

This workflow was developed at an iDigBio workshop in January 2015. The most recent version is available at <https://github.com/iDigBioWorkflows/FlatSheetsDigitizationWorkflows> and <https://www.idigbio.org/content/workflow-modules-and-task-lists>.

Appendix S9. Module 9: Image Archiving

Ideally, images that are saved for archival purposes should be stored and managed in a digital preservation environment. Digital preservation is more than simply "backing up" your images. Ensuring long-term preservation of digital assets is an ongoing process with many steps and considerations and involves medium- to long-term stewardship and accessibility. A large amount of activity has been devoted to the topic of digital preservation, resulting in guidelines, best practices, and standards. Some of these are provided below as resources and were taken into consideration when developing this module.

We strongly recommend that institutions with access to digital preservation experts consult with these experts about developing a policy and/or plan for the preservation of archival images. For example, libraries have been concerned with this topic for a long time and experts in the institutional library should be consulted for advice and guidance. Institutions with a credible digital asset management system (DAMS) should take advantage of it. Current institutional policies governing the preservation of digital assets should be consulted when developing a collections-based policy or plan.

Below are a few guidelines that will help to ensure a basic level of digital image preservation.

| Task ID | Task Description | Explanations and Comments | Resources |
|---------|--|---|---|
| T1 | Develop a digital preservation policy and/or plan. | <p>Consult with local experts and the extensive literature and web-based resources. Steps outlined here would be encompassed in the policy.</p> <p>Some important considerations:</p> <ul style="list-style-type: none"> ● financial cost <ul style="list-style-type: none"> ○ sustainability is a factor ● responsible parties <ul style="list-style-type: none"> ○ who will upload images to the repository? ○ who will curate images in the repository? ○ who will maintain the repository infrastructure? ○ who will fund the preservation efforts? ○ who will administer access to the repository? | <p>Experts and entities at local institution (IT department, library).</p> <p>See: Corrado and Moulaison (2014).</p> <p>CCSDS (2012). http://public.ccsds.org/publications/archive/650x0m2.pdf.</p> <p>Digital Preservation Coalition (2008). http://www.dpconline.org/advice/preservationhandbook.</p> |

| | | | |
|-----------|--|--|--|
| | | <ul style="list-style-type: none"> ● duration of storage (e.g., 5 years, 20 years, in perpetuity?) ● accessibility <ul style="list-style-type: none"> ○ who can access? ○ ease of access? ○ how will the institution ensure that images remain accessible? ○ how is information about an image (e.g., management, administrative, and specimen metadata) stored and how is this information accessed? ● redundancy ● periodicity of fixity checks ● format migration ● hardware migration ● periodic re-evaluation of policy and/or plan | <p>Information Society Technologies (2013). http://www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf.</p> <p>Brown (2013). Practical digital preservation: A how-to guide for organizations of any size.</p> |
| T2 | Identify assets to be archived and/or preserved. | <p>Will the institution preserve:</p> <ul style="list-style-type: none"> ● master only, ● master and derivatives, ● associated specimen, occurrence data with the assets? | |
| T3 | Assess ownership. | Who owns the images? Does the institution have the legal right to store and provide future access to the images? | |
| T4 | Choose a repository. | The repository is the hardware and software environment where images will be preserved and allows for the images to be accessed. The repository should be able to support the institution's preservation policy. | <p>Hardware and software will depend on preservation policy and institutional milieu.</p> <p>Experts and entities at local institution (IT department, library).</p> <p>See: Trusted digital repositories: attributes and responsibilities (2002). http://www.oclc.org/</p> |

| | | | |
|-----------|---|---|---|
| | | | content/dam/research/activities/trusted/repositories.pdf . |
| T5 | Choose format. | <p>Some recommended formats include: uncompressed TIFF, JPEG 2000, DNG.</p> <p>When evaluating file formats, consider these criteria (Park and Oh, 2012.):</p> <ul style="list-style-type: none"> • functionality • metadata • openness (of the format's specifications) • interoperability • independence (avoid formats that need proprietary software) <p>Some herbaria prefer to archive a raw formatted file, as it provides greater flexibility in post-processing to derivative formats or later correction of image quality problems. However, raw files are considered by some to be a less than desirable preservation format and thus suggest using uncompressed TIFFs or DNGs instead (see Sustainability of digital formats).</p> | <p>See: Park and Oh (2012).</p> <p>See: Sustainability of digital formats. http://www.digitalpreservation.gov/formats/content/still_preferences.shtml.</p> |
| T6 | Choose method to identify assets in repository. | <p>Metadata can be embedded in image file or recorded in a sidecar file (See T7, this module). Mnemonic devices such as barcode as filename, etc., are another option.</p> <p>For a more rigorous approach, refer to the OAIS (CCSDS, 2012) Reference Model and the concept of Preservation Description Information (PDI).</p> | <p>CCSDS (2012). http://public.ccsds.org/publications/archive/650x0m2.pdf.</p> |
| T7 | Associate attribution, management, and specimen occurrence metadata with image. | <p>Depending on the image format and term, these metadata can be embedded (e.g., as EXIF, IPTC, XMP fields) in the image file or provided in a sidecar file (e.g., in XML format). Metadata that do not cleanly map to an existing metadata standard may be inserted into the EXIF UserComment field, either as a string or in a standardized format such as XML or JSON. Another approach would be to use Adobe XMP with the appropriate</p> | <p>Application for modifying image metadata (e.g., Adobe Lightroom, EXIFTTool).</p> <p>See: Morris et al. (2013).</p> <p>Audubon Core standard:</p> |

