

TP  
242  
A88  
2010

**PHOTOBLOGS:  
THE IMMEDIATE NEED TO DOCUMENT EARLY  
TWENTY-FIRST CENTURY PHOTOGRAPHY PRACTICES**

by

**Stephanie Atwood  
Bachelor of Fine Arts, The School of the Art Institute of Chicago  
Chicago, Illinois, United States of America, May 2007**

**A thesis  
presented to Ryerson University  
in conjunction with  
George Eastman House International Museum of Photography and Film**

**In partial fulfillment of the  
requirements for the degree of  
Master of Arts  
in the program of  
Photographic Preservation and Collections Management**

**Toronto, Ontario, Canada, 2010**

**© Stephanie Atwood 2010**

I hereby declare that I am the sole author of this thesis.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

## **ABSTRACT**

This thesis attempts to suggest ways in which museums might better understand and make informed decisions about acquiring, preserving, and cataloguing photoblogs, which are an early twenty-first century photography practice. Photographers can now use the World Wide Web to show and share their images, because of the advent of digital cameras, camera phones, and cheap, open-source photo-blogging tools available to the general population. This thesis will help museums to better understand and be comfortable in acquiring digital artefacts, such as photoblogs, that will enrich their photographic collections for future generations. Acquisition tools and preservation methods are defined and discussed. The process of cataloguing photoblogs in current collections-management databases is not much different from cataloguing hard-copy photographs. The “People of Walmart” photoblog is used as an example and an illustration to clearly define the difficult technical jargon separating curatorial and collections management departments from information technology departments.

## **ACKNOWLEDGEMENTS**

I would like to say thank you to my wonderful first reader, Trina Grover, who helped and communicated with me weekly by email and Gmail Chat, and who shared her interest in photoblogs and cataloguing. A thank you to my second reader, Marta Braun, who helped me to finalize this thesis. I will never forget my consultant and newly developed friend (BFF), Ryan Donahue, who patiently explained and creatively demonstrated the complex world of the World Wide Web. To my friends and PPCM colleagues: Mallory Taylor, Loreto Pinochet, Emily McKibbon, Selina Lamberti, Jenn DiCocco, Tasha Lutek, and Heather Pridemore, thanks for your unconditional help, tears, and laughter throughout this program. Thanks to my long-time friend, Laura Bowe, who pushed me to go past the sky. Most of all, I want to say thank you to my parents, Dr. Bruce and Ana María Atwood and my older sister, Irene, for all your support in my endeavours and dreams and madness.

*Begin at the beginning.*  
— Salvador Dalí

## TABLE OF CONTENTS

|   |     |
|---|-----|
| ABSTRACT.....   | iii |
| ACKNOWLEDGEMENTS.....   | iv  |
| LIST OF FIGURES.....  | vii |
| SECTION 1: Introduction.....  | 1   |
| Why Museums Should Acquire, Preserve, and Catalogue Photoblogs..... | 1   |
| SECTION 2: Literature Survey.....                                   | 8   |
| SECTION 3: Research Methodology.....                                | 11  |
| SECTION 4: Photoblogs .....   | 15  |
| Photographically Oriented Web Sites.....                            | 15  |
| Components of a Photoblog.....                                      | 16  |
| The Components of “PeopleOfWalmart.com”.....                        | 18  |
| SECTION 5: Collection Development—What and How to Acquire.....      | 20  |
| Copyright .....   | 21  |
| The Tools: Acquiring a Photoblog .....                              | 21  |
| The Capture: Screenshot .....                                       | 23  |
| The Copy: Saving the Source Code.....                               | 24  |
| The Extraction: Web Scraping .....                                  | 25  |
| The Aggregator: RSS Archiving .....                                 | 26  |
| The Physical Acquisition: Software Archive .....                    | 27  |
| Overview: Copy Versus Scraping Versus Software Archive.....         | 28  |
| SECTION 6: Further Preservation of Acquired Photoblogs.....         | 32  |
| Data Migration .....  | 32  |
| Virtual Machine and Emulation .....                                 | 33  |
| SECTION 7: Creating Catalogue Records for Photoblogs.....           | 36  |
| Levels of Granularity and Fields Included in Descriptions .....     | 37  |
| Accession Numbers .....   | 38  |
| Classification .....  | 39  |
| Date .....  | 40  |
| Constituent(s).....   | 40  |
| Title(s).....   | 41  |
| Medium .....  | 41  |
| Dimensions .....  | 42  |
| Description .....   | 42  |
| Frame Case Mount.....   | 43  |
| Inscription(s).....   | 44  |
| Overview of Cataloguing.....  | 44  |
| SECTION 8: Limitations and Suggested Further Research .....         | 45  |
| Conclusion.....   | 46  |
| REFERENCES.....   | 48  |

## LIST OF FIGURES

|  |    |
|--|----|
| Figure 1. The screenshot.....                              | 23 |
| Figure 2. Saving the source code .....                     | 24 |
| Figure 3. Web scraping.....                                | 25 |
| Figure 4. RSS archiving.....                               | 26 |
| Figure 5. Software archive.....                            | 27 |
| Figure 6. Example of a fabricated photoblog.....           | 28 |
| Figure 7. Example of a browser's script for figure 6 ..... | 29 |
| Figure 8. Example of a server's script for figure 6 .....  | 30 |
| Figure 9. Last page on POW photoblog .....                 | 40 |
| Figure 10. Title on header of browser .....                | 41 |

## **SECTION 1:**

### **INTRODUCTION**

#### **Why Museums Should Acquire, Preserve, and Catalogue Photoblogs**

The landscape of photography is changing. Ever since Daguerre and Talbot showed that it was possible to capture an image of nature on metal and paper with light-sensitive salts, photography has been part of many people's daily lives. With the advent of digital image-capturing and technological improvements on the World Wide Web in the 1990s, photography has continued as a medium for documentation, creativity, information, evidence, reportage, illustration, and novelty. Originally, photographic images were viewed on surfaces of metal or paper, but in the twenty-first century the medium is being viewed on a different platform: the electronic screen. Images are created and viewable by coding and pixels, and lack the chemical makeup and tactile feeling of the photograph. With the evolution of the dial-up modem, and wireless and cellular technology, images can now be shared, viewed, and recaptured through electronic devices (e.g., computers, smartphones, and other personal devices with Internet access). These technological advancements and the increase of consumer-friendly content management systems to build Web sites, blogging services, and image-hosting sites make it possible to store, share, and discuss photographic images on the Internet. Photoblogs are one such kind of Web site; they are growing in numbers and evolving in style.

In 2007, Internet Archive's Wayback Machine (an online public digital library of archived Web sites) crawled and captured 2 billion Web pages; currently, they are at



150 billion pages.<sup>1</sup> Photoblogs are not indexed separately in these 150 billion pages, because Internet Archive can only store text, but the keyword search *photoblog* yielded millions of results on major Internet search engines Google, Bing, and Yahoo! at approximately 2:00 p.m. on June 14, 2010. Google's Blog Search<sup>2</sup> alone retrieved 1,569,893 links using the keyword *photoblog* and Blogpluse.com identified over 140 million blogs on the Web.<sup>3</sup> Content management systems and blogging services such as WordPress, Google's Blogger, Photoblogs.org, and Pixelpost.org make it possible for the average consumer to build their unique photoblogs. Image-sharing sites Flickr and Photobucket provide users with the option of storing and sharing their digital images online. Together, their users upload over 13 million images daily.<sup>4</sup> One of the first major blogging services, Livejournal,<sup>5</sup> and the newest blogging Web site, Tumblr,<sup>6</sup> have over 30 million registered users. Cheezburger.com, a revenue-generating site, owns 53 photoblogs with different themes.<sup>7</sup> News photoblogs Boston.com's The Big Picture and MSNBC.com's PHOTOBLOG have posted thousands of images produced by their own photojournalists and those from Associated Press (AP), Getty Images, and Reuters.

---

<sup>1</sup> Internet Archive, "Wayback Machine," <http://www.archive.org> (accessed May 2010).

<sup>2</sup> Google Blogs, "Blog Search," <http://blogsearch.google.com> (accessed May 2010).

<sup>3</sup> Blogpluse.com, "BlogpluseStats," <http://www.blogpluse.com> (accessed June 2010).

<sup>4</sup> On June 14, 2010 at 2:08 pm, Flickr had 5,831 uploads in the previous minute and Photobucket stated that 4.7 million images were being uploaded every day.

<sup>5</sup> LiveJournal, referred as LJ, was begun in March 1999 by Brad Fitzpatrick. <http://www.livejournal.com/support/faqbrowse.bml?faqid=4&view=full> (accessed June 2010).

<sup>6</sup> CEO David Karp founded Tumblr in 2007. Wikipedia, "Tumblr," <http://en.wikipedia.org/wiki/Tumblr> (accessed June 2010).

<sup>7</sup> Jenna Wortham, "Once Just a Site with Funny Cat Pictures, and Now a Web Empire," *The New York Times*, <http://www.nytimes.com/2010/06/14/technology/internet/14burger.html?scp=1&sq=i+can+has+cheezburger&st=cse> (accessed June 13, 2010).

A large portion of those born in the past 30 years use digital photography, as opposed to analogue,<sup>8</sup> and many older people are also transiting from analogue to digital photography. Both young and old find that digital photography is flexible, easy to use, quick, and cost effective. With analogue photography, people had to purchase film and learn how to load and remove it; then they had to take it to a photo developer, where they would purchase the prints they desired; finally, they would need to organize the prints in albums, boxes, and/or frames. With digital photography, the only requirements are to acquire a digital camera with a memory card and a computer with software to upload the images. An advantage of using digital photography is that the images are measured by file size and do not take up any physical space—the only space required for them is the computer's disk storage capacity and on the blog's provided bandwidth. Digital photographers can produce as many images as their memory card allows, with no need to purchase, load, and develop film. Then they can quickly upload their images onto their computer, delete the images from the camera to free up the camera's memory card, and continue to take more photographs. Many digital photographers post their digital images, publicly sharing them with strangers on their blogging Web site or photo-sharing site; others choose private means to share photographs with family and friends.

