

UK Legal Deposit in the digital age: organisational and operational challenges

Introduction

- The British Library – facts and figures
- Legal deposit in the UK – history and content
- Addressing the digital black hole
 - Legislation
 - Technical infrastructure developments
 - Preservation

The British Library: Facts and Figures

National library of the UK. Serves researchers, business, libraries, education & the general public

Collection fills over 600km of shelving and grows at 11km per year
16 Tb of digital material through voluntary deposit

Collection includes over 2m sound recordings, 5m reports, theses and conference papers, the world's largest patents collection (c.50m)

CA&D processes c.1 million items per annum:
500,000 serial issues
200,000 monographs
150,000 newspaper issues



Helping people advance knowledge to enrich lives

3 main sites in London and Yorkshire. Circa 2,000 staff

The largest document supply service in the world. Secure e-delivery and 'just in time' digitisation enables desktop delivery within 2 hours

Generates value to the UK economy each year of 4.4 times public funding

GIA Funding 08/09: £94.8m operational, £12m capital

Other funding secured 07/08: c.£33m

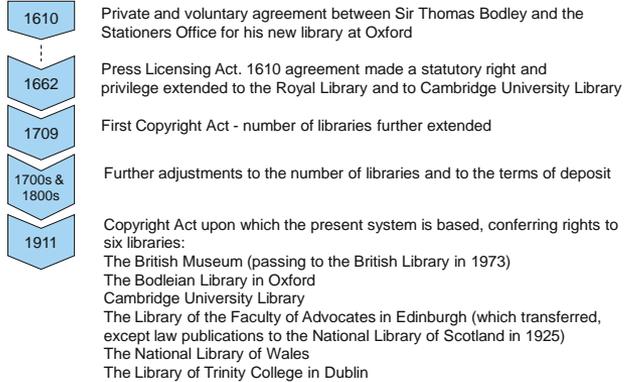
The benefits of legal deposit

In the United Kingdom legal deposit dates back several hundred years and has been a vital element in preserving and making available the published record of previous generations for the researchers of today and the future.

For the British Library legal deposit strongly supports our mission to help people advance knowledge to enrich lives.

The growth in electronic publishing brings with it a whole range of challenges and opportunities for legal deposit libraries both in the UK and across the world.

History of print legal deposit in the UK

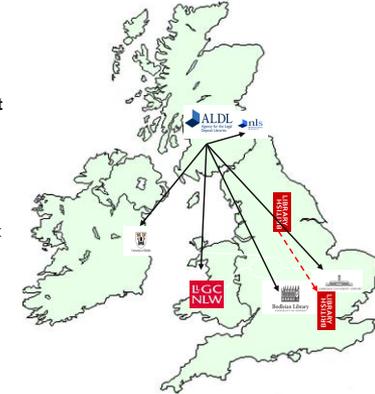


5

Print legal deposit infrastructure

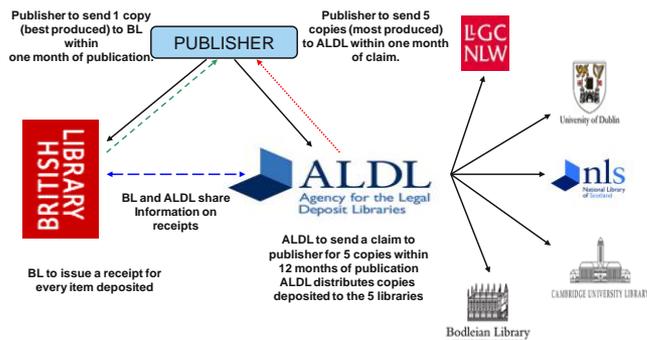
- British Library, St Pancras
- British Library, Boston Spa
- Agency for the Legal Deposit Libraries

The remaining 5 legal deposit libraries (i.e. excluding the BL) in the UK work manage their legal deposit claiming and receipting activities through a single, joint funded agency that operates on their behalf



6

Print legal deposit infrastructure



7

Print legal deposit volumes

Items received on Legal Deposit 2008/09

- 141,755 Monographs
- 259,753 serials issues
- 1,932 Maps and atlases
- 2,041 Music scores
- 144,982 Newspaper issues
- 275 Playscripts



8

The Digital Universe – what is a format?

