

simplyCT: A Lightweight Digital Library Framework

Lighton Phiri
Department of Computer Science
University of Cape Town
Private Bag X3
Rondebosch 7701
Cape Town, South Africa
lphiri@cs.uct.ac.za

Hussein Suleman
Department of Computer Science
University of Cape Town
Private Bag X3
Rondebosch 7701
Cape Town, South Africa
hussein@cs.uct.ac.za

ABSTRACT

Digital Libraries are information systems that store digital objects, and have associated services for accessing, managing, and preserving the digital objects. Digital Libraries began as an abstraction layered over databases to provide higher level services. As the services and tools became more complex, they became more difficult to maintain, extend and reuse. One possible solution is to explicitly avoid formalisms, abstractions and Application Programming Interfaces (APIs) so that Digital Library Systems are more easily adopted and managed.

Categories and Subject Descriptors

H.3.7 [Information Storage and Retrieval]: Digital Libraries—*Collections, System issues, User issues*

General Terms

Design, Experimentation, Human Factors

Keywords

Architecture, Framework, simplyCT

1. INTRODUCTION

Digital Libraries are information systems that store digital objects, and have associated services for accessing, managing, and preserving the digital objects (Arms, 2001). Digital Libraries began as an abstraction layered over databases to provide higher level services.

As the services and tools became more complex, they became more difficult to maintain, extend and reuse. One possible solution is to explicitly avoid formalisms, abstractions and Application Programming Interfaces (APIs) so that Digital Library Systems are more easily adopted and managed.

2. RELATED WORK

A number of Digital Library System architectures have been proposed over the past two-decades, ranging from those specifically designed to handle complex objects (Tansley et al., 2003; Lagoze et al., 2006) to those with an overall goal of creating and distributing collection archives (Millington and J., 2007; Witten et al., 2000).

The core of this research is focused on investigating whether it is feasible to implement Digital Library Systems based on simple architectures. This is the basis of the simplyCT framework and a high level architectural design of the framework is shown in Figure 1. As such, this research will seek to provide a comprehensive solution on how simple architectures should be defined and further explore the advantages and disadvantages that simple DL architectures have when compared to complex DL architectures.

3. RESEARCH APPROACH

A clear demonstration that the underlying principles behind the simplyCT framework will result in DLSs that are effective and efficient is vital for the success of this research. This will be done systematically by first conducting a detailed case study of a few popular existing DLSs, to provide a clear understanding of the principles and theories behind Digital Libraries. A prototype Digital Library System will be implemented and a number of collections will be integrated with the DLS to assess the overall applicability of the tool.

The overall research assessment will be based on the three axes of evaluation strategy and will involve:

- Case studies, where a number of real life collections will be implemented to measure the effectiveness of the framework
- User studies, conducted to assess the relevance of typical DLSs implemented using the simplyCT framework
- Controlled experiments to measure efficiency for typical DL tasks such as search and browse

4. ANTICIPATED OUTCOMES

The following deliverable are expected to be product upon successful completion of this project:

- A set of principles underlying theory behind the simplyCT framework

- Proof of concept systems in form of working digital collections based on the simplyCT framework

It is anticipated that the guiding principles of the simplyCT framework will result in the implementation of DLSs and/or services that will be easily adopted and fairly easy to maintain. This would be especially useful when implementing DLSs such as cultural heritage collections in developing countries.

5. BIBLIOGRAPHY

- [1] William Y. Arms. “Libraries, Technology, and People”. In: *Digital Libraries*. Ed. by William Y. Arms. 2nd. Cambridge, Massachusetts: The MIT Press, 2001. Chap. Chapter 1, pp. 1–20.
- [2] Carl Lagoze, Sandy Payette, Edwin Shin, and Chris Wilper. “Fedora: an architecture for complex objects and their relationships”. In: *International Journal on Digital Libries* 6.2 (Apr. 2006), pp. 124–138. DOI: [10.1007/s00799-005-0130-3](https://doi.org/10.1007/s00799-005-0130-3).
- [3] Peter Millington and Nixon William J. *EPrints 3 Pre-Launch Briefing*. 2007. URL: <http://www.ariadne.ac.uk/issue50/eprints-v3-rpt/> (visited on May 1, 2011).
- [4] Robert Tansley, Mick Bass, David Stuve, Margret Branschofsky, Daniel Chudnov, Greg McClellan, and MacKenzie Smith. “The DSpace Institutional Digital Repository System: Current Functionality”. In: *Proceedings of the 2003 Joint Conference on Digital Libraries*. Houston, United States: Institute of Electrical and Electronic Engineers, May 2003, pp. 87–97. DOI: [10.1109/JCDL.2003.1204846](https://doi.org/10.1109/JCDL.2003.1204846).
- [5] Ian H. Witten, Stefan J. Boddie, David Bainbridge, and Rodger J. McNab. “Greenstone: A Comprehensive Open-source Digital Library Software System”. In: *Proceedings of the Fifth ACM Conference on Digital Libraries*. DL ’00. New York, NY, USA: ACM, 2000, pp. 113–121. DOI: [10.1145/336597.336650](https://doi.org/10.1145/336597.336650).

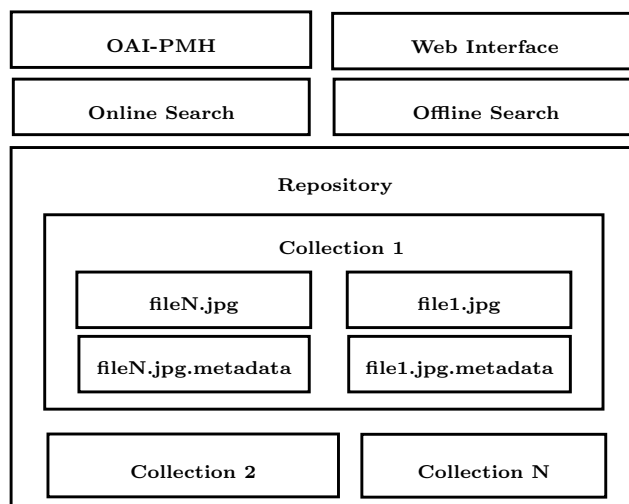


Figure 1: simplyCT architectural elements