

7-2009

Report on the First International Workshop on Innovation in Digital Preservation (InDP 2009)

Frank McCown Ph.D.

Harding University, fmccown@harding.edu

Hannes Kulovits

Vienna University of Technology

Andreas Rauber

Vienna University of Technology

Follow this and additional works at: <http://scholarworks.harding.edu/computer-science-facpub>



Part of the [Computer Sciences Commons](#)

Recommended Citation

McCown, F., Kulovits, H., & Rauber, A. (2009). Report on the First International Workshop on Innovation in Digital Preservation (InDP 2009). *D-Lib Magazine*, 15 (7/8). <http://dx.doi.org/10.1045/july2009-mccown>

This Article is brought to you for free and open access by the Computer Science at Scholar Works at Harding. It has been accepted for inclusion in Computer Science Faculty Research and Publications by an authorized administrator of Scholar Works at Harding. For more information, please contact scholarworks@harding.edu.



HARDING
UNIVERSITY

D-Lib Magazine
July/August 2009

Volume 15 Number 7/8

ISSN 1082-9873

Report on the First International Workshop on Innovation in Digital Preservation (InDP 2009)

[Frank McCown](#)

Harding University

<fmccown@harding.edu>

[Hannes Kulovits](#)

Vienna University of Technology

<kulovits@ifs.tuwien.ac.at>

[Andreas Rauber](#)

Vienna University of Technology

<rauber@ifs.tuwien.ac.at >

Overview

The first workshop on Innovation in Digital Preservation (InDP) was held on June 19, 2009, in conjunction with the Joint Conference on Digital Libraries (JCDL 2009), on the campus of the University of Texas at Austin. The goal of InDP was to provide a forum where researchers could share and discuss the latest innovations in digital preservation (DP) by non-traditional methods or methods that are not normally used by traditional memory organizations.

There were sixteen participants at this full-day workshop, and six presentations were given that highlighted a number of different research projects in the areas of digital preservation (the six papers can be found online at [1]). Each of the presentations sparked a great amount of discussion among the workshop participants. Time was left at the end of the workshop to outline a research agenda that would stake out future directions to be investigated in the field of DP.

All presentations were filmed by Spencer Lee (Virginia Polytechnic Institute and State University) and will be made available in Second Life [2] on the virtual island of Digital Preserve via streaming video. A screenshot of the virtual JCDL and initial construction of InDP are shown in Figure 1.

