



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## What EDINA Does: Ensuring Ease and Continuity of Access

### Citation for published version:

Burnhill, P 2015, 'What EDINA Does: Ensuring Ease and Continuity of Access' Paper presented at 10th International CALIBER 2015, Shimla, India, 12/03/15 - 14/05/15, .

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Peer reviewed version

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# What EDINA Does: Ensuring Ease & Continuity of Access

*Peter Burnhill*

## 1. Introduction

**1.1.** This paper has two parts. The first purpose is to describe EDINA and what it does. EDINA is a centre of digital expertise and service delivery. Playing a designated role for Jisc in providing online services and expertise to benefit research and education in the UK, EDINA forms part of the Information Services Group of the University of Edinburgh. Brief description of Jisc and the University of Edinburgh are included in order to provide the context in which EDINA operates, nationally in the UK and internationally. There is also an explanation of how EDINA came about as it approaches its twentieth anniversary.

**1.1.1.** This is all addressed in Part A of the paper.

**1.2.** The second purpose of the paper is to highlight threats to the integrity and continuing access to the scholarly record and to describe actions being taken. The challenge to the integrity of the scholarly record is one to be shared across research universities. It is an international matter, recognizing that research literature is not limited to what is published in one's own country. Together we need to devise sustainable infrastructure and to prompt action that results in successful, sufficient and timely archiving of all that constitutes the record of scholarship.

**1.3.** Many of the issues raised in this paper were discussed at the Coalition for Networked Informa-

tion (CNI) Membership Meeting in December 2014. That included a summary presentation that covered the activities reported here.

**1.3.1.** This is addressed in Part B of the paper.

## PART A

## 2. EDINA, Jisc and The University of Edinburgh

### 2.1. EDINA

**2.1.1.** EDINA develops and delivers world-class online services and expertise that benefit research and education in the UK and beyond. We seek ways to enhance research and education in the UK. For researchers, students and their teachers this means enhancing their productivity with services that both inspire and save time, helping to make the imagined possible! For universities, colleges and other learning and skills organizations, this means helping them succeed more effectively in their mission, improving outcome and increasing impact within limited budgets.

**2.1.2.** Following designation as a national data centre in 1995, as part of an open competition across all universities, EDINA was launched with a UK remit twenty years ago on 25th January 1996. This drew upon competence gained through the University's Data Library, which was established in 1983/4, and the role played by the University in providing computing infrastructure across Scotland. This experience provided a connection into networks of university libraries both in the UK and internationally.



**2.1.3.** What is written here complements the summaries to be found on the EDINA website.<sup>1</sup> That includes the online version of the EDINA Annual Review which forms part of our formal accountability to Jisc and its stakeholders. EDINA provides a significant part of the contribution that Jisc makes as a champion for the use of digital technologies in UK university and further education and skills sector.

**2.1.4.** A more detailed account of what EDINA does is set out in section 2.4 below.

## **2.2. Jisc**

**2.2.1.** Jisc is a membership body owned by UK universities and colleges, having previously been the Joint Information Systems Committee of the UK higher education funding bodies. As indicated on the Jisc website, it has the remit to be the UK's expert body for digital technology and digital resources on behalf of higher education, further education and research. Since its foundation in the early 1990s it has played a pivotal role in the adoption of information technology to improve research, teaching and the student experience as well as institutional efficiency in the following ways:

**2.2.1.1.** Network and IT services – with access to the SuperJanet Internet backbone, computer security protection and ubiquitous 'eduroam' wi-fi access.

**2.2.1.2.** Digital resources – with a licensing consortium (Jisc Collections, through which institutions procure electronic journals and other information resources used for research and teaching), shared services to support libraries and institutional repositories, resource discovery tools, and access to digital archives, geospatial data, multimedia and open educational resources.

**2.2.1.3.** Advisory services - enabling universities to get the most out of using technology for teaching, research and administration, along with advising the sector on how to respond to new challenges and opportunities.

**2.2.1.4.** Selective research and development activity to help keep the UK apply leading-edge technology, uncovering promising ideas and translating new digital solutions from other sectors and contexts into products and services that benefit education and research communities and add real value.

