

## ARL Bimonthly Report 227

### April 2003

#### Metalist of Open Access E-Print Archives: The Genesis of Institutional Archives and Independent Services

by Steve Hitchcock, Intelligence, Agents, Multimedia Group, Southampton University

Editor's note: The updated version of the following metalist is available at <<http://opcit.eprints.org/explorearchives.shtml>>.

#### Overview

Open access e-print archives are where authors of published research papers and papers destined for peer reviewed publication can self-archive the full texts of their work for all to see. Researchers who self-archive want to improve access to papers while preserving the recognized quality control established by journals ([Harnad 2001](#)). The engine for growth of these archives is the recognition by researchers and policy-makers that the improved impact achieved through open access, demonstrated by [Lawrence \(2001\)](#), is not only desirable but entirely compatible with peer reviewed publication.

What is the scale of open access e-print archives, and of author self-archiving, currently? Despite the rhetoric there are no quantitative studies. The context for such studies is not just the growing scale of open access archives and the sheer number of archives, but the evolving structure of distributed archives and independent services. Web-based open access archives are not simply collections built for browsing but also are open data sources for powerful, automated independent services such as search, aggregation and impact measurement.

The enabling infrastructure for distributed archives and independent data services was introduced by the Open Archives Initiative (OAI) with its Protocol for Metadata Harvesting (PMH) in January 2001 ([Lynch 2001](#)). [Tomaiuolo and Packer \(2000\)](#) provided a checklist of disciplinary preprint archives that, because OAI was then in its infancy, recognized the likely influence of cross-archive services such as search but could not have detected the growth in institutional archives that OAI has subsequently motivated.

So while a new checklist is warranted, a mere list of open access e-print archives, and examination of their contents, is insufficient as a measure of the challenge. It is important to look at archive service providers too.

Thus, this is not a list of individual open access archives of full-text research papers, but instead lists and comments on other lists of individual archives. This list and its categorization gives a broad overview of the structure, size and progress of full-text open access e-print archives.

This list will be maintained and updated as far as is possible, and is intended to assist further quantitative research on the open access e-print phenomenon for those who want to measure the growth and quality of open access e-print archives.

For a chronological view of the development of open access institutional archives in the wider context of free online scholarship (FOS), including many of the services and archives listed here, see [Suber's Timeline of the FOS Movement](#).

The [Budapest Open Access Initiative](#) (BOAI), which supports both open access e-print archives and journals, has reinvigorated the cause and adoption of services providing open access to full-text research papers. While this list covers e-print archives, [Bosc et al.](#) offer an overview of new models of scientific communication (in French) that is more in line with the broader BOAI agenda.

#### Structure of the Metalist

1. [General Lists of Open Access E-Print \(Full-Text\) Archives](#)
2. [OAI Archives](#)
  - 2.1 [OAI Services-Based Lists of Archives](#)
3. [Lists of Institutional Archives](#)
  - 3.1 [Institutional Archives](#)
4. [Eprints.org Archives](#)
5. [Gateways \(Indexes, Unified Search and Browse of Covered Sites\)](#)
  - 5.1 [Centralizing Subject-Based Archive Gateways](#)
  - 5.2 [Decentralizing Archive Gateways](#)
  - 5.3 [The Economics Network \(RePEc\) Example](#)
6. [Open Access Journal Archives](#)

7. [Disciplinary Archives](#)
  - 7.1 [Mathematics](#)
  - 7.2 [Cognitive Science](#)
  - 7.3 [Library and Information Science \(LIS\)](#)
  - 7.4 [Publisher-Supported \(Author Self-Archiving\) Preprint Archives](#)
  - 7.5 [Other Disciplinary Archives](#)

Where the number of archives given in a source is stated, this is an approximate number intended to give an estimate of size. Since the numbers can change on a daily basis these are dated for reference, either by the last-modified date claimed by the resource when viewed, or the date viewed by the compiler of this list.

### 1. General Lists of Open Access E-Print (Full-Text) Archives

There are many different types of archives. One principal distinction is between subject-based, disciplinary archives and institutional archives. Both disciplinary and institutional archives can be preprint (pre-journal publication versions of papers) or e-print archives (which can include successive versions of papers pre- and post-publication, but are primarily distinguished by inclusion of post-publication versions). This difference is often ignored or incorrectly glossed over. Archives of interest in this study are characterized by containing full-text papers that have been self-archived (i.e., deposited) by their authors.

Open Directory Project, Free Access Online Archives (60 archives listed, last update 16 March 2003)

Electronic Archives "providing free and unrestricted access to peer reviewed scientific papers and academic publications."  
<[http://dmoz.org/Science/Publications/Archives/Free\\_Access\\_Online\\_Archives/](http://dmoz.org/Science/Publications/Archives/Free_Access_Online_Archives/)>

HighWire Press, Earth's Largest Free Full-Text Science Archives (20 archives), list produced to highlight HighWire's Free Online Full-text Articles ([see section 6 below on Open Access Journal Archives](#)) as the largest such archive.  
<<http://highwire.stanford.edu/lists/largest.dtl>>

University of Maryland Libraries, Virtual Technical Reports Center: E-prints, Preprints, & Technical Reports on the Web, "Institutions listed here provide either full-text reports, or searchable extended abstracts of their technical reports". Alphabetical by institution name (last updated March 05, 2003). <<http://www.lib.umd.edu/ENGIN/TechReports/Virtual-TechReports.html>>

