HIGHLIGHTED PAPERS: ARCHIVING 2010

**Obsolete Media Award for
Best Interactive Paper**

Investigations on Color Microfilm as a Medium for Long-term Storage of Digital Data

*Christoph Voges, Volker Märgner, and Tim Fingscheidt,
Technische Universität Braunschweig (Germany)*

Abstract: Digital data storage on microfilm is a promising alternative for long-term storage of digital data. Its estimated lifetime of up to 500 years and the availability of reading devices allow entirely new migration-free storage concepts. This paper presents investigations on the suitability of color microfilm as a medium for digital data storage being an alternative to conventional black-and-white film material. The main question we address is whether the advantage of three color channels justifies higher efforts and expenses related to this material. Therefore, an analysis based on several exposed test patterns has been performed. It turned out that the regarded film in combination with the employed exposure setup is very differently ca-pable of storing data points depending on the color layer. Although black-and-white film material has several advantages in our opin-ion, special cases are pointed out where the use of color microfilm for digital data storage is attractive.

Move It or Lose It: Investigating Digital Curation Portability for Access to Government Information

Christopher A. Lee, University of North Carolina; (USA)

Abstract: A fundamental issue of digital preservation is that information resources must often out-live the systems that are used to maintain them at any given time. It is also important to consider sustainability across the boundaries of collection environments. Portability is an essential consideration. The project called "A Model Technological and Social Architecture for the Preservation of State Government Digital Information," administered by the Minnesota Historical Society, is developing strategies and systems to provide enhanced online access to state legislative materials. The project is testing software and strategies to collect and provide access to state legislative documents and associated contextual information. The long-term sustainability of the effort will require interoperability among a various parties, including (1) those who might share responsibility for the preservation of legislative resources from Minnesota, and (2) collecting institutions from other states who would like to make use of the project's methods and software. The author is investigating characteristics of the state legislative information system that are most likely to support or hinder portability of software and digital objects across the boundaries of organizations. The findings from this investigation should be relevant to information professionals responsible for digital collections or collection management systems that must be sustained across the boundaries of specific technical or organizational arrangements.

To view full papers go to
www.imaging.org/ist/publications/reporter/index.cfm

* These papers were presented at the Archiving2010 conference, held June 1-4, 2010, in The Hague, the Netherlands.

INSIDE THIS ISSUE

User Requirements for a Next Generation Digital Preservation Framework: Analysis and Implementation

Brian Aitken,[1] Perla Innocenti,[1] Seamus Ross,[2] and
Leo Konstantelos[1]; [1]University of Glasgow (UK) and
[2]University of Toronto (Canada)

Abstract: The EU-funded SHAMAN (Sustaining Heritage Access through Multivalent ArchiviNg, http://www.shaman-ip.eu/) project is investigating the long-term preservation of large volumes of digital data in a distributed environment, by developing a preservation framework that is verifiable, open and extensible. During the initial stages of the project, a detailed user requirements analysis led by HATII at the University of Glasgow was conducted across three domains: memory institutions, industrial design and engineering, and e-science. This research pinpointed the needs and expectations that end-users and service providers feel should be met by such a preservation framework. This paper gives an overview of the requirements that were gathered, formulated and adopted by this project. It then discusses the outcomes of this empirical research and indicates both how these outcomes are being implemented within SHAMAN and how external parties may also benefit from the findings.

Approaches to digital preservation are often still ad hoc and based on a single institution focus. They frequently do not take into consideration the needs of the variety of actors who will come into contact with a system throughout the preservation lifecycle. This paper provides an insight into the preservation practices that a broad range of real-world organisations would like to follow and provides a discussion of how SHAMAN intends to meet the needs of the identified users.

Demonstrating the Next Generation Preservation Framework

Thomas Wollschlaeger, Attila Zabos, and Denise E. Keifer,
Deutsche Nationalbibliothek (Germany)

Abstract: The goal of the SHAMAN project is the development of fundamental concepts and technologies for the next generation Digital Preservation network system. The German National Library and its partners have created a demonstrator that showcases the distributed instances of the envisaged archival concept. The demonstrator is aligned in accordance with the new archive-centric information lifecycle model and connects new innovative technology of the international partners, especially of notable universities and technology enterprises, to foster the distributed ingest into an enhanced preservation framework and to adapt services that are used to manage, preserve, enrich and access digital data.

