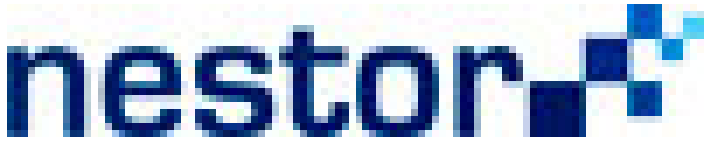


International Conference on
Preservation of Digital Objects
15 - 16 September 2005 in
Göttingen, Germany

Information Life Cycle Management and Long Term Preservation - Technical Aspects Of Workflow Organisation

Reinhard Altenhöner, Die Deutsche Bibliothek, Frankfurt am Main

Agenda

- Brief overview: Die Deutsche Bibliothek and its activities in long-term preservation
- The kopal initiative – our technical approach for LTP
- Present workflow organisation in DDB
- What the others do: Information Lifecycle Management as a strategic approach
- Principles for workflow organisation, technical consequences and first results

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

Activities in DDB – our way to LTP…

- **DissOnline**

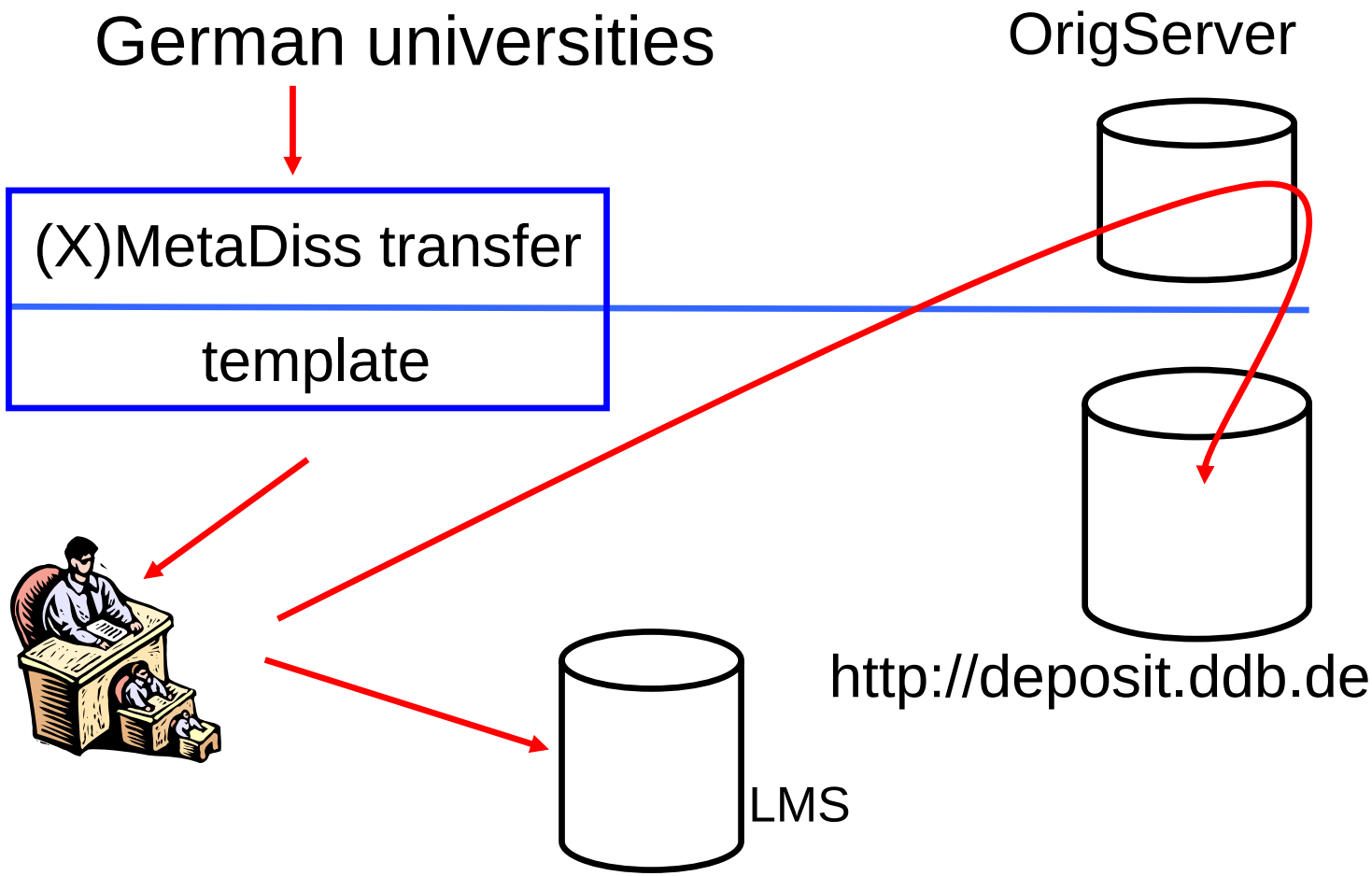
  in close cooperation with universities and university libraries
- **E-journals**

  in close cooperation with publishers
- **Digitization activities**
- **Transfer of Amiga-, Atari-, Commodore applications** into a modern storage technology
- **URN-Service**
- **Exchange-models** for metadata and objects

