

Digital Humanities and Metadata: Linking the Past to the Digital Future

Marina Morgan
Ryerson University,
Canada
marina.morgan@ryerson.ca

MJ Suhonos
Ryerson University,
Canada
mjsuhonos@ryerson.ca

Fangmin Wang
Ryerson University,
Canada
fwang@ryerson.ca

Keywords: metadata schema; metadata mapping; digital collections; digital humanities; digitization.

1. Introduction

The purpose of this poster is to highlight cross-domain metadata uses, metadata mapping, and success measures at the Ryerson University Library and Archives. The Library is highly involved in Ryerson-based proposals for interdisciplinary projects, especially in the Digital Humanities. Designing an online environment for the preservation and analysis of illustrated texts for children and Canadiana is a collaborative effort that involves cataloguing, metadata mapping, digitization, and website design.

2. Background

In 2012 the Centre for Digital Humanities partnered with Ryerson University Library and Archives (RULA) for a project involved in making available to Ryerson and other online research communities some of the more unique holdings of the Children's Literature Archive (CLA) collection. The main goals of the project were the preservation and digitization, as well as making available online over 2000 illustrated texts for children written primarily between the late nineteenth century and the mid-twentieth century. Ryerson library staff, students, and professors collaborated on creating a searchable catalogue of the specific archival materials in the collection. At the same time, students connect their research and writing with a wider community by building scholarly, on-line exhibitions of CLA books. The team has defined the development of innovative pedagogical tools as another key element of this project.

3. Implementation and Challenges

As a result of analyzing various web-publishing platforms, the research team chose Omeka for this project. Omeka incorporates features of Web Content Management, Collections Management, and Archival Digital Collections Systems. Some of the features that met the expectations of the research team were the simplicity of installation, flexible approach to metadata, import and export functionality of specific standard data formats, and ease of various plug-in deployment (Kucsma, Reiss, & Sidman, 2010). The metadata standard used was the Dublin Core Metadata Element Set (DCMES), as it initially offered users the greatest flexibility. While more metadata flexibility was highly desired, the Ryerson team unanimously decided to opt out of specific plugins that would have added the full set of DCTERMS properties, including element refinements and supplemental elements. Once the data was mapped, the researchers decided that better specificity for particular elements that would reduce loss of semantics was highly desired: refining the DC date field using the DCTERMS dateCopyrighted element, refining the DC title field using the DCTERMS alternative element, and refining the DC description field using the DCTERMS abstract element. This change required additional work to migrate existing data to the new fields using the DCTERMS schema.

Additional challenges involved determining which elements to display and how many were required to make the best use of the metadata within the website, whether a brief or full display of

the records would improve searchability and support Digital Humanities research. Rose (2012) asserts that interfaces are trending towards a brief record display, and that there is no consensus on which fields would enable or prevent certain user tasks. However, the research collaborators decided on a full record display, allowing users to distinguish between similar records.

Other challenges were ensuring the consistency of data, especially integrating semantic mappings between existing field names defined by cataloguers in an Excel spreadsheet and Dublin Core metadata elements across the collection. Initial analysis of the catalogued collection revealed that only ten of the fifteen core elements would be used, and some defined elements were not compatible with the DCMES. Extensive corrections of the initial data were done to ensure compatibility with the Omeka software and the DCMES. There was discussion among the research group around the treatment and use of the DC creator and contributor elements to improve specificity. In particular, researchers expressed the desire to indicate specific roles, such as illustrator, editor, etc. as part of a refinement scheme. This level of detail is important for Digital Humanities research, but practices around use of vocabularies such as MARC Relator terms within the DCTERMS schema are not well documented.

In summary, the team was initially constrained in the use of the DC elements that would pose problems in terms of mapping, semantic ambiguity or semantic overlaps (Han, Cho, Cole, & Jackson, 2009). The decision-making was always by consensus and always with the user in mind.

