Analyse d'usage des archives ouvertes dans le domaine des mathématiques et l'informatique

Anna Wojciechowska

▶ **To cite this version:**

Anna Wojciechowska. Analyse d'usage des archives ouvertes dans le domaine des mathématiques et l'informatique. Jan 2006. sic_00001707v2

HAL Id: sic_00001707

https://archivesic.ccsd.cnrs.fr/sic_00001707v2

Submitted on 15 Mar 2006 (v2), last revised 15 Mar 2006 (v3)

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.

Analyse d'usage des archives ouvertes dans le domaine des mathématiques et l'informatique

Anna Wojciechowska, Janvier 2006

Résumé :
Cette étude analyse l'auto-archivage des articles d'une partie de la communauté mathématique et informatique en France liée aux bibliothèques du Réseau National des Bibliothèques en Mathématiques. L'objectif du questionnaire qui a été envoyé via certaines bibliothèques du RNBM visait l'usage des archives ouvertes institutionnelles en France, en l'occurrence, le Hal.

1. INTRODUCTION

Contexte

Le libre accès à l'information scientifique et technique (IST) occupe une place de plus en plus importante. Le mouvement en faveur des «Archives Ouvertes» s'est développé au point que ce principe est admis comme essentiel à la communication des résultats de la recherche et à la diffusion des connaissances.
Ce libre accès à l'IST peut prendre deux formes :

- l'auto-archivage d'articles sur le Web en général, et particulièrement sur les serveurs à accès public (archives ouvertes),
- la publication de revues électroniques librement consultables.

L'objectif du questionnaire visait l'usage des archives ouvertes. En ce qui concerne les journaux libres, quelques questions ont eu pour but juste de signaler les connaissances des chercheurs sur ce sujet.

En 2001, dans le cadre de la collaboration entre le CNRS[1] et l'INRIA[2], en partenariat avec la cellule MathDoc[3], le CCSD[4] a développé le «Hal[5]» un outil de communication scientifique entre chercheurs qui dans le cadre du mouvement OAI (Open Archive Initiative) prévoit une promotion d'un auto-archivage institutionnel.

Pour les mathématiques (entre autres), chaque fois que la sous discipline correspondante existe dans «ArXiv[6]», le serveur Hal y envoie automatiquement une copie du document.

Le CNRS, l'Inserm[7], l'INRA[8] et l'INRIA sont des signataires de la déclaration visant à promouvoir le libre accès à l'information scientifique et technique (Berlin, octobre 2003[9]). Le 22 mars 2005 ils ont signé un accord portant sur une politique commune pour développer des archives ouvertes et la création des archives institutionnelles propres à chaque établissement et compatibles, permettant aux chercheurs de déposer leurs publications.

Le chercheur se retrouve au centre du dispositif d'archivage. Ses pratiques et ses besoins par rapport aux dépôts de publications dans les archives ouvertes ont été visées par cette enquête.

Présentation

Un questionnaire sur l'utilisation des archives ouvertes a été adressé auprès des mathématiciens et informaticiens. Ces derniers font partie des lecteurs des bibliothèques du réseau RNBM[10].
Les domaines (mathématiques et informatique) n'ont pas été choisis par hasard :

1 CNRS : Centre National de la Recherche Scientifique, <http://www.cnrs.fr/>
2 INRIA : Institut National de Recherche en Informatique et en Automatique, <http://www.inria.fr/>
3 MathDoc : Cellule de Coordination Documentaire Nationale pour les Mathématiques, <http://www-mathdoc.ujf-grenoble.fr/>
4 CCSD : Centre pour la Communication Scientifique Directe, <http://www.ccsd.cnrs.fr/>
5 Hal : Hyper Article en Ligne, <http://www.hal.ccsd.cnrs.fr/>
6 ArXiv : <http://arxiv.org >ou <http://ccsd.cnrs.fr/arXiv.html>
7 Inserm : Institut National de la Santé et de laRecherche Médicale, <http://www.inserm.fr/>
8 INRA : Institut National de la Recherche Agronomique, <http://www.inra.fr/>
9Max Planck Gesellschaft, Declaration Berlin 2003, <http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html>
10 RNBM : Réseau National des Bibliothèques de Mathématiques, <http://www.rnbm.org/>

- actuellement on parle beaucoup des archives ouvertes en SHS[11] (en faisant des comparaisons avec STM[12]),
- il s'agit des domaines qui concernent l'auteur de cette enquête qui est documentaliste dans une bibliothèque de mathématiques et informatique et qui a voulu vérifier les connaissances et les usages des archives ouvertes de son entourage.

La première version du questionnaire a été testée au mois de mai 2005 sur certains chercheurs du Laboratoire d'Analyse, Topologie et Probabilités à Marseille (LATP). A partir des discussions qui en sont résultées, certaines questions ont été reformulées, supprimées ou ajoutées.
La version corrigée a été envoyée début juillet 2005 aux membres des laboratoires marseillais. Les résultats obtenus ont apporté encore quelques corrections à la version finale du questionnaire qui a été envoyé par e-mail au mois de novembre 2005 aux mathématiciens et informaticiens via certaines bibliothèques du RNBM.
Les documentalistes du RNBM ont diffusé ce questionnaire auprès de leurs lecteurs, en l'occurrence des mathématiciens et informaticiens. Le nombre de personnes atteintes par ce questionnaire est estimé à 2200. 128 personnes ont répondu.

Participants

Douze centres de recherche ont participé à l'enquête : Besançon, Bordeaux, Clermont Ferrand, Grenoble (Imag et Institut Fourier), Marseille, Nancy, Paris (Jussieu et Orsay), Rouen, Strasbourg.

Des réponses incomplètes, n'ont pas permis de connaître les proportions entre les mathématiciens et les informaticiens. Il nous semble regrettable de ne pas pouvoir faire la distinction. Les différences de publications (revues, conférences, ...) entre les mathématiciens et informaticiens doivent induire des différences dans le dépôt (et l'accès) des articles.

Parmi les 128 personnes qui ont participé à l'enquête il y avait surtout des enseignants-chercheurs et des chercheurs CNRS. La participation selon leur statut est présentée dans le tableau 1 :

Tableau 1

Participants	***Pourcentage***
chercheur CNRS	20,30%
enseignant-chercheur	56,30%
autre	23,40%

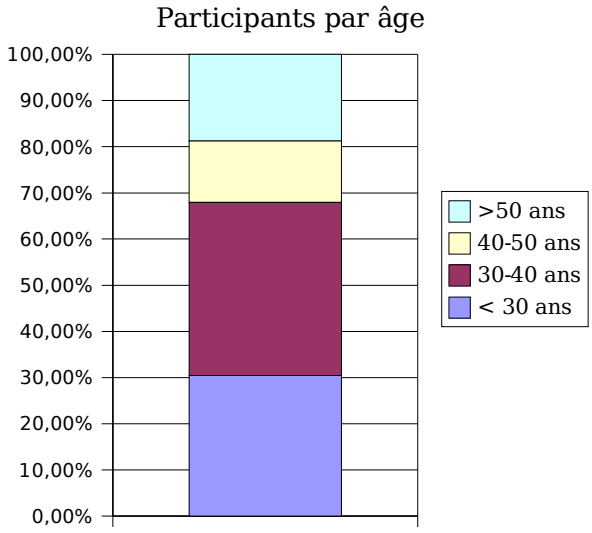
Parmi les « autres » il y avait des doctorants, des post-docs et des retraités.

La productivité d'un chercheur n'est pas la même au début, au milieu ou à la fin de sa carrière. L'utilisation de nouvelles technologies pour accéder à l'information n'est pas la même non plus. D'où les statistiques concernant l'âge des participants pour tenter de dégager des différences dans les pratiques en fonction de ce critère.
Presque 70% de participants ont moins de 40 ans. La participation à l'enquête selon l'âge est présentée dans le tableau 2.

Tableau 2

Âge	***Pourcentage***
< 30 ans	30,50%
30-40 ans	37,50%
40-50 ans	13,30%
>50 ans	18,70%

11 SHS : sciences humaines et sociales
12 STM : sciences, technologie et médecine



Presque la moitié de chercheurs participant à l'enquête disent connaître le terme «archives ouvertes». Un entretien permettra de préciser cette connaissance.

Tableau 3

Est-ce que vous connaissez le terme « archives ouvertes »	***Pourcentage***
oui	47,00%
non	53,00%

Nous avons voulu savoir comment les chercheurs ont appris l'existence des archives ouvertes, quelles sont leurs motivations pour effectuer les dépôts, et ce qu'ils savent sur le droit d'auteur et sur les journaux ouverts.

Les sources probables d'informations sur les archives ouvertes sont présentées dans le tableau 4. 110 personnes ont répondu à la question « Comment avez vous appris la possibilité d'archiver vos publications dans les archives ouvertes institutionnelles ? ».