The number of photoblogs and the sheer volume of photographic images on the Web can easily surpass the number of existing photographic objects in museum

---

<sup>8</sup> John Palfrey and Urs Gasser's *Born Digital* (New York: Basic Books, 2008) describes the young generation who were born after 1980 as *digital natives*. They have used technology since their birth. Palfrey uses the term *digital settlers* to refer a group of older generations that were born in the analogue world, but became part of the digital world later in age. Palfrey calls those who were born and have lived heavily in the analogue world *digital immigrants*.

collections. Not only are photoblogs growing in numbers; like Web sites in general, they are also ephemeral. Studies at the Library of Congress suggest that the average lifespan of a Web site is 44 days.<sup>9</sup> Not only that, the Library of Congress points out, “44 percent of the Web sites found in 1998 could not be found in 1999.” We are confronted daily with small and large changes in the favourite Web sites we visit. Online newspapers have new daily headlines, and blogging Web sites contain new posts and comments. Advertisements on Web sites might change when the Web page is refreshed. Some of us encounter Web sites that contain a 404-error.<sup>10</sup>

Wu, Heok, and Tamsir describe Web sites as “evidence of the cultural activities of contemporary society,”<sup>11</sup> and argue that they should be collected and cared for in a manner that preserves the contextual evidence of the content. This idea can be extended to include photoblogs; as a contemporary practice of photography, they are also evidence of cultural and social activities like their Web site counterparts. Photoblogs should be archived by institutions that have a mandate to collect photography. The flexibility of digital cameras and camera phones has helped more people to take pictures of everything—from mundane subjects to artistic subjects—and then post them online. The sheer volume of images and the ephemeral, volatile nature of the Web mean that curators, archivists, and historians of culture and society risk losing a significant portion of early twenty-first century photography that is on the Web. One of the largest

---

<sup>9</sup> The Library of Congress, “Web Capture and Archiving,” <http://www.loc.gov/acq/devpol/Webarchive.html> (accessed 26 May 2010)

<sup>10</sup> Sometimes known as “domain unknown.” 404 Error Pages.com, “Welcome to 404 Error Pages.Com,” <http://www.404errorpages.com/> (accessed June 13, 2010).

<sup>11</sup> P. H. J. Wu, A. K. H. Heok, and I. P. Tamsir, “Annotating Web Archives—Structure, Provenance, and Context through Archival Cataloguing,” *New Review of Hypermedia and Multimedia* 13, no. 1 (2007): 55–75, <http://dx.doi.org/10.1080/13614560701423620> (accessed May 2010), p. 55.

photographic collections, the George Eastman House International Museum of Photography and Film (GEH), has as part of its mission to “keep and care of images, literature, and technology to tell the story of photography and motion picture in history and in culture.”<sup>12</sup> GEH and similar museums with photographic collections might well look into adding photoblogs in their collections, but how will they accomplish that?

“Minerva,” the U.S. Library of Congress’s project for Web archives, has archived United States election Web sites, 9/11 Web sites, the “Law Library Legal Blawgs Web Archive,” and many others.<sup>13</sup> The University of North Texas’s “CyberCemetery” has archived expired government Web sites<sup>14</sup> and the British Library’s “UK Web Archive” is preserving United Kingdom Web sites.<sup>15</sup> Many institutions, libraries, and archives are actively documenting and preserving Web sites, but they are choosing specific sites according to their mission and policies. As a result, there are millions of other Web sites and photoblogs with the potential to enrich a museum’s photographic collection and future exhibition opportunities that may be missed. While it is labour-intensive to document, index, and catalogue every photoblog on the World Wide Web, if museums step in, they will have an opportunity to preserve a slice of time in the ongoing history of photography.

---

<sup>12</sup> George Eastman House, “Our Mission,” <http://www.eastmanhouse.org/Main/museum/mission.php> (accessed May 2010), para. 4.

<sup>13</sup> The Library of Congress, “Library of Congress Web Archives *Minerva*,” <http://lcweb2.loc.gov/diglib/lcwa/html/lcwa-home.html> (accessed 10 June 2010).

<sup>14</sup> University of North Texas Libraries, and U.S. Government Printing Office, “CyberCemetery Home,” <http://govinfo.library.unt.edu/> (accessed June 10, 2010).

<sup>15</sup> UK Web Archive, “Welcome to the UK Web Archive,” <http://www.Webarchive.org.uk/ukwa/> (accessed June 10, 2010).

There are museums working with Web sites and the Internet, including the Solomon R. Guggenheim Museum in New York. The Guggenheim Museum is one of the partners with Berkeley Art Museum and Pacific Film Archive's "Archiving the Avant-Garde," which is a consortium project for preserving variable media art. These two institutions are focusing on the uniqueness of artists' Web works and other digital media. The Guggenheim Museum commissioned Mark Napier to produce *Net Flag* in 2001. The artwork is a Java applet that is accessible through the Internet.<sup>16</sup> Napier's work is part of the Guggenheim Museum's Variable Media Network initiative, which explores and attempts to "develop the tools, methods and standards needed to implement"<sup>17</sup> its strategy of preserving the museum's collection of digital media. The museum is exploring strategies for preserving the original *Net Flag* and other digital media since computer platforms, architectures, and operating systems evolve, and software and applications upgrade over time. The museum is focusing on methods of preserving the original software, interaction, resolution, browser display, and exhibition.

Photoblogs are susceptible to change from evolving technology and obsolescence, as are artists' variable media art; the process and motivation for preserving photoblogs should not be any different from variable media art. This thesis is a discussion point for museums that may not be familiar with handling, managing, documenting, preserving, and describing photoblogs. It encourages further discussion between museum curatorial staff, archivists, cataloguers, exhibition designers,

---

<sup>16</sup> See The Variable Media Network, "Variable Media Case Study: Mark Napier," [http://www.variablemedia.net/e/case\\_napie\\_netfl.html](http://www.variablemedia.net/e/case_napie_netfl.html) (accessed June 10, 2010). See also Mark Napier, "Net.Flag—a Flag for the Internet," <http://netflag.guggenheim.org/netflag/> (accessed June 10, 2010).

<sup>17</sup> The Variable Media Network, <http://www.variablemedia.net/e/welcome.html> (accessed June 10, 2010), para. 1.

registrars, and information technologists. With artist-commissioned new media works, a museum is secured as a first step and the work is automatically added to the museum's collection. Photoblogs are not created in this fashion, and thus present unique challenges in acquiring and preserving them; this will be discussed in sections 5 and 6, along with suggestions for change. After acquisition, records and descriptions need to be indexed in order to improve access and cross-referencing; section 7 discusses cataloguing possibilities for photoblogs.

## SECTION 2:

### LITERATURE SURVEY

The literature on digital and Web-based media is vast, now containing over twenty years of journals, articles, papers, and conference proceedings. It contains hypothetical approaches to practical projects completed on the acquisition, preservation, and cataloguing<sup>1</sup> of digital media. However, the literature on photoblogging is limited; the majority of references focus on social aspects for researchers in the fields of sociology and technology.<sup>2</sup> Regardless, photoblogs *are* digital media; thus, findings and suggestions presented in the literature for institutions and the Web are relevant and provide valuable insights.

---

<sup>1</sup> E.g., Howard Besser, "The Transformation of the Museum and the Way It's Perceived," <http://besser.tsoa.nyu.edu/howard/papers/garmil-transform.html> (accessed March 10, 2010); Andreas Rauber and Andreas Aschenbrenner, "Part of Our Culture Is Born Digital: On Efforts to Preserve It for Future Generations," *On-line Journal for Cultural Studies* (2001): 1–16, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.20.7640> (accessed March 10, 2010). See also Peter Lyman and Brewster Kahle, "Archiving Digital Cultural Artifacts," *D-Lib Magazine* (July/August 1998), <http://www.dlib.org/dlib/july98/07lyman.html> (accessed March 5, 2010); Howard Besser, "Moving from Isolated Digital Collections to Interoperable Digital Libraries," paper presented at the Victorian Association for Library Automation, Melbourne, Australia, February 8, 2002, <http://besser.tsoa.nyu.edu/howard/papers/vala01.html> (accessed March 10, 2010). Besser's two papers discuss more about access and interoperability.

<sup>2</sup> E.g., Kris R. Cohen, "What Does the Photoblog Want?" *Media, Culture & Society* 27, no. 6 (2005): 883–901, <http://dx.doi.org/10.1177/0163443705057675> (accessed February 17, 2010); Eric T. Meyer, Howard Rosenbaum, and Noriko Hara, "How Photobloggers Are Framing a New Computerization Movement," paper presented at "Internet Generations 6.0," Annual Conference of the Association of Internet Researchers, Chicago, IL, October 6–9, 2005, [http://people.oii.ox.ac.uk/meyer/wp-content/uploads/2007/09/meyerrosenbaumhara\\_photoblog\\_aoir\\_paper\\_v1\\_5.pdf](http://people.oii.ox.ac.uk/meyer/wp-content/uploads/2007/09/meyerrosenbaumhara_photoblog_aoir_paper_v1_5.pdf) (accessed February 2010); Haliyana Khalid and Alan Dix, "From Selective Indulgence to Engagement: Exploratory Studies on Photolurking," *Proceedings of the 20th British HCI Group Annual Conference 2*, ed. B. Fields, T. Stockman, L. Valgerour, & P. Healey, 17–20 (London: Queen Mary, University of London, The British Computer Society), <http://www.hcibook.com/alan/papers/HCI2006-indulgence> (accessed February 2010); and Lisa Gye, "Picture This: The Impact of Mobile Camera Phones on Personal Photographic Practices," *Continuum* 21 (2007): 279–288, <http://dx.doi.org/10.1080/10304310701269107> (accessed January 24, 2010).

Most of the published literature on digital media are available in scholarly online databases (e.g., D-Lib Online Magazine, New Media Society). Papers and conference proceedings are available online as well (e.g., Archives & Museum Informatics' Museum, the Web). Many historians will find these resources useful in their research, however as computer technology frequently upgrades, museum staff will benefit from more up-to-date recent publications.

Because of the abundance of the literature available, my search began online with four focused key terms: *photoblog*, *Web site acquisition*, *digital preservation*, and *cataloguing Web sites*. I filtered the material by keeping it museum-related, since libraries have different practices from museums on acquiring, organizing, and cataloguing artefacts. Each museum collection is unique: museums all collect for different purposes, and many of them organize and describe their collections differently from other museums' collections. Unlike libraries, many do not follow standardized cataloguing procedures, so that interested customers, researchers, and staff may not be able to access sources equally. Museums provide access to their collection by exhibiting the artefacts either in real or virtual space, for which they need to produce policies and guidelines relevant to their particular mission.

My purpose in surveying the literature was to consider how methods of acquisition, preservation, and cataloguing of photoblogs could fundamentally improve the relationship between museums' curatorial and exhibition departments, now and into the future. This focus posed a challenge at the beginning of my research, because the literature presented was intended for libraries or archives. Late in my research, Routledge released Ross Parry's *Museums in a Digital Age*, which contains articles on



information (information technology), space (within or outside museum walls), access, interpretation of digital materials, objects (are digital materials' authority, authenticity, and trust the same as for tangible objects?), delivery, and the future. All of these articles deal with digital materials and some with the Web; and photoblogs are digital material as well as Web material. None of the papers in the book mention photoblogs, but the content in *Digital Age* is devoted to museums and addresses their limits, questions, concerns, interests, and goals; discussions, statements, and suggestions offered in the book can be applied to photoblogs as well. In particular, Matthew Gansallo's "Curating New Media," focusing on Web art, could provide museums with reasons to be more comfortable acquiring a photoblog. Because one of museums' goals is to provide access to their collection, they need to exhibit their collection and Gansallo's article may give insight into how to exhibit a photoblog in the future. Other articles, such as Antonio M. Battro's "From Malraux's Imaginary Museum to the Virtual Museum," could provide a vision for how photoblogs in a museum's collection might be displayed and accessed.

Unfortunately, none of the articles in *Digital Age* mention photoblogs or provide suggestions for their acquisition, preservation, and cataloguing.

