Open Archives Forum
- Technical Validation -

Birgit Matthaei

Humboldt University Berlin, Germany

Computer and Media Service, Electronic Publishing Group
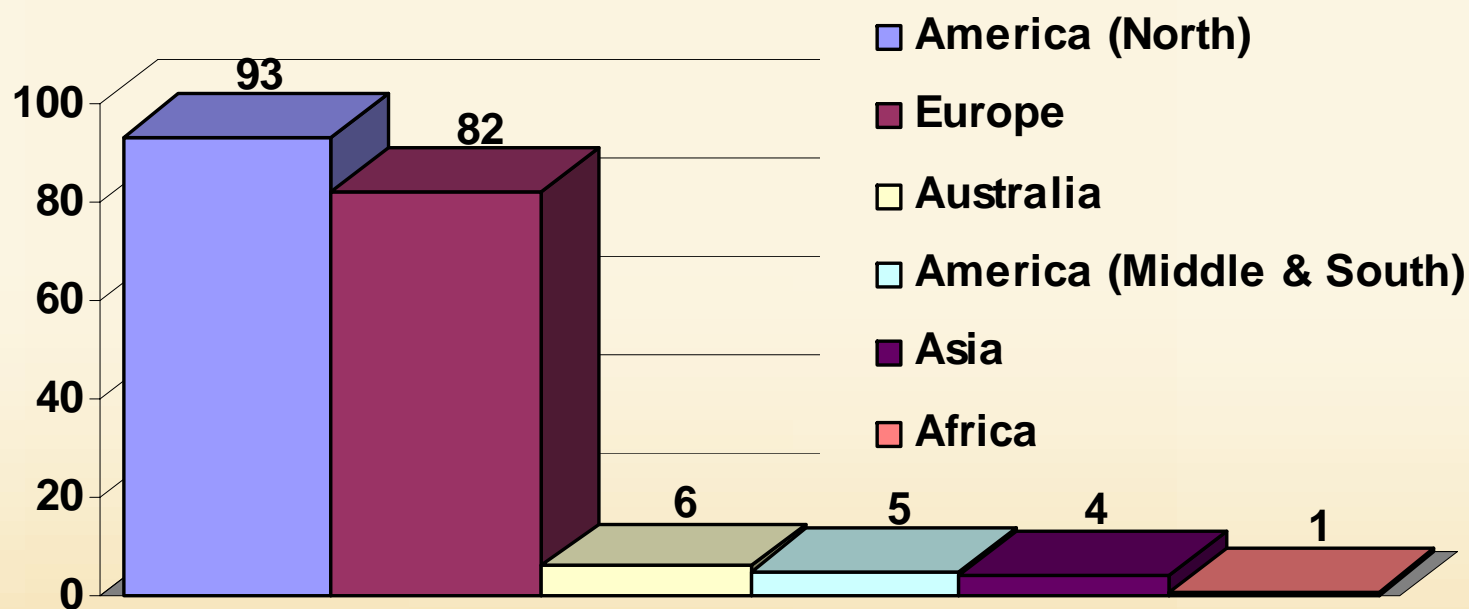
birgit.matthaei@cms.hu-berlin.de

Creating Information Sources

➢ European portal for open archives information

- Information Resource Database (registry of repositories, services, software, projects and associated organisations

➢ Evaluation of status, experiences and future plans regarding European OAI implementations

- Online "Technical Validation Questionnaire"

➢ Systematic inventories

- Repositories, services and tools

➢ Reports on own experiences

- OAI-PMH 2.0 alpha and beta tester,
- Implementation of OAI Services • experiences software tools

➔ Highlight some aspects of european activities on OAI in relation to worldwide activities
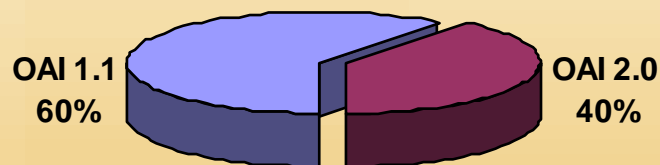
Overview on OAI activity (continents)

Legend:
- America (North)
- Europe
- Australia
- America (Middle & South)
- Asia
- Africa

