

Free Open Source OAI-PMH 2.0 compliant Software [as at 17 November 2004]

Product	Platform	Service Provider	Data Provider	Self-Archiving	Content Management	Simple search	Advanced search	Display	Browse	Other Features
Arc [version: alpha] Old Dominion University	Java (JDK1.4) Tomcat 4.0x MySQL Lucene search engine http://Sourceforge.net/projects/oaiarc/	Yes Examples: http://arc.cs.odu.edu http://americansouth.org/	No See also Kepler - OAI self-archiving platform: http://kepler.cs.odu.edu/	No	No	Keyword Phrase Righthand truncation Cluster results Sort by relevance ranking, accession date. Refine search.	Keyword in specific fields "And" "Or" Limiting options Group records Sort records Search within search results.	Brief display Links to - author browse index, "Similar subject", reference linking (dependent on repository). Full record display	Individual repositories – alphabetical by title.	Annotations.
Archimede [version0.8] Laval University Library http://www.bibl.ulaval.ca/archimede/index.en.html	Java OS independent OCLC OAICat harvester Lucene search engine http://sourceforge.net/project/showfiles.php?group_id=112130	No	Yes	Yes	Yes	Keyword	All fields, fulltext, individual fields Type of search – fulltext, at least one word, phrase Boolean operators Date range Research group Document type Confirmation of search strategy before execution	Title/author – link to full record Collection – browse index Research group – browse index	Browse by research groups, collections, subjects, titles, authors, date	User account Email new resources
ARNO [version 1.1] University of Amsterdam, Tilburg University, University of Twente	Perl 5.6 Oracle Unix or Linux Apache Zebra search engine http://www.uva.nl/arno	No	Yes	Yes Multiple archives Roles – input, editor, administrator (by archive)	Yes	See Keystone				
CDSware [version: 0.3.1] CERN http://cdsware.cern.ch	Python PHP MySQL Apache Linux http://cdsware.cern.ch/news/	Yes Harvests internal and external metadata. Handles non-OAI-PMH repositories using XSL. Records stored in MARC21 format. http://cdsweb.cern.ch/	Yes	Yes Password One field per page, with summary view. User defaults Save partial records. Review records. File transfer - File upload, URL. File linkage by URL. Sends email to referee. Editing – select fields to edit. Submit revised file.	Yes File types – PS, PDF, Word, Latex, JPEG (thumb created on the fly)	Keyword search Limit options Prefix, infix and suffix truncation Phrase search	All words, exact phrase, any of the words, partial phrase, regular expression. Select a field Select Boolean operators Sort option Limit by collection type Fulltext search	Simple search display sorted by collection Option to display in single list. Further sort by author, title, report no. Select no. of records to display. Select display format Jump to record. Full record – all the above plus – full abstract, link to archive, direct link to individual file formats.	Browse keyword index – limit by any field. Browse collection type tree. Displays record nos. at each node. Perform search at any spot in the tree.	Add to basket. Link to canned search of Google or AllTheWeb. My account: Settings – email address, password, interface display. Searches – stores last 30 days. Basket – define specific collections, store records for later use. Alerts – run searches periodically.

Free Open Source OAI-PMH 2.0 compliant Software [as at 17 November 2004]