The paper will outline the developmental goals and focus especially on the development work of the demonstrator. By highlighting these aspects, the innovative approach of SHAMAN and the objective to provide a next generation Digital Preservation framework will become visible.

Visualization for Archival Appraisal of Large Digital Collections Using a Random Noise Target

Weijia Xu, Maria Esteva, and Suyog Jain Dott,
University of Texas at Austin (USA)

Abstract: Our research examines data-driven visualization methods for archival purposes. Using data extracted from a large and heterogeneous digital collection, we created an information visualization that uses RDBMS and treemap to enable archival analysis. Different views present the collection's structure and properties at different levels of aggregation and abstraction, transforming 1,000,000 data points into information that enables observation and decision-making.

Mass CD/DVD Migration: A Novartis Case Study

Daniela Bienz and Rudolf Gschwind, University of Basel, and
Mario F. Pozza and Ludwig Gantner, Novartis Pharma AG (Switzerland)

Abstract: Nowadays computers are an essential part of modern pharmaceutical research and data storage. It is also well known that experimental and observational data is crucial for scientific research. For approximately twelve years now, the pharmaceutical industry is storing its research data on portable storage media, where each research group was obligated to archive their research data on such portable media but they had also much individual responsibility for the archiving process. So far there are over 3000 portable media in the raw data archives of Novartis. Digital data archives are more unstable than paper-based archives basically due to development of computer technology and rapid changes of digital storage media. To guarantee the future readability and to preserve the research data, Novartis has decided to migrate the data from the current portable storage media to a storage type more suitable for long-term preservation while preserving the authenticity and provenance. In this paper, we would like to report on this migration project that nicely illustrates several problems of digital archiving and provides state-of-the-art solutions for them.

Archiving 2010 Held in The Hague

by Peter Burns (Carestream Health, Inc.) and Inge Angevaare (KB)

S&T's seventh annual Archiving conference came to the Koninklijke Bibliotheek (KB), the National Library of the Netherlands, in the first week of June. The conference was chaired by Simon Tanner (King's College London) and Astrid Verheuson (KB). The excellent facilities and support from the KB contributed to a congenial atmosphere for both institutional and private-sector participants. The conference was spread over four days, with the first dedicated to short courses and a welcome reception at the Park Hotel. During the last afternoon there were several tours of interest to conference participants. The conference program included oral and interactive poster presentations, as well as a conference reception at the Foto Museum. Each day started with a keynote speaker who helped set the stage for the day's talks.

ARCHIVING 2010 STATS

Attendees:	113
Oral Papers:	34
Interactive Papers:	8
Short Courses:	9
Exhibitors:	10
Dates:	June 1-4, 2010
Location:	The Hague, the Netherlands

Short Courses

Nine short courses were offered in a program that included: preservation business models, microfilm archiving, project management, JPEG 2000, imaging performance monitoring, and imaging standards. On this last subject an extended overview course on international standards for imaging performance was organized. This was well-attended and continued a theme from previous Archiving conferences: the adaptation of imaging and color-management techniques to international standards for archives and museums.

Keynotes

The conference opened with Bas Savenije, (KB) shedding light on the Dutch approach to digital preservation. The Netherlands Coalition for Digital Preservation (NCDD) has just published "A Future for Our Digital Memory (2): Strategic Plan for Long-term Access." Within the public sector, four domains can be identified with particular digital dynamics: scholarly communications, public records, media, and cultural heritage. Within each domain the NCDD assigns a network leader the task of bringing about the establishment of dedicated storage facilities and, even more importantly, of making sure that roles and responsibilities are clearly defined. The network leaders in the Netherlands include national organizations such as the KB, Data Archiving and Networked Services (DANS), the National Archives (Nationaal Archief), and the Netherlands Institute for Sound and Vision (Beeld en Geluid). Savenije also shared the news that the public archives have developed a plan for a shared services organization for digital archiving which was formally signed on June 15.