Tableau 4

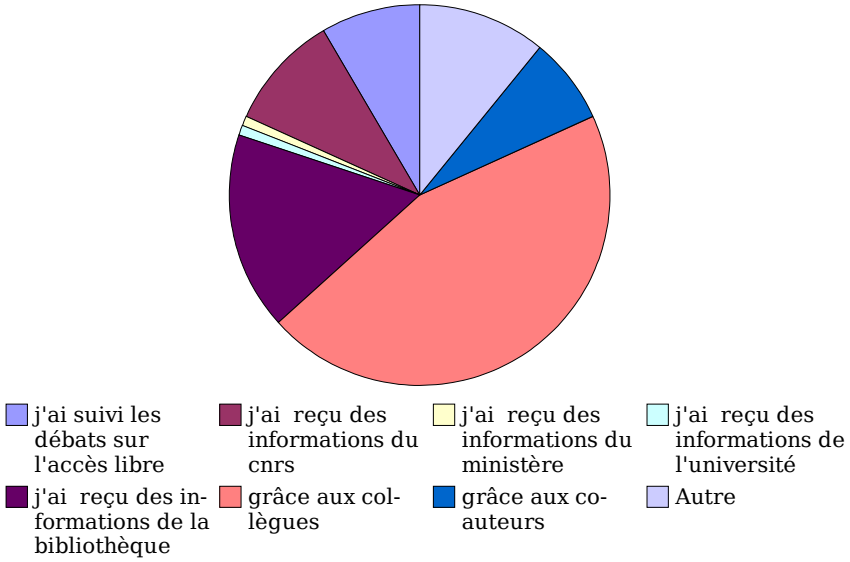
Source d'information sur les archives ouvertes	***Pourcentage***
j'ai suivi les débats sur l'accès libre	7,80%
j'ai reçu des informations du CNRS	9,40%
j'ai reçu des informations du ministère	0,80%
j'ai reçu des informations de l'université	0,80%
j'ai reçu des informations de la bibliothèque	15,60%
grâce aux collègues	42,20%
grâce aux coauteurs	7,00%
autres	10,00%

« Les collègues », c'est la source d'information sur l'existence des archives ouvertes institutionnelles pour 42% de chercheurs.

Autres sources signalées :

- ce questionnaire,
- les informations de l'Inria.

Source d'information sur les Archives Ouvertes



Parmi ceux qui archivent leurs publications dans les archives ouvertes institutionnelles, 55% le font par principe et 25% parce que les archives ouvertes existent (tableau 5).

Tableau 5

Motivation	***Pourcentage***
je suis motivé par principe	54,70%
les articles en accès libres sont cités plus que les articles publiés dans les journaux payants	7,80%
encouragement de la bibliothèque et l'administration	6,30%
encouragement des collègues et co-auteurs	11,00%
existence des archives ouvertes dans votre domaine	25,00%
autres	11,00%

Parmi d'autres motivations signalées :

- l'obligation institutionnelle,
- protection des résultats,
- protection de la date d'un travail non-publié.

Les directeurs des départements scientifiques du CNRS ont envoyé aux laboratoires une demande de dépôt d'ensemble de publications sur le serveur Hal, mais le CNRS ne prévoit pas d'obligation de le faire.
La question « Quelle serait votre réaction si votre employeur (CNRS ou Ministère) exigeait le dépôt de vos publications dans les archives ouvertes ? » (tableau 6) était volontairement provocatrice et pourtant 74% de personnes interrogées seraient d'accord avec cette exigence.

Tableau 6

	Pourcentage
je me soumettrais volontairement	74,20%
je me soumettrais à contre-coeur	7,80%
je ne me soumettrais pas	6,20%

2. RECHERCHE DE L'INFORMATION

Les chercheurs sont aussi bien les auteurs que les lecteurs des articles. Pour cette raison la

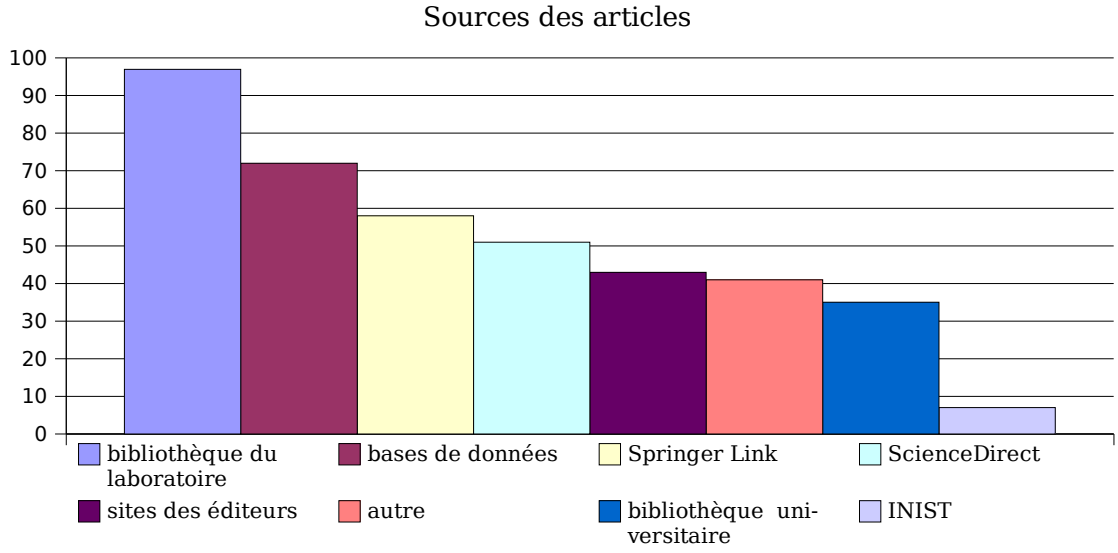
première partie des questions posées a concerné les chercheurs-lecteurs et la recherche de l'information scientifique nécessaire pour leur travail. Il s'agit surtout des recherches de références bibliographiques et de textes intégraux d'articles, aussi bien récents qu'anciens.

Nous avons voulu savoir où les chercheurs les trouvent, comment ils y accèdent et avec quelle facilité, s'ils ont besoin d'aide de professionnels de la documentation, quel est l'âge des articles consultés et la fréquence des consultations des articles électroniques et des prépublications.
A la question : « Où obtenez-vous les articles dont vous avez besoin ? » les chercheurs pouvaient donner plusieurs réponses (tableau 7).

Tableau 7

Source des articles	***Pourcentage***
bibliothèque du laboratoire	75,80%
bibliothèque universitaire	27,30%
bases de données[13]	56,30%
ScienceDirect[14]	39,80%
sites des éditeurs	33,60%
Springer Link[15]	45,30%
INIST[16]	5,50%
autre	32,00%

Même si 76% de personnes interrogées trouvent les articles (ou leurs références) dans la bibliothèque du laboratoire, déjà 56% citent les bases de données comme source d'information, et les journaux en texte intégral (les abonnements payants) sont de plus en plus consultés : 45% de chercheurs trouvent les articles dans Springer Link et 40% dans ScienceDirect. 32% donnent d'autres sources d'articles, comme les pages Web personnelles, ArXiv, Hal, d'autres abonnements en lignes, comme Jstor[17], CiteSeer[18], ou le contact direct avec les auteurs.

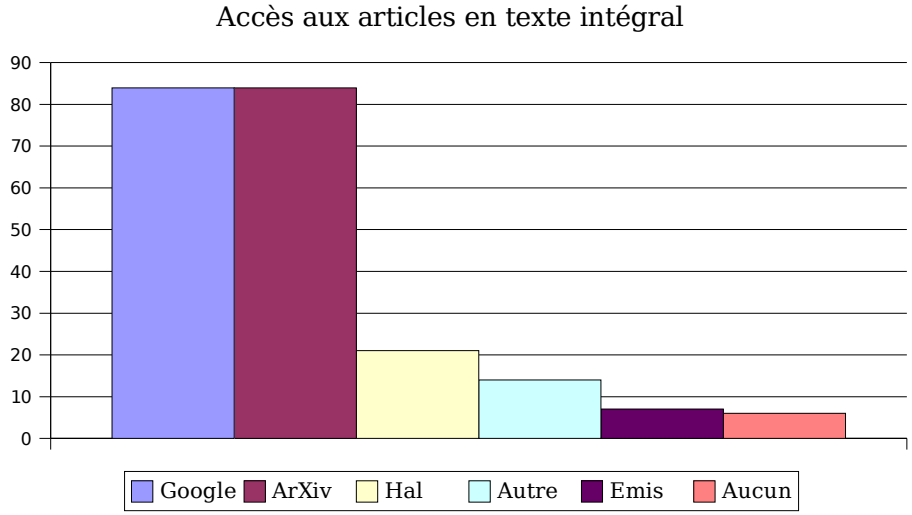


D'après les résultats, l'utilisation des bibliothèques reste importante malgré l'accès au texte intégral des publications. Un autre questionnaire, visant cette fois les professionnels de la

13 Surtout MathSciNet (<http://e-math.ams.org/mathscinet/>) et Zentralblatt (<http://www.emis.de/ZMATH/>), les accès payés par les bibliothèques
14 ScienceDirect : journaux en texte intégral d'Elsevier, l'accès payant dans le cadre du consortium Couperin, <http://www.sciencedirect.com> et <http://www.couperin.org/>
15 SpringerLink : l'accès payant aux journaux en texte intégral de Springer, Birkhauser et Kluwer, <http://www.springerlink.com>
16 INIST : Institut de l'Information Scientifique et Technique, <http://www.inist.fr/>
17 Jstor : The Scholarly Journal Archive : <http://www.jstor.org/>, l'accès payant
18 CiteSeer : Scientific Literature Digital Library, <http://citeseer.ist.psu.edu/>, l'accès payant

documentation, pourra donner des précisions sur l'utilisation des bibliothèques par les chercheurs.

