Metadata to fit your needs...
How much is too much?

Mark Phillips
March 16, 2009
The mission of the University of North Texas Libraries is to **acquire, preserve, provide access to, and disseminate** recorded knowledge in all its forms.

Access will be provided increasingly through **electronic networks** and consortial arrangements.

The Libraries, through traditional methods and through digital information resources, provide bibliographic, reference, and instructional support to assist the university's programs of teaching, research, scholarly and creative production, and public service.
Digital Initiatives at the UNT Libraries

- Government Documents
- Music Library
- Rare Books and Texana Collection
- University Archives
- Media Library
- Portal to Texas History
- Intellectual output of the UNT faculty, staff and students.
Case Studies: A tale of two systems

- UNT Libraries' Digital Collections
  Collections made up of resources from around the library and campus departments

- Portal to Texas History
  Nearly 100 partners from around the state, very small to very large.
What kind of stuff?

- Audio and video of recitals
- Electronic Thesis and Dissertations
- Born Digital Federal Documents
- Digitized material
- Data Sets
- Web Archives
- All sorts of things...
What is metadata?

“Data about Data”

- Descriptive Metadata
- Structural Metadata
- Administrative Metadata
  - Rights Metadata
  - Preservation Metadata
Descriptive Metadata

- Usually domain specific
- Libraries: MARC, MARCXML, MODS, VRA Core, Dublin Core, ONIX …
- Library Catalogs, Journal Databases, Digital Library Systems
- Basis of discovery systems
Structural Metadata

- MPEG-21: Digital Item Declaration Language (DIDL)
- Metadata Encoding and Transmission Standard (METS)
- Usually XML Based
- Typically highly complex
- Interoperability?
Administrative Metadata

- Rights Management:
- Preservation Metadata:
  PREMIS, Metadata for Images in XML (MIX)
Metadata in Libraries

• Central to the libraries way of life.
  – Library Catalogs
  – Electronic Resources
  – Digital Library Collections

• Systems for resource management

• Systems for access to resources
Metadata, How Much?

- Depends...
- Targeted user group?
  - Internal Users?
  - External Users?
  - Subject Specialists?
- Different schemes, different ideas and focus.
  - Using book based metadata for photographs
  - Describing people...
  - Users needs and expectations.
How much...

• Cost benefit
  – Traditional Cataloging
  – Metadata for digital collections

• What are the most important pieces of information for your users? Who? What? Where? When?
Metadata in Search Systems

- Fulltext vs. Metadata based search
  - Use both!
- Users are getting used to the high quality search engines of the world...
- High quality metadata can lead to innovative searching and browsing interfaces on top of your data.
UNT Digital Collections Search:

- Locally qualified Dublin Core Metadata
- Page level OCR
- Extracted “interesting text” from documents
- System considers human created metadata as “top shelf”
- Page level OCR isn't given priority.
- Extracted “interesting text” is somewhere in the middle
Metadata Maintenance

Metadata analysis tools
People make mistakes
Distributed metadata creation
Ideas and rules change

Controlled Vocabularies
Use them
Create them
Share them (in machine readable format!)
End of the line...

Thank You