2015 NDSA Web Archiving Survey Report Highlights

Nicholas Taylor (@nullhandle)
Web Archiving Service Manager
Stanford University Libraries

Archives 2016: Web Archiving Roundtable
August 3, 2016
NDSA Web Archiving Survey Working Group

Jefferson Bailey
Internet Archive / Archive-It

Edward McCain
University of Missouri

Abbey Potter
Library of Congress

Abbie Grotke
Library of Congress

Christie Moffatt
National Library of Medicine

Nicholas Taylor
Stanford University Libraries
NDSA Web Archiving survey background

2011
• 78 respondents
• program info
• tools/services
• access
• policies

2013
• 92 respondents
• program info
  • staff, metrics, skills, content concerns
• tools/services
• access/discovery
  • new options
• policies
  • embargo, social media, robots.txt, resources

2015
• 106 respondents
• program info
  • areas of progress, collaboration interests + barriers
• tools/services
  • replication targets
• access/discovery
  • researchers
• policies
Respondent Characteristics
increasing proportion of universities
SAA WebArch RT tops group affiliations

<table>
<thead>
<tr>
<th>group</th>
<th>2011</th>
<th>2013</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIPC netpreserve.org</td>
<td>8%</td>
<td>7%</td>
<td>8%</td>
</tr>
<tr>
<td>NDSA</td>
<td>31%</td>
<td>33%</td>
<td>39%</td>
</tr>
<tr>
<td>Web Archiving SAA Roundtable</td>
<td>45%</td>
<td>50%</td>
<td></td>
</tr>
</tbody>
</table>
growth in staffing at the middle
shift from pilot to production (all responses)
shift from pilot to production (longitudinal responses)
increased perceptions of progress

- **2013**
  - Some progress: 43%
  - Significant progress: 29%
  - About the same: 15%
  - Slightly worse off: 2%
  - Much worse off: 1%
  - Other: 10%

- **2015**
  - Some progress: 22%
  - Significant progress: 49%
  - About the same: 19%
  - Slightly worse off: 0%
  - Much worse off: 0%
  - Other: 10%
perceived progress on data capture, appraisal, vision
low perceived progress on access, metadata, QA
Archiving Focus

“Ant Farm Media Van v.08 (Time Capsule) in Bellewether at Southern Exposure” by Steve Rhodes under CC BY-NC-SA 2.0
more archiving of own content
access, cost, quality are top valued metrics
tools, appraisal, QA are top valued skills
collaboration interest in QA, capture, metadata

- Risk management: 40.91%
- Capture optimization: 52.27%
- Collection development: 31.82%
- APIs: 27.27%
- Metadata standards: 51.14%
- QA: 64.77%
- Tool development: 46.59%
- Other: 3.41%
Tools and Services

“Photocopier” by Joriel "Joz" Jimenez under CC BY-NC-ND 2.0
more using multiple archiving solutions
flat data transfer from service provider

- Transfer data:
  - 2011: 19.15%
  - 2013: 20.29%
  - 2015: 20.27%

- Do not transfer data:
  - 2011: 80.85%
  - 2013: 79.71%
  - 2015: 79.73%
transferring data locally + to other services

- Local: 58.82%
- External: 47.06%
- Both: 5.88%
trust provider + building locally for data transfer
Archiving Policies
more conditional handling of robots.txt
policies of other orgs key for policy-making

- ARL Code of Best Practices
- Oakland Archive Policy
- Section 108 Study Group
- Legal counsel
- Statutory authority
- Policies of other orgs
- Previous NDSA surveys
- Other

2013 vs 2015
Landscape Summary
overall picture

• generally **positive perceptions** of progress
• more moves to **production**
• growing proportion of **universities**
• growing proportion of **focus on own content**
• staffing **largely remains fractional w/ minor growth**
• majority use external service but **hints of hybrid approaches**
• more comfort w/ **conditional policy approaches**
recurring themes: progress

• data capture
  • highest perceived progress
  • 2nd highest collaboration interest

• appraisal
  • 2nd highest perceived progress
  • 2nd highest valued skill
recurring themes: opportunities

• quality assurance
  • highest collaboration interest
  • 3rd highest desired skill
  • 3rd lowest perceived progress

• access
  • lowest perceived progress
  • tied for highest valued metric

• metadata
  • 2nd lowest perceived progress
  • 3rd highest collaboration interest
Nicholas Taylor
@nullhandle