

Digital preservation in the Die Deutsche Bibliothek and cooperative initiatives in Germany

**Preservation of digital heritage: Basic concepts and main initiatives
CONFERENCE, Madrid, 14th-16th March
2006**

Reinhard Altenhöner
Die Deutsche Bibliothek



Overall agenda

- The national library of Germany and its setting
- The challenge: Long term preservation of our digital heritage – introduction (very short)
- The nestor initiative – a network of competence for LTP
- Different projects in Germany (examples)
- The kopal initiative – our technical approach for LTP

Political system of the Federal Republic of Germany





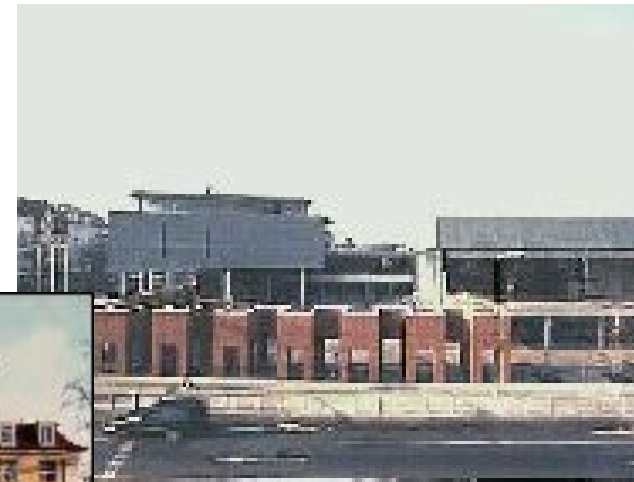
Legal deposit in Germany

- On the national level: DDB
- On the “Länder”-level ca. 35 libraries, mostly financed by the Länder, often in combination with other types of libraries (e.g. university libraries)
- On the communal level: A lot of public libraries have similar tasks for their town or region, often they share this task with the local archive



Die Deutsche Bibliothek

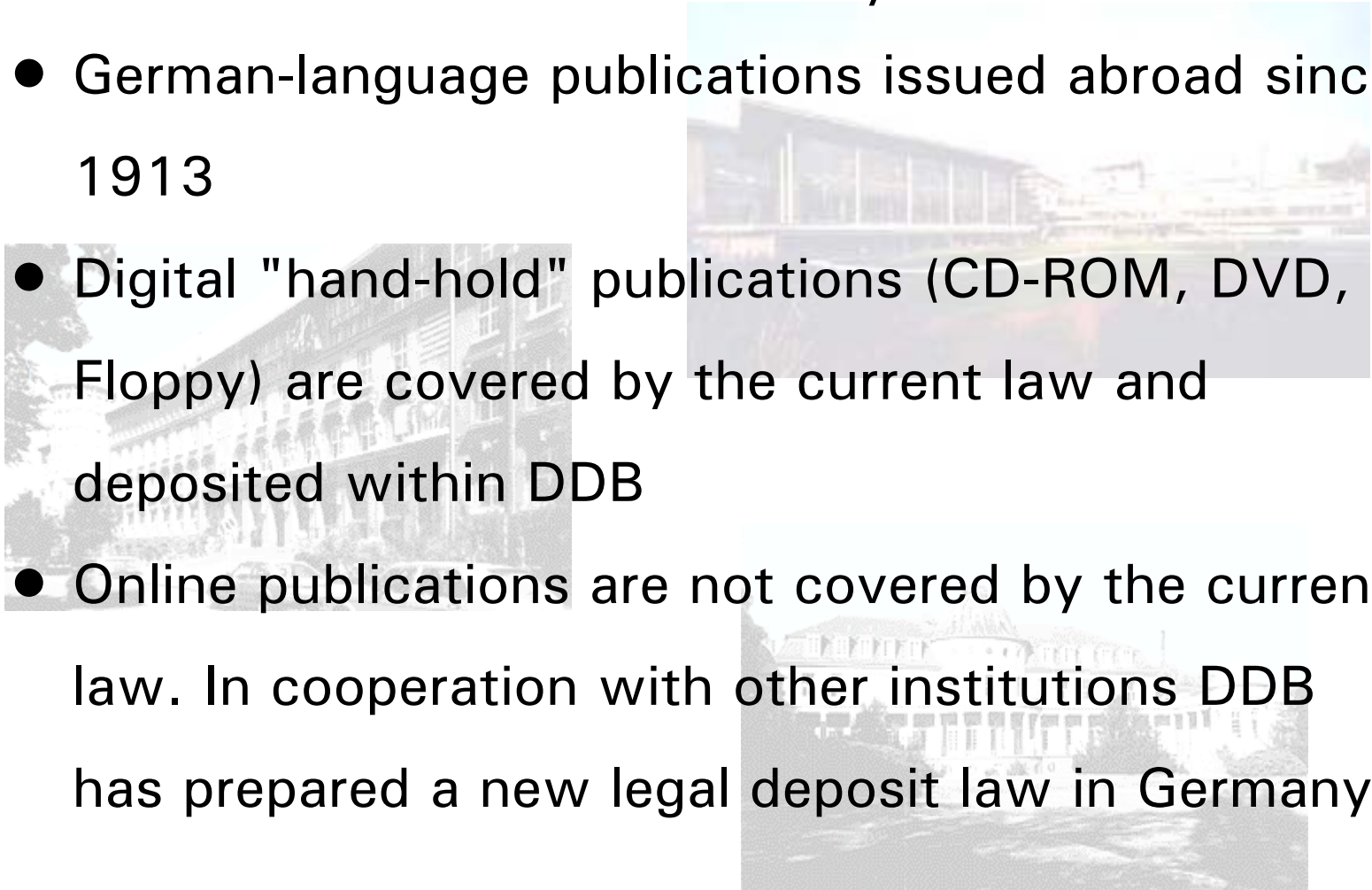
The National Library of Germany



ddb

Our task: Collecting and archiving

- Publications issued in Germany since 1913
- German-language publications issued abroad since 1913
- Digital "hand-hold" publications (CD-ROM, DVD, Floppy) are covered by the current law and deposited within DDB
- Online publications are not covered by the current law. In cooperation with other institutions DDB has prepared a new legal deposit law in Germany



DDB: Functions and services

- German National Bibliography
- Other bibliographic services and data (ZDB)
- Authority files
- Central German authority and co-ordination agency for bibliographic standards
- Represents the German library community at the national and international levels

Items	22.500.000
GNB	240.000 p.a.
Staff	About 760
Budget	38,5 Mil €
Orders	1.000.000 p.a.