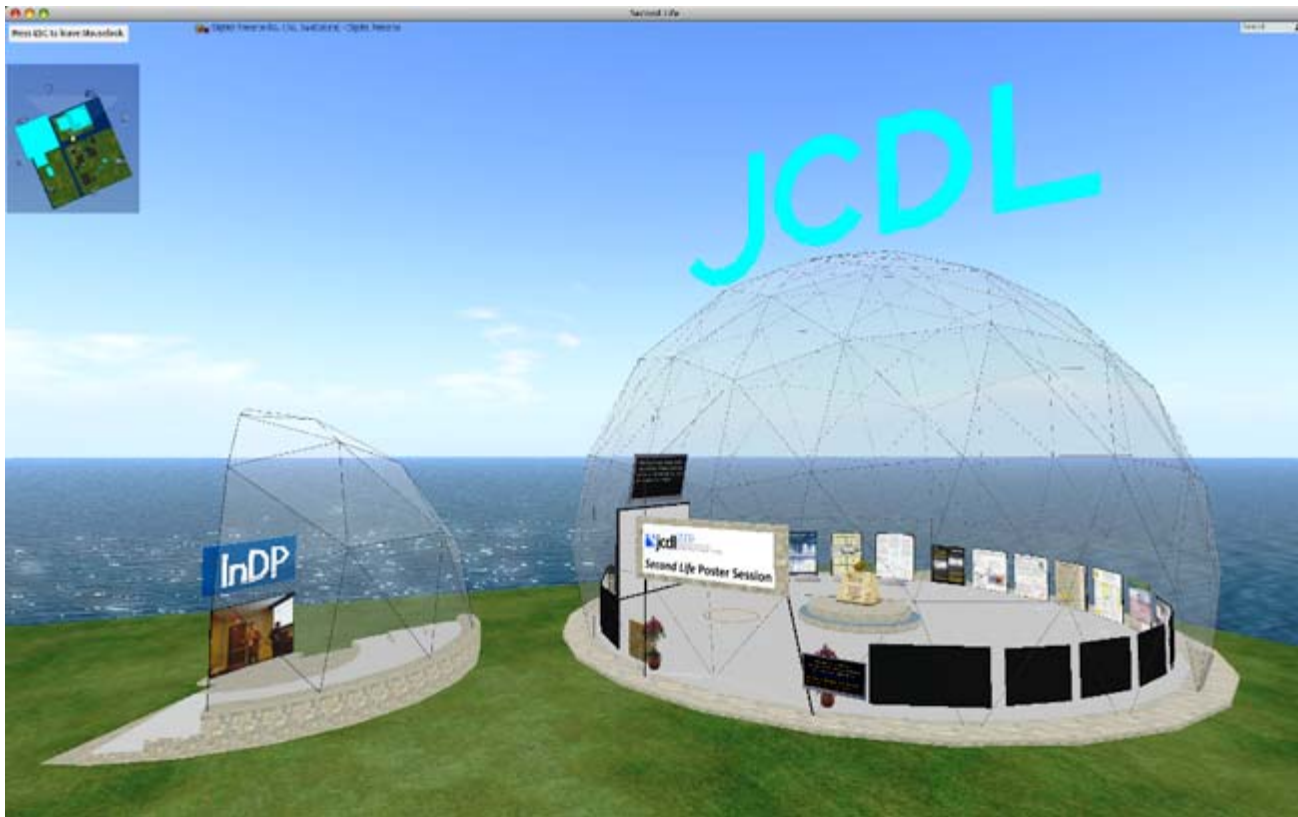


Figure 1. JCDL and InDP in Second Life. Copyright 2009 Spencer Lee. Used with permission.

Presentations

Rudolf Mayer (Vienna University of Technology) gave the first presentation, discussing how context could be established after the fact on a collection of digital objects. Mayer introduced a semi-automatic method that could determine the creation and usage context of digital objects, and he showed how it could be applied to a collection of emails. Another participant noted that such a system could be helpful in her work of sifting through thousands of emails that were turned over to her from a leading researcher.

Martin Klein (Old Dominion University) then presented his study on how web page titles change over time, using historical snapshots from the Internet Archive [3]. Klein's research focuses on finding missing web pages (web pages that are no longer accessible at a once-valid URL) by using search engines and web archives. His detailed study showed that web page titles change very little over time, which aids in using search engines to find web pages that change URLs. Klein is working on building predictive models for showing how titles may change in the future.

The third presentation was given by Dominic Heutelbeck (University of Hagen), who carefully examined the preservation difficulties currently plaguing design and engineering companies. Preservation best practices have not yet made inroads into companies that design and produce electronic equipment and products, and Heutelbeck believes the key is in showing how the economic benefits of integrating such practices can help the bottom line. Developing a preservation system is very challenging, especially considering the breadth of proprietary software and data formats used in the industry.

Josep Lluís de la Rosa (University of Girona) gave the final presentation before lunch on a novel use of agents for performing preservation. de la Rosa presented the findings of an experiment using agents of various abilities to migrate at-risk files. Using the Shout and Act algorithm, the agents pursued at-risk files on a hard drive and called for help to other agents to assist in migrating the files

to more stable formats. His findings showed promise in building distributed-agent preservation systems.

After lunch, Gonçalo Antunes (INESC-ID) argued that DP could be better understood as a specific case of Systems of Systems Engineering (SOSE). By placing DP in an SOSE context, a number of best-practices can be applied from SOSE, and some terminology and concepts can be better expressed. A mapping of DL techniques to SOSE terminology is given in the paper.

The final presentation was given by Edward Fox and Spencer Lee (Virginia Polytechnic Institute and State University). Fox discussed emerging research being conducted by Virginia Tech and the University of North Carolina at Chapel Hill in the preservation of virtual worlds and in using virtual worlds to facilitate educational and working goals of curators and archivists. Fox gave a virtual tour of the digital island of Digital Preserve in Second Life where posters from JCDL 2009 were made available to curious avatars.

Research Directions

The workshop concluded with an open discussion of where the DP field was headed and where it had possibly gone astray. Kulovits opened with some provocative statements designed to illicit discussion. For example, Kulovits stated that DP research is often reactive instead of proactive, that results are needed now, so there is a great lack of creative energy given to problems of the future. DP systems should be better evaluated by using stringent standards of empirical evidence, validation, and benchmarking.

The workshop participants churned some ideas of what was needed in DP research. For example, one participant suggested that a formal framework for DP was needed and that the field should be more clearly delineated. Others argued that it was difficult to determine what success looked like; we often do not know if preservation has succeeded until a sufficient amount of time has passed. One question was asked, "How can we more actively engage lawyers, historians, and archeologists in DP?" Several participants suggested that DP must be made invisible, and our systems must be preservation-friendly from the start.

The authors of this report were excited by the participation and interest shown in this first InDP workshop. Clearly, there is a need for more innovative solutions in the field of digital preservation, and we hope this workshop may continue to illicit more participation in the future at other venues.

References

1. InDP schedule. [Online] <<http://cs.harding.edu/indp/schedule.html>>.
2. *Second Life*. [Online] <<http://www.secondlife.com/>>.
3. *Internet Archive's Wayback Machine*. [Online] <<http://www.archive.org/>>.

Copyright © 2009 Frank McCown, Hannes Kulovits, and Andreas Rauber

[Top](#) | [Contents](#)
[Search](#) | [Author Index](#) | [Title Index](#) | [Back Issues](#)
[Previous Conference Report](#) | [Next Conference Report](#)
[Home](#) | [E-mail the Editor](#)

[D-Lib Magazine Access Terms and Conditions](#)

doi:10.1045/july2009-mccown