



Archives ouvertes – Vers une obligation de dépôt ? Synthèse sur les réalisations existantes, les pratiques des chercheurs et le rôle des institutions

Dominique L'Hostis, Pascal Aventurier

► To cite this version:

Dominique L'Hostis, Pascal Aventurier. Archives ouvertes – Vers une obligation de dépôt ? Synthèse sur les réalisations existantes, les pratiques des chercheurs et le rôle des institutions. INRA; Direction de l'Innovation et des Systèmes d'Informations - Information scientifique et Technique. 2006, 44 p. sic_00115513v3

HAL Id: sic_00115513

https://archivesic.ccsd.cnrs.fr/sic_00115513v3

Submitted on 26 Mar 2007

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.



Open archives Towards a policy of mandatory deposit?

A summary report on current developments, researcher practices and the role of institutions

Update and translation of the original report:

http://archivesic.ccsd.cnrs.fr/sic_00115513

or

http://www.inra.fr/prodinra/pinra/data/2006/12/PROD20062cdda66f_20061206082805610.pdf

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

- (1) INRA Nantes – Research Support Services
Regional Team for Scientific and Technical Information
e-mail: Dominique.L'Hostis@nantes.inra.fr
- (2) INRA Avignon – Research Support Services
Regional Team for Scientific and Technical Information
e-mail: Pascal.Aventurier@avignon.inra.fr

With the collaboration of:

Hélène Bosc

Special Advisor to INRA
Convenor of the Euroscience Scientific Publishing Working Group
http://www.euroscience.org/WGROUPS/SC_PUBLISHING/index.htm

Translation revised by :

Annik Lacombe

(INRA/DISI/IST)

Warning

This study was carried out in the context of the institutional archive project "ProdINRA" and submitted to senior INRA management for approval.

How to navigate within this document

In order to facilitate reading and enable a rapid overview of this relatively lengthy text, a summary version is offered as part of the introduction. Cross references allow the reader to move between the short version and the complete version of the document.

- In the paper version, use the cross references for the pages and table of contents,
- In the electronic version: from the summary version, click on the headings to consult the full text, and vice versa, click on "rapid read" links to return from the full text to the short version.

Utilisation & dissemination



Attribution – Non-commercial – No Derivative Works 2.0
France <http://creativecommons.org/licenses/by-sa/2.0/fr/deed.fr>

You are free to copy, distribute and communicate this work to the public under the following conditions:



Attribution. You must give the original author credit.



Non-Commercial. You may not use this work for commercial purposes.



No Derivative Works. You may not alter, transform or build upon this work.

Table of Contents

Summary of the document (rapid overview).....	5
1- Introduction.....	9
2- Open Access: a recent but now widespread practice	10
3- A reminder of important recent stages in the movement	10
3.1- Self-archiving – a definition of open access and the characteristics of archives	11
3.2- Interoperability and visibility of archives	14
4- Advantages of self-archiving to the various players	17
4.1- For authors	17
4.2- For research institutions	19
4.3- For potential readers: citizens – society	19
4.4- For research funding organisations.....	20
4.5- Enrichment of document repositories: a potential added value for Open Access.	20
5- Policies which are insufficiently followed in practice.....	21
5.1- Reluctance or "wait and see" by authors: good or bad arguments ?.....	21
5.2- The positioning of French research institutions: a key to success.....	23
6- Policies encouraging or mandating deposit	27
6.1- Deliberate policies within institutions.....	27
6.2- Strong encouragement by funding agencies and political institutions.....	29
6.3- Links between deposit and assessment.....	31
6.3.1- The example of the United Kingdom: the IRRA (Institutional Repositories and Research Assessment)	31
6.4- Elements to be taken into account when linking assessment and document deposits	32
6.5- The institutional archive: the best model?.....	33
6.6- Conclusion: towards mandatory deposit.....	33
7- General conclusion.....	35

Glossary.....	37
Bibliography	37
Annex 1 – New publishing models	39
Annex 2 – Addresses of electronic resources	42

Summary of the document (rapid overview)

N.B.: Click on the headings to consult the full version of the document

Chapter 1: Introduction (page 9)

& Chapter 2: Open Access: a recent but now widespread practice (page 10)

The aim of the movement in favour of the open access to scientific documents is free and immediate access to scientific publications via two different pathways:

- the creation of alternative journals instead of the standard "subscriber pays" model, a new "author/institution pays" publishing model (where the author or his institution pays a fee for publication but dissemination is then free of charge and open access).
(see [Annex 1: "New publishing models"](#), p.39)

- the self-archiving of documents, *particularly published journal articles*: the authors deposit their work in "institutional repositories", managed locally by the institution to which they are attached and/or in discipline-driven archives accessible free of charge via the internet.

Chapter 3 - A reminder of important recent stages in the movement (page 10)

This movement was initiated between 1991 and 1997 by communities of scientists (first of all in theoretical physics, then in mathematics and the cognitive sciences) who developed servers to deposit their papers prior to publication (preprints) and then after publication in a scientific journal (postprints). The challenge was the free circulation of scientific knowledge and a resistance to the policies adopted by scientific publishers (exorbitant subscription prices, total abandonment of copyright to publishers, excessive delays in the dissemination of knowledge between submission of the article and its publication; denial of access for those who cannot afford to pay).

The movement then became more widespread and was taken up by organisations funding research, and gave rise to several declarations by scientists and research agencies in different countries meeting in Budapest (2002) and then in Berlin (2003, 2004 and 2005). Researchers can now deposit their documents in institutional repositories sharing a common technological base (the OAI-PMH protocol) enabling the collection by harvesting systems (OAIster) of a description of bibliographical information (metadata).

They can also submit their papers to open access journals: this may entail some cost (the author or institution paying a fee for publication) but access to the document is free of charge, hence the "author/institution pays" model.

In the first instance, scientific publishers were hostile to this trend in favour of open access, but now more than 90% of them (see Sherpa/RoMEO) officially endorse authors depositing their papers in their institutional server. They are also gradually adapting their business models by proposing either an open access option (which may require a publication fee) and also, very recently, by providing authors with the possibility of disseminating (on the publisher's website) the manuscript version (= preprint) of papers submitted for peer review, and the possibility of advance online review of the paper by the scientific community (e.g. "*Atmospheric Chemistry and Physics*" and "*Nature*").

Chapter 4 – Advantages of self-archiving for different players (page 17)

For researchers

- Greater visibility for publications,
- Increased accessibility, and thus a much higher citation rate according to all the studies mentioned (a gain of between 25% and 250%),
- Immediate dissemination, and a greater immediacy factor (papers are cited more rapidly after their publication),
- Increase in the impact factor for most journals providing open access,
- permanence of data,
- the emergence of new impact measurements applicable to papers themselves (independently of journals) should also be noted.

For institutions

- An institutional data repository constitutes a shop window for a scientific organisation (or a shop window for all French research when all open archives of French research agencies will be visible in the future French national archive),
- It enables the monitoring of in-house scientific yield (potential production of indicators) and could be used for scientific assessment,
- the advantages listed above for researchers also apply to institutions,
- Finally, Open Access constitutes an essential tool for the universal and direct dissemination of knowledge towards teachers and the general public, as well as towards developing countries,
- Many potential applications for open access are still to be developed. Thus researchers could deposit other types of documents: experimental findings, texts of documents, etc. and thus create added value for institutional sites.

Chapter 5 – Policies which are insufficiently followed in practice (page 21)

Despite the positioning of numerous institutions in favour of self-archiving at both the national and international levels, the rates of deposit observed remain relatively low: 465 institutional archives are indexed by ROAR, but on average they contain only 15% of the research articles generated annually by their institutions, with major differences between disciplines.

There are several reasons for this:

- Researchers are not yet sufficiently aware of the possibility and benefits of Open Access,
- They find it difficult to emancipate themselves from traditional methods because of concerns about the "author/institution pays" model and worries about peer-review,
- Deposit in institutional archives is too frequently perceived as an administrative constraint and a waste of time,
- Researchers are uninformed about the legal context of deposit.

To be fully operational, deposit rates in institutional archives must be as close as possible to 100% of annual research output. Different surveys have shown that researchers are willing to self-archive, but only if their institution requires them to do so.

The policy of research institutions is thus a critical factor in the success of open archives: A policy of mandatory deposit must be implemented by institutions, backed by a support system for researchers (increasing awareness of the stakes, legal context, usage, citations, etc.).

In the short term, all French research institutions have or will have an open institutional archive, although they will adopt a wide variety of policies.

Some institutions have based themselves on existing databases: this is the case of CEMAGREF, which benefits from an exhaustive inventory of its publications (without full texts) and will use these assets as a foundation to launch its open archive.

Other institutions such as INSERM or INRIA have chosen to use HAL, the tool developed by the CNRS and used for the national platform.

The solution to the problem of multi-institution deposits (in cases where a document needs to be present in several institutional archives because of the affiliations of its authors) is one primary local deposit, plus linking or harvesting of the metadata, or both the metadata and the document, to the other archives.

Chapter 6 – Policies encouraging or mandating deposit (page28)

Both nationally and internationally, mandatory deposit is encouraged at two levels:

- Within some research institutions which require it from their scientists.

The ROARMAP (Registry of Open Access Repository Material Archiving Policies) site lists 56 registered institutional and funder self-archiving policies from Europe, America and the Far East, 20 of them mandates, and five proposed mandates. This is promising but numbers need to be accelerated.

Among the six "pilot" institutions listed, reference can be made to:

- The Queensland University of Technology (Australia) which has achieved 100% of self-archiving of its scientific data,
- The University of Minho (Portugal) which provides financial incentives to laboratories which deposit their publications in the institutional open archive,
- The Electronics & Computer Science Department at Southampton University in the UK which since 2001 has implemented a policy of self-archiving and achieved a 50% deposit rate; they wish to render deposit mandatory.

- By agencies funding public-sector research

Six of the eight Research Councils in the United Kingdom have reaffirmed Open Access policies, five of them self-archiving mandates.

The Wellcome Trust (leading funding agency for biomedical research) requires that any work published with its financial support must be deposited on the PubMed Central site (a NIH project which allows open access to medical research publications).

The invitational policy of the NIH in favour of self-archiving has failed. Only 4% of articles are self-archived.

In January 2006, the European Union published a report comprising ten major recommendations, including open access to work funded by the EU. The Euroscience Workgroup on Science Publishing proposes the mandatory self-archiving of all publications but only deposits from journals that have endorsed Open Access self-archiving (70%) can have access set as Open Access Closed Access immediately upon deposit; the rest of the deposits are Closed Access, in compliance with the embargo periods imposed by publishers (recommended limit: 6 months).

At the French level, a national cooperation framework is gradually being developed (a draft agreement between the EPST (public sector scientific and technology establishments) and the Conference of University Presidents was signed in July, 2006).

Declarations by agencies commissioning research to encourage self-archiving have a limited effect as they do not really have the means of imposing their requirements on research institutions. Research funding agencies must also cover the cost of open access in scientific journals (full or optional open access) so as to support the scientific publishing system.

Documents deposited in institutional archives could be used for collective or individual scientific assessments. In this context, the mandatory self-archiving of documents, without requiring open

access to them, could reach 100% of scientific production and would thus facilitate assessment procedures.

In the UK, the IRRA (Institutional Repositories and Research Assessment) project has rendered this link between institutional archives and assessment a practical proposition. The system envisaged facilitates the work of scientists who deposit their findings with assessment bodies, as well as that of assessment commissions. It will be used for scientific assessment with effect from 2008.

General conclusion *(page 35)*

Institutional repositories enable the storage (archiving) of all scientific work in a digital form for the purposes of access and preservation. Numerous studies have demonstrated that above all they provide **added value in the form of increase in visibility, usage and impact** for scientific work. Deposit rates, however, will remain very low (at around 15%), until institutions adopt mandatory deposit policies.

Mandatory deposit has shown itself to be the only way to ensure the deposit of 100% of annual research output in an open institutional archive, an essential condition for deriving all the benefits anticipated from Open Access. The mandate should be linked to assessment: this will guarantee a high deposit rate and facilitate the work of assessment commissions.

Linking mandatory deposit with the assessment process has two advantages:

- an increase in the number of documents deposited in open institutional archives, thus guaranteeing the regular updating of data,
- a considerable reduction in the work necessary to prepare assessments, for both assessors and researchers.

Annexes

Annex 1 – New publishing models *(page 39)*

Annex 2 – Addresses of electronic resources *(page 42)*

1- Introduction

[Rapid overview]

At a very early stage, development of the Internet (as early as 1990, with the creation of the web) enabled the direct dissemination of scientific information. Over the past ten years, it has revolutionised access to the scientific literature.

Although publishers have adapted their subscription systems to electronic content, new means of supplying open access to publications were reaching maturity:

- Open archives: websites where the author directly deposits publications after peer review (postprints) or before peer review (preprints). However, dissemination is subject to moderation by those managing the open archive,
- Open Access journals: publishers disseminate peer-reviewed journals free of charge via the internet. On the other hand, it is generally necessary to pay to submit a paper.

At the same time, French research agencies (CNRS, INSERM, etc.) moved in the same direction by signing the Berlin declaration on Open Access (October 2003):

http://www.inist.fr/openaccess/article.php3?id_article=38.

In 2005 and 2006, the number of institutional archives increased considerably (deposit of scientific work by researchers affiliated to an institution) as did that of thematic archives (PubMed Central, the NIH project), and more and more research management bodies and parliamentary institutions, etc., were adopting policies in favour of open access.