University of Virginia Science and Engineering Libraries, Preprint Servers and Databases (33 archives, last modified January 13, 2003), pointers to a variety of electronic preprint sources in all areas of science and engineering.  
<<http://viva.lib.virginia.edu/science/guides/s-preprn.htm>>

Tardis (JISC FAIR project 2002- ), E-print and Related Archives with Subject and Institutional Categories Identified (113 archives, first posted January 2003). Institution, multi-institution, subject and multidisciplinary archives.  
<<http://tardis.eprints.org/discussion/eprintarchivessubjecttable9103.htm>>

Aardvark, Asian Resources for Libraries, Free preprint and full text science archives (115 archives, viewed 20 March 2003).  
<<http://www.aardvarknet.info/user/subject19/index.cfm?all=All>>

American Mathematical Society (AMS), Directory of Mathematics Preprint and e-Print Servers. <<http://www.ams.org/global-preprints/>>

- Umbrella servers, which cover all areas of mathematics, e.g.
  - Front for the Mathematics ArXiv, alternative arXiv interface <<http://front.math.ucdavis.edu/>>
  - MPRESS/MathNet.preprints  
([See section 5.1 below on Centralizing Subject-Based Archive Gateways.](#))
  - Preprints by country, from Austria-Bavaria, France, Germany
- Special subject (disciplinary) servers (17 archives, covering mathematics, physics)
- Mathematics departments and institutes (institutional servers, 56 archives, international)

Astronomy Preprints & Abstracts, hosted by National Radio Astronomy Observatory, Charlottesville, VA, linked list of sites, includes institutional preprint servers (56 archives, viewed 20 March 2003). <[http://www.cv.nrao.edu/fits/www/yp\\_preprint.html](http://www.cv.nrao.edu/fits/www/yp_preprint.html)>

### 2. OAI Archives

Until 1999 many institutionally based archives would have had a departmental bias and contained technical reports (TRs), the Guild Model identified by [Kling et al. \(2002\)](#). Since then, the Open Archives Initiative (OAI) has given momentum to a new type of institutional archive that contains e-prints of published (refereed) journal papers produced within research and educational institutions. OAI archives can be disciplinary or institutional, but the OAI's primary contribution has been to motivate new institutional archives. Not all OAI archives serve full-text papers, and it is definitely not a pre-condition of compliance with OAI that the items described by OAI metadata are openly or freely accessible. Study of OAI-compliant sites shows these include portals, software repositories, and test archives, and sites containing metadata about physical objects and collection-level metadata, but not always full texts.

Open Archives Initiative, registered data providers, "conforming repositories" (77 archives, viewed 27 March 2003). Sites found still to

be using OAI 1.1 on 2002/12/01 were purged from this list. <<http://www.openarchives.org/Register/BrowseSites.pl>>

Open Archives Forum, List of Repositories (20 archives, viewed 20 March 2003). No reasons for selection given (OAF is a focus for dissemination of information about European activity related to open archives and, in particular, to the OAI). <[http://www.oaforum.org/oaforum\\_db/list\\_db/list\\_repositories.php](http://www.oaforum.org/oaforum_db/list_db/list_repositories.php)>

### 2.1 OAI Services-Based Lists of Archives

Where TR archives were essentially separate archives that could be indexed (see for example the [Unified Computer Science Technical Report Index \(UCSTRI\) list of sites](#), one of the first TR indexes on the Web) but had to be accessed and searched separately for each institution or department, the OAI Protocol for Metadata Harvesting (OAI-PMH) enables independent services to provide common search and browse interfaces covering many archives. To give users an idea of scope and coverage, these automated services typically provide useful details of the indexed archives.

Celestial, Open Archives gateway that harvests and caches metadata from OAI-PMH repositories and makes these data available for other services to harvest, includes number of records in repository and metadata namespace. <<http://celestial.eprints.org/cgi-bin/status/>>

OAIster, serving 1,093,169 records from 144 institutions (updated 21 February 2003). <<http://oaister.umdl.umich.edu/o/oaister/viewcolls.html>>

Arc, an experimental cross-archive search service, used to investigate issues in harvesting OAI compliant repositories and making them accessible through a unified search interface, List of Existing Archives (140 archives, viewed 4 April 2003). <<http://arc.cs.odu.edu:8080/oa/admin.jsp>>

my.OAI, user customizable search engine covering selected metadata databases from the OAI, see forms-based list of databases in guest search interface (15 archives, viewed 4 April 2003). <<http://www.myoai.com/search/Search.cgi/LoginForm?Login=guest&Password=guest>>

Public Knowledge Project, Open Archives Harvester (12 archives, viewed 20 March 2003). Listed archives have to request harvesting). <<http://www.pkp.ubc.ca/harvester/archives.php>>

Open Archives Initiative--Repository Explorer, Virginia Tech interface to test archives interactively for compliance with the OAI-PMH, see forms-based predefined archive list in Repository Explorer interface (60 archives, viewed 4 April 2003). <<http://oai.dlib.vt.edu/cgi-bin/Explorer/oa2.0/testoai/>>

### 3. Lists of Institutional Archives

Some lists focus on institutional archives as the most likely area for growth of open access, OAI-based e-print archives ([Crow 2002](#), [Young 2002](#)).