## SECTION 3:

### RESEARCH METHODOLOGY

Currently, since there are no known institutions with photographic collections that include photoblogs, I felt the best way to proceed was to address the very phenomenon of photoblogs. Since the literature concerning the acquisition, preservation, and cataloguing of photoblogs is limited, my research approach was to observe, analyze, and describe a photoblog, by using George Eastman House (GEH), International Museum of Photography and Film's collections-management database, The Museum System (TMS), and their collection and cataloguing practices.

Throughout this thesis, a photoblog, "People of Walmart," will be used as an example to illustrate the complexities of working with an immaterial object. It was chosen using the following criteria: it was new; it had changed over a short time (e.g., content had been removed and added); it had dynamic pages, with more than one provenance; and its images were photographs. "People of Walmart" (POW) contained images (produced mostly by camera phones, with some produced by digital cameras) of subjects ranging from actual Walmart shoppers to the flashy automobiles found in the parking lot. It also included a view of the creators, photographers, and visitors of the POW Web site. POW has provided a window into the culture and society of Walmarts in both the United States and Canada. It began as a personal photoblog among three friends in the summer of 2009, but quickly, the Web site took a sharp turn,<sup>1</sup> attracting more than four million visitors a day who have browsed, submitted images, created

---

<sup>1</sup> "The last week in August, 'People of Walmart' went viral, earning mention on Web sites like 'Funny or Die.' Its traffic increased by 700% on Aug. 27 alone. Photo submissions went from one or two to more than 120. The flood of visitors has even crashed the site."  
<http://www.time.com/time/arts/article/0,8599,1919401,00.html>

captions, or left comments.<sup>2</sup> POW's audience base has expanded to include Twitter, Facebook, and the site has offered an RSS feed option.

POW started as a basic photoblog, its layout consisting of traditional chronological posts with captions and a comment section. Images could only be submitted by emailing the file to the site owner.<sup>3</sup> On February 25, 2010, POW revamped their entire site with new hyperlinks, adding a video and stories section, and now images can be uploaded directly on the Web site.<sup>4</sup> Recently, it became revenue-generating collaborative photoblog in conjunction with Three Ring Blogs, a marketing company, and it has begun to run ten other photoblogs.<sup>5</sup>

"People of Walmart" is not and was never intended to be a serious Web site; as the original creators stated, "this is purely for entertainment purposes only." I am using the photoblog strictly as an illustration of the issues and solutions discussed in this thesis. In addition, POW has contemporary sociological research value. The creators of POW call it "Americana," and the site also provides insight and information on contemporary photography practices in America and the Web. POW demonstrates common cell phone camera usage, choice of subject matter, and behaviour and participation of photographers; the photoblog's photographers and visitors show the

---

<sup>2</sup> According to the "Advertise With Us" Web page, POW receives over seven million unique visitors per month. [http://www.peopleofwalmart.com/?page\\_id=15](http://www.peopleofwalmart.com/?page_id=15)

<sup>3</sup> Unfortunately, I have no screenshots or copy of the original site, so this comment is from memory. This is an example why documenting the beginning stages of a photoblog is critical, it changes over time without warning and there are no ongoing projects to archive them.

<sup>4</sup> PeopleOfWalMart.com, "Redesign," <http://www.peopleofwalmart.com/?p=9821> (accessed February 26, 2010).

<sup>5</sup> Three Rings Blogs, "Blog Network: Three Ring Blogs—Humor, Funny, Satirical, Shocking, LOL, WTF, OMG," <http://www.threeringblogs.com/> (accessed June 9, 2010)

activity and participation in photography on the Web. Therefore, POW is suitable for the study of photography in America on the Web of 2010.

With regard to the “People of Walmart” photoblog, I discussed and consulted on acquisition and preservation methods with Ryan Donahue, Manager of Information Systems for GEH. In response to my questions about what a museum would be looking for, Mr. Donahue explained how methods of acquisition and preservation worked in detail. The questions contained “what-ifs,” which forced us to think analytically about which of today’s methods would be the most suitable, such as whether a museum would be satisfied with a screenshot or whether they would want the entire photoblog. One of the goals of working with Mr. Donahue was to be able to explain the complex technology in the simplest terms possible.

As for cataloguing, I used TMS and GEH’s guidelines for describing artefacts. Empirical data I gathered and observed from GEH collecting practices influenced the decisions I made. I also visited other major art museums in Toronto, New York City, Chicago, Boston, Washington, DC, Columbus, Ohio, and Rochester, New York, where it became clear to me that every museum has its own cataloguing practices. I came to understand that catalogue records should be viewed as a reference when researching a museum’s collection.

Again, as stated, literature on photoblogs in museums is limited. I therefore deemed it appropriate to begin at the beginning, since all experiments start with a hypothesis. Making the right choices depends on the ultimate goal. I have watched how the Internet has grown since 1991, and my experiences at major museums, influences

from the literature presented, discussions with museum staff members—all these have played a role in informing the choices I present in this thesis.

## SECTION 4:

### PHOTOBLOGS

#### Photographically Oriented Web Sites

The term *photoblog* is derived from *Weblog*, commonly known as *blog*. A blog is

... a Web page made up of usually short, frequently updated posts that are arranged chronologically—like a new page or a journal. The content and purpose of blogs varies greatly—from links and commentaries about other Web sites, to news about a company/person/idea, to diaries, photos, poetry, mini-essays, project updates and even fiction.<sup>1</sup>

In place of words for thoughts, a photoblog has photographs. In a recently published guide book, *Photoblogging with WordPress*, David Busch and Rebekkah Hilgraves state that a “photoblog is a site devoted to photography. It may or may not include comments and technical information about the photograph.”<sup>2</sup> However, there are photoblogs that are not devoted to the art of photography. Instead, these photoblogs feature photographs as implements—they share pictures from a trip, demonstrate a cooking or painting skill, or show a newly remodelled kitchen. Kris Cohen’s “What do Photoblogs Want?” gives another definition:

A photoblog is a type of blog that is regularly updated with photos. Some photoblogs focus only on photography, while others have photos in addition to

---

<sup>1</sup> Murat Basaran, *Mobile Blogging: Seminar: Data Communication and Distributed Systems SS 2003*, under “What Is a Blog,” <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.72.5120&rep=rep1&type=pdf> (accessed 14 June 2010).

<sup>2</sup> David D. Busch and Rebekkah Hilgraves, *David Busch’s Quick Snap Guide to Photoblogging with WordPress: An Instant Start-up Manual for Creating and Promoting Your Own Photoblog* (Boston, MA: Course Technology, 2010), 7.

other content. All photoblogs, however, consider photos to be an important part of their chronological blogging structure.<sup>3</sup>

However, with the rapid changes in Web site development since blogging was made possible in early 21st century, the definition has become broader. Photoblogging usage has gone from private blogs, shared between friends and family, to marketing blogs that publicly share information worldwide with a large community of users. Photobloggers can—and do—now incorporate other media, such as audio and video clips, and nonphotographic illustrations. Therefore, it seems fitting to use *photoblog* as an umbrella term for any photographically oriented Web site that is chronologically updated with images.

In this thesis, I am concerned not only with the phenomenon of photoblogging but also with how photoblogs are created and what types of files and media they include, because these aspects will also play a role in any decisions made about acquisition and preservation.

### **Components of a Photoblog**

Museums must take into account the underlying structure and encoding of a photoblog when deciding how and what to acquire. From a systems point of view a photoblog is a Web site and Web sites are identified by a URL (uniform resource locator), which acts as an address specific to that resource. The inventor of the World Wide Web, Tim Berners-Lee, describes the URL as “a unifying syntax for the

---

<sup>3</sup> Click Opera, *Essay: Photoblogging*, under “What Is Photoblogging,” <http://imomus.com/photoblogging.html> (accessed 26 July 2010). The original Web site for this quotation, [www.photoblogs.org](http://www.photoblogs.org), no longer exists.

expression of names and addresses of objects on the network.”<sup>4</sup> Web browsers, applications that read and display Web pages, connect to the URL address, and retrieve and display the information found there. Currently the most widely used browsers are Mozilla Firefox, Apple Safari, and Internet Explorer.

The URL may or may not tell the viewer how the Web page is built. If it is visible, the URL will contain a new part after the top-domain name: a file extension code, separated from the main name by a slash. An example of a file extension code, shown here in italics, is *http://www.estamanhouse.org/main/index.php*. Otherwise, if it is not visible, the viewer can obtain the source code from the browser’s “view page source” option. The file extension or the page’s source code will tell the viewer what software was used to built the Web site (e.g. HTML (hypertext markup language), Java, and WordPress, for example). The file extension and the page source code can both be used to identify the blogging technology used and thereby assist in preservation.

The *source code* is the language of the photoblog; it generates HTML, which creates the structure of the Web page. It includes links, text, images, sound, and moving images. It contains coding that links to another Web page of the photoblog or to an external Web site. Included in the source code is the Cascading Style Sheet, or commonly known as CSS, which is responsible for the design and layout of the Web page. It describes the photoblog’s background colour and/or pattern, choice of font, size, and the colour of text, and the placement of text, images, and other designs. Simply put, it tells the script code how to display the Web page.

---

<sup>4</sup> Tim Berners-Lee, “Uniform Resource Locators,” <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.38.5918&rep=rep1&type=pdf> (accessed 16 June 2010).



A basic photoblog will contain photographs, text, and link(s). With the advent of coding and programming, photoblogs are becoming more complex and dynamic. A complex photoblog may include media other than photographic images: video, sound, and moving images (such as GIF and Flash). Interactive voting and commenting sections may be included as well. A briefing paper from Digital Preservation Europe says that blogs “come in all shape and sizes” and that statement applies to photoblogs.

### **The Components of “PeopleOfWalmart.com”**

The “People of Walmart” (POW) photoblog has a blue background with a white table background for posts, links, and other media. POW has a top header containing several links to other Three Ring Blogs. There are 15 major, related Web pages to visit at POW. They are “Home,” “Photos,” “Stories,” “Video,” “Submit,” “Store,” “About Us,” “Advertise with Us,” “FAQ,” “Hate Mail,” “Love Letters,” “News,” “Newsletter,” “Privacy Policy,” and “Terms and Conditions.” The “Photos” Web page has four possible Web pages: “Newest,” “Random,” “By State,” and “Top Rated.” The “Photos” Web page, displaying a traditional blog layout with the most recent post, it has over 300 Web pages with four photographs on each page. The other major Web pages contain more related Web pages and external links. All of POW’s Web pages contain advertisements, which are located at the top, bottom, and right side of the Web page. Each time a POW Web page is refreshed, a new advertisement may appear.

On the “Newest Photos” Web page, above each photograph is the date of the post and the post’s title; below it is a caption and the U.S. state where the photograph was taken. Below that, there is a “Share/Save” option if the viewer wishes to show or keep the post through Facebook, Twitter, Reddit, MySpace, Google Reader, Amazon

Wish List, LinkedIn, Google Buzz, or one of many other applications. It also has an email and bookmark option. All of this is viewable by placing the mouse cursor over the “Share/Save” button, which automatically displays a drop-down list.

