Conference Proceedings

Edited by: Luciana Duranti and **Elizabeth Shaffer**

The Memory of the World in the Digital Age: **Digitization and** Preservation

An international conference on permanent access to digital documentary heritage



Educational, Scientific and • 20th Anniversary Cultural Organization .

United Nations . Memory of the World

Hosted by:



a place of mind THE UNIVERSITY OF BRITISH COLUMBIA

In collaboration with



26 to 28 SEPTEMBER 2012

Vancouver, British Columbia, Canada Sheraton Vancouver Wall Centre





Memory of the World

20th Anniversary

United Nations Educational, Scientific and Cultural Organization





Organisation · Mémoire du monde des Nations Unies · 20^e anniversaire

.

pour l'éducation, .

la science et la culture

Conference Proceedings

Edited by: Luciana Duranti and Elizabeth Shaffer

The Memory of the World in the Digital Age: Digitization and Preservation

An international conference on permanent access to digital documentary heritage

26 to 28 SEPTEMBER 2012 Vancouver, British Columbia, Canada Sheraton Vancouver Wall Centre

UNESCO Memory of the World Programme, Knowledge Societies Division

This book of Proceedings includes most of the papers and posters presented at the International Conference "The Memory of the World in the Digital Age: Digitization and Preservation" held on 26-28 September 2012 in Vancouver, British Columbia, Canada, by the UNESCO Memory of the World Programme, Knowledge Societies Division, and The University of British Columbia in collaboration with the University of Toronto.

The proceedings have been compiled and formatted with minor editing; papers and posters appear as submitted. The authors are responsible for the choice and the presentation of the facts contained in this publication and for the opinions they express, which are not necessarily those of UNESCO and do not commit the Organization.

The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The organizers of this UNESCO Memory of the World Programme Conference would like to sincerely thank everyone who contributed to the Conference in Vancouver and to these proceedings.

Published by UNESCO 2013, with the financial support of the Social Sciences and Humanities Research Council of Canada | Conseil de recherches en sciences humaines du Canada (SSHRC) and the International Research on Permanent Authentic Records in Electronic Systems (InterPARES) Project.



Social Sciences and Humanities Conseil de recherches en Research Council of Canada

sciences humaines du Canada





Preface

This publication presents the proceedings of the international conference 'Memory of the World in the Digital Age: Digitization and Preservation' which was held in Vancouver, Canada, from 26 to 28 September 2012.

More than 500 experts and other interested persons from all regions of the world participated in this knowledge-sharing and policy-driving event to discuss and exchange opinions on how to protect the world's documentary heritage. Although this heritage is the record of knowledge, its physical carriers are extremely vulnerable and can easily disappear without a trace. Whether recorded on a clay tablet or an electronic tablet, our methods of sharing content and knowledge need to be protected.

It is impossible to exaggerate the importance of documentary heritage in our lives. It governs our actions whether these relate to creating the basis of mutual respect between different civilizations and communities or building knowledge societies. Documentary heritage provides the foundation of peace, our identity and knowledge.

UNESCO's interest in this subject matter is as fundamental as its constitution with its mandate to contribute to building peace through the spread of knowledge from improved access to printed and published materials. These core materials, our documentary heritage, have been preserved in archives, libraries and museums for generations.

But while measures needed to maintain access to print materials are globally understood, the newer challenges related to preserving digital information are not keeping pace with technological development. The need for dedicated hardware and software, associated with their rapid obsolescence, hamper our ability to keep invaluable content accessible. Unless timely migration to newer technologies, operating systems and software platforms is assured, we face the risk developing digital Alzheimer's.

UNESCO's expectation from this Conference was to obtain a better definition of our expected role, and our contribution to setting a global digital agenda. The UNESCO/UBC Vancouver Declaration sets out specific recommendations which we will be implementing and incorporating into our digital strategy. Likewise, we expect that our Member States, professional organizations and private sector bodies will also implement the recommendations addressed to them.

Only through collaborative strategic alliances can we overcome the major challenges threatening the preservation of digital information. We believe that the presentations featured in this publication provide the basis for a global commitment to preserving the memory of our world in this digital age.

Jānis Kārkliņš Assistant Director-General for Communication and Information

Preface	. 4
Opening Keynotes Kenneth Thibodeau Wrestling with Shape-Shifters: Perspectives on Preserving Memory in the Digital Age	15
Luciana Duranti Trust and Conflicting Rights in the Digital Environment	24
Anne Thurston Digitization and Preservation: Global Opportunities and Cultural Challenges	31
Intellectual Property Infrastructure Initiatives for Digital Heritage	
Heather Christenson and John P. Wilkin Intellectual Property Rights & the HathiTrust Collection	39
Elizabeth Townsend Gard The Durationator [®] Copyright Experiment	46
Kate Hennessy The Intangible and the Digital: Participatory Media Production and Local Cultural Property Rights Discourse	58
Preservation Infrastructures: Current Models and Potential Alternatives	
Ilaria Pescini and Walter Volpi An Example to Follow: An Infrastructure for Interoperability and Governance in the Tuscan Public System for Digital Preservation	70
Francis G. Mwangi The Road to Providing Access to Kenya's Information Heritage: Digitization project in the Kenya National Archives and Documentation Service (KNADS)	83
Jeremy York A Preservation Infrastructure Built to Last: Preservation, Community, and HathiTrust	92
Hrvoje Stančić, Arian Rajh and Ivor Milošević "Archiving-as-a-Service": Influence of Cloud Computing on the Archival Theory and Practice1(08
The CODATA Mission: Preserving Scientific Data for the Future	
Elizabeth Griffin and the CODATA DARTG Team Recovering the Forgettery of the World12	27
Patrick C. Caldwell Tide Gauge Data Rescue	34
Stephen Del Greco Environmental Data Through Time: Extending The Climate Record	50
Tracey P. Lauriault and D. R. Fraser Taylor The Map as a Fundamental Source in the Memory of the World	60

