

Reinhard Altenhöner

**Zwischen Vision und Alltag:
der neue Sammelauftrag der
Deutschen Nationalbibliothek in
einem veränderten Umfeld von
Publikation und Zugang**

10. InetBib-Tagung in Würzburg

ToC

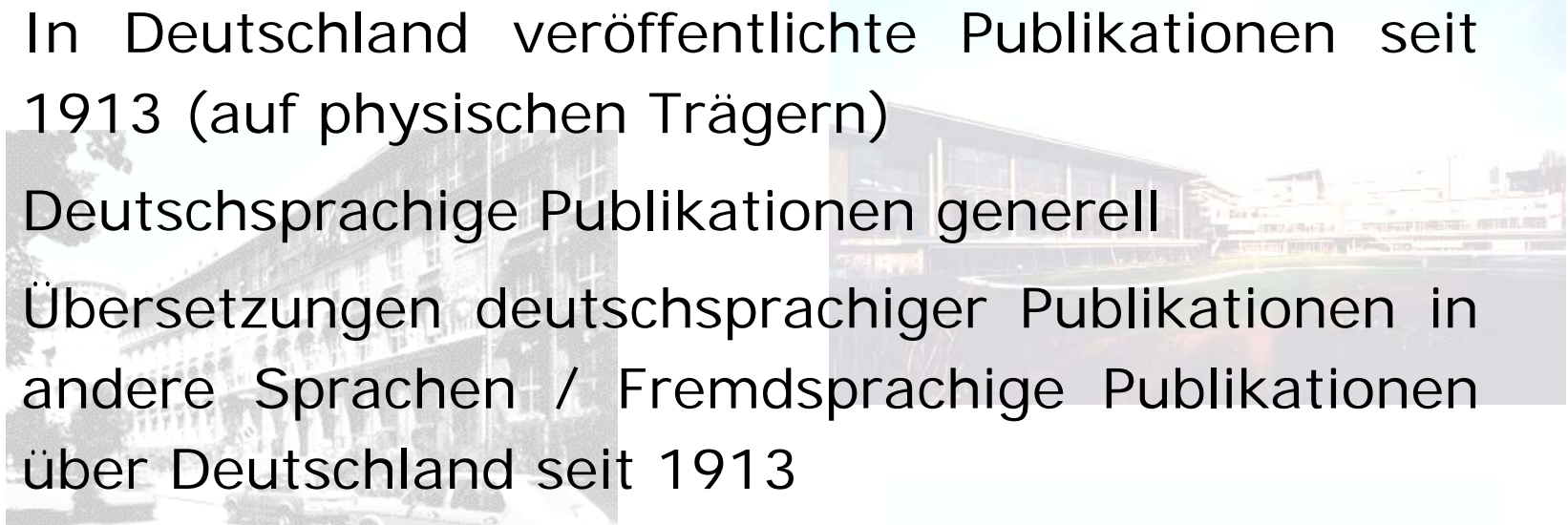
- 1. DNB und der neue Auftrag**
- 2. Umfeld- und Standortanalyse: Was bedeutet das für DNB?**
- 3. Arbeitsschritte in DNB**
- 4. Ausblick**

Deutsche Nationalbibliothek



Sammeln, Erschließen, Archivieren, Bereitstellen

- In Deutschland veröffentlichte Publikationen seit 1913 (auf physischen Trägern)
- Deutschsprachige Publikationen generell
- Übersetzungen deutschsprachiger Publikationen in andere Sprachen / Fremdsprachige Publikationen über Deutschland seit 1913
- Sondersammlungen
- Digitale Publikationen auf physischen Trägern



Medienvielfalt im Übergang



- steigende Anzahl an Ressourcen
- rasant wachsende Wissensdomänen
- „Klassisch“: abgeschlossene Verlagspublikationen (Bücher, Zeitschriften, Filme, Briefe)
- Immer häufiger:
 - Digitale Ressourcen
 - vernetzte Informationsfragmente und dynamische Ressourcen (Websites, Foren, Blogs, E-Mails, ...)
 - Applikationen
- Erschließung wird schwierig
 - Steigende Quantität
 - Neudefinition d. Begriffs "Publikation" / Objekt
- Verfügbarkeit wird zum Problem

