To bin or not to bin? Deselecting print back-runs available electronically at Imperial College London Library

RUTH COOPER
Team Leader
Information Resources Development
Imperial College London Library

DAVID NORRIS
IT Officer (Applications & Systems)
IT Team
Imperial College London Library

Abstract

Increasingly, academic libraries are investing heavily in e-journals which duplicate their print back-runs. For libraries facing acute pressures on space, one solution to their problem is to dispose of or relegate print back-runs which overlap with their electronic holdings. This case-study focuses on work at Imperial College London Library to provide a tool-kit for staff making such deselection decisions. Imperial has established criteria to determine the sustainability of their e-journals. A purpose-built database has been developed to store data on sustainability and enable staff easily to identify the overlap between print and electronic holdings at each Library site. Librarians at Imperial are now using the database and have already identified 700 metres of print journals with sustainable electronic access for disposal.

Background

Imperial College London is a world leading university, specializing in science, engineering, medicine and business. With a staff of 126 and a budget of £8.4m, the library is one of the UK's major research libraries, providing a crucial information service to support our research and learning activities. The Central Library on the South Kensington campus is home to the College's main print journal collection, and a further eight campus libraries and five departmental libraries house the remainder. Our journal collections comprise 17,000 print titles and 23,000 electronic titles, and over 3,500 of these titles currently overlap. Our print journals in the Central Library alone cover 5.3

kilometres of open access shelving and over 4 kilometres of closed access shelving.

There are acute pressures on space throughout the College. Since 2005 this has been a major strategic and operational preoccupation of both the College and the Library service, and this preoccupation is evidenced in the Library's 10-year vision document (2005)[1]. Part of the space-saving strategy outlined in the Vision revolves around withdrawing or relegating print holdings to closed access if we also have electronic access:

'*Some areas once used predominantly for the storage of printed collections will be transformed into flexible, inspirational spaces for student learning … Space occupied by printed collections will shrink, with the visible, open access collections being those that are heavily used and not yet available electronically.*'

To support and build on the Vision, we spelt out the strategy for our print journal collections in more detail in a new collection development policy which we named the Information Resources Management and Access Policy (IRMAP)[2].

The IRMAP recognizes and differentiates between two types of electronic access at Imperial: 'sustainable' and 'unsustainable'. Sustainable access is where we anticipate access to be available to Imperial freely or affordably for as long as required. Where we have sustainable access, the IRMAP recommends disposal; where the electronic access is unsustainable, the IRMAP recommends relegation. The policy received strong endorsement by the College Library Committee in 2006, and since then we have been working hard to assess and record the sustainability of our electronic holdings in order to implement it.

Sustainability criteria: long-term preservation of content

Our first task was to agree what constitutes sustainable electronic access at Imperial. We decided that we would not be held back by uncertainties over long-term digital preservation. Instead we have armed ourselves with two insurance policies which offer protection against future loss of access to valued journal content. These are our involvement in the UK Research Reserve Project (UKRR) and membership of Portico.

The UKRR[3] is an agreement between higher education and the British Library whereby the British Library will store journals no longer required by HE libraries, retain them permanently and make them accessible to researchers and others who wish to consult them. Imperial is the lead institution in Phase 1 of the project and under its aegis we have pledged that we will only dispose of titles where this would leave an agreed number of copies within the research library network and at the British Library.

Portico[4] is an e-journal archiving service which provides long-term trusted digital access. Many of Imperial's 'big deal' publishers (Elsevier, Wiley, OUP, CUP) have already signed archiving agreements with Portico. We expect to see other e-journal archiving services emerge over the next few years and would be very interested to support any that complement Portico's content and are similarly relevant, robust and affordable.

Imperial's three sustainability criteria

Imperial has adopted three sustainability criteria against which all our e-journals are assessed. We found very little in the professional literature to guide us in their development, and they emerged over several months, informed by our experience and knowledge of electronic journal management. It was important to the project that our criteria were fairly quick and easy to apply, as a very limited staff resource of one FTE had only two months in which to assess and record the sustainability of several thousand journal titles. We were also mindful of the long-term maintenance costs of the system we were developing.

We class our electronic access as sustainable when at least one of the following applies:

- Imperial has perpetual access rights to the content, via the web
- the journal is permanently open access for all years or certain years. Hybrid open access journals are not included in this category as the project is not interested in sustainability at the article level
- the content is in one of Imperial's trusted services such as JSTOR, the ACM digital archive or a JISC-funded archive.

Perpetual access rights via the web

Perpetual access is a term used within the library community to describe the ability to retain access to e-journals after the contractual agreement for these materials has passed[5]. A library which does not have such rights over an e-journal will lose access when the contract ceases. Imperial's perpetual access rights are nowhere near as ubiquitous as we would wish; we estimate that less that 50% of our content is covered. In addition, some of our licences specify an unsuitable delivery method for post-termination access. As we no longer support networked CD ROMs and do not have the resources to mount journal content locally, we will consider a journal sustainable only if perpetual access is provided via the web.