## **2.3. The University of Edinburgh**

**2.3.1.** The University of Edinburgh is one of the world's leading universities and one of the largest universities in the UK. It was founded in 1583 by Royal Decree, shortly after the 'Toun Council' of Edinburgh established its first library around which the 'Tounis College' and then the University was formed. Today the University employs over 12,000 members of staff in academic and supporting roles.

**2.3.2.** The University of Edinburgh was proud to honour its first Indian graduate in 1876. With a liaison office in Mumbai, today there are 279 students from India studying at the University of Edinburgh. The University is also home to the Centre for South Asian Studies, the principal academic unit in Scotland dedicated to study of the Indian subcontinent, the new Edinburgh India Institute was established in 2014. In February 2011, the University signed a Memorandum of Understanding with Jawaharlal Nehru University (JNU) in Delhi to collaborate together. There is now a long list of other collaborations.

**2.3.3.** The University offers more than 500 undergraduate and 160 postgraduate courses to over 35,000 students each year. There are about 14,000 international students from more than 140 countries representing over 40% of the total who undertake postgraduate and undergraduate study. Two-thirds of the world's nationalities study at Edinburgh. More recently the University has launched significant online teaching initiatives, with approaching 2,000 students currently studying its online distance learning postgraduate programmes, and a total to date of one million enrolments for Edinburgh MOOCs.

**2.3.4.** The latest UK Research Assessment Exercise highlighted the University's place at the forefront of international research. Research income is £200m representing 27 per cent of total income. It regards its professional services as critical to its success as well as its world-class teaching, research and student facilities. The many achievements of staff and graduates include activities that have explored space, revolutionised surgery, published era-defining books, paved the way for life-saving medical breakthroughs and introduced to the world many inventions, discoveries and ideas from penicillin to cloning Dolly the sheep.

**2.3.5.** Having a broad subject base, the University of Edinburgh has 22 Schools spread over three Colleges:

**2.3.5.1.** Humanities and Social Science: rated 12th in the world (The Times Higher Education 2014/15 Ranking). It comprises the former faculties of Arts, Divinity, Law, Music and Social Science, the Moray House College of Education and Edinburgh Art College, and so has a very broad range of research activity. A leading role is provided by the Institute for Advanced Studies in the Hu-

manities. More than 1,000 scholars from 66 countries have held Institute fellowships since its foundation in 1969; up to 25 Fellows are in residence at any one time.

**2.3.5.2.** Medicine and Veterinary Medicine: rated 1st in the UK for medical research (by the Hospital-based Clinical Subjects Panel) with a long history as one of the best medical institutions in the world. Veterinary Medicine came 1st in the UK having made a joint submission with the Roslin Institute and Scotland's Rural College (SRUC).

**2.3.5.3.** Science and Engineering: one of the highest-ranked science and engineering groupings in the UK and a key player in many European and international research collaborations. For example, the School of Informatics has consistently been assessed to have more internationally excellent and world-class research than any other UK submission in Computer Science and Informatics, confirmed again in the latest REF 2014 results. It holds the Silver Athena SWAN award, in recognition of commitment to advance the representation of women in science, mathematics, engineering and technology.

**2.3.6.** The support that the University gives to EDINA and also to the Digital Curation Centre which also has a national and international remit and funding from Jisc, reflects the national and international role played by the University of Edinburgh.

## **2.4. What EDINA Does**

**2.4.1.** One way to review the services delivered at EDINA is presented on the EDINA website, organized in categories of services geared for use by researchers and students (Reference & Multimedia;

Maps & Data) and those geared for use by librarians, other academic support staff and by developers. A good summary is also set out in the EDINA Community Report which can be downloaded from the Web.

#### **2.4.2. Flagship services include:**

**2.4.2.1.** Digimap, launched in 2000, to provide access to the 'Big Data' of geospatial mapping data, from Ordnance Survey and other national agencies for geological, hydrographic and environmental map data. There are associated geo-portal and geo-tagging services and support for the UK Location Council and Scottish Government as part of the spatial infrastructure.

**2.4.2.2.** SUNCAT, the Serials Union Catalogue for the research community launched in 2006, aggregating holdings information for 100 research libraries across the UK. There are associated services supporting the use of OpenURL interoperability and long term access for the scholarly record. The latter include a global monitor on archiving activity for e-journal content (Keepers Registry) and a key 'Open Access' deposit facility for researchers worldwide without an institutional repository (OpenDepot.org) which is also capable of routing researchers to the webpage for their institutional repository.