Preserving Tradition and Performing Arts in Digital Form	
Ravi Katikala, Kurt Madsen and Gilberto Mincaye Nenquimo Enqueri Life at the Edge of the Internet: Preserving the Digital Heritage of Indigenous Cultures	190
Lekoko Kenosi Digital Madness, Archival Theory and the Endangered Sound Archives of Radio Botswana	206
Jørgen Langdalen Editing Historical Music in the Age of Digitization	212
Lauren Sorensen and Tanisha Jones Developing and Implementing a Digital Video Repository for Legacy Dance Documentation: Dance Heritage Coalition's Secure Media Network	217
Beyond Access: Digitization to Preserve Culture	
Fernanda Maria Melo Alves, José António Moreiro González and José Manuel Matias Safeguarding of the Portuguese Language Documentary Heritage: The Lusophone Digital Library	229
Benoit Ferland et Tristan Müller	-
Le réseau francophone numérique	236
John Van Oudenaren The World Digital Library	246
Strategies for Building Digital Repositories	
Bronwen Sprout and Sarah Romkey A Persistent Digital Collections Strategy for UBC Library	257
Neil Grindley Building the Business Case for Digital Preservation	269
Kevin Bradley Requirements of a Remote Repository	278
Digital Forensics for the Preservation of Digital Heritage	
Wayne W. Liu Accountability for Archival Digital Curation in Preserving the Memory of the World	288
Christopher A. Lee and Kam Woods Automated Redaction of Private and Personal Data in Collections: Toward Responsible Stewardship of Digital Heritage	298
Corinne Rogers and Jeremy Leighton John Shared Perspectives, Common Challenges: A History of Digital Forensics & Ancestral Computing for Digital Heritage	314
Giving a Permanent Digital Voice to the Silenced	
Terry Reilly For the Children Taken: The Challenge to Truth Commissions in Building digital collections for research and long-term preservation	338

National Strategies as the Foundation of Togetherness
Andris Vilks and Uldis Zariņš National Planning as the Key for Successful Implementation of Digitization Strategies
Ivan Chew & Haliza Jailan Preserving the Crowdsourced Memories of a Nation: The Singapore Memory Project
Ernesto C. Bodê Digital Preservation Policy of The Chamber of Deputies: Methodology for its development
Web 2.0 Products as Documentary Digital Heritage: Can We Access and Preserve Them?
Jamie Schleser Unprotected Memory: User-Generated Content and the Unintentional Archive
Heather Ryckman Context 2.0: User Attitudes to the Reliability of Archival Context on the Web
Lisa P. Nathan and Elizabeth Shaffer Preserving Social Media: Opening a Multi-Disciplinary Dialogue
The Role of Culture in Digitization and Digital Preservation
Fiorella Foscarini, Gillian Oliver, Juan Ilerbaig and Kevin Krumrei Preservation Cultures: Developing a Framework for a Culturally Sensitive Digital Preservation Agenda
Tukul Sepania Walla Kaiku and Vicky Puipui Political, Cultural and Professional Challenges for Digitization and Preservation of Government Information in Papua New Guinea: An Overview
Xincai Wang and Yunxia Nie Current Situation, Problems and Prospects of the Digital Preservation of Documentary Heritage in China
Open Archival Information System Reference Model: Answer or Inspiration?
Stefano S. Cavaglieri Digital Archiving Systems Confronted with the OAIS Reference Model
Saeed Rezaei Sharifabadi, Mansour Tajdaran and Zohreh Rasouli A Model for Managing Digital Pictures of the National Archives of Iran: Based on the Open Archival Information System Reference Model
Collaboration in Digital Preservation or Lack Thereof: What Works
Maria Guercio Digital Preservation in Europe: Strategic Plans, Research Outputs and Future Implementation. The Weak Role of the Archival Institutions
Rolf Källman Models for National Collaboration: Coordination of the Digital Cultural Heritage in Sweden
Victoria Reich Building and Preserving Library Digital Collections Through Community Collaboration
Steve Knight National Library of New Zealand, Digital Preservation and the Role of UNESCO

The Economics of Preserving Digital Information	
David S. H. Rosenthal, Daniel C. Rosenthal, Ethan L. Miller, Ian F. Adams, Mark W. Storer and Erez Zadok	
The Economics of Long-Term Digital Storage	513
Ulla Bøgvad Kejser, Anders Bo Nielsen and Alex Thirifays Modelling the Costs of Preserving Digital Assets	529
L.M. Udaya Prasad Cabral Economically Easy Method to Digitize Oversized Documents with Special Reference to Ola Leaf Manuscripts in Sri Lanka	540
Patricia Liebetrau Preserving Our Heritage: An Independent Advantage	549
Is A New Legal Framework Required for Digital Preservation or Will Policy Do?	
Tony Sheppard Is a New Legal Framework Required for Digital Preservation or Will Policy Do? Building a Legal Framework to Facilitate Long-term Preservation of Digital Heritage: A Canadian Perspective	559
Alicia Barnard Development of Policies and Requirements for Ingesting and Preserving Digital Records Into a Preservation System: Where to start?	570
Jason R. Baron and Simon J. Attfield Where Light in Darkness Lies: Preservation, Access and Sensemaking Strategies for the Modern Digital Archive	580
Elaine Goh Strengthening the Regulatory Framework in a Digital Environment: A Review of Archives Legislation	596
Digital Curation: Convergence of Challenges, Institutions and Knowledge	
Sarah Higgins Digital Curation: The Challenge Driving Convergence across Memory Institutions	607
Jackie R. Esposito Digital Curation: Building an Environment for Success	624
Patricia Forget Célébrations institutionnelles : Événement catalyseur propice à l'implantation d'un projet de conservation du patrimoine numérique permettant de réunir les acteurs d'intérêts divergents	636
Jeannette A. Bastian and Ross Harvey The Convergence of Cultural Heritage: Practical Experiments and Lessons Learned	650
Digitization and Digital Preservation Experiences in a Developing Country Perspective	
Elizabeth F. Watson The Conservation and Preservation of Heritage in the Caribbean: What Challenges Does Diaitization Pose?	661
ga	501