Bar chart values: 93, 82, 6, 5, 4, 1

Upgrading to OAI 2.0 (Nov. 2002)

OAI 1.1 60% — OAI 2.0 40%

Upgrading to OAI 2.0 (Aug. 2003)

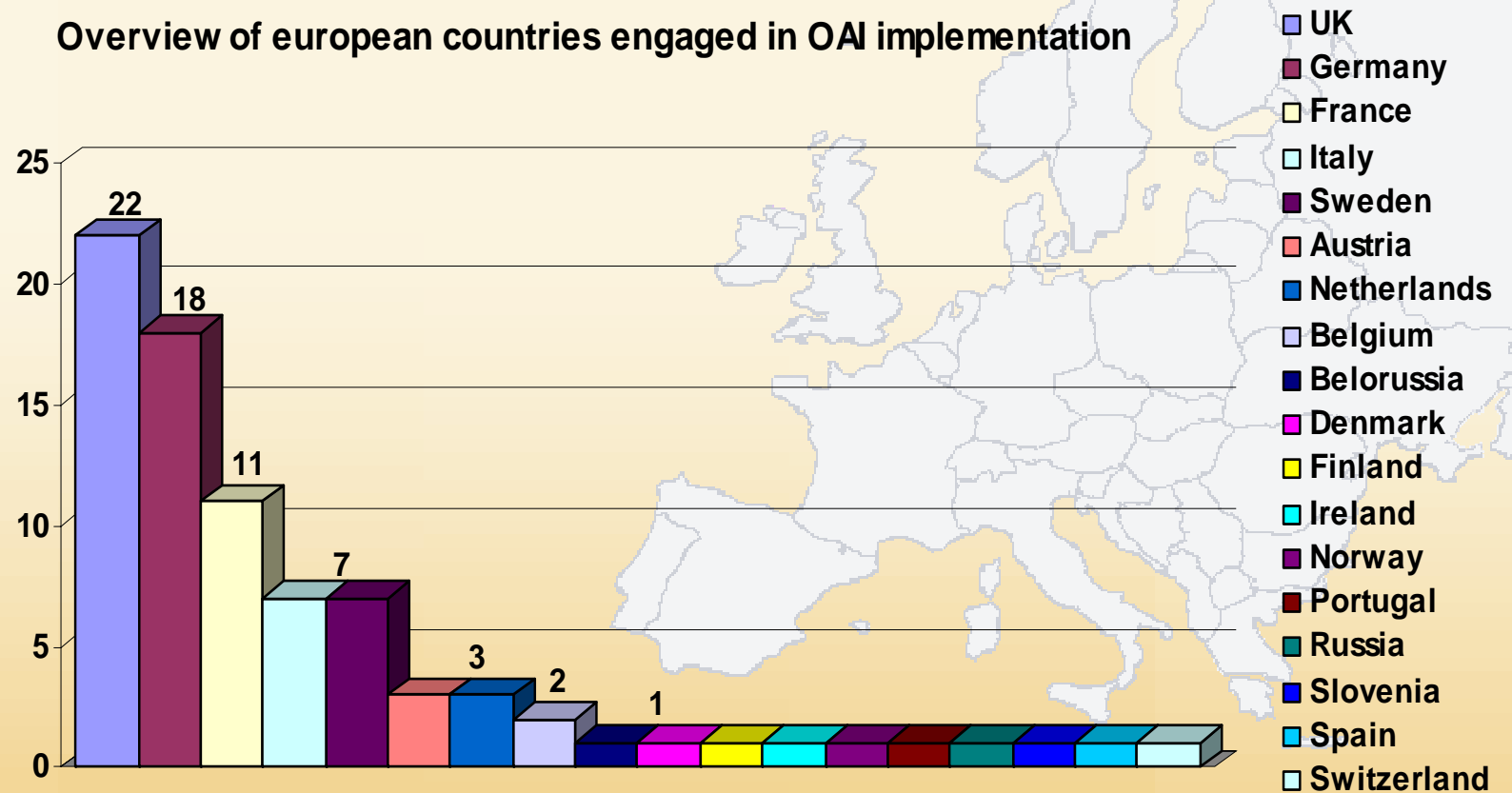OAI 1.1 17% — OAI 2.0 83%
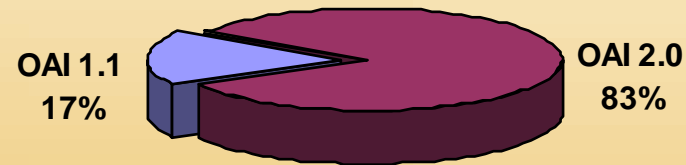
IST- 2001-320015

