

Metadata Quality Assurance: The University of North Texas Libraries Experience

Daniel Gelaw Alemneh and Hannah Tarver

University of North Texas Libraries; Digital Projects Unit,

Post Office Box 305190, Denton, Texas 76203, USA

E-mail: daniel.alemneh@unt.edu E-mail: hannah.tarver@unt.edu

Introduction

Libraries exist to connect users with resources and information. In a traditional library, a card or online catalog is merely one aspect of finding and accessing holdings, albeit an essential one. In the same way, for digital libraries, metadata records allow users to gain access to the digital resources. Successful metadata adds value - allowing users to click and delve deeper than they'd be able to with a traditional static representation. Maintaining usable and sustainable digital collections necessitates maintaining high-quality metadata about those digital objects. The two aspects of digital library data quality are the quality of the data in the objects themselves, and the quality of the metadata associated with the objects. Because poor metadata quality can result in ambiguity, poor recall and inconsistent search results, the existence of robust quality assurance mechanisms is a necessary feature of a well-functioning digital library.

Metadata Quality at the UNT

The University of North Texas (UNT) Libraries participate in a number of collaborative digital initiatives. Recognizing the critical role of quality metadata in digital resource life cycle management, the UNT Libraries employ a number of metadata quality assurance procedures and tools at each (pre-ingest and post-ingest) stage.

Pre-Ingest

From the start, we provide intensive training (both face to face and online tutorials) and supplement it with detailed documentation. While providing continuous support, we also instruct metadata creators and/or editors on proper formatting and tools available for them to check their own work. Among other tools and procedures, the following pre-ingest activities facilitate the metadata creation process:

- A metadata creation template (web-based form for creating records) partially automates data population, validates metadata values, and checks formats, links, etc.
- The UNTL controlled vocabularies and dropdown lists draw different terms and concepts into one single preferred word/phrase to ensure consistency.
- Template readers provide firsthand visual checking capability.

Post-Ingest

Our web-based metadata analysis tools allow us to compare field values across a particular collection or our entire holdings to easily identify errors including misspellings, incorrect formatting, empty (null) values, and other likely mistakes. The following tools, among others, allow us to view, analyze, and check for errors in uploaded records:

- List or Browse metadata values:
 - All Values for each element (refined or enhanced by Qualifiers–Use/Ignore, Highlighter-On/Off)
 - Null Values (e.g. for mandatory elements)
 - Authorities Values
- Other visualization and graphical reporting tools:
 - Clickable Maps by Institution and Collection
 - Word Clouds by elements
 - Records added overtime

Summary

Quality services depend on good metadata. Incorrect information: errors, omissions, or ambiguities in the

metadata affect the consistency of search results and can limit the ability of the service provider to include special functions and creative services. In order for end users to benefit fully from the development of digital libraries, responsible and viable service providers need to address metadata quality issues. Based on the UNT Libraries experiences, this presentation will discuss issues related to metadata quality management and demonstrate a number of tools, workflows and quality assurance mechanisms.