

Repositorien heute und morgen. Ein Überblick

Dr. Helge Steenweg, UB Kassel

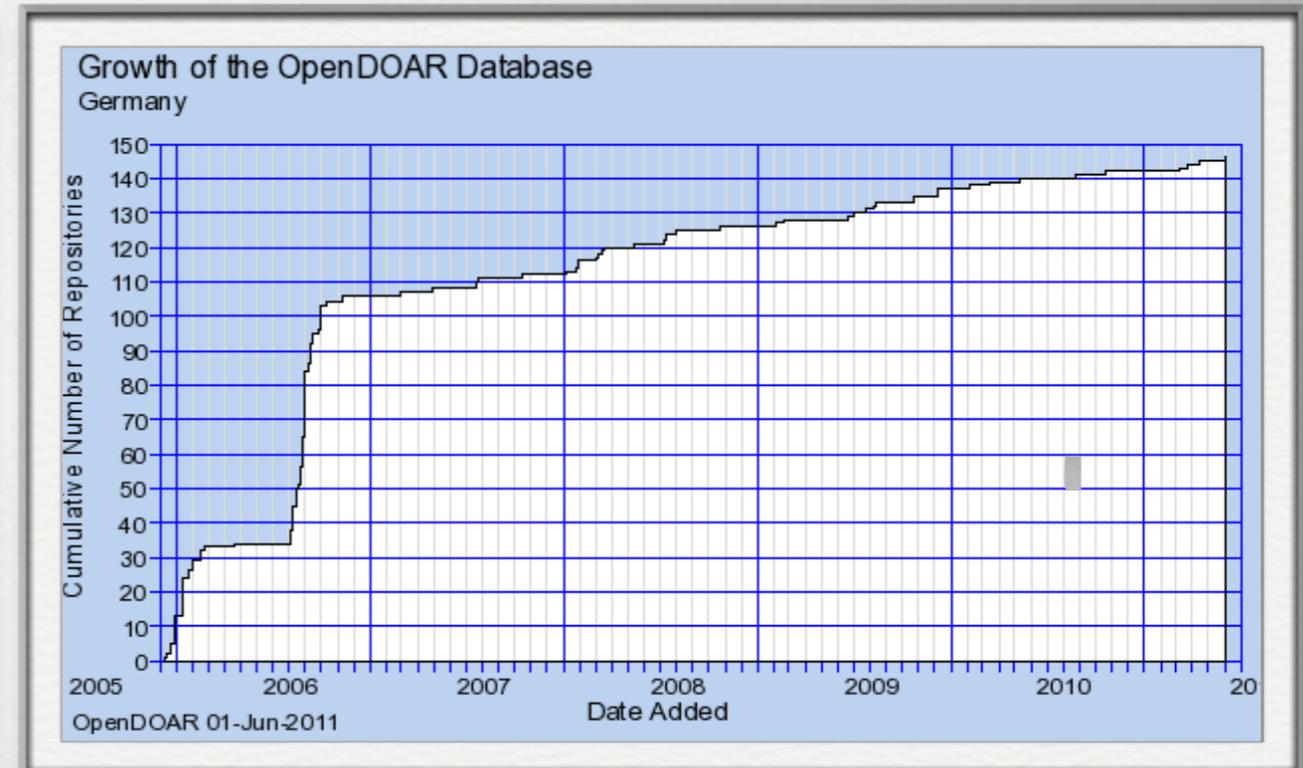
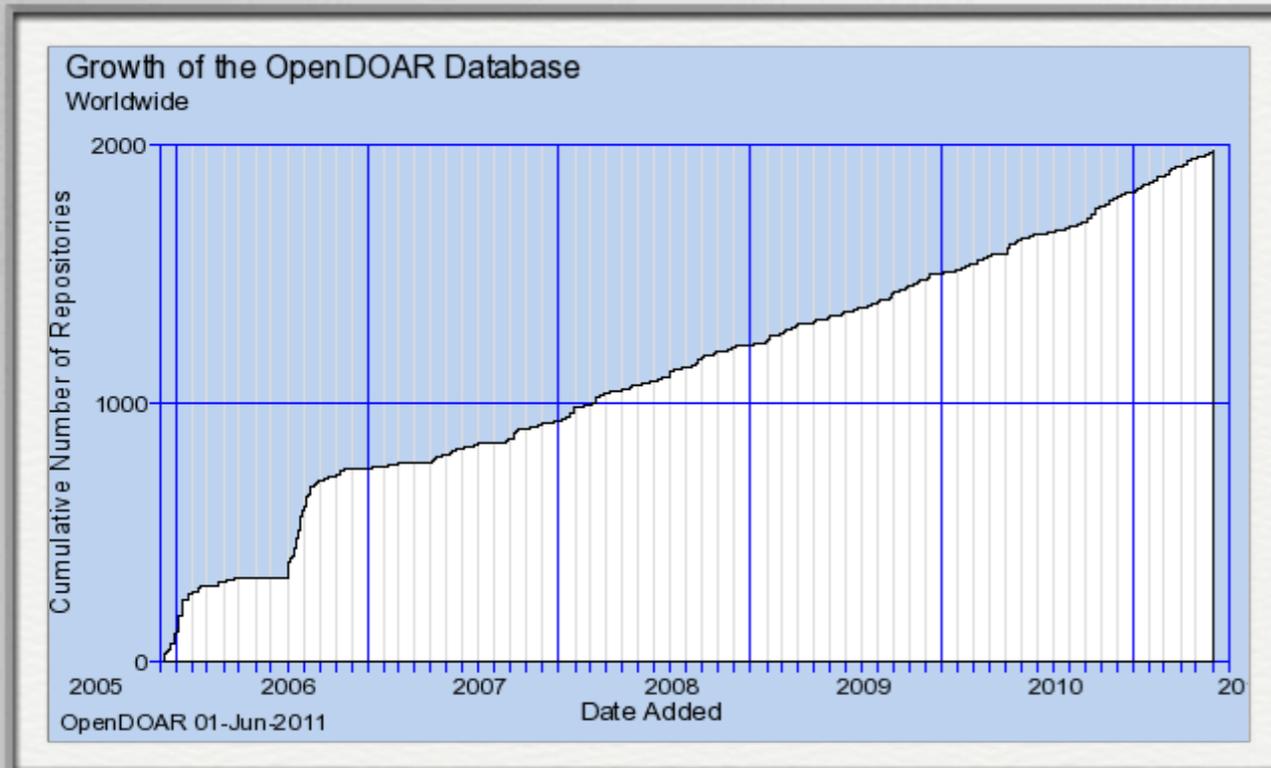