However, it was clear that despite this encouragement in favour of open access, the results did not achieve the levels anticipated: the numbers of documents deposited for open access often represented less than 15% of all publications by the institutions (Hajjem et al. 2005b).

The stakes are considerable, both for institutions (management of their scientific production, generation of indicators, visibility and scientific impact of research) and for scientists (increased consultation of publications causing a rise in citation rates and removing publishing delays, unlike the system where a journal sometimes disseminates an accepted paper several months after acceptance has been confirmed, etc.).

What are the obstacles? The positioning of commercial publishers or learned societies? The inertia of researchers and their institutions? How can we encourage the deposit of scientific information under open access? Should we not be moving towards the mandatory deposit of documents?

Based on recent or ongoing studies, this report tries to clarify the situation regarding French and foreign initiatives to encourage or render mandatory the deposit of documents in open institutional archives.

This document only deals with the self-archiving method for publications in institutional archives.

As a supplement, Annex 1 "[New publishing models](#)" presents current trends in publishing models: the "classic" journal proposing optional open access, and the entirely open access journal which functions on the basis of "author payer".

2- Open Access: a recent but now widespread practice

[[Rapid overview](#)]

Since the 1990s, an international campaign led initially by a few "activist" scientists such as Paul Ginsparg (mainly in the fields of physics and mathematics) and Stevan Harnad (all disciplines) has been pleading in favour of open access to scientific and technical information (Harnad 2003).

A first method was the creation of alternative journals instead of the standard "subscription" model, a new "author/institution pays" publishing model, where the author or his/her institution pays a fee for publication, but dissemination is then free of charge and under open access.

A second method is to encourage authors to deposit their work in "institutional repositories" which are accessible free of charge via the internet and managed locally by their institution, and/or in discipline-driven archives.

Although their implementation has raised numerous questions from the economic, legal, technical and organisational points of view, recent years have seen the upsurge of Open Access solutions (via different stages and based on important foundation texts) and increasing interest has been expressed by universities, the world of research and funding agencies in this global movement.

3- A reminder of important recent stages in the movement

[[Rapid overview](#)]

See also the INIST website "Le libre accès à l'information scientifique et technique", at the page devoted to reference texts: http://www.inist.fr/openaccess/rubrique.php3?id_rubrique=11

- 2002: Budapest Open Access Initiative (BOAI): signatories undertook to encourage Open Access using one of the two complementary recommended approaches:
 - BOAI-1: self-archiving of articles (preprints and postprints),
 - BOAI-2: publication in open access journals.
- 2003:
 - Bethesda Declaration concerning open access publishing: a precise definition of open access publishing (BOAI-2), but individual rather than institutional commitments (the participants mainly being Anglo-Saxon).
 - French version:
http://www.inist.fr/openaccess/imprimersans.php3?id_article=58
 - First Berlin Declaration concerning open access to knowledge in the exact sciences, life sciences, human and social sciences
 - signatories: CNRS, INSERM (October 2003), INRA, INRIA (July 2004),
 - *"The authors grant to all users a free, irrevocable, worldwide right of access to, and a license to copy, use, distribute, transmit and display the work publicly.... subject to proper attribution of authorship,*
 - *The work shall be deposited (and thus published) in at least one online public repository... that offers guaranteed open access, unrestricted distribution, inter-operability (OAI-PMH protocol) and long-term archiving".*
- 2004: Berlin 2, or how to render open access to scientific literature a practical proposition,
- 2005: Berlin 3 or how to accelerate practical implementation: notably by recommending to institutions to encourage their scientific community to self-archive their publications.

- 2006:
Note should be made in France of the signature on July 6, 2006 of a framework protocol defining a coordinated, national approach to the open archiving of French scientific work. In addition to the CNRS, the leading French research agencies (INSERM, INRIA, INERIS, IRD, ADEME etc.) and the Conference of University Presidents (CPU), the Conference of Higher Education Specialised Schools ("grandes écoles") and Institut Pasteur also signed this agreement (text of the protocol on: <http://www.cge.asso.fr/presse/Protocole-d-accord-archivage-ouvert.pdf>).

3.1- Self-archiving – a definition of open access and the characteristics of archives

♦ **Definition of "Open access"** taken from the Budapest Initiative (BOAI) dated February 14, 2002
<http://www.soros.org/openaccess/fr/read.shtml>

"The literature that should be freely accessible online is that which scholars give to the world without expectation of payment" [...]

"By "open access" to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search or link to the full texts of these articles....."

This definition covers both documents (articles) published in **open access journals** and **documents deposited in an open archive**.

For those who wish to learn more about developments in this movement and the definitions of open access over the years: see the report by Peter Suber "Open Access Overview".

<http://www.earlham.edu/~peters/fos/overview.htm>

The remainder of this document only addresses the question of **self-archiving**.

Two types of open archive have developed in parallel over time: discipline-driven and institutional. The "Registry of Open Access Repositories" (ROAR) <http://archives.eprints.org/> site indexes the different open archives existing throughout the world and provides statistics concerning their input.

This site also proposes a filter to identify existing archives on the basis of different criteria (by type, by country, by software system, etc.).

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

Because registration is not mandatory, this list is not necessarily exhaustive, but it nonetheless indexes more than 700 archives or journal sites registered in OAISTER (<http://www.oaister.org/o/oaister>) and around 460 if account is only taken of archives declared as "institutional or multi-institutional archives" (Registry of Open Access Repositories (ROAR) : <http://archives.eprints.org/> - figures for November 2006).

♦ A few examples of archives

International thematic archive

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

Set up in 1991 at the initiative of Paul Ginsparg, this archive contains both preprints and postprints in the fields of physics, mathematics, informatics and quantitative biology, with more than 350,000 articles.

French inter-institutional archive

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

A targeted, international, multidisciplinary "full text" archive generated by the CNRS since 2001, it has now become the federative platform for other institutional archives: INRIA, INSERM, and soon the Institut Pasteur, etc.

♦ **The characteristics of archives:**

- **Deposit by authors**

A deposit system centred around the author is strongly recommended. Thus the scientist is responsible for the quality of the document and remains in control of how his or her results are disseminated.

Deposits only concern documents which have already been validated by the reading committees of journals or scientific meetings (postprints) or those which could be submitted to a reading committee or for discussion during scientific meetings (preprints).

- **Basic moderation**

A basic, non-scientific check is generally ensured: alignment between content and description, verification and/or supplementation of **metadata** (see paragraph 3.2- Interoperability and visibility of archives). Scientific moderation is applied in some archives (verifiers working by discipline at the CNRS).

- **Permanence of deposits**

Once accepted on the server, withdrawal of a document is strongly discouraged (unless in exceptional circumstances), but new versions can be proposed and the metadata accompanying the document can be corrected or supplemented.

Discouraging withdrawals means that authors must take more responsibility for deposits and scientific value, and guarantees the permanence of access to the document.

- **Institutional affiliation of the depositor**

At deposit, the depositor (deemed by default to be the author) must indicate his or her affiliation to a laboratory or institution.

In some non-institutional archives (e.g. ArXiv), a sponsoring system between authors (and even between different institutions) has been implemented.

- **Compliance with author and publisher rights**

The role of journals is to publish, but also and above all to select, control, correct and validate work, notably through peer review procedures. On condition that the respective rights of authors and publishers are respected, any scientist can entrust to the open archive of his or her institution the version accepted by the reading committee of a journal (final, validated version by the author = postprint).

Open archives and traditional journals can thus cohabit and supplement each other (see [Annex 2](#), p. 39).

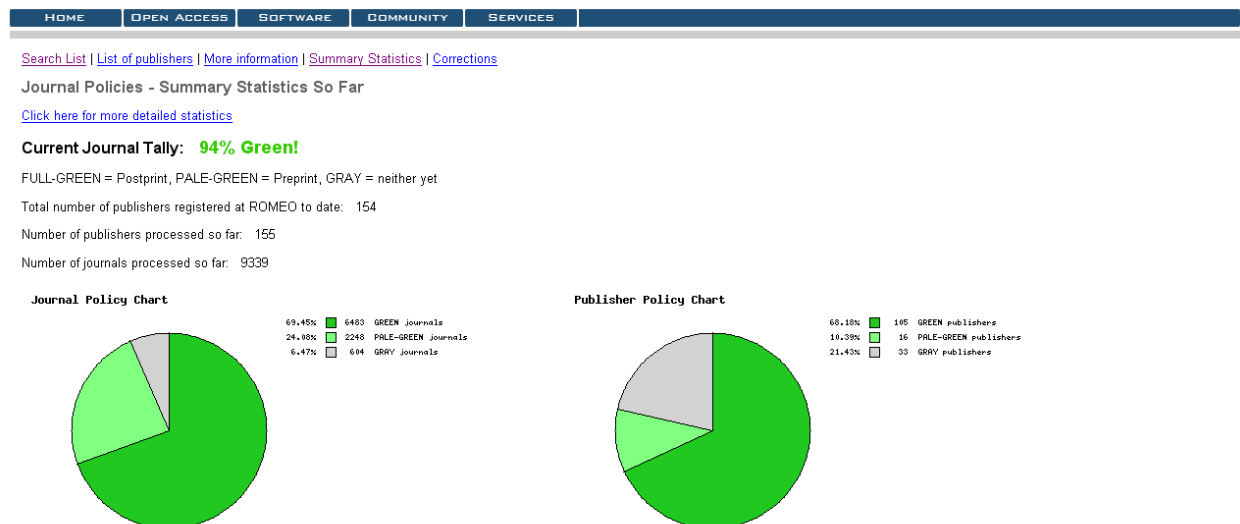
The <http://romeo.eprints.org> site (recommended for its clarity), and the Sherpa/Romeo site <http://www.sherpa.ac.uk/romeo.php> have become reference tools to monitor the policies adopted by scientific publishers with respect to self-archiving.

Extract dated 31/07/06: Statistics concerning the 155 publishers registered to date on the site

RoMEO Colour Code	Archiving policy	Number of journals	%	Examples of publishers
green	Allows the archiving of postprints (i.e. last author version after peer review)	6483	69%	<ul style="list-style-type: none"> Elsevier (1882 journals) CAB International (17 journals) Cambridge University Press (186 journals)
pale green	Allows the archiving of preprints (i.e. author version prior to peer review).	2248	24%	<ul style="list-style-type: none"> Blackwell (698 journals) Marcel Dekker (83 journals) Nature Publishing Group (47 journals) Oxford University Press (180 journals)
grey	Archiving not allowed	604	7%	<ul style="list-style-type: none"> American Chemical Society (35 journals) American Physiological Society (16 journals) Lippincott (287 journals)

In total: approximately **94%** of the journals in this list allow some form of self-archiving by authors.

Journal Policies - Summary Statistics So Far



Initially very reluctant, publishers have now rendered their policies much more flexible. The decision taken by Elsevier in 2004 to join the group of publishers allowing the archiving of preprints and postprints was a notable turning point.

- **Ethical and deontological rules**

The same rules as those in the classic publishing circuit must be applied to documents deposited in an archive. In a recent speech, Franck Laloë (Laloë 2006) recalled these "good practices":

- agreement of all authors and co-authors,
- supply of all scientific data,
- citation of third parties,
- no plagiarism,
- loyalty towards employers and entities providing funds for the research described (no disclosure of confidential results such as patents, and mention of financial support received).

- **Technical guarantees/Server**

Open archives must ensure the permanence of information, at the level of both access to electronic resources (archiving) and reading of the formats of the documents deposited (in most cases, in a PDF or MSWORD format), which implies that the software necessary to read these formats in the correct version must also be conserved.

3.2- Interoperability and visibility of archives

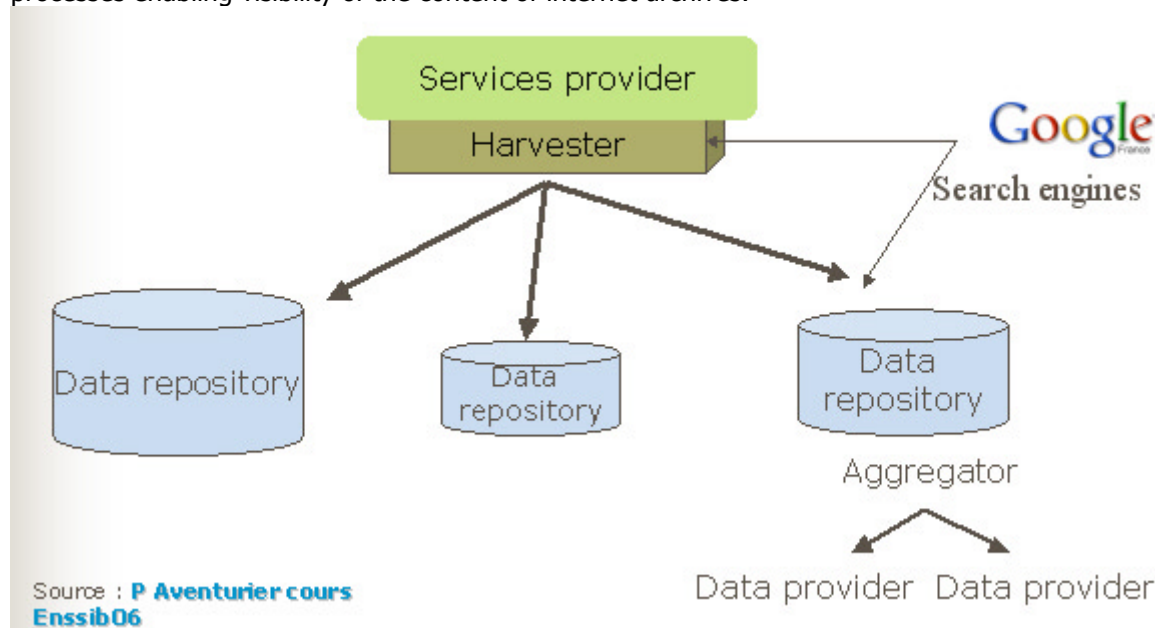
Data repositories can be built up using freely available software programs (the most widespread being EPrints and DSpace) or specifically-developed tools. Their common characteristics are an availability module using the OAI-PMH protocol (Open Archives Initiative Protocol for Metadata Harvesting) and metadata describing the electronic resources (author, title, etc.).