Parmi les sites proposés comme les points d'accès aux articles en texte intégral en accès libre les chercheurs en majorité ont choisi Google (66%) et ArXiv (66%). Parmi d'autres réponses on peut trouver : les pages personnelles ou Numdam[19].



Google permet en effet de trouver les pages personnelles, les pages des laboratoires ou des bibliothèques qui possèdent les serveurs locaux où les publications scientifiques sont stockées, il permet aussi d'accéder directement aux articles.
L'utilisation d'ArXiv signifie surtout l'accès aux prépublications électroniques, même si on y trouve des articles publiés.

La question concernant les points d'accès aux preprints électroniques a confirmé les connaissances du HAL et d'ArXiv (pour 58%), mais 77% de chercheurs accèdent aux preprints électroniques via les pages personnelles.

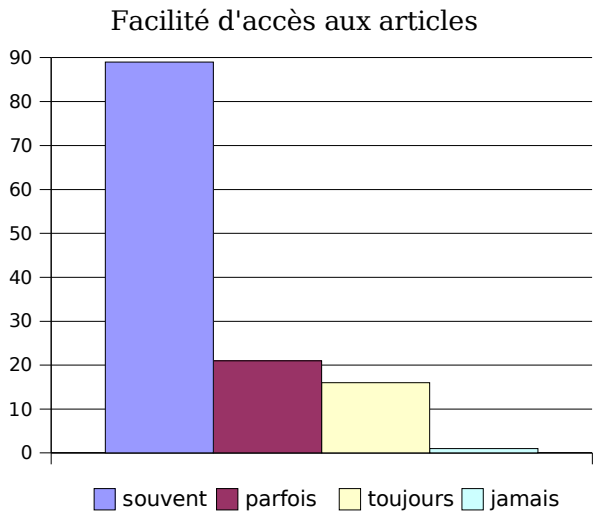
Tableau 8

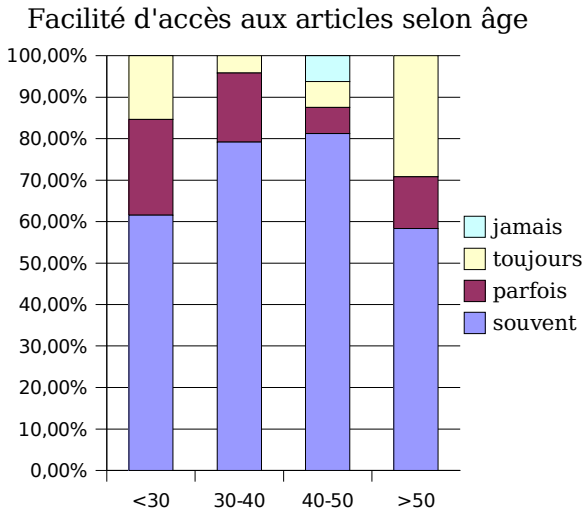
Point d'accès aux preprints électroniques	***Pourcentage***
pages personnelles	77,30%
HAL ou ArXiv	57,80%
sites des laboratoires	18,80%
sites des bibliothèques	14,00%
autres	4,70%

.
L'accès via les sites des laboratoires ou des bibliothèques a été signalé par 18,8%, ce qui signifie l'existence des dépôts locaux.

Les réponses obtenues à la question « Pouvez-vous accéder facilement aux articles dont vous avez besoin pour votre travail ? » montrent que plus de 80% des chercheurs a l'accès facile à la plupart des articles dont ils ont besoin pour leur travail. Et cela ne pose pas de difficultés quelque soit l'âge des chercheurs.

19 Numdam : fonds ancien en mathématiques numérisé, <http://www.numdam.org/>



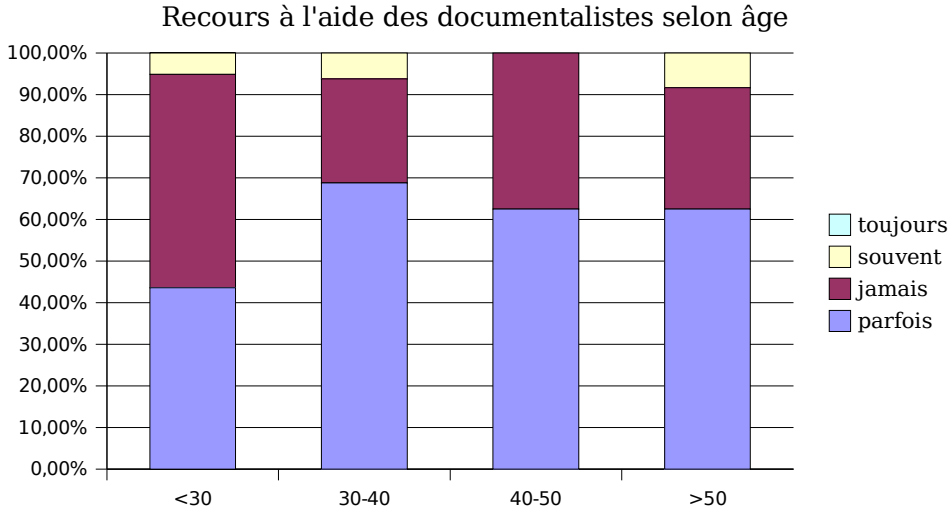


La recherche documentaire, bibliographique fait toujours partie du travail des documentalistes. Les chercheurs s'adressent habituellement aux centres de documentation et aux bibliothèques pour trouver des références ou les publications dont ils ont besoin. Avec la possibilité de trouver l'information scientifique en ligne, ces recherches ont été simplifiées.

Nous avons voulu savoir si cette facilité d'accès aux articles est obtenue grâce à l'aide des documentalistes (tableau 9).

Tableau 9

Demande d'aide	***Pourcentage***
toujours	0,00%
souvent	5,50%
parfois	58,60%
jamais	35,20%



Presque 60% de chercheurs disent utiliser rarement les compétences professionnelles des documentalistes et 35% n'ont pas besoin d'aide pour faire les recherches documentaires et cela concerne toutes les tranches d'âge.

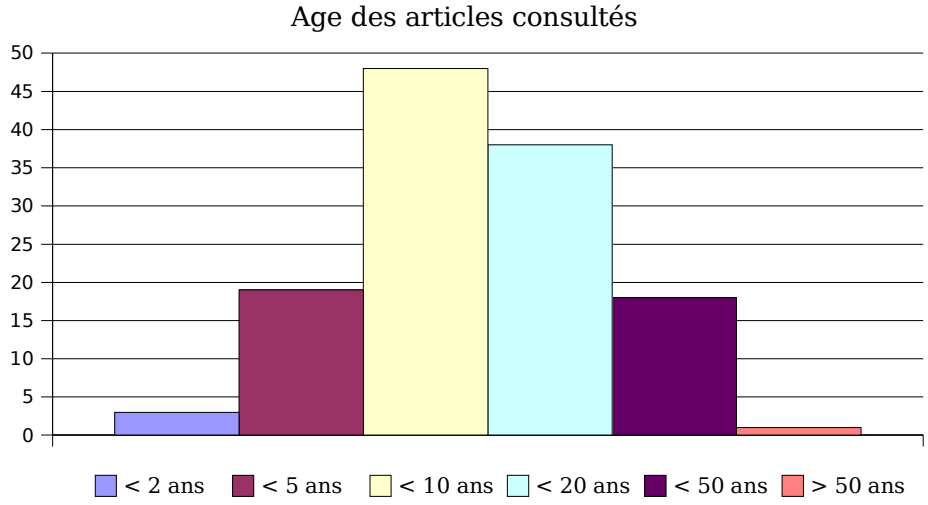
Nous savons que les articles en texte intégral sont plutôt récents (d'habitude à partir de 1995). Il existe aussi de plus en plus d'articles anciens numérisés dans le cadre de divers projets locaux, nationaux ou internationaux qui sont en ligne en accès libre.
Si les chercheurs n'ont pas forcement besoin d'aide pour leurs recherches documentaires, c'est parce que presque 40% de chercheurs utilisent les articles qui ont été publiés pendant les 10 dernières années, cad. les articles qui sont en majorité disponible en texte intégral en ligne ou dans les bibliothèques en version papier.

