



Welcome to the fifth SUNCAT newsletter. The newsletter appears twice yearly, to provide a progress report on SUNCAT.

New Libraries on SUNCAT

.....

SUNCAT now contains 66 libraries, including three new libraries loaded since the last newsletter: the Royal College of Music, the University of Dundee and the University of St. Andrews. The library of the National Institute of Medical Research will also be loaded early in the new year.

In our last newsletter we reported that a decision had been taken early in 2007 to temporarily freeze the addition of new libraries to SUNCAT. The main reason for this decision was to allow the SUNCAT team to focus on getting the existing libraries' data up-to-date and also to address the increasing amount of duplication on the catalogue. Please see the sections on deduplication and updating below for further details.

Updating SUNCAT

.....

The primary focus of the SUNCAT bibliographic team during the past year has been the currency of the database. This means ensuring regular and frequent receipt of update files from all our Contributing Libraries. Phase One libraries were asked for a full file of all their records, as their data was the most out-of-date. Data specifications had to be re-written in light of the new data received, and update specifications written for handling different ways of marking a deleted record. For those libraries unable to supply update files in the form of new, changed and deleted records, a full file update workflow has been designed, whereby the library supplies all the records in their catalogue with every update. Around half of the SUNCAT Contributing Libraries are now sending regular updates, which are being processed and loaded in a matter of days. Updates are sent at varying intervals as agreed with by the Contributing Library. Some updates are received weekly; some, semi-annually. The majority are sent either at monthly or quarterly intervals. If a library sends updates more regularly than monthly, we will endeavour to process these as time permits, but will ensure that they are concatenated and loaded on a monthly basis at the very least.

About half of the remaining Contributing Libraries are in negotiations for sending updates; this leaves a few libraries which have not yet started the update process in SUNCAT. This is normally due to time constraints and workflow conflicts on the part of the library; they will be updating at some stage. Methods of informing users as to the currency

of data in SUNCAT by library are being explored, from RSS feeds with the details of the latest updates being loaded, to easily updateable web-pages with the date of the last load for every library.

SUNCAT has acquired a new full file from the ISSN Centre in Paris. This contains all records with an ISSN – in excess of one million. When SUNCAT first loaded the ISSN file, the data was in ISSNMARC. Now, the ISSN supplies files in MARC21, necessitating a complete re-write of the data specification, with intensive study of the data. This is being converted at present; it is hoped to have the new ISSN file loaded before Christmas. The ISSN file is not visible from the usual searches in SUNCAT, due to the nature of the agreement that SUNCAT has with the ISSN Centre, but the ISSN records will be available to the Contributing Libraries for download.

Pilot MARC Record Download Service

.....

We have recently released a pilot downloading facility for our Contributing Libraries. The pilot service enables Contributing Libraries to download a subset of SUNCAT records in MARC format via Z39.50. The records are from CONSER, the Library of Congress Cooperative ONLINE SERIALS programme, <http://www.loc.gov/acq/conser>. There are over one million CONSER records available and these are of very high bibliographic quality and are particularly strong in their coverage of current Scientific, Technical and Medical (STM) titles. These records are free of charge to our Contributing Libraries enabling them to download records for newly purchased serials but also for those wanting to upgrade their existing serials records.

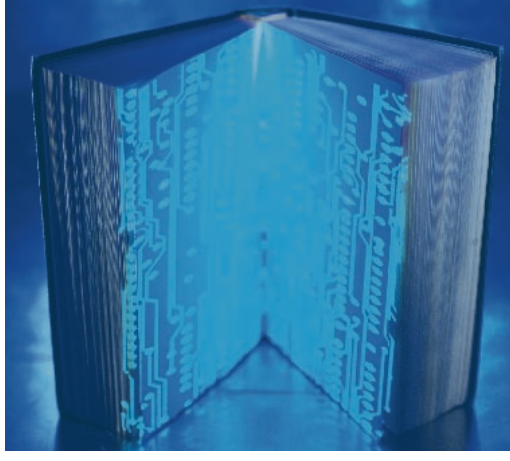
We are currently unable to offer the entire SUNCAT database of records in MARC format due to IPR restrictions imposed by a commercial supplier on records supplied to our Contributing Libraries. However, we are investigating methods of filtering out these records so that we can offer a greater proportion of our Contributing Libraries data for download and have plans to release ISSN records in the near future. Additional plans include developing the download facility from the Web interface.

Details of how to set up and configure the Z39.50 connection are available at <http://www.suncat.ac.uk/support/z-target-auth.shtml>

As this is currently operating as a pilot service, we would welcome your comments on the functionality and would like to work with our Contributing Libraries to resolve any problems they experience with the process. Please contact



basic search screen



the EDINA helpdesk at edina@ed.ac.uk with any comments or queries.

Web 2.0 Developments

The team are currently working on a number of Web 2.0 developments for SUNCAT. The first will enable users to search SUNCAT from within their Facebook profile by providing a simple search box which can be easily added to their profile.

