

OAI in Deutschland – eine Bestandsaufnahme

Silke Schomburg (HBZ)





OAI – eine Bestandsaufnahme

Open Archives Initiative – Bedeutung und Idee

- Mit der Entwicklung der OAI-Technologie ist eine einfache Schnittstelle zum Austausch von Metadaten vorhanden. Die Implementierung ist durch schon vorhandene Tools recht einfach und das lokale Anpassen durch von schon anderswo entwickelten Skripten erleichtert
- OAI-Technologie: Mit Hilfe von sechs Protokollanfragen werden Metadaten eingesammelt, neue Informationen über Dokumente aktualisiert und auf gelöschte Inhalte verwiesen
- Diese Technik eignet sich aufgrund der einfachen Implementierbarkeit besonders gut als Schnittstelle für den Datenaustausch lokaler Dokumenten- und Digitalisierungsserver





OAI – eine Bestandsaufnahme

Open Archives Initiative – Bedeutung und Idee

- Verbessert den Informationsaustausch in den Wissenschaften und unterstützt die Open Access-Bewegung
- Vernetzt elektronische Archive weltweit und ermöglicht flexibel und unproblematisch den Austausch von Inhalten bzw. Metadaten
- Elektronische Archive stellen strukturierte Informationen als Metadaten über eine genau definierte Schnittstelle zur Verfügung (Data-Provider)
- Andere Organisationen sammeln die Metadaten ein, führen sie mit denen anderer Archive zusammen, reichern sie an und stellen sie auf einer einheitlichen Oberfläche zur Recherche zur Verfügung (Service-Provider)
- Nutzt existierende Standards wie http, XML und DC
- Unabhängig von der Software des elektronischen Archivs





OAI – eine Bestandsaufnahme

Open Archives Initiative – Geschichte

- Treffen der Universal PrePrintService (UPS)
Santa Fe Convention (1999): Technische Übereinkunft zum Metadaten Harvesting
- Version 1.0: Jan 2001 OAI-Protocol for Metadata Harvesting (OAI-PMH)
- Version 2.0: Juni 2002, nicht mehr experimentell, sondern stabile Basis zur Produktion von Software
- Große Zahl von Archiven weltweit unterstützt das Protokoll
- Genaue Anzahl nicht klar, siehe Liste unter: www.openarchives.org
- Viele OAI-Szenarien laufen in Intranets, z.B. in Wissensgesellschaften
- Juni 2002: 100 registrierte Archive, inzwischen Zahl verdoppelt