Product	Platform	Service Provider	Data Provider	Self-Archiving	Content Management	Simple search	Advanced search	Display	Browse	Other Features
<p>CWIS v1.1.0 (11/2/2004)</p> <p>Internet Scout Project demo:</p> <p>http://scout.wisc.edu/Projects/CWIS/</p>	<p>PHP 4.0.6 MySQL 2.23+ Linux Apache</p> <p>http://scout.wisc.edu/Projects/CWIS/downloads.php</p>	<p>No</p> <p>Could add a harvester</p>	<p>Yes</p>	<p>Yes</p> <p>Password – email address. Mandatory fields.</p> <p>Administration - web configurable – fields, mandatory, viewable in record display, keyword search, advanced search, search weight, field size, paragraphs/columns, qualifiers.</p>	<p>No</p>	<p>Search: http://scout.wisc.edu/Projects/CWIS/demo/</p> <p>Keyword search on all fields.</p>	<p>Fielded searching – Limit by selecting from dropdown menus - URL Domain, Date Issued, Resource Type, Language, Audience, Format.</p> <p>Phrase searching ""'. Dash – no results that contain that term. Plus sign - results to contain that term or phrase.</p>	<p>Brief display – link to resource, link to full record, rating.</p> <p>Sorted by relevance - combination of word occurrence frequency and a weighted value of the metadata fields being searched.</p> <p>Full record - canned search on - publisher, creator, subject, format, resource type, language; rating, comments.</p>	<p>Browse by subject (2 level tree). No. of classifications or records.</p>	<p>What's new. Individual account - append comments, ratings, discussion list, customise interface, save and re-run searches, SDI via email, recommend resources.</p>
<p>DLESE Catalogue v1.1 (17/9/02)</p> <p>Harvester v2.0.9 (2/9/03)</p> <p>ODL search client 1.0 (9/3/04)</p> <p>Digital Library for Earth System Education:</p> <p>http://www.dlese.org/libdev/dcs_overview.html</p>	<p>Java 2 Tomcat 4.0x Apache Ant 1.6.1 Linux 7.0x</p> <p>http://sourceforge.net/project/showfiles.php?group_id=23991</p>	<p>Yes</p> <p>Harvests internal and external metadata</p> <p>Harvester - Automatic updating and synchronizing. Gzip compression. Monitor harvest. Save files in folders by baseURL path, set and format. Under development – validation of records, indexing and search over sets.</p>	<p>Yes</p> <p>Records stored in XML</p>	<p>Yes</p> <p>No authentication.</p> <p>One field per page. Save records. Edit records via record ID.</p>	<p>Yes</p>	<p>Keyword search title, description, keyword, creator's last name, subject, resource type, URL.</p> <p>Phrase searching ""'. Wildcards *.</p> <p>Automatic righthand truncation. AND/OR</p>	<p>Fielded searching</p> <p>Limit by grade level, resource type, collections, standards.</p>	<p>Brief display – link to full record, link to collection. Sorted by relevance and reviewed collection, then by search terms matching title.</p> <p>Records displayed = 10. Refine search by keyword, grade level, resource type, collections, standards.</p> <p>Full record – technical requirements, cost/copyright, resource creator contact details and link.</p>	<p>Browse by subject (2-level tree), grade level, resource type. Limit by collection. Bar display of record numbers.</p>	<p>Add a teaching tip.</p>
<p>DSpace v1.1.1 (28/9/03)</p> <p>MIT</p> <p>http://libraries.mit.edu/dspace-mit/</p>	<p>Java 1.3+ Tomcat 4.0+ Linux Apache 1.3+ Ant 1.5 PostgreSQL 7.3+</p> <p>http://libraries.mit.edu/dspace-mit/technology/download.html</p>	<p>No</p> <p>Batch load data</p>	<p>Yes</p> <p>Can't be harvested by search engines without conversion: http://dlib.cs.odu.edu/dp9/</p>	<p>Yes</p> <p>https://dspace.mit.edu/index.jsp</p>	<p>Yes</p> <p>Mandatory - must be stored internally</p>	<p>Keyword Case insensitive. Right truncation Unknown characters ? Stop words. Word stemming.</p>	<p>Field search, single or in combination. No subject thesaurus. And, or, not.</p>	<p>Brief display – date, title (link to extended display), authors. Extended record display. Collection link. File name, size, type, link. DC display Multiple files for one item.</p>	<p>Browse by communities/collection, title, author, date. No subject.</p>	<p>New collections Recent submissions My DSpace – email updates, restricted content Usage reports by community.</p>

Free Open Source OAI-PMH 2.0 compliant Software [as at 17 November 2004]