Fundsache

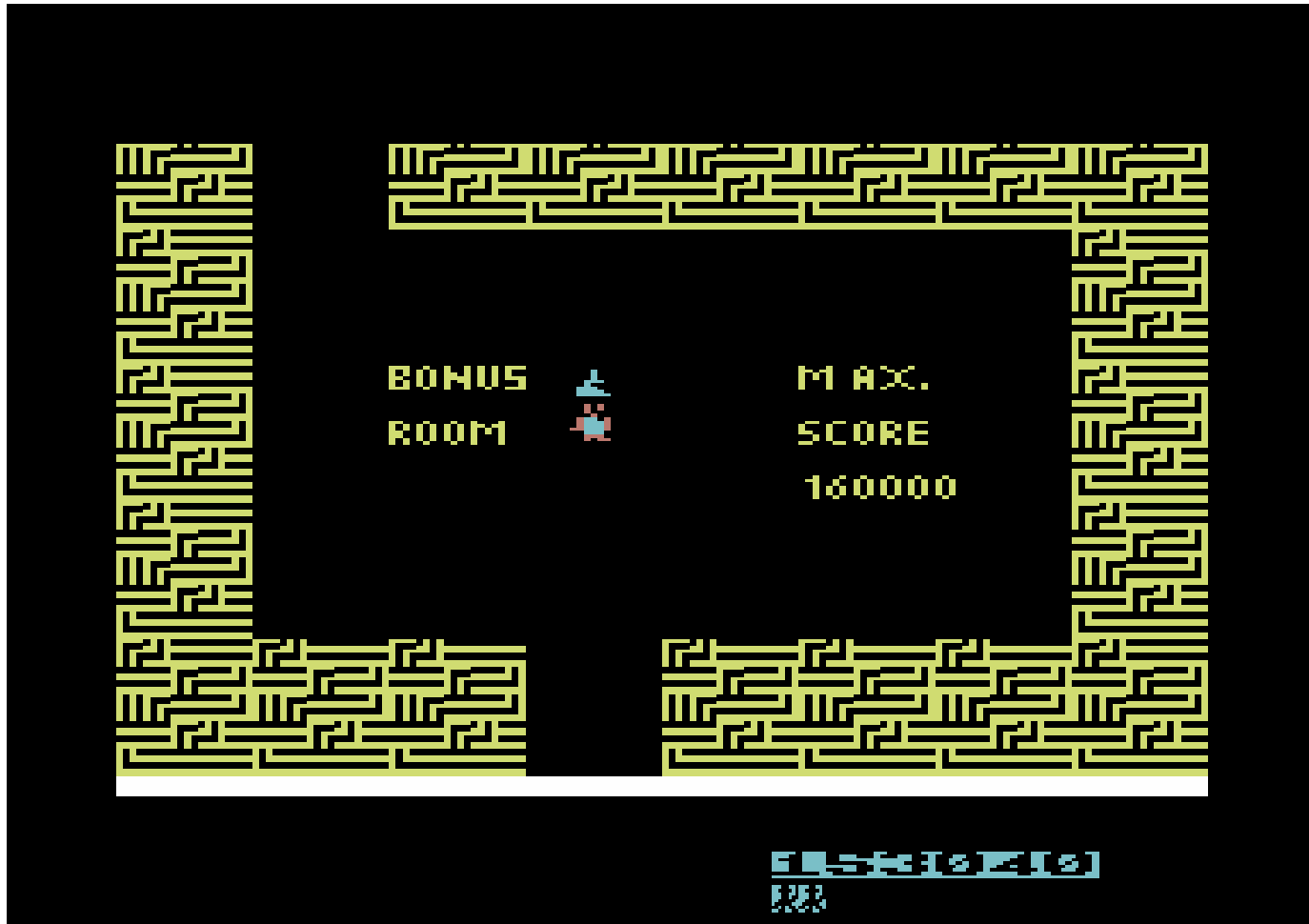


Aus dem Nachlass des Nobelpreisträgers Heinrich Bölls, gefunden am 1.2.2008, enthält vermutlich die Korrespondenz der letzten Jahre, vor allem auch Briefentwürfe

