

The Open Citation Project

Final (Year 3) Report to JISC

Project Manager: Steve Hitchcock
Lead Institution: Southampton University
Duration of Award: 10/99-09/02
Period of Report: 10/01-12/02

Version history of this report

This version 1.1.1 Final draft (public version)

First year report <http://opcit.eprints.org/y1report/y1report-final.pdf>

Second year report <http://opcit.eprints.org/y2report/y2report20.pdf>

1 Background to project

It has always been clear that there is not enough awareness of the importance of open access to published research papers — which means that all users can access the papers free of charge, at any time, anywhere — among the research and academic community. What is needed to increase the provision and use of open access are real tools and services to show that open access works, in forms that are transparently beneficial to authors of research papers. Central to this is the causal connection between research access and research impact: open access increases impact¹.

In the UK there are signs the next Research Assessment Exercise (RAE) will use citation analysis² in its search for a “less burdensome assessment method”³ of measuring the impact of research. Open access would ensure that research is “assessable continuously online”⁴ because **the tools to build open access archives and enable authors to provide open access to their papers, and to measure impact by citation measures as well as usage measures, are already available as a result of the Open Citation (OpCit) Project** (<http://opcit.eprints.org/>).

These tools are GNU EPrints (also known as eprints.org) software for building open-access archives, and Citebase, a citation-ranked search engine for open archives. At the completion of OpCit **these tools are available for others to use, and they will be applied in a number of projects in the JISC FAIR programme**. They will also enhance two major initiatives that the project helped launch to raise awareness of open access, the Open Archives Initiative (OAI) and the Budapest Open Access Initiative.

¹ Steve Lawrence, Free Online Availability Substantially Increases a Paper's Impact, *Nature Web Debate on e-access*, May 2001. <http://www.nature.com/nature/debates/e-access/Articles/lawrence.html>

² Sam Jaffe, Citing UK Science Quality: The next Research Assessment Exercise will probably include citation analysis, *The Scientist*, Vol. 16, No. 22, Nov. 11, 2002 http://www.the-scientist.com/yr2002/nov/prof1_021111.html

³ Mandy Garbner, RAE review: Tired, hard-up RAE seeks saviour with light touch, *Times Higher Education Supplement*, Dec. 6, 2002

⁴ Stevan Harnad, Why I think research access, impact and assessment are linked, *Times Higher Education Supplement*, May 18, 2001

2 Methodology

The principal partners in the Open Citation Project were Southampton University's IAM Group, the Digital Library Research Group at Cornell University, and arXiv, at the outset of the project based at Los Alamos and now hosted at Cornell.

The method used by the project at Southampton has been to build tools to measure and analyse citations from the 200,000+ papers stored by the arXiv physics archives, the largest eprint archive of its type. These data have been complemented, experimentally, with data on how the archives are used, e.g. which papers are viewed most. Collectively the citation and usage data are stored in Citebase, a citation database which provides a user interface for search and discovery, and a machine interface for analysis of this rich data source by other services.

With the emergence of OAI and the consequent emphasis on institutional archives, it was evident there would be a need for large numbers of local, institution-based archives smaller than arXiv, but which would need to operate on similar principles — low cost, largely automated deposit, offering indexing and dissemination of author-archived content. Software used to build CogPrints, a cognitive sciences archive modelled on arXiv, was rewritten to make it OAI-compliant, and then to make it generic. This became the basis of GNU EPrints, which was further developed within the remit of OpCit to generalise the author and management interfaces for open-access archives.

Of most significance, EPrints builds archives that comply with the OAI Protocol for Metadata Harvesting (PMH). This means that any content deposited within an EPrints-based archive will become visible to users of independent OAI services, such as Citebase, immediately enhancing the chances of discovery. Authors depositing papers in an EPrints archive are not required to have any knowledge of OAI metadata, as it is generated automatically.

Connecting papers in open archives and a citation database is a method for automatically extracting metadata and reference lists from the papers. There are many different applications for reference linking. The project at Cornell considered the question "what would be the ideal behavior of a digital object that supported reference linking (both incoming and outgoing)"? Answering this question led to an application programming interface (API) for reference linking.

