

"kopal goes live!"

## Nutzungsmodelle und Perspektiven eines Langzeitarchivs digitaler Informationen

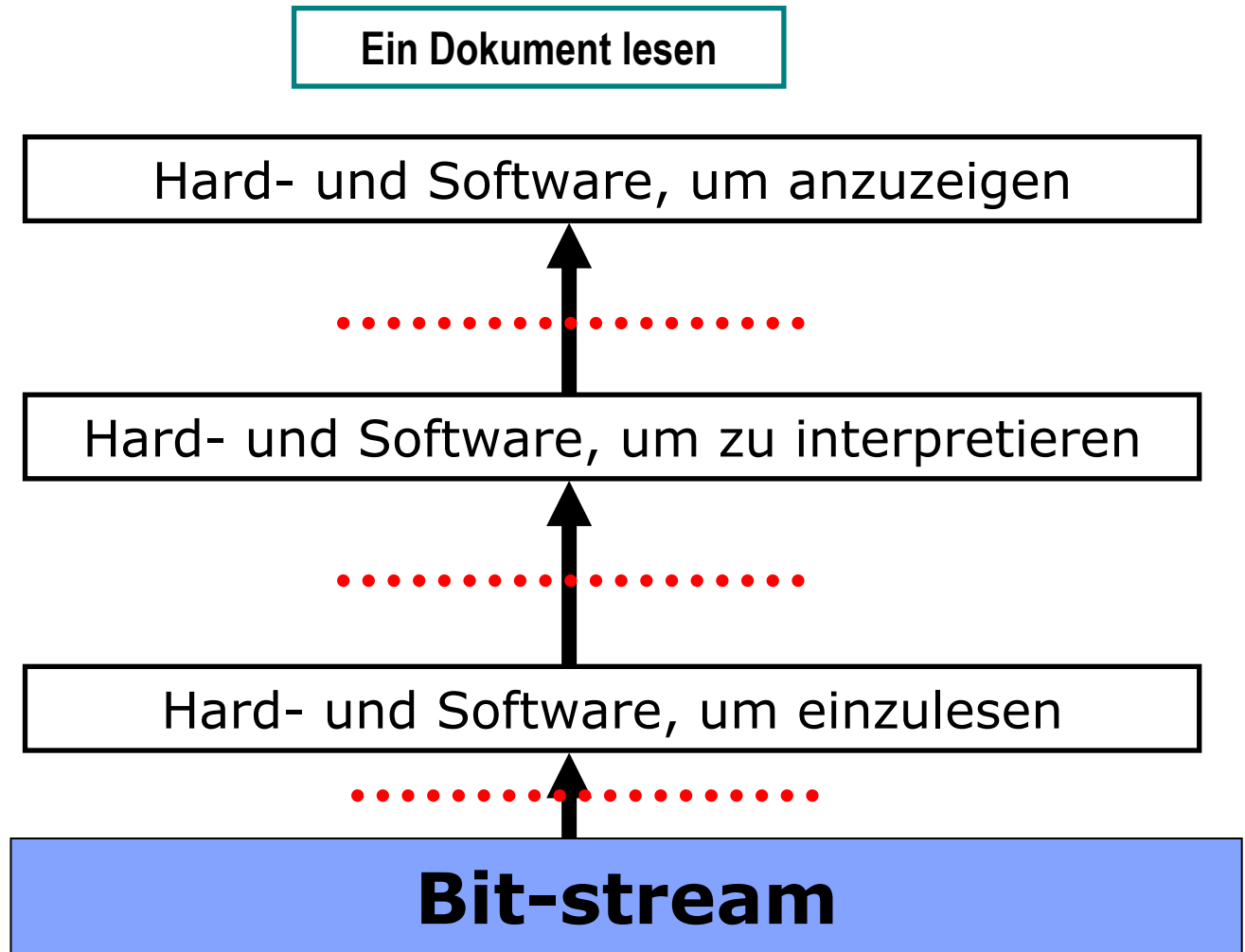
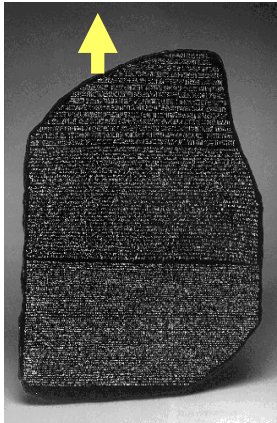
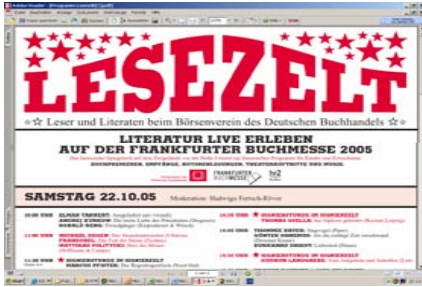


