An Annotated Bibliography of Scholarly Articles on Best Practices for OER Metadata

Compiled by Bobby Bothmann
Catalog and Metadata Librarian
Minnesota State University, Mankato
May 2020

Abeywardena, Ishan Sudeera, Chee Seng Chan, and Choy Yoong Tham. "OERScout Technology Framework: A Novel Approach to Open Educational Resources Search." *The International Review of Research in Open and Distributed Learning* 14, no. 4 (September 30, 2013). DOI: 10.19173/irrodl.v14i4.1505

The authors offer a concept of a search tool based on text mining. They note existing search engines most commonly used to search for OER, such as Google and Bing, are not effective at OER discovery. They cite the lack of a single metadata standard as one of the chief factors that inhibit effective OER discovery. The authors propose "OERScout," a technology framework that uses text mining to cluster text-based OER. Their proof-of-concept testing demonstrates that the method works. The examples used in the paper are limited to broad disciplines, like chemistry, so it is not clear how this system would work for more specialized disciplines. This method also would not work well for non-text-based OER unless OER media have good ingestible text.

Atenas, Javiera, and Leo Havemann. "Questions of Quality in Repositories of Open Educational Resources: A Literature Review." *Research in Learning Technology* 22 (2014). DOI: 10.3402/rlt.v22.20889

The researchers, from the University of London's Department of Science, Technology, Engineering, and Public Policy and IT Services respectively, conducted a meta-analysis of 122 peer-reviewed journal articles to identify the lack of best practices in OER repositories in order to devise a method of evaluating these repositories. Their analysis produced four themes of purpose: search, share, reuse, and collaborate; and 10 quality indicators: featured resources, user evaluation, peer review, authorship attribution, keywords, standardized metadata (chiefly LOM or Dublin Core), multilingual interfaces, social media tools, license specification (particularly for Creative Commons), and the availability of the source code or original files. This article does not directly relate to best practices for metadata, but does provide some insight into how a repository could be evaluated.

Atwell, Claire. “5 Tips for Writing Effective Meta Descriptions (and Why They Matter),” 2013.

Created by a successful web design, development, and marketing agency for mass consumption, this short piece defines simply what makes for good metadata. The page gives five tips: 1) don't overlook them; 2) make them compelling, concise, and informative; 3) Keep them between 150 and 160 characters; 4) Utilize important key words, and 5) Create unique descriptions. In essence, this page summarizes many of the principles of subject analysis and resource description.

Bied Sperling, Barbra. “LRMI Implementation Case Study: MERLOT.” *Open World* (blog), August 15, 2014. https://lornamcampbell.wordpress.com/2014/08/15/lrmi-implementation-case-study-merlot/.

Team leader for technical development, the author briefly describes the success of implementing LRMI in MERLOT and provides a link to an LRMI to MERLOT map. The MERLOT metadata scheme is proprietary. LRMI (Learning Resources Metadata Initiative) is an open scheme developed by Creative Commons and the Association of Educational Publishers. The map is useful for the mapping comparison of OER-related metadata.

Brooks, Christopher, and Cord McCalla. “Towards Flexible Learning Object Metadata.” *International Journal of Continuing Engineering Education and Lifelong Learning* 16, no. 1 (2006): 50–63. DOI: 10.1504/IJCEELL.2006.008917

The computer scientist authors critique the 76-element Learning Object Metadata (LOM) standard as ineffective and propose a flexible method of metadata for learning objects. They note underlying assumptions with metadata assume human creation and consumption, and that the creation of a schema using stakeholders results in either a too-narrow or too-broad focus. They propose that usage information from users, once associated with the objects they use, will provide better discoverability. This approach requires a Semantic Web infrastructure that is not currently present.

Canham, Steve, and Christian Ohmann. “A Metadata Schema for Data Objects in Clinical Research.” *Trials* 17, no. 1 (November 24, 2016): 557. DOI: 10.1186/s13063-016-1686-5

The authors present a proposal for a metadata schema specific to objects in clinical research. Each attribute for the proposal is listed and explained. Much of it is based on identifiers that are important for research grants and digital objects. The schema is based on DataCite with proposed extensions for the source study and rights. The schema provides a basic metadata structure.