Daten auf der Spur



Die Frage bleibt offen

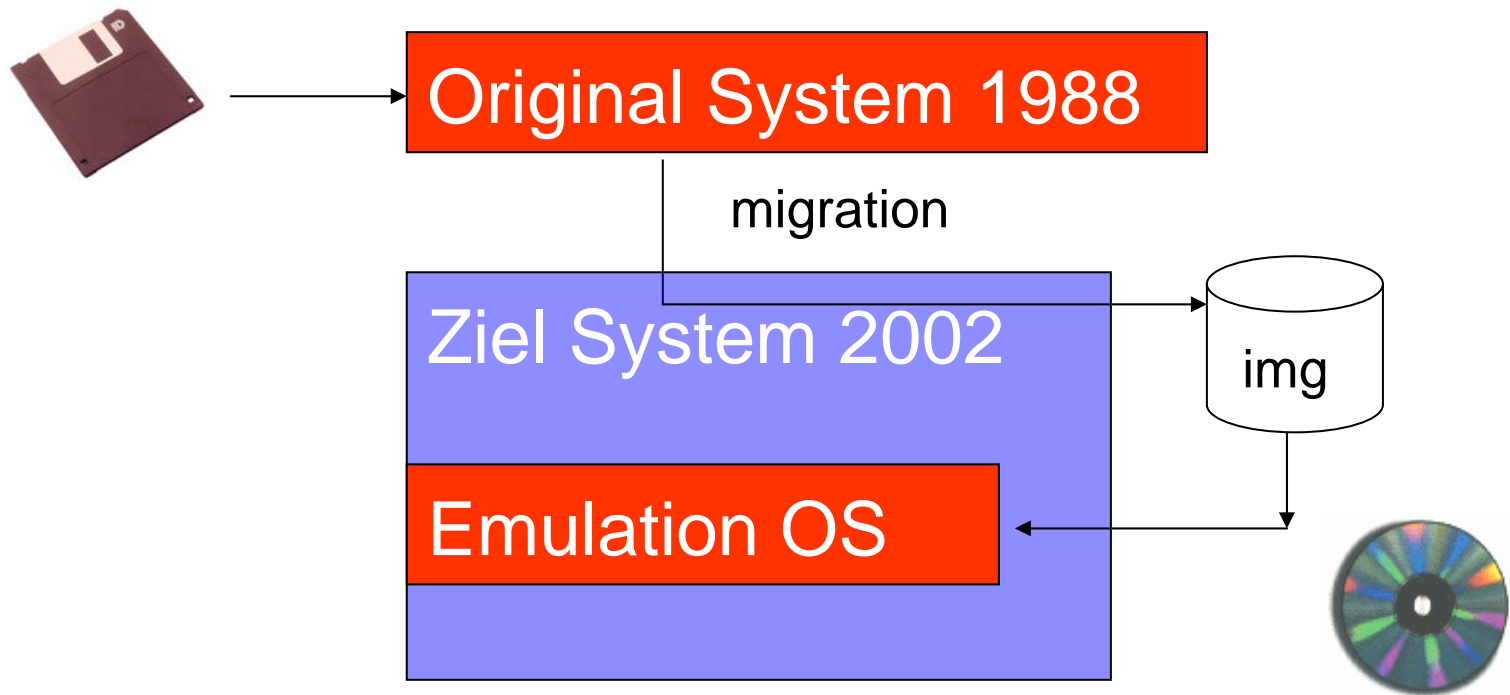


Handfeste Probleme in der DNB

- Von 1971 bis 1998: 18 verschiedene Diskettenformate zwischen 8" Zoll und 3,5".
- Disketten und andere Datenträger verlieren ihre Information, werden unlesbar.
- Die Hardware (Lesegeräte, Verarbeitungsmaschinen, Ausgabegeräte) funktioniert nicht mehr. Ein Computermuseum mit dem Ziel, alles betriebsfähig zu halten, ist unübersehbar und unsicher.
- (System) Software: Steht die noch zur Verfügung? Welche überhaupt? Spezifische Treiber, Erweiterungen, Hardwareanpassungen?
- Dateien: Mit welcher Software wurde die Datei erzeugt? Mit welcher Software kann ich sie nutzen?

In systematischer Perspektive

z.B. C64 / AMIGA



Migration & Emulation



Gesetz über die Deutsche Nationalbibliothek (DNBG) vom 22. Juni 2006



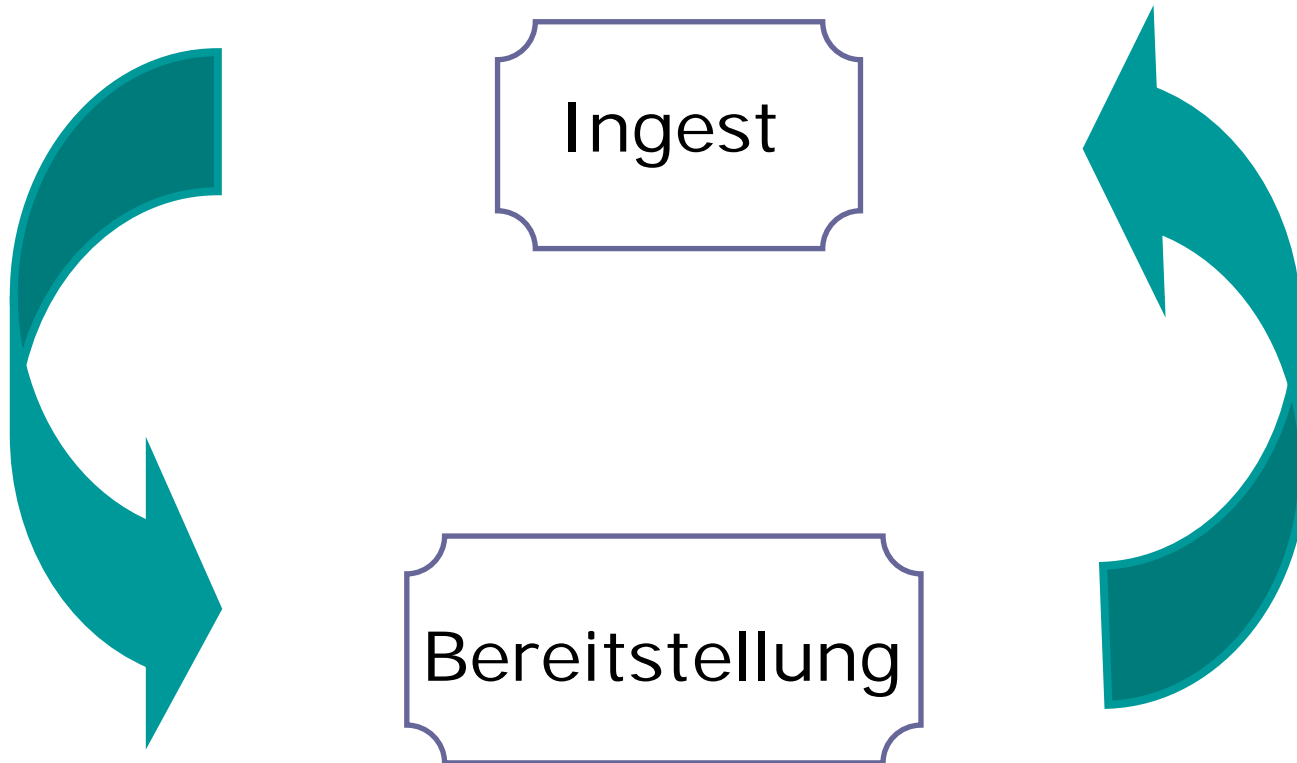
„Die Bibliothek hat die Aufgabe, die ab 1913 in Deutschland veröffentlichten Medienwerke ... im Original zu sammeln, zu inventarisieren, zu erschließen und bibliografisch zu verzeichnen, auf Dauer zu sichern und für die Allgemeinheit nutzbar zu machen ...