Metadata can be defined as "*a series of structured data describing physical or digital resources*" ... *an essential link in the sharing of information and the interoperability of electronic resources. The OAI-PMH protocol ensures that metadata in open archives can be harvested by specialised search engines, called "harvesters". The metadata set managed by this protocol is the Dublin Core*" (Morel-Pair 2005).

Describing electronic resources using common metadata for each French institution is a major challenge to guarantee interoperability and provide a reliable, high-performing common system.

Metadata in open archives are "harvested" by specialised search engines called "harvesters" which can simultaneously interrogate several archives as a single depository of data.

(N.B. unlike some engines, OAI harvesters only deal with metadata). The figure below illustrates the processes enabling visibility of the content of internet archives.



- Harvesters/aggregators:
The best known is OAster (<http://www.oaister.org/o/oaister>) which references 660 data repositories containing more than 8 million documents. The search interface is however limited to information on the title, author and institutions (creator), year, language and type of document.
- Commercial initiatives can also provide indirect access to the content of open archives:
 - SCIRUS: <http://www.scirus.com/srapp/>
A specialised search engine developed by Elsevier Science, containing not only resources published by the Group but also documents from databases, preprint servers and Open Archives.
 - SCOPUS: <http://www.scopus.com/>
Also developed by Elsevier, this is a targeted science and technology database which, in addition to references arising from the databases produced by this publisher also includes PubMed references, references to articles published by Elsevier and websites detected by the SCIRUS search engine. Only validated (peer reviewed) resources are included.
 - GOOGLE SCHOLAR: <http://scholar.google.com/>
A search engine specialised in scholarly literature. Its true coverage is not known, but it indexes articles from peer-reviewed periodicals, doctoral theses, books, preprints, reports from scientific publishers, learned societies, preprint repositories and university servers.
- Associated technologies:
 - OPENURL
"An architecture which serves to create contextual links. In practice, the aim is to link metadata (e.g. the bibliographical references of an article) to the resource itself (the full text of the article).
For a resource (an article in this case), we know the context in which it is cited (the bibliography which refers to it), described (the metadata in the article and who compiled them), used (the reader who is looking for it and the rights available to him) and how the protocol is used (which problem-solver, to obtain which service) to link it to what it describes (the article itself) (definition taken from the site: <http://www.figoblog.org/document207.php>).

These tools also enable:

- Multiform searches: on metadata in articles and/or full text searchers;
- Permanent access to a document: The URI (Uniform Resource Identifier) aims to attribute to each resource (and version or format of the resource) a single, permanent identity (Morel-Pair 2005),
- Using some tools, display of the number of citations and number of downloads,
- A link between references cited and the full text,
- Access to previous documents (before journals were put on line) in some cases of retrodigitalization.

4- Advantages of self-archiving to the various players

[[Rapid overview](#)]

4.1- For authors

The principal aim of scientists is to communicate their results in order to be "read by others and cited by others".

"Authors are never paid for the articles they produce. Their remuneration is acknowledgement of their work by their peers and employers (to obtain funding, career advancement, etc.), hence the importance of the broadest possible dissemination of these publications which is decisive regarding the impact of their research" (Schmitt 2006).

For the purposes of their scientific activities, they also need to monitor and collect, real-time, all scientific work in their field, at an international level.

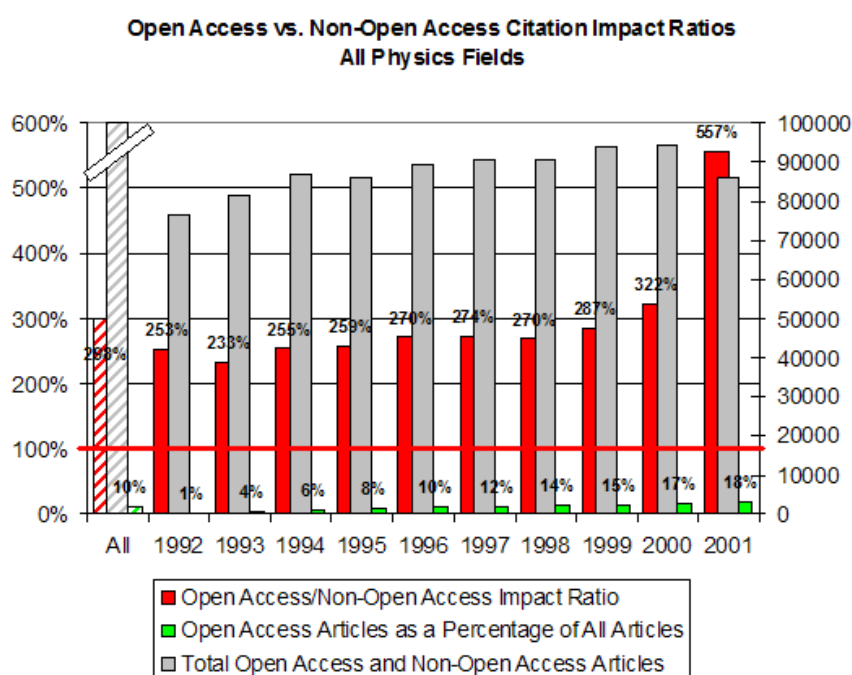
Self-archiving appears to constitute the **optimum means of disseminating research results**, the effects of which in terms of **impact** can now be measured.

The <http://opcit.eprints.org/oacitation-biblio.html> site provides a bibliographical list of the different studies performed to measure the effect of self-archiving on the impact of articles.

♦ Measurement of impact – based on **citation rates**

- Since the article by Lawrence (Lawrence 2001) on the correlation between the citation rate of an article and its availability under Open Access, numerous studies have demonstrated a substantial gain in citations when documents are deposited in an open archive, although there are notable differences depending on the disciplines considered.

The figure below presents a recent study (Harnad and Brody 2004) which compared the impact (citation) of self-archived articles in Open Access (OA), with those in the same journal which were not thus archived. <http://www.dlib.org/dlib/june04/harnad/06harnad.html>.

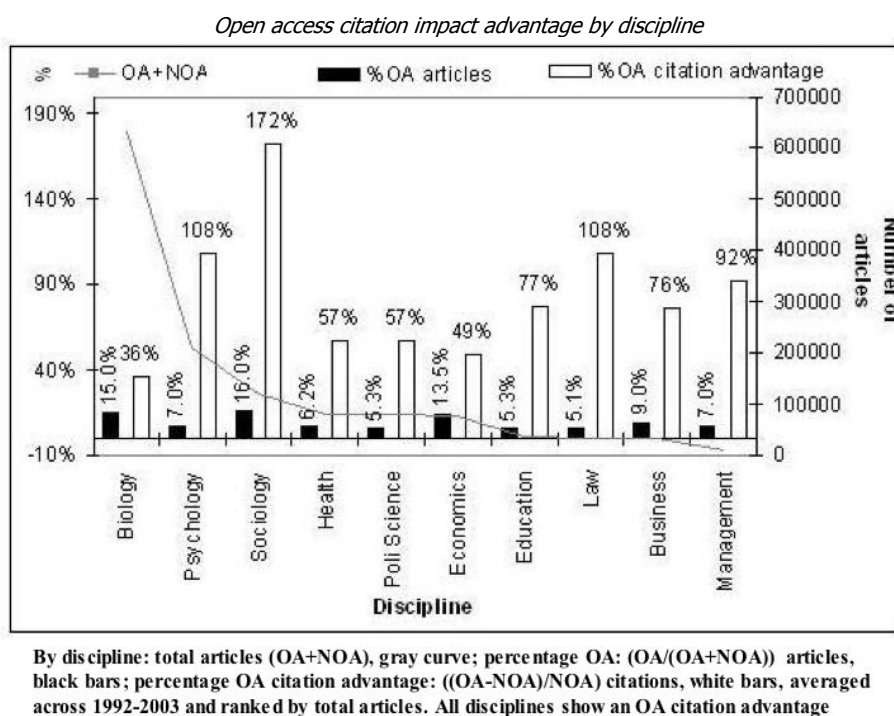


- A recent study in 2005 (Hajjem et al. 2005a), performed on a sample of approximately 1000 journals (indexed by the ISI) representing four disciplines (Biology, Business, Psychology and Sociology) demonstrated notable advantages in terms of citation rates if documents were deposited in parallel in an open archive. **This gain could reach between 25% and 250%, depending on the discipline.**

(<http://eprints.ecs.soton.ac.uk/11687/>)

- Another study by the same authors (Hajjem et al. 2005b), reported by T. Chanier (Chanier 2006) demonstrated converging results in six other disciplines.

<http://sites.computer.org/debull/A05dec/hajjem.pdf>



♦ Measurement of impact – other types of evaluation

Studies have demonstrated a link between consultation rates (downloading) and citation rates. In some fields, a link can also be seen between the rapidity of consultation and the citation rate (Brody et al. 2006; Perneger 2004).

In addition, new models are emerging to evaluate the impact of documents themselves (and not just the journals). Adapted to open access, they allow the automatic calculation of citations in the open access environment.

- Citebase <http://www.citebase.org>

A graphical analysis demonstrating the number of downloads, the latency period between publication and the first downloads and citations of its own citations.

- CiteSeer : <http://citeseer.ist.psu.edu>

Can search for documents and citations in scientific and technical fields. Its database contains more than 700,000 documents.

Google Scholar

A web search engine that indexes the full-text of scholarly literature (precited in page 16). Through its "Cited by" feature, Google Scholar provides access to a list of articles and documents that have cited the document originally retrieved in the search.

- The ISI has recently started developing a new database, the Web Citation Index, which takes account of the content of some archive repositories. Integrated in the Web of Knowledge, this is a multidisciplinary index of the citations of publications accessible via the internet in institutional archive repositories or discipline-based archives. Still in the test phase, it includes journal articles, preprints, doctoral theses, reports, technical reports and other documents in the grey literature. For each document indexed, the interface allows access to the references cited as well as to the number of citations received and the references of the citing articles.

Open archives would thus enable:

- Greater **visibility** of publications,
- **Increased accessibility**,
- **Immediate dissemination**, a greater immediacy factor,
- **An increase in the impact factor** for some journal titles,
- **New measurements of impact applicable to the articles themselves**, independently of journals,
- **Permanence of information.**

4.2- For research institutions

Open archives constitute a potentially unique shop window for the scientific community, for partners in research and for society in general.

They also constitute a precious tool for the in-house monitoring of scientific production:

- continual updating,
- permanent collection and archiving of documents,
- potential production of indicators,
- a possible tool for assessment (documents being available in electronic form in a single system).

Because open archives are constantly evolving, we feel it is important to mention (as did Sale in 2006) (Sale 2006) that in the current context, there is a potential risk for institutions if open archives are not developed, in terms of international visibility and the dissemination of scientific scholarship.

4.3- For potential readers: citizens – society

Open archives may constitute an essential tool for the dissemination of knowledge to teachers and the general public, as well as towards developing countries.

- wider variety of documents, larger databases (harvesting),
- free of charge,
- simplicity of access.

Researchers do not appear to be indifferent to this major principle of disseminating knowledge, even if this idea alone is insufficient to gain their approval. 90% of researchers who publish under open

access do so because they adhere to the principle of "free access" and the right to information (Swan 2005).

4.4- For research funding organisations

Organisations funding research are also influential actors in Open Access (see paragraph 6.2- , p. 29)

The vocation of the authors of scientific papers is not to commercialise their intellectual production to live. The State, which remunerates most of them, as well as research funding organisations, nonetheless have the right to hope for a true return on their investments through the dissemination of research results throughout the scientific world, amongst students and even citizens in general (http://www.edsh.cnrs.fr/spip/article.php3?id_article=976).

In a recent study published in July, 2006, John Houghton and Peter Sheehan (Houghton and Sheehan 2006) questioned the true impact of new models for the dissemination of scholarly communication, in this case open access. They sought to determine whether an improvement in the dissemination of research results, and thus the funds devoted to this research, could also increase returns on investment in R&D. Using an adaptation of a model for economic growth, they calculated series of impact indicators for overall global expenditure and public R&D expenditure for all OECD countries, and thus obtained a practical evaluation of the gains which could be expected from broader access to research results.