All three components have been tested, evaluated and demonstrated to be useful by third-party users, and will continue to be developed and integrated within new projects and products beyond the lifetime of the OpCit project.

3 Activities

The activities of the OpCit project were described by Hitchcock *et al.* (2002a).

3.1 Citebase

Citebase is a citation-ranked search and impact discovery service that measures citations of scholarly research papers that are available on the Web in the larger open access, OAI disciplinary archives - currently arXiv (<http://arxiv.org/>), CogPrints (<http://cogprints.soton.ac.uk/>) and BioMed Central (<http://www.biomedcentral.com/>). Citebase harvests OAI metadata records for papers in these archives, automatically extracting the references from each paper. The association between document records and references is the basis for a classical citation database.

The primary means for users of accessing this database is the Citebase Web interface (<http://citebase.eprints.org/>) (Figure 1). The user can classify the search query terms (typical of an advanced search interface) based on metadata in the harvested record (title, author, publication, date). In separate interfaces, users can search by archive identifier or by citation. What differentiates Citebase is that it also allows users to select the criterion for ranking results by Citebase processed data (citation impact, author impact) or based on terms in the records identified by the search, e.g. date. It is also possible to rank results by the number of 'hits', a measure of the number of downloads and therefore a rough measure of the usage of a paper. This is an experimental feature to analyse the quantitative and the temporal relationship between hit (i.e. usage) and citation data, as measures of impact. Hits are currently based on limited data from download frequencies at the UK arXiv mirror at Southampton only.



Figure 1. Citebase search interface showing user-selectable criteria for ranking results

The combination of data from an OAI record for a selected paper with the references *from* and citations *to* that paper is also the basis of the Citebase record for the paper. A record can be opened from a search results list. The record contains bibliographic metadata and an abstract for the paper, from the OAI record. This is supplemented with four characteristic services from Citebase:

- Graph of this Article's Citation/Hit History
- All Articles Cited by this Article (Reference List)
- Top 5 Articles Citing this Article (option to view All Articles Citing this Article)
- Top 5 Articles Co-cited with this Article (option to view All Articles Co-Cited with this Article)

Another option presented to users from a results list is to open a PDF version of the full paper. This option is also available from the record page for the paper. This version of the paper is enhanced with linked references to other papers identified to be within arXiv, and is produced by OpCit. An earlier evaluation found that arXiv papers are the most appropriate place for reference links because users overwhelmingly use arXiv for accessing full texts of papers, and references contained within papers are used to discover new works. (see <http://opcit.eprints.org/evaluation/v10/v10evaluation.html>).

Prior to a more recent evaluation (section 3.2) Citebase had records for 230,000 papers, indexing 5.6 million references. By discipline, approximately 200,000 of these papers are classified within arXiv physics archives.

3.1.1 Extending analysis of citation data: OpCit e-Services

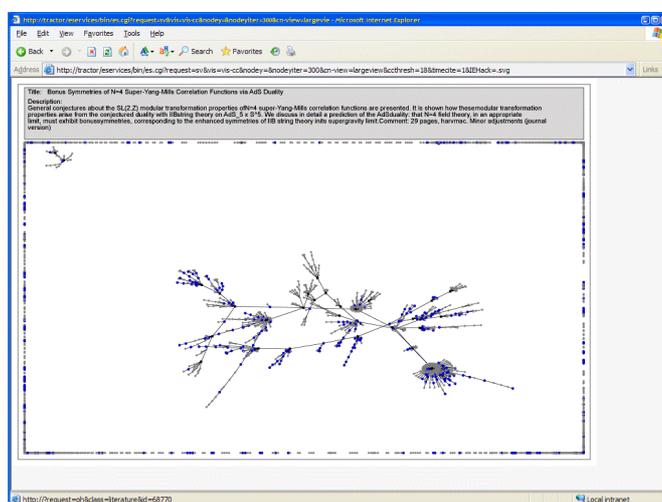


Figure 2. Co-citation map of the entire arXiv collection

A first attempt to extend the analysis and presentation of citation relationships was explored with OpCit e-Services (<http://opcit.eprints.org/eservices/>). Like Citebase, the e-Services framework uses OAI metadata. The case illustrated in Figure 2 uses data from Citebase, for which it then provides advanced services:

- simple visualisations (e.g. number of e-print deposits each year)
- knowledge services (e.g. most significant papers)
- co-citation visualisations (uses the co-citedness of papers as a proximity measure when plotting papers on a graph) (Figure 2)

The approach needs refinement before user interface issues can be tackled. First, the large dataset causes computation to slow significantly. Second, due to erroneous or missing citations, some visualisations may not display convincing or useful patterns.

