Key Issues in the Creation, Delivery, and Preservation of Born Digital Images

Jessica Bushey; InterPARES2, University of British Columbia; Vancouver, BC, Canada.

Introduction

The digital epoch is in full swing. The anticipation of receiving photographic prints from the imaging lab days after a family reunion or commercial shoot is long gone. We have become a global culture of digital image creators and consumers that expect immediate satisfaction and settle for nothing less than the lightening transfer of digital data at all hours of the day and to every corner of the planet. With the touch of a button, teenagers send and receive images on cellular phones, art collectors select and purchase images from online galleries, and scholars interact with 3-d images of ancient artifacts deemed too fragile to touch. The proliferation of digital images will continue to increase as new devices are developed for capture and delivery, and venues are established (both virtual and physical) to share and exhibit 2-d and 3-d images.

As all aspects of society embrace digital imaging, the key issues facing both professional and amateur photographers are: 1) how to create digital images, 2) how to manage and disseminate their expanding collections of images, and 3) how to preserve images to ensure ongoing access to them in the future. The task of preserving access to born digital images that have been created and set-aside for future use (i.e., files that are saved and stored) is complicated by the fact that image files rely on a system of hardware and software to be read and represented. Technological obsolescence is caused by relentless versioning of file formats, software applications and operating platforms without addressing the need for backward or forward compatibility. This paper discusses the problems inherent in the digital medium and presents some of the issues raised by photographers working with digital technology.[1] The findings and recommendations presented in this paper are based on interdisciplinary research into the reliability and authenticity of born digital images.[2] Central to the research is the changing role of the creator (i.e., the photographer), who is becoming responsible for performing activities traditionally carried out by photofinishing and reprographic studios, print houses, art galleries, museums and archives. The recommendations made in this paper will assist individuals in safeguarding their valuable digital image collections for future use. It is important to understand that preservation involves a series of decisions and ongoing actions directed towards maintaining the readability and functionality of born digital images.

Definitions

Investigation into the concepts of reliability and authenticity in relation to photography and digital imagery reveals inconsistencies in the terminology. The transformation of photographic practice from analogue into digital presents an opportunity to revisit the language used to describe the actual images and the process of image creation.

For the purposes of this paper, reliability is viewed as the trustworthiness of an image as a statement of fact and refers to the accuracy of its content. The accuracy of content is determined by the methods employed in the creation of the image. The processing capabilities of digital cameras and image editing programs offer operations that surpass traditional darkroom techniques; therefore, the born digital image is often viewed as a fabrication of reality. Examination of the controls over the procedure of creation and the authority and competency of the persons involved in these activities determines the reliability of an image.

Authenticity refers to the fact that an image is what it purports to be and has not been tampered with or corrupted since it was set-aside. To a great extent digital images are regarded as less trustworthy than their analogue counterparts because of their inherent ability to be seamlessly manipulated and endlessly replicated without any detectable degradation in quality.

In many cases the only difference between the first instantiation of a digital image and the last is the timestamp. Whereas the film emulsion of a negative can be examined for proof of tampering and factual correlation with the photographic print can reveal manipulations introduced in the development process, this is not possible in the digital environment. To ensure authenticity the integrity and identity of a digital image must be established and maintained. Establishing the integrity and identity of a digital image requires additional contextual information to remain linked to the image file. The necessity of metadata (i.e., information about information) to capture and make explicit contextual information about born digital images will be explored further in this paper along with an analysis of existing metadata schemas being utilized by photographers.

Survey on the Record Keeping Practices of Photographers Using Digital Technology

Under the auspices of the International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2 Project: Experiential, Interactive and Dynamic Records, an international and interdisciplinary exploration into the issues related to the creation and maintenance of accurate and reliable records and the long-term preservation of authentic records in the context of artistic, scientific, and e-government activities, a web-based questionnaire entitled "Survey on the Record Keeping Practices of Photographers Using Digital Technology" was mounted online for a period of two months.[3] Professional photographers were invited to participate through postings to digital imaging list services, association websites and direct email to individuals. A total number of 402 photographers responded to the survey. Of these respondents, the majority were located throughout North America and Great Britain, and their industry affiliations included photojournalism, commercial studio, fine art, military and law enforcement, geology,

astronomy, and medical imaging. The defined objectives of the survey were to obtain qualitative and quantitative data on photographers' use and knowledge of digital technology. The survey was contextualized within the larger research goals of InterPARES 2, mainly the investigation of problems surrounding the reliability, authenticity, permanence, and accessibility of digital records. Data analysis was conducted on the basis of qualitative techniques such as tallying the responses to each multiple-choice question and expressing these numbers in percentages and examining the additional textual responses for categories and themes.

