Policy on Acceptable Formats for iDigBio-hosted Images

iDigBio's mission requires that it be able to aggregate and distribute digital images of biological specimens, records, and other objects associated with specimens (e.g., labels and notes) generated by TCNs and other bio- and paleo-collections hosting institutions. This document provides current iDigBio policy as well as recommendations for acquiring, processing, archiving, and distributing still, two-dimensional digital images. We recognize the importance of other image types, including three-dimensional images, and will expand this current policy through time as demand dictates.

The current policy is as follows:

- **Compression:** All images submitted must be internally compressed, and that compression should attempt to maximize compression without sacrificing image quality.
- **Unacceptable formats:** DNG, RAW, PSD, and proprietary formats that are not readable with widely available free or open-source software.
- **Recommended formats:** While iDigBio will accept all formats that are not listed above as unacceptable, the following formats are recommended: JPEG/JPEG2000 with fully preserved Metadata (EXIF and others). We prefer images generated with lossless compression at the camera's native resolution, or very high quality lossy compression if the workflow doesn't permit lossless. (see recommendations below)
- **Archival storage:** iDigBio sees the importance of archival storage and will remain the focal point for data/image storage and is working on long-term storage solutions, recommendations, and guidance. iDigBio currently does not provide archival storage, and hosting of images in iDigBio should not be seen as such. As the current iDigBio NSF agreement does not fund archival storage, within current resource limitations, iDigBio may be able to facilitate storage of images on a case-by-case basis.

Recommendations for the Acquisition, Processing, Storage, and Distribution of Digital Images

iDigBio is pleased to offer the following recommendations for capturing, processing, archiving, and distributing digital images based on work of the Developing Robust Object-to-Image-to-Data Workflows Working Group (DROID), a review of industry standards*, input from the broader collections community, and feedback from iDigBio staff. These recommendations are divided into four clusters: image acquisition, image archiving, image derivatives, and image distribution.

Image Acquisition Recommendations
- Camera RAW is the preferred format for recording camera images.
- Scanner images should be recorded in RAW, if available, or 24 bit or higher TIFF.
- Record at the highest native resolution available for the imaging device.
- Adjust white balance settings to match the light source.
- A visible color checker and scale are recommended, as appropriate.

- Adobe RGB or sRGB are the preferred color spaces and should be stated in the metadata.
  - CMYK color space should be avoided.
- Image file names should be restricted to alpha numeric characters, without spaces or other special characters, underscore (_) and hyphen (–) excepted.

Image Archiving Recommendations

- Archiving images is strongly recommended and is viewed as an institutional responsibility, governed by institutional policy.
  - Archived image files should be permanently retained in a secure, redundant environment on institutional infrastructure or commercial back-up services.
- Uncompressed DNG is the preferred archival format.
  - Images captured in proprietary camera RAW format (e.g., CR2, NEF, PEF, etc.) should be converted to Digital Negative format (DNG).
  - Conversion to DNG can be accomplished via Adobe DNG Converter, available free from Adobe.com, or via Adobe Photoshop or Adobe Lightroom.
- Uncompressed or lossless compressed, unmodified TIFF at 24 bits or higher is an alternative archival format.
- Uncompressed or lossless compressed TIFF at 24 bits or higher is preferred for archiving processed images.

Image Derivative Recommendations

- JPEG and JPEG 2000 images should be saved at their native dimensions (spatial resolution).
  - Manually increasing resolution (e.g., alterations to image size or ppi or dpi settings) beyond that recorded by the camera may insert derived data or unwanted visual artifacts into an image file and is not recommended.
- EXIF, IPTC, and other image metadata (including JPEG 2000 XML data), including that added at time of processing, should be preserved in perpetuity and distributed with all derivative and distributed copies of images.
- Image processing should be carried out only on DNG or other RAW files to prevent accumulated data losses and potential image degradation by repeated processing of JPEG and TIFF image files.
- Excessive image manipulation or compositing, including sharpening, enhancing color saturation or balance, and adjusting contrast should be avoided.

Image Distribution Recommendations

- Images uploaded to aggregators and portals should be fit for display.
  - Lower quality images should be distributed to aggregators and portals only when there is no higher quality image available.
  - When a low quality image is uploaded or distributed, the metadata for that image should indicate that it is the only image available for that specimen.
- JPEG format at native resolution and minimal compression is preferred for distribution via the internet.
  - Lossless compression is preferred.
  - When lossless compression is not an option, lossy compression at the lowest level is preferred.

*Industry References
- Universal Photographic Digital Imaging Guidelines
- The AIC Guide to Digital Photography and Conservation Documentation
  American Institute for Conservation of Historic and Artistic Works
- OpenRAW.org
- Adobe Digital Negative (DNG) Specification