OpCit e-Services shows that, when no single interface can serve all users, open access to data offers the opportunity to produce multiple interfaces to serve different purposes.

3.2 Evaluation of Citebase

This was the first detailed investigation of the impact on users of an open access Web citation indexing service. The evaluation, including details of methodology, design and results, has been reported by Hitchcock *et al.* (2002b).

The following elements of Citebase were the focus of the evaluation:

- The primary search interface
- Services available from a Citebase record
- Linked PDFs
- The effectiveness of the means of navigating between these services
- Support for users when these services fail to produce the required result

Given the wide prospective user base, what was evaluated was not just the current implementation of Citebase, but the *principle* of citation-based navigation and ranking.

The evaluation sought to:

1. evaluate the usability of Citebase (can it be used simply and reliably for the purpose intended)
2. assess the usefulness of Citebase (how does it compare with other services)
3. measure user satisfaction with Citebase
4. raise awareness of Citebase
5. inform ongoing development of Citebase

The evaluation used two methods to collect data:

- A two-part Web-based questionnaire
 - Form 1: about users; a practical exercise - building a short bibliography; views of Citebase
 - Form 2: user satisfaction
- Usage statistics for the Citebase search page

The evaluation was open from June 2002, when the first observational tests took place, to the end of October 2002 when a closure notice was placed on the forms.

Valid submissions to Form 1 were received from 195 evaluators. Although the primary target group were physicists, responses also came from mathematicians, computer scientists, information scientists, cognitive scientists, biologists, health scientists, and others.

3.2.1 Usage of Citebase

The current target user group for Citebase is physicists. The impact being made by OAI should help extend coverage significantly to other disciplines, although because the emphasis of OAI is on promoting institutional archives the impact on disciplines, as measured by services such as Citebase, may take longer to emerge. For this reason

there was a need to target this evaluation at prospective users, not just current users. Citebase should be designed for an expanding user base.

Prior to evaluation Citebase had not been announced generally and was little used. The evaluation was first announced to selected discussion lists targeted at: colleagues in digital library research, advocates of open access to the scholarly literature, and librarians. **The most significant contributor to increased usage was the inclusion of links, on a trial basis, from abstract pages of papers in arXiv to the corresponding Citebase records.** Links from arXiv became active on 20th August.

A notable success of the evaluation has been to increase usage of Citebase, in terms of average daily visits, by more than a factor of 10. There is still considerable scope to increase usage of Citebase by arXiv physicists. According to Paul Ginsparg, founder of arXiv: "(Citebase) is a potentially critical component of scholarly information architecture".

3.2.2 Principal findings of the evaluation

Overall, results of the evaluation show there is much scope for improvement, but as exemplified by Citebase **Web-based citation indexing of open access archives is closer to a state of readiness for serious use than had previously been realised.**

Within the scope of its primary components, the search interface and services available from a Citebase record, it was found **Citebase can be used simply and reliably for resource discovery.** The majority of users were able to complete a task involving all the major features of Citebase. **More data need to be collected and the process refined before it is as reliable for measuring impact.** As part of this process users should be encouraged to use Citebase to compare the evaluative rankings it yields with other forms of ranking.

Citebase is a useful service that compares favourably with other bibliographic services, although it needs to do more to integrate with some of these services if it is to become the primary choice for users.

The linked PDFs are unlikely to be as useful to users as the main features of Citebase. Among physicists, linked PDFs will be little used, but the approach might find wider use in other disciplines where PDF is used more commonly.