Pour savoir plus sur l'âge des articles consultés, nous avons posé la question suivante (tableau 10) : « La majorité des articles que vous consultez ou citez dans vos publications ont été publiés il y a » :

Tableau 10

Âge des articles consultés	***Pourcentage***
moins de 2 ans	2,30%
moins de 5 ans	14,80%
moins de 10 ans	37,50%
moins de 20 ans	29,70%
moins de 50 ans	14,00%
plus de 50 ans	0,80%

Presque 40% de chercheurs utilisent des articles publiés pendant les 10 dernières années, cad. des articles qui sont en majorité disponibles en texte intégral en ligne.



Les publications électroniques sont consultées par presque 60% des personnes interrogées une fois par semaine ou plus et par rapport à la question précédente, il s'agit d'articles publiés pendant les 10 dernières années.

Tableau 11

Fréquence de consultation des articles électroniques	***Pourcentage***
une fois par semaine ou plus	57,80%
une fois par mois ou plus	31,30%
une fois par trimestre ou plus	10,20%

Presque tous (93%) consultent les preprints en ligne, mais, comme on va le voir plus loin, ils sont moins nombreux à déposer leurs prépublications en ligne en accès libre.

Tableau 12

Consultation des preprints électroniques	***Pourcentage***
oui	93,00%
non	7,00%

3. PUBLICATIONS

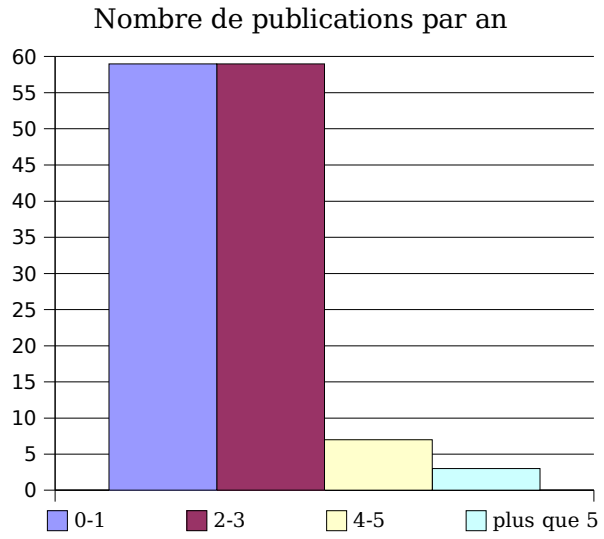
Cette partie analyse le nombre de publications de chercheurs, les façons de les déposer dans les archives ouvertes, le nombre et les types de dépôts.

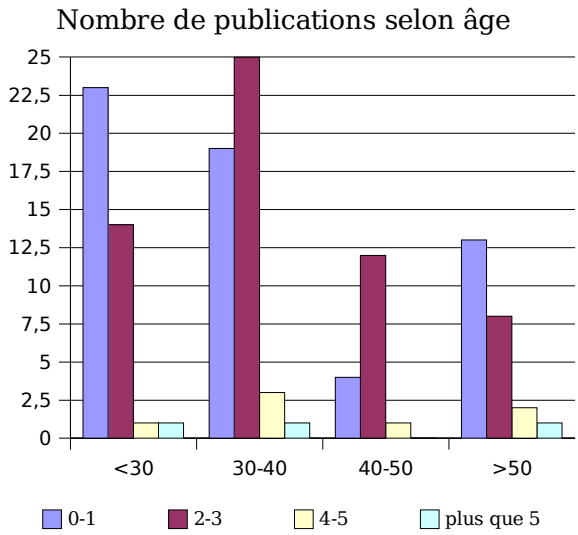
Une publication peut avoir la forme d'un preprint (prépublication), c.à.d. d'un texte écrit, mais pas encore publié (évalué) ou en cours de validation, ainsi que d'un article référé, validé par un comité scientifique et publié.

Il existe plusieurs façons d'auto-archiver un article (preprint ou postprint). Le chercheur peut déposer une copie de son article sur son site Web personnel, dans des archives ouvertes institutionnelles (HAL, par exemple), ou dans des archives ouvertes par domaine (ArXiv, par exemple).

Nous avons voulu savoir combien d'articles en moyenne un chercheur publie par an.

46% de chercheurs interrogés publient au maximum 1 article par an, mais il s'agit ici surtout des doctorants ou jeunes enseignants. 46% de personnes interrogées déclarent publier 2-3 articles par an.



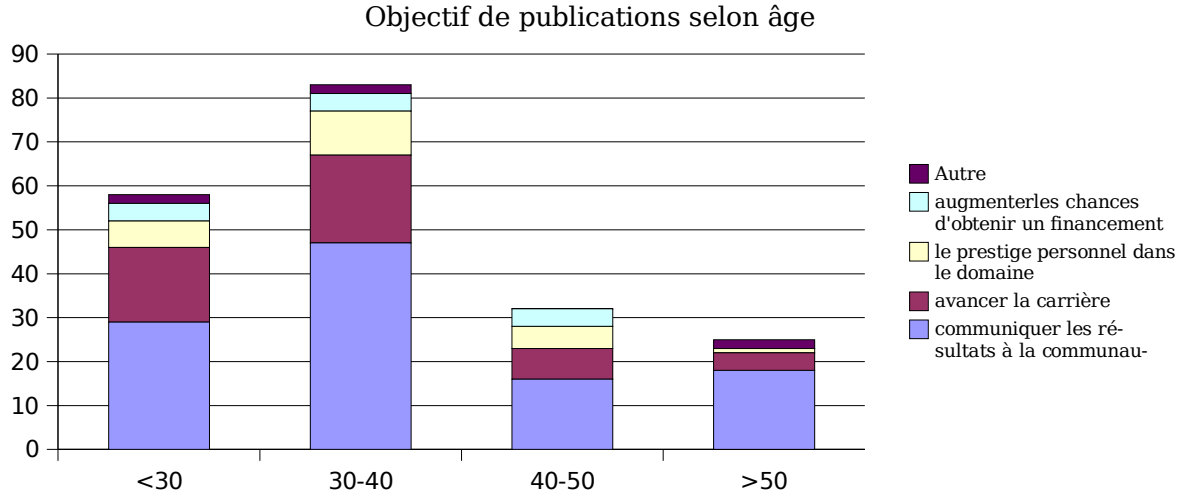


La publication d'un article correspond généralement à une avancée de travaux de recherche. C'est sur la base de ces articles que le chercheur est évalué et financé.
Pour les chercheurs qui ont répondu au questionnaire la priorité est de communiquer les résultats de leur recherche à la communauté scientifique (tableau 13).

Tableau 13

Objectif de publications	*Pourcentage*
communiquer les résultats à la communauté	86,00%
avancer la carrière	37,50%
le prestige personnel dans le domaine	17,20%
augmenter les chances d'obtenir un financement	9,40%
autre	5,50%

Chacun pouvait donner plusieurs réponses, le pourcentage est calculé par rapport à 128 participants.



Parmi d'autre réponses on peut trouver la curiosité scientifique, la volonté de reconnaissance des résultats, la volonté de préciser la formulation des résultats obtenus.

Expérience d'auto-archivage

L'auto-archivage consiste à déposer un document électronique sur un site Web (dans le cas des chercheurs il s'agit de leurs articles) pouvant être consulté gratuitement par tous. Il est effectué afin d'optimiser la visibilité et l'accessibilité à la recherche.

Les publications scientifiques sont en grande partie déposées en ligne par les auteurs et/ou les coauteurs (tableau 14). Certains auteurs publient entre 2 et 4 articles par an, mais ne les auto-archivent pas (les articles sont envoyés directement chez les éditeurs commerciaux).
Les dépôts effectués par les secrétariats des laboratoires d'habitude ne contiennent pas de textes intégraux. Il s'agit ici des dépôts des notices bibliographiques.

Tableau 14

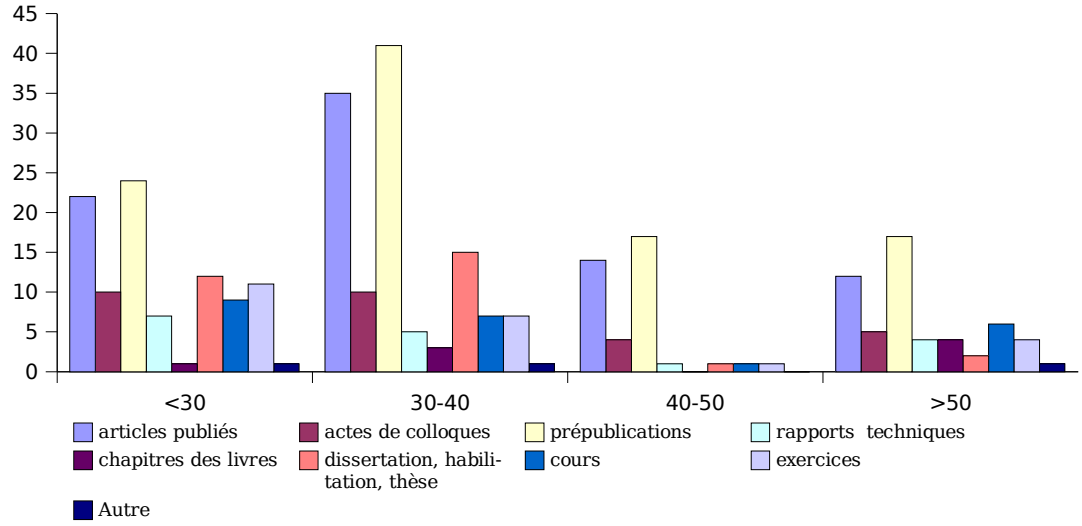
Auteur du dépôt des publications	*Pourcentage*
vous même	74,20%
votre coauteur	10,90%
le secrétariat du laboratoire	11,70%

La question « Quel genre de publications déposez-vous dans les archives ouvertes ? » a permis de faire le point sur le type de publications déposées en ligne (tableau 15).
65% des personnes interrogées ont déposé en ligne des articles publiés et 77% des prépublications. Parmi d'autres types de publications on peut trouver des errata et des fichiers de conférences à projeter (type powerpoint).

