Science Dissemination using Open Access

Open Access means aims to remove restrictions that exist on the access to articles and knowledge to the world-wide scholarly community, in aparticular to those in developing countries. Scientists in these countries still have difficulty in publishing their work due to the lack of access to the network, to their institutional economic difficulties or to the lack of awareness of available Open Access solutions.

The visibility, usage and impact of researchers' own findings can increase with Open Access, as does their power to find, access and use the work of others.

This book aims to guide the scientific community on the requirements of Open Access, and the plethora of low-cost solutions available. A compendium of selected literature on Open Access is presented to increase the awareness of the potential of open publishing in general.

The book also aims to encourage decision makers in academia and research centers to adopt institutional and regional Open Access Journals and Archives to make their own scientific results public and fully searchable on the Internet.

http://sdu.ictp.it/openaccess/book.html

ISBN 92-95003-40-3

Science Dissemination using Open Access

SCIENCE DISSEMINATION using OPEN ACCESS

A compendium of selected literature on Open Access

Editors
Enrique CANESSA ✦ Marco ZENNARO

Science Dissemination using Open Access

A compendium of selected literature on Open Access

Science Dissemination using Open Access

For more information about this book, visit us online at
http://sdu.ictp.it/openaccess/book.html

Editors: Enrique Canessa and Marco Zennaro

Publisher

✦ ICTP - The Abdus Salam International Centre for Theoretical Physics

First edition: July 2008

ISBN 92-95003-40-3

Disclaimer

The editors and publishers have taken due care in preparation of this book, but make no expressed or implied warranty of any kind and assume no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information contained herein.

Table of Contents

Foreword

Our knowledge of the world has been improving more and more rapidly with time and is becoming increasingly reliable. It is clear that the adaptation of societies to the knowledge base has become ever more critical for sustained growth. It is sad to note, however, that access to knowledge is becoming restricted and large fractions of the world are being left behind in this enormously exciting enterprise of our time. It is critical to bridge the knowledge-gap in every possible way, especially because every segment of the world today depends on the other. A means of achieving some semblance of balance is offered by implementing Open Access solutions.

Open Access means aims to remove restrictions that exist on the access to articles and knowledge to the world-wide scholarly community, in particular to those in Developing Countries. Scientists in these countries still have difficulty in publishing their work due to the lack of access to the network, to their institutional economic difficulties or to the lack of awareness of available Open Access solutions. One hopes that Open Access will enhance educational and research opportunities and bring the world together.

Open Access solutions are desirable from another point of view as well. The visibility, usage and impact of researchers' own findings can increase with Open Access, as does their power to find, access and use the work of others. Thus, Universities and Research centers also benefit from their researchers' increased impact, which also increases the return on investment of those who fund the research -such as governments, foundations and scientific societies.

We at ICTP are acutely aware of these dimensions. Our experience with Open Access has been successful, both on the technical and organizational level. One concrete example of Open Access services at ICTP are the on-line Lecture Notes Series available through the ICTP

website www.ictp.it. These notes are formally structured pedagogical material on advanced topics directed to young students and researchers, in particular to those working under less favourable conditions. Another example is the ICTP's African Physical Review (www.aphysrev.org) which offers new possibilities to scientists and helps disseminate research carried out by African colleagues. These experiences form the basis for this book, and offer both general and technical insight into "*Science Dissemination using Open Access*", which is the chosen title for the book.

This book aims to guide the scientific community on the requirements of Open Access, and the plethora of low-cost solutions available. A compendium of selected literature on Open Access is presented to increase the awareness of the potential of open publishing in general. Discussions on open publishing via Academic Webcasting are also included. The book also aims to encourage decision makers in academia and research centers to adopt institutional and regional Open Access Journals and Archives to make their own scientific results public and fully searchable on the Internet.

The book is a collaborative effort between CERN (Switzerland) and ICTP (Italy) enabled by the support of INASP (UK). I thank the Editors and all the Authors of this compendium for their work as well as their decision to make the book freely available both in print and on the Internet, and am pleased to be able to make these prefatory remarks.

K.R. Sreenivasan
Abdus Salam Research Professor
Director, ICTP

About this book

Credits

This book was prepared for the ICTP Workshop on "Using Open Access Models for Science Dissemination" held in Trieste, Italy in July 2008 carried out in collaboration with the European Organization for Nuclear Research (CERN) and the International Network for the Availability of Scientific Publications (INASP).

Editors

Enrique Canessa is a PhD Physicist working as a Scientific Consultant at ICTP. His main areas of research and interest are in the field of Condensed Matter and scientific software applications, with particular focus on disseminating science to and within Developing Countries using open source and state-of-the-art rich-media technologies.

Marco Zennaro received his Engineering degree in Electronics from the University of Trieste, Italy. He is currently working at ICTP in projects involving wireless networking, multimedia and Open Access, and is studying for his PhD at KTH - Royal Institute of Technology, Stockholm, Sweden.

Funding

Development and publication of this book have been made possible with funding support from INASP - International Network for the Availability of Scientific Publications.

Special thanks

Our sincere thanks go to the many contributors/authors of the Open Access literature selected for the book. We have given proper attribution to the author(s) of each section and/or chapter included in this work. In particular we would like to acknowledge the contributions of Philip Bourne (PLoS), Leslie Carr (ePrints.org), Richard D. Jones (HP Labs), Ismael Peña-Lopéz (Univ. Oberta de Cataluya), Kevin Stranack (OJS, Public Knowledge Project), Peter Suber (Earlham College) Imma Subirats (FAO), Jens Vigen (CERN) and everyone else who has made this project possible.

PART 1: SELECTED LITERATURE

'tag cloud' by wordle.net

1

Overview

Based on:

- P. Suber and S. Arunachalam, World-Information City, Oct 2005 (newspaper distributed to delegates at 2005 WSIS meeting in Tunis).

- P. Suber's postings in Open Access News:
http://www.earlham.edu/~peters/fos/2008/04/strong-and-weak-oa.html
http://www.earlham.edu/~peters/fos/newsletter/02-02-06.htm

- Websites:
http://www.eprints.org/openaccess/
http://www.plos.org/oa/
http://openaccess.eprints.org/index.php?/archives/399-guid.html

Since the birth of the scientific journal in 1665 scientists have been publishing journal articles without payment. They may expect royalties for their textbooks and monographs, but they give away their journal articles in exchange for a host of intangible benefits, such as a time-stamp that gives them priority over other scientists working on the same problem, and the prestige, citations, and impact that advance their careers.

For more than 300 years, these author-donated works were distributed in print editions, whose costs were covered by subscription fees. The rise of the Internet, however, meant that the tradition of free offering by authors could finally be matched with free distribution -or Open Access- to readers.

At about the same time that the Internet was born, the price of journals began to grow sharply. The average price of a science journal has risen four times faster than inflation for the past two decades. The result is an access crisis in which no institutions can afford access to the full range of journals. Librarians have responded by canceling subscriptions and cutting into their book budgets. Scientists have responded by working out alternative ways of sharing their research.

Open Access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions. It can be delivered through OA Journals, which perform peer review, or through OA archives or repositories, which do not. One of the achievements of the worldwide OA movement is to persuade 80% of non-OA journals to let their authors deposit the peer-reviewed versions of their work in OA repositories.

OA is gathering momentum around the world. Today there are thousands peer-reviewed OA journals and interoperable OA repositories. In the US, the National Institutes of Health asks all its grantees to provide OA to the results of NIH-funded research within 12 months of publication. The Welcome Trust requires OA to Welcome-funded research within six months of publication, and the Research Councils UK are considering a similar policy with an even shorter delay. Major research institutions in

other countries have committed themselves to provide OA to their research output.

What is Open Access?

About 25,000 peer-reviewed journals are published worldwide, in all disciplines and all languages (http://www.ulrichsweb.com/ulrichsweb/). They publish about 2.5 million articles per year. Most universities and research institutions can only afford to subscribe to a fraction of those journals, so all those articles are accessible to only a fraction of their potential users. That means that research is having only a fraction of its potential usage and impact and is achieving only a fraction of its potential productivity and progress. If 100% of research articles were freely accessible through OA, then the usage, impact, productivity and progress of research would be maximised.

In the paper era there was no way to remedy this, but in the web era there is a way: "Open Access" provides free web-wide access to research journal articles, immediately and permanently.

In 2002 the Open Society Institute initiated the Budapest Declaration, supported by a group of scholars and seconded since then by thousands of signatories. The Declaration stated:

"Open Access to peer-reviewed journal literature is the goal. Open Access to peer-reviewed journal literature is the goal. Self-archiving (I.) and a new generation of open-access journals (II.) are the ways to attain this goal (Budapest Open Access Initiative 2002)."

This set the basis of OA and was later complemented by the Bethesda Statement on Open Access Publishing (2003) and the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003). These all aimed to provide definitions and commitments related to the OA paradigm. The discussion was transposed to a development context through the Salvador Declaration on Open Access: the developing world perspective (Brazil 2005), later revisited in the Bangalore Declaration: A National Open Access Policy for Developing Countries (2006).

There are two roads to OA:

- the "golden road" of OA journal-publishing, where journals provide OA to their articles (either by charging the author-institution for refereeing/publishing outgoing articles instead of charging the user-institution for accessing incoming articles, or by simply making their online edition free for all);
- the "green road" of OA self-archiving, where authors provide OA to their own published articles, by making their own e-prints free for all.

The two roads to OA should not be confused or conflated; they are complementary.

OA self-archiving is not self-publishing; nor is it about online publishing without quality control (peer review); nor is it intended for writings for which the author wishes to be paid, such as books or magazine/newspaper articles. OA self-archiving is for peer-reviewed research, written solely for research impact rather than royalty revenue.

Who benefits from Open Access?

Society as a whole benefits from an expanded and accelerated research cycle in which scientists can advance more effectively because they have immediate access to all the findings they need.

The visibility, usage and impact of researchers' own findings increases with OA, as does their power to find, access and use the findings of others. Universities co-benefit from their researchers' increased impact, which also increases the return on the investment of the funders of the research, such as governments, charitable foundations, and the tax-paying public. A number of funding bodies and research councils are now beginning to mandate that works funded by them should be made freely available using OA repositories. For teachers, Open Access means no restrictions on providing articles for teaching purposes. Only the URL need be provided; Open Access takes care of the rest. Publishers likewise

also benefit from the wider dissemination, greater visibility and higher journal citation impact factor of their articles.

Why is Open Access important?

Published research results and ideas are the foundation for future progress in science and medicine. Open Access publishing therefore leads to wider dissemination of information and increased efficiency in science, by providing:

- *Open Access To Ideas*. Whether you are a patient seeking health information, an educator wishing to enliven a lesson plan, or a researcher looking to formulate a hypothesis, making papers freely available online provides you with the most current peer-reviewed scientific information and discoveries.
- *Open Access To The Broadest Audience*. As a researcher, publishing in an Open Access journal allows anyone with an interest in your work to read it - and that translates into increased usage and impact.

Open Access: "Strong" and "Weak"

The term "Open Access" is now widely used in at least two senses. For some, "OA" literature is digital, online, and free of charge. It removes price barriers but not permission barriers. For others, "OA" literature is digital, online, free of charge, and free of unnecessary copyright and licensing restrictions. It removes both price barriers and permission barriers. It allows reuse rights which exceed fair use. Most of our success stories deliver OA in the first sense, while the major public statements from Budapest, Bethesda, and Berlin (together, the BBB definition of OA) describe OA in the second sense.

Scientists have agreed to use the term "weak OA" for the removal of price barriers alone and "strong OA" for the removal of both price and permission barriers. The new terms are a distinct improvement upon the previous state of ambiguity because they label one of those species weak and the other strong.

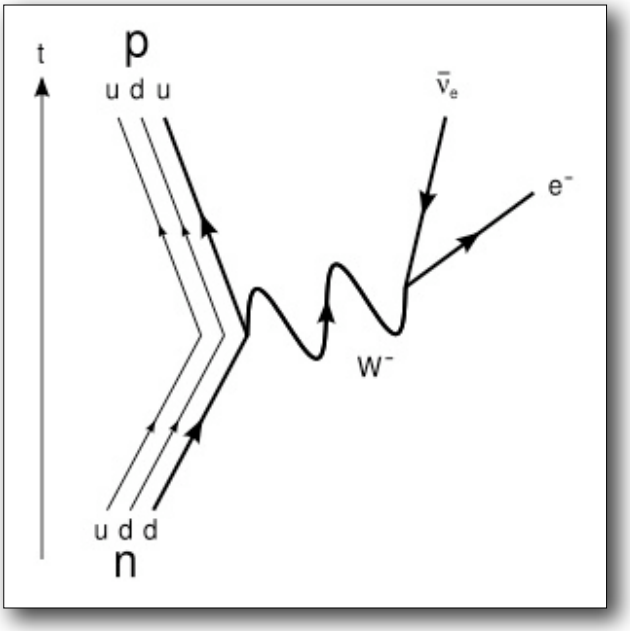


Figure: "strong" and "weak" in particle physics

On this new terminology, the BBB definition describes one kind of strong OA. A typical funder or university mandate provides weak OA. Many OA Journals provide strong OA, but many others provide weak OA.

Weak OA is a necessary but not sufficient condition of strong OA. Weak OA is often attainable in circumstances when strong OA is not attainable and should not be delayed until one can achieve strong OA. Strong OA is a desirable goal above and beyond weak OA. The desirability of strong OA is a reason to keep working after attaining weak OA, but not a reason to disparage the difficulties or the significance of weak OA.

There is more than one kind of permission barrier to remove, and therefore there is more than one kind or degree of strong OA.

Six things that researchers need to know about Open Access

1. What OA journals exist in your field?

There's no excuse not to know the OA Journals in your field. Go to the DOAJ (www.doaj.org) and browse by discipline.

Some of the Journals you find may not meet your standards for prestige or impact. But others might. Nearly every scientific discipline has an OA journal in the top cohort of impact factors.

If you learn what OA Journals exist in your field and decide against each of them, all right. At least you made an informed decision. But check the DOAJ again when you've written your next paper. Things are changing fast. Established OA Journals are growing in prestige; some are getting impact factors; new OA Journals are being launched; non-OA Journals are converting to full OA or OA hybrid models; and non-OA Journals are experimenting with different forms of OA.

If you don't publish in an OA Journal, you can publish in a non-OA Journal and self-archive the peer-reviewed version of your manuscript in an OA repository. About 70% of existing non-OA Journals already permit this. More in #4 below.

2. OA journals are not the whole story of OA. There are also OA archives or repositories

When people hear about OA for the first time, they tend to take away that OA Journals are the way to deliver it. There are two primary vehicles of OA, not just one. OA repositories don't perform peer review; they merely make their contents freely available to the world. But they can contain peer-reviewed postprints as easily unrefereed preprints. You can deposit a preprint at the time you submit it to a journal and then deposit the postprint after it's published. You can deposit your postprint in an OA repository even if you also publish it in a conventional or non-OA journal. Don't let the novelty of OA repositories make them invisible. Don't believe that if the concept is too good to be true then it can't be true.

The best places to look for OA repositories are the Registry of Open Access Repositories (ROAR) at http://archives.eprints.org/ and OpenDOAR (Directory of Open Access Repositories) at www.opendoar.org

3. OA archiving only takes a few minutes

Les Carr and Stevan Harnad (see http://eprints.ecs.soton.ac.uk/10688/) studied two months of log activity at a much-used repository and found that the time required for deposit averaged 10 minutes per paper. Taking into account the rate at which authors had their work archived for them by others (co-authors, librarians, students, or assistants), authors who published one paper per month would spend less than 40 minutes per year on their deposits.

If you haven't deposited papers in a repository yourself and worry about adding one more task to your schedule, at least trust the Carr-Harnad evidence more than any anecdotes you might have heard from colleagues. If you've deposited once but not twice, trust the evidence that the time requirement plummets. (Compare the first time you used endnotes in a word processor with the second time). If you're worrying about adding a new task regardless of the time required, then think about the many more time-consuming jobs you already do to make your work known to the world, such as keeping your c.v. up to date, mailing offprints, and sending your bibliography to deans and department chairs. Self-archiving takes less time and has more impact than any of these.

4. Most non-OA journals allow authors to deposit their postprints in an OA repository

The best current estimate is that 70% of non-OA Journals consent in advance to postprint archiving (see http://romeo.eprints.org/stats.php).

When you publish in one of these Journals, you don't need further permission for self-archiving, even if you've transferred the copyright to the journal. These Journals have already given permission. For this significant majority of peer-reviewed Journals, the obstacle to OA is author failure, not copyright complexity or publisher opposition. Journals have opened the door and authors have to walk through.

SHERPA (http://www.sherpa.ac.uk/romeo.php) and Eprints (http://romeo.eprints.org) both maintain online databases where you can look up a journal and finds its policy on self-archiving.

Three notes on the 70% figure. First, it represents surveyed Journals. Among unsurveyed Journals, there are likely to be Journals that do, and Journals that don't permit postprint archiving. We don't know their proportions yet. Second, the number represents Journals that consent in advance to postprint archiving without requiring case-by-case requests. Many that do not consent in advance will still consent if asked individually, however. Elsevier routinely granted individual requests until mid-2004 when it decided to offer blanket permission instead. Third, it represents the Journals that consent to postprint archiving, not preprint archiving. If we count the Journals that consent to preprint or postprint archiving (or both), the figure rises to 93%.

Note the all-important consequence of this kind of blanket permission. OA archiving is compatible with publishing in most conventional, subscription-based Journals. If the top Journals in your field (by impact or prestige) are not OA, you can go for impact or prestige and still have OA. It's rarely a trade-off.

5. Journals using the Ingelfinger rule are a shrinking minority

Some authors are afraid that depositing a preprint in an OA repository will disqualify it for subsequent publication. It's true that some Journals refuse to publish papers that have previously circulated as preprints or whose results have been publicized. This is called the Ingelfinger rule, named after a former editor at the New England Journal of Medicine. The rule is rare outside the field of medicine and in decline.

There are some very rare Journals, like the California Law Review, that allow postprint archiving but not preprint archiving. But essentially all the Journals that don't allow preprint archiving (i.e. that follow the Ingelfinger rule) also bar postprint archiving. Only 7% of surveyed Journals fall into this category. Don't let groundless fears deter you from preprint archiving.

If you worry about the Ingelfinger rule, check out the policies of the Journals where you intend to submit your work.

6. OA enlarges your audience and citation impact

This is the chief reason for authors to provide OA to their own work. OA increases the audience for a work far beyond the audience of any priced journal, even the most prestigious or popular journal. Studies in many fields show a correlation between OA and citation-count increases from 50% to 250% (see http://opcit.eprints.org/oacitation-biblio.html).

There is almost certainly causation here as well as correlation, though this hasn't been nailed down yet. There are many hypotheses to explain the correlation. Some of it seems to arise from the fact that self-archived articles circulate sooner than journal-published articles (and have a head-start toward citations) and the fact that authors self-archive their best work (biasing the OA sample toward quality). But it's very likely that ongoing studies will show that much of the correlation is simply due to the larger audience and heightened visibility for the work among researchers who find the work useful, relevant, and worth citing in their own work.

These studies bring a welcome note of self-interest to the case for OA. Providing OA to your own work is not an act of charity that only benefits others, or a sacrifice justified only by the greater good. It's not a sacrifice at all. It increases your visibility, retrievability, audience, usage, and citations. It's about career-building. For publishing scholars, it would be a bargain even if it were costly, difficult, and time-consuming.

2

Declarations

Based on:

- Websites:

http://www.soros.org/openaccess/
http://oa.mpg.de/
http://www.public-domain.org/?q=node/60

Budapest Open Access Initiative

The Budapest Open Access Initiative (http://www.soros.org/openaccess/) arises from a small but lively meeting convened in Budapest by the Open Society Institute (OSI) on December 1-2, 2001. The purpose of the meeting was to accelerate progress in the international effort to make research articles in all academic fields freely available on the Internet. The participants represented many points of view, many academic disciplines, and many nations, and had experience with many of the ongoing initiatives that make up the Open Access movement. In Budapest they explored how the separate initiatives could work together to achieve broader, deeper, and faster success. They explored the most effective and affordable strategies for serving the interests of research, researchers, and the institutions and societies that support research. Finally, they explored how OSI and other foundations could use their resources most productively to aid the transition to Open Access and to make open-access publishing economically self-sustaining. The result is the Budapest Open Access Initiative. It is at once a statement of principle, a statement of strategy, and a statement of commitment.

The initiative has been signed by the Budapest participants and a growing number of individuals and organizations from around the world who represent researchers, universities, laboratories, libraries, foundations, journals, publishers, learned societies, and kindred open-access initiatives. We invite the signatures, support, and participation of the entire world scientific and scholarly community. Contact e-mail: openaccess@soros.org

Berlin declaration on Open Access to knowledge in the sciences and humanities

Our mission of disseminating knowledge is only half complete if the information is not made widely and readily available to society. New possibilities of knowledge dissemination not only through the classical form but also and increasingly through the Open Access paradigm via the Internet have to be supported. We define Open Access as a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community. In order to realize the vision

of a global and accessible representation of knowledge, the future Web has to be sustainable, interactive, and transparent. Content and software tools must be openly accessible and compatible.

Definition of an Open Access contribution

Establishing Open Access as a worthwhile procedure ideally requires the active commitment of each and every individual producer of scientific knowledge and holder of cultural heritage. Open Access contributions include original scientific research results, raw data and metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia material.

Open Access contributions must satisfy two conditions:

- The author(s) and right holder(s) of such contributions grant(s) to all users a free, irrevocable, worldwide, right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship (community standards, will continue to provide the mechanism for enforcement of proper attribution and responsible use of the published work, as they do now), as well as the right to make small numbers of printed copies for their personal use.
- A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in an appropriate standard electronic format is deposited (and thus published) in at least one online repository using suitable technical standards (such as the Open Archive definitions) that is supported and maintained by an academic institution, scholarly society, government agency, or other well established organization that seeks to enable Open Access, unrestricted distribution, inter operability, and long-term archiving.

Supporting the transition to the electronic Open Access paradigm

Our organizations are interested in the further promotion of the new Open Access paradigm to gain the most benefit for science and society. Therefore, we intend to make progress by:

- encouraging our researchers/grant recipients to publish their work according to the principles of the Open Access paradigm;
- encouraging the holders of cultural heritage to support Open Access by providing their resources on the Internet;
- developing means and ways to evaluate Open Access contributions and online journals in order to maintain the standards of quality assurance and good scientific practice;
- advocating that Open Access publication be recognized in promotion and tenure evaluation;
- advocating the intrinsic merit of contributions to an Open Access infrastructure by software tool development, content provision, metadata creation, or the publication of individual articles.

We realize that the process of moving to Open Access changes the dissemination of knowledge with respect to legal and financial aspects. Our organizations aim to find solutions that support further development of the existing legal and financial frameworks in order to facilitate optimal use and access.

Governments, universities, research institutions, funding agencies, foundations, libraries, museums, archives, learned societies and professional associations who share the vision expressed in the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities are therefore invited to join the signatories that have already signed the Declaration. Contact e-mail: praesident@gv.mpg.de

Open letter signed by 25 Nobel Prize winners

Dear Members of (the U.S.) Congress:

As scientists and Nobel laureates, we are writing today to express our strong support for the House Appropriations Committee's recent direction to NIH (National Institutes of Health) to develop an open, taxpayer access policy requiring that a complete electronic text of any manuscript reporting work supported by NIH grants or contracts be supplied to the

National Library of Medicine's PubMed Central (www.pubmed.gov). We believe the time is now for all Members of Congress to support this enlightened policy.

Science is the measure of the human race's progress. As scientists and taxpayers too, we therefore object to barriers that hinder, delay or block the spread of scientific knowledge supported by federal tax dollars - including our own works.

Thanks to the Internet, today the American people have access to several billion pages of information, frequently about disease and medical conditions. However, the published results of NIH- supported medical research for which they already have paid are all too often inaccessible to taxpayers.

When a woman goes online to find what treatment options are available to battle breast cancer, the cutting-edge, peer- reviewed research remains behind a high-fee barrier. Families looking to read clinical trial updates for a loved one with Huntington's disease search in vain -because they do not have a journal subscription. Libraries, physicians, health care workers, students, researchers and thousands of academic institutions and companies are hindered by the costs and delays in making research widely accessible.

There's no question, Open Access truly expands shared knowledge across scientific fields -it is the best path for accelerating multi-disciplinary breakthroughs in research.

Journal subscriptions can be prohibitively expensive. In the single field of biology, journals average around US$1,400 and the price is almost double that in chemistry. These already-high prices are rising fast, far in excess of inflation and the growth of library budgets. An individual who cannot obtain access to a journal in a library may buy copies of solo articles they need, but that can cost them US$30 or more for each article.

The National Institutes of Health has the means today to promote Open Access to taxpayer-funded research -through the National Library of Medicine. If the proposal put forth in the House of Representatives is adopted, NIH grantees may be expected to provide to the Library an electronic copy of the final version of all manuscripts accepted for publication, after peer review, in legitimate medical and scientific journals. At the time of publication, NIH would make these reports freely available to all through their digital library archive, PubMed Central (PMC).

Figure: logo of PubMed Central

There is widespread acknowledgment that the current model for scientific publishing is failing us. An increase in the volume of research output, rising prices and static library budgets mean that libraries are struggling to purchase subscriptions to all the scientific journals needed.

Open Access, however, will not mean the end of medical and scientific journals at all. They will continue to exercise peer- review over submitted papers as the basis for deciding which papers to accept for publication, just as they do now.

In addition, since Open Access will apply only to NIH-funded research; journals will still contain significant numbers of articles not covered by this requirement and other articles and commentary invaluable to the science community. Journals will continue to be the hallmark of achievement in scientific research, and we will depend on them.

The trend towards Open Access is gaining momentum. Japan, France and the United Kingdom are beginning to establish their own digital

repositories for sharing content with NIH's PubMed Central. Free access to taxpayer funded research globally may soon be within grasp, and make possible the freer flow of medical knowledge that strengthens our capacity to find cures and to improve lives. As the undersigned Nobel Laureates, we are committed to Open Access. We ask Congress and NIH to ensure that all taxpayers get their money's worth. Our investment in scientific research is not well served by a process that limits taxpayer access instead of expanding it. We specifically ask you to support the House Appropriations Committee language as well as NIH leadership in adopting this long overdue reform.

Signed by Twenty Five Nobel Laureates

August 26, 2004

Name	Category of Nobel Prize Awarded	Year
Peter Agre	Chemistry	2003
Sidney Altman	Chemistry	1989
Paul Berg	Chemistry	1980
Michael Bishop	Physiology or Medicine	1989
Baruch Blumberg	Physiology or Medicine	1976
Gunter Blobel	Physiology or Medicine	1999
Paul Boyer	Chemistry	1997
Sydney Brenner	Physiology or Medicine	2002
Johann Deisenhofer	Chemistry	1988
Edmond Fischer	Physiology or Medicine	1992
Paul Greengard	Physiology or Medicine	2000
Leland Hartwell	Physiology or Medicine	2001
Robert Horvitz	Physiology or Medicine	2002
Eric Kandel	Physiology or Medicine	2000
Arthur Kornberg	Physiology or Medicine	1959
Roderick MacKinnon	Chemistry	2003
Kary Mullis	Chemistry	1993
Ferid Murad	Physiology or Medicine	1998
Joseph Murray	Physiology or Medicine	1990
Marshall Nirenberg	Physiology or Medicine	1968
Stanley Prusiner	Physiology or Medicine	1997
Richard Roberts	Physiology or Medicine	1993
Hamilton Smith	Physiology or Medicine	1978
Harold Varmus	Physiology or Medicine	1989
James Watson	Physiology or Medicine	1962

3

Open Access to science in Developing Countries

Based on:

- P. Suber and S. Arunachalam, World-Information City, Oct 2005 (newspaper distributed to delegates at 2005 WSIS meeting in Tunis).

- K. Stranack, "Starting a New Scholarly Journal in Africa", Public Knowledge Project, 2006.
Available at: http://pkp.sfu.ca/files/AfricaNewJournal.pdf

- News from ICTP Magazine, Winter 2006-2007, No. 119

Open Access is a matter of special concern in Developing Countries, which have less money to fund or publish research and less to buy the research published elsewhere. Most libraries in sub-Saharan Africa have not subscribed to any journal for years. The Indian Institute of Science, Bangalore, has the best-funded research library in India, but its annual library budget is just Rs 100 million (about US$2.2 million).

There are several programs, like HINARI (http://www.who.int/hinari/), AGORA (www.aginternetwork.org) for Libraries and eJDS (www.ejds.org) for single scientists, in which journal publishers donate electronic subscriptions to Developing Countries whose per capita GDP is less than US$1,000. These programs mitigate the access crisis but do not solve it. India is surprisingly excluded even though its per capita GDP is less than US$500. Moreover, insofar as they satisfy demand, they reduce the urgency of deep reforms that will bring about a superior, OA system of scientific communication.

About half the world's OA journals pay their bills by charging upfront fees for accepted papers. The fees are usually paid by the author's research grant or employer, not out of the author's pocket. The Public Library of Science and BioMed Central, the two best-known OA publishers, waive these fees in cases of economic hardship, no questions asked.

There are many successful OA initiatives in the developing world. These include Bioline International, which hosts electronic OA versions of 40 Developing Countries journals; SciELO (www.scielo.br), which hosts more than 80 journals published in Latin American countries and Spain; and African Journals Online (AJOL, www.ajol.info), which provides free online access to titles and abstracts of more than 60 African journals and full text on request. The Electronic Publishing Trust for Development (EPT, www.epublishingtrust.org), established in 1996, promotes Open Access to the world's scholarly literature and the electronic publication of bioscience journals from countries experiencing difficulties with traditional publication.

India is home to many OA journals that charge no author-side fees. All 10 journals of the Indian Academy of Sciences and all four journals of the Indian National Science Academy are OA journals. INSA (www.insaindia.org) has already produced free-access electronic versions of back volumes for all its journals, and the Indian Academy of Sciences has launched a similar digitization project for its back run. The Journal of the Indian Institute of Science is also available in this form back to its very first issue, published in 1914. The Indian Medlars Centre of the National Informatics Centre is bringing out OA versions of 33 biomedical journals and has an OA bibliographic database, providing titles and abstracts of articles from 50 Indian biomedical journals. Medknow Publications (www.medknow.com), a company based in Mumbai, has helped 30 medical journals make the transition from print to electronic Open Access and most of them are doing much better now than before.

For researchers in Developing Countries, OA solves two problems at once: making their own research more visible to researchers elsewhere, and making research elsewhere more accessible to them. OA, if adopted widely, can raise the profile of an entire nation's research output. When Indian research, for example, is published in expensive journals, then all too often it goes unnoticed by other researchers in India. OA journals and archives help to integrate the work of scientists everywhere into the global knowledge base, reduce the isolation of researchers, and improve opportunities for funding and international collaboration.

Although developed countries were the first to encourage OA to publicly-funded research, the model is very appealing in Developing Countries and likely to spread. One direct way is simply to put an OA condition on publicly-funded research grants. Another is to have universities and research laboratories set up institutional archives and adopt policies encouraging or requiring researchers to deposit their research output even if they also publish it in conventional journals.

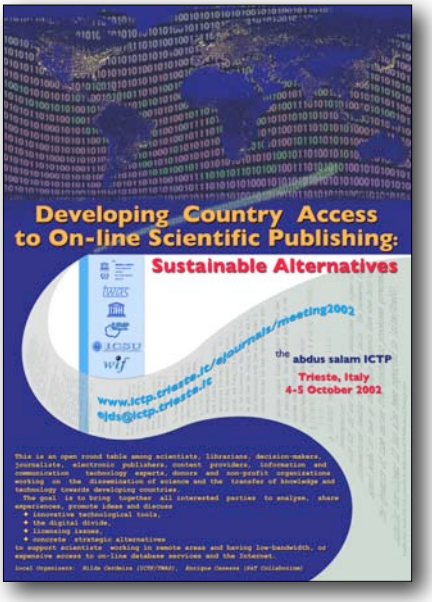


Figure: conference on Developing Country Access to On-line Scientific Publishing

Providing OA to publicly-funded research accelerates research, gives taxpayers (both lay readers and professional researchers) access to the research they funded, and increases the return on their investment in research. As this argument gets traction in Developing Countries, the transformation should be dramatic.