Inhalt

- Anzahl, Inhalt, weltweite Verteilung
- Statistik
- Inhalt
- funktionale Weiterentwicklung
- Vernetzung
- Langzeitarchivierung
- Arbeit, KnowHow, Hosting
- Zertifizierung
- Software



Anzahl der Repositorien 2006 - 2011



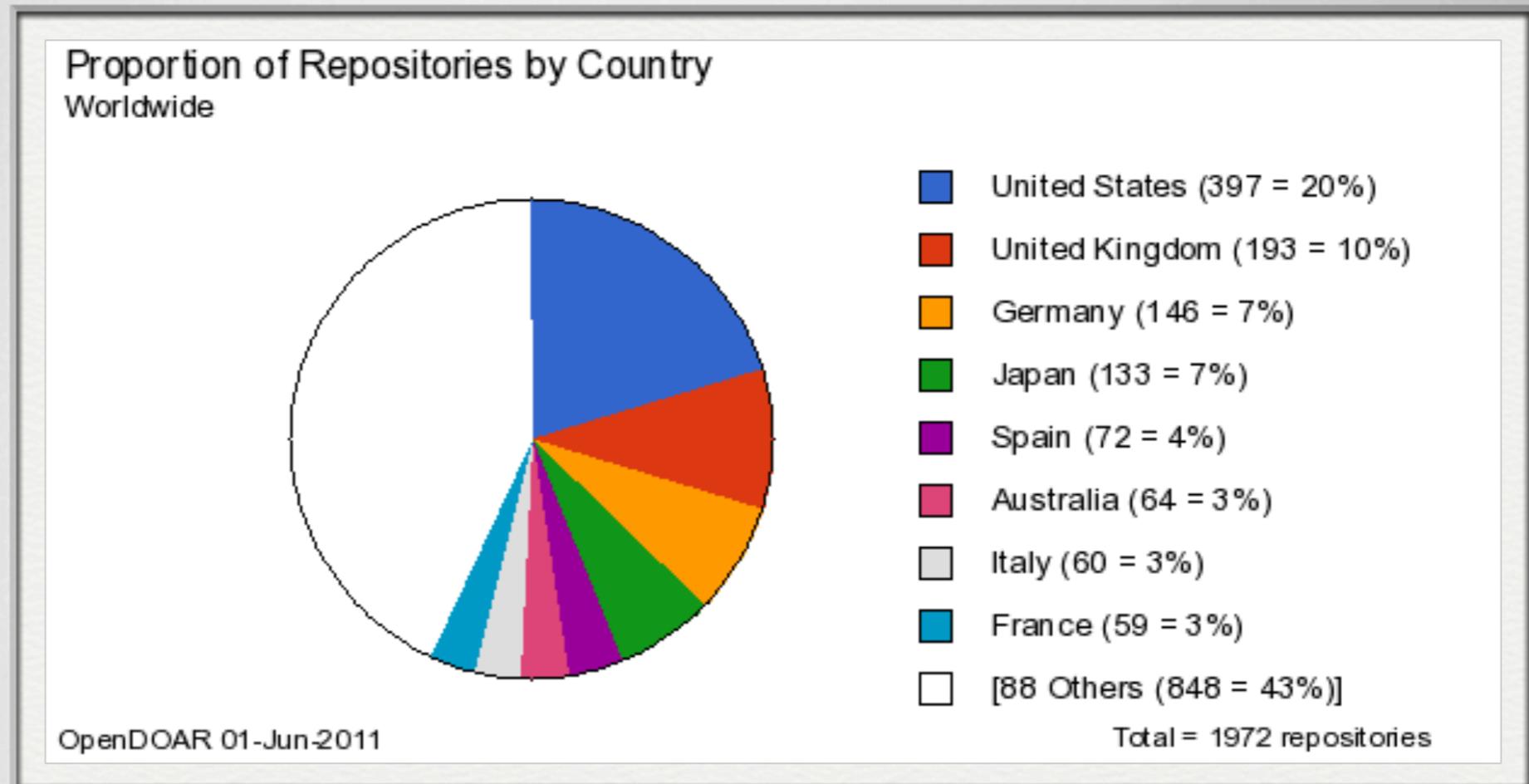
Quelle: <http://www.opendoar.org/>

Trend:

- deutliches Wachstum, weltweit 300%, in Deutschland 40%

Anteile weltweit

4



● Repositorien in Deutschland sind etabliert

Quelle: <http://www.opendoar.org/>



Ranking

- weltweites Ranking von Repositorien 01/2011
- Kriterien:
 - 50% Sichtbarkeit (externe Links-inlinks) (Quelle: Yahoo, Exalead)
 - 20% Größe (Yahoo, Google, Live Search Exalead)
 - 15% Anteil PDF-Dokumente (Quelle: Yahoo, Google)
 - 15% Anteil von aktuellen akademischen Schriften (2001-2008) (Quelle: Google Scholar)

Ranking Web of World Repositories: Top 100 Repositories: Usa & Canada

http://repositories.webometrics.info/top100_continent.asp

Ranking Web of World Repositories
January 2011

home > home > Top USA & Canada

Data

- About Us
- About the Ranking
- Top Repositories
- Top USA & Canada
- Top Europe
- Top Institutionals
- Top Portals
- Directory
- Methodology
- Best Practices

