

Author Posting. (c) 'The North American Serials Interest Group, Inc. , 2010.

This is the author's version of the work. It is posted here by permission of The North American Serials Interest Group, Inc. for personal use, not for redistribution.

The definitive version was published in The Serials Librarian, Volume 58 Issue 1, January 2010.

doi:10.1080/03615261003622742

(<http://dx.doi.org/10.1080/03615261003622742>)

PILOTING AN E-JOURNALS PRESERVATION REGISTRY SERVICE (PEPRS)

by

Peter Burnhill AND Fred Guy (EDINA, University of Edinburgh, Scotland, United Kingdom)

Abstract

It has been estimated that there are approximately 66,000 e-journals in existence with around 15,000 being scholarly journals. There are particular issues appertaining to journals in electronic format which do not apply to journals in printed format. Essentially a library pays for access but that access is not necessarily guaranteed over time. A number of organisations now take responsibility for preserving e-journals but it is recognised that there exists an overall lack of information about where e-journals are archived and what arrangements exist for access. PEPRS is a two year funded project with the objective of exploring the provision of pilot services to fill the information gap. EDINA, which is a national data centre running information services for higher education in the UK, based at the University of Edinburgh, has been funded by the Joint Information Systems Committed (JISC) to carry out the work in partnership with the International Standard Serial Number International Centre (ISSN IC) in Paris. The agencies involved are: British Library; CLOCKSS; E-Depot at Koninklijke Bibliotheek; LOCKSS and Portico. Essentially data from the ISSN IC will be linked to data from the agencies to provide appropriate services. Discussions have been held with staff from the agencies and a first pass pilot demonstrator is scheduled to be available in October to allow time for testing and evaluation and release of a system to undergo external testing in February 2010.

Keywords

Electronic journals; preservation; registry; access;

1. INTRODUCTION

The last 10 years has seen a very rapid expansion, both in the number of journals which are published in electronic form (although it still is only a small, yet increasing, number which are published only in e-format) and the relative usage of e-journals.

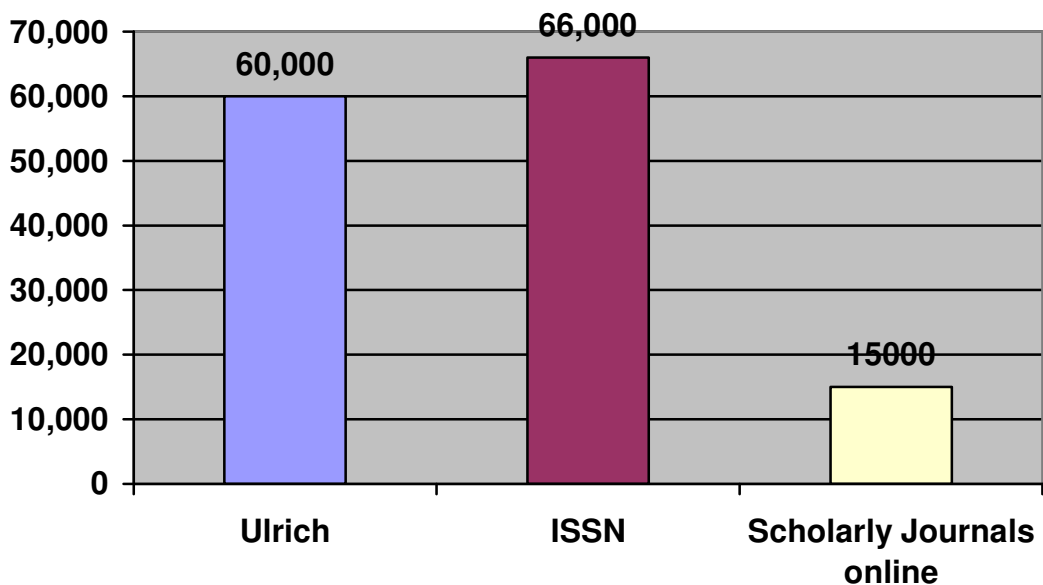


Table 1. E-journals in existence

Table 1 provides some figures on the number of e-journals in existence. Ulrich (June 2007) stated that there are 59,549 in the Ulrich database (1). In June 2009 the ISSN IC had identified 66,000 e-journals in the *Register* (see below). The figure for the number of scholarly online journals is taken from a 2006 report (2).

Table 2 shows the % of journals available online (3).

