# Metadata Quality Control Workflow

The level of quality control (QC) for metadata will vary based on the nature of the project, the project's end goal, and the project's metadata specifications.

Metadata is created at multiple points throughout the digital project life cycle. For example, rights metadata is captured during the planning phase in the *Risk Assessment Matrix*, and later recorded in the project's metadata worksheet, whereas descriptive metadata is captured during the implementation phase directly in the metadata worksheet. The workflow for creating metadata varies from project-to-project. When scanning is done in-house, metadata is recorded simultaneous to scanning. Other times, like when scanning is outsourced to a vendor, it is recorded after scanning has been completed.

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 information about the size of the batch load and resources that can be used to conduct QC.

## Points of Quality Control

There are 4 points for QC: primary QC indicators, secondary QC indicators, and pre- and post- upload indicators. Each point is detailed below. QC indicators are determined in the project-planning phase and are documented in the *Project Profile: Specifications for Metadata Creation*. The required elements and attributes outlined in that document are the primary QC indicators, whereas the recommended elements and attributes are secondary QC indicators.

| 4 Points of Quality Control | Checklist |
|---|---|
| Primary quality control check [Completed in batches by metadata creator] | <ul><li>Completeness of required metadata fields</li><li>Adherence to the *Project Profile: Specifications for Metadata Creation*<ul><li>Correct use of data content, value, and structure standards outlined in project profile</li></ul></li><li>Correct spelling</li><li>Correct formatting, i.e. subjects are separated by semicolons with no spaces between</li></ul> |
| Secondary quality control check <br><br> [Completed by metadata creator or project team member] | <ul><li>Accuracy of metadata<ul><li>Have the values been appropriately used to describe the object?[1]</li><li>Do the values accurately describe the object?</li></ul></li><li>Consistency of metadata<ul><li>Are the values used consistently throughout the worksheet?</li></ul></li><li>Completeness of metadata</li></ul> |

---

[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

| | |
|---|---|
| | ▪ Is there missing or incomplete information about the resource?<br>▪ Does the element set completely describe the objects?[2]<br>▪ Are all relevant elements used for each object?[3] |
| Pre-upload quality control check<br><br>[Completed by UCLA Digital Library and the metadata creator or project lead before upload] | ▪ MODS XML validation<br>▪ Are the MODS XML records accurate?<br>▪ Do the MODS XML records display all of the metadata fields used in the project? |
| Pre-upload quality control check<br><br>[Completed by metadata creator or project lead after upload] | ▪ Is the metadata displaying appropriately in the repository or platform? [Check a sample set of 10%.] |

## 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.[4]

| Levels of Quality Control | | |
|---|---|---|
| **High** | **Medium** | **Low** |
| ▪ Secondary quality control indicator check (70-100% sample)<br>▪ MODS XML validation<br>▪ Primary QC indicator check | ▪ Secondary quality control indicator check (50-60% sample)<br>▪ MODS XML validation<br>▪ Primary QC indicator check | ▪ Secondary quality control indicator check (20-40% sample)<br>▪ MODS XML validation<br>▪ Primary QC indicator check |

---

[2] Ibid.

[3] Ibid.

[4] 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.

For the metadata to be ingested into Islandora, MODS XML records must be created from the digital project's metadata worksheet, typically created in Excel. These XML records must be validated before ingestion.

**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 batch of digitization and metadata creation, 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, such as the project lead or another project team member. It is always preferable to have a fresh pair of eyes review and, if needed, revise metadata before ingest into a digital platform or repository.
4. The QC reviewer will note any QC issues in the metadata worksheet and highlight the change in red, so that the original metadata creator can note the changes made.
5. The worksheet also serves as a way for supervisors to monitor the quality control work.
6. 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.