Information

- Glossary
- Blog
- Links
- Contact Us
- Site Map
- Disclaimer

Search Repositories for:

Submit

Category: All

Top Europe

Repositories 1 to 100

CONTINENT RANK	REPOSITORY	COUNTRY	WORLD RANK	POSITION			RICH FILES	SCHOLAR
				SIZE	VISIBILITY	FILES		
1	CERN Document Server	CH	5	4	20	2	8	
2	HAL Sciences de l'Homme et de la Société	FR	10	57	11	58	36	
3	Munich Personal RepEc Archive	DE	11	49	26	17	29	
4	HAL Institut National de Recherche en Informatique et en Automatique Archive Ouverte	FR	16	35	19	131	26	
5	University of Twente Repository	NL	18	116	16	27	72	
6	Igitur Archive Universiteit Utrecht	NL	19	53	59	5	23	
7	École Polytechnique Federale de Lausanne Infoscience	CH	20	8	36	41	220	
8	Universitat Autònoma de Barcelona Dipòsit Digital de Documents	ES	21	17	74	4	78	
9	Universidade do Minho Repositorium	PT	23	32	44	66	49	
10	Dissertations of the University of Groningen	NL	24	55	63	12	47	
11	Universitat Stuttgartelektronische Hochschulschriften	DE	25	191	10	111	299	
12	Ludwig Maximilians Universität München Dissertationen	DE	29	126	33	69	58	
13	University of Southampton: Department of Electronics and Computer Science	GB	34	90	39	67	142	
14	Universität Heidelberg Dokumentenserver	DE	37	72	64	62	62	
15	Digital CSIC	ES	38	52	70	31	140	

Cubermetrics

http://www.webometrics.info

12	Digital CSIC	ES	38	52	70	31	140
14	Universität Heidelberg Dokumentenserver	DE	37	72	64	62	62
15	Digital CSIC	ES	38	52	70	31	140

Statistik zur Nutzung von Repositorien

6

- zur Durchsetzung von OA werden verlässliche Nutzungsdaten benötigt (-> Zitationen)
- durch Einbindung der OA-Repositorien (Indizierung) in kommerzielle Portalsoftware (Google Scholar, Ebsco, Primo, Summon, WorldCat local) werden verstärkt Zugriffe erfolgen
- Statistik zur Nutzung von Repositorien (Open Access Statistik - OA-S, Berlin, Göttingen, Saarbrücken, Stuttgart)

<http://www.dini.de/projekte/oa-statistik/>



Trend

- Nutzung von Repositorien wird sich erhöhen
- Nutzungsstatistik wird daher immer wichtiger als Nachweis



Inhalte

7

Country matches any of "Germany" - Registry of Open Access Repositories

http://roar.eprints.org/cgi/roar_search/advanced?location_country=de&software=&type=&order=-recordcount

Registry of Open Access Repositories (ROAR)

Home | Search | Search Content | Browse by Country | Browse by Type | Browse by Software | Register in ROARMAP | About

Login | Create Account/Add Repository

Country matches any of "Germany"

Germany (129) | Any Software | Any Repository Type | Sort by number of records | Search

Displaying results 1 to 20 of 129.

Export 129 results as Activity table | Export | RSS 1.0 | Atom | RSS 2.0

Search Content | Graphical analysis

- 1.** **Deposit Activity** **Publishing Network for Geoscientific & Environmental Data (564231 records)**
[Alfred Wegener Institute for Polar and Marine Research - 11 July 2005](#) [Record Details]
- 2.** **Deposit Activity** **Zentrales Verzeichnis Digitalisierter Drucke: Hauptseite (491032 records)**
<http://www.hbz-nrw.de/> - 16 January 2008 [Record Details]
- 3.** **Deposit Activity** **Max Planck Society Edoc Server (165921 records) - 06 June 2002**
[Record Details]