Richard Marcoux, Laurent Richard and Mamadou Kani Konaté Digital Preservation of Demographic Heritage: Population Censuses and Experiences in Mali and the Democratic Republic of the Congo	672
Brandon Oswald Partnership in Paradise: The Importance of Collaboration for Handling Traditional Cultural Expression Material in the Pacific Islands	685
Ensuring That it Won't Happen Again	
Victoria L. Lemieux Financial Records and Their Discontents: Safeguarding the Records of our Financial Systems	700
Myron Groover The White House E-Mail Destruction Scandal of 2007: A Case Study for Digital Heritage	713
Kenneth Thibodeau The Perfect Archival Storm: The Transfer of Electronic Records from the G.W. Bush White House to the National Archives of the United States	724
Trusting Records	
Lorraine Dong The Ethical and Legal Issues of Historical Mental Health Records as Cultural Heritage	735
Marie Demoulin et Sébastien Soyez L'authenticité, de l'original papier à la copie numérique : Les enjeux juridiques et archivistiques de la numérisation	745
Web Archiving as Part of Building the Documentary Heritage of Our Time	
Liu Hua, Yang Menghui, Zhao Guojun and Feng Huiling Chinese Web Archiving and Statistical Analysis on Chinese Web Archives	765
Gustavo Urbano Navarro Implications of the Web Semantization on the Development of Digital Heritage	775
Matt Holden Preserving the Web Archive for Future Generations: Practical Experiments with Emulation and Migration Technologies	783
Technology as the Mediator of Heritage and Its Relations with People	
Ian S. King The Turtle At The Bottom: Reflections on Access and Preservation for Information Artefacts	797
Erik Borglund Challenges to Capture the Hybrid Heritage: When Activities Take Place in Both Digital and Non-Digital Environments	814
Limited Resources or Expertise: Case Studies in Addressing the Issue	
Jean Bosco Ntungirimana La problématique de la préservation de la mémoire collective au Burundi à l'ère des NTIC : Étude de cas menée à la Cour supreme	823
Farah Al-Sabah Digitizing A Survivor's Identity: The Past, Present, and Future of the Kuwait National Museum Archives	838

Wayne W. Torborg, Theresa M. Vann and Columba Stewart The Challenges of Manuscript Preservation in the Digital Age	851
Plenary 3 Keynotes	
Dietrich Schüller Challenges for the Preservation of Audiovisual Documents: A General Overview	863
International Perspectives and Cooperation	
Claudia Nicolai, Rachele Oriente and Fernando Serván One Year of Efforts for Digital Preservation at FAO	871
Peter Burnhill, Françoise Pelle, Pierre Godefroy, Fred Guy, Morag Macgregor and Adam Rusbridge Archiving the World's E-Journals: The Keepers Registry as Global Monitor	880
The World Audiovisual Memory: Practical Challenges, Theoretical Solutions?	
Jean Gagnon	
Treasures That Sleep: Film Archives in the Digital Era	892
Caroline Frick Seeing, Hearing, and Moving Heritage: Issues and Implications for the World's Audiovisual Memory in the Digital Age	896
Edoardo Ceccuti The Digitization of Films and Photos of the Istituto Luce	904
Adam Jansen Challenges and Triumphs: Preserving HD Video at the UBC School of Journalism	909
Mick Newnham, Trevor Carter, Greg Moss and Rod Butler Digital Disaster Recovery for Audiovisual Collections: Testing the Theory	921
Metadata and Formats for Digitization and Digital Preservation	
Joseph T. Tennis Data, Documents, and Memory: A Taxonomy of Sources in Relation to Digital Preservation and Authenticity Metadata	933
Adam Rabinowitz, Maria Esteva and Jessica Trelogan Ensuring a Future for the Past: Long-term Preservation Strategies for Digital Archaeological Data	941
Giovanni Michetti and Paola Manoni It FITS the Cultural Heritage! Formats for Preservation: From Spatial Data to Cultural Resources	955
Lois Enns and Gurp Badesha File Viewers: Examining On-the-Fly File Format Conversion	962
Walter Allasia, Fabrizio Falchi, Francesco Gallo and Carlo Meghini Autonomic Preservation of "Access Copies" of Digital Contents	976
A Methodology Framework to Ensure Preservation	
Anca Claudia Prodan Bias and Balance in the Preservation of Digital Heritage	989

	Giovanni Michetti Archives Are Not Trees: Hierarchical Representations in Digital Environment1	1002
	Göran Samuelsson The New Information Landscape: The Archivist and Architect – Drawing on a Common Map? 1	1011
	Shadrack Katuu Enterprise Content Management and Digital Curation Applications: Maturity Model Connections	1025
	Christopher J. Prom Facilitating the Aggregation of Dispersed Personal Archives: A Proposed Functional, Technical, and Business Model	1042
Di	gital Objects as Forensic Evidence	
	Carsten Rudolph and Nicolai Kuntze Constructing and Evaluating Digital Evidence for Processes	1057
	Aaron Alva, Scott David and Barbara Endicott-Popovsky Forensic Barriers: Legal Implications of Storing and Processing Information in the Cloud	1064
	Michael Losavio, Deborah Keeling and Michael Lemon Models in Collaborative and Distributed Digital Investigation: In the World of Ubiquitous Computing and Communication Systems	1079
	Fabio Marturana and Simone Tacconi Cloud Computing Implications to Digital Forensics: A New Methodology Proposal	1093
	Andrew F. Hay and Gilbert L. Peterson Acquiring OS X File Handles Through Forensic Memory Analysis1	1102
In	stitutional and Inter-Organizational Initiatives in Digitization	
	Anup Kumar Das Digitization of Documentary Heritage Collections in Indic Language: Comparative Study of Five Major Digital Library Initiatives in India1	1126
	Ronald Walker Digital Heritage Preservation - Economic Realities and Options1	1139
	S. K. Reilly Positioning Libraries in the Digital Preservation Landscape	1146
	Heidi Rosen, Torsten Johansson, Mikael Andersson and Henrik Johansson Experiences from Digidaily: Inter-Agency Mass Digitization of Newspapers in Sweden	1153
Pr	eserving Images: What Do We Need to Know?	
	Adama Aly Pam Chemins de la mémoire : Les archives audiovisuelles au secours de l'identité d'une organisation internationale africaine1	1163
	Krystyna K. Matusiak and Tamara K. Johnston Digitization as a Preservation Strategy: Saving and Sharing the American Geographical Society Library's Historic Nitrate Negative Images1	1173
	Jessica Bushey Born Digital Images: Creation to Preservation1	L189