**2.4.2.3.** Jisc MediaHub brings together high-quality video, image and audio online resources that have been copyright-cleared for use in education. This includes multimedia collections from Channel 4, ITN, Films of Scotland, Reuters, Getty Images, Imperial War Museum, Royal Mail Film Classics, Wellcome Collection and many more. Some of these collections are hosted at EDINA, others held on other web sites whose metadata and

'thumb nail images' have been aggregated. Some of these are licensed for educational use in the UK by Jisc, some are made available by other providers; some allow access to UK research and education only, others on 'open access'.

**2.4.3.** All of these services are 'free at the point of use', given the right credentials: some requiring licences and restricted for use within the UK; some are openly available. The uptake and use of services at EDINA continues to grow, as it has done consistently since 1995/96 when EDINA first began its part, leveraging value from the University of Edinburgh in which we are based for the wider UK academic community. As of September 2014, 130 universities within the UK (nearly 85%) and 235 colleges (nearly 60%) use at least one of the services delivered by EDINA, the large proportion of which are on behalf of Jisc.

**2.4.4.** The primary goal for EDINA is to enhance the productivity and quality of research, learning and teaching in the UK and beyond. EDINA does this through the development and delivery of services, digital infrastructure and tools. This means that EDINA is also engaged actively in project activity and in working collaboratively with others, nationally and internationally.

#### **2.5. History of EDINA, 1995/96 -**

**2.5.1.** Origins in Edinburgh University Data Library

**2.5.1.1.** The Data Library was set up in 1983/4. Researchers were looking to the University to acquire data resources they needed for their scholarship. This included small area statistics from the 1981 Population Census and a range of government survey datasets. In response, the Data Library was

formed. Geographic information was a focus from the beginning, not only for the population census but conversion of area-based agricultural census data to grid square estimates with detailed visualisation of land use across the UK. Collaboration with the Department of Geography saw the establishment of the Regional Research Laboratory for Scotland, focusing on quantitative techniques in the Social Sciences.

**2.5.1.2.** Then followed support for the libraries in Scotland's universities, the National Library of Scotland and the two major civic libraries of Edinburgh and Glasgow through SALSER, which could claim to be the world's first web-based union catalogue of serials in 1994/5. SALSER remains heavily used to this day, providing public access to important specialist serials collections.

**2.5.1.3.** Links to the research community were cemented by the RAPID project, which linked research activity to the output of other work funded by the Economic and Social Research Council. Uniquely, RAPID included not only conventional monograph and journal publications, but also new types of research output such as software, datasets and learning materials.

**2.5.1.4.** Such projects built up a wealth of knowledge within the Data Library, and a breadth of expertise which was to serve it well for the most important event in its history – designation as a national data centre in 1995.

**2.5.2.** The decision was taken that the national service should be launched in January 1996 with a new name, EDINA, that began with an E, to sit alongside B for BIDS (University of Bath) and M for MIDAS (University of Manchester).

**2.5.3.** EDINA's new services included online access to bibliographic indexes such as BIOSIS Previews, PCI, Times Index, INSPEC and Ei Compendex, as well as UKBORDERS, which filled a gap in census, political and postal boundary data.

**2.5.4.** EDINA had earlier gained knowledge about digital preservation from the engagement of the Data Library with IASSIST, the international association for data librarians and data archivists serving the needs of the social sciences. Staff at EDINA also played a leading role in establishing the Digital Curation Centre. That was joint activity with the Universities of Glasgow and Bath and the Science and Technology Facilities Council, on the basis of collaboration across this University with the School of Informatics and the National eScience Centre (NeSC).

**2.5.5.** EDINA has represented the University of Edinburgh as a partner in Controlled LOCKSS (CLOCKSS) and continued to support the Data Library, both delivering data services into the University and advancing research data management as part of the University of Edinburgh.

## **PART B**

### **3. Ensuring Continuity & Integrity of the Scholarly Record**

#### **3.1. Our Essential Task**

**3.1.1.** A key task, and perhaps the essential task for research libraries, is to ensure that researchers and students, now and into the future, have both ease and continuity of access to the record of scholarship. There is an associated responsibility to ensure the integrity of the scholarly record.

