Tales from The Keepers Registry: Serial Issues About Archiving & the Web

Peter Burnhill

Causewayside House, University of Edinburgh, EH9 1PR, 160 Causewayside, Edinburgh EH, Scotland, UK

ARTICLE INFO

ABSTRACT

A key task for libraries is to ensure access for their patrons to the scholarly statements now found across the Internet. Three stories reveal progress towards success in that task. The context of these stories is the shift from print to digital format for all types of continuing resources, particularly journals, and the need to archive not just serials but also ongoing 'integrating resources' such as databases and Web sites.

The first story is about The Keepers Registry, an international initiative to monitor the extent of e-journal archiving. The second story is about the variety of 'serial issues' that have had to be addressed during the PEPRS (Piloting an E-journals Preservation Registry Service) project which was commissioned in the UK by JISC. These include identification, naming and identification of publishers, and the continuing need for a universal holdings statement. The role of the ISSN, and of the ISSN-L, has been a key.

The third story looks beyond e-journals to new research objects and the dynamics of the Web, to the role of citation and fixity, and to broader matters of digital preservation. This story reflects upon seriality, as the Web becomes the principal arena and medium for scholarly discourse. Scientific discourse is now resident on the Web. Much that is issued on the Web is issued nowhere else: it is a digital native.

Statistics that indicate the extent of archiving for e-journals to which major university libraries subscribe are also included in the article.

1. Prologue

There are three stories to tell, one tale within another. All the stories relate to the wish to ensure enduring access to scholarly statements and resources in the digital age. The context is the shift from print to digital format for all types of continuing resources, particularly journals, and the need to archive not just serials but also ongoing 'integrating resources,' such as databases and Web sites, that are cited in scholarly statement.

One purpose is to give an account of an international endeavor to monitor the success of arrangements for long-term access to the world's serials literature: a project that led to the creation of The Keepers Registry. To accomplish this, it is necessary to make introductory reference to literature beyond that generally within the sphere of the serials librarian, for example, on digital archiving and on scholarly use of the Web. This is not a review of the serials literature, as others do usefully (Collins, 2011). The purpose here is to bring attention to several threads but without the obligation to tie these together.

The context for a registry of archival activity for digital scholarly statements is dynamic. In this article therefore the reader is provided with both short summaries of what is currently understood and pointers to further reading, often because these matters remain unresolved.

The central proposition that underpins The Registry is summarized in a four-point mantra:

1. Assign an identifier at the 'point of issue' for a stream of digital content,
2. Ensure that (digital) content is archived routinely and, preferably, have others/peers do that for you too,
3. Tell someone what you are doing (and how) and what you hold,
4. Publish the terms of access for the archived content (now and when triggered as orphaned).

This may sound familiar to librarians who manage published print serials. However, the proposition extends beyond such 'well-published' material.

1.1. Monitoring Archiving: Who is Looking After Your E-Journal?

The first story told is about the making of The Keepers Registry, an online and global facility that enables all concerned to discover what serial content is being preserved for the long term. The real heroes of the story are the Keepers themselves, the archiving organizations who have stepped forward to act as stewards for e-journal content. The Registry is a lens onto their archiving activity and the policies that they have adopted.

The driver is recognition and concern, across the world, that libraries no longer have custody of the contents of journals to which they

subscribe: what was once on-shelf locally is now online remotely. It has become vital for scholarship that we know what information has assured availability for the long term and what is at risk of loss.

1.2. Serial Issues

The second and longer story concerns a variety of matters about serials that have been addressed in the making of The Keepers Registry. These include use of the International Standard Serial Number (ISSN), challenges about 'holdings,' publisher information and what to do with digitized content of print journals. The project assisted thinking about changes, of principle and practice, within the ISSN Network. The solutions devised helped in the development of The Keepers Registry, but they also will have a more general impact for those working with serials, whether print or electronic.

1.3. The New Literature of the Web

In many ways the context for the third story is in the phrase 'what is so different about the digital' and also about what is becoming understood about the ever-increasing use of the Web, including its properties over the long term. The third story is, therefore, a rich mix of associated topics that bears upon the purpose of The Keepers Registry. These include notions of fixity, the copy of record, citation and the continuing value of seriality — even as the richness of scholarly statements increases.

The practice of seriality, both through the issue of parts and in recognition of the impact of time upon changes to databases and Web content, is powerful. Whether it can be harnessed and prove useful for many of the unresolved matters that are discussed here is moot. This bears upon the attempt to ensure that there really is some completeness for the continuity of access to the resources that scholarship requires.

2. The First Story: A Monitor for Archiving Activity

2.1. E-Journal Preservation

The shift to journal content that is digital, online and held remotely has challenged the essential responsibility that libraries have in ensuring continuity of access to scholarly content for their patrons. After many studies and projects, including Jones (2003) and Cantara (2003), an increasing number of organizations have stepped forward to act as long-term archives for e-journal content, as noted in Kenney, Entlich, Hirtle, McGovern and Buckley (2006) and in Morrow, Beagrie, Jones and Chruszcz (2008).

The Keepers Registry has emerged as a resource that could address "who is looking after what, how, and what are the terms of access?" These questions have now been re-phrased as an invitation on The Keepers Web site, "Discover who is looking after what e-journals," as illustrated in the screenshot below (Figure 1).

This online facility allows search on a title or an ISSN, as well as browse across publisher and journal title in order to look up the organizations that have committed to a title's long-term continuity of access. Perhaps the best way to appreciate what the Registry does, now and as a full service as of 2013, is to go online (http://thekeepers. org) and use it.

2.2. The PEPRS Project

The Keepers Registry is the product of the PEPRS project, which aimed to Pilot an E-journal Preservation Registry Service (PEPRS). That project was jointly carried out by EDINA (http://edina. ac.uk) and the ISSN International Centre (ISSN IC) (http://www.issn. org) in response to a call by JISC (http://www.jisc.ac.uk), the

innovation and service organization for higher and further education in the UK.

Following positive review in February 2010, JISC authorized a second two-year phase for the PEPRS project in order to transform the prototype Registry into a form suitable for delivery of service, initially as a public (PEPRS) Beta service that was launched formally at the meeting of the ISSN Governing Board held in Paris in April 2010. Re-named as The Keepers Registry, the Beta service was re-launched at the annual Meeting of Directors of ISSN Centres held in Sarajevo, Bosnia, and Herzegovina in October 2010.

Others had done the early thinking about a registry. JISC had noted concerns about long-term access to e-journals being expressed in the literature, in both Jones (2003) and Kenney, Entlich, Hirtle, McGovern and Buckley (2006), with the latter including a call for "clarity of public statement by each agency or through a registry." In 2007 JISC commissioned a study to investigate the feasibility and the perceived need for an e-journal preservation registry. The study by Sparks, Look, Muir and Bide (2008) reported in the affirmative indicating that such a registry could be built around the Serials Union Catalogue (SUNCAT), the national union catalog in the UK developed at EDINA (Burnhill, Halliday, Rozenfeld & Kidd, 2004).

EDINA had been exploring new ways in which union catalogs of serials could play a strategic role in the information environment (Burnhill, Guy & Osborne, 2007; Burnhill & Law, 2005). Recognizing the emphasis given in the feasibility study to the global context, as the assurance of continuity of literature published across the globe is of interest to all regardless of the place of publication, EDINA approached the ISSN International Centre to be a partner in the project. The project began in 2008.

Success was to be judged on the delivery of a network-level facility that could be regarded and used as an authoritative reference source on the preservation status of e-journals. An equally important aim for the project was an advance in knowledge. Moreover, in addition to providing support for libraries, policy makers, publishers, and intermediaries, there was always the prospect that the project would assist the archiving organizations themselves who might thereby discover ways of establishing mutually acceptable division of labor.

Throughout the project there was always an intention to be international. A scholar in one country reads what is written and published in another. There was a need to engage librarians and others beyond national boundaries. This led to presentations across Europe, at the North American Serials Interest Group (NASIG) Annual Conference (Burnhill & Guy, 2010) and at the meetings of the ISSN Network, including China, as reported on the project Web sites (http://edina.ac.uk/ projects/peprs/links.html and http://www.issn.org/2-24096-Projects. php).

In the proposal for the PEPRS project it was important to define a scope that would ensure that the project was manageable and could achieve results. This meant an initial focus upon serial-level (as opposed to volume-level or even article-level) information and upon those e-serials that already had an ISSN present in the metadata. The high priority that the ISSN Network had placed upon assigning ISSN to e-serials meant that there was high expectation that nearly all the e-journals of primary interest (as scholarly journals) would have an ISSN. Later, we would make a principle of our practice, declaring that if something was worth preserving, then it deserved an identifier.

The focus on serial-level metadata allowed the project to defer until the second phase its attention upon the more problematic matter of the 'holdings statement' (volumes and issues held or preserved). However, what would be required before launching The Keepers Registry was some evidential statement about archival activity, i.e., affirmation by each archiving agency regarding the actual volumes of digital stuff they had ingested.

Other project objectives included a requirements analysis (with use cases), a formal statement of the information architecture and a
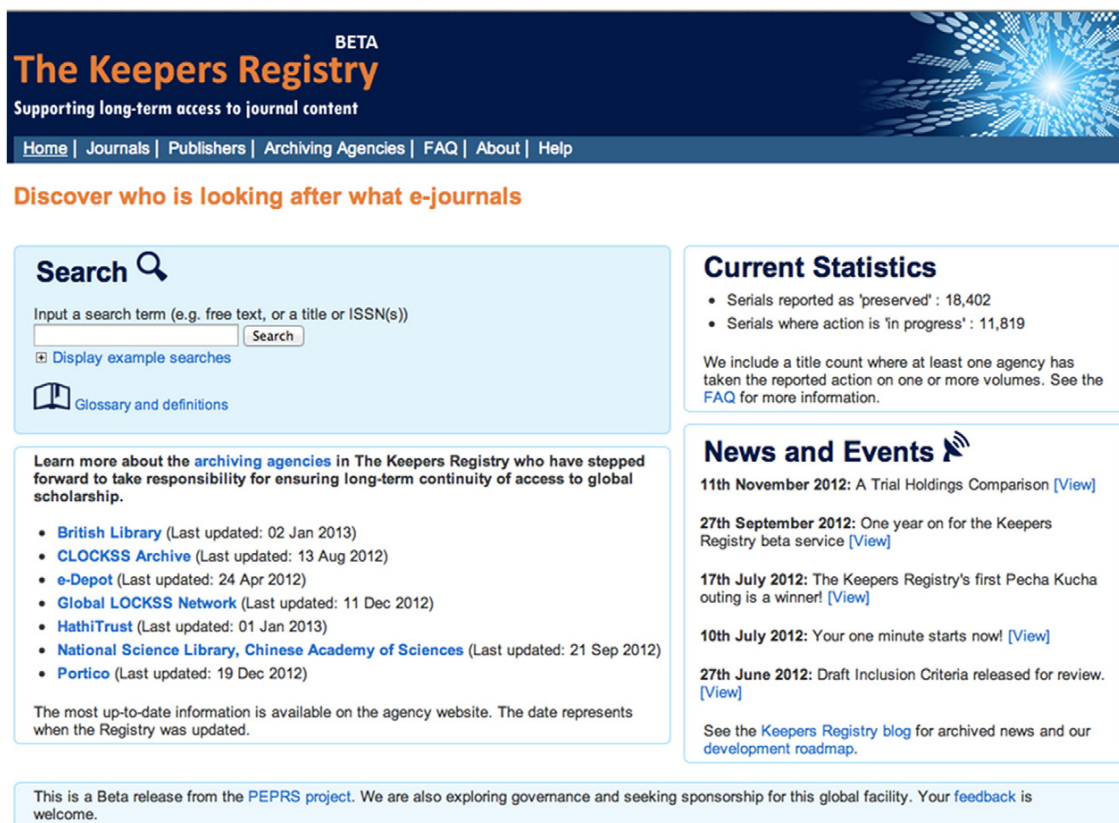
Figure 1. The Keepers Registry [screenshot].
(Taken from the Beta release of The Keepers Registry on January 29, 2013.)

prototype or working demonstrator that was suitable for external evaluation. An important task in the project was to think through what the data dependencies would be. The identified dependencies and the general design are illustrated in the model below (Figure 2), taken from the reference article on PEPRS by Burnhill et al. (2009) published in *Serials*, the journal of the UKSG. (As though to illustrate forthcoming challenges, the title of *Serials* has recently been changed by its issuing body to *Insights*.)

