Emerging trends in metadata research

Heather Lea Moulaison
School of Information Science & Learning Technologies
University of Missouri
303 Townsend Hall
Columbia, MO 65211
+1 573.882-8323
moulaisonhe@missouri.edu

Susan Rathbun-Grubb
School of Library and Information Science
University of South Carolina
1501 Greene Street
Columbia, SC 29208
+1 803.777.0485
srathbun@mailbox.sc.edu

June Abbas
School of Library and Information Studies
The University of Oklahoma

Jane Greenberg
SILS Metadata Research Center
School of Information and Library Science
University of North Carolina at Chapel Hill

Kathryn La Barre
The Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign

Eva Méndez Rodríguez
Associate Professor
Dpto. Biblioteconomía y Documentación
Universidad Carlos III de Madrid

Erik Mitchell
College of Information Studies
University of Maryland, College Park

Alenka Šauperl
Department of Library and Information Science and Book Studies
University of Ljubljana

ABSTRACT
Metadata can be used in a variety of environments, from library catalogs and research databases to data archives, proprietary knowledge bases, and the open web. Multiple sources for metadata creation exist, with metadata being generated by professionals, through automated techniques, and by the public at large. Information professionals are eager to share and leverage data sources to create new knowledge and feed the demand for new interoperable applications. There is great potential for improving discovery by weaving together traditional library metadata with social cataloging metadata to create a new scheme for representing information in library catalogs, institutional repositories, or digital libraries.

KEYWORDS
Metadata, research, trends

INTRODUCTION
Metadata can be used in a variety of environments, from library catalogs and research databases to data archives, proprietary knowledge bases, and the open web. Multiple sources for metadata creation exist, with metadata being generated by professionals, through automated techniques, and by the public at large. Information professionals are eager to share and leverage data sources to create new knowledge and feed the demand for new interoperable applications. There is great potential for improving discovery by weaving together traditional library metadata with social cataloging metadata to create a new scheme for representing information in library catalogs, institutional repositories, or digital libraries.

1 Conveners
2 Panelists, for affiliations, please see the section at the end of this document.
Metadata research embraces certain tensions and creates bridges. Novel uses for old and new data require both theoretical and practical research methodologies. Metadata projects that cater to diverse user requirements necessarily balance the need for encoding standards and controlled vocabularies with the utility of user-defined tags and automated procedures. Emerging streams in metadata research illustrate a variety of “bridge-building” approaches that lead to the rich possibilities that metadata and its study currently represent.

Each of the researchers on this panel represents a unique perspective on the ways in which metadata research is unfolding and will continue to progress over the next few years. Drs. Erik Mitchell, Jane Greenberg, Eva Méndez Rodríguez, June Abbas, Alenka Šauperl, and Kathryn La Barre will discuss their current research streams, speaking to the above issues via an exemplary project they have chosen to describe. Interactivity between the panelists and among the audience members will be encouraged after the panelists have presented their projects for discussion.

**PANEL PRESENTATIONS**

Dr. Mitchell will begin the conversation on automatic metadata generation by discussing the new challenges metadata creation and use are facing in the world of Linked Open Data (LOD), particularly the need for increased attention to complex data models, adherence to metadata syntactic and semantic representation, and a stricter definition of metadata boundaries. His project explores the reconciliation of metadata and evaluation of metadata quality using human and computational methods. In developing computational techniques for evaluating syntactic and encoding metadata quality this research utilized data feeds from MARC and OAI data sources which were then processed using clustering and reconciliation techniques in the Google Refine platform.

Drs. Greenberg and Méndez Rodríguez will continue the discussion of linked data by describing the ways that the growth of linked-data vocabularies has been motivated by new technology and the promise of new means for sharing knowledge. Their discussion emerges from their work on HIVE (Helping Interdisciplinary Vocabulary Engineering) project R&D activities, the extension of HIVE-ES (España), and an exploratory study examining system effectiveness and usability. (“HIVE is an automatic metadata generation approach that dynamically integrates discipline-specific controlled vocabularies encoded with the Simple Knowledge Organisation System (SKOS), a World Wide Web Consortium (W3C) standard. HIVE assists content creators and information professionals with subject cataloging and provides a solution to the traditional controlled vocabulary problems of cost, interoperability, and usability” (http://ils.unc.edu/mrc/hive/)).

