Ensuring Continuity of Access to Resources for Scholarship

Peter Burnhill
Director, EDINA
University of Edinburgh

Ensuring Continuity of Access to Resources for Scholarship

Peter Burnhill
Director, EDINA
University of Edinburgh

Bio: I'm a data person turned into something else

With degrees in Economics and in Statistics, I began at the University of Edinburgh as a survey statistician in research centre in 1979, studying transition from school

Changed career in 1984 to set up the University Data Library
Then combined that with

• Co-director, Regional Research [GIS] Laboratory for Scotland, 1987/93
• Past-President of IASSIST, 1996 - 2001
  – International assoc. for data librarians and archivists
  www.iassistdata.org
• Director, Digital Curation Centre, 2004 - 2006 (Phase 1)
  – Digital preservation [conserving content] * data curation [adding value to data]
  www.doc.ac.uk
• Director, EDINA – JISC/UK national data centre, 1996 - present day
  – Online services for UK universities and colleges
  edina.ac.uk

The screen versions of these slides have full details of copyright and acknowledgements
Re-thinking stewardship for scholarly works

The central task:
• To ensure that researchers, students & their teachers have continuity of access to the online scholarly resources they need
• Digital preservation is crucial but need to keep focus on ‘continuity of access’
• Data is important in its potential for use as evidence

Emergence of digital library

• Mix of the document tradition (signifying objects & their use) and the computation tradition (applying algorithmic, logical, mathematical, and mechanical techniques to information management)
  – “Both traditions are needed. Information Science is rooted in part in humanities and qualitative social sciences. The landscape of Information Science is complex. An ecumenical view is needed”
• The digital library has words, numbers, pictures and sounds
  – Numeric data, online learning & teaching materials, digital pictures and other audio-visual materials
• What do researchers do?
• And what do they want/need of a digital library
  – That they cannot do for themselves?

Infrastructure to support four ‘demand-side’ verbs

<table>
<thead>
<tr>
<th>Verb</th>
<th>Description</th>
</tr>
</thead>
</table>
| Discover | Information object of interest  
  e.g. dataset, article referenced in database, etc. |
| Locate  | Organisation offering service  
  e.g. data centre or document delivery service |
| Request | Use of service  
  via open access, privilege of membership, payment of money |
| Access  | Object of interest  
  via online access, document delivery, personal visit |

Based on MODELS workshops (UKOLN/JISC eLib)
Ensuring Continuity of Access to Resources for Scholarship

Peter Burnhill

Scientific / scholarly record

- Science equates to more than what is published in refereed journals

  But

- Record of Science
  - Does contain what is published in refereed journals
- Record of Scholarship (including Science)
  - Contains what is published in journals (& books)

What's the problem for E-Journal content?

- First, the Good News!
  - Researchers and students now have online access to journal articles
    * To read & download: Any-where, Any-time …

- Next, the Bad News!
  - What is now in digital form may not always be available
  - Frustrates economic benefits of existing investment in digital
  - Not good for libraries, not good for publishers

Why worry about digital preservation?

- All that is now digital may not always be available
  - For a variety of reasons
    - Natural disaster
      - Earthquake, Fire, Flood
    - Computing failure
      - Digital decay
        - Bit rot, Format obsolescence
    - Human folly
      - Criminal/political action
      - Hacking
    - Commercial issues
      - Publishers ceases publication with no transfer
      - Publisher goes out of business with no transfer

The screen versions of these slides have full details of copyright and acknowledgements
Ensuring Continuity of Access to Resources for Scholarship
Peter Burnhill

Some consequences of Web
- Essentials of supply chain have changed
  - Licensed to access, not sale of content
- Libraries no longer take physical custody of much key content
  - Online remotely, not on-shelf locally
- Role of libraries as trusted keepers of information and culture has been disrupted
  - Need assurance of continuity of access
    - Of all content for future generations
    - Of the back copies, post-cancellation of the license
- Scholarly, cultural & intellectual heritage is at risk

How important are E-journals?
- 96.1% of Science journals are online
- 86.5% of Arts and Humanities are online
- 2006-2007 – 102,000,000 downloads
  - Up 21% from previous year
- 17% usage is at the weekend

