Metadata Quality Control Workflow

The workflow for creating metadata varies from project-to-project. When metadata is available prior to the implementation stage, such as when MARC records exist, it is created in a batch through transformation scenarios, but when metadata does not exist, it is created during ingest via forms specific to that project.

Despite the varied workflow for creating metadata, metadata QC should be done in batches and, if possible, by someone different from the individual creating the metadata. The *Digitization Workflow* document notes at what point metadata QC occurs. This document also includes resources that can be used to conduct Metadata QC, as well as definitions of QC levels. The level of quality control and the QC indicators for metadata will vary based on the project's metadata specifications.

Stages of Quality Control

There are 2 stages of metadata QC: primary QC, which occurs prior to ingest and is typically completed by the metadata creator, and secondary QC, which occurs after ingest and is typically completed by someone other than the metadata creator. Each stage is detailed below. QC indicators are determined in the project-planning phase and are documented in the *Project Profile: Specifications for Metadata Creation*.

2 Stages of Quality Control	Checklist
Primary quality control check [Completed in batches by metadata creator or prior to form submission]	▪ Adherence to the *Project Profile: Specifications for Metadata Creation*, including correct use of data content, value, and structure standards outlined in project profile ▪ Completeness of **required** metadata fields ▪ Does the MODS/XML validate?
Secondary quality control check[1] [Completed by metadata creator or project team member]	▪ Have the values been appropriately used to describe the object? ▪ Do the values accurately describe the object? ▪ Are the values used consistently throughout the metadata? ▪ Is there missing or incomplete information about the resource? ▪ Is the metadata displaying appropriately in the repository or platform? [Check a sample set of 10%.]

[1] From: Bruce, Thomas R. and Diane I. Hillman. "The Continuum of Metadata Quality: Defining, Expressing, Exploiting," from *Metadata in Practice*. http://www.ecommons.cornell.edu/handle/1813/7895

Levels of Quality Control

Also defined in *Project Profile: Specifications for Metadata Creation* is the level of QC undertaken for the project. There are three levels: low, medium, and high. These three levels have the metadata creator and the QC reviewer, if assigned, measuring metadata against the QC indicators discussed in the section above. **The only difference between the three levels of QC is the number of records reviewed for the secondary QC indicators.** Refer to the chart below to see what percentage of records are sampled for each QC level.[2]

Levels of Quality Control		
High	**Medium**	**Low**
▪ Secondary quality control indicator check (70-100% sample) ▪ MODS/XML validation ▪ Primary QC indicator check	▪ Secondary quality control indicator check (40-70% sample) ▪ MODS/XML validation ▪ Primary QC indicator check	▪ Secondary quality control indicator check (20-40% sample) ▪ MODS/XML validation ▪ Primary QC indicator check

Note: If the metadata creator encounters an anomaly not covered in the QC guidelines or in the *Project Profile: Specifications for Metadata Creation*, he or she should stop with metadata creation and report the problem to the supervising staff member as soon as possible.

Guiding Principles of Quality Control

1. All primary QC indicators are checked during metadata creation.
2. Secondary QC indicator checks are completed after each ingest batch, whether the collections being digitized are small or not.
3. If possible, the secondary QC indicator check should be done by someone who was not responsible for digitization and metadata creation. It is always preferable to have a fresh pair of eyes review and, if needed, revise metadata.
4. When batches of QC are completed, the reviewer will submit comments to the project team to notify them of the status of QC via the project charter's timeline. The timeline will help the project team track when quality control has been conducted and when it is completed.

[2] Metadata quality control will require a more rigorous review of quality and, thus, a higher percentage sample than is assigned for digitization quality control. It is through complete, accurate, and consistent metadata that users are able to successfully perform searches for materials in a digital repository or platform.