Ian Rowlands and David Nicholas
Thursday 3 November 2011
XXXI Annual Charleston Conference
About the Charleston Observatory
Engaging the library and publishing communities

The Observatory, established in 2009, is a mechanism by which exciting ideas raised at the Charleston Conference can be researched and the results reported back to provide continuity and build.

It is a place where evidence can be collected globally in a robust manner and where all the key information stakeholders (librarians, publishers, agents and academics) can come together and share data for the benefit of all.

The Observatory's first project (2009) looked at the impact of the world-wide recession on libraries. Last year, the Observatory considered social media and how they are impacting on research practice. Both studies received widespread coverage and generated a series of reports and peer-reviewed publications.
Aims of this study

Strategic directions for digital repositories

**Aims**

To understand what library directors (and researchers) see as the goals of digital repositories;

To identify the critical success factors behind successful digital repositories;

To assess the wider impact of digital repositories.

**Digital repositories**

Institutional repositories

Subject-based repositories

Format-based repositories (e.g. e-theses)

**Research design**

Online survey of library directors

Online survey of researchers (next stage)

Focus groups and interviews (next stage)
About the survey
A global survey of facts and opinions

Survey sample
153 library directors had completed the survey by 4 Nov, representing 7.2% of all 2,126 OpenDOAR repositories.
Responses from universities, colleges, medical schools, government and charities in 35 countries.

Research partners
CIBER Research Limited
Emerald
Elsevier
Institute of Physics Publishing
Research Information Network
Does your institution have a digital repository?

Filtered questionnaire structure

- Have a digital repository (73%)
  - Operational facts
  - Goals
  - Progress against goals
  - Benefits
  - Impacts

- At the planning stage (21%)
  - Perceived benefits
  - Perceived impacts

- Not making plans (6%)
  - Why not?
What is your main reason for not having a repository?

Main reasons

- Unconvinced of the benefits: 18%
- Lack of management support: 9%
- Limited resources: 36%
- Not highly research-intensive: 36%

This is really for the report not for public consumption. It's really the greasy chute to get rid of inappropriate respondents - mainly VERY small US liberal arts colleges.

It would be nice to have an idea about what libraries that have digital repositories consider the main purpose (record of institution’s publications, institutional archives? highlight specialised collections or projects?)
Survey findings
What is the main collection focus of your repository?

Percentages of institutions with repositories

<table>
<thead>
<tr>
<th>Collection Focus</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>We collect all or most research outputs</td>
<td>72.5</td>
</tr>
<tr>
<td>We only collect particular formats (e.g. theses)</td>
<td>23.5</td>
</tr>
<tr>
<td>We only collect particular subjects</td>
<td>4.0</td>
</tr>
</tbody>
</table>

- We collect what people contribute
- ... trying to expand, difficult to get faculty to submit
- We collect documents and objects that are valuable for our national heritage
How is your repository funded?
*Percentage split across all institutions with repositories*

<table>
<thead>
<tr>
<th>Source of Funding</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular budget line for your institution’s library</td>
<td>69.0</td>
</tr>
<tr>
<td>Special initiative supported by your institution</td>
<td>14.4</td>
</tr>
<tr>
<td>Departmental budgets</td>
<td>6.5</td>
</tr>
<tr>
<td>Grant from an external source</td>
<td>5.8</td>
</tr>
<tr>
<td>Contributory basis / IR membership</td>
<td>1.2</td>
</tr>
<tr>
<td>Publication budget</td>
<td>0.5</td>
</tr>
<tr>
<td>Other sources</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Recurrent spending on digital repositories averages around 1.8 percent of library operational budgets.
How many staff work on your repository?

Full time equivalent (FTE) staff

- Less than 1: 24.6%
- 1 to 2: 46.4%
- 3 to 4: 22.5%
- 5 or more: 6.5%
What types of content does your repository manage?

Percentages of repositories currently holding these resource types

<table>
<thead>
<tr>
<th>Type of Content</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal articles</td>
<td>78.9</td>
</tr>
<tr>
<td>Conference papers</td>
<td>74.4</td>
</tr>
<tr>
<td>Book chapters</td>
<td>70.7</td>
</tr>
<tr>
<td>Books or monographs</td>
<td>62.4</td>
</tr>
<tr>
<td>Technical reports</td>
<td>60.2</td>
</tr>
<tr>
<td>Working papers</td>
<td>52.6</td>
</tr>
<tr>
<td>Research datasets</td>
<td>51.9</td>
</tr>
<tr>
<td>Computer software</td>
<td>51.1</td>
</tr>
<tr>
<td>Images or photographs</td>
<td>36.6</td>
</tr>
<tr>
<td>Sound recordings</td>
<td>32.3</td>
</tr>
<tr>
<td>Video recordings</td>
<td>30.1</td>
</tr>
<tr>
<td>PhD or Masters' theses</td>
<td>26.6</td>
</tr>
<tr>
<td>Administrative records</td>
<td>23.3</td>
</tr>
<tr>
<td>Learning objects</td>
<td>21.8</td>
</tr>
<tr>
<td>Patents</td>
<td>20.3</td>
</tr>
<tr>
<td>News or press coverage</td>
<td>15.0</td>
</tr>
<tr>
<td>Annual reports</td>
<td>13.5</td>
</tr>
<tr>
<td>Metadata-only records</td>
<td>11.3</td>
</tr>
<tr>
<td>Special collections</td>
<td>8.3</td>
</tr>
</tbody>
</table>

Other formats collected include:
- blog posts, interview notes, student magazines, archives of distinguished faculty, honours theses, scanned herbarium materials ...