As an example, we show the estimates for France and Germany:

Global expenditure						Public expenditure					
France						France					
39,740						6,640					
Per cent change in accessibility and efficiency						Per cent change in accessibility and efficiency					
Return to R&D						Return to R&D					
	25%	40%	50%	60%	75%		25%	40%	50%	60%	75%
Recurring annual gain from move to open access (USDm PPPs)						Recurring annual gain from move to open access (USDm PPPs)					
1%	200	320	399	479	599	1%	33	53	67	80	100
2%	401	642	803	963	1,204	2%	67	107	134	161	201
5%	1,018	1,629	2,037	2,444	3,055	5%	170	272	340	408	510
10%	2,086	3,338	4,173	5,007	6,259	10%	349	558	697	837	1,046
Germany						Germany					
58,688						7,775					
Per cent change in accessibility and efficiency						Per cent change in accessibility and efficiency					
Return to R&D						Return to R&D					
	25%	40%	50%	60%	75%		25%	40%	50%	60%	75%
Recurring annual gain from move to open access (USDm PPPs)						Recurring annual gain from move to open access (USDm PPPs)					
1%	295	472	590	708	885	1%	39	63	78	94	117
2%	593	948	1,185	1,423	1,778	2%	79	126	157	188	236
5%	1,504	2,406	3,008	3,609	4,512	5%	199	319	398	478	598
10%	3,081	4,930	6,162	7,395	9,243	10%	408	653	816	980	1,225

- Germany: for global expenditure of \$58.7 billion and an R&D return to society of 50%, growth of about 5% in the accessibility of scientific publications would result in a gain of \$3 billion per year.

- France: global expenditure of \$39.8 billion, expected gain/same conditions: \$2 billion per year.

4.5- Enrichment of document repositories: a potential added value for Open Access.

Laurent Romary (Romary 2005) noted that data accessible via open archives should not be restricted to publications only, but could take various forms as a function of scientific disciplines: genomic databases, experimental findings, corpus of documents, databases of images or lexicons, clinical results, etc.

This enrichment represents a potential true added value for archive repositories, and an operation which could be achieved through collaboration between authors and IST staff.

5- Policies which are insufficiently followed in practice

[[Rapid overview](#)]

Despite the positioning of numerous organisations in favour of self-archiving at both the national and international levels, relatively low deposit rates are still observed for most repositories.

According to the Registry of Open Access Repositories (<http://archives.eprints.org>), there are about 460 (see p. 11) institutional or multi-institutional open archives (figures for November, 2006), but on average only 15% of the annual production of these institutions is self-archived, with marked differences between disciplines (Hajjem et al. 2005b).

- "Despite its international reputation, Pubmed Central achieves only a very low rate of deposit, in a context which does not define this as being mandatory (depending on the sources: 4% to 15%)." (Groupement Français de l'Industrie et de l'Information 2006).
- During the Round Table discussion which introduced the last meeting of scientific information specialists in June 2006 in Nancy, the same observation was made: the tools are in place at the CNRS and INSERM but deposit rates remain insufficient (Groupement Français de l'Industrie et de l'Information 2006).

Finally, **different practices are observed in different research disciplines and scientific communities**. Thus the life sciences are very poorly represented, only accounting for about 3% (November 2006) of the CNRS HAL archive, for example.

What are the reasons for these relatively low rates?

5.1- Reluctance or "wait and see" by authors: good or bad arguments ?

- Lack of understanding of Open Access

Deposit in an open archive remains **a system which is poorly understood** by a large number of scientists.

Researchers have not always understood that they can self-archive an article which has been accepted for publication. It is a copy of the published document that they tend to deposit in the archive.

The disciplinary context is not neutral: "In some disciplines, researchers are not numerous, know each other and ignore frontiers. They spontaneously create an exchange network and common pool of publications, such as the well-known arXiv.org. By contrast, researchers in other disciplines, sometimes under contract to private companies, may find themselves in competition and do not adhere to this shared knowledge dynamics. Note should also be made of very uneven support for electronic resources. For example, in the social sciences and humanities, printed publications remain preponderant." (Guha 2005).

Many are still not aware of the stakes involved, and several surveys have illustrated this situation:

- The VST information letter in March 2005 (<http://www.inrp.fr/vst/LettreVST/mars2005.htm>) thus reported the contrasted commitment of authors as revealed by a joint report from the JISC (Joint Information Systems Committee, UK) and the OSI (Open Society Institute) published in 2004, "*Journal authors survey*" (http://www.jisc.ac.uk/uploaded_documents/JISCOAreport1.pdf).

- A study commissioned by the JISC in 2005 provided an update on the behaviour of researchers regarding the self-archiving of their publications.

1 296 scientists responded to this cross-disciplinary survey. More than half of authors had deposited at least one article during the past three years. The reasons they gave to explain their reluctance towards self-archiving were of a technical and legal nature (Swan and Brown 2005).

- Fear of deposit compared with the "peer review" system

Peer reviewing is not called into question by self-archiving (cf. page 12) but it is not easy to modify publishing habits. The field of biomedical research is finding it difficult to extricate itself from traditional methods.

- Perception of a further administrative constraint and fear of time wasting.

In many archives, the researcher can simply register a simplified publication, reduced to a few bibliographical fields, which will then be completed and enriched by qualified scientific information staff responsible for the archive.

In a 2005 survey (Carr and Harnad 2005), the authors showed that self-archiving only required about 10 minutes per publication, and estimated that this would only represent less than 40 minutes of a researcher's time each year!

Different surveys have nonetheless shown that researchers are prepared to self-archive if their institution requires them to do so (Swan 2005; Swan and Brown 2005).

- Distrust of, and even indifference towards, the legal framework for deposit.

Not always sufficiently understood by information professionals, and even less by researchers, this is a crucial point which must not be neglected, as it appears to be a fundamental determining factor for an increase in deposits and a constraint which must be overcome if open archives are not to become simple metadata repositories (bibliographical entries), but involve deposits with access to a full text for a maximum of documents.

Anne-Marie Benoit thus described the legal framework for open archives (Benoit 2005)

"Because they have not been the subject of any transfer of author rights, preprints can be openly and legally self-archived". On the contrary, with respect to postprints: "the work of a researcher which, by contract, has been the subject of a transfer of rights, including exclusive copyright to digital versions, must not be self-archived, unless agreed by the publisher."

The legal conditions concern less the deposit of a full text (usually the final, accepted author version, which differs from the publisher version unless otherwise specified) than its free availability and possible copying.¹

Closed access (i.e. reserved for the author and/or his employers or government authorities) always remains an option, with supply of the full text (if the rights so permit) if requests are addressed by e-mail..

The RoMEO site referred to above (page 12) presents the positions adopted by different scientific publishers: 94% of journals already allow their authors to self-archive.

Deposit of a preprint (before it has been peer-reviewed) is perhaps not advisable in the life sciences (even if this document could benefit from new types of collective validation), as this does not

¹ Deposit of the publisher version for the purposes of storage and not dissemination may in some cases be assimilated with the private copy.

correspond to the practices of scientific communities in this field. Deposits of postprints (as soon as they have been peer reviewed and validated) can be widely recommended. Furthermore, researchers should increasingly be informed on how best to manage their rights, and encouraged, at submission of an article, to include clauses in any contract which allow them to retain some or all of their rights <http://repositories.cdlib.org/iber/econ/policies.html> .

5.2- The positioning of French research institutions: a key to success

The interests of scientists are the same as those of their institutions. In the preface to the book by C. Aubry and J. Janik, Laurent Romary (Romary 2005) then in charge of scientific information at the CNRS, insisted on the essential intervention of institutions and their necessary involvement in order to "render systematic the deposit of publications in an institutional archive, while at the same time not complicating the procedure for those who would in any case have made such a deposit, nor to frighten newcomers by presenting an excessively administrative image of the system".

In a 2005 study by the JISC (Swan 2005), Swan reported that 81% of authors would volunteer to self-archive their work if this was required by their institution or if this was a condition of funding, 13% would be reluctant to do so and 5% stated that they would refuse, even if this was mandatory.....

The meeting of Scientific Information Specialists in Nancy in June 2006 provided an opportunity for an inventory of institutional policies amongst French research agencies and for observations on early experience in this area. The "Dépêche du GFII" provided a very explicit summary.

"The data repositories which are currently being built up will only attain their objectives if the rate of deposits (without reaching 100%) is nonetheless close to this limit. But the experiences reported over the past few days in Nancy confirm the "PubMed model": too much energy is currently being devoted in the field to the "deposit battle" (convincing scientists to deposit their work), with very mixed results.

Put simply, we can say that a "laissez faire" policy only results in deposit levels reaching a few percent. Investment by information specialists in the "deposit battle" (raising awareness, managing certain tasks) makes it possible to reach a deposit rate of around 12%. The combination of requesting deposit with a "carrot and stick" policy (for example, making the deposit rate of researchers one of the elements in their annual assessment) could raise levels above 20%, but not much more. By contrast, some institutions (CEMAGREF, which introduced mandatory deposit as early as 1992, and INERIS (the National Institute for Industrial Environment and Risks) where the deposit rate reaches 100%) have, for obvious reasons, fulfilled their contract, in other words, they have developed open access institutional archives which reflect their research activities in an exhaustive fashion."

♦ CEMAGREF

Mandatory deposit in the CEMAGREF Publications database of a paper copy of all publications has been effective since 1992. In addition to this paper copy, a pdf file is also required (in compliance with the respective rights of authors and publishers and according to the contract signed regarding any dissemination).

The CEMAGREF Publications database contains 14,000 scientific and technical publications produced by the institution (<http://cemadoc.cemagref.fr/basepubli.html>). Its workflow system allows for direct deposit by authors. A first assessment has been made of how this system has been appropriated. The database is scheduled to evolve towards an open access institutional archive and allow the inclusion of CEMAGREF publications in the inter-institutional HAL platform. This technical project will be accompanied by campaigns to raise author awareness to open access. Information scientists must improve their technical and legal training in order to be able to provide effective local support and to validate deposits. This project is being carried out by the INST in liaison with scientific management.

(poster presented at the Rencontres 2006 des Professionnels de l'IST - (Giansily et al. 2006).

CEMAGREF reports a deposit rate close to 100%, and sees relatively little reluctance concerning the system amongst its researchers, thanks to the institution's policy and the development of related services provided to authors.

♦ CIRAD

By means of the mandatory **institutional deposit** of all written scientific work (grey literature and true publications) by CIRAD scientists, the Agritrop database constitutes the living **scientific memory** of the establishment, a true heritage reservoir enriched by several thousand references each year, associated with full texts that can be accessed via the intranet, in compliance with author and publisher rights.

The owned domain and technical limitations of the bibliographic information system (Loris/Doris) adopted in 1998 by the CIRAD nonetheless markedly restrict the external accessibility and dissemination of Agritrop data.

In 2005, the poor visibility of CIRAD publications on the internet thus led this institution to approach the CCSD in order to study the feasibility of the self-archiving of CIRAD publications on an open platform, based on the experience, skills and technical expertise of HAL.

Adoption of the **common HAL platform**, confirmed by signature of a national agreement in 2006 and accompanied by signature of the **Declaration of Berlin** made it possible for CIRAD to put its commitments into practice.

However, this has not solved the problem of the reporting and visibility of grey literature from the CIRAD which accounts for 70% of its written scientific production (mission reports, activity reports, contractual missions or expert reports, unpublished communications at meetings, etc.).

The CIRAD "Open Archive" project launched in 2006, and the introduction of a **specific HAL-CIRAD interface**, aims to allow the deposit of any type of publication produced by the CIRAD, it being accompanied by a legal protection system and the transfer of data to HAL or an enriched institutional database.

♦ CNRS

A pioneer in France with the creation of the CCSD in 2000 (Centre pour la Communication Scientifique Directe, or *Centre for Direct Scientific Communication*) and the dissemination of HAL and TEL archives.

HAL, wholly developed by the CCSD, has been adopted by other French institutions and will be the national federative platform for French open archives.

In a letter sent to all Unit Directors on June 21, 2006, the Director General of the CNRS, Arnold Mingus, reaffirmed the approach adopted by the CNRS towards open archives (deposit in HAL) and open access, but without opting for mandatory deposit.

http://openaccess.inist.fr/breve.php3?id_breve=418

♦ IFREMER

In August 2005, IFREMER launched Archimer, its institutional archive, a database of full texts allowing open access to all publications, doctoral theses, congress proceedings and internal reports. A document presented in detail the implementation of this tool (Merceur 2005). In addition to "classic" objectives (support for Open Access, exploitation of scientific production, creation of a new database on marine sciences, improvements to the visibility of the IFREMER internet site) it also emphasised the usefulness of the system to "strengthening the links between research teams and library staff".

In this archive, documents (reported by authors or detected thanks to a watch system on databases) are entered by scientific information staff: verification of distribution rights if necessary, entry of

metadata, classification by subject fields, addition of key words, reformatting of full texts and conversion into pdf format if necessary, transfer of full texts to the server.

The results announced refer to a 70% deposit rate for full text documents during the first year (<http://www.ifremer.fr/docelec/resultat2006.htm>).

♦ INERIS

In order to improve the national and international visibility of its scientific work and thus enhance the impact of its research, INERIS wishes to ensure the deposit of scientific work by its researchers in the HAL open archive.

The policy of this institution is the mandatory deposit of all publications and presentations.

♦ INRA

INRA was a signatory of the Berlin Convention in May 2004 and actions in favour of open access were then initiated:

- the institution joined Biomed Central (2003-2005) to fund publications ("author pays") in open access journals run by publishers,
- incentives in favour of the "open choice" option for 7 of the 8 INRA journals. The authors only pay 300 euros for an article under open access, even though the true publication cost for an article is around 1300 euros. The number of publications under open access ranges from 14% to 18%, depending on the journal,
- introduction of ProdINRA,
- compliance with the protocol for agreement at a national level regarding the open archiving of scientific work.

