INFORMATION SCIENCE PUBLISHING
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.igi-global.com

IGR5195

This paper appears in the publication,
Handbook of Research on Global Information Technology Management in the Digital Economy
edited by Mahesh S. Raisinghani © 2008, IGI Global

Chapter XIX
Digital Preservation by Design

Stacy Kowalczyk
Indiana University, USA

ABSTRACT

Current knowledge is produced, disseminated, and stored in digital format. This data will not be preserved by benign neglect; digital information will be preserved only through active management. This chapter will provide a theoretical foundation for digital preservation, an overview of current best practice for digital preservation, and a research agenda as well as a proscriptive framework by which to design digital preservation into a system.

INTRODUCTION

An ever-increasing percentage of data, information, and expertise is being documented and distributed in an exclusively electronic medium. Our cultural history is now digital. Not only are family histories now being created with digital cameras, almost all current events are captured digitally. Newspapers, magazines, and books are produced digitally. Virtually all published photographs are created digitally—35mm film is nearly extinct. All broadcast media is digital. Radio is recorded digitally—audiotape is nearly extinct. Television is produced and delivered digitally. Movies will be distributed digitally within the next few years. Not just cultural history, but predictions are that scholarly output will be exclusively digital by 2010.

Businesses, as producers of knowledge, create many documents that contribute both to our cultural history and to the scientific record. Contracts, white papers, marketing and promotional materials, audio and video documentaries of important occasions, and people in the company are important not only to the company but to the world. All types of research—from pharmaceutical and medical, to electronics and computer science, to new power fuel and energy technologies—are business activities, and all digital. From buildings to clothing, almost all design is digital. All commerce is digital. Current knowledge is produced, disseminated, and stored in digital format.

What is the problem? Isn't information in digital form more secure than paper? It is not a matter of degree, it is a matter of understanding the differences between a fixed media and digital. While paper can be put in a box on a shelf, digital objects are increasingly expensive to store because the media needs to be replaced on a three- to five-year schedule. Unlike paper which will remain useable if put in a relatively safe place, digital objects require constant and perpetual maintenance. Digital objects depend on and are bound to a technical environment and infrastructure. As the environment changes, so might the objects. And as the digital objects become more complex, the problems just increase. Data will not be preserved by benign neglect. Digital materials will only be preserved by active management.

Because so much information exists only in digital format, concern about our ability to keep this information available and usable for the future is increasing. In August, 2006, a prime example of a digital disaster made headlines. Tapes with very high-quality pictures of the first moon walk in 1969 have been reported as missing. To add insult to injury, the only remaining tape drive that can read these tapes is scheduled to be retired within the year. The missing tapes have a much higher quality picture than what was seen on, and recorded for, television (Macey, 2006). The loss is incalculable.

Businesses, governmental agencies, libraries, archives, and museums all need solutions to these problems. Over the past 10 years, digital preservation has emerged as an important area of research in both computer science and information science, at a national and international scale. Much of the research in digital preservation is coming from digital library programs more than from the traditional academic research community. Digital preservation is defined as:

…the managed activities necessary: (1) For the long-term maintenance of a byte stream (including metadata) sufficient to reproduce a suitable facsimile of the original document and (2) For the continued accessibility of the document contents through time and changing technology. (RLG & OCLC, 2002)

This chapter will provide a theoretical foundation for digital preservation, an overview of current best practice for digital preservation, and a research agenda.

DIGITAL PRESERVATION STRATEGIES

In a research area that is only 10 years old, it seems almost silly to talk of the "early days"; but in the early days, the focus of the conversation and theorizing revolved around preservation strategies. Three strategies emerged: technology preservation, technology emulation, and data migration. These are best defined though example. We will use the trivial example of a CD-ROM application that provides a virtual tour of the Vatican.

The technology preservation strategy proposes to save the technology platform for an application to preserve not only the data, but the "look and feel" as well. In our example, we would need to save the CD-ROM, a computer with a CD-ROM reader with the appropriate read speed as well as the correct operating system for the CD-ROM application software. In order to take the virtual tour, one would need to use that specific computer. Known as the museum-style approach (UKOLN, 2006), one can see the attraction to this model—it seems simple to just keep the hardware functioning. However, its simplicity is also its downfall. Depending on the number of applications to be kept functioning and the number of different platforms, the complexity will escalate until the cost becomes prohibitive. But the most significant failure point is that eventually, due to the increasing age, the hardware and storage media will fail irrevocably. Preserving the hardware as a scientific artifact for the future is a worthy goal. A number of museums are collecting computers. The Science Museum and the Computer Conservation Society in the UK are primary examples (UKOLN, 2006). But as a preservation strategy, because of its obvious weaknesses, it was never considered to be a serious option.

Like technology preservation, technology emulation is a strategy that has as its goal the preservation of the complete functionality of the original system. This strategy is based on the premise that system behaviors can be defined independently from its implementation. With the description of behavior, an engine could be developed that would re-create those behaviors. Alternatively, the source

Related Content

Electronic Health Records (EHR) and Clinical Decision Support Systems: Integrating AI Solutions
G. U. Vasanthakumar, V. Dankan Gowda, Prabhakar S. Manage, K. D. V. Prasadand Venkatesan Hariram (2024). *Handbook of Research on Innovative Approaches to Information Technology in Library and Information Science (pp. 277-302).*
www.irma-international.org/chapter/electronic-health-records-ehr-and-clinical-decision-support-systems/337313

Structuring Information Systems-in-Use: Studying the Replication of an E-Procurement System through a Practice Lens
Angela Linand Shin-Horng Chen (2013). *Journal of Global Information Management (pp. 1-18).*
www.irma-international.org/article/structuring-information-systems-use/73786

The Role of Multinationals in Recent IT Developments in China
Michelle Rowe (2008). *Global Information Technologies: Concepts, Methodologies, Tools, and Applications (pp. 1938-1954).*
www.irma-international.org/chapter/role-multinationals-recent-developments-china/19085

Evaluation of the SCM Performance in Using of Global Logistics Information Technologies: A Research Study in Hong Kong
Pui Yuk Chanand Xinping Shi (2008). *Global Information Technologies: Concepts, Methodologies, Tools, and Applications (pp. 1573-1586).*
www.irma-international.org/chapter/evaluation-scm-performance-using-global/19059

Determinants of Smart Digital Infrastructure Diffusion for Urban Public Services
Bhaskar Choudhuri, Praveen Ranjan Srivastava, Shivam Gupta, Ajay Kumarand Surajit Bag (2021). *Journal of Global Information Management (pp. 1-27).*
www.irma-international.org/article/determinants-smart-digital-infrastructure-diffusion/295976