Traditional categories for deposit



Wide range of digital formats

ARC_IA TIFF
XML JPEG
PDF ZIP
GIF WARC

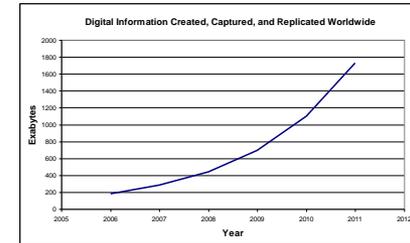
9

The Digital Universe – estimated growth

• Estimated volume of digital information worldwide in 2007: 281 Exabytes

- Estimated growth rate: ca. 60%
- → 700 Exabytes in 2009!

1000	k	kilo
1000 ²	M	mega
1000 ³	G	giga
1000 ⁴	T	tera
1000 ⁵	P	peta
1000 ⁶	E	exa
1000 ⁷	Z	zetta
1000 ⁸	Y	yotta



source:
"The Diverse and Exploding Digital Universe"
IDC White Paper, March 2008
<http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf>

10

The Digital Universe – UK published?

Items received on Legal Deposit 2008/09

- 141,755 Monographs
- 259,753 serials issues
- 1,932 Maps and atlases
- 2,041 Music scores
- 144,982 Newspaper issues
- 275 Playscripts

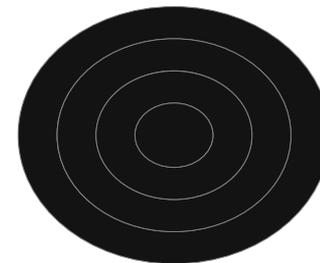


Defining "UK"
and determining
scale of
collecting in a
digital
environment



11

The Digital Black Hole



12

Addressing the digital black hole

Strengths

- Established role of libraries in managing, preserving and providing access to large scale collections
- Reputation as trusted custodians
- Staff expertise
- Longevity
- Vision

Opportunities

- Collaboration (national/international)
- Ability to influence legislation
- Shared costs & infrastructures
- Reducing the publisher burden
- Research funding
- Digital Britain
- Sharing knowledge and experience

Weaknesses

- Lack of digital storage infrastructure
- Lack of digital preservation strategies and tools
- Lack of digital access management systems
- Ability to act fast enough to a rapidly changing external environment

Threats

- Potential loss of digital materials
- Complexity of legal deposit legislation
- Unknown scale of digital publishing
- Complexity of digital publishing formats
- Defining 'UK' in a digital environment
- Resources
- Managing print/digital balance

13

Addressing the digital black hole

To address the black hole we need to turn some of our weaknesses and any threats into strengths and opportunities. We will now look at how legal deposit libraries in the UK are addressing some of the key issues in relation to :

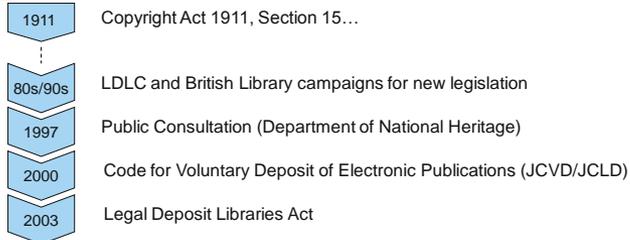
Legislation

Technical Infrastructure

Preservation

14

The path to legislation – electronic publications



However.....