SPARC, Select list of Institutional Repositories, by country, lists type of content (mostly preprints, published papers), software used (13 of 26 repositories listed use EPrints.org, last updated February 13, 2002), URL of repositories. <<http://www.arl.org/sparc/core/index.asp?page=m1>>

Signal Hill, a European partnership for academic publishing set up by the University Libraries of Utrecht and Delft and Firenze University Press, institutional archives by country (34 archives, viewed 20 March 2003). <<http://www.signal-hill.org/archives/institutions.html>>

#### 3.1 Institutional Archives

It is not the intent in this paper to list individual institutional archives extensively, although a few are chosen to highlight different implementation models, described by [Tennant \(2002\)](#), adopted within institutions to motivate the uptake of archive services across the range of cultures and disciplines found within academic institutions. Institutional archives need not be exclusively e-print archives. [Lynch \(2003\)](#) delineated the "all outputs" archiving approach and the research papers output approach, although it can be anticipated that e-prints, which as journal publications are intended for wide dissemination, will form the bulk of institutional archives, at least initially. The perception is that institutional e-print archives, backed by institutional policies on deposit and publication ([Crow 2003](#)), will be able to build higher levels of content, faster, than has been achieved by disciplinary archives, with the exception of arXiv. Institutions might use archives as showcases for their research output, but in building these archives can minimize complexity and cost by recognizing one significant fact of user behavior: institutional archives will rarely be browsed or searched directly. Underpinned by OAI, these functions will devolve to services that can provide disciplinary or some other research focus for users.

University of California, California Digital Library eScholarship Repository, offers faculty a central location for depositing any research or scholarly output deemed appropriate by their participating research unit, center, or department, including working papers and prepublication scholarship. <<http://repositories.cdlib.org/escholarship/>>

Caltech, Collection of Open Digital Archives (CODA), includes more than 10 repositories in production or in development. <<http://library.caltech.edu/digital/>>

US Department of Energy (DOE), the Information Bridge, provides the open source to full-text and bibliographic records of DOE research and development reports in physics, chemistry, materials, biology, environmental sciences, energy technologies, engineering, computer and information science, renewable energy, and other topics. Contains full-text documents produced and made available by the DOE National Laboratories and grantees from 1995 forward. Legacy documents are included as they become available. (See also DOE PrePRINT Network, included in [section 5.1 on Centralizing Subject-Based Archive Gateways](#).) <<http://www.osti.gov/bridge/>>

#### 4. Eprints.org Archives

Institutional archives can be distinguished by the type of software used to build the archives, providing core functions and interfaces for deposit and data management while reducing cost and complexity. As can be deduced from the lists of institutional archives, the software most widely used for this is produced by Eprints.org. Many Eprints.org-based archives are institutional, but not exclusively so. The Cogprints disciplinary archive was built with software that evolved to become Eprints.org. Other types of archive software are becoming available, and no doubt there will soon be lists of archives supported by these packages. Whichever software is chosen, these packages invariably produce archives that are OAI-compliant, so this list will overlap with the OAI list above.

GNU EPrints, software for the development of institutional e-print archives, but can also be used to build other types of archives with other types of content. All the repositories known to have been built using the first two version releases of this software are in these two lists (viewed 20 March 2003):

- E-prints 2 Archives <<http://software.eprints.org/#ep2/>> (37 archives)
- E-prints 1 Archives <<http://software.eprints.org/#ep1/>> (29 archives)

#### 5. Gateways (Indexes, Unified Search and Browse of Covered Sites)

##### 5.1 Centralizing Subject-Based Archive Gateways

OAI services were not the first to introduce unified search and browse interfaces for archives. Various gateway services preceded these. While not e-print archives in their own right, these services are important for the way in which they have enabled the structure of different archives to evolve. Some gateways are based on the largest archives, in this case the physics, mathematics, and computer science archives at arXiv. For example, a number of previously independent mathematics archives merged with arXiv without loss of functionality or focus due to interfaces such as the Front for the Mathematics ArXiv. Other services combine searches on high-energy physics and astronomy in arXiv with bibliographic sources.

##### *ArXiv Search Interfaces*

- Front for the Mathematics ArXiv, alternative arXiv interface <<http://front.math.ucdavis.edu/>>
- NASA, Astrophysics Data System (ADS) ArXiv Preprints Query Form <[http://adsabs.harvard.edu/preprint\\_service.html](http://adsabs.harvard.edu/preprint_service.html)>
- Die Pro-Physik Findemaschine, specialized German search engine, includes arXiv among searchable resources, uses flexible taxonomies to support thematic searching across disciplines <<http://findemaschine.pro-physik.de/?language=e>>

NASA ADS Harvard-Smithsonian Center for Astrophysics Preprints (CfA) Preprints Query Form.  
<<http://adsabs.harvard.edu/cfa/preprints.html>>

The Stanford Linear Accelerator Center (SLAC), SPIRES HEP literature database contains more than 500,000 high-energy-physics-related articles including journal papers, preprints, e-prints, technical reports, conference papers and theses, indexed by the SLAC and Deutsches Elektronen Synchrotron (DESY) libraries since 1974. <<http://www.slac.stanford.edu/spires/hep/>>

Citebase, citation-ranked search and impact discovery for arXiv (also covers CogPrints and BioMed Central).  
<<http://citebase.eprints.org/help/coverage.php/>>

