

Draft Notes for Breakout Session #2 – **Image Production Characteristics**

Session Participants: David Bearman, AMICO
Stephen Chapman, Harvard
Robin Dale, RLG
Don D’Amato, Mitretek Systems
Liping Di, NASA/RITSS
Nick Eiteljorg, Bryn Mawr
John Eyre, De Montfort
Betsy Fanning, AIIM
Carl Fleischhauer, Library of Congress
Erik Landsberg, MoMA
Alan Newman, Art Institute of Chicago
Merilee Proffitt, UC Berkeley
Judy Thomas, UVA
Karla Youngs, TASI

I. DISCUSSION

Several issues were directed to this breakout session from the larger group. Discussion focused on production processes associated with copy digital photography, rather than creating original (“born digital”) images. Erik Landsberg provided a step-by-step description of the production process for digital photography at the Museum of Modern Art. This on-the-spot case study prompted the discussion group to list the system features that contributed to image production and controlled image quality. Decisions of whether or not to document these features as technical metadata elements were based upon the evaluation criteria suggested in the full session:

- What’s the element going to do? For whom?
- Will data elements recorded now be useful later? What are the long-term needs that are to be supported?
- Does the metadata need to be system and human readable? Is it to be used by applications, managers, or users?
- Will the costs of creating this metadata be justified by benefits – either near term or long term? Are tools available to create this metadata?

In addition to beginning a list of proposed data elements (see Section IV), the discussion produced three outcomes: a set of principles (Section II), functional requirements for production metadata (Section III), and general agreement on the categories in which system features should be documented, if not actually classed as metadata elements with controlled vocabulary. We concluded that documentation of systems *and* documentation of rationale (the methodology that informed the production) were both key elements to assessing system performance and controlling image transformations in the future.

The five categories for documenting production metadata are:

- source attributes
- hardware
- software (including the filters/tones curves that were applied to “raw” scanning data)
- the viewing environment
- operator judgment/decisions

On the topic of source attributes, we agreed that these should be limited strictly to the features that are relevant to machine processing of metadata for output – such as h x w dimensions, orientation, or polarity. All other information about the source should reside with the intellectual metadata that is associated with the image.

We also noted that workflow metadata, such as the name of the scanning operator, *might* be relevant, although it would be important to justify the long-term benefit of collecting and retaining this information in electronic form.

II. PRINCIPLES

- ❖ If one can create persistent metadata inside the raster image, it will be possible to resolve a major architectural problem intrinsic to other electronic objects: keeping metadata with files over time. The potential solution to use targets or other embedded metadata is unique to digital images.
- ❖ Point to existing practice in other domains whenever possible. (Note: specific examples were not cited during the discussion; this preliminary list was created by the moderator.)
 - DIG2000 Initiative, http://www.digitalimaging.org/i_dig2000.html
 - ISO JPEG Committee, ISO/IEC JTC1/SC29/WG1, <http://www.jpeg.org/public/jpeglinks.htm>
 - ISO/TC42N, Electronic still picture imaging – Removable memory – Part 2: Image data format – TIFF/EP, <http://www.pima.net/standards/iso/tc42/WG18.htm> -- “**parent image, reduced resolution RGB uncompressed image**”
 - ISO/TC42/WG3/JTG TC130 (ISO/DIS 3664), Viewing conditions for graphic technology and photography, <http://www.pima.net/standards/iso/TC42/17th%20Plenary/N4355.pdf>
 - ISO 16067, Photography – Electronic scanners for photographic images – Spatial resolution measurements: Part 1: Scanners for reflective media,” <http://www.pima.net/it10a.htm#16067>
 - IEC/TC 100/PT 61966, Colour Measurement and Management in Multimedia Systems and Equipment - Part 2-1: sRGB colour space, <http://w3.hike.te.chiba-u.ac.jp/IEC/100/PT1/>; <http://w3.hike.te.chiba-u.ac.jp/IEC/100/PT1/parts/part2/index.html>

- ❖ Document the system and the rationale for how it was used. in order to provide means for assessment (by managers and users) and means for controlled representation and transformations (by systems).