The intention in the design was to specify what was 'minimally-sufficient' for such a registry: access to up-to-date and authoritative data streams for (a) metadata on the objects, namely e-journals (serials), and (b) metadata on the archival action being taken for the issued content.

For serials, the ISSN Register is an accepted authority. For metadata on archival action there is no such accepted authority. Accordingly, metadata in a preservation registry would consist of periodic self-statements by the
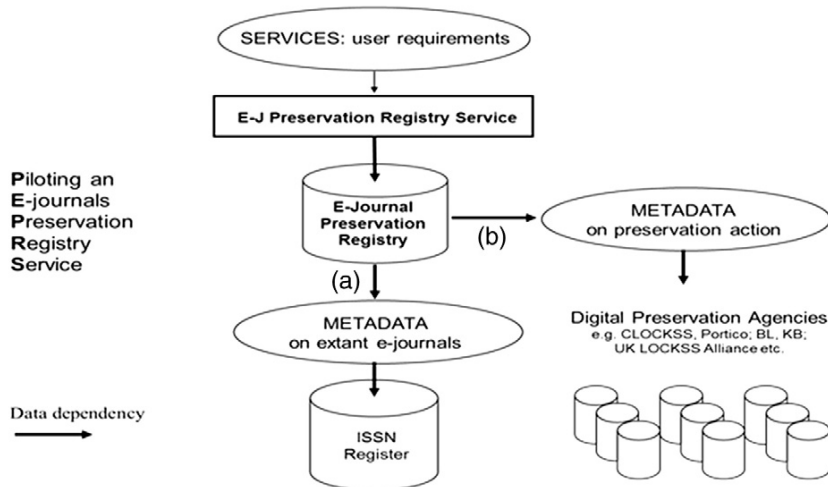

Figure 2. Abstract data model for PEPRS and The Keepers Registry.

archival organizations themselves: on what they held and how, including mention of any reports of audit and certification.

2.3. About the Partners

EDINA and the ISSN IC, the two project partners, had worked together for many years, having been project partners in the late 1990s in an EU-funded project (http://casa.unibo.it/meetings/bull1eng.html) and through EDINA/SUNCAT representation as an Observer at ISSN meetings. Each organization appreciated the complementary strengths brought to the project by the other.

The ISSN IC is an intergovernmental institution governed by statutes in a convention (SC-76/WS/4) made in 1976 between UNESCO and France (the host country) for assignment of the ISSN over "the full range of recorded knowledge" (http://www.issn.org/2-22669-International-Centre-Statutes.php). The IC manages the ISSN Register and acts as coordinator for the multi-national ISSN Network of over 80 national ISSN centers located primarily in national libraries, including the Library of Congress. Partnership with the ISSN IC provided access to first-class expertise about serials, including determination to understand and address key questions about the broad range of electronic continuing resources. The ISSN Network traditionally had not concerned itself with the extent of issues published. While the ISSN Network deemed information about the current publisher important, it had found maintenance of the volatile metadata to be a significant challenge for serials registered with ISSN worldwide.

EDINA, as the JISC-designated national data center at the University of Edinburgh, has extensive experience in the delivery of online services for serials and articles and in project management. This experience provided a connection into networks of university libraries both in the UK and internationally. EDINA had also gained knowledge of digital preservation from its active participation with the Digital Curation Centre (DCC) (http://www.dcc.ac.uk), the UK LOCKSS (Lots of Copies Keep Stuff Safe) Alliance (http://www.lockssalliance.ac.uk/), and through representing the University of Edinburgh as a partner in Controlled LOCKSS (CLOCKSS) (http://www.clockss.org/clockss/Home). With JISC, EDINA also contributes infrastructure for institutional repositories and for access to commercially licensed journal content.

Five leading archiving organizations agreed to act as associate partners. These included CLOCKSS, Portico (http://www.portico.org/), the Global LOCKSS Network (http://www.lockss.org/), Koninklijke Bibliotheek (KB) (whose e-Depot (http://www.kb.nl/hrd/dd/index-en.html) has international significance because of its positive engagement with Elsevier and Kluwer) and the British Library (http://www.bl.uk/aboutus/stratpolprog/legaldep/#elec). During the second phase of the project, two other archiving organizations of note joined this effort: HathiTrust (http://www.hathitrust.org/about) and the National Library of Science of the Chinese Academy of Sciences (http://english.las.cas.cn/). They each provide test data, the intention is to investigate how in the longer term the supply of those metadata might be made available automatically across the Internet.

2.4. Progress Report

One sense of progress with respect to e-journal preservation is the fact that the seven archiving organizations are reporting into The Keepers Registry and that they affirm that about 18,500 unique titles are being preserved. Less positive, as will be shown later, is that the extent archived for a given title, in terms of volumes and issues, is highly variable.

It would be comforting to think that those 18,500 titles represented good coverage of scholarly journals, recalling the initial motivation of the PEPRS project. However, as yet, there is no categorisation of serials within The Keepers Registry that would enable estimate of such

as statistic. That will have to be reported later once additional metadata has been added to the e-journal registry.

The total number of refereed scholarly journals is around 30,000 as indicated by Yvette Diven (SerialsSolutions) in an email to Sally Morris that was forwarded to a closed e-mail list, "As of 3 August 2012, the number of active peer-reviewed journals listed in Ulrichsweb is 28,094 titles." As noted below, ISSN have been assigned to 100,000 e-serials, a number expected to rise as ISSN is increasingly assigned to Web sites and to digitized journals.

Another take on progress with respect to e-journal preservation would the extent to which a given library could have re-assurance about the e-journals that it is keen to provide to its patrons, especially as it considers a switch to some variation of 'digital only', for either current subscriptions or even back copy as part of some program of print archiving. That will of course vary from one library to another, but concern has been urgent for libraries with large expenditures, as indicated by the title of a presentation, 'A Potential Crisis In Electronic Journal Preservation' (http://www.cni.org/topics/digital-preservation/preservation-status-of-eresources/) and by associated activity:

1. Cornell and Columbia University Libraries (2CUL) reported comparison of their electronic journal holdings with titles preserved by LOCKSS and Portico, noting the potential for such a facility when commenting upon the PEPRS project (http://2cul.org/sites/default/files/2CULLOCKSSFinalReport.pdf).
2. There was similar analysis in Germany (Seadle, 2011).
3. The Preservation Officer at Duke University carried out a title-by-title review via the user interface for The Keepers Registry, motivated by the 2CUL Study. This review revealed gaps in the holdings for some titles with volumes missing even across all archiving agencies.

News of these activities led to invitations being made to Duke and 2CUL in late 2012 to help test the holdings comparison facility that was being developed for The Keepers Registry. In all three instances, the files supplied by the libraries were of online serials for which the ISSN was known and present. Each library received back a file that indicated the following: which title (ISSN) was preserved; which were in some sort of process (awaiting preservation); and titles where there was no reported archival action being taken. Volume information was also supplied in order that the libraries could later carry out and report more extensive and revealing analyses. The engagement with Duke is available as a case study (http://thekeepers.blogs.edina.ac.uk/2012/11/05/a-trial-holdings-comparison/).

With their kind permission, the three university libraries have allowed the summary title-level statistics for the exercise to be presented in Table 1. These should be regarded as indicative and ought not to be over-interpreted; the results, in terms of percentages, are based only on those online serials for which an ISSN was known. Nevertheless, a broad picture emerges: about one quarter of the titles of each (with ISSN) are being preserved by one or more of the archival organizations reporting into The Keepers Registry. There is greater assurance of long term preservation indicated for a smaller number (and percentage) of titles that are being preserved by three or more archives. Conversely, about three-quarters of titles (with ISSN) could be said to be at risk, although there may be some titles for which archiving is still 'in progress' or is intended. Recall also that this analysis is limited to those serials for which the ISSN was known, about half the serial titles listed by each library. There is no knowledge of those without an ISSN.

Yet another take on progress might be to carry out analysis that would generate a similar title-level summary by publisher, given that publishers may be required to provide evidence that they have engaged one or more archival organizations to act as stewards for the long term. Similarly, such checks might trigger libraries and archiving organizations to contact publishers whose content appears to be at risk.

Table 1
E-Journal Preservation Status[a] for Three US Research Libraries.

University library	% 'Preserved' by 1 or more	% 'Preserved' by 3 or more	% Not known as 'Preserved'	Total having a valid ISSN
Columbia	26	12	74	58,882
Cornell	28	13	72	54,698
Duke	22	10	78	61,682

[a] When checked against preservation activity of The British Library, CLOCKSS, e-Depot (Netherlands), Global LOCKSS Network, HathiTrust, National Science Library of China and Portico, as reported into The Keepers Registry.

3. The Second Story: A Volume of Serial Issues

In both the initial stages of the PEPRS project and in later stages when preparing for the launch of The Keepers Registry Beta service, the project team encountered a number of challenges to do with metadata about serials and their content. The project's work with the test data provided by the archiving agencies underpins the experiential basis for the second story to be told, a story about both data ingest (do we know what we have?) and about display for users (how do we collate and aggregate the data we have ingested?).

Typically the metadata that each archiving organization receives from the publishers, and, therefore, what they have available to pass on to be aggregated by The Keepers Registry, is at the level of issue or volume and in some instances at the level of the article. With some variability across archiving organizations, the following types of challenges were encountered: uncertainty with respect to identifiers; variants in publisher information; and variability about 'holdings' information relating to issues, volumes, and other buckets of digital stuff.

3.1. Identification

When working with test data from the archiving agencies we encountered the first set of 'issues' listed above: uncertainty with respect to identifiers. The decision was made at the start of the project that although the Registry should be inclusive in scope, content relating to a given title must have an identifier — rather than attempt matching of text in titles. The ISSN Network included electronic resources as in scope for ISSN (http://www.issn.org/2-22638-ISSN-and-electronic-publications.php). For project purposes we required a valid ISSN that could be validated against the ISSN Register. With priority given by the ISSN Network to working initially with the larger publishers, there had been a dramatic increase in the number of ISSN assigned to those journals the project intended to cover. The ISSN General Assembly in April 2012 noted that the ISSN Network had assigned 97,581 (circa 100,000) ISSN to online continuing resources.

Having 100,000 'eISSN' (the shorthand for the ISSN assigned to the e-version of a resource is commonly used but is not an official ISSN Network term) is impressive, but there is strong indication that there is more assignment work to be done. ISSN assigned for electronic continuing resources are highly clustered geographically, reflecting the scale of production by publishers in those countries and also the success of each ISSN national center in ensuring ISSN assignment for serials that are published online. About 20% of those 100,000 were issued by the U.S. ISSN Center at the Library of Congress for serials with a place of publication in the United States. The British Library's UK ISSN Centre assigned about 10% of the total for those published in the UK. ISSN national centers in the Netherlands, Germany, and Brazil have each assigned about 4.5%. The ISSN International Centre assigns about 3.5% on behalf of international bodies and a number of smaller countries. The relatively low numbers of ISSN assignments for titles from India and China, and Hong Kong, given their size, are a reminder

of the number of e-journals and the like that remain hidden. Egypt and the ISSN IC on behalf of South Africa assigned the most ISSN for that continent. Egypt also had the largest assignment for the Arab States. Collectively there are a large number of eISSN assigned in Latin America and the Caribbean.

3.2. Incorrect, Erroneous or Missing (with note on the value of the ISSN-L)

The metadata for the content that the archiving organizations receive from publishers did not always include an appropriate ISSN, indeed sometimes there was none at all. There are a variety of reasons for this, which are discussed below along with the problems and solutions for handling e-journal content with incorrect, erroneous, and missing ISSN. (There were also occasional deviant ISSN detected relating to matters internal to the ISSN Network.)

3.2.1. Incorrect Use of the Print ISSN

The ISSN for the print version is sometimes used instead of the ISSN assigned for the electronic version. This situation is indistinguishable at first glance from the instance where only the ISSN for the print version has been assigned — see Section 3.2.4 below.

Use of the print ISSN for the online version might have been a problem but for the relatively recent implementation of the ISSN-L (http://www.issn.org/2-22637-What-is-an-ISSN-L.php). This is an especially designated and tagged ISSN shared by all medium versions of a serial and present within every ISSN record. The ISSN-L therefore acts as a field for linking equivalent content. There seems little doubt that the ISSN-L now provides the means to rise above the level of the manifestation: regardless of manifestation as print or digital and whatever one thinks of the applicability of the Functional Requirements for Bibliographic Records (FRBR) model to serials given the primacy of seriality over hierarchy (Reynolds, Whitney, Chesler & Beck, 2009; Shadle, 2006).