Overview Repositories - Europe

Overview of european countries engaged in OAI implementation

Legend: UK, Germany, France, Italy, Sweden, Austria, Netherlands, Belgium, Belorussia, Denmark, Finland, Ireland, Norway, Portugal, Russia, Slovenia, Spain, Switzerland

Values: 22, 18, 11, 7, 3, 2, 1

Birgit Matthaei, 4th Sept. 2003, Bath, 4th OAForum Workshop: 'In Practice, Good Practice'
Humboldt University Berlin, Germany – Computer and Media Service – Electronic Publishing Group

Details of Questionnaire

Data Provider

active	19
in development	10
planned	22

0   5   10   15   20   25

Service Provider

active	5
in development	11
planned	16

0   5   10   15   20   25
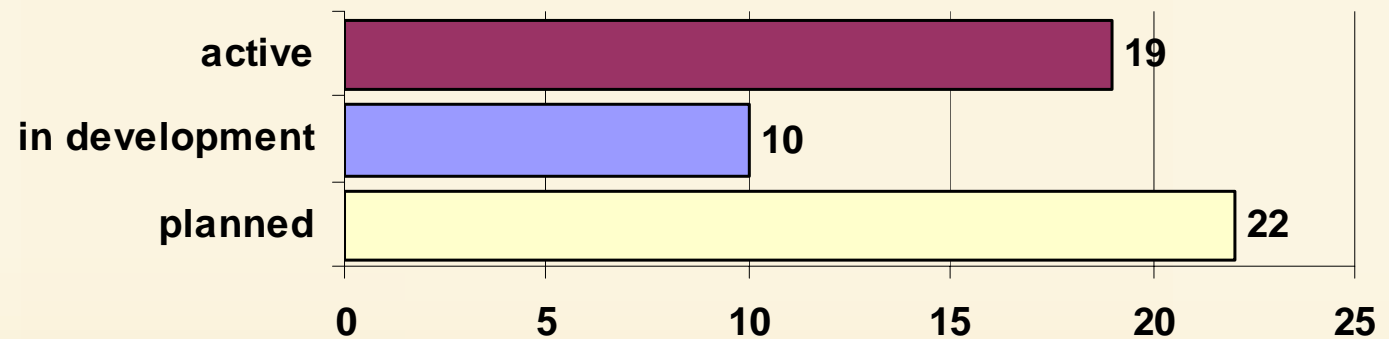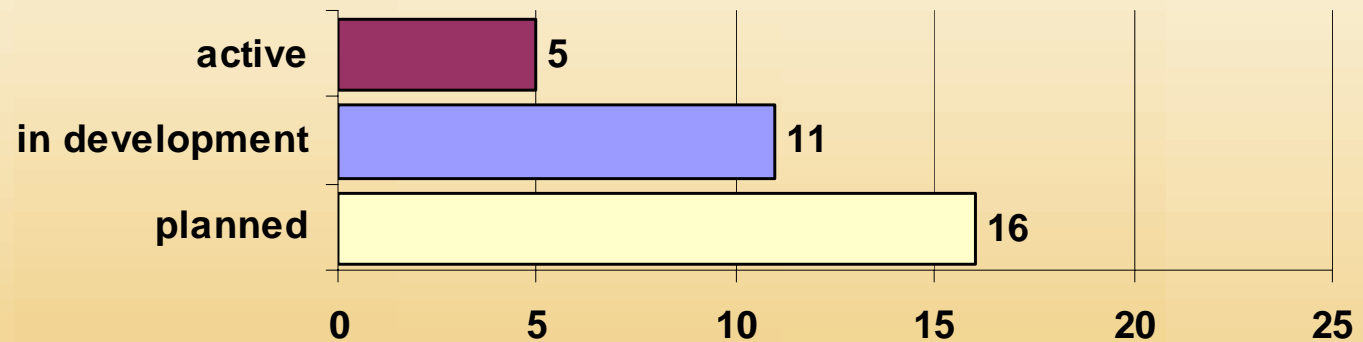
Questionnaire: Community Types