Elsevier, Scirus, "the most comprehensive science-specific search engine on the Internet", covers over 135 million science-related pages, consisting of 120 million Web pages from paid-for sources as well as prominent e-print archives.  
<<http://www.scirus.com/about/#content/>>

CERN Document Server (CDS), searchable Web interface to over 550,000 bibliographic records, including 220,000 full-text documents in particle physics and related areas, covers preprints, articles, books, journals, photographs.... <<http://weblib.cern.ch/>>

Results include reference links (including journal links to publisher site, abstract, summary only, not OpenURL) and cited by, but cannot search or rank by citations

CDS services include:

- HEPDOC Search, alternative CERN service for quick searching from HEP preprint servers:  
<<http://weblib.cern.ch/share/hepdoc/>> (With Mozilla browser does not work fully, cannot search servers singly.) Covers these servers:
  - CERN: Search for Articles & Preprints <[http://weblib.cern.ch/Home/Library\\_Catalogue/Articles\\_and\\_Preprints/](http://weblib.cern.ch/Home/Library_Catalogue/Articles_and_Preprints/)>
  - DESY (=SLAC)
  - KEK
  - SLAC

- Document impact tool, frequency of full-text document download on CDS, how many times files (full-text documents) stored on the CDS are downloaded within a given period of time. <<http://doc.cern.ch/cgi-bin/impact/search.pl>>

PhysDoc--Physics Documents Worldwide--offers lists of links to document sources, such as preprints, research reports, annual reports, and list of publications of worldwide distributed physics institutions and individual physicists, ordered by continent, country and town. <<http://de.physnet.net/PhysNet/physdoc.html>>

MPRESS, the Mathematics Preprint Search System, a searchable index of preprints from 10 servers, mostly covering geographical servers, but also disciplinary mathematics servers including Topology Atlas, Algebraic Number Theory Archives and K-theory Preprint Archives, as well as the mathematics part of the arXiv mirror at Augsburg. <<http://mathnet.preprints.org/>>

US Department of Energy (DOE), PrePRINT Network, searchable gateway to preprint servers that deal with scientific and technical disciplines of concern to DOE: physics, materials, and chemistry, as well as portions of biology, environmental sciences and nuclear medicine. (See also DOE Information Bridge in [section 3.1 above on Institutional Archives](#).) <<http://www.osti.gov/preprints/ppnbrowse.html>>

NTRS, NASA Technical Reports Server, search interface for 18 databases. <<http://ntrs.nasa.gov/>>

## 5.2 Decentralizing Archive Gateways

Gateways have not exerted solely a centralizing influence on deposit processes, and in two notable examples, RePEc (Research Papers in Economics) and NCSTRL (Networked Computer Science Technical Reference Library), can be found forerunners of the distributed OAI model: independent archives, indexes, and databases. The growth and appeal of NCSTRL appears to have been limited by the large administrative, maintenance, and metadata overhead imposed on participating institutional archives, a lesson learnt by the OAI designers who wanted a simpler, more widely accepted standard metadata format describing the contents of archives. NCSTRL is being converted into an OAI-compliant index.

Networked Computer Science Technical Reference Library (NCSTRL) is being developed into a sustainable OAI conformant framework in a collaborative project involving NASA Langley, Old Dominion University, University of Virginia, and Virginia Tech. <<http://www.ncstrl.org/>>

Browse list of participating archives: <<http://www.ncstrl.org:8900/ncstrl/body.html>>

Networked Digital Library of Theses and Dissertations (NDLTD), theses rather than e-prints, but included here as an example of an archive aiming to present open access to full-text research outputs. <<http://www.ndltd.org/>>

Open Language Archives Community (OLAC), creating a worldwide virtual library of language resources, 21 participating archives, three service providers including OLAC Aggregator, Swahili Language Resources, and a virtual service provider. Open Language Archives are repositories of language data, documentation and description, including texts, recordings, dictionaries, grammars and field notes, where there is an intent to make the materials openly available, includes any such repository which has an accessible digital component, even if it is just an online catalog or a few digital holdings (use of "open" is inspired by OAI). Less an e-print archive, more a preservation and rescue service for language resources. <<http://www.language-archives.org/>>

## 5.3 The Economics Network (RePEc) Example

RePEc (Research Papers in Economics) is a large database of working papers, journal articles, and software components, an "Open Library," open to contributions and providing open data for user services ([Krichel 2000](#)). Interpretations vary on the proportion of material available as full texts from the constituent archives of "working papers," but RePEc is claimed to be the "second-largest source of freely downloadable scientific preprints" after arXiv. RePEc records over 177,000 items, over 86,000 of which are available online (27 Feb 2003). <<http://repec.org/>>

The following services provide access to all or part of the RePEc database for browse or search:

- IDEAS (Internet Documents in Economics Access Service): search the complete RePEc database <<http://ideas.repec.org/>>
- EconPapers: provides access to RePEc <<http://econpapers.hhs.se/>>
- NEP (New Economics Papers): e-mail announcement service for new downloadable working papers for over 40 specific fields <<http://netec.wustl.edu/NEP/>>
- Citations in Economics, CitEc: citations from items in the RePEc database <<http://netec.ier.hit-u.ac.jp/CitEc/>>
- Inomics: downloadable working papers integrated with information on economics conferences, job information, and Internet search for economists <<http://www.inomics.com/cgi/show/>>
- LogEc: detailed access statistics for participating RePEc services <<http://logec.hhs.se/>>