15

Legal Deposit Libraries Act 2003

- Essentially an update of 1911 Copyright Act
- Primary legislation which 1) reaffirms print deposit and 2) creates a framework for secondary legislation to cover non-print deposit

Print	Non-print
<p>British Library:</p> <ul style="list-style-type: none"> • Compulsory deposit within 1 month • 'Best' copy • Deposit direct to BL + receipted <p>Other five Libraries:</p> <ul style="list-style-type: none"> • Deposit on request within 12 months • 'Standard' copy • Most deposits delivered via Agency <ul style="list-style-type: none"> • Copyright & defamation protections... 	<ul style="list-style-type: none"> • Requires regulations via 'affirmative resolution' • No distinction between libraries • Access restricted to library premises • Excludes sound and film where they are the primary content • Deposit 1 copy if on multiple media • Copyright & defamation protections...

16

Progress since 2003 Act

Legal Deposit Advisory Panel

The role of the Panel is to advise the Secretary of State on the timing and content of regulations relating to legal deposit and to oversee the implementation of the Legal Deposit Libraries Act 2003.

Members:

5 Independents	5 Publishers	5 Libraries
----------------	--------------	-------------

Support:

Libraries -	LDAP Policy Officer Libraries' Project Manager
Government -	Department for Culture, Media and Sport Department of Business Enterprise and Regulatory Reform

17

Defining categories of e-publication in regulations

- LDAPs first attempt to map the universe of electronic publishing identified 19 categories
- In 2006 LDAP commissioned a study from EPS Ltd. to refine this first map and propose a new and complementary map which could be used as a basis for defining categories of e-publication in regulations

The study proposed a new map based upon four questions (simplified below) about the usual interaction between electronic content and its users:-

1. Is the content enquiry-driven (such as a database) or not?
2. Is delivery offline (e.g. handheld media) or online?
3. Is the content delivered to or collected by its users?
4. Is the content made openly available or protected behind a barrier?

David Powell, EPS Ltd: "Refining the map of the universe of electronic publications potentially eligible for legal deposit"
http://www.culture.gov.uk/images/publications/EPS_Report_to_LDAP_Nov_2006.pdf

18

Categories for recommendations

1. Offline publications on hand held media
2. Material (both websites and documents) which, when published, is made freely available on the internet
3. Online websites and documents that are protected behind a barrier which requires payment, registration or compliance with some other formality
4. Structured, enquiry-driven data sets which require additional software or systems for searching, selecting or displaying the data
5. Content which is 'pushed' or delivered to the user by email or other means, such as a news feed to the computer desktop

19

Options for recommendation

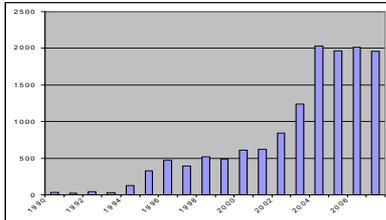
Options :

- 1) Leave to Market
- 2) Voluntary Code
- 3) Self-Regulated Code
- 4) Regulation

20

Progress to date – Offline publications

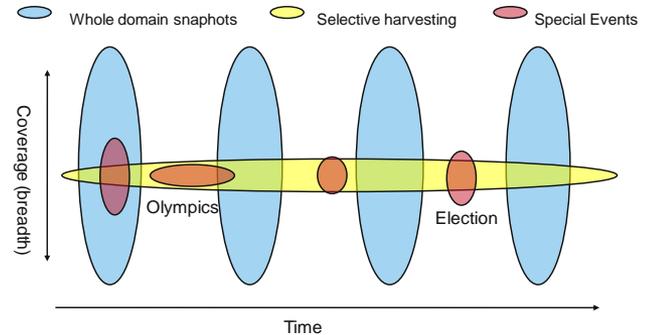
- Existing Code for Voluntary Deposit of Electronic Publications
- 12-month analysis period, with statistics reported monthly
- Case studies and publisher interviews. Cost analysis.
- Trends:



Recommendation - Voluntary Code (submitted Nov 2008 and accepted)

21

Progress to date – Freely available online material (Web archiving)



Adapted from: Netarchive.dk (Denmark)

22

Progress to date – Freely available online material (Web archiving)