Research Questions

The survey addresses the following questions: (1) What kinds of digital images to photographers produce? (2) What are the assumptions of photographers about future access to their images? (3) What is the nature and variety of digital materials used by photographers?

Findings

The results of the survey are divided into two broad sections that address issues contributing to the reliability and authenticity of digital images. The first section explores actions and procedures that affect the reliability of a digital image, such as the choice of file formats and software applications utilized by photographers to create and manage their images. The second section explores actions and procedures that affect the authenticity of a digital image, such as security measures during transmission and the actions taken to ensure access to images once they have been saved and set-aside. The application of metadata to an image file at various stages throughout the process of creation, use, and preservation will be discussed in both sections.

Creation and Use

Photographers select image file formats that have the capacity to fulfill more than one creative objective. The majority of survey respondents re-purpose their digital images to serve a variety of functions that range from simple online viewing to the production of large-scale prints for gallery exhibitions. Although standardized file formats exist, such as the Joint Photographic Experts Group (JPEG) and the Tagged Image File Format (TIFF), many photographers risk the future usability of their digital images by selecting proprietary file formats, such as RAW, for initial capture. The "raw" digital image data captured by the camera's photosensitive detectors, using charge-coupled device (CCD) or complementary metal oxide semiconductor (CMOS) technology, is likened by survey participants to the in-camera analogue negative before chemical development. This raw camera information offers photographers a wider color gamut, larger file size, and greater bit depth than digital capture in the JPEG file format. Drawing upon this analogy, Adobe Systems Inc., released the Digital Negative (DNG) Specification, which describes an image file format for storing raw camera information. The DNG specification provides photographers with the option of converting their proprietary RAW file formats into the more "open" DNG specification.[4] However, it should be noted that DNG is not at this time an open standard. At present, the DNG format is gaining support within the imaging community as common raw format for preservation through collaborative efforts, such as the Universal

Photographic Digital Imaging Guidelines (UPDIG) Working Group.[5]

Nearly half of all survey respondents capture digital images in the JPEG file format. This is especially true for photographers using "point-and-shoot" digital cameras that limit in-camera capture to JPEG. After in-camera capture, the majority of photographers apply naming conventions and then copy their digital masters to compact disc (CD) or digital versatile disk (DVD) and create a digital surrogate to function as a working copy that will undergo alterations with image processing software. As indicated by survey responses, the most common file formats for digital surrogates are JPEG and TIFF.

Control over the in-camera capture process is important because it is at this stage that technical and descriptive metadata are attached to, and/or embedded within, the image. Technical metadata refer to the mechanical settings automatically recorded by the capture device. Technical metadata are used in determining how the image was constructed and the parameters for its digital representation. Descriptive metadata are the explanatory notes that define the contextual information supporting the digital image and are input manually by the photographer at the time of capture or later through image management software (IMS). Descriptive metadata are used to describe or identify images. Unlike analogue photographs, which can be physically browsed, digital images must be retrieved using a computer to view them. The act of retrieval creates the need for information about the images themselves to be captured and made recognizable. It is fundamental to the reliability of digital images to establish a protocol for metadata writers and readers to guarantee that metadata remain linked to the image throughout workflow processes regardless of the file format. Until a metadata protocol is established and supported, file format decisions made at the time of creation will greatly impact the longevity of an image.

Additional responses provided by survey participants identified that a few were familiar with the technical metadata captured by the Exchangeable Image File Format (Exif) for digital still images, a specification launched by the Japan Electronics Industry Technological Association (JEITA) in 1998. The Exif schema is essentially system metadata that are automatically generated by the capture device (i.e., camera and/or scanner), and provide information regarding basic image parameters, presentation, and the technical "how" of digital image creation. The mandatory metadata elements are static and exist as "read-only" data. Read-only metadata that are automatically generated by the hardware or software system are very reliable because they do not require manual input by the photographer. The fact that the schema is recognized as a standard by the imaging industry and written to JPEG (standard) and TIFF (de facto standard) offers a high degree of assurance to creators and preservers of digital images that the schema will be viable and operable in the future. On the whole, the application of the Exif schema for the delivery and protection of images in a collection is limited. Photographers utilize the schema as a source for technical reference (mimicking traditional darkroom notation), and not as a method for image retrieval and delivery.

