Integration of ETD into Topical Digital Library Collections: Facilitating ETD Use and Reuse

Daniel Gelaw Alemneh and Mark Edward Phillips | University of North Texas Libraries, Digital Projects | PO Box 305190, Denton, TX 76203

Introduction

The synergies of numerous emerging trends are shaping creation, access, use, and management of digital information resources. Electronic theses and dissertations (ETD) represent a wealth of scholarly and artistic content created by Master's and Doctoral students in the degree-seeking process.

Good digital collection building involves a number of interrelated activities. Unlike other types of digital resources, ETD are highly specialized collection that demands a more specialized treatment and characterization to better capture the semantic relations of the underlying concepts. Metadata is a systematic representation of an information-bearing object, and the creation of accurate metadata is fundamental to the discovery, use and reuse of digital contents. Accordingly, the University of North Texas (UNT) Libraries actively promotes metadata-based digital resource management, and employs a system of metadata developed in-house (richer than qualified Dublin Core) and add a number of ETD-specific fields that provide optimum searching, discovery, and retrieval, while ensuring long-term preservation and viability of the digital objects. The system also exposes the location of content via OAI-PMH, and (making the UNT ETD open to all) supports the “Open Access” aspirations of UNT. This poster demonstrates the UNT Libraries effort in integrating ETD into topical digital library collection and making existing digital collections more accessible and useful in research processes.

ETD Collection in Digital Library

UNT Libraries offer a single point of access to all UNT theses and dissertations from UNT Digital Library. One of the first five American universities to require ETDs for graduation, UNT began accepting theses and dissertations in electronic format in 1999. Digitizing and providing access to the wealth of information currently held on paper theses and dissertations generates more digital contents that can be and reuse by diverse users for different purposes.

Maintaining usable digital collections necessitates maintaining high quality metadata about those digital objects. Different metadata elements describe different characteristics or aspects of an object. There are many ways to categorize important concepts and possible entries in the metadata record of any digital object. The keywords and subject terms are among the most useful metadata about a digital object that explicitly describes what the item is about.

Given the proliferation of scholarly digital contents, it has become increasingly difficult for researchers to find relevant contents in their own, not to mention related, disciplines. A number of researchers analyzed and described the general behavior of users' informational queries (where the user is seeking information on a particular topic). Zavalina (2010) summarizes collection-level metadata practices and their contribution to subject access and noted that the two major reasons why users experience problems with subject access are the quality and application of subject metadata on one hand, and the complexity of knowledge required for successful subject access on the other. To maintain the consistency of search results and high recall of available resources, it is critical to ensure the quality of the keywords and taxonomies used to describe heterogeneous digital resources within digital libraries. As noted by Peterson (2006) and Spiteri (2007); among others, combining both traditional taxonomies with folksonomies is the solution for delivering a rich user experience on the Web while leveraging the benefits of composite applications, mash ups, and service-oriented architectures.

Summary

While the capacity to create digital content is great and the appetite for it seemingly insatiable, much work remains to make the infrastructure robust. In view of the growing interdisciplinarity and constant changes/development of human knowledge, access to digital resources relies on a seamless discovery process that offers all possible options to users.

Although different metadata fields are often mutually complementing, good subject or key word terms help users find what they need, even when they are not sure themselves what they need. Such (topical and natural) approaches to the subject matter in a digital collection; provide a high-level description with content values that represent a digital collection. Considering the unique value and (mostly) openness of ETDs, it's imperative to start leveraging and treating ETDs as tightly integrated components of Digital Library. Such an aggregation of digital items will add values and of course, enhance relevance, accessibility, and coherence, while ensuring long-term access to heterogeneous digital resources.

References