Often our perpetual access rights have an associated cost, for example one of our big deal licences specifies that we must remain a subscriber to the service in order to maintain access to journals that we previously subscribed to. We have licences which state that there will be a cost but give no indication of what this will be; others state that there will be a 'minimal' cost to maintain perpetual access. We expect our information resources budget to be able to support these extra costs in the future and have not made any attempts to solicit further information from publishers as we doubt that they would be able to provide any at this time.

Permanently open access journals

In the open access (OA) model, research output is made freely and permanently available on the web for anyone to read and use. For the purposes of our project, we consider journals which are fully OA to be

sustainable and hybrid OA journals, where only certain articles are OA, to be unsustainable.

Trusted services

Of our three sustainability criteria, this last, covering services that do not offer perpetual access rights, was the hardest to pin down. The services falling into this category all share two characteristics. The first is a good track record of stability, i.e. they have demonstrated continuity of titles from one year to another for as long as we have subscribed. The second is a history of and reputation for affordability and value for money. In producing this list we have been fairly cautious; had we been committed to a more radical deselection programme it would have been longer. We plan to review the list regularly and hope to see it grow as more services prove trustworthy.

Currently we consider electronic journal content which does not match at least one of these criteria to be unsustainable and we will relegate rather than dispose of the print copies.

Bullet-proofed access?

Does the label 'sustainable' against a title guarantee that we will always have electronic access? Sadly, but perhaps not surprisingly, the answer is 'no'. In the short time since we embarked on this project we have already lost access to a small proportion of our sustainable content with one publisher, and this is currently being queried with them

We expect to lose some of our sustainable content in the future because of journal transfers. Unfortunately, even where perpetual access rights are granted under a licence, they do not necessarily provide insurance against this category of loss. For example, under one of our big deal licences, if a journal is sold or transferred to another publisher, the licensor undertakes to use 'reasonable efforts' to retain copies of the volumes published and make them available through their online service. It follows that if their reasonable efforts fail, access to the journal in question may be lost. Project Transfer[6] is an initiative set up under the auspices of the UK Serials Group to address the challenges caused by the growing number of titles moving between publishers. We hope that one outcome of the Project will be a code of practice which safeguards a library's perpetual access rights in the event that a journal changes hands.

Despite the concerns expressed here, we nevertheless expect the vast majority of our sustainable e-journals to be available and accessible to Imperial for as long as our users require them. In our opinion these are low risk titles whereas our unsustainable journals are medium to high risk titles.

Collecting data on sustainability

Information on perpetual access rights was initially sourced from our licences. By this means we hoped to gather information on the majority of our big deal titles. Where it was not clear from the licence which titles and/or dates of our

holdings were covered by perpetual access rights, we planned to contact the publisher for a list. In fact, we contacted the publishers of all our big deals in any case as we wanted absolute clarity about our access rights on a title-by-title and date-by-date basis, and this was rarely provided in the licences.

Once we had trawled through the terms and conditions of all our electronic journal licences, we turned our attention to the publishers whose e-journals were not covered by licences. Frustratingly, these far outnumbered the publishers with whom we had licences. As time did not permit us to investigate them all, we restricted our investigations to certain categories of publisher, focusing on medical publishers – at that point in the project, our medical journals were strategically the most important – and publishers with whom we had a reasonable number of subscriptions. The websites of these publishers were checked for information on perpetual access rights. More often than not, this information was completely lacking.

The information (the relevant text excerpt from the licence or publisher's website) was entered in an Excel spreadsheet. The journals of the publishers who were not investigated are for now being classed as 'unsustainable'.

Finding OA journals which fulfilled our criteria proved harder than anticipated. The main stumbling block was our need for assurance on the permanency of OA. Although the Bethesda[7] and Berlin[8] Declarations on OA include perpetual access in their definitions[9], we discovered that not all so called 'open access journals' meet this criterion of permanency. For example, there is no guarantee given of permanency for journals listed in the *Directory of Open Access Journals (DOAJ)*[10]. We also discovered that publishers with a stated commitment to OA do not always practise what they preach. The signatories of the Washington DC Principles for Free Access to Science[11] share a commitment to providing free access and wide dissemination of published research findings. Yet early on in our project a Washington DC signatory withdrew its previously free access to 113 years of a journal without apology or explanation. Thus our final list of permanently OA journals ended up disappointingly short.

Finally we considered our aggregated services such as EBSCO Business Source Premier and Factiva. These clearly did not fulfil any of our sustainability criteria and were therefore recorded as unsustainable.