| | | | |
|------------|--|---|---|
| | | <p>namespaces.</p> <p>At the very least include these terms that are designated as required by AudubonCore:</p> <ul style="list-style-type: none"> ● dc:rights AND/OR dcterms:rights ● dcterms:identifier (required for a media collection and strongly recommended for a media item) ● dc:type AND/OR dcterms:type ● ac:metadataLanguage ● ac:metadataLanguageLiteral <p>Some other recommended AudubonCore terms:</p> <ul style="list-style-type: none"> ● xmpRights:Owner ● xmpRights:UsageTerms ● dwc:scientificName ● ac:associatedSpecimenReference (the specimen barcode or dwc:occurrenceID can be placed here). <p>Do not modify technical metadata that is written by the camera.</p> <p>For a more rigorous approach to dealing with metadata that is adequate for preservation, refer to the OAIS Reference Model (CCSDS, 2012) and the concept of Preservation Description Information (PDI).</p> | <p>http://www.tdwg.org/standards/638/.</p> <p>CCSDS (2012). http://public.ccsds.org/publications/archive/650x0m2.pdf.</p> |
| T8 | Transfer image(s) to repository. | <p>Options include:</p> <ul style="list-style-type: none"> ● batch loading ● UI loading ● automated processes | Internet access. Application for uploading images. Repository. |
| T9 | Validate transfer. | <p>Compare source and repository images:</p> <ul style="list-style-type: none"> ● file sizes ● filenames ● checksums <p>For large numbers of images, this process can be performed in batch.</p> | |
| T10 | Other actions specific to preservation policy go into effect here. | <p>Major recommended additional actions for a basic level of digital preservation include:</p> <ul style="list-style-type: none"> ● provide redundancy (local and off- | Preservation policy and/or plan. |

| | | | |
|--|--|---|--|
| | | site) <ul style="list-style-type: none">• perform periodic fixity checks• perform periodic hardware refresh• monitor for format changes• ensure accessibility• monitor sustainability• monitor for changes in technology | |
|--|--|---|--|

Literature Cited

Brown, A. 2013. Practical digital preservation: A how-to guide for organizations of any size. Facet Publishing, London, United Kingdom.

CCSDS (Consultative Committee for Space Data System Practices). 2012. Reference Model for an Open Archival Information System (OAIS). Recommended Practice. CCSDS 650.0-M-2. Washington, D.C., USA.

Corrado, E. M., and H. L. Moulaison. 2014. Digital preservation for libraries, archives, and museums. Rowman & Littlefield, Lanham, Maryland, USA.

Digital Preservation Coalition. 2008. Preservation Management of Digital Materials: The Handbook. <http://www.dpconline.org/advice/preservationhandbook>.

Information Society Technologies. 2013. erpa Guidance: Digitization Preservation Policy Tool. Electronic Resource Preservation and Access Network. <http://www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf>

Morris, R. A., V. Barve, M. Carausu, V. Chavan, J. Cuadra, C. Freeland, G. Hagedorn, P. Leary, D. Mozzherin, A. Olson, G. Riccardi, I. Teage, and G. Whitbread. 2013. Discovery and publishing of primary biodiversity data associated with multimedia resources: The Audubon Core strategies and approaches. *Biodiversity Informatics* 8(2): 185–197. <https://journals.ku.edu/index.php/jbi/article/view/4117>.

Park, E. G., and S. Oh. 2012. Examining attributes of open standard file formats for long-term preservation and open access. *Information Technology and Libraries* 31(4): 46–67.

Sustainability of digital formats. Planning for Library of Congress collections. http://www.digitalpreservation.gov/formats/content/still_preferences.shtml. Accessed 1 May 2015.