One of the most important findings of the evaluation is that **Citebase needs to be strengthened in terms of the help and support documentation it offers to users.**

The first step must be to examine the results of this evaluation to improve Citebase with a view to establishing it as a service used regularly by arXiv users.

There are wider objectives and aspirations for developing Citebase. **Where there are gaps in the open-access literature Citebase will motivate authors to accelerate the rate at which these gaps are filled, especially when it is realised there is a direct correlation, which Citebase will confirm, between open access, increased impact and the outcomes of research assessment exercises.**

3.3 GNU EPrints

EPrints is software for building open-access archives aimed at institutions and special-interest communities, and is now used by nearly 60 archives.

In its current incarnation, the name GNU EPrints reflects that it is open source and freely available under the GNU General Public License and conforms to the strict GNU guidelines for free software. The last major release of EPrints, version 2.0, appeared in February 2002, although it has been updated (now on version 2.2.1) to conform with the latest OAI-PMH (also version 2) announced in June. Features of EPrints version 2, described by Gutteridge (2002), include:

- Internationalised metadata stored as Unicode
- Support for multiple archives on one server
- An improved user interface

EPrints is extending its focus on institutional research papers. It is now configurable for adoption as a journal-archive, e.g. *Behavioral and Brain Sciences* and *Psychology*, by new open access journals or established journals converting to open access, and will include the facility to manage peer review and peer commentary. **It is planned to extend EPrints for structured data handling in, e.g. e-science applications.**

3.4 Reference linking API

The API automatically extracts metadata and reference lists from papers using four principal methods:

1. `getMyData()` - the digital object should emit standard metadata describing that object, i.e., title, authors, year of publication, etc. in Dublin Core format.
2. `getReferenceList()` - the digital object should say what its list of references is (this is the fixed number of references contained in the online document).
3. `getCitationList()` - the object can say what other works the object knows have cited it. (This list grows as more and more items are analyzed.)
4. `getLinkedText()` - returns the original content of the digital object but with link information added to it so that each reference can be used to go directly to an online copy of the referenced work, if an online copy is available.

Each component produced by these methods can be seen in a typical Citebase record, but **this approach is generalisable to other reference linking applications.**

A few Java classes were defined to support reference linking in an object oriented way. These methods can be invoked on the surrogate, a special class in the API that encapsulates data regarding a particular online digital object. To use the API, a new surrogate is instantiated, passing it the URL of the online digital object for which information is to be gathered.

The bulk of the analysis within the API program is done by the surrogate constructor. This call downloads the online work, turns it into XHTML, parses the XHTML, and extracts information, such as citations and references. The next call on the API

invokes the method that returns the references in the form of an XML document, which is then converted to a string and printed.

It is anticipated that repositories will at some point contain reference linking data, so the API was later extended to support persistent storage of surrogates. Once a surrogate is instantiated, it can be saved to a repository, if desired. Thus one could build a repository of surrogates, which could later be re-instantiated and have the basic API methods invoked on them.

3.4.1 API evaluation

The API was used to build several applications against online journals (*D-Lib Magazine*, *Journal of Electronic Publishing*, ACM Digital Library). With five methods (the original four, plus save) the API was found to be sufficiently usable. The main limitation of the software is that not all HTML pages are equally easy to analyse, e.g. some HTML is badly written and cannot be converted into XHTML and, therefore, cannot be parsed. This is likely to remain a problem on the Web for some time. A more complete description of the reference linking API and its evaluation, including the *D-Lib* application, can be found in Bergmark and Lagoze (2001).

4 Outputs

All three components described above, and a new component, Paracite, a software agent and search interface for parsing and locating raw references on the Web, are usable and will continue to be so beyond the conclusion of OpCit. What is available, the means of access, and plans for maintenance of services, are noted below:

