

Solving the E-Journal Problem

What does the Keepers Registry tell us?

Peter Burnhill

EDINA, University of Edinburgh, UK

<http://creativecommons.org/licenses/by/3.0/>

The problem to solve:

Key task for academic & research libraries is to ensure **ease & continuity of access to the scholarly & cultural record**

We all know what has changed ...

- Libraries buying '**ease of access**': What was once available locally now online & accessed remotely, anytime/anywhere
- Good News 😊

But the Bad News! ☹️

- **Academic libraries are no longer the custodians of the scholarly record**
- Are Libraries **ensuring continuity of access?**
 - Where are their digital shelves?



On the Scholarly Web, a larger problem lurks

- **Reference to what is ‘on the Web’ - Web References**
 - Increasing use of HTTP URL/URI to cite resources & work of others
 - a fundamental part of scholarly discourse, in which cited (printed) material sat on a shelf that could, with tedium and delay, be obtained on Inter-Library Loan
 - Content at URI can and does change, or even cease to be
 - ‘citation rot’ and ‘reference rot’ – not only ‘link rot’
- ***Hiberlink***, a joint UoE/LANL/EDINA (Mellon funded-) project to investigate Time Travel for the Scholarly Web 😊
 - Builds on study by Sanderson, Phillips & Van de Sompel (2011)
 - 28% of resources referenced by articles in an IR had been lost
 - 45% (66,096) of the URLs [in arXiv] that were extant had not been archived.

E-journals may be the easier part of the problem:
so is the e-journals problem is being solved?



So, a welcome to this Report from Neil

- Contrasts '**Continuing Access**' & '**Long-term Preservation**'
- Lists 6 Use Cases which illustrate both

1. Library cancels a JOURNAL subscription

2. Library exits a Big DEAL

3. Back issues of journal become unavailable from publisher

4. Journal becomes 'orphan' as publisher goes out of business

5. Journal 'unavailable' as operation of publisher hits disaster

3 to 5 are the 'preservation' Use Cases

6. Library decides to remove / dispose print journals

1,2 & 6 are where librarians focus policy



Is 'preservation' still a problem?

We have had many Reports over past 10 Years

...

- highlighting risks in digital media & formats
- warning against single points of failure

And impressive number of archiving agencies

- ① web-scale not-for-profit archiving agencies
e.g. CLOCKSS Archive & Portico
- ② national libraries (with legal deposit in mind)
e.g. e-Depot (Netherlands); British Library
& National Science Library of China
- ③ research libraries: consortia & specialist centres
e.g. Global LOCKSS Network, HathiTrust,
Scholars Portal, Archaeology Data Service

