

Hoppla - Digital Preservation Support for Small Institutions

Stephan Strodl and Andreas Rauber
 Institute of Software Technology and Interactive Systems
 Vienna University of Technology
 {strodl, rauber}@ifs.tuwien.ac.at

Motivation

- Digital Preservation primarily addressed by large heritage institutions
- Increasing amounts of digital data are held and organized by small institutions and at home
- Data of high value (scientific, business or emotional)
- Current archiving solutions are mostly CMS or secure storage with no support for long term preservation
- Need of practical preservation solution for SME, SOHOs (Small Office / Home Office), private users

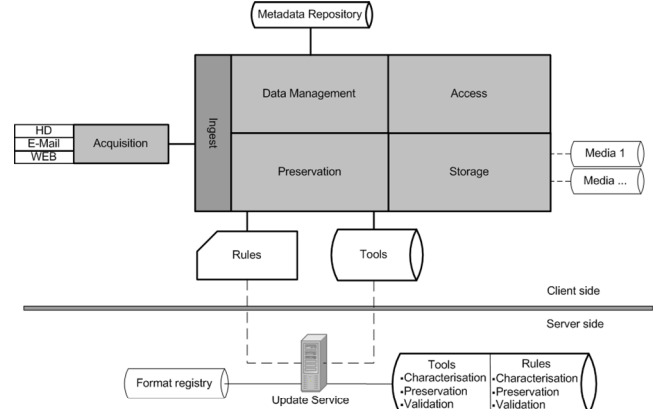
Requirements & Challenges

- Limited or no expertise in digital preservation and IT
- Basic IT-infrastructure
- Full automation of archiving process
- Outsourcing of preservation expertise
- Reliable software solution
- Support of bitlevel and logical preservation and versioning
- Flexible handling of differing individual requirements
- Support for variety of formats and storage media

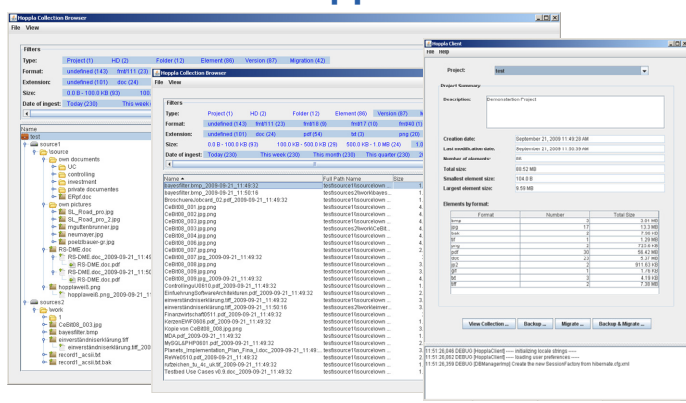
Archiving System

- Client-Server Architecture
- Modular design
- Service model similar to current firewall and antivirus software
- Web services provide preservation rules and tools
- Support of various source media such as file systems, e-mail, wikis and websites
- Multiple backups on storage media (SAN, RAID, ext. HDD, DVD, online storage)
- Compliance to international standards and certification initiatives (OAIS, TRAC, nestor)

Architecture



Hoppla



Application

- Java based development
- Small entry barriers, hiding complexity of preservation
- Adaption and integration of existing modules, for example DROID for format identification, JHOVE for characterisation
- Web service
 - Rule service for characterization, preservation action and migration validation
 - Tool and Service Registry providing portable tools
- Action on objects are performed on the client side
- XML files are stored on backup media describing the collection and the content

Outlook

- Refinement of preservation recommendations for specific user groups
- Heuristics for selection of digital material
- Plug-ins for different source and storage media
- Enhanced access and search functionality
- Logging and compliance checking