Source: E-journals: their use, value and impact. Research Information Network. UK April 2009

* Try searching the Web to update this slide *

What’s the answer?
1. Think: Consider how we ensured continuing access to printed works over the long term
   - Human-readable format; relatively enduring media (paper)
   - Multiple copies held in multiple places (a network of libraries)

2. Think again: Understand what is different about the digital
   - Formats become obsolete; unseen digital decay (‘bit rot’)
   - Can easily be altered (authenticity), copied and transported (theft)

3. Propose: Develop digital preservation policy
   - Including practices that address threats & risks

4. Act: Implement policy & practices for global effect
   - Need to command consensus across stakeholders (Transparency)
   - Need to be sustainable, in organisational, technical & financial terms

5. Reflect: Test, monitor and report: Community & Transparency
In whom or in what should be put our trust?

We need more than faith

- We plan for the future, beyond our own working lives
- The digital is different: we can make ‘copies’, not ‘unique’
- Trust implies risk
- Not (just) trust in people but trust in procedures
  - Importance of fire drill procedure, and of fire drill practice

Life cycle models, provenance, authenticity, diplomacies

- “Life cycle models are shaping the way we study digital information processes... represent the life course of a larger system, such as the research process... Life Cycle Information for E-Literature (LIFE) and the Digital Curation Centre [ ] have been influenced by this approach. Similarly, data archivists are advancing life cycle models to improve practices in preserving research data.” Chuck Humphreys
- “…‘Life cycle’ is different from ‘life span’... implies an environment in which resources are... created, curated, made accessible, and preserved for subsequent research, learning, and policy activity. The challenge is to infuse the digital life cycle with the metadata and services that will enable access, evaluation and re-use over time.”
- “Archival science, with its principles of uniqueness, provenance, arrangement and description, authenticity, appraisal, and its tools sets such as diplomatics... [as] framework for a theoretical foundation for digital libraries.” Seamus Ross, Keynote Address ECDL2007, Budapest

Trust and the avoidance of disaster & deceit

- ‘Trust’ means more than ‘reliance’
  - Failure threatens more than disappointment: significant loss
    “We all make mistakes; the trick is being able to recover”
  - When we trust, or are trusted by others, we have interest in:
    – Methodology, that is explicit, testable and ‘proven’
    – Competence to do what is required
    – Commitment to do it, in a timely fashion
    – “Disclosure” of what is done, and then “audit” of what is said and done
Ensuring Continuity of Access to Resources for Scholarship

Peter Burnhill

Many schemes emerging to meet challenge

- But who is doing what?
  - How can libraries & policy-makers assess which e-journals are being archived, by what methods, and under what terms of access?

The keepers registry & how it came about

Aims to be a global online facility, with dual purpose:
1. Reporting who is looking after what e-journal (& how)
2. A showcase for organisations having archival intent

- Idea of e-journals preservation registry mooted in reports
  - M. Jones, 2003/4
- ‘Clarity of public statement by each agency or through a registry’
  - CLIR Report, 2006
- JISC commissioned scoping study that recommended an e-journals preservation registry should be built, possibly as part of UK Union Catalogue of Serials (SUNCAT)
  - Rightscom & Loughborough U, 2007

The keepers registry & how it came about (2)

Aims to be a global online facility, with dual purpose:
1. Reporting who is looking after what e-journal (& how)
2. A showcase for organisations having archival intent
   - Assurance that a publisher’s content is being preserved
   - Assurance that a library’s stewardship is enacted
- In 2008, EDINA & ISSN-IC became partners in JISC-funded project: Piloting an E-Journal Preservation Registry Services [PEPRS]
  - Phase 1: August 2008 – July 2010
    - ‘Investigate, prototype and build’ [evaluation in Feb. 2010]
  - Phase 2: August 2010 – July 2012
    - ‘Preparing for service & governance’
  - Preparing for Phase 3: August 2012

The screen versions of these slides have full details of copyright and acknowledgements
Ensuring Continuity of Access to Resources for Scholarship