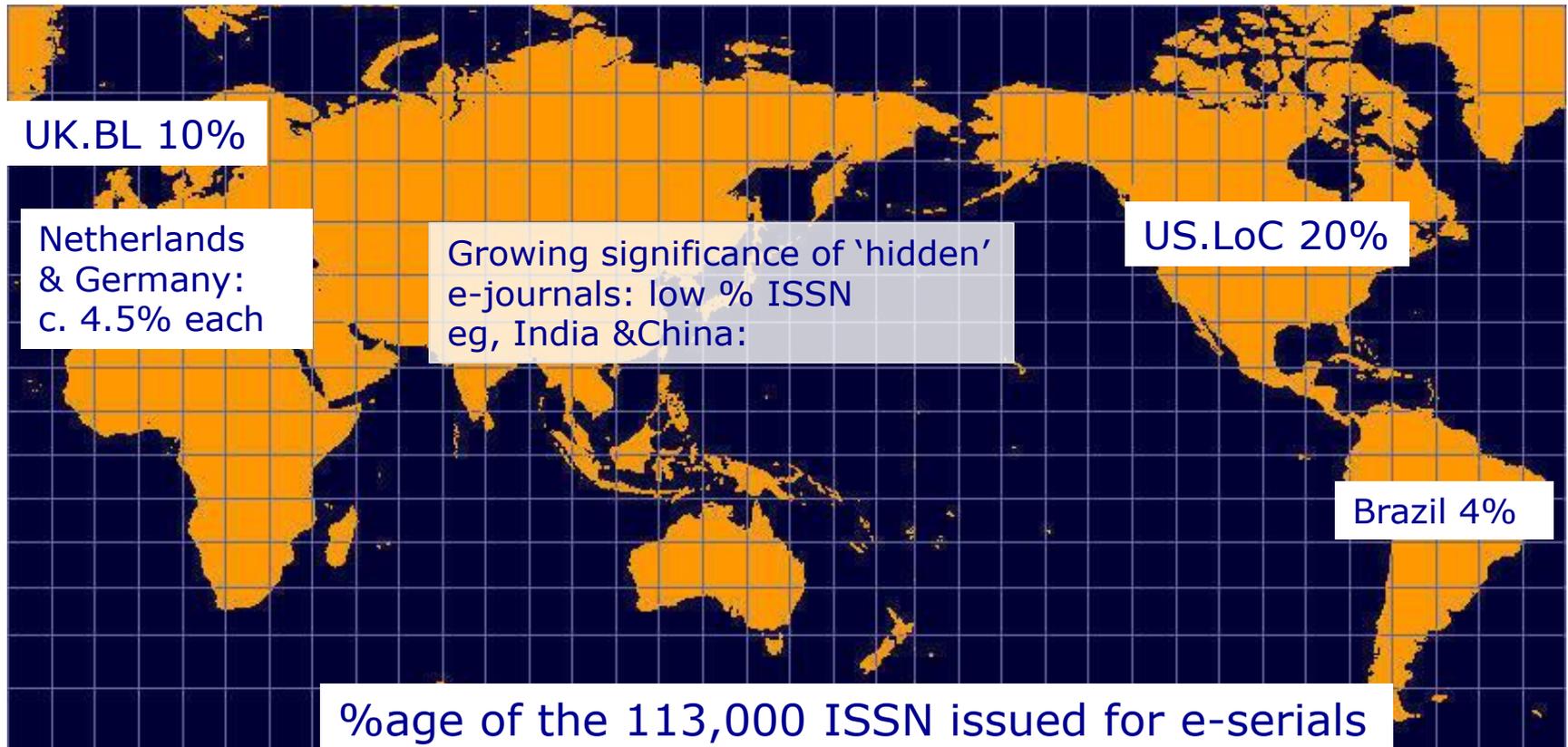
Many archiving organisations is a Good Thing 😊



“Digital information is best preserved by replicating it at multiple archives run by autonomous organizations”