**3.1.2.** Ease of access has been radically improved as so much of that has become available

online. However, as unforeseen consequence, the first important 'take home message' for all librarians is that very little of the scholarly record in digital format is now in the direct custody of academic libraries. What was once on-shelf locally is now online somewhere, without the assurance that it will be available thereafter. Moreover, what may be cited as a resource on the Web at the time of writing may have changed at the time of reading by another, or simply have disappeared without trace.

**3.1.3.** The second 'take home message' is that there are several more threats to the existence of significant parts of the scholarly record. These are being quantified, with evidence presented here.

**3.1.4.** Fortunately, as a third take home message, is that there are actions that are being carried out to provide remedy. Jisc and EDINA are involved in some of these activities on behalf of UK research libraries but necessarily with colleagues in other countries and with the publishing sector. The fourth take home message is that this is a 'trans-national' challenge: researchers and students in any one country are dependent on what is written and published in a country other than their own.

**3.1.5.** The main focus of the actions thus far has been on research literature that is published in e-journals and in related serials. This is partly because this literature is significant but also because there is economy in thinking about 'streams of issued content', noting that ISSN is now assigned to some updating websites (as 'integrating resources'), where content changes over time, as well as those continuing resources where content is issued in parts.

**3.1.6.** This has included the Keepers Registry, a global monitor of digital archiving. It is very important that we can all know who is looking after what,

and what might remain at risk of loss. That can be used as the source of statistics on the progress being made, and the evidence-base for assisting librarians focus attention on what should be prioritized for archiving. Information on the background and on many of the issues involved are included in a review article and a presentation given to the Library of Congress, available in the University's institutional repository.

**3.1.7.** More recent work, in a project called Hiberlink, has been on what is termed 'reference rot' for the web-based content increasingly referred to in citations made in scholarly statement: in theses, as well as in research articles.

### **3.2. The Keepers Registry**

**3.2.1.** The purpose of the Keepers Registry is to record and report what is known of archival activity for online serials across the world. It was the main outcome of a Jisc-funded project carried out by EDINA and the ISSN International Centre (ISSN IC). The ISSN IC forms the hub of the ISSN Network of national centres in individual member countries around the world. The ISSN IC publishes the ISSN Register which records all issued ISSN, including over 132,000 assigned to online continuing resources.

**3.2.2.** Now operated and funded at EDINA as a Jisc Core Service, the Keepers Registry acts as a global monitor with three main purposes:

**3.2.2.1.** To enable librarians and policy makers to find out who is looking after which e-journal, how and with what terms of access

**3.2.2.2.** To highlight the e-journals which are still at risk of loss

**3.2.2.3.** To act as a showcase for the organisations (the keepers) which operate as digital shelves for access over the long term.

**3.2.3.** In practice this means the aggregation and combination of metadata on serials (from the ISSN Register) and of metadata on archiving organizations, including not only the extent of their archival activity with respect to serial content, but also their policies. Provision is made for archiving organizations to state the outcome of third-party audit and certification; the Registry reports what each archiving organization asserts, without itself carrying out audit or certification.

**3.2.4.** There are divergent views about the extent to which digital preservation requires more than archiving of the digital bits, and other matters which have not been discussed here. However, if there is no archiving there can be no prospect of long term preservation.

**3.2.5.** The ten 'keepers' that currently report into the Registry are of three types:

**3.2.5.1.** International web-scale organisations: CLOCKSS Archive & Portico

**3.2.5.2.** National libraries: British Library, e-Depot (Netherlands), Library of Congress, National Science Library, Chinese Academy of Sciences

**3.2.5.3.** Consortia of research libraries: Global LOCKSS Network, HathiTrust, Scholars Portal, Archaeology Data Service.

**3.2.6.** A summary description of the approach to ingest and to digital preservation for each keeper is available on the Registry, together with information on access conditions for the journal content that they hold. There is a wide variety of business model and approach across the archiving organisations.

CLOCKSS and Portico derive income from the publishers who pay to have their content ingested and preserved for the long term, there is also income received from libraries. CLOCKSS maintain an open access policy for triggered content; Portico limit benefit to libraries that subscribe. The economics for national libraries is based upon national mission, sometimes underpinned by legal deposit legislation for voluntary or compulsory deposit of content of what is published in that country.

