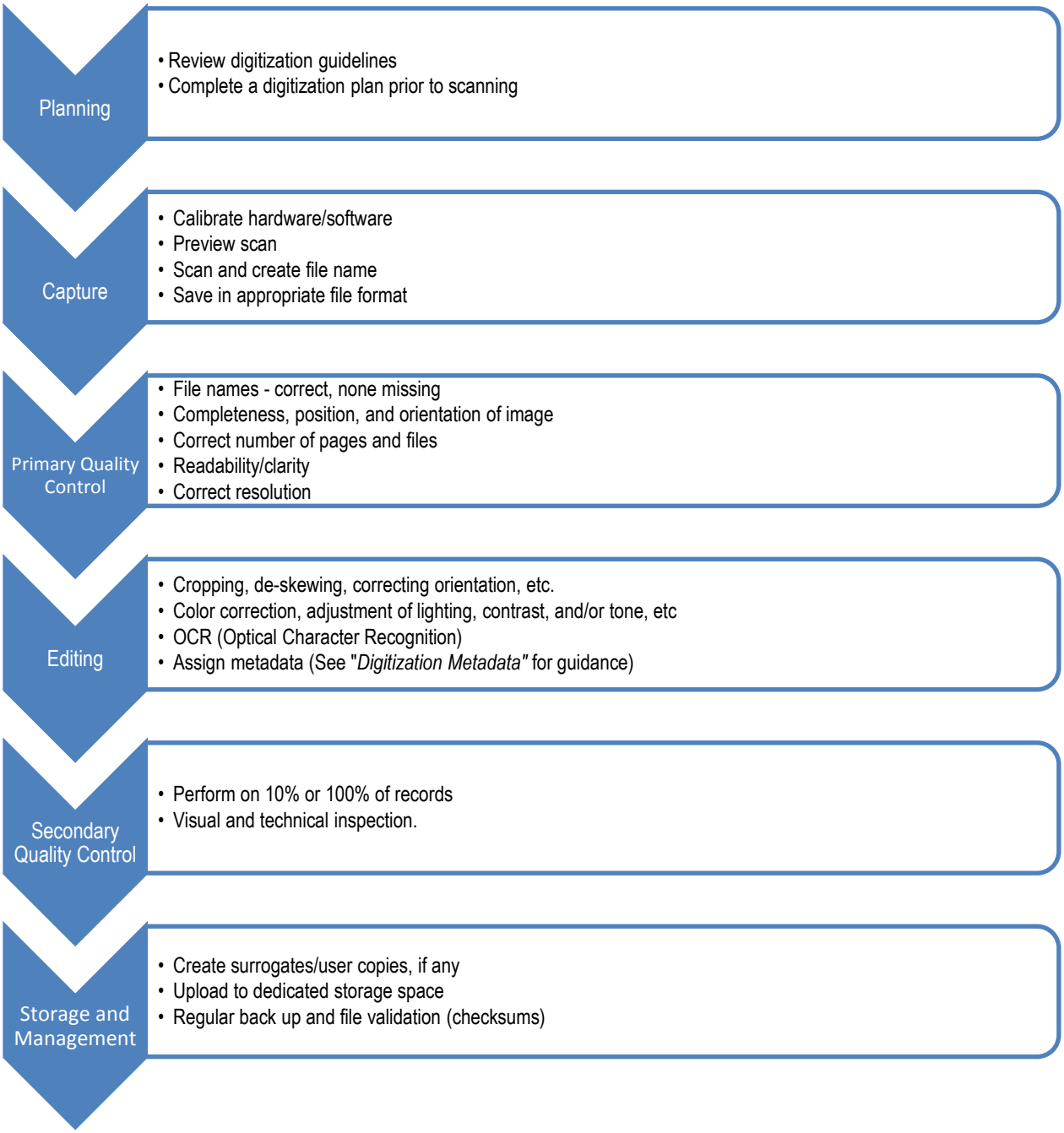
Digitization Workflow

Introduction

A documented workflow streamlines the many processes and tasks involved in a digitization project, facilitating consistency and reliable results, especially when dealing with a large volume of objects to be digitized. Documenting the workflow is also important for tracking material through each stage of the process and identifying errors or problems.

Basic Workflow



Planning

It is important to plan all steps of a digitization project at the outset to maximize efficiency and minimize the risk of making mistakes, potentially costing significant time and resources.

- Complete a "*Digitization Project Worksheet*" before any scanning begins.
- Determine a procedure for documentation in order to track records through each stage of the project. For example, a spreadsheet may be created in which a digital file name, the date an item was scanned, and the name of the person who scanned it may be recorded.

Capture/Scanning

The basic steps of scanning records are listed in the chart on page 1. Please see related sections in this document on hardware/software, minimum recommended standards, technical specifications, file naming conventions and file formats for guidance.

Primary Quality Control

The bulk of quality control tasks are performed during the scanning process, including:

- Confirming all file names follow a naming convention and no files are missing
- Verifying images are complete and positioned/oriented correctly
- Confirming the number of pages and files are correct
- Ensuring the images were scanned at an appropriate resolution. They should be readable and clear on both a computer screen and printed paper.

The "*Quality Control for Digitization Projects*" document will provide you with more detail on steps to be taken during this phase.

Editing

Common standards for the images will be determined during the planning phase, and may include cropping, de-skewing, or correcting the orientation of each image. Color correction or adjustments in lighting, contrast, and tone may be necessary, but it is important to keep in mind that the goal is to create an accurate representation of the item, and not correct permanent defects, such as stains. Conversion to Optical Character Recognition (OCR) may also be performed during this stage, which makes text in the document searchable.

Secondary Quality Control

A final inspection of the files should be conducted to catch any defects that may have been missed previously. If you have a small number of digitized records, it is recommended that you perform a high-level secondary quality control check and review all of the files. For large numbers of digitized records, a low-level secondary quality control check is more common. This entails reviewing 100% of the files at the onset of the project until the proof of concept is complete at which point many projects review only 10% of the remaining files. This includes a visual inspection of the images (e.g. as thumbnails), as well as checking that a file's name, size, resolution, and format are correct. Keep in mind that if you are digitizing records with the intention of disposing of the paper, you may want to plan for a much higher secondary quality check since the digital version will now exist as the legal copy.

The "*Quality Control for Digitization Projects*" document will provide you with more detail on steps to be taken during this phase.

Ongoing Storage and Management

If desired, user copies can be created from the masters and the digitized records will be uploaded to dedicated long-term storage. Ongoing management of digital files requires significant time and resources, which should be considered when planning a digitization project. Regular backups and file integrity testing should be performed in an effort to avoid files being lost or corrupt.

Resources referenced for this document

FADGI - Still Image Working Group. *Technical Guidelines for Digitizing Cultural Heritage Materials*. 2016. http://www.digitizationguidelines.gov/guidelines/digitize-technical.html (accessed June 2017).

University of Exeter. *Digitisation Workflow and Guidelines*. 2009. https://projects.exeter.ac.uk/charter/documents/DigitisationWorkflowGuidev5.pdf (accessed June 2017).