Peter Burnhill

The Keepers Registry

http://thekeepers.org

Soon to add activity of National Science Library of China + 3 others waiting: from Canada, UK and USA

Piloting …

PEPRS

Project: Funded by JISC,

- Over two years, starting August 2008
  - Review after 18 months into prospect for move into service
- Partners: EDINA and ISSN International Centre (Paris)
  - Support of Governing Body and Directors of ISSN Network
- Purpose: Scope, develop & test a registry service
  - Establish and test an Information Architecture
  - Seek consensus across stakeholders
  - Technical & financial sustainability

PEPRS is a project funded by JISC

Joint Information Systems Committee

- JISC manages funding from all the UK government agencies responsible for higher and further education
  - “To provide world-class leadership in the innovative use of ICT to support education and research”
- JISC manages and funds more than 200 projects within 15 programmes; outputs and lessons are made available to the HE and FE community
- JISC also supports 50 Services that provide expertise, advice, guidance and resources to address the needs of all users in HE and FE
  - The three largest services are JANET(UK) - which oversees networking – and two national academic data centres, EDINA and Mimas, based respectively at the Universities of Edinburgh and Manchester

The screen versions of these slides have full details of copyright and acknowledgements
Ensuring Continuity of Access to Resources for Scholarship

Peter Burnhill

E-Journals

PEPRS
Scope: Journal and other serial content in digital format
- Focus on those serials with the ISSN identifier
  * If its worth saving, it should have an ISSN
Multi-level: article is the information object of desire
- Focus on Journal Title-level
- Issued Content, i.e. Volumes (Year), Articles
International:
- Matters for the UK
  * But matters to all countries
- Cannot be resolved in (national) isolation

Digital preservation agencies in the pilot

- Two 3rd Party Organisations
  - CLOCKSS
  - Portico
- Two National Libraries (c.f. legal deposit)
  - British Library (BL)
    British Library e-Journal Digital Archive
  - Koninklijke Bibliotheek (KB e-Depot)
    KB, National Library of the Netherlands
- One library cooperative
  - UK LOCKSS Alliance
Ensuring Continuity of Access to Resources for Scholarship

Peter Burnhill

Legal deposit

- Works well with print via legislation and national libraries
- Countries with legislation enacted (or ‘in train’) for e-materials include: Canada, Denmark, Finland, France, Germany, Iceland, New Zealand, Norway, South Africa, Sweden, UK
- But not all countries (notably USA) and in UK the legislation supports voluntary deposit, with restrictions of mode of access

System architecture

Piloting an E-journals preservation registry service (PEPRS)
JISC-funded project, EDINA & ISSN-IC as partners
Ensuring Continuity of Access to Resources for Scholarship

Peter Burnhill

Ensuring researchers, students and their teachers have ease and continuing access to online scholarly resources

Re-thinking trust: assurance of continuing access

- Stewardship (for digital content) as action on trust
  - The trust put in libraries and archives by present and future children, citizens and scholars
    * Not just budgets and provision of cost-effective access
- Digital Accident: ease of online access also means that libraries no longer hold back copy on their shelves
  - A challenge of stewardship for publishers as well as libraries
    * "What to do with that responsibility: outsource to libraries?"
- The Central Task (for research and university libraries et al.): Ensure that researchers, students & their teachers have continuity of access to the (online) scholarly resources they need

The screen versions of these slides have full details of copyright and acknowledgements
Ensuring Continuity of Access to Resources for Scholarship
Peter Burnhill

Trust: avoiding disaster by sharing the task

1. Methodology, that is explicit, testable & ‘proven’
   – At a local level that should include ‘copies’
   – At a global level that should different approaches
2. Competence to do what is required
   – Mix of education, skill, knowledge and wisdom
3. Commitment to do it, in a timely fashion
   – That means funds to secure and sustain effort
4. Disclosure of what is being done
   – Can the ‘Registry of Keepers’ model be generalised?
5. And then audit and report what was said & done

“We all make mistakes; the trick is being able to recover” ☺

The screen versions of these slides have full details of copyright and acknowledgements