Tableau 15

Types de publications déposées	*Pourcentage*
articles référés	64,80%
actes de colloques	22,70%
prépublications	77,30%
rapports techniques	13,30%
chapitres des livres	6,00%
dissertation, habilitation, thèse	23,40%
cours	18,00%
exercices	18,00%
autres	2,30%

Types de publications selon âge



A la question : « Combien d'articles avez vous déposés pendant les 3 dernières années dans des archives ouvertes », une partie des chercheurs a déclaré le dépôt des articles sur les sites Web personnels. Parmi les chercheurs qui n'ont déposé aucun article, la majorité ne possède pas de pages Web personnelles.
Cette question a été divisée en plusieurs sous-questions permettant d'évaluer séparément les dépôts des prépublications et des articles référés sur les sites personnels et institutionnels.
Certaines personnes n'ont pas répondu à ces questions : le dépôt a été fait peut-être par le coauteur, par le secrétariat, ou tout simplement le chercheur n'a pas donné de réponse.

Les chercheurs pouvaient donner plusieurs réponses, ce qui veut dire que probablement ceux qui ont archivé des articles sur leurs sites Web, l'ont également fait (au moins en partie) dans des archives institutionnelles.
Les résultats globaux (tableaux 16 et 18) n'indiquent donc pas le vrai niveau de l'auto-archivage surtout que les articles pouvaient être déposés par d'autre personnes (coauteurs, etc.).

Nombre de prépublications déposées pendant les 3 dernières années — Tableau 16

Preprints déposées sur :	*Nombre de preprints :*	*1*	*2-3*	*4-5*	*Plus de 5*
sites personnels	%	9,40%	28,00%	7,80%	21,00%
sites des bibliothèques ou des laboratoires	%	11,70%	16,40%	1,60%	7,00%
Hal	%	2,30%	11,70%	2,30%	2,30%
ArXiv	%	7,00%	12,50%	4,70%	9,40%

Prépublications sur les sites personnels
66,4% de chercheurs ont déposé au moins 1 prépublication sur un site personnel.
Parmi les 29,7% de ceux qui n'ont rien déposé, la majorité est constituée de doctorants et de personnes de plus de 50 ans.

Prépublications sur les sites des laboratoires ou des bibliothèques
36,7% de personnes ont déposé au moins 1 preprint sur le site du laboratoire ou de la bibliothèque.

Prépublications sur HAL
18,6% de personnes ont déposé au moins 1 preprint sur HAL.
Parmi eux, 87,5% ont déposé au moins 2 preprints sur HAL.

Prépublications sur Arxiv
33,6% de personnes ont déposé au moins 1 prépublication sur ArXiv.
Parmi eux, 79% ont déposé au moins 2 preprints.

73% ne craignent pas de « pillage » ou d'usage abusif des prépublications en ligne, ce qui correspond (plus au moins) au nombre de déposants des preprints dans les archives ouvertes.

Tableau 17

Est-ce que vous craignez l'usage abusif des prépublications en ligne ?	***Pourcentage***
oui	22,00%
non	73,00%

Nombre d'articles déposés pendant les 3 dernières années

Tableau 18

Articles publiés déposés sur :	***Nombre d'articles :***	***1***	***2-3***	***4-5***	***Plus de 5***
sites personnels	%	5,50%	21,00%	9,40%	27,30%
sites des bibliothèques ou des laboratoires	%	2,30%	5,50%	2,30%	7,00%
Hal	%	2,30%	5,50%	2,30%	3,00%
ArXiv	%	4,70%	7,00%	0,80%	3,90%

Articles référés sur les sites personnels
63,3% de chercheurs ont déposé au moins 1 article référé sur un site personnel.

Articles référés sur les sites des laboratoires ou des bibliothèques
17,2% de chercheurs ont déposé au moins 1 article publié sur le site du laboratoire ou de la bibliothèque.

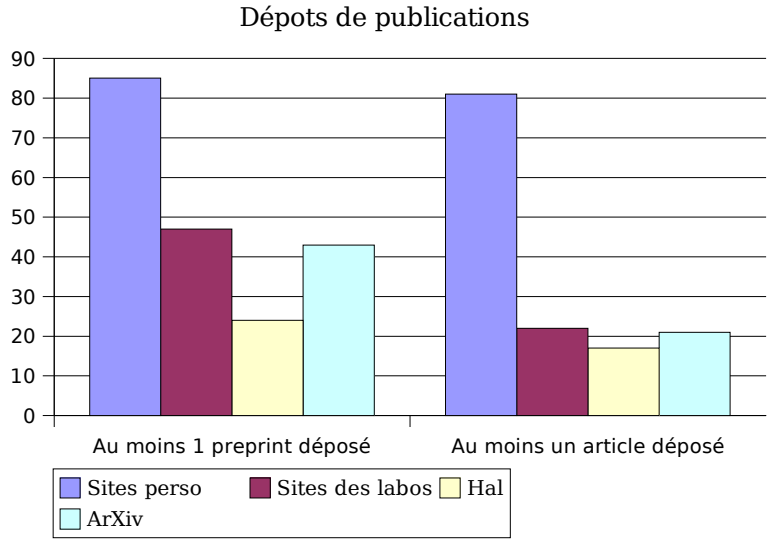
Articles référés sur HAL
13,3% de chercheurs ont déposé au moins 1 article référé sur HAL.
Parmi eux, 82,4% ont déposé au moins 2 articles référés.

Articles référés sur Arxiv
16,4% de chercheurs ont déposé au moins 1 article publié sur ArXiv.
Parmi eux, 71,4% ont déposé au moins 2 articles publiés.

Les réponses ont été données par seulement 128 personnes, mais si on regarde le site du Hal, on s'aperçoit que les publications en mathématiques et informatique déposées ne sont pas aussi nombreuses qu'en physique[20] par exemple.

20 Le 31/12/05 le nombre de publications déposées sur Hal était le suivant : physique : 8479, mathématiques : 2181, informatique : 1736.

Il est impossible de comparer le nombre d'articles publiés par un chercheur par an avec le nombre de dépôts effectués pendant les 3 dernières années, car souvent les chercheurs déposent aujourd'hui les publications plus anciennes.



En ce qui concerne l'absence des dépôts de publications pendant les 3 dernières années, parmi les personnes interrogées :

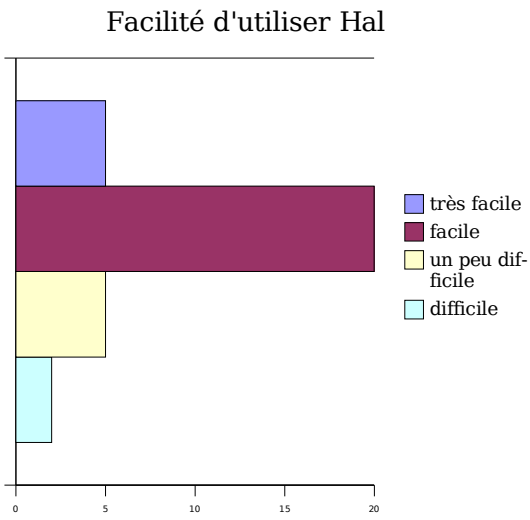
- 33,6% ont déposé 0 prépublication sur les sites personnels,
- 63,3% ont déposé 0 prépublication sur le site de la bibliothèque ou du laboratoire,
- 81,3% ont déposé 0 prépublication sur Hal,
- 66,4% ont déposé 0 prépublication sur Arxiv,
- 36,7% ont déposé 0 article sur les sites personnels,
- 82,8% ont déposé 0 article sur le site de la bibliothèque ou du laboratoire,
- 86,7% ont déposé 0 article sur Hal,
- 83,6% ont déposé 0 article sur Arxiv.