Angelina Altobellis Essential Skills for Digital Preservation: Addressing the Training Needs of Staff in Small Heritage Institutions	1198
Small and Large Scale Digitization: Towards a Shared Conceptual Model	
Peter Botticelli, Patricia Montiel-Overall and Ann Clark Building Sustainable Digital Cultural Heritage Collections: Towards Best Practices for Small- scale Digital Projects	1205
Marco de Niet, Titia van der Werf and Vincent Wintermans Preserving Digital Heritage: The UNESCO Charter and Developments in the Netherlands	1219
Paul Conway Validating Quality in Large-Scale Digitization: Findings on the Distribution of Imaging Error	1233
Lars Björk Lost in Transit: The Informative Capacity of Digital Reproductions	1252
Preservation of Audiovisual Material	
Mike Casey The Media Preservation Initiative at Indiana University Bloomington	1266
George Blood Video CompressionFor Dummies?	1273
Pio Pellizzari, Álvaro Hegewich The Ibero-American Preservation Platform of Sound and Audiovisual Heritage	1289
Trusting Data and Documents Online	
Junbin Fang, Zoe Lin Jiang, Mengfei He, S.M. Yiu, Lucas C.K. Hui, K.P. Chow and Gang Zhou Investigating and Analysing the Web-based Contents on Chinese Shanzhai Mobile Phones	1297
Junwei Huang, Yinjie Chen,Zhen Ling, Kyungseok Choo and Xinwen Fu A Framework of Network Forensics and its Application of Locating Suspects in Wireless Crime Scene Investigation	1310
F.R. Van Staden and H.S. Venter Implementing Digital Forensic Readiness for Cloud Computing Using Performance Monitoring Tools	1329
Yongjie Cai and Ping Ji Security Monitoring for Wireless Network Forensics (SMoWF)	1340
Workshops	
Peter Van Garderen, P. Jordan, T. Hooten, C. Mumma and E. McLellan The Archivematica Project: Meeting Digital Continuity's Technical Challenges	1349
Hannes Kulovits, Christoph Becker and Andreas Rauber Roles and Responsibilities in Digital Preservation Decision Making: Towards Effective Governance	1360
Posters and Presentations	
Collence Takaingenhamo Chisita and Amos Bishi Challenges and Opportunities of Digitizing and Preserving Cultural Heritage in Zimbabwe	1382

Donna McRostie The long and winding road from aspiration to implementation – building an enterprise digitization capability at the University of Melbourne
Asger Svane-Knudsen and Jiří Vnouček Retrieving a part of Danish colonial history: From dust to digital copy
Mitra Samiee and Saeed Rezaei Sharifabadi A Paradigm for the preservation of national digital memory of Iran
Chinyere Otuonye, Tamunoibuomi F. Okajagu, Samuel O. Etatuvie, Emmanuel Orgah, Gift Eyemienbai, Luke Oyovwevotu, Ewoma Borgu, and Janet Ukoha Insights on the Digitization of Traditional Medicine Knowledge in Nigeria
Nader Naghshineh and Saeed Nezareh Crowd-sourced digital preservation: An Iranian model1397
Chris Muller Data at Risk: The Duty to Find, Rescue, Preserve1399
Natalia Grincheva Digital diplomacy: Providing access to cultural content, engaging audiences on a global scale 1401
Rusnah Johare Preserving digital research data1403
Claudia M. Wanderley Multilingualism at the University of Campinas1405
Anne Thurston Open government and trustworthy records1407
Jan Marontate, David Murphy, Megan Robertson, Nathan Clarkson and Maggie Chao Canada – Aural memories: A case study of soundscape archives
Na Cai, Leye Yao and Liu Liu Creating Social Memories of Major Events in China: A Case study of the 5•12 Wenchuan Earthquake Digital Archive
Addendum
Howard Besser Archiving Large Amounts of Individually-Created Digital Content: Lessons from Archiving the Occupy Movement
Nadja Wallaszkovits Digitisation of Small Sound Collections: Problems and Solutions
UNESCO/UBC Vancouver Declaration
The Memory of the World in the Digital Age: Digitization and Preservation

Sponsors

Born Digital Images

Creation to Preservation

Jessica Bushey

School of Library, Archival and Information Studies (SLAIS) The iSchool at University of British Columbia

Abstract

In the last decade professional photographers have adopted born digital photographic processes to fulfill business needs and realize creative endeavours. More recently, the proliferation of born digital images on social media websites as documentation of individual experiences and a primary tool for communication has highlighted the necessity for information professionals to understand the key factors affecting the reliability and authenticity of born digital images as records. Based on findings from the International Research on Permanent Authentic Records in Electronic Systems (InterPARES 2) "Survey on the Recordkeeping Practices of Photographers using Digital images such as image file formats, metadata, workflow procedures, and storage media. In an effort to support cultural institutions responsible for the acquisition, management and preservation of born digital images, this paper provides an overview of recent developments initiated by photographers, the imaging industry and cultural heritage organizations. These developments reflect the dynamic state of digital practice and reveal the importance of collaboration among photographers, software developers, hardware manufacturers, and cultural stakeholders in the creation and preservation of born digital images.

