Blogger Perspectives on Digital Preservation: Attributes, Behaviors, and Preferences

Carolyn Hank
School of Information and Library Science
University of North Carolina at Chapel Hill

Future of Today’s Legal Scholarship
Georgetown Law Library
July 25, 2009
Bloggers

57%

post several posts a week
Bloggers

~5 hours working on blogs a week

(Mode = 1 - 2 hours)
Bloggers

80%

read other blogs daily
Method…

Web Survey
- 46 questions
- Active bloggers

Snowball Technique

33 core group

223 completed surveys
Bloggers

43% have blogged for 3 years or more
41% contribute to two or more single-author blogs

47% contribute to two or more collaborative blogs
Identity

56% 20% 24%

Real name | Both | Pseudonym
Audiences

- Customer: 7.40%
- Family: 27.40%
- Myself*: 40.90%
- General Public: 54.40%
- Friends: 55.80%
- Colleagues/Professional Peers: 63.70%
71% agree that their own blogs should be preserved

36% agree that every blog should be preserved
Hosting

- Blog service provider: 86%
- Own server: 16.10%
- Commercial server: 14.50%
- Server at work: 13.50%
- Server at school: 3.60%
- Social networking site: 7.30%
- Commercial blog provider: 3.60%
- Aggregated blogs: 2.10%
- Other non-commercial server: 1.60%
Often use

**blog editor**

Alternative way:
Word processor (desktop and online)
Email composer
Mobile phone/PDA
Desktop plug-in blog editor
Instant messenger
2% require password to view ALL posts

19% require password to view CERTAIN posts
96% edited content after it was posted

Spelling & grammatical errors
Rephrasing
Updating links
Erroneous information
DELETE Posts

39% deleted content after it was posted

Remove opinions no longer held
Remove info no longer correct
Erroneous information
23% have deleted an entire blog
84% have changed the appearance of their blogs.
Duplication

67% duplicate some posts
(18% duplicate all)

36% duplicate some comments
(12% duplicate all)
Responsibility

- Blog authors/owners
- IT department
- Software providers
- Libraries/archives
- Public trust
- Search engine

Comparison of responsibility distribution:
- Own blog vs. Blogs in general
54% Would not pay to preserve their OWN blog

ANNUAL $1 – 200
Avg. $36.82
Characteristics

- Unique/innovative content
- Cultural significance of bloggers
- Preservation preference
Characteristics

* Number of readers
* Age of blog
* Size of blog

Unimportant
Selection Criteria

- **context**: representative & significant
- **content**: subject & attribute
- **style**: expression & utility
- **format**: elements (next slide)
- **audience**: readership & preference
- **author**: credentials & affiliations
Selection Criteria

context
content
style
format
audience
author

Text Photos
Video Audio Other_multimedia
Comments Keywords Links
Imported_images Bookmarks
Imported_tags Feeds
<table>
<thead>
<tr>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>context</strong></td>
</tr>
<tr>
<td><strong>content</strong></td>
</tr>
<tr>
<td><strong>style</strong></td>
</tr>
<tr>
<td><strong>format</strong></td>
</tr>
<tr>
<td><strong>audience</strong></td>
</tr>
<tr>
<td><strong>author</strong></td>
</tr>
</tbody>
</table>
“[S]hout, plead to my boss not to fire me!”

“I have no sense of identity invested in my Weblog, so I probably wouldn’t do anything, but simply let it go.”

“I would seek back-up files from the sys admin, as I have a hard time believing that a blog would just disappear.”

“[D]rink some coffee and do other hobby.”
Conclusions

Findings
Bloggers are interested
Save some but not all
New content added
Old content altered
Personal responsibility
Defining roles of others

Future
Selection and appraisal
Responsibility
Access scenarios
Versioning
Intellectual property
Changing roles
THANK YOU

Acknowledgements…
Songphan Choemprayong
Laura Sheble

Resources…
http://archivepress.ulcc.ac.uk/