LTSA Grant Team

January 3, 2018

Group 1 metadata overview report

As a member of the Group 1 institutions participating in the ORBIS Cascade initiative, there is a rich depth of digital content to be shared. While some institutions are further along in digitization efforts, the goal of this group is to have a total of 10600 items contributed by July 2018. The LTSA Grant Team is relying on the Group 1 institutions with direct support from the Metadata Analyst to make that goal possible. With that goal in mind, the Metadata Analyst analyzed each institution’s

collection, to determine how best to support this project. Those findings will be sent individually to each institution.

**GROUP 1 SNAPSHOT**

|  |  |  |
| --- | --- | --- |
| **Group 1 institutions:**   * Densho * Everett Public Library * Multnomah County Library * Oregon Historical Society * Richland Public Library * Seattle Public Library, * Spokane Public Library * Whatcom Public Library   **Staffing:** All institutions are committed to this project through some staffing, with staffing levels varying across institutions.  **Metadata experience:** All institutions have some in-house expertise with Dublin Core, with varying levels of experience across institutions. | **Collections:**  **LTSA Grant (10000 items)**  **Group 1 digital items by the numbers:**  Densho (current collection: 14,000), Everett Public Library (current collection: 2,600); Multnomah County Library (current collection: 2,200); Oregon Historical Society (current collection: 8,300); Seattle Public Library (current collection; ; 18,000; project: ~5,000), Spokane Public Library (current collection: 2,320)  WRH participants: Richland Public Library (goal: 100 items); (Whatcom Public Library (800 items goal; 42 (WRH);  **Digital objects by type:** The majority of digital objects created thus far are photographs, with some oral histories and textual items, such as yearbooks.  **Additional resources: Most institutions have objects in un-scanned backlogs; primary format: photographs** | **Workflows & Systems**  **Digitization workflow:** While the majority of institutions have digitization workflows in place, several are either starting workflows or revising legacy workflows.  Additional workflows that may impact this project are archival processing, intellectual control, and rights management.  **Systems:** Half of the institutions use CONTENTdm. Three institutions will be part of the Washington Rural Heritage Project. For the non-CONTENTdm institutions, OAI-PMH harvesting may need some tweaking. |

**GROUP 1 METADATA READINESS**

In determining readiness of Group 1 institutions for participation in the ORBIS Cascade initiative via the LSTA grant, a variety of analyses were conducted including a census of readiness via a survey sent to potential Group 1 institutions, sharing of resources with Group 1 institutions, metadata analysis by the Project Manager and Metadata Analyst, and through dialogue with individual institutions. Resources shared included Dublin Core Best Practices, version 2.2, workshops on rights, RDA, and OpenRefine materials (for data cleanup and enrichment), among others.

**From the readiness analyses and discussions with institutions, the following areas were determined to be of greatest overall need in terms of training and support**

* **Rights statements**
* **Names (RDA format)**

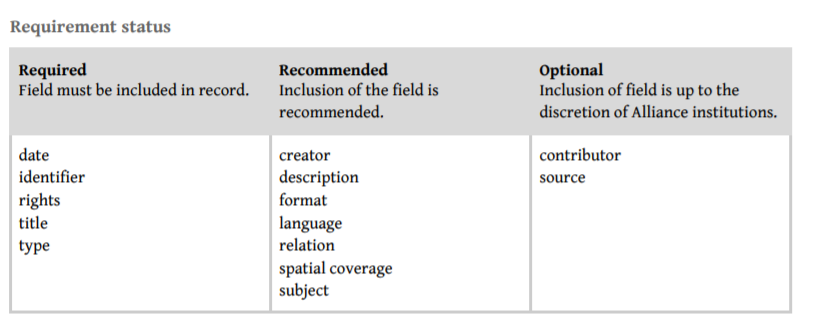
For the majority of Group 1 institutions who are actively digitizing and creating metadata, their metadata meets threshold for required metadata with some exceptions, primarily in the use of the rights statements.

With the recommendation of usage of the Creator field in the [ORBIS’ Dublin Core Best Practices, Version 2.2 document](https://www.orbiscascade.org/dublin-core-best-practices) and the frequent occurrence of names in metadata records in the Creator, Contributor, and/or other local (custom) fields, additional support and guidance in creating names has been requested by Group 1 institutions and as evidenced through metadata analysis. Some institutions also cited a desire for review of the Dublin Core Best Practices, with a focus on rights, dates, and names.

In addition to the rights and name related fields, occasional required elements may be missing for specific collections. All institutions affirmed the need for appropriate data quality control for metadata, citing challenges in staffing, legacy data, and changing standards/practices. The LTSA Grant Team understands the challenges of this type of work and will facilitate quality control workflows, as feasible. Each institution will receive an overview of the metadata analysis for the Group 1 institutions as a whole, in addition to an overview of their metadata analysis, where available.