- Citebase is now up-to-date and indexes arXiv fully. Citebase can be searched by users at <http://citebase.eprints.org/>. A machine interface for data sharing with other services is operational, and Citebase is listed as an OAI 2.0-conforming data provider (<http://www.openarchives.org/Register/BrowseSites.pl>). Researchers at Old Dominion University have harvested Citebase data as part of their Archon federated digital library on physics (<http://archon.cs.odu.edu/>), as has the OAI search engine OAIster (<http://oaister.umdl.umich.edu/o/oaister/viewcolls.html#c>), and arXiv is a possible (re)harvester of Citebase data too. Due to ongoing developments with the data formats, enquiries about the machine interface should be directed at the developer, Tim Brody tdb01r@ecs.soton.ac.uk. **The citation database will continue to be updated and expanded in terms of coverage. Both interfaces to Citebase will continue to be developed and maintained.**
- GNU EPrints is available as open source software and is downloadable from <http://software.eprints.org/>. Machine requirements for running GNU EPrints are other open source components including Linux, Apache Web server, Perl and a MySQL database. **GNU EPrints will continue to be developed and maintained.**
- The Reference linking API was written in Java and is downloadable from the OpCit project site at Cornell <http://www.cs.cornell.edu/cdlrg/Reference%20Linking/>. **The API is not being developed at present.**

- Paracite is still experimental, but can be tried at <http://paracite.eprints.org/>. There are plans to use the reference linking API within Paracite. As well as providing a user interface, Paracite could mediate between data sources (archives) and linking services (citation databases, OpenURL, etc.). **Paracite development is ongoing.**

5 Impacts

The ideas that have characterised OpCit will be taken forward not just in the products of the project, such as Citebase and GNU EPrints, but in new environments, such as the JISC FAIR programme.

Potentially, the biggest impact the project could make rests with the restructuring of the UK RAE. The tools to manage open access and measure impact are ready and demonstrated, and the benefits to the RAE have been noted by those responsible for it. The advantages to the RAE and the cause of open access are mutually reinforcing. By virtue of having a RAE “the UK is uniquely placed to lead the world”⁴ in the uptake of open access and as a consequence to strengthen the impact of UK research. **By association there are clear benefits to JISC: recognition of its role in funding the work that has made this breakthrough possible, and greater impact of its ongoing projects and the JISC Information Environment.**

5.1 JISC FAIR programme

The JISC Focus on Access to Institutional Resources (FAIR) programme includes major projects that will extend the use of EPrints-based archives in UK universities through the provision and targeting of new archives and supplementary services:

- SHERPA (Securing a Hybrid Environment for Research Preservation and Access), will build EPrints-based archives at six UK universities, using this experience to report on the implications for management and quality control of the archives.
- E-Prints UK plans to use Citebase software and citation data to enhance its database for discovery of eprint papers available from open access archives hosted at UK universities and colleges.
- TARDIS (Targeting Academic Research for Deposit and dISclosure) will investigate strategies 'to overcome the technical, cultural and academic barriers', which might be found to be restricting the development of institutional eprint archives, by developing a working model of a multidisciplinary institutional archive based on EPrints.
- RoMEO (Rights MEtadata for Open archiving) will canvas users to identify (mis)perceptions about how rights should be formulated and protected for 'give away' works — "texts from which the author does not seek sales revenue" — promoting practical approaches that can "assigned, disclosed, harvested, and displayed" via the OAI-PMH.

5.2 Enhancing OAI infrastructure

To improve interoperability, scalability and reliability of OAI services, OpCit has worked with a team from Old Dominion University (USA) on infrastructure components such as proxies and caches (Liu *et al.* 2002). Serious errors in OAI data require an intermediate storage approach: aggregation and caching. Celestial, an OAI aggregator, is software developed to act as a buffer between Citebase and the source repositories so that, e.g. it doesn't overload arXiv. Celestial harvests metadata from OAI-compliant repositories and re-exposes that metadata to other services - in effect an *OAI cache*. Celestial software is available from <http://oai-perl.sourceforge.net/>

5.3 OpCit people and products in the news

EPrints software is undoubtedly the better known product of the OpCit project, and this is reflected in coverage in news and feature sources shown below. Citebase or similar services will ultimately have more impact with users, but EPrints is necessary now and plays a critical role in enabling open-access archives to be filled.