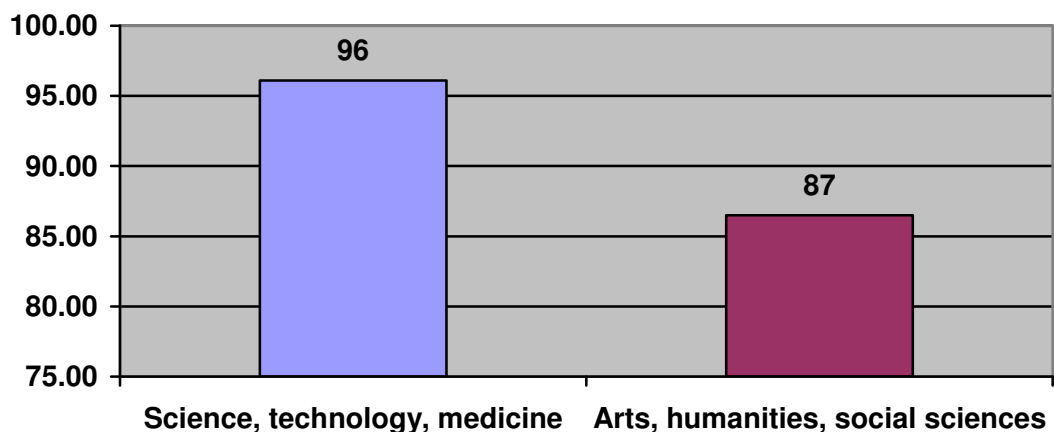


Table 2. % of journal titles available online.

There are issues appertaining to journals in electronic format which do not apply to journals in printed format. When a library takes out a subscription to a printed journal the copies received will be housed in the library and be available to the library users, essentially for ever or until the library decides to dispose of the copies. In other words, access to the copies is under the control of the library not the publisher. It will not

matter that the publisher ceases to publish the title, sells it to another publisher, goes out of business etc. The Library may also cease to subscribe to the particular journal but none of these events will affect the availability of what the library has paid for. All copies will continue to be available as long as the library wishes that to be the case.

The situation with e-journals is different because subscription does not mean that the library hosts the journal in the way it does print journals. Essentially the library pays for access and that access is not guaranteed over time. The events noted in the preceding paragraph (cessation of publishing, transfer, subscription cancellation) can lead to a particular set of users no longer having access to a journal. Clearly a failure to ensure access over time will have a potentially detrimental affect on scholarly communications.

This issue of access over time has taken two aspects. One is concerned with the long term preservation of electronic journals and the other has to do with providing for post-cancellation access. This paper will be dealing with long term preservation in the context of a specific project: *Piloting an E-journals Preservation Registry Service* (PEPRS).

2. BACKGROUND

In recent years a number of key reports have been written on the issues concerned with archiving and electronic journals.

Jones, stating that the National Electronic Site Licensing Initiative had encouraged the take-up of electronic access, discussed how clauses on archiving in the Licence might be implemented pointing out that publishers have never been expected to undertake preservation role in the print environment (4). She made a number of recommendations for further work to be carried out including maintaining a watching brief on initiatives and services being developed to provide archiving of licensed e-journals by trusted third parties (5).

Kenney et al, in a report published by the Council on Library and Information Resources, state that current license arrangements are inadequate to protect a library's long term interest in electronic journals (6). A specific recommendation is that libraries should participate in a registry of archived scholarly publications that indicates which programs have preserved them.

A key report entitled '*Scoping study for a registry of electronic journals that indicates where they are archived*' was commissioned by the JISC in 2007(7). The authors of that study interviewed a range of stakeholders including representatives from national and university libraries, publishers and archiving organizations.

"There is an information gap that needs filling" (8); "Almost everyone agreed that there was ... an overall lack of information about where e-journals were archived, but more particularly, the difficulty of finding the information across a range of sources". (9)

On the proposal for a registry, there was confirmation that there was need for readily-accessible information about whether and where e-journals were archived - although there was not uniform opinion: at least one archiving organization wished funding to go more directly to sustaining archiving.

The scoping study went on to note "We have not encountered anything which indicates that a registry would not be feasible. There are however a number of barriers (in) achieving it" (10) "There is no single view of what constitutes a registry." Differences of view emerged in their interviews and workshop about the context in which such a network-level registry should sit, and matters related to timing, implementation and sustainability.

Apart from costs and resources, the main challenge identified was achieving consensus on the scope of the registry.