Doesn't the digital divide interfere with these plans? Yes and no. First, internet access is improving rapidly in many Developing Countries and equipment costs and connectivity charges are coming down. Second, we should work now on the content side of the divide in order to take full advantage of every increment of progress on the hardware side. Primarily, this means educating scientists about the benefits of OA and persuading universities, libraries, funding agencies, and governments to adopt OA-friendly policies.

OA helps researchers directly, both as authors and readers. It helps the institutions that fund and supervise research, from universities and laboratories to foundations and governments. It widens the distribution of research literature and lowers costs at the same time, and does so without compromising peer review, preservation, indexing, or the other virtues of conventional publishing. Above all, because OA enhances research productivity and accelerates the pace of discovery, it helps everyone who benefits from research advances.

Starting a new scholarly Open Access journal in Africa

There are several important benefits to starting a new journal, for you, your discipline, your institution, and your country. With these benefits come challenges, including the need to find time, money, and people to bring it all together successfully. Despite these challenges, many new journals are successfully established every year, based on a variety of publication types and economic models. This chapter outlines choices and offers suggestions to help you develop a respected and sustainable publication.

Benefits

There are a variety of important reasons for creating a journal, such as providing a new and unique record of scholarly activity. Whether it is in the Health Sciences, Sociology, or Geology, presenting an African perspective on these ideas is a crucial contribution to the academic community. Another reason to publish a journal is for the scholarly recognition it will bring to your institution. As well, your career development will be enhanced through participation in the creation and sharing of new ideas and knowledge. Finally, these new ideas and knowledge will contribute to the economic and cultural development of your community, your country and Africa. Indigenous publishing can help to close the "knowledge gap" between the well-funded and powerful voices from the north and the often-overlooked ideas, innovations, and discoveries from the south.

Another important contribution that your journal can make is the enrichment of your own research area. Already, African journals exist in a

wide range of disciplines, providing a uniquely African perspective on critical research questions. Journals such as SAHARA J (Journal of Social Aspects of HIV/AIDS Research Alliance) are having a significant impact on the wider understanding of the issues they examine, and to the region in which they are situated. When you start a new journal, you join a community of scholars challenging the domination of ideas by International publishing bodies. Journals can, through the production and sharing of local knowledge and a local perspective, make local research more visible throughout Africa and to researchers, students, and scholars around the world. Not only does developing/establishing a journal benefit your research area, but it can also support the goals of your institution, university or research centre. The kind of recognition, both national and international, that a new journal can bring are of utmost important to research administrators. A successful new journal demonstrates research money at work. This is important to the institutional funders, who must make decisions about the allocation of scarce financial resources. Your journal can demonstrate the ability of your institution to compete with other research agencies in the production of knowledge, while also forming the basis of new collaborations, between local, regional, or international researchers, research departments, and institutions.

Challenges

As well as benefits, there are challenges to establishing a scholarly or scientific journal. One of the first to be considered is the commitment of time and money. Adopting basic project management skills, such as work plans with due dates, agreed upon areas of responsibilities and tasks, and a forecasted budget with scheduled times for project evaluation and accountability, can turn this endeavour from an overwhelming challenge to a rewarding success. Another significant challenge faced when starting a journal is finding the right people to participate. There is, however, a long tradition of volunteer labour in journal publication for authoring, reviewing, and editing, with the time donated being part of an individual's professional development, and as a way to keep current and contribute to the profession. This unpaid labour in these key roles makes it possible to start a journal with a very modest budget.

You will first need a committed journal manager and editor to lead the project. Often these two roles are carried out by one person - and it may be you! You will need to form an effective editorial board, made up of respected scholars in the subject areas of your journal. They will play a critical role in developing the policies for the new journal, and establishing its credibility from the beginning. And finally, you will need to encourage submissions from the best and brightest authors in the field, as well as recruit committed reviewers. It is possible, however, to see this as an opportunity to offer the right people involvement in building a highly reputable journal.

Other issues to consider are the high cost of producing and distributing a print journal. You will need to explore a variety of funding sources, but you should consider the substantial savings that can be had through online publishing. For online journals, you must ensure that you, your journal collaborators, and your readers have adequate hardware, software, and bandwidth availability. Without this basic technological infrastructure, publishing any journal is a significant challenge. Lastly, you will need to find ways to promote your journal. Its success will be dependent upon finding an audience, and the audience finding the journal. This booklet provides some answers to these questions, and offers some suggestions in ways to overcome the challenges you will face in setting up your journal.

Personal benefits can also result from starting a new journal. Becoming a managing editor provides outstanding experience and makes a powerful addition to any curriculum vitae, leading to promotions and appointments. This is also true for the community of authors, editors, and reviewers that will develop around your journal, all of whom will also accrue valuable experience and opportunities for career advancement. The recognition that comes from managing a journal can also lead to opportunities, such as invitations to speak at conferences or to collaborate in larger research initiatives. The potential for career benefits are an important factor in why you should consider starting a new journal. Not only can you, your colleagues, your institution, and your research discipline benefit from the establishment of a new journal, but so too can your city or town, your country, and Africa itself. A new African scholarly

or scientific journal will contribute local research knowledge which can be applied to national economic and cultural development goals.

The African Physical Review: An example

As part of a larger effort to build scientific capacity in Africa and to promote international scientific exchange, the Abdus Salam International Centre for Theoretical Physics (ICTP) has launched The African Physical Review (APR, www.aphysrev.org). This is an on-line peer-reviewed international journal featuring articles in all branches of theoretical and experimental physics. In addition to high-quality research articles, the journal will include "literature reviews" and "brief communications" on a broad range of topics of interest to the physics community. Among others, APR is being cosponsored by the African Academy of Sciences in Nairobi, Kenya.

APR publishes articles by physicists not only from Africa but from around the world. A unique feature of the journal is the publication of invited articles on the growing interdisciplinary nature of physics research. There are also be special editions dedicated to a single topic and sections devoted to conference proceedings, particularly conferences that take place in Africa. The goal is to create a first-class journal of value to African physicists and physicists worldwide.

APR seeks to increase interaction between African physicists and the global physics community.

Figure: The African Physical Review

Two major factors led ICTP to sponsor this journal. First, low-cost and easily accessible electronic publishing is rapidly replacing traditional

high-cost print journal publishing. This trend is altering the entire field of peer-reviewed scientific publishing, removing the stigma of inferior quality from electronic publications. And, second, poor nations do not subscribe to conventional print publications simply because they are too expensive. Limited access to the most recent publications have put scientists in Africa and other Developing Countries at a distinct disadvantage, especially in the many fields of physics where new information and discoveries are taking place at an unprecedented pace. APS is designed to address this problem.

APR not only reflects the ongoing commitment of ICTP to the African physics community but also represents an effort to create new and innovative channels for fostering both greater interaction and integration between African physicists and their colleagues from around the world.

4

Types of journals published

Based on:

- K. Stranack, "Starting a New Scholarly Journal in Africa", Public Knowledge Project, 2006.
Available at: http://pkp.sfu.ca/files/AfricaNewJournal.pdf

There are a variety of different kinds of journals produced in the world of scholarly and scientific publishing. Some journals are broadly focused and cover a range of topics from a diversity of contributors. Others are more narrowly focused, on either the research interests of the contributors, all coming, perhaps, from the same academic discipline or from the same institution, or on the subject matter. It is important to select the best type of journal to meet both your immediate and long-term needs.

For a new scientific or scholarly journal, deciding on the type of journal to produce is a significant decision, which will have long-term implications on the direction of your publication. A wide variety of options exist, each with their own strengths and challenges. Deciding on the best choice for your new journal must be based upon your own research interests, the size of the potential audience for your new journal, and any institutional requirements that may need consideration. It is also important to remain flexible, as it is possible to combine different journal types, depending on your own situation. For example, you may wish to start a journal with a national perspective and with more general content, or instead, a Pan-African journal with a more specialized topic might be the most appropriate.

General journals

One option is the general journal, which may be multidisciplinary, broadly focused, and accepting contributions from many fields of research. Some examples of these types of journals include the IFE Journal of Science (http://www.ajol.info/journal_index.php?jid=219&ab=ijs) and the Humanities Review Journal (http://www.ajol.info/journal_index.php?jid=36&ab=hrj).

Both of these journals provide coverage of a variety of topics within their broad areas of science or the humanities. The IFE Journal of Science covers research in the areas of chemical, biological, mathematical and physical sciences, as well as the applied areas of biochemistry, geology, microbiology and such allied fields as biotechnology, genetics, food chemistry, agriculture, medical, and pharmaceutical sciences.

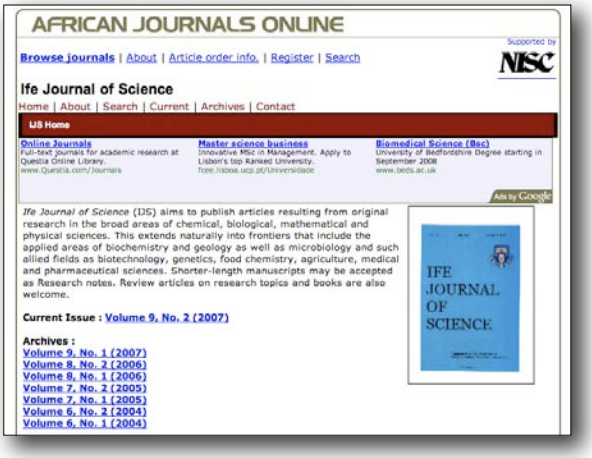


Figure: IFE Journal of Science

The Humanities Review Journal includes contributions from the fields of theatre arts, philosophy, English language, English literature, history, music, communication arts, anthropology and other relevant disciplines.

Figure: Humanities Review Journal

This type of journal is particularly useful for those needing to draw on a wider range of authors, editors, and readers. Because of the breadth of

coverage, participation may be easier to encourage, especially in the early development of a new journal. The main disadvantage of this type of journal can be its lack of focus, which may make it difficult to promote to a new audience.

Specialized journals

The specialized journal is more common than the general journal. It has a more narrowly-defined focus, and is often discipline-specific. Some examples include Obstetrics and Gynaecology Forum (http://www.ajol.info//journal_index.php?jid=202&ab=ogf) and SAHARA J, Journal of Social Aspects of HIV/AIDS Research Alliance (http://www.ajol.info//journal_index.php?jid=197&ab=sahara). This type of journal is an excellent choice when a large enough community of readers and authors exists to sustain it. Its focused content readily appeals to scholars or practitioners in the content area, making marketing and promotion much easier than for the more general journal.

Regional journals

Another option is the regional journal, covering research from a particular geographic area, whether national or international. The Uganda Journal (http://www.ajol.info/journal_index.php?jid=135&ab=uj) is a good example of a national journal interested in all scholarly aspects of the country. A Pan-African example is Africa Insight (http://www.ajol.info/journal_index.php?jid=128&ab=ai), which examines changes happening throughout the continent. This type of journal draws on a community of scholars throughout a country, or even beyond, into several countries across Africa. A regional focus can often be attractive to institutional funders, interested in how the journal is contributing to the nation. The Pan-African journal can also be of interest as an example of international collaboration and cost-sharing.

Institutional journals

In addition to journals focusing on a particular discipline or coming from a specific geographic location, some journals are also centred on an individual institution. Examples include Huria: Journal of the Open

University of Tanzania (http://www.ajol.info/journal_index.php?ab=huria&jid=235) and the Journal of the Cameroon Academy of Sciences (http://www.ajol.info/journal_index.php?jid=93&ab=jcas). Both of these journals primarily contain content concerning their sponsoring institution. Financial support for your journal may require focusing on the issues or contributions from an association or institution. This can often be an excellent source of sustainable income.

Annual reviews

Another possible model has less to do with content than it does with the publication schedule. Annual reviews are produced once each year, and can cover any of the areas discussed above. The main benefit to this format is the reduced publishing schedule requires fewer resources to sustain. For publishers with very limited sources of income, this could provide a way to produce their content. An example of this type of journal is the Annales Aequatoria (http://www.ajol.info/journal_index.php?jid=32&ab=aq), which covers a wide range of subjects including African Linguistics, Cultural Anthropology, Literature in Bantu languages, History, Archaeology and more. Although it only produces one volume each year, each one can be as much as 600 pages in length, depending upon the content that was generated throughout the year. This format provides maximum flexibility, with the possibility of the initial volumes being more modest in length, but able to grow as alternative sources of revenue become available. The Annual Review could also be an option for your first year or two, as your journal develops. With time, you may be able to expand to publishing more frequently each year.

Deciding on a publication type

The most important considerations in determining the type of journal you wish to publish include:

- your research interests;
- your existing or potential collaborators; and
- your institutional assets and requirements.

In terms of your personal research, it is important that your journal will allow you to participate in the scholarly exchange that will be taking place. To sustain your long-term interest, your area of expertise must be accommodated, whether as part of a journal focused specifically in your discipline, or within a broader journal, where your contributions and those of your colleagues will be valued. Another important consideration will be determining the potential audience. This will be the source of your readers, contributors, reviewers, and editors. The larger the pool of talented people to draw upon, the better the chance of your journal's success. If a significant community of researchers in a particular field of interest in your country already exists, you could establish a discipline-specific, national journal. If the community was spread across several countries, you may want to consider an international journal. If researchers in your area of expertise are too limited, you may want to set up a more general journal, providing a more sustainable arena for your work, and the work of other scholars. Lastly, you may need to establish an association or university-based journal, based on the demands of your funders. Winning and maintaining the support of funding institutions will be critical to the sustainability of your new journal.

5

Getting professional

Based on:

- K. Stranack, "Getting Found, Staying Found, Increasing Impact", Public Knowledge Project, 2006.
Available at: http://pkp.sfu.ca/files/GettingFoundStayingFound.pdf

All successful academic journals are based upon a solid foundation of professional recognition. Being recognized and respected by your peers encourages their participation (as readers, authors, reviewers, editors, and board members). It helps to overcome any reluctance to get involved with your journal, easing fears that they may be wasting their valuable time and effort on their road to securing tenure or building their own careers. Professional recognition also builds confidence in the different indexing and database services to include your content, and in libraries willingness to promote your publication.

Challenges for new journals

Fostering professional recognition is always a challenge, and can be especially difficult for a new journal, with no archive of high quality content to point to. There are, however, some important first steps that can be taken to assist with this process:

1. Your journal must develop a professional layout, for both your web site and your articles. Web-based interfaces have become extremely sophisticated and, whether accurate or not, amateur designs tend to reflect amateur efforts. The Open Journal Systems (OJS) software provides a strong template for your journal, but enlisting the services of a professional graphic designer to give it a fresh, unique look can be a very important step in building respect and recognition.
2. Ensure that your journal is being hosted on a secure, professionally managed server. Your content must be online and available at least 99.9% of the time. Frequent downtime reflects an amateur effort.
3. Try to encourage the participation of recognized colleagues as editorial board members or authors, and be sure to make their involvement in your project visible.
4. Make use of the OJS peer review system and let your audience know that your journal is following the established guidelines for ensuring high quality content.
5. Provide Open Access to your content, to allow a much wider audience to freely and immediately see the value of your project.

Once your journal has begun publishing, it is critical that you maintain a reliable publication schedule. If your journal publishes quarterly, you must produce four issues a year, even if this means fewer articles per issue. Anything less reflects a lack of professionalism. If you are unsure about your ability to sustain a quarterly journal, start out with a semi-annual or even an annual journal.

Although it sounds obvious, it is important to remember that you should only be accepting the highest quality submissions to your journal. Again, this may mean fewer articles per issue, but producing and sharing high quality scholarly information is at the heart of your project.

Measuring your impact

Another method of building professional recognition is through the use of different techniques for measuring the use and impact of your journal. This will provide clear evidence of the success of your journal and allow you to regularly evaluate your progress.

With the traditional subscription model, the number of subscribers often formed the basis for understanding the usage of an individual journal. For Open Access journals, without a subscriber base to point to, this can be a challenge. OJS, however, does provide the option of requiring readers to register. A fee is not necessarily required for this registration, but does allow the journal to develop a better understanding of its audience. Some statistical analysis and reports are available to the OJS Journal Manager, including reporting on the number of registered users.

Another way to measure usage is through the analysis of web logs. These are records produced by the server hosting your journal which counts the number of times your journal is visited. Details such as the geographic location of the reader are also available. While web logs cannot measure whether someone briefly visited your site for ten seconds or became involved in some sustained reading for an hour or more, they do provide some understanding of use, can provide comparisons to previous months or years, and can provide some data for evaluating your impact.

Perhaps more important than the question of how many people are accessing your journal is how they are using it. This is known as measuring the "impact" of your journal. Journals with high citation impact are among the most respected and successful academic journals in their fields. Citation impact refers to how often an article, an author, or a journal, is cited by other scholars. While this is not an uncontroversial means of measuring the value a journal is having in the academic community, it is the standard one that most people recognize and operate on the basis of. Readers looking for reliable information will often first look to journals with a high citation impact. Prospective authors, reviewers, and editors will be interested in volunteering their time with journals that have a high citation impact. Indexes and databases will want to include journals in their resources that have a high citation impact. And lastly, libraries will be motivated to promote journals with a high citation impact. All of this can lead to a cyclical pattern, where high impact journals are more likely to be used and supported, leading to more recognition, and higher impact. The challenge for every new journal is to get this process started.

Journal standards and identifiers

International Standard Serials Number (ISSN)

Another way of helping people find your journal, and helping libraries manage and promote it, is to obtain an International Standard Serial Number, or ISSN. An ISSN is an eight digit, "standardized international code which allows the identification of any serial publication, including electronic serials, independently of its country of publication, of its language or alphabet, of its frequency, medium, etc.– ISSNs are widely used by libraries, citation indexes, and the publishing industry to uniquely identify journals, and are often more important than the journal title itself for serials management. Every serious journal has an ISSN, and one can be obtained free of charge from a local ISSN Centre (http://www.issn.org/en/flexinode/table/1). An example of an ISSN is "1544-9173" for the journal Public Library of Science Biology. The OJS Journal Manager can enter the ISSN in the Setup of the journal.

Digital Object Identifiers (DOI)

In addition to an ISSN, you may also consider obtaining a Digital Object Identifier (DOI) for your journal. A DOI is another standardized code, allowing libraries, citation indexes, and the publishing industry to discover your content. DOIs differ from ISSNs, however, in that they only apply to electronic information and that they uniquely identify each of your articles as well as your journal. It is important to note that even if you changed your server, renamed your journal, or even moved off of the OJS system, your DOIs would not change, providing readers with a persistent URL to your content, which is important for reliable linking into course management systems, library-created article lists, and readers' electronic bibliographies. This is an important advantage over URL linking, which is notorious for changing in the online environment.

An example of a DOI is "10.1371/journal.pbio.0040176", for the article "Open Access Increases Citation Rate" in the journal Public Library of Science Biology. When the prefix "dx.doi.org/" is added to the DOI, a persistent URL is created: http://dx.doi.org/10.1371/journal.pbio.0040176 Entering this URL into your Internet browser will always take you directly to this article. Although not as common as ISSNs, DOIs are increasingly being used for advanced library services such as link resolving -also known as citation or reference linking (Link resolving provides a web-based link between a citation in an index and the fulltext of the content referred to in the citation). It is important to note that DOIs require the payment of an annual fee. Applications can be made to CrossRef (www.crossref.org), a not-for-profit network of publishers. Like the ISSN, your DOI can be added to OJS by the Journal Manager in the Setup options. You will also need to regularly submit your journal's metadata to CrossRef and determine a structure for your article-level DOIs. See the CrossRef publisher site for further details: http://www.crossref.org/02publishers/index.html. Although not as crucial as the ISSN, the benefits of using DOIs for persistent linking are worth considering. If you decide against it for your startup journal, consider revisiting this option once you have become more established. DOIs provide yet another powerful tool for connecting your content with readers.

Building reliable and ongoing content

For print journals, reliable access is helped by the production of many physical copies of the journal, and their widespread distribution. If one copy is lost or misplaced, it can easily be replaced by a copy of another. For online content, there is the danger of there being only a single copy, which is shared electronically among all readers over the Internet. If that single copy is lost, it is irreplaceable, and the results can be catastrophic for your project.

Lots of Copies Keeps Stuff Safe (LOCKSS)

Once your journal is set up on a secure and reliable server, you will still want to make sure you have an emergency preservation strategy in place. Unlike print publishing, where multiple copies are produced, distributed, and maintained by libraries, electronic journals often produce only a single electronic file (or set of files), which are accessed by multiple readers over the Internet. If this single file is lost, due to a system failure or human error, and no reliable backup exists, all of your work, and the work of your collaborators could simply disappear –permanently. To help online journal publishers overcome this potential disaster, Stanford University developed the open source LOCKSS project (Lots of Copies Keeps Stuff Safe). LOCKSS consists of geographically distributed servers maintained by libraries, the traditional experts in information storage and preservation. LOCKSS ensures that multiple copies of your content exists on a network of servers, that all of your latest content is collected and securely stored by a specialized web crawler (similar to those used by search engines), is continually examined for lost or damaged content, and makes any necessary repairs. Further details on how this system works is available on the LOCKSS website (www.lockss.org).

For publishers using OJS, the software has LOCKSS-compliance built into the system, and your journal can easily take advantage of this effective preservation strategy through the Journal Manager menu. It is one easy step toward disaster prevention for your journal.

6

Legal framework

Based on:

- Websites:

http://wiki.creativecommons.org/Legal_Concepts
http://creativecommons.org/about/licenses/meet-the-licenses
http://creativecommons.org
http://www.plosone.org/static/license.action
http://sciencecommons.org/about/
http://en.wikipedia.org/wiki/Copyleft

The public domain

Creativity and innovation rely on a rich heritage of prior intellectual endeavor. We stand on the shoulders of giants by revisiting, reusing, and transforming the ideas and works of our peers and predecessors. Digital communications promise a new explosion of this kind of collaborative creative activity. But at the same time, expanding intellectual property protection leaves fewer and fewer creative works in the "public domain" – the body of creative material unfettered by law and, to quote Supreme Court Justice Louis Brandeis, "free as the air to common use".

Until 1976, creative works were not protected by U.S. copyright law unless their authors took the trouble to publish a copyright notice along with them. Works not affixed with a notice passed into the public domain. Following legislative changes in 1976 and 1988, creative works are now automatically copyrighted. It is believed that many people would not choose this "copyright by default" if they had an easy mechanism for turning their work over to the public or exercising some but not all of their legal rights. It is Creative Commons' goal to help create such a mechanism.

Open content

The free software and open source software communities have inspired what is sometimes called "open content." Some copyright holders have made books, music, and other creative works available under licenses that give anyone permission to copy and make other uses of the works without specific permission or a royalty payment. Creative Commons hopes to build on the work of these pioneers by creating a menu of license provisions that people can combine to make their work available for copying and creative reuses.

Intellectual property conservancies

As we help people make their work available with public domain dedications and generous licenses, we will also build an "intellectual property conservancy." Like a land trust or nature preserve, the conservancy will serve to protect works of special public value from

exclusionary private ownership and from obsolescence due to neglect or technological change. We will encourage people to donate their works to Creative Commons to be held in public trust; in some cases, we may purchase important works to help guarantee both their integrity and widespread availability. Our ultimate goal is to develop a rich repository of high-quality works in a variety of media, and to promote an ethos of sharing, public education, and creative interactivity.

Creative Commons (CC) licenses

The following describes each of the six main licenses offered when you choose to publish your work with a Creative Commons license. We have listed them starting with the most restrictive license type you can choose and ending with the most accommodating license type you can choose. It's also helpful to know there are a set of baseline rights all six licenses offer to others and we've prepared a list of things to think about before choosing a license.

1. Attribution Non-commercial No Derivatives (by-nc-nd)

This license is the most restrictive of our six main licenses, allowing redistribution. This license is often called the "free advertising" license because it allows others to download your works and share them with others as long as they mention you and link back to you, but they can't change them in any way or use them commercially.

2. Attribution Non-commercial Share Alike (by-nc-sa)

This license lets others remix, tweak, and build upon your work non-commercially, as long as they credit you and license their new creations under the identical terms. Others can download and redistribute your work just like the by-nc-nd license, but they can also translate, make remixes, and produce new stories based on your work. All new work based on yours will carry the same license, so any derivatives will also be non-commercial in nature.

3. Attribution Non-commercial (by-nc)

This license lets others remix, tweak, and build upon your work non-commercially, and although their new works must also acknowledge you and be non-commercial, they don't have to license their derivative works on the same terms.

4. Attribution No Derivatives (by-nd)

This license allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to you.

5. Attribution Share Alike (by-sa)

This license lets others remix, tweak, and build upon your work even for commercial reasons, as long as they credit you and license their new creations under the identical terms. This license is often compared to open source software licenses. All new works based on yours will carry the same license, so any derivatives will also allow commercial use.

6. Attribution (by)

This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered, in terms of what others can do with your works licensed under Attribution.

Creative Commons licenses: An example

Attribution-NonCommercial-ShareAlike 3.0 Unported

CREATIVE COMMONS CORPORATION IS NOT A LAW FIRM AND DOES NOT PROVIDE LEGAL SERVICES. DISTRIBUTION OF THIS LICENSE DOES NOT CREATE AN ATTORNEY-CLIENT RELATIONSHIP.

CREATIVE COMMONS PROVIDES THIS INFORMATION ON AN "AS-IS" BASIS. CREATIVE COMMONS MAKES NO WARRANTIES REGARDING THE INFORMATION PROVIDED, AND DISCLAIMS LIABILITY FOR DAMAGES RESULTING FROM ITS USE.

License

THE WORK (AS DEFINED BELOW) IS PROVIDED UNDER THE TERMS OF THIS CREATIVE COMMONS PUBLIC LICENSE ("CCPL" OR "LICENSE"). THE WORK IS PROTECTED BY COPYRIGHT AND/OR OTHER APPLICABLE LAW. ANY USE OF THE WORK OTHER THAN AS AUTHORIZED UNDER THIS LICENSE OR COPYRIGHT LAW IS PROHIBITED.

BY EXERCISING ANY RIGHTS TO THE WORK PROVIDED HERE, YOU ACCEPT AND AGREE TO BE BOUND BY THE TERMS OF THIS LICENSE. TO THE EXTENT THIS LICENSE MAY BE CONSIDERED TO BE A CONTRACT, THE LICENSOR GRANTS YOU THE RIGHTS CONTAINED HERE IN CONSIDERATION OF YOUR ACCEPTANCE OF SUCH TERMS AND CONDITIONS.

1. Definitions

a. "Adaptation" means a work based upon the Work, or upon the Work and other pre-existing works, such as a translation, adaptation, derivative work, arrangement of music or other alterations of a literary or artistic work, or phonogram or performance and includes cinematographic adaptations or any other form in which the Work may be recast, transformed, or adapted including in any form recognizably derived from the original, except that a work that constitutes a Collection will not be considered an Adaptation for the purpose of this License. For the avoidance of doubt, where the Work is a musical work, performance or phonogram, the synchronization of the Work in timed-relation with a moving image ("synching") will be considered an Adaptation for the purpose of this License.

b. “Collection” means a collection of literary or artistic works, such as encyclopedias and anthologies, or performances, phonograms or broadcasts, or other works or subject matter other than works listed in Section 1(g) below, which, by reason of the selection and arrangement of their contents, constitute intellectual creations, in which the Work is included in its entirety in unmodified form along with one or more other contributions, each constituting separate and independent works in themselves, which together are assembled into a collective whole. A work that constitutes a Collection will not be considered an Adaptation (as defined above) for the purposes of this License.

c. “Distribute” means to make available to the public the original and copies of the Work or Adaptation, as appropriate, through sale or other transfer of ownership.

d. “License Elements” means the following high-level license attributes as selected by Licensor and indicated in the title of this License: Attribution, Noncommercial, ShareAlike.

e. “Licensor” means the individual, individuals, entity or entities that offer (s) the Work under the terms of this License.

f. “Original Author” means, in the case of a literary or artistic work, the individual, individuals, entity or entities who created the Work or if no individual or entity can be identified, the publisher; and in addition (i) in the case of a performance the actors, singers, musicians, dancers, and other persons who act, sing, deliver, declaim, play in, interpret or otherwise perform literary or artistic works or expressions of folklore; (ii) in the case of a phonogram the producer being the person or legal entity who first fixes the sounds of a performance or other sounds; and, (iii) in the case of broadcasts, the organization that transmits the broadcast.

g. “Work” means the literary and/or artistic work offered under the terms of this License including without limitation any production in the literary, scientific and artistic domain, whatever may be the mode or form of its expression including digital form, such as a book, pamphlet and other

writing; a lecture, address, sermon or other work of the same nature; a dramatic or dramatico-musical work; a choreographic work or entertainment in dumb show; a musical composition with or without words; a cinematographic work to which are assimilated works expressed by a process analogous to cinematography; a work of drawing, painting, architecture, sculpture, engraving or lithography; a photographic work to which are assimilated works expressed by a process analogous to photography; a work of applied art; an illustration, map, plan, sketch or three-dimensional work relative to geography, topography, architecture or science; a performance; a broadcast; a phonogram; a compilation of data to the extent it is protected as a copyrightable work; or a work performed by a variety or circus performer to the extent it is not otherwise considered a literary or artistic work.

h. "You" means an individual or entity exercising rights under this License who has not previously violated the terms of this License with respect to the Work, or who has received express permission from the Licensor to exercise rights under this License despite a previous violation.

i. "Publicly Perform" means to perform public recitations of the Work and to communicate to the public those public recitations, by any means or process, including by wire or wireless means or public digital performances; to make available to the public Works in such a way that members of the public may access these Works from a place and at a place individually chosen by them; to perform the Work to the public by any means or process and the communication to the public of the performances of the Work, including by public digital performance; to broadcast and rebroadcast the Work by any means including signs, sounds or images.

j. "Reproduce" means to make copies of the Work by any means including without limitation by sound or visual recordings and the right of fixation and reproducing fixations of the Work, including storage of a protected performance or phonogram in digital form or other electronic medium.

2. Fair Dealing Rights.

Nothing in this License is intended to reduce, limit, or restrict any uses free from copyright or rights arising from limitations or exceptions that are provided for in connection with the copyright protection under copyright law or other applicable laws.

3. License Grant.

Subject to the terms and conditions of this License, Licensor hereby grants You a worldwide, royalty-free, non-exclusive, perpetual (for the duration of the applicable copyright) license to exercise the rights in the Work as stated below:

a. to Reproduce the Work, to incorporate the Work into one or more Collections, and to Reproduce the Work as incorporated in the Collections;

b. to create and Reproduce Adaptations provided that any such Adaptation, including any translation in any medium, takes reasonable steps to clearly label, demarcate or otherwise identify that changes were made to the original Work. For example, a translation could be marked "The original work was translated from English to Spanish," or a modification could indicate "The original work has been modified.";

c. to Distribute and Publicly Perform the Work including as incorporated in Collections; and,

d. to Distribute and Publicly Perform Adaptations.

The above rights may be exercised in all media and formats whether now known or hereafter devised. The above rights include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. Subject to Section 8(f), all rights not expressly granted by Licensor are hereby reserved, including but not limited to the rights described in Section 4(e).