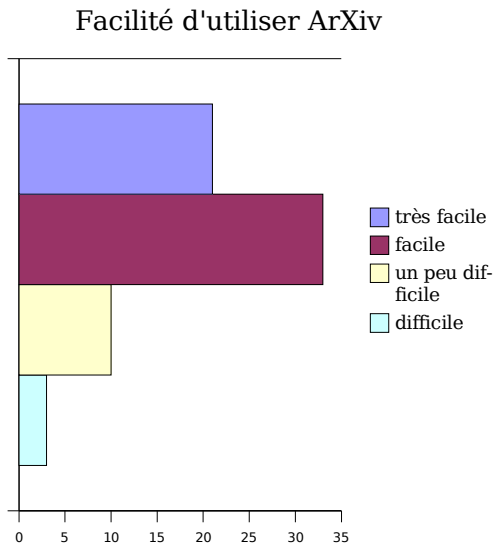
Ensuite, nous avons voulu connaître l'opinion des chercheurs sur l'ergonomie du Hal et d'ArXiv (tableau 19). Le pourcentage est calculé par rapport au 128 participants.

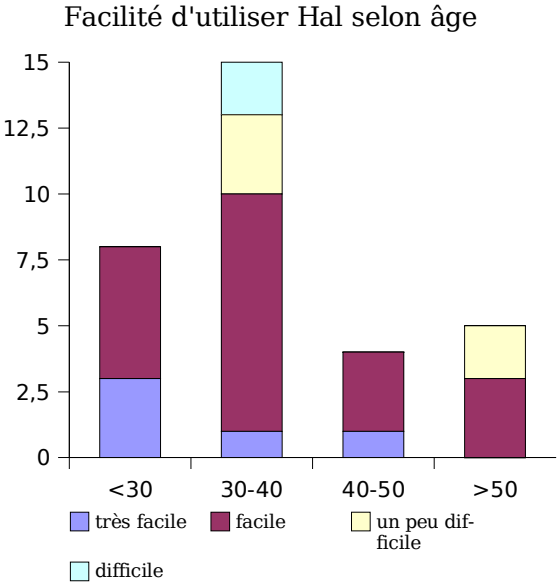
Tableau 19

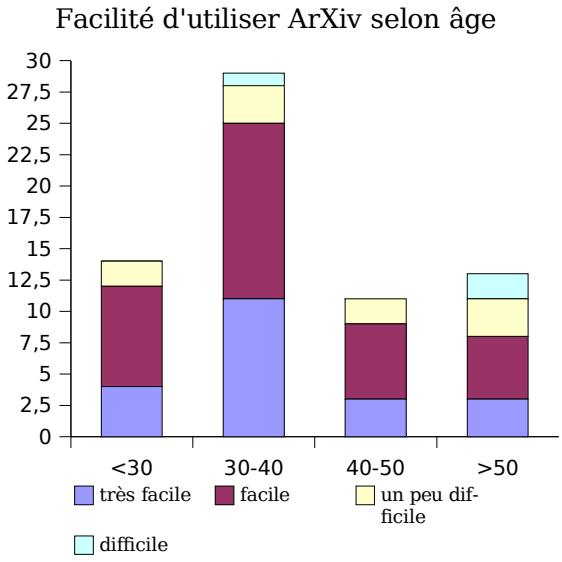
Facilité de dépôt	***Sur Hal***		***Sur ArXiv***	
	Nombre	***Pourcentage***	***Nombre***	***Pourcentage***
très facile	5	3,90%	21	16,40%
facile	20	15,60%	33	25,80%
un peu difficile	5	3,90%	10	7,80%
difficile	2	3,90%	3	2,30%
sans opinion	92	71,90%	60	46,90%

124 chercheurs ont donné une réponse sur l'utilisation de Hal et 127 sur l'utilisation de ArXiv.
Parmi ceux qui ont donné une opinion sur l'utilisation de Hal (32 personnes), 20 le trouvent facile (62,5%).
Parmi ceux qui ont donné une opinion sur l'utilisation de ArXiv (67 personnes), 33 le trouvent facile (presque 50%), 21 : très facile (31,3%). Parmi les réponses « sans opinion », il y a un certain nombre de personnes qui n'archivent pas d'articles.









Le temps nécessaire pour effectuer les dépôts sur Hal et ArXiv est présenté dans le tableau 20. Le pourcentage est calculé par rapport au nombre de personnes qui ont répondu à cette question, c.à.d. 29 en ce qui concerne Hal, et 46 dans le cas d'ArXiv.

Tableau 20

Temps du dépôt	***Hal***				***ArXiv***			
	1er dépôt		***dépôt suivant***		***1er dépôt***		***dépôt suivant***	
	nombre	***%***	***nombre***	***%***	***nombre***	***%***	***nombre***	***%***
< de 15 min	9	31	22	81,5	13	28,3	33	73,3
< de 30 min	10	34,5	5	18,5	19	41,3	9	20
< d'1 heure	4	13,8			10	21,7	3	6,7
> d'1 heure	6	20,1			4	8,7		

Il est difficile de savoir si les réponses ont été données seulement par les personnes qui ont fait des dépôts, ou aussi par celles qui ont tenté de le faire.

En ce qui concerne le temps nécessaire pour effectuer un dépôt sur Hal :
le premier dépôt prend moins de 30 min pour 10 personnes (34,5%) et le dépôt suivant moins de 15 min pour 22 personnes (81,5% par rapport au 27 qui l'ont effectué).

En ce qui concerne le temps nécessaire pour effectuer un dépôt sur ArXiv :
le premier dépôt prend moins de 30 min pour 19 personnes (41,3%) et le dépôt suivant moins de 15 min pour 33 personnes (73,3% par rapport au 45 qui l'ont effectué).

Les chercheurs ont eu la possibilité de proposer des modifications de fonctionnement de Hal (tableau 21). Certaines personnes en ont profité ici pour poser des questions sur l'utilisation de Hal ou pour faire des remarques concernant son fonctionnement. Même si cela n'était pas le but du questionnaire, nous leur avons donné des explications et leurs remarques vont être transmises au CCSD.

Tableau 21

Est-ce que vous voudriez proposer des modifications de fonctionnement de HAL	***Pourcentage***
oui	5,50%
non	78,90%

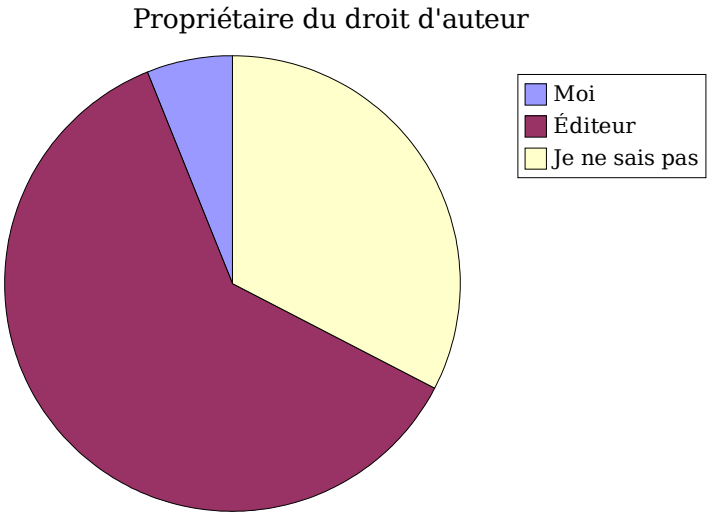
A la question « Quel genre de modifications ? », les réponses ont été diverses :

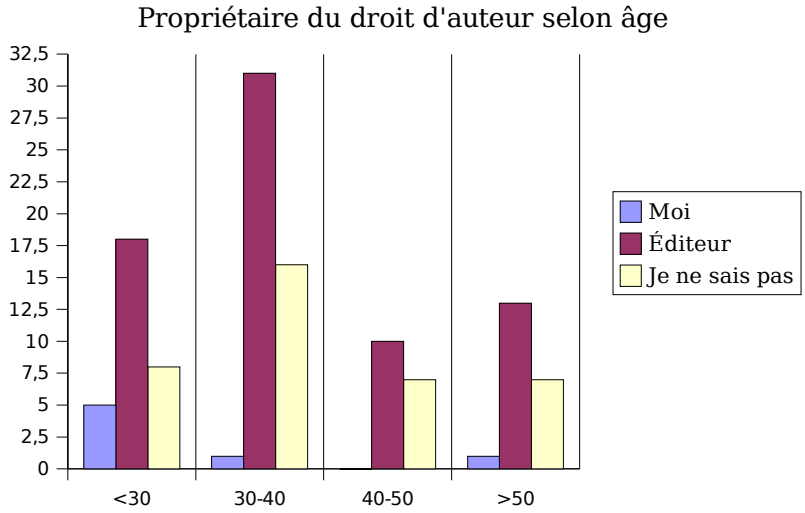
- « donner la possibilité de supprimer le fichier .tex après publication sur ArXiv »,
- « donner la possibilité d'accéder aux statistiques des articles d'autres utilisateurs »,
- « l'import automatique de serveur de publications existant pour les laboratoires et équipes (de type Basilic) »,
- « simplifier : le système veut tout savoir sur l'institution qui accueille chacun des auteurs, et il faut donner beaucoup d'informations à travers des menus déroulants. Ca peut prendre du temps notamment quand le deuxième auteur est dans une institution à l'étranger »,
- « j'ai eu de gros problèmes de compilation LaTex. J'ai certes utilisé un package peu courant (quoique disponible sur le CTAN) pour dessiner des tableaux de Young, mais je n'ai rencontré aucune de ces difficultés avec ArXiv »,
- « accepter sur HAL la version publiée de l'article ».