William Kilbride (UK Digital Preservation Coalition) kicked-off day two with "Digital Preservation in Byte-Sized Chunks: Good Practice, Best Practice and Why We Should Be Careful What We Wish For."
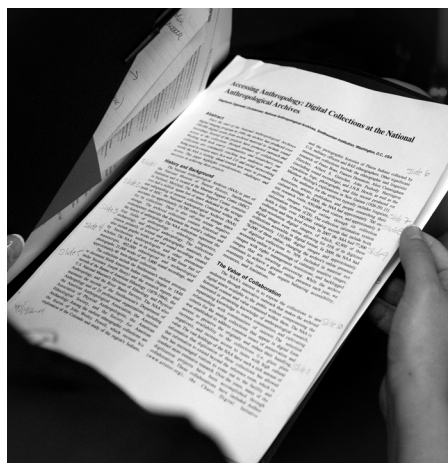
(Above) Bas Savenije, general director of the Koninklijke Bibliotheek, opened the conference with a keynote on digitally preservation efforts in the Netherlands. (Right) Dietmar Wueller (Image Engineering, on left) was one of 10 exhibitors at this year's conference, in addition to being a short course instructor and author.

Photos: Inge Angevaare (KB).

He started by observing that digital preservation (DP) literature makes gloomy reading because all we talk about are problems. This can lead to inactivity, especially in smaller organizations. In addition, this approach can deter potential DP funders. In his inimitable good-spirited fashion, Kilbride challenged the audience to turn this thinking around and to focus on opportunities and the value of the digital assets we preserve. He recommended a new basic assumption, that "information is an asset to [an] organization, like people or estates." He encouraged participants not to focus on data, access, and risks, but rather on people and opportunities. This was illustrated by his reminder to tailor solutions to the size and resources of an institution. "The OAIS model for digital preservation was designed for NASA [the US National Aeronautics and Space Administration]," Kilbride emphasized, "no wonder it does not work for a museum with 5 staff and 10 volunteers. The best advice we might give them is to get the backup off the floor in the store room."

In the third keynote, Martin Jurgens (photograph conservator and recipient of the 2010 HP Image Permance Award) presented "What is (and What Isn't) a Photography Today?" He reminded the audience of the variety of media and processes that are now considered part of a 'photograph', emphasizing the need not only for knowledge of physical processes and equipment, but also a concise language to describe the content of collections. This will facilitate communication between conservators, curators, archivists, and the larger community, he said.

Conference Sessions

The technical papers presented were divided into to three core areas of interest:
- Digital Collections
- Capture, Imaging and Workflow
- Access, Strategies and File Formats

There were many very good papers this year, and a few themes emerged.

Photo: Inge Angevaare (KB).

William Kilbride, executive director of the Digital Preservation Coalition, gave the Thursday morning keynote.

Several presentations were devoted to digital asset management (DAM). Howard Goldstein (Center of Digital Imaging, CDI) stressed that DAM is not the same as a digital repository and should never be taken on in isolation. Rather it should be integrated into regular workflows to be successful. He referred to the example of the Dutch Rijksmuseum, where such integration was accomplished.

Photo: Inge Angevaare (KB).

General Co-chair Simon Tanner (King's College London), explains the details of an interactive paper by a colleague unable to attend the event.

Photographer and dpBestflow Project Manager Peter Krogh recommended use of Adobe's Digital Negative (DNG) in combination with Parametric Image Editing (PIE) for a sound photography DAM workflow—from creation through to archive. DNG has not been widely adopted yet, but some members of the audience indicated that they would have another look at it. Krogh's recommended workflow allows for storing both the original

image and various renderings, with validations at each step.

The audience was impressed with Scott Geffert's (CDI) account of rich data imaging at the Van Gogh Museum. This approach involves capturing more than one still image of an object that can be used for enhanced visualization. The use of a combination of a polarized light, and varying illumination geometry was demonstrated for Van Gogh's paintings. When presented on a touch-enabled device (*e.g.,* iPad), the content is dynamic and intuitive—by touching on the right or left of the image, the light seemed to come from that angle.

As mentioned above, a continuing theme of the imaging element of this annual conference is the development of standard methods and tools for the monitoring and improvement of imaging performance and practice in various institutions. These have resulted in the development and adoption of several international guidelines and standards in this area.

Hans van Dormolen (KB) presented updated information on the Metamorfoze Guidelines. These are being developed by the KB for digital conversion projects, and specify methods and specific results (performance levels). He described how test targets, such as the ambitiously named Universal Test Target (UTT), could be used with commercial software. The approach taken in the US Federal Agencies Digitization Guidelines Initiative (FADGI) is to emphasize common methods and tools, (similar to those of Metamorfoze), but allow specific performance levels to be based on the collection content and project intent.