Same procedure as every day?



© KB, The Hague

Electronic archives at Die Deutsche Bibliothek

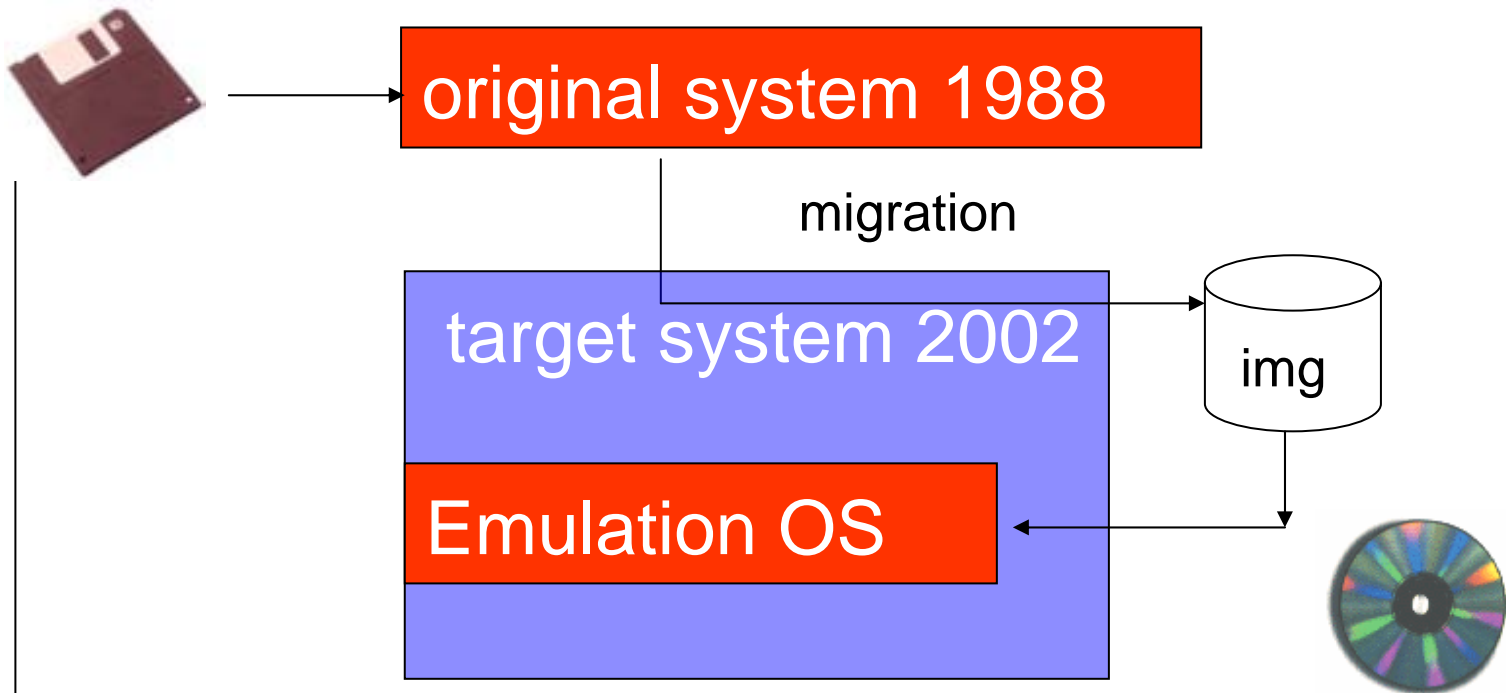
- ~ 600 TByte on physical media

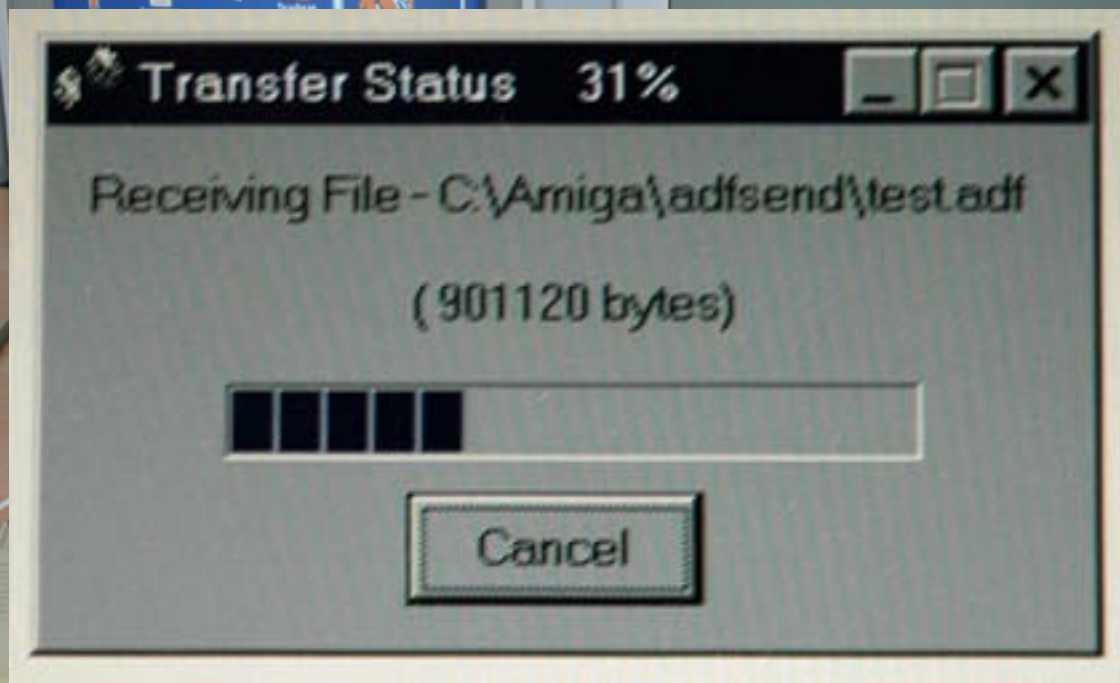
E/net-Publications:

- 43.000 Online-Dissertations
- 454 e-journals and 1.700 monographs from Springer (Heidelberg, Berlin)
(<http://link.springer.de>)
- 60 newsletters (currently)
- 4.500 electronic publications from approx. 210 commercial and non-commercial publishers
- Web-Sites
- → workflows, modules, procedures

Transfer of data from an old platform

eg. C64 / AMIGA





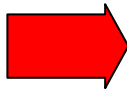


Persistent access via Uniform Resource Name (URN)

Titel: Nitric oxide in the olfactory epithelium [Elektronische Ressource] / von Oliver Schmachtenberg
Verfasser: [Schmachtenberg, Oliver](#)
Erscheinungsjahr: 2001
Hochschulschrift: Hannover, Univ., Diss., 2001
Persistent Identifier: [urn:nbn:de:gbv:089-3321752945](http://nbn-resolving.org/urn:nbn:de:gbv:089-3321752945)
URL: [Archivserver Der Deutschen Bibliothek](#)
Sachgruppe: 33 Medizin ; 32 Biologie



<http://nbn-resolving.de/urn/resolver.pl?urn=urn:nbn:de:gbv:089-3321752945>



<http://edok01.tib.uni-hannover.de/edoks/e01dh01/332175294.p>

Nitric Oxide in the Olfactory Epithelium

Von dem Fachbereich Biologie der Universität Hannover

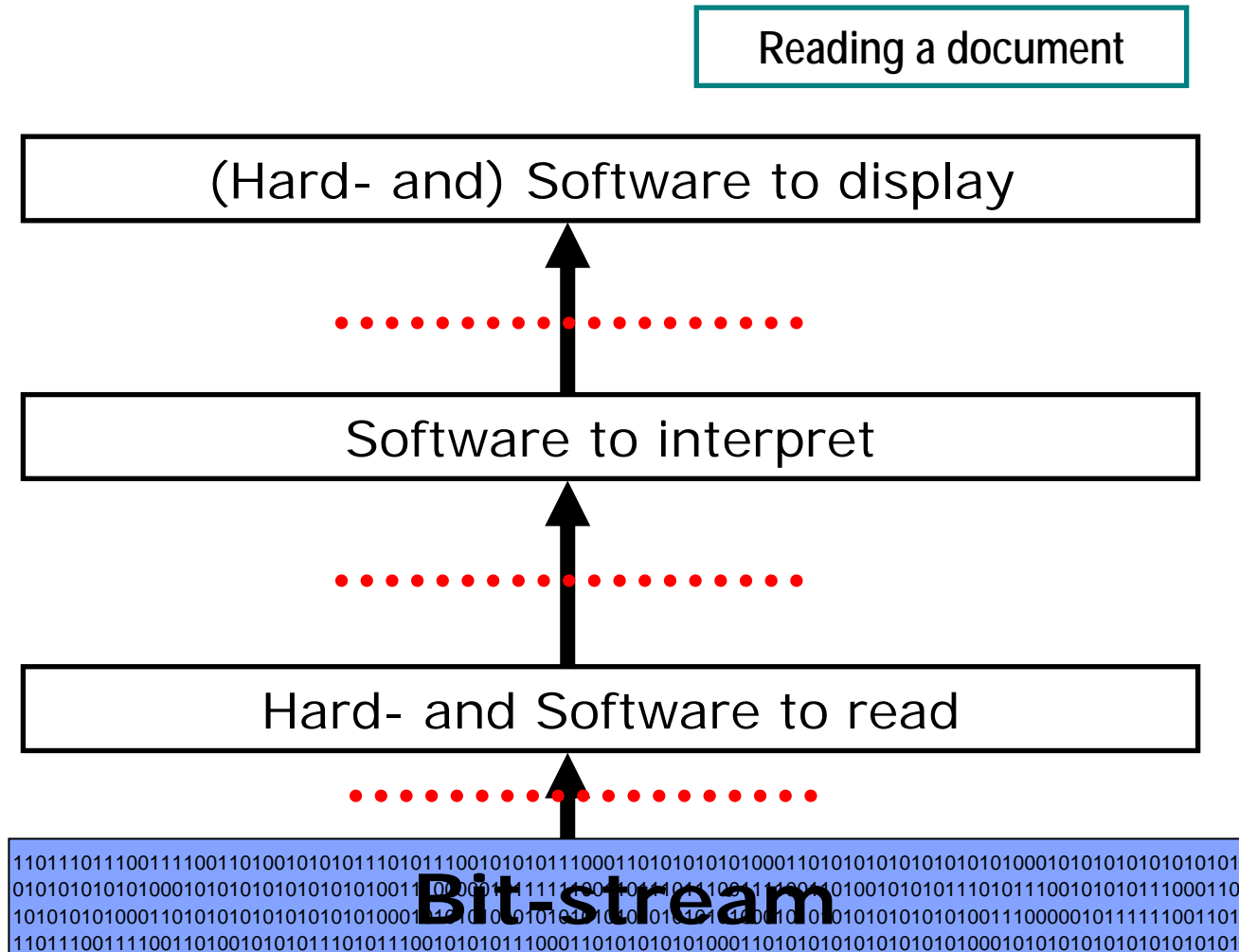
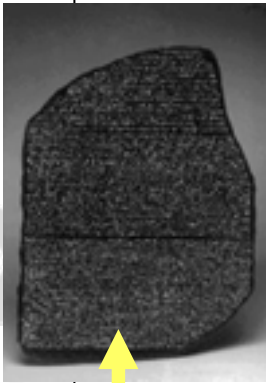
zur Erlangung des Grades eines

Doktors der Naturwissenschaften

Dr. rer. nat.