Dr. Abbas bridges traditional and social sources of metadata in her research, and she will talk about her project that evaluates the metadata schemes and conceptual models present in traditional, discipline-specific (MARC, MODS, FRBR, RDA, etc.), socially-constructed (LibraryThing, Shelfari, Goodreads, Delicious, Flickr, etc.), and personal information management contexts. This work will examine what the scheme might include beginning with the new RDA standard, and build the scheme out from the elements of the standard.

Dr. Šauperl builds on the conversation about descriptive metadata by focusing on the user perspectives of librarians, literary theorists and readers who may share an interest in the same work of fiction and use the same information systems (e.g., library catalogs, reader advisory services). She will talk about her meta-analysis of published reports and her original research that analyzed the typical metadata elements used by each of the user groups. A content analysis of different descriptions of novels (subject headings, class numbers, annotations, and reviews) in different environments (library catalogs, social networks, online bookstores, publishers' catalogs and literary journals) indicate that the story, information about the author, genre, personal experience with reading the novel, and evaluation are most frequently used by the four groups; adding these descriptive elements to other metadata may enhance user satisfaction.

Dr. La Barre’s work focuses on new descriptive cataloging standards, RDA. Since 2005, Library and Information Science researchers have continually found themselves grappling with a dynamic information environment both in terms of the increasing prevalence of full text resources, and in terms of changes in the codes and standards by which bibliographic access is created. User expectations have shifted accordingly. At the heart of many of these changes lies the Functional Requirements for Bibliographic Records (FRBR) final report (IFLA, 1998), and Resource Description and Access (RDA), the new unified cataloging standard based on the FRBR model – soon to be implemented on March 31, 2013. This talk will discuss the task analysis findings from the research project Folktales and Facets that provides a glimpse into a real world user survey of information seeking tasks among a small set of subjects who rely on folk narratives in their professional lives as storytellers (n=6) scholars or instructors (n=4), or librarians (n=2). This briefing will compare and contrast observed user tasks with those formulated by the IFLA study group, which relied on the theoretical and practical backgrounds and experiences of the study group members.

**AUDIENCE INTERACTION**

Each talk will be given in 12-15 minutes, with ample time reserved at the end for questions from the audience and other panel members. To encourage interaction, attendees will all be given small blocks of post-its and will be able to tag talks with brief questions and comments to be addressed by panelists during the question period. To encourage participation, conveners/moderators will also lead a guided discussion.
BIographyal information and contributions, panelists

Dr. June Abbas, Creating metadata schema from the user-down or from the standard-up
Dr. June Abbas, PhD, is an Associate Professor in the School of Library and Information Studies (SLIS) at the University of Oklahoma, Norman campus. She obtained her PhD in Information Science from the University of North Texas in 2001 and taught in the Department of Library and Information Studies at the State University of New York in Buffalo from 2001-2008 before joining the SLIS in August 2008. Her research focuses on the development of user-centered digital libraries, institutional repositories, and other knowledge organization structures. She also conducts research on children and teens and their use of technology, and the intersection between information behavior, information retrieval, and structures for organizing knowledge. Currently she is Chair of the ALA RDA Conference Forums and Programs Taskforce.

Dr. Jane Greenberg, To HIVE: Linked data, a global step
Dr. Jane Greenberg is a professor at the School of Information and Library Science, UNC-CH, and director of the SILS Metadata Research Center.

Dr. Kathryn La Barre, RDA and user tasks: A brief report from the Facets and Folktales project
Dr. Kathryn La Barre is an Assistant Professor at the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. She is the Chair of the Special Interest Group for History and Foundations of Information Science (SIG/HFIS) and President of the Canada/United States regional chapter of the International Society for Knowledge Organization. She conducts research in the area of knowledge organization and access systems (historical and contemporary), and has published on these subjects in the Journal of the American Society for Information Science and Technology, and in the Annual Review of Information Science and Technology.