4. Restrictions. The license granted in Section 3 above is expressly made subject to and limited by the following restrictions:

a. You may Distribute or Publicly Perform the Work only under the terms of this License. You must include a copy of, or the Uniform Resource Identifier (URI) for, this License with every copy of the Work You Distribute or Publicly Perform. You may not offer or impose any terms on the Work that restrict the terms of this License or the ability of the recipient of the Work to exercise the rights granted to that recipient under the terms of the License. You may not sublicense the Work. You must keep intact all notices that refer to this License and to the disclaimer of warranties with every copy of the Work You Distribute or Publicly Perform. When You Distribute or Publicly Perform the Work, You may not impose any effective technological measures on the Work that restrict the ability of a recipient of the Work from You to exercise the rights granted to that recipient under the terms of the License. This Section 4(a) applies to the Work as incorporated in a Collection, but this does not require the Collection apart from the Work itself to be made subject to the terms of this License. If You create a Collection, upon notice from any Licensor You must, to the extent practicable, remove from the Collection any credit as required by Section 4(d), as requested. If You create an Adaptation, upon notice from any Licensor You must, to the extent practicable, remove from the Adaptation any credit as required by Section 4(d), as requested.

b. You may Distribute or Publicly Perform an Adaptation only under: (i) the terms of this License; (ii) a later version of this License with the same License Elements as this License; (iii) a Creative Commons jurisdiction license (either this or a later license version) that contains the same License Elements as this License (e.g., Attribution-NonCommercial-ShareAlike 3.0 US) ("Applicable License"). You must include a copy of, or the URI, for Applicable License with every copy of each Adaptation You Distribute or Publicly Perform. You may not offer or impose any terms on the Adaptation that restrict the terms of the Applicable License or the ability of the recipient of the Adaptation to exercise the rights granted to that recipient under the terms of the Applicable License. You must keep intact all notices that refer to the Applicable License and to the disclaimer of warranties with every copy of the Work as included in the Adaptation

You Distribute or Publicly Perform. When You Distribute or Publicly Perform the Adaptation, You may not impose any effective technological measures on the Adaptation that restrict the ability of a recipient of the Adaptation from You to exercise the rights granted to that recipient under the terms of the Applicable License. This Section 4(b) applies to the Adaptation as incorporated in a Collection, but this does not require the Collection apart from the Adaptation itself to be made subject to the terms of the Applicable License.

c. You may not exercise any of the rights granted to You in Section 3 above in any manner that is primarily intended for or directed toward commercial advantage or private monetary compensation. The exchange of the Work for other copyrighted works by means of digital file-sharing or otherwise shall not be considered to be intended for or directed toward commercial advantage or private monetary compensation, provided there is no payment of any monetary compensation in con-nection with the exchange of copyrighted works.

d. If You Distribute, or Publicly Perform the Work or any Adaptations or Collections, You must, unless a request has been made pursuant to Section 4(a), keep intact all copyright notices for the Work and provide, reasonable to the medium or means You are utilizing: (i) the name of the Original Author (or pseudonym, if applicable) if supplied, and/or if the Original Author and/or Licensor designate another party or parties (e.g., a sponsor institute, publishing entity, journal) for attribution ("Attribution Parties") in Licensor's copyright notice, terms of service or by other reasonable means, the name of such party or parties; (ii) the title of the Work if supplied; (iii) to the extent reasonably practicable, the URI, if any, that Licensor specifies to be associated with the Work, unless such URI does not refer to the copyright notice or licensing information for the Work; and, (iv) consistent with Section 3(b), in the case of an Adaptation, a credit identifying the use of the Work in the Adaptation (e.g., "French translation of the Work by Original Author," or "Screenplay based on original Work by Original Author"). The credit required by this Section 4(d) may be implemented in any reasonable manner; provided, however, that in the case of a Adaptation or Collection, at a minimum such credit will

appear, if a credit for all contributing authors of the Adaptation or Collection appears, then as part of these credits and in a manner at least as prominent as the credits for the other contributing authors. For the avoidance of doubt, You may only use the credit required by this Section for the purpose of attribution in the manner set out above and, by exercising Your rights under this License, You may not implicitly or explicitly assert or imply any connection with, sponsorship or endorsement by the Original Author, Licensor and/or Attribution Parties, as appropriate, of You or Your use of the Work, without the separate, express prior written permission of the Original Author, Licensor and/or Attribution Parties.

e. For the avoidance of doubt:

I. Non-waivable Compulsory License Schemes. In those jurisdictions in which the right to collect royalties through any statutory or compulsory licensing scheme cannot be waived, the Licensor reserves the exclusive right to collect such royalties for any exercise by You of the rights granted under this License;

II. Waivable Compulsory License Schemes. In those jurisdictions in which the right to collect royalties through any statutory or compulsory licensing scheme can be waived, the Licensor reserves the exclusive right to collect such royalties for any exercise by You of the rights granted under this License if Your exercise of such rights is for a purpose or use which is otherwise than noncommercial as permitted under Section 4(c) and otherwise waives the right to collect royalties through any statutory or compulsory licensing scheme; and,

III. Voluntary License Schemes. The Licensor reserves the right to collect royalties, whether individually or, in the event that the Licensor is a member of a collecting society that administers voluntary licensing schemes, via that society, from any exercise by You of the rights granted under this License that is for a purpose or use which is otherwise than noncommercial as permitted under Section 4(c).

f. Except as otherwise agreed in writing by the Licensor or as may be otherwise permitted by applicable law, if You Reproduce, Distribute or Publicly Perform the Work either by itself or as part of any Adaptations or Collections, You must not distort, mutilate, modify or take other derogatory action in relation to the Work which would be prejudicial to the Original Author's honor or reputation. Licensor agrees that in those jurisdictions (e.g. Japan), in which any exercise of the right granted in Section 3(b) of this License (the right to make Adaptations) would be deemed to be a distortion, mutilation, modification or other derogatory action prejudicial to the Original Author's honor and reputation, the Licensor will waive or not assert, as appropriate, this Section, to the fullest extent permitted by the applicable national law, to enable You to reasonably exercise Your right under Section 3(b) of this License (right to make Adaptations) but not otherwise.

5. Representations, Warranties and Disclaimer

UNLESS OTHERWISE MUTUALLY AGREED TO BY THE PARTIES IN WRITING AND TO THE FULLEST EXTENT PERMITTED BY APPLICABLE LAW, LICENSOR OFFERS THE WORK AS-IS AND MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND CONCERNING THE WORK, EXPRESS, IMPLIED, STATUTORY OR OTHERWISE, INCLUDING, WITHOUT LIMITATION, WARRANTIES OF TITLE, MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, NONINFRINGEMENT, OR THE ABSENCE OF LATENT OR OTHER DEFECTS, ACCURACY, OR THE PRESENCE OF ABSENCE OF ERRORS, WHETHER OR NOT DISCOVERABLE. SOME JURISDICTIONS DO NOT ALLOW THE EXCLUSION OF IMPLIED WARRANTIES, SO THIS EXCLUSION MAY NOT APPLY TO YOU.

6. Limitation on Liability.

EXCEPT TO THE EXTENT REQUIRED BY APPLICABLE LAW, IN NO EVENT WILL LICENSOR BE LIABLE TO YOU ON ANY LEGAL THEORY FOR ANY SPECIAL, INCIDENTAL, CONSEQUENTIAL, PUNITIVE OR EXEMPLARY DAMAGES ARISING OUT OF THIS LICENSE OR THE USE OF THE WORK, EVEN IF LICENSOR HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

7. Termination

a. This License and the rights granted hereunder will terminate automatically upon any breach by You of the terms of this License. Individuals or entities who have received Adaptations or Collections from You under this License, however, will not have their licenses terminated provided such individuals or entities remain in full compliance with those licenses. Sections 1, 2, 5, 6, 7, and 8 will survive any termination of this License. b. Subject to the above terms and conditions, the license granted here is perpetual (for the duration of the applicable copyright in the Work). Notwithstanding the above, Licensor reserves the right to release the Work under different license terms or to stop distributing the Work at any time; provided, however that any such election will not serve to withdraw this License (or any other license that has been, or is required to be, granted under the terms of this License), and this License will continue in full force and effect unless terminated as stated above.

8. Miscellaneous

a. Each time You Distribute or Publicly Perform the Work or a Collection, the Licensor offers to the recipient a license to the Work on the same terms and conditions as the license granted to You under this License.

b. Each time You Distribute or Publicly Perform an Adaptation, Licensor offers to the recipient a license to the original Work on the same terms and conditions as the license granted to You under this License.

c. If any provision of this License is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this License, and without further action by the parties to this agreement, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable.

d. No term or provision of this License shall be deemed waived and no breach consented to unless such waiver or consent shall be in writing and signed by the party to be charged with such waiver or consent.

e. This License constitutes the entire agreement between the parties with respect to the Work licensed here. There are no understandings, agreements or representations with respect to the Work not specified here. Licensor shall not be bound by any additional provisions that may appear in any communication from You. This License may not be modified without the mutual written agreement of the Licensor and You.

f. The rights granted under, and the subject matter referenced, in this License were drafted utilizing the terminology of the Berne Convention for the Protection of Literary and Artistic Works (as amended on September 28, 1979), the Rome Convention of 1961, the WIPO Copyright Treaty of 1996, the WIPO Performances and Phonograms Treaty of 1996 and the Universal Copyright Convention (as revised on July 24, 1971). These rights and subject matter take effect in the relevant jurisdiction in which the License terms are sought to be enforced according to the corresponding provisions of the implementation of those treaty provisions in the applicable national law. If the standard suite of rights granted under applicable copyright law includes additional rights not granted under this License, such additional rights are deemed to be included in the License; this License is not intended to restrict the license of any rights under applicable law.

Creative Commons Notice

Creative Commons is not a party to this License, and makes no warranty whatsoever in connection with the Work. Creative Commons will not be liable to You or any party on any legal theory for any damages whatsoever, including without limitation any general, special, incidental or consequential damages arising in connection to this license. Notwithstanding the foregoing two (2) sentences, if Creative Commons has expressly identified itself as the Licensor hereunder, it shall have all rights and obligations of Licensor.

Except for the limited purpose of indicating to the public that the Work is licensed under the CCPL, Creative Commons does not authorize the use by either party of the trademark “Creative Commons” or any related

trademark or logo of Creative Commons without the prior written consent of Creative Commons. Any permitted use will be in compliance with Creative Commons' then-current trademark usage guidelines, as may be published on its website or otherwise made available upon request from time to time. For the avoidance of doubt, this trademark restriction does not form part of this License.

CASE STUDY - PLoS ONE Journal: "No Permission Required"

The Public Library of Science (PLoS) ONE (eISSN-1932-6203) is an international, peer-reviewed, open-access, online publication. PLoS ONE welcomes reports on primary research from any scientific discipline.

PLoS applies the Creative Commons Attribution License (CCAL) to all works they publish (read the human-readable summary or the full license legal code). Under the CCAL, authors retain ownership of the copyright for their article, but authors allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in PLoS journals, so long as the original authors and source are cited. No permission is required from the authors or the publishers.

In most cases, appropriate attribution can be provided by simply citing the original article (e.g., Kaltenbach LS et al. (2007) Huntingtin Interacting Proteins Are Genetic Modifiers of Neurodegeneration. PLoS Genet 3(5): e82. doi:10.1371/journal.pgen.0030082). If the item you plan to reuse is not part of a published article (e.g., a featured issue image), then please indicate the originator of the work, and the volume, issue, and date of the journal in which the item appeared. For any reuse or redistribution of a work, you must also make clear the license terms under which the work was published.

This broad license was developed to facilitate Open Access to, and free use of, original

About Science Commons

Science Commons has three interlocking initiatives designed to accelerate the research cycle -the continuous production and reuse of knowledge that is at the heart of the scientific method. Together, they form the building blocks of a new collaborative infrastructure to make scientific discovery easier by design.

1. *Making scientific research "re-useful"* –We help people and organizations open and mark their research and data sets for reuse. Learn more.
2. *Enabling "one-click" access to research materials* –We help streamline the materials-transfer process so researchers can easily replicate, verify and extend research. Learn more.
3. *Integrating fragmented information sources* –We help researchers find, analyze and use data from disparate sources by marking and integrating the information with a common, computer-readable language.

All three elements have been implemented in the Neurocommons, a "proof-of-concept" project within the field of neuroscience. The Neurocommons is a beta open source knowledge management system for biomedical research that anyone can use, and anyone can build on. See: http://sciencecommons.org/projects/data/

Copyleft

Copyleft is a play on the word copyright and describes the practice of using copyright law to remove restrictions on distributing copies and modified versions of a work for others and requiring that the same freedoms be preserved in modified versions.

Copyleft is a form of licensing and may be used to modify copyrights for works such as computer software, documents, music, and art. In general, copyright law allows an author to prohibit others from reproducing, adapting, or distributing copies of the author's work. In contrast, an author may, through a copyleft licensing scheme, give every person who receives a copy of a work permission to reproduce, adapt or distribute the work as long as any resulting copies or adaptations are also bound by the same copyleft licensing scheme. A widely used and originating copyleft license is the GNU General Public License. Similar licenses are available through Creative Commons –called Share-alike.

7

Institutional Open Access policies and mandates: NIH example

Based on:

- Websites:
http://publicaccess.nih.gov
http://grants.nih.gov/grants/guide/notice-files/NOT-OD-05-022.html

The National Institutes of Health (NIH) Public Access Policy ensures that the public has access to the published results of NIH funded research. It requires scientists to submit journal final peer-reviewed manuscripts that arise from NIH funds to the digital archive PubMed Central (www.pubmedcentral.nih.gov). The Policy requires that these articles be accessible to the public on PubMed Central to help advance science and improve human health.

Figure: PubMed Central website

In accordance with Division G, Title II, Section 218 of PL 110-161 (Consolidated Appropriations Act, 2008), the NIH voluntary Public Access Policy (NOT-OD-05-022) is now mandatory. The law states:

The Director of the National Institutes of Health shall require that all investigators funded by the NIH submit or have submitted for them to the National Library of Medicine's PubMed Central an electronic version of their final, peer-reviewed manuscripts upon acceptance for publication, to

be made publicly available no later than 12 months after the official date of publication: Provided, That the NIH shall implement the public access policy in a manner consistent with copyright law.

Need for the policy

The public comments were largely supportive of the proposed policy to enhance public access to archived publications resulting from NIH-funded research. Comments noted that this policy provides equal and timely access to all via the Internet and that this accessibility should improve individual health outcomes. Many scientists appreciated that the policy would improve the visibility of their work. A large number of comments suggested that publicly-funded research publications should be made accessible to the public in full-text version in a timely manner. Many commenters expressed support for the policy given their concerns about the high and rising cost of subscriptions to scholarly journals, especially in the areas of science, technology, and medicine.

Other commenters questioned the need for the policy and considered it redundant to existing information sources and systems. Some questioned the added value of the policy and noted that journals increasingly are making full-text articles available immediately upon or within one year of publication through a variety of sources. Commenters noted that many of these articles are already linkable through The National Library of Medicine (NLM) PubMed web-based literature retrieval system that contains citations and abstracts from thousands of journals, dating back to 1950. A significant number of comments also questioned why the NLM could not simply provide a link to the publisher's website, or work with existing vendors to broaden offerings to include peer-reviewed publications not associated with NIH funding.

The primary purpose of the NIH Public Access Policy is the creation of a stable archive to ensure the permanent preservation of vital, peer-reviewed research publications resulting from NIH-funded research findings now and for future generations. While links exist to journal articles that are publicly accessible, these are not sufficient because publishers' websites are not permanently available nor consistently maintained. Additionally,

the formatting of journal articles may vary significantly among publishers' websites. The Policy addresses this deficiency in that all articles in PMC (PubMed Central), regardless of their original format, are converted into a single, explicit, and well-specified data format. This format is known as the NLM Journal Article Extensible Markup Language (XML) Document Type Definition (DTD). Further, as new needs arise, and as technology and applications change, there is a single, uniform base upon which to build.

Preservation of the biomedical literature is a responsibility that is specifically mandated in NLM's authorizing legislation, found at 42 U.S.C. 286(b)(1), and one that has successfully been carried out by the NLM since 1836. It is logical in this electronic era to expect libraries, and particularly national libraries, to continue this vital function, including keeping pace with the ever-changing technology surrounding document preservation. Updating the data formats to keep up with the changes in technology and the needs of biomedical research requires an ongoing investment in research and development, which is within the NIH mission. As the electronic article increasingly becomes the authoritative and most useful document for researchers and as scientists are actually computing on the contents of these documents - the text itself as well as the associated data - the impermanence of the publishers' websites presents a substantial risk. Creating such an archive is a historical and necessary NIH responsibility.

NIH believes that the NIH Public Access Policy will effectively advance its stated goals. By storing research publications from diverse sources in a searchable, electronic archive with a common format, PMC facilitates greater integration with related resources in other NLM databases such as DNA and protein sequences, protein structures, clinical trials, small molecules (PubChem), and taxonomy thus providing the opportunity to develop unprecedented scientific search and analysis capabilities for the benefit of science. One of the primary goals of PMC is the creation of a permanent, digital archive of journal literature, which by definition, means the full text must be deposited in PMC. This searchable archive will enable NIH program officials to manage their research portfolios more efficiently, monitor scientific productivity, and ultimately, help set research priorities.

This strategy also will enable NIH to advance its goal of creating an end-to-end, paperless grants management process. Finally, it will make the publications of NIH-funded research more accessible to and searchable for the public, health care providers, educators, and scientists.

A few commenters asked NIH to strengthen the proposed policy to make submission to PMC a requirement instead of a request. We believe that the voluntary nature of the final policy is preferable to a "one size fits all" requirement, as it permits sufficient flexibility to accommodate the needs of different stakeholders and leaves the ultimate decision in the hands of our scientific investigators who are the best to judge the scientific circumstances and the time frame under which their work may be made accessible to the public at large. It is worth clarifying that NIH does not require or expect that PMC be the sole repository for NIH-funded research publications. Others may choose to post and/or archive peer-reviewed publications resulting from NIH-funded research, subject to applicable laws or permission from any copyright holders.

Scope of the policy

The NIH Public Access Policy applies only to peer-reviewed research publications that have been supported, in whole or in part, with direct costs from NIH. Numerous comments reflected misunderstandings about the scope of the policy as it was proposed. Some comments sought to broaden the Policy to include publications from non-NIH-supported investigators, and others asked that it include publications that did not contain original research findings, e.g., book reviews.

The Policy does not apply to contributed book chapters, editorials, reviews, or conference proceedings. Although PMC does contain articles from non-NIH-supported research, the Policy is focused on final, peer-reviewed manuscripts and publications that result from research supported, in whole or in part, with direct costs from NIH.

Potential for public misunderstanding of research findings: NIH prospective

A number of comments questioned the lay public's ability to understand fully original research publications, and expressed fear that potential harm could result from misinterpretation of them.

We believe that individuals who seek to read publications concerning a particular disease, health condition, or treatment should not be denied access because of the possibility that they will misunderstand the publications. Rather, NIH encourages such individuals to become educated consumers about their health care and related research, and to consult with health care professionals for specific guidance. It is important that NIH-supported research publications be made more readily available to provide credible information and to improve public understanding of the benefits of scientific research. The public demand for credible health information is clear. About 93 million Americans searched for at least one of 16 health topics online within the past year (see: http://www.pewinternet.org/pdfs/PIP_Health_Report_July_2003.pdf). In a 2003 survey, 58 percent of Internet users said they brought information obtained from the Internet to their doctor's office (see: Cybercitizen Health 3.0 Survey, Table 10 -Manhattan Research, New York, 2003).

The NIH is strongly committed to conveying the importance of the research it funds to the public. Each NIH Institute and Center has an active staff that produces high-quality educational and informational materials on various health and research topics, many of which highlight the publications of NIH-funded researchers. Institute and Center staff, often with the assistance of third parties and patient advocacy groups, works diligently to develop, review, and disseminate these products. For example, the National Library of Medicine's consumer health site, Medline Plus (http://www.nlm.nih.gov/medlineplus/) houses extensive information on over 650 health conditions. NIH believes that these products effectively advance NIH's strong commitment to improving public health through research.

Figure: web page about the public access policy of NIH

The Policy specifically relates to original research publications. NIH needs to compile these publications into a single archive in order to manage its research portfolio better and monitor its funding choices. NIH recognizes that providing public access to this electronic archive, may also help scientists, policymakers, doctors, patients and the lay public to understand better the research that NIH funds.

Version control and quality of manuscripts

Some commenters raised concerns about potential confusion resulting from differences between the author's final manuscript within PMC and the published version of the corresponding article at journal-sponsored websites. Others questioned how corrections, retractions, and other post-publication changes will be accommodated.

Through this Policy, NIH is requesting that NIH-funded investigators submit an electronic version of the author's final manuscripts resulting from research supported, in whole or in part, with direct costs from NIH, after all changes resulting from the peer review publication process have been incorporated. A growing number of journals are currently posting final author manuscripts to provide timely access to their subscribers prior to final publication of the publisher's copy edited version. In addition, under the Policy, the final manuscript will not be made available to the public through PMC until after the copyedited version is published by the journal. Corrections and other necessary revisions of author's final

manuscripts will be accommodated. Furthermore, when publicly available, the published article on the journal-sponsored website and the author's final manuscript in PMC will be appropriately linked through PubMed. Corrections and post-publication comments referring to a publication are currently identified and linked in PubMed, and this capability will be linked to the corresponding manuscript in PMC. If publishers wish to provide PMC with the publisher's final version, this version will supersede the author's final manuscript in PMC.

Potential for acceleration of medical cures

A few commenters questioned whether the proposed policy, and enhanced access to NIH-funded publications, will facilitate scientific progress and accelerate research for medical cures.

We believe that improved access through PMC to peer-reviewed, final manuscripts of NIH-supported investigators will facilitate scientific progress because it will enable NIH to manage better its research portfolio and funding choices. The NIH encourages the sharing of ideas, data, and research findings to help accomplish its important public mission to uncover new knowledge that will lead to better health for everyone. As such, we envision that the PMC resource will have widespread and varied uses for the research community. It will create a stable, permanent, and searchable archive of peer-reviewed research publications that NIH and the public can access, without a fee, to review scientific productivity, monitor the state-of-the-science, and apply such knowledge in other ways to accelerate medical research. Greater interconnectivity and functional integration between the multiple and large research data bases (e.g., Genbank and PubChem) and an archive of NIH-funded publications has the potential to enhance research in novel ways.

Potential economic impact on journal publishers

Commenters contended that NIH had not carefully considered the potential adverse economic impact of its proposed policy on publishers, in particular, not-for-profit professional and learned societies and associations that rely on subscriptions to cover costs. The consequences of

the proposed policy for many small journals, as well as bimonthly and quarterly journals, were of particular concern to some. Concern also was raised that relative to commercial publishers, not-for-profit publishers would be more disadvantaged because they often support highly specialized areas that tend to draw greater representation by NIH-funded researchers. Others questioned the fairness of allowing publishers to continue to profit by restricting access to health-related information.

Publishing patterns vary from year to year and from one journal to another. Using 2003 data, NLM estimates that, on an annual basis, publications resulting from NIH-funded research represent approximately 10 percent of the articles in nearly 5,000 journals indexed by PubMed. In addition, for only one percent of these journals do NIH-funded articles account for more than half of the total published articles. As such, it is unlikely that scientists and libraries would use the NIH Public Access Policy as the rationale for replacing their journal subscriptions. If they did, they would be able to access only a fraction of a journal's content. It also is important to note that there are many other journal offerings, such as science news, industry information, literature reviews, job announcements, functional websites, and other time-sensitive products that bring value to the reader but are not a part of the PMC archive. Access to journal articles through the NIH archive might increase Internet traffic to those journals, by both the scientific community and the general public.

The NIH supports the current publishing process by providing its funded investigators with an estimated US$30 million annually in direct costs for publication expenses, including page and color charges and reprints. In addition, NIH provides funds, through indirect costs, to research institutions for library journal subscriptions and electronic site licenses. NIH also supports the current process by encouraging publication of NIH-supported original research in scientific journals.

NIH has made modifications to the proposed policy to provide greater flexibility to accommodate the range of business models represented by large commercial publishing houses through the smaller specialized journals of learned societies. The most significant change is to allow

authors to specify the timing of the posting for public accessibility through PMC of their final manuscript. The NIH intends to maintain its dialogue with publishers and professional and learned societies as experience is gained with the Policy.

A NIH Public Access Advisory Working Group of the NLM Board of Regents will be established. The Working Group will be composed of stake-holders that will advise NIH/NLM on implementation and assess progress in meeting the goals of the NIH Public Access Policy. Once the system is operational, modifications and enhancements will be made as needed with the Working Group, or a permanent subcommittee of the Board, providing ongoing advice on improvements.

Potential impact on journal peer review

NIH recognizes the enormous value and critical role that peer-reviewed journals play in the scientific quality control process. Only peer-reviewed articles accepted for publication will be posted in PMC. Some commenters asked if scientific integrity would be compromised if journals were to go out of business, thus significantly narrowing journal options for authors. A few commenters feared that the NIH proposed policy would limit an author's freedom to publish how, when, and where he or she chooses.

We do not believe that the Policy will compromise scientific integrity or significantly narrow journal options for authors. While NIH encourages investigators to publish and share the results of the research that it funds, NIH does not dictate the means of publishing the research it supports. This Policy is designed to preserve the critical role of journals and publishers in peer review, editing, and scientific quality control processes. It is not intended to alter in any way the manuscript submission process, investigator choice of journal for publication, or existing publication process.

NIH highly values traditional routes of research information dissemination through publication in scientific, peer-reviewed journals. Peer review is a hallmark of quality for journals and is vital for validating the accuracy and

interpretation of research results. Publication in peer-reviewed journals is a major factor in determining the professional standing of scientists; institutions use publication in peer-reviewed journals in making hiring, promotion, and tenure decisions. NIH also values the communities of research created by scientific organizations and the journals they publish. By not mandating but instead requesting from our investigators that access be provided to the public within a range of acceptable delays extending from 0 to 12 months, the NIH believes that its Public Access Policy addresses the concerns raised by both for-profit and not-for-profit publishers and will ensure that peer review of scientific articles is preserved. The NIH believes that archiving and making publicly accessible NIH-funded biomedical and behavioral literature after a reasonable time delay can preserve the critical role of journals and publishers in peer review, editing, and scientific quality control. The policy should have no effect on the author's choice of journal. We expect that greater access to research publications will increase the impact of the publicly-funded research. For example, there is emerging evidence that easier access increases impact as measured by the number of times a paper is cited (see: http://opcit.eprints.org/oacitation-biblio.html).

Potential impact on scientists

A number of comments expressed the concern that researchers would be adversely affected by the proposed policy if publishers experienced a decline in subscriptions and subsequently chose to increase charges to authors. It was suggested that higher charges would disadvantage disproportionately researchers with more limited resources. In addition, some researchers were concerned that the proposed policy would create an additional burden on them.

NIH-funded investigators are expected to make the results and accomplishments of their activities available to the research community and to the public at large. Consequently, NIH considers publication costs, which include fees charged by a publisher, such as color and page charges, or fees for digital distribution, to be allowable charges to NIH research awards.

Concerning burden, public access submissions will provide NIH-supported investigators with an alternate means by which they can meet and fulfill the current requirement to provide a copy of each publication in their progress reports and other application and close-out procedures. It is anticipated that investigators applying for new and competing renewal support from the NIH will utilize this resource by providing links in their applications to their PMC -archived information. NIH, therefore, anticipates that this process may reduce, rather than increase, burden for investigators.

It is also worth noting that the development of a searchable archive of published findings from NIH-supported research will be a rich resource for all scientists. Access to such information not only will make it easier to investigate a specific area of research, but also may lead to identification of new research questions.

Open Access publication and the NIH Public Access Policy

Some commenters believed that the NIH Public Access Policy constitutes an Open Access model of publishing. The NIH Policy is not a form of publishing; rather, it creates a stable archive of peer-reviewed research publications resulting from NIH-funded research. In addition, the Policy does not dictate the means of publishing but is compatible with any publishing model that authors and journals choose to employ. For example, some subscription journals already allow free electronic access to published manuscripts directly from their websites after an embargo period. In addition, one survey reports as many as 92 percent of journals allow authors to self-archive either a postprint (79 percent) or preprint (13 percent) of the article on personal websites or on their institution's website (see: http://romeo.eprints.org/stats.php). Copyright to all material deposited in PMC remains with the publisher, individual authors, or awardees, as applicable. PMC currently includes a copyright notice alerting the public to the rights of copyright holders and will continue to post this notice as it has done in the past.

8

Economic models for journal publishing

Based on articles published by:

- K. Stranack, "Starting a New Scholarly Journal in Africa", Public Knowledge Project, 2006.
Available at: http://pkp.sfu.ca/files/AfricaNewJournal.pdf

To determine the best economic model for your new journal, you will need to look at the benefits and challenges offered by the subscription, Open Access, and limited Open Access models. Every journal is different, and what may lead to success for one journal may not be appropriate for yours'. Carefully examining your own situation will help you make the best decision for your new publication. For most new journals, however, the global audience made available through Open Access publishing, and the opportunities for alternative sources of funding, make it an option worth serious consideration.

Subscriptions have long been an important means of financially supporting journal publishing. Readers may access the journal content through a personal subscription, or through a subscribing library or research institution. In either case, the cost of the subscription permits access to the content of the journal. The cost of a journal subscription can, however, act as a barrier between a journal and its potential audience.

Open Access publishing is a new approach to distributing journal content, replacing subscription-based access with universal, online availability. Instead of readers paying for their own subscriptions, or having their library subscribe, the content can be read free of charge. Without any subscription income, however, Open Access journals must have a sustainable source of funding to ensure the long-term viability of their publication.

Subscription-based journals

Benefits

The primary reason for developing a subscription-based journal is to take advantage of the income produced by your readers or their libraries. This income can help to ensure the long-term sustainability of a journal, by providing a reliable source of funding. Another benefit of the subscription model is that it is the traditional method for producing a journal, and may be better understood by your collaborators and supporting institution. This level of familiarity may result in a great level of comfort in participating in your project.

Finally, a large and growing list of subscribers can be used to demonstrate the success of the journal, revealing the existence of people and institutions with enough interest in your publication to pay to for it.

Challenges

Despite the attractiveness of adopting a familiar model that may be able to generate a steady source of income for your journal, there are some important drawbacks to the subscription model that should be kept in mind. Perhaps the most important is that charging for your journal will limit the number of readers that you will be able to reach. Many people or institutions that would be interesting in reading your publication may simply be unable to afford another subscription, and will pass your journal by. You will want to think carefully about the audience you need to reach, and what their resources will be for accessing your content.

A further consideration often overlooked by new journals is the cost of managing a subscription-based journal. Requests for subscriptions must be processed, payments must be managed, accounts must be tracked, contact addresses must be kept current, renewals notices must be sent out, overdue payments must be pursued, and many other associated tasks can all bite deeply into the income generated by the subscriptions themselves!

Open Access journals

Open Access publishing has grown steadily in the past number of years, in response to the skyrocketing prices of many traditional journal subscriptions. The increasing cost of academic journals undermines the ability of scholars from less wealthy regions or institutions from accessing the information and knowledge required to conduct their own research. Open Access is an important publishing alternative developed to help solve this problem.

Benefits

The most important benefit of making your new journal Open Access is the connection you will instantly have to readers around the world. A further advantage to the Open Access model is the elimination of the need for the time consuming and costly subscription management functions

outlined above. Without subscriptions to register or payments to collect, more time can be dedicated to producing the high-quality content critical for your success.

While Open Access journals do not generate revenue from subscriptions, many alternative sources of income are available. Some journals receive funding in exchange for advertising on their web site (see, e.g., Section on "Funding Scientific Open Access"). Other Open Access journals require payment from their authors for article submissions, to help offset the costs of publication. Government subsidies may also be available to assist with the cost of producing freely available information. International funding agencies are also becoming increasingly interested in supporting Open Access journals, which contribute to the free exchange of scholarly information globally. One example of international collaboration is Bioline International (www.bioline.org.br), an online, Open Access publishing service operated by the University of Toronto Libraries in Canada, the Reference Center on Environmental Information in Brazil, and Bioline UK in the United Kingdom.