genehmigte Dissertation

The challenge of LTP





National initiatives

- nestor

Network of Expertise in Long-Term Storage of Digital Resources for Germany



- kopal

Co-operative development of a long-term digital information archive



ddb

nestor

**Network of Expertise in Long-Term
Storage of Digital Resources**

Kompetenznetzwerk Langzeitarchivierung
und Langzeitverfügbarkeit Digitaler
Ressourcen für Deutschland

nestor 2003 - 2006

- Die Deutsche Bibliothek (lead)
- Bavarian State Library
- State and University Library Goettingen
- Computer and Media Service of Humboldt University, Berlin
- Bavarian State Archive – Head Office
- Institute for Museum Studies, Berlin
- Federal State Archives

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Die Deutsche Bibliothek

BSB Bayerische
StaatsBibliothek



S M
B Institut für Museumskunde



Das
Bundesarchiv

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Goals of nestor

- 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
- establish a permanent organisation which coordinates and represents the concerns of long-term storage

Information and communication

- www.digitalpreservation.de - cumulation and integration of German language information resources on long-term preservation
- bridge to information resources in other languages (PADI)
- newsletter
- reviews
- calendar

Kompetenznetzwerk Langzeitarchivierung

nestor network

- Home
- Project Information
- Studies
- Guides
- Events
- Workgroups
- Glossary

Functions

- Settings
- Logout

Information

- Subject Gateway
- Who Where What
- What's new
- Calendar
- Projects
- Newsletter

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.

Project kopal implements Universal Object Format

Standardisation is a critical factor for success in long-term preservation of digital objects. Object format, transfer protocols and system interfaces need standardisation in order to cope with large ... [\[more\]](#)

Posted by: Hans Liegmann (2006-02-17)

Invitation to nestor-discussion on a German Policy for long-term preservation of digital resources

It belongs to the tasks of the project nestor - network of expertise in long-term storage to formulate recommendations regarding the long-term preservation of digital resources in Germany and bring ... [\[more\]](#)

Posted by: Philipp Höhler (2006-02-17)

Calendar

ECURE 2006

From 2006-02-27 to 2006-03-01 in Tempe, Arizona, United States

95th German Librarian Congress 2006: "Network Library"

From 2006-03-21 to 2006-03-24 in Dresden

Museums and the Web 2006 - the International Conference for Culture and Heritage On-Line

From 2006-03-22 to 2006-03-25 in Albuquerque, New Mexico, USA

more dates in [calendar](#)

nestor Mailing list

Name

Email

[Go](#)

Public relations and awareness building

- professional PR and media campaign
- presence at events like book fair, DMS Expo, ARCHIVISTICA
- presentations at conferences, workshops
- dissemination of information via articles, interviews

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
- implement discussion lists
- organise workshops
- mechanisms for surveys, statistics

Promote LTP strategies and models

- „hub“ for German LTP project information
- transfer knowledge about LTP activities to and from Germany:
 - kopal, ARCHISIG, SUNCoE, ARCHE...
 - ← DPC, NDIIP, PANDORA, PRONOM, ...
- interconnect information about LTP projects - between German and international level

Trusted Repositories - Certification

- nestor Working Group since 12/2004
- based on existing international efforts
- discuss requirements and specifics of digital repositories
- define evaluation criteria for digital repositories
- develop and implement a procedure for certification

nestor studies 1



[Legal framework for
long-term
preservation]
2004

nestor studies 2



[Digitization
and preservation
in German
museums]

2005

nestor studies 3



[Evaluation criteria
for
state-of-the-art
archiving systems]
2005

nestor studies 4



[Long-term
Preservation
of Online Publications
(E-Journals)]
2005

nestor studies 5



[Long-term
Preservation
of Multimedia
Objects]
2006

nestor studies 6



[Long-term
Preservation
of Scientific Data
Sets]

2006

nestor studies 7



[Development of a
National Long-term
Preservation
Policy]
2006

nestor guides 1



[Not in perpetuity -
little guidebook for
digital data
preservation in
museums]

2005

Training activities

- 1st nestor Seminar "Introduction into Long-term Preservation" (12/2005)
- 2nd nestor Seminar "Archives and Long-Term Preservation" (01/2006)
- nestor handbook

Recommendations on Long-term Preservation of Digital Information in Germany

- Task of national relevance
- financing by public budgets
- co-ordination of responsibilities
- need for sustainable infrastructure
- coordinated selection of objects
- including all types of digital objects
- integrated access to digital and non-digital objects
- integration of digitisation activities
- promotion of open, non-proprietary formats
- foster standardisation
- professional training

nestor's Vision

- importance of long-term preservation efforts is in the minds of all stakeholders and players in the field
- www.digitalpreservation.de 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



Different projects in Germany 1: CASHMERE-int

Objectives:

- to produce active participation in selected areas of developing standards in the context of the Semantic Web
- to transmit these standards to universities in Germany.
- focal point: participation in the developing of preservation standards. (METS, PREMIS working group)

<http://www.iwi-iuk.org/cashmere/index.en.shtml>



Different projects in Germany 2:

Database Supported Long-term Archiving of Digital Objects

Objectives:

- Explore basics and techniques to achieve archiving systems which are more integrated, open, and flexible.
- Analyse the potential of database systems for the implementation of archiving systems
- basic issues for modeling representation information (this kind of information is necessary for interpreting archived bitstreams)
- A framework for the classification of archiving scenarios provides the theoretical basis

<http://ist.unibw-muenchen.de/Inst2/Research/LZA/index.html>

Different projects in Germany 3: Functional long-time-archiving



Objectives:

- develop a functional model of emulation (emulator in a emulator system) of software and hardware environments, which enables a web-based access to the emulators
- By now it offers a prototype of a web-based accessible emulation environment, which offers emulations of commodore 64, ZX Spectrum, Atari 800, a DOS-Box and some computer games

<http://www.ks.uni-freiburg.de/fla/>



Different projects in Germany 4: ARCHE: ArchiveLaser Project

Objectives:

- new perspectives in accurate and safe long-term storage by using Laser recording technology of digital data on color microfilm offers.
- Development of a novel laser recording system – the so called ArchiveLaser
- Embedding the ArchiveLaser technology in a versatile workflow, the digital and analog world is combined into a long-term storage concept.