Single most important resource?
- Journal articles (44.1%)
- PhD or Master’s theses (35.3%)
- Special collections (11.8%)
What is your policy for including copyright-protected materials?

*Institutions with repositories: all resource types (tick as many as apply)*

<table>
<thead>
<tr>
<th>Policy</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>No copyright materials deposited</td>
<td>11.9</td>
</tr>
<tr>
<td>Accept preprints only</td>
<td>23.7</td>
</tr>
<tr>
<td>Include but password protected</td>
<td>24.6</td>
</tr>
<tr>
<td>Obtain permission</td>
<td>55.9</td>
</tr>
</tbody>
</table>

So, a majority accept copyright materials and address the issue by either putting password controls in place or seeking publisher permissions.

We ask depositors for assurances they have obtained permission.

Clearance is the responsibility of the depositor.

We add a citation and a link to the full version.
How do you maintain quality?
Institutions with repositories: all resource types

- Accept everything, without review: 69.5%
- Use peers to review quality: 14.7%
- Post-acceptance review or rating: 15.8%

Seems a rather passive approach on the part of librarians!

- All deposited content is curated by library staff.
- [Academic] communities establish their own criteria for what is acceptable to deposit.
- Content is mostly solicited, so review is built in.
Which versions of journal articles do you allow to be deposited?

Institutions that accept journal articles (tick as many as apply)

- Accepted MS before publication: 88.6%
- Final published version: 86.4%
- Author MS after peer review: 52.3%
- Author MS before peer review: 47.7%
Does your institution provide financial support for author pays?

<table>
<thead>
<tr>
<th>Institutions that accept journal articles</th>
<th>Yes</th>
<th>No</th>
<th>I’m not sure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>22.0</td>
<td>60.2</td>
<td>17.8</td>
</tr>
</tbody>
</table>
Do you actively support submission to subject-based repositories?

*Institutions that accept journal articles*

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Yes</strong></td>
<td>34.7</td>
</tr>
<tr>
<td><strong>No</strong></td>
<td>41.5</td>
</tr>
<tr>
<td><strong>I'm not sure</strong></td>
<td>23.7</td>
</tr>
</tbody>
</table>
Does your institution operate an institutional mandate?

Institutions that accept journal articles

- Yes: 28.0
- No: 55.9
- No but we plan to: 14.4
- I'm not sure: 1.7
What services do you offer to your users?

_Institutions with repositories (tick as many as apply)_

<table>
<thead>
<tr>
<th>Service</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assistance with deposit</td>
<td>87.2</td>
</tr>
<tr>
<td>Assistance with metadata creation and management</td>
<td>80.0</td>
</tr>
<tr>
<td>Assistance with copyright clearance</td>
<td>60.8</td>
</tr>
<tr>
<td>Download statistics</td>
<td>55.2</td>
</tr>
<tr>
<td>Personalised web page or cv</td>
<td>28.8</td>
</tr>
</tbody>
</table>

*Other services include:*
Institutional CVs and personalised web pages, publication lists, automatic addition of citation data, assistance with digitising materials, help with data management plans, information on most downloaded articles.
What do you think are the main advantages?

Mean ratings, where 0=Not important, 3=Very important

- Long-term preservation of your institution’s digital materials: 2.55
- Providing maximal access to the results of publicly funded research: 2.54
- Enhancing the external prestige of your institution: 2.41
- Better services to students inside your institution: 2.4
- Better services to learning communities outside your institution: 2.37
- Better services to researchers outside your institution: 2.33
- Maintaining control over your institution's intellectual capital: 2.26
- Contributing to the reform of scholarly communication and publishing: 2.22
- Contributing to the changing library culture more digital: 2.18
- Reducing the time between discovery and dissemination: 2.17
- Provision of identifiers for easier citability of digital materials: 2.01
- Registration of new ideas: 1.67
Tenure and promotion have become a very important reason that our faculty deposit ... and this will become even more of a focus as junior faculty (who are increasingly committed to sharing their outputs) become established and as what counts for tenure and promotion shifts to away from the published literature to a wider range of object content.
What do you think are the main disadvantages?