Electronic archives at Die Deutsche Bibliothek

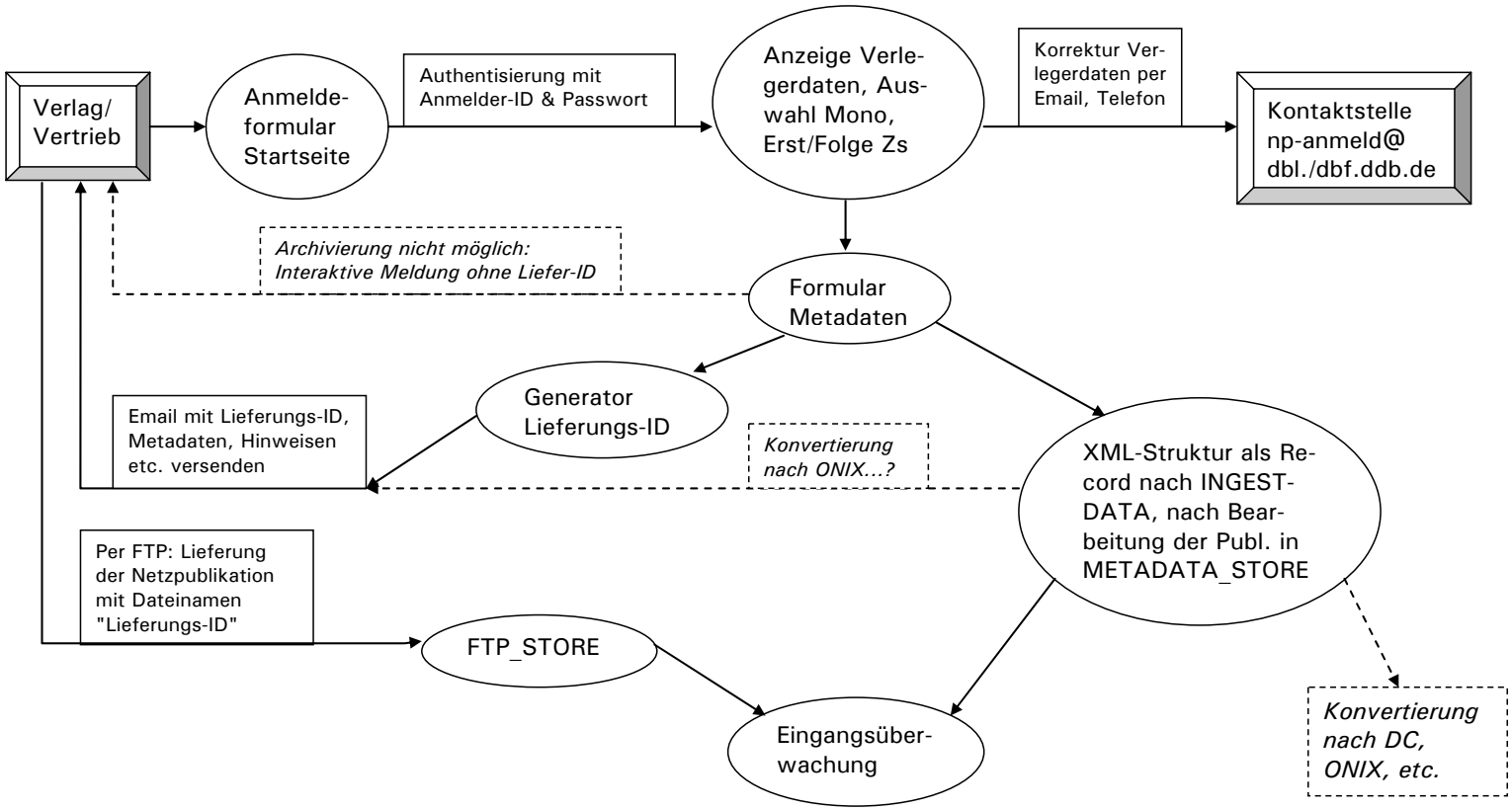
- 37.000 Online-Dissertations
- 454 e-journals and 1.400 monographs from Springer (Heidelberg, Berlin) (http://link.springer.de)
- 60 newsletters
- 2.600 electronic publications from approx. 190 commercial and non-commercial publishers
- Web-Sites
- → workflows, modules, procedures

Online theses: DissOnline

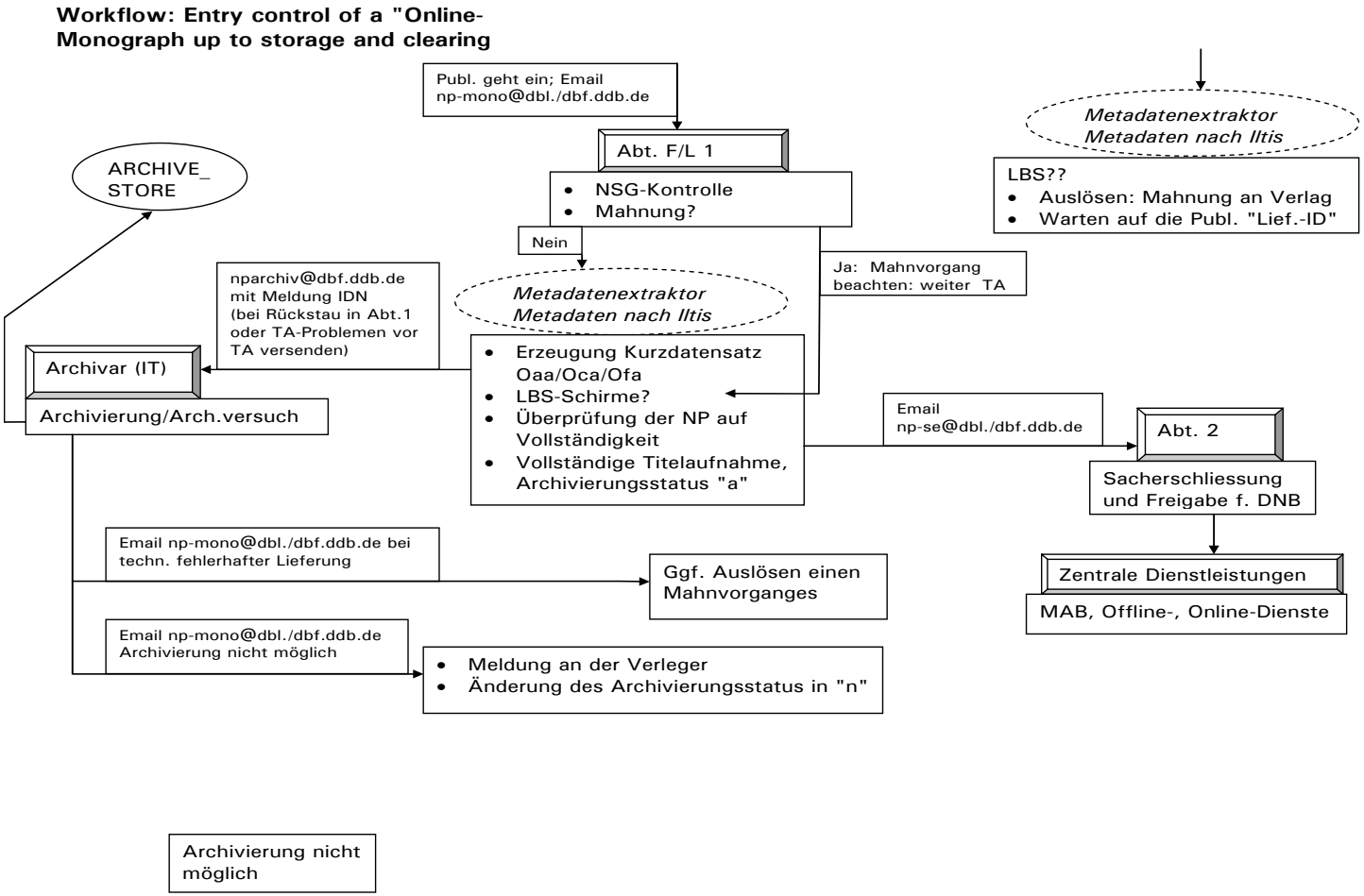


Workflow-organisation I

Workflow: registration of a "Online-Monograph" respectively initial registration or follow-up registration of a "Online journal"



