
SUNCAT as a national serials' facility for researchers and librarians



Fred Guy
Project Manager¹, Edinburgh University Data Library, George Square, Edinburgh EH8 9LJ
Tel: 0131 651 3875
E-mail: f.guy@ed.ac.uk



Peter Burnhill
Co-Director and Director of EDINA, Edinburgh University Data Library, George Square, Edinburgh EH8 9LJ
E-mail: peter.burnhill@ed.ac.uk

INTRODUCTION

Journals and other periodicals continue to play an extremely important role in scholarly communication. They have been the favoured means of reporting research work outcomes in all disciplines for a very long period and continue to be so even in an age when technology has revolutionised the means of production. Technology has helped streamline the process of publication but electronic journals manifest the same characteristics as paper journals in publishing collections of papers, usually in a particular subject area for scholarly record, at defined time periods. At the same time, technology has also encouraged the growth of periodical publishing, both formal and informal. Search engines, especially Google, in all its new forms, also play a part in encouraging periodical usage by retrieving periodical titles which students expect libraries to hold.

The need for scholars, researchers and students to consult periodicals, in paper form or electronic, is as great as it ever was. Indeed, in some disciplines, such as biological sciences and medicine, it is even greater, with the requirement to be working with the most up-to-date knowledge and best practice. Libraries, though, have been challenged by the increasing numbers of periodicals being pub-

lished and the increased costs of purchasing them. Serials have increased in price much more than monographs with the result that libraries have been forced to reduce the number of subscriptions they take out.

What if a particular periodical is not available for reading in an institution to which a researcher belongs? This ought not to be a major problem if it is available in another institution. How, though, does the user find the location(s) and means of access to the journal whether on-shelf or on-line? Clearly it would be possible to search the OPACs of other institutions but that could be very time consuming and would result in hit and miss results due largely to inconsistency of search and catalogue display styles across different library management systems (LMSs).

The answer put forward is a national, UK Serials **UN**ion **CAT**alogue (SUNCAT). Managed by EDINA at the University of Edinburgh, the national serials union catalogue is providing access via a comprehensive series of search options within an easy to use interface to the serial holdings of major UK research libraries. By searching a single title, users can readily find which institutions hold that title, what holdings they have and whether or not an online version is available, and under what privilege. At the time of writing, SUNCAT has over 4 million bibliographical records and serials holdings for 22 major UK research libraries².

SUNCAT has also been designed to be a source of high quality bibliographical records for serials which can be downloaded to library OPACs. To this end the Co-operative Online Serials (CONSER) database comprising around 1 million bibliographical records and the ISSN Register also comprising 1 million records have been loaded and will be available for download to those libraries which are participating in SUNCAT.

The design of SUNCAT is discussed by Burnhill, Halliday, Rozenfeld and Kidd³. The role of SUNCAT in the evolving information environment is explored by Burnhill and Law.⁴

THE ORIGINS OF SUNCAT

SUNCAT has arisen out of a series of preparatory projects carried out on behalf of the Joint Information Systems Committee (JISC) and Research Support Libraries Programme (RSLP) in conjunction with the British Library. In 2000/2002 the bodies funded a feasibility study for a national union catalogue (UKNUC). The key findings, as

far as the serials element of a union catalogue was concerned, were that the catalogue should:

- include the records and holdings from the larger research and university libraries, including the British Library, the national libraries of Scotland and Wales, as well as smaller, specialised libraries
- be a centralised catalogue of high quality bibliographic records
- use customised software to import serials records from local catalogues, match them with existing high quality records from external sources (e.g. CONSER), upgrade records where appropriate and upload them into SUNCAT with standardised holdings information
- be able, where appropriate, to copy back upgraded records to local catalogues
- be so designed to facilitate linkages and enable the development of added-value services

Consultation on the findings was carried out with the user community and the outcome was strong support for a serials' union catalogue but not a monographs' union catalogue. The funders then agreed to fund a scoping study and specification of a serials union catalogue. From this an invitation to tender was issued and, in February 2003, following a successful bid, the University of Edinburgh and its partner, Ex Libris, started work on the SUNCAT project. EDINA has provided the project management and much of the staff effort; Ex Libris has supplied and installed the Aleph 500 library management system on hardware at EDINA to act as the database for SUNCAT.