ProdINRA, which will be accessible at the end of 2006, will become the open institutional archive for INRA, but will also contain bibliographical references imported from the previous publication database, PubInra. The aims of ProdINRA are:

- the preservation of scientific heritage (the conservation of documents),
- the production of in-house indicators,
- the valorization of publications (open access to full texts).

Information campaigns have been conducted in most INRA centres by scientific information professionals. Furthermore, Mrs Hélène Bosc (INRA Tours) is an acknowledged specialist in Open Archives at the European, if not the world level. She created the first French information site on the subject in 2000

(http://www.tours.INRA.fr/prc/internet/documentation/communication_scientifique/comsci.htm) and is actively involved in internal and external efforts at INRA towards the development of Open Access.

However, this voluntary policy has had little impact as yet on researchers. Obviously, the launch of ProdINRA provides an opportunity to increase awareness and "train" scientists. But according to the experience referred to in paragraph 5- page 21, "the considerable investment by information specialists in the "deposit battle" (raising awareness, managing certain tasks) only makes it possible to reach a deposit rate of around 12%..."

♦ INRIA

The HAL-INRIA open archive <https://hal.inria.fr/> was launched on April 27, 2005.

The trend towards open archives is "culturally" better accepted by those working in informatics (than those in the life sciences, for example).

- Deposits are made without scientific moderation,
- The researcher is responsible for all his or her scientific results, which are rendered visible (successive or final versions: in compliance with any confidentiality clauses which might lead to contracted scientific collaboration),
- Technical and bibliographic checks are made (enrichment of metadata) a posteriori,

- The documents deposited include preprints or published documents (postprint version, i.e. that submitted to the publisher), depending on publisher authorisations, such as: articles from scientific journals/general press; articles presented at congresses/workshops; larger publications – books and congress proceedings; chapters of books; reports; lectures and tutorials; filed patents; software files; doctoral theses, dissertations; HDR (Habilitation à Diriger les Recherches), unpublished documents.
- Authors lie at the centre of the system: the deposit rate appears to be good or very good, depending on the type of document (nearly 100% for congress papers, the proceedings of which have been managed in HAL-INRIA - Jacques Millet – RPIST Nancy 2006 / Opening Round Table: http://rpist.inist.fr/article.php3?id_article=29)

The quality of data pages in the HAL-INRIA system on the institute's website <http://www.inria.fr/publications/archiveouverte/index.fr.html> clearly testifies to the institution's commitment to this approach.

◆ INSERM

Since the end of 2005, INSERM has been depositing its publications in the HAL-INSERM platform: <http://www.hal.inserm.fr>. Metadata, references and the type of moderation have been adapted to the specific needs of INSERM.

The involvement of managers and scientific information staff is very strong, in order to raise the awareness of researchers to the deposit approach. At present, this is only recommended, but mandatory deposit may be envisaged as from 2007 (http://rpist.inist.fr/IMG/pdf/Nicole_Pinhas.pdf).

◆ IRD

For more than 60 years, the IRD has been pursuing policies regarding the archiving and dissemination of scientific publications by its researchers, and since 1995 has been digitizing data in order to render more than 35,000 documents accessible via the internet through the Horizon /Pleins textes system. The IRD has decided to adopt HAL for the deposit of researcher publications and to integrate this new tool in the existing system. (Cavet 2006).

"The deposit rate is an essential parameter for development of the use value of these "open" repositories of sophisticated scientific information. Whether these open archives are assigned objectives concerning the simple communication of public research results, or more ambitious objectives regarding the international visibility of research entities, as an exhaustive reflection of public funded R&D activities or links between this R&D and innovative business (OECD perspective), the repositories which are being built up will only achieve these targets if the deposit rate (although it may not reach 100%) is very close to this limit" (Groupement Français de l'Industrie et de l'Information 2006).

All public-sector scientific and technology establishments (EPST) (and soon the universities) are now committed to the creation of institutional archives. In the longer term, the results of these initiatives will be visible in the future national database, in this case the HAL platform.

At present, it is important to encourage the deposit of documents in archives, even if they are then found in several places. When a publication has been written by several authors from different institutions or belonging to Joint Research Units, the problem of multiple document deposit will then be raised.

- Should this publication be deposited in all the institutional archives corresponding to the authors' affiliations?
- Will authorisations for dissemination have to be requested several times from the authors (once for each archive)?
- Is it planned to remove double entries from the national archive?

These are important questions, but to date they have not been discussed in the literature available.

6- Policies encouraging or mandating deposit

[[Rapid overview](#)]

A convergence is currently developing between:

- The development of prescriptive policies which introduce true, **mandatory deposit** within institutions (emerging trend),
- and
- **Encouragement from the bodies funding** public sector research.

6.1- Deliberate policies within institutions

The ROARMAP website (Registry of Open Access Repository Material Archiving Policies: <http://www.eprints.org/signup/fulllist.php>) collates declarations on archiving policy at an international level. Thus:

- 56 self-archiving policies (institutions or funding agencies) are listed;
- mandatory self-archiving is now adopted by 20 institutions (*) or funding agencies (**) and proposed by 5 more:
 - * Australia - Queensland University of Technology
 - * Europe - CERN
 - * India - Nat Inst Tech Rourkela
 - * Portugal – University of Minho
 - * SWITZERLAND – University of Zurich
 - * UK - University of Southampton ECS
 - ** UK BBSRC
 - ** UK ESRC
 - ** UK MRC
 - ** UK Wellcome Trust.

The section below analyses the results achieved by a few "pioneer" institutions which require the mandatory self-archiving of publications and state their policies on the ROARMAP website (source: <http://www.eprints.org/openaccess/policysignup>).

♦ Electronics & Computer Science Department at Southampton University UK: 2001

- mandatory deposit of all preprint and/or postprint articles. All lists requested for administrative or promotional reasons are extracted from the archive,
- for books and chapters of books: only metadata are mandatory,
- declared efforts to minimize data entry work (cf. study referred to above by Carr and Harnad: evaluating the effort required for deposit at about 10 minutes' work).

Results:

- The number of deposits has increased consistently since the archive opened in 2001. The deposit rate for full texts reached 50% of annual output in 2004 and is continuing to rise. This growth is supported by a policy of mandatory deposit for the entire university, announced for the autumn of 2006, and the policies declared by three national fund-providing agencies (MRC, BBSRC, ESRC) as well as one private funder (Wellcome Trust).
(cf. speech by Alma Swan – ESOF Munich 2006 : <http://www.isn-oldenburg.de/~hilf/vortraege/esof06/esof06-swan.pdf> slides 16-17).

◆ Queensland University of Technology: 25 Feb 2004

- mandatory deposit,
- wide variety of documents: preprint or postprint articles, doctoral theses, grey literature, conference presentations, etc., but not including any documents which involve any notion of confidentiality.

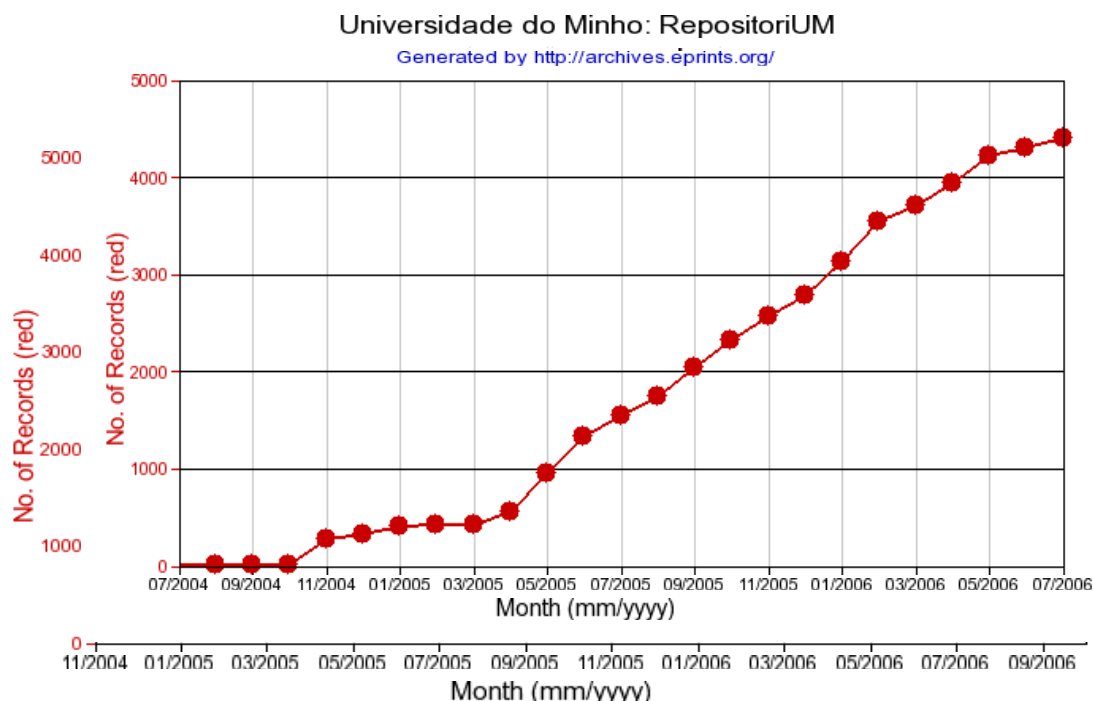
Results: coverage of annual production is close to 100% (full text documents).

◆ University of Minho: 21 Dec 2004

- The first university to require self-archiving since 2004, it differs by having introduced the notion of financial incentives: research centres and schools receive grants from the university which are proportional to their contribution to the archive and calculated as follows: number of documents deposited, date of documents deposited, policy of the laboratory or research department with respect to self-archiving.

Results:

- A deposit rate for full text documents of around 70%
<http://archives.eprints.org/?url=https%3A%2F%2Frepositorium.sdum.uminho.pt>
and an extremely positive analysis in terms of consultations, with impressive chronological growth since the archive was opened: 60% of documents have been consulted at least 5 times. A 25% increase in the consultation rate has been recorded in one year (cf. slide n°24 – (Bosc 2006)). Presented very recently at the ESOF meeting in Munich (July 2006) by Eloy Rodrigues, the most recent results are spectacular (<http://www.isn-oldenburg.de/~hilfe/vortraege/esof06/esof06-rodrigues2.pdf>) and according to the author mainly linked to mandatory deposit, accompanied upstream by a strong awareness and information campaign.



◆ CERN: 25 Feb 2004

Following its signature of the Berlin 3 declaration, the CERN became actively involved in the two roads to open access:

<http://cdsweb.cern.ch/search.py?sc=1&ln=en&p=cern-open-2005-006&f=reportnumber>

- self-archiving on the one hand: mandatory deposit of all scientific documents,
- active support for open access journals: the CERN encourages researchers to integrate the cost of publication by preferring open access, electronic journals, if the quality is equivalent,
- the CERN undertakes to ensure that articles published in classic journals and those deposited in open archives (peer-reviewed) will be considered on an equal footing by assessment commissions,
- For some "popular" journals, it is envisaged that the institution will pay a global fee so that CERN researchers can publish free of charge ("Biomed central" model – "author pays", see Annex 1 - p. 39).

6.2- Strong encouragement by funding agencies and political institutions

"Institutional initiatives are increasing in number as the phenomenon of open archives spreads throughout the world and, at the same time, ministries and governments are becoming aware of the fact that open archives provide them with an opportunity to retain control over scientific work that they have most often funded, either directly or indirectly, and dissemination of which had become the prerogative of private publishers."

(http://fr.wikipedia.org/wiki/Open_Archives_Initiative_-_Les_premiers_fr.C3.A9missements_institut)

According to the OECD report published on September 2, 2005 on "Digital Broadband content: scientific publishing" and available at: <http://www.oecd.org/dataoecd/42/12/35393145.pdf>, governments should ensure open access to findings from publicly funded research.

The question of the mandatory nature of deposit has been the subject of animated debate in the USA and UK, and a very recent event also highlighted the importance of this trend: SHERPA launched [JULIET](#), a site which reviews the policies of research funders towards open archives. At present, it contains details on the eight British Research Councils, the Wellcome Trust and the NIH (Press release dated 29/06/06 Launch of JULIET : <http://www.sherpa.ac.uk/news/julietrelease.html> website: <http://www.sherpa.ac.uk/juliet/>).

♦ In the United Kingdom

Research Councils UK (RCUK)

During the summer of 2005, the UK Research Funding Councils proposed that the research they funded should be the subject of the systematic self-archiving of the scientific publications it generated, in open archives freely accessible via the internet. This proposal aroused a strong reaction from the Association of Learned and Professional Society Publishers, which took the form of an open letter, sent to the President of RCUK. This letter in turn triggered a response from a certain number of scientists, including Tim Berners-Lee, Steven Harnad, Peter Suber, Charles Oppenheim, etc. Beyond the polemics, this dialogue of open letters shows how the new technical conditions for the dissemination of scientific information over the internet can also call into question the economic models which exist in this sector.