Author

Jessica Bushey is a doctoral student at the School of Library, Archival and Information Studies (SLAIS) The iSchool at University of British Columbia (UBC). Her research interests address the trustworthiness of born digital media and authorship in the online environment. Ms. Bushey is a Systems Analyst for Artefactual Systems Inc., the lead software developer for open-source archival applications ICA-AtoM and Archivematica. After receiving her MAS in 2005, Ms. Bushey developed and implemented the digitization program at the Museum of Anthropology at UBC. This four-year project involved born digital capture of 35,000 artefacts in order to increase public access to the collection, encourage scholarly research, and support preservation activities.

"It took 50 years for us to reach 14 million analog images in the Still Picture Unit, but we'll reach 10 million digital images in approximately 10 years."

> -- Billy Wade, Archivist Still Pictures Unit, U.S. National Archives¹

1. Introduction

In the last decade we have witnessed the transformation of photography from a film-based medium to pixel-based. The switch to digital began with professional photographers working in fields such as

¹ Billy Wade, "From Analog to Digital," *Media Matters: The Blog of the National Archives' Special Media Archives Services Division*, April 18, 2012, http://blogs.archives.gov/mediamatters/2012/04/18/from-analog-to-digital/.

medicine, journalism, geospatial, and law enforcement. Supported by new technologies for capturing, transmitting and storing born-digital images, the practice of digital photography grew quickly and soon spread beyond professional boundaries and throughout the general populace. In fact, recent advances in cell phone/camera technology have managed to couple quality with convenience and as a result, unprecedented numbers of born-digital images are being created and shared on a global scale.

Unlike their analogue predecessors, digital images require software and hardware for viewing, sharing and storing. The convenience of traditional photofinishing services that added numbers to slide frames, provided protective housing for film and prints, and offered organizational materials at point-ofpurchase are no-longer available. Contemporary photographers working with digital media must be selfreliant and perform the majority of management tasks with commercially available software (often a suite of applications) to the best of their abilities. Passive approaches to digital image management and preservation place valuable image collections at risk. With the exception of nitrate film, the storage requirements of early photographic materials provided generous margins for neglect; whereas, unnamed and forgotten born-digital images are unlikely to survive beyond the next upgrade. In light of emerging responsibilities for photographers to manage digital images and provide ongoing access to digital files it is necessary to re-articulate the role of the photographer as both author/creator and preserver.

This paper discusses the findings of the *InterPARES 2* general study entitled "Survey on the Record-Keeping Practices of Photographers using Digital Technology" in the context of industry initiatives aimed at the creation, management, and preservation of born-digital images.² An analysis of the InterPARES 2 findings will establish the framework in which photographers are approaching digital practice and provide the foundation for an exploratory discussion of recent studies and current practices within the archival community aimed at preserving born-digital cultural heritage. Of particular importance is an examination of metadata schema standards used by professional photographers and supported by software vendors, and storage providers (including cloud services), to determine their efficacy to ensure born-digital images as reliable and authentic records.

2. Terminology³

The term "born-digital image" refers to a digital image that never physically existed before becoming a digital file. The most common example is an image created with a digital camera. An existing photograph

² InterPARES 2 Project, "General Studies," accessed on August 31, 2012,

http://www.interpares.org/ip2/ip2_general_studies.cfm. A general overview and the final report of the GS-07 Survey on Recordkeeping Practices of Photographers Using Digital Technology, can be downloaded from the project site, see http://www.interpares.org/ip2/ip2_general_studies.cfm?study=29. The IP2 GS-07 Survey was undertaken by Marta Braun (Ryerson University) and research assistant Jessica Bushey (University of British Columbia) and conducted under the auspices of the InterPARES 2 Project. For a more in-depth analysis of the findings of the "Survey on the Record-Keeping Practices of Photographers Using Digital Technology" see Jessica Bushey, "Born Digital Images as Reliable and Authentic Records," MAS Thesis, University of British Columbia, Vancouver, 2005. An earlier version of the findings of the Survey was presented in the article, "He Shoots, He Stores: New Photographic Practice in the Digital Age," *Archivaria* 65 (Spring 2008): 125-149. For the purposes of the current paper, the discussion has been significantly revised and updated to reflect changes introduced by both the photographic and archival communities in the past five years and to address the context of online networks including social media sites.

³ Parts of this section originally appeared in Bushey, "He Shoots, He Stores," pp. 131-132.

or document that is scanned or digitally reproduced to create an image file is not considered to be born digital but to have been digitized.⁴

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. 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 ensure authenticity the integrity and identity of a digital image must be established and maintained. Establishing the integrity and identity of a digital image file.

Lastly, the relationship between a born-digital image and a record must be addressed. From its nascence, photography has been associated with the act of recording an event. In the context of this approach, the photograph is received as a visual account of something and an aid to memory. The process of naming, saving, and setting-aside a digital image file for long-term storage makes explicit the creator's intent to carry forward visual information about an event for future use and/or reference. Additional capture of technical and descriptive metadata about the born-digital image further supports its capacity to function as a record.⁵

3. Survey on the Record-Keeping Practices of Photographers Using Digital Technology

Under the auspices of InterPARES2, the "Survey on the Record-Keeping Practices of Photographers Using Digital Technology" was launched as a web-based questionnaire in the Fall 2004. The survey targeted photographers who were known to create digital images and use digital technology to manage and store their images. An invitation to participate was posted to professional online fora and photographic association web sites that foster a community of photographers using 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. The survey questions were formulated to gather information regarding the principles and procedures that contribute to the creation, use, and preservation of digital images as reliable and authentic records; however, the terms "reliability" and "authenticity" were omitted from the survey to avoid confusion resulting from individual and disciplinary interpretation of their meaning. 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.