Dr. Eva Méndez Rodríguez, To HIVE: Linked data, a global step
Dr. Eva Méndez Rodríguez is an associate professor at Dpto. Biblioteconomía y Documentación, and Associate Vice-rector of Infrastructures and Environment, Universidad Carlos III de Madrid, Spain.

Dr. Erik Mitchell, Measuring metadata quality via human and computational techniques
Erik Mitchell, PhD is an Assistant Professor at the College of Information Studies at the University of Maryland, College Park. In addition to studying information technology adoption and use in libraries, Erik examines metadata issues and professional development in library and information science. Before joining the University of Maryland Erik served as the Assistant Director for Technology Services in the Z. Smith Reynolds Library at Wake Forest University where he worked for twelve years.

Erik is a columnist for the Journal of Web Librarianship and has published and presented on library IT, metadata use and pedagogical approaches. He holds a PhD in Information and Library Science from the University of North Carolina at Chapel Hill, a Masters in Library Science from the University of South Carolina and a Bachelor of Arts in Literature from Lenoir-Rhyne College.

Dr. Alenka Šauperl, Descriptive metadata related to fiction: Similarities of different user groups
Dr. Alenka Šauperl received her PhD from the UNC School of Information and Library Science in 1999. Her doctoral research focused on subject description in library catalogs. She received the ACRL/ISI doctoral dissertation award in 1999. A part of the dissertation was published with the title Subject determination during the cataloging process by Scarecrow in 2002. She has published several papers in the international ILS journals and participated at international conferences. She teaches organization of information and indexing and abstracting at the Department of Library and Information Science and Book Studies at the Faculty of Arts, University of Ljubljana. Her research continues to focus on subject representation of information objects in different environments, including libraries, archives and museums.

Biographical Information, Conveners
Heather Lea Moulaison, PhD, is an Assistant Professor at the University of Missouri’s School of Information Science and Learning Technologies. Dr. Moulaison’s teaching and research interests include organization of information (including cataloging, metadata) and emerging technologies in libraries. Dr. Moulaison earned her PhD at Rutgers’ School of Communication and Information, an MSLIS at the University of Illinois, Urbana-Champaign GSLIS, and an MA in French from UIUC. Before coming to the University of Missouri, Dr. Moulaison taught in Canada and in Morocco.

Susan Rathbun-Grubb, PhD, is an Assistant Professor at the University of South Carolina’s School of Library and Information Science where she teaches courses in library technical services, subject analysis and classification, indexing and abstracting, and social science information services. Her research interests include career and workforce issues in LIS, organization of information, and data curation. Dr. Rathbun-Grubb earned her PhD and MSLS from the University of North Carolina’s School of Information and Library Science, and a MAT and AB from Duke University.
PANELISTS AND AFFILIATIONS

June Abbas, Ph.D.
Associate Professor
School of Library and Information Studies
College of Arts and Sciences
The University of Oklahoma
401 W. Brooks, Bizzell Library
Norman, OK 73069
+1 405-325-3921
jmabbas@ou.edu

Dr. Jane Greenberg
Professor and
Director, SILS Metadata Research Center
School of Information and Library Science
University of North Carolina at Chapel Hill
205 Manning Hall, CB # 3360
Chapel Hill, NC 27599-3360
Phone: (919) 962-8066
janeg@email.unc.edu

Dr. Kathryn La Barre
The Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign
501 E. Daniel Street, MC-493,
Champaign, IL 61820-6211 USA
+1 217) 244-4449
klabarr@illinois.edu

Dr. Eva Méndez Rodríguez
Associate Professor
Dpto. Biblioteconomía y Documentación
Associate Vice-rector
Infrastructures and Environment
Universidad Carlos III de Madrid
Madrid, Spain
(+34) 91 624 86 20
emendez@bib.uc3m.es

Dr. Erik Mitchell
Assistant Professor
College of Information Studies
University of Maryland, College Park
4121G Hornbake Building, South Wing
+1 (301) 405-7622
erik@umd.edu

Dr. Alenka Šauperl
Department of Library and Information Science and Book Studies
Faculty of Arts
University of Ljubljana
Aškerčeva 2
SI - 1000 Ljubljana
SLOVENIA
sauperl@ff.uni-lj.si