**Recommendations for Improving Patron Access to SCRC Digital Collections**

Prepared by Tricia Mackenzie, Head, Resource Description & Metadata Services &

Peggy Griesinger, Metadata & Cataloging Services Librarian

Importance of Clear, Consistent Metadata

The primary purpose of descriptive metadata creation is to provide access to information and resources for users. Professional literature on metadata quality makes it clear that the key to making digital collections available to our patrons and outside researchers is the creation of high quality shareable metadata. Metadata should conform to standard formats and descriptive content standards that make records more predictable, conform to standard controlled vocabularies, encoding standards, and technical standards.[[1]](#footnote-1) Metadata quality is essential for the reuse of data, especially if it is to be shared outside of an institution, with the Digital Public Library of America (DPLA), for example. Below are suggested areas of improvement in existing SCRC collections.

LUNA



<http://images.gmu.edu/luna/servlet/browseByCategory> – Luna: Browse by Category

This example has a number of errors that create duplication in the subject list:

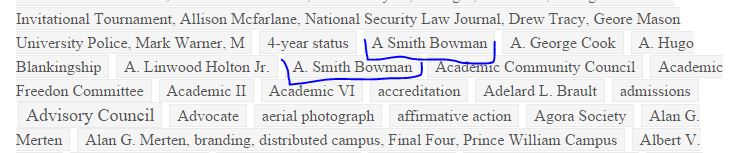
* Punctuation errors mean “Baseball--Photographs” is split into two groups because of a period at the end of some usages of the term and not in others. The same is true of “Basketball—Photographs”.
* The photographs subheading should be used consistently for all items considered photographs. This is an issue because, for example, clicking on “Baseball” also brings up three baseball-related photographs. This means that, taking into account the aforementioned punctuation errors, a user has to view three separate collections ("Baseball", "Baseball--Photographs", "Baseball--Photographs.") to see all of the baseball images stored in Luna.

OMEKA

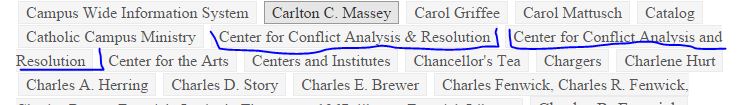


<http://ahistoryofmason.gmu.edu/items/tags> - Browsing tags in Omeka collection “George Mason: A History”

This is a recurring error in this collection: terms were not entered into the system properly, so all of the tags entered for an item are considered one single tag. In this example, “2009, Fenwick Library, Fairfax buildings, construction, Shepley Bullfinch, Lukmire Partnership” is considered one term, when each comma should denote the beginning of a separate term. This problem occurs regularly throughout this collection. This means that items that should be displayed when selecting, for example, “Fenwick Library,” are not showing up in the results because they are lost in this mixed term.



Names are not entered uniformly, so people reappear multiple times. This means that anyone interested in researching A. Smith Bowman would have to know to look for both of these terms and look at both of them.

