Long-term preservation activities of the Bavarian State Library

Latest challenges and developments
The Bavarian State Library
Some facts & figures

- 680 employees
- ~ 10 million volumes
- ~ 150,000 incoming volumes p.a.
- more than 40,000 current periodical titles
- 15,687 current e-journals (licensed), 600 databases
- 450,000 e-books (including digitised material)
Long-term Preservation: Three intersecting Spheres of Activity

Routine operation of the existing infrastructure
- BABS Library Archiving and Access System

Further development and consolidation of the required structures
- Profile and concepts, workflows
- Organisation
- Technology
- Rosetta Project (to be mentioned later on)
  → Technical support + library-related business

Research and development
- Evaluating future technologies through innovative projects
Long-term Preservation Milestones

- **1999** First Long-term preservation project

- **2004** Switch from CD to tape storage (Leibniz Supercomputing Centre)

- **2005** Implementation of the **BABS** Library Archiving and Access System

- **2006** First successful hardware migration
Lots of Digitisation Activities

Maps 16th Century Books
Newspapers Biographic 100(0) Key documents on German History
Literary & Scientific Remains Latin Manuscritps Incunabula Digi20 German Manuscripts Exhibition Catalogs Monumenta Germaniae Historica Reichstag Session Protocols
100(0) Key Documents on Russian History Encyclopaedia

Google Music Supplies Digitisation on Demand Block Books Incunabula Journals on Bavarian History Bavarica
Lots of Legal Deposit Activities

eBooks Biographic Databases Online
Dissertations Sound Storage Media Scientific Databases Digitised Maps Subject Portals Bavarian Official Governmental Gazettes Electronic Newspapers Scientific Journals Photographic Collections Web Resources Literary & Scientific Remains Digitised Books
Major Stakeholders of Long-term Preservation

Electronic Media Department
Centre For Electronic Publishing
Munich Digitisation Centre
BSB’s special collection fields
Bavarian Library Network
Bavarian Publishing houses
Public authorities
Present Responsibilities

Digital Objects from the (Mass-)Digitisation Programme
- ~1.3 mio. volumes, >300 mio. pages

Born-digital materials
- Electronic journals: ~15,687 titles (≠ articles!)
- Databases: ~600
- E-books: ~8,200

- Legal Deposit of Public authorities (legal obligation since 01/2009),
- Deposit of regional publishers (voluntary)
- BSB‘s special collections fields
Long-term Preservation: Major Activities

BABS Library Archiving and Access System
- Concept and Setup (2005-2007)
- Self-evaluation and improvement in terms of trustworthiness and scalability (2009-2010)

Nestor
- German Network of expertise for the long-term preservation (since 2003)

Scientific Studies, e.g.
- Organisational and business models for the long-term preservation of digitised materials

LOCKSS
- Worldwide pilot project together with LoC and other partners
LRZ: The backbone of BSB 's LTP strategy

The Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities

- IT-services for the research institutions in Munich
- Munich Scientific Network
- backup and long-term preservation
- performance Supercomputing Centre for all German universities
Status Quo 2010

Several large scale digitisation projects

Intensified legal deposit activities

→ Growing electronic archive
  ▪ 294 M Files
  ▪ 181 TB
  ▪ +10 TB/month
File Formats and Metadata in LTP

File Formats
PDF, PDF/A, TIFF, JPEG 2000, XML

Metadata
- Bibliographical
- Technical
- Preservation
- Rights
How it is done: Existing Architecture
Rosetta at the Bavarian State Library

A new Digital Preservation System for the Bavarian State Library

- improving existing preservation workflows
- implementing additional workflows
- intensifying preservation planning activities
Why Rosetta?

Comprehensive and specific functionality for all key areas of digital preservation:

- **Flexible configuration**, parameterisation and automation for different submission and ingest workflows / for internal and external stakeholders

- **Digital Asset Management**: checks for integrity / authenticity, format conversion, generation of metadata
Why Rosetta?

- Preservation planning and action: risk analysis and risk management, planning and implementation of preservation action (*with V. 2.0*)
  
- Support of standards (OAIS, METS, Dublin Core, PREMIS, OAI-PMH)

- **Open Platform**: APIs and software development kit support integration of additional software

- Easy exchange/exports of Archival Information Package (AIP)
Origins of Rosetta

Joint development of a digital preservation system between the National Library of New Zealand and ExLibris

- **2004** initiation of the National Digital Heritage Archive Program in New Zealand
- **2005/2006** substantial gap analysis for existing products by NDHA → detailed requirements specification
- **July 2007** decision to develop of a commercial solution together with ExLibris (2007-2010)
- **Nov. 2008** DPS Version 1.0 goes live at the National Library of New Zealand
- **Jan. 2009** General release of DPS (Version 1.0) as *Rosetta* and start of Charter Group Program
Rosetta at the Bavarian State Library: Partners

- Munich Digitisation Centre (Digital Library Dept. of the BSB)
- Head Office of the Bavarian Library Network at the BSB
- Leibniz Supercomputing Centre Munich (LRZ)
- Selected partner libraries out of the Bavarian Library Network (to be named)
- ExLibris
Integration into the existing software architecture

- Web site
- Web site
- Web site
- WCT (web curator tool)
- ExLibris Preservation
- ZEND
- Legal and voluntary deposit; Licensed media
- Digitization projects
- Digitization on Demand
- Google Books
- Journal loader
- Journal loader
- Journal loader
- LRZ
- Permanent Repository
- Deposit SIP
- AIP
- Publishing Delivery
- Search Tools
- Library Catalog
- Content aggregator

BSB Bayerische StaatsBibliothek
Information in einer Linie
Rosetta: Development Roadmap

- Q1/2010: General release of version 2.0
- 2010: Planned start of general migration from DigiTool to Rosetta’s Digital Asset Management Module
- 2010/2011: Beta-release of Version 3.0
Implementing Rosetta at the BSB: Roadmap

A. Pilot stage 2009/2010

- Specification, implementation, testing and consolidation of a pilot-system for 2-3 designated workflows / media types (BSB, LRZ, ExLibris)
- Integration into the existing software infrastructure at BSB & customizing according to BSB‘s requests
- Training for administrators and disseminators
- Participation in the Rosetta Charter Group for general enhancement of the product
- Documentation, review and evaluation
Implementing Rosetta at the BSB: Roadmap

B. Second stage

- Starting of routine operation
- Implementing additional workflows at BSB
- Adopting existing workflows
- Migrating existing data
- Providing a platform for the data of other Bavarian libraries
Long-term Preservation at the BSB
Overall Objectives

- To be a trusted partner in a cooperative solution for the preservation of the digital heritage of Germany
- To preserve digital media from different sources
- To keep everything accessible and available (= No dark archive)
Thanks for coming & watching!

thomas.wolf-klostermann@bsb-muenchen.de