

Automatic Exposure Project Meeting

Sponsored by RLG and NISO

The Venetian
Las Vegas, NV

5 November 2003

Introductions

Günter Waibel, RLG
Robin Dale, RLG
Ben Blackwell, University of California, Berkeley Art Museum
Mikki Carpenter, Museum of Modern Art
Stephen Chapman, Harvard University Libraries
Robert Gregory, Kirtas Technologies
Nancy Hoebelheinrich, Stanford University Libraries
Roger Howard, Getty
Greg King, Sinar Bron
Gunar Penikis, Adobe Systems
Tim Au Yeung, University of Calgary

Brief overview of initiative

Günter Waibel began the meeting with a brief overview of the initiative, its origins, and its ties to the Museum Computer Network (MCN) conference.

The concept that underpins the *Automatic Exposure* project came out of Günter's work with Ben Blackwell at the University of California, Berkeley Art Museum & Pacific Film Archive. In creating images, they were trying to obtain the technical metadata from the digital cameras and found that the cameras did not provide this information, as expected. In an attempt to remedy this, Günter and Ben approached BetterLight, the manufacturer of the in-house camera and asked if modifications could be made to allow the automatic capture of the required technical metadata.

About the same time, Robin Dale spoke at the 2002 MCN conference panel chaired by Günter Waibel. The topic of her presentation, the NISO Z39.87: Technical Metadata for Digital Still Images (Draft Standard for Trial Use), spoke to the exact issue that Waibel was trying to address via conversations with BetterLight. During that same session, it became clear that many attendees were unclear how to access TIFF technical metadata or any other kind of technical metadata. Very few were aware that if this kind of information was recorded by the capture device, it had the potential of being automatically extracted, saving a great deal of time, money, and lessening the potential for error caused by human input of great amounts of metadata. After a brief meeting following the panel, it was determined that this technical metadata activity needed to

move forward and plans were made. During Günter's visit to MOMA NY after the conference, the resolve to launch an initiative was solidified in conversations with Mikki Carpenter and Erik Landsberg.

In January 2003 Günter joined RLG, enabling Günter and Robin to start what would become the Automatic Exposure initiative. It was clear that the first step would be to determine current practice for technical metadata acquisition within the library, archives, and museum community. Before we could begin a dialog with device manufacturers, we needed to determine the types and brands of devices in use, as well as what institutions were doing related to technical metadata capture. Were institutions capturing the technical metadata about for their digital still images? If so, how? Were institutions utilizing the anticipated time-consuming and manual process? What kinds of capture devices were being used in cultural heritage institutions? Predominantly digital cameras? Predominantly scanners? What were the most common devices used? Answers to these and other questions would help us determine the shape and the direction the initiative would take.

By July 2003, a survey had been created and was distributed via relevant distribution and discussion lists. The survey was designed to help us gather the information we believed to be critical, but also served to confirm community interest in an initiative like Automatic Exposure. At the time of distribution, we estimated responses from 25-30 institutions. Surprisingly, we received responses from over 100 institutions before the end of the response period and could quickly confirm that the capability to capture technical metadata for digital still images – especially in some automated way – was of huge interest to the cultural heritage community.

The data gathering process allowed us to determine the “top ten” list of manufacturers in the cultural heritage community (i.e., the top ten manufacturers of devices most commonly owned). The top ten manufacturers make digital cameras, digital camera backs, and scanners, and were identified as:

- Nikon
- Epson
- Microtek
- PhaseOne
- BetterLight
- UMAX
- HP
- Canon
- Creo
- Agfa

The survey results also identified the current practices of institutions capturing technical metadata for images they created, and in some cases, images acquired from partner institutions. In all cases, respondents expressed concern about the availability of technical metadata from the images, the amount of technical metadata required, and how the cost of the traditionally labor-intensive process could be lowered.

Invitations to the 5 November 2003 meeting

The survey responses identified manufacturers with whom we need to work to affect the changes the community desires. Based on the responses, the top ten manufacturers were targeted as participants, as well as several specialty (high-end) device manufacturers which are well represented in cultural heritage institutions, though not in the large numbers of the top ten. [Note: though Agfa made the top ten, they no longer produce scanners and therefore were not targeted to participate in this meeting.]

Discerning the correct representatives in industry was difficult and led us to use a variety of resources to come up with a list of possible invitees. Some contact information came from survey responses, though in general, most surveys provided contact information for sales representatives, rather than the technical people who needed to attend the meeting. In following up these leads, we often hit dead ends or never received replies to our email and telephone queries. Other contacts came through referrals by experts such as Franziska Frey of the Image Permanence Institute, or by contacting industry representatives who had represented device manufacturers on standards activity like DIG35 and JPEG2000. Steven Puglia (National Archives and Records Administration) also suggested that we invite Adobe, since so many organizations that digitize material use Adobe Photoshop as a part of the image post-processing cycle. This recommendation was taken and contacting Adobe led to their participation in the meeting.