**3.2.7.** As noted, the Registry aggregates metadata to record the journal titles being preserved by each archiving organisation using the ISSN Register as a common data source and identifier for journal titles. This enables search on title and ISSN to discover archival status, including information on which volumes and issues are held by each archiving organisation.

**3.2.8.** There is also a Members Area in which a file of the ISSN may be uploaded for matching against information held in the Registry. This can be used by libraries to check the archival status of e-journals that they regard as important.

**3.2.9.** There is additional mechanism to enable archival status to be displayed in union catalogues and the like via machine-to-machine interface.

### **3.3. Evidence on what is archived, and what is not**

**3.3.1.** One role for the Keepers Registry is to provide some statistics on the progress that is being made to make sure that what is issued online (with an ISSN) is being kept safe. The good news is that, as at January 2015, over 26,700 serials are recorded as having volumes reported as 'ingested and archived' by at least one keeper of digital content.

**3.3.2.** That good news needs to be qualified. Inspection of information about individual titles will reveal that there are many 'missing' volumes. Moreover, over 132,000 ISSN that have been assigned to online serials and so that 27,000 represents only 20% success, even allowing that many are not for 'scholarly journals' but for other online continuing resources.

**3.3.3.** The serial lists from three large US university libraries were cross-checked against the information held in the Registry in 2011, in order to narrow the focus upon the scholarly record. The outcome of that exercise is presented in Table 1 (taken from the Serials Review article previously cited).

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

<sup>a</sup> 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.

**3.3.4.** A broad picture emerges: about one quarter of the titles of each (with ISSN) were 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. (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.)

**3.3.5.** Opportunity was also taken to examine the problem from the reader point of view, using the

logs of usage of the UK OpenURL Router which is middleware operated by EDINA for the UK's digital library, routing some 8 to 10 million requests for online resources from discovery facilities to publisher web sites (via what are called OpenURL resolvers). Of the 53,000 online titles requested by researchers and students from 108 UK universities in 2012, only 15% were being kept safe by three or more 'keepers'; over two-thirds were being held by none. Over 36,300 online titles appear to be at risk of loss.

**3.3.6.** Plans are being drawn up to repeat this analysis for usage data relating to 2013 and 2014, with a view to publishing this routinely on the Keepers Registry homepage.

**3.4.** The summary conclusion is that much remains to be done in order to ensure that the online resources needed by researchers and students are being kept safe. There is growing belief, which requires some systematic confirmation, that the titles that are being archived are those from the larger publishers and that (smaller) publishers of one or two online titles are not being engaged by archiving organizations. The long tail of very many journals issued by the many small publishers are at risk of loss. New strategies need to be devised to tackle this problem of the 'long tail'. This is not confined to Open Access journals, rather it is the lack of financial ability required to engage with the business models of CLOCKSS and Portico.

**3.5.** There may be an enhanced role for national libraries with the support and encouragement from consortia of research universities and their librarians; in the absence of action from a national library, universities may have to act on their own

especially as their interests extend beyond national boundaries.

### 3.6. Reference Rot

**3.6.1.** The scholarly record is under threat in another way that is being investigated in joint activity between the University of Edinburgh (EDINA and the Language Technology Group in the School of Informatics) and the Los Alamos National Laboratory Research Library. The Hiberlink project began in March 2013, funded by the Andrew W. Mellon Foundation, to examine a vast corpus of scholarly publication in order to assess and quantify the scale of the threat of 'Reference Rot'.

**3.6.2.** Reference Rot is the combined effect for a URL of 'link rot' (404 - Page Not Found) and 'content drift' (when what is rendered in a browser is dynamic and changes over time, or has changed completely as the URL now directs to a completely different website).

**3.6.3.** This recognises that resources on the Web that are cited as a supporting reference at the time of writing (or even when note-taking) are liable to have changed or even to have disappeared by the time the scholarly statement is read later by other researchers or acted upon for some policy or practical purpose.

**3.6.4.** Citations to material on the Web are now commonly found side by side with traditional references back into the scholarly literature. Citations are also made to software, datasets, websites, ontologies, presentations, blogs, videos, etc.

**3.6.5.** The initial focus of the project has been on 'short length' scholarly statement, with research activity resulting recently in the publication of a journal article. There is parallel work reported by re-

searchers at Harvard Law School who note reference rot in legal citations.