⁴ The digitized image acts as a surrogate to the analogue original. A born-digital image refers to an image that is generated entirely from digital hardware and/or software, and which has no analogue genesis.

⁵ A record is a residue of activity retained by its creator for reference or use in later activity. See Luciana Duranti and Kenneth Thibodeau, "The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES," *Archival Science* 6 (2006): 13-68, p. 66.

4. Research Questions

The survey addresses the following questions: (1) What kinds of digital images do photographers produce? Of these digital images, which constitutes the "original?" (2) What are the assumptions of photographers about future access to their images? Additionally, what is the intention of photographers for the dissemination and presentation of their digital images? (3) What is the nature and variety of digital materials used by photographers? Specifically, what hardware and software do photographers use, and what methods or materials do they select for long-term storage?

5. Survey Findings⁶

Results of the survey are presented in two broad areas that reflect the method in which photographers approach digital image creation, use, and preservation. The first section addresses the actions and procedures that photographers use to create digital images, such as the selection of capture hardware and software, image file formats and their characteristics, and the automatic and manual addition of technical and administrative metadata. The choices made by photographers at this stage of their digital practice affect the reliability of the born-digital image. The second section addresses the steps taken by photographers to store and preserve their digital images, such as security measures, procedures for transmission and dissemination, selection of storage media, and the manual addition of administrative and preservation metadata. The choices made by photographers at this stage of their digital practice for transmission and dissemination, selection of storage media, and the manual addition of administrative and preservation metadata. The choices made by photographers at this stage of their digital practice affect the authenticity of the born-digital image.

Throughout both sections the critical role of metadata, and its essential contribution to establishing the reliability and proving the authenticity of a born-digital image, will be discussed because metadata are automatically generated and manually input throughout the life cycle of a record. This paper is concerned with metadata schemas that are accepted standards and currently available to photographers via the functionality of hardware and software products. Where applicable, developments instigated by the photographic community will be discussed, including changes in practice and adoption of new metadata schemas.

6. Creation and Use

The majority of survey respondents identified their practice as "completely digital" and provided additional comments that date their transition to digital as commencing in 1998. The ability to re-purpose digital images (i.e., re-format and share the same image to serve different creative and business needs) influences most photographers' choice of image file formats at the time of digital capture. Many professional photographers select proprietary formats, such as *RAW*, for initial capture because it offers the highest quality data with the most potential for re-purposing. The RAW format is exclusive to *Digital Single Lens Reflex (DSLR)* cameras marketed to professional photographers. In cases where photographers use RAW format to capture the scene digitally, the RAW file is equated with the 'original' image and treated as such throughout subsequent procedures for use and preservation. The fact that RAW image formats are proprietary, often differing for each successive camera model from the same

⁶ Parts of this section originally appeared in Bushey, "He Shoots, He Stores," pp. 133-141.

manufacturer, and that the practice of encryption to conceal the RAW specification from users is widespread, present an enormous risk to the future usability of born-digital images created and stored in the RAW format. In 2004 Adobe Systems Inc., launched the *Digital Negative (DNG)* format as an alternative RAW format aimed at supporting image preservation. The DNG is based on the *Tagged Image File Format (TIFF)* specification and provides photographers a file format that is self-contained (image and metadata intact) and cross-platform interoperable.⁷ In 2005 the OpenRAW initiative was launched by photographers to raise awareness of the risks posed by proprietary RAW formats to digital image preservation. In 2006 OpenRAW conducted a survey of over 19,000 photographers to gather data about professional practices and concerns regarding RAW formats. One of their key findings is the "increased probability that as time passes a RAW file will be unreadable or cannot be used to reproduce the photographer's original interpretation."⁸

Respondents to the survey identified *Joint Photographic Experts Group (JPEG)* as the most common file format for in-camera capture. Designated an ISO standard in 1994, JPEG has gained industry-wide support regardless of capture device (e.g., camera, cell phone, pads etc.). As an open standard, the JPEG specification is made available to the public, and it is cross-platform operable, which means that the image and its metadata should remain intact when they are transmitted across systems and software applications. The drawback to the format is its use of lossy compression, which enables the file to be transmitted quickly (making it ideal for email attachments and social media sites), but results in loss of information each time the file is saved. Generational loss due to re-compression results in degradation of the image quality, making the JPEG format insufficient for preservation purposes. Thus, the survey findings in response to research question (1) reveal that photographers produce images that fulfill their creative and business needs, while providing the opportunity for future re-use. The original born-digital image is equated with the in-camera capture format, which in the case of this survey, is either RAW or JPEG, formats that have known risks for preservation.

During the procedure of creation, technical and descriptive metadata are attached to and/or embedded within the digital image. Technical metadata refer to the settings that are automatically recorded by the capture device (i.e., camera), such as pixel width and height of the image, colour space, and image compression. Technical metadata are used in determining how the image is constructed and the parameters for its digital representation. Survey respondents' comments regarding the variety of information recorded about digital images, identified their knowledge and use of the *Exchangeable Image File Format (Exif)* for digital still image metadata, which is a specification that was launched by the *Japan Electronics Industry Technological Association (JEITA)* in 1998 and is now backed jointly with the *Camera & Imaging Products Association (CIPA)* to encourage interoperability between imaging devices.⁹ As long as hardware and software support the information model promoted by Exif, the technical metadata are properly exchanged and retained along with the digital image.

⁷ Adobe Systems Incorporated, "In depth: Digital Negative (DNG)," 2012,

http://www.adobe.com/ca/products/photoshop/extend.displayTab2.html. The Digital Negative (DNG) Specification v1.3, 2009.

http://wwwimages.adobe.com/www.adobe.com/content/dam/Adobe/en/products/photoshop/pdfs/dng_spec.pdf.

⁸ OpenRAW, "The Problem with Proprietary RAW files," 2006, http://www.openraw.org/info/index.html.