The second will be a SUNCAT Blog which will keep you up to date with new libraries added to SUNCAT, updates to libraries on SUNCAT, new developments and items of news as they occur. It will also be possible to receive this information as an RSS feed.

Deduplication

Since reporting in the last newsletter that we are exploring ways to reduce the effects of duplication on SUNCAT, we have conducted an online survey with our Contributing Libraries to assess their experience of searching SUNCAT and their tolerance of duplicate records.

Participants were asked to search for three different journals using a title or ISSN search, to note the number of duplicates and to record their tolerance of the number of duplicates.

The survey highlighted the problems of poor data quality in a large number of the records on SUNCAT, the main reason for the high level of duplication. However, it also highlighted the difficulty of distinguishing between actual duplicate records for the same title and records which, although they look like duplicates at first glance, are not, in fact, duplicates but describe similar but not identical titles.

The results showed that duplicate records were displayed in the results' sets of 92% of the searches conducted.

It also revealed that in over one quarter, 28%, of the searches carried out, respondents found the number of duplicate records displayed unacceptable, with 38% finding the number acceptable and the remaining 36% found the number workable but excessive.

The results suggested that:

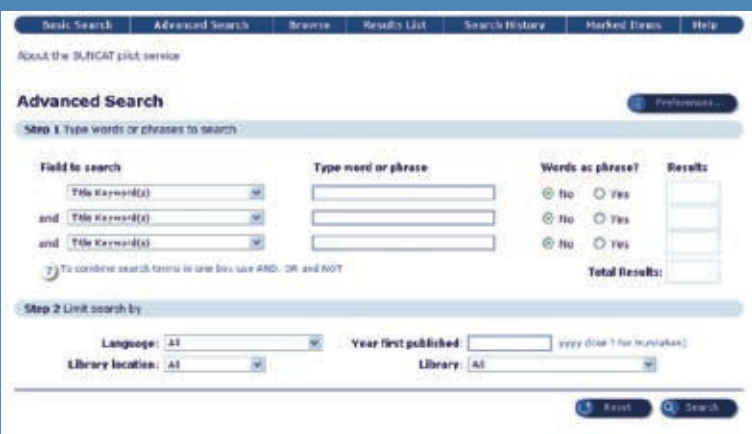
- It would be desirable to reduce duplication to a maximum of three records per title.
- A maximum of six or seven records would still be considered workable.

In addition to reducing the number of duplicate records, the results suggested a number of modifications which should be investigated as means of improving users' experience of searching SUNCAT and dealing with the display of duplicate records. These include testing new Title and ISSN indexes which would return smaller results sets by only covering the core title and ISSN fields and not, for example, the linking fields to previous and next titles. Secondly, presenting more information about titles on the Results Display screen to help users differentiate between similar titles and decide whether to enter the full record display.

We will be testing the new indexes over the next few months and if the results are favourable we will release new search options to the interface next year.

We have also asked our project partners, Ex Libris, to develop functionality to enable libraries holding a title to be displayed on the initial Results Display screen. Users will then be able to tell at a glance if a title is held in a location or locations of interest thus reducing the number of records they need to enter.

Finally the team are also investigating a secondary matching algorithm to supplement the existing algorithm which struggles to make appropriate matches when faced with records which contain little information bar the title. The original algorithm has already been tweaked to try and accommodate this; however, this has not proved sufficient to reduce the amount of duplication significantly. Therefore, SUNCAT has been in discussion with the School of Informatics within the University of Edinburgh (see: <http://www.inf.ed.ac.uk> for further information on the department), and one of their staff members will be working with the SUNCAT team to create an additional algorithm to help improve the matching. Work on this is due to start in February. In the meantime, the Bibliographic team are gathering examples for this new algorithm to "learn" from. This should prove to be a useful and productive exercise, and it is hoped that the initial 3-month long proof-of-concept project will reap its rewards and improve matching in the database.



advanced search screen



Map of Contributing Libraries

We have added a map of our Contributing Libraries to the SUNCAT website at http://www.sunecat.ac.uk/description/suncatmap/Suncatsubs_final.htm



The map provides details of how to connect to, contact, access and find libraries, linking to directions on Google Map.

Please take the time to check your institution's entry and contact the Edina helpdesk edina@ed.ac.uk if any amendments are required.

Bibliographic Quality Advisory Group Report

The SUNCAT Bibliographic Quality Advisory Group met in the EDINA offices on the 11th October, 2007. This group meets on a semi-annual basis, and is made up of a representative of serials cataloguing from some of our Contributing Libraries. Their collective wisdom is very beneficial to SUNCAT. Topics covered included z39.50 and the downloading of records (more elsewhere in the newsletter on this topic), duplication in the database, and how SUNCAT is trying to solve this, updating progress and associated issues. Peter Burnhill, SUNCAT's co-director, and the director of EDINA, led a discussion on data quality metrics. As usual, the discussions throughout the day were most useful. The next meeting will be held sometime in the late spring.

