

Architektur Digitaler Bibliotheken

- Anforderungen und Aufgaben
- Funktionsmodell
- Komponenten
- Digitale Objekte
- Systemarchitekturen / Realisierung
- **Ansatz** (Vorlesung insgesamt): Top-down
 - Identifikation von Leistungen/Diensten --- Komponenten, Verbindungen, Protokollen
 - Anforderungen an einzelne Komponenten
 - Grundlagen und Verfahren
 - ...Algorithmen und Datenstrukturen

G. Görz, FAU, Informatik 8

Anforderungen und Aufgaben

- Globale Sicht - drei Ebenen:
 - Lieferanten (Autoren, Verlage,)
→ *Produktionsperspektive*
 - „System Bibliothek“
 - Nutzer
→ *Rezeptionsperspektive*
- Architektur
 - Funktionsmodell
 - Objektmodell
- Struktur: Implementierung

G. Görz, FAU, Informatik 8

Funktionsmodell

- Funktionalität des Systems für die Nutzer
 - Wozu verwendbar?
 - Welche Leistungen?
 - Welche Anpassungen?
- „Architektur“ als Anwendungsarchitektur
 - Invarianten sind konstitutiv
 - Was sich häufig ändert, wird zu Parametern
- vs. Struktur:
 - Komponenten
 - Algorithmen und Datenstrukturen, ...

G. Görz, FAU, Informatik 8

Aufgaben des „Systems Bibliothek“

- Sammeln
- Archivieren
- Recherchieren
- Bestand erschließen
- Bestand präsentieren
- Bestände verbinden
- Neue Dienstleistungen
 - Benachrichtigung
 - Personalisierte/gruppenspezifische Arbeitsplätze
 - Dokumenten-Management: Konvertierung, Klassifikation, Inhaltserschließung,...

G. Görz, FAU, Informatik 8

Rezeptions- und Produktionsperspektive

- „Geschäftsprozesse“ als Anwendungsfälle
 - Suchen
 - Bestellen
 - Liefern
 - Abrechnen
 - Nutzer betreuen
 - Lieferanten betreuen
- Objektorientierte Modellierung: Klassen und Operationen (UML-Diagramme s. Endres/Fellner, Kap.14)

G. Görz, FAU, Informatik 8

Nutzer-Erwartungen: Integration

- Erwartungen (nach Kaletta)
 - Informationen sollen unmittelbar am Arbeitsplatz verfügbar sein; Option für zeit- und ortsunabhängigen Zugang
 - Zugriff soll sich an Bedarf und Arbeitsweisen des Benutzers orientieren
 - Globalisierter Wissenschaftsbetrieb
- Integration mit Arbeitsformen wie
 - eMail, Telekonferenz, eLearning, eCommerce, Textverarbeitung, Surfen,...

G. Görz, FAU, Informatik 8

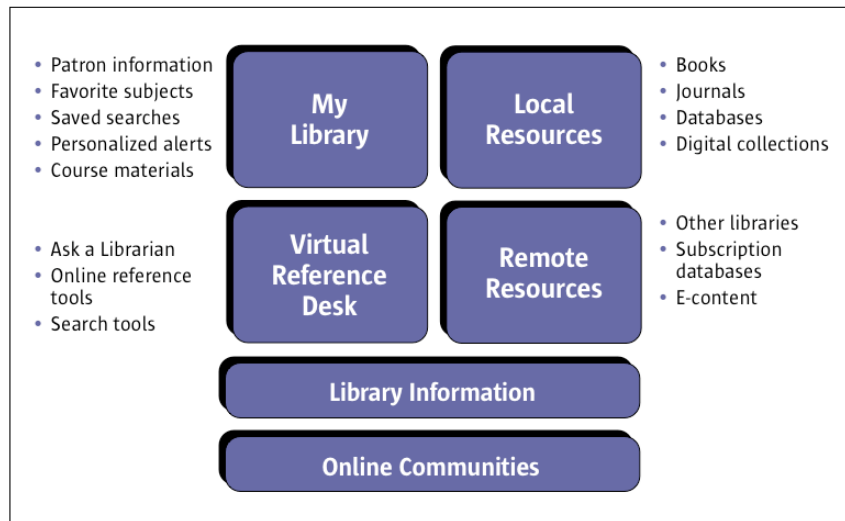
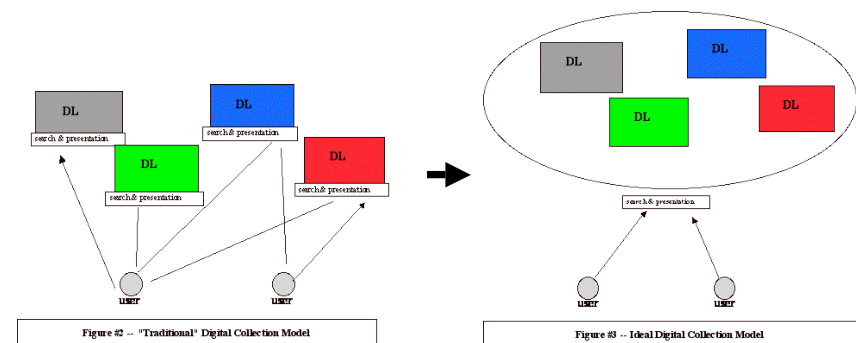


