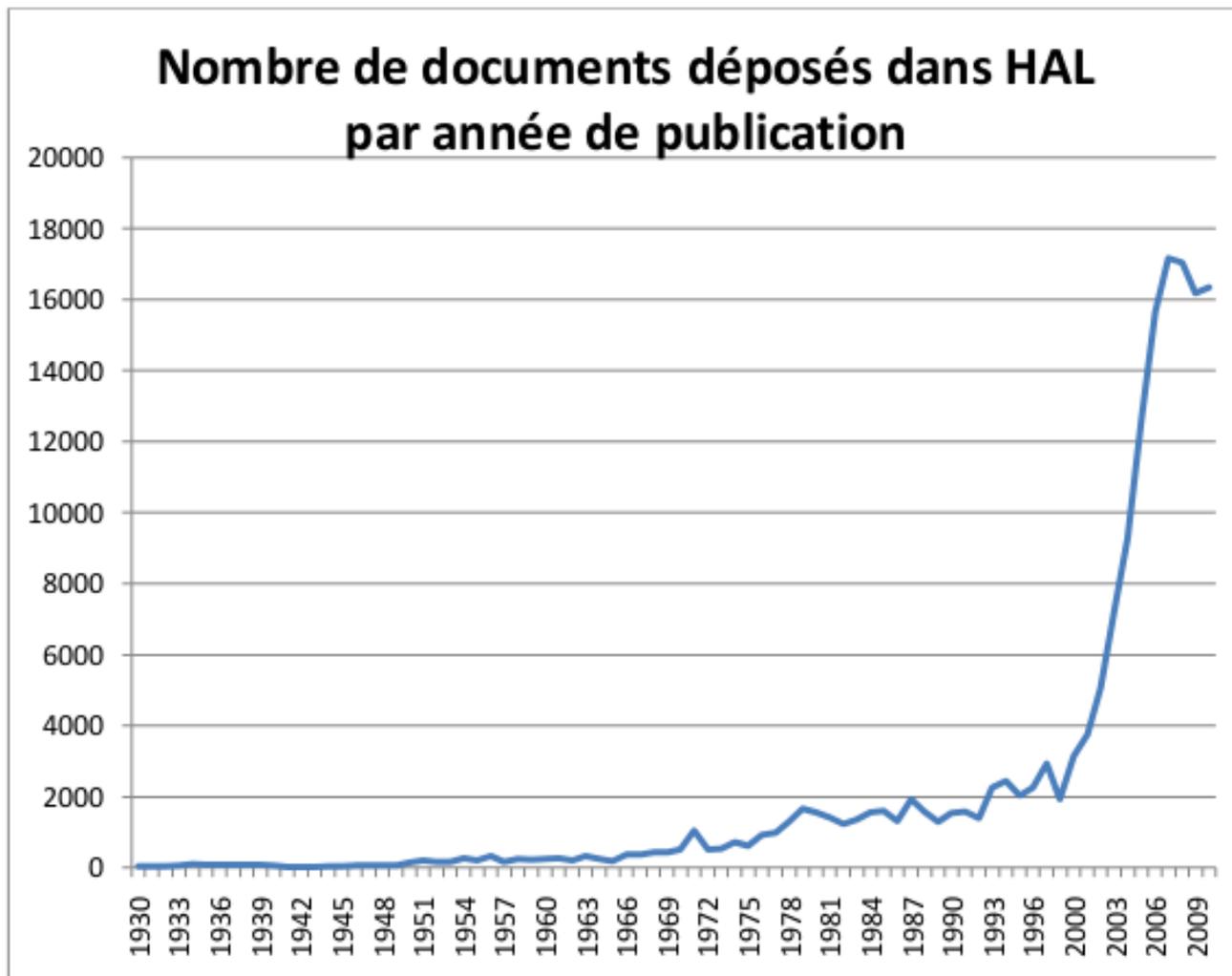


Figure1 : Graphique du nombre de documents déposés dans HAL par année de publication

Au regard du graphique et du tableau ci-dessus, on voit une évolution significative du nombre de dépôts. On remarque trois tendances :

- une évolution par petits « pics » de 1965 à 1998 : cela signifie qu'il y a très peu de dépôts de documents datant de 1965 ou moins.
- une évolution continue de 1998 à 2007 : aujourd'hui, on dépose beaucoup plus les documents datant de cette période, car à cette époque, il n'y avait pas de système d'archivage et les chercheurs désirent garder une trace de leurs documents quelque part par une sauvegarde.
- une légère baisse du nombre de dépôts à partir de 2007 : on peut relier cela à la création progressive et à l'augmentation des archives institutionnelles qui ne dépendent pas forcément de l'interface HAL. Par exemple HAL-INRIA a été créée en 2005 et d'autres institutions ont suivi le mouvement.

On sait pourtant que la création du site date de 2001. Ici, il y a seulement la date de publication du document, mais aucune date référent à la date de dépôt précise.

En ayant eu la date de dépôt des documents, on aurait pu voir une évolution débutant en 2001. Il faut donc relativiser ces évolutions. Elles montrent seulement qu'il est possible aujourd'hui de déposer des documents numériques antérieurs à 1990. La date de publication ici importe très peu, le but étant l'archivage des documents.

Par ce graphique, on peut constater qu'il y a eu une volonté de publier en libre accès. Étant un serveur central, HAL rassemble différentes disciplines, ce qui explique cette évolution immédiate. De même, HAL est un serveur rassemblant différents sites institutionnels et il impose le dépôt sur HAL et sur le site institutionnel, ce qui donne un nombre de dépôts plus important. On peut parler d'un « gonflement du nombre de dépôts ».

De même, cette évolution quasi-instantanée peut s'expliquer par un phénomène de mode comme peut le montrer l'enquête 4, et par une diffusion libre et ouverte à tous qui n'est pas « verrouillée ». Derrière cela, il y a aussi l'esprit de partage (partager ses travaux) :

Enquête 4 : « Enfin, on a du mal maintenant à savoir comment ça pourrait marcher maintenant mais si on n'avait pas accès aux travaux des autres, eh ben en général on ne les citerait pas ou beaucoup moins ou on les étudierait moins donc aujourd'hui le dépôt sur HAL accélère la recherche et puis c'est aussi naturel quand on est payé par des pouvoirs publics, c'est aussi normal que ce soit accessible à tout le monde. On n'est pas payé par les éditeurs donc c'est normal que ce ne soit pas verrouillé.»

2.2.2) Le dépôt par type de documents

Le tableau et le graphique ci-dessous illustre la volonté que j'avais à identifier les types de documents que ces dépôts contiennent.

Type de document	Nombre de documents déposés	Nombre de documents déposés en %
Rapport de recherche	9398	5,09 %
Articles dans les revues avec comité de lecture	92305	49,96 %
Articles dans les revues sans comité de lecture	1551	0,84 %
Conférences invitées	1257	0,68 %
Communications avec actes	27092	14,66 %
Communications sans actes	4391	2,38 %
Ouvrages scientifiques	508	0,27 %
Chapitres d'ouvrages scientifiques	4431	2,40 %
Directions d'ouvrages scientifiques	27	0,01 %
Autres publications	3118	1,69 %
Thèses	21720	11,76 %
HDR	1388	0,75 %
Cours	309	0,17 %
Préprints	17272	9,35 %
Total	184767	100 %

Tableau2 : Nombre de documents déposés sur HAL par type de publications

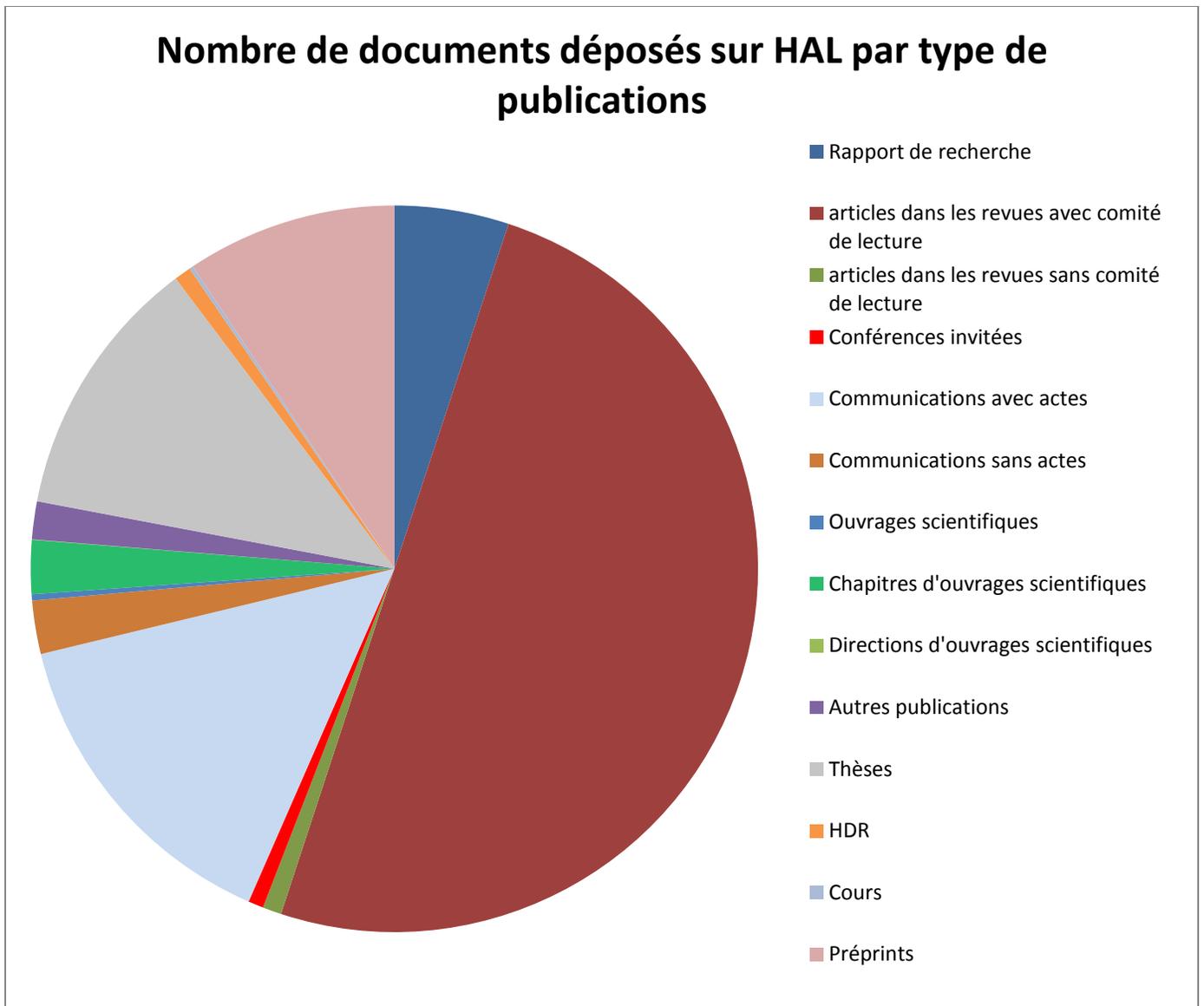


Figure2 : Nombre de documents déposés sur HAL par type de publications

De toute cette classification, il en ressort une réelle diversification des types de documents que l'on peut déposer. De plus, les chercheurs déposent non pas un seul type de documents mais plusieurs types distincts. On le peut le voir notamment dans les deux citations ci-dessous :

« Enquêteur : Que déposez-vous comme type de publication ?

Enquêté 3 : On dépose les articles dans les conférences, les articles dans les revues, les démonstrations, les rapports de recherche, les thèses je pense que c'est à peu près tout.

Enquête 4 : Donc ma thèse et mes publications qui sont soit des articles dans des journaux, soit des travaux présentés dans des conférences ».

Sur le graphique, on remarque également une disparité entre les dépôts. Le site mentionne le fait de pouvoir trouver plusieurs types de documents, mais en regardant de plus près, il semble évident qu'on ne trouve principalement que des articles de revues avec comité de lecture, des communications avec actes et des thèses. Ces trois types de documents concentrent 76,38 % et les autres documents se répartissent sur 23,62 %.

Dans les entretiens que j'ai recueillis, j'ai repéré le même phénomène. La plupart des chercheurs interviewés me disent déposer principalement des articles à la question « quel type de document déposez-vous ? » :

Enquête 1 : « Je dépose des articles scientifique souvent en texte intégral mais pas tout le temps»

Enquête 2 : « Des articles scientifiques principalement, pas de communications ».

Anna Wojciechowska [12] nous indique que « 65 % des enquêtés déposent des articles référés. »

A l'inverse, les dépôts concernent très peu les ouvrages, et directions d'ouvrages, ni les cours et HDR. En effet, le pourcentage de ces quatre types de dépôt est égal 1,21%.

2.2.3) Le dépôt par discipline

Sur HAL, il y a un classement par discipline et par sous-discipline. Pour avoir une meilleure visibilité des données et pour éviter tout problème de compréhension, j'ai décidé de ne représenter que les disciplines et pas les sous-disciplines.

Discipline	Nombre de documents déposés	Nombre de documents déposés en %
Chimie	2433	1,17 %
Économie et Finance Quantitative	160	0,08 %
Informatique	35653	17,21 %
Mathématiques	17113	8,26 %
Physique	74136	35,79 %
Planète et Univers	16856	8,14 %
Sciences non-linéaires	671	0,32 %
Sciences Cognitives	1849	0,89 %
Science de l'environnement	2835	1,37 %
Sciences de l'Homme et de la Société	29857	14,41 %
Sciences de l'ingénieur	15870	7,66 %
Sciences du Vivant	9732	4,70 %
Statistiques	1785	0,86 %
Total	207165	100 %

Tableau3 : Nombre de documents déposés sur HAL par discipline

Nombre de documents déposés dans HAL par disciplines

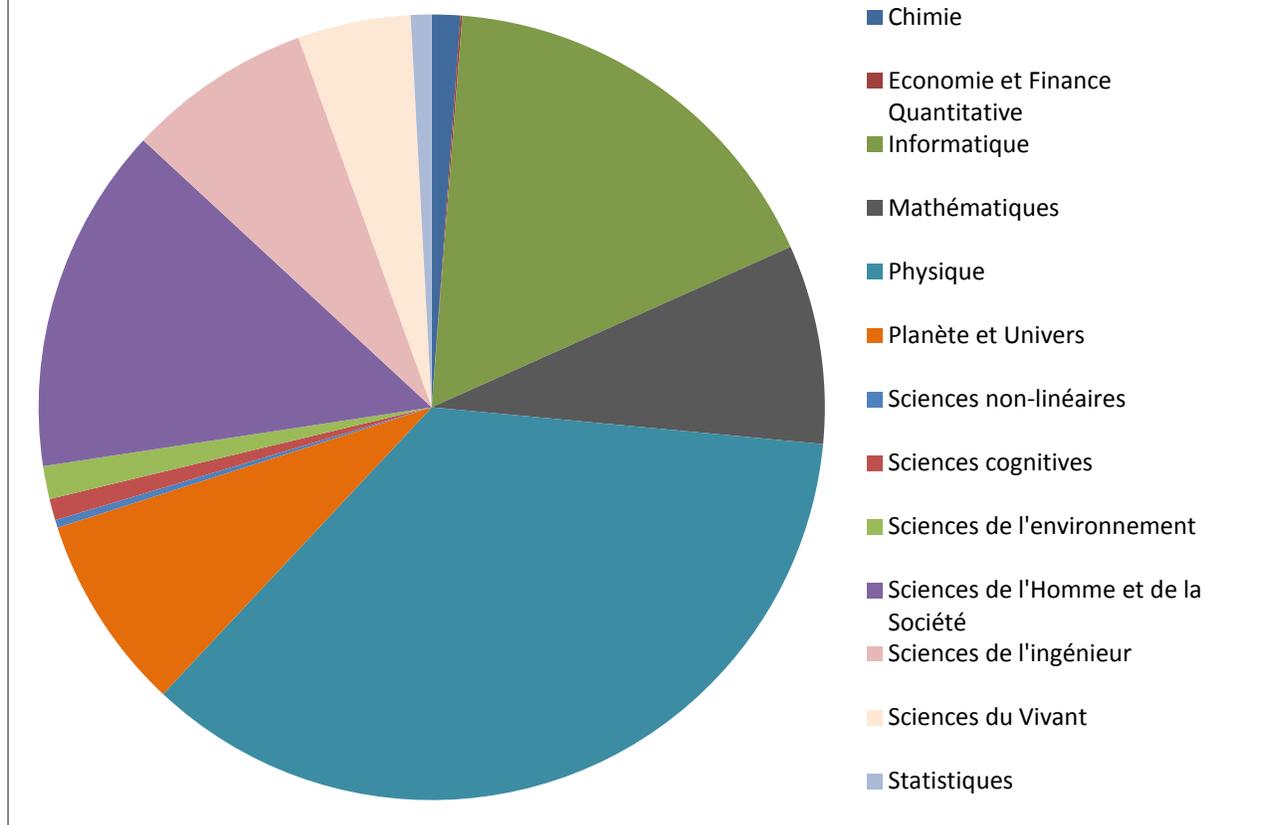


Figure3 : Nombre de documents déposés sur HAL par discipline

Il faut savoir que le dépôt par discipline concerne principalement la discipline de recherche. Si les chercheurs enseignent à mi-temps, leurs cours sont souvent liés de près à leur discipline de recherche ou à des disciplines connexes. Il est rare qu'ils déposent des documents de disciplines connexes à la leur.

Ici, sur ce graphique, on peut remarquer une répartition inégale des dépôts par discipline. La majorité des dépôts des chercheurs concerne la Physique (35,79 %), l'informatique (17,21 %) et les Sciences de l'Homme et de la Société (14,41 %).

Par contre, le dépôt est très faible pour l'Économie et Finance Quantitative (0,08%), les Sciences non-linéaires (0,32 %), les Sciences cognitives (0,89 %) et les Statistiques (0,86 %). Il s'agit surtout de disciplines où les chercheurs ont entre eux une pratique d'échange de documents très faible.

2.3) Le dépôt des chercheurs de Lille 1 sur HAL

2.3.1) Le dépôt des chercheurs de Lille1 par année de publication

Année de publication	Nombre de documents déposés dans HAL (Lille1)
1971	1
1974	1
1975	1
1976	2
1980	2
1981	4
1982	9
1983	5
1984	9
1985	9
1986	6
1987	6
1988	4
1989	8
1990	6
1991	13
1992	9
1993	6
1994	17
1995	7
1996	16
1997	24
1998	30
1999	36
2000	427
2001	542
2002	699
2003	766
2004	1197
2005	1917
2006	2192
2007	2303
2008	1721
2009	1594
2010	1605

Tableau4 : Dépôts des chercheurs de Lille 1 sur HAL par année de publication

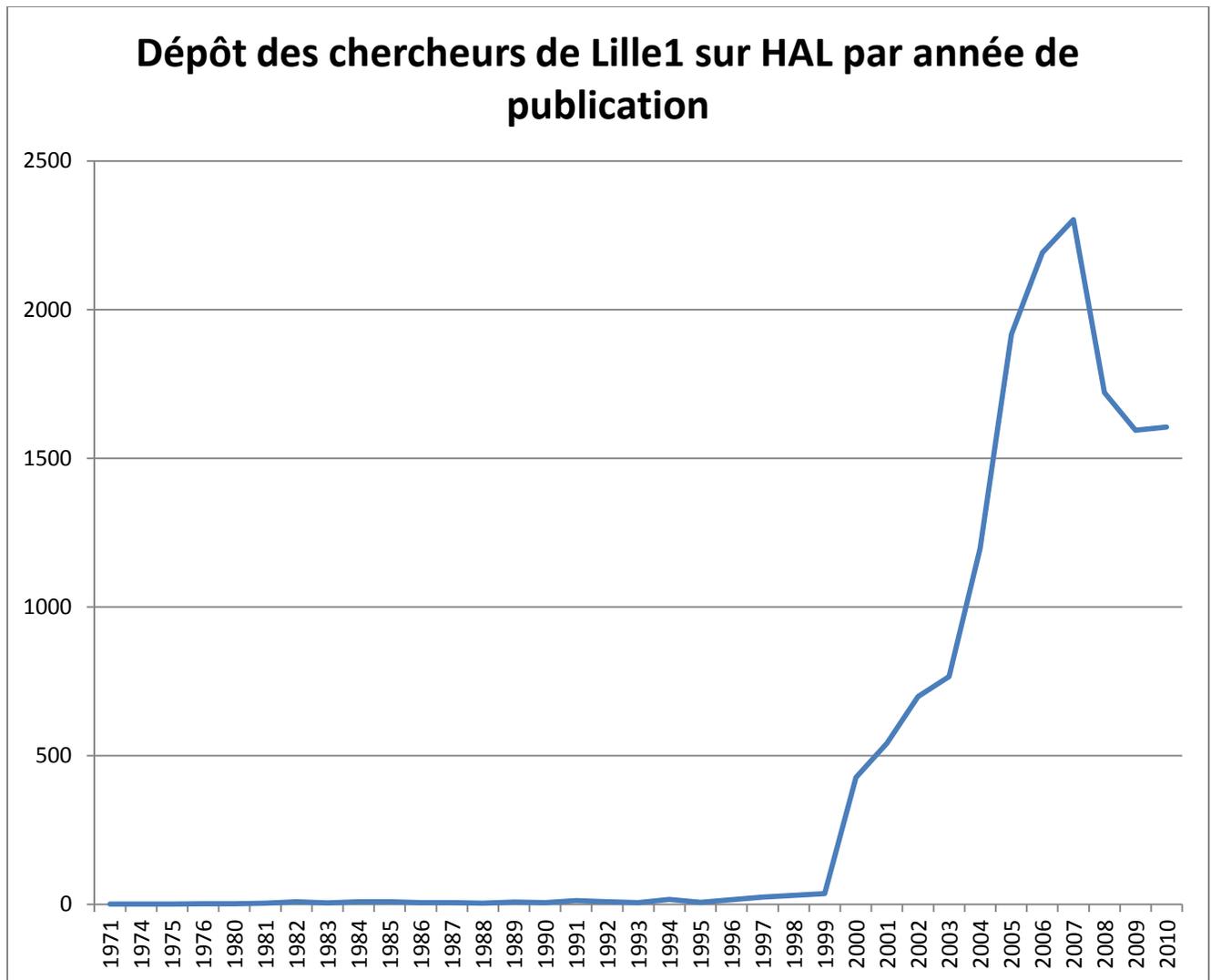


Figure 4 : Dépôt des chercheurs de Lille1 sur HAL par année de publication

Ici, on peut faire la même remarque que pour le dépôt de documents des chercheurs dans leur globalité : on retrouve sur ce graphique les dates de publication des documents et non pas des dates de dépôt. Cela peut prêter à confusion puisqu'on remarque que le site donne accès à une consultation « par date de publication, rédaction et dépôt ».