Medienwerke sind alle Darstellungen in Schrift, Bild und Ton, die in körperlicher Form der Öffentlichkeit zugänglich gemacht werden. ...

Medienwerke in unkörperlicher Form sind alle Darstellungen in öffentlichen Netzen. ... „

Ausgenommen: *Filmwerke (Ausnahmen im Musikbereich), Rundfunkproduktionen, reine Öffentlichkeitsarbeit, Bestandsverzeichnisse, Warenangebote, Serviceangebote, Betriebsprogramme oder Anwendersoftware ohne Inhalt, Spiele ...*

Ingest und Archivierung/Bereitstellung



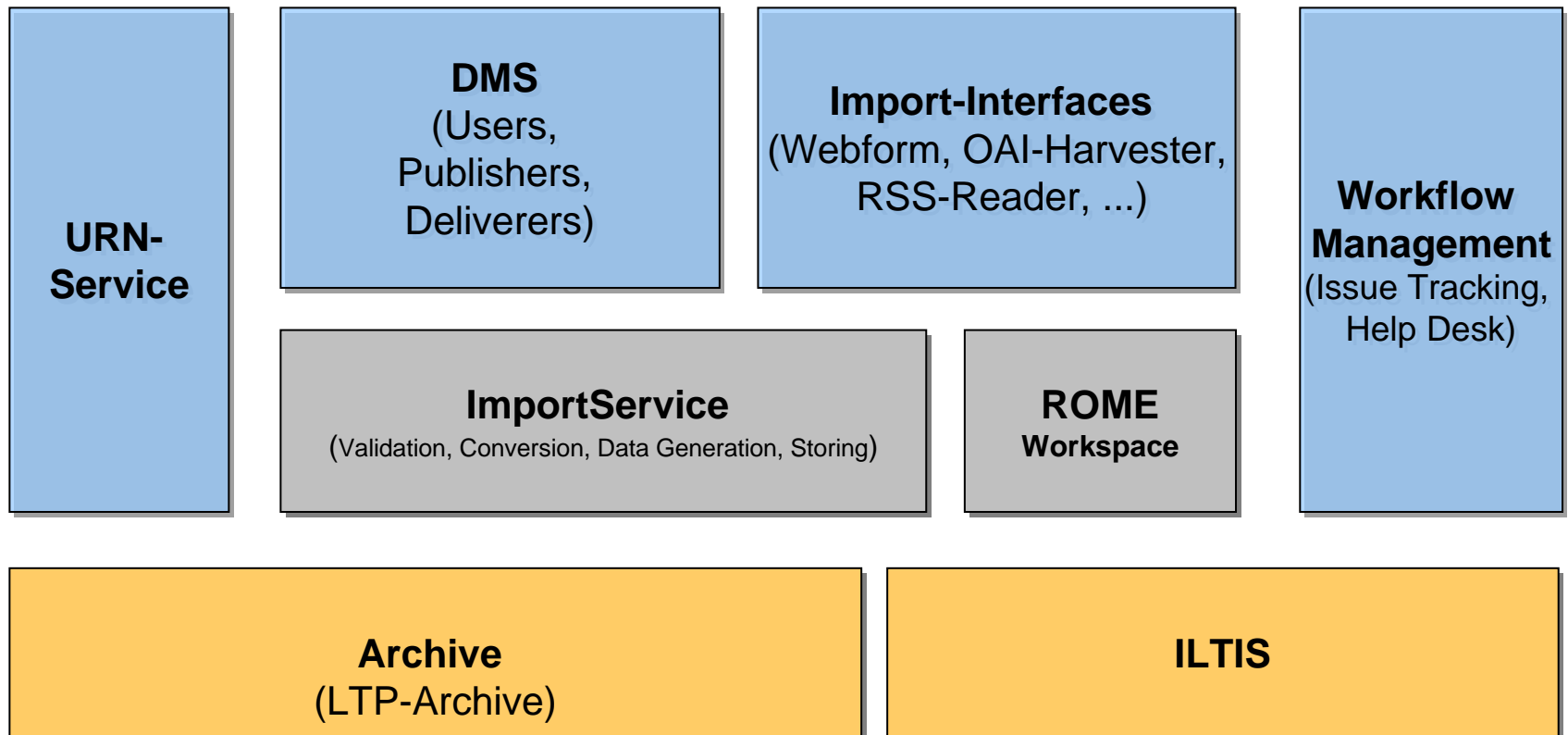
Sammlung * Erschließung

- Festlegung Sammelumfang und –strategie
→ Pflichtablieferungsverordnung / Sammelrichtlinien
→ push / pull ; objektbezogen / flächig
- Aufbau einer Harvesting-Infrastruktur