During his presentation van Dormolen appeared to acknowledge this diversity of needs by introducing Metamorfoze Light into the guidelines, for material that is less demanding. An example was provided by Edwin Klijn (KB), who described an eight-million-page newspaper digitization

A course on International Standards, taught by a cadre of experts, including Hans van Dormolen (KB standing) and W. Scott Geffert (Center for Digital Imaging; far right) offered a chance for hands-on information sharing.

Photo: Inge Angevare (KB).

project the is executed under other simplified image quality assurance monitoring.

Another way to tailor workflow testing to collection material is to specify color test targets with characteristics in common with collection material. In "Capturing the Color of Black and White," Don Williams (Image Science Associates) and coauthor described how this could be done for near-neutral materials using a set of 19th century photographic prints. During the following week, at the IS&T CGIV Conference in Joensuu, Finland, Giorgio Trumpy (Inst. of Applied Physics) presented a related paper, on the digital reproduction of small-gamut objects based on custom color targets.

The session on imaging performance monitoring continued with a description of the methods used for genealogical-record imaging at FamilySearch International by Richard Laxman. He described the digital pipeline for the capture of image- and metadata for microfilm and digital camera digitization. FamilySearch's scanners have to capture all of the information from old, handwritten documents, and they must be presented to genealogists and volunteers who are often senior citizens with poor eyesight and limited computer skills for further tanslation. Paul Jones (CertifiMedia) presented, Image Quality for the Real World, which emphasized the used of automated high-speed analysis, and Adobe's Extensible Metadata Platform (XMP) to store analysis results into image headers.

The format of digital archives was also addressed in several ways. Robert Sharpe (Tessella) presented "Managing 'Visually Lossless' Compression with JPEG2000." This lead to discussion of criteria for visual image differences. The use of raw image formats discussed at previous conferences was addressed by Michael Bennett (University of Connecticut Libraries) in collaboration with Barry Wheeler (US Library of Congress). Thomas Zellman (LuraTech) then presented, PDF/A-2: The New Part of PDF/A. Recent developments include JPEG2000, transparency, ICC (color) profiles, Open Type fonts, and collections (portfolios). A different type of format was discussed by Florian Müller (University of Basel) in "The Digital Time Capsule and Other Applications of Microfilm." He described the permanent visual archive approach to migration-free archiving.

The IS&T 2010 Conference has shown once again that it is a great place to meet to develop preservation strategies and imaging technologies. The range of speakers and the subjects covered provides an excellent overview of the field. Papers from the conference proceedings book are available at imaging.org; there is no charge to download PDFs for IS&T members.

Archiving 2011

Next year's Conference will be held 16-19 May in Salt Lake City, Utah (US). The general chairs are Wayne Metcalfe (FamilySearch) and Kate Zwaard (US Government Printing Office). ▲

A New Approach to Accessing Images forEducation

Kalliopi Vacharopoulou[1], David Penfold[2],
Lindsay MacDonald,[2] and Pandora Mather-Lees[1];
[1]The Bridgeman Art Library and [2]University of Arts London (UK)

Abstract: The paper presents the results of a research project aimed at enhancing a digital image resource for educational users which contains fine art, cultural and historical images that are individually keyworded and catalogued. The project aimed at identifying the needs of the sector in terms of image resources and directed the management and the development of the image collection towards addressing these needs. In the process, the definite need for high-quality (metadata and resolution) images

in teaching, research and studying was identified, together with the requirement of exploring are who the users and what are their image needs in accordance with the level and field of study. However, what also emerged was a lack of awareness of quality resources (either free or upon payment of a fee) that provide copyright-cleared images and what they can offer within education institutions for students and academic staff.

Users of the resource confirmed that they are sometimes not convinced about how learners and educators would use it since they are not accustomed to online resources. If images cannot be found, students and academic staff will settle for what they can find and not what they need. Academics also confirmed their lack of time to search for images in a wide range of databases. Visual resource professionals suggest a variety of ways (from wikis to user tagging to blogs) in which searching and retrieving images can be effectively achieved. As a result, the focus of the second part of the project became the

STANDARDS UPDATE:

David Q. McDowell, Editor

This issue of Standards Update is a potpourri of a variety of issues from ISO TC 130 (Graphic technology) and TC 42 (Photography).