The Keepers Registry can lay claim to one of the earliest, if not the first, application of the ISSN-L where it has certainly proved its value. By deploying the ISSN-L as the key field for matching both in the design of the database and in the user interface, the potential for confusion between the ISSN assigned to print and the eISSN is addressed and solved. If the print ISSN were (incorrectly) presented by either the archiving agency or the end user, the ISSN-L would link to the ISSN in the ISSN Register for the electronic version.

The ISSN-L has potential for much greater strategic value and could benefit the full range of aggregation services for the purposes of search, delivery, and counting, including union catalogs and the variety of services that use the OpenURL syntax in which the ISSN is a key component. The success of a link requires the same ISSN in both the OpenURL request and the target (i.e., the full-text resource), but when the title in question is available both in print and online digital format, there is more than one valid ISSN to consider. The ISSN-L enables a link resolver knowledge base to manage the link successfully, as noted by the National Information Standards Organization (NISO) and UKSG's Knowledge Bases and Related Tools (KBART) working group (http://www.niso.org/workrooms/kbart).

There is also huge potential for the ISSN-L to be the linchpin within Linked Data statements, as the basis for the uniform resource identifier (URI) for the predicate between the ISSN for print and electronic versions, for example.

3.2.2. Erroneous Use of an ISSN: Use of Some Completely Different Number

The involvement of the ISSN IC as a full partner in the project meant that there was expertise available for direct cross-check of the metadata supplied by the archiving agencies against the ISSN Register, the database that generates the ISSN Portal product. In addition

to containing metadata for all fully registered ISSN, the Register documents the "suppressed" ISSN (those assigned but later found to not have been published) and the "free ISSN" (those ISSN that had been allocated to national centers but for which metadata had not yet been entered into the ISSN Register) or those ISSN not yet allocated at all. Use of erroneous, "suppressed," or "free" ISSN were few and far between, and when encountered, the corrections could be fed back to the archiving agencies and the publishers that had supplied them with the metadata.

3.2.3. Missing ISSN by Omission: Where the ISSN has been Assigned

As stated above, the decision was made that The Keepers Register would use the ISSN Register as the authority file for each continuing resource being archived, and therefore, the ISSN had to be present in the metadata presented by an archiving organization.

The solution to the 'missing' problem – where the ISSN exists but has been omitted either by design or by lack of awareness on the part of the archiving agency – is a practical matter requiring commitment of cost and effort. However, there are pricing and cost implications related to use of the ISSN Register and thus to the provision of a text matching facility to enable archiving organizations to enter the missing metadata prior to acceptance within The Keepers Registry.

3.2.4. Missing ISSN Because an ISSN had not yet been Assigned

The presence of an ISSN is more than a project convenience. If a resource is worth preserving then it should have an identifier, and if it is a continuing resource and is eligible for an ISSN, then it should have an ISSN and be in the ISSN Register.

The simple explanations for this type of 'missing' ISSN (an archived title doesn't have an ISSN or has an ISSN for the print version only) are that the particular continuing resource in electronic format has not been brought to the attention of the ISSN Network or is in a queue waiting upon assignment by the national ISSN center of the country in which the organization responsible for the issue of the content is based. The ISSN IC is investigating appropriate workflows that would assist and lead to the assignment by the ISSN Network of an ISSN for serials reported as being preserved.

3.3. Extending the Scope to Include Digitized Journals

At the start of the PEPRS project there were open questions about ISSN assignment for content that had been digitized from printed serials — typically this was older, out-of-print, material that sat on shelves in research and university libraries. The ISSN Network had a long-standing policy dating from the days of microform and print reproductions that the ISSN of the print original was applicable to these 'reproductions,' as they were termed. Accordingly, the initial intention in the first phase of the project was to limit the project scope to material that was issued electronically by publishers.

During the second phase of the project, after 2010, contact was made with HathiTrust to investigate how its archiving activity could be included within the preservation registry. With access given to a log of ingested content, members of the project team carried out preliminary analysis of the metadata for the volumes of digitized content. This suggested that there were as many as 250,000 serials represented, of which the vast majority (largely because the issue of the material pre-dates the assignment of ISSN, did not have an ISSN present for even the print version) were presumed unassigned. This was a potential roadblock for enrolling HathiTrust into The Keepers Registry.

As a way forward, a small initial subset of 3,820 serials was defined for report through The Keepers Registry. These included content which had an ISSN present in the metadata but excluded journals that were still being published and those not thought to be in the public domain. This gave HathiTrust confidence in the serials for which it could potentially provide access and allowed the project and The Keepers Registry to remain on track. However, knowledge of what was excluded also provided a stimulus to tackle head-on the matter of journals digitized from print within the project.

A debate across the ISSN Network began. Was the appropriate ISSN to be an ISSN for the electronic serial, an ISSN for the printed serial, or even a separate and new ISSN for the digitized serial? Were these 'digital reproductions' to be regarded as mere surrogates of the print or, being made digital and malleable, were they to be regarded as equivalent digital representations to the 'born digital?' What should and, in practice, could be done to implement the assignment of identifiers, both to assist knowledge of what was being preserved and to enable users to exchange and find content?

Digitization was not itself the only consideration. What also seemed to be significant was whether the current publisher was involved. When current publishers carried out retrospective digitization of their earlier issues it was generally impossible for third parties to distinguish the digitized from born digital. When digitization was carried out by libraries or by third parties (e.g., JSTOR or Google), several matters became problematic. Not least of these was, and is, that of huge scale, including the variety and multiplicity of organizations that might request assignment, not just a current publisher. This bears upon such practical questions as which national center would be delegated responsibility by the ISSN Network for assigning the ISSN.

The outcome of deliberation within the ISSN Network was recorded in an update in the ISSN Manual, as announced in the June edition of the ISSN Newsletter (http://www.issn.org/2-24139-The-newsletter.php?id=40). The 2012 version of the ISSN Manual, which for the first time contains examples of ISSN records, incorporates new rules and new criteria for ISSN assignment, in particular for electronic resources and digitization of long-ceased publications. The scope of ISSN assignment was widened to the following: "reproductions in a medium different from the medium of the original edition, including digital media, are eligible for ISSN assignment" (Section 2.2.6). Also included is the rule that "A single ISSN is assigned to identify all online versions made available under the same title including: versions digitized from print [and] born digital versions" (Section 2.2.3).

This means that the same ISSN applies to all digital versions of an online resource. Separate ISSN are assigned to continuing resources published in different media, such as print. This seems implicit recognition that the digital is becoming the mainstream.

The rules on the assignment of ISSN to digital reproductions of ceased print serials are in Section 0.6.4 of the ISSN Manual. "As a general policy, National Centres should assign at the same time an ISSN to both the digital reproductions and to the original print versions when the latter are not already identified." The rules include circumstances when an institution such as a library or an archives provider offers a digitized version.

Those rules are interpreted here as meaning that if an ISSN has already been assigned to a serial issued as an online version, then the digitized content would take that ISSN. If no ISSN were to exist for an online version of that serial, a new ISSN would be assigned, whether or not an ISSN had been assigned to the serial for the printed issues.

Even with ISSN already assigned, all was not smooth sailing when processing the metadata on digitized content from HathiTrust. The good news, as already noted, was that the ISSN-L could be used to recognize those ISSN relating to the printed editions of the serial. Moreover, those serials with print ISSN but no ISSN for the online version could be among the first candidates for testing the new assignment

Title	ISSN	Publisher	Current extent of archiving ℹ	Archiving agency
1. Serials review	1879-095X (Online);	Elsevier	Preserved: v. 35-present	CLOCKSS Archive
	0098-7913 (Print)	Elsevier Science	Preserved: v. 1-20, 22-37	e-Depot
		Elsevier	Preserved: v. 7-30, 32-37	Portico

Page 1 of 1 ⬅ ➡

Figure 3. Archived content for *Serials Review* as reported to The Keepers Registry.
(Taken from the Beta release of The Keepers Registry on 12/09/12.)

rules. What remained challenging was the unstructured metadata for holdings attached to the digitized content; this had been transcribed directly from what was written upon the spines of bound physical volumes.

3.4. Recording the Extent Preserved

It was decided at an early stage in the project not just to record which serial titles were receiving the attention of the archiving organizations but to go a step further and require and record the actual volumes of content that were being preserved.

Perhaps a good way to illustrate some of the challenges of 'holdings' addressed in the second phase of the project is to show how the e-journal issues of *Serials Review* are being archived. This is a work in progress both for the archiving organizations and for the team behind The Keepers Registry (Figure 3).

Three archiving organizations are engaged with the publisher of *Serials Review*, ensuring that the content issued in digital format is being kept for the future. This provides a high level of confidence especially as the three have very different technical and organizational approaches. However, this also highlights the extent to which a given e-journal is not 'well preserved.' Volumes 1 through 37 are all being preserved but volumes 1 through 6 and volume 21 may only be in the care of one archiving agency; a volume-level update from CLOCKSS is awaited at the time of writing. Looking forward to the future, collaboration between the archiving organizations might lead to an arrangement to complete the run of archived volumes at each archive, thereby mitigating risk by providing greater resilience in preservations.

3.5. Issues About Issues

The finer granularity below that of volumes is at the issue level. This level reveals more about the extent of a serial being preserved and about the problems of format for recording and displaying that extent.

As illustration, the unsorted 'holdings statement' in Figure 4 displays a number of missing issues within the (as yet unsorted) volumes. Examples are issue 4 of volume 12 and issue 2 of volume 13. This prompts the following thoughts: the possibility that these were 'never issued' (plausible because of war, etc. but unlikely); that

these had not yet been ingested by the archiving agency; or the chance of a processing error either at the archiving organization or at The Keepers Registry. As stated earlier, this is a Beta release, so not too much should be made of the actuality of the screen shot shown in Figure 4.

It should be noted that the metadata that comes from archiving organizations ingesting digital content from publishers does not necessarily contain information about dates. Volume and issue seem to be considered the essential elements of serial metadata for processing and noting 'who holds what.' The date may have to be sourced from some external list in order to provide what is generally regarded as a very valuable attribute.

On the matter of format, another point of note is that the entry '3–4' for volume 27 is a transcription of that particular 'value' in the metadata. This value may be for a 'dual issue' volume but it also highlights the more general topic of the buckets of digital content that are ingested and reported by the archiving agencies. This is particularly significant for the organizations that are managing long-term archiving for digital content from digitization of print copy. The following report of the analysis of the metadata from the digital content ingest by HathiTrust illustrates that challenge further (Figure 5).

What is being preserved at HathiTrust is the digitized content of printed serials that were bound and sat upon shelves. Sometimes this means an odd collection of issues, even including more than one serial. The bound material was taken down, taken apart, and digitized. What is recorded as the metadata for a given bucket of digital content is what was written on the spine of the bound material. For example, the spine label can relate directly to a given volume, as with "v.33 1901," but as seen in "v.2 (1860–70)," there is a strong suggestion that more than one volume is included.

3.6. Holdings Statements

Even in days of print, many associated with serial catalogs would tell of the horrible history of the holdings statement. "Highly variable and mostly poor" was the judgment when holdings statements for the leading research libraries in the UK were viewed from the perspective of the UK union catalog of serials (Burnhill et al., 2004). Not easily read by anyone other than a trained cataloger, their

Preserved: v.33(1, 2, 3, 4), v.34(1, 2, 3), v.35(3, 4), v.36(1, 2, 3), v.37(1, 2), v.23(1, 2, 3, 4), v.27(1, 2, 3-4), v.7(1, 2, 3, 4), v.8(1, 2, 3, 4), v.9(1, 2, 3, 4), v.10(1, 2, 3, 4), v.11(1, 2, 3, 4), v.12(1, 2-3), v.13(1, 3, 4), v.14(1-2), v.15(1, 2, 3, 4), v.16(1, 2, 4), v.17(1, 2, 4), v.18(1-2, 3, 4), v.19(2, 4), v.20(1, 2, 3, 4), v.21(1, 2, 3, 4), v.22(1, 2, 3, 4), v.24(1, 2, 3-4), v.25(1, 2, 3, 4), v.26(1, 2, 3, 4), v.28(1, 2, 3, 4), v.29(1, 2, 3, 4), v.30(1, 2, 3, 4), v.32(1, 2, 3, 4)

Figure 4. Report of Issues of *Serials Review* Archived in The Keepers Registry (unsorted volumes).

