Digital preservation activities at the German National Library
nestor / kopal

Entretiens de la BnF, Paris Dec. 8th, 2006
Our digital heritage consists of …

… text information such as books, journals, newspapers etc.

… still images or films

… audio materials

… databases (static or interactive)

… primary data facts and figures etc.

… files

… e-learning services

… e-museums, electronic works of art

… and a lot more
Institutions dealing with digital resources

- libraries
- archives
- museums
- scientific institutions
- commercial enterprises
- ICT companies
- etc.
nestor
Kompetenznetzwerk
Langzeitarchivierung und
Langzeitverfügbarkeit Digitaler Ressourcen für Deutschland

Network of Expertise in Long-Term Storage of Digital Resources
Goals of nestor (1)

- create a network for information and communication about present and future LTP activities in Germany
- establish a cross-sectoral community to promote and support LTP activities and to raise awareness in society
- trigger synergies between on-going activities in Germany and cooperate with international partners and projects
Goals of nestor (2)

- establish a permanent organisation which coordinates and represents the concerns of long-term preservation
Network of Expertise in Long-Term Storage of Digital Resources

The project's objective is to create a network of expertise in long-term storage of digital resources for Germany. As the perspective of current and future archive users is central to the project, the emphasis is put on long-term accessibility. Within the project the following offers will be created: a web-based information forum, a platform for information and communication, criteria for trusted digital repositories, recommendations for certification procedures of digital repositories, recommendations for collecting guidelines and selection criteria of digital resources to be archived, guidelines and policies, the concept for a permanent organisation form of the network of expertise in digital preservation. The long-term goal is a permanent distributed infrastructure for long-term preservation and long-term accessibility of digital resources in Germany comparable e.g. to the Digital Preservation Coalition in the UK.

European Conference on the Preservation of Digital Resources, April 20-21, 2007 in Frankfurt

To coincide with the German EU Presidency, the Commissioner for Cultural and Media Affairs of the German Federal Government, and the Deutsche Nationalbibliothek, in conjunction with the partners of nestor - Network Expertise in Long-Term Storage - are extending an invitation to experts and all interested parties in the European member states to come to the German National Library in Frankfurt am Main to discuss current issues and assess the prospects for long-term digital preservation.

Further information you will find here -->> [German] [English] [French]

October Issue of RLG DigiNews online

The newest issue of RLG DigiNews is published. It contains an article by Kevin Gilik, Eliet Wilczek and Robert Docking about "Fedora and the Preservation of University Records Project" on a second ... [more]

Posted by: Stefan Strathmann (2006-11-16)

2007 Digital Preservation Award - call for entries

The Digital Preservation Coalition has just announced the call for entries for the third Digital Preservation Award. The award will be presented at the British Museum on the 27th of September 2007. ... [more]

Posted by: Stefan Strathmann (2006-11-09)

Web Curator Tool released

The National Library of New Zealand and The British Library have recently announced the...
Establish a network of competence

- infrastructure to address and involve institutions, organisations, companies and individuals and to make their competence available to others
- facilitate communication between interested parties
- initiate necessary projects
- implement discussion lists
- organise workshops
- mechanisms for surveys, statistics
nestor's success

- importance of long-term preservation efforts is in the minds of all stakeholders and players in the field
- www.langzeitarchivierung.de is accepted as a specialised LTP "focus", ready for international cooperation
- LTP development projects benefit from the community building efforts of nestor
- a durable business model for a cross-sectoral competence network community will be established
Sustainability of LTP activities

- information and communication between experts and institutions
- public relations and awareness building
- education and training
- initiation of technical and organisational projects
- standardization
- international cooperation
German national initiative „kopal“

- Co-operative development of a long-term digital information archive
- 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 as a facilitator for co-operation between libraries and other institutions / companies
koLibRI Retrieval Component

Selection

Collection

User

Presentation components

XML + Data

koLibRI Ingest Component

Metadata Extraktion

Metadata Generation (JHOVE)

UOF Creation (SIP with METS)

koLibRI Retrieval Component

Cache

Deutsche Nationalbibliothek

XML + Data

koLibRI

Ingest Component

koLibRI

Ingest

Preservation

Data Manag.

Access

Archival Storage

Admin

(UOF (SIP))

(UOF (DIP))

DIAS

(OAIS Compliant)
Kopal preservation strategy

- **Migration strategies**
  - Migrate object with a specific identifier into a new format
  - Migrate all objects
    - of format xxx and/or
    - that have been ingested before a certain date and/or
    - that are larger than xxx MB

  into a new format (e.g. from TIFF to PNG)

- Implementation of **emulation** view paths

- **No restriction** as of file size or file format / type – all known and unknown file formats are being accepted (text, pictures, video, audio, executables, ... etc.)
Present ingest

- Productive system was installed and made available to SUB and DNB in June 2006
- Several tests conducted
- Productive ingests of all DNB theses & dissertations started early August 2006
  - About 40,000 dissertations processed
  - Everything ingested to DIAS was processed correctly
  - koLibRI software (kopal Library for Retrieval and Ingest) works as designed
- Next to process:
  - Electronic journals & serials from Springer publisher
  - Exemplary number of CD-ROM's / DVD's as (ISO) images
  - Exemplary digitised materials from DNB collections
Data ingest for kopal with ETD's as start

Planned Ingest for kopal
(except WEB harvesting)

- Alternative Ingest
- Planned Ingest
- Local Cache (1.4 TB)
- Pre-Production (6 TB)
- Productive System (185 TB)

Ingest in Terabytes (TB)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.35</td>
<td>0.22</td>
<td>0.25</td>
<td>0.15</td>
<td>5</td>
<td>1.5</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>0.35</td>
<td>0.22</td>
<td>0.25</td>
<td>0.15</td>
<td>5</td>
<td>1.5</td>
<td>28.6</td>
<td>9.6</td>
<td>545</td>
</tr>
</tbody>
</table>

Timescale

01.06.2006 01.07.2006 01.08.2006 01.11.2006 01.02.2007 01.01.2007 01.04.2007 01.06.2007 01.10.2007
Challenge: Preservation Planning + Access

- In face of rising data amounts and large single objects (e.g. digitised DVD-ROM images with ~8 GB):
  - Guarantee a sufficient performance of the system
  - Implementation of suitable access systems
  - Fast Internet connections, user support

- Implementation of a functioning Preservation Planning mechanism
  - Functioning international File Format Registry
  - Performant migration of large data amounts
  - Successful implementation of emulation mechanisms

- Information, support & encouragement of ETD producers towards a format & preservation awareness
Thank you for your attention

Ute Schwens
mailto: u.schwens@d-nb.de