**3.6.6.** Opportunity has also been taken to examine the much longer form e-thesis that is prepared for award of a doctorate. The results of this have yet to be written up but is available as a presentation. (Arguably the e-thesis has greater importance for its author as well as taking longer in preparation and so giving greater opportunity for reference rot to occur.)

**3.6.7.** The empirical evidence of the threat of reference rot in various forms of scholarly statement, and therefore the scholarly record, is overwhelming. The longer the time lapse the greater the probability that the referenced content will no longer be at the end of the cited URI. There is less than a 50:50 chance that the cited content will have been archived by routine web archiving.

## 4. Actions to Meet the Challenge

**4.1.** Research literature is international – a researcher in any one country is dependent upon that which is written or published in another. Actions are required at a variety of levels. Jisc have made available funds over the next two years to support activity to facilitate cooperation between the archival organisations, to reach out nationally (to SCONUL and RLUK ) and internationally to encourage collaboration between other consortia of research libraries and their respective national libraries. This is assisted by links that EDINA and the Data Library have forged and by access to the ISSN Network in which over 80 countries participate. There has been comment that there should be outreach beyond the library community to learned societies, International Scientific Unions and associations of research-intensive universities.

**4.2.** An example of international web-scale archiving: CLOCKSS

**4.2.1.** The University of Edinburgh plays a leading role in supporting the CLOCKSS Archive Network. This is a globally distributed 'dark archive' with replicated copies of e-journal content held for the long term. EDINA plays a support role for the release of 'triggered' open access content.

**4.2.2.** The University was involved from the project stage in 2006. Vice Principal Helen Hayes was approached to join other leading research universities to engage with several of the world's largest scholarly publishers to build a global archive for the very long term. This was recognition of mutual responsibility to preserve digital scholarly assets for the good of the entire community.

**4.2.3.** The University's Information Services subsequently took on a role to act as one of twelve (12) geographically-distributed archival nodes in the CLOCKSS Archive on behalf of the wider international community. This is carried out by EDINA and IT Infrastructure, incurring minimum costs of hosting a closed server in secure conditions and contributing a subscription of \$15,000 per year. As a founding participant, the University is represented on the CLOCKSS Board of Directors, comprising twelve research libraries of long standing and twelve of the world's leading academic publishers, which oversees the not-for-profit company.

**4.2.4.** In the CLOCKSS business model (as with Portico) publishers pay to have their content ingested into a 'dark archive network'; signing over rights that their content may be triggered for access in the event that they can no longer supply. In the case of CLOCKSS the triggered content is made available for free, under a creative commons license, on one

or both of 'open access' platforms maintained by Stanford University Library and EDINA. A formal super majority vote by the CLOCKSS Board is required before such triggered release.

**4.2.5.** CLOCKSS deploys the 'Lots of Copies Keeps Stuff Safe' (LOCKSS) technology that automatically checks and repairs any detected 'bit rot' which inevitably will occur over the long term. Earlier this academic year the CLOCKSS Archive was certified according to the Trusted Repository Audit Criteria (TRAC), using a finding aid structured according to ISO16363. CLOCKSS reached the previous highest score and gained the first-ever perfect score for Technology.

**4.3. An Example of National Archiving: SafeNet (UK)**

**4.3.1.** This is a two-year 'service-in-development' project at EDINA commissioned by Jisc Futures. It is being carried out in collaboration with Jisc Collections and RLUK. The aim is two-fold: to clarify continuing access rights through use of an entitlement registry that will record subscription history and to develop the foundations of national archive infrastructure to host a UK collection of archived e-journals. This builds upon prior work. One is support for the UK LOCKSS Alliance, in which libraries (including the University's Library & Collections) collaborate to host copies of some e-journals; the other was a project to scope what would occur with respect to digital back copy should a library decide to cancel a subscription.

**4.3.2.** The project involves both consultation with the UK research library community in order to identify priorities with respect to content and, via Jisc Collections, negotiation with selected publishers to ensure clarity about terms of licence. The in-

tended outcome is to be a Jisc shared service for the UK research community, which will provide the means for permanent and trustworthy access to subscribed journal articles, post-cancellation. The objective is simplicity for the end user, underpinned by the ease of interoperability, should there be a cancellation or some other trigger relating to lack of access at the publisher site, with related tools and services in the digital library environment. To some extent this has already been achieved with trigger events by CLOCKSS in which EDINA provides an open access platform. Here the challenge is to restrict access according to a given library's 'entitlement'.