The authors commented: "It is very important for all users to be clear about what the registry is and is not able to achieve at various stages of its implementation; in particular it is vital that librarians are not given the impression that because a title is in the registry, they will necessarily have access to it". (11)

Based on the report and its recommendations, the decision was taken by the JISC (see below) to fund a two year project to put into place the proposals for a registry as outlined in the report.

3. THE INVOLVED ORGANISATIONS

The funding for the project has been provided by:

Joint Information Systems Committee (JISC) (12)

The JISC was established in 1993 and supports United Kingdom post-16 and higher education and research by providing leadership in the use of ICT (Information and Communications Technology) in support of learning, teaching research and administration. It manages and funds more than 200 projects within 16 programmes and also supports 49 Services that provide expertise, advice, guidance and resources to address the needs of all users in HE and FE. The funding provided by the JISC for the development of innovative uses of ICT has helped ensure that a wide range of UK institutions have been able to become involved in research and development projects and that the international community has also been able to benefit.

The two partners on the Project are:

EDINA (13)

EDINA is a JISC national academic data centre based at the University of Edinburgh, Scotland. There is also an office in Cheshire, England with four staff. The main activity is the delivery of online services to staff and students in UK universities and colleges. With market coverage of over 90% of universities and two thirds of all colleges, EDINA is an established part of the UK digital library. In addition staff also undertake a range of research projects to inform future service.

The full range of EDINA activities is detailed on the web site and an annual community report is available for downloading. (14)

International Standard Serial Number International Centre (ISSN-IC) Centre International d'ISSN (15)

The ISSN IC is responsible for maintaining the ISSN Register which is the most comprehensive and authoritative source for the identification of serial publications worldwide.

Each of the bibliographic records in the Register contains, besides the unique authority elements (ISSN - key title - abbreviated key title), supplementary information, indispensable or useful for the identification of the publication concerned (frequency of publication, language, other forms of the title, place of publication, publisher...) and of the links which may associate it to other serial publications, which are also identified by different ISSN (former title(s), other language editions, other media editions...). It currently contains approximately 1,500,000 records of which some 66,000 are e-journals.

For the purposes of the Project the ISSN IC is providing the authoritative data (ISSN Register). EDINA has overall responsibility for the Project and will be developing the pilot service and hosting it.

4. THE AGENCIES

Kenny et al examined a dozen different e-journal archiving initiatives which the authors were aware of when the report was written and which all met their seven indicators of viability (6). Morrow et al selected six which they examined in some detail (16). The following is a brief introduction to the initiatives. Section 7 in the report contains a detailed analysis and comparison of the solutions (17).

- **LOCKSS**

LOCKSS (Lots Of Copies Keep Stuff Safe) enables libraries to play an active part in preservation. LOCKSS libraries maintain "LOCKSS boxes" which are used to store copies of all e-journal material that the library subscribes to (subject to publisher agreement). LOCKSS is therefore a closer analogue to paper distribution than any of the other solutions.

- **CLOCKSS**

CLOCKSS (Controlled LOCKSS) uses LOCKSS software to create a "dark archive". Access is enabled when the publisher authorises it, or when there has been substantial disruption to access via the publisher and the Board determines that the content can become freely available.

- **Portico**

Portico is specifically designed as a third-party electronic archiving service. Portico's role is as a permanent dark archive. Access is only permitted when there has been substantial disruption to access via the publisher.

- **e-Depot**

e-Depot was established in 2003 by the government-funded Dutch national library, Koninklijke Bibliotheek (KB). Content includes Dutch university repositories, websites, and other national large-scale digitisation programmes (eg Dutch newspapers). It originally included only Dutch e-journals, but later added an international dimension including worldwide scientific publications. Access is by agreement with the publisher, but is generally restricted to on-site viewing for private research.

A very welcome recent development has been the announcement that Lund University, host of the Directory of Open Access Journals (currently there are 4,000 journals in the collection) and Koninklijke Bibliotheek have started a co-operation to secure long-term preservation of open access journals. This material will be freely available. The Swedish Library Association is sponsoring the initiative (18).

- **OCLC Electronic Collections Online (ECO)**

ECO provides access to wide range of titles and publishers. ECO promises long-term access to subscribed content, subject to the library continuing to pay OCLC's access fee.