27.	Journal of the American Geographical and Statistical Society	1931-0854 (Online); 1536-0393 (Print)	John H. Schultz & Co.	Preserved: *Summary of extent not yet available*	HathiTrust
			American Geographical Society of New York [etc.]	Preserved: *Summary of extent not yet available*	HathiTrust

	Earliest date of coverage	Most recent date	Publisher	Current extent of archiving ⓘ	Archiving agency	Updated
1.			John H. Schultz & Co.	Preserved: v.2 (1860-70)	HathiTrust	01 Jul 2012
2.			American Geographical Society of New York [etc.]	Preserved: v.33 1901; v.26 1894; v.13 1881; v.4 1872; v.1 1859; v.37 1905; v.34 1902; v.23 1891; v.21 1889; v.15 1883; v.10 1878; v.5 1874; v.42 1910; v.29 1897; v.11 1879; v.14 1882; v.8 1876; v.35 1903; v.18 1886; v.39 1907; v.25 1893; v.7 1875; v.32 1900; v.16 1884; v.9 1877; v.17 1885; v.22 1890; v.30 1898; v.36 1904; v.8 1876 2nd edition; v.5 1873; v.3 1870-1871; v.27 1895; v.6 1874; v.12 1880; v.19 1887; v.28 1896; v.24 1892; v.40 1908; v.41 1909; v.20 1888; v.31 1899	HathiTrust	01 Jul 2012

Figure 5. Report of 'Volumes' Archived by HathiTrust in The Keepers Registry.

variable format meant that there was little prospect that the content of a given holdings statement could be read and acted upon by machine algorithm, thus making the prospect of interoperability nigh on impossible. As often as not these statements did not even reflect the reality of missing volumes that had been consumed by mouse or student.

In the immediate aftermath of online access, the defects of the holdings statement may have seemed irrelevant as the library acted as a customer for serial content. Searchers only needed to know the bibliographic summary information of the article sought; the start and finish dates of the library's license were sufficient in order to access the full text online. The presumption was that the publisher/platform had it all.

Perhaps what is being observed now is re-focus upon the responsibility of libraries for custody and stewardship for online digital content, not just a role as customer on behalf of their immediate patrons. Whether or not that responsibility is exercised directly and locally or via dependence upon an archiving organization, it is necessary to have reassurance of the extent of digital content archived — with metadata that indicates easily what volumes and issues may be missing.

This concern continues to be relevant as indicated from the renewed attention being given to print archiving, either to support de-duplication of holdings for economic purposes, as with the UK Research Reserve (UKRR) (http://www.ukrr.ac.uk/), or for preservation purposes, as with work associated with the print archiving preservation registry created by the Print Archives Preservation Registry (PAPR) project (http://papr.crl.edu/). In either case there is comparable necessity for reassurance of the extent of print content archived — with metadata that can also be read and acted upon by computer algorithm for management and reporting purposes.

3.7. Interoperability Standard for Issues and Volumes

Given the existence of rich metadata and identifiers at the serial level (e.g., the ISSN) and now for the article (e.g., the digital object identifier — DOI), the holding statement continues to suffer from "middle child syndrome" — overlooked with a sense of neglect. There needs to be a better way to record and represent what is being held, whether digital or in print format.

A key question is how to represent issued content in ways that are friendly for the machine as user (the need to parse for algorithm) in addition to the requirement to be amenable for consumption by humans. This provides a special challenge given the vagaries of how serial content can be issued. Initial thoughts favor a simple matrix

or arithmetic approach with the expected periodicity of a serial expressed in matrix notation supplemented by additions and subtractions about that norm to allow for 'special issues' and 'not issued.' This approach requires prototyping and testing. Then there are the challenges of representing issued content (and holdings) of serials that spread across title changes, especially for serials that also have a long history. These may not be recognized (or recognized in the same way) by all the different organizations involved and, therefore, variously represented according to the different cataloging rules and practices by libraries.

Of course, the problem of holdings is a large problem and not one that the project team working to deliver The Keepers Registry could hope to solve on its own. An insight on the extent to which this challenge is both shared and unresolved was given in presentations made at the American Library Association's (ALA) Annual Holdings Update Forum in, 2011, including one on PEPRS and The Keepers Registry (Burnhill, 2011).

As part of the effort to simplify and standardize the data ingest process, the project partners approached EDItEUR to see whether the ONIX (Online Information Exchange) for Serials family of Extensible Markup Language (XML) formats (http://www.editeur.org/17/Serials/) could be extended to handle specific e-journal preservation information requirements. This would have the advantage of providing publishers with ways to provide metadata that would assist the archiving agencies — and hence The Keepers Registry. It would also provide a standards-based framework for establishing the interoperability between the archiving organizations and The Keepers Registry.

EDItEUR agreed and set up a cross-sector working group that also included representatives from a number of the archiving agencies participating in The Keepers Registry. The ONIX for Preservation Holdings (ONIX-PH) standard has now been made available in draft form (v0.21) to offer a standard for notification and exchange of such metadata. It is undergoing testing (http://www.editeur.org/127/ONIX-PH/). The structure of the message format can be shown in a simple sample file (http://thekeepers.org/ONIX-PH/ONIX-LOCKSS-Sample.xml). Other organizations are encouraged to make similar files available.

3.8. Variability in Publisher Information

Among the many elements of metadata for a serial that can change over time, the change in publisher is not problematic for successful ingest of data from the archiving organizations into The Keepers Registry. However, good information about publisher would be very useful for the prospective users of the Registry. The information now

Title:	Journal of the Royal Statistical Society. Series A, Statistics in society
ISSN:	1467-985X (Online); 0964-1998 (Print)
First publisher:	Blackwell.
First year of issue:	1997 (Online)
Place of publication:	United Kingdom
Frequency of publication:	Three times a year
Other publisher information:	Blackwell Publishing Ltd; John Wiley & Sons, Inc.

Title:	Journal of autism and developmental disorders
ISSN:	1573-3432 (Online); 0162-3257 (Print)
First publisher:	Kluwer
First year of issue:	2002 (Online)
Place of publication:	Netherlands
Frequency of publication:	Quarterly
Other publisher information:	Springer Science+Business Media; Springer; Springer US; Kluwer Academic/Plenum Publishers; Kluwer Academic Publishers-Plenum Publishers

Figure 6. Two examples of publication information in The Keepers Registry.

available to The Keepers Registry is displayed for two e-journals in Figure 6.

For both examples, all but the last field comes from the ISSN Register. ISSN records must include the publisher at the time of the assignment of an ISSN. Until the relatively recent implementation of a repeatable publisher field, the ISSN Register did not have a mechanism for providing both the original publisher and up-to-date information on the current publisher. And, even with the new capability, keeping up with publisher changes will be a challenge for many ISSN centers. What is shown as "Other publisher information" comes from the archiving organizations. In the first example, there is a simple explanation in name changes due to mergers and acquisitions: in 2001 (Blackwell Publishing) and in 2007 (Wiley). In the second example there is transfer in ownership, from Kluwer to Springer, and a variety of name expressions. This is illustrated in Figure 7.

First, thoughts were divided amongst the project team about providing and maintaining publisher information: on the one hand, there was a wish to completely ignore the problem; on the other, there was a determination to establish our own 'publisher authority file.' However, neither was thought sensible although the problem could not be ignored if The Keepers Registry was to offer a browse facility on publisher.

There is no solution defined as yet, but there have been some positive signs. The TRANSFER Code of Practice (http://www.uksg.org/transfer/) is one possible future solution, which aims to establish a set of standards that apply whenever a journal is transferred from one publisher to another.

Publishers have also contacted staff working with The Keepers Registry to request changes in how their name is represented, both in respect to what they regard as their 'proper name' and the appropriate imprint

	Title	ISSN	Publisher	Current extent of archiving ⓘ	Archiving agency
1.	Journal of autism and developmental disorders	1573-3432 (Online); 0162-3257 (Print)	Kluwer Academic/Plenum Publishers	Preserved: v. 27-34	e-Depot
			Kluwer Academic Publishers-Plenum Publishers	Preserved: v. 27, 35-37	e-Depot
			Springer US	Preserved: v. 37-42	e-Depot
			Springer Science+Business Media	Preserved: v. 1-41	Global LOCKSS Network
			Springer	Preserved: v. 27-42	National Science Library, Chinese Academy of Sciences
			Springer	In progress: Not yet processed.	Portico

Figure 7. Differences in publisher information.

for their journal. This was important input but given that costs need to be kept low, the team behind The Keepers Registry could not afford to deal bilaterally with each and every publisher that wants some consolidation across all its different name expressions reported by the archiving organizations, often because the publishers differed in how they described themselves in the metadata supplied to each archiving organization. A more systematic approach is required.

3.9. Identifiers for Publishers

Until recently there was no universally accepted identification scheme for publishers, neither of books nor of serials. The continued concern about name variants for publishers was seen as one aspect of the more general quest for authority files for personal and corporate names. However, there is now a prospective solution with the emergence of the International Standard Name Identifier (ISNI), an ISO (International Organization for Standardization) Standard (ISO 27729), whose scope is identification for Public Identities (http://www.isni. org/). This builds upon work done for the VIAF, the Virtual International Authority File (http://viaf.oclc.org/) as the result of a joint project to lower the cost and increase the utility of library authority files by matching and linking widely-used authority files and making that information available on the Web.

Online Computer Library Center (OCLC), which implemented and hosted the VIAF facility, has had long association with investigating how to resolve International Standard Book Number (ISBN) prefixes to names of publishers of books, relating variant names and capturing various attributes of individual publishers by data mining (http://www. oclc.org/research/activities/publisherns.html) (Connaway & Dickey, 2011). This work gained leverage from the inclusion in the ISBN number assigned to each book edition of a publisher code.

As noted, serials change between publishers, and the ISSN contains no special coding: it is a 'dumb number' conveying no meaning other than to act as a reference to what appears in a list or database. ISNI is also a dumb number. However, although not all publishers are yet present in the ISNI database, there is large overlap between the publishers of books and the publishers of serials, so the inclusion of knowledge from VIAF and the OCLC work should benefit the world of serials through the operation and use of ISSN.

The purpose of ISNI is to assist disambiguation of the public identities involved throughout the creation, production, management, and content distribution chain. Use of the ISNI should avoid the tedium and error associated with matching and disambiguating text, thereby providing ease and accuracy of matching and also improving browse and search operations. Note that as with title, a variety of name expressions are to be expected for publisher, not just because of language differences.

Identification of a public identity is distinct from the roles that might be recognized in the ISNI database record. There is indication that EDItEUR might provide coordinating input for a domain-specific data model for the assignment of role based upon various use cases that illustrate the functional divisions of labor for the information value chain.

For example, it is not clear at present whether the role of 'issuing body' will be distinguished from that of 'publisher,' nor what other functional roles will be specified with respect to how the content of serials are made available digitally.

The governance for ISNI is provided by the International Agency (ISNI-IA), a not-for-profit UK company managed by a consortium comprising the Confédération Internationale des Sociétés d'Auteurs et Compositeurs (CISAC), the International Federation of Reproduction Rights Organisations (IFRRO), the International Performers Database Association (IPDA), the Online Computer Library Center (OCLC), ProQuest, and the Conference of European National Librarians (CENL) — represented by the Bibliothèque nationale de France and the British Library. ISNI-IA appoints Registration Agencies to provide ISNI-related services, such as the allocation of ISNI.

The ISNI is for more than publishers; it is for natural persons (like an author) and even fictional characters (like Peter Pan). For example, searching on the ISNI Portal at the ISNI Web site (http://isni.oclc.nl/) not only returned the ISNI number for the 'Nature Publishing Group' (0000 0001 2180 3855) but also for 'The University of Edinburgh' (0000 0001 2108 4395) and even for 'Edinburgh University Library' (0000 0001 0664 9204).

3.10. Metadata for Serials and their Issued Content

Many of the matters about serials related in this second story are for others to address, which once resolved will enable improvement to be made to The Keepers Registry in future years as service enhancements. Figure 8 shows a data model for serials, taken from project activity at EDINA, that attempts to include a minimally sufficient set of elements upon which good metadata are required.

