Using a dumb identifier to do smart things

Peter Burnhill
Director, EDINA, University of Edinburgh, UK
Overview for talk

0. Introduction
   – EDINA, JISC national datacentre at University of Edinburgh

1. Opportunity to share a digital library journey
   – and say hello to old friends

2. SUNCAT: union catalogue of serials in UK (& OpenURL Router)

3. Keepers Registry: e-journal preservation (& ISSN-L)
   – Digitised journals & print archiving

4. Entitlement Registry & KnowledgeBase (KB+)

5. Forward Look
research, learning & teaching in UK universities & colleges

Formed in 1995/96, based upon Edinburgh University Data Library (1983/4)

Digital Content & Metadata

National Data Centres

EDINA

MIMAS

JISC Collections

JISC Sub-Committees

UK funding councils for HE & FE
Overview

1. Internet and emergence of the Web 1994 -
   - Digital access to metadata about Serial Content
     *Content was mostly PRINT ; becoming digital

2. SUNCAT & OpenURL Router, 2003 -
   - for Serial Content that is available as Print & Digital

   - focus on DIGITAL, Digitised (and maybe print)

4. Entitlement Registry & KnowledgeBasePlus, 2011/12
   - for Serial Content that is DIGITAL (and subscribed)

5. Forward Look, with 2020 Vision!!
   - General scheme for DIGITAL & PRINT(?) Content
Spotlight upon essential role for the ISSN

and upon useful changes made by ISSN Network:

1. Invention and use of the ISSN-L linking field

2. ISSNs for digitised journal content (d-journals)

3. Growth in ISSNs assigned to e-serials
   * circa 100,000 ISSNs for electronic ‘continuing resources’
<table>
<thead>
<tr>
<th>COUNTRIES</th>
<th>Online code (cr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL</td>
<td>97581</td>
</tr>
<tr>
<td>UNITED STATES</td>
<td>20208</td>
</tr>
<tr>
<td>UNITED KINGDOM</td>
<td>9672</td>
</tr>
<tr>
<td>CANADA</td>
<td>5911</td>
</tr>
<tr>
<td>NETHERLANDS</td>
<td>4584</td>
</tr>
<tr>
<td>GERMANY</td>
<td>4578</td>
</tr>
<tr>
<td>BRAZIL</td>
<td>4471</td>
</tr>
<tr>
<td>SPAIN</td>
<td>3884</td>
</tr>
<tr>
<td>INTERNATIONAL</td>
<td>3515</td>
</tr>
<tr>
<td>FRANCE</td>
<td>3134</td>
</tr>
<tr>
<td>AUSTRALIA</td>
<td>3017</td>
</tr>
<tr>
<td>DENMARK</td>
<td>2570</td>
</tr>
<tr>
<td>ITALY</td>
<td>2368</td>
</tr>
<tr>
<td>FINLAND</td>
<td>2133</td>
</tr>
<tr>
<td>NEW ZEALAND</td>
<td>1918</td>
</tr>
<tr>
<td>INDIA</td>
<td>1789</td>
</tr>
<tr>
<td>CZECH REPUBLIC</td>
<td>1683</td>
</tr>
<tr>
<td>POLAND</td>
<td>1527</td>
</tr>
<tr>
<td>NORWAY</td>
<td>1490</td>
</tr>
<tr>
<td>JAPAN</td>
<td>1395</td>
</tr>
<tr>
<td>ROMANIA</td>
<td>1360</td>
</tr>
<tr>
<td>HUNGARY</td>
<td>1357</td>
</tr>
<tr>
<td>RUSSIAN FEDERATION</td>
<td>1258</td>
</tr>
<tr>
<td>ARGENTINA</td>
<td>960</td>
</tr>
<tr>
<td>SWITZERLAND</td>
<td>900</td>
</tr>
</tbody>
</table>
1. Some Consequences of The Web/Internet

- Essentials of supply chain have changed

- Libraries do not take physical custody of key content
  * online remotely, not on-shelf locally

- Role of libraries as trusted keepers of information and culture has been disrupted
  - Need assurance of continuity of access

- Can re-state the role of libraries graphically ...
Our Central Task: To ensure researchers, students and their teachers have ease and continuing access to online scholarly resources.

Classic use case: libraries secure provision of journal content

P. Burnhill, Edinburgh 2009
EDINA

began in 1996 as a JISC National Datacentre at the University of Edinburgh

- Art Abstracts, BIOSIS Previews, Compendex, EconLit, INSPEC, MLA, PAIS, etc

In 1994, we had launched SALSER, (one of) the first national union catalogues on the Web:

- union catalogue of serials:60 libraries in Scotland
  * 14 universities, National Library of Scotland, special libraries
  * Emphasis on assisting the searcher and visits to other libraries
    - early use of WWW and also use of Z39.50 for federated searching

In late 1990s, we gained knowledge of distributed architecture and Z39.50 in the JOIN-UP Programme: Table of Contents & British Library Document Supply, with federated searching using Z39.50 and then the use of OpenURL
EDINA