Preprint/ Science 13%

Museum 8%

Publisher 2%

Archive 19%

Others 14%

no specification 13%

Library 31%
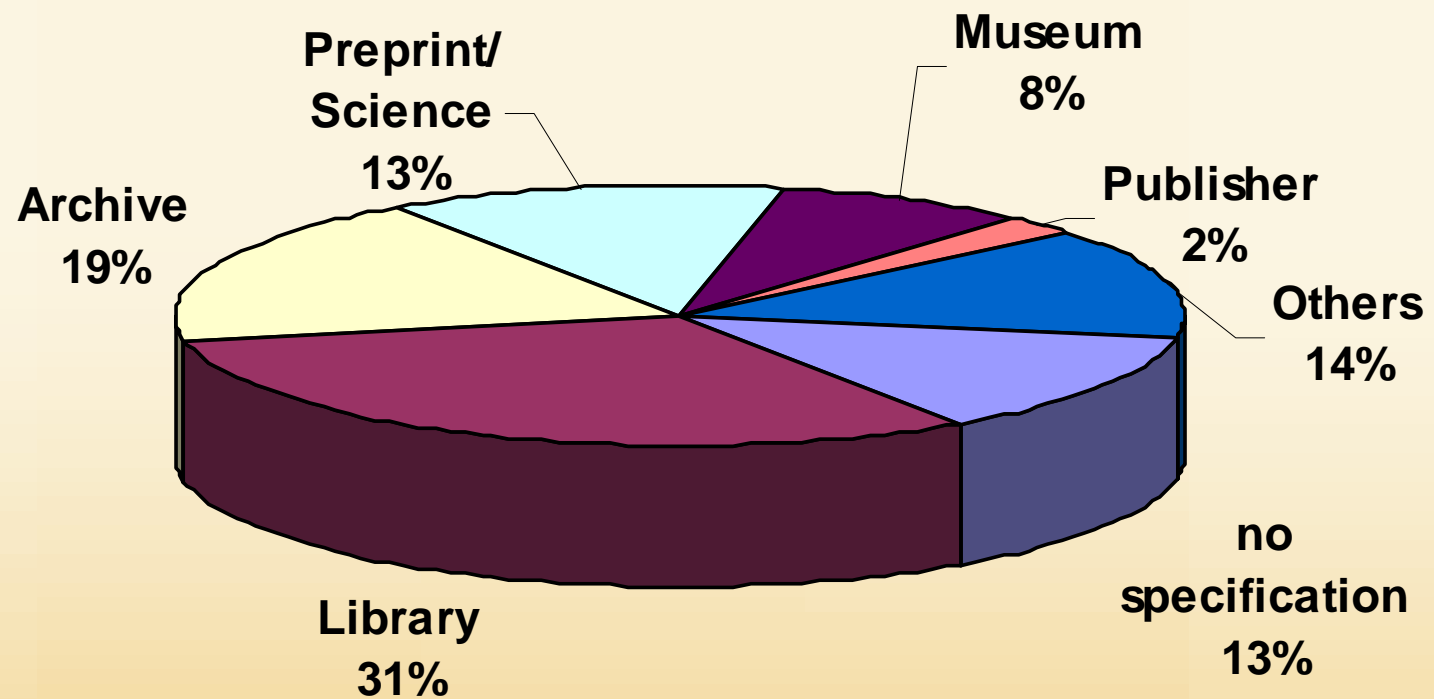
Multiple answers possible

Questionnaire: Object Types

Multiple answers possible

Object Type	already OAI compatible	not yet OAI compatible
Others	2	
Raw/Statistic Data	1	1
Audio	1	1
Software	1	3
Video/Streams	5	
Images - Vector graphics		1
Images - digitised mat.	6	12
Abstracts	8	9
Fulltext documents	17	14
Metadata	23	15