OAI – eine Bestandsaufnahme

Open Archives Initiative – Technik

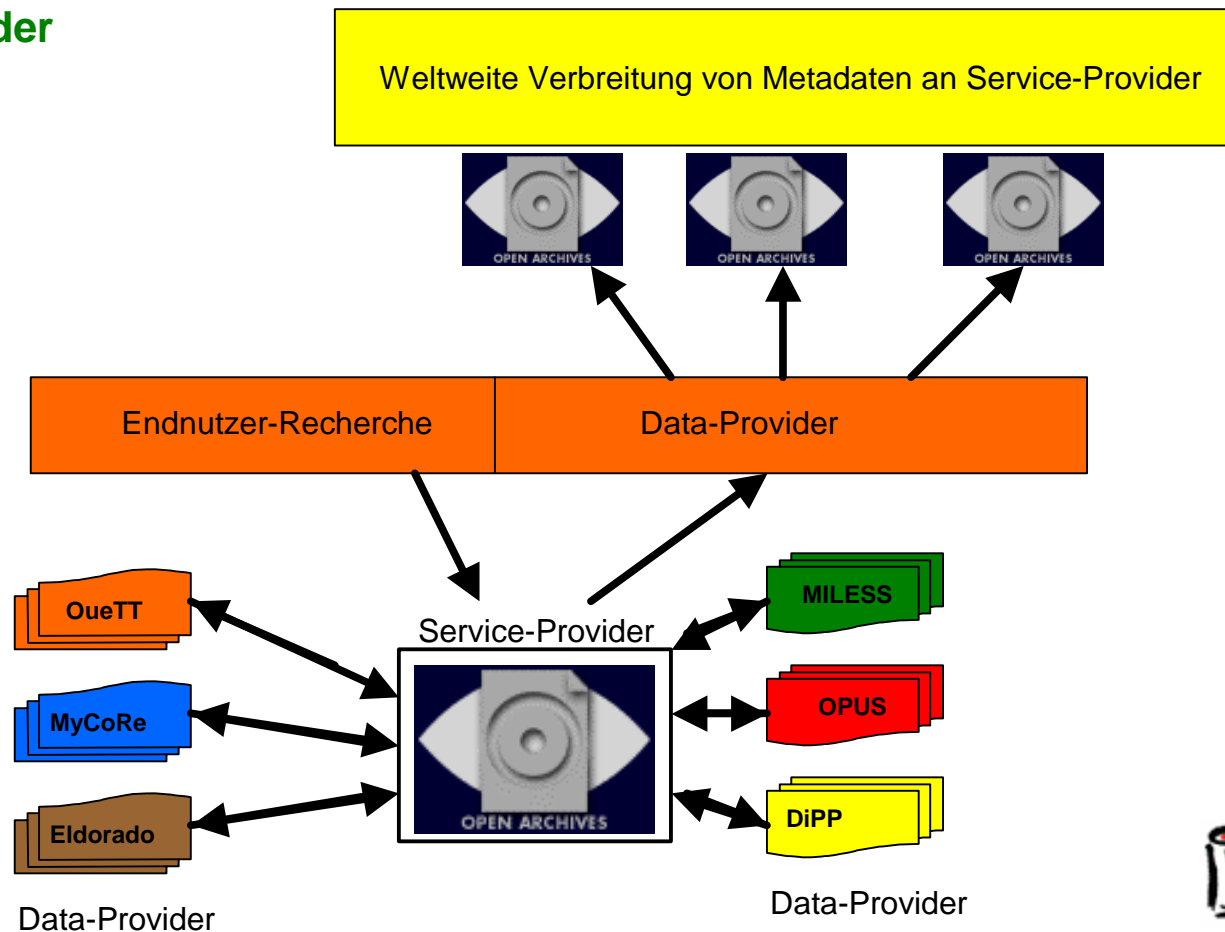
- Protokoll basiert auf http, XML als Ausgabeformat, sechs Anfragen an Data-Provider: Identify, ListMetadataformats, ListSets, ListIdentifiers, ListRecords, GetRecord
- Protokoll unterstützt jedes Metadatenformat, jedoch mindestens Dublin Core unqualified, besser Dublin Core qualified
- Optional: Sets, Protokoll unterstützt Einsammeln von Teilmengen
DINI-Arbeitsgruppe Empfehlung zur Verwendung von Sets, z.B. Publikationstyp (Artikel, Vorträge...) oder inhaltliche Sets nach DDC (Chemie, Biologie...)
- Wichtig: Sets, eindeutige Identifier in den Archiven, flexible Metadaten Formate, selektives Harvesting (z.B. nach Datum und Sets)





