TEXAS DIGITAL LIBRARY
METADATA WORKING GROUP
(TDLMWG)

Update on the TDL Metadata Working Group’s activities for 2009–2010
What does TDLMWG do?

- Provide Texas Digital Library (TDL) with general metadata expertise. In particular, the Metadata Working Group will address the following TDL needs:
  - Develop best practices.
  - Educate stakeholders on all aspects of metadata.
  - Provide guidance on current status and future trends in metadata development in ways that facilitate compliance at TDL.
Who makes up TDLMWG?

- Daniel Alemneh, University of North Texas
- Jee-Hyun Davis, University of Texas–Austin
- Jason Thomale, Texas Tech University
- Mingyu Chen, University of Houston
- Holly Mercer, Texas A&M University
- Jeanne Hazzard, Texas State University
- Amanda Harlan, Baylor University
Overview

- **Project 1: Create best practices for images, datasets**
  - Images: Update
  - Datasets: Update

- **Project 2: Create metadata courses for TDL training**
  - Survey Results and Follow-Up

- **Project 3: Design a metadata information website**
  - Website: Update
Project 1

Create Best Practices for Images & Datasets
Update: IMAGES

- Draft of recommended image format types
  - **Access**
    - .jpg (JPEG)
    - .jp2 (JPEG2000)
  - **Thumbnails**
    - .gif (GIF)
    - .png (PNG)
  - **Preservation**
    - .jp2 (JPEG2000)
    - .tif, .tiff (TIFF)
  - **Not Recommended**
    - .svg (SVG 1.1 – no Java binding)
      - This format type is too specific and not applicable to all image types.

- Descriptive metadata for images
  - Still reviewing and compiling best practices for images.

- Technical metadata for images
  - Waiting for the TDL Preservation Working Group to make recommendations for us to look at.
Update: DATASETS

- Draft of a generic definition of a dataset
- List of datasets specific to a field of study
- Descriptive metadata for datasets
  - Still gather and compiling best practices for datasets
- Technical metadata for datasets
  - Waiting for the TDL Preservation Working Group to make recommendations for us to look at.
Project 2

Create Metadata Courses for TDL Training
Outline

- Why metadata training needs assessment
- The Survey
- Data collection
- The participants and responses
- Survey results
- Follow-up
- Summary
Why Training Needs Assessment

- At the time of the reactivation of TDLMWG, there were no metadata training courses offered by TDL.
- As the group discussed this we realized there was a wide variety of training needs we could create courses around. To narrow these possibilities down we all agreed that a survey was the quickest way to identify the most pressing training needs in Texas.
- The goal is to take the data gathered from the survey and create training courses that will expand participants’ current knowledge about metadata and digital asset management.
Purpose of the survey was to gather information from TDL member institutions and friends of TDL about their metadata needs.

A web-based survey questionnaire was developed based on the draft questions outlined during the Fall 2009 Working Group meeting:

- Survey questions were created and revised using GoogleDocs.
- Finalized questions and layout of survey were created online using Baylor University’s survey tool, Snap.
The Survey...

- Survey focused on identifying metadata training needs in two main sections:
  - Section 1. Gathering demographic information regarding institutional affiliation and position of the respondents and respondents’ current knowledge of metadata; and
  - Section 2. Discovering respondents’ expectations on what kind of metadata courses should be offered.

- Questionnaire consisted of open-ended questions that allowed participants to express their opinions.
  - Including contact information for follow-up and in-depth feedback.
More than 1300 potential participants were identified via various professional associations and list groups in Texas:

- 145 (11%) Academic Institutions,
- 549 (42%) Public Libraries, and
- 608 (47%) Other (Museums, Centers, Etc.)
Invitations to participate were sent in the first week of February.