0 5 10 15 20 25 30 35 40

□ already OAI compatible ■ not yet OAI compatible
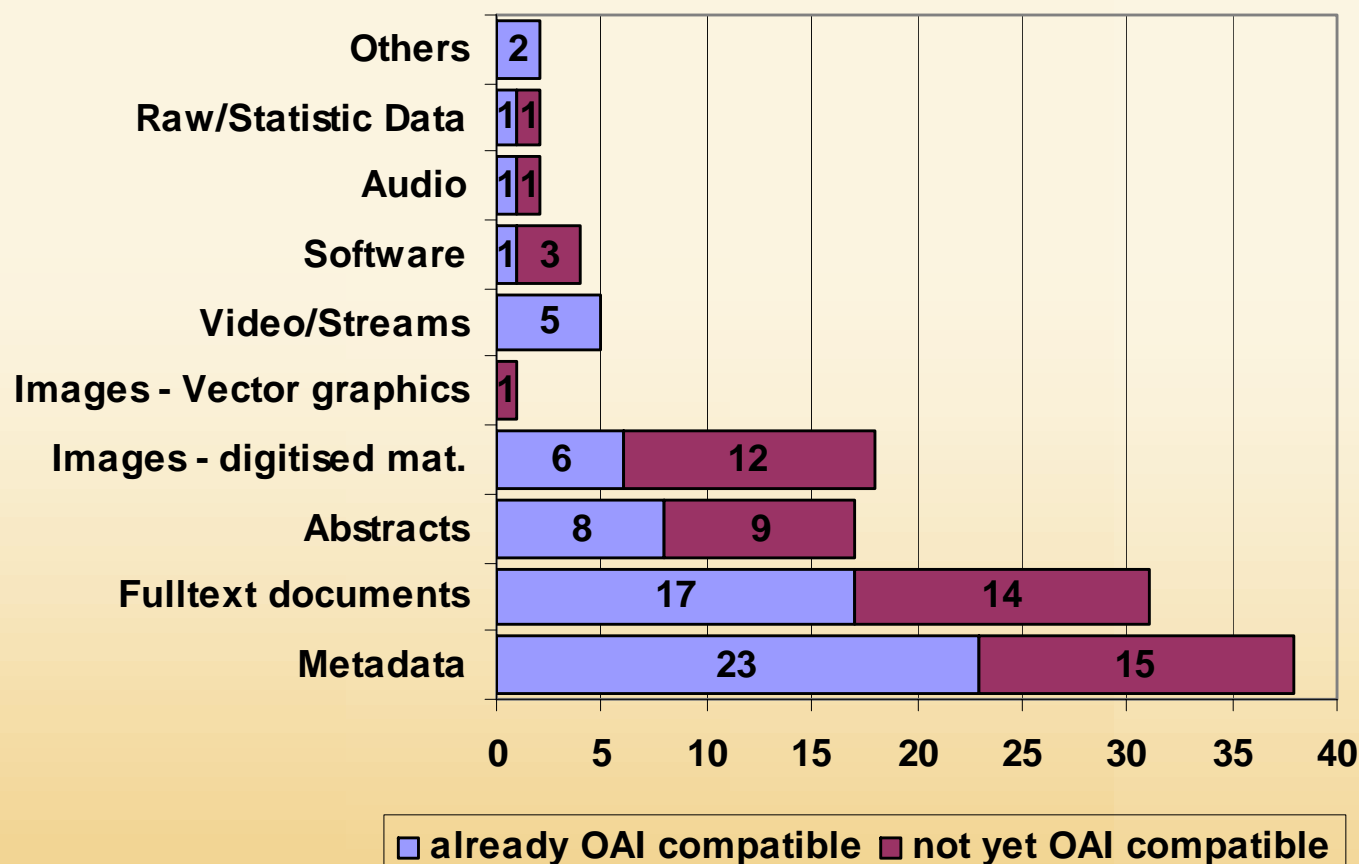
Questionnaire: Content Types

Multiple answers possible

Content Type	already OAI compatible	not yet OAI compatible
Others	15	11
Recordings	2	2
Lectures	6	4
Conference Proceedings	6	6
Preprints	9	8
Journal Articles	13	8
Dissertations	14	8