Creators of the POW photoblog used WordPress, an open source content management system (CMS), to develop the photoblog’s content: text, links, images, videos, and other media. WordPress generates the script that is visible on the browser display.

The components of the “People of Walmart” photoblog will be briefly described, as they illustrate how the proper acquisition tools, preservation, and cataloguing methodologies can be determined.

## SECTION 5:

### COLLECTION DEVELOPMENT—WHAT AND HOW TO ACQUIRE

Museums need to decide how deeply they wish to document the photoblog. Museums who decide that photoblogs will enrich their collections should ask themselves the question Anne Laforet posed in her paper on “Preservation of Net Art in Museums”: “Which elements of Net art works should be emphasized, described, documented and kept?”<sup>1</sup> Are advertisements important? Are comments in posts important? Is the star rating or share feature important? Is the interactivity of photoblog important? Are related Web pages, such as “About,” “FAQ,” “Store,” “Contact,” and other possible pages that do not contain posts or images, and even external Web sites, important? The answer is: all of these are important, because a photoblog’s content describes its cultural and social context. To remove any content means the loss of a photoblog’s full history, material, construction, design, and function. Recently, the Library of Congress acquired Twitter. *The New York Times* online interviewed Amy Murrell Taylor, associate professor of history at the State University of New York in Albany about the importance of Twitter. Taylor confirmed that while individual Tweets might look like junk, looking through them collectively could be valuable.<sup>2</sup>

---

<sup>1</sup> Anne Laforet, “Preservation of Net Art in Museums,” in *Digital Visual Culture: Theory and Practice*, ed. Anna Bentkowska-Kafel, Trish Cashen, and Hazel Gardiner (Bristol, UK: Intellect, 2009), 110.

<sup>2</sup> Randall Stross, “When History Is Compiled 140 Characters at a Time,” *The New York Times* (April 30, 2010), [http://www.nytimes.com/2010/05/02/business/02digi.html?\\_r=1&ref=library\\_of\\_congress](http://www.nytimes.com/2010/05/02/business/02digi.html?_r=1&ref=library_of_congress) (accessed July 27, 2010).

## Copyright

Museums need to follow their country's current copyright laws in their acquisition policies. In many cases, permission must be obtained from the Web site owner. Current U.S. copyright law does not specifically state that Web sites fall under the protection of copyright; however it does state that "Copyright protects 'original works of authorship' that are fixed in a tangible form of expression. The fixation need not be directly perceptible so long as it may be communicated with the aid of a machine or device."<sup>3</sup> Literary, pictorial, and graphic works are copyrightable works.

For copyright protection and privacy reasons, a letter, email, or a phone call to the original site owner(s) is recommended and customary. Web scraping (software that uses a robot to copy a Web sites or Web pages) requires using the bandwidth of the photoblog's owner, which could raise suspicion and concerns about the Web site being hacked. Reproducing and displaying screenshots for other than fair use may violate copyright. One should confer with the museum's registrar (or similar department) and use current practices to obtain copyright and transfer of titles as suitable.

### The Tools: Acquiring a Photoblog

E. McClung Fleming's "Artifact Study: A Proposed Model"<sup>4</sup> has examined how an artefact's information provides cultural analysis and interpretation. Fleming wanted scholars to recognize that an artefact "implements, express, and documents a particular way of life," and suggested five possible acquisition methods that a museum can choose

---

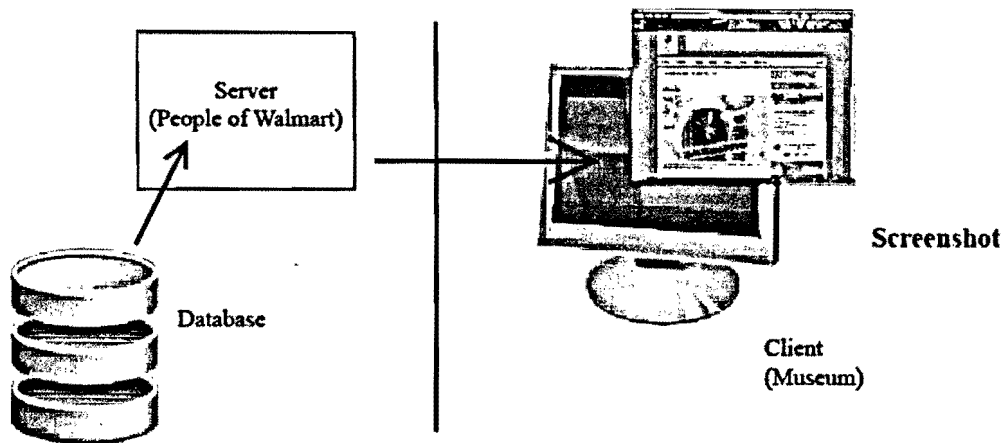
<sup>3</sup>United States Copyright Office, "Copyright Basics," under "What Works Are Protected?" <http://www.copyright.gov/circs/circ1.pdf> (accessed June 2010).

<sup>4</sup> E. McClung Fleming, "Artifact Study: A Proposed Model," *Winterthur Portfolio*, 9 (1974): 153–173, <http://www.mat.mat.fc.ul.pt/~jnsilva/Marta2.pdf> (accessed July 26, 2010).

from: capturing, copying, extracting, documenting, and archiving. This section is intended to serve as a reference guide for museum staff unfamiliar with the terminology. Museum staff should work with their information technologists who are familiar with Web development, databases, systems, and networks.

The sections below discuss acquisition tools from the most basic to the richest, including their advantages and disadvantages. The illustrations given for each acquisition tool give a visual example of how a digital object is acquired. On the left-hand side of each illustration is a photoblog's database, which contains records and information of the photoblog's text, links, images, and other media. The database provides the information to the photoblog's content management system (CMS) application (e.g., WordPress), which will generate the code once a URL command has been set on the user's browser. The server and the client (e.g., the museum) are divided by physical space. The server is located in one place while the client can be located anywhere where there is a computer and Internet access. In order for the client to view the desired photoblog, the client must insert a URL so that the browser will speak to the server. This will generate the source code with the database and then return the requested result on the browser's display.

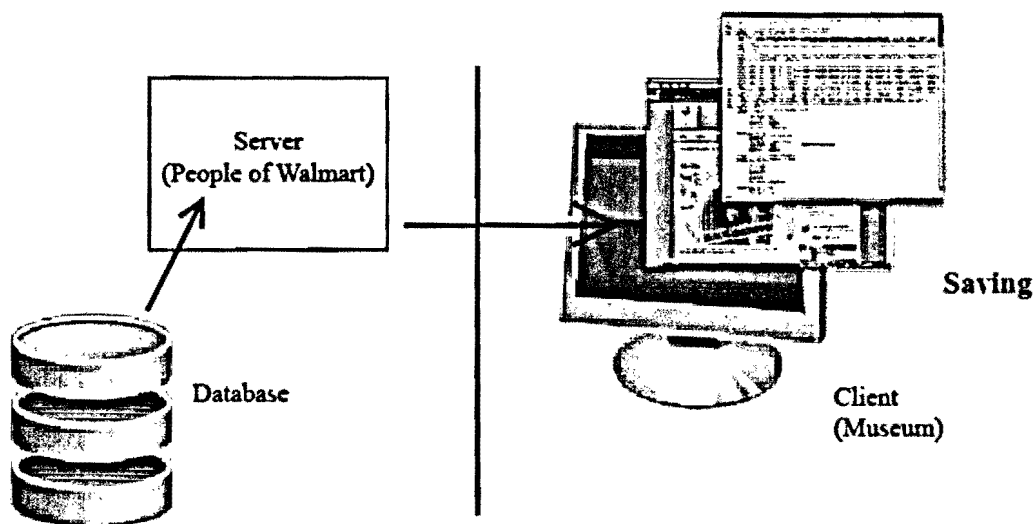
## The Capture: Screenshot



*Figure 1. The screenshot*

Screenshots—images documenting items displayed on the computer’s monitor—can be taken right on the museum’s computer. The advantages of taking screenshots are that no software installation is required and the museum can selectively choose what to document. However, there are quite a few disadvantages for taking screenshots from a curatorial standpoint. Depending on the museum’s computer operating system, screenshots are saved as BMP, PNG, or JPG files and the capture image is limited to the museum’s browser choice. Screenshots are static and there is a loss of the photoblog’s interactivity and dynamic features.

## The Copy: Saving the Source Code



*Figure 2. Saving the source code*

Museum staff can copy the source code by visiting their browser's view source option, and then save the source code as a Web archive or HTML source, depending on the computer's operating system. This can be viewed online, offline, or the user can host it online. Saving the source code is similar to taking screenshots, except it is more interactive. However, some dynamic features and links may be lost during the copying process, such as new comments, advertisements, and flash-operated media (moving images). Only one Web page can be saved at a time, requiring the museum to visit every Web page in order to save them.

## The Extraction: Web Scraping

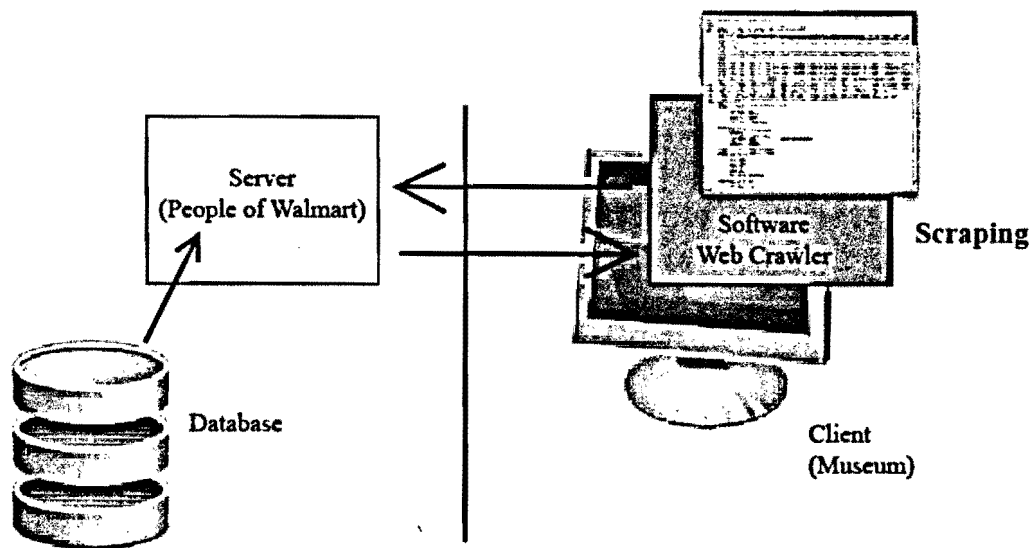


Figure 3. Web scraping

Web scraping, also known as Web harvesting and Web data extraction, is more complete than manual copying (see “Saving the Source Code”). It uses a software application program that retrieves data or information from Web sites.<sup>5</sup> One of the benefits of using a Web scraper is that it contains a crawler or a robot that will retrieve all possible Web pages requested. In acquiring a larger photoblog, however, one disadvantage is that the user has to think of all possible commands in order for the software to obtain all the pages the museum wishes to acquire. The Web scraper application goes directly to the Web site’s database and acquires the HTML. The museum can request the crawler to periodically crawl the photoblog, and either update their current files or add a new file folder. Web scraping software application saves data

---

<sup>5</sup> R. F. Sari and Aja Kurniawan, “Implementation of Indonesian Electronic Citation System Based on Web Extraction Techniques,” paper presented at the Third International Conference on Knowledge Discovery and Data Mining, 2010 (WKDD), Phuket, Thailand ( January 9–10, 2010), <http://dx.doi.org/10.1109/WKDD.2010.104> (accessed June 2010).



in files, which the user has to convert to make them viewable online. Often, the software may produce duplicate pages, may not get every page, and may block some pages. Also, owners of photoblogs who are concerned with privacy and copyright may insert a blocking code, a *robots.txt* file on their Web server, which prevents any crawler from retrieving or indexing the pages.