Contrairement à la Figure 1, le nombre de dépôts par année de publication est beaucoup moins important puisqu'ils ne concernent que les dépôts de l'université de Lille1 soit seulement « un pan de HAL ». Les premières publications datent de 1971 alors que sur la figure 1, elles dataient de 1930.

Par contre, on constate les mêmes tendances que sur la Figure 1. Globalement, on observe que l'on dépose peu de publications antérieures à 1999. Ensuite, il y a un nombre très important de publications datant de 1999 à 2007.

On trouve sur le site 2303 publications datant de 2007. Ce qui illustre la volonté de mettre en ligne et de préserver les documents des chercheurs. Enfin, il y a une diminution du nombre de publications datant de 2007 à 2010.

2.3.2) Le dépôt des chercheurs de Lille1 par type de document

Type de document	Nombre de documents déposés (Lille1)	Nombre de documents déposés en % (Lille1)
Articles dans des revues avec comité de lecture	6660	42,74 %
Articles dans des revues sans comité de lecture	243	1,56 %
Conférences invitées	419	2,69 %
Communications avec actes	2459	15,78 %
Communications sans actes	2879	18,48 %
Ouvrages scientifiques	205	1,32 %
Chapitres d'ouvrages scientifiques	806	5,17 %
Directions d'ouvrages	87	0,56 %
Thèses	289	1,85 %
HDR	22	0,14 %
Rapports de recherche	173	1,11 %
Brevets	108	0,69 %
Autres publications	1014	6,51 %
Documents sans référence de publication	217	1,39 %
Total	15581	100 %

Tableau5 : Dépôts des chercheurs de Lille1 sur HAL par type de document

Nombre de dépôts par type de publication pour l'université de Lille1 (date: 19 mai 2011)

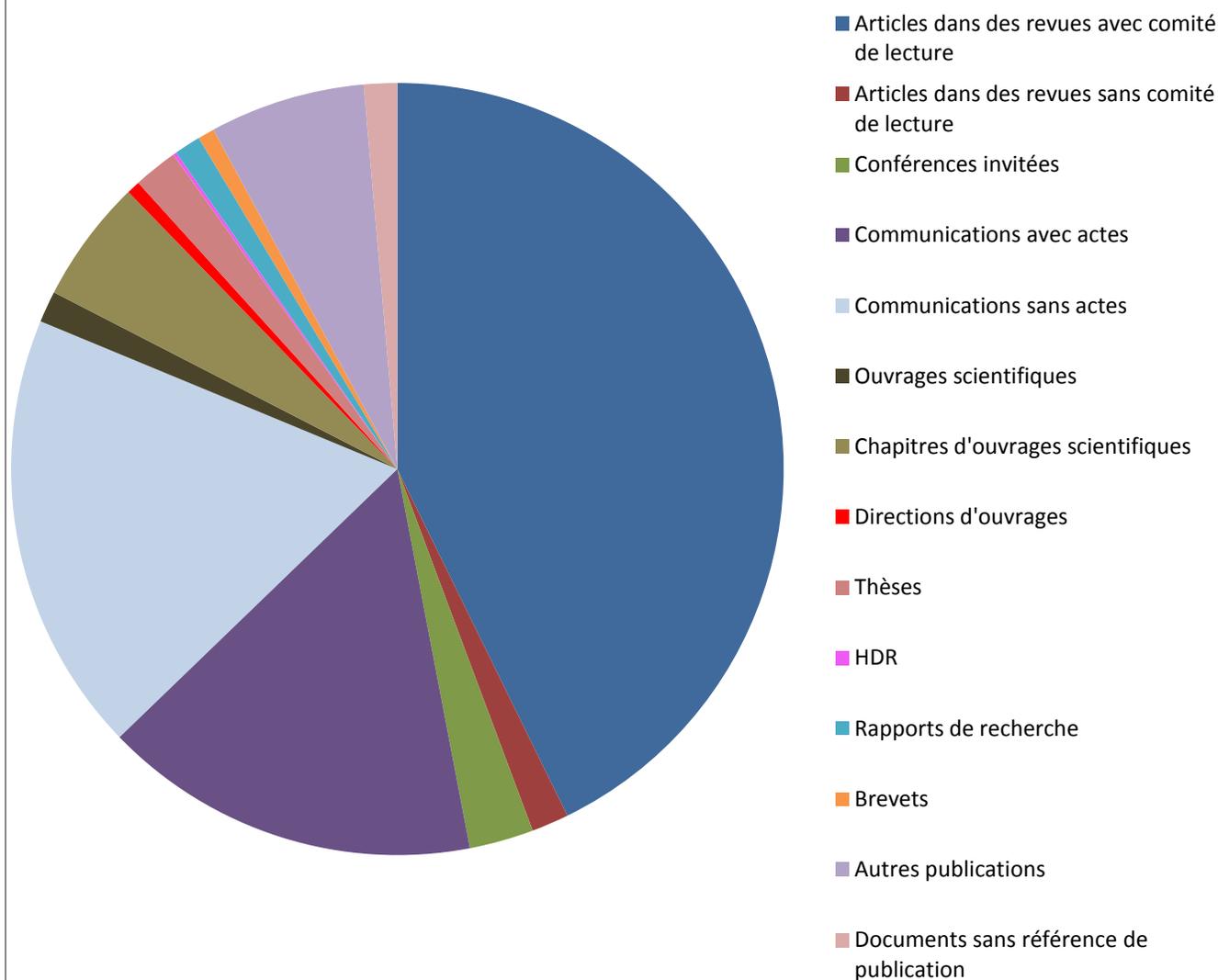


Figure5 : Dépôts des chercheurs de Lille1 sur HAL par type de document

Comme pour le graphique 2 (dépôts par type de documents des chercheurs en France), on voit bien que les dépôts des chercheurs de Lille1 concernent pour la grande majorité les articles publiés dans les revues avec comité de lecture (pour 42,74%).

On avait pu observer sur la figure 2 que les dépôts des chercheurs français dans HAL étaient des dépôts d'articles, de communications avec actes, de thèses et préprints.

A l'inverse, une grande partie de chercheurs de Lille 1 déposent des communications avec ou sans actes. Le dépôt de ce type de document correspond à un total de 34,26 %.

Soit si on comptabilise les dépôts d'articles de revues avec comité de lecture, et les dépôts de communications, on obtient 11998 documents sur le total global de 15581. Ils représentent alors 77,0% de la masse globale des documents.

De même, il y a un dépôt quasi-nul pour certains types de documents : les brevets, les HDR, les Directions d'ouvrages.

Il y a donc une réelle inégalité de répartition des documents. Cela peut s'expliquer par le fait que la communication des chercheurs à Lille1 passe d'avantage par les publications d'articles, de conférences, séminaires et communications surtout dans les sciences et technologies.

L'enquête 4 démontre bien cette idée :

« La communication entre chercheurs dépend aussi des disciplines. Je sais qu'en Informatique on se retrouve souvent dans les conférences, donc c'est notre moyen principal de communiquer et les séminaires aussi. En mathématiques par exemple, on ne publie presque pas dans les conférences, c'est dans les journaux (revues). »

2.3.3) Le dépôt des chercheurs de Lille1 par disciplines

Discipline	Nombre de documents déposés	Nombre de documents déposés en %
Chimie	1362	15,09 %
Économie et Finance Quantitative	3	0,03 %
Informatique	1643	18,21 %
Mathématiques	483	5,35 %
Physique	601	6,66 %
Planète et Univers	338	3,75 %
Sciences non-linéaire	24	0,27 %
Sciences Cognitives	15	0,17 %
Science de l'environnement	101	1,19 %
Sciences de l'Homme et de la Société	3202	35,49 %
Sciences de l'ingénieur	426	4,72 %
Sciences du Vivant	735	8,15 %
Statistiques	89	0,99 %
Total	9022	100 %

Tableau6 : Dépôts des chercheurs de Lille1 sur HAL par discipline

Nombre de dépôts de Lille 1 par disciplines dans HAL

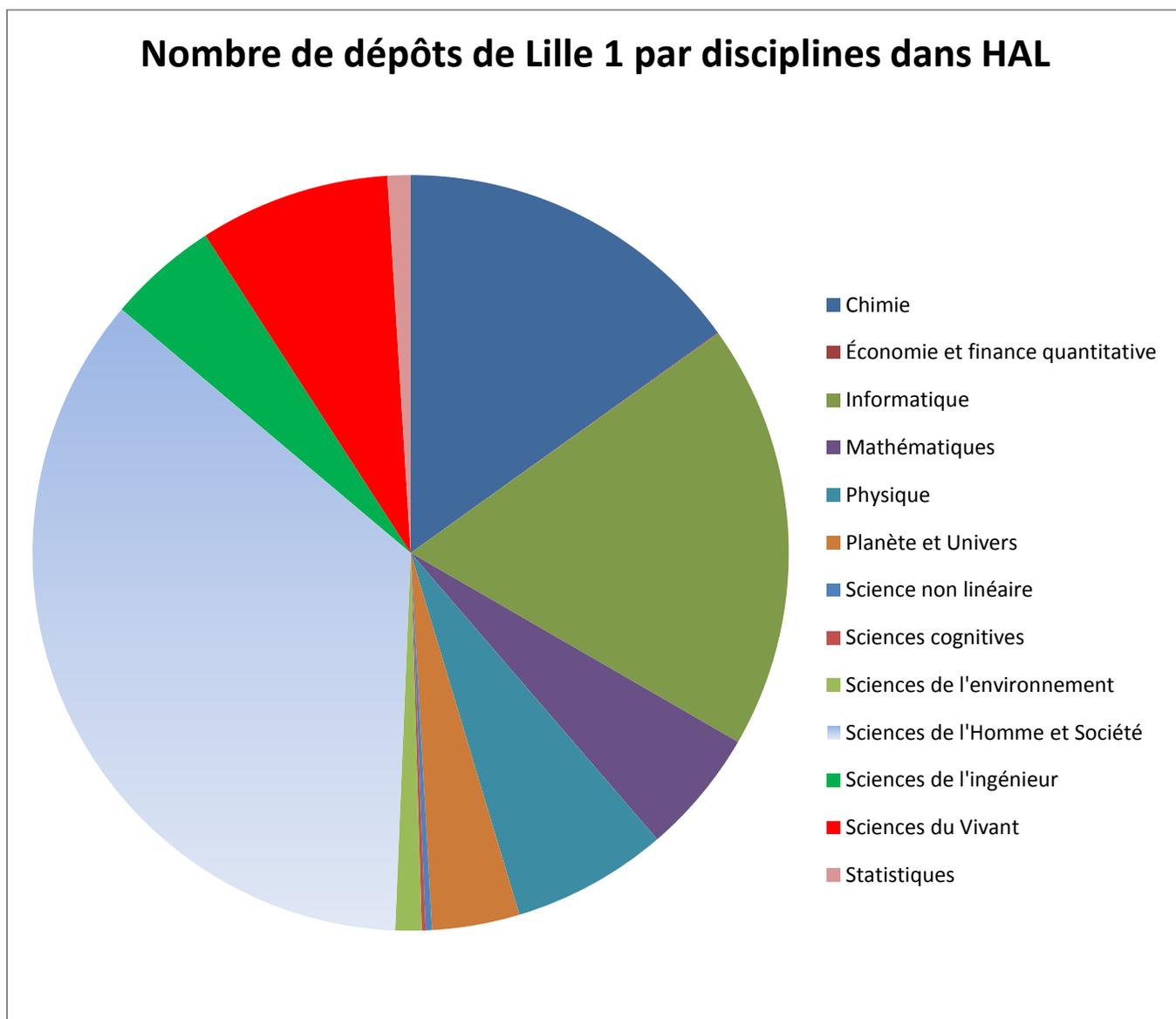


Figure6 : Dépôts des chercheurs de Lille1 sur HAL par discipline

Si l'on compare le tableau 6 avec le tableau 3, on peut entrapercevoir un nombre total de dépôts nettement inférieur au nombre de dépôts de l'ensemble. Ainsi, pour savoir la part de documents de l'établissement de Lille 1 dans l'ensemble de documents contenus dans HAL, j'ai procédé à un calcul simple : $9022 / 207165 * 100$. On obtient ainsi 4,35 %. Cela signifie donc dans cette partie que les documents de Lille 1 représentent 4,35% de la masse globale. On peut donc dire qu'ils représentent une partie minimale du site HAL.

Ensuite, voulant aller plus loin, j'ai procédé à une comparaison des disciplines les plus représentées et des disciplines les moins représentées. Il faut se souvenir

que dans la figure 3 pour les disciplines les plus représentées, on retrouvait la physique (35,79 %), l'informatique (17,21%) et les Sciences de l'Homme et de la Société (14,41 %) tandis que pour les disciplines les moins représentées on avait l'Économie/Finance Quantitative(0,08 %), les Sciences non-linéaires (0,32 %), et les Sciences Cognitives(0,89 %).

Pour l'Université de Lille 1, on voit clairement d'une part que les Sciences de l'Homme et de la Société occupent une grande partie de l'espace. Ils représentent en effet 35,49 % des dépôts effectués à Lille1. Ensuite, on peut aussi distinguer que les chercheurs de Lille1 déposent beaucoup en informatique (18,21%) et en Chimie (15,09%). Ces trois secteurs ont donc des pratiques de dépôts mais aussi des pratiques d'échange et de communication entre eux très poussées.

D'un autre côté, si on regarde les disciplines les moins bien représentées dans la figure 6, on peut dire qu'on retrouve les mêmes disciplines que pour la figure 3 : Économie/Finance Quantitative (0,03%), Sciences Cognitives (0,17%), Sciences non-linéaires (0,27%). On voit donc des pratiques de dépôts très disparates selon la discipline.

On peut donc affirmer que ces trois dernières disciplines, aussi bien à Lille 1 qu'en France en général, ont des pratiques de dépôts très peu développées. La question est de savoir la raison de ces faibles pratiques ? Est-ce parce que ce sont des disciplines ayant un faible effectif de chercheurs, et du coup qui sont moins bien représentées ? Ou est-ce parce qu'antérieurement, dans ces disciplines l'échange systématique des documents papiers et des prépublications était très peu pratiquée ? On peut rapprocher cela avec l'échec constaté par la mise en place du site Cogprints²¹, échec qui fût expliqué par le « manque de culture dans le partage des prépublications en sciences cognitives et par l'absence de raison d'être une archive centrale et disciplinaire ou même par l'absence d'une politique institutionnelle d'auto-archivage ». Peut-être peut-on aussi rapprocher ces idées à un manque d'intérêt de publier en archives ouvertes par les chercheurs en Sciences Cognitives. Quoiqu'il en soit, la question mérite d'être posée et d'être approfondie, question que je n'aurai pas le loisir de traiter ici.

2.4) Facilité d'utilisation de HAL corrélé à l'âge des chercheurs

Ici, j'en suis venue à me poser la question s'il y avait une corrélation entre la facilité d'utilisation d'un site et de dépôt par les chercheurs et l'âge des chercheurs.

A ce titre, A. Wojciechowska [13] analyse la « facilité d'utilisation de HAL » en fonction de l'âge des chercheurs. Ici, on remarque que quel que soit l'âge, il y a une certaine facilité d'utilisation de HAL. Dans la tranche d'âge des 30-40ans et des plus de 50ans, ils éprouvent quelques difficultés d'utilisation. Cependant, ils

²¹ Pour rappel, Cogprints a été créée en 1997 et il s'agit de la première archive permettant de déposer des publications en Sciences Cognitives

représentent une infime partie de l'ensemble des répondants, c'est pourquoi, on ne peut pas vraiment dire qu'il y a une corrélation entre l'âge des chercheurs et la facilité ou la difficulté à utiliser un site comme HAL.

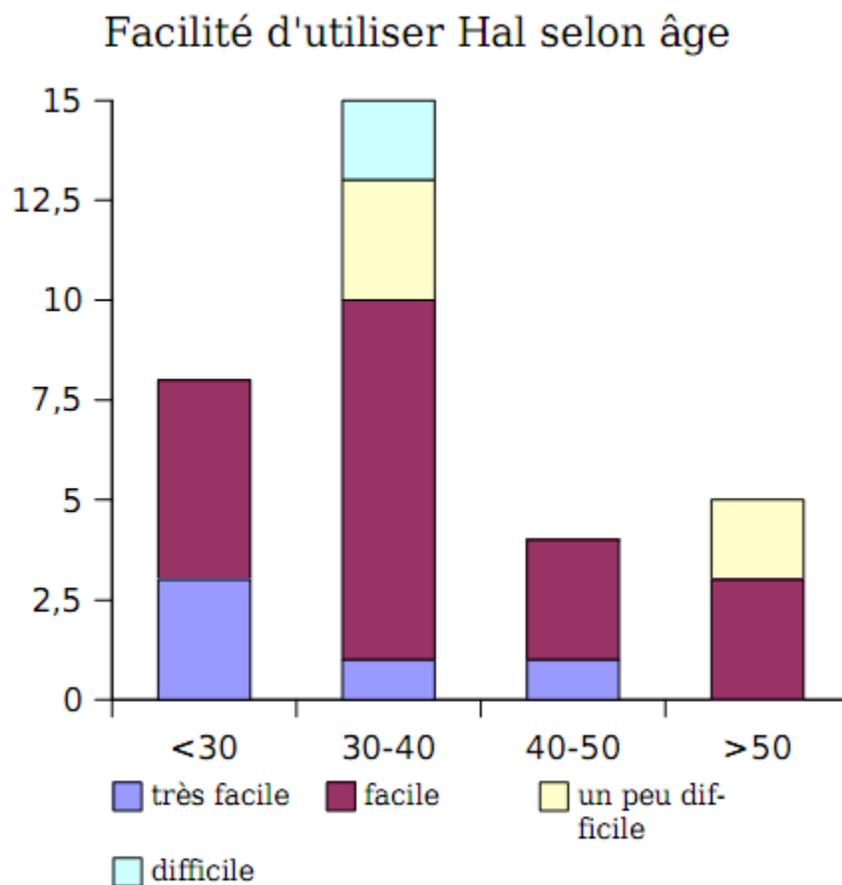


Figure7 : Schéma montrant la facilité d'utilisation de HAL selon l'âge des personnes

Effectivement, HAL dispose d'une interface graphique assez simple d'utilisation et de couleurs assez attrayantes qui permettent aux utilisateurs de se rendre plus facilement sur le site.

2.5) Le dépôt des chercheurs corrélé au nombre de dépôt et au temps de dépôt

2.5.1) Chercheurs et nombre de dépôts

Dans cette sous-partie, je m'intéresse au nombre de dépôts par chercheur. En posant cette question, je voulais voir si tous les chercheurs déposaient le même

nombre d'articles, ou s'il y avait un nombre plus important selon le niveau du chercheur.

En réalisant cela, je me suis rendue compte d'une part que le nombre de dépôt variait selon le chercheur, et d'autre part que les dépôts n'étaient pas forcément des dépôts individuels. A la question « quel est le nombre de dépôts que vous effectuez par an ? », voici les réponses des quatre enquêtés :

« Enquêté 1 : Par an peut-être 10 à 20, quelque chose comme ça, du moins pour l'équipe.

Enquêtée 2 : Alors là, ça dépend de la production, pas beaucoup, un ou deux »

Enquêtée 3 : Ben, ça va beaucoup dépendre du taux d'acceptation qu'on a, donc euh(...), comme on dépose tout ce qu'on a d'accepté. On a des années où nos articles sont plus acceptés que d'autres.

Enquêté 4 : Environ 3 où je suis co-auteur. »

De ces 4 réponses, il en ressort trois types de dépôts :

- un dépôt individuel
- un dépôt par équipe
- un dépôt en tant que co-auteur

De même, le nombre de dépôts varie de 1 à 4 au maximum par personne.