Quelle: <http://roar.eprints.org/>

Trend

- Inhalte verändern sich
- Statistik ist stets inhaltlich zu hinterfragen
- bei Open Access (OA) Repositorien ist der Anteil der Pre- oder Postprints wichtig

fachspezifische - institutionelle Repositorien

8

- Durchlässigkeit und Verbindung
- Fachwissenschaftler nutzen eher fachspezifische Repositorien als Ersteinstieg
- DFG-Projekt UB Konstanz: Open-Access-Fachrepositorien

<http://www.ub.uni-konstanz.de/bibliothek/projekte/open-access-fachrepositorien.html>

- Verknüpfung von IR-Dokumenten in fachliche Repositorien

● Durchlässigkeit und Verknüpfung wird sich verbessern



Trend



IR: institutionelle - offene Repositorien (IR-OR)

- Frage der Definition
- idealistisch: IR nur für die digitale Werken einer Institution
- realistisch: Produkte aus anderen Arbeitsgebieten (Digitalisierung), Arbeiten aus anderen Beschäftigungszeiten und -orten, wer urteilt über Qualität (z.B. bei Qualifizierungsarbeiten)?

Differenzierungen werden sich annähern



Funktionale Weiterentwicklung

10

- Bibliometrie-Problematik und Publikationsmanagement an Hochschulen
 - > Hochschulbibliographie
 - > Forschungsinformationssystemen (FIS)
- mögliche Varianten: Erweiterung (EPrints) oder Frontend (PUMA, Publister) (Sword2.0)
- Forschungsumgebung (VRE, Forschungsdaten)
- Vereinfachungen in der Handhabung (Apps von EPrints - Bazaar)



Trend

- Neue Aufgaben und Weiterentwicklungen
- Mehrwertdienste schaffen



Medienformate

- Formate für mobilen Zugriff (epub)
- E-Books (Problem: Originalformatierung fällt ev. weg, da sich „epub“ an Reader anpasst)
- IR- Authentizität mit Original?
- Multimediale Medien, z.B. 3-D-Daten, Verknüpfung Audio-Daten; Probado-Projekt: <http://www.probado.de/>
- Weiternutzbarkeit von Objekten OAI-ORE (Object ReUse) - bedingt auch Zugriffsfragen, Rechtsfragen, Architektur der Datenbanken