■ already OAI compatible ■ not yet OAI compatible
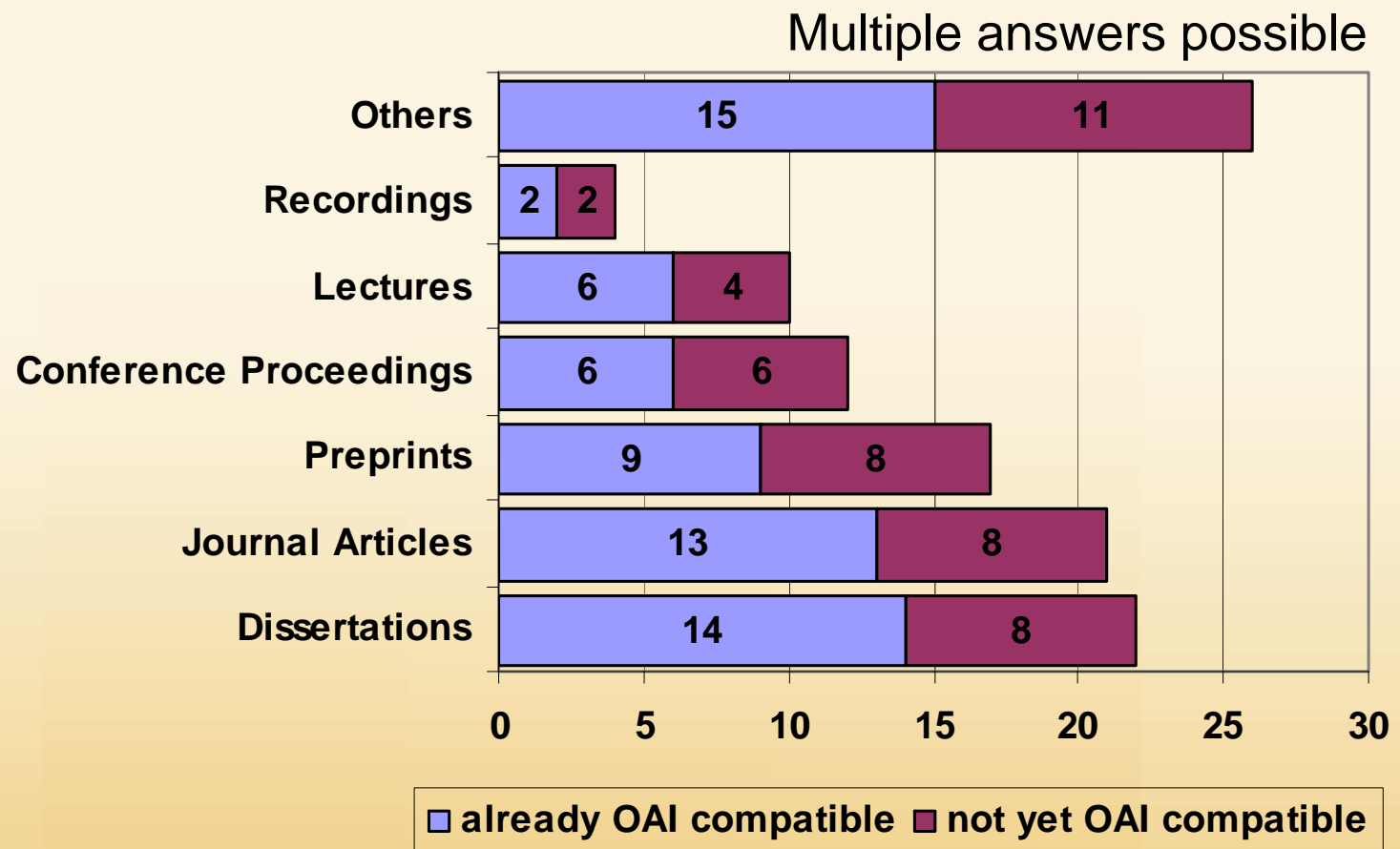
Questionnaire: Used Software

➤ Many different tools were mentioned:
  ✖ ADLIB, ARNO, CDSware, DIENST, DSpace, Elektra, EPrints, PERL implementations, OAI Cat, OAI Harvester, VT-ETB-db

➤ Today: 50 % use self developeded tools
  One year ago: 80 % used self-developed tools

