Slide 1

Ensuring Continuity of Access to Resources for Scholarship

Peter Burnhill
Director, EDINA
University of Edinburgh

ISSN   JISC   EDINA
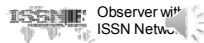
1

Slide 2

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.dcc.ac.uk

• Director, EDINA – JISC/UK national data centre, 1996 - present day
  – Online services for UK universities and colleges
     edina.ac.uk

ISSN   Observer with ISSN Network

2

Slide 3

EDINA®   *providing resources for staff and students in higher and further education in the UK and beyond*

Services [Select] Go   Projects [Select] Go

Reference & Multimedia

Film, Images & Sound
JISC MediaHub
including NewsFilm Online
Journal Catalogues
SUNCAT (UK)
SALSER (Scotland)
Article References
Land Life Leisure
CAB Abstracts
Historical Materials
The Statistical Accounts of Scotland
Sharing Research Papers
OpenDepot.org

Maps & Data
Geo-data Discovery
GoGeo
Maps & Datasets
Digimap - Ordnance Survey
Geology Digimap
Historic Digimap
Marine Digimap
UKBORDERS
agcensus
For Schools
Digimap for Schools
Sharing Data
ShareGeo Open
Geodoc
Research Data MANTRA

Developers & Librarians
Middleware & Infrastructure
UK federation
Unlock
Digimap OpenStream
OpenURL Router
EDINA Developer Corner
Repositories & Preservation
The Keepers Registry
UK LOCKSS Alliance
CLOCKSS
Repository Junction Broker
UK RepositoryNet+

News, Training & Events
Welcome to Will's World: A Shakespeare registry
23 April 2012
Geoforum 2012: Programme & Booking Information
20 April 2012
Latest Statistics of E-Serials Preserved with The Keepers Registry
19 April 2012
UKLA Roundup: April 2012
18 April 2012
fig*share* and Pensoft join CLOCKSS
16 April 2012
all news
Quarterly Newsletter

What EDINA Does: Community Report (pdf)

Jobs at EDINA

Follow us
Presentations and Publications

Projects & Expertise

EDINA is a JISC National Data Centre based at the University of Edinburgh | Privacy Policy | Accessibility | Acknowledgements

JISC

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

4

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"
      *M.Buckland, *Journal of American Society for Information Science*, **50** p970-74 1999

- 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?

5

Infrastructure to support four 'demand-side' verbs

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

6   Based on MODELS workshops (UKOLN/JISC eLib)

Scientific / scholarly record

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

<u>But</u>

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

7

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

8

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

9

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

10

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 *

11

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

12

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

13

**Life cycle models, provenance,
authenticity, diplomatics**

- "**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 data life cycle with the metadata and services that will enable access, evaluation and re-use over time"
Ann Green http://www.iassistdata.org/blog/conceptualizing-digital-life-cycle

- "When Documents Deceive: **Trust and Provenance** as New Factors for Information Retrieval in a Tangled Web" *Clifford A. Lynch, Journal of the American Society for Information Science and Technology,* **52**(1):12–17, 2001

- "Archival science, with its principles of uniqueness, **provenance**, arrangement and description, **authenticity**, appraisal, and its tool sets such as **diplomatics** …, [as] framework for a theoretical foundation for digital libraries" *Seamus Ross,* Keynote Address ECDL2007, Budapest

14e underlying concept is Trust

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

15

HENRY STEWART TALKS

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?

16

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

17

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

18

The Keepers Registry BETA

http://thekeepers.org

Supporting long-term access to journal content

Home | Journals | Publishers | Archiving Agencies | FAQ | About | Help

Discover who is looking after what e-journals

Search 🔍

Input a search term (e.g. free text, or a title or ISSN(s))

[   ] [Search]

⊞ Display example searches

📖 Glossary and definitions

The Keepers Registry allows you to search on title or ISSN

Learn more about the archiving agencies in The Keepers Registry who have stepped forward to take responsibility for ensuring long-term continuity of access to global scholarship.