- Art Abstracts, BIOSIS Previews, Compendex, EconLit, INSPEC,
  - We had SALSER, a national union catalogue for Scotland on the Web
  - We had knowledge of distributed architecture and Z39.50
  - We were learning the importance of the ISSN

And then …

... in 1995/6, CIB (Bologna) made contact to propose an EU project: the CASA Project ...
CASA: ‘Co-operative Archive on Serials & Articles’

- EU 4th Framework: *Telematics for Libraries Programme*
- Prof. J. Di Cocco, CIB (Bologna, It); [A. Bollini, A. Citti, V. Verniti]
- Partners: CIB, ISSN-IC, ICCU (It), Ariadne (It), EDINA (UK) + many others (It)
  - F. Pelle
- 3-year, 2-phase project that began in 1997

- By 2nd phase, focus of project was switching from centralised ‘Archive’ to distributed ‘Activity’, hence new ‘descriptive title’:
- **CASA: ‘Co-operative Activity on Serials & Articles’**
  - (although ‘title proper’ remained unchanged)

[From a PPT presented to EU Commission in Luxembourg, November 2000]
The information chain

need to ‘verify’ the object: ISSN as ‘identifier’ to provide the linchpin

Verify
identification of article/serial

Discover
article of interest

Locate
article/serial service

Request
access/delivery of article/serial

Access
article of interest

not all serial lists have ISSN present

Based on the MODELS verbs (UKOLN, UK)

[From a PPT presented to EU Commission in Luxembourg, November 2000]
UK decides to create union catalogue of serials

- [www.SUNCAT.ac.uk](http://www.SUNCAT.ac.uk), 2005 -
  - v. many titles, mostly print; CONSER + ISSN + DOAJ
    * British Library, Oxford, Cambridge etc; 85 libraries
  - ISSN used to enhance OPAC records & assist matching, but missing/unassigned ISSNs

Now funded as a JISC Core Service
a Locate facility for researchers & students on the Web

www.SUNCAT.ac.uk
and for researchers & students ‘on the go …’

- **Service enhancement:** mobile App/geo-locate

**SUNCAT - ‘who holds what *where*?’**
SUNCAT – extra functions

1. Source of catalogue records for Contributing Libraries, including ISSN records

2. Support for researchers to access to text, via:
   - Tables of Content (ToC)
   - the UK OpenURL Router
     * a register of OpenURL Routers used by Institutions
   - the UK Access Management Federation (Shibboleth)

3. Support for decision-making by libraries who want to remove little-used journals from their shelves
   - UKResearchReserve (UKRR)
     * de-shelving; print archiving
     ‘if it is worth archiving it should have an identifier’
3. E-journal Preservation: Who looks after what?

• Switch of focus:
  – Digital/online/electronic only
  – Digital preservation
Our Central Task: To ensure researchers, students and their teachers have ease and continuing access to online scholarly resources

"ease"

usability

Licence

restricted

open

access

to content & services

"continuing"

access to back content

long-term preservation

Classic use case: libraries secure provision of journal content

P. Burnhill, Edinburgh 2009
E-journal Preservation: Who looks after what?

• EDINA had gained some experience through
  – CLOCKSS (EDINA represents U. of Edinburgh) and
  – UK LOCKSS Alliance (Digital Curation Centre & EDINA)

• In 2007, JISC had commissioned scoping study for a e-journal preservation registry which suggested basing this on SUNCAT

• EDINA & ISSN-IC combined as partners in the JISC-funded PEPRS Project in 2008.

• The design of the E-journal Preservation Registry had the ISSN the centre from the start ...
SERVICES: user requirements

E-J Preservation Registry Service

Piloting an E-journals Preservation Registry Service

DATA DEPENDENCY

E-Journal Preservation Registry

(a)

METADATA on extant e-journals

(b)

METADATA on preservation action

Digital Preservation Agencies e.g. CLOCKSS, Portico; BL, KB; UK LOCKSS Alliance etc.

m2m/API

ISSN Register

Abstract Data Model: Figure 1 in reference paper in *Serials, March 2009*
The outcome of the PEPRS project is the Keepers Registry, http://thekkeepers.org
- Launched as a full Beta service in October 2011, at the ISSN Directors’ Meeting in Sarajevo

Aiming to be a global facility, operating at web-scale, with international governance ...
The Keepers Registry allows you to search on title or ISSN * and with the ISSN-L as the key matching field, ISSNs for print can be tolerated *

Also acts as showcase for archiving organisations

3 CLOCKSS Nodes in Europe:
- Humboldt University, Germany
- Università Cattolica del Sacro Cuore, Italy
- University of Edinburgh, Scotland UK
Progress To Date: [ISSN-L as the basic unit] 17,000 unique titles now ‘preserved’ by the 7 archiving organisations registered as Keepers; ‘progress’ for a further 10,000. Some overlap, titles being preserved by more than one agency: good 😊

But
1. The coverage of volume/issues held is incomplete 😞
2. Still many more titles to be preserved:
   - The 17,000 ‘preserved’ includes both the ‘born-digital’ e-journals from publishers and those ‘born-again’ digitised serials.
** as bonus for EDINA & ISSN-IC, the Keepers report the current publisher in their metadata **

This helps the searcher but, as noted later, this could help build a Publisher Information Database linked to the ISSN.
4. E-Licensing: Who has entitlement to what?