Figure 4. Library Portal Components

G. Görz, FAU, Informatik 8

© Sun

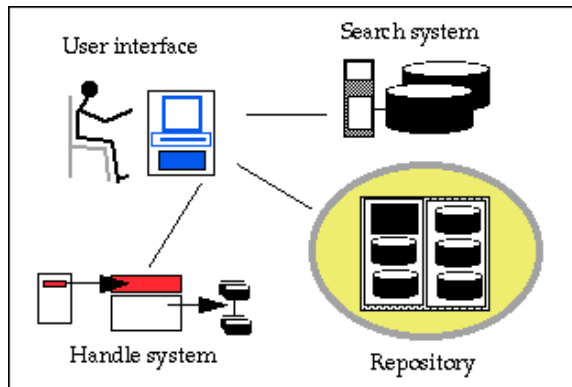
...integrierte Sicht



© Howard Besser

G. Görz, FAU, Informatik 8

Realisierung



© Arms

G. Görz, FAU, Informatik 8

Digitale Bibliotheken: Komponenten - **Rekapitulation**

- Akquisition
- Objektspeicherung- und -zugriff
- Metadaten-Modellierung
- Katalog- und Thesauruskonstruktion
- Anfrage
- Suchmaschinen für
 - strukturierte Daten
 - unstrukturierte Daten
- Metadatenzugriff und Objektauslieferung

G. Görz, FAU, Informatik 8

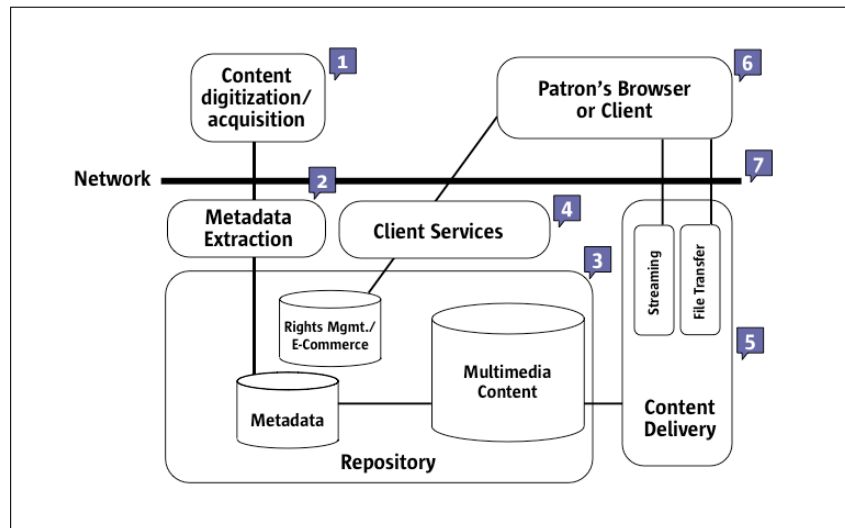
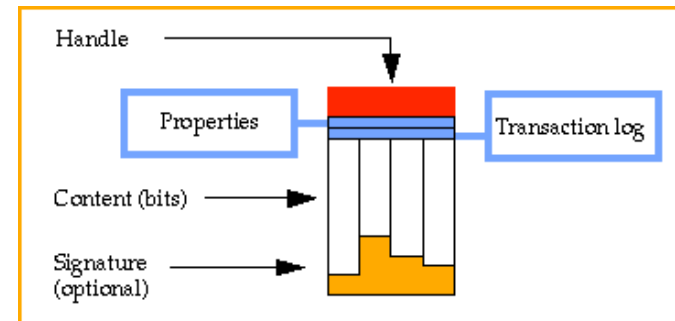