### The Aggregator: RSS Archiving

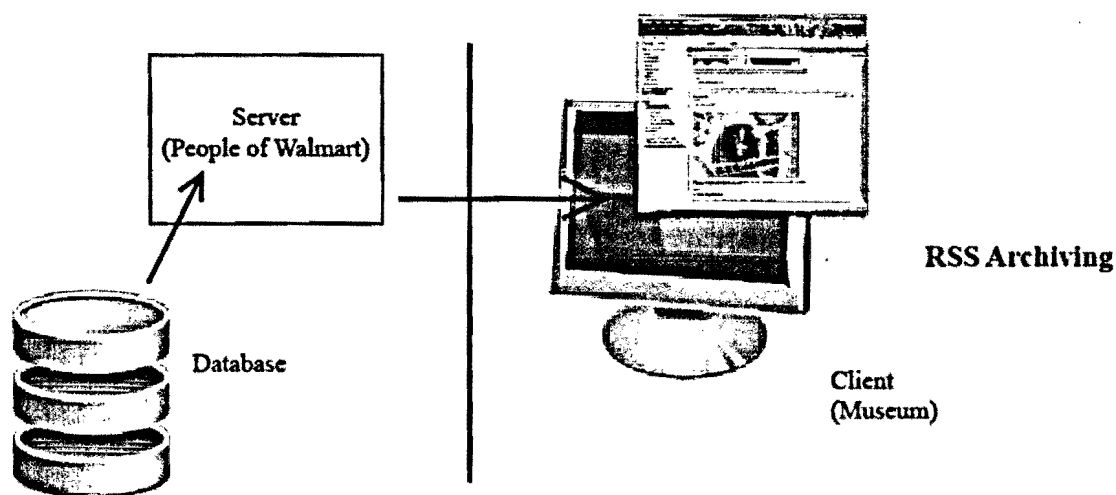


Figure 4. RSS archiving

Started in 1999,<sup>6</sup> Really Simple Syndication (RSS)

... works by allowing content distributors to syndicate brief snippets of content and post it as an RSS (XML) file on the Web. Most RSS files include a title, brief description, and a link where the user can follow up to retrieve the “full-story.” Those who wish to receive RSS content use special applications called RSS aggregators to “subscribe” to RSS feeds.<sup>7</sup>

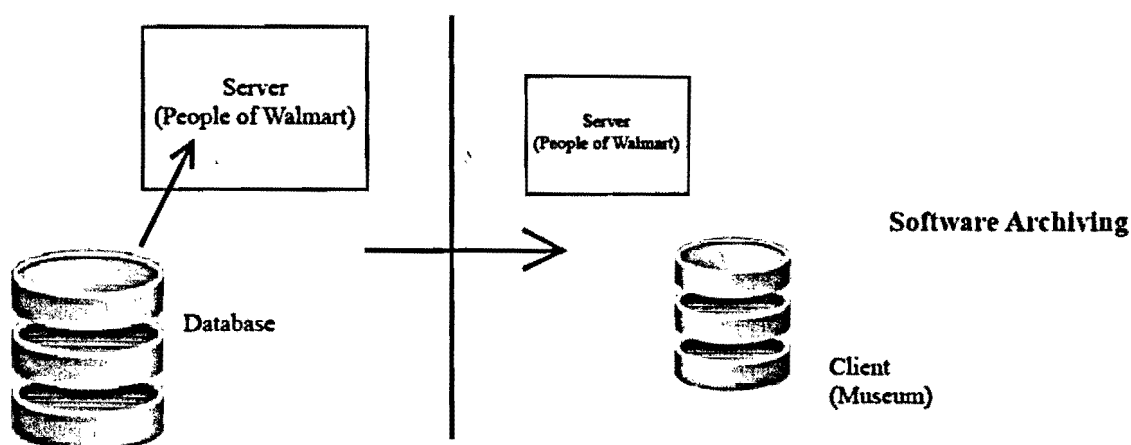
---

<sup>6</sup> Wikipedia, “RSS,” <http://en.wikipedia.org/wiki/RSS> (accessed June 2010).

<sup>7</sup> Ronald J. Glotzbach, James L. Mohler, and Jaime E. Radwan, “Really Simple Syndication (RSS): An Educational Approach,” *Journal of Interactive Media in Education* 3 (December 17, 2009). <http://jime.open.ac.uk/2009/03/jime-2009-03.html> (accessed June 13, 2010).

A benefit of using RSS is that a museum can set up an account with a Web site that extracts information from favourite Web sites, such as Google Reader.<sup>8</sup> By subscribing to a photoblog's URL, Google Reader can retrieve past and current posts. However, the photoblog must have subscribed and allow RSS to obtain the headlines from their Web site. Many personal photoblogs do not have that option.

### The Physical Acquisition: Software Archive



*Figure 5. Software archive*

The richest method of acquisition is to obtain the source code that generates the HTML and the database. By contacting the owner(s) of the site directly, the museum can acquire the physical files that generate and produce a functional and dynamic photoblog. Museums can go as far as purchasing or transferring the domain name of the Web site, if they wish to preserve the whole essence of the photoblog.

---

<sup>8</sup> Google Reader. "Welcome to Google Reader!" <http://reader.google.com> (accessed June 13, 2010).

## Overview: Copy Versus Scrapping Versus Software Archive



Figure 6. Example of a fabricated photoblog

```

Source of http://www.ryandanhue.net/isa.php?id=0041
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
<head>
<style type="text/css">
body {
background: #6CD9EA;
color: #33B9FF;
font-family: Helvetica;
}
p,meta {
color: #333333;
font-size: 11px;
font-style: italic;
}
h1 {
font-size: 40px;
border-bottom: 3px dashed #333333;
}
h2 {
margin: 2px 0;
font-size: 36px;
}
.sidebar {
position: absolute;
right: 20px;
top: 150px;
margin-right: 10px;
background: rgba(255,255,255,.66);
border-radius: 20px;
width: 320px;
height: auto;
padding: 20px;
}
.footer {
font-size: 12px;
}
hr {
border: 0;
height: 3px;
color: #333333;
background-color: #333333;
}
</style>
<meta http-equiv="content-type" content="text/html; charset=utf-8" />
<title>This is a Photoblog | 2 Peoples.</title>
</head>
<body>
<div class="sidebar">
<h2>Photoblog Posts</h2>
<div>
<a href="isa.php?id=0040">EFF Drinks.</a><div>2 Peoples.</div></div>
</div>
<div>
<h2>This is a Photoblog.</h2>
<h2>2 Peoples.</h2>

<div class="meta"></div>
<div class="buzz">Photo Scoop; 2010 HiTops</div>
</div>
<div class="footer">This Photoblog Scoop; 2010 HiTops. Code Scoop; 2010 Bartis.
Footer <a href="#">Link</a></div>
</body>
</html>

```

Figure 7. Example of a browser's script for figure 6

```

1 <?php
2
3 $posts = array('0040','0041');
4
5 if (in_array($_GET['id'], $posts)){
6     $id = $_GET['id'];
7 } else{
8     $id = $posts[0];
9 }
10
11 //--
12 $title['0040'] = "BFF Drinks.";
13 $meta['0040'] = "Posted 3/3/2010 in category 'Random' by Hi-Tops.";
14 $blurb['0040'] = "Photo &copy; 2010 Greatness";
15 //--
16 $title['0041'] = "2 Peoples.";
17 $meta['0041'] = "Posted 3/5/2010 in category 'Random' by Hi-Tops.";
18 $blurb['0041'] = "Photo &copy; 2010 HiTops";
19 //--
20
21 <?DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
22 "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
23 <html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
24 <head>
25 <style type="text/css"></style>
26 <meta http-equiv="content-type" content="text/html; charset=utf-8" />
27 <title>This is a Photoblog | <?=$title[$id]></title>
28 </head>
29 <body>
30 <div class="sidebar">
31 <h2>Photoblog Posts</h2>
32 <p>
33 <?php foreach ($posts as $p){
34     if ($id == $p){
35         echo "<?=$title[$p]><br/>";
36     } else {
37         echo "<a href='isa.php?id=$p'><?=$title[$p]></a><br/>";
38     }
39 }
40 </p>
41 </div>
42 <h1>This is a Photoblog.</h1>
43 <h2><?=$title[$id]></h2>
44 
45 <p class="meta"><?=$meta[$id]></p>
46 <p class="blurb"><?=$blurb[$id]></p>
47 <hr />
48 <p class="footer">This Photoblog &copy; 2010 HiTops. Code &copy; 2010 Bertie.
49 Footer <a href="#">Link</a></p>
50
51 </body>
52 </html>
53
54

```

Figure 8. Example of a server's script for figure 6

Figure 6 illustrates what a photoblog looks like on the Safari browser. By comparing the script retrieved from the browser (figure 7) to the server's script (figure

8), the reader will see that some coding is missing or rearranged. Saving the script (figure 7) through the browser is limited to the output that the server has given the viewer, thus limiting the viewer to certain contents of the Web page. For example, when saving Web page script from the “People of Walmart” photoblog, the advertisements and flash objects were not able to load. Web scraping will retrieve all the information from the server’s script or source code (figure 8), and display the photoblog more faithfully than a copy. However, with the physical acquisition of a software archive, the museum will acquire the database that is responsible for all the information that each item contained in the photoblog and it will run and display the photoblog faithfully than a saved script. Web scraping is a suggested tool for photoblogs that are live: new posts and comments are continuously added and the content changes over time. Web scraping software provides an option for museums to periodically scrap and index Web pages automatically. A software archive physically removes the photoblog from the owners and therefore institutions could choose to continue the life of the photoblog (such as by allowing new comments) or stop any future posts, comments, and/or other features from being added. Acquisition methods should be discussed between the curatorial and IT departments to decide on appropriate tools that will benefit the collection in both the short term and the long term. Each acquisition method has its advantages and disadvantages. Some institutions may find screenshots on their own suitable to demonstrate examples of photoblogs, while other institutions may wish to acquire the software archive of the photoblog because they want to preserve the content and the context of the blog.