B. Cooper and H. Garcia-Molina (2002)

Scale & International dimension: ISSN



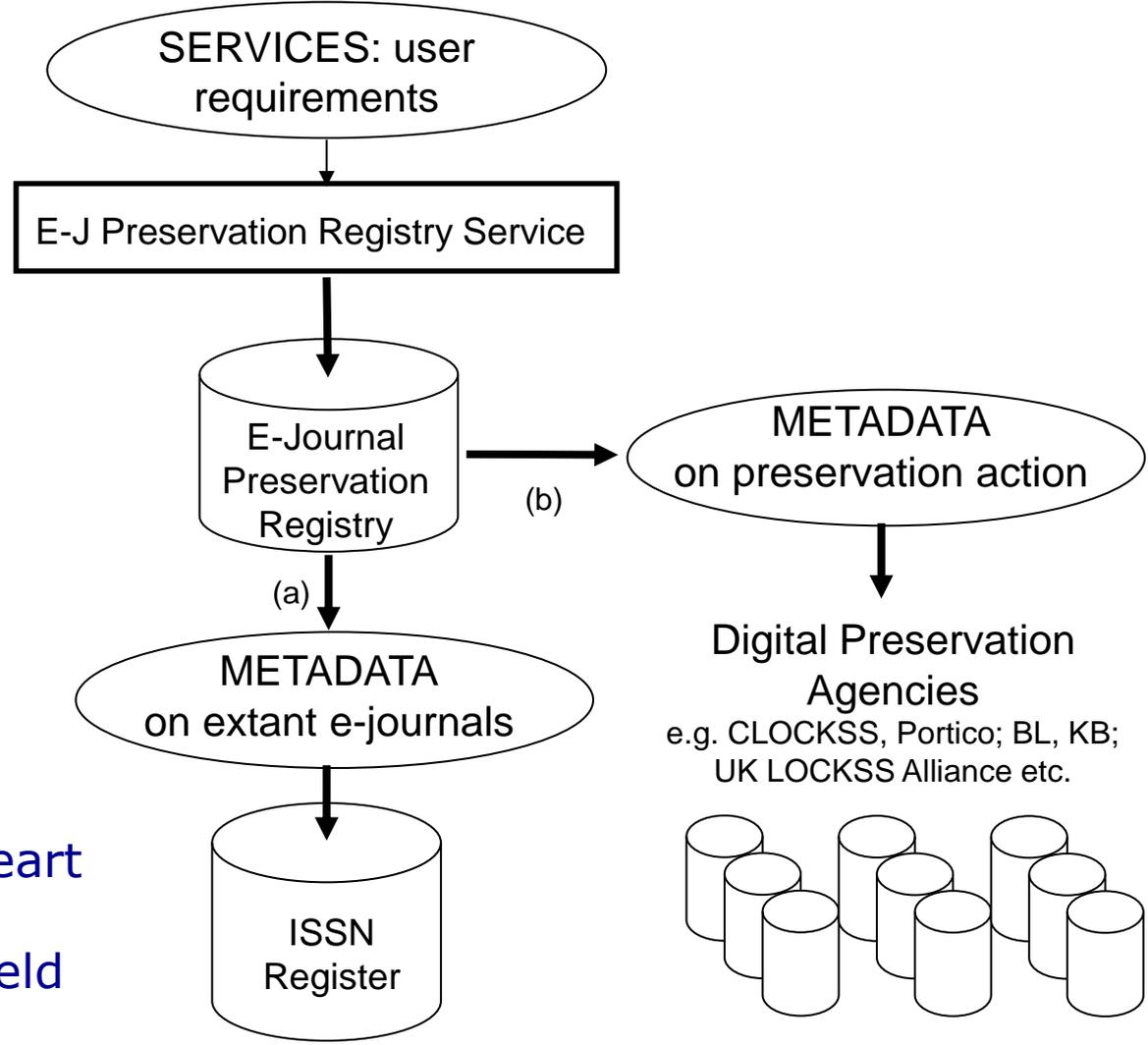
UK researchers & students (and therefore libraries) are dependent upon content written and published in other countries

* That is true in every country *

How to know who is looking after what & how? (and uncover what is still at risk)

Piloting an
E-journal
Preservation
Registry
Service

Data
dependency
→



ISSN Register at heart
of the Data Model;
ISSN-L as kernel field

(Taken from Figure 1 in reference paper in [Serials](#), March 2009)

Now have a global Registry of e-journal archiving

BETA
The Keepers Registry ... to discover who is looking after what
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))

[Display example searches](#)

[Glossary and definitions](#)

Enter title
or ISSN

Current Statistics

- Serials reported as 'preserved' : 20,987
- Serials where action is 'in progress' : 11,263

We include a title count where at least one agency has taken the reported action on one or more volumes. See the [FAQ](#) for more information.

News and Events

15th May 2013: Two new Keepers report e-journal archiving [\[View\]](#)

ers Registry [\[View\]](#)

KSG 2013 [\[View\]](#)

ation at UKSG 2013 [\[View\]](#)

at the "Digital Future & You" series [\[View\]](#)

gs Comparison [\[View\]](#)

[See the keepers registry blog](#) for archived news and our [development roadmap](#).

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.

- [Archaeology Data Service](#) (Last updated: 27 Jun 2013)
- [British Library](#) (Last updated: 02 Jul 2013)
- [CLOCKSS Archive](#) (Last updated: 13 Aug 2012)
- [e-Depot](#) (Last updated: 24 Apr 2012)
- [Global LOCKSS Network](#) (Last updated: 25 Jun 2013)
- [HathiTrust](#) (Last updated: 01 Jul 2013)
- [National Science Library, Chinese Academy of Sciences](#) (Last updated: 28 Jun 2013)
- [Portico](#) (Last updated: 28 Jun 2013)

to search across metadata
reported by leading
archiving organisations

The most up-to-date information is available on the agency website. The date represents when the Registry was updated.

