

Dokumentenserver: Praxisfragen

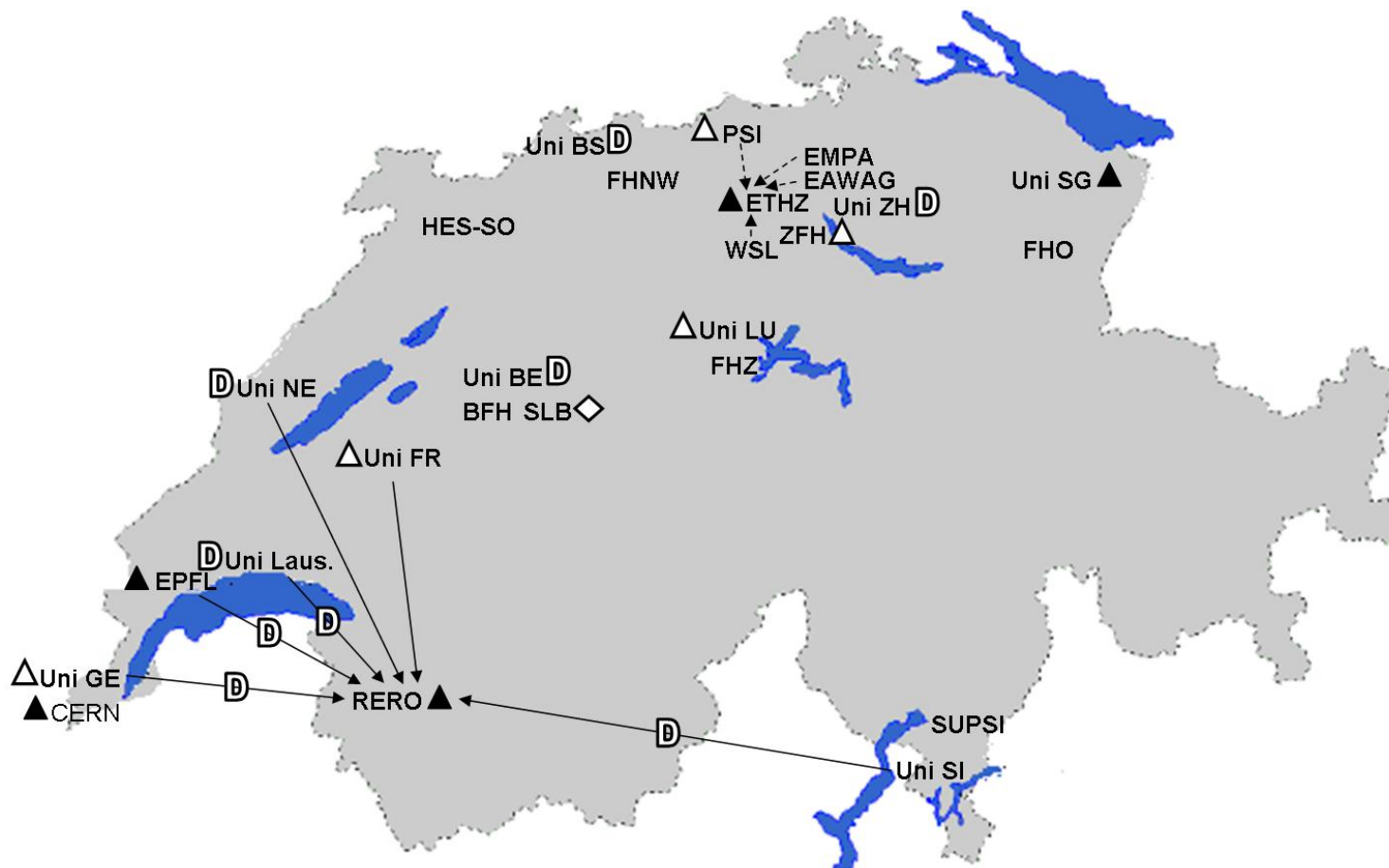
Workshop in Bern, 18. Mai 2006

Dr. Matthias Töwe
Konsortium der Schweizer Hochschulbibliotheken

Gliederung

- Situation in der Schweiz
- Offene Fragen zum Betrieb
 - Lokale Server vs. zentrale Dienste (z.B. RERO DOC)
 - Sind die nötigen Ressourcen dauerhaft verfügbar?
 - Akzeptable Formate
 - Metadatenstandards und Erfassung
 - Weltweite Sichtbarkeit
 - Suchmaschinen
 - OAI-PMH (Harvesting der Metadaten durch Service Provider)
- Fazit
- Perspektiven

Dokumentenserver CH, Juni 2005



Situation

- Alle Unis haben eigene Server oder nutzen zentrale Dienste
- Situation der Fachhochschulen unterschiedlich und wenig bekannt
- Es überwiegt bei weitem die Nutzung für Dissertationen, d.h. auch die grösseren Server werden noch wenig als Werkzeuge für Open Access im engeren Sinne genutzt, weil Ressourcen fehlen
- Nur wenige Schweizer Server sind bei internationalen Service Providern registriert und lassen ihre Metadaten harvesten
- Autorinnen/Autoren nutzen die bisherigen Angebote wenig – einzelne sind aber sehr interessiert
- Es werden bei genügend umfangreichen, sichtbaren und relevanten Inhalten erhebliche Zugriffszahlen registriert

Offene Fragen: Lokale Server vs. zentrale Dienste

- Eigene Server
 - + volle Kontrolle und Flexibilität bei technischer und formaler Gestaltung
 - + stärkere lokale Präsenz, grössere Kundennähe
 - ganzer Aufwand fällt lokal an, Kosten oft wenig transparent
- Nutzung eines zentralen Dienstes
 - + technischer Betrieb mit klaren Konditionen ausgelagert, u.U. mehr Ressourcen für Kundenbetreuung verfügbar
 - + Synergien mit anderen Einrichtungen, reduzierter Gesamtaufwand