In June 2006, the Research Councils of the United Kingdom adjusted their position, allowing each Research Council to publish its own directives appropriate to the specific needs of its scientific community. Three of them have already announced a policy of mandatory self-archiving. <http://www.rcuk.ac.uk/access/2006statement.pdf>

As at October 1st 2006, four of the eight RCUK councils had already adopted an official, mandatory self-archiving policy: <http://www.rcuk.ac.uk/access/index.asp>

WellCome Trust

http://www.wellcome.ac.uk/doc_wtd002766.html

The Wellcome Trust is the leading supporter of biomedical research in the United Kingdom and the first of the world's research charities to require open access to publications arising from its funded research; this concerns around 3500 articles each year. This policy, announced in November 2004, applied to all research funded as from October 1st of that year. With effect from October 1st, 2006, it will apply to all funded research, whatever the date the funds were made available. Articles must be deposited in PubMed Central (PMC), at the latest six months after their publication, with authors being responsible for reaching an agreement with their publishers. Furthermore, the Wellcome Trust is envisaging the implementation of a UK PMC in partnership with other institutions.

[information via [SPARC Open Access Newsletter](#) and the [Liblicense](#) list]

Review / note by Annaig Mahé – URFIST blog.

http://urfistinfo.blogs.com/urfist_info/2005/10/le_wellcome_tru.html

♦ In the USA

NIH

The NIH (National Institutes of Health) require the researchers they fund to deposit a copy of their articles published in peer-reviewed journals on PubMed Central, but with no restrictions concerning the embargo period, which is left at the discretion of publishers (a period of between 0 and 12 months at most). This policy has been applicable since May 2005 to all current or future NIH-funded work.

Objectives:

- to create a stable archive of publications generated thanks to NIH funding, with the aim of conservation and permanence,
- to create a document repository to be used as research management support (document searches, productivity evaluation and monitoring, and aid to setting priorities for research),
- to broaden access to documents published in the context of NIH research.

Examples of open access policies via PMC : <http://www.earlham.edu/~peters/fos/newsletter/06-02-05.htm#nih>. It should be noted that this voluntary policy has totally failed. The self-archiving rate only reaches 4%, which is even lower than the spontaneous self-archiving rate (15%). The NIH are soon going to review their policy and render self-archiving mandatory, within a maximum period of 6 months. At the same time, two US Senators have proposed the "Federal Research Public Access Act" to the American Congress, which would decree that all major research institutions receiving American funds should adopt a policy of mandatory self-archiving: http://cornyn.senate.gov/doc_archive/05-02-2006_COE06461_xml.pdf.

♦ "In developing countries"

A recent Workshop on Electronic Publishing and Open Access organized by the Indian Institute of Science was held in Bangalore (India) in November 2006. The participants (India, China, Brazil and South Africa) have convened to share their experiences in the field of open archives for promoting the open access movement in their own country. A key output from the meeting was an agreement from participants on a model National Open Access Policy for Developing Countries.

♦ At the level of the European Union

The European Commission report published in January 2006: "Study on the economic and technical evolution of the scientific publications markets in Europe" includes some ten recommendations. Recommendation A1 guarantees "public access to public research results shortly after publication" highlighting the need to "Establish a European policy mandating published articles arising from EC-funded research to be available after a given time period in open access archives".

The "Euroscience Workgroup on Scientific Publishing" has proposed certain modifications to this recommendation with the aim of optimizing it, suggesting mandatory deposit of the full text of any article paid for from EC funds, **as soon as it is accepted, without any exceptions** (cf. 2006 Final Report of the Workgroup:

http://www.euroscience.org/WGROUPS/SC_PUBLISHING/eurorec_final_210506.pdf)

More specifically, it suggests the alternative which follows:

- mandatory **deposit** of the full text, as soon as it is accepted, and without any exceptions,
- mandatory open access to all bibliographical metadata in articles at the time of deposit,
- **open access to the full text itself is simply encouraged whenever possible**, but is not mandatory. The full text may become accessible at a later stage, with reference to copyright or other proprietary rights,
- the software systems for open archives must also allow for the inclusion of a " **E-print request** " function, so that reprints can be requested and sent by e-mail, because the metadata will be visible to all.

The European Research Advisory Board (EURAB) has recommended that the European Commission should promote open access publication policies for all their publicly funded research. EURAB was invited by the Commission to examine the issue of scientific publication with particular reference to policy recommendations regarding open access for Framework Program 7 (FP7). It has recommended that a clear policy at European level is required which sets out a number of key high level principles. The Commission can play a role in three respects: as a funding body, as a policy body, as a supporting body.

http://ec.europa.eu/research/eurab/pdf/eurab_sci pub_report_recomm_dec06_en.pdf

♦ In France

At present there is no framework document concerning the introduction of institutional archives, either on the site for the Agence Nationale pour la Recherche (*National Research Agency*) nor on that for the Ministry for Higher Education and Research.

However, the agreement signed in July 2006 between the EPST (text of protocol: http://www.godoc.cnrs-gif.fr/docspdf/protocoleAO_pour_signature.pdf) and the Conference of University Presidents by representatives from each institution provided a framework for national cooperation (constitution of working groups, and of a national tool).

6.3- Links between deposit and assessment

It is important to remember that the deposit of documents for assessment **does not always imply that they will become openly available**.

However, if documents are deposited, they may gradually be rendered accessible to a wider and wider audience (assessors, researchers, general public), while complying with publishing contracts and copyright requirements.

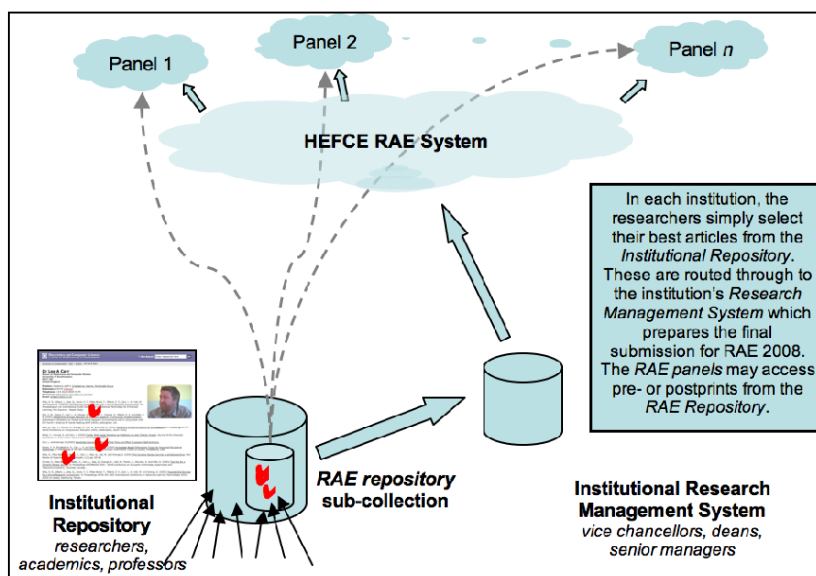
To our knowledge, there is only one system which wholly links assessment and research output repositories: that of the IRRA (Institutional Repositories and Research Assessment) in the United Kingdom (<http://irra.eprints.org/>).

6.3.1- The example of the United Kingdom: the IRRA (Institutional Repositories and Research Assessment)

Software developments for the Eprints and Dspace platforms will allow researchers themselves, as well as archive managers in different British universities, to deposit and generate data appropriate for assessment procedures on British research (RAE: Research Assessment Exercise). This procedure will be in place for assessments in 2008, and take account of the six previous years.

The principle is as follows:

- Each researcher, in each specialized area, deposits his or her documents in a repository (such as an institutional archive). A special interface is available so that he/she can select the publications and any other documents that he/she wishes to be submitted for assessment, with the possible addition of comments,
- Each institution then assembles the documents and assessment dossier for researchers, and transmits it in a single submission by discipline to the assessment body,
- This body is then responsible for distributing dossiers to the different panels of experts involved in assessment.



"IRRA workflow diagram" – Taken from: <http://irra.eprints.org/white/>

Practical applications are being developed at the universities of Southampton and Edinburgh.

This system is intended to function with repositories of different types, containing:

- either a mixture of metadata deposits only and metadata-full text deposits,
- or deposits of metadata + full text only.

Experiments by the University of Southampton have shown that the use of institutional repositories considerably reduces the workload of both assessors and researchers (<http://irra.eprints.org/white/>).

6.4- Elements to be taken into account when linking assessment and document deposits

We feel that the IRRA project in the United Kingdom (see page 31) is a very interesting approach to linking mandatory deposit and an assessment system.

- The system in fact comprises two separate systems: a document repository and a specific system for assessment based on this repository.
- Those responsible for assessment:
 - o define the information necessary to carry out the assessment, which is then translated into the form of descriptive data (metadata associated with documents),

- base themselves on the exhaustive deposit of all types of documents (preprints and postprints) in the document management system.
- The IRRA system means that through a specific interface, researchers can simply signal the documents they wish to be submitted for their collective or individual assessments, and supply those additional elements they wish to reach the assessment bodies (comments, reports, reading notes, etc.).

Linking mandatory deposit with assessment procedures has two advantages:

- a considerable reduction in the work necessary to prepare assessments, for both assessors and researchers,
- an increase in the number of documents deposited in open institutional archives, thus guaranteeing the regular updating of data.

6.5- The institutional archive: the best model?

In a 2005 study (Swan et al. 2005) performed in the United Kingdom, the authors compared three functional models for self-archiving:

- the centralised model: publications are deposited in a national archive as a priority and are then immediately accessible on-line,
- the "distributed" model: deposits are made in a network of compatible OAI archives of different types (institutional archives, thematic archives, open access journals, etc.), the metadata from which are harvested and then rendered accessible to users,
- the "harvested" model, derived from the above but under which the metadata are processed (enrichment and standardisation) before they are put on line.

The study recommended adopting the third model, and demonstrated the advantages of the distributed model over the centralised model: it was more economical, performed better and enabled more progressive implementation (the processing of metadata opens the way to improving the performance and functionalities of research), and a better "deposit rate". Researchers and institutions have a shared interest in optimising the impact of their respective scientific work: **potential contributors will supply more data to an institutional archive (and more rapidly) than to a centralised repository.**

6.6- Conclusion: towards mandatory deposit

The objectives declared by research institutions, the national and international context and the benefits for both researchers and their institutions, lead us to recommend the **mandatory deposit** of scientific work in institutional archives, which may take different forms (ranging from institutional encouragement to mandatory deposit, linked or not to assessment):

1- The "encouraged" solution, or the CNRS model

Action	Description
Type	Mandatory self-archiving of all work at the level of metadata (each principal author enters the document in the appropriate system).
Archiving of document	Immediate deposit strongly encouraged (even if the document is not disseminated).
Open access/confidentiality	Immediate open access if possible; if not, documents remain confidential for the necessary period of time.

	As soon as barriers to open access are lifted (embargo period), dissemination becomes immediate.
Links with assessment	Allow in the deposit system for a field which enables classification of the publication according to the reference system used by assessment bodies (a simple tool to be developed for researchers and/or the heads of publications of scientific departments but Links with assessment cannot really be envisaged because the coverage of scientific work is not exhaustive.

2- Solution involving mandatory deposit and dissemination under open access whenever legally possible, in line with the recommendations of the Euroscience Workgroup on scientific publishing (http://www.euroscience.org/WGROUUPS/SC_PUBLISHING/eurorec_final_210506.pdf)

Action	Description
Type	Mandatory self-archiving of all work at the level of metadata (each principal author enters the document in the appropriate system).
Archiving of document	Mandatory immediate deposit (even if the document is not disseminated)
Open access/confidentiality	Immediate open access if possible; if not, documents remain confidential for the necessary period of time. As soon as barriers to open access are lifted (embargo period), dissemination becomes immediate.
Links with assessment	Allow in the deposit system for a field which enables classification of the publication according to the reference system used by assessment bodies (a simple tool to be developed for researchers and/or heads of publications of scientific departments). Links with assessment can be envisaged because the coverage of scientific work is more exhaustive, and documents are accessible to assessment bodies.

3- Mandatory solution based on assessment.

Action	Description
Type	Mandatory self-archiving of all work at the level of metadata (each principal author enters the document in the appropriate system).
Archiving of document	Mandatory deposit (even if the document is not disseminated)
Open access/confidentiality	Immediate open access if possible; if not, documents remain confidential for the necessary period of time. As soon as barriers to open access are lifted (embargo period), dissemination becomes immediate.
Links with assessment	Allow in the deposit system for a field which enables classification of the publication according to the reference system used by assessment bodies (a simple tool to be developed for researchers and/or heads of publications of scientific departments).

	publications of scientific departments). Assessment commissions base themselves solely upon scientific work which is present and archived (having the right of access to the full text of documents) in the institutional archive.
--	---

Under all these solutions, researchers must be actively encouraged to deposit their publications:

Institutional support

- Public declaration of the policy adopted by an institution,
- Strong encouragement from the institution, and mobilisation of the scientific hierarchy,
- Training and communication activities coordinated at the level of the institution, organisation of in-house symposia in units or centres,
- Involvement of a network of information professionals (awareness and training activities, assistance with the management of copyright, etc.).

Interactive feedback on the effects of Open Access

- "E-print Request " functionality.
When a document is not openly accessible, individual requests for the full text can be transmitted via semi-automatic e-mail ("e-print request " button included in the archive software). The principal author receives an e-mail to authorise sending the document by a single click, and can rapidly distribute his documents on request (he can thus see concrete proof of the impact of open archive sites on the scientific community).
- Consultation statistics
When a document is openly accessible, each author should be able to access the consultation statistics concerning the documents he has deposited and thus monitor the impact of open access on the visibility of his own work.