How does SUNCAT work?

Overall model

The Aleph software allows for the creation of 4 types of union catalogue. These are:

- 1 Central catalogue
- 2 Central catalogue and local catalogue
- 3 Union view catalogue
- 4 Union catalogue

1 Central catalogue

The characteristics of this model are that there is a single bibliographic record for each title. The central catalogue is shared by multiple administrative libraries with each one controlling its own acquisitions, circulation and serials.

2 Central catalogue and local catalogues

Whilst having a single bibliographic record (as in the above) participating libraries copy records for material held by them from the central database to local catalogues. If the central record is updated all libraries with copies of the central record are notified.

3 Union view catalogue

There may be multiple bibliographic item records for a single title, with no central cataloguing authority having the right to control the bibliographic record. There is no deduplication of records for the same title in the catalogue. Rather, deduplication is carried out for the purposes of viewing; that is, the user interface supports title-level searching.

4 Union catalogue

A single unified database is compiled from the records sent by participating libraries. As with the *Union view catalogue* records are not physically deduplicated and deduplication is done for view.

The SUNCAT system follows the Model 3, the Union view catalogue.

APPLICATION OF THE MODEL

All incoming records are converted into MARC21 format, by no means a trivial task now largely automated. Those received from contributing libraries are at the 'item-level'. The approach taken in support of Model 3 is to form sets of matched records that are equivalent with respect to key attributes. The ISSN, LCCN, BNB and Short Title play a role in the initial stage, with subsequent use of weights on the Date of Publication, Place of Publication, Main Entry and Added Entry.

It is often not possible to distinguish between print and electronic format from the data obtained. Matching is therefore carried out above the level of the 'manifestation', at the 'title' level. All equivalent records within a set are assigned the same SUNCAT ID. When a search is performed all records sharing the same SUNCAT ID are displayed as though one record. The administrative records for each separate record are linked making it possible, thereby, to view a complete integrated list of all holdings of all libraries for that title.

POPULATING SUNCAT

The priority set for the project was to build a critical mass of serial titles from the outset. This was done for the holdings of the twenty-two

largest research libraries in the UK, in three waves. The records from the associate partner libraries (Cambridge, Oxford, Edinburgh and the National Library of Scotland) provided the testbed in the first wave.

OVERVIEW OF PHASES

SUNCAT has evolved and continues to evolve in a series of phases related to funding periods but with particular series of activities happening. These are:

PHASE 1: This ran from February 2003 to December 2004. The key areas of activity were:

Organisational issues

- Development of a governance structure

Liaison activities

- Contact with candidate libraries
- Use of focus groups to assist in development of the user interface
- End user testing of the interface
- Production of training and other materials
- Publicity and dissemination of information about SUNCAT

Development of processes and procedures

- Software installation (initially Aleph v.15)
- Establishment of processes
- Data conversion
- Data loading

Quality assurance and evaluation

- Quality assurance by EDINA staff
- Evaluation of the SUNCAT Serials Union Catalogue by Centre for Research in Library and Information Management (February 2005)

Investigations

- Survey of e-journal management in Phase 1 libraries
- Investigation of emerging standards and initiatives with reference to electronic resources

PHASE 2: (January 2005 – December 2006). This is the current phase. The key areas of activity are:

Organisational issues

- Review of governance structure

Liaison and promotional activities

- Contact with all participating libraries
- Reach agreement about downloading of records with third party organisations
- Extension of promotional activities

Implementation of processes and procedures

- Data conversion of new libraries
- Implementation of streamlining process
- Data updating of participant libraries

Hardware upgrading

- Installation of new processor
- Installation of new disks

Development

- Implementation of Librarian's Interface to assist with deduplication and allow controlled record downloading

Preparations to become a service

- Develop service level definitions (SLDs) for testing
- Test SLDs
- Consideration of funding options for Phase 3.