### RePEc Archives

Current archive providers to RePEc. <<http://ideas.repec.org/archives.html>>

Participating institutions provide over 1,000 RePEc series (many of the top series are journal series or smaller databases). LogEc list of the top 25 RePEc series of the past month. <<http://logec.hhs.se/scripts/seriesstat.pl>>

### Working Papers in Economics

WoPEc, all papers in WoPEc are downloadable but not necessarily free (contains over 80,000 documents in electronic format: 53,035 Working Papers, 41,895 Journal Articles, last updated 23 March 2003). <<http://netec.mcc.ac.uk/WoPEc.html>>

- WoPEc Working Papers Series (list of indexed archives) <<http://netec.mcc.ac.uk/WoPEc/data/PaperSeries.html>>

Among the largest contributing RePEc archives are the following working paper archives:

- DEGREE: full text electronic working papers on economics and business economics from nine universities in the Netherlands <<http://drcwww.uvt.nl/dbi/instructie/degree/>>
- International Monetary Fund (IMF) Working Papers by Date <[http://www.imf.org/external/pubs/cat/wp1\\_sp.cfm?s\\_year=2003&e\\_year=2003&brtype=default](http://www.imf.org/external/pubs/cat/wp1_sp.cfm?s_year=2003&e_year=2003&brtype=default)>
- EconWPA (USA), Economics Working Paper Archive <<http://econwpa.wustl.edu/>>
- Cowles Foundation, Yale University, Discussion Papers <<http://cowles.econ.yale.edu/P/cd/cfdpmain.htm>>
- Center for Mathematical Studies in Economics and Management Science, Northwestern University (USA), Discussion Papers <[http://www.kellogg.nwu.edu/research/math/discussion\\_papers.htm](http://www.kellogg.nwu.edu/research/math/discussion_papers.htm)>
- Caltech Division of Social Sciences (USA), Working Papers On-Line <<http://www.hss.caltech.edu/SSWPLinks.html>>
- UCLA Department of Economics (USA), Working Papers <<http://econweb.sscnet.ucla.edu/research/workingp.html>>
- S-WoPEc, Scandinavian Working Papers in Economics <<http://swopec.hhs.se/>>

*RePEc-Modeled Archives, Not Economics*

Documents in Information Science (DoIS) is a database of articles and conference proceedings published in electronic format in the area of Library and Information Science, holds about 10,042 articles and 3,045 conference proceedings, 6,928 of them are downloadable (28th February 2003). <<http://dois.mimas.ac.uk/>>

A more broadly based database, rclis (Research in Computing, Library and Information Science) is in development. <<http://rclis.org/about.html>>

## 6. Open Access Journal Archives

A notable development in the wider context of full-text e-print archives is the growth of open access journal archives. Papers in these archives are not deposited by authors but by journal publishers. Mostly this is focused on biomedical journals, and was initiated by PubMed Central, the U.S. National Library of Medicine's site, which has grown significantly, and makes copies of subscription-based journals available some time after publication. HighWire Press, a large producer of biomedical e-journals, similarly makes delayed copies of journal papers available free. Unlike PubMed Central and HighWire, the publisher BioMed Central has pioneered a new business model of original open access journals funded through author and institutional payments for review and publication. For some in this field the progress represented by these examples is not enough, as they will be joined by new open access journals from the [Public Library of Science](#) (PLoS). The model adopted by PubMed Central and PLoS has been endorsed by the [Budapest Open Access Initiative](#) (BOAI). There are other distinctive and successful journal-archive models, such as *Advances in Theoretical and Mathematical Physics*, a journal "overlay" of some arXiv physics archives that has published high-impact papers.

BioMed Central (120 journals at 20 Feb. 2003). <<http://www.biomedcentral.com/start.asp/>>

PubMed Central (PMC) is the U.S. National Library of Medicine's digital archive of life sciences journal literature (52 participating journals at 20 Feb. 2003). <<http://pubmedcentral.nih.gov/>>

HighWire Press Free Online Full-text Articles (list limited to journals published online with the assistance of HighWire Press). At 28 Feb. 2003, 472,871 full-text articles were available free from 1,358,713 total articles. <<http://highwire.stanford.edu/lists/freart.dtl>>

Free Online Full-text Articles is the top entry in Earth's Largest Free Full-Text Science Archives, a list produced by HighWire Press. (See [section 1 above on General Lists of Open Access E-print \(Full-Text\) Archives.](#))

*Advances in Theoretical and Mathematical Physics* is an overlay of the arXiv archives. All papers are archived at LANL and its mirror sites. ATMP maintains only links to the above archive, thus realizing one of the first e-journals as an overlay to the global e-print archives. <<http://www.intlpress.com/journals/ATMP/>>

BBS Prints Interactive Archive of the journal *Behavioral and Brain Sciences* containing original refereed 'target' papers, open peer commentary and responses (OAI compliant, Eprints.org journal archive). <<http://www.bbsonline.org/>>

Psycoloquy, articles and peer commentary in all areas of psychology as well as cognitive science, neuroscience, behavioral biology, artificial intelligence, robotics/vision, linguistics and philosophy (Eprints.org archive). <<http://psycprints.ecs.soton.ac.uk/>>

Open access journals per se, without an archive connection, are not included here.