Workflow-organisation II



Transfer of data from an old platform

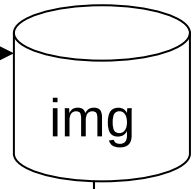
eg. C64 / AMIGA

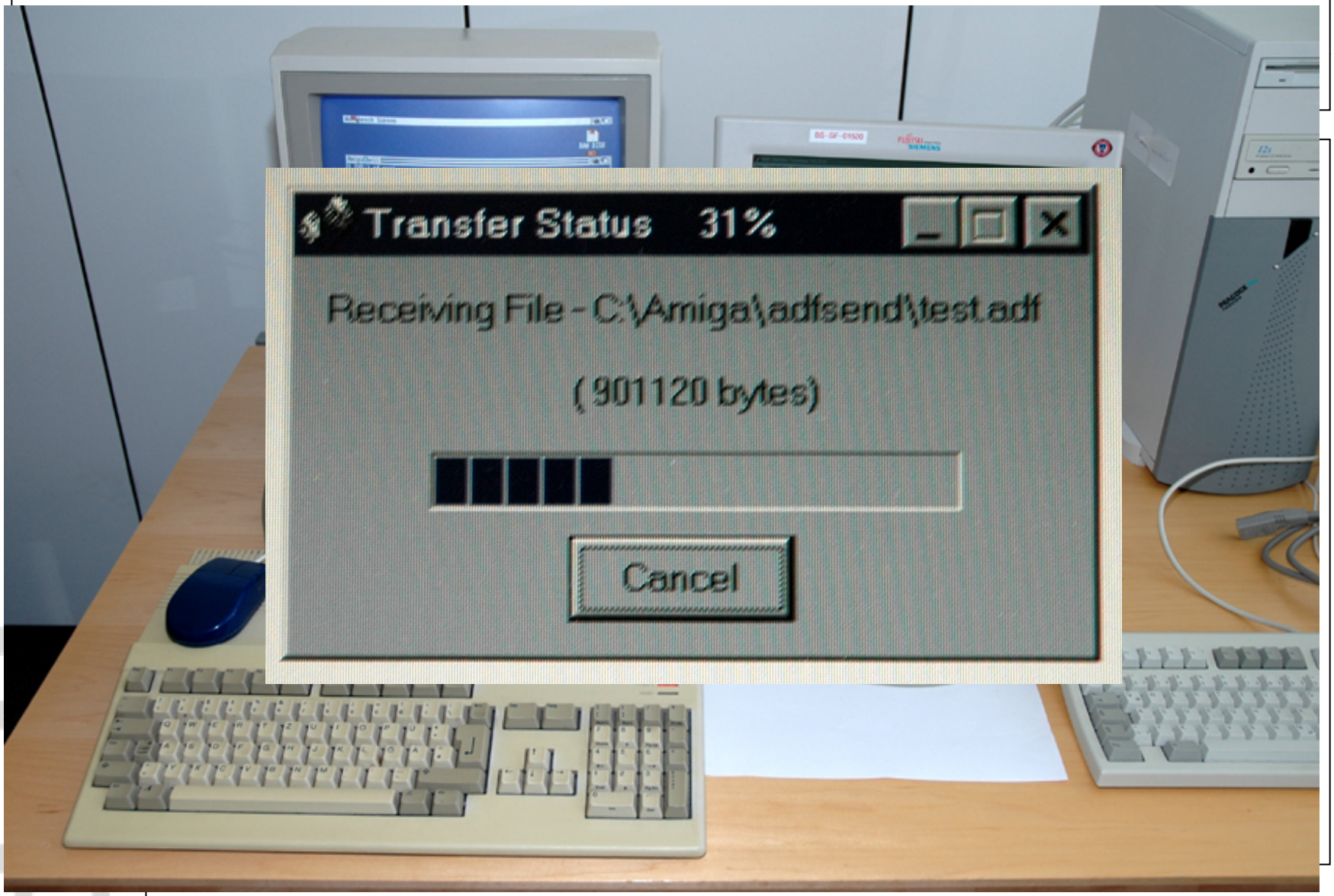
original system 1988

migration

target system 2002

Emulation OS




Transfer Status 31%
Receiving File - C:\Amiga\adfsend\test.adf
(901120 bytes)
Cancel


VideoClips
ANIMATION-MELSCHLE
Commodore
AMIGA
VideoClips
ANIMATION-MELSCHLE

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
URL:	Archivserver Der Deutschen Bibliothek
Sachgruppe:	33 Medizin ; 32 Biologie

URN

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

Co-operative development of a long-term digital information archive - „kopal“

Die Deutsche Bibliothek

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

- funded by the Federal Ministry for Education and Research
- Financial volume: 4,2 Mio € + self-financed activities of all partners, duration: 1.7.2004 – 30.6.2007
- 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

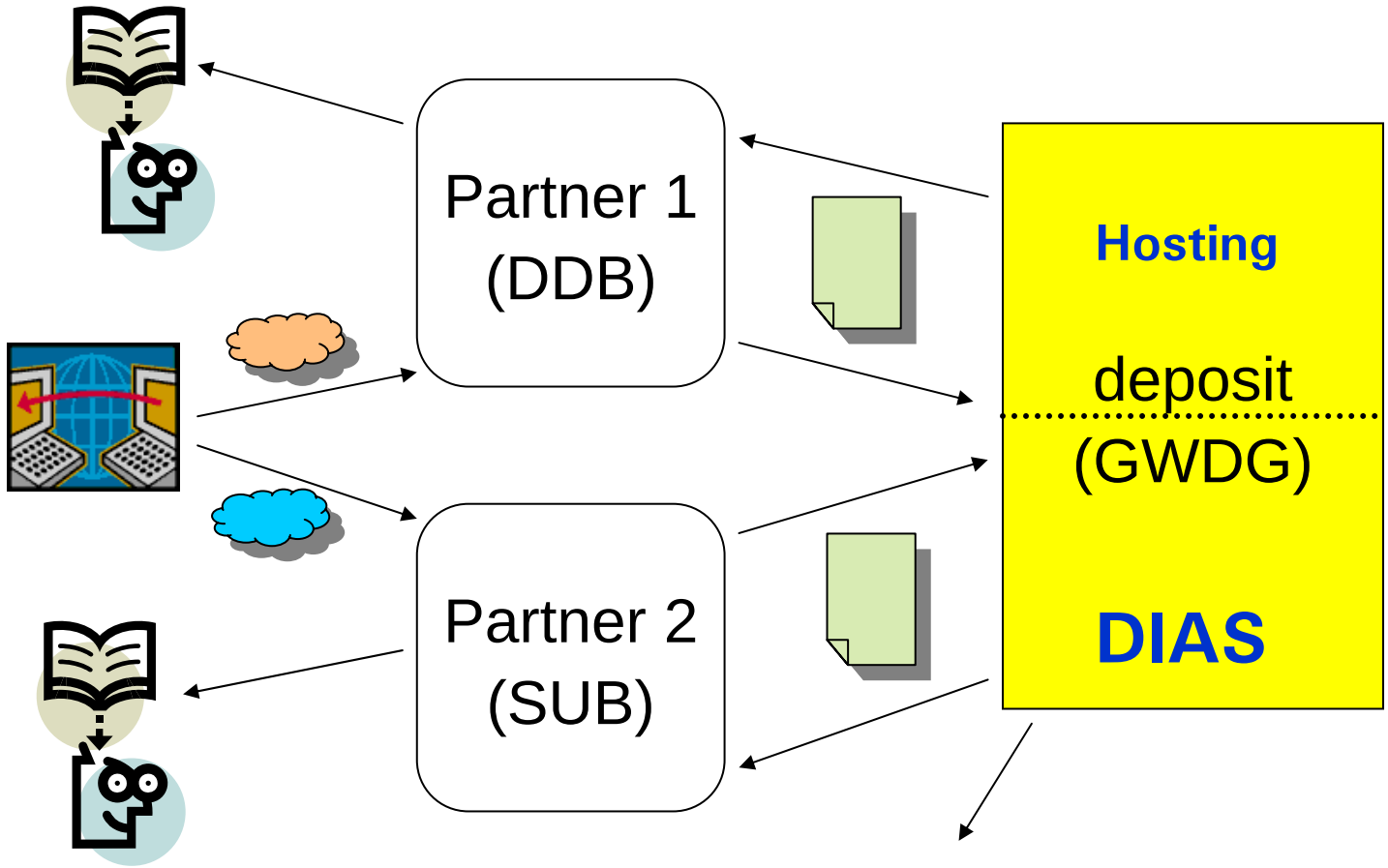
kopal: Partners

- Die Deutsche Bibliothek (leader)
- Staats- und Universitätsbibliothek Göttingen
- Industrial Business Machines (IBM) Germany
- Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen (GWDG)