Data storage and retrieval: the stock management database

Whilst our Information Resources Development Team were busy gathering data on sustainability, the Library's IT Team were designing and developing a tool for storing and retrieving the data. The goal was to build a single repository for all the information needed to make decisions on sustainability, relegation and disposal by combining data from a number of sources:

- print holdings and site restrictions from our Unicorn LMS
- electronic holdings from our openURL resolver (SFX)
- sustainability information from our Excel spreadsheet.

We required the ability to easily identify overlap between print and electronic holdings, highlighting relevant sustainability information and site restrictions.

For the solution we briefly looked at the market to try to identify a suitable product off the shelf but there was nothing available that exactly matched our requirements without the need for major customization. We decided to implement our own solution in-house using Microsoft Access, drawing on our Library IT Team's existing database development and programming expertise. This meant that design and implementation would be carried out by staff who clearly understood the purpose of the database and were in a better position to respond to inevitable changes in specifications.

The basic database design followed the structure and relationships within the data, and identified three main tables: Title, Print Holding, and Electronic Holding. Each title would have one or more related print holdings and zero or more electronic holdings. A fourth table was identified to contain archived electronic holdings.

The Title and Print Holding tables were populated using data from our LMS. A script was run for each separate item location to extract serial records with MARC location fields 971, 972 or 973. The Electronic Holding table was populated using data exported from SFX.

Our LMS had undergone a number of migrations and mergers over the years and this has resulted in some duplication of serial titles with minor punctuation and spelling variations. Added to this is the need to accurately match the electronic holding title to its print equivalent. A process was developed to assign a short version of each full title which is used when matching between print and electronic holdings.

After the initial data load the database contained over 17,000 unique print journal titles, almost 21,000 print holdings, and over 7,300 matching electronic holdings.

Forms were created which would allow data to be viewed, added, edited or exported. The main form for data entry has undergone a number of revisions as new functionality has been added, the latest version is shown in Figure 1.
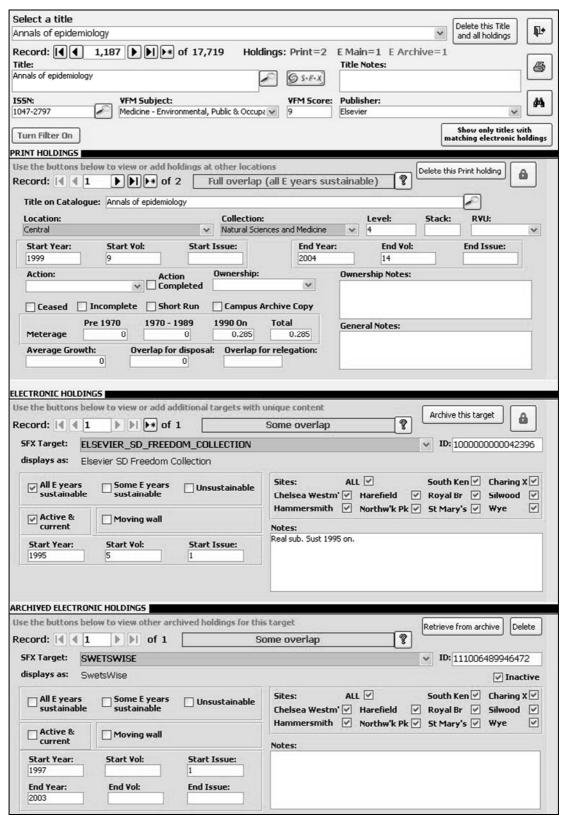
Figure 1. Screenshot of stock management database main data entry form.

In Figure 1 we can see that the journal *Annals of Epidemiology* has two print holdings. The copy at Central is fully overlapped by the electronic holding which is sustainable for all years. Not shown in the screenshot is the copy at the second location which is only partially overlapped. The electronic holding section shows one holding from Elsevier which provides some overlap (full at one location, part at the second). There is also an electronic holding in the archive since it is no longer active in SFX.

The form provides navigation buttons to scroll through the titles, or a title can be selected from a drop-down list. Two filtering options make it possible to limit the scrollable range or drop-down list to only titles that have matching electronic holdings, or to titles with print holdings at specific locations.
The title can be viewed in the OPAC using the magnifying glass icon, and its current electronic holdings displayed in an SFX menu using the SFX icon.

The print overlap indicator shows the overlap for that particular print location by all of the associated electronic holdings. The electronic overlap indicator shows the overlap for that particular target over all the print locations for the title. The overlap indicator on each holding is non-editable but is dynamically updated each time a change is made to any of the elements that contribute to the calculation.

Examples of actions which cause the overlap indicator to be recalculated are:

- adding or deleting a print holding
- adding or archiving an electronic holding
- marking a print holding as having been disposed.