⁹ Standardization Committee, "Exchangeable image file format for digital still cameras: Exif Version 2.3," CIPA DC-008-2010/ JEITA CP-3451B, April 26, 2010, http://www.cipa.jp/english/hyoujunka/kikaku/pdf/DC-008-2010 E.pdf.

The drawbacks to the use of Exif metadata standard are that not all of its elements are mandatory, which means that not all devices or software are required to write and read the majority of Exif tags, and not all social media sites preserve image metadata after upload.¹⁰ The lack of universal write/read support for all Exif tags means that critical system metadata that can assist in uniquely identifying the born digital image or provide assurance of data integrity in the future may be stripped from the image during routine procedures over creation and use. Analysis of the Exif schema in terms of its record-keeping capabilities shows that it fails to provide information about digital image context, hierarchical information about relationships between images in an aggregation of images, and processing history about the image. The schema describes the technical aspects of the image itself and the capture device, but it does not give contextual information regarding external agents such as photographer name, or the business activities or management processes it supports.

The most common type of information that survey respondents record about their digital images is descriptive. The descriptive information identifies the context of image creation (i.e., who created the image, when and where it was taken, and why), and explains the content of the image (i.e., persons, locations, and subject matter represented in the image), for purposes of access and retrieval. Essentially, descriptive metadata are explanatory notes that photographers add to active images, with the aid of commercially available software, in order to identify the persons, actions, and matters related to image creation and use. The value of descriptive metadata is its capacity for establishing record identity; however, metadata must remain persistently linked and be managed along with the image to ensure maintenance of authenticity. Since 1991, the International Press Telecommunications Council (IPTC) has maintained a metadata standard to transfer a data object, which may be an image file or a combination of text and image, along with its pertinent information, such as creator's name, location, subject matter, and copyright/usage notice, between systems. In the past ten years IPTC has worked with Adobe Systems Inc., and professional photographers to devise an XML based metadata schema (IPTC Core) that offers photographers a reliable and convenient method of applying descriptive metadata to their digital images using templates made available in Adobe software products. In July 2010 IPTC released a comprehensive document entitled "Photo Metadata: Core 1.1 and Extension 1.1", which presents the IPTC Core specification and the more recent IPTC Extension specification-a combination of descriptive and rights based metadata.¹¹ Over ninety-percent of survey respondents believe it is important that their images can be proven to be theirs and are properly credited to them. The addition of metadata is one of the methods photographers use to protect their digital images. By implementing a standardized metadata profile via image management software a control is exercised over the procedure of creation and use, which greatly assists in establishing the identity and integrity of born-digital image files.

Unlike the Exif schema, the IPTC metadata schema is not read-only but has dynamic fields of information that may be changed throughout the lifecycle of the image, depending upon the management

¹¹IPTC, "IPTC Standard Photo Metadata: IPTC Core 1.1/ IPTC Extension 1.1, July 2010,

¹⁰ David Riecks, "Social media Photo Metadata use Survey," 2009-2012. Controlled Vocabulary website, https://spreadsheets.google.com/pub?key=tceeIYNw8ZDC0N52UgRcgnA&single=true&gid=0&output=html.

http://www.iptc.org/std/photometadata/specification/IPTC-PhotoMetadata-201007_1.pdf; The rights based metadata is structured according to the *Picture Licensing Universal System*, (*PLUS*) developed in 2006 for worldwide use. PLUS recognizes the importance of standardized image metadata to ensure the long-term preservation of image content and context across networked systems. The initiative works closely with *Digital Object Identifier (DOI)*, IPTC and Adobe Systems Inc., see The PLUS Coalition, "Picture Licensing Universal System," 2011, http://www.useplus.com/index.asp.

and use of the image. The schema provides photographers with a method of capturing attributes of the digital image's creation and use that contribute to uniquely identifying the image; therefore, it is more effective than Exif for the classification and retrieval of digital images in collections. The IPTC metadata schema does not have the capacity to present hierarchical levels and relationships within an aggregate, or an image processing history that would provide a greater understanding of the changes made to an image throughout its lifecycle. Thus, in response to research question (2) the survey findings reveal that photographers are actively engaged with capturing technical and descriptive metadata in an effort to ensure the identity and integrity of their images. The majority of respondents intend to have their images accurately displayed and credited.

7. Preservation and Transmission

The majority of survey respondents is concerned with the longevity of their digital images and incorporates a procedure for long-term storage into their workflow. Image preservation activities include (in order of frequency): selecting storage media (i.e., CD-R, DVD-R, and external drives), designating file format for originals and surrogates (i.e., RAW, TIFF and JPEG), unique file naming that identifies relationships between originals and surrogates, and using software and capture hardware with specific attributes (i.e., batch metadata capabilities and cataloguing offline CDs.) Survey respondents described procedures for transferring and/or copying in-camera images (i.e., originals) to CDs and DVDs immediately following a shoot and then creating digital surrogates to function as working files which undergo edit operations. Comments made by respondents express concern with the longevity of optical storage and seek advice regarding the best brand of "archival" CD and DVD.

The survey findings show that the measures photographers currently take to protect their image files involve making back-ups and refreshing optical storage media by making 'read-only' copies of CDs and DVDs on a regular basis. No mention was made by respondents of providing file verification after transfer processes.¹² The practice of migrating older image file formats is less common. Protecting digital images from loss and corruption due to technological obsolescence and media fragility is only one part of a preservation strategy. Activities aimed at protecting digital images from unauthorized access and destruction assist in ensuring integrity. The survey findings regarding security measures to protect digital images held within systems and stored on removable storage media show that less than half of respondents apply any type of security measures. The transmission of born digital images outside the personal workspace to facilitate client review, assignment submission, and personal promotion present an opportunity for unauthorized access. More than half of survey respondents present their images on the World Wide Web, and of that group, the majority manages access using a custom built database or a vendor management package. In 2009 a study into the preservation of photo metadata by social media websites was initiated to determine the degree of image metadata support provided by social media sites

¹² In early 2010, tools for file verification using MD5 checksums started being discussed by members of the photographic community interested in digital image preservation. Reasons for file verification include ensuring data integrity of transfers and integrity of stored collections on removable media. David Riecks, "The Trouble Transporting Tribbles (or File Verification using MD5 Checksums)," July 2010,

and services including Facebook, Flickr, Twitter and Blogger.¹³ Preliminary results of the study reveal inconsistencies in the way image metadata are handled; however, routine tasks such as uploading and resizing digital images were found to remove both Exif and IPTC metadata from the digital image. Thus, in response to research questions (2) and (3) the survey findings reveal that photographers use a variety of commercial software and storage media to manage and preserve their digital image collections. Survey respondents are willing to use products that adhere to digital preservation standards; yet, many photographers shape their digital practice to meet the growing demands of clients for faster turnaround times and more versatile images, which inevitably results in adopting new technologies.