Aussi, A. Wojciechowska [13] montre dans son article que les chercheurs déposent sur HAL en majorité 2 à 3 articles publiés. Elle indique que 82,4 % des chercheurs ont déposés au moins 2 articles référés sur HAL.

Pour certains chercheurs, le dépôt sur HAL est devenu une évidence, un phénomène de mode comme peut le démontrer l'enquêté 4 :

« Enquêté 4 : Enfin, on aurait du mal maintenant à voir comment ça pourrait marcher autrement, mais si on n'aurait pas accès aux travaux des autres, eh ben en général on le citerait pas ou beaucoup moins, on les étudierait moins. »

2.5.2) Chercheurs et temps de dépôt

Ici, je voulais savoir combien de temps les chercheurs mettaient à déposer un article. Je n'ai pas pensé à intégrer cette question dans ma grille mais il en est quand même ressorti une remarque très pertinente de la part d'un des enquêtés :

« Enquêté 4 : Il faut mettre en ligne un papier, il faut tout renseigner, enfin quand on l'a fait une fois, ça va, mais au début, il faut essayer et y arriver, ça peut prendre du temps. »

Par ailleurs, A. Wojciechowska dans son enquête [13], nous affirme que le premier dépôt prend plus de 30 minutes pour 34,5 % des personnes et plus d'une heure pour 20,1%. Au contraire, le dépôt suivant prend moins de 15 minutes pour 81,5 % des personnes. On peut donc constater que le dépôt prend de moins en moins de temps au fil des dépôts successifs du fait qu'on commence à savoir quels éléments il faut renseigner et qu'il faut plus trop chercher ces éléments.

3) La Consultation des chercheurs

3.1) Lieu de consultation et de recherche d'information privilégié par les chercheurs

D'après l'enquête réalisée par B. Bégault [14], les chercheurs des sciences de l'ingénieur consultent en grande majorité des revues numérisées (30 % des chercheurs interrogés) et des bases de données bibliographiques.

Par contre, B. Bégault [14] a pu constater une utilisation moindre des serveurs d'archives ouvertes qu'elle appelle « Bases de prépublications ». Cela serait dû selon elle à une faible connaissance des outils mis à la disposition des chercheurs mais aussi à un manque d'informations.

On peut se demander si on observe le même phénomène dans les autres disciplines et rechercher dans ce cas la raison.

Par opposition, A. Wojciechowska [12] montre que 58 % des chercheurs interrogés (en Mathématiques et Informatique) ont recours à des articles ou documents venant du site HAL.

Parmi les personnes que j'ai interviewées, deux personnes m'indiquent qu'elles consultent sur HAL, un chercheur me dit qu'il y consulte rarement sur HAL, un autre m'affirme qu'il ne consulte jamais sur HAL ou sur d'autres sites d'archives ouvertes.

Dans tous les cas, les personnes interrogées m'affirment préférer consulter directement à partir d'un moteur de recherche « qui indexe bien les documents » : « Enquêté 4 : Non en fait, moi je cherche toujours par moteur de recherche et HAL est bien indexé, donc quand c'est sur HAL, c'est souvent sur le moteur de recherche. Il y a juste à taper un mot-clé sur le moteur de recherche et on regarde ce qui en ressort »

Au vu de mes entretiens, le chemin privilégié pour accéder à un document est bien souvent le moteur de recherche GOOGLE.

« - Enquêté 1 : je consulte rarement sur HAL, rarement car je pense que ça va plus vite, ça dépend si les articles sont indexés ou pas ou s'ils sont disponibles ailleurs.

- Enquêteur : S'ils sont disponibles ailleurs, vous allez les voir sur quels sites ?

- Enquêté 1 : Ben ça dépend de ce que Google me propose. »

J. Schöpfel [15], dans le rapport DUAO partie 2 qu'il a réalisé, étudie les pratiques des chercheurs qui utilisent IRIS. Concernant les chemins d'accès utilisés, il montre que « 44% du trafic vers IRIS arrive via GOOGLE contre seulement 26% qui accèdent directement au site.»

A. Wojciechowska [12] dans son article sur les usages des archives ouvertes montre que « parmi les sites proposés comme points d'accès aux textes intégraux en accès ouvert, la majorité des chercheurs répondent GOOGLE (66%) »

De même, quand les chercheurs citent un moteur de recherche, ils citent directement GOOGLE comme si c'était le seul chemin d'accès possible.

On voit donc bien qu'il y a une omniprésence de GOOGLE : il occupe une grande part du champ visuel des chercheurs.

3.2) Les caractéristiques de la consultation utilisée par les chercheurs

Dans cette partie, je me suis intéressée à la manière dont les chercheurs consultent sur HAL. Je voulais ainsi comprendre quel critère de sélection ils utilisaient sur HAL, quelle navigation préféraient-ils. Pour comprendre cela, il faut d'abord expliquer les services de consultations mis en place par HAL

3.2.1) Les Services de consultation disponibles sur HAL

L'interface de HAL donne la possibilité aux utilisateurs d'accéder à un moyen immédiat de consultation de leur choix. On peut répertorier 9 classements distincts comme on peut le voir sur l'image 1.

Ainsi, on peut choisir de consulter « par domaine » (ou discipline). Dans ce type de classification, chaque domaine est répertorié dans un dossier, et les sous-domaines sont eux-mêmes classés dans des sous-dossiers. Pour arriver au sous-domaine, il suffit de cliquer sur le « + », et pour accéder aux documents, il suffit de cliquer sur le nom de la discipline.

Ensuite, il est possible de consulter les trente derniers dépôts. Ici l'opération est toute simple puisqu'il suffit seulement de cliquer sur le classement « les trente derniers dépôts ».

Outre cela, on trouve aussi un classement « par année de publication, rédaction, dépôt » qui permet d'accéder aux publications d'une année voulue. Ici,

comme pour le premier type de classement, le fonctionnement est le même, et le rangement est réalisé sous forme de dossiers principaux. Ces dossiers prennent la forme d'un calendrier faisant ainsi référence à la date de publication.

La classification « par type de document » fonctionne de la même manière. La seule différence, c'est le symbole utilisé pour représenter cette classification particulière (livre ouvert).

Par contre les classifications « par collection », « les portails de l'archive ouverte HAL », par établissement fonctionne sous forme de liens indirects, chaque lien est classé dans un ordre alphabétique. Pour accéder au lien, il suffit simplement de cliquer dessus.

En outre, il y a aussi des liens directs vers les sites d'archives ouvertes « ArXiv » et « TEL ».

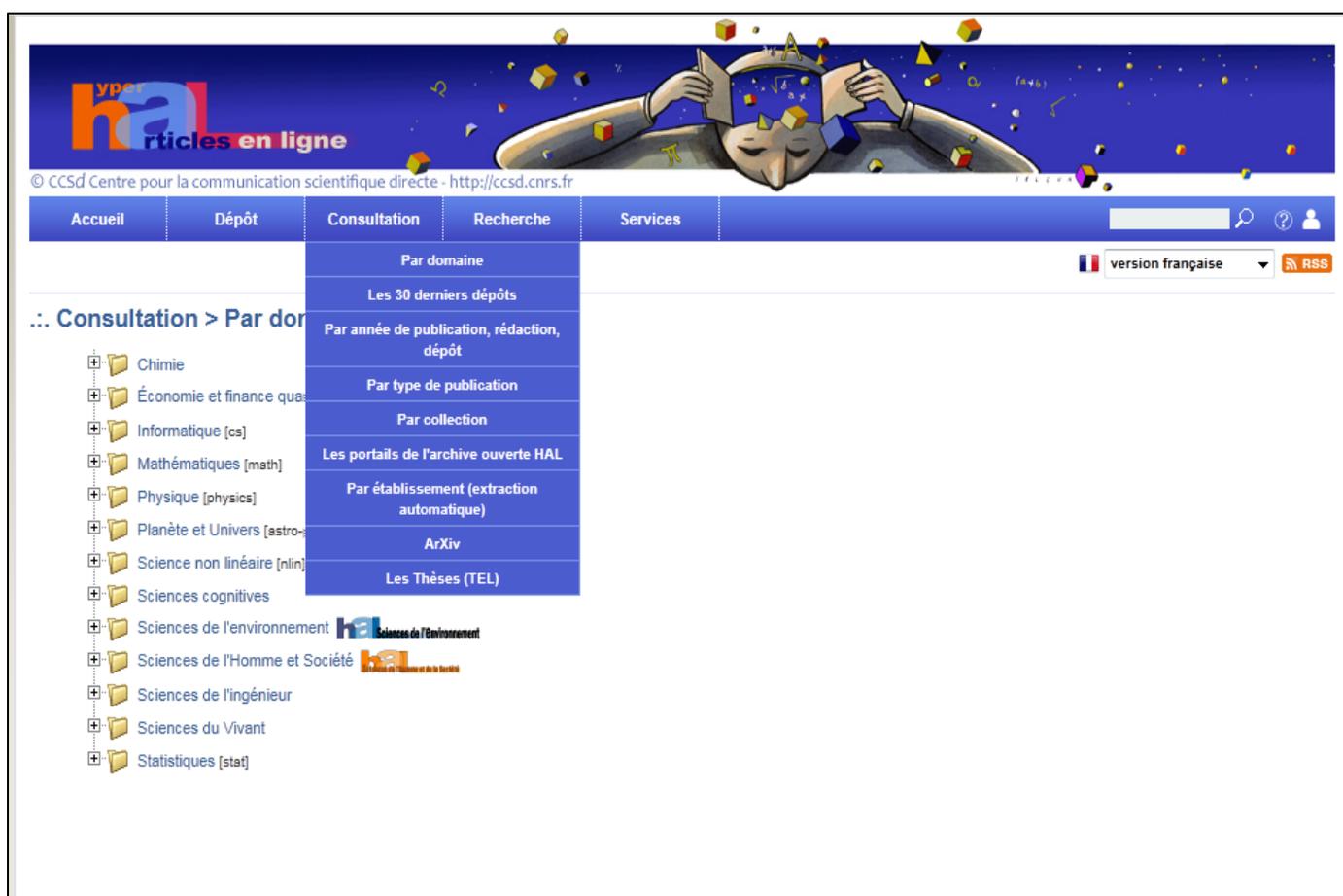


Image1 : Les différents types de consultation possible sur HAL

3.2.2) Quel type de consultation est privilégié par les chercheurs ?

Voulant en apprendre plus sur la stratégie de consultation (et donc de navigation) utilisée par les chercheurs pour arriver à leur fin, je leur ai donc posé directement la question « Comment consulter vous sur HAL ? ».

En leur posant cette question, je m'attendais à ce qu'ils me citent le classement de HAL. Pourtant, après 4 entretiens, j'ai pu constater qu'il y avait une méconnaissance du classement prodigué par HAL concernant la consultation. En effet, si la plupart me disent utiliser HAL pour la consultation et le dépôt, ils m'affirment ensuite consulter « par auteur, par titre, par mots-clés, ou par thème ». Or on voit bien que HAL n'utilise aucun de ces quatre types de classement. En fin de compte les types de classement qu'ils citent sont des classements « simples » et ces classements sont le plus souvent utilisés dans les outils de recherche simple et avancé, ou dans les catalogues des bibliothèques.

En cherchant à comprendre un peu plus la raison de cette utilisation partielle, je me suis rendue compte que les personnes interviewées « consultent un article portant le logo HAL ». Cette consultation « directe et rapide » se fait via les moteurs de recherche et surtout via GOOGLE. On comprend alors très bien que ces personnes se disent avoir utilisé l'outil de consultation. Leur consultation des articles HAL se fait par GOOGLE, et par un critère de recherche comme par exemple l'auteur, mais il est rare que ces personnes consultent des publications directement sur l'onglet « Consultation » du site HAL. Quand on leur demande la raison de l'utilisation directe de GOOGLE, les personnes répondent que c'est un outil du web plus simple et que HAL est bien indexé. Aussi en allant sur GOOGLE, elles se disent « perdre moins de temps », puisque l'accès aux PDF est « direct » la plupart du temps. Il y a donc derrière cela l'idée de « prendre le chemin le plus court ».

3.2.3) La discipline de consultation

A cette question (question « quelle discipline consultez-vous), les chercheurs ont répondu avec plus de facilité. Si le dépôt se faisait principalement dans « sa propre discipline de recherche », la consultation, elle est beaucoup plus libre et beaucoup plus ouverte. On consulte toujours sa propre discipline pour voir les « avancées », pour voir « ce que les autres ont fait, ont découvert afin de ne pas refaire la même chose ». La consultation de sa propre discipline a ici plus un objectif de vérification, ou « d'activité de veille » comme peut le montrer l'enquête 4 : « la partie veille, c'est quotidien, la partie recherche de papiers par moi-même deux fois par semaine. »

De même, si le dépôt n'était autorisé que dans sa propre discipline, la consultation laisse « libre court aux idées des personnes et à leur goûts ». Chacun

choisit en fonction de son intérêt pour un sujet comme on peut le voir dans le court dialogue ci-dessous :

-Enquêteur : Quelle discipline consultez-vous ?

- Enquêté 2 : « La Physique, la mécanique, là, c'est parce que ça m'intéresse, je m'intéresse à la mécanique des fluides, donc ça dépend de la thématique »

4) La Recherche

Dans son rapport J. Schöpfel [15] démontre que 37 % des utilisateurs qui accèdent à IRIS, effectuent des recherches. Suite à cela, il s'est alors posé la question quel modèle de recherche privilégiaient-ils ? Une recherche avancée ou une recherche simple ?

A cette question, il montre que la majorité des personnes lancent leurs recherches à partir de la recherche simple (99%). On voit donc que peu de personnes utilisent la recherche avancée.

J'ai voulu savoir si c'était également le cas sur HAL. Ne pouvant pas accéder aux fichiers log, j'ai donc demandé directement aux interviewés le type de recherche qu'ils utilisaient. A cette question, 3/4 des personnes interrogées m'ont indiqué qu'elles préféraient utiliser une recherche simple. Une des personnes m'a expliqué que « cela ne voulait rien dire pour elle » puisqu'elle utilise GOOGLE pour consulter et rechercher des documents prioritairement.

La recherche avancée est utilisée par une des personnes qu'après l'utilisation d'une recherche simple pour affiner les résultats, en cas vraiment de besoin.

III) Les Besoins des chercheurs en matière d'archives ouvertes

Par besoin, j'entends « tout dispositif technique ou service permettant à l'utilisateur d'utiliser au mieux un site ». Pour comprendre les besoins concrets des chercheurs, j'ai d'abord regardé ce qui se faisait sur les sites d'archives ouvertes en matière de services proposés. Il apparaît que tous les sites donnent un accès réservé aux personnes autorisées et authentifiées à un centre de dépôt et à un espace personnel. De même, chaque plateforme d'archives ouvertes propose un espace de consultation avec divers critères. Cependant, selon les sites, ces espaces n'ont pas les mêmes outils.

1) L'exemple de HAL

Outre l'aspect attrayant de la page d'accueil et du menu défilant, HAL²² donne accès à une multitude de services.

En effet, comme on a pu le voir un peu plus haut, le menu donne accès à un **espace de consultation** permettant de naviguer dans les documents déposés par un critère de son choix : par domaine ; par année de publication ; par type de publication ; par établissement ; par collection ... Elle donne aussi un accès aux 30 derniers dépôts ou aux portails « partenaires ».

Aussi, le site permet également de suivre les dépôts par le fil RSS. Ce fil RSS et la visualisation des 30 derniers dépôts sur la page d'accueil permet donc de se tenir au courant des dernières nouveautés.

Outre cela, elle fournit un outil de **recherche** permettant aux utilisateurs d'effectuer une recherche simple, recherche avancée, recherche par identifiant de l'article, ou une recherche de thèses sur TEL.

Le site HAL dispose aussi d'un onglet « **Dépôt** » permettant aux chercheurs de s'identifier et de déposer. Une fois authentifié, cet onglet donne accès à un « espace personnel » nommé « **mon espace** ». Comme on peut le voir sur l'image 2, cet espace permet d'accéder directement à ses articles, ses notices bibliographiques, ses documents confidentiels. Outre cela, cet onglet permet d'obtenir ses propres statistiques de dépôt, ou de retrouver ses alertes ou encore d'accéder à « son compte ». L'accès à ce compte permet de modifier ses données personnelles.

²² Voir le site suivant pour plus d'informations : <http://hal.archives-ouvertes.fr/>

Sous l'onglet « **Services** », elle donne la possibilité à ses utilisateurs de créer une page web, d'exporter une liste de publications sous différents formats (PDF, doc, tex, ...), de consulter les différents laboratoires connus de HAL, ou encore de tester la compilation « TeX/LaTeX ».

Cependant elle propose aussi aux usagers de télécharger ou visionner un manuel d'informations, une documentation usager/ WebServices via l'onglet d'aide.

Outre cela, elle donne aussi le choix aux usagers de choisir entre deux langues : le Français et l'Anglais.

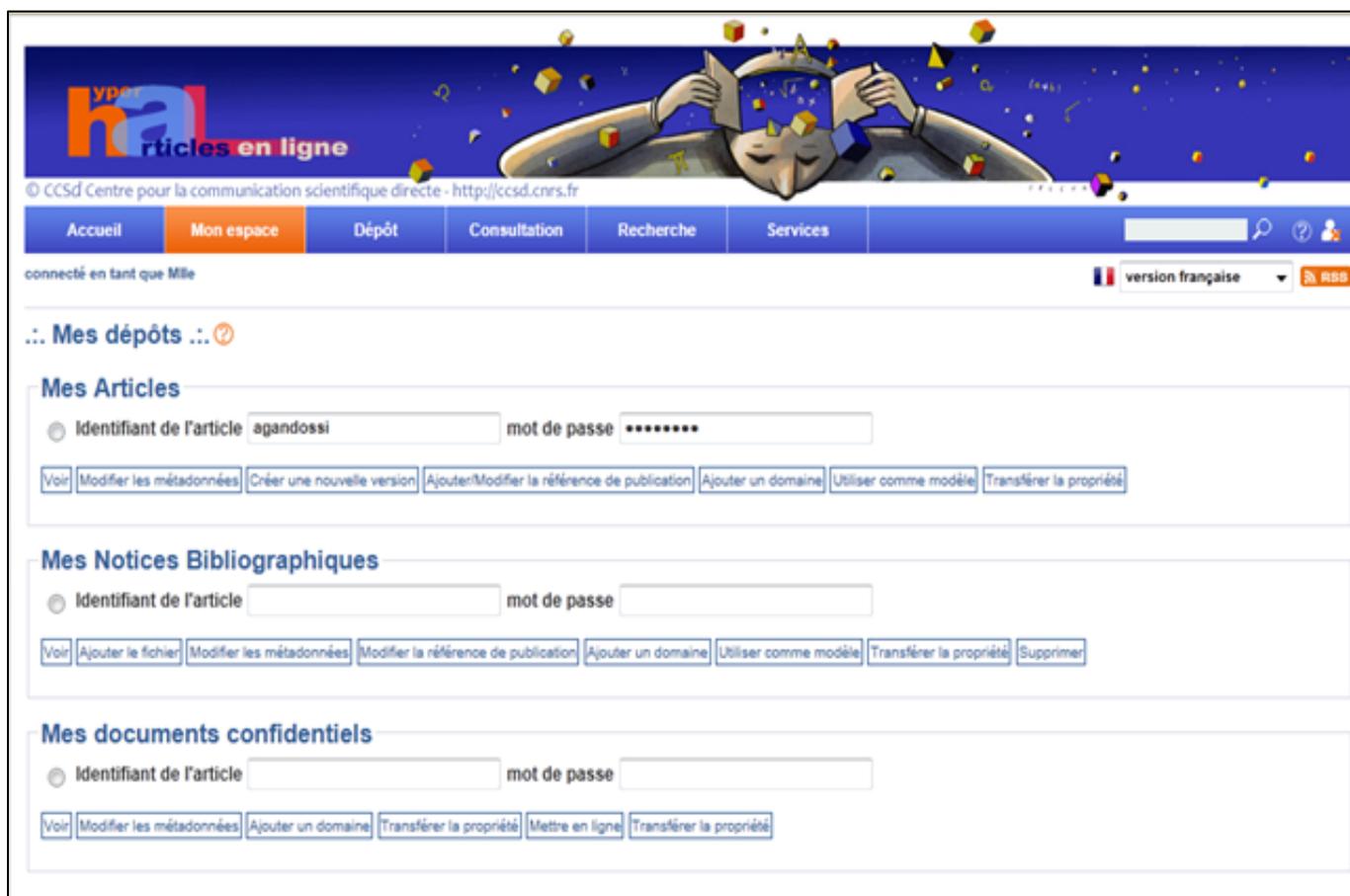
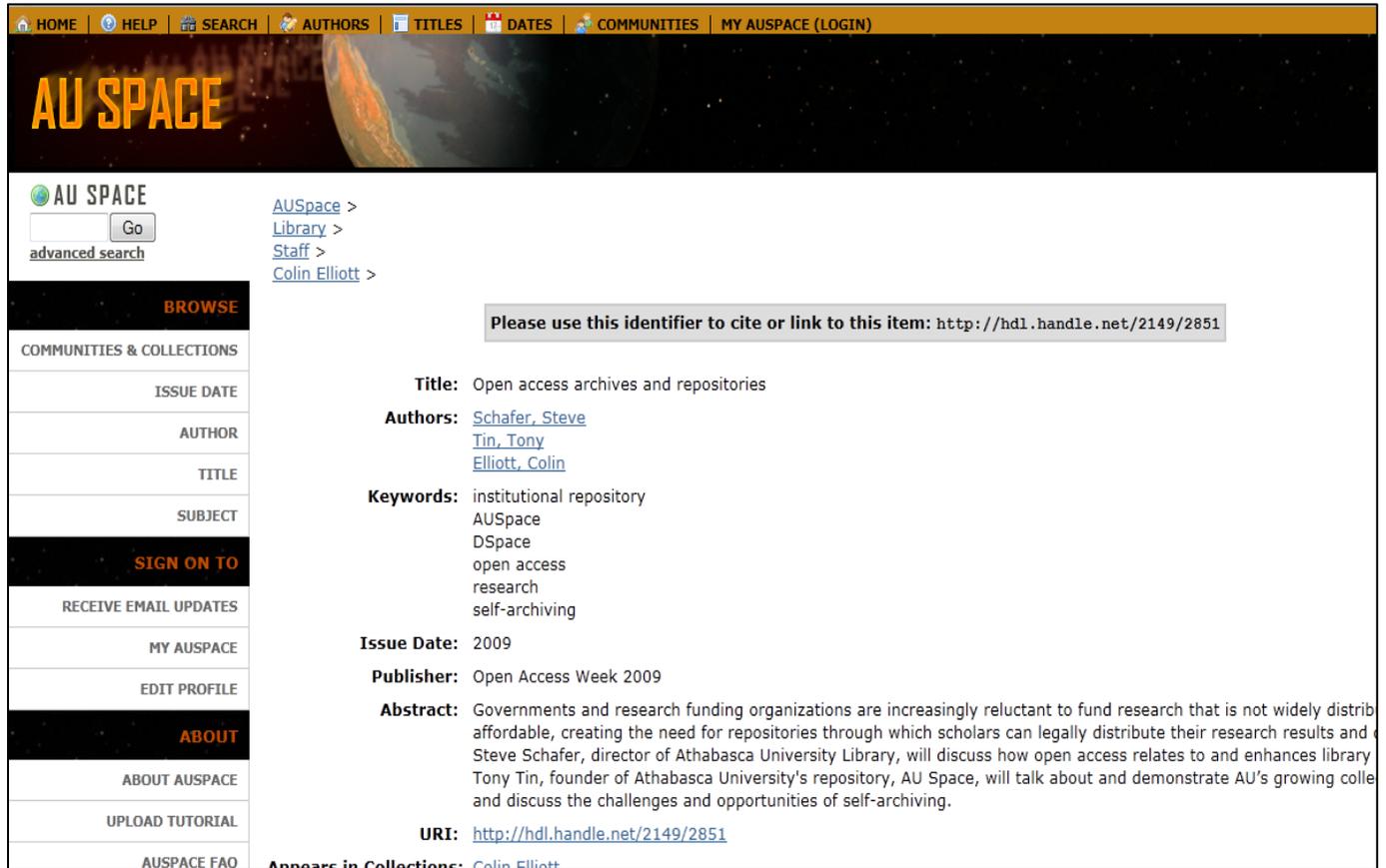


