Co-operative Development of a Long-term Digital Information Archive

iPRES2006 conference, Cornell

Reinhard Altenhöner / Tobias Steinke, DNB
- Long-term preservation: Raising awareness and creating a network of expertise: The nestor initiative
- Pressure comes from a new legal deposit in Germany
- Need for a practical solution for heterogeneous groups of users and data-sets
- Need for a reliable and reusable solution
www.langzeitarchivierung.de

Page Impressions 2004 – 2006 (IVW)
Kopal – Project aims & conditions / background information

- Long-term preservation: Raising awareness and creating a network: The nestor initiative
- Pressure from a new foundation for a legal deposit in Germany
- Need for a practical solution for heterogeneous groups of users and data-sets
- Need for a reliable and reusable solution
Funded by the Federal Ministry for Education and Research


Task: Development of a standardized long-term preservation solution to facilitate long-term preservation for other libraries / industries

Solution should be a facilitator for co-operation between libraries

Cooperation - Standardization - Reusability
kopal: Collaborated work

GWDG: Hosting

IBM: Archiving SW

Deutsche Nationalbibliothek

DNB: Ingest/Acess SW

SUB: Ingest/Acess SW

Common activity: Preservation Planning
Workpackages

DIAS-CORE

Basic development
Basic installation incl. multi-user support
Format support and administrations functionalities

Preservation planning
Planing / Concept
workflow definition
development and integration in DIAS-Core

Local systems

Basic connection
Builder:
universal / individual
Access:
universal / individual

Exemplary data transfer
material and development from Deutsche Nationalbibliothek
material and development from SUB Göttingen
Time schedule: data ingest up to the end of the project

Development data ingest for DNB
(without Web Harvesting)

- Amount
- Total
- Local cache

Needed capacity in TB

<table>
<thead>
<tr>
<th>Time scale</th>
<th>Dissertations</th>
<th>Springer Serials</th>
<th>Springer Artikel</th>
<th>Exilpresse</th>
<th>Digitalisate DBSM</th>
<th>Digitalisate HAB</th>
<th>CD-Roms</th>
<th>DVD-Roms</th>
</tr>
</thead>
<tbody>
<tr>
<td>August 06</td>
<td>0.3</td>
<td>0.22</td>
<td>0.22</td>
<td>0.15</td>
<td>5</td>
<td>1.5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>September 06</td>
<td>0.3</td>
<td></td>
<td>0.25</td>
<td>0.15</td>
<td>5</td>
<td>1.5</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>Oktober 06</td>
<td>0.52</td>
<td></td>
<td>0.25</td>
<td>0.15</td>
<td>5</td>
<td>1.5</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Dezember 06</td>
<td>0.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Januar 07</td>
<td>5.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>März 07</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Juni 07</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oktober 07</td>
<td>11.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Establishing service models for kopal

- We want to provide a modular service in a customer-based perspective
- We offer general consultation, support, services
- And we release know-how (standards), partnership/cooperative use, full service
- Our model must guarantee the refunding for our offers
  - One can keep guidance and aid
  - One can take a partial stewardship
  - One can order a complete system
Technical concepts of kopal

- Global core (DIAS) and local software (koLibRI)
- Open interfaces: Universal Object Format
- Preservation Planning: Migration Manager and File Format Registry
Global core and local software

Deutsche Nationalbibliothek (Frankfurt)
koLibRI (local software)

SUB Göttingen
koLibRI (local software)

GWDG (Göttingen)
DIAS by IBM
Account 1
Account 2
DIAS and koLibRI

**DIAS**
- IBM development for the National Library of the Netherlands
- Conforming to the OAIS reference model
- Based on IBM standard software (Content Manager, TSM, etc.)
- Enhancements for kopal: open interfaces (Universal Object Format), multi-user support, query interface to Data Management

**koLibRI**
- kopal Library for Retrieval and Ingest
- Open source development by DNB and SUB Göttingen
- Tools to generate technical metadata (using JHOVE), create ingest packages and migrate objects
Universal Object Format

- Generic description of packages for archiving and exchanging of digital objects
- Based on METS (special METS profile) and LMER (Long-term preservation Metadata for Electronic Resources)
- One packed file (e. g. ZIP or tar) per logical object, containing one METS file with descriptive metadata (e. g. Dublin Core) and technical metadata for long-term preservation (LMER)
- In kopal used as SIP and DIP (OAIS); metadata (Dublin Core and LMER) is stored in Data Management of DIAS
Preservation Planning

- DIAS: Query interface to Data Management (technical metadata and optional Dublin Core)
- koLibRI: Migration Manager to perform queries, call conversion tools and re-ingest migrated objects
- Planned support for a File Format Registry: Database about file format characteristics, view paths, migration paths and emulation environments
Visit [http://kopal.langzeitarchivierung.de](http://kopal.langzeitarchivierung.de)

- Download of koLibRI (beta)
- kopal demonstrator (Flash animation)
- Specifications of the Universal Object Format