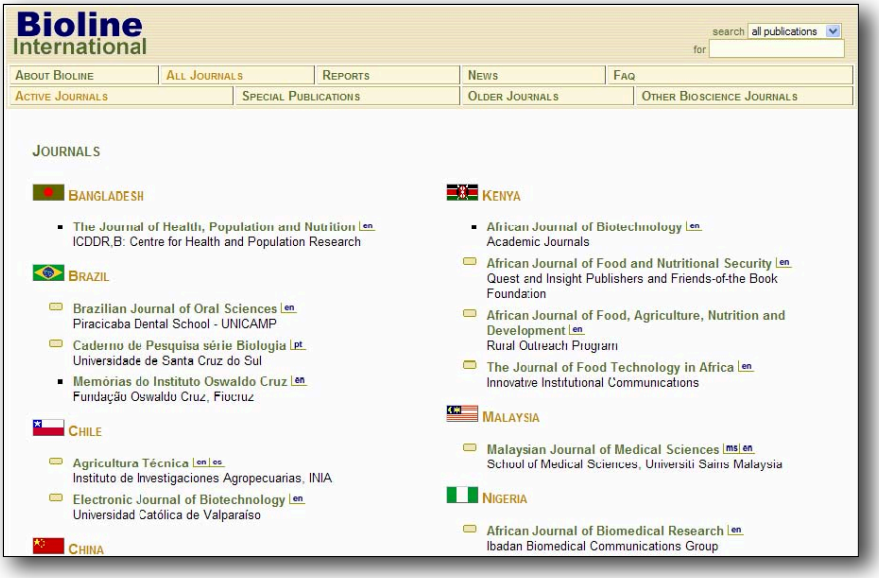


Figure: the website of Bioline International

These three organizations have come together to support over 50 Open Access journals from more than 15 countries in the developing world.

Challenges

The main challenge faced by Open Access journals is finding innovative ways to operate in the absence of any subscription income. Some examples of alternative funding sources were described above, but this remains a crucial decision for any journal to make. One of the advantages of publishing software such as the free OJS - Open Journals System, is that they can significantly streamline the entire publishing process, to help reduce costs to an absolute minimum. Another potential challenge for a new Open Access journal could be resistance from some members of the scholarly community or supporting institutions, which may be unfamiliar with Open Access as a viable alternative to the traditional economic model. Careful planning and research into additional sources of sustainable funding may be required to reassure these important stakeholders. In addition, be sure to point out that large and influential indexing services such as Thomson and Elsevier are now accepting Open Access content into their Web of Science (http://scientific.thomson.com) and Scopus products (http://info.scopus.com), lending considerable legitimacy to the Open Access option.

Limited Open Access journals

Benefits

For some journals, interested in Open Access publishing, but not quite prepared for abandoning the income from subscriptions, the option exists for providing limited Open Access. Limited Open Access restricts a portion of the journal's content (often the most current issue or two), but makes the remainder (the back issues or archives) freely available. This provides a compromise, which allows for a continued source of subscription income, but also opens the content to a wider audience.

Challenges

Although limited Open Access may appear to provide the best of both worlds for new journals looking to maximize the benefits of both Open Access and subscription-based publishing, some important limitations

should be considered. First, the restrictions placed on a portion of the content may discourage many in your potential audience, forcing them to wait months for access to a particular article they need immediately. In this way, limited Open Access may still not be open enough for all of your readers.

Second, by providing Open Access to a portion of your content, such as the back issues of your journal, you may be undermining the very subscriber base you hope to maintain. Your subscribers may be less willing to pay if the content of your journal will be freely available in few months. As attractive as it may seem, in some cases the limited Open Access model can result in the worst of both worlds.

9

Funding scientific Open Access

Based on:

- *E. Canessa, M. Zennaro and C. Fonda, "Information for Development Magazine" (i4d), Jun 2006.*
  Website: www.i4donline.net

- *P. Suber, "SPARC Open Access Newsletter", issue #94, Feb 2006.*
  Website: http://www.earlham.edu/~peters/fos/newsletter/02-02-06.htm

- *SCOAP³ Website: www.scoap3.org*

- *SCOAP³ Working Party, CERN-OPEN-2007-009 Report, Mar 2007.*
  Website: http://cdsweb.cern.ch/record/1020110

Open Access Journals need some source of revenue or subsidy to have a chance at viability. Those with two or more sources have a greater chance. Plural revenue streams not only bring in more money, but insulate journals against fluctuations in one of their revenue streams. If OA journals charge author-side fees (and some do), then plural revenue sources let them lower their fees and increase the number of fee-waivers they grant, and thereby increase the number of submissions and the quality of accepted articles.

Financial sustainability via advertising: A proposal

To finance all costs to run a scientific electronic Journal, there is reason why to sell advertising space within an OA Journal by associating ads to given keywords within every article.

OA Journals are more expensive to sustain financially than OA Archives due to higher running costs, such as referee process, editorial and secretarial board, article processing, etc. Advertising can communicate an effective message to the targeted scientific audience.

Each article could carry an advertisement within its layout related to its contents (not generic Ads). The ads can be dynamically created advertising some related products, events like conferences, instruments, scientific services etc. A charge is to be levied to the advertiser and not to the author or its institution. The payment received from these targeted ads could well be used to fund a Physics OA Journal. To the best of our knowledge such an approach has not yet been fully adopted by the available on-line scientific scholarly e-Journals (www.doaj.org). Multiple benefits is meant for all. Authors can get the possibility to publish their results for free after the referee and editorial acceptances. In turn, institutions may reduce their expenses, publishers of OA Journals can make profits, and advertisers themselves can reach a bigger scientific market, bringing readers to their websites.

Target specific advertisement

To implement these ideas, it is necessary to have an appealing way to get advertisers to pay for the advertising of single scientific articles. To set out,

how to get advertisers, is in the nature of doing science itself. Advertising to a scientific audience, which is notoriously exclusive and therefore difficult to reach, will persuade advertisers on the benefits from placing their ads. Most companies not only buy ads for competitive reasons, but also to communicate something. The later implies to potential product or to let scientists know about new services provided for doing research. Even more important advertisers can advertise to help a community of scientists to save some money.

Advertised OA Journals offers to companies more exposure and most cost effective forms of (targeted) advertising to the right audience. For example, in doing experiments and measurements, trusted information sources are needed, as well as updated information and technical advice on the right scientific instruments and products. Scientists are not random, but very specialised readers. Thus the advertising must be selected for their relevance to each published article. The number of advertisements on a scientific hot topic may guide to set the prices of advertisements.

Ads by Google

Serving ads that are relevant to what is being read on-line is not new. “Ads by Google” (http://services.google.com) is an example of this implementation. Just as a search on Google connects to the information one is looking for, “Ads by Google” provides links to products and services that are relevant to the search results. It uses technologies that grasp the nuances of language and closely matches or targets ads to the specific content of web pages. “Ads by Google” has been adopted by some OA Journals: the Free Software Magazine (www.freesoftwaremagazine.com) in the technical area and AmericanScience (www.americanscience.org) in the scientific field.

Google AdSense ads for Open Access journals

For Journals that already have a source of revenue or subsidy, Google AdSense ads can provide a critical back-up (http://www.google.com/adsense/).

- Journals pay nothing for them. There is a trivial cost in placing them on the page, but it only takes some standard HTML, which could easily become i part of a Journal's template. They take a few minutes of time, once, not an advertising staff or marketing department.

  The amount of money they generate is a function of how many readers click on them, which depends on how many view them, which obviously varies hugely across journals and even across articles within the same journal. But some is better than none, and even a trickle is all gain when AdSense ads are essentially costless to use.

- The journal doesn't know in advance what ads Google will select for a given page, only that the ads will be keyword-relevant to that page. Google's algorithm decides what ads to put on a given page based on the keywords that Google's crawler detects on the page and the Google AdSense customers that have signed up to be associated with those keywords. The ads picked for a given page might change from day to day.

  Journals cannot warp their objectivity in order to ingratiate advertisers if they don't know, at the time their editorial decisions are made, whom to ingratiate. Journals that really want to sell their souls to advertisers would not know how, at least not without shifting back from Google ads to traditional ads.

  AdSense ads won't compromise editorial decisions or peer review, and for readers who understand how they work, they won't even present the appearance of a conflict of interest. Of course, not all readers will understand how they work and some will object that the ads threaten the Journal's objectivity. But if these objections arise, they are much

easier to answer than similar objections to conventional ads. AdSense ads are immeasurably less compromising than conventional ads from known companies solicited, cultivated, or negotiated by journal staffers. While journals don't pick advertisers or ads, Google does give them the option to block ads of certain kinds.

- Google ads are much more likely to fit the interests of the readers of a given page than conventional ads. Because they are selected for their keyword relevance, they are narrowcast to readers of a given page, not broadcast to readers of varying interests. The better the fit between ad and reader, the less readers will complain about the presence of the ads and the more they will click-through, increasing the revenue for the journal.

- Readers can help the journals they read without making donations, simply by clicking through on some of its ads.

In principle OA repositories could use AdSense ads as readily as journals, though repositories could probably only place ads on general pages, not on author-deposited article pages. On the whole, the benefits for OA Journals are the same as for non-OA journals. While OA Journals are usually in greater need of an additional source of revenue, non-OA Journals that add revenue through AdSense ads could cut prices or increase their OA experiments.

A free fully-hosted Open Journal systems platform

Scholarly Exchange (www.scholarlyexchange.org) offers its free platform for OA journals financed in part through Google ads. Start-up is easy with the free use of a fully hosted and ready-to-use installation of the open-source Open Journal Systems (OJS developed by the Public Knowledge Project) for the first year.

Scholarly Exchange (SE) uses the chosen domain name for the journal and performs secure daily backups to two separate sites. The platform is complete, offering web-based editorial management (manuscript collection, peer review, editorial decision-making) and online display.

SE offsets its costs with contextually appropriate advertising and shares excess revenue with participating journals. It provides the OJS software and use of the Scholarly Exchange service at no cost for the first year in conjunction with relevant on-screen advertising (currently supplied by Google) and a reader-donation 'button'. SE retains the advertising revenues to defray costs. Second and subsequent years: SE continues the service for an annual fee, with all advertising revenue split 50/50 between the Journal and SE.

SCOAP^3

Sponsoring Consortium for Open Access Publishing in Particle Physics

Open Access tenets of granting unrestricted access to the results of publicly-funded research are in contrast with current models of scientific publishing, where access is restricted to journal customers. At the same time, subscription costs increase and put considerable strain on libraries, forcing them to cancel an increasing number of journals subscriptions. This situation is particularly acute in fields like High-Energy Physics (HEP), where pre-prints describing scientific results are timely available online. There is a growing concern within the academic community that the future of high-quality journals, and the peer-review system they administer, is at risk.

To address this situation for HEP and, as an experiment, Science at large, a new model for OA publishing has emerged: SCOAP^3 (Sponsoring Consortium for Open Access Publishing in Particle Physics, see www.scoap3.org), headed by CERN (European Organization for Nuclear Research, www.cern.ch). In this model, HEP funding agencies and libraries, which today purchase journal subscriptions to implicitly support the peer-review service, federate to explicitly cover its cost, while publishers make the electronic versions of their journals free to read. Authors are not directly charged to publish their articles OA.

SCOAP^3 will, for the first time, link quality and price, stimulating competition and enabling considerable medium- and long-term savings.

Today, most publishers quote a price in the range of 1,000-2,000 € per published article. On this basis, it is estimated that the annual budget for the transition of HEP publishing to OA would amount to a maximum of 10 M€/year, sensibly lower than the estimated global expenditure in subscription to HEP journals.

Each SCOAP^3 partner will finance its contribution by canceling journal subscriptions. Each country will contribute according to its share of HEP publishing. The transition to OA will be facilitated by the fact that the large majority of HEP articles are published in just six peer-reviewed journals. Of course, the SCOAP^3 model is open to any, present or future, high-quality HEP journal aiming at a dynamic market with healthy competition and broader choice.

HEP funding agencies and libraries are currently signing "Expressions of Interest" for the financial backing of the consortium. A tendering procedure will then take place. Provided that SCOAP^3 funding partners are prepared to engage in long-term commitments, many publishers are expected to be ready to enter into negotiations.

The example of SCOAP^3 could be rapidly followed by other fields, directly related to HEP, such as nuclear physics or astro-particle physics, also similarly compact and organized with a reasonable number of journals.

Benefits of SCOAP^3

The SCOAP^3 model will initiate a significant shift of paradigm for the dissemination of results from scientific research, with new benefits and clearly defined roles for all stake-holders in the publication process:

- Readers will benefit from unrestricted access to all relevant literature in their field of research.

- Authors will benefit from a wider dissemination of their results, thus from better opportunities for recognition and career evolution. Their transition to OA will be transparent: they can continue to publish in the

same journals as before. However, the increased visibility of their results will be a strong incentive to give preference to OA journals.

- Publishers will benefit from a more sustainable business model than the traditional subscription scheme, becoming increasingly fragile in the Internet era. Their prime responsibility will be to ensure quality of the highest standards through independent editorial boards and peer review. Publishers will also continue to meet demands for print subscriptions, reprints, color plates, and other premium services, outside the scope of SCOAP[3].

- Funding agencies will profit from increased visibility of their research results in high-quality OA journals. They will benefit from improved stability of publication costs and possible long-term savings generated by a competitive publication market.

- Libraries will benefit from solving the problem of spiraling subscription costs of HEP journals. Further, access to the published literature will be offered through library gateways without barriers.

10

Getting found, staying found, increasing impact

Based on:

- *K. Stranack, "Getting Found, Staying Found, Increasing Impact", Public Knowledge Project, 2006.*
  Available at: http://pkp.sfu.ca/files/GettingFoundStayingFound.pdf

- *K. Stranack, "Starting a New Scholarly Journal in Africa", Public Knowledge Project, 2006.*
  Available at: http://pkp.sfu.ca/files/AfricaNewJournal.pdf

Getting found: Building the visibility of your journal

The success of your journal depends on developing a regular readership, who will become part of your scholarly community, cite your content in their own work, and tell others about the value of your publication. To do this, however, they will first need to be able to find you. This chapter examines a variety of ways to increase the 'findability' of your journal, through the use of commercial indexes, open databases, libraries, the media, professional networks, and professional recognition.

What are commercial indexes?

Commercial indexes are collections of journal citation details (such as author names, article title, journal title, volume and issue numbers, abstracts, etc. – also known as "metadata") maintained in a central, searchable database. As commercial services, these indexes are only available with a paid subscription, and are often accessed by readers through their library. A significant portion of any academic library budget goes toward making these commercial products freely available to their faculty and students. One of the most influential indexes is Thomson Scientific's Web of Science (http://scientific.thomson.com).

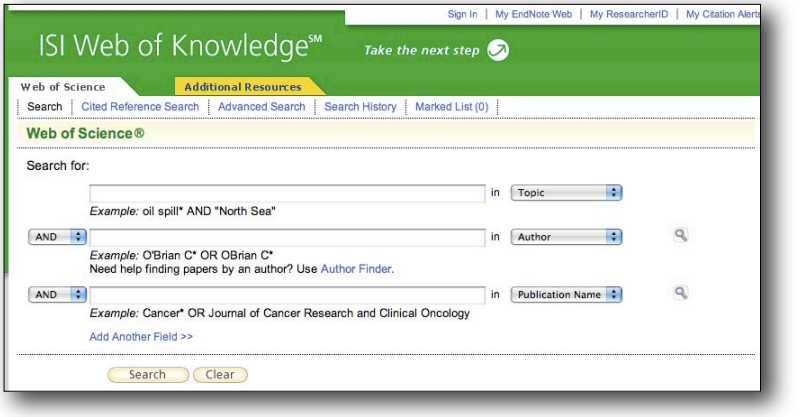


Figure: Web of Science

Some indexes may be focused on a single discipline, such as PsycInfo (http://www.apa.org/psycinfo/) for psychology, while others are

multidisciplinary (Elsevier's Scopus - www.scopus.com). Some combine information from hundreds of journals, and others may only include the metadata from a few. Some indexes are produced by scholarly societies or non-profit organizations and others by for-profit businesses. Commercial indexes are often the most important way for readers to find your content, and getting included in one or more of them is important for your journal's success.

How to get indexed

To take advantage of these powerful research tools, and "be recognized as an authoritative, high-quality source of information" ("Getting Your Journal Indexed"), you will need to have your journal's citation metadata included in the most appropriate indexes for your discipline. Each index will have its own set of criteria for inclusion, but it generally includes:

- High quality content
- Peer-reviewed
- Compatible subject matter (for indexes focused on a single discipline)
- Recognized editorial board
- Stable publishing history

For a new journal, it is impossible to demonstrate a history of stable publishing, and may make getting into many of the established, commercial indexes challenging for your first few years. The next section will look at some of the available alternatives while you build your journal's professional recognition.

What are open databases?

Open databases are similar to commercial indexes, in that they aggregate citation metadata into a single, searchable database or listing. The main types of open databases include open indexes, directories, search engines, and open archive metadata harvesters. One of the principle advantages of open databases is that they are freely available on the Internet for anyone to use, including individual readers and libraries.

Many open databases will also be more willing to include content from new journals, placing more emphasis on the quality of your content and your Open Access policy than on a large archive of published material. In addition, your content can often be included more quickly in open databases.

Open databases are becoming increasingly important to researchers. While they may not yet have the same prestige or influence as some of the commercial indexes, becoming part of one or more of them will significantly raise your journal's profile with a wider audience of readers.

Open indexes

PubMed Central (www.pubmedcentral.nih.gov) is one of the most recognized and respected open indexes. As with commercial indexes, PubMed Central collects metadata from different journals (all in the field of medicine, of course) and combines them into a single, searchable database.

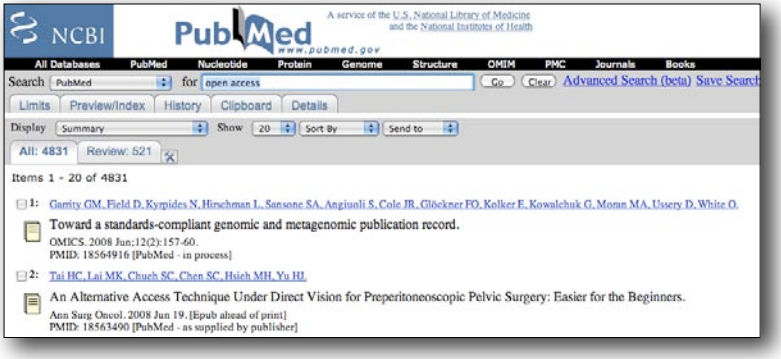


Figure: PubMed Central

The main difference is that open indexes are often publicly-funded and made freely available. OJS medical journals should certainly work to get included in PubMed Central. To facilitate this, OJS includes an exporting tool, which will produce a file of all of your journal's metadata, suitable for sending directly to PubMed. Some open indexes, such as BioMed Central or Chemistry Central, only include their own published content.

For journals not published by BioMed Central or Chemistry Central, it is not an option. Examples of other open indexes include Agricola (http://agricola.nal.usda.gov/), from the U.S. National Agriculture Library and ERIC (www.eric.ed.gov) sponsored by the U.S. Department of Education, Institute of Education Sciences.

How to get indexed

Like commercial indexes, open indexes are also looking for high quality content, peer review, compatible subject matter, and evidence of stability and sustainability. Some, however, may be willing to accept submissions from new journals lacking an established history of publication. If you do not know the best open indexes for your journal, contact your library. They will be able to guide you in the appropriate direction.

Directories

Directories are primarily lists compiled by individuals or organizations, organized into subject areas, and made freely available on the Internet. Often directories do not offer any article-level searching capabilities, but instead are simply browsed. Like commercial and open indexes, directories are popular with libraries, because they are highly useful tools for connecting readers to appropriate content.

Examples of open directories include the Directory of Open Access Journals (www.doaj.org), NewJour (http://gort.ucsd.edu/newjour), the Librarians Index to the Internet (http://lii.org), and the Open Directory Project (http://dmoz.org).

Yahoo (http://dir.yahoo.com) was one of the original Internet directories, but has increasingly moved toward becoming a search engine. It does, however, continue to also be a directory.

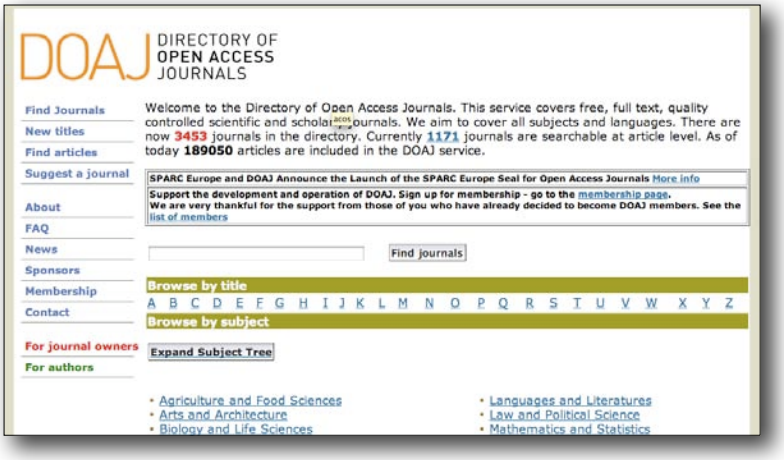


Figure: An example of open directory - DOAJ

Contact the organizer of the directory to get your journal included. The criteria for inclusion will be similar to commercial and open indexes, but may be somewhat less demanding and more immediate. Directories are an excellent way for new journals to raise their profile, both with readers and with libraries.

Contact your local library to get included in any lists that they maintain and to find out about other directories that might be appropriate for your journal.

Search engines

Search engines are tools used to help people find information on the World Wide Web. Using programs known as 'spiders', which 'crawl' the 'web', search engines accumulate descriptive metadata and keywords from the fulltext of various web sites and build massive, searchable indexes. Some search engines, such as Google (www.google.com) attempt to cover the entire web, while others focus on a particular kind of content. Scirus (www.scirus.com), for example, only indexes scientific web sites and Google Scholar (http://scholar.google.com) only covers academic information. Microsoft's Live Academic (http://academic.live.com)

provides a similar service. Your library will be able to point you toward the best search engines for your journal to work with.

Figure: Google Scholar

Although no where near as sophisticated or focused as the resources previously discussed, search engines are becoming the first research choice not only for students, but for an increasing number of scholars as well. Getting effectively indexed in a number of search engines will go a long way toward widening your audience.

One of the main challenges for your journal is to ensure your site is 'crawled' and indexed by the major search engines, and to make sure the indexing terms collected will be useful for helping people connect to your content. With many search engines returning hundreds of thousands of results for almost any search, you want your journal to appear as high as possible in the result list. Searchers rarely look beyond the top ten results, so improving your ranking is very important. Each search engine will have its own unique system of relevance ranking. Google Scholar, for example, "aims to sort articles the way researchers do, weighing the full text of each article, the author, the publication in which the article appears, and how often the piece has been cited in other scholarly literature". Some important suggestions to improve your journal's position in a variety of

search engines includes the proper use of metatags, positioning of keywords, and actively submitting your site for search engine indexing. Valuable details on these tips and more are available from Search Engine Watch (http://searchenginewatch.com/webmasters/). Help for indexing in specific search engines is also available, such as Google's Webmaster Central (http://www.google.com/webmasters) and their support page for scholarly publishers (http://scholar.google.com/intl/en/scholar/publishers.html).

Open Archive metadata harvesters

As the name implies, open archive metadata harvesters are tools which collect freely available metadata, and aggregate it into a single, searchable, open database. Usually operated by universities and their libraries, harvesters generally gather metadata from institutional repositories and participating journals. Examples include OAIster (http://oaister.umdl.umich.edu) from the University of Michigan, the Canadian Association of Research Library harvester (http://carl-abrc-oai.lib.sfu.ca), and the PKP harvester (http://pkp.sfu.ca/harvester2).

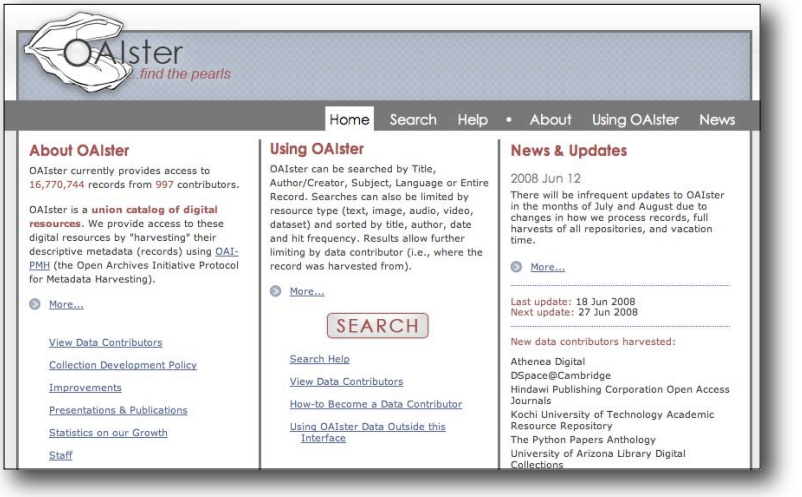


Figure: OAIster

Harvesters are valuable for scholars as open resources of high quality research material. The easiest way to get indexed into one of these tools is to submit your metadata to the PKP Harvester. This is built right into your OJS software.

Libraries

One of the fundamental objectives of libraries is to connect readers to high quality content. As a journal publisher, you will want to develop a relationship with your local library, to ensure they know about your valuable work, include your journal in their collection, and promote it to their users.

In the traditional print environment, libraries subscribe to a journal and add each physical issue to their collection. Readers visit the library, browse the journal shelves, and make choices based on what is available to them. In the online environment, with literally thousands of electronic journals that come and go, one of the challenges is getting your journal recognized and included in the library's 'virtual' collection. The virtual collection is usually made available through the library's online catalogue or electronic journal database.

Figure: Electronic Journal Database

Online catalogues and journal databases allow readers to browse or search for appropriate content for their research, and allow for direct linking (to your journal or your articles) into online educational resources such as course web pages, or course management systems, including WebCT (www.webct.com) or Moodle (http://moodle.org). The most

straightforward way to get your journal included in a library's virtual collection is to simply ask, letting them know about your project, and the audience you are trying to reach. Libraries are very open to collaboration and would welcome your request and the information you could provide.

Another way to ensure you are part of many library collections is to become part of a larger journal entity, such as one of the open directories mentioned earlier. Most libraries provide links to the DOAJ, for example, and the DOAJ provides libraries with a list of titles to download and repurpose on their own web sites. Participating in any local metadata harvesting initiatives (such as the Synergies and Erudit projects in Canada, www.synergies.umontreal.ca), institutional repositories, or directory projects could also help your journal get recognized as part of a library's virtual collection.

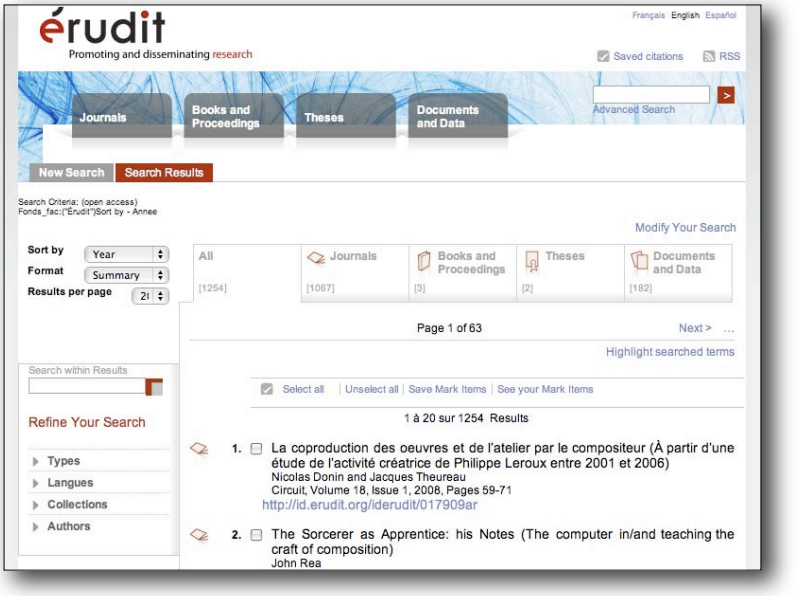


Figure: www.erudit.org

Another reason to work with your local library is to access information you will need to widen the reach of your journal. In the previous sections, this booklet outlined many ways in which libraries can help you find out

about indexes, directories, repositories, harvesters, and other tools and services that can let prospective readers find you more easily.

The media

Communicating with the media is another important way of getting the word out about your journal. A press release is "a written or recorded communication directed at members of the news media for the purpose of announcing something claimed as having news value". Writing an effective and professional press release will help the media to understand your message, putting it into a format they can use. To be effective, a press release should be concise and have a clear focus. The majority of press releases are 500 words of text organized into four or five, short two to three sentence paragraphs. All of the essential information (who, what, where, when, and why) should appear in the first paragraph. It should also follow an established format, including:

1. Add the words "FOR IMMEDIATE RELEASE" in the top left-hand margin in all caps. Follow this line with relevant contact information: name, title, address, phone number, e-mail address. If you have a logo for your journal, include it as a letterhead.
2. Create a headline and center it in bold type just above the first line of the body of the press release. Headlines typically highlight the most significant fact in the release.
3. Include a dateline - the first line of the body of your press release – with the city where the release is generated and the date.
4. Use block style, without paragraph indentations.
5. The last paragraph should end with a "for additional information" line, indicating a location to find more details. Your web site is an obvious source of information for this.
6. Finally, centre these marks: "# # #" or "-30-", at the bottom of the page to indicate the end of your release.

How to distribute a press release

To most effectively distribute your press release, you will want to target the most appropriate media for your message. For example, a press release announcing the launch of your journal would be welcomed by other related professional or scholarly publications, but would probably not be used by the popular media (television, radio, magazines, or newspapers). However, if you are publicizing a major new discovery or research finds on a topic of more general interest, the popular media may be interested. Targeting your press releases is the key to the successful distribution of your press releases. Once again, be sure to make use of your library to find out about all of the potential media outlets for your press releases.

11

Web 2.0 and Open Access

Based on:

- Ismael Peña-López, " The Personal Research Portal: web 2.0 driven individual commitment with Open Access for Development", 2007, Knowledge Management for Development Journal 3 (2007) 35.
Website: http://www.km4dev.org/journal

- Website:
http://www.ictlogy.net

Open Access can be considered a way to achieve universal reach of research diffusion at inexpensive and immediate levels. Most OA efforts have been aimed at the institutional level, devoting little energy to what the individual can do to contribute to this goal. Even though there are some valid reasons for this imbalance, there is ample opportunity for the individual to make a difference.

This chapter aims to explore how individuals can contribute to the diffusion of research in the OA paradigm by means of social software and web 2.0 technologies. The example of the Personal Research Portal (PRP) "a concept more than an artifact" can contribute to making knowledge more accessible to researchers in developing and developed countries alike, but also provides a model by which international research networks might be fostered. In detail, it analyzes how the PRP can contribute to creating an "online identity", how this identity can help to create a network and how digital publishing is the currency of this network.

The personal research portal (PRP)

One problem facing developing nations is that "access to high-quality research information has historically ranged from being extremely limited to altogether non- existent"; even so, this does not mean that its citizens do not produce high-quality research information themselves and that they cannot act individually to close "the existing gap between those countries that have ample access to electronic research information and those that do not". This can be a complementary trajectory to initiatives emerging at the institutional level.

The approach presented here is closely related to the concept of 'e-portfolio':

An e-portfolio is a digitized collection of artifacts, including demonstrations, resources, and accomplishments that represent an individual […]. This collection can be comprised of text-based, graphic, or multimedia elements archived on a Web site or on other electronic media.

E-portfolios are usually associated with students and teaching rather than with researchers, their main goal being for students to gather and present their work for assessment; therefore the term 'personal research portal' (PRP) is introduced here as an alternative, whereby its main goal is to act as a knowledge 'gatherer', contributing to

1. Enhanced access to international research output;
2. Access to research generated in Developing Countries;
3. Promotion of institutional research output;
4. Improved citation and research impact;
5. Improved access to subsidiary data; and
6. A strongly facilitated peer review.