## Vorstellung der Projektergebnisse

Frankfurt am Main 18.6.2007, Reinhard Altenhöner, DNB

- 11 Uhr Empfang & Anmeldung
- 11:30 Uhr Begrüßung – Frau Schwens, DNB
- 11:40 Uhr Vorstellung der Projektergebnisse -- Hr. Altenhöner, DNB
- 12:15 Uhr I. Kooperation & Nachnutzung der kopal-Lösung
- SUB Göttingen, Hr. Dr. Lossau, Hr. Klaproth
- 13 Uhr Mittagspause
- 13:45-15:00 Uhr II. Technik & Entwicklung
- 1. Technik der kopal-Lösung -- Hr. Steinke, DNB
- 2. DIAS-Betrieb und -Hosting -- Fr. Ulrich, GWDG
- 3. DIAS-Entwicklung und -Lizenzierung – Herr Schwarz, IBM
- 4. koLibRI: Workflows und Tools -- Hr. Ludwig, SUB
- 15:00-15:30 Uhr Abschlussdiskussion

- Rasanter Technologiewechsel erschwert Zugriff auf ältere Datenformate
  - Problem 1: Erhalt der binären Daten (0 und 1)
    - kein existierender Datenträger ist ewig haltbar
    - Lösung: Regelmäßige Erhaltungsmaßnahmen
  - Problem 2: Zugriff auf die Inhalte
    - zahlreiche Formate, ständig neue, alte verschwinden
    - Abhängigkeit von aktueller Soft- und Hardware
    - Lösungen:
      - Migration (regelmäßiges Konvertieren),
      - Emulation (Nachstellen benötigter Systeme)



```

11011101110011110011010010101011101011100101010111000110101010101000110101010101010101000101010101010101
01010101010100010101010101010100111000001011111001101110111001110011010010101011101011100101010111000110
10101010100011010101010101010100010101010101010101010101010101010101000101010101010100111000001011111001101
11011100111100110100101010111010111001010101110001101010101010001101010101010101010101000101010101010101
    
```

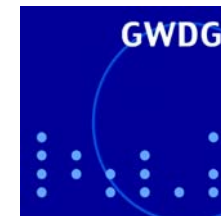
- → Wir brauchen Informationen über die Produktionsumgebung des gegebenen Objekts und über das Objekt selbst
  - Erforderliche Software / Formatanforderungen
  - Technischen Entstehungskontext
  - Eingebettete Objekte
- → Entscheidend: Metadaten  
(übrigens auch selbst langzeitverfügbar zu halten)

- Langzeitarchivierung ist teuer und aufwendig: technisch wie personell
- Einheitliche Standards: Sie bilden sich und leben nur in kooperativen Arbeitszusammenhängen
- Die LZA muss national und international koordiniert werden
- Sie muss Institutionen- / Spartengrenzen überspringen
- Braucht komplementäre Nutzungszusammenhänge
- Sie bedarf eines klaren Organisations- und Betriebsmodells

