ETD 2022
Novi Sad
Sep 7 – 9, 2022

25th International Symposium on Electronic Theses and Dissertations
– ETD 2022, Novi Sad, Serbia
September 7 - 9, 2022

Empirical Evaluation of ETD-ms Compliance in Institutional Repositories

There are presently a number of academic institutions that have set up Institutional Repositories (IRs) aimed, in part, to archive, manage and facilitate access to Electronic Theses and Dissertations (ETDs). Increasingly, descriptive metadata associated with the ETDs produced by most of these institutions are now automatically harvested by National ETD portals and entities such as The Networked Digital Library of Theses and Dissertation (NDLTD) Union Catalog. However, there are growing concerns regarding the quality of ETD descriptive metadata originating from most of these institutions, in particular, adherence to the ETD-ms metadata standard, the de-facto standard used to describe ETDs. Compliance with well-established metadata standards such as ETD-ms (Hickey et al., 2021) can arguably improve the quality of metadata.

Suleman highlights that errors in metadata records harvested by the NDLTD Union Catalog are a recurring problem and that while widely advertised, ETD-md compliance is a major challenge (Suleman, 2012).

The quality of metadata is known and cited as being crucial to the effective preservation of digital content and, more significantly, aids in the discoverability of digital content (Park, 2009). There are a number of studies that have been conducted in order to investigate metadata quality. Studies such as the one conducted by Currier et al. have been used to identify errors in learning object repositories, object metadata, by untrained resource creators and the lack of use of authority control and subjects as being the major issues associated with quality of metadata (Currier et al., 2004). In addition, bibliometric analyses of the Dryad repository have identified major problems with Creator, Data and Type metadata elements (Rousidis et al., 2014).

This paper presents an empirical analysis of ETD-ms compliance of the metadata associated with the 5,949,744 ETDs in the NDLTD Union Catalog. ETD metadata records were harvested from the NDLTD Catalog using the OAI-PMH protocol and, subsequently, an analysis was conducted to determine compliance to the ETD-ms metadata standard. In addition, to provide contextual overview of the potential root causes of non-compliance to ETD-ms, a case study was conducted at a Higher Education Institution, in order to understand current practices and procedures employed during ingestion and association of metadata quality. The study could potentially provide direction on issues to be addressed during the ingestion of ETDs into IRs

References

Currier, S., Barton, J., O'Beirne, R., & Ryan, B. (2004). Quality assurance for digital learning object repositories: issues for the metadata creation process. In *Research in Learning Technology* (Vol. 12, Issue 1). https://doi.org/10.3402/rlt.v12i1.11223

Hickey, T., Pavani, A., & Suleman, H. (2021, April). *ETD-MS v1.1: an Interoperability Metadata Standard for Electronic Theses and Dissertations*. https://ndltd.org/wp-content/uploads/2021/04/etd-ms-v1.1.html

Park, J.-R. (2009). Metadata Quality in Digital Repositories: A Survey of the Current State of the Art. In *Cataloging & Classification Quarterly* (Vol. 47, Issues 3-4, pp. 213–228). https://doi.org/10.1080/01639370902737240

Rousidis, D., Garoufallou, E., Balatsoukas, P., & Sicilia, M.-A. (2014). Data Quality Issues and Content Analysis for Research Data Repositories: The Case of Dryad. In *Let's Put Data to Use: Digital Scholarship for the Next Generation* (pp. 49–58). IOS Press.

Suleman, H. (2012). *The NDLTD Union Catalog: Issues at a Global Scale*. http://hdl.handle.net/10757/622568