## 7. Disciplinary archives

Although the primary intent of this paper is not to list individual archives, disciplinary archives are significant enough to be included in their own right. These archives demonstrate a range of types, from the ubiquitous arXiv, to publisher-sponsored preprint collections, as well as smaller, specialized archives. [Guédon \(2001\)](#) describes a context in which publisher preprint archives based on author self-

archiving independently of submission to a specific journal, although formative, may be more significant than the mere size of such archives currently suggests. The large Citeseer autonomously indexed collection of computer science papers, mostly cached from authors' personal Web pages, shows how many e-prints are available outside managed archives, reflecting personal practices that are likely to be seen as characterizing the early history of author self-archiving on the Web but which shows no sign of diminishing yet. Some publisher copyright agreements seek to exploit the distinction between "personal" Web sites and managed archives. As this list shows, the distinction is often untenable, and is wholly untenable for institutional self-archiving (e.g., Eprints.org) where authors administer a personal space within a managed framework.

arXiv (1991- ), main administration site at Cornell University, multiple mirrors worldwide, manages access to over 230,000 papers, abstracts include links to citation analysis for the paper by SLAC Spires and Citebase. <<http://www.arXiv.org/>>

Citeseer (1998- , a.k.a. ResearchIndex), developed at NEC Research Institute, NJ, USA, caches openly accessible full-text research papers on computer science found on the Web in Postscript and PDF formats for autonomous citation indexing, it is claimed to index over 500,000 papers. Not yet OAI compliant, but planned to become so. <<http://citeseer.nj.nec.com/cs/>>

ebizSearch (2001- ), administered by the eBusiness Research Center at Pennsylvania State University, based on Citeseer software, autonomously creates citation indexes of e-commerce literature. The search engine crawls Web sites of universities, commercial organizations, research institutes and government departments to retrieve academic articles, working papers, white papers, consulting reports, magazine articles, and published statistics and facts. Not all documents are stored by eBizSearch, which performs a citation analysis of all articles accessed. <<http://gunther.smeal.psu.edu/>>

## 7.1 Mathematics

- K-theory Preprint Archives\* (papers from 1995), managed at Mathematics Department, University of Illinois at Urbana-Champaign. <<http://www.math.uiuc.edu/K-theory/>>
- Topology Atlas\* preprint server (papers from 1995), was most active in 1996 and 1997 and still accepts submissions but suggests using the Mathematics Archive (arXiv.org or its Front) for distributing and finding preprints, hosted at York University, North York, Ontario. <<http://at.yorku.ca/topology/preprint.htm>>
- Algebraic Number Theory Archives\* (papers from 1996, frozen since Jan. 2003) hosted at the Mathematics Department, University of Illinois at Urbana-Champaign. <<http://www.math.uiuc.edu/Algebraic-Number-Theory/>>
- Mathematical Physics Preprint Archive, mp\_arc (papers from 1991), hosted by Mathematics Department, University of Texas at Austin. <[http://rene.ma.utexas.edu/mp\\_arc/](http://rene.ma.utexas.edu/mp_arc/)>
- Hopf Topology Archive (papers from 1997), hosted by the Department of Mathematics, Purdue University. <<http://hopf.math.purdue.edu/>>
- Preprints on Conservation Laws (papers from 1996), administered at Department of Mathematical Sciences, Norwegian University of Science and Technology (NTNU), Trondheim. <<http://www.math.ntnu.no/conservation/>>
- MCMC (Markov Chain Monte Carlo) methodology Preprint Service (papers from 1993), administered at the Statistical Laboratory, University of Cambridge. <<http://www.statslab.cam.ac.uk/~mcmc/>>
- Jordan Theory Preprint Archives (papers from 1996), hosted at Institut für Mathematik, Universität Innsbruck. <<http://mathematik.uibk.ac.at/mathematik/jordan/>>
- Groups, Representations and Cohomology Preprint Archive (papers from 1995), managed at Department of Mathematics, University of Georgia, USA. <<http://www.math.uga.edu/archive.html>>
- Field Arithmetic Archive, located at Ben Gurion University in Be'er-Sheva, Israel, stores electronic preprints on the arithmetic of fields, Galois theory, model theory of fields, and related topics. <<http://www.cs.bgu.ac.il/research/Fields/>>
- MGNet preprints (papers from 1991, last paper deposited 2001), Department of Computer Science, Yale University, repository for information related to multigrid, multilevel, multiscale, aggregation, defect correction, and domain decomposition methods. <<http://casper.cs.yale.edu/mgnet/www/mgnet-papers.html>>

\* Searchable via MPRESS. (See section 5.1 on Centralizing Subject-Based Archive Gateways.)

The [International Mathematical Union adopted a resolution \(May 2001\)](#) encouraging mathematicians to make their work available online: "Open access to the mathematical literature is an important goal....Our action will have greatly enlarged the reservoir of freely available primary mathematical material, particularly helping scientists working without adequate library access."