TC130 Graphic Technology

New activities

Carbon footprint At the recent TC 130 meetings in Switzerland (April 2010) two new activities were initiated. These were TC1130 Task Force 1 whose purpose is "to produce a report for TC 130 outlining options for a possible New Work Item (NWI) to explore alternatives to harmonise methods of carbon footprint calculation. This NWI would investigate ways to establish a common metric for measuring and sharing carbon footprinting data, based on a common methodology, such as PAS 2050 or ISO draft 14067. Options to achieve this include a guidance document for a specific framework methodology; Product Category Rules (PCRs) that printers and publishers could use in conjunction with methodologies; an implementation document for a framework methodology, combined with PCRs; or a unique bespoke methodology."

Postpress standards At the 2009 Plenary meeting in China, TC 130 established a new Working Group on "Postpress requirements" provided a NWIP on this topic passes the NWIP ballot. The convenorship and secretariat for the new WG will be assigned to China. Brazil has offered to assist in the support of this working group. This group met for the first time in Switzerland and recommended that a Task Force (TF2) be created to develop an overall vision of the standards that are needed in the postpress area and a plan for the initiation of work in this area.

XMP Standardization Representatives of DIN and Adobe Corporation attended the meeting of TC 130 WG2 in St. Gallen to request TC 130 to consider accepting the Extensible Metadata Platform (XMP) Specification as an ISO document using the "fast-track" process. Agreement was reached to proceed in this standardization effort. The current document is being revised by Adobe to create a new draft that incorporates the data model, serialization, and core properties into a single document and also makes use of the ISO language style and formatting. Published by Adobe in early July, this document and will be placed in the fast-track balloting process.

The application-specific parts of the XMP specification (both those currently documented by Adobe and new applications) will be reviewed by a series of Joint Working Groups (JWGs) of the concerned ISO Technical Committees with TC 130. These JWGs will prepare additional parts of the XMP document series.

Ongoing projects

Process Agnostic Workflow ISO 15339-1, *Graphic technology—Printing of digital data—Part 1: Basic principles,* which is currently in CD ballot, has been colloquially referred to as a "process agnostic" standard. It has as a basic premise that a limited number of color gamuts, and associated characterization data, can form the basic reference printing conditions for the production of printed material from digital data, *regardless of printing process to be used* for final reproduction.

For many, it describes the color managed workflow that has become the norm for much of the commercial color printing world. Although it can be viewed as a replacement for much of the ISO 12647 *(Graphic technology—Process control for the production of half-tone colour separations, proof and production prints)* series, both documents will need to co-exist for some time as the industry transitions between the traditional process control based definition of printing and the new color managed characterization data world of printing.

SCID Images The family of Standard Colour Image Data standards, the ISO 12640 series, centimes to grow. ISO 12640-4, *Graphic technology—Prepress digital data exchange—Part 4: Wide gamut display-referred standard colour image data [Adobe RGB (1998)/SCID]* is in the final stages of publication at ISO.

The newest member of he family, ISO 12640-5, *Graphic technology—Prepress digital data exchange—Part 5: Scene referred standard colour image data (RIMM/SCID)* is starting CD ballot in July. RIMM/SCID represents a significant departure from the other SCID standards and significantly extends the SCID family of images. As its title says it represents scene referred data. Within the standard, the scene-referred image state is defined as the "image state associated with image data that represents estimates of the colour-space coordinates of the elements of a scene". Since virtually all image data that we normally deal with is rendered data of one sort or another, this is a significant departure. It will provide a set of standard data for evaluating image transforms for applications such as camera raw data, image capture, film simulations, scene compression, etc.

ICC Specification *ISO 15076, Image technology colour management—Architecture, profile format and data structure—Part 1: Based on ICC.1:2010,* is in the final stages of publication at ISO. It will also be published simultaneously by the ICC and will be available through both organizations. This latest revision of the ICC specification incorporates all of the changes approved by the ICC since the 2005 version and also represents a significant effort by the ICC to insure consistent use of terminology and consistency between tag requirements in various parts of the specification.