**GROUP 1 METADATA ANALYSIS OVERVIEW**

**Metadata required element list**



(from Dublin Core Metadata Fields, from [Dublin Core Best Practices, Version 2.2](https://www.orbiscascade.org/dublin-core-best-practices))

**Required Dublin Core Elements**

In reviewing metadata contributions thus far, all Group 1 institutions met the required elements in the Dublin Core Best Practices version 2.2 including date, identifier, title; type was missing in some collections; rights statements encompassed a variety of practices and usage.

**Recommended Dublin Core Elements**

In reviewing metadata contributions thus far, all Group 1 institutions implemented recommended elements with consistency: Creator, Description, and Subject were generally present. Creator while present, was not typically formed following RDA guidelines. Format was included to a lesser extent; language and spatial coverage were rarely included. Spatial coverage was sometimes included as part of a Subject, especially when Library of Congress (LCSH) headings were used. When included as part of the Subject, it was most commonly used as a subdivision.

**Optional Dublin Core Elements**

While optional under the Dublin Core Best Practices, Group 1 institutions did often use contributor and/or source. However, contributor was frequently used to cite the collection or institution. Source was often used interchangeably with contributor for that purpose.

**PROJECT WORK SNAPSHOT**

**Collaborative solutions**

Moving forward will involve a multi-prong approach to best meet the needs of each institution, allowing for more customized and tailored support.

The following resources will be available

* Direct support to the Metadata Analyst through scheduled consultations, email, and open office hours
* Workshops - Best Practices Overview, Names in RDA, Rights management. Metadata Cleanup
* Resources - Sharing of resources for questions, RDA workshops, etc.
* Documentation - Review of documentation and tailoring of documentation to specific systems, where feasible
* Communication - Communication of potential issues (e.g., Contentdm move to https:) , project progress, and more

**Timeline**

* **January 2018**
  + Direct support for Group 1 and Alliance members by Metadata Analyst
    - Meetings with each Group 1 institution to discuss metadata analysis, identify needs, and strategic planning
  + LSTA Team creates and plans workshops based upon needs
  + Resources shared to Group 1 members
  + Metadata review of Alliance metadata by Metadata Analyst
  + LSTA Team internal communication and project documentation
* **February**
  + Direct support for Group 1 and Alliance members by Metadata Analyst
  + Workshops – First Round – LSTA Team Trainer
  + Metadata review of Alliance metadata by Metadata Analyst
  + LSTA Team internal communication and project documentation
* **March** 
  + Direct support for Group 1 and Alliance members by Metadata Analyst
  + Workshops – Second Round – LSTA Team Trainer; Metadata Analyst
  + Metadata review of Alliance metadata by Metadata Analyst
  + LSTA Team internal communication and project documentation
* **April** 
  + Direct support for Group 1 and Alliance members by Metadata Analyst
  + Workshops – Third Round – LSTA Team Trainer; Metadata Analyst
  + Metadata review of Alliance metadata by Metadata Analyst
  + LSTA Team internal communication and project documentation
* **May** 
  + Direct support for Group 1 and Alliance members by Metadata Analyst
  + Workshops – Second Round – LSTA Team Trainer; Metadata Analyst
  + Metadata review of Alliance metadata by Metadata Analyst
    - Review all continuing or unaddressed issues
  + Provide specific solutions (“copy and paste”) for remaining issues
  + LSTA Team internal communication and project documentation
  + DPLA initial harvest feedback and discussion
* **May** 
  + Direct support for Group 1 and Alliance members by Metadata Analyst
  + Metadata review of Alliance metadata by Metadata Analyst
    - Review all continuing or unaddressed issues
    - Provide specific solutions (“copy and paste”) for remaining issues
    - One to one work to resolve issues
    - Continue to work on all unaddressed issues; several issues with harvesting tool caused new issues
  + LSTA Team internal communication and project documentation
  + DPLA initial harvest feedback and discussion
    - Began work to address revisions to rights best practices to meet needs of DPLA
* **May** 
  + Direct support for Group 1 and Alliance members by Metadata Analyst
  + Metadata review of Alliance metadata by Metadata Analyst
    - Review all continuing or unaddressed issues
    - Provide specific solutions (“copy and paste”) for remaining issues
    - One to one work to resolve issues
    - Continue to work on all unaddressed issues including harvester generated issues
    - Began to work with institutions to get metadata to 100% BP compliance
  + LSTA Team internal communication and project documentation
  + DPLA initial harvest feedback and discussion
    - Began work to address revisions to rights best practices to meet needs of DPLA
* **June** 
  + Direct support for Group 1 and Alliance members by Metadata Analyst
  + Metadata review of Alliance metadata by Metadata Analyst
    - Review all continuing or unaddressed issues
    - Provide specific solutions (“copy and paste”) for remaining issues
    - One to one work to resolve issues
    - Continue to work on all unaddressed issues including harvester generated issues
    - Continue to work with institutions to get metadata to 100% BP compliance
    - Compile list of remaining issues for post-grant work
    - Work with PM on grant end transition
  + LSTA Team internal communication and project documentation
  + DPLA initial harvest feedback and discussion
    - Work with institutions to resolve metadata issues prohibiting harvest by DPLA.
    - Second DPLA harvest attempt to occur after grant end.