Working relationship:

Royal Dutch Library, The Netherlands

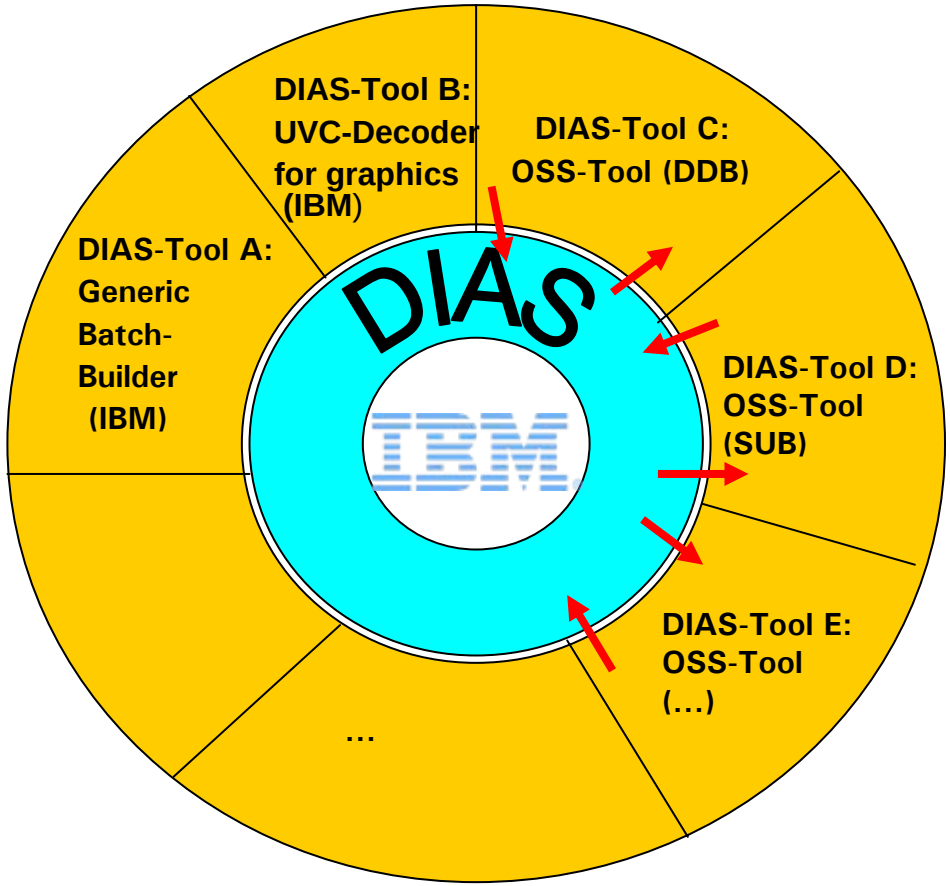
kopal: structure & concept



kopal: Concept

- 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

Software development & sharing



Condition:

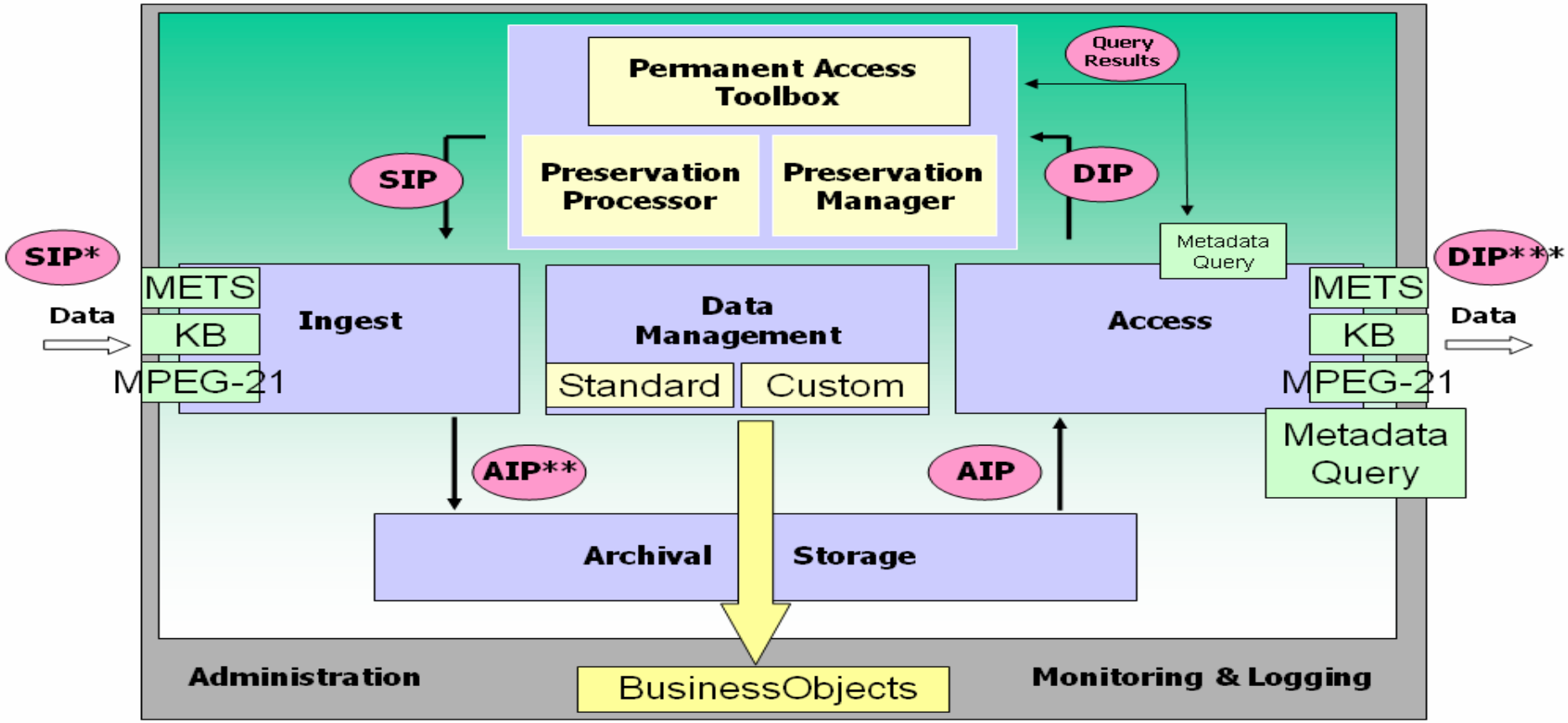
Well defined interfaces (SIP + DIP)