III. Functional Requirements

- ❖ Production/technical metadata must facilitate digital-to-analog conversions (by systems) to meet desired objectives for representation. This is essential for images to be distributed widely and to be retained for long periods in repositories. It must be assumed that devices will change over time in unforeseeable ways. When controls were used at input (during photography), it is highly desirable to avoid “artifacts” such as color shifts, elimination of highlight/shadow detail at output – to all analog formats (screen, print, film).
- ❖ Production/technical metadata must facilitate image assessment (by managers and users). Rationale is a key element to assess system performance and image quality.

In this category, members of the breakout session saw a great deal of potential in recommending two practices other than the traditional method to record metadata as data elements with controlled vocabulary. (We did not, however, have time to establish whether this recommendation should be phrased: *instead of* or *in addition to* recording technical metadata as data elements.) These two practices which deserve fuller investigation are: to photograph targets with the artwork, and to document project specifications as a free-text narrative that would be associated with the image. During the discussion, we found targets to be an extremely useful tool to document the performance of the complete system (all variables working in concert). What also makes targets compelling is that they can be included within the image, thus fulfilling the objective articulated in our first principle above.

IV. ELEMENT LIST

Drawing upon the MoMA case study, the list of issues recommended to the breakout session, and several conference handouts,¹ as well as a suggestion by Carl Fleischhauer, we drafted data element lists in two categories: metadata that are image-specific and project-specific.

¹ RLG Preservation Metadata Element Set (from the RLG Working Group on Preservation Issues of Metadata), MOA2 Technical Metadata Tables (Category #2), the LC-CNRI experiment (Coolidge-Consumerism Experiment), and, from Robin Dale’s notes, a list of elements from the National Library of Australia.

Image Production Element List (**Image-Specific**)

(R) = required

(O) = optional

In-image target(s):

Name (R)

Defined values (O)

Calibration values (O)

System target(s), associated with object:

Name (R)

Defined values (O)

Calibration values (O)

Responsible agent

Rationale:

Free text (see also, Project Narrative section below)

Hardware:

Make, model, serial # of scanner

CCD

Camera lens settings (aperture, focal length)

Camera/scanner color space information (e.g., ICC profile)

Illuminant (type, temperature, filters)

Software:

Driver, model, version

Filters (R)

- discussion about cases where these might be saved as files external to the image

Shadow (black) point

Highlight (white) point

Threshold

Histogram (?)

Image Production Narrative² (Project Specific)

We discussed recording project- or collection-level production metadata and agreed that free prose might be the most cost-effective format to document why imaging was done in a particular way. Carl Fleischhauer used the analogy of scope and content notes as a way of thinking about how to present this information to the user/manager in an ordered, if not actually fully controlled way. Again, time did not allow us to flesh out the contents of such image production narratives, so the moderator drafted the following list for further review and discussion:

Description of the source material

Entire object scanned, or only a portion?

- e.g., covers and blank pages excluded from books

Single type of material or an aggregate of multiple types

- e.g., text only, text + illustrations

Description of the approach that was used in scanning

Were materials categorized into different production workflows or scanned in the same way regardless of format and content?

Which agent(s) was responsible for scanning?

When were materials scanned?

What system(s) were used?

- name(s) of scanner(s), description of its use, quality control
- name of processing software, description of its use, quality control
- description of filters/enhancements used during or after scanning

Were targets used? Which ones and why?

Was the viewing environment controlled? Calibrated to a standard?

How many digital images were produced for each item scanned?

Were texts converted to machine-readable form?

- to what level of accuracy?

Was text marked up? According to guidelines or standards?

Was there an intended output for each copy?

- examples include: 1:1 publication-quality print, 8x10 photographic print, facsimile reprint, full-screen view, thumbnail view, etc.

² Suggested by Carl Fleischhauer. See, for example, the “Digitizing the Collection” essays associated with American Memory Collections, including “California as I Saw It . . .,” <http://memory.loc.gov/ammem/cbhtml/cbedtec.html>