7- General conclusion

[Rapid overview]

The situation concerning open archives is changing very rapidly (adoption of policies by institutions, recommendations from publishers, initiatives by research institutions and universities).

Open archives and open access journals today represent 8 million documents deposited in nearly 700 sites throughout the world (source: OAISTER) and are becoming an increasingly important supplementary source of scientific information.

Institutional archives enable the storage (archiving) of all scientific output in a digital form for the purposes of conservation and preservation. Numerous studies have proved that above all they enable an improvement in **exploitation, and increase the visibility** of scientific work. Their development now seems unavoidable.

But despite the ideal of universal open access to scientific output , encouraged by organisations and different institutions (French Academy of Sciences, etc.), and the apparent agreement of publishers to

the development of institutional open archives, deposit rates remain very low (a maximum of 20% of the production of an institution).

- Researchers feel little concerned by open access, and many of them consider that the procedures to obtain dissemination rights are too restrictive,
- They are under considerable pressure to publish in journals with a strong reputation, the impact factor of the ISI being the only one recognised at present. Citation tools on open archive sites are not yet sufficiently developed: they are poorly understood and under-utilised when searching for information or depositing documents,
- Encouragement from research institutions remains insufficient to achieve the results expected,
- Traditional scientific publishers are developing open access options for documents, either after an embargo period (generally between 6 months and a year) or with a payment system. They are thus demonstrating their ability to adapt their services and business models, to create added value on their publishing platforms and become innovative actors in open access.

Government policies (for example in the UK and USA) in favour of open access appear to be revealing their limitations as they do not really have the resources to impose their position on research institutions.

Mandatory deposit appears to be the only method which will enable full coverage in an open institutional archive, the only guarantee of the satisfactory representativeness of the production of an institution and an inevitable condition to draw all the benefits anticipated from Open Access.

The link between open archives and assessment, although not a national system, is markedly dependent on the in-house culture of an institution and its assessment process, and thus implies strong collaboration between all actors concerned by this subject. It would constitute an effective means of achieving this objective of exhaustive coverage and thus ensure optimum exploitation of the benefits expected from creating an institutional archive.

Glossary

Examples of glossaries:

ENPC dossier

http://www.enpc.fr/fr/documentation/doc_electronique/open_access_6.html

IFREMER / ARCHIMER lexicon

<http://www.ifremer.fr/docelec/lexique.htm - openaccess>

INRIA lexicon

<http://www.inria.fr/publications/archiveouverte/lexique.fr.html>

Bibliography

Benoit, A.-M. Archives ouvertes : approches juridiques. *Les archives ouvertes. Enjeux et pratiques. Guide à l'usage des professionnels de l'information*. Paris, 2005. pp. 235-236.

Bosc, H. Le libre-accès par l'auto-archivage : enjeux d'une politique forte. Présentation orale. *Réunion CEPIA-INRA*. Paris, 2006/06/23.

Brody, T.; Harnad, S.; Carr, L. Earlier Web Usage Statistics as Predictors of Later Citation Impact *Journal of the American Association for Information Science and Technology (JASIST)*, 57 (8): 1060-1072, 2006.

<http://eprints.ecs.soton.ac.uk/10713/>

Carr, L.; Harnad, S. Keystroke Economy: A Study of the Time and Effort Involved in Self-Archiving., 2005.

<http://eprints.ecs.soton.ac.uk/10688/>

Cavet, D. D'une logique de numérisation à une politique d'archives ouvertes : l'exemple de l'IRD. *Les Rencontres 2006 des Professionnels de l'Information Scientifique et Technique*. Nancy, 2006/06/19-21. Communication orale - Atelier.

Chanier, T. Les archives ouvertes : récents développements du libre accès à l'information scientifique et technique. Partie 1. *Journée de formation URFIST*. Rennes, 2006/04/03.

http://lifc.univ-fcomte.fr/~chanier/DocPerso/chanier_rennes_OA_1.ppt

Giansily, C.; Achard, A.-L.; Regolini, A.; Jannès-Ober, E. Evolution de Cemagref Publications en archive ouverte institutionnelle. *Les Rencontres 2006 des Professionnels de l'Information Scientifique et Technique*. Nancy, 2006/06/19-21. Poster.

Groupe Français de l'Industrie et de l'Information. Les grands établissements de recherche français signent un accord opérationnel sur l'archivage ouvert de la production scientifique. *La Dépêche du GFII*, 2006.

<http://www.gfii.asso.fr>

Guha, G. Enjeux de l'Open Access pour l'information scientifique et technique en France, 2005.

http://www.defidoc.com/publications/ds_libre/OpenAccess.htm

Hajjem, C.; Gingras, Y.; Brody, T.; Carr, L.; Harnad, S. Open Access to Research Increases Citation Impact., 2005a. p.

<http://eprints.ecs.soton.ac.uk/11687/>

Hajjem, C.; Harnad, S.; Gingras, Y. Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact
IEEE Data Engineering Bulletin, 28 (4): 39-47, 2005b.

<http://eprints.ecs.soton.ac.uk/12906/>

Harnad, S. Ciélographie et ciélolexie: Anomalie post-gutenbergienne et comment la résoudre. In: Origgi, G.A., N., ed. *Le texte à l'heure de l'internet*.: Bibliotheque Centre Pompidou, 2003. pp. 77-103.

http://www.text-e.org/conf/index.cfm?ConfText_ID=7

Harnad, S.; Brody, T. Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals. *D-Lib Magazine*, 10 (6), 2004.

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

Houghton, J.; Sheehan, P. The Economic Impact of Enhanced Access to Research Findings, Centre for Strategic Economic Studies. Victoria University. July, 2006. 31 p.

<http://www.cfses.com/documents/wp23.pdf>

Laloë, F. Archives ouvertes (AO) : déontologie scientifique et aspects juridiques. *Les Rencontres 2006 des Professionnels de l'Information Scientifique et Technique*. Nancy, 2006/06/19-21.

http://rpist.inist.fr/IMG/pdf/Franck_Laloe_2.pdf

Lawrence, S. Free online availability substantially increases a paper's impact. 2001.

<http://www.nature.com/nature/debates/e-access/Articles/lawrence.html>

Merceur, F. Archimer, ou la mise en place d'une archive institutionnelle à l'Ifremer, 2005. 32 p.

<http://www.ifremer.fr/docelec/doc/2005/rapport-657.pdf>

Morel-Pair, C. Des métadonnées pour les ressources électroniques. In: Aubry, C., and Janik, J., eds. *Les archives ouvertes. Enjeux et pratiques. Guide à l'usage des professionnels de l'information*. Paris: ADBS, 2005. pp. 201-211.

Perneger, T.V. Relation between online "hit counts" and subsequent citations: prospective study of research papers in the BMJ. *BMJ*, 329: 546-547, 2004.

<http://bmj.bmjournals.com/cgi/content/full/329/7465/546>

Romary, L. Préface. In: Aubry, C., and Janik, J., eds. *Les archives ouvertes. Enjeux et pratiques. Guide à l'usage des professionnels de l'information*. Paris: ADBS, 2005. pp. 5-8.

Sale, A. Generic Risk Analysis Open Access for your institution, 2006.

http://eprints.comp.utas.edu.au:81/archive/00000266/01/Risk_Analysis-v1.0.pdf

Swan, A. Open access self-archiving: An Introduction., Technical Report, JISC, HEFCE, 2005. p.

<http://eprints.ecs.soton.ac.uk/11006/>

Swan, A.; Brown, S. Open access self -archiving : an author study. Technical report, external collaborators, key perspectives. 2005.

<http://eprints.ecs.soton.ac.uk/10999/>

http://www.jisc.ac.uk/uploaded_documents/Open%20Access%20Self%20Archiving-an%20author%20study.pdf

Swan, A.; Needham, P.; Proberts, S.; Muir, A.; Oppenheim, C.; O'Brien, A.; Hardy, R.; Rowland, F.; Brown, S. Developing a model for e-prints and open access journal content in UK further and higher education
Learned Publishing, 18 (1): 25-40, 2005.

<http://eprints.ecs.soton.ac.uk/11000/>

Annex 1 – New publishing models

[[Rapid overview](#)]

The open access initiative has given rise to the evolution of publishing models: the classic publishing model ("reader pays") remains, but novel solutions have emerged:

- The "author or institution pays" model: under this model, the author or his institution pays for publication of an article once it has been accepted. The publication is then disseminated under open access, free of charge.

Biomed Central (<http://www.biomedcentral.com/>): approx. 1000 euros/article

Public Library of Science – PloS (<http://www.plos.org/index.php>): between \$2000 and \$2500/article

- mixed solutions: for example, open access to articles authorised after an embargo period (usually 6 months), or a hybrid model of "optional open access" where if authors pay a supplementary fee they can disseminate their article under open access on the journal's website.

In a recent article, Jean-Philippe Schmitt (Schmitt 2006) reviewed the initial reasons for the emergence of open access, the famous "serials crisis". He analysed the now diverging interests of these two communities and compiled a table showing the different models currently in use for open access:

Type of open access	Actors	Operation	Comments
"moving wall"	Publishers	The subscription system is retained, but the publisher undertakes to ensure open access for published articles after a fixed embargo period (e.g. 6 months); a subscription base can thus be maintained.	This is the compromise model, often claimed as being the minimum acceptable.
preprints	Researchers Institutions	Open access to scientific articles before their publication	This model enables access prior to the registration period, but although it often usefully complements publications, particularly in certain fields, it in no case constitutes an alternative: validation is not ensured.
(Self-)archiving of published articles (postprints)	Researchers Institutions	Personal sites, institutional or discipline-driven servers	Increases the visibility of work by a researcher or institution.
"Author or institution pays" model E.g. Biomed Central, PLoS	Publishers Researchers/ institutions	The author or institution (global subscription) pays for publication. Articles are deposited immediately and are accessible to all, free of charge.	The funding modes for the Biomed Central model changed recently, meaning that some research organisations had to cancel their global subscriptions.
Hybrid model, free choice for the author	Publishers Researchers Institutions	The author is free to choose the mode of dissemination for the article he is publishing: - in the form of a subscription - under OA, on condition that he pays for dissemination	After a few, isolated initiatives, this path is being explored by some major commercial publishers: - Springer: Open Choice - Blackwell: Online Open - Oxford University Press: Oxford Open
Peer-reviewed journals under OA, funded by the "author/institution pays" model	Publishers	Numerous journals are already available, either those which have converted to the new business model, or new journals	The launch of a new journal is difficult, but the gain in visibility procured by OA contributes to the rapid recognition of new journals by the scientific community.

The table below shows for some publishers and journals which do not adhere to open access the additional cost per article deposited using the Open Access model.

Publishers	Title of journal	Cost to author	Comments
US Natl. Acad. of Sciences	PNAS	\$1,000	750 euros for subscribers Article available in open access 6 months after publication
EDP Sciences	INRA journals (Annals of Forest Science, Apidologie) not including Dairy science and technology.	300 euros	Major grants to cover costs. True cost:1200 euros
Oxford Univ. Press	Journal of Experimental Botany	\$ 400	Financial support from JISC
Elsevier	Some journals	\$3,000 USD	
Springer	All journals (1582)	\$3,000 USD	Copyright is transferred by contract to authors.

- Economic aspects are important: before the development of open archive sites reaches a level such that it will ultimately enable a significant cancellation of subscriptions, research institutes, libraries or even research units will have to manage simultaneously both standard subscriptions, publication costs and the additional costs linked to new models: "author/institution pays" or a hybrid model of the "Open Choice" type.
- For the moment, publishers (still) benefit from a virtual monopoly over "prestigious" journals with a high impact factor which are widely used by researchers. Although some of them are adopting more flexible policies towards self-archiving (see the RoMEO site), they continue to apply aggressive commercial policies involving very high subscription rates which continue to rise despite the creation of consortia to regulate negotiations (cf. Couperin).
In addition, e-mail subscriptions allow them to obtain statistics concerning the true use of journal content, which they can use to calculate the costs billed to their customers (for example, the increase requested by the NPG group from INRA).
- "Author/institution pays" models of the Biomed Central type, or new hybrid models (e.g. Open Choice) may, at the level of a unit or institution, constitute very high costs which need to be covered in addition to subscriptions and the costs of publishing in classic journals.

Very recently, journals such as "Atmospheric Chemistry and Physics" (http://www.copernicus.org/EGU/acp/acpd/recent_papers.html) have proposed that authors place their submitted and thus non-validated manuscripts (preprints) in an open access site for external comments. This system functions in parallel with the standard peer review process.

Since July 2006, "Nature" has also been experimenting with this type of system.

- In October 2004, Thomson Scientific published an article on "Open Access Journals in the ISI Citation Databases: Analysis of Impact Factors and Citation Patterns."
<http://www.thomsonscientific.com/media/presentrep/essayspdf/openaccesscitations2.pdf>

The VST Newsletter summarised it as follows:

(No 8 – March 2005: <http://www.inrp.fr/vst/LettreVST/mars2005.htm#theme>)

"The author, Marie E. McVeigh, highlights the growing number of journals published under open access and indexed in ISI Citation Databases. This increase, of 239 in June 2004 versus 196 in February 2004, has been more marked in certain disciplines such as medicine or the life sciences. Furthermore, more than 55% of journals and 65% of articles indexed in 2003 in the Web of Science came from publishers which allowed their authors to deposit their articles in institutional archives.