Figure 2. Functional Components of a Digital Library

G. Görz, FAU, Informatik 8

Aufbau digitaler Objekte

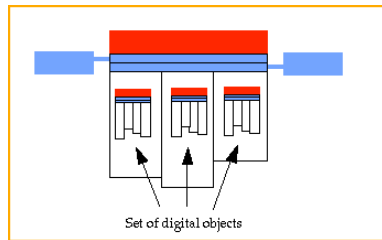


G. Görz, FAU, Informatik 8

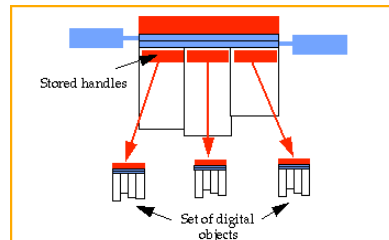
© Sun

© www.cnri.reston.va.us

Aggregierte digitale Objekte



Composite object

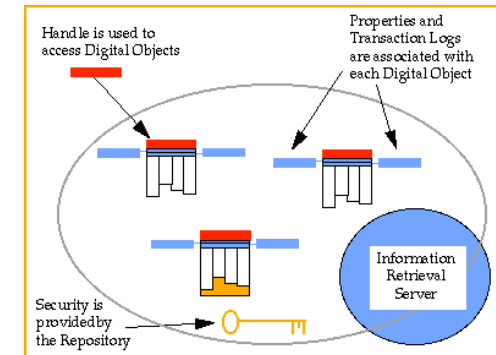


Meta-object

© www.cnri.reston.va.us

G. Görz, FAU, Informatik 8

„Repository“ (Objektspeicher)



© www.cnri.reston.va.us

G. Görz, FAU, Informatik 8

Digitale Objekte: Eigenschaften

- Datenvolumina

- Seite (ASCII-) Text	5 KB
- Seite Image Scan	25 MB
- Buch mit 200 Seiten	5 GB
- Minute Musik	10 MB
- Stunde Musik	600 MB
- Stunde Video	15 GB
- Kinofilm (Masterkopie)	150 TB

G. Görz, FAU, Informatik 8

Digitale Objekte: Eigenschaften

- Kosten

- Digitalisierung
 - Eine Seite (Image Scan): 1,50 Euro
 - 1000 Bücher, je 200 Seiten: 300.000 Euro
- Archivierung (bez. auf 10 Jahre)
 - 1 Terabyte (= 10^{12} Bytes): 20.000 Euro/Jahr
 - 5 Terabyte (= 1000 Bücher): 100.000 Euro/Jahr
- *Fazit*
(Mindest-) Kosten eines digitalen Buchs in 10 Jahren: **1300 Euro**

G. Görz, FAU, Informatik 8

Konsequenzen: Technologie??

- „Moore'sches Gesetz“
 - Verdopplung der Transistoren-Anzahl auf einem Halbleiterelement alle 18 Monate
 - Damit:
 - Rechenkapazität steigt in 10 Jahren 100-fach
 - in 20 Jahren 10.000-fache Steigerung
- Vermutung: optisch-technische Basis
 - Optische Speicher (3D, Petabytes = 10^{15} Bytes)
 - Optische Netze (Terabits/sec)
 - Parallele/optische Rechner (TeraMIPS)
- **Aber:** Architekturen skalierbar?