As in The Keepers Registry, the serial with ISSN takes center stage, noting here the significance of the ISSN-L, the field in ISSN records that brings together various manifestations of the same title. However, it is not enough to know that a serial is being archived. There needs to be some statement by the archiving organizations on the volumes and issues that are actually being archived. What is special about the serial is the issue of parts. What is valuable is the metadata that enumerates volume and/or issue and hopefully provides information about date. It would be good to know how this metadata compares with what was 'ever issued;' sadly such information is not generally available, but perhaps it could be compiled using past data from Abstracting and Indexing (A&I) databases or indeed from electronic tables of contents.

As remarked, the object of desire for future readers, as with present readers, is the article, complete with DOI. There is presumption in The Keepers Registry that all articles for a given issue and volume are safely gathered in by each archiving organization. It might be helpful to have re-assurance from each Keeper that what was gathered corresponded to the table of contents for each issue and volume.

To the left in Figure 8 are the issuing body and the publisher. The ISNI will come into its own for organizations playing each role. For
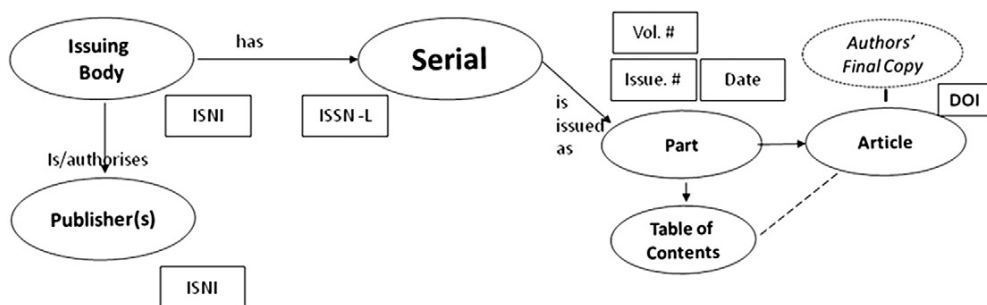

Figure 8. Scholarly statement via the traditional business model.

some serials the two roles are played by different organizations, as when a learned society outsources certain processes to a publisher. For other serials, such as *Serials Review*, the same organization plays both roles. Clearly it would be good to have an ISSN/ISNI table available as a Web-scale resource that represented the history of publishers for each serial.

4. The Third Story: Seriality of Literature on the Web

The thinking that led to the launch of The Keepers Registry had a sharp focus upon monitoring the archiving of e-journal content. This third story revisits the wider aspect of the shared purpose and task for libraries: to ensure continuity of access to the resources that scholarship requires, now and into the future. There is a need for thinking about other electronic content still at risk of loss: to consider whether there is value in the concept of seriality for that purpose; to reflect upon the utility of broadening the scope of The Keepers Registry, directly or implicitly as the ISSN might be assigned at the point of issue.

Serials are but constructs designed to serve everyday life; everyday life has been profoundly transformed by information and communication technology as it enables interaction with flows of linked digital objects accessed from afar. In many ways the context for this third story is set by historic game-changers: the significant properties of the digital medium, particularly its malleability, and what is understood about the 'telematic properties' of the Internet and the Web. The term 'telematics' (*tele*communications infor*matics*) gained currency as translation of 'telematique' in a report commissioned in 1976 by the President of France (Nora & Minc, 1978, 1980). Once a 'best seller,' it drew attention to a paradigm shift, "the crisis of civilization" in what still seems prescient: profound societal implications of this technology for a global economy: legal structures, deregulation, and new organizational models of production.

Thus far not much has been said in the two previous stories about digital preservation. In the first story, about the making of The Keepers Registry, the reader is directed to the Keepers themselves as the organizations with knowledge (and sometimes differing opinions) about archiving e-journal content. The second story was really about the metadata that helps the Registry be an effective monitor.

What The Keepers Registry seeks to do is provide evidence on the extent to which this holds good, although the Registry does not operate as an audit and certification scheme nor does it claim to be an authority on methods of digital preservation. The Registry is merely a lens onto policies that they have adopted and the practical outcome of their archiving activity.

The scope of this third story goes beyond what is now included in The Keepers Registry but may be indicative of what needs to be included. There is, therefore, need of a different frame of reference in order to appreciate what may be required for the Web sites and databases that section 0.3.2 of the ISSN Manual refers to as an "ongoing integrating resource [...] updated over time with no predetermined conclusion."

The story starts with some basics about digital preservation. This is followed by some comments about the notion of fixity and the 'copy of record' for a medium that is characterized by its malleability. This should serve to highlight at least two essential differences for serials librarians. What was traditionally acquired as 'published' was 'the copy of record,' and this had fixity by default through virtue of the marks on paper. All else was gray.

Citation and the search for what is cited is an essential task for authors and readers, for researchers and librarians, for students and for their teachers. This is visited at least twice in this third story: once in relation to citation and linking from within one e-journal article to another and again for what is cited from a journal article with intention later to link to a source on the Web. There is a third type of linkage in the compound character of new forms of scholarly statements and resources as these have been appearing on the Web.

4.1. Preserving the Digital

There is widespread agreement that "Digital information is best preserved by replicating it at multiple archives run by autonomous organizations" (Cooper & Garcia-Molina, 2002), as this is explicit planning for the inevitability of failure. This applies generally.

There is also general agreement that the Open Archival Information System (OAIS) reference model provides useful consensus in terminology for such an archive, an organization that undertakes responsibility for making information available for the long term. The term 'long' is regarded as long enough for there to be a concern about the impact of changing technologies. The OAIS model itself is technology-agnostic.

There is a good introductory guide to the earlier release of OAIS (ISO 14721:2003) by Brian Lavoie (2004); there is commentary on the recent (2012) revision as ISO 14721 in a blog post by Barbara Sierman (2012).

The OAIS model originates and is issued by the Consultative Committee for Space Data Systems at whose Web site is issued all official documents, (Consultative Committee for Space Data Systems, 2012a & 2012b) but it has received wide attention going well beyond space sciences. It is intended to facilitate understanding and increased awareness of archival concepts needed for long-term preservation and access to digital information. This includes digital information of all types and all types of encoding.

There are variants of the information package concept: the Submission Information Package (SIP), the Archival Information Package (AIP), and the Dissemination Information Package (DIP). The first is what is transferred and ingested from the 'Producer;' the second is the (AIP) version that is stored and preserved; the third (DIP) is the version of the information package that is delivered to the consumer in response to an access request. These may all differ from one another: in particular, what is rendered to the consumer (the DIP) and what is stored in an archive (AIP) may differ with respect to the format or amount of the content and the amount of metadata supplied alongside the content. On the ground, the AIP is sometimes called the preservation copy or the copy of record, and the DIP is generally called the access copy.

Digital information is easily lost or corrupted. For that reason several digital preservation functions are required. The Content Data Object within an Archival Information Package can be of any type (text, numeric coded data, video, computer program, etc.) and may comprise multiple files (e.g., a Web site with HTML files and JPEG files). The other part of AIP is the 'Representation Information' necessary for members of a given 'Designated Community' to render and understand the bit sequences constituting the Content Data Object. This last concept signals that only a specialist group may understand some of this digital content.

Preservation of digital information is critically important and arrangements for success are sufficiently different such that the basics should be re-visited from time to time. Some new understanding will emerge with new concepts and new recommendations for policy and practice. However, it is likely that OAIS will continue as a reference. It is also interesting to re-read some of the earlier thinking and synthesis (Waters & Garrett, 1996).

4.2. Digital Fixity & Copy of Record

Fixity was defined in the DCC Digital Curation Manual's section on preservation metadata as "the property of being unchanged between two points in time" (Caplan, 2006). This definition sums up a longer explanation of fixity in Waters and Garrett (1996) that includes the following: "The process of identifying and preserving a digital information object as a whole and singular work [...] also depends, for

instance, on the way that the content is fixed as a discrete object. If an object is not fixed, and the content is subject to change or withdrawal without notice, then its integrity may be compromised and its value as a cultural record would be severely diminished."

What is plain is that there is some difference between the fixity of marks on paper and the malleability of the digital medium. It has been argued that the print format of content had intrinsic fixity. "Easily read on paper and stored on library shelves, the printed journal has long served as the archival record of intellectual discourse" (Abrams & Rosenblum, 2003).

There are counter arguments against this claim for fixity of print and non-fixity of digital. Printed information could have been changed without easy detection. Change to information maintained in digital format can be accounted for through tools for tracking, systematic validation, or database schema that allow detection and evidence of change — whether by accident or design, authorized or malicious. Indeed, within the language set of OAIS, the phrase 'fixity of information' has restricted and specific usage for the validation of the authenticity or integrity of the Content Information: for example, a check sum, a digital signature, or a digital watermark.

The print in hand served two purposes, access and preservation. The conflation of purpose for printed matter, for access (being picked up and read by human) and for preservation (on a library shelf), may have attracted several to the Portable Document Format (PDF). This mimics the fixity of print by encapsulating the presentation as well as the content within a single digital object that is then assigned an identifier (e.g., a DOI). The PDF format, perhaps liked equally by publisher and librarians for its ready surrogate character, is less liked within the scholarly community as it has severe limitations for its re-use by software, both now and into the future. XML became much preferred as the format for the access copy because it enabled use by machines as well as rendering for use by humans.

Digital is not so easy to grasp. Among the several essential characteristics of digital content is its malleability: the ease, via software, with which digital content can be copied and edited, and how such content (and format) can be subject to manipulation and analysis. In particular, digital information can be rendered in a variety of formats for a variety of devices. This underlies different viewpoints about what might be the copy of record, how it should be identified, and what should be preserved for the long term: some specific source file or one of the many renderings and its format. The formats used for particular types of mobile devices are examples of current rendering and format that step beyond what is now mainstream on the Web. They will no doubt become redundant and replaced by others.

Abrams and Rosenblum (2003) were early in arguing the "use of XML as an archiving form [to] facilitate the long-term preservation and retrieval of e-journal content." There is a lot of literature about formats and preservation strategies that involve emulation, migration, and transformation of the Archival Information Package, more than can be adequately summarized here. However, it is worth noting a different approach, asserting "format obsolescence [to be] a rare problem that happens infrequently to a minority of unpopular formats" and that "the sine qua non is that the original bits be preserved" (Rosenthal, 2010).

What is important for this story is how this concept of fixity might be applied for the purpose of preservation and how that is recorded in metadata — in ways that would support a cultural record that serves the purposes of present and future scholarly and scientific enquiry.

4.3. Citation in Serials of Serial Content

The content that is cited as references within the pages of serials, whether in print format or electronic, presently does not itself form part of that serial content. Nevertheless, there seems little doubt that access to that referenced content is important and arguably it is vital for scholarship.

For access to cited content, scholars once looked to libraries, both at the time of publication of an article within a serial and subsequently. When libraries collected print serials, not all the referenced material that was cited in what a scholar read was conveniently on a shelf locally. Scholars, staff, and students had to resort to visits to distant libraries and requests to librarians for inter-library loan in order to track down what they wanted.

Terminology and habits change with the enabling technology: those scholars now visit by HTTP URL and make requests to an online user interface for download. Supply-side initiatives such as CrossRef (http://www.crossref.org/) greatly assisted these efforts by delivering the technical means, using DOI assigned at article-level at the point of issue, whereby online content from one publisher could be accessed when referenced from within the online content of another.

There used to be a presumption that a physical library somewhere had that printed copy (of reference). There is now implicit presumption that by archiving e-journal content there will be continuity of access — to the content that is cited. This presumption has two aspects: that the digital information will continue to exist (be preserved) and that there will be permission to access it (entitlement via license). There is additional and special doubt, examined in Section 4.6 below, now that so much of what is cited within an e-journal is from some other type of electronic continuing resource on the Web — illustrated perhaps by consulting the references in the articles of this issue of *Serials Review*.

This third story has started to put the spotlight again upon what is different about the digital and how these differences from the print have implications for continuity of assured access and long-term preservation. Some of these differences are generic to all digital content, and some are specific for electronic continuing resources, aka serials. However, the first part of this story has kept within the confines and comfort of the e-journal, including reflection on the cross-referencing from one e-journal article to another and confidence that cited content will exist (preserved) and can be accessed (entitlement).

The arrival of the e-journal has however triggered some differences for the mode of access by readers, as illustrated in the graphic taken from Burnhill (2009) shown here as Figure 9.