## SECTION 6:

### FURTHER PRESERVATION OF ACQUIRED PHOTOBLOGS

When a museum acquires a photographic print, the print is then housed in a suitable environment for long-term care. However, to view a photoblog, a functioning computer with a monitor display is needed. Intel co-founder Gordon E. Moore's 1965 prediction that computer technology's performance would double every twelve months for the next ten years was right, and the trend continues. Commonly known as Moore's Law, computers are shrinking in size, growing more powerful, and expanding memory capacity as each year goes by. In this ever-changing environment, digital artefacts require continuous care and preservation. They require stable technology in order to ensure their survival. Technicians and conservators familiar with technology—including the hardware used to build the platform that runs the software—can ensure this survival, but museums are not usually dedicated repositories for housing computers. Photoblogs require a browser that can read their script in order to display the content correctly, but platforms and HTML upgrades become obsolete over time, and as new colour codes, plug-ins, and so forth are added, browsers must be made to adapt. The changes may cause them to ignore the original HTML. This section discusses possible preservation methods that do not require preserving the original hardware.

#### Data Migration

*Data migration* is the most basic preservation strategy. It is suitable for nondynamic, static, and noninteractive photoblogs (e.g., screenshots, copy) that will be read on a computer, because BMP, PNG, and JPG files are documented file formats that are open and available for anyone who wishes to include the file plug-in on their

computer. In data migration, a file is moved from one storage platform to another, for example by saving the file on a CD, memory card, and/or external drive. The task is not work-efficient. As files accumulate over time, a larger file space will be required to move them from one place to another. The task is also not cost efficient in the long term, because it would require the museum to request, build, purchase, and install any older software and applications that would function in the newest platform. Because museums may also find that they do not upgrade their platform as rapidly as the average customer, the next preservation method, the virtual machine, would be the most beneficial. As well, data migration is not beneficial for dynamic photoblogs that have been acquired through scraping or software archive because some of the photoblog's original content and interactivity might be lost. To prevent that would require migrating the original software, applications, and plug-ins to operate the photoblog, another reason why the method is not work-efficient for long-term preservation.

### **Virtual Machine and Emulation**

A virtual machine<sup>1</sup> (VM, also known as VirtualBox) is a guest that runs a desired operating system on a modern computer. Like a computer within a computer, VM builds an environment that allows users to run desired programs and software. For example, if a user wants to be able to run a program that is only compatible on MAC OS X, but the user has a PC with Windows 7, the user can use VM to run and display MAC OS X with the program by using his or her PC. This method of preservation can protect the original content of the photoblog, run the original operating system, browsers (because some Web sites display differently on Safari than on Firefox), and

---

<sup>1</sup> Sari and Kurniawan, "Implementation of Indonesian Electronic Citation System."



any other original software and applications. With a virtual machine, if the POW photoblog had been acquired by software archive or by Web scraping, the museum staff could have their current computer run and display the 2010 operating system, the browser that contains the original HTML coding, and other applications (such as WordPress). This would be beneficial from the standpoint of exhibition and research, since it would be able to preserve and display the original 2010 content into the future.

Another strategy in preservation is emulation, which allows the user to run the original program on a modern computer. For example, in twenty years, browsers will have changed HTML coding and display features, but programmers in 2030 could write a new program that would run and display the Mozilla Firefox or Safari from 2010. Jeff Rothenberg states that emulation “promises predictable, cost-effective preservation of original documents, by means of running their original software under emulation on future computers.”<sup>2</sup> In a non-photoblog example of emulation, a programmer can write a program for Xbox 360 (video game hardware) that mimics the 1991 Super Nintendo’s hardware and operating system, thus allowing Xbox users to view and play all Super Nintendo games (such as Super Mario). David Thomas believes that emulation “is a means of overcoming technological obsolescence of hardware and software by developing techniques for imitating obsolete systems on future generations of computers.”<sup>3</sup>

---

<sup>2</sup>Jeff Rothenberg, “Avoiding Technological Quicksand,” Council on Library and Information Resources, under “Introduction,” <http://www.clir.org/pubs/reports/rothenberg/criteria.html> (accessed June 2010).

<sup>3</sup> Digital Preservation Coalition, “Introduction: Definitions and Concepts,” under “Emulation,” <http://www.dpconline.org/advice/introduction-definitions-and-concepts.html> (accessed 26 July 2010).

Unfortunately, data migration, VM, and emulation do not help with physical platform obsolescence. This will require a conservator specialized in preserving, restoring, and building the original computer, its original hardware, and parts. However, VM and emulation offer an opportunity to display and run photoblogs as they were originally.

## SECTION 7:

### CREATING CATALOGUE RECORDS FOR PHOTOBLOGS

Cataloguing is integral to a museum's collection, as it guarantees access to the collection. While museums may store their acquired photoblogs in digital asset management (DAM), virtual machine, or other service (e.g., an online repository), creating catalogue records for photoblogs in a collections management database such as The Museum System (TMS) is recommended.

TMS can provide flexibility in locating, researching, and accessing photoblogs. TMS catalogue records store information that an artefact may not visibly show, such as the provenance, date, and maker. TMS also makes it easy to search and cross-reference other artefacts in the collection. This is beneficial from a curatorial and exhibition standpoint, since if a museum wished to do an exhibition on camera phone usage in 2010, for example, TMS could provide all possible documented records. In contrast, other search interfaces may be more limited. Internet Archive's Wayback Machine, for example, allows searching only by URL; results are then narrowed by year and month.<sup>1</sup> The British Library's UK Web Archive is browseable by title, by name of the special collection, and by a limited number of subjects.<sup>2</sup>

Catalogue records also provide the location of the artefact and act as a preservation tool, allowing the researcher to review details about the artefact before handling the actual artefact itself. In providing the location of the photoblog, the catalogue record also would prevent the need to pull out or require constant usage of the

---

<sup>1</sup> Internet Archive, "Wayback Machine," <http://www.archive.org> (accessed May 2010).

<sup>2</sup> UK Web Archive, "Blogs," <http://www.Webarchive.org.uk/ukwa/collection/100698/page/1/source/collection> (accessed June 10, 2010).

various hard drives, computers, or other storage mechanisms the museum may have chosen for their collection of photoblogs.

Using a collection management database other than DAM would provide an electronic paper trail of the collection and its artefacts. This section discusses whether incorporating descriptions of photoblogs (e.g., POW) into the existing database is appropriate.

### **Levels of Granularity and Fields Included in Descriptions**

Before beginning a catalogue record of any kind, a museum cataloguer must first determine the record's appropriate level: whether it should contain the entire photoblog in one description, a separate description for each post, or a separate description for each photograph. Depending on the acquisition method the museum has chosen, each level will provide slightly different results.

Many museums use their own in-house standards and cataloguing methodologies to reflect their collection's unique purpose. The following section discusses the core fields that GEH uses in their collection database, TMS. Under each heading below are questions and discussion about how each of these fields would be applicable to a photoblog, using POW as an example at the blog and post level. The reason for cataloguing at the blog and post level, and not at the image level, is that few museums have the resources and time to create individual records for each image. In the case of POW, the image description level would be the same as the post description level, because only one image is used per post. Other photoblogs may contain more than one image per post (e.g., BOSTON's Big Picture).

## Accession Numbers

Accession numbers are assigned to the artefact for identification purposes. These numbers are cross-referenced with other lots acquired in the same year or in the same collection. Assigning an accession number to a tangible artefact is more than identification—in the GEH system it also indicates the year the artefact was acquired, the collection it came in with, and the number of objects in the lot. If the POW photoblog were to be catalogued, an accession number would have to be assigned during its acquisition. POW's accession number might be 2010:0020:0001-1200, indicating that it was acquired in 2010, that it was the twentieth collection lot or photoblog the museum acquired, and that it had 1,200 photographs in the photoblog.

Interestingly, accession numbers are usually inscribed or taped on the artefact, raising the pertinent question of how a museum could write an accession number on a nontangible object. Several online scholarly databases use the digital object identifier (DOI) method, which (similar to an ISBN) provides a unique, stable number and URL for each article.<sup>3</sup> A DOI embedded in the photoblog and used in file naming could act as an accession number. If the museum acquires a software archive, including the domain name(s), the original characters from the URL identifier—such as the italics in <http://www.peopleofwalmart.com/?ph=9798>—might be usable as an accession number. However, these numbers do not signify anything other than location and identification.

---

<sup>3</sup> Wu, Heok, and Tamsir.

## Classification

Categorizing artefacts into classes identifies the artefact's function or type within an organizational scheme. Depending on the museum's choice of scheme, classification terms range from broad to specific. The GEH collection is large and nearly all photographic. Thus they need to be more specific—*stereocard* vs. *photograph*, for example. For other museums with various collection departments, *photograph* would be acceptable on its own. Photoblogs may fall into the category of *blogs*, since they include images that are “posted regularly with the most recent entry appearing first.”<sup>4</sup> The Library of Congress recognizes *blogs* as a valid subject heading<sup>5</sup> and the Oxford English Dictionary Online recognizes *Weblog* and *blog* as nouns in the English language.<sup>6</sup> Further scholarly research is needed to determine the taxonomy of photoblogs if museums wish to standardize digital vocabulary in the long term.

GEH classifies photographs in albums as *album*, indicating that the photograph is part of an album rather than its own entity. For photoblogs, cataloguing at the blog or post level should not matter—after all, photographs in a photoblog are what make it a photoblog. To be able to identify whether the catalogue record is at blog or post level, the accession number and the title description field will clarify whether the catalogue record is for a photoblog or for a post/image.

---

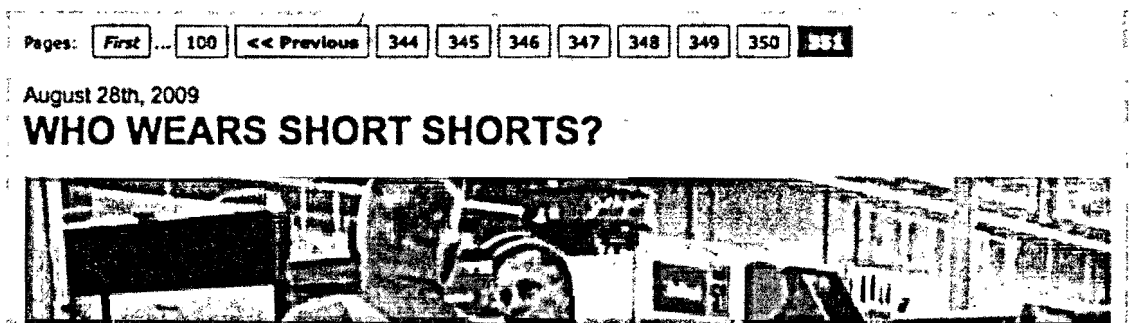
<sup>4</sup> Art & Architecture Thesaurus® Online Full Record Display, “Blogs,” under “Note,” [http://www.getty.edu/vow/AATFullDisplay?find=blog&logic=AND&note=&english=N&prev\\_page=1&subjectid=300265722](http://www.getty.edu/vow/AATFullDisplay?find=blog&logic=AND&note=&english=N&prev_page=1&subjectid=300265722) (accessed June 18, 2010).