To achieve this, the PRP should be a low cost, highly flexible virtual space, which supports:

- hosting a repository for personal production, with public aim, with past and present (work in progress) information and documentation;
- gathering digital resources, news, general information and materials on the same platform, accessible from each computer with an Internet connection;
- self-archiving and self-publishing research results, in terms of ongoing research, reflections, doubts, findings – avoiding waits and delays;
- informing the broader public what one knows and that one knows; and
- increasing one's visibility, enabling networking and knowledge sharing.

All in all, a PRP tracks the 'read-think-write' routine performed by scholars and scientists involved in research. The big difference with publishing is that a PRP not only keeps record of *stock* knowledge –formal knowledge that lasts or should last– but also of *flow* knowledge –non-structured, flexible knowledge devoted to fostering exchange.

As it happens with Personal Learning Environments, there is not such a thing as *the* PRP, because it can be built from a mesh of different

applications (a practical PRP guide is included below), which the user can choose and adapt as needed, for instance:

- a static web site with personal and professional information, drawing the researcher's profile;
- a blog, where to note news, reflections and 'flow' knowledge arising from readings, research results and hypotheses;
- a blogroll, understood as both a live reader for the researcher and a live bibliography of bookmarks for the community;
- a wiki, where 'stock' knowledge is stored but allowed to evolve along time and with the collaboration of third parties;
- a bibliographic manager, with online access to all or most records;
- a personal repository to (self-archived) published papers and (self-published) preprints, working papers, presentations, syllabuses, etc.;
- other tools, such as social bookmarking, file stores (image, sound, video), and so forth; and
- RSS feeds.

In other words, the PRP can be imagined as a lifetime personal web space "magnificently equipped (with software, communication, search, and multimedia tools), beehive[ly]-configured that possesses sufficient organizational plasticity to accommodate the user's developmental capacities and needs across a lifetime". These capacities and needs are related with a researcher's inputs (readings, conversations), transformation processes (reflections, peer reviews), and outputs (communications, preprints, papers).

There are nevertheless two caveats to be made: first, this individual publishing alternative is in no way a complete substitute to the stated ways of institutional OA publishing, but a complementary one that has some exclusive characteristics only attainable by this means. Second, along the same line, this is in no way a substitute for mainstream ways of publishing and validating scientific outcome but, again, a complementary one. The applicability of this tool for researchers in Developing Countries is explored below.

Social software, Web 2.0 and DIY web technologies

In the last years new, user-friendly web tools have appeared which moreover are often interconnected in such a way that communication and collaboration can take place. Such 'social software' –blogs and content management systems, wikis, message boards– is embedded in a wider concept, the 'web 2.0'. This is a model whereby peers contribute to the development of tools, content and communities, all taking place on the Internet. Such technologies are designed to simplify online publishing, simultaneously creating a network of both content and authors.

An important feature of these 'do-it-yourself' (DIY) web technologies is that they are usually licensed under free software licenses, so they can be installed and used for free, or are hosted by a provider that allows free usage, sponsoring the service through advertising. In either case, the cost for the user is restricted to a personal computer connected to the Internet, while the benefits are significant.

- The tools and technologies provide a way for researchers, who might otherwise have more trouble finding appropriate dissemination outlets, to easily share, make public and diffuse findings;
- Equally, information published using these technologies is made easily accessible;
- The more everyone joins a community, the richer it becomes.

The high level of economic sustainability of the proposed PRP model is one of its main highlights. Besides the required tools, the cost of hosting services for those aiming to install free software applications to be run under their own domain is constantly decreasing. In fact, some universities provide basic hosting, and many technology providers offer free hosting in exchange for locating advertisements on a site. One of the major problems that researchers face in poor countries is lack of necessary funding; decentralized web 2.0 tools as described above can contribute to alleviating this aspect, by providing an alternative means for researchers to circumvent costly infrastructures and formal institutions, yet allowing them

join international research communities, access relevant information and make results known.

Figure: ICTlogy

Three barriers stand in the way of widespread usage of this model in a development context. First, infrastructure: while affordable and easy access to ICTs and the Internet are pending issues around the globe, public libraries or civic centers increasingly provide free or low cost access, as do private telecentres. Although an in-depth analysis of these issues goes beyond the scope of this chapter, it is worth stressing that web 2.0 technologies demand relatively low computing power and connectivity quality, and therefore can be considered an interesting knowledge sharing framework for a development context. A second major barrier is user capacity, which is often limited in Developing Countries, in part due to the limited exposure to ICTs as described above. Computer skills are however increasingly addressed in development programs, and moreover, web 2.0 applications and social software are designed for non-technological users. Thus, even with a relatively low level of digital, technological and informational literacy can a user achieve interesting results and foster a 'conversation' among peers and scholars. Third, dissimilar cultural backgrounds and different mother tongues affect the ease of knowledge flow on online fora, but this aspect extends beyond ICT-enabled interaction; moreover, precisely the adaptability of web 2.0 technologies can stimulate the formation of local communities, providing a way by which this problem can be circumvented.

All in all, social software can contribute to the development of a network of peers. It is a form of technology stewardship by means of encouraging participation in conversation: wikis, fora, blogs, and many other tools provide perfect companions for newcomers to make their way into the web 2.0 arena.

'How to' hints: a PRP prototype

So what does a PRP look like? What does it involve? The underlying principle is that 'instead of building new applications from scratch, it makes sense to concentrate in the future on systematic combinations of existing Open Source tools for learning and competency development'. In this light, the design and implementation process is as interesting as the goal.

The combination of e-portfolios, social networks and weblogs may have immense benefits for the learner. These tools and the ethos behind them enhance the prospect for deep learning. Creation of a learning landscape where learners engage in the whole process both academically and socially should increase the opportunity to build one's learning instead of just being the recipients of information.

Some examples of PRPs in practice

On education:

George Siemens	http://elearnspace.org
Stephen Downes	http://downes.ca
Helen Barrett	http://electronicportfolios.com

On ICT for development (ICT4D):

Victor Mbarika	http://www.vmbarika.com
Ismael Peña-López	http://ICTlogy.net

A 'one minute handbook' on how to build a Personal Research Portal would include the following components:

- *Domain and hosting*: A domain name is automatically associated with specific content and its managers and contributes to the 'digital identity' of the owner, as discussed above. Hosting allows autonomous tools to be integrated into the portal, in terms of services, shape, contents and so forth.
- *Content management*: Static pages and most of the dynamic ones can be built using a content management system (CMS). Drupal (www.drupal.org) or Joomla (www.joomla.org) are open source varieties of such systems, with the advantage that they also feature blogs. Reversely, WordPress (www.wordpress.org) is a blog engine that can also be used as CMS. Alternative tools are e-portfolio applications such as OSPI and Elgg.
- *Collaborative tools*: In terms of collaborative tools, the options are clear: if the expected output is content, a wiki is probably the best option. If the goal is the process, the debate itself, then discussion fora are required. Appropriate applications might include Mediawiki (www.mediawiki.org) –for the wiki– and phpBB (www.phpbb.com) –for the message board.

- *Bibliographical tools*: While different bibliographic managers are available, there is little consensus in terms of the best bet. However, Refbase (http://refbase.sourceforge.net) and BibCiter (http://bibciter.sourceforge.net) fit the PRP purpose: both are web based and

have RSS output. EPrints (www.eprints.org) and Open Journal Systems (http://pkp.sfu.ca/ojs) work well for self-archiving and self-publishing, respectively.

- *Social software*: Many other applications exist to share bookmarks, photos and slideshows, podcasts, vodcasts, etc. Most of them are online services provided –and hosted– by third parties. An important consideration when choosing such tools is their capacity to import and export a user's data and the ease by which they can be linked in a PRP.
- *RSS*: 'Really Simple Syndication' (RSS) is an alternative means of accessing the vast amount of information that now exists on the world wide web. Instead of the user browsing websites for information of interest, the information is sent directly to the user. In any case, RSS output, as the glue of such portals is a must.

When connectivity is not available and a user intends to work predominantly 'locally', XAMPP (http://www.apachefriends.org/en/xampp.html) makes it possible to (re)install all these social software applications –in fact the whole PRP– on a hard drive or a USB pen drive. Indeed, it can work as a backup for our PRP and/or make it portable across different operating systems.

PRP and the knowledge divide

Digital identity

One of the main problems that researchers face in Developing Countries is invisibility to the broader research community. This invisibility has at least two major consequences:

- Minimum awareness and recognition of findings, fields of work, interests and existence;
- Difficult access to mainstream publishing circuits.

In order for researchers and their work to be recognized in academic and practitioner circles at the international level, their visibility needs to be enhanced. Setting up a PRP can thus be understood, at a primary level, as the creation of a personal home page, 'building a virtual identity insofar as

it flags topics, stances and people regarded by the author as significant'. This digital identity, or the researcher's presence on the Internet, is complementary to the academic identity shown by authorship in academic journals and conference proceedings. While the latter are strongly tied to the development of the researcher's concepts and his or her contribution to the development of formal knowledge, the digital identity builds on these, providing additional information on:

- the owner's identity (who am I);
- the owner's activities and interests (what do I do);
- the owner's achievements (what have I done);
- the owner's contact details (where am I).

If mainstream systems –congresses, journals, seminars– act as diffusion hubs for offline identities, search engines, portals, third parties' blogs and institutional pages, signature files in e-mails (specially when placed in discussion lists and message boards) act as diffusion hubs for online identities.

Nevertheless, there are two main advantages of online media versus 'offline' dissemination mechanisms:

- t higher potential reach;
- more up to date information. If managed properly, PRPs can show the latest news about a researcher's institutional affiliation, can include recent research trends and so on. In fact, if updated pages use RSS feeds and are correctly meta-tagged, human intervention is not necessary for the changes to be echoed in specific search engines and feed aggregators.

Overall, the main component of a PRP should be evolving, up-to-date information of one's own. Search engines are web 2.0-friendly and award high rankings to dynamic pages with rich and focused content. Descriptions of one's research and interests, as well as providing documents, other relevant materials and links to and from other people

with similar interests, enhance the possibility of being found under specific keywords. This information can be created by means of simple HTML documents or, better, using a CMS –or CMS-like features from other applications such as blogs, which can play a significant part in terms of linking and networking.

Reading, live storing and the public notebook: reinforcing the digital identity

The research process generally involves extensive note taking, as highlights of what has been read, reflections that arise after the reading or simply as a record of the fact that something has been read. Social software empowers researchers in such a way that their notes can be 'published to the World Wide Web as a way to 'display and reflect on their learning' to an audience that is broader than just their classmates'. Moreover, "knowledge only works if each person makes links as he or she browses, so writing, link creation, and browsing must be totally integrated. If someone discovers a relationship but doesn't make the link, he or she is wiser but the group is not".

Such a digital notebook –in the shape of a blog, an important part of the PRP– allows the process of reading, writing, analysis, reflection and learning to be fully public: "Eventually, there will be publications in scholarly outlets, but there are both more immediate and more long lasting benefits. In the near term, ideas can be more readily implemented, data automatically collected". Another immediate consequence of this way of working is that "less knowledge is left behind", as a live digital store is created each day, a store that is categorized, searchable and fully accessible. The PRP here "represents a space where the relationship between memory and promise, the link between past and future is made possible". Hence, a factually driven dynamic identity evolves by tracking the researcher, creating new knowledge in the framework of his community.

This identity is reinforced by the fact that content is categorized –tagged– according to specific keywords. And, besides the fact that categorization (and 'searchability') can be useful to the researcher, full accessibility is the

key: not only data and information are accessible everywhere to the owner or creator of the PRP, but also to other researchers. In view of enhancing accessibility to knowledge and visibility for people in Developing Countries, this can make a difference. Through its inherent characteristic of immediacy, a PRP provides access to knowledge without filters and without waits: the PRP becomes a digital store of resources, news and current events, general information, academic materials and state of the art research. It should be noted that in some countries Internet censorship can obscure this aspect; however, this is a political problem rather than a technological or conceptual one, and so goes beyond the scope of this chapter.

As a collector of 'flow' knowledge, contributing to a wiki, joining a blog or uploading files to a server can add to the 'stock' knowledge of a PRP. In this context, bibliographic tools are also worth exploring. Their purpose is to organize one's references and to ease the task of citation. Some varieties of bibliographic tools are web applications, installed on a web server and run on web browsers. This allows not only managing but publishing one's references and bibliographies. This feature reinforces one's digital identity by allowing cross-referencing in a body of knowledge, and providing more rigour to the content shared on a PRP.

Writing and participating in conversation: Network building

Social software is all about meeting colleagues, exchanging impressions and collaborating. Interconnecting PRPs capitalize on this capacity by taking advantage of automated linking methods. Of the different software varieties and perhaps even more than search engines, RSS feeds, a part of the XML family, enable knowledge sharing and foster community building in real time, for instance through include pingbacks and trackbacks. 'Pingback' is a method for Web authors to request notification when somebody links to one of their documents. This enables authors to keep track of who is linking to, or referring to their articles.

Some weblog software, like WordPress, support automatic pingbacks where all the links in a published article can be pinged when the article is published (source: http://en.wikipedia.org/wiki/Pingback). Trackback is

essentially the opposite, allowing a blogger to see who has seen the original post and has written another entry concerning it. The system works by sending a 'ping' between the blogs, and therefore providing the alert. (source: http://en.wikipedia.org/wiki/TrackBack). While these methods contain implicit technological linking, pingbacking and trackbacking require an explicit conceptual linkage in terms of the body of knowledge which researchers share interest in or are contributing to.

Social networking can further be reinforced by comments on others' PRPs or the creation of 'Friend of a friend' (FOAF) files and blogrolls: 'FOAF' is an XML standard that allows website owners to define who they are as well as their relationships with other website owners – effectively creating a wide area social network.

Inclusion of such tools contribute to shaping a virtual research network around the PRP and in fact around its creator. In terms of development impact, the PRP can potentially "seamlessly link individuals to larger communities, thereby facilitating interpersonal connectivity versus fostering social isolation". By reinforcing this behavior among scholars the presence of 'invisible' researchers can be improved.

Web pages are a form of asynchronous communication, but social software makes it possible that "my web page... mediatively interacts with other people in my absence". Indeed, collaboration can occur, "reducing contact time while also increasing the quality of contact time". From this perspective, PRPs "can help people to define their own success through reflection with evidence often *enhanced with peer or mentor commentary*", through the exchanges and linkages which can be initiated through such virtual exchanges. Although these are of quite a different nature than the habitual double-blind review that most journals follow, open peer exchange boosts networking and collaboration to an effect reaching far beyond that of anonymous readers. Moreover, the immediacy of finding posted on PRPs allows

hypotheses [to be] more easily tested, thus reducing the cost associated with research ventures and increasing productivity. Similarly new

researchers can quickly be integrated into ongoing projects and make contributions to the research […]. In the long term, the external visibility of the web-based research engine will promote a shift in organizational culture toward a more open and cooperative environment where knowledge augmentation and sharing are instrumental to individual learning and organizational development. In such a culture research engine participants will benefit from increased collaboration with qualified colleagues both within and outside the institution.

Overall, to take part in a conversation one must speak –and blogs (individual or collective, supplemented by feedreaders), wikis or online office suites are tools by which to make one's voice heard. Contributing, commenting and linking (directly or through pingbacks and trackbacks) are the way to let others know their work is acknowledged.

Self-archiving, self-publishing

"Whilst the fundamental technical difference between the medium of speech and that of writing is that writing is automatically recorded, web pages introduce another key feature: what is written on a web page (and stored on a web-server) is automatically published". As mentioned above, researchers in Developing Countries often face tough barriers to do such publishing. A tool like the PRP can help address this problem in different ways.

First of all, self-archiving of preprints and published works in a personal repository is an evident purpose for the PRPs to fulfill. This is only a complementary track to journal publishing, but is a way to provide access to published works which would otherwise remain more obscure, at the same time acting as a repository of the owner's (academic) output. "This complete openness may be an anathema to archivists and cataloguers as it abandons all attempts to control the system, but it was suggested that such an approach could greatly facilitate short term uptake" [of knowledge].

'Anathemic' or not, self-publishing goes one step further still in terms of challenging the faculty establishment, because it avoids peer review. Even so, self-publishing has its value, providing an opportunity for publication

of interesting work that might otherwise remain unpublished. Moreover, works that need no review such as newsletters, bulletins, opinion columns, working papers, datasets, or works that have already been reviewed such as working papers, theses and other kind of dissertations can obtain formal identifiers (ISSN or ISBN) and be published on a PRP without violating academic standards or other publishing norms. Under an open license, such publishing contributes to increasing the visibility of the author, shaping a digital identity, enriching the content of the site, making it more appealing to users and search engines and, all in all, helping research to have its rightful place in the academic arena.

In the long run, an increased legitimacy of Open Access science can be expected. The benefits such as higher exposure and easier dissemination make it easier for work to be accessed and reviewed but more difficult to plagiarize, while generally enhancing the advancement of knowledge to researchers in developing and developed countries alike.

Remarks

There is a place for individual initiatives, complementary to institutional efforts, to bridging the research biases and imbalances in the international arena. These initiatives find a perfect companion in social software tools. Some of these tools, such as wikis, social bookmarking, social networking, file sharing, RSS feeds, discussion forums and blogs, are already used for diverse research purposes either directly; of these tools, the blog is the most important.

Despite the digital divide which still restrains researchers in Developing Countries from capitalizing fully on the possibilities provided by these tools, virtual communities have demonstrated their potential for bridging capacity divides, whereby technology stewardship take place naturally, non-hierarchically yet non-chaotically.

More than just a matter of being published, or participating in knowledge communities, collaborative or 'discourse' technologies can play a role in empowering the individual with (digital) means to master a learning and research process within a cultural framework relevant to his or her needs.

The creation of a digital identity is a means of empowerment, contributing to gaining control over one's life, but also participating equally in a globalized knowledge society. As such, the PRP is, overall, an e-inclusion device, contributing the creation of a vast and public body of knowledge for progress.

12

Open academic webcasting

Based on:

- *Open Book "How To Accelerate Your Internet", Chapter 8.*
  Website: http://bwmo.net

- *E. Canessa, M. Zennaro and C. Fonda, "Information for Development Magazine" (i4d), Feb 2008.*
  Website: www.i4donline.net

- *Website: http://ocw.mit.edu*

- *J. Lynn Fink and Philip E. Bourne, CTWatch Quarterly, Vol.3, Issue 3, Aug 2007.*
  Website: http://www.ctwatch.org/quarterly/

- *Website: http://www.scivee.tv*

For the production and delivery of open webcasting, it is necessary to adopt low-bandwidth compliant applications that keep the video and audio quality as high as possible and keep the size of all synchronised files (video, audio, slides) as small as possible. Also relevant to consider, is to automate as much as possible the recordings production and to reduce any manual post-processing and editing. The latter becomes especially relevant when carrying out massive recordings in different rooms across a large campus.

Lectures given in a classroom can be more complex in form than relatively simpler seminar presentations. They usually include the simultaneous use of a standard chalkboard, projected transparencies plus the display of PowerPoint (PPT) or Keynote presentations with the use of some pointers. All such features need to be considered when publishing these lectures online in order to recreate a learning experience that creates the closest possible classroom reality for remote audience without excluding any bit of information. This is still a technological challenge. To produce such type of synchronised recordings, both the audience and their available computer/networking facilities need to be considered.

Bandwidth consuming technologies

The ultimate vision of the Internet is a truly ubiquitous network where information flows freely to wherever humans can make use of it. It is likely that we will eventually build a network where there is sufficient bandwidth for all. But until this is achieved, we must continue to carefully manage network resources, ensuring that this limited resource is equitably accessible to everyone.

Figure: example of bandwidth-intensive scientific web: scivee.tv

The technologies that allow us to effectively manage Internet resources are evolving as quickly as the Internet. New techniques and tools are being developed every day that help us to squeeze even more performance out of our overburdened network connections.

The streaming of audio and video is becoming increasingly popular, and uses large amounts of bandwidth in comparison to traditional text-only media. Downloading live television broadcasts from Web sites (such as BBC, CNN, etc) is also becoming common. In addition to causing a heavy load on the network, streaming technologies need a stable, reliable connection. As webcasting usage is likely to increase to the point of being an essential service, this will place a great demand for effective bandwidth management.

Connectivity trends in Developing Countries

A definite trend is continuing towards multimedia websites, which contain bandwidth-hungry images, video, animations, and interactive content. This puts increasing demands on Internet connections, resulting in slower downloads and decreased usability on saturated connections. Bandwidth management technologies will have to adapt to this change, employing new caching technologies to reduce bandwidth demand and improve user experience. Although policy may ban or discourage a certain type of Internet use, in practice it can be difficult to completely enforce that policy purely by technical means. For example, an institution may have a policy of limiting the use of file sharing. However, determined users can tunnel file sharing over other protocols like HTTP, which is difficult or impossible for the institution to detect.

As the number of Internet users in underdeveloped parts of Africa and Asia expands, there will also be a growing need to provide more local services. It will become more important for copies of large files (e.g., open source software) to be stored on servers closer to users, which will enhance the speed and reliability of downloads. This technique, called ‘mirroring’, is already widely used in the developed world, but there are no known public mirrors of popular software on the African continent. Even today, 59% of African universities have limited or no bandwidth management,

according to reports. As small Local Area Networks (LANs) are created and extended, it is likely that more and more people will become de facto network administrators, despite having little or no training. If this training shortage problem is not addressed, the situation can only become worse. Organisations with tighter budgets, such as universities, colleges, and research institutes, will continue to suffer with respect to the private sector in terms of the speed and reliability of their Internet connections. Unfortunately, these are the same institutions which cannot afford the best systems administrators, and where the public benefit that would arise from unfettered access to information is greatest.

As the knowledge of cheaper phone calls over the Internet spreads in Developing Countries, coupled with gradually increasing bandwidth and gradually spreading networks, users will start to demand faster and better Internet connections. Newer bandwidth management software and network hardware, with better support for guaranteed network Quality of Service (QoS), will spread beyond the best equipped networks. Thus administrators will find themselves under pressure to implement these systems on their networks. Conversely, shared bandwidth connections such as ADSL will continue to grow in popularity, at the expense of guaranteed bandwidth connections such as leased lines. While these seem cheaper, and often offer better performance on average, shared connections make it very difficult to manage bandwidth since the usable bandwidth is unknown and constantly varying.

Enhance your audience (EyA)

One concrete possibility to achieve open academic webcasting of traditional classroom lectures is given by the recording system named "*Enhance your Audience*" (EyA, www.ictp.tv). This is an automated recording system in use at the Abdus Salam International Centre for Theoretical Physics in Trieste, Italy (ICTP, www.ictp.it) to record mathematics and physics lectures, and international conferences and workshops. The automated EyA is an innovative system developed to archive and share scientific webcastings carried out using either digital (PPT, PDF, etc.) presentations or the old and more traditional chalkboard lectures without any human intervention. Video and audio are recorded in

slots of one hour on a "producer" computer with a webcam and USB microphone fixed on the wall. High quality photos (seven megapixels or higher) are taken every 15 seconds with a digital camera (controlled via software) and immediately downloaded from the camera to the computer via USB. Images are compared together to drop duplicates in order to decrease the space needed for storage and download of the recordings. By automatically synchronising the images with the audio/video recordings, the viewer can zoom in regions of a large screen, podium or chalkboard to visualize a presentation more effectively. EyA recordings also offer the benefit of seeing the physical gestures, body languages (of the people present before the camera) and the like which are inherent in classroom lectures.

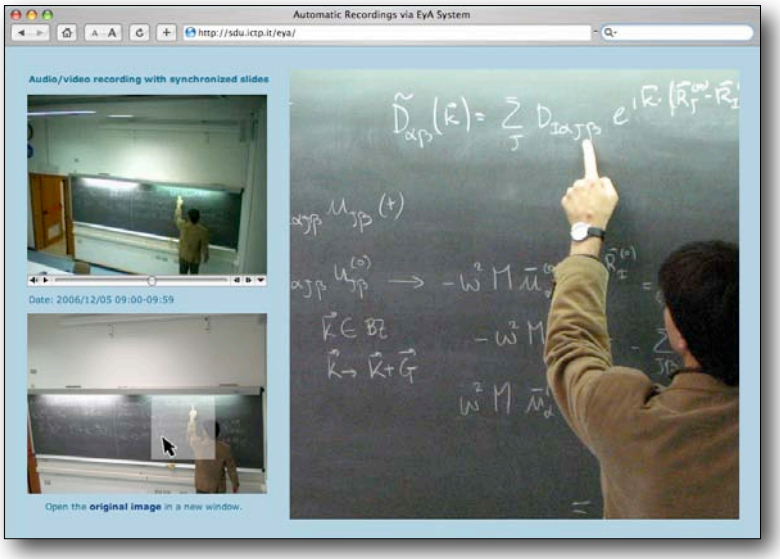


Figure: physics lecture recorded with EyA

Digitization of open course content

Starting from September 2007 onwards, all lectures given within the ICTP Diploma Course of the ICTP are being automatically recorded using the automated EyA system. The access to this digital material is made freely available on the web (www.ictp.tv) to the public i.e., students at our

Campus and beyond. The "ICTP Diploma Course On-Line" project aims to enhance ICTP's mandate of transferring knowledge to scientists from Developing Countries. These public recordings can also be useful for lecturers beginning their teaching careers in science around the world.

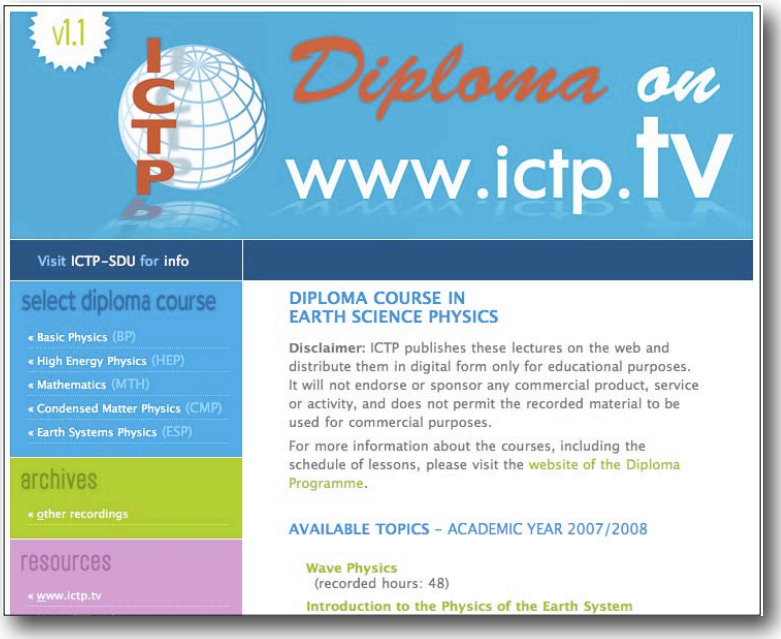


Figure: ICTP Diploma Course on-line

It is possible to follow the open synchronised recordings (in Flash and/or QuickTime formats) from any computer connected to Internet and to download the compressed zip files within an hour of recordings or to watch it using any web browser, including those available on public cybercafés. So far as July 2008, more than 2,140 hours of recordings, which include more than 54 different courses, are being openly published on the web. ICTP publishes these lectures on the web and distributes them in digital form only for educational purposes. It does not endorse or sponsor any commercial product, service or activity, and does not permit the recorded material to be used for commercial purposes.

Evaluation and assessment

ICTP has carried out a recent study of the assessment and evaluation of the EyA use in the ICTP Diploma Programme (for a whole Semester: Sept-Dec 2007). In the following, a students" survey was carried out to evaluate the EyA recording system when applied as an educational technology tool.

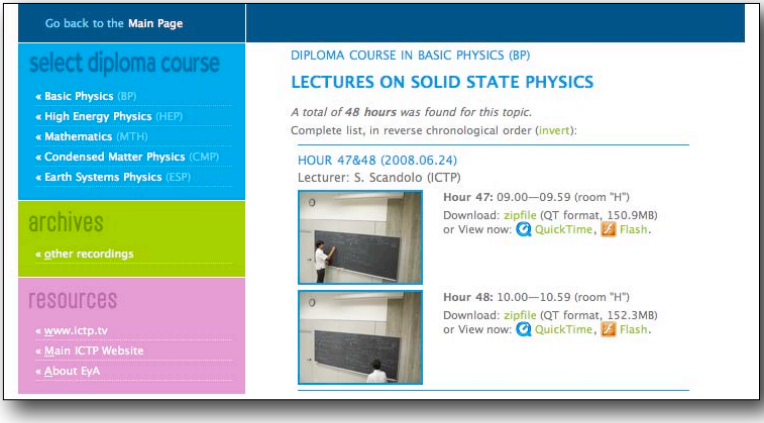


Figure: On-line lectures recorded with EyA

This survey consisted of a questionnaire focusing on the use, assessment, divulgation and overview of the available recorded lectures for each of the five running Diploma Courses: i) Condensed Matter Physics (CMP), ii) High Energy Physics (HEP), iii) Earth System Physics (ESP), iv) Basic Physics (BP) and v) Mathematics (MTH). The number of compiled questionnaires received represents about 70% of the actual number of ICTP Diploma students. When asked about the efficacy of recording system, 97.1% of the students replied "helpful" or "very useful". Only one did reply "not very useful" (because the person seldom watched the recordings). The survey revealed that open webcastings are helping students to:

- review/revise missing points and concepts during a lecture (even after a long time);
- clarify handwritten notes misplaced (taken by the students themselves);

- reminisce a class (missed for good reasons such as administrative duties or illness);
- understand concepts when sometimes lecturers go fast;
- prepare for exams (with the following percent of students): CMP: 100%, ESP: 100%, HEP: 100%, MTH: 75%, BP: 63.6%;
- relax when tired of reading;
- adapt to an Anglophone environment;
- review the class anytime wanted/needed at and understand all details;
- avoid writing lecture notes (enables to concentrate the Blackboard);
- mark it for possible long-term research.

Remarks

Students from Basic Physics ICTP Diploma Course mainly review the most recent Digital Lectures, whereas those from the Condensed Matter Physics Course review those further back on time (more than a week). The maximum time dedicated to watch digital lectures is about 13 hours per week on average, especially in the field of Mathematics. The open online publications are mainly accessed by students from the Earth System Physics Diploma Course. About 50 percent of the students have downloaded more than 25 zip files containing one hour of lecture each. The goodwill and "word of mouth" approach passed down to inform colleagues seem to have applied for some Diploma courses. Accordingly, it was reported that sharing is indeed being done with Vietnam, Kenya, Philippines and Bangladesh. In view of all these results it can be concluded that the adoption of Open Webcasting via EyA in the Diploma Course has been very rewarding as it stands. This project has only been possible because of the active participation between, and close collaboration of the multiple actors involved in the Diploma Programme. These include all Diploma Course Students (about 50) and Lecturers and Tutors (about 80), two Secretaries and the project ICTP Directorate - together with the Science Dissemination Unit (SDU) Team.

It is also a matter of pride to state that the project undertaken by ICTP has been quite successful in its unique endeavor towards achieving the goal of transferring knowledge to scientists and scholars of Developing Countries.

There is probably no such commendable effort around the world that attempts to carry out automatically low-cost, massive and complete recordings to disseminate high-quality, pre-PhD science programs on the Internet. As a difference with, for example, the MIT Open Courseware, the EyA system does not require any video operator for recording and/or editing the on-line lectures.

Overall, the feedback and experiences with regard to online Diploma Course in the last couple of months, encourage ICTP to proceed further in this direction. It is foreseeable that some technical improvements (such as audio quality, rooms luminosity, DVD production, etc.) plus the addition of state-of-the-art developments for the automated EyA system are required. Finally, surprisingly enough it was found that nearly 50 percent of the total students are now following voluntarily even the complementary lectures which do not belong to their own Diploma Courses. This means that open webcasting "also gives the opportunity to students to follow parallel courses in other fields". This opens a new era for science students across the globe and supports them with extra digital material and help.