- The survey had a two-week turn-around from February 2, 2010 to February 16, 2010.
- Total responses received at the close of the survey were 110.
  - Majority of respondents were from academic institutions (56, 51%) followed by public libraries (43, 39%) and other institutions (11, 10%) in that order.
  - Out of which 73 (66%) respondents agreed to participate in the follow-up email survey.
The overall institutional status regarding metadata-related project activities does not seem encouraging:

- Even though academic institutions are among the early adopters of metadata-related projects, almost two thirds of public libraries in Texas have no plans at all to implement metadata-related projects.
  - Overall, about 41% of Texas’ institutions have no short-term plans to undertake any metadata-related projects.
Texas Academic Institutions: Metadata Implementation Status

- Currently implementing: 57%
- Planning to start: 23%
- No plans at all: 20%

Texas Public Libraries: Metadata Implementation Status

- Currently implementing: 74%
- Planning to start: 14%
- No plans at all: 12%

Texas Museums and other Information Organizations: Metadata Implementation Status

- Currently implementing: 55%
- Planning to start: 27%
- No plans at all: 18%
Responses on respondents’ metadata–related experiences ranged from extremely proficient to no knowledge at all, and some had only heard of metadata but did not understand it.

“Always experimenting with repositories and discovery of contents.”

“Implementing Content DM for with Dublin Core for digitized images, student newspaper, oral histories.”

“Limited implementation in our cataloging. We are a small operation.”

SOME METADATA SCHEMAS I KNOW QUITE A BIT ABOUT, BUT OTHERS ONLY A PERIPHERAL KNOWLEDGE.

“My knowledge of metadata-related projects is very low.”

Member of CDP's Dublin Core Best Practices Working Group, Co–PI for the RMOA EAD database.

“Not sure how our members can/do utilize metadata.”

OUR LIBRARY STAFF AND I WOULD LIKE TO LEARN MORE ABOUT Metadata.

“UNT at Denton is currently digitizing our school annuals and newspapers.”

“Have the knowledge, but no opportunities.”

TCDL–2010
Metadata Implementation Experience
(Texas Academic Institutions, Public Libraries, Museums and other Information Organizations)
(1 = no metadata experience to 10 = tremendous amount of metadata experience)
Respondents ranked several possible metadata course topics from least important (1) to most important (5).

- **Course 1**: General Overview of Metadata
- **Course 2**: Hands on Descriptive Metadata Creation
- **Course 3**: Introduction to Specific Metadata Standards and Applications (e.g., DC, METS, etc.)
- **Course 4**: Intermediate level metadata creation, storage, and management
- **Course 5**: Advanced metadata management, quality assurances, and interoperability
- **Course 6**: Tools for Metadata Management
- **Course 7**: Programming for Metadata
- **Course 8**: Other Training Options
Expectations for Metadata Courses

Some respondents gave additional comments on what they felt should be offered, ranging from advanced programming courses to introduction to metadata courses.

“I think my cataloger colleagues would benefit from introductory/overview courses, and I would like the opportunity to learn/discuss/explore ways to introduce and implement DAM processes among traditional library workflows.”

“I would definitely like to see more programming applications taught like XSLT, XPATH, XML, XFORMS, relational databases, use of WEB 2.0 applications, etc.”

“It would be good to have a non-expert understanding of metadata (purpose and standards) so that I can better perform my job.”

“I would like to have hands-on metadata conversion practices, including a little bit of programming or using xml. I think a course that covers several aspects of xml regarding metadata crosswalk and management would be very helpful.”

“Since I know nothing about this Program an introduction to it is most important.”

“The courses would have to be clearly worth my time and money and specifically applicable to my particular digital project, which I would think would make the courses difficult to plan.”

“Tools for interoperability, discovery, and for semantic additions to traditional metadata, for repositories of research data.”

“What it is and how it’s used; how it can be used in small, rural libraries, suburban libraries and larger libraries.”

TCDL-2010
A follow-up survey was emailed to the 73 participants who agreed to be contacted for follow up.

- Out of that 73, we had only 32 people (about 44%) reply.

We asked five open-ended questions about:

- Type of course (in-person, online, etc.)
- Teaching style (lecture, hands-on, etc.)
- Platforms/Formats (CONTENTdm, Dublin Core, etc.)
- Topic Specificity (Management, schema, etc.)
- Other Issues & Aspects
Follow-Up...