Product	Platform	Service Provider	Data Provider	Self-Archiving	Content Management	Simple search	Advanced search	Display	Browse	Other Features
<p>Fedora [1.2.1 20/4/2004] University of Virginia, Cornell University http://www.fedora.info/</p> <p>Designed as a set of building blocks rather than out-of-the-box solution. Additional scripts required.</p> <p>No working sites as yet</p>	<p>Java J2SDK 1.4+ mySQL 3.23.x+ http://www.fedora.info/release/1.2.1/?download=Download+Fedora+1.2.1</p>	<p>No Batch loading via SOAP-enabled utility</p>	<p>Yes</p>	<p>No - under development Batch loading</p>	<p>Yes Internal (batch-loading) or external Object versioning Thumbnails Objects encoded in XML/METS+</p>	<p>Keyword searching across selected fields</p>	<p>Field-specific searching</p>			<p>Shibboleth support in a future release.</p>
<p>GNU EPrints [2.3.3] Southampton University http://www.eprints.org/</p>	<p>Perl MySQL Apache Linux http://Platform.eprints.org/</p>	<p>No see Citebase</p>	<p>Yes</p>	<p>Yes</p>	<p>Yes</p>	<p>http://eprint.uq.edu.au/ Quick search – keyword. Simple search - Keyword in title/abstract/keyword, Author/editor. Select “and” or “or” for each field. Publication year range. Record may fill all or any of the above criteria. Select order of results – oldest first, most recent, author, title. Limit search by archive (sub-views in latest release)?</p>	<p>Match all in any order, match any of: Title, authors, abstract, keywords, subject (select from a drop-down menu), type, conference title, department, editors, status (unpublished, in-press, published), refereed (no preference, yes, no), publication title, publication year range, Record may fill all or any of the above fields. Select “and” or “or” for each field Order of results – oldest first, most recent, author, title. Limit search by archive (sub-views in latest release)?</p>	<p>Brief display - title (link to full record). Records displayed – 10. Further limit search by any of the original search options. Full record – link to local fulltext, subjects (link to subject browse index), depositor (link to contact details). Fulltext searching (latest release)? Link to author record with contact details?</p>	<p>Browse by year, subject or author. Subject – displays tree with no. of entries. Allows links to empty terms. Standard brief display. Browse by archive (sub-views in latest release)?</p>	<p>Latest additions. Statistics – authors, papers.</p>
<p>Greenstone [3] University of Waikato http://www.greenstone.org</p>	<p>Linux or Windows mySQL GDBM, MG indexing) Perl (building) C++ (runtime) Tomcat servlet</p>	<p>No Can import files and metadata</p>	<p>Yes</p>	<p>Yes 4 levels of expertise – assistant librarian, librarian, systems specialist, expert</p>	<p>Yes Collections and sub-collections Unicode Text, graphics, audio, video Plugins to handle web pages, PDF, MS Word, email</p>	<p>Keyword on pre-defined fields Fulltext searching Some/all the words</p>	<p>Search preferences eg case, word endings, Boolean operators, search history, no. of records to display. Phrase searching Switch languages Search all or pre-defines collections</p>	<p>Record summary Support for thumbnails Title link to full record</p>	<p>Browse on selected indexes, e.g. title, subject, people, dates</p>	
<p>i-Tor version 1.1.10 Netherlands Institute for</p>	<p>Java mySQL</p>	<p>Yes</p>	<p>Yes Also a web</p>	<p>Yes Creating</p>	<p>Primary web content</p>	<p>Keyword search of fulltext</p>	<p>Title, subject, summary, author</p>	<p>Title (link to full record), author, date, type of file,</p>		<p>Email alert of new and</p>

Free Open Source OAI-PMH 2.0 compliant Software [as at 17 November 2004]

Product	Platform	Service Provider	Data Provider	Self-Archiving	Content Management	Simple search	Advanced search	Display	Browse	Other Features
Scientific Information Services http://www.i-tor.org/en/	Linux Lucene search engine http://sourceforge.net/projects/i-tor		content management system	permissions for reading and writing	management Produce and edit new objects, link to exist objects Fine-tuning design using templates		Limit by date range Search all or selected archives Automatic "and" Fuzzy logic Search for similar documents	repository		changed content
iVia University of California, Riverside http://infomine.ucr.edu/Via/	Debian Linux Apache Most code in C++ Some interface code in Java, PHP MySQL database Berkeley DB database (for fast indexing functions) http://infomine.ucr.edu/iVia/download.php	Yes Administration manual http://infomine.ucr.edu/docs/admin_manual/ Infomine http://infomine.ucr.edu/	Yes	Yes	No	Keyword	Rich, full-text and/or fielded metadata searchable Boolean and proximity operators supported Search modification on results page Spell check and spelling suggestions for queries Limit searching	Titles only, Regular (ti, su, creator, description, expert/robot created), All <i>Results ranking:</i> relevance or title	All general categories or combinations of general categories (e.g., Gov. Info. and/or Bio./Ag./Med) Subject indexes (e.g., LCSH or LCC) Keywords Specific research disciplines hierarchy	Email alert service for new resources added to the collection New resource suggestions Comment on this resource
Keystone Digital Library Suite Index Data, Copenhagen http://www.indexdata.dk/keystone/	Perl Linux (Debian) Apache 1.x+ Sablotron XSLT processor http://ftp.indexdata.dk/pub/tkl/ Database not included	Yes Support for sets Harvester requires a lot of tweaking to work. Examples: 1. Danish Electronic Research Library http://deff.dk/index_eng.html 2. Library of Texas http://libraryoftexas.org/	Yes	Yes	Yes	Keyword search Boolean operators	All fields, title, author, abstract Boolean operators * wildcards, righthand truncation Limit to fulltext Limit by collection	Total records, display 10 per page One line per record – title and author Click on title to view full record	Browse by subject – no. of entries, 6 per page Sort by title (default), resource type	Meta searching – Z39.50, SOAP Clusters Save personal selections
MyCoRe Version 0.9 University of Essen, University of Bonn, IBM http://www.mycore.de/e	Java Tomcat mySQL Apache Lucene search	Unknown	Yes	Yes Workflow and workbasket	Yes	Keyword	Fulltext			

