Oh my, how much they have grown!

Looking at Open Access Repositories from 3 Different Perspectives

2012 Census of Open Access Repositories in Germany
What the ...?

- Let’s look at the development and status quo of Open Access Repositories in Germany
- Open Access Netzwerk
- Bielefeld Academic Search Engine (BASE)
- 2012 Census of Open Access Repositories in Germany
- Let’s find new solutions and approaches
- Let’s discuss what we’ve found
Open Access Network (OAN)

Initiated by: DINI

Funded: DFG

Project Partners:

Computer- und Medienservice
Institut für Bibliotheks- und Informationswissenschaft (from 2009)

Fachbereich Mathematik/Informatik

NIEDERSÄCHSISCHE STAATS- UND UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN
Aims and Services

Aims

• Strengthen the visibility of OA repository content in Germany (Vision)
• Build a national aggregation node (Technical Approach)
• Support „DINI Certificate of Document and Publication Services“ (Organisational Approach)
• Cooperation network of OA Projects (Organisational, Technical Approach)

Services

• Value-added Services (Users)
• Data (Repository Operators)
• Development (Developers)
Results

- Network of 40 DINI certified repositories (09/12)
- Dataworkflow
  - Aggregation
  - Harmonisation
  - Fulltext Link Finder & Indexing
  - Classification & Deduplication
- Webbased Administration Tool (Scheduling, Service-Control etc.)
- OAN Validator

Data Base

- 268.000 Metadata sets
- ~ 150.000 Open Access Full Texts, via OAN web interface
- Index of 80.000 Documents, update shortly (09/12)
Daten-Workflow
Die modulare OA-Netzwerk-Architektur
• Sustainability
• Aggregation node DDB (DNB), Cooperative Library Networks
• Computer and Media Service (HU Berlin)
• Extension of OAN Data Base
• Linked Open Data
• Metadata in RDF format (Triple Store, Sparql-Endpoint)
• Linked with LOD cloud (author names, PND DNB)
Total Growth of Open Access Items in DINI certified Repositories (2004-2012)

Sep 2004: 9,7k
Average growth: 2.7k/month
Sep 2012: 268k

Source: Netzwerk von Open-Access-Repositorien

Created by Paul Vierkant, Sammy David and Maxi Kindling
National Distribution of Open Access Items
(DINI certified Repositories, September 2012)

111k Nordrhein-Westfalen
45k Schleswig-Holstein
20k Sachsen
18k Berlin
32k Baden-Württemberg
8k Bayern
7k Hessen
6k Brandenburg
5.5k Saarland
2.8k Niedersachsen
2k Bremen
2.5k Rheinland-Pfalz
1.2k HH

Source: Netzwerk von Open-Access Repositorien

Created by Paul Vierkant, Sammy David and Maxi Kindling
Languages of Open Access Items
(DINI certified Repositories, September 2012)

138k German

112k English

13.5k Unknown

4.3k Other

Source: Netzwerk von Open-Access-Repositorien

Created by Paul Vierkant, Sammy David and Maxi Kindling
Type of Open Access Items
(DINI certified Repositories, September 2012)

- 59k Articles, Journals
- 39k Theses
- 30k Other
- 30k unknown or incorrect
- 16k Books
- 1.2k Reviews
- 0.43k Images
- 0.35k Videos
- 0.058k Sheet Music

Source: OAPEN

Created by Paul Vierkant, Sammy David and Maxi Kindling
Subjects of Open Access Items
(DINI certified Repositories, September 2012)

Technology
- 64.3k

Science
- 63.3k

Social Sciences
- 26.3k

Computer Science
- 18k
- Arts and Recreation 3.8k
- History and Geography 3.6k
- Philosophy and Psychology 3.5k
- Religion 0.8k
- Literature 1k
- Language 1.6k

82.1k unknown or incorrect

Source: Netzwerk von Open Access Repositorien

Created by Paul Vierkant, Sammy David and Maxi Kindling
DINI Issues on a Global Scale?