PDF/X Standards Both PDF/X-4 and PDF/X-5 (ISO 15930-7 and ISO 15930-8) have undergone minor revisions to

improve their ability to be used in the packaging industry and to correct some minor mis-interpretations contained in the original publications. The revised documents were published in July 2010.

In addition, ISO 16612-2, *Graphic technology—Variable data exchange—Part 2: Using PDF/X-4 and PDF/X-5 (PDF/VT-1 and PDF/VT-2)* will be published by ISO in late July.

The publication of these documents completes the currently scheduled family of PDF-based graphic arts standards. Obviously, additional members of these families will be needed eventually, but for now the standards community is taking a breather in this area while the applications catch up to the standards.

Soft proofing A new standard, ISO 14861, *Graphic technology – Requirements for colour proofing systems using electronic displays,* is in CD ballot. This is accompanied by a New Work Item (NWI) and CD ballot for the revision of ISO 12646, *Graphic technology—Displays for colour proofing—Characteristics and viewing conditions.* The intent is that these two documents will compliment each other by having the monitor requirements totally contained within ISO 12646 and the soft proofing system requirements defined in ISO 14861. This will insure that there will not be a divergence between the general display requirements and those used for soft proofing within the graphics industry.

Future Meetings

The TC 130 2010 Plenary will be held in Sao Paulo, Brazil October 11 to 16. The 2011 meetings are tentatively scheduled for April 11-15 in Berlin, Germany and in the Sept/Oct time frame in Tokyo, Japan.

TC 42 Photography

Several imaging related standards are being worked on in TC42 WG18 and related JWGs.

The ISO 22028, *Extended colour encodings for digital image storage, manipulation and interchange,* series in particular is receiving considerable attention.

ISO 22028-2 *[Part 2: Reference output medium metric RGB colour image encoding (ROMM RGB)]* A NWI proposal and CD is currently being balloted to elevate this document from a Technical Specification to an International Standard.

ISO 22028-3 *[Part 3: Reference input medium metric RGB colour image encoding (RIMM RGB)]* Work is underway to prepare a revision to Part 3 that will add floating point encoding to enhance the flexibility of this colour space in recording scene based imagery.

ISO 22028-4 *[Part 4: European Colour Initiative RGB colour image encoding (eciRGB [2008])]* Part 4 is in CD ballot as a Technical Specification.

In addition revisions are in preparation for:

ISO 12234-1 *(Electronic still-picture imaging—Removable memory—Part 1: Basic removable-memory model)* The intended revision will add JPEG 200 data compression and also new metadata capabilities.

ISO 15739 *(Photography—Electronic still-picture imaging—Noise measurements)* The planned revision will make the visual noise measurement section normative and will update the sections on dynamic range, viewing conditions, and will update the luminance CSF to preserve the DC level.

Future Meetings

The next TC 42 Plenary is schedule for June 2011 in Rochester, NY.

Information from ISO

ISO and benefits of standards in "CEO speak" ISO has just publicly released a well-received communication package specially targeted at business leaders on the benefits of using its standards and participating in their development. Up to now, Today's state-of-the-art solutions for CEOs has remained for the exclusive use of ISO's national members in enhancing their communication with top management in their countries. Reactions from the ISO members in countries as diverse as Germany, New Zealand, South Africa, and the US have been positive and the package, first published a year ago, now goes on general release. For more information: www.iso.org/iso/pressrelease.htm? refid= Ref1330.

CD collection of ISO standards on mechanical vibration, shock and condition monitoring ISO has just published a CD-ROM compilation of 202 standards and related documents addressing the field of mechanical vibration, shock, and the condition monitoring of machines, including vehicles, and structures, such as bridges and buildings. The documents accessible on the CD-ROM include the entire portfolio of ISO technical committee ISO/TC 108, *Mechanical vibration, shock and condition monitoring,* as well as a selection of other related ISO standards.

For more information: www.iso.org/ iso/pressrelease. htm?refid=Ref1322.