<http://www.ub.uni-stuttgart.de/wirueberuns/projekte/arche/>

Different projects in Germany 5: ArchiSig & TransiDoc

Objectives:

- allow secure and conclusive long-term archiving of digitally generated and signed data for 30 years or more.
- development of system architectures with new technical components and organizational concepts to guarantee the security of digital signatures. Continued by „ArchiSafe“
- TransiDoc – Legally Secure Transformations of Signed Documents
- Based on use cases of the application domains "public administration", "health care", and "notaries"
- analyze problems, develop proposals, implement and evaluate prototypes

<http://www.archisig.de/>

<http://www.transidoc.de>



- Co-operative development of a long-term digital information archive
- funded by the Federal Ministry for Education and Research
- Financial volume: 4,2 Mio € + self-financed activities of all partners, duration: 1.7.2004 – 30.9.2007
- 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

- Basis: DIAS (Digital Information and Archiving System) of the Royal Dutch Library
 - Developed by IBM → reliable (hopefully)
 - Implementation of the OAIS standard
 - Further development of a suitable long-term preservation component (emulation, migration)
- Enhancement for cooperative usage
- Development of a universal object scheme
- Hosting outside the library (remote access)
- Extension of DIAS-Core with peripheral open-source based software tools to broaden its usability
- Development of mass routines for digital objects
- Platform for preservation planning activities

- Die Deutsche Bibliothek (leader)
 - Objects: Electronic theses, eJournals, CD-ROM's, scanned books, eBooks, music files, web pages
 - Role: Project leader, user, software development
- Staats- und Universitätsbibliothek Göttingen
 - Objects: Electronic theses, eJournals, scanned books, videos
 - Role: User, software development
- Industrial Business Machines (IBM) Germany
 - Role: Core software development
- Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen (GWDG)
 - Role: Service provider, hosting
- Working relationship: Royal Dutch Library, The Netherlands