This is a Beta release from the [PEPPS project](#). We are also exploring governance and seeking sponsorship for this global facility. Your [feedback](#) is welcome.

[Terms & Conditions](#)

Supported by
JISC

news
Scholars Portal (Ontario) & Library of Congress (USA)
in process of joining the Keepers Registry

NA

Progress, but still not 'job done' ?

The Keepers Registry <thekeepers.org> reports:

- A. 21,557 e-serial titles are being 'Preserved'
i.e. ingested by organisations with archival intent
 - (Many 'missing volumes and issues')
 - B. 113,092 ISSN assigned to 'online serials in ISSN Register'
- Progress with a key indicator: ratio of A/B = 19%
– was 17% at close of 2011 (16,558 / 97,563)

Do we need to agree a 'priority list' of titles?

1. Should we only be interested in the c.30,000 'peer-reviewed' scholarly journals?
2. Do we look only at on what individual libraries list?

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

c.75% 'at risk'
c.11% held by 3 or more

- Two key indicators: **%age (& number) of titles that are 'at risk of loss'**
%age (& number) of titles that are 'preserved by 3 or more Keepers'.

3. Should we ask the audience?

- The researchers and students who read online serials

Looking from the user's point of view ...

... with usage logs for the UK OpenURL Router

- 10.4m full text requests in 2012; ISSN-L to de-duplicate ISSN
- 53,311 online titles requested by researchers & student from 108/160+

Analysis using the Keepers Registry:

- Only 15% (7,862) are being kept by 3+ Keepers
- Over two thirds (68%) held by none
- 36,326 titles 'at risk' of loss ☹️
- Check robustness with UK logs for 2011 & 2013; Request logs for other countries (WorldCat)

➤ **So 'preservation really is still a problem!**

- **Best Case scenario for IFLA 2020**
 - Libraries (& Publishers) have acted to reduce that alarming 80% figure to near to zero 😊
 - They have ensured that all the e-journal content used by their researchers in 2013 has been preserved and can be successfully used in 2020, and assuredly beyond. 😊
- **Worst Case scenario for IFLA 2020**
 - **Libraries (& Publishers) have failed to act 😞**
 - **Important literature has been lost 😞**
 - **Citizens & scholars complain of neglect!**

Not too late for policy & priority: towards a plan?

- ◆ Stewardship?
 - ◆ Entrust that stewardship to others and/or engage collaboratively?
 - Re-state *a set of values* that **complement & underpin cost-effective access (& budgetary matters)**
 - ◆ **How to deliver on the trust put in libraries & archives!**
 - implement with stories & conversation, not commands:
- ‘Research libraries saved the day, because they acted to ...**
- Always a mix of ‘Access Libraries’ & ‘Holdings Libraries’
 - *Today hardly any library providing e-connections can really call them e-collections*

Good News & Main Challenge?

Good news?

- Most of the big publishers engage with archiving initiatives
- Keepers Registry often show titles held by 3+ 'Keepers'
 - typically CLOCKSS, e-Depot and Portico.

Main challenge?

- The long tail of smaller publishers
 - regardless of business model.
- It is not about Open Access *per se*
 - DOAJ for content of 10,000 e-journals from 4,000 publishers
- Lots of other (important/priority?) e-journal
- Role of national libraries or library consortia?