MIT OpenCourseWare

MIT OpenCourseWare (OCW, ocw.mit.edu) is a web-based publication of virtually all MIT (Massachusetts Institute of Technology) course content. OCW is open and available to the world and is a permanent MIT activity.

There is no registration or enrollment process because OCW is not a credit-bearing or degree-granting initiative. MIT OpenCourseWare is a publication of the course materials that support the dynamic classroom interactions of an MIT education. There are no prerequisites to use MIT OpenCourseWare materials (with more than 1800 Courses so far).
Each course published requires an investment of US$10,000 to 15,000 to compile course materials from faculty, ensure proper licensing for open sharing, and format materials for global distribution. Courses with video

content cost about twice as much, but feedbacks about the significant value of these video materials helps to justify the cost.

- MIT OpenCourseWare averages one million visits each month; translations receive 500,000 more.
- MIT OpenCourseWare is being successfully used for a wide range of purposes:

Use Scenario		% of Use
Educators	Enhance personal knowledge	25.0%
	Develop a course	22.8%
	Prepare for a specific course	17.8%
	Enhance research	13.8%
	Other	20.6%
Students	Complement a course they are currently taking	43.7%
	Enhance personal knowledge	32.1%
	Plan a course of study	12.4%
	Other	11.9%
Self Learners	Enhance personal knowledge	58.1%
	Keep current with developments in their field	17.9%
	Plan a future course of study	10.8%
	Other	13.2%

Video communications with SciVee

Due to the increasing availability of high bandwidth and consumer-level video recording equipment, internet video is now becoming wildly popular. One can take advantage of this trend and use this medium to communicate science more effectively. It is important, however, to bear in mind the need for quality content. To this end, SciVee (www.scivee.tv)

allows authors to upload an Open Access article they have already published with a video or podcast presentation (about 10 minutes long) that they have made that describes the highlights of the paper. The author can then synchronize the video with the content of the article (text, figures, etc.) such that the relevant parts of the article appear as the author discusses them during the video presentation. The result is called a pubcast.

The figure shows a typical SciVee pubcast. It can bee seen how a video presentation is integrated with a published article. While the speaker is discussing a point from the article, the relevant figure or text is highlighted. The viewer can also download the original paper as a companion to the pubcast.

Figure: SciVee.tv pubcast

Anyone can visit SciVee and view the pubcast. It is similar to attending a conference to hear a particular speaker, except that the pubcast is available on demand, can be viewed any number of times, and explicitly refers to the content of the original article. Another important feature of

SciVee is the ability of any user to add or read comments on pubcasts. This allows a community to be established around an article and encourages discussion about the results and their impact on the field. We believe this activity will transform what has traditionally been a static document into a dynamic exchange.

SciVee makes it easier and faster to keep up with current literature by delivering the key points of articles in a portable and enjoyable medium. A reader can interact with several articles using the SciVee website in the time it would take to read a single full article in the traditional way.

Figure: typical pubcast workflow

PART 2: SOFTWARE

'tag cloud' by wordle.net

This second part of the book is about software for setting up an Open Access Journal and an Open Access repository. It is designed to help selecting the software system that best satisfies institutions' or scientists' needs. These needs will be driven by each institution's content policies and by the various administrative and technical procedures required to implement those policies. Therefore, these chapters are designed for institutions already familiar with the various administrative, policy, and related planning issues relevant to implementing an institutional repository or to setting up an Open Access journal.

While this document describes these solutions, it does not attempt to identify the 'best' system or to recommend one system over another. In each institution's case, the best software will be that which aligns well with the institution's particular requirements. This document can only provide an overview of the available software. Further, these systems are evolving rapidly. Readers should also refer to the additional information on system features and functionality available directly from the software providers themselves. Links to this information are provided with each system description.

For publishers with access to computer technology and the Internet, publishing management software offers an effective method of managing the workflow of the entire publishing process. Publishing management software can facilitate the submission of new articles, the peer review process, the editorial life-cycle of the article, and the ultimate publication of the final version. The following chapters will describe two software packages for setting up an Open Access Journal: the Open Journals System and Topaz.For setting up a repository, there are number of software packages available, of which the most popular worldwide are EPrints and DSpace. EPrints is currently the most widely used in the UK. Both of these are described in the following chapters.

Based on:

- Open Society Institute: "A Guide to Institutional Repository Software"
Available at: http://www.soros.org/openaccess/pdf/OSI_Guide_to_IR_Software_v3.pdf

13

EPrints

Based on:

- Websites:
http://wiki.eprints.org/w/Introduction
http://wiki.eprints.org/w/Configuration_orientation

EPrints (www.eprints.org) is a generic repository building software developed by the University of Southampton. It is intended to create a highly configurable web-based repository. EPrints is often used as an open archive for research papers, and the default configuration reflects this, but it is also used for other things such as images, research data, audio archives –anything that can be stored digitally. The EPrints series began in early 2000 and is in use by over 200 sites.

To get a feel for what the software does, start by looking at:
http://demoprints.eprints.org/

You can get the system up and running quite easily. You'll need a UNIX-like machine (linux is good), and a root password is helpful. Thanks to support from Microsoft, it also runs on Windows Vista and XP. The task which will take longest is actually deciding what you want your repository to do. EPrints creates a repository with a sensible default, but many sites want to make significant customizations.

EPrints doesn't require any unusual hardware. It's slightly easier to run on a dedicated machine, but that's not essential, and should not affect performance. Don't forget to budget for a backup system, your data is valuable!

EPrints live CD

The EPrints3 Live CD is designed to help you experience EPrints without the need for a machine to install it on. The CD comes complete with a fully installed EPrints3 distribution which also has a recent version of the training documentation and exercises as well as example material. From the environment the CD provides you can harness the full power of EPrints3 to setup your own archives for testing and also have an easy way to test new EPrints functionality without putting your active archive at risk.

For new repository maintainers the Live CD can also be installed permanently with one click onto your hard drive. This will enable you to preserve changes to your archives between reboots (without needing to use a memory stick). The Live CD really does provide the all in one solution for experiencing EPrints repository management software.

The CD is currently only available for the x86 (32-bit) platforms and is based upon the Ubuntu Linux Live CD with Open Office removed to make available the space to install EPrints and all its dependencies whilst still remaining small enough to fit on a single CD.

Using the live CD

1. Download the image from:
   http://www.eprints.org/files/eprints3/livecd_v3.0-x.iso
   and burn it to a CD using your favorite CD-writing application;
2. boot PC from the CD;
3. to prepare EPrints to run from the live CD open a terminal (available from Applications → Accessories) and type the following:
   ```
   sudo setup_eprints_live
   ```
4. open Firefox or similar browser and navigate to:
   http://training.eprints.org (which is linked to your local machine, no net connection required).

To install from the live CD

From a clean boot into the Live CD click the install icon.
Once installed and running from your hard drive execute the following from within a terminal:

```
sudo touch /etc/apache2/httpd.conf
sudo apache2ctl stop
sudo apache2crl start
```

To enable the training archive:

```
sudo echo “127.0.0.1 training.eprints.org training” >> /etc/hosts.conf
```

Storing your archive on a memory stick

By using the live CD, any changes made will be lost when the system is shut down to avoid this we have created some scripts which backup your archive onto a memory stick.

These scripts should be stored somewhere on the memory stick and the files are stored in a folder called 'live_cd' relative to the position to the scripts.

Extract http://www.eprints.org/files/eprints3/mem_key/mem_key.zip onto your memory stick. You can then:

- archive only backup, this will only backup your archive, it will not include any of the global plugins or settings which apply to all repositories.

```
root@ubuntu$ bash archive_to_usb.sh archive_name
```

- complete EPrints Tree Backup, this will backup everything in your eprints3 tree.

```
root@ubuntu$ bash all_to_usb.sh archive_name
```

Don't forget to safely unmount the memory stick, on the live CD close any windows which popped up when the memory stick was inserted and then as root:

```
root@ubuntu$ cd /
root@ubuntu$ umount /media/disk
```

Script details

The script backs up both your archive and the related database by doing the following, it assumes you have not set a root password on the mysql database:

archive_to_usb.sh:

```
#!/bin/bash
if [ “$1” = “” ]
then
 echo “No Archive Specified”
else
 FOO=`pwd`
 mkdir “$FOO/live_cd”
 rm -fR “$FOO/live_cd/$1”
 mkdir “$FOO/live_cd/$1”
 tar -cf “$FOO/live_cd/$1/$1.tar” /usr/share/eprints3/archives/$1
 mysqldump -u root $1 > $FOO/live_cd/$1/$1.sql
 echo “DONE”
fi
```

all_to_usb.sh:

```
#!/bin/bash
if [ “$1” = “” ]
then
   echo “No Archive Specified”
else
FOO=`pwd`
mkdir “$FOO/live_cd”
rm -fR “$FOO/live_cd/eprints3”
mkdir “$FOO/live_cd/eprints3”
rm -fR “$FOO/live_cd/$1”
mkdir “$FOO/live_cd/$1”
tar -cf “$FOO/live_cd/eprints3/eprints3.tar” /usr/share/eprints3/*
mysqldump -u root $1 > “$FOO/live_cd/$1/$1.sql”
echo “DONE”
fi
```

Restoring your archive from memory stick

Don't forget to follow the step above to configure ePrints to run from the live CD first! Locate the place where you extracted the files to previously on your memory stick and as root:

- archive only

```
root@ubuntu$ bash usb_to_archive.sh archive_name
```

- complete EPrints Tree

```
root@ubuntu$ bash usb_to_all.sh archive_name
```

Script Details

Pretty much the opposite of the other script:

usb_to_archive.sh:

```
#!/bin/bash
if [ "$1" = "" ]
then
 echo "No Archive Specified"
else
 FOO=`pwd`
 rm -fR /usr/share/eprints3/archives/$1
 tar -xf live_cd/$1/$1.tar -C /
 echo "drop database $1" | mysql -u root
 echo "create database $1" | mysql -u root
 mysql -u root $1 < live_cd/$1/$1.sql
 echo "DONE"
fi
```

usb_to_all.sh:

```
#!/bin/bash
if [ "$1" = "" ]
then
echo "No Archive Specified"
else
FOO=`pwd`
rm -fR /usr/share/eprints3/
tar -xf live_cd/eprints3/eprints3.tar -C /
echo "drop database $1" | mysql -u root
echo "create database $1" | mysql -u root
mysql -u root $1 < live_cd/$1/$1.sql
echo "DONE"
fi
```

14

DSpace

Based on:

- Website: http://www.dspace.org

- DSpace Manual Software Version 1.5 May 2008

A groundbreaking digital repository system, DSpace (www.dspace.org) captures, stores, indexes, preserves and redistributes an organization's research material in digital formats. Research institutions worldwide use DSpace for a variety of digital archiving needs –from institutional repositories (IRs) to learning object repositories or electronic records management, and more. DSpace is freely available as open source software you can customize and extend. An active community of developers, researchers and users worldwide contribute their expertise to the DSpace Community.

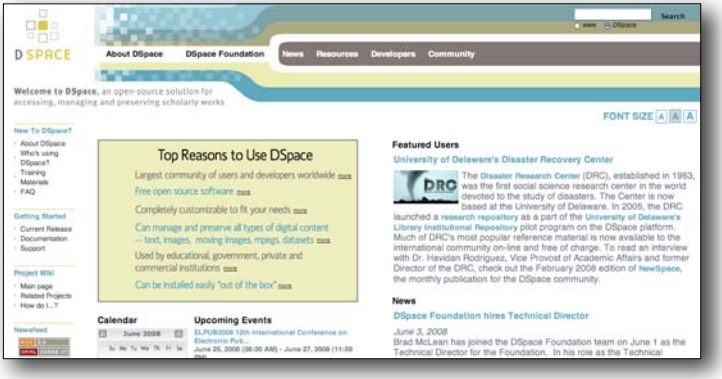


Figure: DSpace

DSpace FAQ

1. Who can join the DSpace community?

Anyone who uses DSpace can get involved, in a number of different ways: programming, defining feature requirements, writing documentation, testing new features, sharing your design or marketing expertise. Get involved by joining the DSpace mailing lists, adding your projects, experiences, and comments to the DSpace Wiki, and collaborating with other DSpace community members.

2. What are DSpace Communities and Collections?

Each DSpace service is comprised of Communities –groups that contribute content to DSpace– and Communities in turn each have Collections,

which contain the content items, or files. In a University environment, for example, Communities might be departments, labs, research centers, schools, or some other administrative unit within an institution. Communities determine their own content guidelines and decide who has access to the community's contributions. An administrator on the DSpace team, usually the DSpace User Support Manager, works with the head of a community to set up workflows for content to be approved, edited, tagged with metadata, etc. Collections belong to a community or multiple communities (for example, research collaborations between two communities may result in a shared collection) and house the individual content items and files.

3. What is a DSpace Early Adopter?

In a DSpace Early Adopter program, you run a trial period for your DSpace service with a few hand-picked communities before launching a full DSpace service. While it's not essential to run an early adopter program, it can be enormously helpful in gathering feedback, working out the kinks, and building support for DSpace at your institution.

4. Who built DSpace?

The MIT Libraries and Hewlett-Packard (HP) jointly developed DSpace. The system is now freely available to research institutions world-wide as an open source system that can be customized and extended.

5. How is DSpace different from other digital repositories?

DSpace is the first digital repository to address the myriad issues inherent in a multi-disciplinary archive, including:

- ✦ differing policies, practices and cultures established by individual disciplines;
- ✦ the variety of digital formats produced in today's multi-media research environments;
- ✦ the complexity of metadata standards needed to accommodate and maintain access to the digital formats supported by the system;

- DSpace is designed with a flexible storage and retrieval architecture adaptable to a multitude of data formats and distinct research disciplines, known as 'Communities'. Each Community has its own customized user portal that can use the Community's own practices and terminology.

6. Is DSpace free?

Yes. The DSpace system is freely available as open-source software, under the terms of the BSD distribution license. We have also tried to find good open-source tools to package with the DSpace application, all freely available under an open-source license (although not all the same license as the one for DSpace itself), so that you get a complete system along with the part that we created.

7. Who can download the software?

Open-source systems like DSpace are available for anyone to download and run at any type of institution, organization, or company (or even just an individual). Users are also allowed to modify DSpace to meet an organization's specific needs. The BSD distribution license describes the specific terms of use.

8. Where can I download the DSpace open-source software?

DSpace is freely available as open-source software from SourceForge, at http://sourceforge.net/projects/dspace/.

9. Can I change the DSpace system?

Yes, you can customize and extend the system to suit your organization's needs. DSpace was designed to make adapting it for individual organizations as easy as possible. Each application is different, but most organizations need to customize the authentication system, for example, to work with existing systems. Some organizations may want to substitute the open-source tools supplied with DSpace with different ones (for example, replacing postgreSQL with mySQL or Oracle).

10. What kind of content does DSpace support?

DSpace accepts all manner of digital formats. Some examples of items that DSpace can accommodate are: documents, such as; articles, preprints, working papers, technical reports, conference papers, books, theses, data sets, computer programs: visualizations, simulations, and other models, multimedia publications, administrative records, published books, overlay journals, bibliographic datasets, images, audio files, video files, reformatted digital library collections, learning objects, web pages.

11. Can I export my digital material out of DSpace?

Yes. Currently DSpace supports exporting digital content, along with its metadata, in a simple XML-encoded file format. The DSpace developers are working on migrating this export capability to use the METS standard, but are waiting for some necessary extension schemas to emerge (such as one for qualified Dublin Core metadata, and one for minimal technical/ preservation metadata for arbitrary digital objects).

12. Will DSpace interoperate with other systems running at my organization?

Yes, DSpace has documented Java APIs you can customize to allow interoperation with other systems an institution might be running (for example, a department's web document system auto-depositing in DSpace, or a campus data warehouse).

13. What sort of persistent identifiers does DSpace use?

DSpace uses the Handle System from CNRI to assign and resolve persistent identifiers for each digital item. Handles are URN-compliant identifiers. The Handle resolver is an open-source system used in conjunction with DSpace. The developers chose to use handles instead of persistent URLs to support citations to items in DSpace over very long time spans – longer than we believe the HTTP protocol will last. Handles in DSpace are currently implemented as URLs, but can also be modified to work with future protocols.

14. How does DSpace preserve digital material?

DSpace identifies two levels of digital preservation: bit preservation, and functional preservation. Bit preservation ensures that a file remains exactly the same over time –not a single bit is changed– while the physical media evolve around it. Functional preservation goes further: the file does change over time so that the material continues to be immediately usable in the same way it was originally while the digital formats (and physical media) evolve over time. Some file formats can be functionally preserved using straightforward format migration, such as TIFF images or XML documents. Other formats are proprietary, or for other reasons are much harder to preserve functionally. No one can predict the formats all users will choose for their research material. They use the best tools for their purposes, and research institutions will get whatever formats those tools produce. For this reason, there are three levels of preservation for a given format: supported, known, or unsupported. Supported formats will be functionally preserved using either format migration or emulation techniques. Examples include TIFF, SGML, XML, AIFF, and PDF. Known formats are those that we can't promise to preserve, such as proprietary or binary formats, but which are so popular that third party migration tools will likely emerge to help with format migration. Examples include Microsoft Word and Powerpoint, Lotus 1-2-3, and WordPerfect. Unsupported formats are those that we don't know enough about to do any sort of functional preservation. This would include some proprietary formats or a one-of-a-kind software program. For all three levels, DSpace does bit-level preservation so that "digital archaeologists" of the future will have the raw material to work with if the material proves to be worth that effort.

15. Where can I find DSpace technical documentation?

You can find DSpace system documentation on the SourceForge project web site (http://sourceforge.net/projects/dspace/) or on the DSpace Technology page (http://www.dspace.org/technology/).

16. I've installed DSpace and I have questions/problems/comments. What should I do?

The DSpace Community of developers support one another and exchange ideas and solutions on the DSpace mailing lists. Before you post a question or problem, check to see if your question has been answered already. Start by searching the DSpace mailing list archives. Then read the FAQ and check the technical documentation. If you still haven't found an answer or solution, post your questions to DSpace-tech, where members of the DSpace community will offer their assistance.

17. Who supports DSpace?

DSpace has a very active community of developers to contribute expertise and support through the DSpace listserv at SourceForge, DSpace-Tech, and the project wiki. To work with the DSpace system you'll need local technical resources (hardware, technical experts, and so on) to really take advantage of the system. The DSpace web site offers technical documentation, and you can join the DSpace listserv, DSpace-Tech, to ask questions or post solutions.

18. What sort of hardware does DSpace require? What about sizing the server? How much disk space do I need?

There are no specific server requirements for DSpace except UNIX. (Because the application is written in Java, in theory it will run on other platforms as well.) DSpace is built on top of free, open-source tools, such as the Apache Web server, the Tomcat Servlet engine, and the postgreSQL relational database system. For your convenience, we package the necessary JDBC and other drivers and libraries together with DSpace. This set of tools should run on any UNIX-type OS, such as Linux, HP/UX, or Solaris, and you can substitute other libraries if you need to run on another platform. The system runs on anything from a laptop to a US$ 500K server, but there are a few general recommendations for hardware architectures. For a research university, DSpace requires a reasonably good server and a decent amount of memory and disk storage.

19.What is Metadata?

Metadata is literally "data about data." It is descriptive information used for querying. Some metadata can be done mechanically, such as file sizes, checksums, and full-text indexing, for instance. Other metadata is a higher order of human-made description such as titles, authors, unique identifiers, and abstracts. DSpace uses a qualified version of Dublin Core metadata across all content. Some communities or collections may also have tailored metadata available (such as MARC records for book collections, or FGDC records for geographic datasets). But even where that's available for some items, we crosswalk more detailed metadata records into our Dublin Core vocabulary to ensure a common layer of descriptive specificity for browsing and searching across everything.

20. What metadata standards does DSpace support? Can I create metadata using the [SCORM or VRA or FGDC or MARC or myOwnSchema]?

In this context support for a given metadata schema means that metadata can be entered into DSpace, stored in the database, indexed appropriately, and made searchable through the public user interface. This currently applies mainly to descriptive metadata, although as standards emerge it could also include technical, rights, preservation, structural, and behavioral metadata. Currently DSpace supports only the Dublin Core metadata element set with a few qualifications conforming to the library application profile. The DSpace team hopes to support a subset of the IMS/SCORM element set (for describing education material) in the coming year. HP and MIT also have a research project called SIMILE that is investigating how to support arbitrary metadata schemas using RDF as applied by the Haystack research project in the Lab for Computer Science and some of the Semantic Web technologies being developed by the W3C.

21. Does DSpace support OAI?

DSpace supports the Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH) v2.0 as a data provider. OAI support was implemented using OCLC's OAICat open-source software to make DSpace item records available for harvesting. DSpace@MIT is registered as a data

provider with the Open Archives Initiative. Other institutions running DSpace may choose to turn on OAI or not, and to register as a data provider or not.

22. More info on DSpace can be found here:

- *http://www.ideals.uiuc.edu/bitstream/2142/1043/3/DSpaceHowToGuide.pdf*

This short booklet is intended to introduce the commonest non-obvious customization-related tasks for newcomers to DSpace administration. It has been written against the stable version 1.4.2 of DSpace and the Manakin user interface, version 1.1. The authors have tried to include instructions for different operating systems as required; most customizations, however, work identically cross-platform.

- *http://cadair.aber.ac.uk/dspace/handle/2160/565*

DSpace Live CD (version 1.5) Burn this .iso file to a CD or DVD, and boot your computer using it. Log in using the username 'dspace' and the password 'dspace'. The DSpace Live CD allows to receive *'hands on'* experience of installing and configuring a DSpace repository.

15

Self-Archiving FAQ

Based on:

- Website: http://www.eprints.org/openaccess/self-faq/

What is self-archiving?

To self-archive is to deposit a digital document in a publicly accessible website. Depositing involves a simple web interface where the depositer copy/pastes in the *'metadata'* (date, author-name, title, journal-name, etc.) and then attaches the full-text document. Self-archiving takes only about 10 minutes for the first paper and even less time for all subsequent papers. Some institutions even offer a proxy self-archiving service, to do the keystrokes on behalf of their researchers. Software is also being developed to allow documents to be self-archived in bulk, rather than just one by one.

1. What is the Open Archives Initiative (OAI)?

The Open Archives Initiative (OAI) has designed a shared code for metadata tags (e.g., 'date', 'author', 'title', 'journal' etc.). The full-text documents may be in different formats and locations, but if they use the same metadata tags they become '*interoperable*'. Their metadata can be '*harvested*' and all the documents can then be jointly searched and retrieved as if they were all in one global collection, accessible to everyone.

2. What is OAI-compliance?

OAI-compliance means using the OAI metadata tags. A document can be OAI-compliant and an Eprint archive can be OAI-compliant. All OAI-compliant documents in OAI-compliant archives are interoperable. This means distributed documents can be treated as if they were all in one place and one format.

3. What is the purpose of self-archiving?

The purpose of self-archiving is to make the full text of the peer-reviewed research output of scholars/scientists and their institutions visible, accessible, harvestable, searchable and useable by any potential user with access to the Internet. The purpose of thus maximizing public access to research findings online is that this in turn maximizes its visibility, usage and impact –which in turn maximizes its benefits to research itself (and hence to the society that funds it) in terms of research dissemination,

application and growth, hence research productivity and progress. This is why Open Access is both optimal and inevitable.

4. What is the difference between distributed and central self-archiving?

All OAI-compliant archives are interoperable. This means their contents are harvestable by cross-archive search engines into global virtual archives. Hence OAI has eliminated the difference between self-archiving documents in one central archive or many distributed archives. Users need not know where documents are located in order to find, browse and retrieve them (any more than they do when they are using commercial indexing or abstracting services); and the full texts are all retrievable.

5. Who should self-archive?

The Budapest Open Access Initiative is focussed specifically on the refereed research literature, across all disciplines. It is the authors of these articles who should self-archive them, in order to maximize the visibility, accessibility, uptake and impact of their work.

6. What should be self-archived?

All significant stages of one's work, from the pre-refereeing preprint to the peer-reviewed, published postprint, to postpublication updates should be self-archived. The OAI tags keep track of all versions. (Note that the postprint need not be the publisher's proprietary PDF: there should always be a link to the publisher's official version, however, for scholarly purposes.)

7. Is self-archiving publication?

Self-archiving is definitely not publication. For purposes of establishing priority and asserting copyright, anything that is made public, even on a single piece of paper, meets the legal definition of '*publication*'. Hence so does self-archiving. But for scholarly and scientific purposes, only meeting the quality standards of peer review, hence acceptance for publication by a peer-reviewed journal, counts as publication. Self-archiving should on no account be confused with self-publication.

8. What about copyright?

The author holds the copyright for the pre-refereeing preprint, so that can be self-archived without seeking anyone else's permission. Sixty-eight percent of journals already give their green light to postprint self-archiving. With the remaining 32%, the author can either try to modify the copyright transfer agreement to reserve the right to self-archive the postprint, or, failing that, can append or link a corrigenda file to the already self-archived preprint.

9. What if the publisher forbids preprint self-archiving?

The right to self-archive the refereed postprint is a legal matter, because the copyright transfer agreement pertains to that text. But the pre-refereeing preprint is self-archived at a time when no copyright transfer agreement exists and the author holds exclusive and full copyright to that draft. So publisher policy forbidding prior self-archiving of preprints is not a legal matter, but merely a journal policy matter.

10. What can researcher/authors do to facilitate self-archiving?

Make sure that your university or research institution has installed OAI-compliant Archives. Self-archive your pre-peer-review preprints in your institutional (or central) Archives. Self-archive your post-peer-review postprints in your institutional (or central) Eprint Archives.

11. What can libraries do to facilitate self-archiving?

Digital librarians are the natural candidates for maintaining the Eprint Archives, their institution's outgoing collection of peer-reviewed research output.

- Offer trained digital librarian help in showing faculty how to self-archive their papers in the university Eprint Archive (it is very easy).
- Offer trained digital librarian help in doing '*proxy*' self-archiving, on behalf of any authors who feel that they are personally unable (too busy or technically incapable) to self-archive for themselves. Authors need only supply their digital full-texts in word-processor form: the digital

archiving assistants can do the rest (usually only a few dozen key/mouse-strokes per paper).

The proxy self-archiving will only be needed to set the first wave of self-archiving reliably in motion. The rewards of self-archiving –in terms of visibility, accessibility and impact– will maintain the momentum once the archive has reached critical mass. And even students can do for faculty the few keystrokes needed for each new paper thereafter.

- Digital librarians, collaborating with web system staff, should be involved in ensuring the proper maintenance, backup, mirroring, upgrading, and migration that ensures the perpetual preservation of the university Eprint Archives. Mirroring and migration should be handled in collaboration with counterparts at all other institutions supporting OAI-compliant Eprint Archives.

12. What can research funders do to facilitate self-archiving?

Mandate that the research that is publicly funded must not merely be published but it must be publicly accessible online (whether through self-archiving, open-access journals, or both) as recommended by the Berlin Declaration. Make it part of grant applications that CVs and bibliographies citing the applicant's prior work should contain links to the online full-text (whether self-archived or in open-access journals, or both). Sign the Declaration of Institutional Commitment to Providing OA .

16

Open Access archives: Examples

Based on:

-Websites:

http://arxiv.org http://en.wikipedia.org/wiki/ArXiv
http://hal.archives-ouvertes.fr/index.php?langue=en
http://www3.imperial.ac.uk/library/digitallibrary/digitalrepository

ArXiv e-Print archive

arXiv (www.arxiv.org) is an e-print service in the fields of physics, mathematics, non-linear science, computer science, quantitative biology and statistics. The contents of arXiv conform to Cornell University academic standards. arXiv is owned, operated and funded by Cornell University, a private not-for-profit educational institution. arXiv is also partially funded by the National Science Foundation. As of June 2008, arXiv.org contains over 483,000 e-prints, with roughly four thousand new e-prints added every month.

The arXiv was originally developed by Paul Ginsparg and started in 1991 as an archive for preprints in physics and later expanded to include astronomy, mathematics, computer science, nonlinear science, quantitative biology and, most recently, statistics. It soon became obvious that there was a demand for long term preservation of preprints. The term e-print was adopted to describe the articles. Ginsparg was awarded a MacArthur Fellowship in 2002 for his establishment of arXiv.

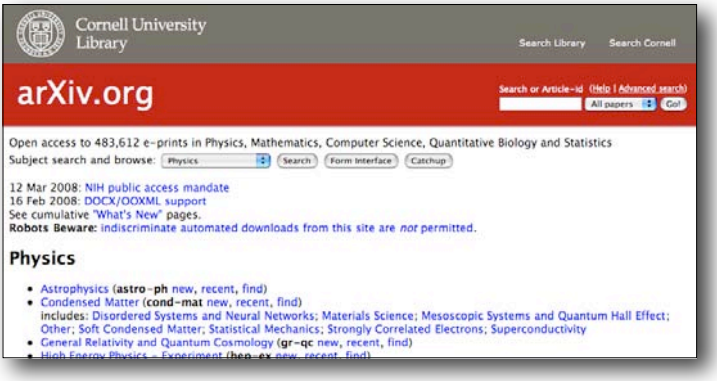


Figure: arXiv website (www.arxiv.org)

It was originally hosted at the Los Alamos National Laboratory (at xxx.lanl.gov, hence its former name, the LANL preprint archive) and is now hosted and operated by Cornell University, with mirrors around the world. Its existence was one of the precipitating factors that led to the current revolution in scientific publishing, known as the Open Access

movement, with the possibility of the eventual disappearance of traditional scientific journals. Professional mathematicians and scientists regularly upload their papers to arXiv.org for worldwide access and sometimes for reviews before they are published in peer reviewed journals.

Although the arXiv is not peer-reviewed, a collection of moderators for each area review the submissions and may recategorize any that are deemed off-topic. The lists of moderators for many sections of the arXiv are publically available but moderators for the mathematics section and for most of the physics sections remain unlisted. Additionally, an '*endorsement*' system was introduced in January 2004 as part of an effort to ensure content that is relevant and of interest to current research in the specified disciplines. Endorsement comes from either another arXiv author who is an endorser or is automatic, depending on various evolving criteria, which are not publicly spelled out. Endorsers are not asked to review the paper for errors, but to check if the paper is appropriate for the intended subject area.

The lack of peer-review, while a concern to some, is not considered a hindrance to those who use the arXiv. Many authors exercise care in what they post. A majority of the e-prints are also submitted to journals for publication, but some work, including some very influential papers, remain purely as e-prints and are never published in a peer-reviewed journal. Papers can be submitted in several formats, including LaTeX, and PDF printed from a wordprocessor other than TeX or LaTeX and DOCX from MS Office. For LaTeX, all files needed to generate the article must be submitted, in particular, the LaTeX source and files for all pictures.

Open Access services at ICTP: Scientific publications

The Publications Office of the Abdus Salam International Centre for Theoretical Physics (publications.ictp.it) processes research papers produced by authors during their stay at ICTP. These scientific technical papers, in diverse fields, are prepared in the form of preprints or reports. The office also manages the ICTP Lecture Notes Series (LNS), a series containing unpublished material presented at ICTP meetings. The sources

of these preprints and complete volumes of the LNS, are placed online for the benefit of the scientific community. It is hoped that this formally structured pedagogical material in advanced topics will be helpful to young students and researchers, in particular to those working under less favorable conditions.

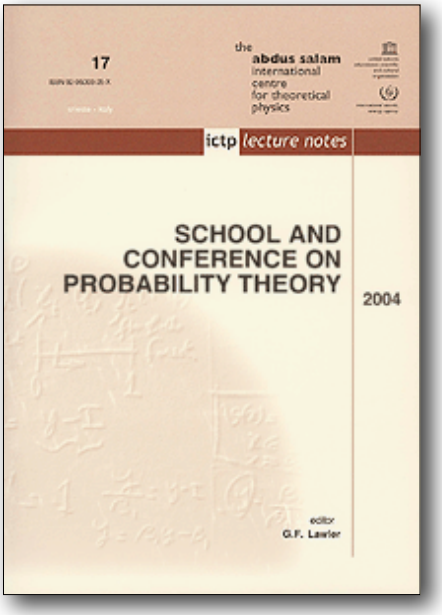
✦ *Sources of Preprints issued at ICTP*

Starting May 1995, the Publication Office has been offering an "electronic report" service containing the Title, Author(s), Abstract and ICTP Serial code of recent preprints issued at ICTP. Starting 1996 most sources are available on-line.

