Digital Preservation and Digital Library Development at the London School of Economics

Ed Fay, Digital Library Manager

e.fay@lse.ac.uk | @digitalfay
LSE Library

NATIONAL RESEARCH LIBRARIES

Designated Outstanding Collection

Library Services
British Library of Political and Economic Science
“Build and preserve distinctive collections to support research and learning, and represent a record of thought in the social sciences”

“Develop our digital library so that we are able to acquire, preserve and provide access to digital collections which match the strength of our print collections”

“...information repository services to support new forms of scholarly communication and enable the School to manage, disseminate and preserve these intellectual assets”
Risk model (DRAMBORA)

- Insufficient backups
- Media degradation or obsolescence
- Inadequate staff skills
- Loss of essential characteristics
- Infrastructure cannot support requirements

Activity overlooked or under resourced

- Failure of authenticity, integrity, provenance
- Cannot implement preservation plans

Loss of trust or reputation
Digital collections

There are known knowns; there are [digital collections] we know that we know.

There are known unknowns; that is to say there are [digital collections] that, we now know we don't know.

But there are also unknown unknowns – there are [digital collections] we do not know we don't know.
## LSE Library Collections

<table>
<thead>
<tr>
<th>Local</th>
<th>£</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Digital</strong></td>
<td></td>
</tr>
<tr>
<td>• Research outputs (publications, data)</td>
<td>• Journals</td>
</tr>
<tr>
<td>• Digitisation</td>
<td>• Books</td>
</tr>
<tr>
<td>• Public lectures</td>
<td>• Newspapers</td>
</tr>
<tr>
<td>• Web harvesting (websites, blogs, tweets)</td>
<td>• Statistics / data</td>
</tr>
<tr>
<td>• Archives (institutional, personal)</td>
<td></td>
</tr>
<tr>
<td>• Theses</td>
<td></td>
</tr>
<tr>
<td>• Official publications</td>
<td></td>
</tr>
<tr>
<td><strong>Physical</strong></td>
<td></td>
</tr>
<tr>
<td>• Archives (institutional, personal)</td>
<td>• Monographs</td>
</tr>
<tr>
<td>• Theses</td>
<td>• Journals</td>
</tr>
<tr>
<td>• Official publications</td>
<td>• Pamphlets</td>
</tr>
<tr>
<td></td>
<td>• Newspapers</td>
</tr>
<tr>
<td></td>
<td>• Statistics</td>
</tr>
<tr>
<td></td>
<td>• Microfilm</td>
</tr>
</tbody>
</table>
## Preservation responsibility

<table>
<thead>
<tr>
<th>Local</th>
<th>£</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital</td>
<td>• LSE Digital Library</td>
</tr>
</tbody>
</table>
| Physical | • Archives Services  
• Print Collections |
| • Consortia  
  • LOCKSS  
  • Portico |
| • Print Collections  
• Consortium  
  • UKRR |
Digital collections

• Increasing volume and diversity
  o Text, image, audio/video, Web/social, data

Institutional repository: additions per month

Archives: new hybrid or digital additions per year
Digital collections

• Increasing volume and diversity
  ○ Text, image, audio/video, Web/social, data

Physical collections: metres added per year
Digital collections

- Increasing volume and diversity
  - Text, image, audio/video, Web/social, data

Size (GB)

(Known) formats

- Public lectures
- Digitised
- Repository
- Digital archives
Digital collections

• Archives: forensic imaging (legacy media)
Digital collections

- Archives: forensic imaging (legacy media)

60 collections (hybrid and digital archives), total size 70GB
14,829 files, average 247 files per collection
Digital collections

• Archives: format profiling (legacy media)

- 87% known puid
- 13% unknown puid

- 81% known mimetype
- 19% unknown mimetype

60 collections (hybrid and digital archives), total size 70GB
14,829 files, average 247 files per collection
Digital collections

- Archives: format profiling (legacy media)

60 collections (hybrid and digital archives), total size 70GB
14,829 files, average 247 files per collection
LSE Digital Library

Who? Bibliographic Services, Archives
What? Existing systems/interfaces
How? Skills in place, resources in place
Who?  Print Collection Preservation, Archives
What?  Open source tools, new interface
How?  Need training, resources in place
LSE Digital Library

Who?  Print Collection Preservation
What?  Open source tools, new interface
How?  Need training, resources in place
Who? Library IT
What? Storage, backups, network
How? Skills in place, need resourcing
**Who?**
External consultants, DL team

**What?**
IA/UX/design

**How?**
Need skills, need resourcing
The Iceberg Model of Digital Libraries

- interfaces
- collections/objects
- workflows
- systems
- storage
- digital preservation
The Iceberg Model of Digital Libraries

- interfaces
- collections/objects
- workflows
- systems
- storage
- digital preservation
Roles and responsibilities

Senior Management
- Strategy
- Resources

Collection Services
- Preservation
- Description
- Infrastructure

Archive Services
- Collection development
- Description
- Preservation

Digital Library Team
- Policy
- Skills / expertise
- Innovation / projects

Academic Services
- Collection development
- Information skills training
Users

students (UG/PGT/PGR), staff (academic/research/teaching/support), visitors, alumni, external students, family/local historians, members of public, commercial users, media, school teachers, other information pros, biographers, depositors/donors, prospective staff/students, funders, picture researchers, competitors
Users

• Digital preservation awareness
  o Researcher/academic support librarian communities
  o Enhanced information literacy training (DICE project)

• IA/UX/design
  o Requirements for online/mobile access + user testing

• Content licensing policy
  o CC-BY-NC-SA for digitised/institutional collections

• Updated archive deposit agreement
Digital Communications Enhancement project (DICE)

Start Early | Explain It | Store It Safely | Share It

Training materials to raise awareness of digital preservation in researcher community and library training providers

Promotional leaflet, self-study training course, presentations for trainer-led sessions

http://lasedice.wordpress.com/
What lessons have we learnt?

• Demonstrating value is ongoing process
• Build room to manoeuvre, use outcomes to demonstrate value
• Core focus is important (innovation), but embedding is essential (service development)
• Community is the only way to do this
Thank you!

LSE Digital Library: digital.library.lse.ac.uk

PhoneBooth: ise.ac.uk/PhoneBooth

DICE: Isedice.wordpress.com

me: twitter.com/digitalfay or e.fay@lse.ac.uk
Image credits

*Egosiliqua malusymphonicus Guts* © Christopher Locke
(used with permission) http://heartlessmachine.com/

*Iceberg* (Public Domain) Wuzur

http://commons.wikimedia.org/wiki/File:Iceberg.svg

*National Research Libraries logo* © [unknown]

All other images (CC-BY-NC-SA) LSE Library