Trend

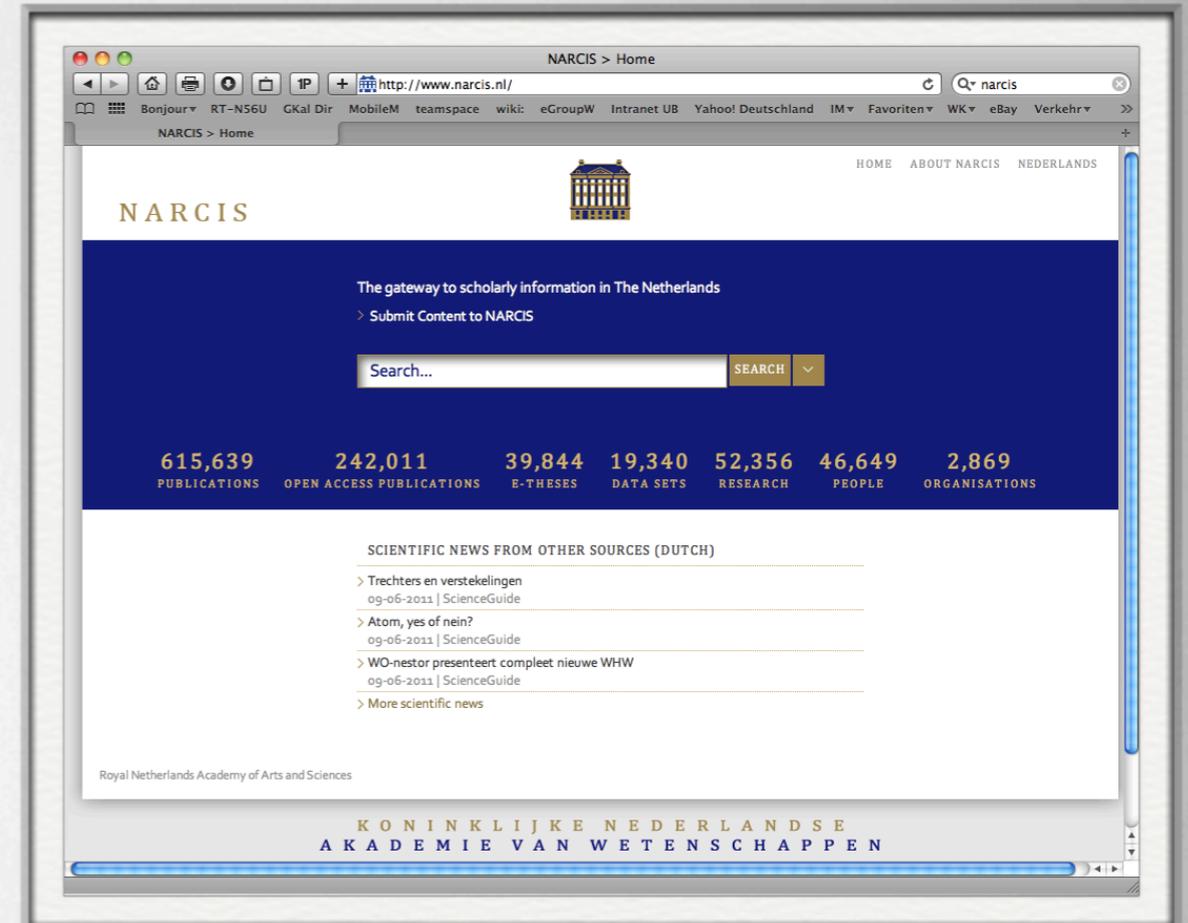
- es werden neue Datenformate wichtig
- Integration dieser Formate



Vernetzung von Repositorien

12

- Beim Aufbau vom Index:
Problem Duplikate
- maschinell nicht immer auszuschliessen
- Ansatz: Autorenidentifikation als ein-eindeutiges Zuordnungsmerkmal (Probleme: Formate, Namensänderungen, etc.)
ORCID-Initiative <http://www.orcid.org/>
- Beispiel: NARCIS <http://www.narcis.nl>



● Autorenidentifikation als Zuordnungsmerkmal



Vernetzung von Repositorien

13

- derzeit unverzichtbar OAI-PMH
- Projekt DRIVER - Ziel: Verbesserung des Zugangs zu Forschungsergebnissen - BASE-Index Bielefeld
- wichtig zukünftig: noch stärkere Vernetzung von Inhalten (Metadaten und Medien)
- durch die Trennung von Metadaten und Medien in manchen Datenbanken ist ein Backup nach einem Crash schwierig, wenn die Datenbank mit den Verknüpfungen nicht zu restaurieren ist.
- OA-Netzwerk
- OpenAIRE (Open Access Infrastructure for Research in Europe)
- COAR (Confederation of Open Access Repositories)



Trend

- stärkere Vernetzung, mehr Informationen in den Indizes
- kooperative Sicherung von Paketen (Metadaten und Medien)
- internationale Zusammenarbeit



Langzeitarchivierung

- Voraussetzung: Persistente Identifier (URN/NBN, DOI, Handle) und deren einfache Auflösung in jedem Browser
- die genannten Formatänderungen und Backup-Probleme verschärfen die ohnehin nicht leicht zu praktizierende Langzeitarchivierung
- kooperatives Backup zwischen Repositorien z.B. mittels Sword-Paketen wäre sinnvoll, wenn eine zentrale Lösung nicht möglich ist.