Many of the survey respondents identified themselves as photojournalists and routinely transmit JPEG files and descriptive metadata to news service bureaus. The Information Interchange Model (IIM) released in 1991 by the International Press Telecommunications Council (IPTC) and the Newspaper Association of America (NAA) is the metadata schema used to

transfer a data object, which may be an image file or a combination of text and image, along with its pertinent editorial and technical information. In more recent years, a collaborative effort between IPTC, Adobe Systems Inc., and the International Digital Enterprise Alliance (IDEAlliance) resulted in the release of a new schema entitled IPTC Core, which is based on Adobe's technical Extensible Metadata Platform (XMP) framework, and read/ write compatible with any Extensible markup Language (XML) compliant device. The elements of the IPTC Core schema enable photographers to capture key information regarding the context surrounding the digital image such as, creator name, image location, and copyright notice. IPTC Core functions as administrative and descriptive metadata that document the content and context of a digital image and define the legal and regulatory parameters of its use. The survey findings reveal that over ninety-percent of photographers believe it is important that their images can be proven to be theirs and are properly credited to them. Additional comments made by respondents qualified the necessity of attaching copyright information to their digital image files to facilitate recognition and payment for their work.

The majority of photographers in the survey use commercial-off–the-shelf (COTS) software applications to display, edit, and manage their digital images collections. When asked about the type of information they record about their digital images, survey respondents selected descriptive information about who, when, where and why and descriptive information about the image itself to allow access and retrieval.

Preservation and Transmission

Most photographers in the survey indicated that they rely on their own knowledge and the recommendations they receive from others to shape their born digital image preservation procedures. The most common method they employ to preserve digital images and associated metadata for the long term is to copy the in-camera digital file and/or the final digital file to a CD or DVD. The issue of media fragility is a concern among photographers that rely on optical discs for long-term storage. In response, more than half of photographers surveyed take active measures to protect their digital image files form becoming obsolete, outdated, and irretrievable. Prevalent measures involve a combination of more than one method. Some of the methods implemented include selecting TIFF for archiving digital image files, making back-up files of digital images and copying them onto CD and DVD and refreshing optical storage media on a regular basis. Photographers are aware that the preservation of digital images involves monitoring and maintenance; yet the task of overseeing large accumulations is daunting and time-consuming.

Less than half of the photographers surveyed indicated that they apply security measures to protect their digital image files from access and accidental destruction. Among the minority that does apply security measures, most maintain their images offline, which is achieved by saving images onto CD, DVD, or an external drive. A handful of respondents store copies of their image collection off-site.

Most photographers are aware that transmitting their digital images outside of their personal computer space via email and online may lead to unauthorized use, but continue to do so anyway. Protecting images from unauthorized use is an ongoing challenge for photographers and Web masters. Of the small percentage of photographers that do protect their digital images, their methods include the addition of digital watermarks, metadata, and copyright statements to assert the ownership and usage rights associated with each image. When photographers mount their digital image collections onto the Web (which half of the survey respondents do), they manage access to these digital images using IMS with a dynamic publishing component. Photographers also provide access to their images through vendor management packages that provide point-of-purchase options and gallery browsing.

Conclusion

In general the survey findings indicate that most photographers are aware of the issues of media fragility and technological obsolescence; however, their decisions regarding creation, use, and preservation are made in direct response to business needs and artistic intentions, which lead to choices that may place the longevity of their digital images at risk. This situation is further compounded by the rapid introduction of alternative devices for capturing images digitally that do not support standardized file formats and metadata schemas utilized throughout the imaging industry. The majority of respondents indicated that they would follow a standard for digital image creation and file maintenance to ensure the longevity of their digital images if the standard was available and applicable to their practice. From the findings we may draw the following conclusions:

1. Photographers keep their digital images for re-use and reference, as demonstrated by their choice of in-camera file format and the fact that they keep multiple versions of a digital image. Photographers' preference for RAW file captures in-camera, combined with the nature of the technology industry in which innovation fuels an endless stream of new products and new versions increases the risk of obsolescence and data corruption.
2. Photographers are generally concerned with authenticity and reliability, as proven by the routine capture of metadata and routine preservation procedures; however, many of the measures taken by photographers to protect their digital image files from loss and corruption, such as refreshing older media and migrating obsolete file formats, alter the image file and, in effect, change it from what it was when it was first set-aside.
3. Photographers have begun to understand the challenges to continuing access and long-term preservation presented by the use of proprietary file formats and software applications.