Investigations

- Implement a capacity to accept ONIX for Serials messages.

PHASE 3: (January 2007 -) The key aspect of this will be the running of a sustainable service.

PHASE 1 (FEBRUARY 2003 – DECEMBER 2004)

Organisational issues

The SUNCAT Project was overseen by the SUNCAT steering committee, which had come into being as part of the Research Libraries Support Programme. The committee, chaired by Professor Derek Law (University of Strathclyde), comprised representation from libraries and staff from the JISC Executive. EDINA staff attended as invited.

The SUNCAT steering committee also established a sub-committee, chaired by Peter Burnett, University of Oxford. This sub-committee played a key role in advising the project team on appropriate libraries and groups of libraries who should be invited to participate in SUNCAT.

The SUNCAT Project team comprised staff drawn from EDINA and the University Library responsible for direction and management, bibliographical staff, systems staff and staff responsible for user testing and dissemination. Ex Libris designated expert staff based in Israel.

From the outset, EDINA established the Bibliographic Quality Advisory Group (BQAG) comprising representatives from the Associate

Partners, library experts well qualified to advise and comment on technical matters related to data quality⁵. The BQAG played a key role in guiding the project team and providing support.

Liaison activities

Working with contributing libraries was critical to the project's success. That there was ready acceptance to participate, thereby involving a degree of local effort, bears witness to the recognition by librarians of the importance of creating such a central resource. Information about the contributing library, its data and the library management was captured in two data questionnaires completed by the libraries. Special attention was given to ensuring that libraries were well informed about what the likely commitments would be and what the likely benefits would be, through the document, *Contributing libraries: benefits and expectations*, routinely issued to all potential contributors.

There was also special focus on developing an attractive and highly usable user interface, drawing upon experience within EDINA for its other bibliographic services. A focus group of nine people, comprising cataloguers, serials acquisitions staff, subject librarians, reference enquiry staff and interlibrary loans staff was established and aspects considered were the layout of each screen, the wording of the on-screen text and the labeling of functions.

User testing was carried out in the period August to October 2004. This involved end users from a number of disciplines carrying out a set of defined tasks as well as a range of volunteers from Phase 1 contributing libraries, BQAG members, SUNCAT Steering committee members and members of the sub-committee. The outcomes from all the testing were collated and led to changes being made to the interface.

Effort also went into the production of instructional materials, some freely available from the project web site and some restricted to contributing libraries. The key dissemination activity was the SUNCAT web site, launched in July 2003. Information about the project was provided at a number of professional conferences including those of UK Serials Group and CILIP's annual conference.

Development of processes and procedures

The ingest of data from libraries having different library management systems represented a large challenge. The approach adopted involved initially working with the partner libraries (Universi-

ties of Oxford, Cambridge, Glasgow, Edinburgh and National Library of Scotland). These libraries are all characterised by having large numbers of serials records, as well as considerable experience and expertise in the operation of library management systems. The task proved extremely time consuming and labour intensive. From the initial work with the partner libraries, however, there emerged a series of processes and procedures which were adopted with the later libraries. These were:

- a) Libraries were approached by SUNCAT staff about becoming a contributing library (CL). (The steering committee advised SUNCAT on potential libraries). A document entitled *Contributing libraries: benefits and expectations* was sent to all potential contributors. This document outlines the process which is followed as well as listing what CLs will be expected to do.
- b) If the library agreed to become a CL, staff were asked to complete a *preliminary questionnaire*. This included questions about the specific Library Management System used, number of serial records held, format, number of records added or amended each month and the way in which e-journals are managed.
- c) A submission date was agreed and by that date the CL would have sent a file of all serial records together with a completed *data questionnaire*. This questionnaire included detailed questions particularly on the way holdings and location information has been handled by the CL.
- d) To ensure that the data is normalised for coherent display in SUNCAT, manipulation was required. A *data specification* showing the proposed mappings was prepared by the SUNCAT Team and submitted to CLs.
- e) Once the *data specification* had been agreed by the CL and SUNCAT, the data was converted according to the specification and loaded to the pre-production database.
- f) Further checks were carried out and once completed the data was copied to the production database.