In this traditional business model for the e-journal, the author of the article and the reader of the article are connected via intermediation of the publisher of serials and licenses that the library had acquired, and paid for, as part of institutional arrangement. A number of additional value-added services emerged to enhance access, supplanting document delivery and inter-library loan (ILL). The point of public issue was through the publisher, so that role holds, but as noted earlier, libraries no longer take custody of the content of e-journals, and the reader goes online to get access to the article — the information object of desire. This was the context in which The Keepers Registry was designed and implemented, with the ISSN Register at its heart.

Now is time to consider other changes and their implications in the modes of scholarly communication of those who have become resident on the Web.

4.4. New Models of Scholarly Communication

Exchanges between peers in and across the disciplines have always taken place beyond institutional arrangement, reaching out into what is regarded as the invisible college of subject space. The journal of a learned society is a typical example of such a space. The author and the reader sometimes knew each other and met for face-to-face scholarly communication in seminars and conferences and the exchange of drafts and working papers. There grew a formal economy mediated by publisher and librarian. There was institutional payment by the library for license to the publisher, both for the outsourced role for journals of learned societies and increasingly for journals of the publisher's own design.
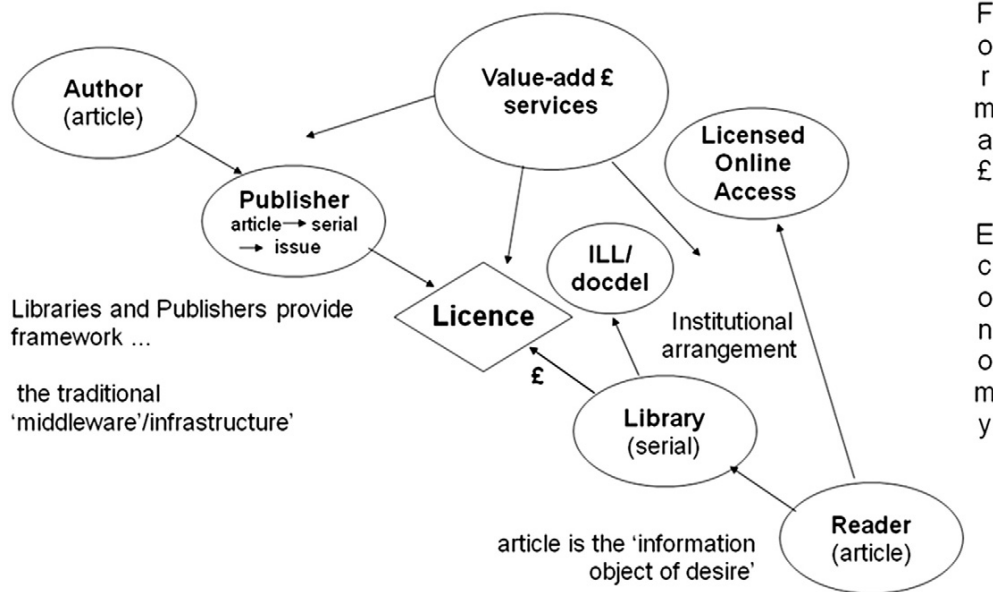
Figure 9. Changes in the issue of scholarly statement.

With desktop computers and the Web, the ease with which the author can issue scholarly statement and the ease with which the reader can discover and access statements that have not been mediated by publisher and librarian have increased remarkably over the past 15 years. This has included the ease with which the author assigns license, explicitly or implicitly, to the article which remains the object of desire for the reader, with the author continuing to seek reward by recognition among readers and peers, indicated variously and indirectly through downloads by readers and citation by subsequent authors. The following graphic (Figure 10), taken from Burnhill (2009), attempts to paint this into a picture.

Peer to peer communication increasingly occurs via subject repositories, which for some disciplines as places where content is first issued and made available. This is complemented by the provision of institutional repositories. What is on the Web is referenced and cited in support of scholarly analysis and statement in e-journals, as commented upon above. Scientific discourse is resident on the Web; it is a digital native. Much that is issued on the Web is issued nowhere else.

There can be little doubt that the Web is becoming a principal arena for scholarly communication as well as a dominant means to access the resources that are critical to scholarship. This includes greater facility for the reader to comment upon statement made upon the Web and therefore prospect of greater interaction between author and reader without having to be in the same place and the same time.
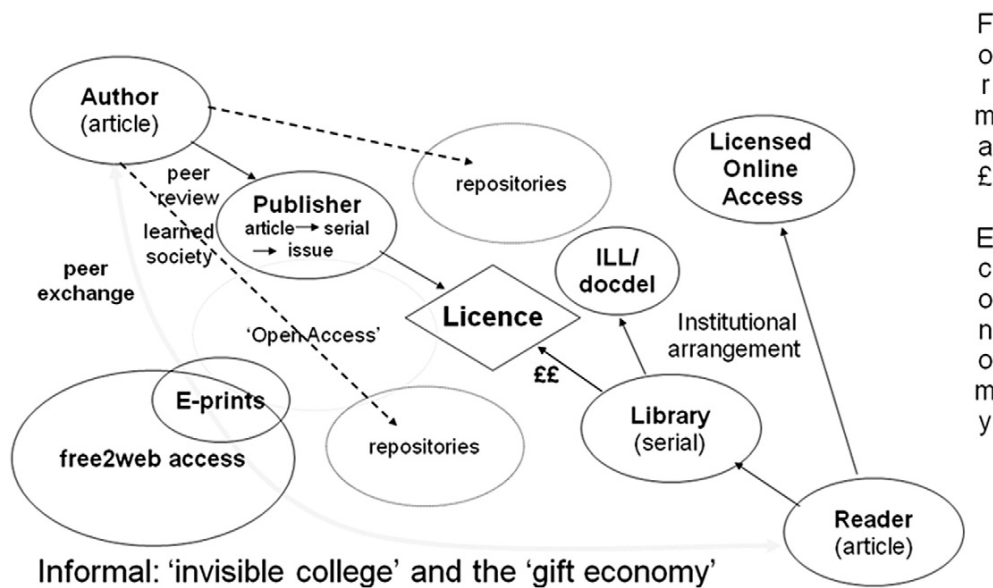
Figure 10. Basic bibliographic elements for serials.

4.5. Cataloging and Archiving the Web

The idea of cataloging what is on the Web to the same extent that print material was cataloged seems absurd. However, it is important for many purposes that the content that is issued on the Web is subject to some form of 'cataloging,' or at the very least there is assignment of identifiers for the points of issue. Assignment of ISSN for e-journals and for other electronic continuing resources that were issued in successive parts is well underway, and although there is no presumption that all that content is being successively preserved for the long-term, the design and operation of The Keepers Registry exists to monitor progress.

There is other content issued on the Web that is not issued in parts. Instead what is issued changes over time. It does not have fixity in either meaning of the term, both malleable in form and changing in informational content. The ISSN Manual notes, as part of the scope of resources eligible for ISSN, Web sites (strictly Web sites that update their content) and databases as examples of an "ongoing integrating resource [...] updated over time with no predetermined conclusion" (section 0.3.2). Editorial responsibility is stated as part of the inclusion criteria for the assignment of an ISSN.

It is imagined that those integrating resources that are assigned an ISSN will fall within the scope of The Keepers Registry. This means some prospect of monitoring especially of the extent of issued content that has been preserved. Note that this content is not issued in parts; there will be no enumeration and counting of volumes and issues. Time will be of the essence.

In many ways this shift to a broader focus should not be regarded as new. "Formal recognition of the need for bibliographic control over computerized information has slowly been evolving within the library and information science profession over the past several years," wrote Sue Dodd thirty years ago in the pre-Web era of the Internet (Dodd, 1982). She was writing about AACR2 Chapter 9 on 'Machine-Readable Data Files,' published in 1978. AACR2 Chapter 9 was renamed 'Computer Files' in the revision published in 1988. Her work briefly came to the surface in the UK (Burnhill and Templeton, 1989). (Then came an interlude that focused upon the new item in the hand, the microcomputer file.) Complete revision of chapter 9 saw it become 'Electronic Resources' in the 2001 amendments that were confirmed in AACR2 2002.

Of course the evolution of AACR2 Chapter 12 is also central to bibliographic description. AACR2 2002 also saw Chapter 12 on Serials renamed 'Continuing Resources,' driven by a wish to harmonize across AACR2, International Standard Bibliographic Description for Serials and Other Continuing Resources (ISBD(CR)), and ISSN but also by common belief in the usefulness of 'seriality' for what was, following widespread adoption of the Web, being recognized as important points of issuance of content. The term 'integrating resources' was used to signify what was updated over time (differing from serials that are issued in separate discrete parts).

The literature on Web archiving is considerable and increasing, as is the dimensionality of the problem space. The assignment of ISSN is itself a challenge. However, no one should presume that what is assigned an ISSN is being kept. On that matter of what is being kept, there is a useful, recent, and international overview by Ball (2010), with note due for the international leadership in the library world being offered by the International Internet Preservation Consortium (IIPC) (http://netpreserve.org/about/index.php).

The classic place for much large-scale Web archiving is the Internet Archive (http://archive.org). It is interesting to note that the Internet Archive not only seems to operate as a library, but in order to receive a particular kind of federal funding, it has been classified as a library (McCoy, 2007).

There are two interesting facilities that provide means of access to what has been archived in the Internet Archive: the Wayback Machine (http://archive.org/web/web.php) and Memento (http://www.mementoweb.org/about/). The latter, described in Van de Sompel, Sanderson, Nelson, Balakireva, Ainsworth and Shankar (2009), enables a Web browser to interface and navigate across the former. But one can only find that which was archived, and by that time the moment may have passed for such archiving.

4.6. New Scholarly Objects, with Links

Publishers and some archiving organizations are turning their attention to e-books, some of which could be regarded as issued in series, deserving of an ISSN and within the scope of The Keepers Registry. The same will probably be said of much that is marketed for mobile devices. Meanwhile, there are different forms of scholarly statement becoming increasingly common that may or may not fit our traditional way of understanding serials; whether they will come within the scope of The Keepers Registry depends upon whether they will qualify for an ISSN.

A useful summary pointer to more radical development is given in an article by Bechhofer, De Roure, Gamble, Goble and Buchan (2010) on 'Research Objects.' The authors ask, "If not traditional papers and volumes, what, then, should researchers be publishing?" They set out "a position paper or manifesto, outlining the principles and features of [their] approach" for these Research Objects: rich aggregations of linked content that have a number of essential properties including suitability for re-use by machine. These include two classes or "stereotypes" of direct relevance: Archived Objects and Publication Objects that "are intended as a record of activity, and should thus be immutable" and citable.

This builds upon an attempt "to distill some core characteristics of a future scholarly communication system" (Van de Sompel, Payette, Erickson, Lagoze & Warner, 2004) with both *registration* (and ultimately preservation) of a scholarly asset being central to its success within a workflow or pathway through various service hubs. Seriality of issuance as such is not utilized. Rather, there is determined focus in subsequent activity upon a "unit of scholarly communication" that is not "journals and their contained articles." This evokes what are referred to as *Compound units*. These "are aggregations of distinct information units that, when combined, form a logical whole" and can be represented in a manner (OAI-ORE) that enables them to be accessed and processed by machines and agents (Van de Sompel & Lagoze, 2007).

The manifesto described by Bechhofer, De Roure, Gamble, Goble and Buchan (2010) is also reminiscent of work by Hunter and Choudhury (2006) and Hunter (2006) that focus, respectively, upon the preservation of composite digital objects using Semantic Web Services and the use of Scientific Publication Packages (SPPs) for linking the raw data, their associated contextual and metadata on provenance, and derived information and knowledge such as publications. Focus areas for these authors are also on publishing and dissemination of scientific results and selective preservation of scientific data.

Going further back, context is offered by Michael Buckland (1998) who wrote of "two traditions, or mentalities, even cultures, that co-exist in the area of Information Science:

(i) Approaches based on a concern with documents, with signifying records: archives, bibliography, documentation, librarianship, records management, and the like; and

(ii) approaches based on finding uses for formal techniques, whether mechanical (such as punch cards and data-processing equipment) or mathematical (as in algorithmic procedures)."

These new scholarly objects have growing significance. Arguably, these objects have become the most important parts of the new information infrastructure for scholarship enabled by the Web. Further investigation is required. Much depends upon how their issuance is regarded and how the point of issue might be regarded with respect to seriality and the assignment of ISSN.