In the end, invitations were issued to:

- Adobe (Andrew Salop, Gunar Penikis)
- BetterLight (Mike Collette)
- Canon (Todd Newman, Craig Brown)
- Creo (Alan Comeau, Ilan Carmi)
- Epson (Parker Plaisted)
- HP (Jack Holm, Bob Gann, Gary Dispoto)
- Kirtas Technologies (Lotfi Belkhir)
- Kodak (Robert Reisch)
- Nikon
- Sinar Bron (Greg King)
- PhaseOne (Soren Krog)

We had responses from everybody except for:

- PhaseOne
- Nikon

Some manufacturers who had great interest in the project could not attend the meeting because their expert staff had to attend a conflicting IS&T Color Imaging conference in Scottsdale, Arizona (HP, Canon, Epson). Still others wanted to participate in the meeting, but could not attend for a variety of reasons. Many of these people expressed interest in working on the project, even though they could not attend the meeting. They included:

- **BetterLight.** Mike Collette wrote e-mail essentially committing BetterLight to the outcomes of this project. According to Collette, “The inclusion of Technical Metadata has already been an informal part of our image capture systems for many years, and we welcome the emergence of a broadly-accepted standard for including this information in a more universally-recognized format. Given a clearly-defined objective, we should be able to implement the necessary additions to our software as part of our ongoing development.” In addition, he stated that BetterLight has a copy of the draft NISO Z39.87 standard, and have started to review this document. They do not foresee any technical problems with implementing a standard set of metadata within our industry-standard TIFF image files.
- **Kodak.** Rob Reisch provided us with the results of several surveys conducted by Kodak. The survey results revealed the degree of metadata captured by various levels of devices (consumer, “prosumer”, and professional levels) as judged using a subset of TIFF tags, EXIF IFD camera capture tags, GPS tags, and thumbnail tags. Copies of the survey results were distributed at the meeting as an appendix to the white paper and give the initiative a starting point in determining what metadata tags are currently being captured by devices.
- **HP.** Jack Holm and his colleagues were unable to attend, but have invited Robin and Günter to meet with them at the HP labs in Palo Alto to discuss outcome of this meeting.
- **Creo.** Ilan Carni called us back from the company headquarters in Israel and while they are interested in participating in *Automatic Exposure*, travel from Israel to Las Vegas for a half-day meeting was obviously impossible.

Between conflicts, travel restrictions, and the difficulty of pinpointing correct technical contacts within huge manufacturers like Nikon, we had only a few manufacturer representatives able to join us at the meeting (Adobe, Kirtas Technologies, and Sinar Bron). Those that accepted the invitations played critical roles in the meeting and we look forward to involving other device manufacturers in the initiative in the coming months, especially at the meeting of manufacturers to be held in conjunction with AIIM in March 2004 (New York City).

Goals of the meeting

Given its small size, the group was able to quickly move through and accomplish many of the goals for the in-person meeting. Utilizing the expertise of the representatives present, we set out to accomplish the following:

- Review the white paper’s assessment of the “State of the state”
 - Were the assessments on track?
 - Were current practices adequately represented?
 - Did other initiatives or projects exist which were already solving some of the problems outlined within the white paper?
- Review the requirements / parameters stated by the white paper
 - Were the needs of cultural heritage institutions adequately described?
 - Did other needs exist that were not adequately represented?
- Map these requirements / parameters to existing initiatives

- Discern where holes still exist
- Could some current initiatives be leveraged in some collaborative fashion to accomplish their particular goals as well as the needs of other institutions and organizations?
- Revise the white paper accordingly
 - This white paper will need to be the basis of our conversation with industry representatives. A clear expression of needs, possible solutions, preferred solutions, etc., need to be presented in March. Clear connections to the needs of other industries are crucial if the white paper is to provide the kind of persuasion we hope will spur device manufacturers to design and market affordable devices that capture and expose technical metadata to a certain standardized level.

“State of the State”

The Automatic Exposure white paper

Robin Dale and Günter Waibel authored the paper on the basis of the survey results and additional research as a communication outlining the initiative to both the cultural heritage community and the manufacturers of digital capture devices or relevant software. As the executive summary from the white paper states:

“The RLG initiative *Automatic Exposure – Capturing Technical Metadata for Digital Still Images* seeks to minimize the cost of technical metadata acquisition and maximize the cultural heritage community's capability of ensuring long-term access to digital assets. The goal of the initiative is to lower the barrier for institutions to capture the data elements proposed by NISO Z39.87: Technical Metadata for Digital Still Images [Draft Standard for Trial Use]¹. NISO Z39.87 defines a standard, comprehensive set of data elements key to an institution's ability to manage and preserve its digital images.