Quality assurance and evaluation

Quality assurance was an important element within the project. This was supervised by a

member of staff of EDINA and involved external (to the project) assessment of each of the work packages. The main areas reviewed were:

- Project documentation
- User guides
- User testing
- Interface accessibility

The commitment to internal QA assisted the external evaluation of the SUNCAT Serials Union Catalogue carried out by the Centre for Research in Library and Information Management (CERLIM) in the latter part of 2004. The report submitted to the SUNCAT steering committee in February 2005 concluded that SUNCAT was a successful project and should continue to be developed; it also stated that there needed to be better communication with the wider community about the aims and intentions regarding SUNCAT. The report contained recommendations about multiple entries of records and a number of matters relating to the web site and its adherence to W3C Web Accessibility Guidelines. It was also suggested that more Help information was required. The steering committee approved the recommendations at their October 2005 meeting; Phase 2 has involved implementing all of the recommendations.

Investigations

Three separate 'R&D' work packages were also carried out. The first, a *Survey of electronic-journals holdings for the UK National Serials Union Catalogue (SUNCAT)*, was a survey of practices and procedures adopted by the 22 libraries in Phase 1⁷. It revealed that the libraries surveyed had adopted a myriad of practices and procedures for the management of e-journals. For example some libraries included e-journals in their main LMS whilst others had completely separate systems. Sometimes libraries had single records for print and electronic journals and sometimes they had separate records. There was evidence of moves towards the use of OpenURL resolvers and proprietary management systems such as Serials Solutions and TNet but by no means did all libraries have definite plans to implement such systems.

The second work package was entitled *Study of NISO/EDItEUR developments*⁸. The emerging ONIX for Serials message formats for the transmission of information across the journal supply chain (publishers/aggregators/agents/libraries) were examined in the context of potential use by SUNCAT to keep holdings data up to date.

The third work package entitled *Requirements analysis for SUNCAT* considered four categories of issues concerned with electronic journals in the SUNCAT context⁹. These categories were: data issues; licensing (access) issues; linking issues and third party services issues. The first three categories of issues were considered in the context of existing systems and services: ONIX for Serials; Electronic Management initiative of the Digital Library Federation; OpenURL Router; GetCopy; ZETOC; Directory of Open Access Journals and Scoping Study into Institutional Profiling and Terms and Conditions Services.

PHASE 2 (JANUARY 2005 – DECEMBER 2006)

Organisational issues

Accepting that the governance structure created for Phase 1 reflected SUNCAT's status as a project, the Steering Committee has been considering how to modify itself to be appropriate to SUNCAT, initially as a pilot service and then as a full service. No definite decisions have been reached as yet and for the time being the Phase 1 governance structure continues.

Liaison and promotional activities

Moving to a service requires focus on promotion and dissemination. Much of this work will be carried out as part of EDINA's user support activity, in association with the JISC Communications and Marketing Team. A detailed plan is to be presented at the next SUNCAT steering committee meeting in January 2006.

Liaison is also required within the project, both to increase the number of contributing libraries and to extend the functionality of SUNCAT to allow the download of records into OPACs.

Considerable effort has been made to increase the number of participating libraries. Thirty seven additional libraries have agreed to participate and, as well as more academic libraries, this figure includes a number of libraries with specialist and unique collections.

The facility to download high quality bibliographic records is requiring attention to three issues. One is to do with workflow at contributing libraries; the second is technical; the third relates to licence. The CONSER file contains records which can be made freely downloadable for libraries. Records in the ISSN Register, on the other hand, required that a licence has been acquired - the facility dependent upon the development of an authentication facility (see below).

There are, of course, other records in the SUNCAT database which have been supplied to libraries by third parties. Discussions are currently being held with the third parties to reach agreement on downloading of those records.