für „fokussiertes“ und „flächiges Harvesten“
(in Kooperation mit anderen Nationalbibliotheken und dem
Internet Archive)
und Optimierung der objektbezogenen Technologien
- Aufbau eines generischen Import-Verfahrens für Objekte +
Metdaten
- Angemessene Erschließungsautomatismen / Verknüpfung
zu existierenden Informationsquellen

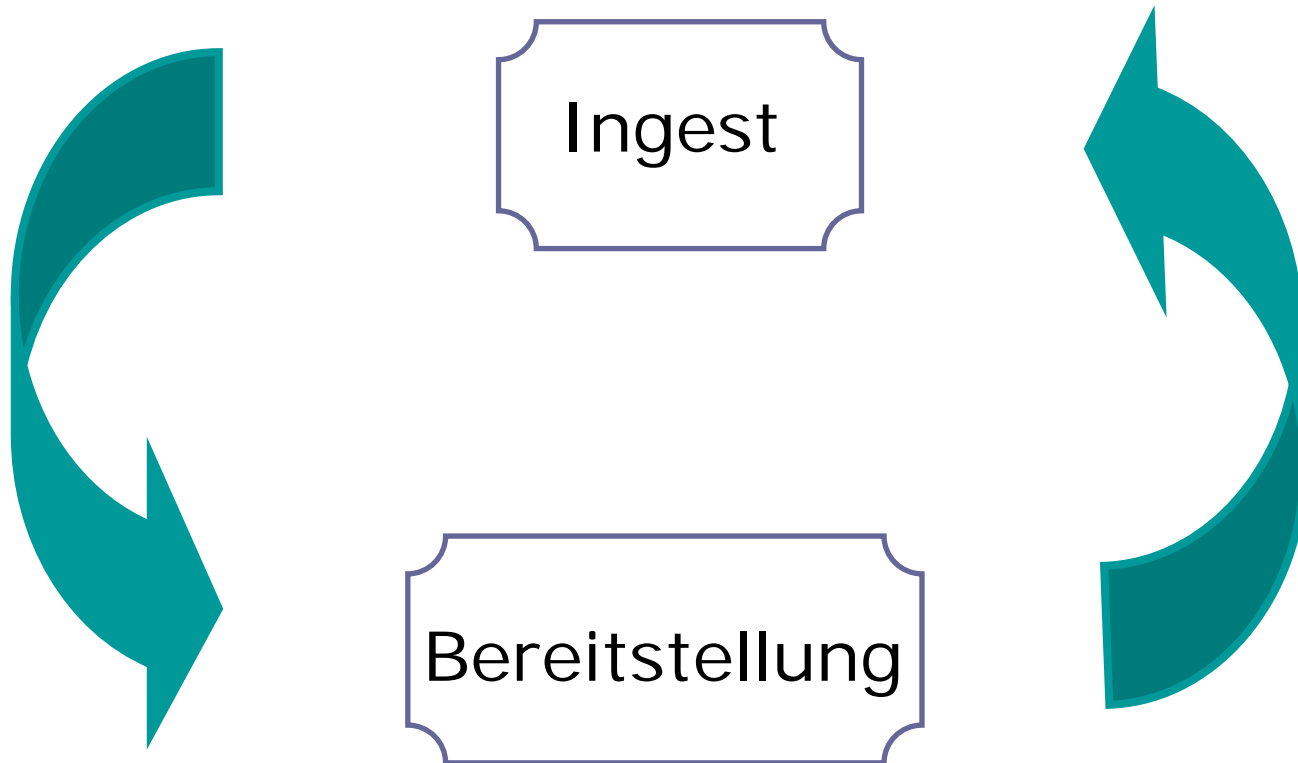
Lieferverfahren: Arbeitsschritte

- Definitionen: Metadatenkernset (bibliografische, rechtliche, technische Angaben)
- Ablieferung: formularbasiert oder über akzeptierte Protokolle und Formate
- Identifizierung der abliefernden Stelle / Verzeichnung des Ablieferers → Deliverymanagementsystem
- Tracking- und Workflowmanagementsystem
- Validierungsmechanismen / automatisierter Ablieferungsdialog