- **British Library e-Journal Archiving Programme**

The British Library has been building a Digital Object Management System (DOM) which can store and manage all BL's digital content. BL began ingesting content from selected publishers in 2007

The PEPRS project is particularly concerned with archiving initiatives and will be working directly with:

LOCKSS; CLOCKSS; Portico; e-Depot and British Library

5. LEGAL DEPOSIT

The favoured method, for ensuring that copies of all printed publications in a particular country are preserved and available in perpetuity, has been to enact legislation, requiring publishers to deposit copies of every publication in at least one library, usually a national library. In the UK, for example, legal deposit, as the process is called, has been in existence since 1610.

Legal deposit of electronic publications is proving to be a complex and difficult challenge. Kenney et al argue that whilst legal deposit can become an important component of a digital preservation programme it will not obviate the need to establish e-journal archiving programmes. Four reasons for this statement are mentioned (19). They are:

- Access issues with the likely requirement that users will have to visit the national library
- The ability of national libraries, despite the successful efforts of KB (Netherlands) and National Library of Australia, is, as yet, largely untested

- There is legislation in some countries including Canada, Australia, New Zealand and Norway but in some countries, such as the UK, there is only a voluntary scheme, and other countries, especially USA, have, at present, no scheme at all.
- The concept of national publications is not a straightforward one. Elsevier, while based in the Netherlands, has many journals with a primarily US-based editorial board and uses servers housed in a third country.

6. PEPRS project: key essentials

PEPRS is two year funded project with the defined aim to create a pilot registry service which will be assessed 18 months through the Project (February 2010) to ascertain the feasibility of the approach adopted being developed further into a sustainable service. Two publications on PEPRS have recently been published (20, 21)

The project is funded by the JISC and the project development work including hosting of the pilot service is being carried out by EDINA. The other project partner, the ISSN IC is providing authoritative data on electronic journals sourced from the large database (ISSN Register) created by the ISSN IC from data supplied by 87 national centres.

This authoritative data will be linked to data created by the agencies.

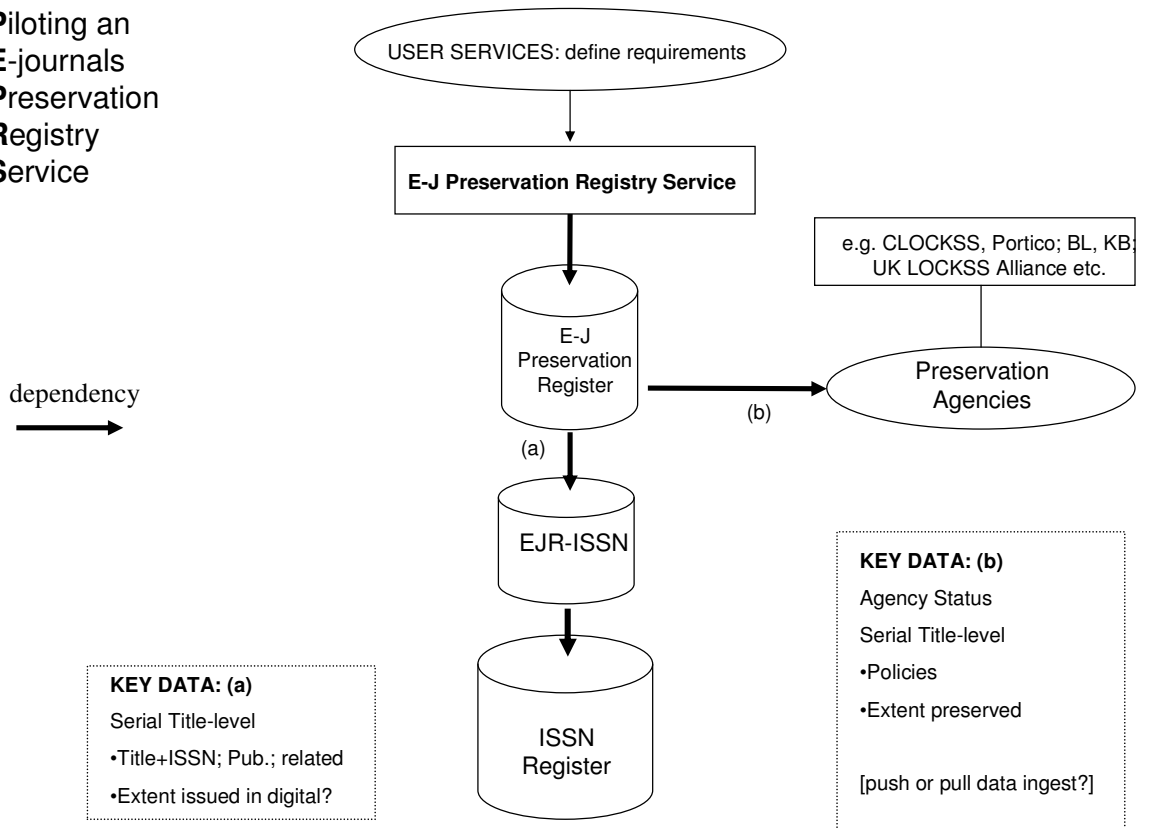
The key component elements within the PEPRS theoretical model are as follows:

- Authoritative data (identifier, descriptive data) on e-journals
- Data which describes what is being preserved and by whom
- Linkages between authoritative data and preservation data
- Services based on the registry
- User requirements

Table 3 below identifies all the key components noted above and the arrows indicate the dependency relationships.

Table 3. Components of PEPRS

**Piloting an
E-journals
Preservation
Registry
Service**



Data fields

The determination of a suitable minimal, yet adequate, set of data fields will be a key challenge within the project and at the time of writing this work is still under discussion.

The ISSN Register uses MARC 21 fields for the Register and a list of some of the fields most likely to be used within PEPRS are listed in Appendix A.

All the agencies involved have developed sets of data fields to reflect their particular approaches. Appendix B lists a set of possible data fields.

7. ISSUES AND QUESTIONS

There are a number of issues which have arisen in the course of the work to date. They include the following:

- a) What should be done about an e-journal which is in a preservation programme but has not been assigned an ISSN?
- b) Should print journals which are digitised retrospectively (many of which will not have been assigned an ISSN) be included in PEPRS?
- c) How should holdings information be collected, recorded and displayed?
- d) Should the registry be an international one and, if so, will it scale?
- e) If attention switches from preservation to post cancellation access should PEPRS try to adapt?

As far as a) is concerned assigning an ISSN is an important task for the ISSN IC so, accordingly, a workflow to identify this material will be developed so that this material can be assigned an ISSN quickly and efficiently.

For b) ISSN current practice has been not to assign a separate ISSN for the digital surrogate and this material has been deemed not to be in scope for PEPRS. Clearly, the situation needs to be monitored and it may well be that for the hoped for service to succeed PEPRS the material will be included.

Issues c), d) and e) will be discussed and debated during the Project and it is to be hoped that at the completion of the Project the ways ahead on each of them will be clearer.

8. NEXT STEPS

Contact and discussions have already taken place with the agencies on a one-to-one basis but it is planned to explore with each agency, in more depth, a series of issues which will include identification of fields, definition of fields and vocabulary use. Information from these discussions will be used in the preparation of a demonstrator.

A critical stage of the project is due to be reached in February 2010. The pilot demonstrator is to be assessed as to its capability of being developed into a service. Prior to this it is planned to have a first pass demonstrator available in October to allow time for testing and evaluation and release of the model which will undergo the external evaluation.

NOTES

1. <http://www.ulrichsweb.com/ulrichsweb/>
UK scholarly journals: 2006 baseline report. An evidence-based analysis of data concerning scholarly journal publishing. <http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/uk-scholarly-journals-2006-baseline-report>
2. (p.5)
3. Research Information Network. *E-journals: their use, value and impact*. A Research Information network report. April 2009. (p. 14). Available at:
[http://www.rin.ac.uk/files/E-journals use value impact Report April2009.pdf](http://www.rin.ac.uk/files/E-journals%20use%20value%20impact%20Report%20April2009.pdf)
4. Jones, M, *Archiving E-Journals Consultancy – Final Report*. Report commissioned by the Joint Information Systems Committee, October 2003; available from:
http://www.jisc.ac.uk/uploaded_documents//ejournalsfinal.pdf
5. Jones, 30
6. Kenney, A R, et al *E-Journal Archiving Metes and Bounds: A Survey of the Landscape*. Washington DC.: Council on Library and Information Resources, September 2006; available from:
<http://www.clir.org/pubs/reports/pub138/pub138.pdf>
7. Loughborough University and Rightscom. *Scoping study for a registry of electronic journals that indicates where they are archived*, Rightscom Ltd. and Loughborough University, January 2008; available from
[http://www.jisc.ac.uk/media/documents/programmes/preservation/ejournalregstudy.p
df](http://www.jisc.ac.uk/media/documents/programmes/preservation/ejournalregstudy.pdf)
8. Loughborough and Rightscom,30
9. Loughborough and Rightscom, 13
- 10.Loughborough and Rightscom, 30
- 11.Loughborough and Rightscom, 3
- 12.<http://www.jisc.ac.uk/>
- 13.<http://edina.ac.uk/>
- 14.What EDINA does: Community report: March 2009.
<http://edina.ac.uk/about/annrep/communityreportmarch09.pdf>
- 15.<http://www.issn.org>
- 16.Morrow, T, Beagrie, N, Jones, M and Chruszcz, J, *A Comparative Study e-Journal Archiving Solutions*; a JISC funded investigation. Final Report May 2008; available from:
[http://www.jisc-
collections.ac.uk/media/documents/jisc_collections/reports/e journals arc
hiving %20solutions report final 080518.pdf](http://www.jisc-collections.ac.uk/media/documents/jisc_collections/reports/e_journals_archiving_%20solutions_report_final_080518.pdf)
17. Morrow et al, 17
- 18.<http://www.doaj.org/doaj?func=loadTempl&templ=090401>
- 19.Kenny et al. 21-23
- 20.Burnhill, P. *Piloting an e-journals preservation registry service*. Serials 22(1) March 2008. Available at:
<http://uksg.metapress.com/link.asp?id=350487p5670h0v61>