- BASE as the global source for Open Access items
- An Open Access repository is not a matter of size but a matter of content!

Source: Open Access Repositories

Sep 2004
0.68m
average growth
0.43m/month

May 2012
35.79m

Created by Paul Vierkant, Sammy David and Maxi Kindling
Global Distribution of Open Access Items (June 2012)

16.46m Americas
14.97m US
12.84m Europe
1.98m GB
2.39m France
1.88m Spain
1.76m Germany
3m International Organisations
2m Asia
1m Japan
1.46m Oceania
0.16m Africa

Source: BASE
Created by Paul Vierkant and Najko Jahn
Type of Open Access Items (July 2012)

- 1.22m Reports, Papers, Lectures
- 1.2m Books
- 1.43m Theses
- 0.48m Primary Data
  - 0.05m Reviews
  - 0.049m Maps
  - 0.048m Sheet Music
  - 0.029m Audio
  - 0.02m Videos
  - 0.003m Software
- 2.9m Images
- 6.63m Article, Journals
- 22m unknown or incorrect

Source: BASE
Created by Paul Vierkant
Subjects of Open Access Items (July 2012)

33.1m

Unknown or Incorrect

0.91m Technology
0.8m Science
0.73m Social Sciences
0.27m Arts and Recreation
0.2m History and Geography
0.22m Computer Science
0.14m Philosophy and Psychology
0.14m Language
0.09m Religion
0.04m Literature

Source: BASE
Created by Paul Vierkant
Publication Years of Open Access Items (June 2012)

- Unknown or incorrect date: 11.43m
- 2000s: 10.9m
- 1990s: 4.29m
- 2010s: 3.72m
- 1980s: 2.77m
- 1970s: 1.53m
- 1960s: 1.01m

Source: BASE
Created by Paul Vierkant and Maria Vierkant
• small Data Base (268k)
  • Metadata
    • 5% unknown Languages
    • 11% unknown Doctypes
    • 30% unknown DDC

• large Data Base (36m)
  • Metadata
    • 34% unknown Languages
    • 61% unknown Doctypes
    • 92% unknown DDC

DINI Certificate makes Sense!
One more thing...

2012 Census of Open Access Repositories in Germany

Experience & Knowledge

GUESS

vs.

Numbers & Facts

DISCOVERY
Some Facts...

Repositories in Germany

Amount of Items in the Bundesländer

2012 Census of Open Access Repositories in Germany
Software & OAN Validator

Score

<table>
<thead>
<tr>
<th>System</th>
<th>0-70</th>
<th>70-80</th>
<th>80-100</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPUS</td>
<td>17</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>DSpace</td>
<td>5</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>EPrints</td>
<td>14</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2012 Census of Open Access Repositories in Germany
2012 Census of Open Access Repositories in Germany

METADATA FORMATS

- **99%** Simple Dublin Core
- **4%** Qualified Dublin Core
- **30%** XM eta Diss
- **43%** XM eta Diss-Plus
- **13%** METS
- **66%** Epicur

**DFG**

Institutions offering an Open Access publication fund supported by the DFG: 17

**COAR**

Institutions being a Member of the Confederation of Open Access Repositories (COAR): 17
Solutions

You should consider supporting initiatives:

You should consider support (metadata) standards!
Solutions

New tools to support repository management:

metadata analyzer for repositories

create report
paste OAI base URL:

or choose OAI base URL:

get report

review report

http://edoc.hu-berlin.de/OAI-2.0

review item

beginning: <url>http://edoc.hu-berlin.de/OAI-2.0</url>

end: </url>
Have they really grown?

• OAN shows the importance of DINI and the DINI Certificate in Germany
• BASE shows OAR issues on a global scale
• 2012 Census reveals that even more issues in Germany have to be tackled
Some repositories have grown, others need more support than ever to catch up.

Bottom line:
Thank You for Your Attention!

Let's discuss!