OAI – eine Bestandsaufnahme

Modell Data-Provider – Service-Provider





OAI – eine Bestandsaufnahme

Open Archives Initiative – Data-Provider

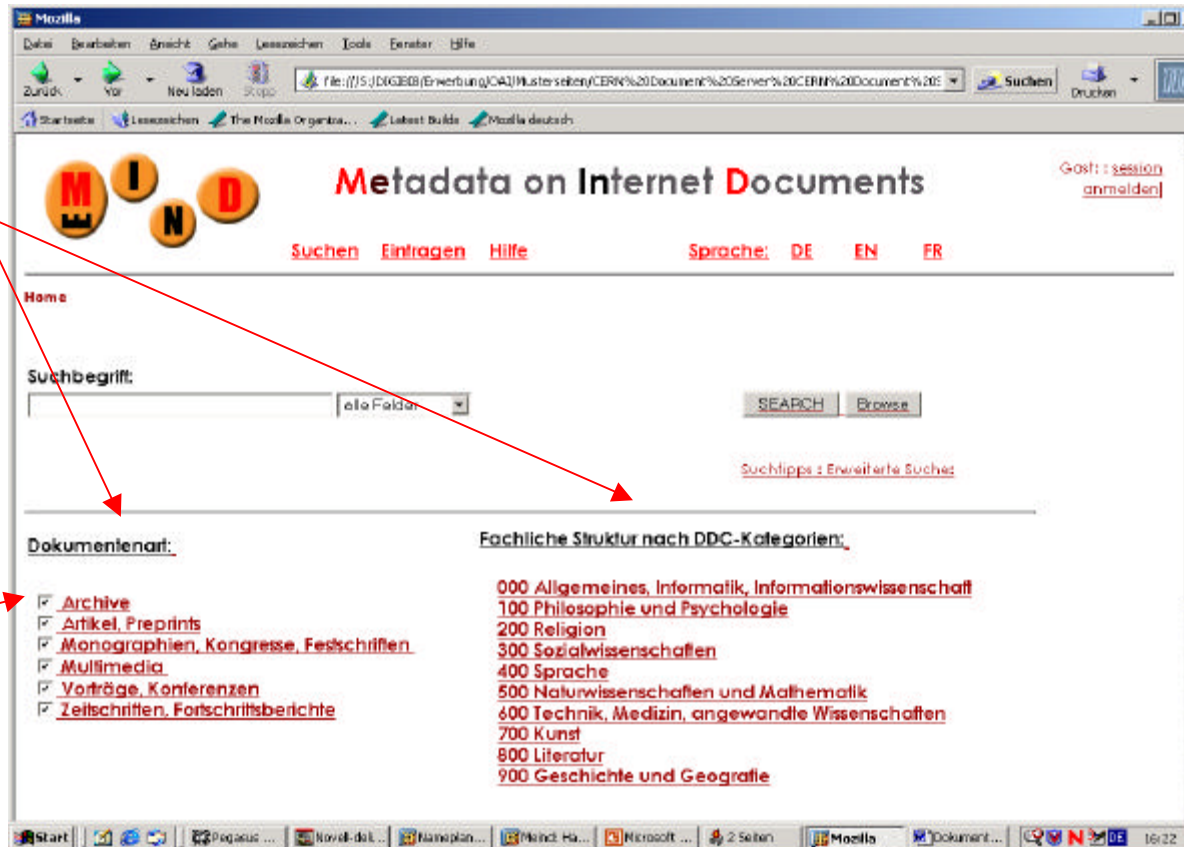
- Zahl noch immer nicht genau zu ermitteln, da freiwillige Registrierung (derzeit: 200)
- Unterschiedliche Dokumentenserversoftware: z.B. OPUS, MyCoRe
- Jedoch nicht alle Server haben OAI-Schnittstellen aktiviert, Verwendung des Servers nur im Intranet der Hochschule
- Umfrage über die InetBib, fast keine Resonanz
- Direkte Anfrage per Mail an einzelne Hochschulen: noch geringere Resonanz
- OAI-Schnittstelle: 28 Standorte sicher, bei weiteren 20 nehmen wir sie wegen der verwendeten Software an
➡ bitte Mail an HBZ für Fertigstellung der Übersicht





OAI – eine Bestandsaufnahme

- Struktur: gezielte Suche oder Browsing nach Dokumentenart (Beispiel: Dissertation) oder nach DDC-Kategorien möglich
- Per Mausklick kann auch in den tieferen Ebenen von DDC gesucht / oder ein Browsing durchgeführt werden
- Alternativ kann die Recherche auch auf deutsche oder weltweite Datenprovider ausgeweitet werden



www.meind.de





OAI – eine Bestandsaufnahme

The screenshot shows the MeIND search interface in a Mozilla browser window. The search term 'Hund' is entered in the search box. Below the search box, there are options for 'Suchen Kollektionen' (Archive, Monographien, Zeitschriften) and 'Sortieren nach' (das Letzte zuerst). The search results section shows 'Übersicht von Ergebnissen: 54 Datensätze wurden in 0.10 Sekunden gefunden.' and 'Dissertationen, 54 Datensätze gefunden.' The first result is a dissertation titled '1. Etablierung von Laktat- und Pyruvatmessung im Plasma und Liquor cerebrospinalis zur Diagnostik von mitochondrialen Erkrankungen beim Hund'.