As at Feb 2009: 7.5 million .uk websites, growing net 11% per annum plus .sco .cym etc.
plus UK material in .com, .edu, .net, etc

Library perspective	Leave to market	Permission-based harvesting by BL, NLS and NLS (UKWAC*)	Regulation based harvesting by BL, NLS and NLW
Cost per annum	£0	£0.5M	£1.13M
Data capture	0 Terabytes (0%)	77 Terabytes (0.5%)	5.3 Petabytes (79%)
Cost per terabyte	Not applicable	£6,476	£215

Recommendation – Regulation (updated submission May09, publication consultation Dec09)
This will allow archiving in breadth as well as depth

*<http://www.webarchive.org.uk>

23

Progress to date – commercial and protected publications

- Voluntary scheme underway for scholarly e-journals
- Scoping study of e-universe of UK e journals undertaken
- Outcomes of scoping study and voluntary scheme to inform options for regulation.

Powell, David J. 2004. Publishing Output to 2020. In *The Future of Print and Electronic Publishing Output Worldwide*, a report by Electronic Publishing Services Ltd. to the British Library, January 29, 2004. Extract available at <http://www.bl.uk/electronic/pubs/epsreport.pdf>
David Powell, EPS Ltd. "Refining the map of the universe of electronic publications potentially eligible for legal deposit" http://www.culture.gov.uk/images/publications/EPS_Report_to_LDAP_Nov_2006.pdf

24

Progress to date – commercial and protected publications

Particular consideration given to:

- Preferred methods for deposit (push/pull, automatic/claim based)
- Preferred file formats
- Preservation requirements (e.g. copying, format shifting)
- Metadata standards, requirements, formats and use
- Access and use
- Migration from print to electronic
- Recommendations expected to be submitted to the Secretary of State in March 2010

25

Addressing the digital black hole

To address the black hole we need to turn some of our weaknesses and threats into strengths and opportunities. We will now look at how legal deposit libraries in the UK are addressing key issues in relation to:

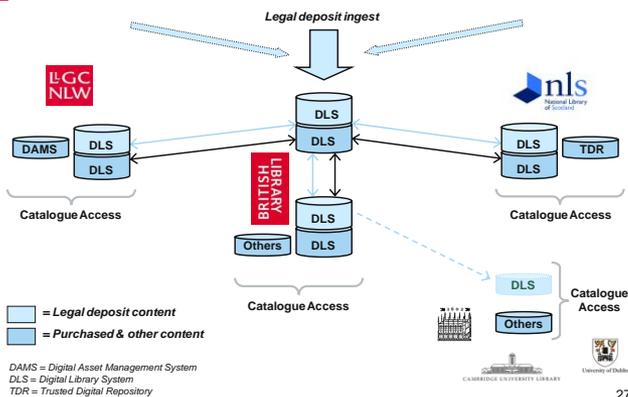
Legislation

Technical Infrastructure

Preservation

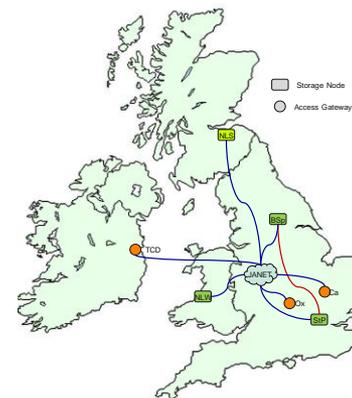
26

Libraries proposals for shared infrastructure



27

Libraries proposals for shared infrastructure



28

Benefits of developing a shared infrastructure for digital content

- Reduces burden on publishers (deposit of 1 distributed copy rather than 6)
- Avoids duplication of effort across legal deposit libraries in the UK across digital life cycle
- Allows costs to be shared
- Standardises access – same legal deposit content accessible across all legal deposit libraries
- Pooling of expertise and knowledge

29

Addressing the digital black hole

To address the black hole we need to turn some of our weaknesses and threats into strengths and opportunities. We will now look at how legal deposit libraries in the UK are addressing key issues in relation to:

Legislation

Technical Infrastructure

Preservation

30

Preservation - Issues

Issues

Media obsolescence (bit rot)
 Hardware obsolescence
 Software obsolescence
 Format obsolescence
 Loss of context (metadata)
 Volume of digital material
 Authenticity

Questions

What is worth preserving?
 What are the priorities for preservation?
 How to preserve?
 How to preserve so much?
 How to ensure quality?
 What does it cost to preserve?