G. Görz, FAU, Informatik 8

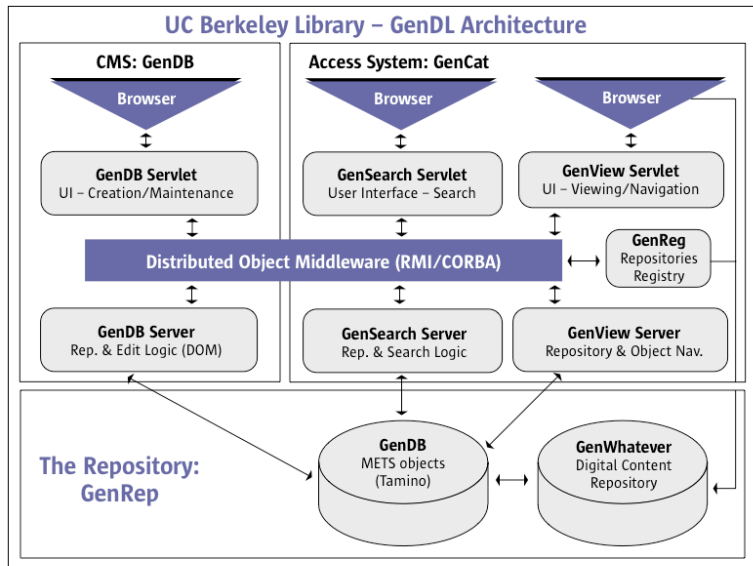


Figure 5. GenDL Architecture

Beispiele für Systemarchitektur / Realisierung

- Berkeley
- Harvard
- Stanford
- Dspace (MIT/HP) [open source]
- Greenstone DL Toolkit [open source]
- Integration in „Web Services“