- Förderprojekt des Bundesministeriums für Bildung und Forschung (BMBF)
- Charakter: Forschungs- und Innovationsprojekt
- Fördervolumen: primär 4,2 Mio € + Nachförderung und Eigenmittel, Laufzeit: 1.7.2004 – 30.6.2007
- Projektziel: Aufbau einer von Gedächtnisorganisationen nachnutzbaren technischen und organisatorischen Infrastruktur zur Sicherung der Langzeitverfügbarkeit elektronischer Publikationen
- Leitbegriffe: Kooperation, Nachnutzbarkeit, Nachhaltigkeit

- Deutsche Nationalbibliothek (DNB) (Gesamtprojektleitung)
- Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB)
- IBM Deutschland GmbH (IBM)
- Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)

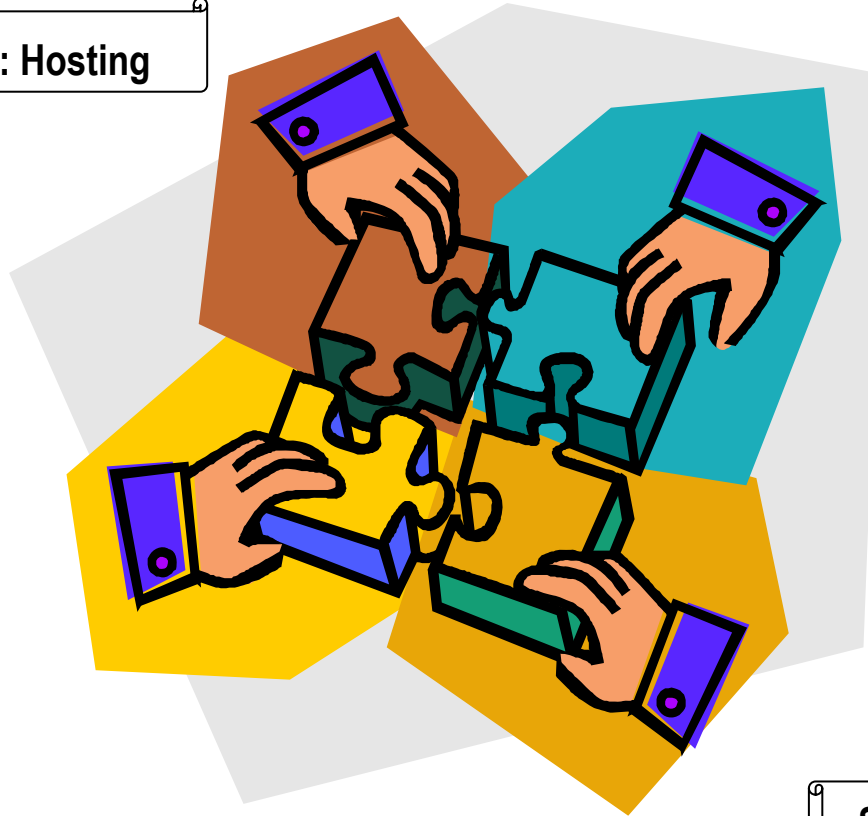
**Deutsche Nationalbibliothek**





GWDG: Hosting

IBM: Archivierung SW

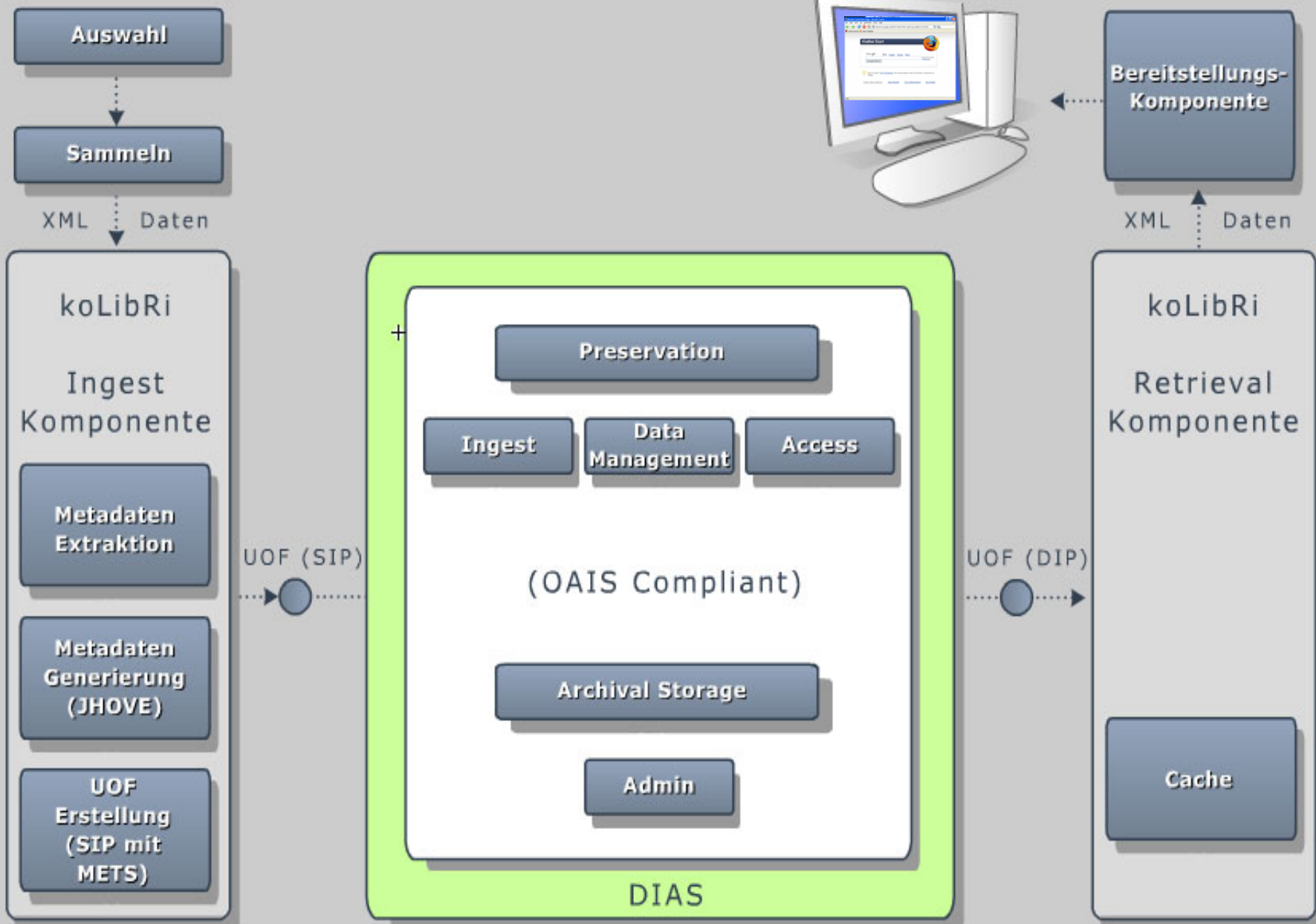


SUB: Ingest/Access SW

DNB: Ingest/Access SW

Gemeinsam: Preservation Planning

- Langzeitarchivierung digitaler Informationen (>100Jahre)
- Entwicklung einer standardisierten Lösung
  - nachnutzbarer Transfer für andere Institutionen
  - vereinfachte Kooperation zwischen Bibliotheken
- Erweiterung des IBM Assets DIAS-Core um Open-Source basierte Komponenten
  - flexibler Einsatz durch Modularität der Lösung
- Entwicklung einer erweiterbaren User Group (KB, DNB, SUB...) zur kooperativen Weiterentwicklung der Lösung
  - Niedrigere Kosten
  - Zeitlich schnellere Weiterentwicklung
  - Stärkere Standardisierung möglich
- Prozessintegration z.B. Automatisierung der Einlieferrountinen, Optimierung für Massenverfahren



- Kernsystem DIAS von IBM
  - Basiert auf bewährter Standardsoftware
  - Zukunftsfähige Trennung von Speicherkonzept und Datenverwaltung
  - Erfahrungs- und Technologieaustausch mit der Nationalbibliothek der Niederlande und anderen internationalen Partnern
- Lokale Software koLibRI von DNB und SUB
  - Generische freie Softwarebibliothek zur Ein- und Anbindung unterschiedlicher Mandanten
  - Automatische Erstellung von technischen Metadaten
  - Konfigurierbare Workflows
- Universelles Objektformat
- OAIS-Referenzmodell konsequent implementiert

