Present and Future Services for the Digital Library

Searching, Linking, and Meta Tags

By Professor Greg R. Notess
Montana State University-Bozeman
Opportunities & Challenges

- New access to library material
- Reach a larger audience
- Internet access
  - World-wide audience
  - Local access at home and work
- Difficulties
  - Access and discovery
  - Changing technology
Search Engines

- Primary finding tool
  - But miss many digital collections

- Ignore most meta tags

- Dependent on link patterns
Digital Library Services Should Consider

- Actual current meta tag handling
- Potential future meta tag handling
- Exploring search engines limitations
- Implement alternate finding strategies
Searching Digital Libraries

- Literature focus on
  - Internal search
    - Gustman 2002
  - Specialized search of groups
    - Detlor and Arsenault 2002
    - Petinot, et. al. 2003
  - Distributed digital library
    - Nicholson and Shiri 2003

- Very important for local users
  - But very different from general Web search engines
Search Engine Accessible

- No password or IP limit
- Not blocked by robots.txt
- No noindex meta tag
- Linked from elsewhere
No Robots

- `robotstxt.org/wc/robots.html`
- `/robots.txt` file (root directory)
  - User-agent: *
  - Disallow: /
- No robot meta tag
  - In header

```html
<META NAME="ROBOTS" CONTENT="NOINDEX, NOFOLLOW">
```
Consider two access points

1. Home page
   Include metadata for collection

2. Every page for each individual record
   Many more, so may be found more
   Each page may become the entry point to the collection
Example 1: Emma Goldman Papers

EMMA GOLDMAN (1869-1940) stands as a major figure in the history of American radicalism and feminism. An influential and well-known anarchist of her day, Goldman was an early advocate of free speech, birth control, women’s equality and independence, and union organization. Her criticism of mandatory conscription of young men into the military during World War I led to a two-year imprisonment, followed by her deportation in 1919. For the rest of her life until her death in 1940, she continued to participate in the social and political movements of her age, from the Russian Revolution to the Spanish Civil War.

PROJECT NEWS:

HOLIDAY LETTER. Read the 2003 holiday letter from the Emma Goldman Papers.

Letter from Harry Weinberger to Emma Goldman

Harry Weinberger to Emma Goldman, July 12, 1919

[New York] July 12th, 1919

My dear E.G.:-

What your nephew suggests can be done, but it will not gain any results or be of any avail. I hope father correct, that is, first 89, second 94, but the 94 date is the one that counts, and minors, twenty-one are only benefitted by it. I think, however, that we can rely on Kersner. I have spent three days this week, finishing up the law on that end, and though the thread is weak, it may have to take into consideration the prejudice that we will meet, though that itself may even be snapping the thread will appear like persecution and not like law.
Example 2: Pali-English Dictionary
In or Out of Search Engines?

- Home Page In
Individual Records Not

Appatissa (p. 57) (& appatissa) (adj.) [a + pati + srũ] not doc agarava A II.20; III.7 sq., 14 sq., 247, 439. Appatissa--vasa an unappatissa.
Linking

- Important for getting included
- Search engine optimization
  - Aims to raise relevance scores
  - Ecommerce mostly
- Some strategies can be used by digital libraries
For example

- Pali-English dictionary entries
- Not just search form
- Link each record separately from list
Titles

- Important in linking & relevance rank
- Some collections: same title on each record
- Instead, use metadata in title
  - Derived from individual record
Picture Australia Collection

- Access to several picture collections
- Some use picture title as HTML title
- Others just have

“PICMAN Database – Full Image Display”
Metadata Importance

- Internal navigation
- Internal searching
- Resource discovery
- Different from meta tags
Meta Tags

- Old Advice: Use to help ranking
- Now: Make no difference
- Hidden in header tag of HTML
  
  ```html
  <meta name="description" content="Text">
  ```
- Dublin Core counterpart
  
  ```html
  <meta name="DC.Description" content="Text">
  ```
Key Point:

Most search engines ignore and do not index any meta tags.
Why?

- Most use was spam
- Meta tags did not reflect content
- So . . .
  - Use metadata to create visible descriptions
  - Use meta tags for internal needs
Searching, Linking, Meta Tags

- Understanding these helps digital libraries
  - Help users find resources
  - Save the time of the searcher
  - Avoids wasted effort