Image2 : "Mon Espace"

2) L'exemple de AUSpace



The screenshot shows the AUSpace website interface. At the top, there is a horizontal navigation menu with links for HOME, HELP, SEARCH, AUTHORS, TITLES, DATES, COMMUNITIES, and MY AUSPACE (LOGIN). Below this is a large banner with the text "AU SPACE" and a background image of a planet. On the left side, there is a vertical navigation menu with sections: BROWSE (containing COMMUNITIES & COLLECTIONS, ISSUE DATE, AUTHOR, TITLE, SUBJECT), SIGN ON TO (containing RECEIVE EMAIL UPDATES, MY AUSPACE, EDIT PROFILE), and ABOUT (containing ABOUT AUSPACE, UPLOAD TUTORIAL, AUSPACE FAQ). In the center, there is a search bar with a "Go" button and a link to "advanced search". To the right of the search bar, there are links for "AUSpace >", "Library >", "Staff >", and "Colin Elliott >". Below the search bar, there is a grey box with the text "Please use this identifier to cite or link to this item: http://hdl.handle.net/2149/2851". The main content area displays the following metadata:

- Title:** Open access archives and repositories
- Authors:** [Schafer, Steve](#), [Tin, Tony](#), [Elliott, Colin](#)
- Keywords:** institutional repository, AUSpace, DSpace, open access, research, self-archiving
- Issue Date:** 2009
- Publisher:** Open Access Week 2009
- Abstract:** Governments and research funding organizations are increasingly reluctant to fund research that is not widely distrib affordable, creating the need for repositories through which scholars can legally distribute their research results and Steve Schafer, director of Athabasca University Library, will discuss how open access relates to and enhances library Tony Tin, founder of Athabasca University's repository, AU Space, will talk about and demonstrate AU's growing colle and discuss the challenges and opportunities of self-archiving.
- URI:** <http://hdl.handle.net/2149/2851>
- Appears in Collections:** [Colin Elliott](#)

Image3 : Site d'archives ouvertes AUSpace

AUSpace²³, Site d'archives institutionnelles de l'Université d'Athabasca (CANADA) contient les publications de cette université et de ses différents établissements et/ou facultés (facultés des sciences humaines et sociales, ...).

Du point de vue de l'aspect graphique et esthétique, la page d'accueil du site est assez simpliste : elle dispose d'un menu horizontal avec plusieurs modes de consultations possibles, avec un espace d'aide... mais aussi d'un menu vertical qui reprend dans ces différents services plus en détail.

En terme de services proposés, AUSpace donne accès à :

- un outil de recherche simple et de recherche avancée
- un outil de consultation permettant de rechercher par collection / établissement (communities & collections) ; par date de publication (issue date) ; par auteur (author), par titre (title) ou encore par sujet (subject)

²³ Voir le site suivant pour plus d'informations : <http://auspace.athabascau.ca:8080/dspace/>

- deux types de fiches catalographique au choix (simple/minimale ; détaillée)
- un espace personnel pour les déposant qui contient la liste de ses articles déposés ou des articles qui ont été soumis ; ou une liste de tâches à effectuer sur un (des) document(s) comme par exemple éditer, revoir ou contrôler un document.
- un espace de dépôt permettant ou déposant de soumettre leurs articles
- la possibilité de recevoir des alertes RSS ou de s'inscrire à une collection et de recevoir des e-mails lorsque de nouveaux ajouts sur le site sont effectués
- la possibilité d'éditer son profil d'utilisateur pour les personnes inscrites sur le site
- la possibilité de télécharger des fichiers sous des formats multiples (pdf, html, ...)
- une interface d'aide et de foire aux questions (espace « ABOUT »)
- un guide d'utilisation du site et les droits associés
- d'autres ressources par des liens vers le site de la bibliothèque Universitaire d'Athabasca, ou à la « [Canadian Association of Research Libraries](#) ».

Ici, sur ce site, on voit tout de même, que l'utilisateur n'a pas la possibilité de choisir la langue du site, l'anglais lui est imposé d'office.

3) L'exemple de Papyrus (Université de Montréal)

The screenshot shows the homepage of the Papyrus digital repository. At the top, there is a navigation bar with links for Répertoires, Facultés, Bibliothèques, Actualités, Sites A-Z, and Recherche sites UdeM. The main header features the University of Montreal logo and the title 'Papyrus : Dépôt institutionnel numérique'. Below the header, there is a search bar with a search button and a language selector for 'FRANÇAIS' and 'ENGLISH'. The left sidebar contains navigation options such as 'Chercher dans le dépôt', 'Recherche avancée', 'Accueil', 'Nouvel utilisateur?', 'Parcourir le dépôt par', and 'Services personnalisés'. The main content area includes a welcome message, a 'Communiqués' section with dates, and a 'Communautés de recherche et d'enseignement représentées dans Papyrus' section with a list of departments and their document counts. On the right, there is a 'À propos de Papyrus...' section with links to 'Questions fréquentes', 'Droit d'auteur et politiques d'éditeurs', and 'Comment interroger Papyrus à distance?'. Below this is a video player showing an introduction to open access, and an 'Author's Rights' section with logos for Creative Commons and SPARC.

Université de Montréal

Répertoires | Facultés | Bibliothèques | Actualités | Sites A-Z | Recherche sites UdeM:

Papyrus : Dépôt institutionnel numérique

FRANÇAIS ENGLISH

Chercher dans le dépôt:

Recherche avancée

Accueil
Nouvel utilisateur?

Parcourir le dépôt par:
Communautés et collections
Date
Auteur
Directeur de recherche
Titre
Sujet
Programme (thèses et mémoires)

Services personnalisés:
Espace personnel
utilisateurs autorisés
Déposer un document
Alertes courriel
Modifier mon profil
Statistiques

Aide
Contacts et commentaires

Bienvenue dans Papyrus, le dépôt institutionnel de l'Université de Montréal. Vous trouverez ici des travaux de recherche de nos professeurs et chercheurs ainsi que des thèses et mémoires de nos étudiants.

Communiqués
18 au 24 octobre 2010 : Semaine internationale du libre accès
1^{er} octobre 2010 : Thèses et mémoires électroniques dans Papyrus : un an déjà!

Communautés de recherche et d'enseignement représentées dans Papyrus

- Centre de recherche en éthique de l'UdeM [282]
- Département de kinésiologie [13]
- Direction des bibliothèques [78]
- École d'optométrie [4]
- Faculté de droit [741]
- Faculté de l'aménagement [79]
- Faculté de médecine [628]
- Faculté de médecine dentaire [10]
- Faculté de médecine vétérinaire [37]
- Faculté de musique [18]

À propos de Papyrus...
Questions fréquentes
Droit d'auteur et politiques d'éditeurs [avril 2009]
Comment interroger Papyrus à distance?

Introduction à la notion de libre accès dans les universités.

Author's Rights

SPARC

Image4 : Accueil du site d'archives ouvertes Papyrus

Papyrus²⁴, archive ouverte de l'Université de Montréal (CANADA), offre de nombreux services à ses utilisateurs :

- une page d'accueil donnant la possibilité à ses utilisateurs de choisir la langue du site (anglais, français)
- des vidéos en accès libre sur le site visant à expliquer la notion de « libre accès » ou de « droit d'auteur »
- un espace de consultation permettant de « parcourir le dépôt » par Communautés et Collections ; Date ; Auteur ; Directeur de recherche ; Titre ; Sujet ; Programme (thèses et mémoires)
- un outil de recherche simple et de recherche avancée
- la recherche possible sur d'autres sites de dépôts (OAIster, OpenDOAR, BASE, CARLABRC, GOOGLE SCHOLAR)
- un programme d'aide comprenant un « tutoriel utilisateur »
- une page « à propos » expliquant ce qu'est Papyrus, donnant la possibilité d'accéder à une brochure simplifiée du site, à une présentation détaillée du site, à des démonstrations animés (sous formes de vidéos Flash ou papier PDF), et une FAQ (Foire aux Questions)
- un espace personnel pour les utilisateurs autorisés et inscrits (« ma zone personnelle ») leur permettant d'accéder à leur liste de documents soumis et acceptés, leur liste de soumissions non-achevées, leur liste de documents qui attendent leur intervention
- la possibilité de mettre à jour son profil
- recevoir les avis de nouveautés, recevoir par fil RSS les nouveaux dépôts
- l'accès pour tous les utilisateurs à des statistiques d'utilisation, de visite, de téléchargement, de dépôt...
- l'accès à un annuaire téléphonique (personnel, étudiant) via l'onglet « Répertoires »
- l'accès aux coordonnées des facultés et des départements de l'Université via l'onglet « Facultés »

²⁴ Voir le site suivant pour plus d'informations : <http://papyrus.bib.umontreal.ca/jspui/>

- un lien direct vers le site des bibliothèques de l'Université de Montréal via l'onglet « Bibliothèques »
- un lien direct vers le site d'actualités (nouvelles) de l'Université de Montréal via l'onglet « Actualités »
- une liste des sites « utiles » pour les étudiants et le personnel (avec leurs liens) comme par exemple des sites d'une Faculté, d'un bureau d'aide... via l'onglet « site A-Z »

4) Les besoins demandés par les chercheurs

Il semble évident que les chercheurs ne disent pas d'eux-mêmes qu'ils ont des besoins en terme de service. Il est donc nécessaire de leur poser directement la question et de leur proposer différents types d'outils afin de voir si ceux-ci pourraient les intéresser. C'est donc de cette manière que j'ai pu recueillir les éléments qu'ils aimeraient retrouver sur un site d'archive ouverte. Bien-sûr, il ne s'agit que des besoins cités par quatre interviewés, donc je n'ai pas la prétention de représenter la totalité du monde des chercheurs.

A.Wojciechowska [13] a choisi de poser la question des besoins aux interviewés sous cette forme : « Est-ce que vous voudriez proposer des modifications de fonctionnement de HAL ». Elle a ensuite montré que seulement 5,5 % des personnes interrogées répondent « oui ». Pour une meilleure compréhension, elle a cherché savoir « quelles sont les modifications qu'aimeraient obtenir les interviewés ». Elle n'a pas obtenu de statistiques mais seulement des idées de changement. Parmi ces idées, on peut retenir le besoin de « retrouver les statistiques des articles d'autres utilisateurs », ou la volonté de « simplifier le système ».

Lors des entretiens, j'ai repéré différents types de besoins.

En effet, le besoin le plus important est que le site soit bien fait et bien structuré mais surtout qu'il soit simple d'utilisation et que l'interface soit en Français.

Ensuite, un deuxième besoin « utile » ressort de ces quatre entretiens : la mise en place d'un « générateur » qui recherche directement dans le fichier déposé toutes les métadonnées associés, ce qui permettrait de gagner du temps et de ne pas avoir à retaper toutes les métadonnées déjà renseignées sur le fichier.

Troisièmement, les quatre enquêtés m'ont affirmé avoir besoin de liens sous une forme d'onglet par exemple. Ces liens doivent être dans tous les cas clairement identifiables comme sur le site AUSpace « Others Resources » où on délimite clairement que l'on aura des liens vers d'autres ressources ou comme sur le site Papyrus où l'on dispose de liens utiles à la vie étudiante « Sites A-Z ».

Ces liens pour certains des interviewés doivent fournir un accès vers « la page personnelle des auteurs, la liste de tous leurs papiers », vers des sites d'archives ouvertes de disciplines connexes, ou encore « vers des sites comme SlideShare ». Dans tous les cas, il faut obligatoirement que ces sites aident à trouver l'information qu'on recherche.

Quatrièmement, j'ai pu entrevoir qu'une des personnes interviewées aimerait avoir « des suggestions de papier quand elle se connecte ». De même, il serait utile selon ces personnes d'avoir une vision claire et précise de l'identité des personnes citant un article.

Enfin, en termes de préférences, la majorité aimerait avoir une consultation « par auteur », « par titre », « par thème » avec une fiche catalographique contenant le fichier en texte intégral (format PDF).

On peut aussi relever la volonté d'accéder à des multimédias et pas seulement à des documents écrits.

IV) La Perception des serveurs d'archives ouvertes

1) Avantages et Bienfaits à l'utilisation des sites d'archives ouvertes ?

Lors des entretiens que j'ai réalisés, j'ai voulu savoir comment les interviewés, leurs collègues et/ou leur équipe percevaient les archives ouvertes. Dans l'ensemble, la réponse est positive, c'est un bon outil « qu'il faut continuer à développer ». Les archives ouvertes sont perçues comme étant « le futur ».

Je leur ai ensuite posé la question des avantages et bienfaits. Mon but était alors de savoir si les avantages qu'ils citeraient se recouperaient avec les objectifs cités plus haut dans la partie 2.1²⁵.

Ainsi, les principaux avantages qu'ils m'ont donnés étaient des objectifs cités plus haut :

- objectif intellectuel : la diffusion de la connaissance
 - favoriser l'accès libre aux documents
 - Ouverture généralisée : « on touche plus de monde »
 - L'information n'est plus aux mains d'une minorité de personnes ou d'une élite particulière : la publication devient alors une information publique et non plus privée
- « C'est ce que chacun faisait sur sa page web, ça permet de centraliser, justement de classer et de trouver plus facilement. »
- Objectif communicationnel : cela permet d'améliorer la communication scientifique, « cela permet d'échanger et d'expliquer ce que l'on fait »
 - Objectif visuel : être beaucoup plus visible sur la toile qu'est Internet

De même, j'ai pu distinguer deux autres avantages « ressentis » par les interviewés. D'abord, il y a l'idée que « l'effet d'obligation ou d'incitation force les gens à mettre leurs publications sur Internet », et c'est cet effet qui rend possible cette diffusion des travaux et des connaissances. Ensuite, le deuxième propos très clair est qu'il y a un objectif politique consistant à se débarrasser des éditeurs justement parce que ceux-ci ne font plus rien aujourd'hui dans la chaîne de production : c'est le chercheur qui réalise l'article, la mise en page, les corrections...

²⁵ Voir 2.1 : Les Objectifs identifiables des archives ouvertes pour les chercheurs et pour le lecteur

2) Freins et Réticences à l'utilisation des archives ouvertes?

Même si il existe des bienfaits pour les utilisateurs, on peut tout de même constater que des freins émergent. Il faut alors se demander « quels sont ces freins et d'où viennent-ils ?

Le frein le plus « vif » est la crainte du plagiat. L'interviewée 2 parle d'une peur que d'autres « piquent ses idées », d'une « peur d'un pillage d'idées » comme elle le montre ci-dessous :

« Enquêtée 2 : C'est vrai que si on prend quelqu'un sur un site de préprints, il faut quand même citer d'où ça vient, c'est déjà une trace de référencer correctement. Donc si tout le monde jouer le jeu, il n'y aurait pas cette idée de piquer les idées des autres et de dire "ah, je l'ai écrite et publiée avant toi". Mais bon, cela pourrait être une dérive qui pourrait être risquée. »

D. Bourrion [10], dans une étude qu'il a réalisé sur les archives ouvertes, montre que 31% des répondants à son questionnaire craignent que leurs travaux déposés en archives ouvertes soient plagiés et que « les étudiants utilisent leur travaux à mauvais escient sans les citer ».

Ce frein vient principalement du fait que les archives ouvertes par leur démocratisation massive, s'ouvre à un grand nombre d'individus mais aussi que les articles déposés ne sont pas toujours publiés en revue. L'interviewé 2 nous explique que ce frein viendrait du fait qu'il y a un temps d'attente entre le moment où on envoie son texte dans une revue et le moment où l'article est publié dans une revue de prestige. Elle nous affirme que « Le prestige est toujours là et le temps de publication n'est pas toujours corrélé ».

A ce premier frein est lié un autre frein vu par D. Bourrion [10] : « la question des droits d'auteur ». Les chercheurs pensent alors « qu'une publication en Archives ouvertes ne garantit pas que leurs droits (moraux) sur leurs productions soient respectés ».

Le deuxième frein que j'ai pu relever c'est la moindre confiance dans les sites utilisés. Cette perte de confiance provient du fait que les auteurs ne sont pas certains de la pérennité d'un site. Ils n'ont aucune certitude concernant l'avenir des archives ouvertes : est-ce un outil qui perdurera dans le temps, ou est-ce un outil qui s'écroulera. Si un de ces sites ferme, l'interviewé 4 craint de perdre tous ces travaux, et qu'ils ne soient pas sauvegardés. A ce titre, D. Bourrion [10] repère ce même frein et montre que celui-ci est perçu par une minorité de personnes, comme on peut le voir dans la citation ci-dessous : « l'expression d'inquiétudes concernant la moindre pérennité du support électronique par rapport au papier, sur lequel on dispose, en effet, de quelques siècles de recul ».

Conclusion

En conclusion, on peut constater que les chercheurs ont une utilisation assez moindre des serveurs d'archives ouvertes et une connaissance partielle des services proposés par ces serveurs. L'usage des archives ouvertes se résume globalement au dépôt et il apparaît que les chercheurs occultent le service de « consultation directe sur le site ». Cela est lié à l'usage quasi-exclusif de GOOGLE qui monopoliserait la recherche d'informations et la consultation de références effectuées par les chercheurs. Mais cela est aussi lié à la recherche directe sur les sites personnels de chercheurs.

Bibliographie

[1] Joachim Schöpfel & Hélène Prost. *Développement et usage des archives ouvertes en France 1^{ère} partie : Développement*. Juillet 2010. Rapport de recherche DUAO en Sciences de l'Information et de la Communication. 49 pages. Université Charles-de-Gaulle Lille 3.

[2] Marion Sabourdy, « Un learning center sur l'innovation à Lille1 », Knowtex Blog, 29/12/2010, <http://www.knowtex.com/blog/un-learning-center-sur-l%E2%80%99innovation-a-lille-1/>

[3] C. Aubry & J. Janik. *Les archives ouvertes, enjeux et pratiques : guide à l'usage des professionnels de l'information*. ADBS. 2005.

[4] Yvon Lemay, « Archives ouvertes et archivistique ». In *Archivage et stockage pérennes, enjeux et réalisations*, sous la dir. de Corinne Leblond, p 93-117, Paris : Lavoisier. 2009.