To that end, the project pursues a three-fold inquiry:

- *Automatic Exposure* engages manufacturers of high-end scanners and digital cameras in a dialog about what technical metadata their products currently capture, and what technical metadata is required for digital archiving.
- It examines how existing industry efforts can be leveraged to satisfy the data set proposed by NISO Z39.87: Technical Metadata for Digital Still Images
- It identifies and evaluates the existing or emerging tools for harvesting technical metadata developed at individual institutions or by vendors, and explores how those tools could scale to serve the entire community.”

¹ For more information on NISO Z39.87, please see http://www.niso.org/standards/standard_detail.cfm?std_id=731

Feedback on the white paper

Günter Waibel asked the group to provide feedback on the assumptions and conclusions put forth by the paper.

Stephen Chapman noted that the document leads readers to believe that all technical metadata can be automatically recorded, whereas some of it (such as the lighting and its color temperature, who the capture station operator is, etc) simply resides beyond the ready grasp of the capturing device / software. Ben Blackwell added that this was of particular concern to photographers, who do not want to assume the burden of manual metadata capture. He suggested capture station profiles as a solution to this problem. The device software should have an option to profile a particular capture station (for example, describe the lights, name the operator, etc). These profiles could automatically be applied to the files shot during a particular session.

Furthermore, the group unanimously recommended that the white paper be more precise in specifying that technical metadata is a subset of preservation metadata, and not a conclusive preservation metadata set. Robin Dale suggested that we could add a reference to the OCLC-RLG Preservation Metadata Implementation Strategies (PREMIS) Working Group, which is charged with defining a conclusive data dictionary of core elements for preservation metadata.

While the bulk of the white paper makes an argument for the importance of technical metadata and introduces a number of industry and community initiatives grappling with different aspects of the problem, it culminates in a position statement of what precisely the cultural heritage community asks of the industry:

“In summary: *Automatic Exposure*’s preference in terms of encoding technical metadata is NISO MIX XML, external and/or internal to the digital file described by the metadata. Our main criteria for evaluating existing industry initiatives specifying an element set is whether it covers all the elements – or at least the required elements – proposed by NISO Z39.87. Our main criteria for evaluating specific transfer mechanisms for technical metadata are ease of access to the metadata, and broad availability and support for the mechanism itself. Our last resort is to employ manipulation tools such as metadata harvesters (in case the metadata is otherwise inaccessible or sits in header tags) or Extensible Stylesheet Language Transformations (XSLT) in order to transform the offered dataset into NISO Z39.87 elements / MIX XML. Obviously, even this last strategy only succeeds if the industry standard the data derives from maps conclusively to NISO Z39.87.”

The group felt that the statement could be worded even stronger – we prefer NISO Z39.87 as MIX; as a second best option, we’d consider a format which maps conclusively (meaning all elements, not just the required elements) to NISO Z39.87. The elements should be automatically captured whenever possible. Capture or processing software should make allowances for manual entry of elements which can’t be automatically populated.

Greg King, representing the manufacturers, backed up the group's resolve to ask for the full element set and not offer compromises from the get-go. He testified that pressure on manufacturers to deal more comprehensively with metadata is mounting. The main issue for manufacturers is that they have to try to satisfy multiple different communities with their metadata solutions--a challenge which largely has stalled the process. Other communities have mainly asked for mechanisms to transport descriptive metadata with the images. According to Greg King, the key to success lies in convincing manufacturers that the NISO metadata set is applicable beyond the cultural heritage community (a key sector from which to gain support is the advertising sector).

Steve Chapman also emphasized that access to this kind of information is increasingly important as cultural heritage institutions continue to acquire and or purchase digital images. In order to preserve these assets, we will need to make sure commercially produced images contain the technical metadata required for their long-term maintenance and use. For that reason, we need to make sure that the Automatic Exposure initiative addresses technical metadata not only from images being produced locally, but also the metadata available when images are acquired. We need to look beyond production and make sure the acquisition and packaging of other images include technical metadata.

Learning about a New Technology Option: Introduction to Adobe XMP

The research for the *Automatic Exposure* white paper turned up Adobe's Extensible Metadata Platform (XMP) as a technology option which promises to address some of the issues at hand. Adobe XMP is a technology which allows capturing, processing and exchanging metadata by embedding it into digital files. Adobe hopes that this open source technology will become an industry-wide standard for sharing metadata across applications, file formats, and devices. It has been implemented across most of the Adobe product line, including Adobe Photoshop.

