

## 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.<sup>1</sup> 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.

<sup>1</sup> Shreeves, Sara, Jenn Riley, and Liz Milewicz. "Moving Towards Shareable Metadata." *First Monday* 11, no. 8 (7 August 2006).

## OMEKA

Building 1966 1967 1967, Fenwick, Charles Rogers Fenwick, Fenwick Library, library 1968  
1968 George Mason College Master Plan 1969 1970 1971 1972 1973 1974 1975  
1977 1978 1979 1980 1981 1982 1983 1985 1986 1987 1989 1990 1991,  
basketball, CAA, Colonial Athletic Association, Steve Moran, 1992 1993 1994 1995 1996  
1997 1998 2001 2002 2004 2005 2006 2007 2008 2009, Fenwick Library, Fairfax,  
buildings, construction, Shepley Bullfinch, Lukmire Partnership 2010 2011 2012 2012, Angel  
Cabrera, Vision, Mason Vision, Board of Visitors, BOV, Daniel Clemente 2013 2013, Broadside, A-10,  
Atlantic 10 Conference, Student Government, Mason Players, Passagio, basketball, College Basketball  
Invitational Tournament, Allison Mcfarlane, National Security Law Journal, Drew Tracy, Geore Mason

<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.

Invitational Tournament, Allison Mcfarlane, National Security Law Journal, Drew Tracy, Geore Mason  
University Police, Mark Warner, M 4-year status A Smith Bowman A. George Cook A. Hugo  
Blankingship A. Linwood Holton Jr. A. Smith Bowman Academic Community Council Academic  
Freedom Committee Academic II Academic VI accreditation Adelard L. Brault admissions  
Advisory Council Advocate aerial photograph affirmative action Agora Society Alan G.  
Merten Alan G. Merten, branding, distributed campus, Final Four, Prince William Campus Albert V.

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.

Campus Wide Information System Carlton C. Massey Carol Griffie Carol Mattusch Catalog  
Catholic Campus Ministry Center for Conflict Analysis & Resolution Center for Conflict Analysis and  
Resolution Center for the Arts Centers and Institutes Chancellor's Tea Chargers Charlene Hurt  
Charles A. Herring Charles D. Story Charles E. Brewer Charles Fenwick, Charles R. Fenwick,

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

Venice extension Extension Division faculty senate Fairfax Fairfax Campus Fairfax Chamber  
of Commerce Fairfax County Fairfax County Parkway Fairfax County Schools Fairfax High  
School Fairfax Va Falls Church Fanny Fern Davis Farr Farr tract Federal Theatre Project  
Fenwick Fenwick Library Festival of Exposure Fidel Castro film Final Day Exercises

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<sup>2</sup>

[age of consent](#) [3]  
[Atlanta](#) [1]  
[Black neighborhood](#) [1]  
[Catholic](#) [3]  
[child development](#) [2]  
[childhood](#) [3]  
[Committee of Fourteen](#) [1]  
[Cotton mills](#) [1]  
[crowdsourcing](#) [1]  
[Depression](#) [1]  
[Digital Harlem](#) [1]  
[digital humanities](#) [1]  
[divorce](#) [1]  
[Eugenics](#) [1]  
[eugenics](#) [1]  
[family life](#) [1]  
[gender](#) [1]  
[Gender and surveillance](#) [1]  
[Harlem](#) [4]  
[historical method](#) [1]  
[historical methodology](#) [1]  
[Historiography](#) [1]  
[History](#) [1]  
[History of sexuality](#) [1]  
[history of sexuality](#) [8]  
[home](#) [1]  
[homosexuality](#) [3]  
[Humanities](#) [1]

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

---

<sup>2</sup> Please note that the Digital Publishing Production Lead and the Metadata Librarian are currently working on a metadata cleanup project for MARS metadata.

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.”<sup>3</sup> 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.”<sup>4</sup>

---

<sup>3</sup> Tani, Alice, Leonardo Candela, Donatella Castelli. “Dealing with metadata quality: The legacy of digital library efforts.” *Information Processing and Management*, 49 (2013) p. 1194.

<sup>4</sup> *Ibid.*, 1195.