[5] « Définition d'Open Access », ASSH, <http://www.sagw.ch/fr/sagw/laufende-projekte/open-access/oa-definitonen.html>

[6] Kumar Guha, « L'Open Access, L'engagement des scientifiques scientifiques », 16/01/2006, http://www.enpc.fr/fr/documentation/doc_electronique/dossier_openaccess.htm

[7] Gabriel Gallezot, « Les Archives Ouvertes en France ». In *archives ouvertes, enjeux et pratiques : guide à l'usage des professionnels de l'information*, sous la dir. de Christinne Aubry & Joanna Janik, p 97-115. ADBS. 2005

[8] Thérèse HAMEAU, <http://openaccess.inist.fr/spip.php?mot35>

[9] Hélène Bosc, « Archives ouvertes : Quinze ans d'histoire ». In *archives ouvertes, enjeux et pratiques : guide à l'usage des professionnels de l'information*, sous la dir. de Christinne Aubry & Joanna Janik, p 31-58. ADBS. 2005

[10] Daniel Bourrion et al, *Les chercheurs en lettres et Sciences Humaines et les archives ouvertes*, Mémoire de Recherche, Enssib, Juin 2006, <http://www.enssib.fr/bibliotheque-numerique/document-649>

[11] Qu'est ce qu'une revue en accès libre ?, <http://faq-formadoct.ueb.eu/a.php?qid=33114>

[12] Anna Wojciechowska, usage des archives ouvertes dans les domaines des mathématiques et de l'informatique, p294-302, Documentaliste, Science de

l'information 2006, vol 43, n°5-6,
www.cairn.info/load_pdf.php?ID_ARTICLE=DOCSI_435_0294

[13] Anna Wojciechowska, usage des archives ouvertes dans les domaines des mathématiques et de l'informatique, Janvier 2006,
http://archivesic.ccsd.cnrs.fr/docs/00/06/27/23/PDF/sic_00001739.pdf

[14] Béatrice Bégault, « Archives ouvertes : Réflexions sur leur implication dans les pratiques des chercheurs en sciences de l'ingénieur », 7^{ème} Colloque du chapitre français de l'ISKO. In Intelligence collective et organisation des connaissances. Université Jean-Moulin, Lyon 3. 2009

[15] Joachim Schöpfel & Hélène Prost. *Développement et usage des archives ouvertes en France 2^{ème} partie : Développement*. Juillet 2010. Rapport de recherche DUAO en Sciences de l'Information et de la Communication. 49 pages. Université Charles-de-Gaulle Lille 3.

Annexe 1 : Grille d'entretien utilisée :

Partie 1 : Identité

Nom, Prénom, âge

Formation (études réalisées)

Profession

Date à laquelle il a commencé à enseigner ? à faire de la recherche ?

Disciplines d'enseignement

Discipline(s) de recherche, thématiques de recherche

Temps de travail pour l'activité de recherche et pour l'activité d'enseignement (par jour/ mois)

Partie 2 : Utilisation des outils fournis par le web

Quels outils utilise-t-il ? (traitement, collaboration (ex : google doc), communication, ...) Pourquoi ?

Pour activité privée, de recherche ou d'enseignement ?

Quels type de documents réalise-t-il avec ? (graphiques, modification de photos, blog, page html, ...)

Au niveau de l'activité de recherche, quel type de documents sont produits ? (articles scientifique, communication, rapport, conférence, ...). Sont-ils mis en ligne ?

Fréquence d'utilisation ?

Utilité de ces outils ?

Utilité des outils de collaboration ?

Ou Préférence pour les outils standards fournis sur l'ordinateur ?

Partie 3 : Pratiques actuelles et antérieures

- Utilisation des sites d'archives ouvertes ? OUI ou NON, Pourquoi ?

- Si **oui**, quels sites utilise-t-il ? (serveur Institutionnel (d'un laboratoire, d'un centre de recherche, d'une université ?), serveur central (HAL, ...), serveur thématique (ArchiveSic, Halshs...).

- Consultation d'une archive en particulier ? (pourquoi ?) ou plusieurs ?
- Dépôt dans une ou plusieurs archives ? (pourquoi ?)
- Justification du choix du site ?
- Depuis combien de temps utilise t-il les archives ouvertes (ou un site en particulier) ?
- Quels outils ou services utilise t-il sur ces sites ? (dépôt, consultation, statistiques, ...)

- Le dépôt

- Y a-t-il une incitation au dépôt, une obligation ?
- Est-ce que c'est lui, ou un laboratoire qui dépose les documents ? Pour quelle raison ?
- Que dépose t-il comme type de document ? (article scientifiques, ...)
- Texte intégral, notice, lien vers un autre site ?
- Contenu du document ? (photos, graphiques, tableaux) + but (prouver, illustrer une idée, ...)
- Dépose-t-il des publications déjà publiées dans des revues ? (post-print) Ou que des préprint ?
- Dans quelle discipline dépose t-il (axe de recherche, discipline)?
- Nombre de documents déposés par type de document, par thématique?
- Fréquence de dépôt ? A quelles occasions ?
- Nombre de pages moyen par type de dépôt ?
- Dépôt simultané dans une revue (si oui, laquelle ?)

- La consultation

- Comment recherche t-il ? (par type de documents, discipline, par date, par auteur, par identifiant de l'article, par laboratoire, par institution) Pourquoi ? (accès plus rapide, recherche d'un document spécifique, ...)

- ➡ Quel type de document consulte t-il (article scientifique, ...)
- ➡ Quel thématique (discipline) consulte t-il ?
- ➡ Consultation de la notice, du résumé
- ➡ Consultation du texte intégral (survol rapide : regarde les grandes idées, les titres, le sommaire ; lecture des premiers paragraphes, lecture complète)
- ➡ Fréquence de consultation ?
- ➡ Temps de consultation ? (temps seconde)
- ➡ Téléchargement du document ? Oui/non ?
De manière systématique ? Occasionnelle ?
Après consultation « rapide » ou avant ?

- la recherche (recherche avancée, recherche simple, quels outils de la recherche utilise-t-il ?)

- Enjeu / intérêt que revête pour lui le dépôt ? La consultation ?

- OUI ET NON, demander les **Pratiques antérieures** =>

Où déposait-t-il ? Pourquoi ? (publication dans une revue ? laquelle ?)

Y a-t-il eu un changement en cours de route (dépôt dans telle archive puis dans telle autre par non-satisfaction, par envie d'essayer une nouvelle...)

(si oui, idée, est-ce qu'il continue à cette ancienne pratique)

Quand a-t-il commencé à publier ?

Fréquence de dépôt à l'époque ? (occasionnelle, fréquente,...)

Nombre de documents publiés / déposés à l'époque

Intérêt / But de la publication scientifique ?

Partie 4 : Les Besoins

- Les points forts et points faibles du site utilisé ? qu'est ce qu'il aime, qu'est ce qu'il aime pas ?

- Que voudrait-il retrouver sur un tel site ?

- ➡ Langue du site

- ➡ Architecture, ergonomie (interface conviviale, choix des couleurs,...)
- ➡ Besoins en termes de services/ outils
- ➡ Éléments déjà mis à disposition (+ choix des titres)
- ➡ Éléments manquants
- ➡ Type de documents archivés
- ➡ Site multidisciplinaire ? Site pour une discipline ou thématique en particulier ?
- ➡ Site contenant un seul type de document ? Plusieurs ?
- ➡ Type de consultation qu'il aimerait retrouver (ex : par auteur ?)
- ➡ Volonté d'avoir accès à des données chiffrées pour connaître l'ampleur des dépôts, à des graphiques...
- ➡ Volonté d'avoir que des documents en texte intégral ?
- ➡ Avoir accès à des liens vers le texte intégral
- ➡ Avoir accès à des liens vers d'autres sites
- ➡ Avoir accès à des liens vers les sites personnels de chercheurs
- ➡ Importance de la communication entre chercheurs
 - ✓ Donner accès à un forum, à un espace de communication entre chercheurs pour faciliter la communication entre chercheurs
 - ✓ Espace de partage entre plusieurs chercheurs
 - ✓ Donner accès sur le site à un moyen immédiat pour commenter des documents, apporter des corrections... (du même style que google doc)

Partie 5 : Pratiques et perceptions des collègues en matières d'archives ouvertes

Qu'est ce que les collègues utilisent ? Pk à votre avis ? Pratiques de publications ?

Utilité des serveurs d'archives ouvertes? Des services proposés ?

(Intérêts/ apports)

Avantages/ Inconvénients par rapport aux modes de publications traditionnels ?

Freins/ Réticences ?

Les Bienfaits ?

Autres remarques ?

Représentation que revête pour eux les archives ouvertes ?

Annexe 2 : Retranscription de l'entretien 1

L'entretien suivant a été réalisé dans le bureau de l'interviewé à Lille1. Avant de commencer l'entretien, je lui ai rappelé le principe d'un entretien semi-directif et je lui ai demandé si je pouvais l'enregistrer. L'entretien a commencé à 17h le 30 mai 2011. Voici donc la retranscription de cet entretien :

Enquêteur : Bonjour, est-ce que vous pouvez me parler de votre formation initiale ?

Enquêté : J'ai fait une thèse et puis HDR et puis voilà.

Enquêteur : Quelle est votre profession ?

Enquêté : Je suis directeur de Recherche.

Enquêteur : Dans quelle discipline ?

Enquêté : En informatique à l'INRIA.

Enquêteur : Enseignez-vous en même temps ou non ?

Enquêté : Un petit peu.

Enquêteur : Dans quelle matière ?

Enquêté : Informatique

Enquêteur : Quel est le temps de travail que vous mettez pour l'activité de recherche et d'enseignement ?

Enquêté : Ben, c'est toute la semaine.

Enquêteur : Quels outils utilisez-vous sur le web principalement ?

Enquêté : Les outils sur le web ? Pour quoi faire ?

Enquêteur : Outil de collaboration, de traitement de texte par exemple...

Enquêté : Le minimum. Donc un peu Google Docs de temps en temps mais c'est tout. Mais sinon, si c'est de la collaboration, nous on a nos propres outils de synchronisation de documents. Donc avec des systèmes de versions comme SVN... C'est des systèmes de versions de contrôle. Mais c'est fait pour gérer du code, mais ça marche très bien pour les articles.

Enquêteur : Vous réalisez quels types de documents avec principalement ?

Enquêté : Des articles de recherche.

Enquêteur : Quel type de contenu y mettez-vous ?

Enquêté : Euh, oui, donc je mets des graphiques en plus du texte.

Enquêteur : Quelle est la fréquence d'utilisation de ces outils ?

Enquêté : Tous les jours.

Enquêteur : Quelle utilité ont pour vous ces outils ?

Enquêté : Euh, Google Docs, là, c'est juste un truc pour me synchroniser avec les gens qui n'ont pas accès à ce que nous on utilise. Mais sinon, je ne l'utiliserai pas.

Enquêteur : Utilisez-vous un outil d'archives ouvertes ?

Enquêté : Oui, nous on a une archive ouverte qui s'appelle HAL, donc voilà, et tous nos documents sont dedans.

Enquêteur : Quels services utilisez-vous sur HAL ?

Enquêté : Juste ce qu'on doit mettre dedans.

Enquêteur : Juste le dépôt ?

Enquêté : Ouais... Et c'est très mal foutu donc de toute façon...

Enquêteur : Pourquoi utilisez-vous ce site là ?

Enquêté : Parce que là, l'INRIA nous oblige à le faire !

Enquêteur : Depuis combien de temps l'utilisez-vous ?

Enquêté : 2007

Enquêteur : Et avant, vous publiez en revue ?

Enquêté : Non, pas forcément. Avant, on n'avait pas d'archives ouvertes, moi je n'utilisai pas d'archives ouvertes mais par contre mes papiers je les mettais sur le web. Pour moi, les archives ouvertes c'est juste que ça institutionnalise un peu le processus, c'est tout.

Enquêteur : Y a-t-il une incitation au dépôt ou une obligation ?

Enquêté : Oui, il y a une obligation : nous, tous nos articles doivent être dans HAL. Donc c'est ça.

Enquêteur : Qui dépose ? Vous ou le laboratoire ?

Enquêté : C'est nous, l'INRIA a son propre HAL. Donc c'est, on sait que c'est fait par l'INRIA de toute façon. Tous les articles sont(...), il y a le laboratoire du(...), des auteurs, donc de toute façon on a tout suivi par rapport à ça quoi.

Enquêteur : C'est quel type de document ?

Enquêté : PDF, plutôt des articles scientifiques.

Enquêteur : C'est du texte intégral ou non ?

Enquêté : Ça dépend, souvent non.

Enquêteur : Quel type vous mettez dans vos documents ?

Enquêté : Ah oui, de temps en temps, on met des captures d'écran, mais nous, on ne met pas de photos. Je sais qu'il y a des gens qui mettent des photos pour leurs expériences. Nous c'est(...), le maximum qu'on peut avoir, c'est des graphiques.

Enquêteur : C'est plutôt des postprints ou des préprints ?

Enquêté : Euh(...), Moi je mettrai tout et après par contre je ne mets pas la version de l'éditeur.

Mais de toute façon je suis contre le modèle éditorial à l'heure actuelle, on fait tout le travail et c'est eux qui vendent les choses. Je pense que le monde va mal, je marche sur la paix.

Enquêteur : Est-ce que vous publiez des articles qui sont déjà publiés en revue ?

Enquêté : Les préprints.

Enquêteur : Dans quelle discipline vous déposez ?

Enquêté : Ben, en informatique.

Enquêteur : Combien de documents sont déposés par an ?

Enquêté : Par an peut-être 10 à 20, quelque chose comme ça, du moins pour l'équipe.

Enquêteur : Et à vous tout seul ça fait combien ?

Enquêté : Je ne sais pas. Ça doit être du même ordre peut-être une dizaine. Si on écrit un article par mois, c'est déjà pas mal.

Enquêteur : Quel est le nombre moyen de pages par dépôt ?

Enquêté : Une vingtaine.

Enquêteur : Par rapport à la consultation, est-ce que vous consultez des documents sur HAL ?

Enquêté : Rarement. Rarement parce que je pense que ça va plus vite, ça dépend si les articles sont indexés ou pas ou s'ils sont disponibles ailleurs.

Enquêteur : S'ils sont disponibles ailleurs, vous les consultez où ?

Enquêté : Ben ça dépend de ce que Google me propose.

Enquêteur : Quand vous allez voir sur HAL, comment consultez-vous ?

Enquêté : Généralement, moi je pense que je recherche par titre ou par auteur tout simplement.

Enquêteur : Quand vous consultez, est-ce que vous consultez la notice, le résumé, ... ?

Enquêté : Euh non, en général c'est plus la totalité du document, donc si c'est possible.

Enquêteur : Quelle est votre fréquence de consultation sur HAL ?

Enquêté : Elle est très faible.

Enquêteur : Combien mettez-vous à consulter un document ?

Enquêté : Ca je pense que ça varie. C'est pas(...), c'est pas ça le facteur.

Enquêteur : Est-ce que vous téléchargez aussi des documents ?

Enquêté : Ben pour le moment, je ne m'en rappelle pas vraiment, je dirais que non. Mais c'est pareil après, ça veut dire qu'à partir du moment où c'est intégré dans un moteur de recherche, ben ça sera fait quoi.

Enquêteur : Est-ce que vous utilisez des outils de recherche avancée, recherche simple sur HAL ?

Enquêté : Non pas vraiment, pour nous ça ne veut pas dire grand-chose. Nous, on arrive à bien identifier les choses, les articles de recherche sont suffisamment assez barbare pour que ça soit un bon identifiant.

Enquêteur : Quel enjeu revête pour vous le dépôt, la consultation ?

Enquêté : C'est un enjeu politique, ce n'est pas un enjeu pratique de mon point de vue. Donc ça veut dire que ce n'est pas normal que l'État nous paye pour produire des informations qui soient privées. Ce n'est pas normal. Donc moi, ben(...), j'ai écrit pas mal de livres, ils sont tous en accès libre. Je crée des logiciels, ils sont en

accès libre aussi. Et ce n'est pas normal que les gens ne puissent pas accéder à ce qu'on a écrit parce qu'on a fait tout le boulot. Parce que les éditeurs ne font rien quand même dans le travail : Zéro ! Moi je pense, à mon avis, ils ne servent à rien. Donc, moi je pense que le futur c'est les archives ouvertes.

Enquêteur : Vous déposiez où avant ?

Enquêté : Sur ma page personnelle.

Enquêteur : Quand avez-vous commencé à publier en ligne ?

Enquêté : 97

Enquêteur : Quel est le nombre de documents que vous mettiez en ligne à l'époque ?

Enquêté : Je ne sais pas, peut-être une trentaine. Maintenant tous mes articles sont en ligne de toute façon.

Enquêteur : Quels sont selon vous les points forts ou les points faibles du site ?

Enquêté : Les points forts ben à la limite, c'est que ça permet d'avoir toute la production de l'INRIA dans un point. Les points faibles c'est que c'est mal fait, mais bon. Après je pense qu'il ne faut pas généraliser, c'est juste une implémentation d'une archive. Moi, ce qui m'intéresse c'est le concept.

Enquêteur : Qu'est-ce que vous aimeriez retrouver sur un site comme celui-ci ?

Enquêté : Euh...

Enquêteur : Du point de vue de l'architecture ou de l'ergonomie par exemple ?

Enquêté : Que l'entrée de documents soit facile !

Enquêteur : Par rapport à la langue du site ?

Enquêté : Il n'y a pas d'importance.

Enquêteur : Et les outils que vous aimeriez retrouver sur ce site ?

Enquêté : Non pas vraiment besoin d'outils en particulier. Potentiellement, si ça génère une entrée Bibtex automatiquement, pourquoi pas. Là par exemple HAL, ça permet aux chercheurs de générer automatiquement leur page web. Mais moi je le fais automatiquement aussi, donc bon. En tant que programmeur, j'ai fait ça il y a longtemps, ça marche donc je n'ai pas besoin de HAL pour faire ça, ni pour mettre à jour ma base. Donc HAL c'est bien pour les gens qui n'avaient aucune infrastructure. Donc après ils avaient aucune infrastructures, ils entrent leurs

données dedans, ah le site génère pleins de trucs pour eux, donc ça c'est super. Moi j'avais déjà tout.

HAL c'est plus "Oh tu dois le faire en plus, donc c'est un doublon", et c'est pas demain la veille que je ne vais pas utiliser mes propres outils pour faire ce que j'ai fait parce que moi mes outils je les contrôle.

Et nous, on est un peu spéciaux parce qu'on est des informaticiens donc après, j'imagine que la personne qui fait de l'histoire ou autre utilisera l'infrastructure de HAL et ce sera très bien, nous on est un peu spéciaux de ce point de vue là !

Enquêteur : Préférez-vous des sites multidisciplinaires ou non ?

Enquêté : Disons qu'à la limite, je ne préfère pas un site qui soit multidisciplinaire, comme ça ça m'introduit pas trop de bruit quand je cherche sur un moteur. Parce qu'un biologiste qui s'appelle pareil que moi, ben ça fait du bruit.

Enquêteur : Quel type de consultation aimeriez-vous retrouver ?

Enquêté : Moi c'est par auteur, par titre en général.

Enquêteur : Des documents, vous préférez en avoir en texte intégral seulement ou non ?

Enquêté : Moi je préfère en PDF que soit lisible. Moi, avoir le TeX et le fichier généré, moi ça m'intéresse pas. Si on ne peut pas lire le document, à mon avis, ça ne sert pas à grand-chose !

Enquêteur : Aimeriez-vous avoir accès à des données chiffrées, à des statistiques d'utilisation du site, de dépôt ?

Enquêté : Non, ou peut-être un truc qui serait rigolo, ça serait de savoir qui est-ce qui fait référence à des articles. Mais c'est tout.

Enquêteur : Est-ce que vous aimeriez avoir accès à des liens vers d'autres sites ?

Enquêté : Oui, pourquoi pas. Moi je pense que si ça aide à trouver des infos, alors oui. Mais sinon, pas spécialement. En fait, Google le fait pour nous, donc voilà.

Enquêteur : Aimeriez-vous avoir accès à des liens vers des sites personnels de chercheurs ?

Enquêté : Euh, c'est la même réponse : ça c'est toute la question de savoir si le site est bien indexé. S'il est bien indexé, on n'a pas besoin de le faire en plus.

Et la question, c'est que c'est important à ne pas forcer les gens à mettre plus d'informations sur le site parce qu'on dit "ah oui, ça serait bien qu'il y ait des pages web, alors tout le monde doit rentrer sa page web, et quand vous changez, vous changez la page web, ok, ouais".

C'est vraiment la partie de rendre l'utilisation la plus simple possible.

Enquêteur : Quelle est l'importance pour vous de la communication entre chercheurs ?

Enquêté : Ben c'est fondamental.

Enquêteur : Comment communiquez-vous entre chercheurs ?

Enquêté : Ben ça dépend, ça dépend si on les connaît ou pas. Si on les connaît par mail ou par discussion. Si on ne les connaît pas, par mail ou par papier.

Enquêteur : Aimerez-vous retrouver des outils tels que des forums, des espaces de communication entre chercheurs?

Enquêté : Non !

Enquêteur : Et des espaces de partage ?

Enquêté : Non plus.

Enquêteur : Des moyens pour commenter des documents, apporter des corrections ?