See: http://publications.ictp.it/preprints.html

✦ *Open Lecture Notes Series*

Since unpublished material presented at the meetings might prove of great interest also to scientists who did not take part in the schools, the Centre has made them available through: http://publications.ictp.it/lns.html



HAL: Hyper Article en ligne

HAL (http://hal.archives-ouvertes.fr) is designed for authors to deposit and thus make publicly available scholarly documents from all academic fields. These documents should be uploaded either by one of the authors with the consent of the others or by an authorized person on their behalf (information specialist or librarian, for example).

Figure: HAL

HAL is a tool for direct scientific communication between academics, managed by the *Centre pour la Communication Scientifique Directe* (CCSd, http://www.ccsd.cnrs.fr), a unit of the National Council of Scientific Research of French. A text posted to HAL should describe completed research work and should meet current scientific requirements in the field concerned, i.e., the content should be comparable to that of a paper that an investigator might submit for publication in a peer-reviewed scientific journal, conference proceedings, etc. A document deposited in HAL will not be subjected to any detailed scientific evaluation -simply a rapid overview, to ensure that it does indeed fall within the category defined above. The CCSd, however, remains free to decide whether to place the document online or not, without necessarily justifying its decision. An uploaded document does not need to have been published or even to be intended for publication -it may be posted to HAL as long as its scientific content justifies it. But should the article be published, contributors are invited to indicate the relevant bibliographic information and DOI.

HAL will ensure the long term preservation of a deposited document –this latter will be stored there permanently and will receive a stable web address. Thus, like any publication in a traditional scientific journal, it can be cited in other work.

The free online access to these documents provided by HAL is intended to promote the best possible dissemination of research work; the intellectual property remains that of the author(s). Contributors must abide by the rules of good usage prevailing in scientific publications -respect and citation of original work, no intellectual plunder, etc. Any document uploaded to HAL which is relevant to a scientific field that exists also in arXiv will be automatically duplicated in this latter base (unless the depositor specifically states otherwise).

Finally, HAL also offers services of a more administrative nature, such as the easy extraction of lists of publications (for an author, laboratory, institution, etc.) in various formats. When a full text file of a published article is not available and in order for these lists to be as complete as possible, a simple 'note' that contains merely the bibliographical references may be added to HAL.

Spir@l: Imperial College digital repository

Spir@l (http://spiral.imperial.ac.uk) is a College-wide digital repository designed to hold the College's research publications. The project, led by the Imperial College Library has been funded to carry out the preliminary tasks of set up, design and configuration of the system to hold electronic copies of academics' publications.

Spir@l uses software called DSpace for the management and accessibility of its digital repository, although users need to use the College Publications system for uploading files. The digital repository's content is Open Access and freely available to all.

The aim of the Repository is:

- to be a central point for collecting research publications across the College;
- to increase the College's research publications' visibility;
- to allow Open Access to all deposited work;
- to provide links to academics' professional web pages.

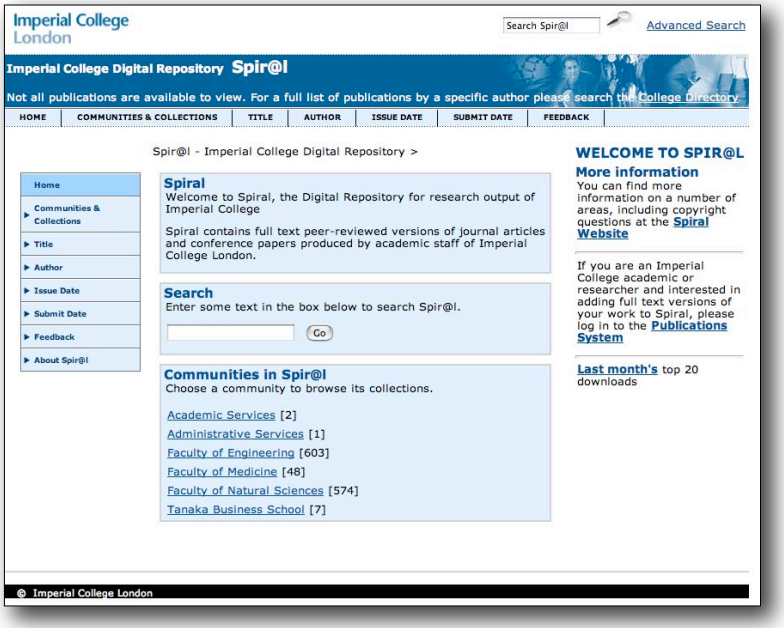


Figure: Spir@l

PubMed Central

PubMed Central (PMC, www.pubmedcentral.nih.gov) is a free digital archive of biomedical and life sciences journal literature at the U.S. National Institutes of Health (NIH), developed and managed by NIH's National Center for Biotechnology Information (NCBI) in the National Library of Medicine (NLM). With PubMed Central, NLM is taking the lead in preserving and maintaining unrestricted access to the electronic literature, just as it has done for decades with the printed biomedical literature. PubMed Central aims to fill the role of a world class library in the digital age. It is not a journal publisher. NLM believes that giving all users free and unrestricted access to the material in PubMed Central is the best way to ensure the durability and utility of the archive as technology changes over time.

PMC follows in the footsteps of other highly successful and useful services that NCBI has developed for the worldwide scientific community: GenBank, the genetic sequence data repository, and PubMed, the database of citations and abstracts to biomedical and other life science journal literature. GenBank, and the tools provided by NCBI for searching and manipulating its contents, have been a boon to molecular biologists and have helped advance developments in the field. PubMed (which encompasses Medline) is the database of choice, for researchers and clinicians alike, to locate relevant articles and, in many cases, link directly to a publisher's site for the full text.

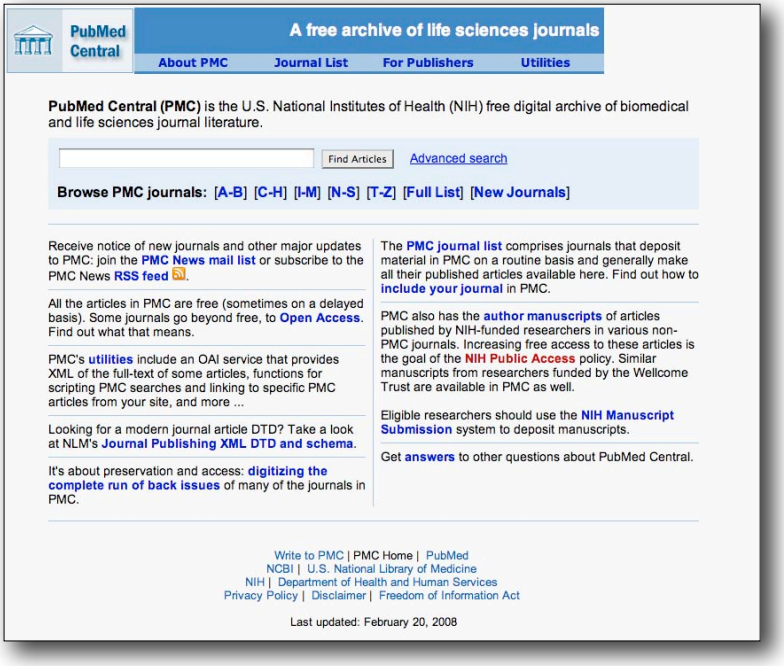


Figure: PubMed Central

Participation by publishers in PMC is voluntary, although participating journals must meet certain editorial standards. Journals are encouraged to deposit all their content (and not just research papers or other selected material) in PMC so that the archive becomes a true digital counterpart to NLM's extensive collection of print journals. In line with this objective,

NLM is digitizing earlier print issues of many of the journals already in PMC. Although immediate free access to all content is most desirable, a journal may delay release of its full text in PMC for some period of time after publication.

A journal is guaranteed access to a copy of its deposited data upon request, at no cost. PubMed Central does not claim copyright on any material deposited in the archive. Copyright remains with the journal publisher or with individual authors, whichever is applicable.

The value of PubMed Central, in addition to its role as an archive, lies in what can be done when data from diverse sources is stored in a common format in a single repository. GenBank has proven the advantages of collecting DNA sequences in a central repository with a common format. You get more rapid searching, manipulation, and cross-linking of the complete collection, and all the benefits that derive from that. Similarly, with PubMed Central, one can quickly search the entire body of full-text articles and locate relevant material regardless of its source. It also makes it possible to integrate the literature with a variety of other information resources such as sequence databases and other factual databases that are available to scientists, clinicians and everyone else interested in the life sciences. The intentional and serendipitous discoveries that such links might foster excite us and stimulate us to move forward.

PMC Open Access Subset

The PMC Open Access Subset is a relatively small part of the total collection of articles in PMC. Articles in the PMC Open Access Subset are still protected by copyright, but are made available under a Creative Commons or similar license that generally allows more liberal redistribution and reuse than a traditional copyrighted work.

17

An international open archive: E-LIS

Based on:

- *A. De Robbio and I. Subirats Coll, "E-LIS: an International Open Archive Towards Building Open Digital Libraries" High Energy Physics Libraries Webzine, issue 11, August 2005.*
  Website: http://library.cern.ch/HEPLW/11/papers/1/

Established in 2003, *e-Prints in Library and Information Science* (E-LIS, eprints.rclis.org) is an international, Open Access archive related to librarianship, information science and technology, and related disciplines, in keeping with the objectives of the e-prints movement (see: http://www.ecs.soton.ac.uk/~harnad/Tp/nature4.htm) and the Free Online Scholarship (FOS) movement (see: http://www.earlham.edu/~peters/fos/fosblog.html).

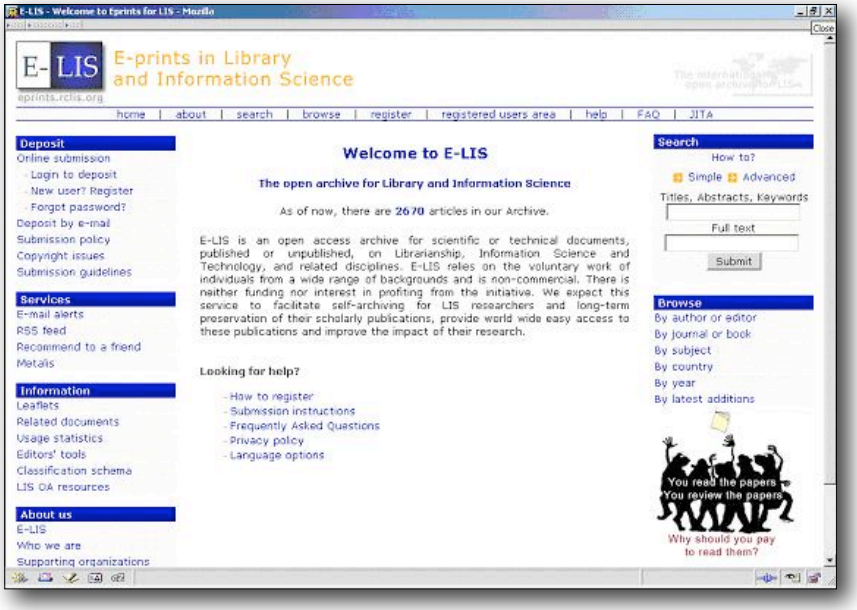


Figure: E-LIS

To date E-LIS is the biggest repository in LIS and after two years contains over 2200 papers. The E-LIS model is based on community standards which provide the mechanism for enforcement of proper attribution and responsible use of published works in line with the two above movements. On the technical level, it is based on the Open Archives Initiative (OAI) and shares its standards and protocols. E-LIS is the first international e-server in this area and is part of the RCLIS (Research in Computing, Library and Information Science) project (see: http://rclis.org).

Overview

The purpose of the E-LIS archive is to make full-text documents visible, accessible, harvestable, searchable and useable by any potential user with access to the Internet. It also aims to support individuals who wish to publish or make their papers available worldwide and it can be used by LIS communities in any country.

Works can be deposited in any language and format; authors can self-archive, and a proxy service supports depositors. The archive is open to proposals from new partners and it has agreements with institutions and library associations of various countries.

E-LIS is organized, managed and maintained by an international team of librarians working on a voluntary basis. The flexible architecture of the OAI is sub-divided into data provider and service provider. A data provider maintains one or more repositories (web servers) that support the OAI protocol as a means of exposing metadata. A service provider issues requests to data providers and uses the metadata as a basis for building value-added services. By self-archiving, authors can store a copy of their documents in a disciplinary or an institutional archive. As the OAI architecture enables documents to become swiftly available worldwide, self-archiving helps to maximize the visibility and accessibility of refereed research, and hence to maximize its usage by researchers and its impact on research.

Self-archiving systems can be either centralized or distributed and are usually subject or institution-based. At present there are several subject-based archives, including ArXiv (http://arxiv.org), CogPrints (http://cogprints.org) and E-LIS, which are centralized, and RePEC (www.repec.org), which is distributed.

There are also many institution-based archives but the advantage of subject-based archives, like E-LIS, is that they are specific to discipline needs and requirements. Many repositories use the OAI protocol and tools to facilitate interoperability between repository servers.

The main characteristics of a centralized archive such as E-LIS are:

- storage and distribution of data from a single location;
- centralized access control over the supply and re-use of data;
- checking, cleaning and processing of data according to standard criteria;
- centralized support service for describing the contents of the data, defining the principles and practices governing the collection of the data, and other relevant properties of the data;
- cataloguing of the technical and substantive properties of the data for information retrieval and,
- user support following the supply of data.

E-print repositories are complementary to, rather than a replacement for, scholarly journals. The peer-review process provided by journals is of critical importance to scholarship. However, the proliferation of institution- and discipline-based e-print repositories could accelerate changes that are currently taking place in the scholarly communication process which in turn could increase the number of journals moving from toll access (reader pays/subscription-based model) to Open Access (author-pays/free access model).

The evolving academic discourse surrounding the concept of Open Access provided the underlying motivation for the establishment of E-LIS. The library and information world is highly integrated with the areas of computing science and technology and it was felt that the LIS discipline should set an example to other communities by providing a state-of-the-art model for the OA movement and digital libraries, particularly in relation to the open archive model, within which E-LIS is a disciplinary repository.

The extension of the OA concept to LIS works and the dissemination of material within the LIS community will contribute to the development of an international LIS network; E-LIS is mutually beneficial. For librarians, metadata creation is costly and the growing trend of authors to self-archive

in the OAI framework is proving an effective way to reduce some of those costs. For librarians as authors, archiving their work in E-LIS gives them an increased understanding of the process of self-archiving and the E-LIS archive ensures data preservation and a wide data visibility in addition to facilitating active participation in the international librarian community.

E-LIS defines Open Access as a property of individual works and it adheres to the Bethesda Statement which states that publication must meet two conditions:

1. The author(s) and copyright holder(s) grant(s) to all users a free, irrevocable, worldwide, perpetual right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship, as well as the right to make small numbers of printed copies for their personal use.
2. A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in a suitable standard electronic format is deposited immediately upon initial publication in at least one online repository that is supported by an academic institution, scholarly society, government agency, or other well-established organization that seeks to enable Open Access, unrestricted distribution, interoperability, and long-term archiving.

The E-LIS organizational model

Authors who contribute to an e-print archive are participating in a global effort by universities, researchers, libraries, publishers, editors, and readers, to redefine the mechanisms of scholarly communication. E-LIS will make LIS research more visible and accessible, which in turn will increase its status and public value.

E-LIS is divided into three sections: administrative, editorial and technical. A discussion list for each section provides the basis for action:

- the administrative section deals with strategic issues including the future direction of the initiative, its policies and their impact on the user community, and its connection with other scientific communities;
- the editorial section is devoted to metadata quality and guidelines;
- the technical section concentrates on the software -its implementation, enhancement, development, added-value functionality, and its operation within the OAI framework.

These sections provide the over-arching structure within which the methods and procedures followed by E-LIS staff are established, maintained and developed.

Users perceive two main structures around which all the contents are organized; both structures have two levels. The first structure is the classification scheme for LIS (called JITA see: http://eprints.rclis.org/jita.html) divided into main subject divisions and subdivisions; the second is a geographical ordering based on continent and country. Users can navigate these structures to retrieve the required information. They can also view documents by browsing by author or year. Information can also be accessed via the search interface through a basic or an advanced search. In order to improve the search mechanisms, E-LIS has enhanced the full-text search furnished by the e-prints software it uses.

The submitter of a document is required to assign subjects and keywords from the E-LIS classification scheme: JITA. JITA has recently opened a second-level, browse-by-country view which has over 120 sub-sections. This simple scheme, which has also been adopted by other repositories, is the result of the fusion and re-arrangement of the NewsAgentTopic classification scheme and the Review of Information Science (RIS) classification scheme. The JITA classification scheme is not intended to be a comprehensive classification scheme, but to facilitate document retrieval through the archive's browsing facility. It is divided into twelve blocks (categorized alphabetically from A-L) which have been created around the three following implicit (virtual) areas:

1. *Theory and generalities (general level)*. This is divided into: theoretical and general aspects of libraries and information; information use and the sociology of information.
2. *User-oriented, directional, and management functions (intermediate level)* –socio-economic and legal issues are included here. This divides into: users, literacy and reading; libraries and information repositories; publishing and legal issues; management; industry, profession and education.
3. *Objects, pragmatic issues and technicalities (on a specific level)*. This covers: information sources, supports and channels; information treatment for information services; technical services in libraries, archives and museums; housing technologies; information technology and library technology.

Since April 2005 it has been possible to browse by country. This gives a truly international aspect to the archive and is particularly aligned with the organization of the editorial board whereby work is channelled through international staff on an individual country basis.

Strategic issues

The core of the organizational model is the administrative section which also has responsibility for determining the international and future vision of the archive, keeping in mind an understanding of national and international needs. In some countries, librarians want to create national archives for LIS instead of inserting papers into an established international archive, such as E-LIS. This is understandable because there are many technical difficulties associated with the use of different languages, alphabets, and non-alphabetic languages and with the consequent problems of input, output and the sorting of data and author names. E-LIS co-operates with each country individually to decide the best solution to the technical and non-technical barriers so that international visibility can be promoted whilst national interests are served. One idea, currently being debated, is to create a service provider with a harvester to gather national metadata from the national LIS archives and which could be part of the RCLIS infrastructure.

Many countries are currently involved with E-LIS and each of them works towards a common vision while maintaining their own cultural individuality. The editors are entrusted with the responsibility of promoting E-LIS within their own country and every editor has a different approach which is tailored to the specific needs of that country. The status of librarianship, which differs significantly from country to country, is the result of several factors such as the role of the professional association or of the LIS School, institutional assets, initiative on the digital library front, the impact of LIS in the social environment and how much OA has become a part of scholarly communication. However, all the editors share a common vision, and bring their own disciplinary and personal experience to bear on the position.

In E-LIS, papers can be viewed on an individual country basis, which highlights the internationality of the project, in addition to facilitating user access. In the LIS community most published papers traditionally came from the US, UK and Europe, but in E-LIS, where OA encourages a wider audience, it has been noted that high quality papers now come not only from the countries believed most innovative, but also from Developing Countries, such as India, and others in Africa, Asia, Eastern Europe, the Russian States and also Central and South America. This has been a very welcome, if unforeseen, development. The sense of this inclusiveness is highlighted by the recent collaboration with Cuban librarians, who have deposited papers in E-LIS from ACIMED, the Journal of Information Professionals in Health. The Open Access vision of E-LIS can be juxtaposed against the blockade imposed by the United States on Cuba, which is designed to prevent the publication of Cuban papers in the main U.S. journals. The Cuban case illustrates that E-LIS knows no borders, and negotiations are in progress regarding the involvement of Arabic countries, Israel and China. Collaboration with other countries has been very beneficial for E-LIS. In particular, Indian librarians have demonstrated a very high competency in bibliometric issues which has proved very stimulating for the E-LIS community as a whole.

E-LIS policies

E-LIS is driven and directed by its policies which determine its identity, quality and direction. It is not sufficient to create an archive merely by putting software on a machine, particularly in the case of E-LIS: an archive's organizational model is the sum of its policies and an archive without policies is like a library without a librarian. The principal policy concerns for E-LIS, which are discussed and democratically agreed upon by the editorial staff, are its:

1. *mission*: Its aims and objectives, what it is and where it is going, its target audience and the communities involved. Its international mission, (previously discussed) is clearly outlined on its website;
2. *submission policies*: who can deposit material and how such deposits should be made (described below);
3. *copyright policies* (described below) are fundamental to any repository and are assigned due importance in E-LIS;
4. *organizational model* (described above), which is the core of any repository and determines its institutional or disciplinary nature.

Submission policy

Every librarian in the world can deposit papers in E-LIS as it promotes intellectual freedom and maximization of impact in the LIS discipline. With E-LIS, as with arXiv in the physics field, experience has shown that this freedom has not led to the submission of low quality papers because authors are aware that their papers will ultimately be judged on a virtual basis by a very wide community of peers. The technical submission requirement is that authors who wish to submit a document must register in order to obtain a user I.D., which is also the basis for obtaining author-view browsing. Librarians, libraries, research institutes, organizations, and individual researchers involved in LIS and related fields are encouraged to make use of, and contribute to, the expansion of the archive and in turn to the critical mass of information available and useful for the building of digital libraries.

The E-LIS submission policy states that the archive accepts any scientific or technical document, published or unpublished, on librarianship, information science and technology or related activities. In this context, categories for different types of material have been created with respective sets of metadata. The criteria for acceptance are that the e-prints are relevant to research in LIS fields, and that they have the form of a finished document ready to be entered into a process of communication. Publications may include preprints, postprints, conference papers, conference posters, presentations, books, book chapters, technical reports/ departmental working papers, theses, and newspaper and magazine articles.

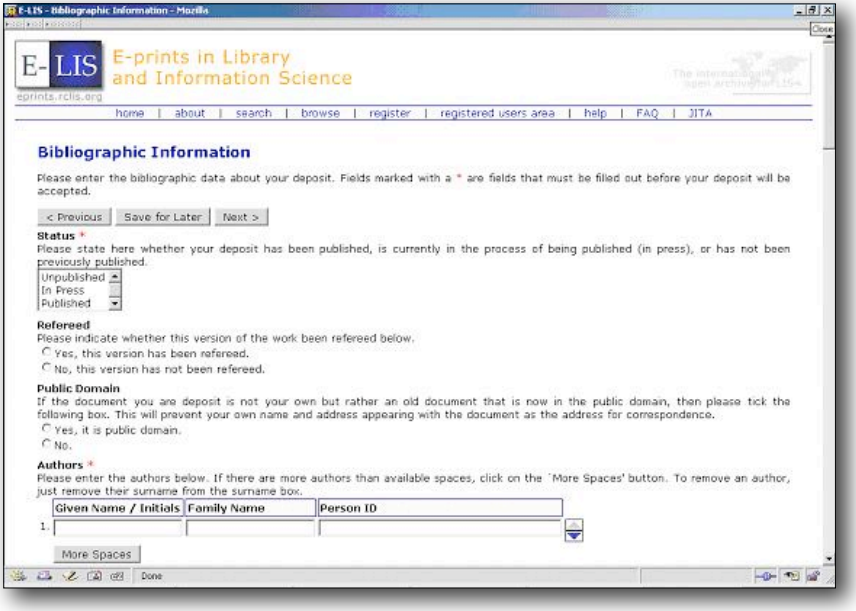


Figure: Bibliographic information in E-LIS

Submitted documents are placed in the submission buffer for approval or rejection by E-LIS staff. Rejection is on the basis of pertinence to the archive. E-prints may also be returned to the author for metadata modification or if there are problems with the electronic file or format. Editors may make formal corrections but they do not make substantial

modifications to the text or contents of the registered data. Documents in the submission buffer are manually reviewed to ensure they conform to E-LIS policy. E-LIS staff control the metadata quality of the document and are allowed to make changes if the metadata is incorrect. Generally, a paper becomes publicly accessible two working days after its deposit in E-LIS. The approval process is conducted by the editor from the country of the depositor. Once submitted, a document cannot be removed from the repository. Even when changing institution, Stevan Harnad argues (see: http://www.ecs.soton.ac.uk/~harnad/Hypermail/Amsci/4112.html) that "Wanting to remove one's work from the [repository of the] old Institution is [as] absurd as wanting to remove it from the shelves of one's old library -or any library". Indeed, removal would seriously hamper the communication process and users would be unable to chart the evolution of the idea or debate. Furthermore, because documents remain in the archive for posterity, the submission of inferior work is thus discouraged.

As an international open archive, E-LIS supports all languages; however, if a document is in a language other than English, it must include an English abstract and English keywords. If the English abstract is missing, the editor inserts it on behalf of the author. The following document formats are accepted: PDF, PostScript, TeX, LaTeX (DVI), HTML, XML, ASCII (text), PowerPoint, MS Word DOC and RTF. The use of HTML and PDF formats are strongly recommended.

Copyright policies

E-LIS does not want to infringe copyright. OA is encouraged where possible, but authors can restrict access to their papers if necessary. Access can also be restricted to the group of registered users of the E-LIS archive, which is a limited and known group of people. Furthermore, access can, in special cases, be restricted to only the depositor and archive staff (the archive administrator and any selected editors, evaluators, etc. in particular countries and organizations).

All work residing on the E-LIS server remains the property of the author. The author holds the copyright for the pre-refereed preprint and therefore, it can be self-archived without any other permission being sought. An

author's works are that author's own intellectual property and they therefore own copyright and other proprietary rights until and if they grant otherwise. Authors submitting to the repository are responsible for ensuring the documents they archive do not have any restrictions on their electronic distribution imposed by a third party (such as a publisher). A pre-refereed preprint can be self-archived at a time when no copyright transfer agreement exists and so the author holds exclusive and full copyright; the author may no longer have the right to self-archive a refereed postprint if a copyright transfer agreement has been signed granting all rights to the publisher.

In general, when an article is published in a journal, copyright is transferred to the publisher. Most journals permit self-archiving of the preprint and sometimes also the postprint, but it depends on the publisher's copyright policy. To avoid infringing any copyright, authors can deposit a postprint inside the archive with restricted access. Another way round this problem is for the author to request that the publisher allow them to retain certain rights, e.g. the right to deposit the postprint in an Open Access archive, or to place a copy on their homepage. Some publishers have stated that they grant these rights as a standard procedure. Alternatively, authors can replace the full text of the preprint with a link to the published version, if it is freely accessible.

Copyright law gives to the creator of a work exclusive rights, which may be both segmented and transferred to others. Publishers have adopted various policies to facilitate author self-archiving. To help authors and editorial staff establish the copyright situation for deposit of a particular item, E-LIS uses the SHERPA (www.sherpa.ac.uk) database of publisher copyright policies and self-archiving. SHERPA is a project investigating key issues in creating, populating and maintaining e-print collections, including specifically: Intellectual Property Rights (IPR), quality control, collection development policies, business models, scholarly communication cultures, and institutional strategies. In addition, E-LIS is aligned with the Rights Metadata for Open Archiving (RoMEO) project which cites the E-LIS policy on metadata (http://eprints.rclis.org/copyright.html). E-LIS's metadata policy permits third parties to collect

metadata from the archive via mechanisms that create end-user services to support the discovery and presentation of the archive content. The general policy is to allow harvesting of metadata but not the harvesting of full content.

Whatever the policy, the authoritative document for a published paper is the copyright agreement signed with the publisher. Fortunately, many publishers are adapting to the changing environment of electronic publishing. Nevertheless, some publishing polices contain ambiguities or have different approaches to self-archiving.

In an OA repository such as E-LIS, a user may find an exact copy of a paper from a journal which is normally only sold by a subscription. When a publisher's archiving policy is quite open it is possible to place a version (preprint or postprint) of a published paper in a repository. A user who knows how to use repositories can find high quality material in a free-access database, instead of acquiring it via a payment. Occasionally, a paper can be found in an aggregator site through a pay-per-view model (on average at US$30 per article), but in some cases papers are only found in electronic journals, some of which are "Open Access" and some "toll access". Some very high quality papers published in toll access LIS journals are available freely within an Open Access repository but it is imperative that archiving authors have a clear understanding of the different archiving policies adopted by publishers.

Editorial section

All work performed by the editorial section is developed by the editorial staff from discussion on a mailing list. Topics include metadata issues, guidelines for cataloguing, promotion of E-LIS and OA in general, and questions which arise from international co-operation. The principal editorial aim is to reflect the best practices of librarianship in each country by inviting the top scholars in the discipline to contribute to E-LIS. Therefore, the choice of editor for a particular country is crucial as they must be thoroughly conversant with the debates and personalities involved in the LIS disciplines in their country and they must also have the dynamism to promote E-LIS and the commitment, talent and patience for

organizing people, events and documents. Again, it must be stressed that this activity is currently done on an entirely voluntary basis, but it is evident that it has been of immense benefit to the LIS discipline as a whole. International co-operation can facilitate debate on current issues on many levels and provides the editors, on a personal level, with new professional experiences and expertise. Editorial tasks include:

- creating and maintaining contact with university-based LIS academics and researchers and LIS-related course personnel;
- approaching LIS publishers, with a view to obtaining permission to add articles from journals (both those already freely available on the web and those still using authentication systems) to E-LIS;
- sending publicity e-mails to national mailing lists and submitting press releases to web-based publications;
- writing literature reviews regarding Open Archive activities and writing journal articles on the E-LIS initiative.

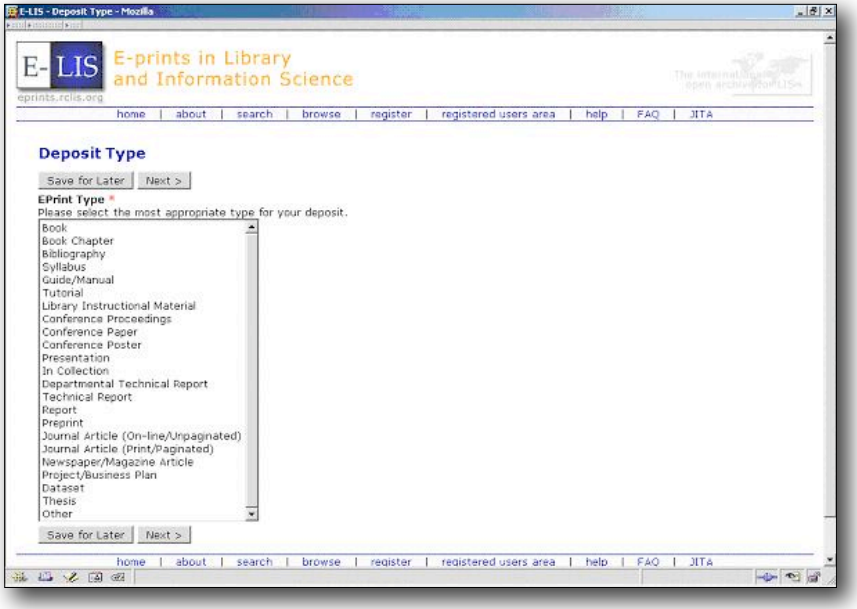


Figure: E-LIS, selection of e-print type

Fundamental to obtaining high metadata quality are the repository's '*guidelines for cataloguing*' which were created to facilitate editorial work and standardize metadata. The purpose of having the guidelines is to ensure that E-LIS adopts recognized, established practices for the creation of records. These guidelines assist the editors in assigning the correct value for each metadata field and were created taking into account the following points:

- types of documents accepted;
- variety of contents accepted;
- who is submitting the material;
- metadata standards;
- advice concerning file formats and preservation of documents.

Each of the twenty-three document types has its own set of metadata which is checked in accordance with the editorial guidelines set by the international editorial committee. The twenty three document types are shown in the figure. In addition to a common metadata core (Dublin Core) there is also a specific configuration of metadata for each category of these document types.