#### 4.4. Finding Remedy for Reference Rot

**4.4.1.** The Hiberlink project has also given attention to devising remedy for reference rot, to prevent or to minimize its occurrence. This has involved identifying opportunities for productive intervention in three basic workflows: for the author when note-taking and writing prior to submission; for the organization, typically an editor working for a publisher, overseeing review and issue; for the library in receipt of published/issued work. Thus far, using development work on annotating URIs, software engineering at EDINA has resulted in prototypes of tools which proactively archive 'DateTime-stamped' snapshots of webpages which are in one or more archives, returning an annotated URI for inclusion in citations. Thereby, what was cited at the time of writing (or at the moment of issue) both exists and can be read at some later time.

**4.4.2.** The Hiberlink project is due to complete at the end of June 2015, with focus now on dissemination of results and plans for moving from prototype to production quality tools to help avoid reference rot.

**4.4.3.** Reference rot features in an article in *The New Yorker*, first sub-titled 'What the Web Said Yesterday'; amusingly its content (and date) changed from first moment of issue.

**4.4.4.** As practical example of the remedy for reference rot being proposed by the Hiberlink project, here are three augmented references that include 'DateTime-stamped' snapshots of webpages:

```
<a href="http://edina.ac.uk/" data-versionurl="http://web.archive.org/web/20150219094431/http://edina.ac.uk/" data-versiondate="2015-02-19T09:44:32">
```

```
<a href="http://www.newyorker.com/magazine/2015/01/26/cobweb" data-versionurl="http://web.archive.org/web/20150219094636/http://www.newyorker.com/magazine/2015/01/26/cobweb" data-versiondate="2015-02-19T09:46:36">
```

```
<a href="http://blog.dshr.org/2014/12/talk-at-fall-cni.html" data-versionurl="http://web.archive.org/web/20150219094645/http://blog.dshr.org/2014/12/talk-at-fall-cni.html" data-versiondate="2015-02-19T09:46:45">
```

These augmented 'Robust Links' can be represented as:

Original URI: <http://edina.ac.uk/> <<http://edina.ac.uk/>>

Archive URI: <http://web.archive.org/web/20150219094431/http://edina.ac.uk/> <<http://web.archive.org/web/20150219094431/http://edina.ac.uk/>>

Archive timestamp: 2015-02-19T09:44:32

OriginalURI: <http://www.newyorker.com/magazine/2015/01/26/cobweb>

ArchiveURI: <http://web.archive.org/web/20150219094636/http://www.newyorker.com/maga->

zine/2015/01/26/cobweb

Archive timestamp: 2015-02-19T09:46:36

OriginalURI: <http://blog.dshr.org/2014/12/talk-at-fall-cni.html>

ArchiveURI: <http://web.archive.org/web/20150219094645/http://blog.dshr.org/2014/12/talk-at-fall-cni.html>

Archive timestamp: 2015-02-19T09:46:45

## 5. Concluding Remarks

**5.1.** The challenge is to devise sustainable infrastructure and to prompt action that results in successful, sufficient and timely archiving of all that constitutes the record of scholarship.

**5.2.** This challenge of digital preservation extends beyond serials to include ongoing 'integrating resources' such as databases and web sites. The record of scholarship has a fuzzy edge. This begs the question on the extent to which data generated by the research process are themselves part of the record of scholarship. It also involves considerations of what constitutes the copy (or copies) of records and notions of digital fixity.

**5.3.** Fortunately a start has been made over the past fifteen years with a growing number of organizations stepping forward to act as our digital shelves. EDINA is playing its part in developing and delivering technical infrastructure and online services. There has been benefit, as illustrated in the activities described in the paper, in working with an international agency for assigning identifiers for serial content and in successful collaboration with some of the leaders in the field, including the architects of LOCKSS and of Memento.