Les connaissances du droit d'auteur :

Un auteur peut déposer dans les archives ouvertes tout type de document dont il possède la propriété intellectuelle. Cela concerne les documents déjà publiés ou en cours de publication, les documents en cours de validation scientifique (prépublications) ou les documents de travail.
Les auteurs ont le droit de mettre en accès libre leurs propres résultats de recherche même si, lors de la publication dans une revue, un chercheur a signé un contrat d'édition.
Il faut une interdiction explicite dans un contrat de cession de droits (qui n'autorise l'exploitation électronique du document que par l'éditeur) pour être obligé de demander à l'éditeur le droit de déposer ce document dans les archives ouvertes[21].
Nous avons voulu vérifier les connaissances des chercheurs concernant les aspects juridiques de publications scientifiques et savoir si les chercheurs lisent les contrats signés avec les éditeurs commerciaux.
A la question « Qui a le droit d'auteur (copyright) de votre dernier article publié ? »56,2% (sur 128 personnes) disent qu'il s'agit de l'éditeur, 30% ne le savent pas. Seulement 5,5% confirment avoir gardé le droit d'auteur sur le dernier article publié.

21 « Guide juridique du CNRS » en cours de publication (http://publicnrs.inist.fr/)





Les chercheurs qui ont auto-archivé (surtout sur les pages personnelles) leur dernier article publié, ne savent pas en majorité si la permission de l'éditeur était nécessaire pour le faire (tableau 22).

Tableau 22

Demande de permission pour auto-archiver	***Pourcentage***
oui	7,80%
non	18,00%
je ne sais pas	54,00%

10 personnes ont demandé à l'éditeur la permission d'auto-archiver et 7 disent avoir le copyright sur le dernier article publié. Nous ne pouvons pas savoir si dans les 3 cas il s'agit d'un refus de la part de l'éditeur.
La réponse « non » (23 personnes) ou « je ne sais pas » (69 personnes) peut signifier que dans les contrats de cession de droits l'auto-archivage n'était pas interdit ou que les articles ont été publiés sans la permission de l'éditeur.
Il est possible aussi qu'il s'agisse des chercheurs qui ne déposent pas d'articles dans les archives ouvertes.
La réponse « oui » est intéressante, parce qu'elle signifie que 10 personnes voulaient déposer les articles référés dans les archives ouvertes légalement et qu'elles connaissent la politique

éditoriale actuelle en matière de libre accès .
On comprend aussi qu'une partie des articles sur les sites personnels a été déposée avec le copyright de l'éditeur.

A la question « Est-ce que vous savez que vous pouvez négocier avec les éditeurs la possibilité de déposer le même article simultanément dans les archives ouvertes ? »

Tableau 23

	Pourcentage
oui	19,00%
non	77,30%

77,3% (par rapport au 128 personnes) disent de ne pas connaître cette possibilité et si on tient compte du nombre de réponses à cette question (123), on peut parler de 80,5%.

Les connaissances des journaux ouverts

La publication de revues électroniques librement consultables constitue la deuxième forme du libre accès à l'information scientifique.
A la question « Avez-vous soumis un manuscrit dans un journal libre pendant les 3 dernières années ? » seulement 22 personnes (17%) donnent une réponse positive.

Tableau 24

Avez-vous soumis un manuscrit dans un journal libre	***Pourcentage***
oui	17,00%
non	78,00%

Les chercheurs ont été invités à indiquer les raisons pour publier dans un journal libre (tableau 25).

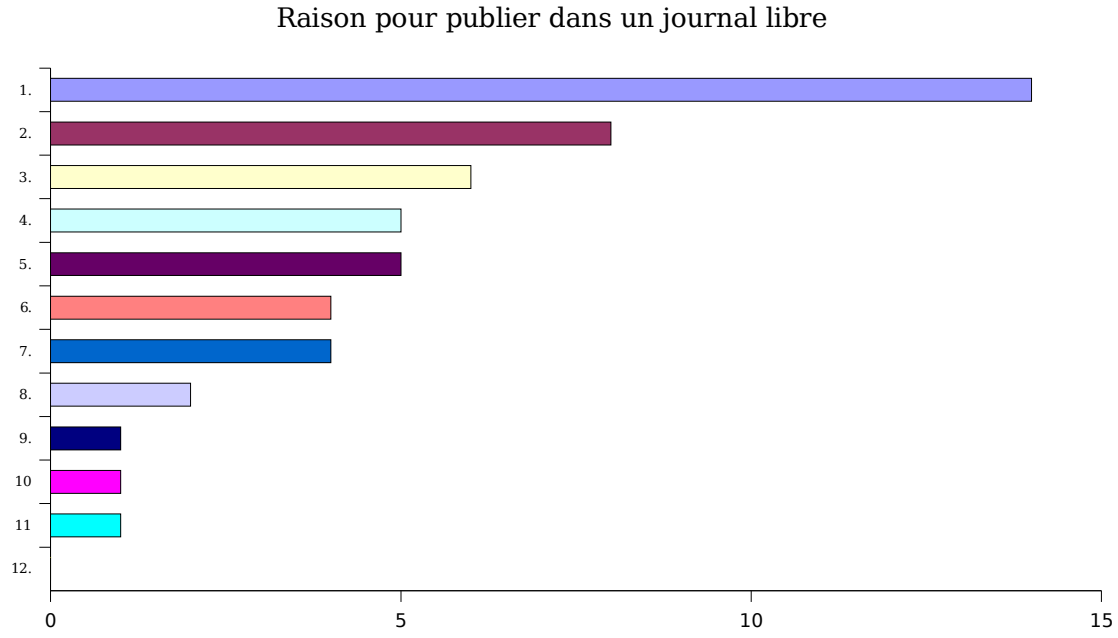


Tableau 25

Raisons pour publier dans un journal libre	*Pourcentage*
1. l'accès libre pour tous les lecteurs	63,60%
2. ce journal est prestigieux dans mon domaine	36,40%
3. ce journal est édité par les gens que je connais	27,30%
4. je me sens concerné par les frais des abonnements de ma bibliothèque	22,70%
5. la publication des articles est plus rapide	22,70%
6. je suis contre les publications dans des journaux commerciaux	27,30%
7. j'étais attiré par l'éditeur ou le comité éditorial	18,20%
8. le lectorat est plus large que dans le cas d'un journal normal (commercial)	9,00%
9. je pense que mon article va être cité plus souvent	4,50%
10. autre	4,50%
11. j'étais influencé par mes collègues	4,50%
12. j'étais influencé par mon laboratoire	0,00%

Le pourcentage est calculé par rapport au 22 personnes qui ont déclaré avoir publié dans un journal libre. Les auteurs qui y publient sont généralement motivés par le principe d'accès libre et par une bonne réputation de ce journal dans leur domaine.

Nous avons aussi demandé aux gens qui n'ont pas publié dans de journaux libres d'indiquer les raisons de ne pas l'avoir fait (tableau 26).

Raisons de ne pas publier dans un journal libre

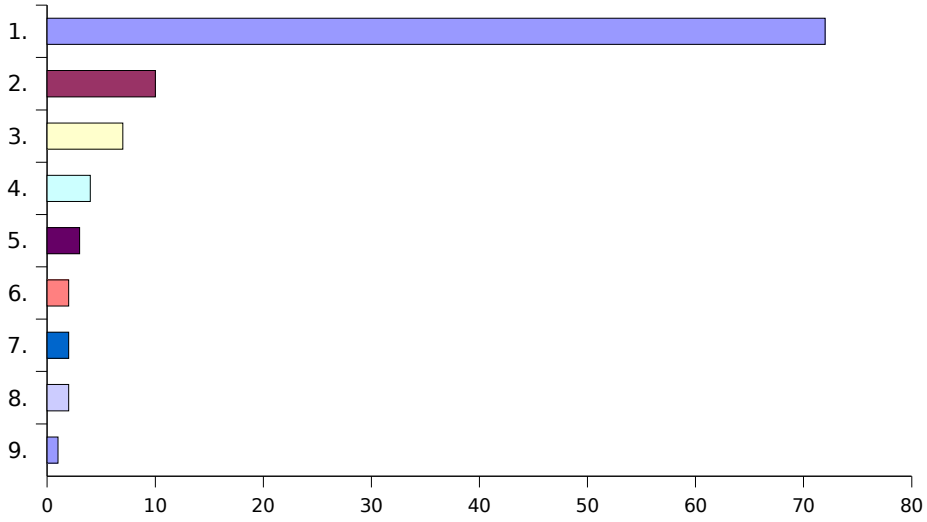


Tableau 26

Raisons de ne pas publier dans un journal libre	*Pourcentage*
1. je ne connais pas assez bien des journaux libres dans mon domaine pour y déposer l'article	72,00%
2. les journaux libres dans mon domaine ne sont pas considérés comme assez prestigieux	10,00%

Raisons de ne pas publier dans un journal libre	*Pourcentage*
3. je suis contre le principe auteur – payeur	7,00%
4. je n'ai pas trouvé de journaux libres dans mon domaine	4,00%
5. je publie toujours dans les même journaux et je suis satisfait	3,00%
6. ma décision a été influencé par mes collègues	2,00%
7. le lectorat d'un journal libre est moins important	2,00%
8. autre	2,00%
9. ma décision a été influencé par mon laboratoire	1,00%
10. la publication des articles est moins rapide que dans des journaux traditionnels	0,00%
je n'ai pas trouvé de financement pour publier dans un journal libre	0,00%

Le pourcentage a été calculé par rapport aux 100 personnes qui ont déclaré de ne pas avoir publié dans des journaux libres.