8. Future Directions

The Survey provided a unique opportunity to engage with a specific "creator community" at a key point in their transition from analogue to digital practice. As a result, we gained valuable insight into procedures conducted by creators throughout the creation, use and preservation of born-digital images, which led us to re-evaluate the role and responsibilities of the creator in the digital environment. Ongoing efforts by the photographic community and imaging industry to increase awareness and support for open formats, standardized image metadata and interoperability across networks contribute to the long-term preservation of born-digital images as reliable and authentic records.

The survey findings are relevant to archivists and cultural heritage professionals that are responsible for preserving and creating access to born-digital images. As individual photographers, cultural organizations and government agencies continue to embrace digital practices, the result will be an increase in born-digital accessions to archival repositories. In the past three years, the United States National Archives has accessioned over 700,000 digital images (in addition to analogue photography) and of that amount, approximately 100,000 of them are already available through their Online Public Access (OPA) system.¹⁴ There is evidence of the growing number of born-digital archives everywhere, for example the City of Vancouver hosted the 2010 Olympic and Paralympic Winter Games and as a result, the Vancouver City Archives acquired twenty-five terabytes (25 TB) of born-digital materials including videos, images and office files.¹⁵ In the traditional paper-based environment archivists often had years to accession (i.e., the physical and legal act of transferring materials from a donor to an archival repository) and process (i.e, the physical and intellectual activities of arranging and describing) the archival materials; however, the fragile nature of digital media and the expectations of scholars to have immediate access to digital archives combine to place immense pressure on repositories. As archives acquire born-digital and hybrid collections many of the challenges to preserve these materials and make them available to the public provide an opportunity to re-visit archival theory and practice.

Since the work begun by InterPARES, archivists and cultural heritage professionals have conducted research and executed pilot projects aimed at understanding the complexity of born-digital collections. Recent publications including *Digital Forensics and Born-Digital Content in Cultural Heritage*

¹³ David Riecks, "The Controlled Vocabulary Survey regarding the Preservation of Photo Metadata by Social Media Websites," 2009-2012. Controlled Vocabulary website,

http://www.controlledvocabulary.com/socialmedia/index.html.

¹⁴ Wade, "From Analog to Digital."

¹⁵ Courtney C. Mumma, Glenn Dingwall, and Sue Bigelow, "A First Look at the Acquisition and Appraisal of the 2010 Olympic and Paralympic Winter Gamers Fonds: or, SELECT * FROM VANOC_Records AS Archives WHERE Value='true'," *Archivaria* 72 (Fall 2011): 93-122.

Collections and findings of the *Born Digital Collections: An Inter-Institutional Model for Stewardship (AIMS)* project (2009-2011) present the preservation community with valuable case studies conducted in a variety of organizational and institutional settings, which significantly extend the research products created by earlier studies such as the *Personal Archives Accessible in Digital Media (Paradigm)* project (2005-2007).¹⁶ Oddly absent from these studies is a direct discussion about born-digital image collections and the role of photo metadata to support archival activities including acquisition, arrangement and description, preservation and access.¹⁷ As demonstrated in this paper, photo metadata contribute to the identity and integrity of born-digital images; therefore, a better understanding of self-describing digital images within the context of archival practice could contribute to current projects being conducted within the archival community on visual literacy, online discovery tools, digital records forensics, and digital preservation systems.

¹⁶ Mathew Kirschenbaum, Richard Ovenden, and Gabriela Redwine, "Digital Forensics and Born-Digital Content in Cultural Heritage Collections," *Council on Library and Information Resources*, 2010,

http://www.clir.org/pubs/abstract/reports/pub149; AIMS, "AIMS-Born Digital Collections: An Inter-Institutional Model for Stewardship," January 2012, http://www2.lib.virginia.edu/aims/whitepaper/AIMS_final.pdf; Paradigm, "Project Overview," 2008, http://www.paradigm.ac.uk/about/index.html. Also see Susan Thomas, "A Practical Approach to Preservation of Personal Digital Archives," Final Report to the Joint Information Systems Committee, March 2007, v1.0, *Paradigm Project*, http://www.paradigm.ac.uk/projectdocs/jiscreports/index.html.

¹⁷ The Digital Images Archiving Study, published in 2006 conducted by the Arts and Humanities Data Service (AHDS) and funded by the Joint Information Systems Committee (JISC) of the Higher and Further Education Funding Councils and the Arts and Humanities Research Council explored digital image content, user expectations and a life-cycle model for preservation. In the final report, the role of metadata (technical, administrative and discovery) is highlighted as an area for further research and development, especially in light of the different metadata standards supported by the digital photography industry (i.e., photojournalists, professional imaging software and hardware, online photo sharing networks) and the standards supported by cultural heritage institutions for the purposes of management and preservation.

Gold Sponsors



Silver Sponsors



Ministry of Education, Culture and Science



UNIVERSITY OF ALBERTA



Library

Supporters



CANADIAN COMMISSION FOR UNESCO COMMISSION CANADIENNE POUR L'UNESCO WWW.UIRESCO.CB











École des sciences de l'information School of Information Studies