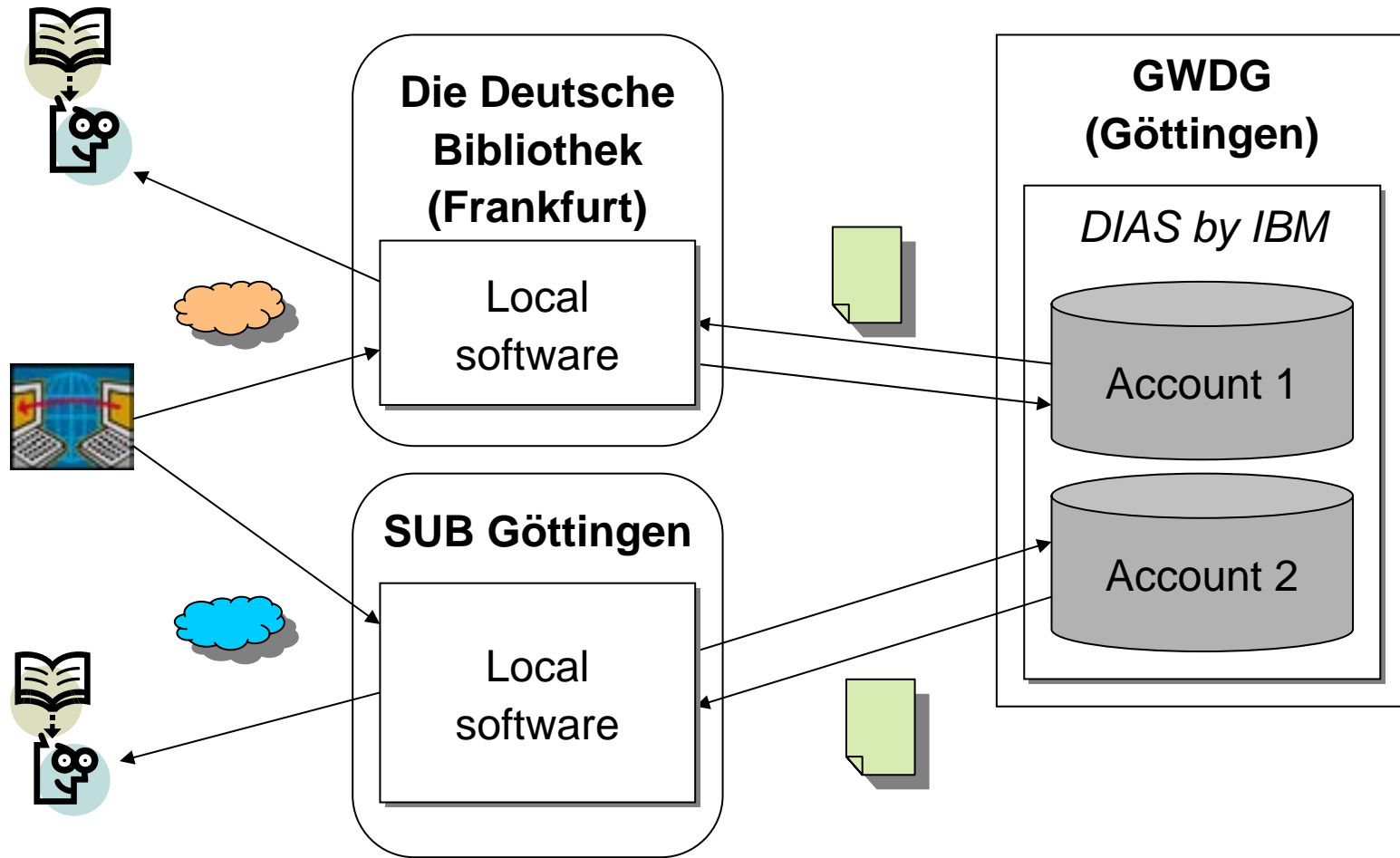
GWDG: Hosting

IBM: Archiving SW

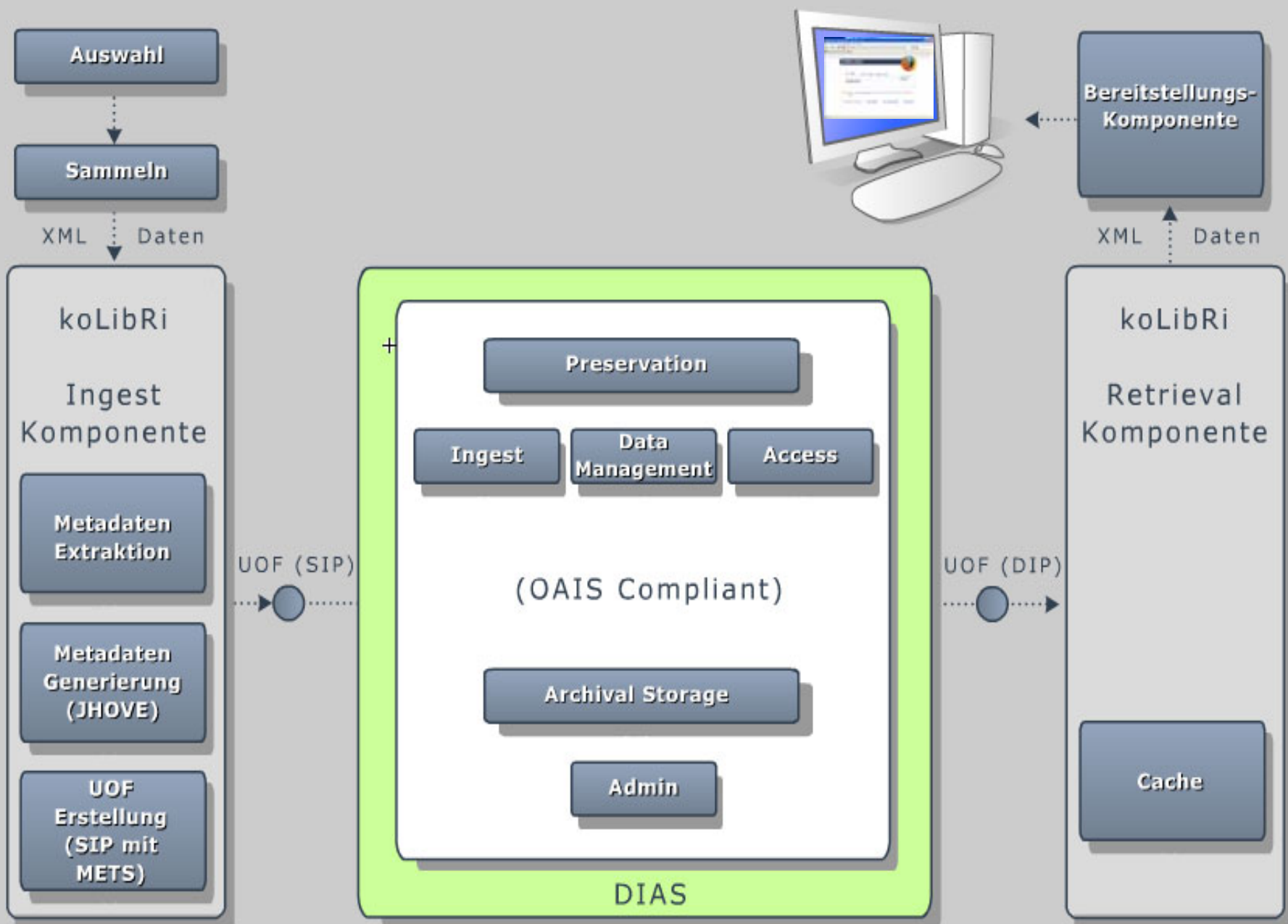
DDB: Ingest/Acess SW

SUB: Ingest/Acess SW

Common activity: Preservation Planning



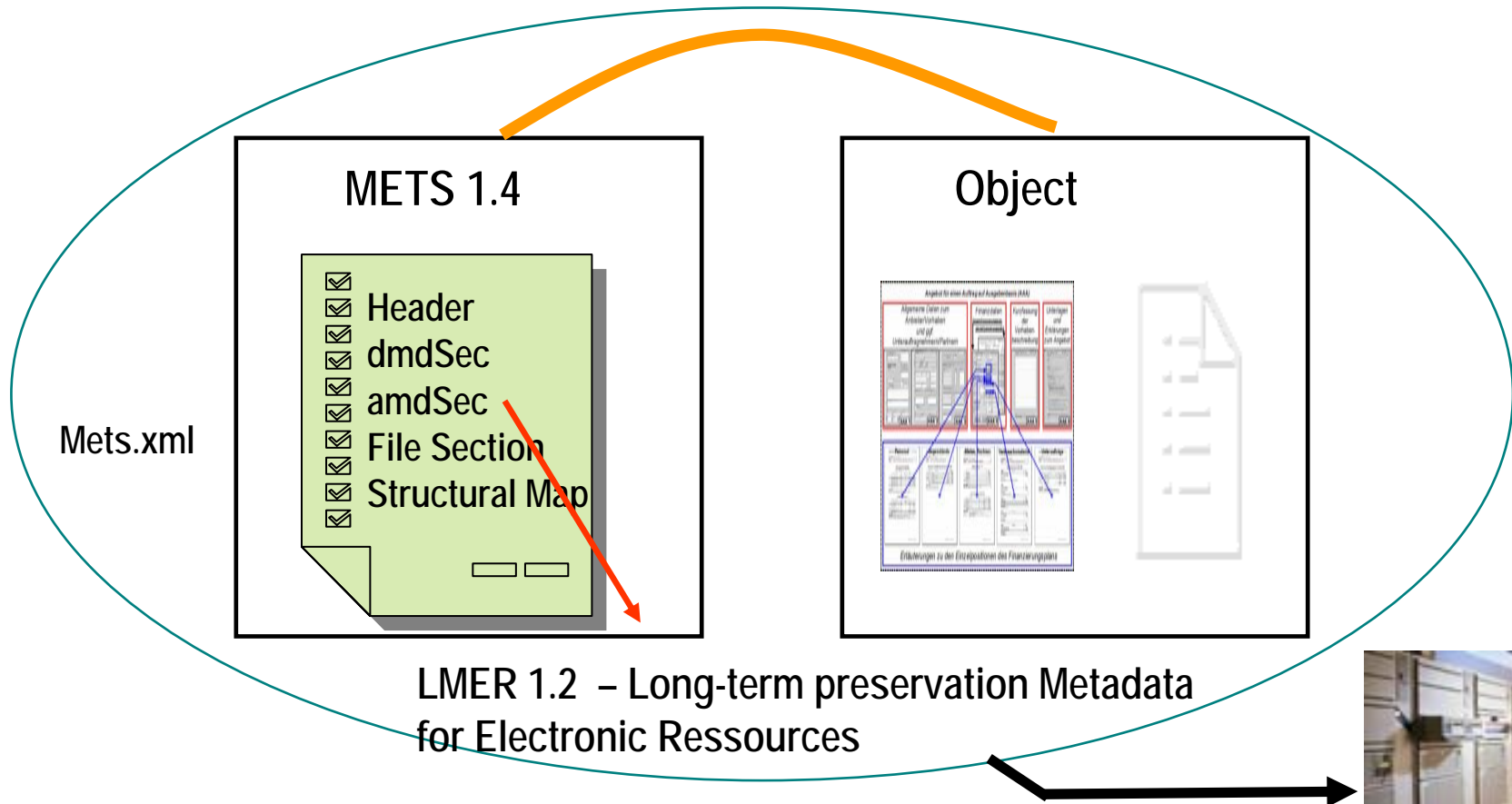
Partners nn



- Basis: OAIS
- Technologies: Java, XML, HTTP, IBM standard products (DB2, Content Manager, TSM, WebSphere)
- (Technical) Metadata: METS (Metadata Encoding and Transmission Standard) as a wrapper, LMER (Long-term preservation Metadata for Electronic Resources)
- Specification for:
 - Submission Information Package (SIP): Ingest
 - Archival Information Package (AIP): Archive
