



Digital Imaging Guidelines

I. Purpose

This document contains ECAI-established guidelines for digitizing 35mm slides, 4x5 transparencies, and reflective objects. By following these guidelines, one can expect that their digital images will be of high quality and will be readable by ECAI shared interfaces. These guidelines are also intended to reduce the chance that content providers will have to re-digitize their materials in the near future. Please note that these guidelines are suggested minimum standards only. In some cases, such as with large maps, it may be important to exceed the recommended file size. Compression issues for such large format items are not addressed here.

II. High Resolution Files

The final file size of digitized images will vary depending on the image's original format and the type of scanner used. In addition, the digitization process, itself, will vary according to the equipment used. However, a degree of quality control and consistency can be maintained if minimum file sizes are established. The following minimum file sizes are recommended by ECAI.

Data Type	File Size (Uncompressed .tif)	Scan Resolution (dpi/ppi)
35mm color slide	18MB	360dpi +
4 x 5 color transparencies	35MB	360dpi +
5 x 6 color print	35MB	360dpi +
6 x 10 color print	35MB	360dpi +
35mm black & white negative	8MB	360dpi +
5 x 6 black & white print	10MB	360dpi +
8 x 10 black & white print	15MB	360dpi +

Note: Some scanners will only allow digitization using color settings. (Some color management specialists recommend always digitizing using color settings.) In such cases, digitize all materials according to the color image guidelines (e.g., 35mm black & white negatives can be scanned on color settings to yield an 18MB file.) Once the scanning is complete, the color can be dropped from the digital file, making it a grayscale file. Except under special circumstances (e.g., for digitizing line art), always work in 'grayscale mode,' rather than 'black and white' or 'line art' mode.

III. Naming High Resolution Files

The ways to store, track and recover digital files continues to evolve and improve. In the face of such change and the proprietary nature of most tracking software, ECAI recommends using a simple file naming convention and low-cost storage process that will help content providers track, store, and recover high-resolution files. Database programmers often recommend automatically generated numeric—rather than alphanumeric—file names. The appropriate file name is then attached to the end of the numeric file names. Since it has no 'human readable' elements, this

system may or may not be challenging to use during the digitization and image correction process, depending on project's technical sophistication, the integrated use of an image database, staff expertise, and other factors. A 'human readable' component is intended to help collection managers during the digitization process, as well as to help them locate and retrieve archived high-resolution files in relatively short order, regardless of their infrastructure's level of technical sophistication.

Given that project situations vary greatly, ECAI makes the following recommendations regarding naming conventions:

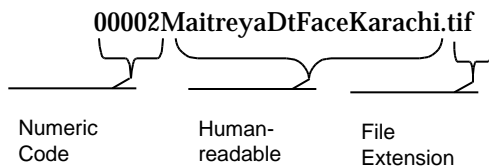
- If possible, use only sequentially numbered numeric codes, with the same number of digits (e.g., 000001.tif, 000010.tif)
- If a 'human readable' component is necessary, it should be placed only after the unique number code (e.g., 000001london_stpauls_w_face_dt.tif)

In this way, each digital item is uniquely coded, its subject matter is identifiable through the alpha character portion of the file name, and the alpha characters can be easily removed during redirection through automated processes. ECAI recommends consulting with ECAItech and your database programmer and system administrators (if available) before making final decisions.

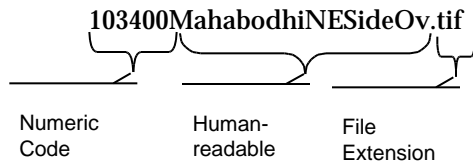
Naming a File

After the numeric code has been added to the file name, the remaining portion of the file names can be 'human-friendly.' It can contain information about the scan resolution, reference to the item's accession number, or descriptive information. This portion of the file name should be entered as systematically as possible. For example, a descriptive section can be entered as general to specific information.

Example of naming convention as used for an object:



Example of naming convention as used for architecture:



Note: For beginning projects that do not yet have a database in place to track their image files, it may be necessary temporarily to adopt a slightly more descriptive file code (i.e., one that includes both alpha and numeric elements). Before adopting such a coding system, feel free to consult with ECAItech about the best way to implement a system.

IV. Storing High Resolution Files

Once the high-resolution, uncompressed .tif files have been created, they can be burned to CDs, DVDs, hard disk, or possibly tape robots for storage. This process is generically termed "archiving," although the appropriateness of this term with respect to digital data has not yet been fully explored by international standards committees.

If CDs are used and the information is digitized in coherent groupings, it is recommended that the .tif files be burned directly to the storage media, without the creation of additional subdirectory folders. This makes recovery much faster. However, subdirectory files may be necessary if digitization cannot be undertaken in a systematic way. The CD/DVD covers and spines should the image code range and any other 'human readable' information that may be useful. For example:

1st CD for Bodhgaya	2nd CD for Bodhgaya
Bodhgaya, India 002200 - 002220	Bodhgaya, India 002221 - 002240

Note: When storing data, it is useful to make natural breaks in the content, even if additional storage space is available on the last CD. For example, if 300 images from the site of Yungang, China, are being prepared at 30MB per image, it will take just under 14 storage CDs to store the image files. Although space is available on the last CD for additional images, the next image set, "Chinese gardens in Suzhou," for example, can be placed on the next set of CDs. This may not be necessary, if one has a very well developed database, storage and backup system. Again, consultation with ECAItech may be helpful in determining the best way to proceed for your circumstances.

V. Image Correction

It is beyond the scope of this document to discuss image correction. However, ECAI would like to recommend that once the original high-resolution images have been burned to CD or stored otherwise, color correction should take place on high-resolution copies of the files. These corrected high-resolution images can then be archived in the same manner as the originals, preserving both the original scans and the high-resolution corrected image files. This will ensure preservation of the original scans *and* that image correction will only need to be done once. Image data redirection (see *section VI*) can be done using the high-resolution corrected images as the starting file.

VI. Redirecting Image Files

In many contexts, such as the World Wide Web and CD-ROM development, it is not appropriate to use uncompressed, high-resolution .tif files. ECAI makes the following recommendations for image data redirection. Once the data is selected for redirection (e.g., 50% of the total images scanned for Japanese Zen Monasteries), the archived, color corrected, high-resolution scan (see *section V*) can be copied and altered as appropriate to the delivery mechanism and audience.

Compression processes, such as Lizardtech's <www.lizardtech.com> MrSid, Zoomify, and the forthcoming JPEG 2000 (IS015444), created by the Digital Imaging Group <www.digitalimaging.org>, allow large files (e.g., 200MB) to be delivered over the Web with ease. Some, such as MrSid, require platform-specific server software. However, new technologies, such as Zoomify, are eliminating this need and are free to download and use for educational purposes.

One can also create smaller static files using copies of the corrected high-resolution files and batching capabilities in software programs like Adobe's Photoshop. However, it may be easier to create a single large image file, which is stored outside of the image database. Through simple programming, a thumbnail or other size file can then be generated 'on the fly' when the database is queried.

Bibliography

Solinet has provided an extensive, simply annotated bibliography on digital imaging. The resource <<http://www.solinet.net/presvtn/leaf/imaging.htm>> includes information on:

- General Sources
- Archiving / Storage
- Intellectual Control
- Scanning Microforms / Hybrid Approach
- Serial Publications
- Electronic Resources
- Special Formats

REVISED 01.10.2000