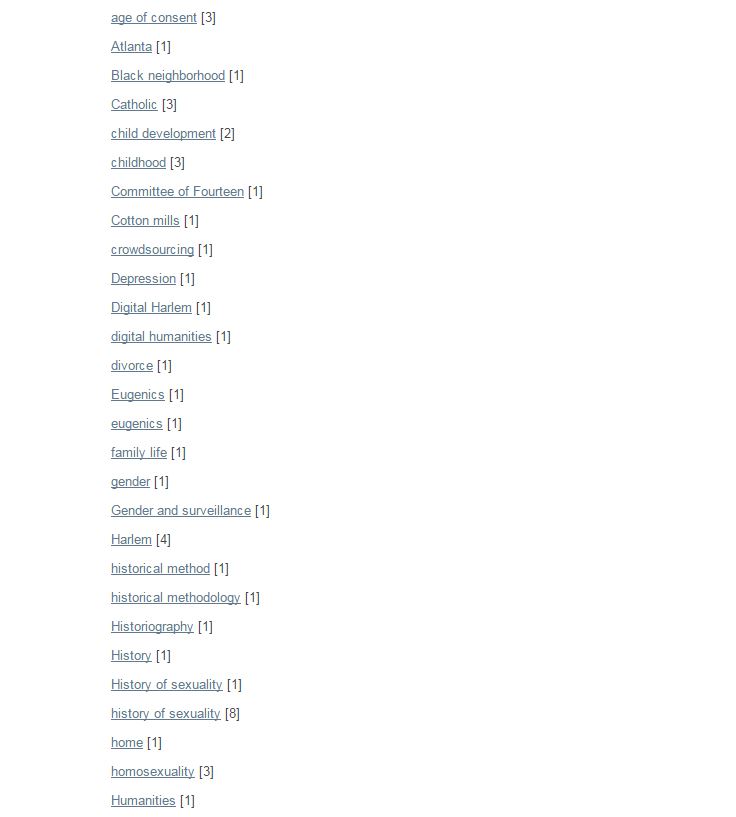


The usage of special characters is not regulated, so both “&” and “and” are used interchangeably, causing duplication of terms.



This example demonstrates how the same (or similar) concepts are represented using different terms. Work needs to be done to determine if these refer to the same concepts and how those terms should be consolidated. The highlighted terms may not all refer to the same concept, but if so, that needs to be clarified.

MARS[[2]](#footnote-2)



<http://mars.gmu.edu/handle/1920/3358/browse> - Browsing Department of History and Art History by Subject

This list of subjects demonstrates the problems that occur when a controlled vocabulary is not used for generating subject headings.

* The same words or phrases are considered separate entities if they are not capitalized the same way, such as “History of sexuality” and “history of sexuality”.
* Two versions of the same word are separated, although they refer to the same concept. See: “historical method” and “historical methodology”.
* Adjectives are used in this list, confusing the semantic meaning of the list. “Catholic” is not a complete term and does not convey its meaning clearly.

All of these problems increase the number of clicks a user needs to make to ensure they’re seeing all the resources on the subject they’re researching. They also make the browsable subject list look messy and unorganized.

Conclusion

Given the importance of metadata to researchers, the quality of metadata is fundamental. “Being conscious of data and metadata quality aspects is a primary need in environments supporting and promoting sharing and reuse of data and metadata like modern data infrastructures do.”[[3]](#footnote-3) Per the information DPLA provides for prospective partners, they suggest that metadata should be as ‘error-free’ as possible and that elements are applied consistently according to the metadata rules that partners employ. It is also important that the metadata will work in a global context and make sense to users that may not be familiar with the subject, field, etc. Additionally, if the implementation of metadata standards varies across collections, those metadata fields may not be included in the DPLA ingest.

Moving forward, we recommend the implementation of a metadata cleanup project for digital collections across our various digital platforms. Using an open source application such as OpenRefine, the Metadata & Cataloging Services Librarian can efficiently address many of the metadata problems in Luna, Omeka, and MARS. She can then tackle any remaining issues on an individual basis. Additionally, we strongly advise that Metadata & Cataloging Services Librarian review and perform quality control on metadata before it is published. “Metadata quality is a characteristic that is directly associated with the digital library value and effectiveness e.g., if metadata quality is poor so is the discovery of the digital library information objects.”[[4]](#footnote-4)

1. Shreeves, Sara, Jenn Riley, and Liz Milewicz. "Moving Towards Shareable Metadata." *First Monday* 11, no. 8 (7 August 2006). [↑](#footnote-ref-1)
2. Please note that the Digital Publishing Production Lead and the Metadata Librarian are currently working on a metadata cleanup project for MARS metadata. [↑](#footnote-ref-2)
3. Tani, Alice, Leonardo Candela, Donatella Castelli. “Dealing with metadata quality: The legacy of digital library efforts.” Information Processing and Management, 49 (2013) p. 1194. [↑](#footnote-ref-3)
4. Ibid., 1195. [↑](#footnote-ref-4)