4. Research Significance

First, this project demonstrates the positive results in Digital Humanities research when librarians, faculty, and students work together on a project that involves collaboration on data migration, DC element identification, cataloguing troubleshooting and correction, and website design and functionality refinement. Second, the project tested and demonstrated limitations and capacity of the Omeka system for online content management using Dublin Core metadata to build a searchable database of archival materials, and a user interface that is accessible to faculty, students, and the general public. Third, the potential of the Omeka Exhibit Builder as a publishing platform has been prototyped using the work provided by undergraduate students. In the future, undergraduate students and researchers, as well as faculty, will use the Exhibit Builder to showcase the CLA holdings and augment the existing item descriptions.

The availability and accessibility of a wide range of educational resources is a high priority for universities, libraries, and other institutions (Koutsomitropoulos, Alexopoulos, Solomou, & Papatheodorou, 2010). As Ma (2009) asserted, educating a wider audience on metadata, collaborating with parties who share a common goal, and researching new projects in the field is beneficial for both parties involved. A strong collaborative partnership between the Ryerson University Library and Centre for Digital Humanities was developed. Digital Humanities research works best with an interdisciplinary team committed to an iterative process of prototyping and refinement. Digital Humanities work also empowers undergraduate students as archival researchers and knowledge producers in the public sphere.

4. Conclusions

The key goals of the Children's Literature Archive online collection were to promote research, teaching and learning, and to connect with a broader community by building scholarly, online exhibitions. Since this was a collaborative project between the Ryerson University Library and Centre for Digital Humanities, it was very important to keep good communication among all the project team members. It is normal for faculty, librarians, programmers, and students to have different perspectives while working together. Establishing goals and objectives that everyone could agree with was critical.

Looking forward, Digital Humanities students will continue building on this project, growing a searchable catalogue of the materials in the collection, and developing specific archives around particular strengths in the collection holdings. Cataloguers' training to support a more consistent

application of the Dublin Core element semantics in the records creation, as well as training students in a range of archiving skills, bibliographic practices, and digitization and curatorial methods were key aspects in the success of the project, as well as information integration from the end users' point of view.

This project connected the traditional archive with cutting-edge digital methodologies, metadata creation, and integrated teaching, learning, and research. Through the creation of an online scholarly publishing platform based on the Children's Literature Archives holdings, this project has successfully brought various kinds of expertise together to explore innovative pedagogical practices.

Acknowledgements

We wish to thank the principal investigators, students involved in the project, research collaborators, and current previous research assistants in the Ryerson Centre for Digital Humanities, particularly Lorraine Janzen Kooistra, Dennis Denisoff, and Andrew O'Malley for their ongoing support and input.

References

- Dublin Core Metadata Initiative. (2012). *DCMI Metadata Terms*. Retrieved from <http://dublincore.org/documents/dcmi-terms/>
- Han, M-J, Cho, C., Cole, T., Jackson, A. Metadata for Special Collections in CONTENTdm: How to Improve Interoperability of Unique Fields Through OAI-PMH. *Journal of Library Metadata*, 9, 231-238. doi: 10.1080/19386380903405124
- Koutsomitropoulos, D., Alexopoulos, A., Solomou, G., & Papatheodorou, T. (2010). The Use of Metadata for Educational Resources in Digital Repositories. *D-Lib Magazine*, 16(1/2). doi:10.1045/january2010-koutsomitropoulos
- Kucsma, J., Reiss, K., & Sidman, A. (2010). Using Omeka to Build Digital Collections: the METRO Case Study. *D-Lib Magazine*, 16(3/4). doi:10.1045/march2010-kucsma
- Ma, J. (2009). Metadata in ARL Libraries: A survey of Metadata Practices. *Journal of Library Metadata*, 9 (1-2), 1-14. doi: 10.1080/19386380903094977
- Mourton, A. (2011). Digital tools: Zotero and Omeka. *The Journal of American History*, 98(3), 952-953. doi:10.1093/jahist/jar520
- Rose, M. (2012). The Ship Has Sailed and We Aren't On It: How Cataloguers Could Support User Tasks and We Won't. *Journal of Library Metadata*, 12(2-3), 127-139. doi: 10.1080/19386389.2012.699828.