- Colin Steele, E-prints: the future of scholarly communication? *InCite*, October 2002 <http://www.alia.org.au/incite/2002/10/eprints.html>
- Konrad Lischka, Der Geist, der aus der Flasche kam, *Telepolis magazine*, 16th March 2002 (in German) <http://www.heise.de/tp/deutsch/special/copy/12031/1.html>

Citebase

- Belinda Weaver, Open archives citation tool, *InCite*, October 2002 <http://www.alia.org.au/incite/2002/10/weaver.html>

EPrints

- Roy Tennant, Institutional Repositories, *Library Journal*, 15th September 2002 <http://libraryjournal.reviewsnews.com/index.asp?layout=articlePrint&articleID=CA242297&publication=libraryjournal>
- Georg C. F. Greve, Brave GNU World - GNU EPrints, *Linux Magazin*, September 2002 (in German) <http://www.linux-magazin.de/Artikel/ausgabe/2002/09/bgw/bgw.html>
- Raym Crow, The Case for Institutional Repositories: A SPARC Position Paper, The Scholarly Publishing & Academic Resources Coalition, August 2002 <http://www.arl.org/sparc/IR/ir.html>
- Kendra Mayfield, College Archives 'Dig' Deeper, *Wired News*, 3rd August 2002 <http://www.wired.com/news/school/0,1383,54229,00.html>
- Jeffrey R. Young, 'Superarchives' Could Hold All Scholarly Output, *Chronicle of Higher Education*, 5th July 2002 <http://chronicle.com/free/v48/i43/43a02901.htm>
- Anon. The ghost is out of the bottle, *Higher Education & Research Opportunities in the UK*, 29th March 2002 http://www.hero.ac.uk/inside_he/the_ghost_is_out_of_the_b1365.cfm
- Ivan Noble, Boost for research paper access, *BBC Online News*, 14th February 2002 <http://news.bbc.co.uk/1/hi/sci/tech/1818652.stm>
- Ed Sponsler and Eric F. Van de Velde, Eprints.org Software: A Review, *SPARC E-News*, August-September 2001 <http://www.arl.org/sparc/core/index.asp?page=g20#6>
- Kendra Mayfield, The Science of E-Publishing, *Wired News*, 19th October 2000 <http://www.wired.com/news/culture/0,1284,39323,00.html>

6 Future priorities

Citebase

1. Establish full arXiv service, with unqualified links from arXiv to Citebase.
2. Extend coverage to include RePEc, which indexes distributed archives of working papers on economics, and GNU EPrints-based OAI institutional archives, applying citation analysis to institutional archives for the first time.
3. Improve support documentation.

GNU Eprints

1. Extend support for structured data handling in, e.g. e-science applications, enabling subparts of very large data archives to be used in analysis without the need to copy contents in their entirety to the processing site.
2. Add peer review management for journal applications.

Paracite

1. Retrieve a reference list from an online article using the RefLink API.
2. Provide a free OpenURL resolver that combines the parsing functionality of Paracite with the OpenURL syntax.
3. Reference checking by taking reference metadata and a template and rebuilding the reference (e.g. with underlining to show the user whether the parsing worked).

OpCit publications 2001-2

The following papers and reports **were produced by the project during the final year of its work from September 2001**. All publications covering the project back to 1999 are listed at <http://opcit.eprints.org/opcitpapers.shtml>

Bergmark, D. and Lagoze, C. (2001) "An Architecture for Automatic Reference Linking". *5th European Conference on Research and Advanced Technology for Digital Libraries (ECDL)*, Darmstadt, September
<http://www.cs.cornell.edu/cdlrg/Reference%20Linking/tr1842.ps>

Bergmark, D., Phempoonpanich, P. and Shumin Zhao, S. (2001) "Scraping the ACM Digital Library". *SIGIR Forum*, Vol. 35 No. 2, Fall
<http://www.acm.org/sigir/forum/F2001/bergmarkFinal.pdf>

Brody, T., Carr, L and Harnad, S. (2002) "Evidence of Hypertext in the Scholarly Archive". *Proceedings of HT'02, the 13th ACM Conference on Hypertext*, University of Maryland, June 2002
<http://opcit.eprints.org/ht02-short/archiveht-ht02.pdf>