Free Open Source OAI-PMH 2.0 compliant Software [as at 17 November 2004]

Product	Platform	Service Provider	Data Provider	Self-Archiving	Content Management	Simple search	Advanced search	Display	Browse	Other Features
nql/index.html	engine http://www.mycore.de/download.html									
Open Digital Library /Digital Library in a Box – Virginia Tech	Java 1.2.2+ Cocoa 0.5. Apache Tomcat MySQL with JDBC driver http://oai.dlib.vt.edu/odl/ http://dlbox.nudl.org/downloadlog/downloadqs.html	Yes Harvests internal and external metadata Search: http://www.citidiel.org/ http://oai.dlib.vt.edu/~etdunion/cgi-bin/index.pl	Yes Search: http://www.cstc.org/cgi-bin/search.pl	Yes	Yes	Keyword search on most fields. Supports natural language searches. Boolean and = +, not = - Limit by field using colon eg "creator:" Weighted fields eg title=5.	Search by author Limit by (dropdown menus) – date from, resource type, subject, institution?	Default brief display – author (canned search). Links to – full record, document, find similar documents. Full record – DC simple record, identifier links to document.	Browse by –type, institution, subject, status. Sort by – date, author, type, title. Subject – 3 level tree, includes no. of records. 2 nd level – display record summary, plus links to 3 rd level and no. of records, search within subject. Institutions – brief display of all records	Recent additions. Annotate records. Recommend additions. Ratings.
Opus Stuttgart University http://elib.uni-stuttgart.de/opus/	PHP Ht://Dig fulltext search engine	No Supports metasearching	Yes Also via <i>dlmeta</i> XML format	Yes	Yes	Fulltext Match – all, any boolean Display –short, long Sort – score, time, title, reserve	Title, author, abstract, date, subject (multiple classifications) Keyword or phrase Boolean operators No. of records to display Sort by title, year	Advance search brief display – title (link to full record), author, date Full record includes URN	By faculty, publication type	Metasearching of other Opus sites?

Free Open Source OAI-PMH 2.0 compliant Software [as at 17 November 2004]

Product	Platform	Service Provider	Data Provider	Self-Archiving	Content Management	Simple search	Advanced search	Display	Browse	Other Features
Public Knowledge Project University of British Columbia	PHP MySQL Linux http://pkp.ubc.ca/pkp-harvester/download.html	Yes	Yes	No	No	Search: http://www.pkp.ubc.ca/harvester/ Free text search of any field Restrict search to one of – author, title, abstract, index term	Free text search limited to a specific archive Search by – author, author affiliation, title, abstract, date range, index terms – discipline, subject, approach/method, coverage, language, sponsor, type	Brief display – link to view full record, link to original Full record – title, archive (hotlink), url (hotlink), author, date, abstract, index terms, publisher, contributors, source, language, relation, type, format, copyright info. Quick search available from results page. “Research support tool” – link to related resources: (1) For this item – view metadata, capture cite (2) Context – author bio, define terms, ejournals, related theory, related studies, online forums, instructional, media reports (3) Action – email author/others, add to portfolio	By archive Brief display, excluding author, but in author order! Option to sort by title or date. Option to search within selected archive.	Save and edit records in personal research portfolio.
RDN and UC Berkley	TCL/TK scripting Linux Apache http://cheshire.lib.berkeley.edu/ [search engine] http://www.rdn.ac.uk/publications/tools/ [Perl harvester]	Yes Plugin harvests external metadata Removes deleted records	No	No	No	Search: http://cheshire.lib.berkeley.edu/webinterface.html http://www.rdn.ac.uk/ Keyword search Automatic “and” Boolean “or”, “not”	Keyword in field search – author, title, subject, topic (searches a number of fields and returns relevance-ranked results) Select no. of records, short/medium and long display	Can't toggle from selected display to others) Display includes canned searches on author, subject	No	
University of Illinois at UC OAI Metadata Harvesting Project	JRE (1.3+) MySQL Linux Apache + Tomcat (4.0+) http://uilib-oai.sourceforge.net	Yes	Yes	Yes	Yes	Keyword Auto “and” Phrase “” Limit to online primary resources	Keywords – all fields, author, title/subject Limit by date range, type	Brief display – search terms highlighted Links to - collection website, full record, item Limit by type Full record – link to item, return to results	Browse by type – displays all records, then offers limit by type! No browsing by collection	Search history Book bag