<sup>5</sup> The Library of Congress, “Library of Congress Authorities,” [http://authorities.loc.gov/cgi-bin/PWebrecon.cgi?Search\\_Arg=blogs&Search\\_Code=SHED\\_&PID=1070Ba35a7wzHv5161OkmAA1ahWY0&SEQ=20100727120244&CNT=100&HIST=1](http://authorities.loc.gov/cgi-bin/PWebrecon.cgi?Search_Arg=blogs&Search_Code=SHED_&PID=1070Ba35a7wzHv5161OkmAA1ahWY0&SEQ=20100727120244&CNT=100&HIST=1) (accessed June 18, 2010).

<sup>6</sup> Oxford Dictionaries, <http://oxforddictionaries.com> (accessed June 18, 2010).

## Date

The date field indicates the time of execution or creation of the artefact, usually at completion. With a photographic object, GEH documents the date the photograph was printed, whether exact or circa. GEH's TMS date field allows the cataloguer to insert multiple dates, such as the day it went online and/or offline, for a blog level record. Finding the day it went online is searchable, since generally a photoblog's creator will post his or her first photograph or post the same day as the photoblog is activated. The last page on the POW photoblog (figure 9) states that its first post was on August 28, 2009; therefore, the photoblog went online that day. POW posts have a date stamped for each post, facilitating any cataloguing being done at a post level.



*Figure 9. Last page on POW photoblog*

## Constituent(s)

A constituent is a person(s) and/or organization(s) that has had a role in the creation or production of the artefact. According to GEH TMS, a variety of roles can provide clear authorship for an artefact: original photographer, donor, maker, subject (e.g., name of the person in the photograph), and organization or persons affiliated. Finding the constituents for POW required a little bit of searching; at the blog level the

maker would be AJK, ADK, and LDW,<sup>7</sup> the original photographer would be *various* or *unidentified*, and Three Ring Blogs could fall into the affiliated constituent role. Again, indicating constituency would vary, depending on the museum's cataloguing policies. Because all of POW's images are submitted anonymously, they would therefore be classified as *unidentified*. A museum's current practices on constituents cannot change simply because the format is digital.

### Title(s)

A title is an identifying name given to the artefact. According to the GEH TMS User Guide, if a title is not given, a descriptive title can be assigned as long as one clarifies that it is descriptive title. GEH's TMS has various title types, including *title on object*, *published title*, and *descriptive title*. The "People of Walmart" photoblog has various titles: on the header of the browser (figure 11), Web page titles (e.g., *Newest Photos*), and titles on posts (figure 9). Adding new title types for photoblogs will clarify where the title was obtained.

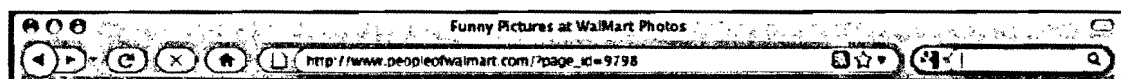


Figure 10. Title on header of browser

### Medium

The medium is a subclassification of the artefact's material makeup. As opposed to a classification, which tends to be broad (e.g., *photograph*), the medium descriptive field is specific (e.g., *silver gelatin*, *albumen*). Photoblogs are similar to albums; they

---

<sup>7</sup> People of WalMart.com, [http://www.peopleofwalmart.com/?page\\_id=12](http://www.peopleofwalmart.com/?page_id=12) (accessed January 2010).



are both repositories for a collection of photographs in one binding place. When cataloguing albums, GEH describes the types of photographs that are in the album and the same should be applied for photoblogs. Most of the POW photoblog's images were taken with a camera phone, thus creating a digital photograph. Other photoblogs may contain images with analogue counterparts that have been converted digitally. There is no need to clarify whether the image was once a colour photograph or albumen print and then digitized; if a museum wishes to be specific, it could be included in the appropriate description field.

### **Dimensions**

Dimensions are the measurement of the artefact's length, height, and depth (if applicable). It is not physically possible to measure the size of the photoblog with a ruler, since many photoblogs adapt to the size of the user's browsers. However, the "People of Walmart" photoblog has a set size, requiring viewers to set their browser in a size that will allow them to view the content of the page. File sizes and image sizes would be important to curators and exhibition designers because they need to determine how much space they need on the computer and how big the monitor needs to be in order to fully display the photoblog. This information would be useful for future researchers as well, especially since computer technology is expanding.

### **Description**

A description of the artefact can be provided if the title seems vague. Since the titles "Funny Pictures at WalMart" and "Who Wears Short Shorts?" can be interpreted differently, descriptions could include "Back view of an unidentified man in jean shorts and yellow tank-top." The description field includes a further account of the artefact's

medium and techniques. Future museum staff and researchers may be interested in how the photoblog was built, what types of plug-ins were used, if there are flash, GIF, or video files, what kind of fonts or styles were used for text, or if the image was digitally converted. This requires the cataloguer to have some knowledge of Web building and programming. Many nineteenth-century photographic portraits have hand-applied colour on the sitter's face and clothing; this information is added into the GEH's TMS description field. POW's photoblog was built with using WordPress, and this information should be entered as part of the description. I cannot stress enough that technology changes rapidly. Recording the information on how the photoblog was built will provide needed insight for future researchers into which technological methods were used, how they were used, and why.

### **Frame Case Mount**

GEH uses this field to describe a unique frame, case, or mount. The field is not often filled out unless its frame, case or mount sets it apart from other photographic objects. Many photoblogs have unique borders, designs, and backgrounds. These are called Web page *templates*, defined in one Web site as including "global navigation and examples of how different types of content, such as headings, pull-quotes, tables and form elements should be treated."<sup>8</sup> Photobloggers can download templates for free or for purchase.<sup>9</sup> Some photobloggers build their templates with plug-ins offered from their blogger service or through their Web site software. Adding descriptions about the

---

<sup>8</sup> The Motive Internet Glossary, "Template," <http://www.motive.co.nz/glossary/template.php> (accessed June 2010).

<sup>9</sup> Blogger Buster, "12 Free Photoblog Templates for Blogger," <http://www.bloggerbuster.com/2008/08/12-free-photoblog-templates-for-blogger.html> (accessed June 2010).

photoblog's template could be beneficial to future researchers who are interested in the types of layout on photoblogs or for a curator who would like to show some diversity or similarity in photoblog layouts for an exhibition.

### **Inscription(s)**

Inscriptions are written, printed, stamped, or engraved words, characters, and/or symbols on the surface of the artefact. While it is not possible to physically write on the photoblog, inscriptions can include symbols, comments, and captions. As described earlier in this thesis, each POW post contains a caption, the name of the U.S. state where the image was taken, and comments. All of these could be considered as inscriptions, and it would help the researcher or museum staff in determining that is the object they wish to evaluate, research, or exhibit before looking for the photoblog.

### **Overview of Cataloguing**

A museum's cataloguing practices with photographic collections need not change dramatically for cataloguing photoblogs. More features would need to be added, such as new vocabularies for classification, medium, techniques, and other attributes; new constituent roles would need to be created, as well as date and title types for clarification, such as *template*, *background*, or *frame case mount*. The Museum System and other collection management database software can provide the interface and operability to document, store, arrange, and search the collection. The most important part is to clarify where the information came from. The only difference in cataloguing a photoblog is its immateriality; the content and the context is not far removed from tangible photographic objects.

## SECTION 8:

### LIMITATIONS AND SUGGESTED FURTHER RESEARCH

Many questions, concerns, and debate have arisen throughout the writing of this thesis. More focus on cataloguing is warranted. In particular, a controlled vocabulary list, and a taxonomy of photoblogs and other photographically oriented Web sites need to be developed. Further exploration of and research into other museums, libraries, historical societies, and archives that are collecting and curating digital and Web media would provide insights into interpretation, authenticity, and metadata preservation.

A section devoted to the exhibition of photoblogs would be highly beneficial, since museums provide access and education to their collection through exhibitions. Originally, this thesis was to include such a section but it was eliminated because it would require its own focus, from theory to practical solutions. Exhibitions today will be different from the exhibitions that will take place fifty years from now. There is much literature and discussion concerning digital artefacts. One in particular, Matthew Gansallo's "Curating New Media," in *Museums in a Digital Age*<sup>1</sup> talks about issues surrounding exhibiting Web art in a room. Gansallo, a curator at London's Tate Museum, started an online project of commissioned work for the Tate's Web site. He wanted to preserve the tradition of entering a physical gallery and viewing artefacts in a space, or in his words, "How do you encourage interaction in both the gallery and online environment?" He believes that contemporary art museums will be able to exhibit Web art within and without museum walls.

---

<sup>1</sup> Ross Parry, ed., *Museums in a Digital Age* (London: Routledge, 2009), Part 5.

## Conclusion

There are five possible acquisition methods, each with their own advantages and disadvantages. In the end, however, to acquire something is better than to do nothing at all. Many museums are digitizing their collections, some of which are available online, and museum staff and information technology departments have already familiarized themselves with the preservation and organization of digital objects. Therefore, these museums are not new to digital and Web media, so they should already be comfortable in methods of preserving digital objects, like photoblogs. Cataloguing methods discussed and described in this thesis show that collection-management databases such as TMS could function appropriately for photoblogs, with some added fields and enhanced vocabularies.

Throughout this research and writing process, the “People of Walmart” photoblog changed every week with new features (and a few new images daily). Since its launch in the summer of 2009, it has been basic: there are no ads, no flash or moving images, viewers have had to submit images through email, and there were only a few major Web pages for “Photos,” “About,” and “FAQ.” Now, the photoblog has attracted so many visitors that the owners have changed the photoblog: they have connected their visitors through social networking sites like Facebook and Twitter, making the photoblog more interactive and a place for visitors to develop their online community. Online activity clearly evidences changes in communication and social activity practices, as well as changes in contemporary photography practices in the early twenty-first century. The rapidly changing “People of Walmart” photoblog, and the surge of photoblogs and photographs being added daily online both illustrate how

museums and other institutions may not have the opportunity to acquire the first few years of this exciting culture if action is not taken quickly.

