APSR - Preserving the past: maintaining the future

Margaret Henty
National Services Program Coordinator, APSR
An Innovative Action Plan for the Future

The Systemic Infrastructure Initiative

Supported by:

An Innovative Action Plan for the Future

The Systemic Infrastructure Initiative
What APSR is...

- A Partnership to Promote Excellence in Digital Repositories
  - Document best practice
  - Address strategic issues
  - Consider the Australian context
  - Stimulate and share experience
Priority Projects S2 2005

- Sustainable models for large data sets
- Cost models and business models for digital repositories
- Good practice for byte stream storage
- Risk assessment models
Repository software

- Archimede
- bepress
- CDSware
- CONTENTdm

- DSpace
- Fedora
- Greenstone
- Open Repository
What is an institutional repository?

“An institutional repository is a database with a set of services to capture, store, index, preserve and redistribute a university’s scholarly research in digital formats.”

Source: LEADIRS Workbook, MIT

...digital repositories
Table 1: Academic institutional Repositories; state of the art in 13 countries - June 2005

<table>
<thead>
<tr>
<th>Countries</th>
<th>Number of IRs</th>
<th>Number of universities</th>
<th>Percentage of universities with an IR</th>
<th>Average number of documents per IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>37</td>
<td>39</td>
<td>95</td>
<td>n.r.</td>
</tr>
<tr>
<td>Belgium</td>
<td>8</td>
<td>15</td>
<td>53</td>
<td>450</td>
</tr>
<tr>
<td>Canada</td>
<td>31</td>
<td>n.r.</td>
<td>-</td>
<td>500</td>
</tr>
<tr>
<td>Denmark</td>
<td>6</td>
<td>12</td>
<td>50</td>
<td>n.r.</td>
</tr>
<tr>
<td>Finland</td>
<td>1</td>
<td>21</td>
<td>5</td>
<td>n.r.</td>
</tr>
<tr>
<td>France</td>
<td>23</td>
<td>85</td>
<td>27</td>
<td>1000</td>
</tr>
<tr>
<td>Germany</td>
<td>103</td>
<td>80</td>
<td>100</td>
<td>300</td>
</tr>
<tr>
<td>Italy</td>
<td>17</td>
<td>77</td>
<td>22</td>
<td>300</td>
</tr>
<tr>
<td>Norway</td>
<td>7</td>
<td>6</td>
<td>100</td>
<td>n.r.</td>
</tr>
<tr>
<td>Sweden</td>
<td>25</td>
<td>39</td>
<td>64</td>
<td>400</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>16</td>
<td>13</td>
<td>100</td>
<td>3,000 / 12,500</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>31</td>
<td>144</td>
<td>22</td>
<td>240</td>
</tr>
<tr>
<td>United States of America</td>
<td>n.r.</td>
<td>261</td>
<td>-</td>
<td>n.r.</td>
</tr>
</tbody>
</table>

Scholarly Communications

• Changing information seeking practices
  ▪ The Amazooglebay phenomenon
• eResearch
  ▪ National and international research collaborations
  ▪ Grid Computing/Cyberinfrastructure
• New epublication types, outlets and economics
‘Born Digital’ Resources

• Data sets
• Communications (email, discussion forums etc.)
• New publications (eTheses, ePrints, eBooks, etc.)
• New environments (Wiki, Blogs, eLearning Systems, Digital Video, Computer Games, Virtual Reality Environments etc.)
What does this mean for researchers

• Open Access (OA) articles are cited between 2.5 and 5.8 times MORE than the non-OA articles

• Overall, citations to online articles are 2.6 times greater than offline articles. On a year by year basis, online articles are cited 4.5 times more often than offline articles

• Restrictive access policies cut readership of electronic research journal articles by a factor of two

Source: Danny Kingsley, July 2005
Research Data Collections Survey - formats

<table>
<thead>
<tr>
<th>ANU and USyd</th>
<th>Open</th>
<th>Proprietary</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formats used to access data</td>
<td>Jpeg, html, xml, tiff, txt, png, seed, sac, zdf, fits, gif</td>
<td>Pdf, bwf, wav, doc, mps, mpeg2, xls, gcm</td>
<td>Application specific community developed format for physics</td>
</tr>
<tr>
<td>Formats used to store data</td>
<td>Tiff, xml, ascii, html, png, seed, sac, zdf, fits, gif</td>
<td>Pdf, bwf, wav, doc, mps, mpeg2, xls, gcm</td>
<td>Application specific community developed format for physics</td>
</tr>
</tbody>
</table>
Case study: Oral Epics

- Varied data - 6 epics, each 37 hours of tape and 1000 pages of printed text
- Varied products arising from the data
- Formats - analogue and digital, open source and proprietary
- Software - proprietary, discipline specific
- Copyright
- Privacy
Case study: Tiger Hunt
Open Repositories 2005

- Open Repositories 2006 - an interlinked set of events collaboratively convened by APSR with the ARROW project, MAMS, Macquarie University, and the University of Technology Sydney.
- DSpace User Group Meeting
- Digital Repositories Forum: The Well-integrated Repository
- Workshop: Managing Openness in Digital Repositories
- January 31 to February 3
- University of Sydney
Keep in touch