 - Abhängigkeit von Dritten
 - Eigenleistung der Institution weniger sichtbar (→ die Hauptleistung sollte allerdings nicht der technische Betrieb sein)
- Mischform: Zentrales Angebot der *Metadaten* von lokalen Servern (→ Teilprojekt des E-Archiving-Projekts)

Metadatenserver

- Zentrales Angebot der *Metadaten* von lokalen Servern
- Voraussetzung:
 - OAI-Fähigkeit der beteiligten Server und Verfügbarkeit von Metadaten, aus denen sich ein relevanter Datensatz erzeugen lässt
- Ziel:
 - Dokumente können aufgefunden werden, auch wenn ihr Speicherort unbekannt ist und sie nicht von Suchmaschinen erfasst werden können
 - Sofern die Metadaten dies ermöglichen, kann ein Zugang nach Sachgebieten über alle Institutionen hinweg ermöglicht werden
 - Es entsteht eine Metadatensammlung, die gross genug für das Harvesting durch internationale Service Provider ist
- Aber: Grundsätzlich ist diese zentrale Sammlung nicht nötig, wenn
 - Jeder Server international sichtbar ist und geharvested wird
 - Auf einen eigenen Schweizer Auftritt verzichtet werden kann

Offene Fragen: Ressourcen

- Sind die nötigen Ressourcen dauerhaft verfügbar?
 - Oft werden Dissertationenserver von Personen nebenbei aufgebaut und betreut, die eigentlich andere Hauptaufgaben haben.
 - Hauptaufwand ist *nach dem Aufbau* nicht technisch, sondern wird für Information und Akquisition bei Autorinnen und Autoren benötigt. Er hängt ab von der Anzahl potentieller Autoren, von der Anzahl erwarteter Dokumente sowie von der Erschliessungs- und Bearbeitungsintensität und der erreichten Stabilität der Abläufe und ihrer Automatisierung.
 - Politische/strategische Entscheidung zur Positionierung der eigenen Einrichtung (Hochschule und Bibliothek)