Recherche nach „Hund“

54 Treffer, nach Dokumenttypen sortiert in der Kurzanzeige ausgegeben





OAI – eine Bestandsaufnahme

Open Archives Initiative – Verzeichnung der Inhalte / Metadaten

Deutsche Bibliothek: freiwillige Ablieferung von elektronischen Dissertationen, nicht alle Hochschulen liefern bzw. liefern aktuell (manchmal Lücken von bis zu einem halben Jahr), Änderung zur Pflichtablieferung erst mit dem neuen Gesetz der DDB

Service-Provider / Suchmaschinen / verteilte Suche

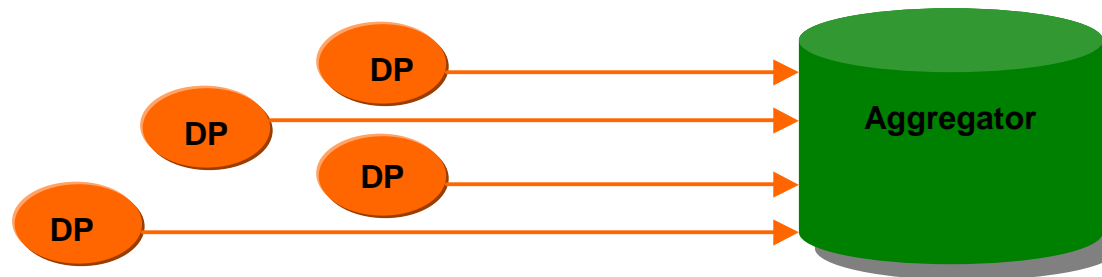
- Humboldt-Universität: 12 deutsche Data-Provider
- Digitale Bibliothek Thüringen: 6 regionale Hochschulen
- Virtueller Medienserver (BSZ): 11 Standorte, haupts. Baden-Württemberg
- MeIND (HBZ): alle deutschen Dissertationen, NRW
- Base (UB Bielefeld): 20 deutsche Anbieter
- Forschungsportal.net: angeblich alle Dissertationen verzeichnet
- TheO: 46 Standorte indexiert
- OASE (Karlsruhe): 21 deutschen Standorte





OAI – eine Bestandsaufnahme

Open Archives Initiative – Recherche nach Dissertationen (2004)



Ergebnisse der Erhebung: Service-Provider / Suchmaschinen / verteilte Suche

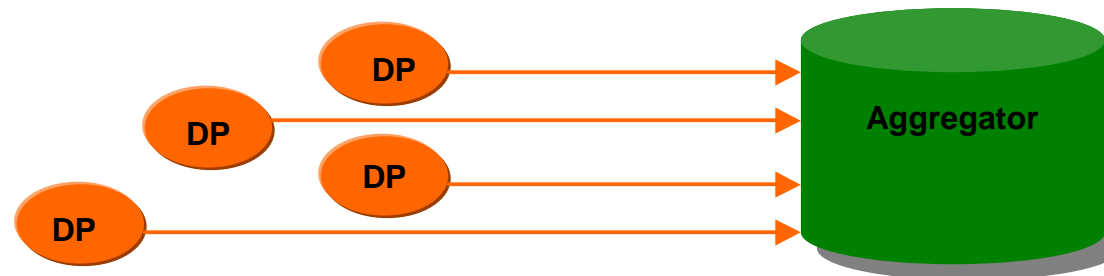
1. Platz TheO: sehr gute Ergebnisse, Fehler bei 10 Standorten