31

Preservation – Life span of media

Parchment: 1000 years

Microfilm: 500 years

Paper: 50 – 200 years
 ■ high levels of acid can cause paper to disintegrate

Magnetic Tape: 100 years

■ the binder that holds magnetic particles to the tape can decompose and cause the layers of tape to stick together in a reel

CD-ROM: 10 years

■ poor manufacturing processes allow the reflective aluminum layer to oxidize

Is this progress?



Slide courtesy of A Farquar, British Library, PLANETS

32

BRITISH LIBRARY

Preservation – addressing the challenge

Planets (Preservation and Long-term Access through NETworked Services)



planets

- Four –year €15 million project
- Co-funded by the European Union under the framework to address core digital preservation challenges

Co-ordinated by the British Library, the Planets consortium is made up of 16 members bringing together the expertise of European National Libraries and Archives, leading research universities and technology companies

The project's goal is to deliver a sustainable framework to enable long-term preservation of digital content, increasing Europe's ability to ensure access in perpetuity to its digital information



<http://www.planets-project.eu/>

33

BRITISH LIBRARY

Preservation - addressing the challenge

Questions	Preservation solutions
What is worth preserving?	Planets Preservation Planning tool – PLATO
What are the priorities for preservation?	Preservation Actions tools (to transform and emulate obsolete digital assets)
How to preserve?	Characterisation tools
How to preserve so much?	Planets Testbed (to verify and test results)
How to understand Digital Objects?	LIFE ¹ and LIFE ² projects: http://www.life.ac.uk
How to ensure quality?	
What does it cost to preserve?	

34

BRITISH LIBRARY

Conclusions - filling in the digital black hole

<p>Strengths</p> <ul style="list-style-type: none"> • Established role of libraries in managing, preserving and providing access to large scale collections • Reputation as trusted custodians • Staff expertise • Longevity • Vision 	<p>Opportunities</p> <ul style="list-style-type: none"> • Collaboration (national/international) • Ability to influence legislation • Shared costs & infrastructures • Reducing the publisher burden • Research funding • Digital Britain • Sharing knowledge and experience
<p>Weaknesses</p> <ul style="list-style-type: none"> • Lack of digital storage infrastructure • Lack of digital preservation strategies and tools • Lack of digital access management systems • Ability to act fast enough to a rapidly changing external environment 	<p>Threats</p> <ul style="list-style-type: none"> • Potential loss of digital materials • Complexity of legal deposit legislation • Unknown scale of digital publishing • Complexity of digital publishing formats • Defining 'UK' in a digital environment • Resources • Managing print/digital balance

35

BRITISH LIBRARY

Conclusions

The digital age is one of increased collaboration and sharing of knowledge and expertise between libraries. Opportunities for collaboration are made possible not only by the nature of digital content itself, but also because of the scale of the task faced by many international, national and institutional libraries.

"Collaboration was critical to the foundation of society [It dates back to] when prehistoric man was hunting mammoths. If you didn't collaborate, the mammoth trampled you and you didn't eat." (Ward 2000)

Access to knowledge is a critical foundation of society today. Through collaboration we can ensure that the scale of the digital challenge does not overwhelm us and that the published output of the world continues to be collected and preserved for the consumption of future generations.

[1] Lewis Ward, senior research analyst at CollaborativeStrategies/Collaboratory in San Francisco
Salopek, J.J. (2000) Digital Collaboration, Training and Development.

36

Thank you for your attention!

Contact information:

Sharon Johnson

sharon.johnson@bl.uk