Mean ratings, where 0=Not important, 3=Very important

Confusion caused by different versions of the same material: 2.26
Fragmentation of access to the literature ('islands of content'): 2.17
Confusion and uncertainty over copyright issues: 2.04
Threatens society and commercial publishers: 1.96
Fear of plagiarism: 1.96
Not comprehensive: lack scale and critical mass: 1.71
Lack of awareness by users: 1.57
Long term funding and support for repositories uncertain: 1.52
Variable quality of material: no consistent peer review: 1.52
Lack of interoperability between repositories: 1.44
Costs of long-term preservation and digital curation: 1.20
Software difficult to use: 0.82

Sloppy repositories can be harmful, as they lower standard for scholarly communication.
Standardisation, clarification and simplification of publishers’ policies on repositories and open access is needed. Current situation very confusing for researchers to understand precisely what they are allowed to do with their research and this is a big barrier to use.
To what extent do you agree or disagree with these statements?

*Mean ratings, where -2=Strongly disagree and +2=Strongly agree*

**Institutional repositories ...**

- Impact negatively on publishers’ revenues: -0.58
- Become redundant as more material goes OA: -0.51
- The first steps to universities becoming digital presses: 0.15
- Raise visibility and are the publishers’ friend: 0.37
- IRs should be regional, country-wide or subject-based: 0.4
- Replace researchers linking from their own websites: 0.59

*I tend to disagree*  
*In the middle*  
*I tend to agree*
What library directors think repositories are for

Mean ratings, where 0=Not important and 3=Very important

- Make the literature more openly available: 2.36
- A research shop window for your institution: 2.31
- Long term preservation and curation: 2.30
- Change the library culture moving it into the digital age: 2.00
- Provide information on research productivity: 1.48
- Manage informal literature better: 1.42
- Provide access to rare materials: 1.37
- First steps on the path to becoming a digital publisher: 0.99
While some on both sides of the open access debates do tend to see repositories only as tools for OA, I think most in the repository communities see a much more nuanced role for institutional and subject repositories as part of a much larger system that includes publishers, societies, etc.
Progress against goals

Mean ratings, where 0=Not important or Not successful and 3=Very important or Very successful.
Progress against goals

Mean ratings, where 0=Not important or Not successful and 3=Very important or Very successful
Progress against goals

Mean ratings, where 0=Not important or Not successful and 3=Very important or Very successful
Priorities for repository development

Mean ratings, where 0=Not on the agenda and 3=High priority

- Attracting more use: 2.63
- Raising awareness of repository: 2.63
- Getting more stuff in: 2.49
- Integration with other systems: 2.27
- Greater interoperability: 2.07
- An institutional mandate: 2.04
- Better usage statistics: 1.99
- Preservation and curation: 1.98
- Raising awareness of preservation: 1.74
- Managing research datasets: 1.74
- More coherent policies: 1.71
- Extend scope of repository: 1.69
What are the most important critical success factors?

**Mean ratings on a 5-point scale**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>The repository should have a clear purpose</td>
<td>4.2</td>
</tr>
<tr>
<td>People should want to put material in</td>
<td>3.2</td>
</tr>
<tr>
<td>People should want to get material out</td>
<td>2.8</td>
</tr>
<tr>
<td>The repository should have a strong sense of ownership</td>
<td>1.8</td>
</tr>
<tr>
<td>The scope of the repository should be clear</td>
<td>1.7</td>
</tr>
<tr>
<td>The community using the repository should be clear</td>
<td>1.4</td>
</tr>
</tbody>
</table>

The importance of robust business plans and sustainability through institutional support and formal policy making cannot be over-emphasised.
How much impact has your repository made?

Percentages of institutions

<table>
<thead>
<tr>
<th>Impact Level</th>
<th>Within your organisation</th>
<th>Outside your organisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>No impact</td>
<td>5.0</td>
<td>11.9</td>
</tr>
<tr>
<td>Small impact</td>
<td>42.6</td>
<td>43.6</td>
</tr>
<tr>
<td>Significant impact</td>
<td>35.6</td>
<td>32.7</td>
</tr>
<tr>
<td>Highly significant impact</td>
<td>16.8</td>
<td>11.9</td>
</tr>
</tbody>
</table>
Compared with today, will repositories become more important?

Percentages of institutions

Collaboration among smaller institutions will become more important for financial and staffing reasons.

They are the future and libraries should lead the way - it’s what we’ve always done but with digital resources not print.

I am personally unconvinced of the value of institutional repositories ... I think academics are more comfortable with subject repositories.

<table>
<thead>
<tr>
<th>Much less important</th>
<th>Less important</th>
<th>About the same</th>
<th>More important</th>
<th>Much more important</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Institutional repositories</strong></td>
<td>2.3%</td>
<td>3.0%</td>
<td>19.7%</td>
<td>53.8%</td>
</tr>
<tr>
<td><strong>Subject repositories</strong></td>
<td>57.6%</td>
<td>18.9%</td>
<td>20.5%</td>
<td>5.3%</td>
</tr>
</tbody>
</table>