These new compound objects could be thought of as a multi-part article, but they go beyond the 'enhanced publication' with its 'supplementary data' (the data behind the graph), published within the 'traditional' journal article and addressed in a variety of ways by the archiving organizations. This is being addressed by the NISO/NFAIS Supplemental Journal Article Materials Project, which has been reporting in 2012. The technical draft regards Supplemental Materials "as a series of objects related to the article" (http://www.niso.org/workrooms/supplemental/).

4.7. Web Citation

The Web is dynamic. What was at the end of a given Uniform Resource Locator (URL) at the moment of citation can and does change. Indeed, it can also cease to be. What is moot is whether that cited content is now resting in an archive or has gone forever. One useful step forward is the editorial convention that when material is cited, the references contain the date at which the referenced content was accessed and was known to be current. But if that content is not to be archived then such convention has much less value.

An engaging overview of what is termed 'citation rot' is given in a talk by Van de Sompel (2011). Using Memento, referenced above, it is possible to access prior versions of Web resources available from Web archives and content management systems by using their original URI and a constructed 'date-time stamp' for the desired version, a bit like 'Time Travel for the Web.'

Memento was also used in a recent study by Sanderson, Phillips and Van de Sompel (2011) to examine the survival of Web-based content cited in articles in two scholarly repositories. The purposes of the study were to explore the extent to which Web-based content had been archived and what was no longer current. The study found that 28% of the resources referenced by the articles in the institutional repository had been lost, and 45% (66,096) of the URLs (in arXiv) that were found to still exist had not been archived.

There are plans being laid for work between a team at the University of Edinburgh, including EDINA, and the Memento Team at Los Alamos National Laboratory to carry out a large-scale investigation to inform recommendations for archiving cited Web content in a project funded by the Andrew Mellon Foundation called 'Time Travel for the Scholarly Web' (TT4SW). At present there is no mechanism in The Keepers Registry for the archiving organizations to report on the continuing existence of the cited content referenced in the text of the e-journals, and it is challenging to think how that might be done.

Van de Sompel et al. (2011) also notes that those who are responsible for the architecture of the World Wide Web promote independence between an identifier and the state of the identified resource. This encourages the re-use of the same URI even when content changes in order to avoid massive numbers of broken links. This resembles the decision-making about serials assigned an ISSN where the metadata for a serial has progressively had all of its parts replaced over time. Should it be regarded as the same serial and thereby maintain the same identifier? Perhaps what is sometimes known as 'the problem with grandfather's axe' (or Theseus' Paradox, in which every part of the Ship of Theseus is replaced and yet it is regarded as the same ship) is at the heart of the problems that must be addressed.

4.8. Points of Issue

The initial focus in the project was deliberately narrow: a preservation registry for e-journals that have an ISSN. As noted there is a larger scope to be reckoned. The qualities that are key for The Keepers Registry, so that the archiving of a given stream of content can be monitored and reported to the scholarly and library world, include seriality of issuance as well as the assignment of an identifier (e.g., ISSN) and the prospect of a copy of record being archived.

What is particularly interesting about the article on Research Objects cited above was how it was made available; it was issued as a reviewed conference paper in *Nature Precedings*. At first sight, *Nature Precedings* resembles a journal, but it is not. Launched in 2007 and closed in 2012, it acted as an open access preprint repository for the Life Science community. It was an integrating resource and as such assigned an ISSN, 1756–0357. Had it been archived, it could have been checked against the ISSN Register and so monitored by The Keepers Registry. Unfortunately, a search of The Keepers Registry reveals that the contents of *Nature Precedings* are not being archived. There is strong argument that having closed its contents are now more at risk than those that are still active.

There are many subject repositories emerging although they are not comprehensive of all disciplines and topic areas. One subject repository that is very active is the long running *arXiv* (http://arxiv.org/), first as the LANL (Los Alamos National Laboratory) preprint archive then moving to Cornell University with its founder, Paul Ginsparg. It was started in 1991 for physicists to share pre-prints. The subject coverage has since "broadened, first to cover most active research fields of physics, then to mathematics, nonlinear sciences, computer science, statistics and, more recently, to host parts of biology and finance infiltrated by physicists" (Ginsparg, 2011). This is a vehicle for the issue of significant content, as it "receives 75,000 new texts each year" and with "close to 700,000 full texts" [it] "serves roughly 1 million full-text downloads to about 400,000 distinct users every week" (Ginsparg, 2011).

Many of the texts within *arXiv* and other subject repositories and related facilities are pre-prints that are also submitted to journals for publication — working papers, for example. However, those that remain purely as e-prints and are never published in a peer-reviewed journal can be highly cited as scholarly statements. The level of citation is one measure of the significance of an article, with aggregation to the journal serving as some indicator of impact. Google Scholar Metrics can be used to compare the number of citations for 'venues' for issued content. On September 8, 2012, (prompted by personal e-mail sent the previous day by Herbert Van de Sompel) the h5-index (for articles published in the last five complete years) was observed to place *arXiv* fifth in the list for top venues — ahead of many well-known journal titles and just behind another subject repository, Research Papers in Economics (*RePEc*), for economics (http://repec.org/). This makes plain how significant the contents of these repositories are to peers that are authors.

With recent announcement of sustainable funding until 2017 (Cornell University Library, 2012), *arXiv* would appear to have become even more of a fixture. As yet, it does not have an ISSN. Its contents are currently archived at Cornell University Library with experimental use of the WayBackMachine in Internet Archive. This prompts a question about how its content should feature within The Keepers Registry. Assurance of archival activity for the content of those repositories is plainly very important.

What may also be problematic, or at least needs clarity, is the relationship and relative archival value of the publisher's final copy and that which is being cited in a repository as the authors' (sic) final copy — the default being more than one author. This prompted a presentation at a large international conference on open repositories on the topic, entitled "E-journal Preservation and the Archival Value of the Authors' Final Copy," delivered as a Pecha Kucha performance with 20 slides (each shown for 20 s) (Burnhill & Andrew, 2012). Originally intended as a conventional and serious conference paper, and perhaps one that that still needs to be written, the result can be accessed via The Keepers Registry blog (Guy, 2012a, 2012b) with links to *SlideShare* and an embedded video on *YouTube*. Of course, the transient character of those two online venues prompt much thought about long-term availability, although the intention of this citation is as ephemeral material for comic effect. But who can tell what might be of scholarly significance in some future time?

There is also the appearance of repositories, such as *figshare* (http://figshare.com/), that exist to make new forms of research output publically available, especially research data. And then there is the content of institutional repositories.

The moral of this third story is that while the short-term scope for The Keepers Registry upon e-journal content lends the simplicity of focus, there is a need to re-think strategy beyond the short-term: both for what should be done and how.

5. Epilogue

The three stories having been told, it is clear that their threads intertwine and each story has an open ending. A lot of ground has been covered, and yet some omissions must be noted. Thus far, there has been limited mention of another change that is occurring for scholarly literature and implications that might have for assured continuity of access: the Open Access (OA) agenda. The Directory of Open Access Journals (DOAJ) ensures presence of an ISSN as a condition of entry, and the content of those open access journals are candidates for inclusion by the archiving organizations and reporting via The Keepers Registry. There is also attention given in the Registry to 'triggered orphan content' categorizing the terms of access. The story about The Keepers Registry and open access must be for another time.

There is also omission in these stories of arrangement for post-cancellation access to back copy, even though that is part of the task of librarians when ensuring continuity of access. The focus here, and thus far for The Keepers Registry, has been on digital preservation and assurance for the medium and long-term with provision made in the Registry for report on terms and conditions of access.

5.1. Serial Issues

The second and longer story has been of a variety of matters about serials that have been encountered, and to some extent addressed, in the making of Keepers Registry. These have included identification (e.g., the use of the ISSN and the ISSN-L), challenges about holdings (to record the extent issued and archived), publisher identifiers and names (and their mapping onto serials), and what to do with digitized content of print journals.

A happy ending depends upon progress by many others and on several fronts. A lot of useful methodological work is being done, but much remains. The shift in focus from e-journals to serials in general and then to other forms of continuing resources and integrating resources will doubtless continue.

5.2. Archiving the New Literature of the Web

The third story told was about a new reality and a sense that the common notion that 'All will be digital and accessed from afar' is a half-truth that urgently needs to be taken seriously. It was about the land of research and development (R&D) in which several threads are left hanging.

A retreat to the higher ground with focus only upon what is published in e-journals is tempting but that would neglect the larger intent of providing assurance of access to scholarly resources. Not all is known and ready to be codified for what section 0.3.2 of the ISSN Manual refers to as an "ongoing integrating resource […] updated over time with no predetermined conclusion," on Web sites and in databases. The focus is widening. A shift in focus to include continuing resources and integrating resources is a likely but daunting prospect for The Keepers Registry.

Not everything on the Web changes, but there seems little doubt that the Web is changing everything. Archiving practice must surely regard that as the norm, as must identification and bibliographic fields for description and availability.

5.3. Mapping the Road Ahead for The Keepers Registry

These tales from The Keepers Registry began with a first story about the PEPRS project, which started in August 2008. The Keepers Registry Beta service resulted and was launched in October 2011. The Registry has been running successfully as an online service since then. In 2012 it received strong endorsement from the e-Journal **A**rchiving **I**mplementation **G**roup (JARVIG), which was established by JISC to determine community response in the UK (http://www.jisc.ac.uk/whatwedo/programmes/preservation/jarvig.aspx).

At the time of writing, The Keepers Registry remains a Beta service. It is under review for funding by JISC so that it might make the transition into a fully operational service.

The challenge now is to ensure that The Keepers Registry emerges as a robust and sustainable global monitor of archiving agencies, a lens onto archiving of digital serial content. The intention is to establish a Board of Governance with international representation and with representatives of research libraries and other parts of the serials business. The world's literature is of global interest, and this motivated presentation at the United Nations Organization for Education, Science and Culture (UNESCO) Conference on 'The Memory of the World in the Digital age: Digitization and Preservation' held September 26–28, 2012, in Vancouver, British Columbia, Canada, (http://www.unesco.org/new/en/communication-and-information/events/calendar-of-events/events-websites/the-memory-of-the-world-in-the-digital-age-digitization-and-preservation/). Proceedings will be published later, but the presentation itself is available now (http://www.issn.org/2-24096-Projects.php), which mixes metaphor by alluding to serials as streams of content and urging that not all eggs are placed in a single basket.

It is also necessary to respond positively to matters raised in the telling of the second and third stories. This requires a twin focus. On one hand, there must be renewed focus upon service quality and upon service enhancement, determining how to add more 'keepers' and enhanced functionality and engaging attention in multinational, multi-lingual, and multi-sectorial ways. The focus on service quality means refining processes and agreeing on common workflows and the basis for interoperability with each of the Keepers, the participating archiving organizations, with adoption of machine-readable format standards such as KBART and ONIX-PH. There is commitment to research and development activity by working with others best placed to contribute.

Some service enhancements, with development underway and envisaged for the future, are written into a 'roadmap' (http://thekeepers.blogs.edina.ac.uk/development-roadmap/). The roadmap is a plan and not a schedule — a sense of direction and a list of things that can be agreed as priority. It is published so that those who are interested may comment and have influence.

One milestone marked on the development roadmap for the Registry is the functionality, already mentioned, that would allow any library, or indeed any publisher, to upload a file with a list of the e-journals. That would allow a library to check the state of archiving for e-journals what are subscribed–or are available as Open Access – and to be able to run a match of the ISSN against the archival activity being reported into the Registry. Without The Keepers Registry and a bulk upload service, the Preservation Officer at Duke University had noted that it '*would have been an enormous undertaking*' … being able "*to include HathiTrust and CLOCKSS Archive, services in which we have memberships, in our review of agencies*" providing "*opportunity to tell publishers that we think this is an important part of the subscription service*" (taken from an interview with Winston Atkins, Preservation Officer for Duke University in Durham, NC on August 21, 2012).

