The Death of Metadata

by Jeffrey Beall
Outline

• How I came to be invited here
• Provide a background on metadata, specifically the Dublin Core and MARC schema
• Explain what a good metadata implementation needs to do
• Explain why full-text searching doesn’t work very well for serious information seeking needs
• Explain why Dublin Core doesn’t work very well
• Talk about points of comparison for metadata schema
• Why MARC is successful; MODS
Auraria Library, University of Colorado at Denver and Health Sciences Center, Downtown Denver Campus
Dublin Core: An Obituary

- **Abstract**: The Dublin Core Metadata Standard was conceived illegitimately, had a troubled life, and has finally met its demise. Developed as a tool for online resource discovery, the standard waned after the arrival of Google. Because its fields were designed with a lack of specificity, Dublin Core suffered from nonstandard data elements and poor interoperability. Also, the poor organization behind the initiative contributed to its failure. Dublin Core will likely soon be replaced by an emerging standard, the Metadata Object Description Schema.
“While it's always nice to be cited, I'd hate for anyone to think I agree with Jeffrey Beall's article.”

“Anyone who is familiar with the DCMI process, or with the people who led the DCMI or the people who spent, and continue to spend, their time and talents defining and promulgating the standard, could not honestly write some of the comments you quote below.”
“I, too, read this article with raised eyebrows.”

In my article, I said this:

"The designers were managers rather than practitioners and held organizational meetings in far-flung parts of the world, such as Finland and Australia, ensuring that day-to-day professionals who actually worked with metadata would be excluded."
And one person responded:

“Ironically, most people deeply involved with Dublin Core are at this moment in "far-flung" Shanghai for this year's meeting, but I do wonder how others see the DCMI and if this author expresses a common distrust/distaste of the DC activities.”
“I'm at the DC2004 conference in Shanghai, and as you can imagine, this article has created a great deal of irritation.”

“First, I'm surprised that the editors of Library Hi Tech News would even allow such an article to be published in the first place. The author is extremely ignorant about DC and DCMI. I could not begin to list the untrue statements that he asserts in the article, there are so many of them!

(continued)
The author fails to provide an [sic] empirical evidence to support any of his statements and he is hopelessly misinformed, misguided, and ignorant about the history, purpose, use, and utility of DC.”

“In addition to the inaccuracies about DCMI, the author is clearly misinformed about Google, OCLC, the workings of search engines, CORC, and MODS. It is no wonder how [sic] he could produce a ridiculous article that is based on fiction, not facts.”
The History of Metadata

• Mid 1990s: birth of metadata
• Sherry Vellucci: bibliographic data becomes metadata
• Who invented metadata?
• Metadata = Dublin Core
• Dublin Core: the anti-library standard
Tower of Babel of Metadata Schemes
What Metadata Ought to Do

- Add value to information
- Work with a system
- Provide collocation
1 Colorado History --> Authority Record ............... 1 entry
2 Colorado History ........................................ 146 entries
3 Colorado History 1876 1950 --> Authority Record .... 1 entry
4 Colorado History 1876 1950 --> See also MILK CREEK, BATTLE OF 1 entry
5 Colorado History 1876 1950 ............................. 20 entries
6 Colorado History 1876 1950 Biography .................. 1 entry
7 Colorado History 1876 1950 Pictorial Works ............ 1 entry
8 Colorado History 1951 --> Authority Record ........... 1 entry
9 Colorado History 1951 ..................................... 2 entries
10 Colorado History Anecdotes ............................... 5 entries
11 Colorado History Anecdotes Juvenile Literature ........ 1 entry
12 Colorado History Audiotape Catalogs .................... 1 entry
13 Colorado History Bibliography ........................... 1 entry
14 Colorado History Bibliography Catalogs ................ 1 entry
15 Colorado History Chronology ............................. 1 entry
16 Colorado History Civil War 1861 1865 --> Authority Record 1 entry
17 Colorado History Civil War 1861 1865 .................... 8 entries
18 Colorado History Curricula Standards .................... 1 entry
19 Colorado History Exhibitions ............................. 1 entry
20 Colorado History Fiction ................................. 2 entries
21 Colorado History Juvenile Literature .................... 11 entries
22 Colorado History Local .................................... 28 entries
23 Colorado History Local Bibliography .................... 1 entry
24 Colorado History Local Exhibitions ....................... 1 entry
Colorado Historical Society
The Colorado Historical Society is a state agency that collects, preserves and interprets the history of Colorado and the West through museum exhibits, ...
www.coloradohistory.org - 5k - Cached - Similar pages