## 7.2 Cognitive Science

- Cogprints (1997- ), an electronic archive for self-archived papers in any area of Psychology, Neuroscience, and Linguistics, and many areas of Computer Science, Philosophy, Biology, Medicine, Anthropology, as well as any other areas pertinent to the study of cognition, initially a project in the JISC Electronic Libraries (eLib) Programme, administered by the IAM Group, University of Southampton. <<http://cogprints.ecs.soton.ac.uk/>>

## 7.3 Library and Information Science (LIS)

- E-LIS, E-Prints in Library and Information Science. <<http://eprints.rclis.org/>>
- DList, Digital Library of Information Science and Technology (October 2002- ), managed by School of Information Resources and Library Science and Arizona Health Sciences Library, University of Arizona. <<http://dlist.sir.arizona.edu/>>

#### 7.4 Publisher-Supported (Author Self-Archiving) Preprint Archives

- NetPrints Clinical Medicine and Health Research (December 1999- ), BMJ Publishing Group and HighWire Press, a repository of non-peer-reviewed original research. <<http://clinmed.netprints.org/home.dtl>>
- Chemistry Preprint Server (August 2000- ), ChemWeb.com, Elsevier, "CPS is registered as a compliant data provider for the OAI protocol" (but not listed as v2.0 compliant) (640 papers at 3rd March 2003). <<http://www.chemweb.com/preprint/>>
- Computer Science Preprint Server (November 2001- ), Elsevier (409 papers at 24th March 2003). <<http://www.compscipreprints.com/comp/Preprint/show/>>
- Mathematics Preprint Server (May 2001- ), Elsevier, a registered data provider, supporting OAIv2.0 (842 papers at 25th March 2003). <<http://www.mathpreprints.com/math/Preprint/show/>>

Elsevier appears a little shy of associating itself with the latter two preprint servers. The connection is not indicated on the home pages of the Computer Science and Mathematics servers, but is made clear on the 'About' pages within the respective services (although even that has not always been the case, as attested to by e-mail correspondence between the author and David Solomon of Michigan State University <<http://www.topica.com/lists/fos-forum/read/message.html?mid=904133189>>). The servers are not linked from the Elsevier Science home page <<http://www.elsevier.com/>>, nor can they be found easily if at all by browsing from this page, and a search returns no results for 'preprint servers' (tried 27 March 2003). All services are searchable from Scirus ([see section 5.1 above on Centralizing Subject-Based Archive Gateways](#)), and the Mathematics preprint server is linked from Elsevier Science's Mathematics Web portal <<http://www.elseviermathematics.com/mathematicsweb/show/>>.

Many journals operate a preprint archive, making electronic copies of papers available prior to print publication. These are typically not based on author self-archiving nor are they open access, and so are not covered here.

#### 7.5 Other Disciplinary Archives

- HTP Prints, the History & Theory of Psychology E-print Archive (September 2001- ), administered at York University, Toronto. <<http://htpprints.yorku.ca/>>
- Education-line (1997- ), a freely accessible database of the full text of conference papers, working papers and electronic literature which supports educational research, policy and practice, initially a project in the JISC Electronic Libraries (eLib) Programme, administered by the Brotherton Library, University of Leeds. <<http://www.leeds.ac.uk/educol/>>
- Social Science Research Network (SSRN), Social Science Electronic Publishing, Inc., working papers and abstracts are provided by journals, publishers, and institutions for distribution through SSRN's eLibrary, which consists of two parts: a database containing abstracts on over 49,200 scholarly working papers and forthcoming papers, and an Electronic Paper Collection containing over 30,800 (27 March 2003) downloadable full-text documents. SSRN is composed of specialized research networks/journals in the social sciences: Accounting, Economics, Financial Economics, Legal Scholarship, Management, Negotiations.  
From an e-print perspective this is a curious amalgam, not a pure e-print archive at all, more a subscription-based service. The business model and purpose are not clear. Are downloadable papers freely downloadable? Clearly some are, but what proportion, if not all, is not clear. Networks can be browsed separately, but not searched separately, it appears. It does not seem to be possible to search only for freely downloadable papers. <<http://www.ssrn.com/>>
- ArchiveSIC (open archive on Sciences de l'Information et de la Communication), full-text papers on information and communication science (bilingual site in French/English). <<http://archivesic.ccsd.cnrs.fr/>>
- Electronic Colloquium on Computational Complexity (papers from 1994), led by the chair of theoretical computer science and new applications at the University of Trier. Research reports, surveys and books in computational complexity. <<http://www.eccc.uni-trier.de/eccc/>>
- Cryptology ePrint Archive (2000- ), maintained by the International Association for Cryptologic Research (IACR), incorporates contents of the Theory of Cryptology Library 1996-1999. <<http://eprint.iacr.org/>>
- The Digital Library of the Commons (DLC), Indiana University, contains a Working Paper Archive of author-submitted papers, as well as full-text conference papers, dissertations, working papers and preprints. (The commons is a general term for shared resources in which each stakeholder has an equal interest. Studies on the commons include the information commons with issues about public knowledge, the public domain, open science, and the free exchange of ideas.) <<http://dlc.dlib.indiana.edu/>>
- Organic E-prints (September 2002- ), established by the Danish Research Centre for Organic Farming (DARCOF), open access archive for papers related to research in organic agriculture. <<http://orgprints.org/>>
- University of California International and Area Studies (UCIAS) Digital Collection (October 2002- ), partnership of the University of California Press, the California Digital Library (CDL), and internationally oriented research units on eight UC campuses, publishes articles, monographs, and edited volumes that are peer-reviewed according to standards set by an interdisciplinary UCIAS Editorial Board and approved by the University of California Press. <<http://repositories.cdlib.org/uciaspubs/>>
- Formations, Faculty of Arts, University of Ulster, hosts e-prints in Media Studies and participative 'eLearning Forums' based on short discussion papers. Initially a project in the JISC Electronic Libraries (eLib) Programme. <<http://formations2.ulst.ac.uk/>>
- Ecology Preprint Registry (papers from July 2001), hosted at the National Center for Ecological Analysis and Synthesis, dissemination of new research results destined for publication (i.e., not white papers or gray literature), only preprints with a theoretical basis can be submitted, the scope may be expanded to include submissions from the entire discipline of ecology. <<http://www.nceas.ucsb.edu:8504/esa/ppr/Query>>
- PhilSci Archive (January 2001- ), hosted at the Department of Philosophy and the Department of History and Philosophy of Science, University of Pittsburgh, preprints in the philosophy of science. <<http://philsci-archive.pitt.edu/>>