Les raisons principales de ne pas publier dans un journal à l'accès libre sont les suivantes :

- les chercheurs ne connaissent pas des journaux libres dans leur domaine (72%)
- ou ils disent que les journaux libres dans leur domaine ne sont pas considérés comme assez prestigieux (10%)
- ils sont contre le principe auteur – payeur (7 %)

Autres raisons : les articles refusés.

Les journaux libres sont mal connus et pourtant en mathématiques il y a plus de 80 titres (http://www.doaj.org/).
Le modèle économique « auteur-payeur » qui existe dans d'autres pays n'est pas populaire en France et, probablement, il est mal connu[22].

Nous avons enfin voulu connaître les intentions des chercheurs par rapport aux publications dans des journaux libres. D'où la question : »Si vous n'avez jamais publié dans un journal libre, est-ce que vous envisagez de le faire ? » (tableau 27).

Tableau 27

	Pourcentage
très probablement	12,50%
probablement	21,90%
non	4,70%
je ne sais pas	38,30%

34,4% de chercheurs envisagent de publier dans un journal libre dans le futur et 38,3% ne savent pas encore s'ils vont le faire.

CONCLUSIONS

- La majorité des chercheurs trouvent les articles (ou leurs références) nécessaires pour le travail dans les bibliothèques, mais les journaux en ligne en texte intégral (les abonnements payants des bibliothèques) sont souvent consultés.

- Pour accéder aux articles en texte intégral en accès libre, les chercheurs utilisent surtout ArXiv (pour accéder aux prépublications électroniques) et Google (pour la recherche des pages Web personnelles).

22 par exemple, le nouveau groupe Springer-Kluwer a lancé l'Open Choice : un auteur peut choisir de payer (3500 dollars) pour publier, de façon à ce que son article soit accessible sans abonnement

- Les sources des articles en accès libre ne sont toujours pas très bien connues, comme par exemple les journaux libres.

- Plus de 80% des chercheurs déclarent ne pas avoir de difficultés pour trouver la plupart des articles dont ils ont besoin pour leur travail. Les recherches documentaires ont été simplifiées avec la possibilité de trouver l'information scientifique en ligne. Presque 60% des chercheurs utilisent seulement parfois les compétences professionnelles des documentalistes et 35% de répondants disent ne plus avoir besoin d'aide des documentalistes pour les faire.

- Les articles disponibles en ligne en texte intégral les plus consultés (au moins une fois par semaine) ont été publiés pendant les dix dernières années.

- Pratiquement 50% des chercheurs publient entre 2 et 3 articles par an. La majorité les archivent sur les pages Web personnelles et 28% déclarent le faire depuis au moins 5 ans.

- Une partie des chercheurs qui déposent des publications sur les sites personnels le font simultanément dans les archives ouvertes institutionnelles. Mais les chercheurs déposent toujours beaucoup plus d'articles sur leurs pages personnelles (63%) que dans Hal (12%) ou ArXiv (16%).

- Les personnes qui ont déjà effectué des dépôts de publications dans Hal ou ArXiv trouvent leur utilisation facile et disent qu'il faut moins de 30 minutes pour le premier dépôt et moins de 15 minutes pour faire le dépôt suivant.

- Les auteurs affirment être sensibilisés au libre accès et ils ont appris l'existence des archives ouvertes institutionnelles surtout grâce à la communication entre eux. Les informations officielles provenant du CNRS ou du Ministère ont été plutôt non-aperçues.

- Ceux qui archivent déjà les publications dans les archives ouvertes le font par principe et en raison de l'existence de ces archives.

- Les chercheurs en majorité ne regardent pas les contrats signés avec les éditeurs, ils ne savent pas qu'ils ont le droit d'archiver les articles publiés et qu'ils peuvent négocier cette clause dans le contrat d'édition.

- Même si cela n'est pas le but du CNRS ou du Ministère, la majorité des déposants accepteraient l'exigence du dépôt des publications dans les archives ouvertes institutionnelles.

<u>En ce qui concerne l'auto-archivage d'articles dans des entrepôts électroniques à accès public :</u>

Pour augmenter le nombre de publications déposées dans les archives institutionnelles il faudrait inciter les chercheurs et les aider à adopter ce nouveau mode de publication afin d'améliorer la diffusion de leur production intellectuelle.
Les difficultés du développement des archives ouvertes ne sont pas techniques, mais sociales. Les pratiques d'archivage font déjà partie des coutumes des chercheurs en mathématiques et informatique, mais il s'agit ici du réflexe « google », cad. de déposer les publications sur les sites personnels en s'assurant de pouvoir les trouver via le moteur de recherche le plus utilisé. L'utilité des archives ouvertes institutionnelles n'est pas encore bien comprise.

Le développement des archives ouvertes est fondé sur l'auto-archivage des publications scientifiques par les auteurs. Il faudrait prévoir des formations sur l'utilisation du Hal destinées aux chercheurs et, dans certains cas, envisager le dépôt des publications avec l'aide des documentalistes.
L'évolution actuelle du mouvement des archives ouvertes demande aussi une plus grande visibilité des informations officielles provenant du CNRS ou du Ministère.

Il apparaît indispensable de diffuser de manière plus large les aspects juridiques des publications scientifiques et de sensibiliser les chercheurs à la vérification des contrats signés

avec les éditeurs.

En ce qui concerne la publication de revues électroniques librement consultables :

Selon une étude menée par le Centre for Information Behaviour and the Evaluation of Research (CIBER)[23], de plus en plus de scientifiques publient dans des journaux en libre accès.
En ce qui concerne les mathématiciens et les informaticiens, en espérant pouvoir confirmer cette tendance, il faudrait refaire un questionnaire dans l'avenir proche.

En attendant, il faudra diffuser plus d'information sur des journaux libres qui ne sont pas assez exploités (selon les résultats du questionnaire). Il s'agit en grand partie d'une mission pour les bibliothèques de diffuser des informations provenant des éditeurs de ces journaux.

Ce questionnaire généraliste est une première étape dans la définition des pratiques et des besoins des chercheurs relatifs au mouvement « archives ouvertes ». Pour avoir une image plus fine, il serait nécessaire de refaire le questionnaire dans l'avenir et de le compléter par des entretiens avec les scientifiques et les documentalistes.

Bibliographie

1. Allen, James. Interdisciplinary differences in attitudes towards deposit in institutional repositories (2005). http://eprints.rclis.org/archive/00005180/

2. Aubry C., Janik J. (dir.). 2005. Les archives ouvertes : enjeux et pratiques. Guide à l'usage des professionnels de l'information. Paris: ADBS, 332 p.

3. Centre for Information Behaviour and the Evaluation of Research (CIBER), http://www.ucl.ac.uk/ciber/ciber_2005_survey_final.pdf

4. Chanier, T. 2005. Archives ouvertes et publication scientifique. Comment mettre en place l'accès libre aux résultats de le recherche ? L'Harmattan, Paris, 188p.

5. Fily, Marie-Françoise. 2005. Introduction au concept d'archive ouverte. http://archivesic.ccsd.cnrs.fr/sic_00001523.html

6. Gallezot, Gabriel. Le Libre Accès (Open Access) : partager les résultats de la recherche, Colloque international : L'information numérique et les enjeux de la société de l'Information - Tunis, 14-16 Avril 2005 - ISD. 14 avril 2005. http://archivesic.ccsd.cnrs.fr/sic_00001416.html

7. Guide juridique du CNRS, en cours de publication, http://publicnrs.inist.fr/

8. JISC Disciplinary Differences Rapport (2005). http://www.jisc.ac.uk/uploaded_documents/Disciplinary Differences and Needs.doc

9. Libre accès à l'information scientifique et technique, www.inist.fr/openaccess

23<http://www.ucl.ac.uk/ciber/ciber_2005_survey_final.pdf >