18

Open Journals System

Based on:

- K. Stranack, "Starting a New Scholarly Journal in Africa", Public Knowledge Project, 2006.
Available at: http://pkp.sfu.ca/files/AfricaNewJournal.pdf

The Open Journals System (OJS, http://pkp.sfu.ca/?q=ojs) is free, open source publishing management software, being used by over 1400 journals (as of March 2008) in ten languages around the world. While primarily developed to support Open Access, electronic publishing, it can be used for subscription-based, print publications as well. OJS was designed to minimize the technical skills required to operate a journal, and can be easily used by anyone familiar with word processing and sending email. It seeks to improve the scholarly and public quality of journal publishing through a number of innovations, from making journal policies more transparent to improving indexing.

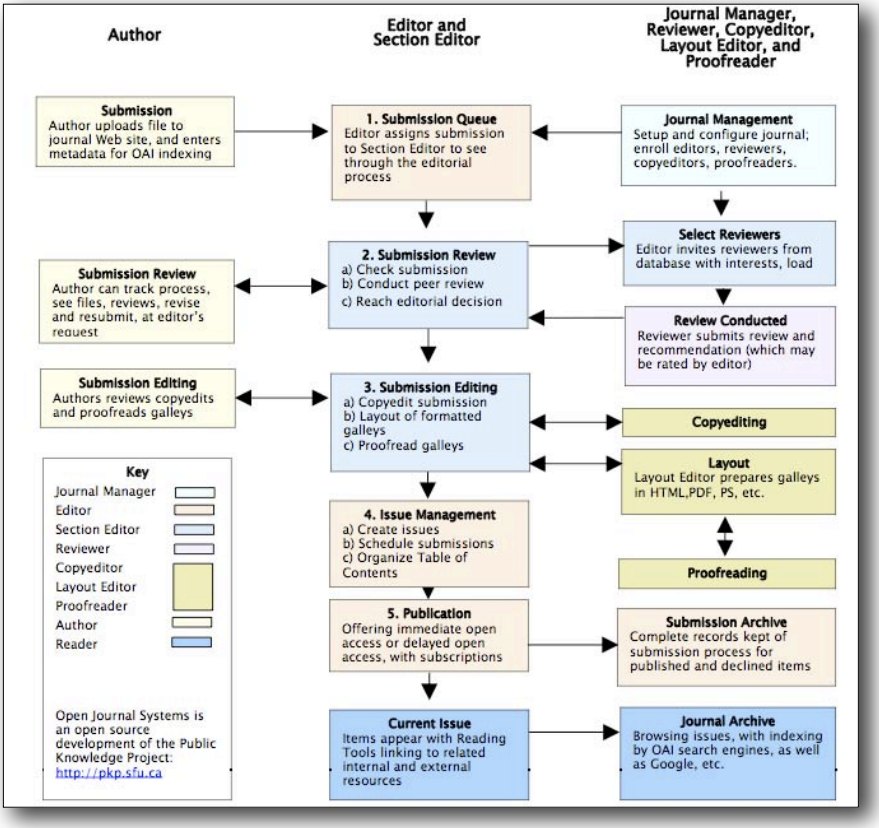


Figure: Structure of OJS

OJS is based on several roles (or responsibilities), including the Journal Manager, the Editor, the Section Editors, Subscription Manager, Layout editors, Copyeditors, Proofreaders, Reviewers, Authors, and Readers. One person can, and often does, take on more than one role. For example, the Editor may also be a Reader of the journal, as well as a contributing Author. An important feature of software like OJS is that it allows a community of editors, authors, reviewers, and readers to collaborate on a journal without ever needing to be in the same location. A journal's production team can share the responsibilities of publishing despite being located in different offices, institutions, cities, or even countries. This chapter will briefly follow a submitted article throughout the system, outlining the various roles performed, until the article reaches its ultimate publication. For more detailed information on using OJS, consult OJS in an Hour, available at http://pkp.sfu.ca/files/OJSinanHour.pdf

Step 1: The Journal Manager

All aspects of journal management are controlled by the Journal Manager, in consultation with the editors, from setting up and configuring the software to overseeing the overall operation of the journal. Journal management begins with setting up the journal. This includes configuring the software, establishing the web site, creating user accounts, and assigning them to their various roles. None of this, however, requires any advanced technical skills, but primarily involves filling out online forms and uploading files. Much of the work flow in OJS is laid out in easy to follow steps. To create the journal's web site, the Journal Manager will follow the "Five Steps to a Journal Web Site". These steps include:

1. details, which includes adding the journal's name, ISSN, mailing address, and other related information;
2. policies, such as the peer review and privacy policies of the journal;
3. submissions, including the author guidelines for being published;
4. management, involving the workflow of the journal, and the publication schedule; and
5. the Look, allowing you to add customized logos, headers and footers, navigation items, and change the background colour and other visual elements of the site.

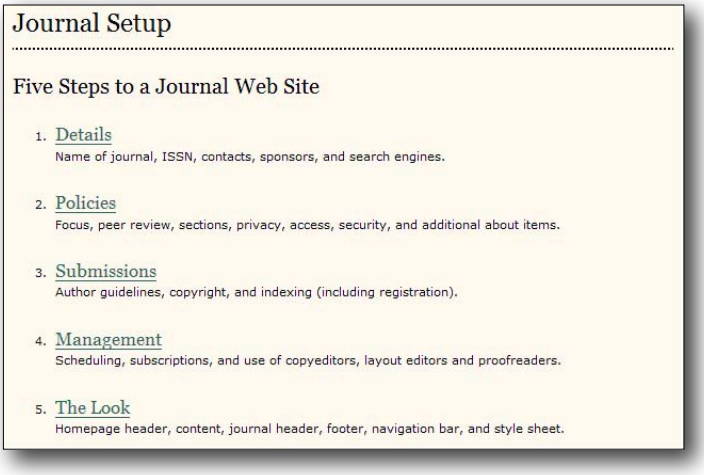


Figure: Five steps to a journal web site

The Journal Manager can also create new sections for the journal (such as Reviews, Articles, Commentaries, etc.), edit the text of the default set of email templates the system uses for communication between various participants, manage the Reading Tools that are available with this journal, and view the Statistics the system can generate.

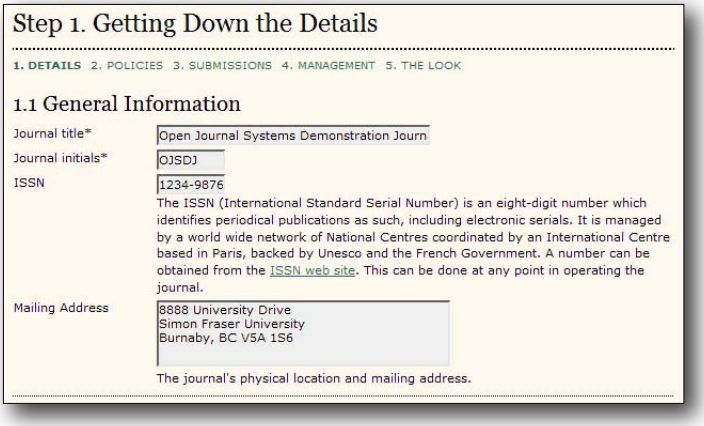


Figure: Filling in the forms

If the new journal is not Open Access, the Journal Manager may be assisted by, or also take on the role of the Subscription Manager, to set up the subscription policies of the journal, and to keep track of accounts, payments, and renewals.

Figure: Managing Subscriptions

Once the journal has been set up, the role of the Journal Manager is minimal, largely limited to making any future configuration changes. Often the Journal Manager will then take on the role of the Editor.

Now that the new OJS journal is in place, the Journal Manager or the Editor should put out a call for submissions, inviting authors to send in their papers, and begin promoting this new publication. The call for submissions can be added to the front page of the journal's web site, but should also be announced as widely as possible, including on relevant professional mailing lists and other scholarly forums.

Step 2: The Author

Authors are able to register and submit items to the journal directly through the journal's web site.

The Author is asked to upload their article, agree to a checklist of requirements as determined by the Journal Manager, and to provide metadata or indexing information on the article, such as their name and the names of any co-authors, the article title, an abstract, subject terms, etc. (The metadata improves the ability of others to search for the article).

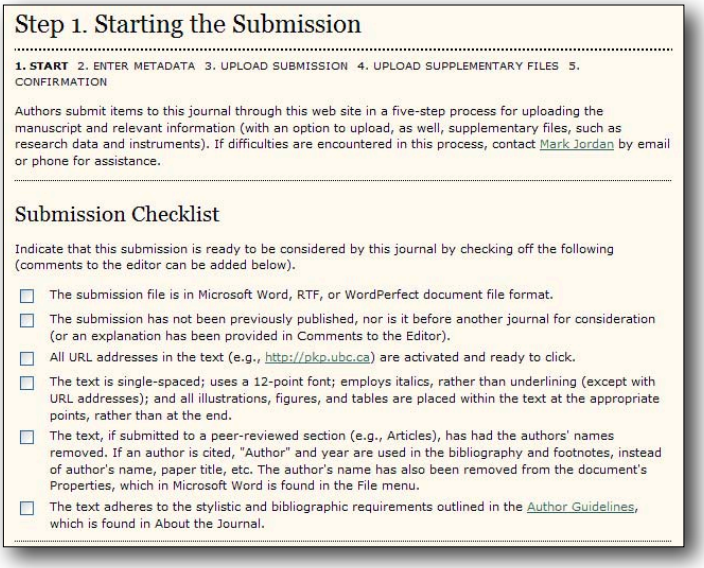


Figure: Submitting an article

The Author can upload Supplementary Files, such as data sets, research instruments, or source texts that will enrich the understanding of the article, as well as contribute to more open and robust forms of research and scholarship. The Author is able to track the submission throughout the editorial process –as well as participate in the copyediting and proofreading of submissions that have been accepted for publication– by logging in to the journal's web site.

Step 3: The Editor

The Editor oversees the entire editorial and publishing process. Working with the Journal Manager, the Editor typically establishes the policies and procedures, which are used in setting up the journal. In the Editorial Process, the Editor assigns submissions to the Section Editors to see through Submission Review and Submission Editing.

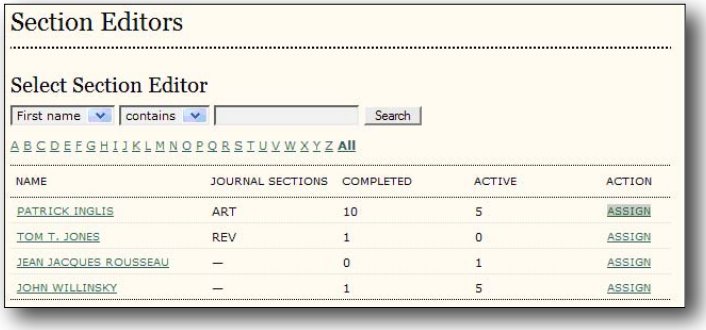


Figure: Assigning a section editor

The Editor keeps an eye on the submission's progress and assists with any difficulties in the process. Often, the Editor also plays the role of Section Editor in the Editing process (seeing accepted submissions through copyediting, layout, and proofreading). Larger journals with many staff or volunteers, however, may have separate Section Editors responsible for each section of the journal, such as for Peer-Reviewed Articles, Book Reviews, Commentaries, etc.

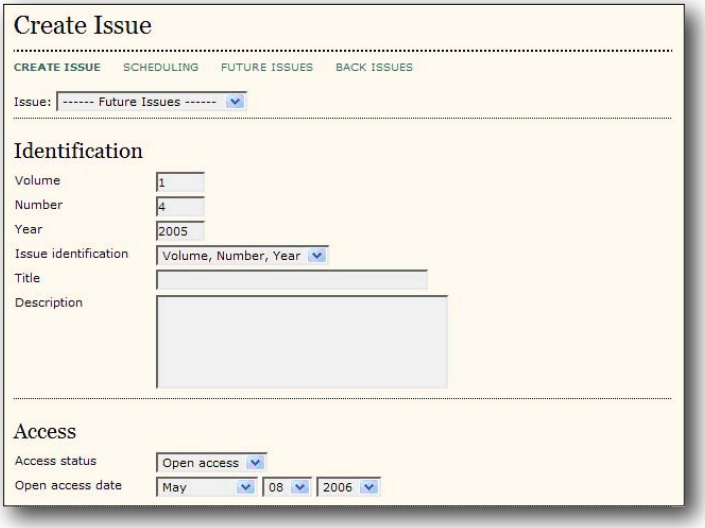


Figure: Creating an issue

The Editor also schedules the completed submissions for publication, arranges the Table of Contents and creates each new issue as part of the Publishing Process.

Step 4: The Section Editor

Having received responsibility for a new submission, the Section Editor assigns it to one or more reviewers and manages the Review and Editing process.

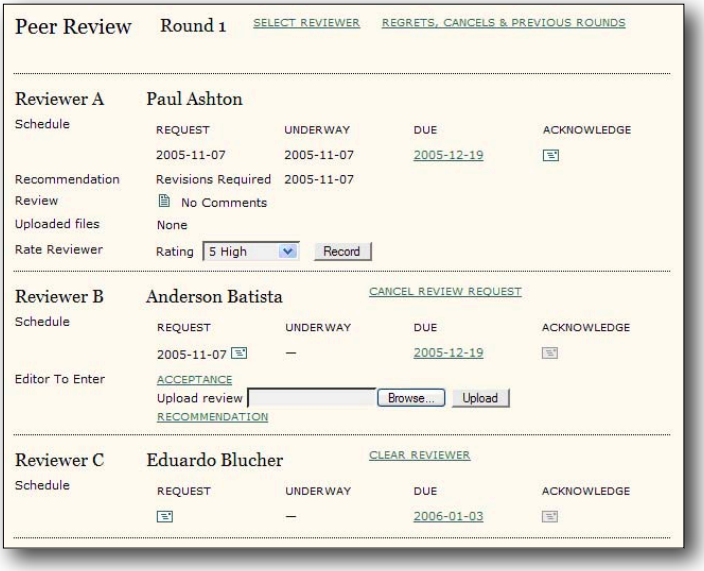


Figure: Managing the review process

Once the Review Process is completed, and if the submission is accepted for publication, the Section Editor will also be responsible for seeing the submission through the Editing Process (that is, through copyediting, layout, and proofreading). In some journals, however, Section Editors only work with the Review process, and an Editor, acting in the role of Section Editor, sees the submissions through the Editing process. Each journal will have a policy on how the tasks are divided, depending on the number of people participating, and based on their own unique best workflow practices.

Step 5: The Reviewer

To begin the Review Process, the Section Editor invites one or more Reviewers to review the submission. The Reviewer can decide whether to accept or decline the invitation from the Section Editor. Reviewers are asked to submit their reviews using the step by step OJS publishing system.

OJS does also allow for alternative methods of reviewing articles. Some journals opt for an email review policy, which allows for the review process to take place using documents attached to email messages. This can also be adapted to make use of photocopying and regular mail when electronic means are not available.

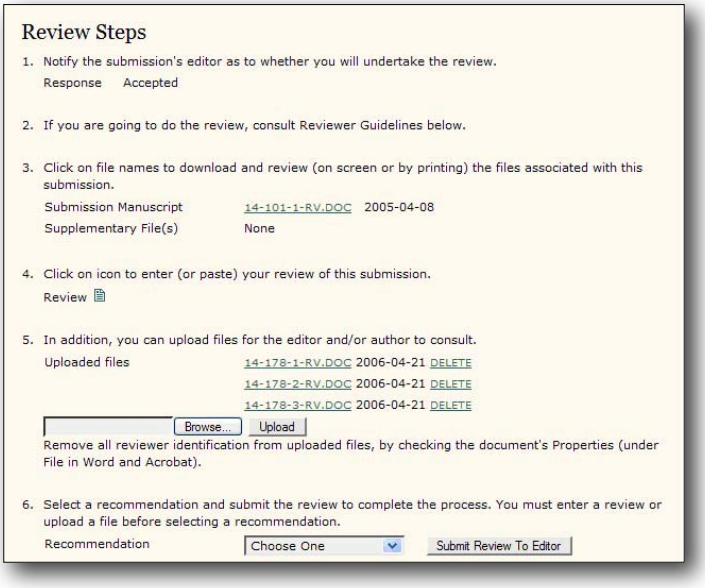


Figure: Review steps

In Step One of the standard OJS review process, the Reviewer accepts or declines the invitation to review, and informs the Section Editor of their intent. Step Two allows the Reviewer to access the submission file. This may only be available once they have agreed to do the review. Step Three

opens a dialogue box for recording the review comments. In Step Four, the Reviewer can upload an edited copy of the original submission. Finally, for Step Five, the Reviewer indicates whether the submission should be accepted, either as submitted or with revisions.

Reviewers may be rated by Section Editors, depending on the policies for this journal. Ratings scores are based on factors such as the quality of the review provided and the timeliness of their work. Ratings help the Section Editor to choose the best Reviewers.

Step 6: The Copyeditor

Once the Review Process has been completed and the submission accepted for publication, a Copyeditor is selected by the Section Editor. Some journals have an Editor or Section Editor play this role, or hire the services of a private copyeditor.

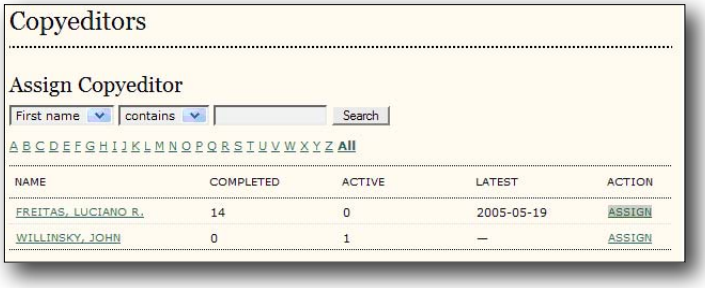


Figure: Assigning a copyeditor

The Copyeditor improves the grammar and clarity of the article, works with the author to ensure everything is in place, ensures strict adherence to the journal's bibliographic and textual style, and produces a clean, edited copy for the Layout Editor to turn into the files that will be in the published format of the journal (known as galley files).

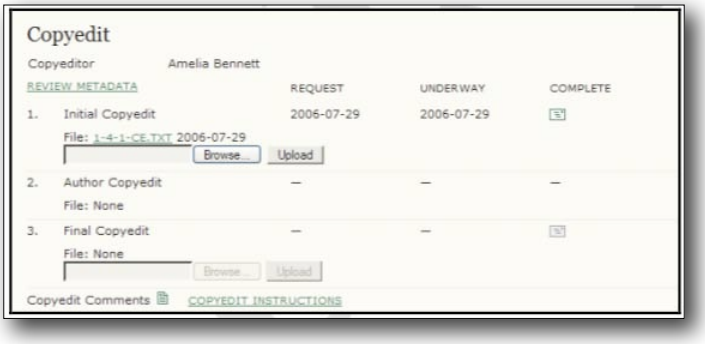


Figure: Copyediting Steps

Step 7: The Layout Editor

The Layout Editor transforms the copyedited versions of the article into galleys, usually in HTML or PDF format for online journals, or QuarkXPress® or Adobe InDesign® for print publications.

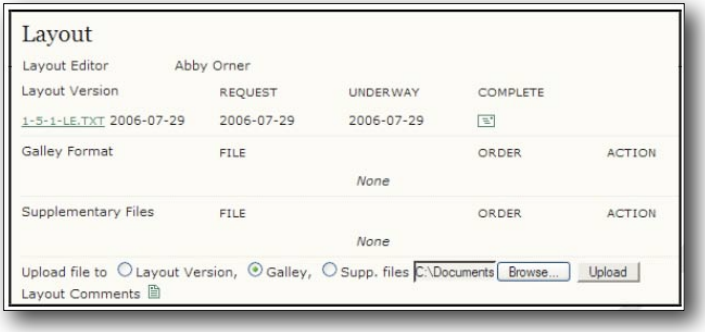


Figure: Layout Editing Steps

OJS does not provide software for converting word processed documents to galley formats, so the Layout Editor should have access to and be able to use third-party software packages for creating galleys (with Adobe Acrobat® or PDF Creator for PDFs, for example; or Dreamweaver® or Nvu® for HTML). These galleys present the articles with a well-formatted and readable layout, in the manner of scholarly journals, and with an eye to this new publishing medium (by consulting the layout design used by

other publishers of online journals, such as Highwire Press in the life sciences or Project Muse in the humanities).

As with the Copyeditor, some journals have an Editor or Section Editor fulfilling this role, or hire the services of a private layout editor.

Step 8: The Proofreader

With the initial galleys completed, the Proofreader carefully reads them over in the various formats in which the journal publishes (as does the Author).

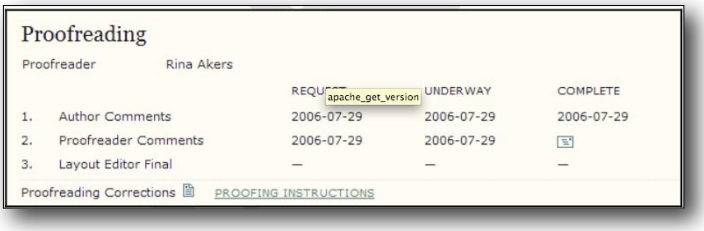


Figure: Proofreading

The Proofreader and Author record any typographic and formatting errors for the Layout Editor to fix. Again, in the case of some journals, the Editor or Section Editor will also serve as Proofreader.

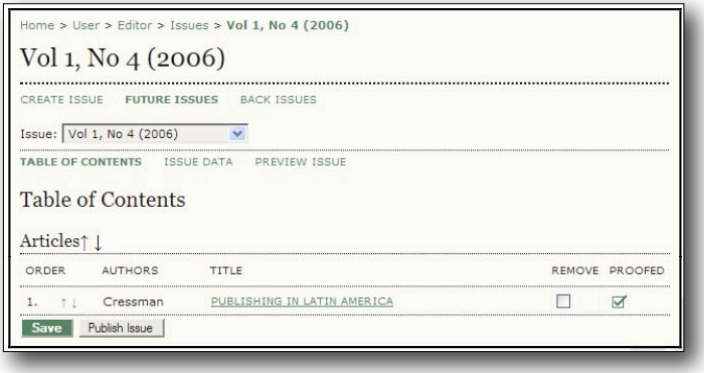


Figure: Publishing an Issue

When the Layout Editor has received all of the necessary changes for the article, the final galleys are produced. These galleys are uploaded into the system, and the Editor is informed that they are ready for publication. The Editor has the option of adding the article to the Table of Contents for the next issue of the journal, or postponing its publication in a future issue. When the Editor has enough content for a new issue, it can be created with the push of a button.

Step 9: The Reader

Readers include subscribers for paid journals and readers who choose to register for Open Access journals. Registered Readers received a notification with the publication of each issue that includes the Table of Contents from the journal. When registering with the journal, Authors are given the option of enrolling as Readers as well.

OJS provides a number of other services that enhance the overall reading experience, such as the ability to post comments on the article, to share the article with colleagues, to email the author, and the use of the OJS Reading Tools from a wide range of academic disciplines. The Tools are intended to assist both expert and novice readers of the journal in building a context for interpreting, evaluating and utilizing the research they are reading. The Tools enable Readers to look up words in the item (by double clicking on any word in the HTML version of the item). The Tools are also designed to take the item's keywords and feed them into the search engines of Open Access databases and other resources grouped into categories such as Research Studies, Author's Other Works, Press and Media, Government Websites, etc., depending on the set of Tools selected. Readers are also able to access background information on each of the selected resources. In each category, whether Studies, Media, or Instruction, the Tools provide multiple choices or databases to consult, while allowing the Reader to learn more about each database by providing a link to an About page for the resource.

19

Topaz

Based on:

- Websites:

http://www.topazproject.org
http://www.escholarlypub.com/digitalkoans/2006/11/15/under-the-hood-of-plos-one-the-open-source-topaz-e-publishing-system/
http://www.plosone.org/static/information.action

What is TOPAZ?

The core of TOPAZ is a digital information repository called Fedora (Flexible Extensible Digital Object Repository Architecture). Fedora is an Open Source content management application that supports the creation and management of digital objects. The digital objects contain metadata to express internal and external relationships in the repository, like articles in a journal or the text, images and video of an article. This relationship metadata can also be searched using a semantic web query languages. Fedora is jointly developed by Cornell University's computer science department and the University of Virginia Libraries. Topaz is a powerful object to RDF (Resource Description Framework) persistence and query service. Based loosely on the ORM family of software, Topaz lets you develop persistent classes following object-oriented concepts such as inheritance, composition, association, etc. Besides allowing use of the underlying RDF store's native query language, Topaz provides its own language (OQL), which applications can utilize to query based on defined objects. Topaz is written in Java and is available under the ECL till 0.8.3 and Apache open source license for subsequent releases. The following diagram provides a conceptual view on where Topaz can be used by an application:

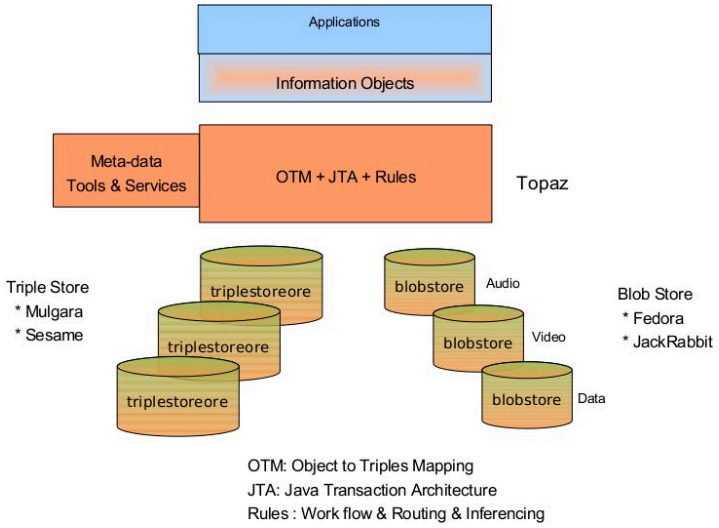


Figure: Topaz architecture

The first application that has been built using Topaz is the Ambra Publishing System and is being used by PLoS (www.plos.org) to host some of their journals (e.g. PLoS ONE). The PLoS ONE web interface was built with AJAX. Client-side APIs will create the community features (e.g. annotations, discussion threads, ratings, etc.) for the website.

Case study: PLoS ONE journal

PLoS ONE (www.plosone.org) is an international, peer-reviewed, open-access, online publication. PLoS ONE welcomes reports on primary research from any scientific discipline. It provides:

- open-access, freely accessible online, authors retain copyright;
- fast publication times;
- peer reviewed by expert, practicing researchers;
- post-publication tools to indicate quality and impact;
- community-based dialogue on articles;
- worldwide media coverage.

PLoS ONE is published by the Public Library of Science (PLoS), a nonprofit organization.

PLoS ONE is run as a partnership between its in-house PLoS staff and international Advisory and Editorial Boards, ensuring fast, fair, and professional peer review.

Scope

PLoS ONE features reports of original research from all disciplines within science and medicine. By not excluding papers on the basis of subject area, PLoS ONE facilitates the discovery of the connections between papers whether within or between disciplines.

Peer review

Each submission is assigned to a member of the PLoS ONE Editorial Board. They are responsible for managing the peer review for each submission - a process which concentrates on technical rather than

subjective concerns and may involve discussion with other members of the Editorial Board and/or the solicitation of the opinions of other experts in the field. If published, papers are made available for community evaluation and discourse involving the addition of online Notes, Comments, and Ratings.

Open Access

The Public Library of Science (PLoS) applies the Creative Commons Attribution License (CCAL) to all works we publish. Under the CCAL, authors retain ownership of the copyright for their article, but authors allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in PLoS journals, so long as the original authors and source are cited. No permission is required from the authors or the publishers.

Publication charges

To provide Open Access, PLoS journals use a business model in which our expenses –including those of peer review, journal production, and online hosting and archiving– are recovered in part by charging a publication fee to the authors or research sponsors for each article they publish. For PLoS ONE the publication fee is US$1250. Authors who are affiliated with one of our Institutional Members are eligible for a discount on this fee.

We offer a complete or partial fee waiver for authors who do not have funds to cover publication fees. Editors and reviewers have no access to payment information, and hence inability to pay will not influence the decision to publish a paper.

About the Public Library of Science

The Public Library of Science (PLoS, www.plos.org) is a non-profit organization of scientists and physicians committed to making the world's scientific and medical literature a freely available public resource.

20

CDS Invenio

Based on:

- Website: http://cdsware.cern.ch

CDS Invenio

CDS Invenio (formerly CDSware, http://cdsware.cern.ch), the integrated digital library system, is a suite of applications which provides the framework and tools for building and managing an autonomous digital library server. The software is readily available to anyone, as it is free software, licensed under the GNU General Public License (GPL, http://www.gnu.org/licenses/gpl.html). The technology offered by the software covers all aspects of digital library management. It complies with the Open Archives Initiative metadata harvesting protocol (OAI-PMH) and uses MARC 21 (http://www.loc.gov/marc/) as its underlying bibliographic standard. Its flexibility and performance make it a comprehensive solution for the management of document repositories of moderate to large size.

Figure: Extensible metadata representation

CDS Invenio is developed by, maintained by, and used at, the CERN Document Server. At CERN, CDS Invenio manages over 500 collections of data, consisting of over 800,000 bibliographic records, covering preprints, articles, books, journals, photographs, and more. Besides CERN, CDS

Invenio is currently installed and in use by over a dozen scientific institutions worldwide (including EPFL Infoscience - EPFL, Lausanne, Switzerland and PADIS - Università La Sapienza, Rome, Italy).

Contact CDS Invenio e-mail: cds.support@cern.ch

Key features

- Configurable portal-like interfaces for hosting various kinds of collections;
- Powerful search engine with Google-like syntax, including parallel searching of external collections;
- Extensible metadata representation (MARC XML) to handle virtually any kind of document (articles, books, photos, videos and more);
- Flexible document type submission and approbation workflow;
- User personalization, including document baskets and email notification alerts;

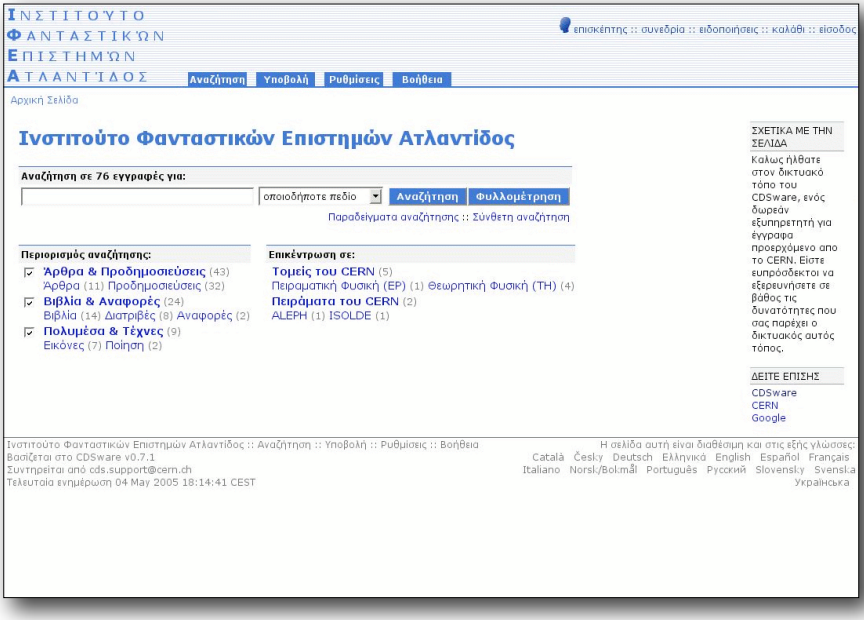


Figure: User personalization

- User collaboration, reviews, comments, knowledge sharing;
- Multilingual interface available in 20 languages, Unicode compliant (UTF-8);
- Compliant to Open Archive Initiative protocol for metadata harvesting;
- Free software (GNU GPL).