Personal & SOHO Archiving

Stephan Strodl
Vienna University of Technology

JCDL 2008, June 17
Motivation

- Increasing amount of electronic material is stored and organised on home PCs and internet services
- SOHOs manage their business by using PCs
- Data have high value (economic and emotional)
- Common practice backups on external devices (CD, DVD, HD)
- Missing long term solution & awareness of digital preservation
State of the art

- Number of tools for backups
- Managing personal information
  - MyLifeBits
  - Apple Time Machine
  - Microsoft Home Server

- Digital Preservation important issue for heritage institutions and large businesses
- Number of research projects worldwide
- Services and models are developed for preservation in professional settings (e.g. OAIS, Fedora, DSpace)
Home & SOHO Archiving vs. Institutional Archiving

- Environment (hardware, infrastructure) & amount of data
- Level of expertise
- Understanding of challenges
- Requirements (e.g. decision making, automation, authenticity)
- Legal obligation
Challenges

- Data are highly distributed (e.g. web services, e-mails, CDs or DVDs from camcorders)
- Variety of formats of different age
- Estimation of future value
- Access
- Confidence and privacy requirements
Personal & SOHO Archiving

- Reduction of the complexity of digital preservation
- Simple and automated preservation services
- Small entry barrier and high degree on automation
- Based on established best practice
- Combination of bit preservation and logical preservation
  (current migration in the SW development)
Personal & SOHO Archiving
Acquisition & Ingest

Acquisition
- Capture data from different sources
- Data and additional information

Ingest
- Selection of digital objects
- Support by characterisation services
- Creation of collection profile
Preservation Management

- Control of logical preservation
- Preservation rules and services provided by Update Service based on collection profile (defined level of detail)
  - Preservation rules
  - Preservation action services
  - Validation services
- Multiple rules for object type
- Validation of migration for correctness and completeness
Update Service

- Provide rules and services to the Personal & SOHO Archiving System
- Critical part
- Selection of preservation rules and services (e.g. Planets Preservation Planning approach)
  - Based on collection profile
  - Consideration of requirements and goals
  - Experience and practise of professional settings
- Watch services for technology
Data Management

- Enrichment of the objects with metadata provided by ingest module and characterisation services
- Documentation of actions in archive system
- Management of versioning of objects
- Metadata repository only used operationally
Access & Storage

Access

- Services to access data in the archive
- Information about dependencies and versioning
- Restore functionality

Storage

- Bitstream preservation
- Multiple copies on various storage media
- Data including all metadata information
Implementation Hoppla Software

- Prototype implementation of Personal & SOHO Archiving
- Acquisition of HD and emails
- Plugin interfaces for acquisition and storage
- Backup and migration rules
- Versioning of objects
- DROID for identification
- Metadata in XML
  (recovery without Hoppla Software)
Implementation Hoppla Software

- All actions (identification, characterisation, preservation and validation) are performed on client side
- Defined level of detail for collection profile
- Potential extension, selection of trusted online services
Outlook

- **Web Update Service**
  - Best practice rules based on collection profile
  - Collection of portable services
    - Action
    - Characterisation
    - Validation

- Support of acquisition sources and storage media

- Heuristics for selection of objects
Conclusions

- Requirements of private users and SOHO
  - Different environment, skills, experience
  - Hide technical complexity, automation
  - Various sources
  - Access & Usability

- Hoppla Software
  - Combination of bit preservation and logical preservation
  - Multiple migration pathways
  - Update Service for rules and tools
Thank you

More information:
www.ifs.tuwien.ac.at/dp
www.planets-project.eu