Marc Richard
McGill University Library, Montreal
Eun G. Park
McGill University, Graduate School of Library & Information Studies, Montreal

An Assessment of e-Theses Metadata in CARL Institutional Repositories

Abstract: The goals of this project are: 1) to analyze metadata element sets used by CARL (Canadian Association of Research Libraries) institutional repositories to describe electronic dissertations and theses; and, 2) based on the analysis, to suggest a prototype for metadata schemes to serve as a standard model for further implementation.

Résumé : Les buts de ce projet sont : 1) analyser les éléments de métadonnées utilisés par les dépôts institutionnels de l'ABRC (Association des bibliothèques de recherche du Canada) pour décrire les mémoires et thèses électroniques ; et 2) d'après cette analyse, suggérer un prototype de schéma de métadonnées pouvant servir de modèle normalisé et être éventuellement mis en œuvre.

1   Project Background

An institutional repository (IR) is a collection of various intellectual works in electronic format that are created by faculty members and other researchers at academic institutions. It provides a central depository space to preserve and provide access to academic publications, including journal articles, preprints, reprints, course-related materials and relevant data. The recently published ARL SPEC Kit on institutional repositories indicates that 43% of the respondents had an operational IR and 31% were planning to have one by 2007 (Association of Research Libraries, 2006).

In Canada, the Canadian Association of Research Libraries (CARL) / Association des bibliothèques de recherche du Canada (ABRC) launched the Institutional Repositories Pilot Project (IRPP) (CARL, 2005; Jordan, 2006; Mark and Shearer, 2006). This project has assisted Canadian academic libraries to build IRs and has hosted harvesting software to provide a central search portal for OAI-PMH compliant metadata. The CARL Metadata Harvester and Search Service provides access to e-content from the following thirteen libraries: University of Calgary, Université Laval, University of Manitoba, McMaster University, Université de Montréal, University of New Brunswick, Université du Québec à Montréal, Queen's University, Simon Fraser University, University of Toronto, University of Victoria, University of Winnipeg, and York University.

A significant proportion of the e-content in Canadian IRs consists of electronic dissertations and theses (EDT). As this reflects the trend that is identified in the ARL SPEC Kit, methods of digitizing theses, providing full-text access to their content, and enhancing access to them through the use of metadata are of interest to academic librarians (Association of Research Libraries, 2006). The reasons of this phenomenon are summarized by Jean-Claude Guédon in the abstract to his presentation on EDT and repositories:

"Theses and dissertation repositories do not always locate themselves easily within institutional repositories…. Furthermore, such documents can be submitted to various forms of evaluation that would help enhance the attractiveness of institutional repositories. Finally, dissertations provide an excellent testbed to design new forms of data and knowledge mining. In other words, dissertations on-line helps us move toward the Web 2.0" (Guédon, 2006).

Metadata harvesting is one of the techniques which enables libraries to organize electronic resources efficiently and effectively, and which provides access points through which users can retrieve these resources in an easy and rapid manner. There is a demand within the academic library community for guidance on best practices and formats for providing EDT metadata which would facilitate such harvesting.

2 Project Goals
The goals of this project are:

a) to analyze the e-thesis metadata element sets currently used by Canadian academic institutional repositories to describe electronic dissertations and theses; and,

b) based on the results of this analysis, to suggest a prototype for metadata schemes that could serve as a standard model for other Canadian academic institutional repositories.

3 Data Collection and Analysis
Through the supportive help of the CARL, sample records of electronic dissertation and thesis resources with metadata descriptions will be collected from participant institutions. Each institution's metadata elements will be examined in order to determine what metadata schemes are being used and to identify the most commonly used and least used data elements. Sample records drawn from the Library and Archives Canada Thesis Portal will be included for purposes of comparative analysis. Based on this analysis, a set of metadata elements will be suggested for other CARL academic libraries. For further development and research, important issues relevant to metadata harvesting will be discussed.

4 References

Association of Research Libraries. *ARL SPEC Kit 292: Institutional Repositories*. July 2006. http://www.arl.org/spec/SPEC292web.pdf (Accessed January 11, 2007).

Canadian Association of Research Libraries. CARL Institutional Repository Project: Online Resource Portal. 2005. http://www.carl-abrc.ca/projects/institutional_repositories/institutional_repositories-e.html (Accessed January 7, 2007).

Guédon, Jean-Claude. "EDT and repositories." ETD 2006: The 9th International Symposium on Electronic Theses and Dissertations, June 7 - 10, 2006, Quebec City, Quebec. http://www6.bibl.ulaval.ca:8080/etd2006/pages/prog.jsf (Accessed January 7, 2007).

Jordan, Mark. "The CARL Metadata Harvester and Search Service." *Library Hi Tech* 24, No. 2 (2006). pp.197-210.

Mark, Timothy and Shearer, Kathleen. "Institutional Repositories: A Review of Content Recruitment Strategies." 72nd IFLA General Conference and Council, August 20-24 2006, Seoul, Korea. http://www.ifla.org/IV/ifla72/papers/155-Mark_Shearer-en.pdf (Accessed January 7, 2007).

5   Relevant Web Sites

CARLCore Metadata Application Profile. http://carl-abrc-oai.lib.sfu.ca/carlcore.

CARL Metadata Harvester and Search Service. http://carl-abrc-oai.lib.sfu.ca.

Library and Archives Canada. Theses Canada Portal. http://www.collectionscanada.ca/thesescanada/index-e.html