## References

1. <http://edina.ac.uk/>
2. <http://www.jisc.ac.uk/>
3. <http://www.ed.ac.uk>
4. <http://www.cni.org/events/membership-meetings/past-meetings/fall-2014/project-briefings-breakout-sessions/>
5. <http://www.slideshare.net/edinadocumentationofficer/long-cni2014-burnhill%20>
6. These are equivalent to the earlier University Grants Council and its Computer Board.
7. <http://www.jisc.ac.uk/content>
8. The University of Edinburgh is ranked 17th in the world by the 2013/14 and 2014/15 QS rankings
9. The list of formal partners in India is growing and includes: University of Calcutta, Commonwealth Veterinary Association, Department of Biotechnology Ministry of Science and Technology, University of Delhi, Indian Institute of Technology Madras, Indian Council for Cultural Relations, Indian Institute of Technology Delhi, Indian Institute for Science Bangalore, Jawaharlal Nehru University, Kerala Veterinary and Animal Sciences University, Maulana Azad Medical College, Tata Institute for Fundamental Research, United Theological College, National Centre for Biological Sciences.
10. A massive open online course (MOOC) is an online course aimed at unlimited participation and open access via the web.
11. <http://www.nutshell-videos.ed.ac.uk/category/colleges/college-of-humanities-social-science/>
12. <http://www.dcc.ac.uk/>

13. <http://edina.ac.uk/about/commrep/>
14. The author was appointed to take over full-time responsibility in 1984.
15. <http://edina.ac.uk/salser/>
16. The name EDINA comes from a poem by Robert Burns whose birthday is celebrated on 25th January each year. 'Address to Edinburgh', begins 'Edina! Scotia's darling seat!' The University's Library Collection includes a copy in his fair hand, <http://edina.ac.uk/about/addresstoedinburgh.html>
17. BIDS subsequently became ingenta in 1998.
18. MIDAS subsequently changed its name to Mimas, allegedly due to a trademark dispute, and has now been merged into Jisc.
19. <http://iasistdata.org/>
20. <http://www.dcc.ac.uk>. The author was its first director in 2004.
21. [http://www.clockss.org/clockss/ Home](http://www.clockss.org/clockss/Home)
22. <http://www.ed.ac.uk/schools-departments/information-services/research-support/data-library>
23. <http://datalib.edina.ac.uk/mantra/>
24. <http://www.issn.org/understanding-the-issn/assignment-rules/the-issn-for-electronic-media/> An ongoing integrating resource must meet certain inclusion and exclusion criteria to be eligible for ISSN assignment (ISSN Manual 0.3.2).
25. <http://thekeepers.org>
26. Burnhill, P. (2013). Tales from The Keepers Registry: Serial Issues About Archiving & the Web. *Serials Review* 39 <https://www.era.lib.ed.ac.uk/handle/1842/6682>
27. <http://hiberlink.org/>
28. <http://edina.ac.uk/projects/peprs/>
29. <http://thekeepers.blogs.edina.ac.uk/2013/10/28/generating-actionable-evidence-on-e-journal-archiving/>
30. <http://dx.doi.org/10.1371/journal.pone.0115253>
31. <http://harvardlawreview.org/2014/03/perma-scoping-and-addressing-the-problem-of-link-and-reference-rot-in-legal-citations/>
32. [http://edina.ac.uk/presentations\\_publications/reference-rot-and-e-theses.pptx](http://edina.ac.uk/presentations_publications/reference-rot-and-e-theses.pptx)
33. Society of College, National and University Libraries, <http://www.sconul.ac.uk/>
34. Research Libraries UK, <http://www.rluk.ac.uk/>
35. <http://clockss.org>
36. Klein M, Van de Sompel H, Sanderson R, Shankar H, Balakireva L, et al. (2014) Scholarly Context Not Found: One in Five Articles Suffers from Reference Rot. *PLoS ONE* 9(12): e115253. doi:10.1371/journal.pone.0115253
37. <http://www.newyorker.com/magazine/2015/01/26/cobweb>
38. <http://www.mikejonesonline.com/contextjunky/2015/01/21/rotters/>
39. <https://www.era.lib.ed.ac.uk/handle/1842/9394>
40. <http://blog.dshr.org/2014/12/talk-at-fall-cni.html>
41. <http://www.slideshare.net/hvdsomp/memento-101>

#### About Author

**Mr. Peter Burnhill**, Director, EDINA, Jisc/UK centre for digital expertise and service delivery at the University of Edinburgh (Scotland, UK).  
Email: [peter.burnhill@gmail.com](mailto:peter.burnhill@gmail.com)