Enquêté : Non, non chez nous on a tout ça déjà, et ça fonctionne bien. Parce que par exemple, qu'est ce que ça veut dire, est-ce que ça doit être privé, public, est-ce que vous vous pouvez voir ce que j'ai écrit ? Est-ce que si j'écris quelque chose, est-ce que je dois écrire "modifier", est-ce que ça insulte l'autre ?... Non, non, non, ça on le gère. On sait comment gérer ça, on sait qui est-ce qui a lu,... C'est toute une partie humaine de la chose.

Enquêteur : Quelle est la perception de vos collègues des archives ouvertes ?

Enquêté : Nous on est pour. Nous on est pour ce système-là.

Enquêteur : Quelle est pour votre équipe, l'utilité des archives ouvertes ?

Enquêté : Pour le moment, les archives ouvertes sont peu utiles, parce que je pense qu'on ne publie assez dedans, et que HAL est pas trop mal indexé par Google. Mais sinon, ouais, je pense que c'est très important politiquement parlant.

Enquêteur : Y a-t-il des freins à l'utilisation pour vous et vos collègues ?

Enquêté : Non, non, non. Le véritable problème à l'heure actuelle, c'est que dans notre domaine, il y a actupuelly, I3E, ACM. C'est notre entité qui gère toutes les connexions, et donc c'est eux ont le pouvoir quoi. Vu qu'à l'INRIA, il n'y a pas beaucoup de gens qui travaillent sur mon domaine, ben on peut le voir en France, ben j'ai très peu de probabilité de trouver un article qui m'intéresse. Donc maintenant, ça veut dire, s'il y a une plus grosse communauté qui utilise une archive ouverte, ben ils ont une plus grande possibilité de trouver de l'information dessus.

Enquêteur : les bienfaits pour vous de ces outils ?

Enquêté : c'est politique, me débarrasser des éditeurs !

Enquêteur : Et pour vos collègues ?

Enquêté : Ben, c'est la même chose, on est bien synchronisé par rapport à ça, on fait tout le travail : on crée les idées, on les met en page, on fait l'évaluation, on fait pratiquement tout : la copie, la vérification, la mise en forme. Pourquoi on devrait payer, pourquoi l'État français devrait payer Elsevier, ou je ne sais pas quoi pour produire cette information ?

Annexe 3 : Retranscription de l'entretien 2

L'entretien suivant a été réalisé dans le bureau de l'interviewée. Avant de commencer l'entretien, je lui ai rappelé le principe d'un entretien semi-directif et je lui ai demandé si je pouvais l'enregistrer. L'entretien a commencé à 17H30 le 31 mai 2011. Voici donc la retranscription de cet entretien :

Enquêteur : Bonjour, est-ce que vous pouvez me parler de votre formation initiale ?

Enquêtée : Alors, moi, j'ai fait un master en Italie et après une thèse de doctorat en France, euh(...) à Paris 11 et après, j'ai eu un poste ici à Lille1. Donc je suis maître de conférences ici à Lille1 Mathématiques, les applications des Mathématiques.

Enquêteur : Quelle profession faites-vous actuellement ?

Enquêtée : Donc maître de conférences.

Enquêteur : A quelle date avez-vous commencé à faire de la recherche ?

Enquêtée : Ben euh, à peu près au début de ma thèse donc en 1991, 1992.

Enquêteur : Et vous enseignez en même temps ?

Enquêtée : Pas à l'époque, là maintenant j'enseigne. Donc je suis en poste ici à Lille depuis 1998, et j'enseigne, et je fais de la recherche. Les deux.

Enquêteur : Et vous enseignez dans quelle discipline ?

Enquêtée : Donc euh, Mathématiques appliquées plus en particulier, c'est l'analyse numérique et les équations des partiels.

Enquêteur : Combien de temps de travail mettez-vous pour la recherche et l'enseignement ?

Enquêtée : Oh ! En théorie, ça doit être moitié-moitié. Après il y a aussi l'administration. Donc on va dire un tiers, un tiers, un tiers.

Enquêteur : Quelle utilisation faites-vous des outils fournis par le web ?

Enquêtée : Euh, on regarde surtout au niveau des articles de recherche s'ils sont déjà publiés et donc soit(...). Donc voilà au niveau des éditeurs qui permettent de regarder ça. Donc les bibliothèques, on est un réseau de bibliothèques, les bibliothèques de mathématiques. Soit sur les pages personnelles des chercheurs quand on connaît spécifiquement quand il y a quelqu'un qui s'occupe d'un domaine spécifique. Euh, c'est surtout ça. Après, c'est plus(...), après contextualiser le(...), le thème donc plus des articles de vulgarisation.

Enquêteur : Quel type de documents réalisez-vous avec les outils que vous utilisez sur le web ?

Enquêtée : Euh, c'est surtout des(...), soit c'est des articles de recherche pour des publications, soit c'est pour des conférences, quelquefois des extraits pour l'enseignement. Donc côté enseignement, ça peut venir(...), enfin ça vient surtout des livres classiques effectivement, mais quelquefois, il y a des sources aussi sur le web.

Enquêteur : Quelle est votre fréquence d'utilisation de ces outils ?

Enquêtée : Là, c'est ponctuel, ça peut-être assez intense mais sur des périodes ponctuelles quand il y a... Voilà donc on ne passe pas(...), quand on fait une recherche bibliographique, on essaye de passer quelque temps là-dessus, et après, c'est du matériel qu'on exploite par la suite donc on ne va pas tout le temps(...), c'est vraiment des utilisations ponctuelles assez intenses.

Enquêteur : Quelle utilité revêtent pour vous ces outils ?

Enquêtée : C'est beaucoup plus pratique que ceux qu'on avait auparavant qu'il fallait attendre les publications, de les voir se publier, donc ça commençait à être un peu vieux. Parfois, ça prend 2 ou 3 ans avant de sortir dans certaines revues. Donc c'est une utilisation pas de(...), des ressources qui sont beaucoup plus à jour et je dirai que c'est surtout cet aspect-là de l'immédiat que(...), qu'on peut avoir.

Enquêteur : Est-ce que vous utilisez des outils de collaboration ?

Enquêtée : Un peu, pas beaucoup, pour(...) dans le sens qu'on est toujours un nombre assez limité, donc on s'envoie par mail le plus souvent. Quelquefois, on a des outils de partage, oui effectivement si vous voulez garder un historique donc en particulier par exemple, il y a un produit de l'INRIA qui s'appelle GForge qu'on utilise pour avoir les versions de codes, de calculs qu'on fait, ou des versions de papiers. Mais étant un nombre assez limité, parfois, on s'envoie aussi par mail, ça va plus vite.

Enquêteur : Par rapport à l'utilisation des archives ouvertes, en utilisez-vous une ?

Enquêtée : C'est GForge pour ce qui est du dépôt des différentes versions, par exemple un code, euh(...), après je vais plutôt utiliser des archives(...), des archivages de preprints, donc ArXiv, HAL, les moyens de(...), en fait de dépôt de publications de l'INRIA, du CNRS, c'est surtout ça.

Enquêteur : Est-ce que vous déposez dans une ou plusieurs archives ?

Enquêtée : Non, je dirai, voilà il y a un type pour taper des codes de calcul par exemple, donc c'est plutôt GForge que je vais utiliser. Par contre quand c'est

l'autre, c'est plutôt des publications, soit c'est une recherche de publications très récentes qui ne sont pas encore traitées et publiées, soit on dépose nous les tous derniers travaux pour qu'ils soient connus assez vite des gens. Donc c'est plutôt côté publications : HAL. Côté développement de codes de calcul, c'est GForge. C'est les deux principaux.

Enquêteur : Depuis combien de temps utilisez-vous les archives ouvertes ?

Enquêtée : Quelques années, déjà. Pour les archives de publications ça existe depuis une dizaine d'années avec ArXiv, il y avait d'autres choses avant, voilà. Pour les codes de calcul, c'est relativement récent, c'est trois ou quatre ans. J'avais entendu parler un peu avant, mais c'est depuis 3 ans que j'utilise ça.

Enquêteur : Quels outils ou services utilisez-vous sur ces sites ?

Enquêtée : Dépôt, Consultation. Parce que toujours, on est un petit nombre, donc il n'y a pas des statistiques d'utilisation, des choses comme ça.

Enquêteur : Est-ce qu'il y a une incitation au dépôt ou une obligation sur ces sites-là ?

Enquêtée : Une incitation je dirais, c'est conseillé.

Enquêteur : Est-ce que c'est vous qui déposez ou un laboratoire ?

Enquêtée : Non, non, c'est le chercheur qui dépose. Mais parfois, il faut avoir des connaissances au niveau du laboratoire, appartenir à un laboratoire, parce que ce n'est pas ouvert à tout le monde. Par exemple pour les produits CNRS, il faut être dans un laboratoire CNRS, pour les produits de l'INRIA, il faut être dans un laboratoire de l'INRIA.

Enquêteur : Quel type de documents déposez-vous ?

Enquêtée : Des articles scientifiques principalement, pas de communications.

Enquêteur : Est-ce que vous déposez du texte intégral, des notices ?

Enquêtée : Euh, pas pour l'instant mais c'est prévue. Ça demande du développement et tout ça, donc euh. Après, oui, j'avais oublié : il y a aussi un produit au niveau de l'enseignement en fait plus local, c'est(...), c'est la plateforme Moodle en fait qui est développé sur l'Université ici sur Lille1. Donc euh(...), quelquefois je l'utilise, mais pas souvent car j'ai des petits groupes donc je donne directement les documents aux étudiants mais quelquefois, c'est vrai que des feuilles d'exercice, des corrections, je peux aussi utiliser ça. Mais ça c'est quelque chose d'un peu plus local, c'est pour l'enseignement.

Enquêteur : Quel contenu mettez-vous dans vos documents ?

Enquêtée : C'est(...), c'est du texte, des articles, des photos, c'est plutôt des images, des graphiques et quelquefois des petits films mais les films c'est toujours des simulations numériques pas des films...

Enquêteur : Quel est le but pour vous de ces outils ?

Enquêtée : Ben, c'est la diffusion de la connaissance le plus rapidement possible.

Enquêteur : Est-ce que vous déposez des postprints ou des préprints ?

Enquêtée : Non, c'est la(...), le(...), l'archivage HAL qui permet de faire ça, les préprints, après quand les articles sont publiés, ils sont publiés, on les trouve dans les journaux.

Enquêteur : Et, dans quelle discipline déposez-vous ?

Enquêtée : Mathématiques, mathématiques appliquées.

Enquêteur : Combien de documents déposez-vous par an ?

Enquêtée : Alors là, ça dépend de la production, pas beaucoup, un ou deux par an, c'est pas...

Enquêteur : Quel est le nombre de pages moyen que vous réalisez par dépôt ?

Enquêtée : Un article, c'est entre 20 et 30 pages.

Enquêteur : Donc au niveau de la consultation, comment recherchez-vous ? Par type de document, par discipline... ?

Enquêtée : Par discipline, je dirai même par sous-discipline, par thématique, euh(...) par auteur certainement, par mots-clés, au niveau du titre et du mot-clé. Quelquefois, on a entendu parler de certaines références, donc la date, mais c'est moins souvent.

Enquêteur : Quel type de document vous consultez principalement ?

Enquêtée : Des articles scientifiques, quelquefois des exposés, mais ça on les trouve plus sur les pages personnelles des gens. Là les archivages, ils ne les déposent pas là-dedans.

Enquêteur : Quelle discipline consultez-vous ?

Enquêtée : La Physique, la mécanique, là, c'est parce que ça m'intéresse, je m'intéresse à la mécanique des fluides, donc l'interaction de la physique avec la mécanique, c'est surtout ça. Donc ça dépend de la thématique.

Enquêteur : Quand vous consultez, vous consultez la notice, le résumé... ?

Enquêtée : Ben, quand il y a(...), ben ça dépend un peu de la bibliographie, donc certainement le résumé, et après si c'est intéressant, on va plus loin. On démarre par le résumé.

Enquêteur : Quelle est votre fréquence de consultation ?

Enquêtée : Ben là encore, c'est ponctuel, c'est des pics, c'est quand on regarde quelque chose, qu'on a besoin, qu'on n'a pas assez de matériel donc c'est difficile à évaluer. Je dirais plusieurs fois par an, plusieurs fois par mois peut-être aussi, mais il peut aussi passer des mois où il y a le travail spécifique à faire, et on a déjà le matériel, c'est plutôt du développement de codes. Par contre il y a des mois où on sèche, on ne sait pas comment faire et donc on fait beaucoup de recherche.

Enquêteur : Quel est le temps de consultation ?

Enquêtée : Euh...(Rires) Entre quelques minutes et quelques heures, quelques jours, ça dépend. Si c'est en pic, c'est quelques jours, plusieurs fois par jour, on revient souvent à chercher des choses. C'est en fonction après de la cible, si on a trouvé des choses.

Enquêteur : Est-ce que vous téléchargez le document ?

Enquêtée : Quand c'est intéressant oui.

Enquêteur : Du point de vue de la recherche, est-ce que vous utilisez les outils de recherche avancée et de recherche simple ?

Enquêtée : Non, c'est plutôt simple.

Enquêteur : Quel est l'enjeu, l'intérêt que revête pour vous le dépôt, la consultation ?

Enquêtée : Je pense que c'est assez fondamental, c'est vraiment un changement dans la façon de chercher des informations par rapport à il y a 10 ans ou 5 ans même.

Je vais très peu à la bibliothèque pour chercher physiquement un article, c'est-à-dire que la plupart des fois, c'est sur le web. Il est stocké sur l'ordi, voir il est imprimé.

Enquêteur : Avant les archives ouvertes, qu'utilisiez-vous comme moyen de communication des résultats ?

Enquêtée : C'était plutôt des publications dans des revues, donc fallait attendre qu'elles soient sorties, aller dans une bibliothèque qui ait cette revue, aller chercher...

Enquêteur : Quand avez-vous commencer à publier en ligne sur HAL ?

Enquêtée : Hum, il y a quelques années, je dirai 3 ou 4 ans, ça existait même avant, mais je l'utilisais moins. Oui, donc depuis 4 ou 5 ans.

Enquêteur : Quel est l'intérêt, le but de la publication scientifique pour vous ?

Enquêtée : Ben, c'est la diffusion de la connaissance dans un sens très général, entre les résultats que nous on trouve et qu'on met à disposition des autres, et donc quand on cherche niveau bibliographique, on regarde toujours les derniers, c'est quand même des thématiques où les évolutions sont assez importantes et donc du coup il faut toujours être un peu à l'affût de ce qui se passe ailleurs pour ne pas faire des choses qui ont déjà été faites avant. C'est important de voir où on est nous par rapport aux autres.

Enquêteur : Quels sont les points forts et points faibles du site que vous utilisez ?

Enquêtée : Euh, je ne vois pas trop quoi répondre. Disons qu'il faut déjà avoir une idée de ce que l'on cherche. Par mot-clé, c'est toujours trop, il y a trop de choses qui sortent. Donc après, on risque vite un peu de se disperser. Il faut déjà avoir un peu une idée précise, c'est très bien quand on a un auteur, quand on a(...) ou alors un groupe de recherche, ou des citations des uns des autres. Quand on met simplement des mots-clés, parfois, c'est trop vague. Parfois, on a énormément de réponses possibles, et après il faut affiner donc du coup, il faut consulter un par un.

Enquêteur : Quels éléments aimeriez retrouver sur ces sites-là ?

Enquêtée : Les sites me conviennent assez, c'est pas important.

Enquêteur : Y aurait-il des types de documents que vous aimeriez retrouver en plus ?

Enquêtée : Pas actuellement, pas pour le besoin actuel.

Alors, il y a aussi, ils essayent de développer des archivages de codes, donc comment ça s'appelle "Ciel", il y a quelque-chose du CNRS pareil, donc on a du dépôt de codes, de calcul, mais j'avais réfléchi un moment, mais c'est trop de travail ce genre de développement là. Pour moi, je pense que quand c'est un groupe, effectivement de plusieurs personnes, et qui veulent faire un produit pour après le mettre à disposition d'autres, c'est vrai. Pour la recherche que je fais moi, non.

Donc là c'est vraiment des sites où la diffusion est massive. Donc, je ne suis pas encore à ce niveau là, du moins pas pour mes besoins personnels. Mais c'est vrai que ça peut devenir intéressant et certainement, il y a des groupes qui sont(...), qui crée un produit et le but c'est après pas simplement étudier un problème pour faire de la publication, c'est vraiment le produit en lui-même qui est quelque-chose de valorisant.

Enquêteur : Est-ce que vous aimeriez retrouver un site multidisciplinaire ou non ?

Enquêtée : Peut-être de disciplines connexes, parce que parfois on ne connaît pas trop, et effectivement, on n'a pas beaucoup l'accès... Donc on a par exemple énormément de revues de Mathématiques même sur les sites communs aux différentes bibliothèques, mais pas beaucoup au niveau de la physique, au niveau des disciplines qui sont un peu connexes.

Quand on fait un travail interdisciplinaire, parfois, ça peut être difficile de trouver des choses qui sont dans d'autres disciplines. Donc faut vraiment connaître, mais ça c'est une question assez compliquée quand même... C'est vrai que des gros sites multidisciplinaires, ça pourrait être intéressant.

Enquêteur : Quelle type de consultation aimeriez-vous retrouver ?

Enquêtée : Je pense qu'il faudrait une consultation par auteur. Par titre, ça serait vraiment trop vague. Ou alors par mot-clé.

Enquêteur : Est-ce que vous aimeriez avoir accès à des données chiffrées par rapport au nombre de dépôts..

Enquêtée : Ça à ce moment là, en Math, parce qu'il y a un produit qui s'appelle MathSyneth (outil de recherche) qui permet de consulter et finalement de trouver les différents articles, et là ça permet de mettre aussi les auteurs, les citations, et tout ça. Ça ça existe déjà en Mathématiques, « MathSyneth » ça s'appelle. Mais après je ne sais pas si ça existe déjà sur HAL... Mais c'est vrai que savoir si quelqu'un d'autre consulte un article, disons simplement le consulter, le regarder, n'est pas un signal qu'après il est utilisé. Donc MathSyneth, c'est vraiment quand un article est cité

Enquêteur : Aimeriez-vous avoir accès à des docs en texte intégral, ou plutôt à des fiches détaillées, notices ?

Enquêtée : Moi, quand c'est un article, je préfère avoir le texte intégral.

Enquêteur : En quel format ?

Enquêtée : PDF, c'est plus facile à utiliser. PDF plus que des formats qu'après il faut être connecté à Internet. Ça permet de l'archivage sur ordinateur, et après, on peut continuer à consulter même en ayant pas internet.

Enquêteur : Aimerez-vous avoir accès à des liens vers d'autres sites ?

Enquêtée : Oui, peut-être des sites similaires dans d'autres disciplines. C'est l'histoire de l'interdisciplinarité.

Enquêteur : Aimerez-vous avoir accès à des sites personnels de chercheurs ?

Enquêtée : Ça, en général c'est les gens qui... Oui, mais ça, on fait une recherche sur Internet, sur Google, on met le nom de la personne Homepage, et en général on tombe dessus si ça existe. A partir(...), oui éventuellement un lien à partir des auteurs de l'article.

Enquêteur : Quelle est l'importance pour vous de la communication entre chercheurs ?

Enquêtée : Là, je pense qu'elle est primordiale dans notre discipline.

Enquêteur : Comment communiquez-vous entre vous ?

Enquêtée : Ben quand on connaît les chercheurs, c'est plutôt par mail ou on peut demander des renseignements à quelqu'un d'autre, donc c'est plutôt par mail.

Enquêteur : Est-ce que vous utilisez des outils de collaboration entre vous pour corriger un document entre vous, vous évaluer ?

Enquêtée : Non.

Enquêteur : Est-ce que vous aimeriez que ces sites donnent accès à des forums, à des espaces de communication entre chercheurs,... ?

Enquêtée : C'est trop trop vaste tout ça. La communication est toujours entre un groupe relativement restreint, donc euh, faire des forums, je vois déjà une utilisation massive.

Enquêteur : Et, à un espace de partage entre chercheurs ?

Enquêtée : Là, ça dépend toujours de la taille, oui, éventuellement.

Enquêteur : Est-ce que vous aimeriez avoir accès à un moyen immédiat pour commenter des documents, pour corriger des documents,... ?

Enquêtée : Non, pas dans l'utilisation que j'en fait. Ça serait très ponctuel, sur des choses... Non, parce que là encore, c'est une question de version différente, c'est

comme ça qu'on se partage, sur le document lui-même, c'est un archivage plutôt des différents documents parce qu'après on se perd sinon.

Enquêteur : Quelle perception que vous avez des archives ouvertes ?

Enquêtée : Du positif, je pense que c'est quelque chose qu'il faut développer.

Enquêteur : Quelle est la perception de vos collègues ?

Enquêtée : C'est pareil, c'est partagé comme un moyen de diffusion.

Enquêteur : Quels sont les avantages et inconvénients par rapport aux modes de publication traditionnelle ?