 - Dissemination Information Package (DIP): Access
- Packages as ZIP, GZ, TAR.GZ, TAR
- File-structure is open: (mets.xml file on root-level)

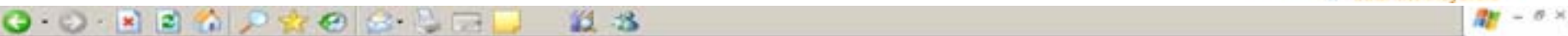
UniversalObjectFormat

Submission Information Package

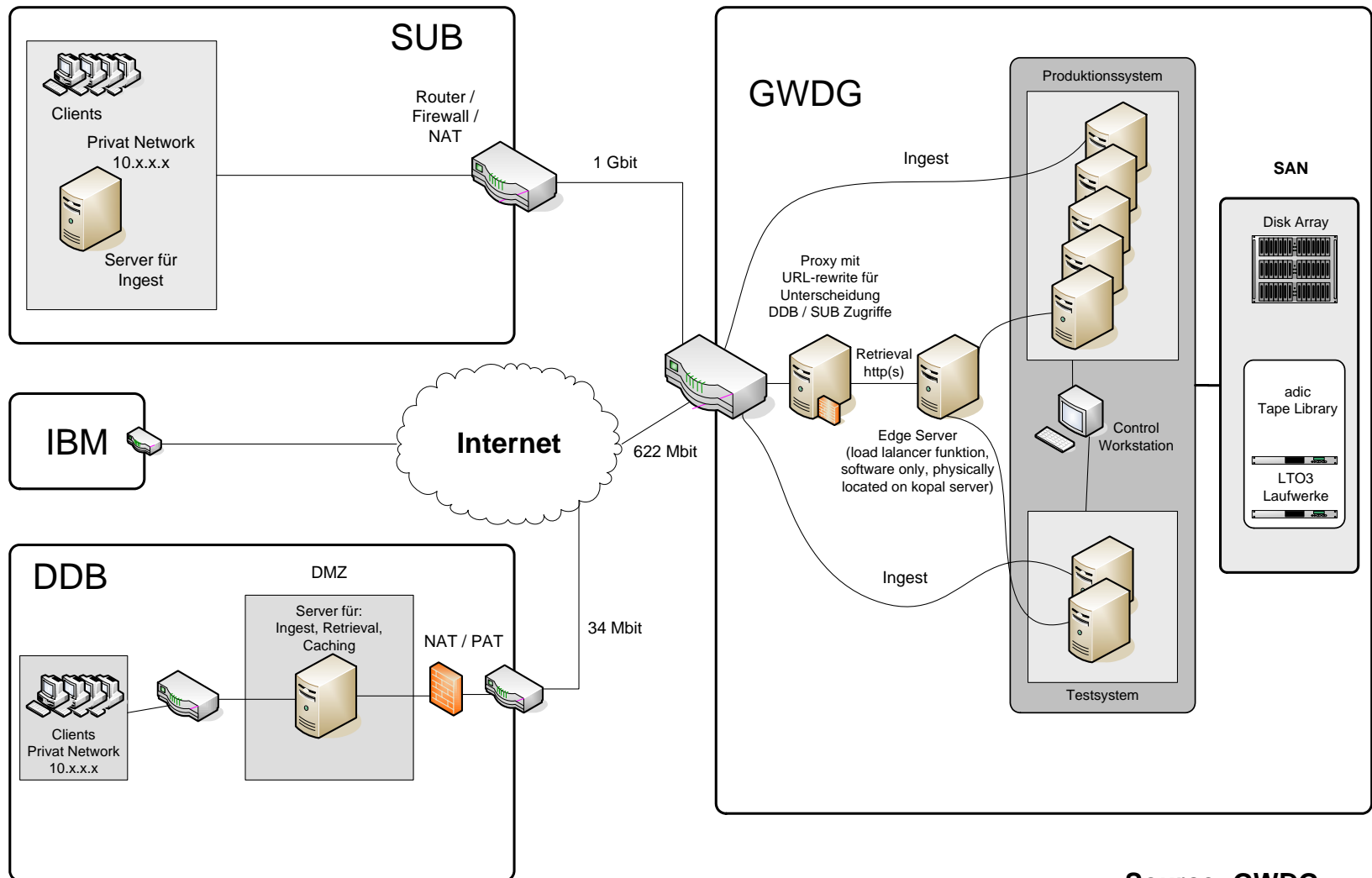


- LMER = „Long term archiving metadata for electronic resources“
- Technical metadata with emphasis on information for long term preservation
- Aim: XML scheme with separate name space as supplement to any given XML metadata (etc. MAB-XML) and joint transfer in one METS package
- Based on the data model of the National Library of New Zealand