## DIAS-CORE

### Basisentwicklung E1

Basisinstallation inkl.  
Mandantenfähigkeit

Objektformatunterstützung und  
Administrationsfunktionalitäten

### Preservation Planning E3

Planung und Konzept,  
Workflowdefinition

Entwicklung und Integration in  
DIAS-Core

## Lokale Systeme

### Basisanbindung E2

Builder:  
universell / individuell

Access:  
universell / individuell

### Exemplarische Übernahme E4

Material und Entwicklung von  
Deutscher Nationalbibliothek

Material und Entwicklung von  
der SUB Göttingen

- Pilotphase (07 – 10/2004) Go!-Entscheidung
- Entwicklungsphase (11/2004 – 06/2007)
- Entwicklung UOF, Feinspezifikation der SIP / DIP – Interfaces
- Entwicklungsarbeiten DNB/SUB & IBM
- Testphase zur Abnahme des DIAS-Systems vom 17.10.2005 bis 16.11.2005
- Positive Entscheidung über Abnahme des Systems durch Projekt kopal
- 12.2005 – 01.2006: Aufsetzen des Produktivsystems
- Februar 2006: Beta-Release der koLibRI-Software
- März – Juni 2006: Updates, Storage-Konfigurationen, Probe-Ingests, Bugfixes und Endabnahme Produktivumgebung
- Ab Juli 2006: Beginn des produktiven Material-Einspielens
- 03/2006 – 06/2007: Spezifikation Administration Interface & Umsetzung, Einführung

- Online-Dissertationen (~ 53.000, ab 1997, ~ 400 GB)
- Netzpublikationen (SpringerLINK), 445 Zss.,
- Monog. ~ 300 GB)
- CD-ROMs (~ 60.000, 42.000 GB), exemplarisch
- Quellensammlungen, Multimedia-Anwendungen, Filme, Spiele, Datensammlungen (tw. wiss. Primärdaten), sonst. Publikationen (~ 14.300 GB), exemplarisch
- Digitalisate, Audiodaten (~ 558.000 GB), exemplarisch

- Workflow von elektronischen Materialien auf Datenträgern
  - Erstellung ausgewählter repräsentativer CD/DVD-Images inkl. Problemanalysen und Einspieltests
- Workflow für (online-only) Netzpublikationen
  - Einbeziehung der Anforderungen der LZA in Gestaltung des neuen Workflows für NP
  - Design des Workflows auf die Schnittstellen des Archivsystems
- Fortlaufende Publikationen (v.a. Zss-Artikel)
  - Festlegung von Erschließungsvarianten
  - Mapping von Archivobjekten auf Katalogobjekte
- Rechtemanagement
  - Formulierung von Anforderungen
  - Definition möglicher Varianten für die Dokumentanzeige
  - Information wird in ILTIS hinterlegt
  - Einbindung in Zugriffssystem der DNB
- Erweiterung des URN-Managements



- Klärung der Storage-Policy der DNB- und SUB-Daten bei der GWDG, inkl. Durchführung eines Storage-Workshops
- Erarbeitung eines Betriebskonzepts für kopal, speziell für den Projektzeitraum (befindet sich in Abstimmung)
- Erstellung von Service- und Lizenzmodellen für den Weiterbetrieb von kopal bzw. die Aufnahme neuer Partner
- Vorbesprechungen mit interessierten Institutionen über Material-/Einspieltests auf der kopal-Testumgebung (u.a. HWWA, FIZ Karlsruhe, BSZ Konstanz, SLUB Dresden, TIB Hannover, DIPF...)
- Proof of concept für migration process



koLibRI

[kopal.langzeitarchivierung.de](http://kopal.langzeitarchivierung.de)

Reinhard Altenhöner

<mailto:r.altenhoener@d-nb.de>

<http://www.d-nb.de>