This metalist was compiled as follow-up work to the Open Citation project, which was funded by the Joint NSF-JISC International Digital Libraries Research Programme.

## References

Bosc, Hélène, Simone Jérôme, and Jean-Philippe Schmitt. "La communication scientifique revue et corrigée par Internet." (2003). <<http://www.tours.inra.fr/tours/doc/comsci.htm>>.

Budapest Open Access Initiative. <<http://www.soros.org/openaccess/>>.

Crow, Raym. "The Case for Institutional Repositories: A SPARC Position Paper." Washington, D.C.: SPARC, 2002. <<http://www.arl.org/sparc/IR/ir.html>>.

----- . "SPARC Institutional Repository Checklist & Resource Guide." Washington, D.C.: SPARC, 2002. <[http://www.arl.org/sparc/IR/IR\\_Guide.html](http://www.arl.org/sparc/IR/IR_Guide.html)>.

Guédon, Jean-Claude. *In Oldenburg's Long Shadow: Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing*. Washington, D.C.: ARL, 2001. <<http://www.arl.org/arl/proceedings/138/guedon.html>>.

Harnad, Stevan. "The Self-Archiving Initiative." *Nature* 410 (2001): 1024-1025. <<http://www.nature.com/nature/debates/e-access/Articles/harnad.html>>.

International Mathematical Union. "Call to All Mathematicians." May 15, 2001. <[http://www.mathunion.org/IMU\\_Committees/call\\_authors.html](http://www.mathunion.org/IMU_Committees/call_authors.html)>.

Kling, Rob, Lisa Spector, and Geoff McKim. "Locally Controlled Scholarly Publishing via the Internet: The Guild Model." Center for Social Informatics Working Paper No. WP-02-01. Bloomington: Indiana University, School of Library and Information Science, 2002. <<http://www.slis.indiana.edu/csi/WP/WP02-01B.html>>.

Krichel, Thomas. "RePEc, an Open Library for Economics." March 2000. <<http://openlib.org/home/krichel/papers/salisbury.html>>.

Lawrence, Steve. "Free Online Availability Substantially Increases a Paper's Impact." *Nature Web Debates* (May 31, 2001). <<http://www.nature.com/nature/debates/e-access/Articles/lawrence.html>>.

Lynch, Clifford A. "Metadata Harvesting and the Open Archives Initiative." *ARL Bimonthly Report*, no. 217 (August 2001): 1-9. <<http://www.arl.org/newsltr/217/mhp.html>>.

----- . "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age." *ARL Bimonthly Report*, no. 226 (February 2003): 1-7. <<http://www.arl.org/newsltr/226/ir.html>>.

Public Library of Science, Journals. <<http://www.publiclibraryofscience.org/journals/>>.

Suber, Peter. "Timeline of the Free Online Scholarship Movement." 2002. <<http://www.earlham.edu/~peters/fos/timeline.htm>>.

Tennant, Roy. "Institutional Repositories." *Library Journal* (September 15, 2002). <<http://libraryjournal.reviewsnews.com/index.asp?layout=article&articleid=CA242297&display=Digital+LibrariesNews&industry=Digital+Libraries&industryid=3760&verticalid=151>>.

Tomaiuolo, Nicholas G. and Joan G. Packer. "Preprint Servers: Pushing the Envelope of Electronic Scholarly Publishing." *Searcher* 8, no. 9 (October 2000): 53-. <<http://www.infotoday.com/searcher/oct00/tomaiuolo&packer.htm>>.

Unified Computer Science Technical Report Index (UCSTRI). <<http://www.cs.indiana.edu/ucstri/sitelist.html>>.

Young, Jeffrey R. "'Superarchives' Could Hold All Scholarly Output: Online Collections by Institutions May Challenge the Role of Journal Publishers." *Chronicle of Higher Education*, July 5, 2002. <<http://chronicle.com/free/v48/i43/43a02901.htm>>.

--Copyright © 2003 Steve Hitchcock

[Back to top](#)

---

[Table of Contents for Issue 227](#) | [Other Scholarly Communication Articles](#)



[ARL Bimonthly Report Home](#) [ARL Home](#)

ARL policy is to grant blanket permission to reprint any article in the *Bimonthly Report* for educational use as long as full attribution is made. Exceptions to this policy may be noted for certain articles. This is in addition to the rights provided under sections 107 and 108 of the Copyright Act. For commercial use, a reprint request should be sent to ARL Director of Information Services, [Julia Blixrud](#).

© The Association of Research Libraries

Maintained by: [ARL Web Administrator](#) Site Design Consultant: [Chris Webster](#) Last Modified: June 18, 2003