mets.xml - example



```
<?xml version="1.0" encoding="utf-8" standalone="no" ?>
<mets OBJID="Kopal Submission Information Package" PROFILE="DDB" xmlns="http://www.loc.gov/METS/" xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:mets="http://www.loc.gov/METS/" xmlns:lmerFile="http://www.ddb.de/LMERfile"
  xmlns:lmerObject="http://www.ddb.de/LMERObject" xmlns:lmerProcess="http://www.ddb.de/LMERprocess" xsi:schemaLocation="http://www.loc.gov/METS/
  http://www.loc.gov/standards/mets/mets.xsd http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/version14/mets.xsd
  http://www.ddb.de/LMERfile http://www.ddb.de/standards/lmer/lmer-file.xsd http://www.ddb.de/LMERObject http://www.ddb.de/standards/lmer/lmer-
  object.xsd http://www.ddb.de/LMERprocess http://www.ddb.de/standards/lmer/lmer-process.xsd">
  <metsHdr CREATEDATE="2006-01-12T14:50:16" RECORDSTATUS="PRODUCTION">
    <agent ROLE="ARCHIVIST" TYPE="ORGANIZATION">
      <name>Die Deutsche Bibliothek</name>
      <note>Automatisch generierte Metadaten. Für weitere Informationen: kopal.langzeitarchivierung.de</note>
    </agent>
  </metsHdr>
  <amdSec ID="AmdSec-0001">
    <techMD ID="TechMD-LMER-Object">
      <mdWrap ID="TechMD-LMER-Object-MdWrap" MIMETYPE="text/xml" MDTYPE="OTHER" OTHERMDTYPE="lmerObject" LABEL="LMERObject">
        <xmlData>
          <lmerObject:name>Eisele, Bettina - Universelle Funktionen mit Lücken-Potenzreihen</lmerObject:name>
          <lmerObject:persistentIdentifier:urn:nbn:de:hbz:385-2902</lmerObject:persistentIdentifier>
          <lmerObject:objectVersion>1</lmerObject:objectVersion>
          <lmerObject:masterCreationDate>2006-01-12T14:50:14</lmerObject:masterCreationDate>
          <lmerObject:metadataCreationDate>2006-01-12T14:50:14</lmerObject:metadataCreationDate>
          <lmerObject:metadataRecordCreator>KOPAL DIAS</lmerObject:metadataRecordCreator>
          <lmerObject:numberOfFiles>3</lmerObject:numberOfFiles>
        </xmlData>
      </mdWrap>
    </techMD>
    <techMD ID="TechMD-File-1">
    <techMD ID="TechMD-File-2">
    <techMD ID="TechMD-File-3">
  </amdSec>
  <fileSec>
    <fileGrp ID="ASSET" ADMID="TechMD-LMER-Object">
      <file ID="FILE0001" MIMETYPE="application/pdf" SIZE="330118" CREATED="2006-01-12T14:50:16" CHECKSUM="17673a8df8fc8ddd101fc0fb6ed69d416608571"
        CHECKSUMTYPE="SHA-1" ADMID="TechMD-File-1">
        <FLocat LOCTYPE="URL" xlink:type="simple" xlink:href="file:///d1/973908904.pdf" />
      </file>
      <file ID="FILE0002" MIMETYPE="text/plain; charset=US-ASCII" SIZE="48" CREATED="2006-01-12T14:50:16"
        CHECKSUM="e54dbd60ec4507768eb7e4a663e19c1045e38595" CHECKSUMTYPE="SHA-1" ADMID="TechMD-File-2">
        <FLocat LOCTYPE="URL" xlink:type="simple" xlink:href="file:///d1/ddbmd5.txt" />
      </file>
      <file ID="FILE0003" MIMETYPE="application/octet-stream" SIZE="13531" CREATED="2006-01-12T14:50:16"
        CHECKSUM="4507ffdb6b0a47bc2d4e80d1b132776359e77db4" CHECKSUMTYPE="SHA-1" ADMID="TechMD-File-3">
        <FLocat LOCTYPE="URL" xlink:type="simple" xlink:href="file:///meta.txt" />
      </file>
    </fileGrp>
  </fileSec>
  <structMap TYPE="ASSET">
  <structMap TYPE="LOGICAL">
</mets>
```

Source: GWDG

- Implementing the new version DIAS V.2 in the productive environment
- Testing of new SW-Components (OSS-Tools)
- Mass procedures for collected materials and new files
- Establishing workflows
- Implementing different access Systems
- Special activities:
 - format registry
 - automatic extraction of technical information

Next steps within the project

- Preservation planning procedures, focal point: migration
- Publishing of kopal Library for Retrieval and Ingest (koLibRI) to create archival objects based on UOF
- Special activity: automatic extraction of technical information: JHOVE
- Graphical User Interface for Workflow tool
- Monitoring functionalities for the system
- Interarchival exchange facilities
- Expression of business models
- www.kopal.langzeitarchivierung.de



- Become a partner and use your own account within DIAS in Göttingen
 - License fee
 - Free use of local software
- Buy the DIAS Solution and be your own service provider
 - Cost of DIAS Solution (not yet a product)
 - Free use of local software
- Build your own OAIS deposit, but support the same interfaces (universal object format)
 - Free use of local software



koLibRI



www.kopal.langzeitarchivierung.de

Reinhard Altenhöner

<mailto:altenhoener@dbf.ddb.de>

<http://www.ddb.de>