How then to end? There are plenty of titles still to be archived. The big publishers are being engaged and often play a leading role in assisting archiving. There are very many volumes and issues, both

born digital and now digitized, that still need to be ingested. Both libraries and publishers need to prompt archiving organizations to be alert to what is missing. It is time perhaps to re-state a central proposition, summarized in a four-point mantra:

1. Assign an identifier at the 'point of issue' for a stream of digital content.
   - If it is worth preserving for the long term then it should have an identifier.
2. Ensure that (digital) content is archived routinely.
   - Have others/peers do that for you too; lots of copies keeps stuff safe(r).
3. Tell someone what you are doing (and how) and what you hold.
   - So all know what has been preserved and what is still at risk of loss.
4. Publish the terms of access for the archived content (now and when triggered as orphaned).
   - The purpose of preservation today is assurance of access tomorrow.

This is stated as a generality for scholarly resources. If it is worth preserving for the long term, then it should have an identifier. There should be routine arrangement for its archiving. There may be benefit in renewed thinking on the interpretation and usefulness of seriality for these data streams of issued content and economy in assigning identifiers for the source of the issue for Web content.

As repeatedly stated, the real heroes of these tales are the archiving organizations, those that are trusted as peers to act as stewards for scholarly content. There is prospect of a 'safe places network' in and around The Registry with associated social media arrangement for both peer-to-peer discussion by the Keepers and for wider engagement and communication. The Keepers Registry plays the necessary role of the international lens onto the archival activity for serial content — currently with the scope limited to that with an assigned ISSN. One is left to wonder how different will be the telling of the tales from The Keepers Registry in three, five, and twenty years' time.

Acknowledgments

References

Abrams, S. L., & Rosenblum, B. (2003). XML for e-journal archiving. *OCLC Systems and Services, 19*(4), 155–161. http://dx.doi.org/10.1108/10650750310508126.

American Library Association (ALA) (2011). ALA annual holdings update forum, New Orleans. http://connect.ala.org/node/151049

Ball, A. (2010). *Web archiving.* Edinburgh, UK: Digital Curation Centre (Retrieved from http://www.dcc.ac.uk/resources/briefing-papers/technology-watch-papers/web-archiving)

Bechhofer, S., De Roure, D., Gamble, M., Goble, C., & Buchan, I. (2010). Research objects: Towards exchange and reuse of digital knowledge. *Nature Precedings.* http://dx.doi.org/10.1038/npre.2010.4626.1.

Buckland, M. (1998). The landscape of information science: The American Society for Information Science at 62. [Preprint]. *Journal of the American Society for Information Science.* Retrieved from http://people.ischool.berkeley.edu/~buckland/asis62.html [The text varies slightly from the published version, Buckland, M. K. (1999). The Landscape of Information Science: The American Society for Information Science at 62. *JASIS, 50*(11), 970–974, found in a Special Topic Issue for the 50th Anniversary of the Journal of the American Society for Information Science, Part 1: The Journal, Its Society, and the Future of Print.]

Burnhill, P. (2009). Morning after the night before: Is the future what it used to be? *Association of subscriptions agents (and other intermediaries) annual conference* (Retrieved from http://www.subscription-agents.org/conferences/2009-conference)

Burnhill, P. (2011). PEPRS: Recording the extent preserved. *Presentation given at the ALA Holdings Update Forum, Universal and repurposed holdings information — Emerging initiatives and projects, June 25, 2011, New Orleans* (Retrieved from http://connect.ala.org/node/151049)

Burnhill, P., & Andrew, T. (2012). E-journal preservation & the archival value of the authors' final copy. *Open Repositories 2012 Conference, Edinburgh, Scotland* (Retrieved from http://www.slideshare.net/edinadocumentationofficer/burnhill-andrewor2012)

Burnhill, P., & Guy, F. (2010). Piloting an E-Journals Preservation Service (PEPRS). *The Serials Librarian, 58*(1–4), 117–126. http://dx.doi.org/10.1080/03615261003622742 Author Posting at http://edina.ac.uk/projects/peprs/docs/Author_version.pdf

Burnhill, P., Guy, F., & Osborne, N. (2007). Scholarly communication and national union catalogues: A strategic role for SUNCAT in the UK information environment. *New Review of Information Networking, 12*(1), 1–21. http://dx.doi.org/10.1080/13614570601133039.

Burnhill, P., Halliday, L., Rozenfeld, S., & Kidd, T. (2004). SUNCAT: A modern serials union catalogue for the UK. *Serials, 17*(1), 61–67 (Retrieved from http://uksg.metapress.com/content/c6y0ltjtxlhrgfr9/)

Burnhill, P., & Law, D. (2005). SUNCAT rising: UK serials union catalogue to assist document access. *Interlending & Document Supply, 33*(4), 203–207. http://dx.doi.org/10.1108/02641619510635669.

Burnhill, P., Pelle, F., Godefroy, P., Guy, F., Macgregor, M., Rees, C., et al. (2009). Piloting an E-journals Preservation Registry Service (PEPRS). *Serials, 22*(1), 53–59. http://dx.doi.org/10.1629/2253.

Burnhill, P., & Templeton, R. (Eds.). (1989). *Cataloguing computer files in the UK: A practical guide to standards. [Review draft]. Joint report to the Computer Files Cataloguing Group.* UK: Economic and Social Research Council.

Cantara, L. (Ed.). (2003). *Archiving electronic journals: Research funded by the Andrew W. Mellon Foundation.* Washington, DC: Digital Library Federation, Council on Library and Information Resources (Retrieved from http://old.diglib.org/preserve/ejp.htm).

Caplan, P. (2006). Preservation metadata. In S. Ross, & M. Day (Eds.), *DCC digital curation manual* (Retrieved from http://www.dcc.ac.uk/resource/curation-manual/chapters/preservation-metadata)

Collins, M. (2011). Open access literature review 2008–9: A serials perspective. *Library Resources and Technical Services, 55*(3), 138–147.

Connaway, L. S., & Dickey, T. J. (2011). Publisher names in bibliographic data: An experimental authority file and a prototype application. *Library Resources & Technical Services, 55*(4), 182–194.

Consultative Committee for Space Data Systems (2012a). *Reference model for an Open Archival Information System (OAIS): Recommended practice CCSDS 650.0-M-2.* (Issue 2. Retrieved from http://public.ccsds.org/publications/archive/650x0m2.pdf)

Consultative Committee for Space Data Systems (2012b). *Reference model for an Open Archival Information System (OAIS): Recommended practice CCSDS 650.0-M-2.* (Issue 2. Retrieved from http://public.ccsds.org/publications/archive/650x0m2.pdf)

Cooper, B. F., & Garcia-Molina, H. (2002). Peer-to-peer data trading to preserve information. *ACM Transactions on Information Systems (TOIS), 20*(2), 133–170. http://dx.doi.org/10.1145/506309.506310.

Cornell University Library (2012). *Ensuring arXiv's future.* (Ithaca, NY. Retrieved from http://news.library.cornell.edu/news/120828/arXiv)

Dodd, S. A. (1982). Towards integration of catalog records on social science machine-readable data files into existing bibliographic utilities: A commentary. *Library Trends, 30*(3), 3235–3361 (Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/7215/librarytrendsv30i3d_opt.pdf?sequence=1)

Ginsparg, P. (2011). ArXiv at 20. *Nature, 476*, 145–147. http://dx.doi.org/10.1038/476145a.

Guy, F. (2012a). The Keepers Registry's first Pecha Kucha outing is a winner! [Web log post] Retrieved from http://thekeepers.blogs.edina.ac.uk/2012/07/17/the-keepers-registrys-first-pecha-kucha-outing-is-a-winner/

Guy, F. (2012b). Pecha Kucha: E-journal preservation & the archival value of the authors' final copy. [Video file] Retrieved from. http://www.youtube.com/watch?

Hunter, J. (2006). Scientific publication packages — A selective approach to the communication and archival of scientific output. *International Journal of Digital Curation, 1*(1), 33–52. http://dx.doi.org/10.2218/ijdc.v1i1.4.

Hunter, J., & Choudhury, S. (2006). PANIC — An integrated approach to the preservation of composite digital objects using semantic web services. *International Journal on Digital Libraries: Special Issue on Complex Digital Objects, 6*(2), 174–183 (Retrieved from http://www.springerlink.com/content/p480171432571j25/?MUD = MP)

Jones, M. (2003). Archiving e-journals consultancy: Final report. *Report commissioned by the Joint Information Systems Committee (JISC)* (Retrieved from http://www.jisc.ac.uk/uploaded_documents//ejournalsfinal.pdf)

Kenney, A. R., Entlich, R., Hirtle, P. B., McGovern, N. Y., & Buckley, E. L. (2006). *E-Journal archiving metes and bounds: A survey of the landscape.* Washington, DC: Council on Library and Information Resources (Retrieved from http://www.clir.org/pubs/reports/pub138/pub138.pdf)

Lavoie, B. (2004). The open archival information system reference model: Introductory guide. *DPC technology watch series report 04–01, digital preservation coalition and online computer library center* (Retrieved from http://www.dpconline.org/docs/lavoie_OAIS.pdf Retrieved 2012-09-02)

McCoy, A. (2007). The internet gives birth to an 'official' online library. Post-Gazette.com. Retrieved from. http://www.post-gazette.com/stories/ae/cybertainment/the-internet-gives-birth-to-an-official-online-library-490707/

Morrow, T., Beagrie, N., Jones, M., & Chruszcz, J. (2008). A comparative study of e-journal archiving solutions. Tee Em Consulting. JISC collections. Retrieved from. http://www.jisc-collections.ac.uk/Reports/e-journal-archiving-comparative-study/

Nora, S., & Minc, A. (1978). *L'Informatisation de la société. Rapport à M. le Président de la République.* Paris: La Documentation Française.

Nora, S., & Minc, A. (1980). *The computerization of society. A report to the President of France.* English translation. Cambridge, Massachusetts: MIT Press.

Reynolds, R., Whitney, M., Chesler, M., & Beck, M. (2009). The tao of serials: 101 things non-catalogers should know about serials … or is it continuing resources? *The Serials Librarian*, *56*(1–4), 44–53. http://dx.doi.org/10.1080/03615260802662123.

Rosenthal, D. S. H. (2010). Format obsolescence: Assessing the threat and the defenses. *Library Hi Tech*, *28*(2), 195–210. http://dx.doi.org/10.1108/07378831011047613.

Sanderson, R., Phillips, M., & Van de Sompel, H. (2011). *Analyzing the persistence of referenced web resources with Memento.* Open Repositories 2011 ([Arxiv preprint]. Retrieved from http://arxiv.org/abs/1105.3459)

Seadle, M. (2011). Archiving in the networked world: By the numbers. *Library Hi Tech*, *29*(1), 189–197. http://dx.doi.org/10.1108/07378831111117001.

Shadle, S. (2006). FRBR and serials: An overview and analysis. *The Serials Librarian*, *50*(1–2), 83–103. http://dx.doi.org/10.1300/J123v50n01_09.

Sierman, B. (2012). OAIS 2012 update. Digital preservation seeds. [Blog post]. Retrieved from. http://digitalpreservation.nl/seeds/standards/oais-2012-update/

Sparks, S., Look, H., Muir, A., & Bide, M. (2008). *Scoping study for a registry of electronic journals that indicates where they are archived.* Loughborough, UK: Rightscom & Loughborough University (Retrieved from http://www.jisc.ac.uk/media/documents/programmes/preservation/ejournalregstudy.pdf)

Van de Sompel, H. (2011). *Time travel for the scholarly web.* [Video file]. Talk given at STM Innovations Seminar 2011 (Retrieved from http://river-valley.tv/time-travel-for-the-scholarly-web/)

Van de Sompel, H., & Lagoze, C. (2007). Interoperability for the discovery, use, and re-use of units of scholarly communication. *CTWatch Quarterly*, *3*(3) (Retrieved from http://www.ctwatch.org/quarterly/articles/2007/08/interoperability-for-the-discovery-use-and-re-use-of-units-of-scholarly-communication/)

Van de Sompel, H., Payette, S., Erickson, J., Lagoze, C., & Warner, S. (2004). *Rethinking scholarly communication: Building the system that scholars deserve. D-Lib Magazine*, *10*, 9 (Retrieved from http://www.dlib.org/dlib/september04/vandesompel/09vandesompel.html)

Van de Sompel, H., Sanderson, R., Nelson, M. L., Balakireva, L., Ainsworth, S., & Shankar, H. (2009). Memento: Time travel for the web. [ArXiv preprint]. Retrieved from. http://arxiv.org/abs/0911.1112

Waters, D., & Garrett, J. (1996). *Preserving digital information.* Report of the task force on archiving of digital information (no. 63). Washington, DC: The Commission on Preservation and Access & The Research Libraries Group.