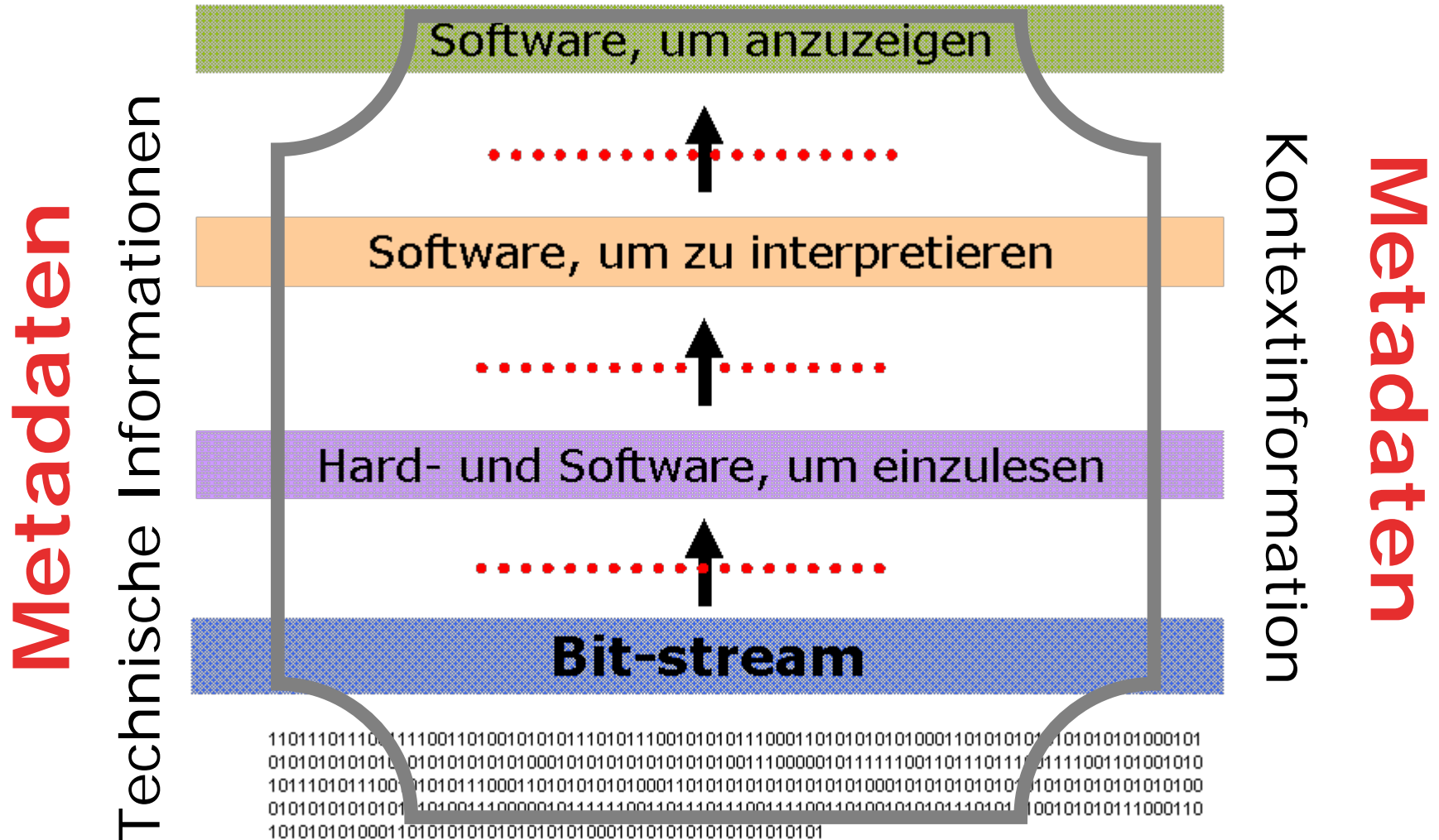
Einliefer-Service



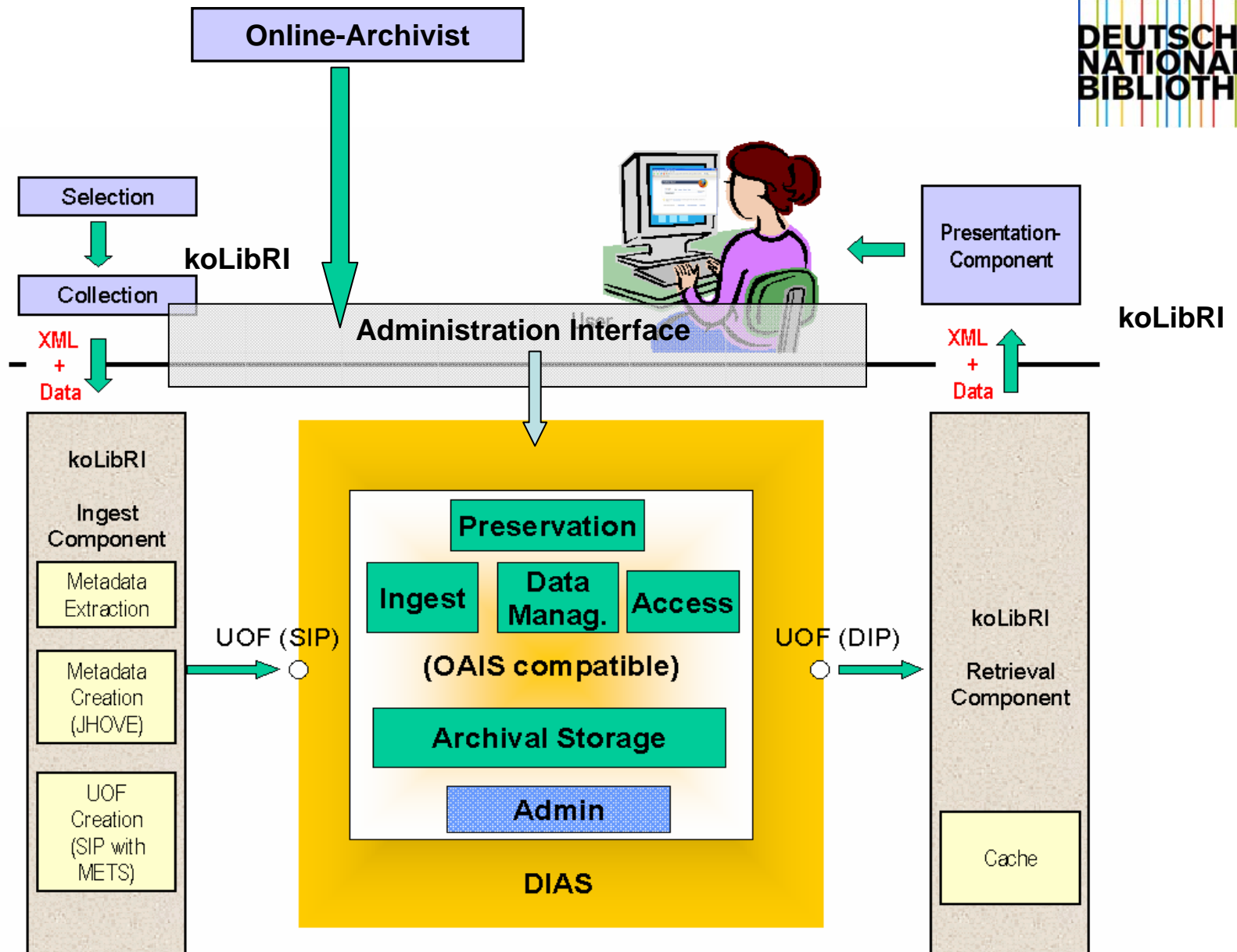
Ingest und Archivierung/Bereitstellung



Ein digitales Objekt anzeigen



- Forschungs- und Innovationsprojekt des BMBF, Fördervolumen: 4,6 Mio €, Laufzeit: 1.7.2004 – 30.6.2007
- Projektziel: Aufbau einer von Gedächtnisorganisationen nachnutzbaren technischen und organisatorischen Infrastruktur zur Sicherung der Langzeitverfügbarkeit elektronischer Publikationen, Perspektive: > 100 Jahre
- Entwicklung einer standardisierten Lösung
 - nachnutzbarer Transfer für andere Institutionen
 - vereinfachte Kooperation zwischen Bibliotheken Leitbegriffe: Kooperation, Nachnutzbarkeit, Nachhaltigkeit
- Erweiterung des IBM Assets DIAS-Core um Open-Source basierte Komponenten: flexibler Einsatz durch Modularität der Lösung
- Prozessintegration z.B. Automatisierung der Einliefererroutinen, Optimierung für Massenverfahren