Enhanced DIAS

Integration of the univeral object format into DIAS

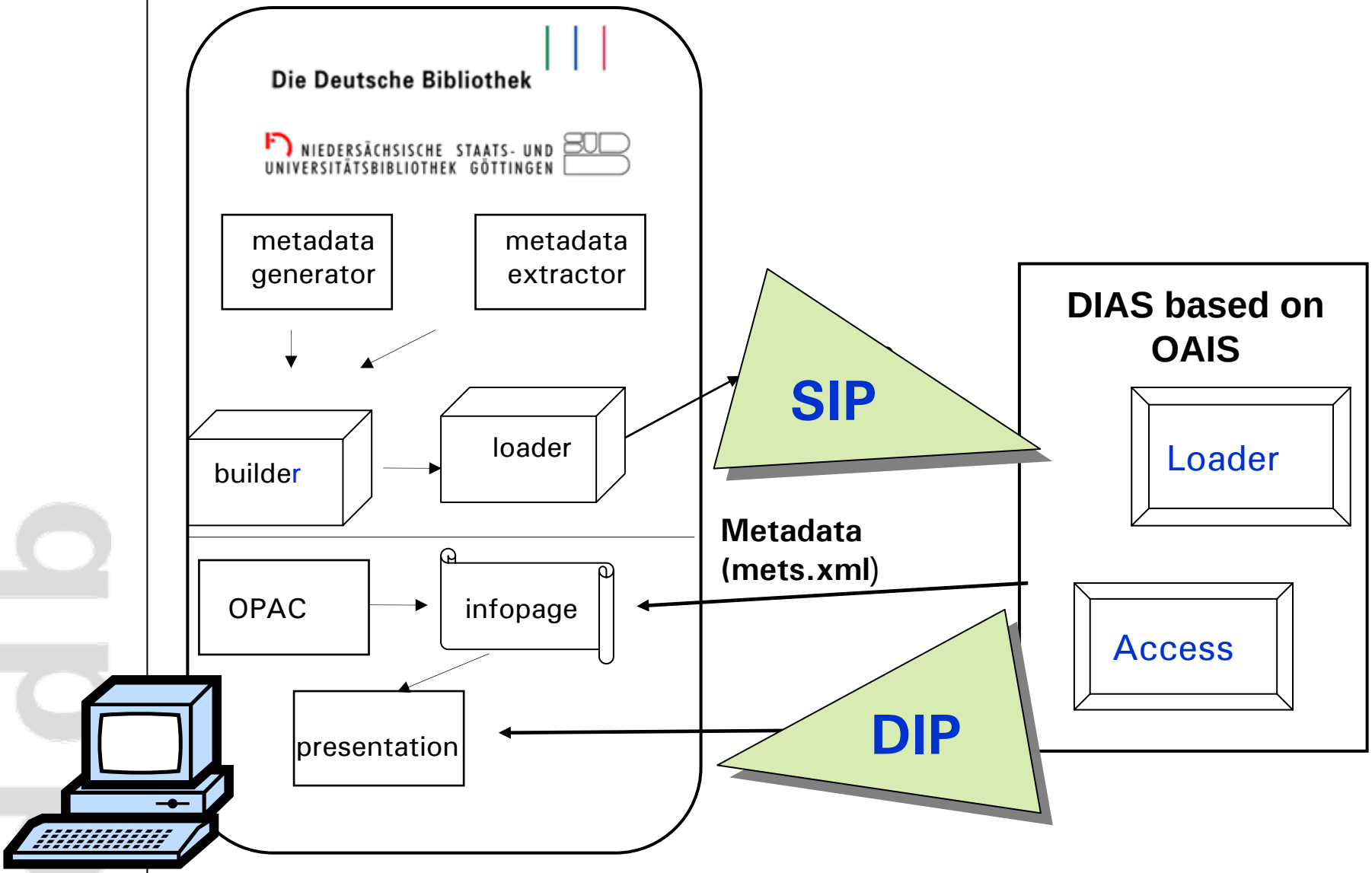
Source:IBM

kopal
Daten für die Zukunft

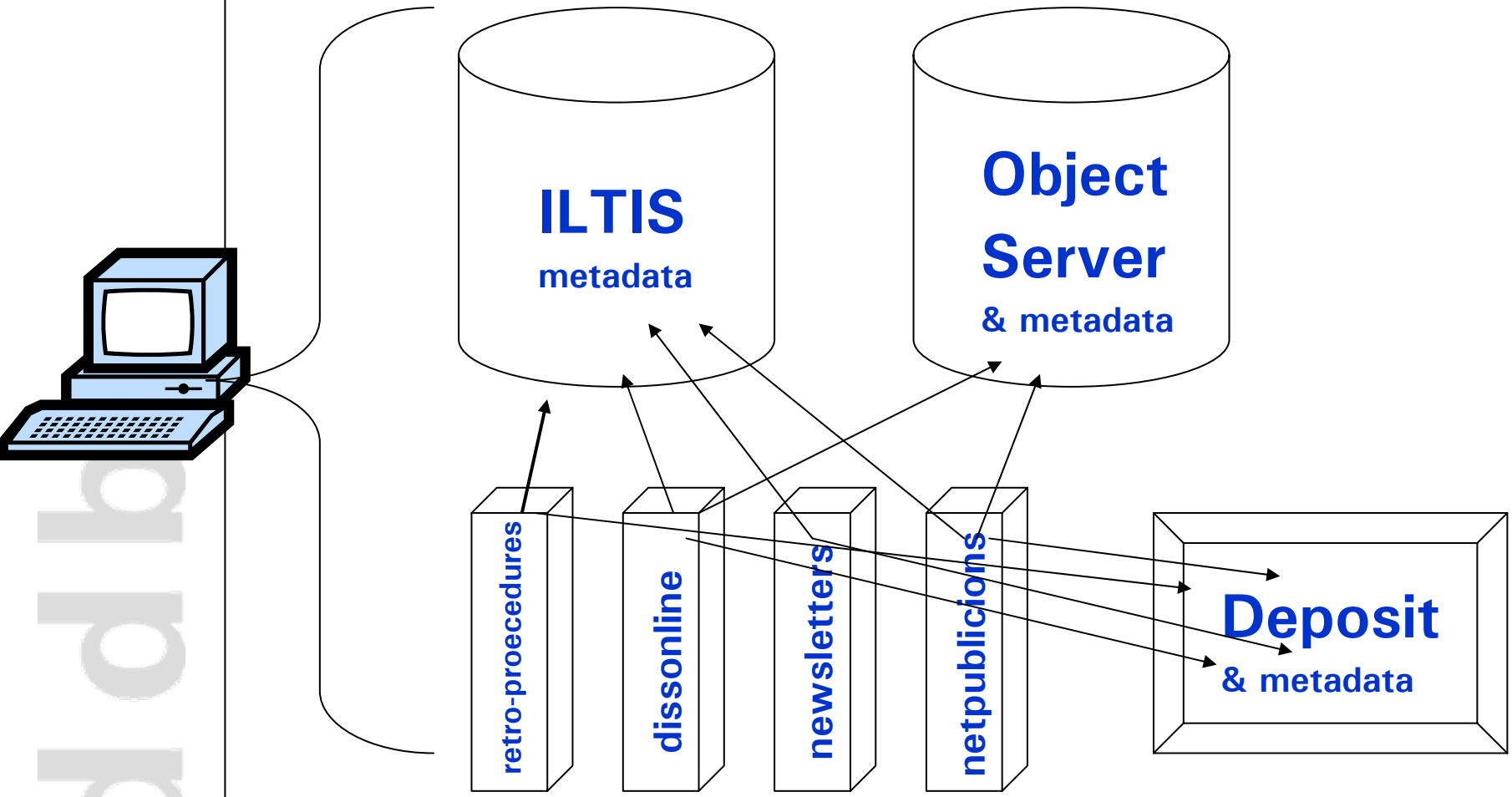


* = SIP: Submission Information Package
** = AIP: Archival Information Package
*** = DIP: Dissemination Information Package