Implementation of processes and procedures

The SUNCAT database was completely rebuilt in December 2004 as part of the Ex Libris standard practice for moving into service. This allowed data currency to be improved and revisions to the matching and de-duplication processes to be introduced.

There are major challenges in keeping an union catalogue up to date. One reason for this is the considerable variance in the form in which updates are presented and in some cases it has not been possible for a library to create a file of updates. There have also been some issues associated with the way updates are handled by the software. This has involved detailed discussions both with contributing libraries and with Ex Libris. It is expected that the main issues will be resolved to allow updating to be completed, in large part, by early 2006.

Processing holdings data for each library was found to be the most time consuming activity during Phase 1. Priority was therefore given to developing a set of procedures to streamline the process. The approach adopted was to mark holdings by location only and omit summary holdings statements. This reduces, quite dramatically, the time required to add new libraries. To date, two libraries have been loaded using this process.

Hardware upgrade

Phase 1 of the Project operated on older hardware, deploying equipment owned by the University of Edinburgh as part of a quick start for the project. A dedicated server was specified once more was known about load and performance, and this, together with a new array of disks, was installed during the summer of 2005, shortly to provide the platform for the SUNCAT service in the Autumn.

Development

The system was set up using Aleph 500 Release 15 in the Union view along the lines of the implementation in Melvyl, union catalogue of the California Digital Library.¹⁰ It was quickly recognised that the system had some limitations at least partly arising from the more variable quality of data for the UK. The matching algorithm, for example, which seemed to work well in the Californian context where incoming data was richer

and more consistent, did not work so well in the UK context. Aleph Release 16 provided additional functionality to improve matching and this was implemented in late 2004.

There are three aspects to the planned developments in Phase 2. They are:

- a) Development of a Librarian's Interface
 - b) Implementation of a title level identifier (SUNCAT ID)
 - c) Z39.50 improvements
-
- a) A number of components were identified for the Librarian's Interface. They include a facility to assist manual matching of a contributing library's records as well as controlled downloading of records from SUNCAT. Two other components, notification to changes to the preferred record and manual registration of holdings are currently being considered by EDINA and Ex Libris.
 - b) The SUNCAT ID has been developed given recognition that many titles in SUNCAT do not have associated ISSNs and that a unique number (at title level) needs to be in existence. The SUNCAT ID will be assigned to all record sets.
 - c) Testing carried out by the M25 Systems Team in 2003 on Library Management Systems targets, including Ex Libris targets, revealed some issues over implementation of the standard. Additionally some developments specifically concerned with SUNCAT requirements were identified.

Preparations to become a service

An important part in operating a service is the definition of service levels and performance indicators for all aspects of the service including operation, updating, supporting users and supply of documentation in an agreed Service Level Definition document. In the pilot service phase the opportunity is taken to implement and report on the measures to assess their validity. A draft Service Level Definition document is currently in the process of consideration with the JISC Services Team.

Consideration of funding options

Work has commenced to consider what funding options might be possible for full service running from January 2007. There is strong commitment to a service that is free at the point of use, and there are doubts that universities can be persuaded to subscribe for such a national facility. Nevertheless, an agreed revenue model has to be established

that will provide maximal benefit as a productivity tool for researchers and librarians.

Investigations

The need for 'R&D' continues, especially for the development of functionality to meet the demands of electronic journals and allied licence information. On behalf of SUNCAT, EDINA was successful in bidding for funding under the JISC Publisher Metadata & Interoperability Projects II. The project entitled *Automating Ingest of Metadata on Serials Subscriptions* (AIMSS), and which will run from October 2005 – May 2006, will take forward proposals explored in Work Package 2 of Phase 1¹¹. The project partner is Serials Solutions¹². EDINA will develop a capacity to accept real-world data from Serials Solutions, process it and update the SUNCAT test database accordingly. Data will be transmitted using ONIX for Serials (Serials Online Holdings) message format.¹³ This outcome of this project should have value more generally for research libraries.