Some of the comments included were

- **Type of Course**
  - “Webinar or in-person all day.”
  - “Anything that would not require more than 1 overnight stay would be best.”
  - “Either online training or onsite training over 2–3 days would be preferable. Half-days in-person would not be cost-effective.”
  - “Having options is good. Considering budgets at this time, I would lean more toward the online course, webinar, or distance learning options.”

- **Teaching Style**
  - “Lecture with hands on.”
  - “Not lecture. Hands on learning working through real life documentation and processes.”
  - “Group training with some hands-on.”
  - “A hybrid would be best.”
Follow-Up...

- **Platforms/Formats**
  - “Metadata in IRs, METS, MODS, XML and EAD.”
  - “Dublin Core, DSpace harvesting.”
  - “Am not familiar with metadata formats, so any would be fine.”
  - “ContentDM Dublin Core metadata, EAD.”

- **Topic Specificity**
  - “Both! We are still new enough to this that tweaks that improve our workflows are most welcome. But we also really need specific training so that more of our staff are up-to-speed.”
  - “Training on particular schemas and tools.”
  - “Management/Workflow issues.”
  - “Particular schema and tools.”

- **Other Issues & Aspects**
  - “The application profile; harvesting and tools; MARC and non-MARC platforms in relation to metadata mapping/crosswalks; name authority in Irs and other non-MARC platforms.”
Based on the final analysis and findings, we will recommend and create appropriate metadata training courses that address the expressed training needs of the majority of Texas institutions.

We would like to use this opportunity to thank you all, who supported the TDLMWG initiatives, in one way another.

Please make sure that you have written your name and email address on the sheet that is going around the room so that we can add you to the TDL’s Texas General Metadata Discussion List which we will be creating after TCDL where you can participate in various metadata related discussions.
Project 3

Design a Metadata Information Website
Why did TDLMWG feel this was an important enough to do?

- The group knew there was a lot of websites already on metadata, but they were all very specific to the topic they were covering. We wanted to compile all these websites and information about them into one site where people could link from.
- In addition to above, TDL doesn’t have a resource page for metadata at the moment for its members.
Areas the Website will Cover

- Metadata Standards (descriptive, technical, digital provenance, rights, preservation, structural)
- Professional Organizations
- Conferences
- Training
- Informed Discussions (Discussion Lists, Blogs)
- Bibliography
- Metadata Projects
- Metadata Guidelines & Procedures
- Metadata Tools
- Metadata Software
- Semantic Web
- Blog
- History/About TDLMWG
Blog vs. Wiki: which is better?
  ◦ At first the group decided on a wiki website for ease of keeping the information up-to-date.
  ◦ Then at the group’s Spring 2010 meeting, we decided on making the website’s framework a blog website because
    • We could easily keep the site up-to-date as with a wiki.
    • The group wanted an interactive presence with TDL’s members and we felt the Blog was the best way to do this by making it the focus on the website.
    • Since the decision was made to go with a blog, the group decided that there had to be a few ground rules the group had to follow to make the blog site a success.
      • Members of group needed to take the blog training TDL offers if they haven’t already done so.
      • Each member needed to participate on the blog once a month.
      • Someone would need to moderate the comments.
Lab instance of website was created in April 2010.
Compiled information was arranged on the website.
Feedback from group was initiated in May 2010.
Finalize design, layout, and information on website.
Goal to make Live: Summer 2010
Link: TDL Metadata Website
What is on the agenda for 2010–2011

**Project 1**
- Finalize and define core descriptive metadata elements.
- Write TDL descriptive guidelines for images & datasets.
- Start work on technical metadata in conjunction with the Preservation Working Group.
- Write TDL technical guidelines for images & datasets.

**Project 2**
- Start creating some metadata training sessions based on the data from the survey.
- Start teaching metadata training sessions either in the Fall 2010 or Spring 2011.

**Project 3**
- Make the website officially live this Summer 2010.
Any Questions?