Trend

- praktische Umsetzung von Langzeitarchivierungsstrategien
- kooperatives Backup



Arbeit

15

- nicht zu vergessen: Arbeit und Personal
- nicht nur bezüglich technischer Realisierung und KnowHow
- Umfeld wichtig:
 - Aquise und Marketing
 - Metadatenbearbeitung
 - Formatbearbeitung und-konvertierung

- Repositorien kosten Arbeit und Personal
- Hosting als Alternative



Hosting von Repositorien

- DSpace - DuraCloud

<http://www.duraspace.org/duracloud.php>

- EPrints - EPrintsServices

<http://www.eprints.org/services>

- OPUS 3/4 - 31 Instanzen beim KOBV, ?Instanzen beim BSZ

http://www.kobv.de/fileadmin/opus/workshop/2011_05_05/Conrad-Rempel_Migrationsplanung.pdf



- Hosting wird weltweit stark zunehmen
- Cloud ist derzeit „Hype“
- Frage von KnowHow und Personal
- Langzeitarchivierung



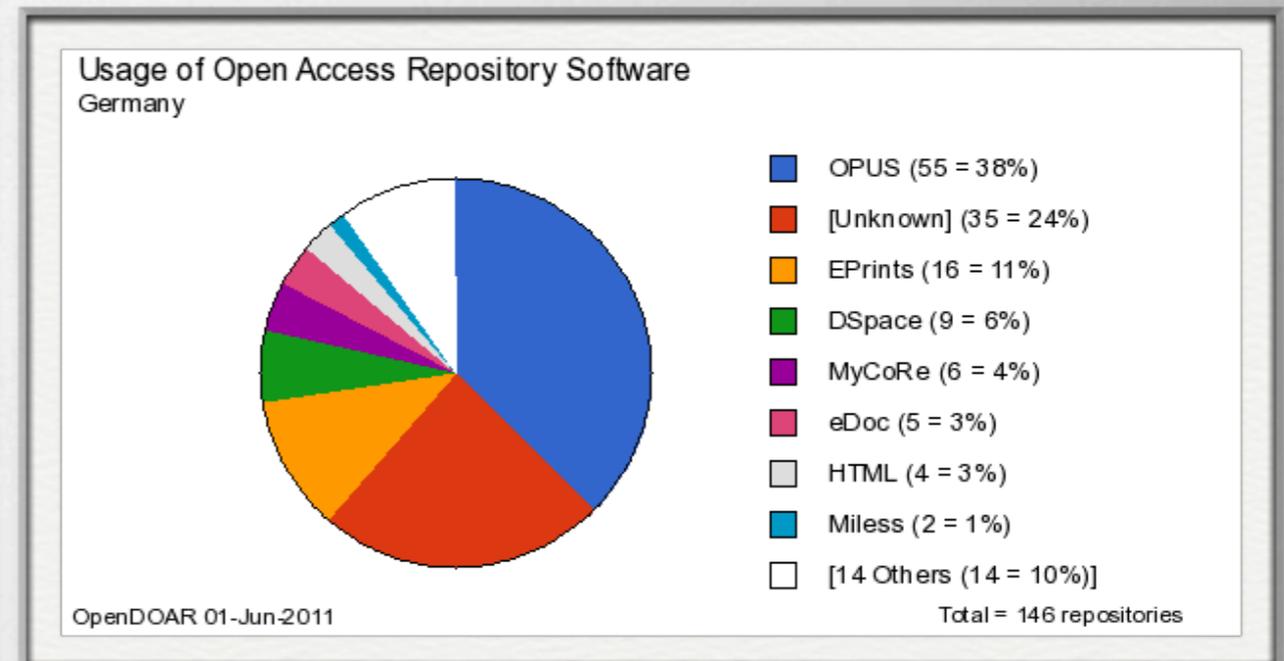
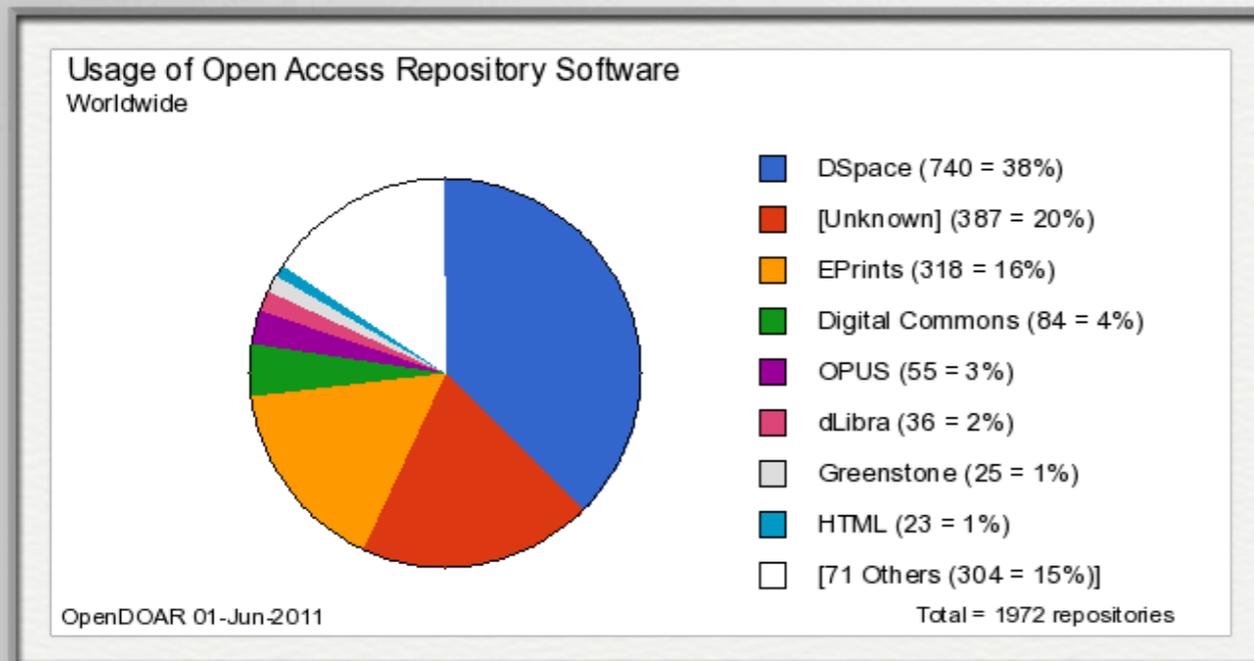
Dini-Zertifikat