For suggestions for (or input to) future updates, or standards questions in general, please contact the editor at mcdowell@npes.org or mcdowell@kodak.com.

classification of the available images into sets, to facilitate their searchability and retrievability by users. Consequently, sets of images and search indexes were created for subjects not strongly related to the arts. The completed sets relate to Graphic Design and History of Science. Work is currently being undertaken to create an Archaeology set and a Fashion and Costume set, together with two sets on Asian and Indian Art. The Graphic Design Search Index is the result of a project undertaken to investigate the use of art and design images in higher education graphic design courses. The History of Science Search Index was created in consultation with students on a related course and evaluated by academic staff.

The creation of these sets included consultation with experts in their respective fields and in art periods and movements, themes and categories. Despite the fact that the image metadata is consistent and uses the VRA Core 4.0 satisfactorily, as well as the ULAN artists' birth and death dates together with nationality, the realisation that keywords and classification should be in accordance with cataloguing standards adopted by the image library has been integral to this process. This paper discusses the ways in which the images were classified and retrieved in order to form the specific sets. In the future, more search indexes are planned to apply to a wide variety of fields to guide users through the image collections.

Image Validation in End-to-End Workflows

Peter Krogh, American Society of Media Photographers (USA)

Abstract: One of the most vexing issues in digital imaging is the challenge of image validation. Images may be corrupted at any point in the handling chain, from capture through transfer, image editing, storage and migration. Our research into image validation workflow has led to a lifecycle-based set of recommendations, used in conjunction with the DNG file format. These techniques can create an end-to-end workflow that is validated at each step, once an initial visual verification has been performed. This leads to both an increase in security, as well as a reduction in the resources needed to maintain the integrity of image files.

The traditional approach to data validation is to make a database of checksums of stored files, and to run a periodic validation sampling. While this approach does provide some good protection for static archived files, it is only appropriate for files that are completely static - any alteration to the file, such as image adjustment or the addition of embedded metadata produces a mismatch. Moreover, a traditional checksum approach typically relies on an external database of checksums, which creates a difficult workflow as files are transferred between different computers.

This paper provides an alternate methodology for a fully validated image file workflow, from initial image creation through to archive. It makes use of two tools to accomplish this: the Adobe DNG file format, as well as Parametric Image Editing (PIE) software. In a Parametric Image Editing environment, source image data is never modified, but instead is reinterpreted. This allows preservation of the original image, even as the image may be re-rendered according to different parameters. The Adobe DNG file format includes an area to store the source image data, rendering settings, metadata, as well as one or more fixed renderings of the image. One of the metadata fields that is part of the specification is an open source MD5 Checksum that refers only to the unchanging source image data.

Effective use of the DNG file, therefore, creates a portable validation key that can be assigned very early in the lifecycle, and travel with an individual file. The checksum remains viable even as a file is changed, or as the image is readjusted. Adobe has also released several free software tools that can check on the integrity of large collections of image file data automatically and reliably. These tools can, for instance, reliably identify a single changed bit in a file inside a multi-terabyte archive.

While the DNG can provide the majority of data validation needs for a digital image library, other validation tools are needed to fill in the gaps. Visual validation is still required at the start of workflow, and transfer validation should also be used regularly. When images must be converted to a rendered filetype, it becomes necessary to rely on the more traditional data validation tools. ▲

UPCOMING IS&T EVENTS

September 19–23, 2010; Austin, Texas
**NIP26: 26th International Conference on Digital Printing Technologies/
Digital Fabrication Processes 2010**
General Chairs: Gerhard Bartschar (NIP26) and Reinhard Baumann (DF2010)

November 8-12, 2010; San Antonio, Texas
Eighteenth Color Imaging Conference (CIC18)
General Chairs: Francisco Imai and Erno Langendijk

January 23-27, 2011; San Francisco Airport Hyatt Regency
Electronic Imaging 2011
Symposium Chairs: Sabine Süsstrunk and Majid Rabanni

May 16-19, 2011; Salt Lake City, Utah
Archiving 2011
General Co-chairs: Wayne Metcalfe, FamilySearch, and Kate Zwaard, US Government Printing Office

October 2–6, 2011; Minneapolis, Minnesota
**NIP27: 27th International Conference on Digital Printing Technologies/
Digital Fabrication Processes 2011**
General Chairs: Xavier Bruch (NIP27) and Shinri Sakai (DF2011)

To learn about all upcoming IS&T meetings, go to www.imaging.org/ist/Conferences/.
For a complete list of imaging-related meetings, go to
www.imaging.org/ist/conferences/events.cfm