While it merits pointing out that Automatic Exposure will not limit its investigations to Adobe's XMP but will also consider other technologies with the potential of solving the technical metadata issue, at this particular meeting, XMP received a lot of attention.

Gunar Penikis (Product Manager, Adobe) introduced Adobe's Extensible Metadata Platform (XMP) to the group (presentation available for download on the *Automatic Exposure* web site).

To summarize Gunar's presentation, with special emphasis on the issues most relevant to *Automatic Exposure*: Adobe's Extensible Metadata Platform (XMP) allows embedding metadata as XML RDF into binary files, and viewing the metadata in applications supporting XMP (such as Adobe Photoshop). A free, open source license, XMP SDK is available for download from Adobe.com for developers who want to use XMP in non-Adobe applications.

Adobe XMP contains a number of pre-defined, built-in namespaces (among them DC, EXIF2.2 and IPTC). Adobe encourages using the elements provided by these namespaces if suitable. Additional namespaces may be defined through the use of custom panels. The additional panels can be shared among a community of users.

The elements defined in custom panels are reflected in the Photoshop CS user interface in the File Info dialog. Photoshop CS allows customization of the File Info panel so XMP metadata can be displayed in your preferred format, using the elements from your custom defined namespace alongside elements from the built-in namespaces. The user interface can also be manipulated to display your preferred names for the units of information.

Adobe XMP populates the namespaces by writing pre-existing legacy metadata from formats such as TIFF fileheaders or EXIF2.2 to the native XMP format. In this way, the technology makes legacy metadata available for viewing and editing.

The embedded metadata can be extracted as RDF XML using a packet scanner that is available in the SDK. Furthermore, metadata can be exported from the File Info dialog in Photoshop CS. Photoshop can also store XMP data separate from the image as a so-called "sidecar" file.

How Adobe XMP applies to Automatic Exposure

Our debate of the technology concluded that Adobe XMP addresses a number of the problems raised by RLG's *Automatic Exposure* initiative. The initiative aims to make a specific set of technical metadata (NISO Z39.87) available for import into preservation systems such as Digital Asset Management databases. Two main issues need to be addressed for the initiative to succeed: first, capture devices have to record as much NISO Z39.87 technical metadata as automatically possible. Second, the metadata has to be exposed for editing (so that information which can not be recorded automatically can be added) and for transfer into digital preservation systems.

Adobe's XMP technology could be harnessed to address the second issue if:

- Automatic Exposure creates and shares a Photoshop CS custom panel reflecting the complete NISO Z39.87 metadata set
- Automatic Exposure creates and shares guidelines for exporting the metadata set out of Photoshop CS, for example as the XMP native RDF XML file
- Automatic Exposure creates and shares an XSLT stylesheet for transforming RDF XML into the NISO Z39.87 reference implementation NISO Metadata in XML Schema (MIX)

The discussion returned to Adobe XMP a number of times throughout the rest of the meeting. Stephen Chapman argued that creating and disseminating a custom panel reflecting the NISO Z39.87 metadata set would demonstrate impressively what metadata we have access to, and where the gaps are. The group agreed that while a custom panel wouldn't net the community more metadata elements, it would serve to highlight the issues. Günter Waibel pointed out that it would be crucial that the export and subsequent

manipulation of the metadata set through XSLT can be performed at the individual file level as well as at a batch level. Roger Howard pointed out that the Getty already uses Adobe XMP to access existing metadata, and signaled an interest in creating a custom panel for the NISO Z39.87 metadata set. Gunar Penikis stated that Adobe could offer help in reviewing and promoting such a panel. Gunar also mentioned that Adobe has been actively courting camera manufacturers to write metadata as native XMP, so far without success.

Next Steps

Towards the end of the meeting, Robin and Günter raised the question of how to move the Automatic Exposure Initiative forward after the Las Vegas meeting. Robin noted that so far, the effort has advanced through the efforts of Günter and herself, but as the complexities of the issues at hand became clear, it also became more apparent that the initiative warrants broader stewardship. Most of the invitees present volunteered to become part of a more formal working group (Greg King, Tim Au Yeung, Roger Howard, Stephen Chapman, Ben Blackwell, Nancy Hoebelheinrich, Mikki Carpenter). Other members for the group should be solicited. As a concrete next step, this group could work on a demonstration project implementing an Adobe XMP panel. More work ahead consists in finalizing authoritative mappings from NISO Z39.87 to various industry standards (such as DIG35), and identifying which data fields cannot be automatically captured. Some community representative could also attend a project meeting with more manufacturer representatives at the Association for Information and Image Management (AIIM) tradeshow in New York on March 7, 2004.