## REFERENCES

- Archives & Museum Informatics. "Consulting, Publishing and Training for Cultural Heritage Professionals." <http://www.archimuse.com/index.html> (accessed March 10, 2010).
- Art & Architecture Thesaurus® Online Full Record Display. "Blogs." [http://www.getty.edu/vow/AATFullDisplay?find=blog&logic=AND&note=&english=N&prev\\_page=1&subjectid=300265722](http://www.getty.edu/vow/AATFullDisplay?find=blog&logic=AND&note=&english=N&prev_page=1&subjectid=300265722) (accessed June 18, 2010).
- Basaran, Murat. "Mobile Blogging." *Seminar: Data Communication and Distributed System SS 2003*. [http://www-i4.informatik.rwth-aachen.de/content/teaching/seminars/sub/2003\\_2004\\_ws\\_docs/moblogging.pdf](http://www-i4.informatik.rwth-aachen.de/content/teaching/seminars/sub/2003_2004_ws_docs/moblogging.pdf) (accessed June 14, 2010).
- Berners-Lee, Tim. "Uniform Resource Locators: A Unifying Syntax for the Expression of Names and Addresses of Objects on the Network" (October 14, 1993). <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.38.5918&rep=rep1&type=pdf> (accessed June 16, 2010).
- Besser, Howard. "Moving from Isolated Digital Collections to Interoperable Digital Libraries." Paper presented at the Victorian Association for Library Automation, Melbourne, Australia, February 8, 2002. <http://besser.tsoa.nyu.edu/howard/papers/vala01.html> (accessed March 10, 2010).
- Besser, Howard. "Longevity of Electronic Art." Paper submitted to the International Cultural Heritage Informatics Meeting, February 2001. <http://besser.tsoa.nyu.edu/howard/papers/elect-art-longevity.html> (accessed March 10, 2010).
- Besser, Howard. "The Changing Role of Photographic Collections with the Advent of Digitization." Draft of a chapter to appear in Katherine Jones-Garmil, ed., *The Wired Museum: Emerging Technology and Changing Paradigms*, Washington: American Association of Museums, 1997. <http://besser.tsoa.nyu.edu/howard/papers/garmil-eastman.html> (accessed March 10, 2010).
- Besser, Howard. "The Transformation of the Museum and the Way It's Perceived." Draft of a chapter to appear in Katherine Jones-Garmil, ed., *The Wired Museum: Emerging Technology and Changing Paradigms*, Washington: American Association of Museums, 1997. <http://besser.tsoa.nyu.edu/howard/papers/garmil-transform.html> (accessed March 10, 2010).

- Blogger Buster. "12 Free Photoblog Templates for Blogger."  
<http://www.bloggerbuster.com/2008/08/12-free-photoblog-templates-for-blogger.html> (accessed June 2010).
- Blogpluse.com. "Blogpluse Stats." <http://www.blogpluse.com/> (accessed June 2010).
- Busch, David, and Rebekkah Hilgraves. *Photoblogging with WordPress: An Instant Start-Up Manual for Creating and Promoting Your Own Photoblog*. Boston: Course Technology/Cengage Learning, 2010.
- Click Opera. *Essay: Photoblogging*. <http://imomus.com/photoblogging.html> (accessed 26 July 2010).
- Cohen, Kris R. "What Does the Photoblog Want?" *Media, Culture & Society* 27, no. 6 (2005): 883–901. <http://dx.doi.org/10.1177/0163443705057675> (accessed February 17, 2010).
- Digital Preservation Coalition. "Introduction: Definitions and Concepts."  
<http://www.dpconline.org/advice/introduction-definitions-and-concepts.html>  
 (accessed July 26, 2010).
- Fleming, E. McClung. "Artifact Study: A Proposed Model." *Winterthur Portfolio*, 9 (1974): 153–173. <http://wwmat.mat.fc.ul.pt/~jnsilva/Marta2.pdf> (accessed July 26, 2010).
- 404 Error Pages.com. "Welcome to 404 Error Pages.Com."  
<http://www.404errorpages.com/> (accessed June 13, 2010).
- George Eastman House, International Museum of Photography and Film. "Our Mission." <http://www.eastmanhouse.org/Main/museum/mission.php> (accessed May 2010).
- Glotzbach, Ronald J., James L. Mohler, and Jaime E. Radwan. "Really Simple Syndication (RSS): An Educational Approach." *Journal of Interactive Media in Education* 3 (December 17, 2009). <http://jime.open.ac.uk/2009/03/jime-2009-03.html> (accessed June 13, 2010).
- Google Blogs. "Blog Search." <http://blogsearch.google.com> (accessed May 2010).
- Google Reader. "Welcome to Google Reader!" <http://reader.google.com> (accessed June 13, 2010).
- Gye, Lisa. "Picture This: The Impact of Mobile Camera Phones on Personal Photographic Practices." *Continuum* 21 (2007): 279–288.  
<http://dx.doi.org/10.1080/10304310701269107> (accessed January 24, 2010).
- Internet Archive. "Wayback Machine." <http://www.archive.org> (accessed May 2010).



- Khalid, Haliyana, and Alan Dix. "From Selective Indulgence to Engagement: Exploratory Studies on Photolurking." *Proceedings of the 20th British HCI Group Annual Conference 2*, edited by B. Fields, T. Stockman, L. Valgerour, & P. Healey, 17–20. London: Queen Mary, University of London, The British Computer Society. <http://www.hcibook.com/alan/papers/HCI2006-indulgence> (accessed February 2010).
- Laforet, Anne. "Preservation of Net Art in Museums." In *Digital Visual Culture: Theory and Practice*, edited by Anna Bentkowska-Kafel, Trish Cashen, and Hazel Gardiner, 109–14. Bristol, UK: Intellect, 2009.
- LiveJournal: Discover global communities of friends who share your unique passions and interests. "Frequently Asked Question #4." <http://www.livejournal.com/support/faqbrowse.bml?faqid=4&view=full> (accessed June 2010).
- Lyman, Peter, and Brewster Kahle. "Archiving Digital Cultural Artifacts." *D-Lib Magazine* (July/August 1998). <http://www.dlib.org/dlib/july98/07lyman.html> (accessed March 5, 2010).
- Meyer, Eric T., Howard Rosenbaum and Noriko Hara. "How Photobloggers Are Framing a New Computerization Movement." Paper presented at "Internet Generations 6.0," Annual Conference of the Association of Internet Researchers, Chicago, IL, October 6–9, 2005. [http://people.oii.ox.ac.uk/meyer/wp-content/uploads/2007/09/meyerroosenbaumhara\\_photoblog\\_aoir\\_paper\\_v1\\_5.pdf](http://people.oii.ox.ac.uk/meyer/wp-content/uploads/2007/09/meyerroosenbaumhara_photoblog_aoir_paper_v1_5.pdf) (accessed February 2010).
- Napier, Mark. "Net.Flag—a Flag for the Internet." <http://netflag.guggenheim.org/netflag/> (accessed June 10, 2010).
- Palfrey, John G., and Urs Gasser. *Born Digital: Understanding the First Generation of Digital Natives*. New York: Basic Books, 2008.
- Parry, Ross, ed. *Museums in a Digital Age*. London: Routledge, 2009.
- PeopleOfWalmart.com. "Advertise with Us: Unique Visitors." [http://www.peopleofwalmart.com/?page\\_id=15](http://www.peopleofwalmart.com/?page_id=15) (accessed January 2010).
- PeopleOfWalmart.com. "Redesign." <http://www.peopleofwalmart.com/?p=9821> (accessed February 26, 2010).
- Rauber, Andreas, and Andreas Aschenbrenner. "Part of Our Culture Is Born Digital: On Efforts to Preserve It for Future Generations." *On-line Journal for Cultural Studies* (2001): 1–16. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.20.7640> (accessed March 10, 2010).

Rothenberg, Jeff. "Avoiding Technological Quicksand." Council on Library and Information Resources. <http://www.clir.org/pubs/reports/rothenberg/criteria.html> (accessed June 2010).

Sari, R. F., and Aja Kurniawan. "Implementation of Indonesian Electronic Citation System Based on Web Extraction Techniques." Paper presented at the Third International Conference on Knowledge Discovery and Data Mining, 2010 (WKDD), Phuket, Thailand, January 9–10, 2010. <http://dx.doi.org/10.1109/WKDD.2010.104> (accessed June 2010).

Smith, J. E., and Ravi Nair. "The Architecture of Virtual Machines." *Computer* 38, no. 5 (May 2005). <http://dx.doi.org/10.1109/MC.2005.173> (accessed June 2010).

Stross, Randall. "When History Is Compiled 140 Characters at a Time." *The New York Times*. [http://www.nytimes.com/2010/05/02/business/02digi.html?\\_r=1&ref=library\\_of\\_congress](http://www.nytimes.com/2010/05/02/business/02digi.html?_r=1&ref=library_of_congress) (accessed July 27, 2010).

Suddath, Claire. "That Viral Thing: People of Walmart." TIME.com. <http://www.time.com/time/arts/article/0,8599,1919401,00.html> (accessed January 2010).

The Library of Congress. "Library of Congress Authorities." [http://authorities.loc.gov/cgi-bin/PWebrecon.cgi?Search\\_Arg=blogs&Search\\_Code=SHED\\_&PID=1070Ba35a7wzHv5l6lOkmAA1ahWYo&SEQ=20100727120244&CNT=100&HIST=1](http://authorities.loc.gov/cgi-bin/PWebrecon.cgi?Search_Arg=blogs&Search_Code=SHED_&PID=1070Ba35a7wzHv5l6lOkmAA1ahWYo&SEQ=20100727120244&CNT=100&HIST=1) (accessed June 18, 2010).

The Library of Congress. "Library of Congress Web Archives Minerva." <http://lcWeb2.loc.gov/diglib/lcwa/html/lcwa-home.html> (accessed May 10, 2010).

The Library of Congress. "Web Capture and Archiving." <http://www.loc.gov/acq/devpol/Webarchive.html> (accessed May 26, 2010).

The Motive Internet Glossary. "Template." <http://www.motive.co.nz/glossary/template.php> (accessed June 2010).

Three Rings Blogs. "Blog Network: Three Ring Blogs—Humor, Funny, Satirical, Shocking, LOL, WTF, OMG." <http://www.threeringblogs.com/> (accessed June 2010).

UK Web Archive. "Blogs." <http://www.Webarchive.org.uk/ukwa/collection/100698/page/1/source/collection> (accessed June 10, 2010).

UK Web Archive. "Welcome to the UK Web Archive."

<http://www.Webarchive.org.uk/ukwa/> (accessed June 10, 2010).

United States Copyright Office. "Copyright Basics."

<http://www.copyright.gov/circs/circ1.pdf> (accessed June 2010).

University of North Texas Libraries, and U.S. Government Printing Office.

"CyberCemetery Home." <http://govinfo.library.unt.edu/> (accessed June 10, 2010).

The Variable Media Network. <http://www.variablemedia.net/e/welcome.html> (accessed June 10, 2010).

The Variable Media Network. "Variable Media Case Study: Mark Napier."

[http://www.variablemedia.net/e/case\\_napie\\_netfl.html](http://www.variablemedia.net/e/case_napie_netfl.html) (accessed June 10, 2010).

Wikipedia. "RSS." <http://en.wikipedia.org/wiki/RSS> (accessed June 2010).

———. "Tumblr." <http://en.wikipedia.org/wiki/Tumblr> (accessed June 2010).

Wortham, Jenna. "Once Just a Site with Funny Cat Pictures, and Now a Web Empire." *The New York Times*.

[http://www.nytimes.com/2010/06/14/technology/internet/14burger.html?scp=1&sq=i can has cheezburger&st=cse](http://www.nytimes.com/2010/06/14/technology/internet/14burger.html?scp=1&sq=i%20can%20has%20cheezburger&st=cse) (accessed June 13, 2010).

Wu, P. H. J., A. K. H. Heok, and I. P. Tamsir. "Annotating Web Archives—Structure, Provenance, and Context through Archival Cataloguing." *New Review of Hypermedia and Multimedia* 13, no. 1 (2007): 55–75.

<http://dx.doi.org/10.1080/13614560701423620> (accessed May 2010).