Enquêtée : Ben la diffusion plus rapide, mais de l'autre côté que ce n'est pas officialiser avec un article publié et tout ça. Si ça traîne, ça risque de laisser à d'autres des idées qui ne sont pas encore publiées. C'est vrai que si on prend quelqu'un sur un site de préprints, il faut quand même citer d'où ça vient, c'est déjà une trace de référencer correctement. Donc si tout le monde jouer le jeu, il n'y aura pas cette idée de piquer les idées des autres et de dire "ah, je l'ai écrite et publiée avant toi". Mais bon, ça sa pourrait être une dérive qui pourrait être risquée, et surtout dans les journaux d'un certain prestige mais qui demande beaucoup de temps pour être publiés. Le prestige est toujours là et le temps de publication n'est pas toujours corrélé.

Enquêteur : Quels sont les freins selon vous ?

Enquêtée : Ben, c'est les inconvénients de tout à l'heure.

Enquêteur : Quels sont les bienfaits selon vous ?

Enquêtée : Je pense que tout ce qui va vers une amélioration de la communication, de la diffusion de la connaissance, c'est positif.

Annexe 4 : Retranscription de l'entretien 3

L'entretien suivant a été réalisé au sein de l'IRCICA, dans le bureau de l'interviewée. Avant de commencer l'entretien, je lui ai rappelé le principe d'un entretien semi-directif et je lui ai demandé si je pouvais l'enregistrer. L'entretien a commencé à 9H30 le 31 mai 2011. Voici donc la retranscription de cet entretien :

Enquêteur : Pouvez-vous me parler des que vous avez réalisé ?

Enquêtée : Les études, alors euh moi j'ai fait un BAC S, spécialité Math et derrière j'ai intégré une école d'ingénieurs qui recrutait juste après le BAC. Donc j'ai fait 5 ans en école d'ingénieur donc j'ai fait l'INSA de Lyon et en spécialité j'ai fait "Télécommunications" donc il y avait un département à Lyon qui s'appelait la Télécommunications. Et en dernière année, j'ai fait une double formation en école d'ingénieurs et en parallèle un DEA donc le master recherche maintenant, à l'époque ça s'appelait DEA. Et à la suite de mon DEA j'ai commencé une thèse, thèse en informatique dans les réseaux de capteurs donc ça a duré trois ans. A l'issue de ça j'ai passé les concours d'État pour obtenir un poste de chercheur que j'occupe actuellement.

Enquêteur : Et vous avez commencé à enseigner et à faire de la recherches a quelle année a peu près ?

Enquêtée : J'ai commencé la recherche en stage de DEA, c'était fin, dernier(...), dernier semestre de 2003. Pendant ma thèse j'ai fait 3 ans de recherche. Et puis, depuis je suis chercheur à temps plein.

Enquêteur : Quelles sont vos disciplines de recherche ?

Enquêtée : Donc moi je suis dans les réseaux sans fils. Donc je travaille sur des protocoles de communications pour des objets contraints comme les capteurs ou comme les RFID.

Enquêteur : Le temps de travail pour vos activités de recherche ?

Enquêtée : Oh beaucoup (rire). En fait, comme je suis chercheur à temps plein, tout ce qui n'est pas recherche c'est juste gestion administration des projets donc je dois faire bien 90% de recherche.

Enquêteur : Ensuite pouvez-vous me parler des outils que vous utilisez sur le web ?

Enquêtée : Les moteurs de recherches en priorité. Euh(...), après des outils qui sont uniquement sur le web pas tellement. Après on va aller télécharger des logiciels purement des outils de recherche ou alors les sites qui contiennent des articles(...), des articles de recherche et des bibliographies.

Enquêteur : Pouvez-vous me parler des pratiques par rapport aux types d'archives ouvertes ?

Enquêtée : HAL

Enquêteur : Oui.

Enquêtée : Quel genre de pratiques vous voulez dire, comment nous on travaille avec HAL ?

Enquêteur : Oui

Enquêtée : Ben, en fait, nous toutes nos publications on les enregistre sur HAL, du moins on essaye. Et euh(...), ça c'est pour la partie qu'on fournit dans HAL donc les publications c'est en terme d'articles dans les conférences, dans des revues, dans des journaux même les démonstrations, les thèses on les renseigne dans HAL. On essaye le plus souvent de mettre le document en soit, pas seulement la notice mais également le document. Et ensuite on fait nos recherches également sous HAL soit directement en allant sur HAL mais plus souvent on va plus dans Google ou dans d'autres. Et puis comme HAL est bien répertorié on(...), Google nous renvoie sur HAL.

Enquêteur : Vous utilisez que HAL ou d'autres archives ?

Enquêtée : Principalement HAL, parce que il y a tout un pan qui est INRIA/CNRS. Donc on a des outils supplémentaires dessus. Euh(...), par exemple on a des outils de statistiques pour voir à quelles fréquences un document est téléchargé, s'il est lu, s'il est juste consulté. Je veux dire, soit on regarde juste le résumé soit on consulte l'article en soit. Tout ça(...), on a des outils qui permettent de le mesurer. Donc on fait principalement HAL parce qu'ensuite l'INRIA derrière fait des connexions entre HAL et d'autre moteurs à archiver.

Enquêteur : Depuis combien de temps utilisez-vous HAL ?

Enquêtée : Pas depuis très très longtemps en fait. Avant on l'utilisait mais de façon sporadique. Mais de l'utiliser de telle façon à avoir l'automatisme de tout renseigner dedans ça doit faire un an peut être.

Enquêteur : Par rapport à ces sites-là vous utilisez quelles services ?

Enquêtée : Ben le dépôt justement parce qu'on renseigne tout ça, on consulte, c'est à peu près(...) à peu près tout.

Enquêteur : Par rapport au dépôt, y a t-il une incitation au dépôt ou une obligation ?

Enquêtée : On n'a pas d'obligation. On nous recommande fortement de le faire. Mais c'est vrai que derrière, on est gagnant de renseigner HAL parce que justement comme je vous le disais HAL est bien répertorié par Google. Nos articles vont être plus facilement trouvés. Et puis, déjà directement et en plus parce que l'INRIA derrière a connecté son site web à HAL. Ce qui fait que quand on veut juste avoir les publications d'une équipe par exemple, on va chercher dans HAL. Donc si on n'a pas renseigné on n'est pas visible donc ... disons on n'est pas obligé : on nous dit juste ça serait bien que vous le fassiez et puis derrière on est gagnant si on le fait.

Enquêteur : C'est vous ou le laboratoire qui met en ligne ?

Enquêtée : Oui, c'est nous, enfin les articles c'est nous, les rapports de recherche c'est l'INRIA.

Enquêteur : Quel type de document déposez-vous ?

Enquêtée : On dépose les articles dans les conférences, les articles dans les revues, les démonstrations, les rapports de recherche, les thèses... euh(...) je pense que c'est à peu près tout.

Enquêteur : Les documents que vous déposez sont du texte intégral ou juste la notice?

Enquêtée : On essaye le plus souvent de mettre le texte intégral.

Enquêteur : Sous quel format ?

Enquêtée : PDF

Enquêteur : Quel contenu mettez-vous dans vos documents ? (graphiques, photos...) ?

Enquêtée : Dans le PDF ?

Enquêteur : Oui

Enquêtée : Oui en général, il y a nos résultats, donc c'est souvent sous forme de courbes, beaucoup de schémas.

Enquêteur : Le but que revêt pour vous le dépôt ?

Enquêtée : Le but de déposer ?

Enquêteur : Oui

Enquêtée : Comme je disais c'est plus pour être visible derrière, pour que nos articles soient disponibles facilement, pour que les gens les trouvent facilement parce que plus ils les trouvent facilement plus on a des chances d'être connu, enfin que nos travaux soient connus. Et puis également ça permet de faire une bibliothèque de tout ce qu'on peut faire, pas seulement individuellement mais également au niveau de l'équipe, c'est pas mal.

Enquêteur : Dans quelle discipline déposez-vous ?

Enquêtée : Ça doit être marqué informatique.

Enquêteur : Quel est le nombre de documents que vous déposez ?

Enquêtée : Ben, ça va beaucoup dépendre du taux d'acceptation qu'on a, donc euh(...), comme on dépose tout ce qu'on a d'accepté. On a des années où nos articles sont plus acceptés que d'autres. Je ne pourrais pas vous dire un chiffre précis.

Enquêteur : Le nombre de page par dépôt ?

Enquêtée : Là aussi ça dépend de l'article vu que suivant les conférences on a des nombres limités de pages, c'est en général entre 8 et 12.

Enquêteur : Est ce qu'il vous arrive de déposer dans une revue en même temps.

Enquêtée : Euh non en général.

Enquêteur : Par rapport à la consultation, comment consultez-vous ? Par type de document, par discipline ?

Enquêtée : Souvent par le titre soit par les auteurs.

Enquêteur : Quel est le type de documents que vous consultez ?

Enquêtée : Des articles de conférence principalement.

Enquêteur : Quelle discipline consultez-vous ?

Enquêtée : Ben la mienne.

Enquêteur : Quand vous consultez, est-ce que vous consultez la notice ou bien le résumé... ?

Enquêtée : Je commence par le résumé avec la notice pour voir si ça correspond à ce que je recherche et si ça correspond ben je télécharge le papier s'il est disponible bien-sûr.

Enquêteur : Quelle est votre fréquence de consultation ?

Enquêtée : Ça c'est variable, quand je fais des états de l'art en fait je ne saurais pas vous dire.

Enquêteur : Le temps de consultation environ ?

Enquêtée : Ça dépend si le papier correspond à ce que je veux ou pas (Rires). Donc, euh(...), si ça correspond pas, euh(...), c'est juste quelques secondes alors que si ça correspond, ben je télécharge le papier, et celui là je l'imprime généralement.

Enquêteur : Par rapport à la recherche, est ce que vous utilisez les outils de recherche avancé, et/ou de recherche simple ?

Enquêtée : Euh, d'abord simple, et si j'ai vraiment plusieurs critères pour faire des filtres : avancées.

Enquêteur : L'enjeu que sa revête pour vous la consultation ?

Enquêtée : Au niveau de la consultation ?

Enquêteur : Oui

Enquêtée : C'est trouver des travaux similaires aux miens qui peuvent m'aider soit pour m'inspirer, soit pour que je me compare à quelque chose pour évaluer mon travail.

Enquêteur : Et pour le dépôt ?

Enquêtée : Pour le dépôt c'est plutôt la visibilité.

Enquêteur : Donc avant de déposer dans HAL vous déposiez où ?

Enquêtée : On ne déposait pas.

Enquêteur : Vous ne déposiez pas ?

Enquêtée : On les mettait sur nos sites personnels.

Enquêteur : Donc vous avez commencé à publier sur votre site personnel à partir de quel moment.

Enquêtée : Ben, je pense quasiment dès le début quand j'ai eu des papiers. Peut être pas tout le début, sur la fin de ma thèse peut-être.

Enquêteur : Quelle était votre fréquence de dépôt à l'époque ?

Enquêtée : Ben ça dépendait de mes papiers acceptés donc là encore, elle était peut être un peu plus faible que maintenant, pareil c'était variable.

Enquêteur : Quel est l'intérêt, le but de la publication scientifique pour vous ?

Enquêtée : Ben, c'est de pouvoir partager ses résultats. Donc pour moi, la recherche c'est avant tout un échange, parce qu'on ne peut pas tous avoir toutes les idées, il faut qu'on s'inspire ce qui a été fait ailleurs pour essayer, ben, soit de trouver ce qui ne va pas pour améliorer soit pour se comparer pour être sûr que ce qu'on fait ça a du sens et pour avancer tout simplement.

Enquêteur : Quels sont les points forts et les points faibles du site utilisé donc de HAL ?

Enquêtée : Alors, point forts, ben c'est assez intuitif pour déposer, ce n'est pas trop long, enfin, on a une réactivité plutôt bonne quand on dépose quelque chose puisqu'il y a toujours une vérification avant que ce soit vraiment(...), vraiment accepté donc là c'est assez rapide. C'est donc bien traquer c'est visible. Et sinon en défaut, je n'ai pas trouvé comment supprimer quelque chose qu'on enregistre, sinon à part sa...

Enquêteur : Est-ce que vous savez qui vérifie derrière ?

Enquêtée : Je ne sais pas.

Enquêteur : Qu'est ce que vous aimeriez retrouver sur un site comme HAL ?

Enquêtée : Peut-être quand on se connecte des suggestions de papiers qui viennent d'être déposés.

Enquêteur : Et par rapport à la langue du site est-ce que vous avez une langue en particulier que vous aimeriez retrouver sur le site ?

Enquêtée : Il y a le français, donc ça me va. (Rires)

Enquêteur : Les éléments, le type de documents archivés, est ce qu'il y a des documents qui manqueraient ou pas ?

Enquêtée : Pas que ... de problème que j'ai rencontré, j'ai toujours trouvé ce que j'avais besoin donc ...

Enquêteur : Est ce que vous préféreriez un site multidisciplinaire comme HAL ou un site avec une seule discipline ?

Enquêtée : Euh(...), moi, je me suis jamais posé la question peut être qu'effectivement si c'est qu'une discipline ça cible déjà directement. Mais comme HAL fait des filtres très efficaces donc quelque part ça revient au même. De mon point de vue c'est pareil.

Enquêteur : Le type de consultation que vous aimeriez retrouver ?

Enquêtée : En plus de celle qui existe ?

Enquêteur : Oui

Enquêtée : Moi je trouve ce dont j'ai besoin, donc euh...

Enquêteur : Est ce que vous aimeriez avoir accès à des données chiffrés comme sur d'autre site ?

Enquêtée : Pas nécessairement.

Enquêteur : Actuellement il y a des documents qui sont sans le texte intégral, est ce que vous aimeriez avoir le texte intégral avec ?

Enquêtée : Euh ben, c'est vrai qu'en général quand on cherche un document c'est qu'on veut le lire. Mais, si les auteurs ont fait le choix de ne pas déposer le document, enfin, je respecte leur choix et si j'ai déjà la notice ça me permet de le retrouver ailleurs. Ce n'est pas gênant.

Enquêteur : Est ce que vous aimeriez avoir un onglet « lien » sur le site HAL par exemple ?

Enquêtée : Oh, oui ça peut être utile.

Enquêteur : Et quels liens aimeriez vous retrouver ?

Enquêtée : Eh ben, peut-être la page personnelle des auteurs par exemple, ou la liste de tous leurs papiers.

Enquêteur : Quelle est pour l'importance de la communication entre chercheurs ?

Enquêtée : Ha très importante c'est ce qui permet d'avancer, et d'améliorer ce qui est en place.

Enquêteur : Elle se fait comment ici la communication entre chercheurs ?

Enquêtée : En local ?

Enquêteur : Oui

Enquêtée : Beaucoup de visu parce qu'on est assez proches les uns des autres donc on n'hésite pas à aller se rencontrer. Pour les gens qui sont un peu plus loin par mail principalement, sinon par échange de fichiers.

Enquêteur : Et l'échange de fichiers se fait par mail aussi ?

Enquêtée : Ça dépend de la taille du fichier. (Rires) Non, en général on utilise des services(...), des services de transfert de fichiers ou alors des SVN, des dépôts SVN sécurisés.

Enquêteur : Aimerez vous avoir sur ces sites là un accès à des forums ou a des espaces de communication entre chercheurs ?

Enquêtée : Sur un site comme Hal, ça me paraît pas nécessaire.

Enquêteur : Et un espace de partage entre chercheurs ?

Enquêtée : Non

Enquêteur : Aimerez vous avoir un accès à un moyen immédiat pour commenter les documents, les corriger, apporter des corrections ?

Enquêtée : Ça pourrait être utile, je ne dis pas que je m'en servirai, mais je pense que ça peut être utile.

Enquêteur : La perception que revêt pour vos collègues les archives ouvertes ?

Enquêtée : Je ne sais pas trop.

Enquêteur : Et pour vous ?

Enquêtée : Moi je trouve ça(...), je trouve ça important, c'est pour ça aussi que je mets(...), je mets mes documents toujours pour que nos travaux soient quand même à disposition, toujours pour pouvoir échanger et puis y expliquer ce qu'on fait.

Enquêteur : Les avantages et inconvénient par rapport au mode de publication traditionnel ?

Enquêtée : Avantages : on touche plus de monde, Inconvénient : moi j'en vois pas.

Enquêteur : Quels sont les freins, les réticences ?

Enquêtée : Moi, non je n'en vois pas

Enquêteur : Quels sont les bienfaits selon vous ?

Enquêtée : Une meilleure visibilité et une recherche plus rapide de papier.

Enquêteur : La représentation que revête pour vous les archives ouvertes ?

Enquêtée : C'est un bon outil qui permet d'être plus visible.

Annexe 5 : Retranscription de l'entretien 4

L'entretien suivant a été réalisé à Lille 1. Avant de commencer l'entretien, je lui ai rappelé le principe d'un entretien semi-directif et je lui ai demandé si je pouvais l'enregistrer. L'entretien a commencé à 17H30 le 31 mai 2011. Voici donc la retranscription de cet entretien :

Enquêteur : Bonjour, est-ce que vous pouvez me parler de votre profession ?

Enquêté : Donc, oui, actuellement je suis post-doc à l'Université de Mons. Euh, donc j'ai fait ma thèse ici à Lille et puis je viens d'être recruté à Bordeaux, donc en septembre, je commence à Bordeaux en tant que maître de conférences et dans le secteur de l'informatique. Je suis enseignant-chercheur.

Enquêteur : Quelles sont les disciplines d'enseignement ?

Enquêté : Ouais, donc c'est surtout l'informatique, donc les matériels informatiques, et j'enseignais un peu à Lille 3, à Lille1, là à Mons

Enquêteur : Quelles sont les disciplines de recherche ?

Enquêté : L'informatique, un peu les mathématiques aussi, c'est à la frontière.

Enquêteur : Quel est le temps de travail pour les activités d'enseignement et de recherche ?

Enquêté : En proportion ?

Enquêteur : Oui

Enquêté : Ça dépend des années, donc là je suis post-doc donc je suis payée que pour faire de la recherche donc euh, cette année, j'ai fait un petit peu d'enseignement, mais j'ai fait 10% d'enseignement on va dire. Mais en général, même pendant ma thèse, j'étais payé que pour faire de la recherche, donc c'est plutôt des dérogations pour pouvoir faire de l'enseignement et qui ont en général une limite. Donc j'ai fait 100 heures d'enseignement dans ma thèse pendant 3 ans et puis, cette année, j'ai fait une vingtaine ou une trentaine d'heures d'enseignements.

Enquêteur : Est-ce que vous pouvez me dire les outils que vous utilisez sur le web ?

Enquêté : Euh, j'en utilise pas mal. Euh, enfin, j'utilise beaucoup le web pour justement les publications c'est-à-dire rechercher les publications soit dont je connais déjà les références, soit pour chercher des références. Donc il y a tous les sites comme Science Direct, CiteSeer, CiteYouLike, euh(...), ben, les sites des éditeurs comme Springer, I3E, etc. Euh, j'utilise quoi comme outil ? Ca me sert

aussi un peu de veille c'est-à-dire euh(...), j'ai aussi des alertes qui me disent quand des sites sont susceptibles de m'intéresser, et des publications sont susceptibles de m'intéresser. Il y a aussi des sites comme euh(...) Archivix, mais ça s'appelle, ça s'appelle aussi « Corr ». En fait ce n'est pas des publications, c'est plutôt des gens qui(...), qui envoient leur recherches en cours. Donc, ce n'est pas soumis à revue, c'est-à-dire que c'est accepté directement mais ça a un statut de rapport technique en quelque sorte. Ça permet à chacun de voir où en est(...), où en sont les autres. Et puis, j'utilise beaucoup Internet, mais...

Enquêteur : Utilisez vous des outils de traitement comme Google Docs, ... ?

Enquêté : Non. Mais j'utilise beaucoup de(...), mais c'est pas web, c'est plus internet. Des systèmes de version de contrôle c'est-à-dire(...), non des systèmes de contrôle de versions c'est-à-dire que quand on écrit un papier à plusieurs par exemple, on centralise les sources, c'est un peu comme Google Doc, mais euh, ce n'est pas via un site internet. Et donc y a ça, hum... Et tout ce qui est mail bien-sûr.

Enquêteur : Est-ce que vous produisez des documents sur le web ?

Enquêté : Ouais.

Enquêteur : Quels types de documents ?

Enquêté : Donc euh(...), euh oui, moi j'ai l'habitude d'y mettre mes publications au format PDF et euh, parfois et même souvent mes exposés, c'est-à-dire quand je fais un exposé, un séminaire, quelque chose, je mets mes slides sur Internet. Euh(...), ah oui, j'utilise aussi des wikis, c'est ça que je voulais dire. Ben en fait, comme on travaille comme ça à plusieurs on utilise tous les outils collaboratifs. Mais, pour mes publications, ouais c'est surtout ça.

Enquêteur : Quelle est votre fréquence d'utilisation de ces outils du web ?

Enquêté : Euh, tout ce qui est collaboratif, j'utilise quotidiennement. Ce qui est(...) enfin tout ça c'est quotidien, tout ce qui est même recherche de publications c'est vraiment...

Enquêteur : Quelle est l'utilité de ces outils pour vous ?

Enquêté : Ben, c'est devenu central, c'est-à-dire qu'on a même perdu l'habitude d'aller dans les bibliothèques : mais, on les utilise toujours sauf(...), donc ce qui est publication hein, sauf les vieux papiers où là il faut aller chercher dans des bibliothèques. Mais en gros ouais. Et puis tout ce qui est euh(...) collaboratif on ne saurait pas faire sans.