1. Platz MeIND (HBZ): sehr gute Ergebnisse, Fehler bei 7 Standorten





OAI – eine Bestandsaufnahme



2. OASE: durchsucht derzeit nur 21 Server / Technik des KVK (http-Abfragen) jedoch häufig nicht von allen Standorten Rückmeldungen zu den Recherchen
2. Base (UB Bielefeld): bisher unzureichend, da nur 20 deutsche Anbieter
3. Humboldt-Universität: OAI Service-Provider, bisher nur 12 Standorte (Einbindung von weiteren OAI Daten-Providern wünschenswert)
3. Virtueller Medienserver (BSZ): nur 11 Standorte, Schwerpunkt Baden-Württemberg
4. Digitale Bibliothek Thüringen: nur 6 regionale Hochschulen, Schwerpunkt regional
-Forschungsportal.net: Suchmaschine, Treffer nur in Ausnahmefällen (z.B. Oldenburg





OAI – eine Bestandsaufnahme

Open Archives Initiative – Verzeichnung anderer Inhalte

- Im Fokus stehen nicht nur Dokumentenserver, sondern auch Zeitschriften-, E-Learning- und Digitalisierungsserver
- Leider haben diese Server nur in Ausnahmefällen OAI-Schnittstelle realisiert
- Beispiel: DFG-geförderte deutsche Zeitschriftenprojekte
 - GAP: OAI-Schnittstelle
 - GMS: keine entsprechende Schnittstelle
 - zepl (Zentrum für elektr. Publizieren in den Literaturwissenschaften):
OAI-Schnittstelle geplant
- Beispiel: DiPP (zentral in NRW gefördert): OAI-Schnittstelle





OAI – eine Bestandsaufnahme

Open Archives Initiative – Verzeichnung anderer Inhalte

- E-Learning-Inhalte liegen teilweise auch auf den Dokumentenservern, jedoch in der Regel kein Zugang, sondern IP- oder Passwort geschützt
- Originäre E-Learning-Server haben ebenfalls keine OAI-Schnittstelle
- Ähnlich bei Digitalisierungsservern, keine entsprechenden Schnittstellen (Ausnahmen: Z.3950-Schnittstellen) und oft nur nach Projektende nicht weiter gepflegte Webseiten
- Lösung: DFG-Projekt ab 2005 zur postkoordinierenden Zusammenführung von Digitalisierungsprojekten (Partner: Verbundzentrale GBV, HBZ, AG Sammlung Deutscher Drucke)
Aufforderung zur nachträglichen Implementierung von OAI-Schnittstellen, ansonsten im Projekt aufwändiger Einsatz von Suchmaschinentechnologie





OAI – eine Bestandsaufnahme

Open Archives Initiative – neue Tendenzen

- Qualified Dublin Core anstatt unqualified DC ermöglicht bessere Dienstleistungen des Service-Providers
Ziel: Einhalten der Metadaten Sets, vgl. DINI-Empfehlungen als allgemeine Richtlinien
- Nicht nur DC, sondern auch MAB-Generierung über OAI und Austausch von technischen und Strukturmetadaten
- Neue URN werden von DDB über die OAI-Schnittstelle geharvestet
Prototyp: OPUS und Dokumentenserver der HU
- LZA: nicht nur Metadaten werden ausgetauscht, sondern auch die elektronischen Objekte werden über den Link in den Metadaten übertragen (Projekt Landesbibliotheken - DDB)





OAI – eine Bestandsaufnahme

Vielen Dank für Ihre Aufmerksamkeit!

Für weitere Fragen und Auskünfte:

Dr. Silke Schomburg, Tel. 0221-40075-176, schomburg@hbz-nrw.de

Guido Pelzer, Tel. 0221-40075-153, pelzer@hbz-nrw.de

Katharina Dähling, Tel. 0221-40075-188, daehling@hb-nrw.de

HBZ Service-Provider: www.meind.de