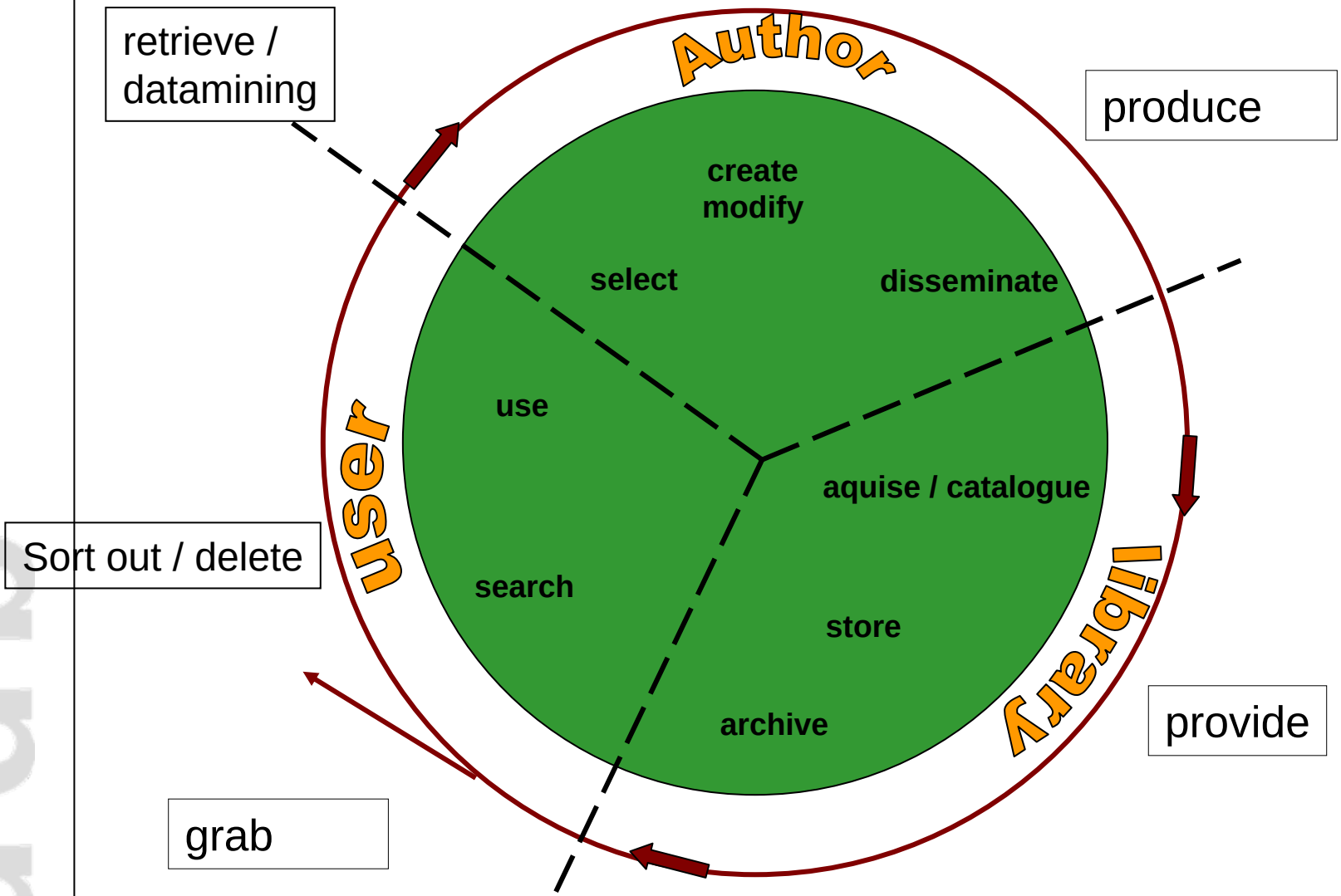
kopal – DIAS *and* DDB/SUB-tools



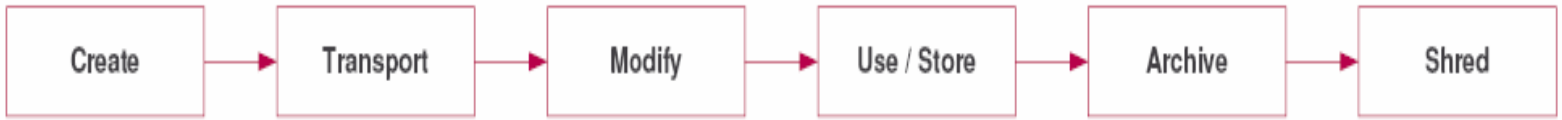
Integration - Coordination



Information Cycle



Information/Data Life Cycle



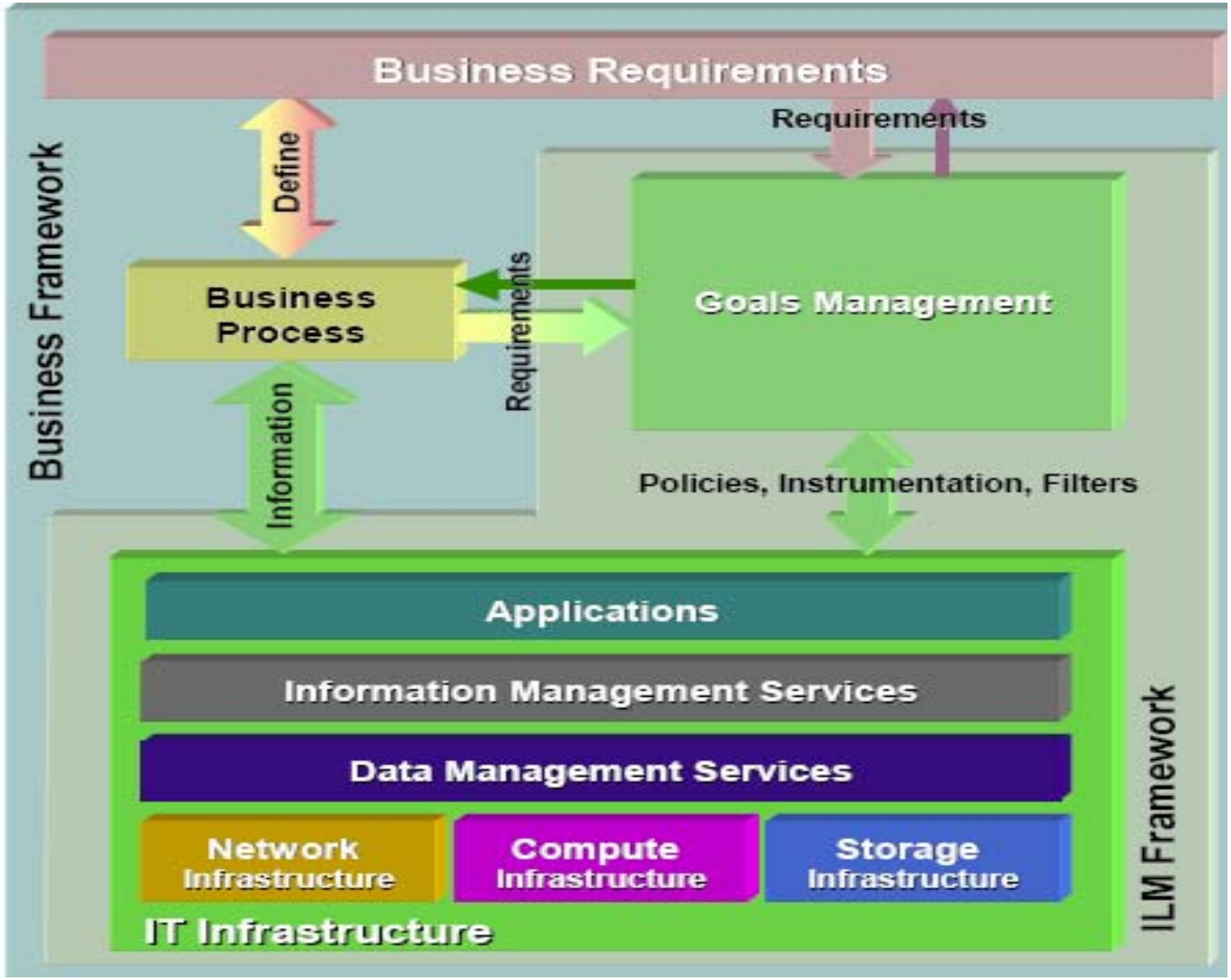
- **Create:** production of data (6 Exabytes worldwide p.a.).
- **Transport:** volume is 18 Exabytes worldwide p.a.
- **Modify:** Ca. 10 % of the data gets changes.
- **Use / Store:** 30 days after storing the data only 20% are in use.
- **Archive:** the archiving depends on the regulations/rules of an enterprise.
- **Shred:** planned and organised destruction of data

Information Lifecycle Management

Information Lifecycle Management is comprised of the policies, processes, practices, and tools used to align the business value of information with the most appropriate and cost effective IT infrastructure from the time information is conceived through its final disposition. Information is aligned with business requirements through management policies and service levels associated with applications, metadata, and data.