Enquêteur : Quelles sont vos pratiques actuelles en matière d'archives ouvertes

Enquêté : Moi depuis 2 ou 3 ans, toutes mes publis, je ne les mets pas directement sur mon site, mais directement sur HAL. Il y a HAL et une branche de HAL qui est pour l'INRIA où je travaillais avant. Quand j'étais en thèse, on mettait tout là-bas, et puis j'ai toujours un co-auteur qui est français, donc on met tout sur HAL-INRIA, et puis il y a aussi ma thèse que j'ai mis sur TEL.

Enquêteur : Vous déposez dans deux archives si je comprends bien ?

Enquêté : Enfin, ouais, enfin surtout dans HAL, et juste ma thèse dans TEL qui est dédié aux thèses.

Enquêteur : Depuis combien de temps utilisez-vous les archives ouvertes ?

Enquêté : Ça doit faire euh(...), ça doit faire 3 ans.

Enquêteur : Quand vous allez sur HAL ou sur TEL, vous utilisez quels outils principalement ?

Enquêté : Donc, à la fois j'envoie mes publications et puis, ça m'arrive de trouver des publications là-bas. En général c'est juste pour ça. Quand c'est(...), il y a deux systèmes d'archives comme Corr où euh(...), ou Archivix, ou euh(...), où je suis abonné aux flux RSS par exemple. Tu sais c'est découpé par domaine donc j'ai vraiment un flux pour mon domaine, et je suis alerté quand un papier est mis en ligne quoi.

Enquêteur : Et y a-t-il une incitation au dépôt ou une obligation ?

Enquêté : Euh, oui, à l'INRIA, il y avait ça pendant ma thèse, en fait ça facilite par exemple les rapports d'activité des laboratoires et puis ça facilite aussi la mise en place d'indicateurs pour les rapports annuels et en général.

Enquêteur : Est-ce que c'est vous qui déposez ou un laboratoire ?

Enquêté : C'est moi, ou un de mes co-auteurs.

Enquêteur : Et en types de documents, vous déposer quoi ?

Enquêté : Donc ma thèse et mes publications.

Enquêteur : C'est quels types de publications ?

Enquêté : C'est soit des articles dans des journaux, soit des travaux présentés dans des conférences.

Enquêteur : C'est plutôt du texte intégral ou des notices ?

Enquêté : Non, nous c'est toujours le texte intégral.

Enquêteur : Quel contenu mettez-vous dans vos documents ?

Enquêté : Ben c'est(...), donc nous dans le secteur informatique, en gros, il y a des conférences qui sont des papiers d'une dizaine de pages avec en général les preuves. Après il y a aussi les versions longues qui sont publiées dans les journaux où là on s'étale un peu plus : ça fait une trentaine de pages en général avec les preuves et plus de détails.

Enquêteur : Est-ce que vous mettez des photographies, tableaux, ... ?

Enquêté : On fait des schémas, beaucoup de schémas mais pas de photos euh ouais.

Enquêteur : Est-ce que vous publiez des publications qui sont déjà publiées dans des revues ?

Enquêté : Oui. Enfin oui et non, c'est-à-dire qu'en général on n'a pas le droit de les mettre. Donc ce qu'on fait, c'est que ce qu'on met en ligne c'est les préprints, les versions qu'on a envoyé avant la mise en forme finale.

Enquêteur : Dans quelle discipline déposez-vous ?

Enquêté : Informatique.

Enquêteur : Quel est le nombre de documents que vous déposez par an ?

Enquêté : Par an, ça va de(...), je sais pas 3 quelque chose comme ça, enfin 3 où je suis co-auteur. Donc si on voudrait faire le bilan global, oui, il faudrait diviser, mais oui, je dirais 3.

Enquêteur : Par rapport à la consultation, comment recherchez-vous ?

Enquêté : Alors, c'est soit j'ai des outils comme RSS qui m'alertent ou soit(...), ou même sur Scopus, par exemple, on peut avoir des alertes quand on est cité par quelqu'un d'autre, en général, c'est des articles intéressants pour nous. Soit je travaille sur un sujet et j'ai besoin de savoir s'il y a déjà eu des travaux et là je recherche par un moteur de recherche ou alors euh... En général, c'est de deux façons : soit on a une idée et on veut savoir si ça a déjà été fait, soit on sait déjà que ça a été fait et on recherche. Enfin, c'est plus dur quand on ne sait pas si ça a été fait parce que des fois les problèmes sont formulés dans d'autres termes et tout ça.

Enquêteur : Est-ce qu'il vous arrive de rechercher par exemple sur HAL par exemple par type de document, par discipline, ... ?

Enquêté : Non, en fait, moi je recherche toujours par un moteur de recherche et HAL est bien indexé, donc euh... Quand c'est sur HAL, c'est sûrement sur un moteur de recherche.

Enquêteur : Quels types de publications consultez-vous ?

Enquêté : Ben des publications, des fois des thèses, euh, parfois plus rarement des présentations, euh(...), des rapports techniques. Ouais c'est tout je pense. Euh, je suis des blogs par exemple aussi de chercheurs pour m'informer.

Enquêteur : Quand vous consultez, est-ce que vous consultez seulement la notice ou le résumé, ... ?

Enquêté : Là aussi ça dépend, c'est-à-dire si c'est pour voir ce qui est fait sur une idée plus vague, je regarde plutôt juste les résumés, mais si c'est vraiment un papier technique où je veux voir le détail, là je lis le papier en entier.

Enquêteur : Quelle est votre fréquence de consultation ?

Enquêté : Euh(...), ben quotidiennement, enfin, la partie veille, c'est quotidien, la partie recherche de papiers par moi-même je dois dire deux fois par semaine.

Enquêteur : Et, vous mettez combien de temps à consulter un document ?

Enquêté : Pour un document, euh en moyenne, moi je dirais 5 minutes. Parce qu'en fait on en voit beaucoup. Il y en a beaucoup qu'on ne lit pas et après quand on en lit un, là ça prend quelques heures, mais... En général, enfin je dis 5 minutes mais peut-être 10 minutes. En lire un, forcément, ça prend beaucoup plus de temps et comme là hier, j'en ai lu 3 entièrement, ça fait à peu près 2 ou 3 heures par document, et pour arriver à sélectionner ces 3 là, avant, on en a lu beaucoup, on a choisi, donc euh... peut-être 20 minutes, ouais, je dirais 20 minutes.

Enquêteur : Est-ce que vous téléchargez systématiquement un document ou est-ce que vous le visionnez seulement ?

Enquêté : Euh, je télécharge toujours, enfin des fois ça dépend des sites, des fois on prend l'abstract et on sait tout de suite si ça va nous intéresser ou pas et si ça m'intéresse, je télécharge.

Enquêteur : Par rapport à la recherche, est-ce que vous utilisez les outils de recherche avancée et recherche simple ?

Enquêté : Souvent la recherche simple, parfois... Enfin moi j'utilise le moteur de recherche... avec(...), enfin des fois j'explique des mots, des choses comme ça, et je ne vais pas forcément sur HAL pour faire ça. Parce qu'en fait, il y a trop de sites comme HAL alors si je devais faire une recherche par site ça serait trop long. En

général je passe par un moteur de recherche et j'essaie d'affiner directement dans le moteur.

Enquêteur : Quel intérêt revêt pour vous le dépôt et la consultation ?

Enquêté : Euh(...), dans les deux cas, ça permet de laisser son travail accessible à tous, donc euh... Enfin, on aurait du mal maintenant à voir comment ça pourrait marcher autrement, mais si on n'aurait pas accès aux travaux des autres, eh ben en général on les citerait pas ou beaucoup moins, on les étudierait moins. Aujourd'hui ça accélère la recherche... Et puis c'est aussi naturel, enfin(...) on est payé par des pouvoirs publics, par tout le monde donc c'est aussi une manière de(...), c'est aussi normal que ce soit accessible à tout le monde. On est pas payé par les éditeurs, c'est normal que ce ne soit pas verrouillé.

Enquêteur : Et avant les archives ouvertes, vous utilisiez autre chose ?

Enquêté : Euh, non, enfin j'ai commencé ma thèse il y a pas très longtemps donc euh... Il y a 3 ans, c'était justement le moment où tout le monde commençait à vraiment utiliser(...), à utiliser massivement HAL. Sinon, avant ça, ce qu'on faisait, c'est qu'on le mettait sur nos pages web.

Enquêteur : Vous avez commencé à mettre sur vos pages web en quelle année à peu près ?

Enquêté : Ben 2007, ouais.

Enquêteur : Quelle était la fréquence de dépôt sur vos pages web à l'époque?

Enquêté : Ben c'est pareil, c'est 2-3 par an.

Enquêteur : Quel est l'intérêt pour vous de la publication scientifique ?

Enquêté : Euh, ben c'est une bonne question. Euh un des gros intérêts c'est de mon point de vue, c'est d'être revu par les pairs parce que ça valide, enfin ça évite le fait que chaque lecteur doive vérifier les travaux des autres. Une fois que c'est vérifié, même si c'est jamais vérifié à 100% mais ça donne du crédit à nos publications et(...), ça c'est un bon point donc ça génère du travail en tant que reporter, mais en général c'est intéressant quand même. Après, l'intérêt de publier, c'est que ça nous force à écrire les choses une fois pour toutes. Sinon, il n'y aurait que les rapports techniques qui évolueraient dans le temps ou on n'écrirait pas(...), souvent on n'écrit pas toujours les choses, on ne publie pas toujours tout ce qu'on trouve parce qu'on n'a pas le temps. Mais c'est vrai que ça nous force à écrire et rendre les choses accessibles.

Enquêteur : Quels sont selon vous les points forts et points faibles de HAL ?

Enquêté : Euh, l'un des points faibles à une époque c'était par exemple d'avoir un rendu en disant « Je veux que HAL me fournisse une page avec toutes mes publications et ça me sert de page de publications. Donc au niveau individuel et par équipe par exemple. Je sais que par équipe, maintenant on le fait, à Lille, là ils le font mais ça a mis un peu de temps à se mettre en place. Mais maintenant c'est bon. Je ne serais pas trop quels seraient les points faibles de HAL. Ah si, il y a un point faible que je trouve, c'est que par exemple HAL rajoute sur les PDF des informations sur le côté, ça dégrade un peu le PDF. En même temps ça l'identifie clairement mais bon enfin... De toute façon comme la personne va le récupérer par HAL, euh, je ne vois pas trop l'intérêt de mettre ce truc en plus. Y a ça que je trouve un peu dommage, mais... Sinon, c'est assez pratique, ce qui est bien c'est qu'on peut versionner, donc quand on a des erreurs qu'on corrige dans le papier, on peut redonner une version et c'est clairement indiqué que c'est une nouvelle version.

Enquêteur : Et ça marche comment ça ?

Enquêté : Ben, en fait, euh(...), on envoie une version, donc la première version, donc on dit où ça a été publié, etc, et ensuite, simplement on dit "je re-upload" le PDF qui est la nouvelle version et ça crée une nouvelle version, un nouveau tag (PDF). Et puis quand on va sur la page de la publication, en haut, on voit toutes les nouvelles versions, on voit toujours la dernière version par défaut et on peut cliquer pour voir la version d'avant,... Donc ça permet d'être honnête dans le sens où s'il y avait une erreur, qu'elle n'y reste pas. Après, euh... Après, quand on cite un papier, il faut toujours citer un papier précis, s'il évolue dans le temps, on ne sait pas ce qu'on cite vraiment. Donc ça c'est bien et par contre enfin voilà ça permet de corriger en temps voulu, ce qui est normal, c'est ce qui est fait dans les journaux... Donc ça c'est pas mal.

Enquêteur : Quels éléments aimeriez-vous retrouver sur un tel site ?

Enquêté : Euh, je ne sais pas peut-être par rapport à Corr par exemple, peut-être ce qu'il manque mais je crois que c'est déjà un peu fait, c'est une classification assez fine des travaux c'est-à-dire comment dire, dire « cette publication-là elle est dans tel domaine précis » et que ça permette de suivre tout ce qui est uploadé dans une discipline. Mais je crois que c'est fait. Je pense que c'est un peu trop large par rapport à ce que je voulais, il y a trop de papiers qui ressortent.

Enquêteur : Y a-t-il un type de document archivé que vous aimeriez retrouver ?

Enquêté : Ben moi, c'est surtout les publications, enfin en général ce qui est en journaux, et conférences. Ce qui est rapports techniques général, c'est uploadé sur les sites des universités ou des laboratoires plutôt. Donc euh, je ne pense pas qu'on puisse demander de les mettre sur HAL, ça en ferait trop. Et, même les

présentations, on pourrait les mettre sur HAL. On peut mettre l'article et le PDF. Ce qu'on pourrait rajouter, c'est du multimédia, c'est-à-dire quand on a une vidéo de la présentation, la mettre ou des sites comme SlideShare où on peut faire défiler les slides et puis mettre une bande audio, ... mais il ne faudrait pas trop surcharger le site, il ne faudrait pas que ça pas "trop". Après on peut toujours rajouter des liens vers ces sites...

Enquêteur : Est-ce que vous préféreriez avoir un site multidisciplinaire ou avec une seule discipline ?

Enquêté : Alors, moi je pense que oui mais nous on utilise toujours la HAL-INRIA qui est juste pour l'Informatique. Pour moi ça devrait être multidisciplinaire car toutes les disciplines utilisent en gros les mêmes demandes par rapport à ce genre de site, ça ne servirait à rien de le dupliquer et puis ça facilite les recherches. Donc une fois que les publications sont bien classées par catégorie, je trouve ça plus intéressant de mettre ça sur un seul site.

Enquêteur : Quel type de classification aimeriez-vous retrouver ? Par auteur, ... ?

Enquêté : Ouais, moi en général, enfin ça dépend, je fais soit par auteur, soit par thème comme je disais avant « thème de recherche », soit par auteur pour voir(...) parce qu'en général quand il y a un article qui nous intéresse il y en a quelques uns qui nous intéressent aussi, et là récemment j'ai fait aussi par équipe parce que quand on fait un exposé dans une équipe, on aime savoir ce qu'ils font, donc ça c'est intéressant aussi.

Enquêteur : Est-ce que vous aimeriez avoir accès à des données chiffrées, à des graphiques, ... ?

Enquêté : Non, je sais qu'il y en a sur la HAL il y en a. Mais ça ne m'intéresse pas trop.

Enquêteur : Préféreriez-vous avoir des documents en texte intégral seulement ou non ?

Enquêté : Euh... A l'INRIA, on trouve toujours nos documents en texte intégral, c'est-à-dire que ce soit des conférences ou des versions articles, c'est toujours en version intégrale, j'ai jamais eu trop de problèmes. Je préfère que ce soit le texte intégral.

Enquêteur : Aimeriez-vous trouver un onglet lien sur ce site ?

Enquêté : Pour ajouter des liens relatifs à une publication ? Ouais, c'est ce que je disais avant, par exemple, il y a SlideShare, ...

Enquêteur : Est-ce que vous aimeriez avoir des liens de sites personnels de chercheurs ?

Enquêté : C'est-à-dire via HAL ? C'est-à-dire que via des publications on aurait des liens vers leurs sites ?

Enquêteur : Oui par exemple.

Enquêté : Ça, en fait ce n'est pas tellement un problème. Enfin ça serait plutôt un problème de l'ajouter parce que les liens évoluent tellement vite, les chercheurs changent de labo et tout ça. En général c'est plus rapide de chercher par moteur de recherche, quand on connaît le nom du chercheur. Je pense que ça ce n'est pas forcément une bonne chose, mais par contre des liens vers d'autres articles.

Enquêteur : Quelle est l'importance de ma communication entre chercheurs pour vous ?

Enquêté : Eh ben, c'est compliqué parce que... Disons que c'est sûr qu'on a toujours intérêt à communiquer, à diffuser ce qu'on a fait. Par contre en gros ça a un prix, ça prend du temps d'aller présenter et typiquement la recherche est mondialisée donc si on doit parcourir le globe sans arrêt, ce n'est pas forcément rentable. Ça dépend aussi des disciplines, je sais qu'en informatique on se retrouve souvent dans les conférences donc c'est notre moyen principal de communiquer et les séminaires aussi. En mathématiques par exemple, on ne publie presque pas dans les conférences, c'est dans les journaux. Moi, je serais plutôt sur une approche « journaux » car les conférences ça prend beaucoup de temps, il faut organiser le voyage,... Ça permet de se voir, ça permet de connaître les gens mais souvent on a pas beaucoup de temps pour démarrer quelque chose de constructif dans une conférence alors que dans un séminaire, il y a plus de discussions et souvent c'est plus bénéfique.

Enquêteur : Aimeriez-vous avoir accès à un forum ?

Enquêté : Euh je ne sais pas, souvent il y a déjà des sites qui font ça comme PUBZONE par exemple j'ai vu, où entre chaque papier, il y a un moyen de laisser des commentaires, etc. de noter, même dans SiteSeer, on peut noter quelques publications. Sur le fond c'est intéressant, après il faudrait qu'un site fasse délimiter. Après, je pense pas que HAL le fera car c'est un site français. Les publications qui y sont, c'est surtout des publications françaises. Faut que ça reste constructif, faut pas qu'il y ait des tensions. Je ne sais pas trop comment ça évoluerait, c'est un peu difficile à dire. Mais c'est sûr que on aimerait bien avoir l'avis des gens sur nos papiers, mais après c'est vrai que notre avis évolue dans le temps.

Enquêteur : Aimeriez-vous avoir accès à un espace de partage ?

Enquêté : Ça moi je le fais avec les co-auteurs, j'utilise GFORGE INRIA, c'est un outil pour ça et les wikis en général. Sinon, souvent ce qui est technique c'est soit des réunions, soit par mail, voilà, c'est comme ça qu'on avance dans les travaux. Et après ce qui est rédaction, c'est un outil collaboratif de type SVN ou CVS pour mettre en commun les documents retravaillés.

Enquêteur : Quelle est la perception que vous avez des archives ouvertes ?

Enquêtée : C'est Positif parce qu'en gros c'est ce que chacun faisait sur sa page web, ça permet de centraliser, justement de classer et de retrouver plus facilement.

Enquêteur : Quel est l'avantage pour vous des archives ouvertes ?

Enquêté : En gros, c'est un peu ce que je viens de dire. C'est de pouvoir classer par labo, équipe, auteur, de pouvoir retrouver des papiers plus facilement. Ça centralise les publications des chercheurs en un seul endroit et ça permet de trouver des documents qui sont liés par l'auteur, le thème par exemple...

Et puis ça force les gens à mettre leurs papiers car souvent les labos nous poussent à ce que tout soit mis sur HAL. Enfin je dis souvent, mais nous dans notre domaine c'est souvent. Comme ça, les rapports d'activités on les fait facilement, simplement en regardant tous les papiers du labo sur une année par exemple et à mon avis, la ligne de force, c'est que ça va de mieux en mieux renseigner. Même les thèses sur TEL, je sais qu'on incite les gens qui ont passé des thèses même il y a longtemps à mettre leur thèse sur TEL. C'est pareil, les thèses pour les consulter, ce n'est pas toujours en ligne, ce qui est dommage parce que c'est beaucoup plus pratique.

Enquêteur : Quels sont les inconvénients de ces sites là ?

Enquêté : Euh, je n'en vois pas beaucoup. Peut-être qu'un inconvénient c'est(...), enfin il y a toujours une petite ambiguïté : sur le site d'un éditeur, on est sûr que la version conférence d'un article, conférence de journal, c'est-à-dire la version finale d'un article celle qu'il y a sur l'éditeur soit bien celle qui est publiée alors que sur une archive ouverte comme HAL, on n'est pas sûr c'est-à-dire que c'est l'auteur qui a envoyé ça, il y a peut-être déjà eu des modifications, corrections... Ça c'est un inconvénient, mais bon pour moi ce n'est pas très grave parce que si ça a changé, logiquement c'est en mieux donc euh... Peut-être qu'on ne regarde pas ce qu'on pense citer mais ce n'est pas un gros inconvénient.

Enquêteur : Quels sont les freins pour vous ?

Enquêté : Ben, les freins, c'est qu'il faut mettre en ligne un papier, il faut tout renseigner, enfin quand on l'a fait une fois, ça va, mais au début, il faut se faire un peu la main, y arriver, ça peut prendre du temps. Après, il faut avoir confiance dans le site. Comme il y a beaucoup de changement au niveau politique des instituts

etc., si là, HAL viendrait à fermer, on ne sait pas trop ce que ça deviendrait, on n'a pas trop d'assurance sur le long terme, que tout va être gardé, etc. Il faut quand même garder nos papiers localement, pas loin en cas où on doit passer à un autre site. Moi j'ai assez confiance mais on ne sait jamais.

Enquêteur : Quels sont les bienfaits ?

Enquêté : C'est accessible à tout le monde : chercheurs et autre. Pour HAL, l'histoire des versions et tout ce qui est diversification des papiers, le bon référencement de HAL et de ses articles,...

Peut-être un des inconvénients aussi que je n'ai pas dit avant, c'est que c'est Français, ça aurait été pas mal d'avoir un site mondial pour ça, mais en même temps il y a toujours un risque que tout soit centralisé et si ça ferme,...

Donc globalement pour moi, ça reste un site web assez simple, avec une base de données assez bien derrière.