Finally, this study proposes a breakdown by geographical region and discipline of journals, as well as an alphabetical list of open access journals indexed in the ISI database."

This study also showed that the Impact Factor of Open Access journals indexed by the ISI was generally a little below average (with some exceptions, such as Plos Biology), while the Immediacy Index (citations received during the first year after publication) was much higher.

- In an article published in PLoS Biology in 2006, Eysenbach (Eysenbach 2006) showed that within the same journal (in this case, the PNAS) articles benefiting from Open Access ("author pays") had greater impact: on average, they were cited more rapidly and more often than articles published under a traditional system.

Sample studied: 1492 articles published in PNAS, between June and December 2004, including 212 articles (or 14.2%) in OA and 1280 articles (85.8%) in traditional publication = NOA

	Date	OA (n=212. 14.2%)	NOA (n=1280. 85.8%)
Articles not cited	December 2004	170 (80.2%)	1056 (82.5%)
	April 2005	78 (36.8%)	627 (49.0%)
	October 2005	11 (5.2%)	172 (13.6%)
Average citation rate	December 2004	0.9	0.7
	April 2005	1.5	1.2
	October 2005	6.4	4.5

Open access to research articles thus results in significant effects in terms of readership and citations. Recent periodicals (created since 2000) have in a short time acquired a significant impact factor (in the ISI sense), such as PLoS Biology with 14,672 in 2005 for a journal less than two years old, or numerous titles under Biomed Central.

(<http://www.biomedcentral.com/info/about/pr-releases?pr=20050623> et <http://www.biomedcentral.com/info/libraries/BMClibupdate15>)
(Lardy 2006a)

In a document published in October 2006, J.P. Lardy, URFIST Lyon (Lardy 2006b), reviewed these new hybrid models and emphasised their divergences from certain fundamental principles concerning access, and particularly concerning author copyright.

Bibliography - Annex 1

Eysenbach, G. Citation Advantage of Open Access Articles. *PLoS Biology*, 4 (5), 2006.
http://biology.plosjournals.org/archive/1545-7885/4/5/pdf/10.1371_journal.pbio.0040157-L.pdf

Lardy, J.P. Journal of Citation Reports en profondeur, 2006a. pp. 3.
<http://urfist.univ-lyon1.fr/JCR-en-profondeur.pdf>

Lardy, J.P.- Le modèle de publication hybride : lecteur payant / auteur payant. 2006b
http://urfist.univ-lyon1.fr/publication_hybride.pdf

Schmitt, J.P. Open Access. L'avenir des publications scientifiques ? *Guide pratique Archimag : les publications électroniques. 3ème ed.*: Archimag, 2006. pp. 9-12.

Annex 2 – Addresses of electronic resources

[[Rapid overview](#)]

I- News on open access and essential sites

- ❑ Peter Suber's blog: <http://www.earlham.edu/%7Epeters/fos/fosblog.html>
SPARC Open Access Newsletter: <http://www.earlham.edu/~peters/fos/newsletter/archive.htm>
- ❑ Hélène Bosc's web pages:
INRA: http://www.tours.inra.fr/prc/internet/documentation/communication_scientifique/comsci.htm
Euroscience : http://www.euroscience.org/WGROUPS/SC_PUBLISHING/index.htm
- ❑ S. Harnad's blog
<http://openaccess.eprints.org/>
- ❑ CCSD website: Centre pour la Communication Scientifique Directe du CNRS
<http://www.ccsd.cnrs.fr/accueil.php3?lang=fr>

A few interesting reports and articles

- ❑ ENPC dossier: http://www.enpc.fr/fr/documentation/doc_electronique/dossier_openaccess.htm
- ❑ The "Open Access debate", to be followed in "Nature"
<http://www.nature.com/nature/focus/accessdebate/index.html>

II- Founder texts on Open Access

- ❑ INIST website – Libre accès à l'Information Scientifique et Technique. Special page on reference texts:
http://www.inist.fr/openaccess/rubrique.php3?id_rubrique=11
- ❑ Agreement signed in 2006 between EPST, universities, Institut Pasteur and the Conference of Advanced Specialised Schools (signed on 06/07/06)
text of protocol: http://www.godoc.cnrs-gif.fr/docspdf/protocoleAO_pour_signature.pdf.

III- Inventories and detection of archive repositories

- ❑ Registration of archives and input statistics
<http://archives.eprints.org>
- ❑ Filter / different criteria (by type, country, software employed)
<http://archives.eprints.org/?country=fr&version=&type=institutional&order=recordcount&submit=Filter>

IV- Open access and publishers

Policies adopted by scientific publishers / open access

- ❑ <http://romeo.eprints.org> or <http://www.sherpa.ac.uk/romeo.php>

Author and publisher rights

- ❑ CNRS website:
<http://www.sg.cnrs.fr/daj/propriete/droits/droits.htm>
- ❑ Example – Elsevier
<http://www.elsevier.com/wps/find/authorshome.authors/authorsrights>

Publishing models

- ❑ Atmospheric Chemistry and Physics :
http://www.copernicus.org/EGU/acp/acpd/recent_papers.html
- ❑ Biomed Central : <http://www.biomedcentral.com>
- ❑ Blackwell –« Oline Open » model : <http://www.blackwellpublishing.com/static/onlineopen.asp>
- ❑ Nature –2006 peer review trial : <http://blogs.nature.com/nature/peerreview/trial/>
- ❑ Oxford University Press - « Oxford Open » model : <http://www.oxfordjournals.org/oxfordopen/>
- ❑ Public Library of Science – PloS : <http://www.plos.org/index.php>
- ❑ Springer –« Open Choice » model : <http://www.springer.com/dal/home/open+choice?SGWID=1-40359-0-0-0>
- ❑ Directory of Open Access Journals <http://www.doaj.org>
- ❑ Lardy, J.P.- Le modèle de publication hybride : lecteur payant / auteur payant. Octobre 2006
http://urfist.univ-lyon1.fr/publication_hybride.pdf

V- Tools for open access

Metadata

Search engines – associated services

Harvesters/aggregators

- ❑ OAISTER : <http://www.oaister.org/o/oaister>
- ❑ ARC : <http://arc.cs.odu.edu/>

Commercial tools

- ❑ SCIRUS : <http://www.scirus.com>
- ❑ SCOPUS : <http://www.scopus.com>
- ❑ GOOGLE SCHOLAR : <http://scholar.google.com>
- ❑ CROSSREF : <http://www.crossref.org/index.html>

VI- Open access and researchers

Open access and measuring the impact of publications

- ❑ List of different studies performed to measure the effects of open access on the impact of articles
<http://opcit.eprints.org/oacitation-biblio.html>
- ❑ CITEBASE : <http://opcit.eprints.org>
- ❑ CITESEER : <http://citesser.ist.psu.edu/>

Open access and assessment

- ❑ Example of the IRRA (Institutional Repositories & Research Assessment)
 - <http://irra.eprints.org/white/> (White Paper)
 - http://www.city.ac.uk/inau-vb/nc/rc/136/rc136_item5.doc (mai 2006)
- ❑ Australian document
<http://ausweb.scu.edu.au/aw06/papers/refereed/richardson/paper.html>

VII- Policies in favour of open access

- ❑ **Reference site / declaration of policies**
 ROARMAP (Registry of Open Access Repository Material Archiving Policies) :
<http://www.eprints.org/signup/fulllist.php>
- ❑ List of signatories of the Declaration of Berlin
<http://www.zim.mpg.de/openaccess-berlin/signatories.html>

Research agencies and institutions – positioning and/or address of archive

- ❑ CEMAGREF
 archive : <http://cemadoc.cemagref.fr/exl-php/cadcgp.php?MODELE=PRESENTHTML&query=1>
- ❑ CNRS : http://openaccess.inist.fr/breve.php3?id_breve=418
 HAL archive: <http://hal.ccsd.cnrs.fr/>
 federative platform: <http://www.ccsd.cnrs.fr/accueil.php3?lang=fr>
- ❑ IFREMER : <http://www.ifremer.fr/docelec/doc/2005/rapport-657.pdf>
 archive : <http://ifremer.fr/docelec>
- ❑ INRA : <http://www.nancy.inra.fr/extranet/doc/diffusion/libreaccesOAI-OHologneNancy2004.pdf>
 archive : <http://www.inra.fr/prodinra/pinra/index.xsp>
- ❑ INRIA : <http://www.inria.fr/publications/archiveouverte/index.fr.html>
 archive : <http://hal.inria.fr/>
- ❑ INSERM : <http://www.inserm-actualites.com/index.php?id=514>
 archive : <http://www.hal.inserm.fr/>

Funding agencies and research partners

- ❑ UK BBSRC : The Biotechnology and biological Science Research Council <http://www.bbsrc.ac.uk/>
- ❑ UK ESRC : Economic and Social Science Research Council <http://www.esrc.ac.uk/>
- ❑ UK MRC : The Medical Research Council <http://www.mrc.ac.uk/>
- ❑ Wellcome Trust <http://www.wellcome.ac.uk/>
- ❑ EURAB : http://ec.europa.eu/research/eurab/index_en.html

Institutions practising mandatory deposit

- ❑ Australia - Queensland U. Technology : <http://www.qut.edu.au/>, Open Access Policy
http://www.mopp.qut.edu.au/F/F_01_03.html
- ❑ Europe - CERN <http://doc.cern.ch/archive/electronic/cern/preprints/open/open-2005-006.pdf>
- ❑ India - Nat Inst Tech Rourkela : <http://dSPACE.nitrkl.ac.in/dSPACE/>
- ❑ Portugal – University of Minho <https://repositorium.sdum.uminho.pt/>
- ❑ Switzerland- University of Zurich : <http://www.unizh.ch/>
- ❑ UK - University of Southampton ECS : <http://www.ecs.soton.ac.uk/>

VIII- Recent congresses

- ❑ Conferences and Workshops Related to The Open Access Movement
<http://www.earlham.edu/~peters/fos/conf.htm#past>
- ❑ Dassa, M. ; Guha, K.- Journée d'étude ADBS Recherche (12 mai 2006) : Archive ouverte, libre accès : quelle mise en place ? Documentaliste – Sciences de l'information 2006, vol. 43, n°2, 150-154.
- ❑ « Archives ouvertes, libre accès : questions et enjeux de la publication scientifique » Journée d'étude du CERSIC-ERELLIF et des URFIST de Nice et Rennes, 28 juin 2006 – Université de Rennes 2 : http://www.uhb.fr/alc/erellif/cersic/spip/article.php3?id_article=19
- ❑ [Open Access: threat or blessing](http://www.isn-oldenburg.de/%7Ehilfe/vortraege/esof06/) – Euroscience Seminar - Munich juillet 2006
<http://www.isn-oldenburg.de/%7Ehilfe/vortraege/esof06/>
- ❑ « Workshop on Electronic Publishing and Open Access » - Indian Institute of Science, Bangalore, 2-3 November 2006. <http://scigate.ncsi.iisc.ernet.in/OAworkshop2006/>
- ❑ 2. Conference Open Access - Conferência sobre o acesso livre AO conhecimento – University of Minho (Portugal) 27-28 November 2006, http://www.sdum.uminho.pt/confOA/index_en.htm

IX- Lectures, books, powerpoint presentations

- ❑ Aubry, C.; Janik, J., eds. *Les archives ouvertes. Enjeux et pratiques. Guide à l'usage des professionnels de l'information*. Sciences et Techniques de l'Information. Paris (FRA): ADBS Editions, 2005. pp. 332.
- ❑ Bosc, H. - « Les archives Ouvertes dans le monde, le modèle à suivre et pourquoi ? » Journées d'Archives Ouvertes , Université de Toulouse , 13/04/06
<http://www.urfist.cict.fr/seminaire/bosc.ppt>
- ❑ Chanier, T. Les archives ouvertes : récents développements du libre accès à l'information scientifique et technique. Partie 1. *Journée de formation URFIST*. Rennes, 2006/04/03.
http://lilc.univ-fcomte.fr/~chanier/DocPerso/chanier_rennes_OA_1.ppt
- ❑ Chanier, T. Les archives ouvertes : récents développements du libre accès à l'information scientifique et technique. Partie 2. *Journée de formation URFIST*. Rennes, 2006/04/03.
http://lilc.univ-fcomte.fr/~chanier/DocPerso/chanier_rennes_OA_2.ppt
- ❑ Documents INRA Avignon
http://www.avignon.inra.fr/internet/unites/ecologie_des_invertebres/documentation/Open%20Access-commentaires.pdf
- ❑ Jacobs, N., ed. *Open access : Key strategic, technical and economic aspects*. Chandos Information Professional Series. Oxford (GBR): Chandos Publishing, 2006. pp. 243.
- ❑ Lardy, J.P. – Université Claude Bernard – URFIST Lyon
Le libre accès à l'information scientifique et technique – août 2006
http://urfist.univ-lyon1.fr/ppt/OA_Acces-ouvert_JPL.ppt
- ❑ Mahé, A. - URFIST Paris :
« Libre accès à l'information scientifique : contexte et enjeux » - Juin 2005
<http://www.ext.upmc.fr/urfist/ArchiveOuverte/OA.html>

« Les archives ouvertes en pratique » – Janvier 2006
<http://www.ext.upmc.fr/urfist/ArchiveOuverte/OApratique.html>
Document complémentaire : Sources et liens