Developing the UK infrastructure for discovery to delivery service

This upcoming JISC funded project is designed to create more explicit links and machine-to-machine interfaces between JISC funded resource discovery services to help provide seamless access and delivery. It is anticipated

that the Project will commence in December 2007 and be completed by March 2009.

A November 2006 RIN report (<http://www.rin.ac.uk/researchers-discovery-services>) highlighted the difficulties which researchers face in moving between discovery and delivery and in accessing the full versions of all the content whose existence they have discovered.

A subsequent RIN seminar (<http://www.rin.ac.uk/union-catalogues>) attended by a wide range of stakeholders, brought out the need for:

- moving from discovery to delivery;
- building on Copac and SUNCAT;
- sustaining and developing UK-focused services.

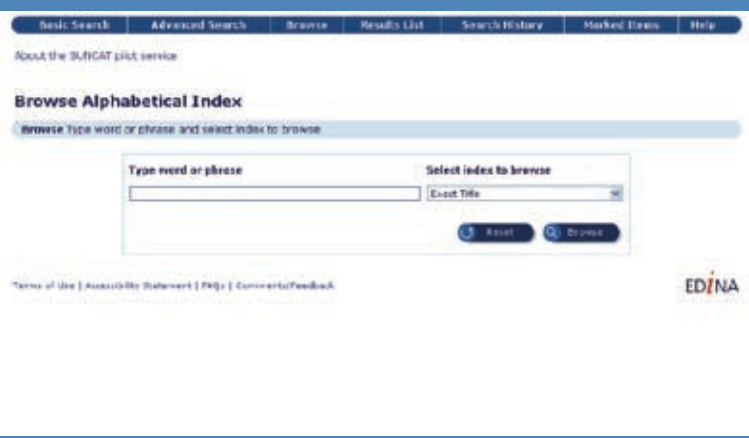
Project scope

The key aim of project will be to deliver the most effective methodology for:

- i) integrating, from a user perspective, the functionality which is currently separately delivered by Copac, SUNCAT, Zetoc, Intute (in particular the Intute repository search) and, where identified, other relevant British Library services and the UK Research Reserve
- ii) enhancing these services to move from discovery to delivery, seamlessly, as far as this is possible within the context of local environments e.g. Library Portals or VLEs

Issues to be considered include:

- moving from discovery to delivery (using OpenURL and similar technologies for this)
- the possibility of improving the quantity of TOC and other data currently available, by incorporating the services of other aggregators and publishers
- harvesting metadata and the attached content from institutional repositories of all kinds, across sectors (libraries, archives, etc)
- making appropriate connectivity with other JISC-funded resource discovery initiatives
- possibilities for presenting different views (e.g. regional and local) from national aggregations, and how to deal with the lack of a unified interface.
- the integration of authentication services, and a consistent approach to the presentation to end users of licensing restrictions



browse alphabetical index screen



- linkages with interlending functionality for material which only exists in a physical format, and with local circulation control systems
- the incorporation of web 2.0 functionality to allow users to personalise and share data, and to encourage innovative uses of content, including the possible development of subject based wikis
- enhancement of current awareness functionality (exploit technologies such as RSS for this)
- the need for simple and intuitive interfaces
- development of collaborative collection management strategies (e.g. the UKRR).

An initial meeting of all involved parties, including EDINA, is due to be held on 11th December 2007.

Promotional Material

EDINA has produced flyers and posters for SUNCAT. These are available free of charge and can be used to promote the service to your local user community. If you would like copies of these for your institution please send your contact details, including name, email, telephone and postal address to edina@ed.ac.uk

There are also plans to create online and print user scenarios to promote SUNCAT to both end-users and Inter-Library-Loan staff. More information about these will be made available shortly.

Meetings and Events

- July 2007 - Meeting with Janifer Gatenby of OCLC
- October 2007 Meeting with British Library at Boston Spa
- October 2007 SUNCAT Bibliographic Quality Advisory Group Meeting
- December 2007 Online Information -- Stand
- December 2007 Discovery to Delivery Project meeting, London
- December 2007 Personalisation for EDINA & MIMAS meeting, Manchester

SUNCAT Team



Back row from left: Christine Rees (Co-director of SUNCAT), Moira Whitson (Assistant Project Officer - Bibliographic), Zena Mulligan (Project Officer – User Requirements), Fred Guy (Project Manager).

Front row from left: Natasha Aburrow-Jones (Project Officer - Bibliographic), Nicola Osborne (Assistant Project Officer - Bibliographic).

Not included in the picture are Peter Burnhill the second Co-director of SUNCAT and Director of EDINA, Morag Macgregor (Systems Librarian) and Liz Stevenson, University of Edinburgh Electronic Resources Librarian (p/t Advisor to SUNCAT).

Season's Greetings

The SUNCAT team would like to wish all our users and Contributing Libraries



Image from morguefile.com

****A Happy Festive Season and All the Best for 2008****