IJDL focussed issue on persistent archives

Dominic Heutelbeck · Claus-Peter Klas

Published online: 26 June 2012
© Springer-Verlag 2012

Digital preservation is a major concern in the digital age. The classical challenges we face in this area are multifaceted and range from data medium conversion to software emulation for documents and multimedia objects. As we generate massive amounts of highly complex data, new challenges arise to keep resources usable, meaningful, and understandable over long time periods.

The four papers collected in this issue look at different aspects of persistent archives, i.e., classification of collections, different distributed architectures for archival, as well as the state-of-the-art in the domain of engineering which has been and is still driving a lot of the requirements for persistent archives.

In the paper “Using wavelet analysis for text categorization in digital libraries: a first experiment with Strathprints” by Darányi/etal the focus is on handling large amounts of documents and to categorize them automatically for later access. It combines different research areas to form an innovative approach.

In “Extending OAI-PMH over structured P2P networks for digital preservation” Everton/etal explore the possibility to distribute the responsibility of keeping copies and therefore prevent loss of information to a network of OAI-enabled libraries. They also take the reliability into account and calculate the best distribution.

To investigate complex objects Brunsmann explores in “State-of-the-art of long-term preservation in product life

cycle management” the current expertise on preservation in industry, especially in product life cycle management. In such an environment, we have to handle not only a document, but a set of information objects like CAD drawings and the connected personnel, project information, customer requirements etc.

In Bing et al.’s “Digital repository: preservation environment and policy implementation” a flexible rule-based and policy-driven system called iRODS is presented. It enables a peer-to-peer server middleware, provides the requisite dynamic extensibility needed to manage time-varying policies, automate validation of assessment criteria, manage ingestion processes, manage access policies, and manage preservation policies.

Research in the domain of persistent archives is a continuous research with new challenges appearing as technologies develop. Preservation must keep pace and include new developments to avoid the loss of critical knowledge. One example is the increasing distributed and dynamic nature of information resources. Paradigms like cloud-based and service-oriented architectures increase the degree of distribution and heterogeneity of complex resources. These trends result in the need for integration of preservation throughout the entire life cycle of digital resources. These new requirements will continue to make research on persistent archives exciting and challenging for the years to come.

D. Heutelbeck (✉) · C.-P. Klas
University of Hagen, 58084 Hagen, Germany
e-mail: Dominic.Heutelbeck@fernuni-hagen.de