Offene Fragen: Formate

- Akzeptable Formate
 - Grundsätzliche Ablehnung von Formaten in der Praxis schwierig
 - Aber: Es soll klare Empfehlungen für die Verwendung möglichst weniger Standardformate geben. Zusätzlich ist anzugeben, welche Verantwortung für unterschiedliche Formate übernommen wird. Wo es möglich ist, sollte eine Umwandlung in ein Standardformat erfolgen.
 - PDF ist aus Sicht der Archivierung streng genommen nicht das ideale Format, aber es sollte mit gewissen Einschränkungen akzeptiert werden. Der Standard PDF/A definiert diese Einschränkungen und soll in der Praxis genutzt und unterstützt werden.
 - Weitergehende Standardisierungen wie die Speicherung in XML sind wünschenswert, aber im Normalfall nicht zu bewältigen. Wenn sie gewünscht sind, sollten Dokumentvorlagen bereitgestellt werden.
 - Es gibt Objekte, die sich nicht sinnvoll bzw. mit vernünftigem Aufwand in ein Standardformat bringen lassen.

Offene Fragen: Metadaten

- Metadatenstandards
 - Beschreibende Metadaten: Minimum Dublin Core, kein Grund zur Abweichung von Katalogisierungspraxis.
 - Technische/administrative Metadaten: Minimum Dublin Core. Standards im Hinblick auf Archivierung noch nicht stabil (z.B. PREMIS, LMER), aber nutzbar.
 - Sachliche Erschliessung als Problem bei Zusammenführungen
- Erschliessung
 - An welcher Stelle?
 - Abhängig von inhaltlicher Qualität und formalem Typ der Dokumente?
 - Grundsätzlich im Katalog mit späterer Extraktion für andere Zwecke?
 - Durch wen?
 - Autorin mit/ohne Nachbearbeitung durch Bibliothekspersonal
 - Durch Bibliothekspersonal analog zur sonstigen Katalogisierung

Offene Fragen: Weltweite Sichtbarkeit

- Suchmaschinen
 - Benötigen in aller Regel statische HTML-Seiten zum Auffinden und Indexieren
- Metadatenharvesting via OAI-PMH
 - Betreiber des Dokumentenservers muss bestimmte Informationen bereitstellen, damit Harvesting möglich ist.
 - Dokumentenserver (als Data provider) und externer Service provider müssen OAI-PMH einrichten. Der Data provider bestimmt, welche Daten er bereitstellen will.
 - Teilweise ist eine aktive Anmeldung eigener Server bei Service providern möglich, die dann allenfalls Metadaten harvesten.

Fazit und Perspektiven

- Die technischen Voraussetzungen sind gegeben, um den Routinebetrieb im Dienst von Open Access auszubauen.
- An verschiedenen Standorten in der Schweiz ist Know-how vorhanden, das genutzt und vernetzt werden kann. So könnte auch die Vielfalt der verwendeten Lösungen mittelfristig verringert werden.
- Die inhaltliche Zielrichtung der Bibliotheken für ihre Dokumentenserver ist anscheinend nicht immer klar definiert. Unter anderem fehlen Kriterien, welche Dokumente aufgenommen werden sollen und wie ihre Qualität geprüft und gekennzeichnet wird.
- Vom Konsortium her können lediglich zusammen mit den Partnern Empfehlungen erarbeitet werden, die lokal umgesetzt werden oder nicht. In bestimmten Fällen möchte die KUB möglicherweise ein einheitliches Vorgehen vereinbaren.

Vielen Dank!

Dr. Matthias Töwe
Konsortium der Schweizer Hochschulbibliotheken
c/o ETH-Bibliothek
Rämistrasse 101
CH-8092 Zürich
0041-(0)44 632 60 32
matthias.toewe@library.ethz.ch
<http://lib.consortium.ch>