Currier, Sarah, Jane Barton, Ronan O’Beirne, and Ben Ryan. “Quality Assurance for Digital Learning Object Repositories: Issues for the Metadata Creation Process.” *Research in Learning Technology* 12, no. 1 (2004). DOI: 10.3402/rlt.v12i1.11223

These British scientists examine metadata quality using three example repositories from the UK. The authors limited their examination to metadata needed for discovery. They point out that

discovery metadata can also be used for resource selection, which is usually inherent in conceptual metadata such as reviews or user comments. A limitation of their study is that none of their evidence addresses how users discover digital objects. Across the three example repositories, they authors found that consistency and spelling were the major contributors to poor metadata quality. metadata creators applied classification inconsistently, had difficulty with the application of the Rights and Relation elements, and mixed up the concepts of form with content. They suggest metadata creators make use of spell check, employ authority control, assign classification and format consistently employing metadata creators with that expertise, and providing accessibility properties. They conclude with the suggestion that metadata creation models be limited to the resource creator, metadata specialists, or a collaboration of the two. This article contributes considerations for the quality control of a best practice.

Dichev, Christo, and Darina Dicheva. "Open Educational Resources in Computer Science Teaching." In *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education*, 619–24. Raleigh, North Carolina: ACM, 2012. DOI: 10.1145/2157136.2157314

Focused on the needs of computer science educators, Dichev and Dicheva from the Computer Science Department, Winston Salem State University, report on a survey to understand computer science (CS) instructors' needs related to OER. Referencing an earlier study that examined how easy it was to find CS OER, the aim of this study was to identify why OER adoption is so low. In their discussion the authors criticize existing metadata standards as not supporting the expectations of instructors. Merlot is named as having the richest metadata schema, with MIT Open Courseware a close second. Everything else is dismissed as limited.

Dimitrova, Aneliya. "Learning Object Metadata Workflows for Description, Findability and Reusability Improvement." MSc Thesis. Delft University of Technology, 2018.

Dimitrova's master's thesis in computer science examines metadata generation for OER in MOOC platforms and seeks to solve issues with inconsistencies and gaps in metadata. The author looks at industry practices and proposes a metadata taxonomy. She critiques LOM's lack of metadata to describe recommendations and learning and teaching styles. She makes an interesting recognition that some metadata are intrinsic (directly extracted), and others semantic, requiring human intervention to select meaningful data or extracted from encoded content (e.g., abstract, keywords). Dimitrova proposes several design patterns that group common processes into easy steps to improve metadata interoperability.

Glogoff, Stuart J., and Garry J. Forger. "Metadata Protocols and Standards." *Internet Reference Services Quarterly* 5, no. 4 (December 1, 2000): 5–14. DOI: 10.1300/J136v05n04_03

Both librarians from the University of Arizona, Glogoff and Forger cite their work to identify metadata needed for a database in the College of Agriculture as the premise for this article. Their apparent intent is to make the case that metadata quality for digital objects was poorly planned (probably because so many libraries that began digital repositories in the 1990s eschewed the work of trained catalogers as too intractable to work with alternate mark-up languages). They introduce the Dublin Core schema and briefly mention MERLOT. Judging by the intended

audience of the journal, which are reference librarians, many of whom have some oversight of digital repositories, and with the hindsight of two decades, the article feels like a subtle rejoinder for not including catalogers as metadata specialists. This article is helpful for the historical setting of the beginnings of digital object repositories.

Gómez-Zermeño, Marcela Georgina, and Lorena Yadira Alemán de la Garza. "Temoa: An Open Educational Resources Portal to Seek, Investigate and Inquire." *Open Praxis* 7, no. 3 (July 1, 2015): 211–26. DOI: 10.5944/openpraxis.7.3.211

The authors, both doctors in education at Technológico de Monterrey (Mexico), describe the development of Temoa, a multilingual catalog of collections of OER. They provide an explanation of the concepts of metadata and open educational resources (OER), and explain the ingestion of OER objects into the Temoa repository. The article provides a list of the core metadata attributes used for description and the summary level of the classification system. As a working repository, it provides examples of metadata categories that can inform best practices.

Markantonatou, Stella, Panagiotis Minos, Katerina Tzortzi, and George Pavlidis. "Metadata for the Learning Objects That Contain Cultural Objects." *Mediterranean Archaeology and Archaeometry* 16, no. 5 (2016): 53–61. DOI: 10.5281/zenodo.204966

The authors, all from the Athena Research & Innovation Center in Information Communication & Knowledge Technologies, present their work on an ontology that unifies LOM metadata with CIDOC-CRM. The IEEE/LOM (Learning Object Metadata) is designed to describe the resource for discoverability, whereas CIDOC-CRM (Conceptual Reference Model) describes the information and concepts in a cultural heritage object. The marriage of these two schema would allow for discovery based on a question such as "Alpha is a LO (learning object) made by X and contains a digital image of Joconda by DaVinci." The article quickly becomes quite technical, but the summary is that the authors were able to successfully unify the two data models and build a proof-of-concept database that demonstrates the viability of a complex semantic query.