- British Library (Last updated: 30 Dec 2011)
- CLOCKSS Archive (Last updated: 12 Apr 2011)
- e-Depot (Last updated: 11 Apr 2011)
- Global LOCKSS Network (Last updated: 11 Apr 2011)
- HathiTrust (Last updated: 01 Jan 2012)
- Portico (Last updated: 06 Jan 2012)

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

19

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

20

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

21

22

E-Journals

P**E**PRS

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

23

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

24

HENRY STEWART TALKS

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
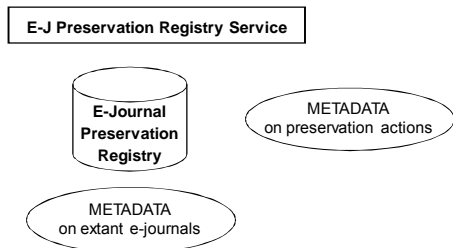
25

System architecture

26

Piloting an E-journals preservation registry service (PEPRS)
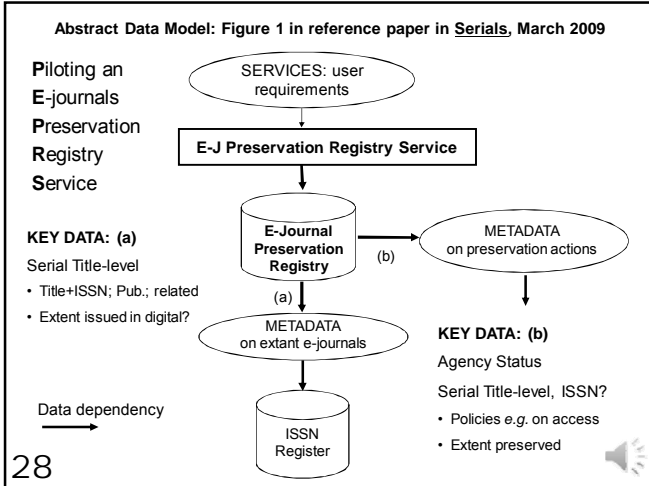**JISC-funded project, EDINA & ISSN-IC as partners**

E-J Preservation Registry Service

E-Journal Preservation Registry

METADATA on preservation actions

METADATA on extant e-journals

27

Abstract Data Model: Figure 1 in reference paper in Serials, March 2009

Piloting an
E-journals
Preservation
Registry
Service

SERVICES: user requirements

E-J Preservation Registry Service

E-Journal Preservation Registry

(b) → METADATA on preservation actions

KEY DATA: (a)

Serial Title-level
• Title+ISSN; Pub.; related
• Extent issued in digital?

(a) →

METADATA on extant e-journals

KEY DATA: (b)

Agency Status

Serial Title-level, ISSN?
• Policies *e.g.* on access
• Extent preserved

Data dependency ────▶

ISSN Register

28

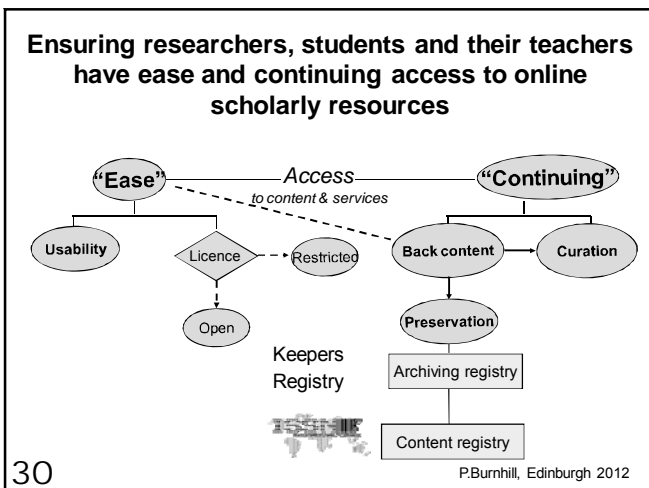**Re-thinking trust: assurance of continuing access**

- *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

29

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

"Ease" ---- *Access* ---- **"Continuing"**
to content & services

Usability   Licence - -▶ Restricted   Back content → Curation

Open   Preservation

Keepers Registry   Archiving registry

Content registry

30   P.Burnhill, Edinburgh 2012

Trust: avoiding disaster by sharing the task

- When we trust, or are trusted by others, the steps are:

  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

31 *"We all make mistakes; the trick is being able to recover"* ☺

32