Realisierte Konzepte in kopal : Modularisierung + Standards

- 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

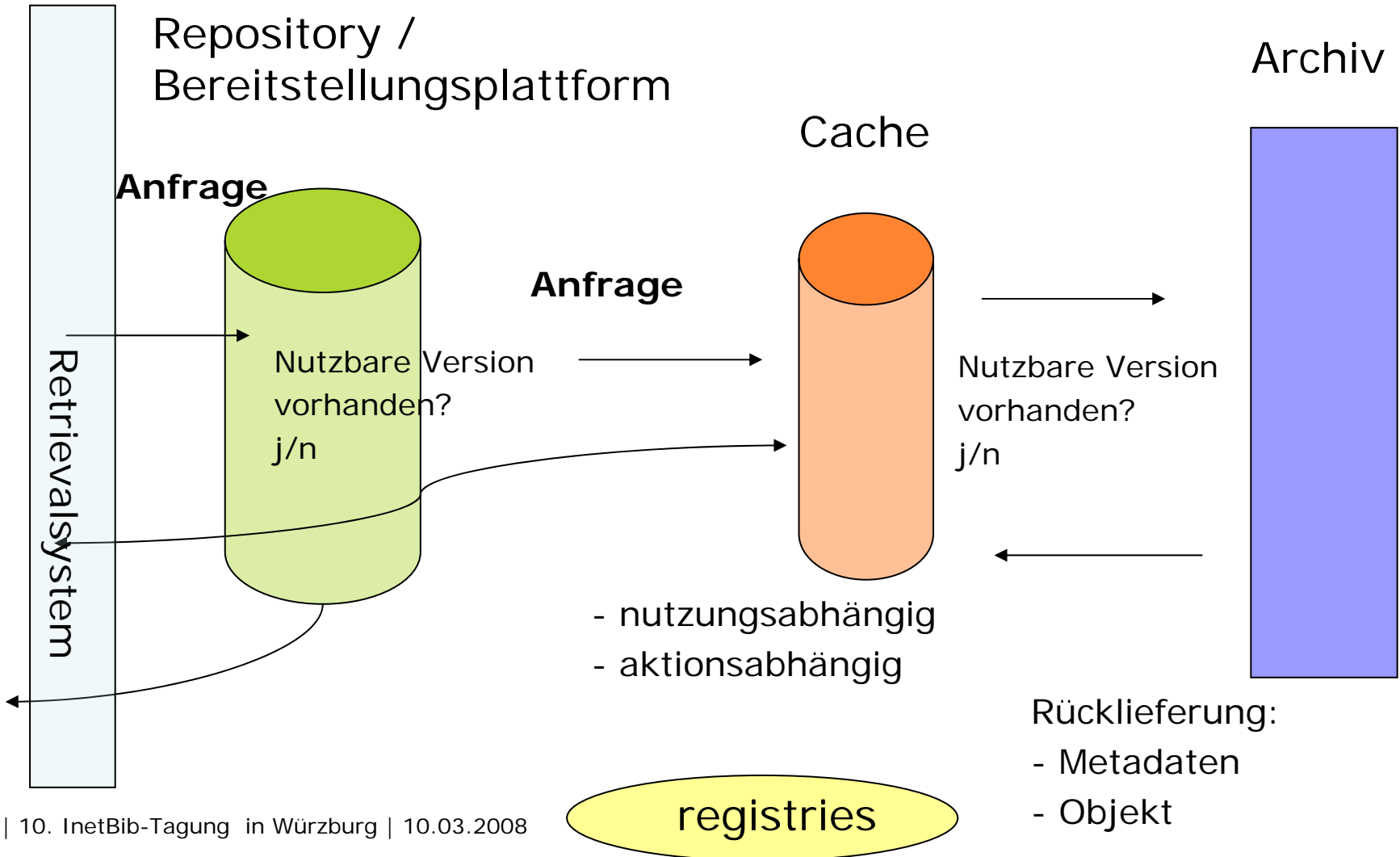
- Universelles Objektformat

- OAIS-Referenzmodell konsequent implementiert

Migrationsszenario

- Ein laufender Dienst reagiert auf eine Migrationsempfehlung von einer externen Quelle
- Die Anforderungen für die Migration werden in eine passende Abfrage für das Archivsystem transformiert (z. B. konvertiere alle nicht-animierten, schwarz-weißen GIF-Dateien zu JPEG2000)
- Jedes Objekt der Antwortliste von DIAS wird als DIP im Universellen Objektformat abgefragt
- Die METS-Datei im UOF wird analysiert, Abhängigkeiten der Dateien werden erkannt und aufgelöst (z. B. GIF - HTML)
- Das vorgesehene Tool zur Konvertierung wird für jede betroffene Datei ausgeführt
- Neue METS-Datei wird erzeugt inkl. Historie der Migration in den Metadaten
- Neues SIP wird gepackt und in das Archivsystem eingespielt

Datenfluss schematisch



Archivierung * Bereitstellung

- Integration der vorhandenen Module
- Bereitstellungssystem in Stufen
 - Statische Formate wie PDF
 - Webarchiv(e)
 - Applikationsgebundene Objekte

DNB: Existente und neue Funktionen

- Deutsche Nationalbibliographie: Metadatenproduktion und –vertrieb
- Normdaten
- Standardisierung

- Content: Automatisierte Verknüpfungsverfahren
- Authority-files
- WEB-Standardisierung
- Langzeitverfügbarkeit digitaler Publikationen



Reinhard Altenhöner

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

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