Other forms in the database provide options for selecting and editing data in tabular views rather than single items, and for globally editing electronic holdings at the same target.

Reports provided by the database

The database has a reporting tool which enables users to tailor their own selections, display fields and method of export. The report tool can be used to supply a variety of outputs, such as:

- show all print titles at a specific location which are overlapped by fully sustainable electronic holdings, highlighting any site restrictions, and include the stack where shelved and meterage figures
- show all electronic holdings that do not overlap any print holdings and can therefore be archived in the database
- show all print titles at a specific location that are marked for disposal but have not yet been disposed of.

Maintaining the database

New print holdings must be added to the database manually. This is a straightforward task and the low number of new print titles added to our collection does not justify developing an automated alternative.

Changes to the electronic holdings are dealt with using data exported from SFX. We follow a monthly cycle where we update the SFX knowledge base with the latest amendments from the vendor's server, we review the changes in SFX, then we export the active full-text holdings to a file ready for importing into the database.
We considered fully automating the process of updating the database with any changes required as a result of the SFX update, but decided not to since decisions would need to be made about sustainability before some of the

changes were made. Instead, the database is automatically updated only with minor changes to volume, issue and moving wall time period information. All other changes required can be viewed in a spreadsheet from within the database.
The spreadsheet identifies six types of change:

- new holdings for titles that have no electronic holdings in the database
- new holdings for titles that have existing electronic holdings in the database
- holdings in the main table that have threshold changes
- holdings in the main table that are no longer active in SFX
- holdings in the archive table that have threshold changes
- holdings in the archive marked as inactive that are now active in SFX.

Staff can use the spreadsheet reports to manually update the database, although adding new holdings is simplified by, in most cases, requiring only the SFX target name to be selected for all the other fields to be populated automatically. The spreadsheet is dynamic in that when a change is actioned, it will no longer appear when the spreadsheet is regenerated.

Future plans for the database

Major development work on the database has finished and it is regularly being used by staff to aid decision making on disposing and relegating.

In the future we plan to merge the stock management database with two other Access databases developed in-house; one that manages periodical subscriptions and one that manages site licences. The resultant database will then be migrated to SQL Server to provide better security and data integrity. We also plan to convert the front-end of the database, including the forms and reports, to a web-based interface.

Conclusions

We are now 21 months into the project and consider it to be a resounding success. We have developed a robust and easy to use tool-kit for print back-run deselection decisions. With the aid of the database we have identified 700 metres of sustainable stock for disposal from one site, and the deselecton exercise will be rolled out to other sites in the near future. Although it is still early days, our sustainability criteria seem to be working. The only sustainable content that we have lost comes from four journals from the same publisher, and we are in the process of challenging this loss. This has proved to be an added benefit of the database; without it we would not have been aware that content over which we have perpetual access rights had been lost. We expect the stock management database to be guiding our print back-run management decisions for many years to come, as an increasing amount of this journal content becomes available to us electronically.

References

1. A World Class Library Service for the 21st Century: Vision for Imperial College Library:
http://www.imperial.ac.uk/Library/pdf/a_world_class_library_service_for_the_21st_century_vision_for_imperial_college_library_july_2005.pdf

2. Information Resources Management and Access Policy:
http://www.imperial.ac.uk/library/aboutthelibrary/policystrategicdirections/irmap

3. UK Research Reserve (UKRR) Project:
http://www.curl.ac.uk/projects/CollaborativeStorage/Home.htm

4. Portico: http://www.portico.org/

5. Wikipedia: http://en.wikipedia.org/wiki/Perpetual_access

6. Project Transfer: http://www.uksg.org/transfer

7. Bethesda Statement on Open Access Publishing:
http://www.earlham.edu/~peters/fos/bethesda.htm

8. Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities: http://oa.mpg.de/openaccess-berlin/berlindeclaration.html

9. Berlin declaration refers to 'irrevocable' access to the work.

10. *Directory of Open Access Journals (DOAJ)*:   http://www.doaj.org/

11. Washington DC Principles for Free Access to Science:
http://www.dcprinciples.org/index.htm

Article © Ruth Cooper and David Norris

▪ **Ruth Cooper**
Team Leader
Information Resources Development
Central Library
Imperial College London
E-mail: r.cooper@imperial.ac.uk

▪ **David Norris**
IT Officer (Applications & Systems)
IT Team
Central Library
Imperial College London
E-mail: d.norris@imperial.ac.uk

To view the original copy of this article, published in *Serials*, the journal of the UKSG, click here:
http://serials.uksg.org/openurl.asp?genre=article&issn=0953-0460&volume=20&issue=3&spage=XXX

For a link to the full table of contents for the issue of *Serials* in which this article first appeared, click here:
http://serials.uksg.org/openurl.asp?genre=issue&issn=0953-0460&volume=20&issue=3