McClelland, M. "Metadata Standards for Educational Resources." *Computer* 36, no. 11 (November 2003): 107–9.

McClelland, a professor of computer information systems at North Carolina Central University, provides an overview of standards for OER in this trade journal article. She describes Dublin Core as a schema designed for use by untrained metadata creators. She also describes the IEEE LOM, which employs 60 metadata elements in a structuralist approach. Both schema make use of controlled vocabularies for such elements as classification and subject content. The article is meant to be an awareness piece, and so could be useful as an introductory reading for those new to metadata.

Open Knowledge Foundation. "Open Metadata Handbook." OER Commons, 2015.

The Open Metadata Handbook is an OER created by two working groups from the Open Knowledge Foundation, a non-profit organization promoting open access. The handbook gives a

passing attempt at defining the concept of metadata, and an overview of the concept of "open" with respect to metadata. The technical overview section glosses over the common metadata elements (e.g., title, language, date) and gives a somewhat misinformed explanation of common metadata standards. Overall not a reliable resource for an explanation, but it does provide a somewhat comprehensive list of example schema.

Pons, Daniel, José Ramón Hilera, Luis Fernández, and Carmen Pagés. "A Learning Quality Metadata Approach: Automatic Quality Assessment of Virtual Training from Metadata." *Computer Standards & Interfaces* 45 (March 1, 2016): 45–61. DOI: doi.org/10.1016/j.csi.2015.12.001

The authors, form the University of Alaciá (Spain) Department of Computer Science present LQM (Learning Quality Metadata), a schema for virtual education based on IEEE LOM (Learning Object Metadata). A variety of new data types are listed, some of which define authenticity, some address accessibility, and several refer to technology needs. The authors also discuss the meaning of quality and propose a data structure for information storage. The writing is poor and often rambles, making it difficult to pull out useful information.

Roy, Devshri, Sudeshan Sarkar, and Sujoy Ghose. "A Comparative Study of Learning Object Metadata, Learning Material Repositories, Metadata Annotation & an Automatic Metadata Annotation Tool." *Advances in Semantic Computing* 2 (2010): 103–26

The authors list some open metadata standards, including Dublin Core, IEE Learning Object Metadata, IMS Global Learning Consortium, the Advance Distributed Learning Initiative, and the CanCore Learning Resource Metadata Initiative. A full list of data elements are provided in Table 1 on page 107. Further discussion of learning object metadata-based repositories provides more insights into potential metadata elements including didactical context, course level, difficulty level, interactivity level, semantic density, pedagogical duration, resource type, grade, and discipline, which can draw out unique qualities of learning objects. The rest of the paper discusses the feasibility of automatic annotation.

Saundry, A. "Institutional Repository Digital Object Metadata Enhancement and Re-Architecting." In *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, 1–3, 2017. DOI: 10.1109/JCDL.2017.7991603

This conference abstract by a librarian from the University of British Colombia (UBC) presents work on UBC's "Open Collections" (OC) institutional repository portal that interfaces with four different platforms. They have enhanced their metadata, preferring a scholarly information architecture over an administrative one. They state that the enhancements, which are not named or described in this abstract, have improved object findability. The lack of any specific information in this two-page abstracts limits its usefulness, but does provide a lead for future investigation.

Sobotka, Clare, Holly Wheeler, and Heather White. “Leveraging Cataloging and Collection Development Expertise to Improve OER Discovery.” *Oregon Library Association Quarterly* 25, no. 1 (August 14, 2019): 17–24. DOI: 10.7710/1093-7374.1971

All catalogers at academic libraries in Oregon, the authors provide a short newsletter item that addresses bibliographic control issues they have grappled with for OER they have added to their collections or that has been produced by their faculty. They indicate that best practices for metadata for OER description were scarce. They identify a number of specific MARC Bibliographic tags for use with descriptive cataloging of OER to provide better precision to OER search results in their catalogs.

Steinacker, Achim, Amir Ghavam, and Ralf Steinmetz. “Metadata Standards for Web-Based Resources.” *IEEE MultiMedia* 8, no. 1 (2001): 70–76.