G. Görz, FAU, Informatik 8

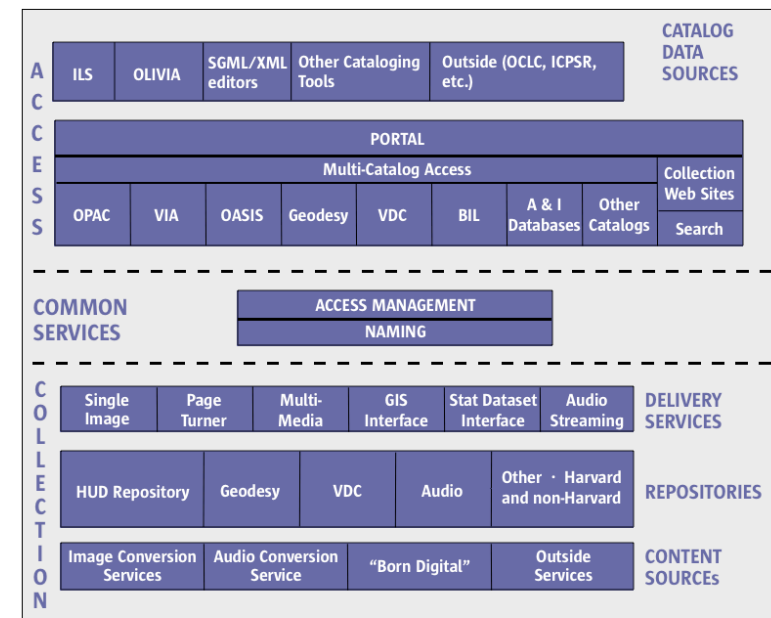


Figure 6. Harvard University Digital Library Architecture

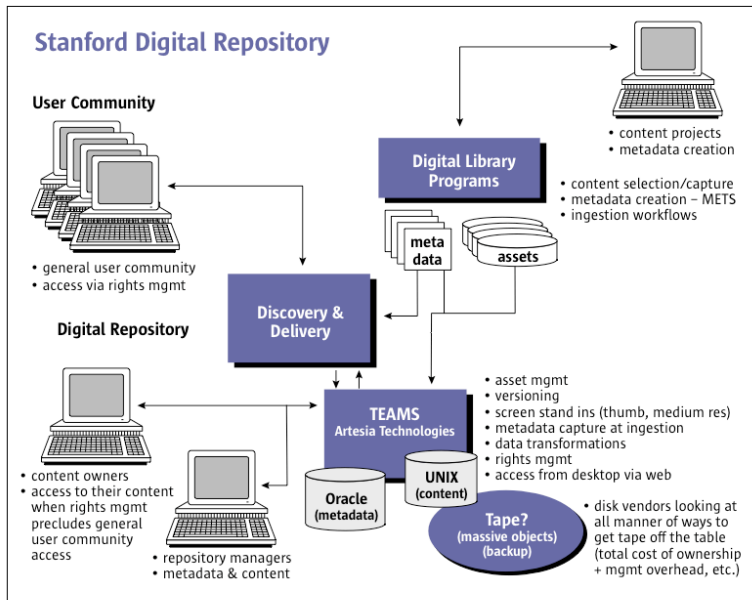


Figure 7. Stanford Digital Repository Architecture

© Sun

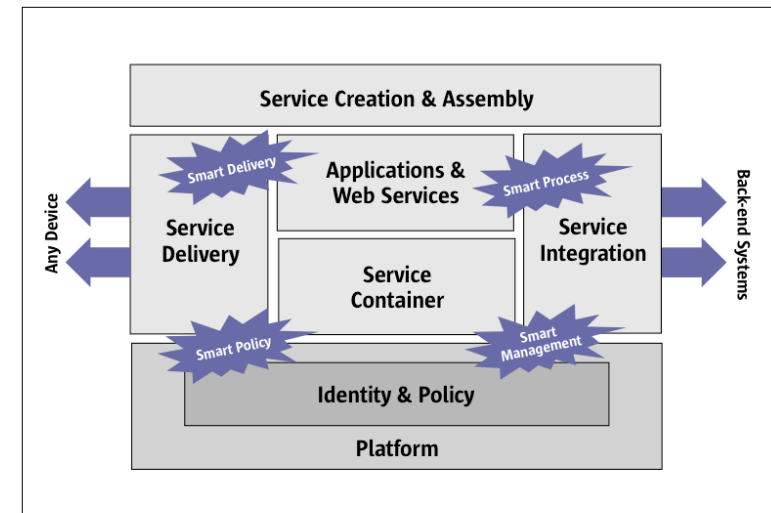
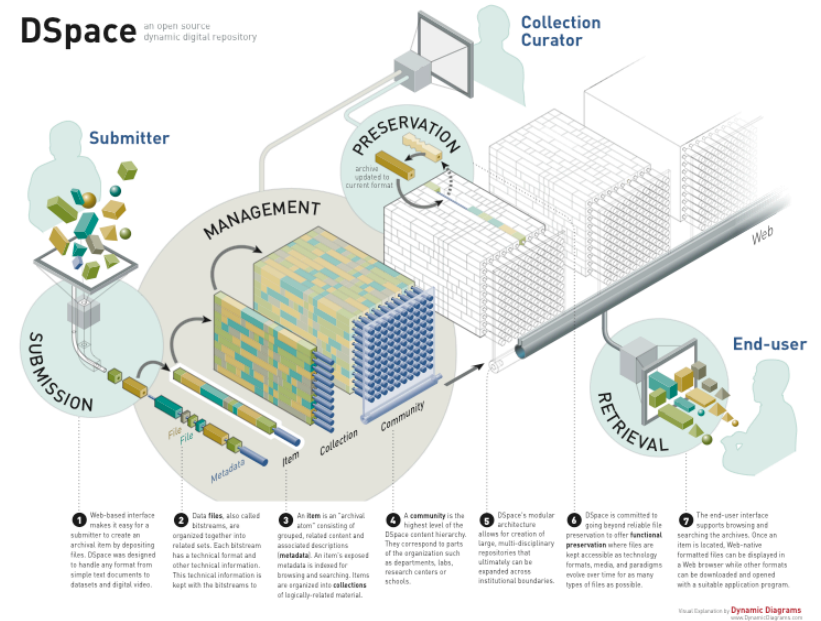


Figure 8. Components of the Sun ONE Smart Web Services Architecture

© Sun

Greenstone DL Toolkit

- Baukasten: Organisation und Publikation digitaler Informationsobjekte in voll durchsuchbarer, metadaten-gesteuerter digitaler Bibliothek (=> Bainbridge/Witten)
- Open Source für alle gängigen Systemplattformen
- Interoperabilität: standardkonform (METS!), OAI Protocol for Metadata Harvesting, DSpace-Austausch
- Benutzer- und Administrator-Schnittstellen
- Alle gängigen Metadatenstandards und Objektdaten-Formate enthalten