PHASE 3 (AUGUST 2006 – DECEMBER 2006) AND PHASE 3 (JANUARY 2007 ON)

SUNCAT is scheduled to operate as a service from August 2006 onwards. However, further development of SUNCAT is envisaged. Most obvious is the extension in the number of contributing research libraries, obliging firmer definition of what is a UK research library and what is not. There is also the matter of electronic subscription information, and the integration of SUNCAT into the larger UK digital library environment, linking to article supply. There is much to be defined as part of the agenda for Phase 3.

There is already background work reflecting upon the appropriate architecture for a modern serials union catalogue, given the successful set-up of SUNCAT, and as interoperability standards and practices mature.

CONCLUSION

The SUNCAT project has been, and is, both ambitious and complex, reflecting the importance of journals and other serials for scholarly communication; there is considerable challenge in understanding and meeting expectations in the digital era within the context of the legacy position inherited by the UK serials community, and the commitment to access to the scholarly record, in both print and electronic format.

There is still work to be done to create a top quality invaluable resource for researchers and librarians. From a standing start, though, much

has been achieved and a solid platform is in place to allow SUNCAT to play its part in the evolving information environment by serving the needs of researchers and librarians.

ACKNOWLEDGEMENTS

We wish to acknowledge the efforts of all members of the SUNCAT Team (past and present). They are:

Direction and management

Christine Rees (January 2005 –) co-director
Leah Halliday (February 2003 – February 2005)
project manager

Systems

Morag Macgregor (February 2003 -) systems
librarian

Bibliographic

Nathalie Schulz
(March 2003 – March 2005) project officer
Natasha Aburrow-Jones
(October 2003 – April 2004) senior library assistant
(May 2004 -) project officer
Maira Whitson
(July 2004 -) senior library assistant
Nicola Osborne
(August 2005 -) senior library assistant

User requirements

Liz Stevenson (February 2003 -) (electronic
resources librarian Edinburgh University Library)
Zena Mulligan (August 2003 -) project officer

Quality assurance

Peigi MacKillop (EDINA)

Special advisers

Tony Kidd (University of Glasgow)
Slawek Rozenfeld (Paris)

- 1 Project Manager since June 2005
- 2 <http://www.suncaat.ac.uk/about/contributing.html>
- 3 Peter Burnhill and Leah Halliday 'SUNCAT: a modern serials union catalogue for the UK', *Serials* Vol. 17 No. 1 March 2004 pp. 61-67 [http://uksg.metapress.com/\(yci2cmvu3paqjz55i2wnwrvy\)/app/home/contribution.asp?referrer=parent&backto=issue,15,21;journal,6,18;linkingpublicationresults,1:107730,1](http://uksg.metapress.com/(yci2cmvu3paqjz55i2wnwrvy)/app/home/contribution.asp?referrer=parent&backto=issue,15,21;journal,6,18;linkingpublicationresults,1:107730,1)
- 4 Peter Burnhill and Derek Law, 'SUNCAT Rising: UK Serials Union Catalogue to assist document access'. [To be published in *International review of library interlending*].

- 5 Members of the BQAG in Phase 1 were:
Alason Roberts (Chair and University of
Edinburgh); Hugh Taylor (University of
Cambridge); Hugh Croll (University of Edin-
burgh); John Nicklen (National Library of
Scotland); Susan Miles (University of Oxford);
Tony Kidd (University of Glasgow); Slawek
Rozenfeld (special adviser)
- 6 <http://www.suncat.ac.uk>
- 7 This work was carried out by the M25
systems team. The project officer was Fraser
Nicolaidis
- 8 This work package was carried out by Fred
Guy, as consultant
- 9 This work package was carried out by Fred
Guy, as consultant
- 10 [http://www.cdlib.org/news/background.
html](http://www.cdlib.org/news/background.html)
- 11 [http://www.jisc.ac.uk/index.
cfm?name=programme_pals2](http://www.jisc.ac.uk/index.cfm?name=programme_pals2)
- 12 <http://serialssolutions.com/home.asp>
- 13 [http://www.editeur.org/onixserials/ONIX_
SOH1.0.html](http://www.editeur.org/onixserials/ONIX_SOH1.0.html)