Ask a librarian in 2020: 3 possible answers

1. "Yes, we have it (we've checked recently, both in the catalogue and in actuality), and you can access it now"
2. "No, but we know some body that does (we trust),
 - so we can point you to (or arrange access to) it now/soon-ish"
3. "Sorry, we don't know ...
 - perhaps no body has it
 - it may be lost forever, altho' perhaps somebody somewhere ..."
 - That was true for the print world
 - Unfortunately, we could allow the 3rd answer to become the common one for a lot of e-journal content

1. To provide libraries & archiving organisations with lists of titles that seem to be at risk of loss
 - With (what we know of) publisher information
2. To discover and recruit more archiving organisations as Keepers into the Registry
 - The Registry is not an audit / certification authority but there are eligibility checks for integrity of ‘archival intent’
3. To keep a close focus on volumes & issues
 - The extent of a serial that is held by each of the Keepers
4. To assist collaboration between Keepers with ‘a safe places network’: many met at iPres 2013 in Lisbon this September
5. To assist the ISSN Network assign more ISSN
 - If it is worth preserving, it really should have an identifier

Monitoring Preservation Progress: Serial Content

Thank you for
listening

<http://thekeepers.org>

<http://thekeepers.blogs.edina.ac.uk/>

<http://www.flickr.com/photos/shinez/5000985919/>

Sidebar note on National Libraries

Should we wait upon Legal Deposit?

– 94% of libraries have some form of legal deposit *for print*.

- Only 44% national libraries had legislation in 2011 for e-books or e-journals; expected to rise to 58% by June 2012.

from presentation, CENL 2011 Survey by Lynne Brindley
to CDNL Annual Meeting Puerto Rico, 15/8/11

- Only 27% [expected to rise to 37% by June 2012] actually ingesting via legal deposit
 - Total national libraries collecting = those 14 via legal deposit + 9 by other means (Netherlands, UK/BL, Switzerland *voluntary deposit*)
 - Only KB e-Depot, BL, NSLC (+ LoC) in The Keepers Registry
 - Only when the other 19 join will all know about their activity
 - **Key point is not about call for ‘legal deposit’ but that this is taking too much time**

Early Archiving Initiatives

- eDepot at Koninklijke Bibliotheek
 - national role (for the Netherlands) with international significance because of Elsevier and Kluwer
- the LOCKSS project at Stanford University
 - from which came CLOCKSS
- the Electronic-Archiving Initiative at JSTOR
 - from which came Portico

The Idea of a **Registry** mentioned in 2003/4, and again in 2006:

“by which it **would be plain what content was being archived, and therefore what was not** (US CLIR Report).

In 2009, a Scoping Study commissioned by Jisc from Loughborough University (UK) & Rightscom confirmed the need:

“overall lack of information about where e-journals were archived
... the difficulty of finding the information across a range of sources”.

... and discover details of its 'archival status'

Origins of life

1 hit found
Current search: Origins of life (Key Words)

This e-journal is being archived by 5 archiving agencies ...

Title	ISSN	Publisher	Current extent of archiving 	Archiving agency
1. Origins of life and evolution of the biosphere	1573-0875 (Online); 0169-6149 (Print)	Springer	Preserved: v. 36, 40-41	CLOCKSS Archive
		Kluwer Academic Publishers	Preserved: v. 27-37	e-Depot
		Springer Netherlands	Preserved: v. 37-41	e-Depot
		Springer Science+Business Media	In progress: v. 1-present	Global LOCKSS Network
		Springer	Preserved: v. 27-42	National Science Library, Chinese Academy of Sciences
		Springer	Preserved: v. 39-41	Portico

... but coverage of volumes is partial & patchy

Example search: 'Origins of Life'



Pushing at an (IFLA) open door?

IFLA Key Initiative 1 - Digital Content Programme:

- working collaboratively to build a legal, technical & professional base that enables libraries to play a major role in collecting, **preserving**, and offering wide access to all types of physical and digital materials:
 - **digital preservation** – with a focus on legal mechanisms [*] for harvesting and **preserving born digital information and local content hosted on websites** and in social media

Activities: Develop and endorse the IFLA Statement on Legal Deposit [*]

The Strategic Programme on Preservation and Conservation (PAC) has one major goal : **to ensure that library and archive materials**, published and unpublished, **in all formats, will be preserved in accessible form for as long as possible** according to the following principles :

- preservation is essential to the survival and development of culture and scholarship;
- international cooperation is a key principle;
- each country must accept responsibility for the preservation of its own publications.