Gutteridge, C. (2002) "GNU EPrints 2 Overview". Author eprint, Dept. of Electronics and Computer Science, Southampton University, October, and in *Proceedings 11th Panhellenic Academic Libraries Conference*, Larissa, Greece, November
<http://eprints.ecs.soton.ac.uk/archive/00006840/>

Gutteridge, C. and Harnad, S. (2002) "Applications, Potential Problems and a Suggested Policy for Institutional E-Print Archives". Author eprint, Dept. of Electronics and Computer Science, Southampton University, September
<http://eprints.ecs.soton.ac.uk/archive/00006768/>

Harnad, S. (2001) "Skyreading and Skywriting for Researchers: A Post-Gutenberg Anomaly and How to Resolve it". *text-e virtual symposium*, 14 – 30 November
http://text-e.org/conf/index.cfm?ConfText_ID=7

Harnad, S. (2003) "Electronic Preprints and Postprints". *Encyclopedia of Library and Information Science* Marcel Dekker, Inc.
<http://www.ecs.soton.ac.uk/~harnad/Temp/eprints.htm>

Harnad, S. (2003) "Online Archives for Peer-Reviewed Journal Publications". *International Encyclopedia of Library and Information Science*. Edited by John Feather and Paul Sturges. Routledge
<http://www.ecs.soton.ac.uk/~harnad/Temp/archives.htm>

Hitchcock, S., Bergmark, D., Brody, T., Gutteridge, C., Carr, L., Hall, W., Lagoze, C. and Harnad, S. (2002a) "Open Citation Linking: The Way Forward". *D-Lib Magazine*, Vol. 8, No. 10, October
<http://www.dlib.org/dlib/october02/hitchcock/10hitchcock.html>

Hitchcock, S., Woukeu, W., Brody, T., Carr, L., Hall, W. and Harnad, S. (2002b) "Evaluating Citebase, an open access Web-based citation-ranked search and impact discovery service". Evaluation report, IAM Dept., University of Southampton
<http://opcit.eprints.org/evaluation/Citebase-evaluation/evaluation-report.html>

Liu, X., Brody, T., *et al.* (2002) "A Scalable Architecture for Harvest-Based Digital Libraries - The ODU/Southampton Experiments". *D-Lib Magazine*, Vol. 8, No. 11, November
<http://www.dlib.org/dlib/november02/liu/11liu.html>
First available as *arXiv Computer Science cs.DL/0205071*, May 2002
<http://arxiv.org/abs/cs.DL/0205071>

OpCit conference presentations 2002

OpCit-related presentations were given at the following meetings during the final year of the project in 2002. The full list of presentations, including presentations from previous years, with a link to keynote presentations by Stevan Harnad, can be found at <http://opcit.eprints.org/opcitpapers.shtml>

November 6-8 "Academic Libraries of Open and Continuous Access", *11th Pan Hellenic Conference of Academic Libraries*, Larissa, Greece

October 17-19 "Gaining independence with e-prints archives and OAI", *2nd Workshop on the Open Archives Initiative (OAI)*, CERN, Geneva

September 13 "Open Access Journals - will they fly?" *ALPSP/OSI round table meeting*, London

June 24-25 *JISC/NSF Digital Libraries Initiative (DLI) All Projects Meeting*, Edinburgh

May 29 "Applications of Metadata", a one-day conference organised by the BCS Electronic Publishing Specialist Group, London

May 13-14 *First Workshop of the Open Archives Forum*, Pisa

April 12 "We can't go on like this: the future of journals", *ALPSP International Learned Journals Seminar*, London

March 22 *The Future of Journal Publishing*, Nottingham University

March 4 *CURL ePrints workshop*, Glasgow

January 24-25 *JISC All-Projects Synthesis Meeting*, Manchester

OpCit people

The following **researchers** were involved with the Open Citation Project during the year reported:

Stevan Harnad, Carl Lagoze (Principal Investigators), Wendy Hall (Management Chair), Les Carr (Project Technical Director), Steve Hitchcock (Project Manager), Donna Bergmark (Linking API), Tim Brody (Citebase), Christopher Gutteridge (EPrints), Mike Jewell (Paracite), Zhuoan Jiao, Simon Kampa (e-Services), Arouna Woukeu (evaluation)