The authors, all information specialists from German and Canadian universities, describe what metadata are and how they are used in this trade journal article. They further describe RDF (Resource Description Framework), the data model that allows metadata to be interoperable across various applications and serves as the engine for the semantic web. They are setting up the idea that RDF can be used with LOM and Dublin Core to create complex statements about learning objects. They conclude by pointing out that most metadata schema attempt to be both specialized and generic, which limits their application. Similar to McClelland (2003), this article serves as a good introductory explanation of the concepts.

Steiner, Elisabeth, and Carina Koch. “A Digital Archive of Cultural Heritage Objects: Standardized Metadata and Annotation Categories.” *New Review of Information Networking* 20, no. 1–2 (July 3, 2015): 255–60. DOI: 10.1080/13614576.2015.1112171

Steiner and Koch from the Centre for Information Modeling at the University of Graz (Austria) define indispensable, mandatory metadata--all basic information--for their digital archive of cultural heritage objects. The authors also identify challenging categories, like Material, Measurements/Format, Keywords, and Transcription and propose methods for them: controlled vocabulary; units with uniform separator; controlled vocabulary or lists; and subject vs. genre, respectively. Standards the institutions decided upon are the German national authority file, Geonames, and Getty's AAT (art & architecture thesaurus). Their explanations of how they thought through problems and resolved them provide insights for development of best practices.

Vidal-Castro, Christian, Alejandra Andrea Segura Navarrete, Victor Menendez-Dominguez, and Claudia Martinez-Araneda. “Towards a Holistic Model for Quality of Learning Object Repositories: A Practical Application to the Indicator of Metadata Compliance.” *The Electronic Library* 35, no. 5 (January 1, 2017): 953–76. DOI: 10.1108/EL-10-2015-0202

The authors, information specialists and a mathematician from universities in Chile and Mexico, address metadata quality needs for learning objects. Much of their literature review revolves around various uses and interpretations of quality, however, the authors do not provide any explicit definition of "quality" for their research. They do provide a complex table listing a host

of quality characteristics culled from three cited papers and two unnamed repositories. However, the utility of the table is not apparent. Much of the paper addresses technical and mathematical aspects of their research, demonstrating algorithms for identifying and ranking quality indicators. The mathematical complexity of this article and its failure to clearly define concepts of quality place it outside the realm of usability in the context of a best practice for OER metadata.

Addendum: Selected Examples of OER-related Metadata Schema

Creative Commons. “LRMI/Properties/1.1 - Creative Commons,” 2013. https://wiki.creativecommons.org/wiki/LRMI/Properties/1.1.

-----. “Marking Works Technical - Creative Commons,” 2006. https://wiki.creativecommons.org/wiki/Marking_Works_Technical.

Digital Public Library of America. “Metadata Application Profile.” Digital Public Library of America, 2012. https://pro.dp.la/hubs/metadata-application-profile.

-----. “DPLA Metadata Quality Guidelines.” Digital Public Library of America, 2016. https://docs.google.com/document/d/1dITqEYEWsMX1a2pLPmkL78k1LN2b4im03spn8_QFscY/edit?usp=sharing.

IEEE Computer Society. Learning Technology Standards Committee. “IEEE Standard for Learning Object Metadata.” *IEEE Std 1484.12.1-2002*, September 2002, 1–40. https://doi.org/10.1109/IEEESTD.2002.94128.

IMS Global Learning Consortium. “IMS Learning Resource Meta-Data Best Practice and Implementation Guide.” IMS Global Learning Consortium, 2001. https://www.imsglobal.org/metadata/imsmdv1p2p1/imsmd_bestv1p2p1.html.

OER Commons. *OER Commons Learning Object Metadata Element Set Based on IEEE Learning Object Metadata Information Model OERCommons.Org / ISKME.Org*. V. 4. OER Commons, 2015. http://www.mccvlc.org/~staff/uploads/ckeditor/files/OERCv4-MetadataApplicationProfile.pdf.

SPARC OER Discovery Workgroup. “Cataloging and Metadata Templates - Google Drive,” 2019- https://drive.google.com/drive/folders/1rpiLVb6awmPWGXRfMFcR7igpSvM8SVLK.

-----. “OER Discovery: Best Practices for Academic Libraries,” 2019- https://docs.google.com/document/d/1YbQiBnuoPxaJXobckrLZ5hgupfnjbK2D1PeyyWaOnvk/edit?usp=embed_facebook.

United States National Archives. “Lifecycle Data Requirements Guide.” National Archives, 2016- https://www.archives.gov/research/catalog/lcdrg.