21. Burnhill, P. *Tracking e-journal preservation: archiving registry service anyone?* Against the Grain. 21(1) February 2009. pp. 32,34,36

CONTRIBUTOR NOTES

Peter Burnhill

EDINA,
University of Edinburgh
Causewayside House
160 Causewayside
Edinburgh EH9 1PR
Tel: + 44 131 650 3301
Fax: + 44 131 650 3308
E-mail: p.burnhill@ed.ac.uk

Peter Burnhill is Director of EDINA and Head of the Edinburgh University Data Library.

Fred Guy

EDINA,
University of Edinburgh
Causewayside House
160 Causewayside
Edinburgh EH9 1PR
Tel: +44 131 651 3875
Fax: + 44131 650 3308
E-mail: f.guy@ed.ac.uk

Fred Guy is a project manager at EDINA.

APPENDIX A. ISSN MARC21 Fields

Data Source	MARC 21 Tag	Name	Rationale
ISSN Data	008	Fixed-Length Data Elements	Contains information including "Start date", "End date", "Country code", "Country of publication", and "Frequency".
ISSN Data	022	International Standard Serial Number	ISSN and ISSN-L
ISSN Data	222	Key Title	Formal name for title
ISSN Data	210	Abbreviated Title	Abbreviated version of formal name
ISSN Data	245	Title Proper	Title of serial generally used for identification
ISSN Data	246	Varying Form of Title	
ISSN Data	710	Added Entry - Corporate Name	The 210 and 222 are specifically ISSN forms of the title, may be more useful to have a field which contains the issuing body (like the 710) to distinguish titles with generic names such as "Annual report" or "Bulletin", but it will depend how the preservation agencies store such information.
ISSN Data	260	Publication, Distribution, etc. (Imprint)	Publisher from first assignment of ISSN (in event of title transfer, may be wrong?)
ISSN Data	362	Dates of Publication and/or Sequential Designation	First known year of publication
ISSN Data	776	Additional Physical Form Entry	The "Additional physical form entry" can be used to link sibling print and e-journal titles and ISSNs.
ISSN Data	780	Preceding Entry	This field describes prior instances, I.e. the serial before it changed publisher. Records are related via Record Control Number (or ISSN)
ISSN Data	785	Successor Title	The 785 "Successor title" could be useful in addition to the 780 "Former title".

APPENDIX B: SOME POSSIBLE AGENCY FIELDS

Agency	Archiving Agency	Name under which agency is known
Agency	e-ISSN	Use ISSN data for public view. Where possible use e-ISSN as primary link back to ISSN data
Agency	Print-ISSN	Use ISSN data for public view
Agency	Title	Use ISSN data for public view, probably Title Statement
Agency	Publisher	Use enumerated list of publishers?
Agency	Preservation Status	Preserved, Queued, Committed/Negotiated
Agency	Holdings (Volume, Issue)	Summary record of all volumes and issues held (e.g. V1 issue 1,2,3,4, V2 issue 1,2 etc). May wish to record both Volume and Year. Not to be included until second development iteration
Agency	Start Date of Committed Titles	Not to be included until second development iteration
Agency	End Date of Committed Titles	Not to be included until second development iteration
Agency	Start Date of Processed Titles	Not to be included until second development iteration
Agency	End Date of Processed Titles	Not to be included until second development iteration