Recommendations

As demonstrated, the longevity of authentic and reliable born digital images for future use rests on the decisions and actions of creators. The following recommendations support the creation, use and maintenance of born digital images in a manner that would allow for their long-term preservation. A more encompassing set of Guidelines will be made available on the InterPARES 2 Project Website in Spring 2007. The Guidelines have been developed for individuals and small organizations creating digital materials as a result of their professional and personal activities.

1. Select hardware and software components that are interoperable and provide a high degree of assurance that digital images will remain accessible over time. Interoperability refers to the ease in which devices and applications share digital information. As mentioned earlier, software should be backward and forward compatible to access older and future image file formats. When selecting a camera or an image management application, make certain that imaging industry standards for image file exchange are supported. Every change in a photographer's technology base should be undertaken with the understanding that interoperability among system components must be maintained to ensure accessibility. Therefore, documentation that identifies the system infrastructure including software specifications and functionality should be compiled to assist in preventing loss due to component failure and/or upgrades made to the system.

2. Ensure that digital image files are properly identified. Adherence to a standard metadata schema to describe the properties or attributes of a digital image is a critical step in conveying the identity of an image file. Essential information includes: name of the photographer, title and/or subject of the image, dates of creation and transmission, expression of documentary context (i.e., the broader group of files in which the image belongs, such as client name or job title), and indication of copyright. If multiple versions of an image file exist, such as drafts, online and/or print versions and finals, then a version indicator should be included as a standard file naming convention.

3. Changes made to digital image files should be documented. Currently available COTS IMS and existing image metadata schemas do not offer photographers a reliable method for attaching information regarding changes made to the content of digital image files. To demonstrate that the integrity of a digital image is intact and uncorrupted after it has been created and set-aside, photographers should attach documentation that explicitly states the person(s) responsible for managing the image file(s), any technical changes made to the image file(s), and the rules governing deletion or removal of image files from the system.

4. Organize digital images collections. The days of filing cabinets filled with color-coded binders and index lists of slide sheets may be over, but the importance of logical groupings should not be dismissed. Grouping image files that you want to save for future use according to a schema, such as client names, shoot locations or chronologically will assist in the efficient management and retrieval of digital files. Maintaining image files within logical groups also eliminates confusion regarding file destruction and retention.

5. Protect digital image files against unauthorized use. Access to computers and storage devices that contain digital image collections should be restricted. The accuracy and authenticity of an image file cannot be presumed if controls over the collection, such as user passwords and permission profiles, are not in place.

6. Protect digital image files from accidental loss and corruption. Backing up the image collection and the operating system should be routinely scheduled. Backups can be achieved manually, automatically initiated by software applications and provided by a third party service.

7. Photographers should make regular upgrades to operating systems and hardware and software components as a preventive measure against technological obsolescence. Structure technological systems to accept and adapt to changes in both hardware and software without requiring extensive reconfiguration. Commit to maintaining the digital image collection and its functionality, not the mechanical apparatus that houses it.

References

[1] M. Braun, and Jessica Bushey, "Survey of the Record-Keeping Practices of Photographers Using Digital Technology." InterPARES 2 (2004), http://www.interpares.org/gs07/login.cfm.

[2] Jessica Bushey, "Born Digital Images as Reliable and Authentic Records," Master's thesis, University of British Columbia, 2005.

[3] InterPARES 2.,"Project Summary." (2004), http://www.interpares.org/ip2/ip2_index.cfm

[4] Adobe Systems Inc., "Digital Negative (DNG)," Adobe Systems Inc. (2004), http://www.adobe.com/products/dng/license.html

[5] UPDIG Working Group, "Universal Photographic Digital Imaging Guidelines, Version 2.0" (2006), http://www.updig.org/guidelines/index.php

[6] IPTC, "Software supporting IPTC photo metadata standards IIM and 'IPTC Core'." (2007), http://www.iptc.org/photometadata/softwaresupportlist1.php

Author Biography

Jessica Bushey is a professional photographer and received her Master's Degree in Archival Studies from the School of Library, Archival and Information Studies at the University of British Columbia (2005). Jessica continues to work with the InterPARES 2 Project on issues related to the authenticity and reliability of born digital images. She is currently the Head of Digital Imaging at the Museum of Anthropology, UBC.