- „deutscher TÜV“ für Repositorien (Harnap)
- Anleitung und Hilfe (Leitlinie) für (neue) Betreiber von Repositorien
- beim Abarbeiten am eigenen Repository erkennt man Schwächen und Fehlstellen
- verdeutlicht nochmals die Wichtigkeit einzelner Komponenten
- wird ständig aktualisiert



Verteilung der Repositorien - 18 Software

Quelle: http://www.open_doar.org/



- deutliche Dominanz weniger Systeme
- in Deutschland andere Verteilung
- zu erwarten: Internationalisierung

Repositorien-Software

19

Fragen allgemein:

- Frage nach der Software? Welche? Warum?
 - Welche Art von Repository möchte ich auf- oder ausbauen? (IR-OR-fachspez.R)
 - Aufwand/Personaleinsatz
 - welche Funktionalität zusätzlich
(Publikationsmanagement samt FIS, Versionierung, Forschungsdaten, Weiterverwendung von Elementen (ORE), etc.)
 - welche Medienformate sollen abgedeckt werden
 - gibt es Ansätze für Langzeitarchivierung



Repositorien-Software

Fragen:

- Zukünftige Entwicklung von OPUS
- Weiterentwicklung DSpace/DuraSpace
- Erweiterung des Funktionsumfangs bei EPrints
- Multifunktionalität von MyCoRe
- Variabilität von Fedora/eSciDoc PubMan