Colorado's Historic Sites - Colorado History Museum
Colorado History Museum 1300 Broadway Denver, Colorado 80203 ... Discover the diversity of Colorado from its earliest inhabitants to its pioneer families. ...
www.coloradohistory.org/hist_sites/CHM/Colorado_History_Museum.htm - 6k - Cached - Similar pages
[ More results from www.coloradohistory.org ]

Colorado History: Your best source for Colorado History ...
Colorado History: Your best source for Colorado History information, as well as information about each county, museums, historical societies, ...
www.coloradohistory.com/ - 21k - Cached - Similar pages

Colorado History: Poker room, Faro & gambling in the old Saloon.
Colorado History: Poker room, Faro & gambling in the old west Saloon.
www.coloradohistory.com/pokerroom.asp - 21k - Cached - Similar pages
[ More results from www.coloradohistory.com ]

Colorado's Kids and Students Page - History - Past and Present People
Information kids can use about Colorado, including places to go and things to do.
www.state.co.us/kids/history.htm - 4k - Cached - Similar pages

Sponsored Links

Colorado History
Learn about the history of your state in historical newspapers.
Ancestry.com

Colorado History
Products and info about Colorado History.
eBay.com

Colorado History
PriceGrabber.com

Colorado history
You'll find books on this topic and more from Dover Publications.
www.DoverPublications.com

Pikes Peak History DVD
"Legends of Pikes Peak" DVD shows the rich history at this mountain.
www.LittleDogCinema.com
What Metadata Ought to Do (continued)

- Search precision
- Be easily sharable
- Provide consistency (authority control)
Consistency?

- Montgomery, L. M. (Lucy Maud), 1874-1942
- Montgomery, Lucy Maud
- Montgomery, L. M.
- Montgomery, L. M., b. 1874
- Montgomery, L. M., 1874-
- Montgomery, Lucille, d. 1942
- Montgomery, Lucy, 1874-1942
- Montgomery, L., n. 1942
- MacDonald, Lucy Maud Montgomery, 1874-1942
Full-text Searching: An Obituary

• False hits
  – Mercury (Planet)
  – Mercury
  – Mercury (Roman deity)
  – Mercury automobile
Full-text (continued)

- Different language or spelling
- Different term
- Term not used
- Term not readable
Why Dublin Core Doesn’t Work Well

• DC tries to split the difference between full-text searching and standard cataloging
• Creator-produced metadata doesn’t work
• No good systems to exploit DC metadata
• DC metadata hidden in objects it describes
Why Dublin Core Doesn’t Work Well

• DC tries to split the difference between full-text searching and standard cataloging
• Creator-produced metadata doesn’t work
• No good systems to exploit DC metadata
• DC metadata hidden in objects it describes
Comparing Metadata Schemes

1. Level of description / specificity
2. Connection to content standards
3. Availability of systems that can handle metadata created by the schema
4. Degree of community specificity
5. Interoperability
6. Proven success / reputation / popularity
Comparing Metadata Schemes (continued)

7. Amount of training needed
8. Viability of the organization
9. Schema’s ability to handle a specific function
10. Adaptability to local needs
11. Scalability
The MARC Model

1. Many elements of description
2. Connected to content standards that provide consistency
3. Numerous vendors that sell systems that accommodate MARC data
4. Is used by diverse communities of practice
5. Is highly interoperable
6. It is popular and its future prognosis is strong.
7. MARC Metadata is generally created by professionals who are experts at metadata creation.
8. Has a dynamic organization behind it.
10. Is adaptable to local needs and is scalable.
MODS

- Metadata Object Description Schema
- Language-based rather than numeric tags
- Expressed in XML
- Bright future, already in use in various applications
<mods xmlns:xlink="http://www.w3.org/1999/xlink" version="3.0"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="http://www.loc.gov/mods/v3"
xsi:schemaLocation="http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/v3/mods-3-0.xsd">
  <titleInfo>
    <title>Sound and fury :</title>
    <subTitle>the making of the punditocracy /</subTitle>
  </titleInfo>
  <name type="personal">
    <namePart>Alterman, Eric</namePart>
  </name>
  <role>
    <roleTerm type="text">creator</roleTerm>
  </role>
  <typeOfResource>text</typeOfResource>
  <genre authority="marc">bibliography</genre>
  <originInfo>
    <place>
      <placeTerm authority="marccountry" type="code">nyu</placeTerm>
    </place>
    <place>
      <placeTerm type="text">Ithaca, N.Y</placeTerm>
    </place>
    <publisher>Cornell University Press</publisher>
    <dateIssued>c1999</dateIssued>
    <dateIssued encoding="marc">1999</dateIssued>
    <issuance>monographic</issuance>
  </originInfo>
</mods>
Conclusion

• Metadata is not dead
• Valuable content merits quality metadata
• Metadata helps turn information into knowledge
References

5. Beall, p. 41.