- In days of print, what was paid for was put on the shelf, so the back-copy was there for perpetual access and use.

Reminder
Ensuring Continuity of Access means 2 things:
1. Preservation
Our Central Task: To ensure researchers, students and their teachers have ease and continuing access to online scholarly resources
Reminder: Consequences of Web/Internet

• Essentials of supply chain have changed
  * licensed to access, not sale for delivery of goods (content)
  * But is sale of good = access service + sale of entitlement??

• Libraries do not take physical custody of key content
  * online remotely, not on-shelf locally

• Role of libraries as trusted keepers of information and culture has been disrupted
  – Need assurance of continuity of access
    * of all content for future generations
    * of the back copies, especially if post-cancellation of the licence
E-Licensing: Who has entitlement to what?

• In 2010, EDINA & JISC Collections began to investigate how best to support libraries (and their patrons) through access to e-journal content **post-cancellation**.

• **We have been looking to see how we can learn from the design for the Keepers Registry**
Thoughts about a UK Private LOCKSS Network as dark store
E-Licensing: Who has entitlement to what?

- About the same time, 2010/11, the UK Standing Committee of University & Research Libraries (SCURL) was investigating the cost-benefits of an infrastructure for ‘shared services’.

- This is all coming together for a shared service for **Electronic Resource Management (ERM)** in what is called **KnowledgeBasePLus (KB+)**

All work-in-progress, but some (more) diagrams that may suggest what is going on ...

... and the importance and potential role of ISSN
"Two Systems" Dempsey
Basic bibliographic elements:

- Is/authorises
- has
- belongs to
- Date
- is issued as
- Vol. #
- ISSN
- DOI
- 'as RDF Triples'

<table>
<thead>
<tr>
<th>Entity</th>
<th>Predicate</th>
<th>Entity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Publisher</td>
<td>Is (authorized by)</td>
<td>Issuing Body</td>
</tr>
<tr>
<td>2 Serial</td>
<td>Belongs to</td>
<td>Issuing Body</td>
</tr>
<tr>
<td>3 Issue</td>
<td>Is part of</td>
<td>Serial</td>
</tr>
<tr>
<td>4 Article</td>
<td>Is part of</td>
<td>Issue</td>
</tr>
<tr>
<td>5 Volume</td>
<td>Is attribute of</td>
<td>Issue</td>
</tr>
<tr>
<td>6 Table_of_Contents</td>
<td>Is part of</td>
<td>Issue</td>
</tr>
<tr>
<td>15 Patron</td>
<td>Is (authenticated) member of</td>
<td>Library/Institution</td>
</tr>
</tbody>
</table>

Library:

- Org/ID
- belongs to

Patron:

- Ukamf

Org/ID:

- Institution
Subscription Taken may differ in Licence (Terms & Conditions) and Content:
(a) needed for digital library
(b) and by librarian ‘what serials are in Package?’
which Serials in which Packages?

- Need very good metadata that relates Serials (ISSN) with current Publishers

Licence to buy/access Serials is often offered in Packages, and negotiated by Consortia.
5. Forward Look, with 2020 Vision!!

1. General scheme for **DIGITAL & PRINT(?)** Content

2. Semantic Web, RDF Triples, URIs and all that jazz

1. Print will be on shelves for a long while yet, and the visit to a library elsewhere may be as (or more) important as (than) a visit to your own library

1. Hope to be sitting on a beach in Puglia in 2020
In the UK I recall the JOIN-UP Programme in which 4 Projects tried to become an integrated set of National Services but failed 😞
Another interesting graphic. CASA WP4 made reference to XML/RDF + Z39.50.

But did not go beyond the prototype 😞

Bollini, Burnhill & Di Cocco, 1998
To return to the essential role for the ISSN & useful changes being made by ISSN Network

1. Invention and use of the ISSN-L linking field
   * Makes a ‘family’ of ISSN (print & for electronic) for easier linking

2. ISSNs for digitised journal content (d-journals)
   * (in my opinion) Makes the electronic/digital the general case
     – as print becomes the special case – ‘fixity’ of content
   * and do not need (or want) a different ISSN for iPads etc

3. ISSNs for e-serials will keep on growing
   * as digitised versions of (unassigned) print journals emerge (libraries)

4. Wish to have ISSN integrate within Semantic Web

5. Need to recognise other identifiers (eg DOI, those created in union catalogues, and a muddle of URIs)
   * They will not go away
   * Need for ‘cross-walk’ between different identifiers for same Serial, as ‘same as’ in RDF Triples
Time for Q&A

Information about all projects and services, including ‘The Community Report: What EDINA Does’
http://edina.ac.uk

A splendid Helpdesk: edina@ed.ac.uk

Peter Burnhill: p.burnhill@ed.ac.uk