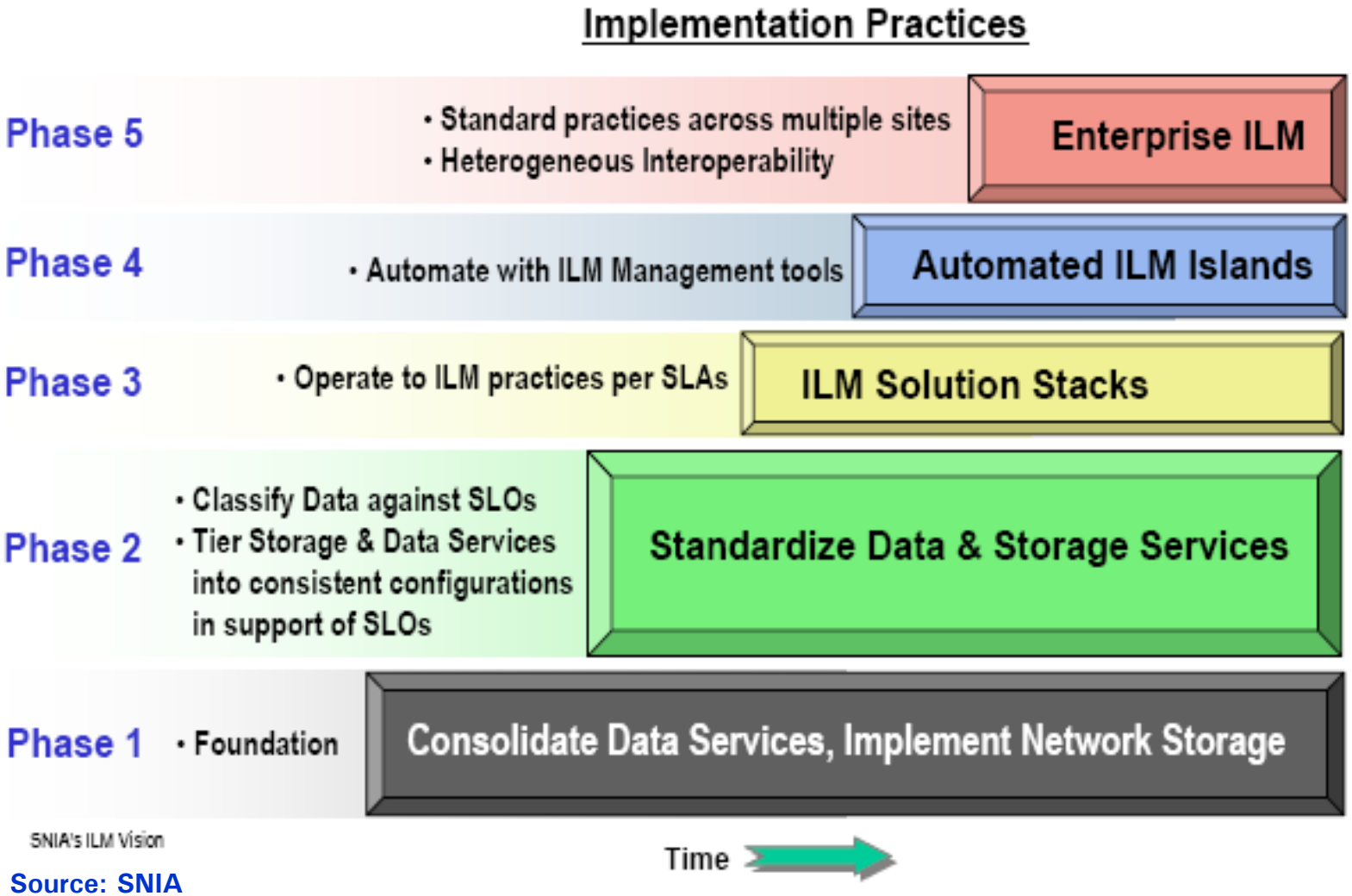
Storage Networking Industry Association

ILM



Source: SNIA

ILM Roadmap



Source: SNIA

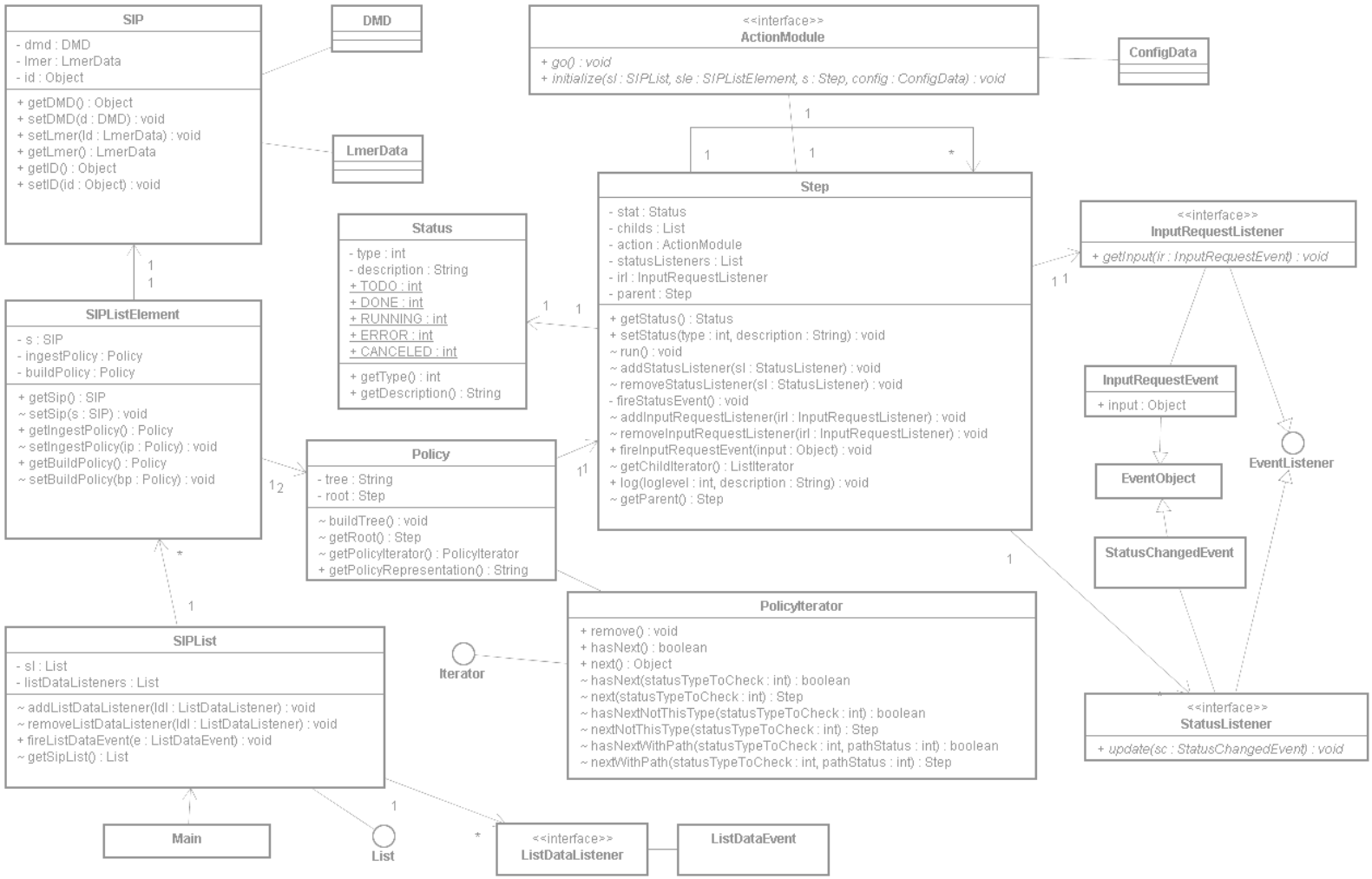
Characteristics ILM

- Comprehensive, sustainable strategy
- Storage strategy including history
- Refers to single objects, not to classes or pathes
- Information value measurement per single object Clear & well defined processes
- High automation process
- Uniform storage areas in central administration
- Avoids media breaks
- Documentation / quality management

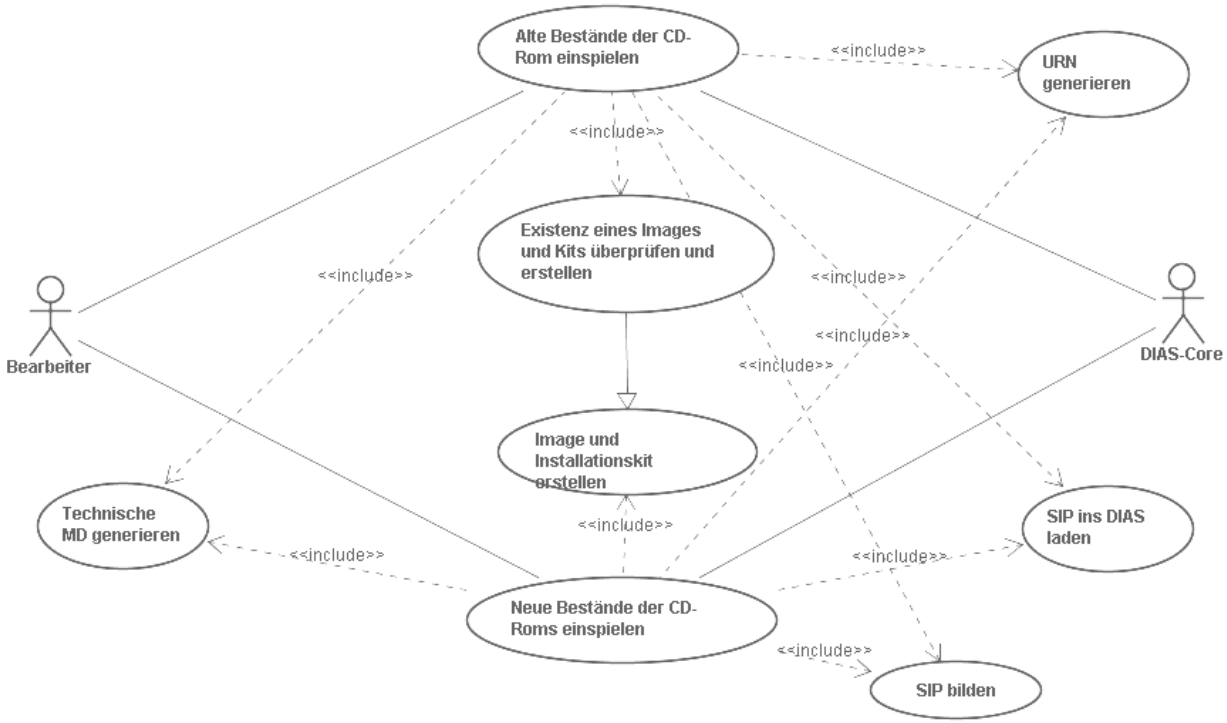
Consequences for DDB

- Complete life cycle for information objects
- Special metadata set
- Metadata per one single object
- Migration history & documentation
- Generic archive infrastructure
- Multi-client model
- Specific operation model
- Export / import possibility
- Technical methods for the legal transformation of objects
- Uniform, homogeneous workflow for heterogeneous objects

Structure of classes for kopal



Object classes in SW-classes



1998 - 2000

Source: nedlib-project

What's our aim ?

DDB-products

DNB

OPAC

Portals

Index

Index

Metadata

Metadata

Cooperations

Authority files

enrichment

printed publications

abstracts

fulltext

archive

collections

web-Sites

Enhenced information

Scans

Multimedia-

Audio-files

Digitasations

LZA: kopal

Reinhard Altenhöner
Die Deutsche Bibliothek
Head of IT-Department
Adickesallee 1
D-60322 Frankfurt am Main
Telefon: +49-69-1525-1700
Telefax: +49-69-1525-1799
mailto:altenhoener@dbf.ddb.de
http://www.ddb.de