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

post several posts a week

57%
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

223 completed surveys

core group
33
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

Real name | Both | Pseudonym

56% | 20% | 24%
Preservation

71% agree that their own blogs should be preserved.

36% agree that every blog should be preserved.
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
DELETE Entire Blog

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

Graph showing the responsibility distribution for different roles in both 'Own blog' and '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
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
Exclusion Criteria

- **context**: ephemeral & insignificant
- **content**: personal & inappropriate
- **style**: hate speech
- **format**: splog, storage concerns
- **audience**: popularity
- **author**: personal
“[S]hout, plead to my boss not to fire me!”

“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.”

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