➤ Trend: Need of tools which are user-friendly complete solutions
  ✖ cover typical functionalities
  ✖ to be installed by relatively small expenditure
  ✖ adaptable to special requirements if necessary
  ✖ little expenditure with the further care of the data

➢ **GNU Eprints -** developed at the Electronics and Computer Science Department of the University of Southampton, UK

➢ **DSpace -** newly developed as a joint project of the MIT Libraries and the HP Company, USA
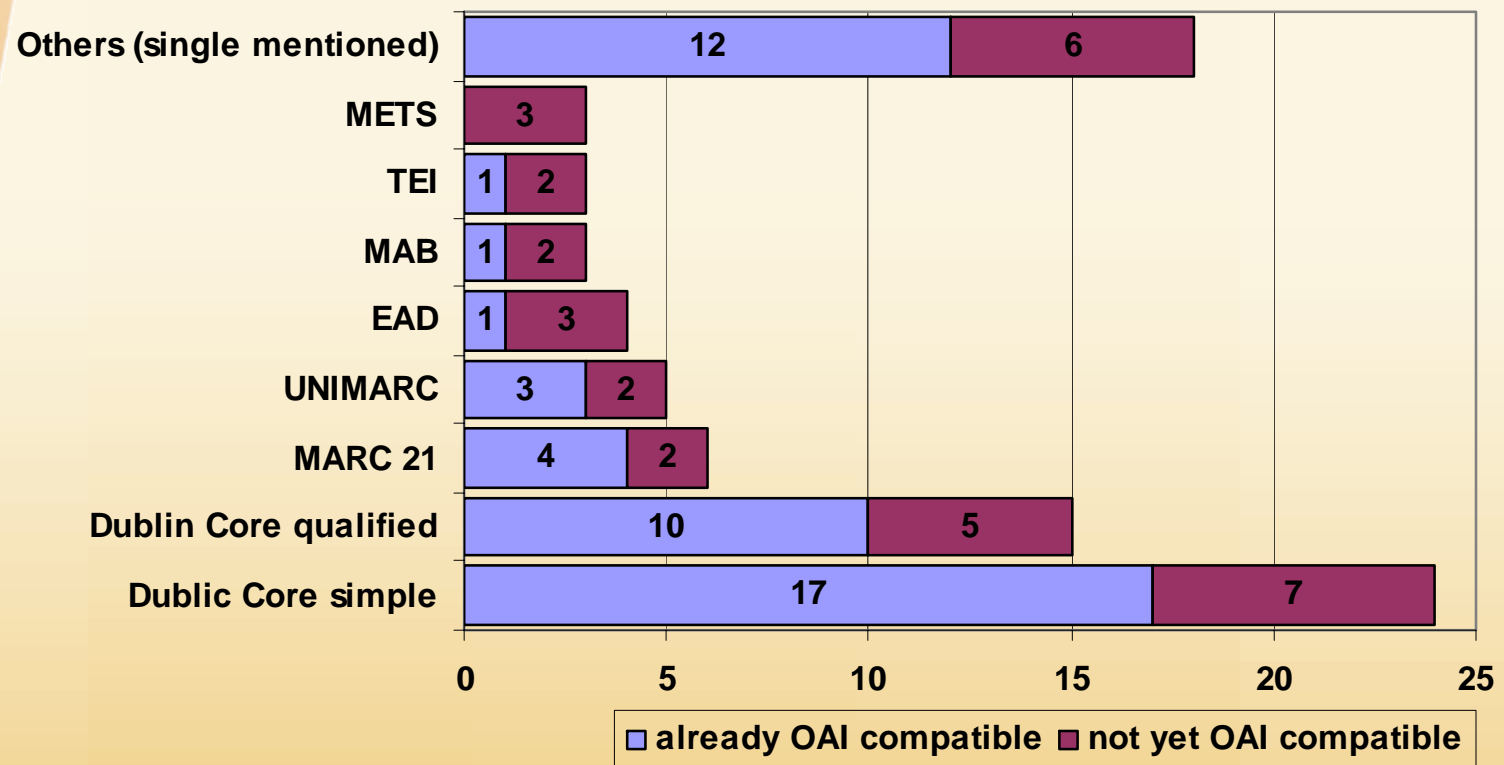
➢ Some numbers of the inventory of repositories:

Nov 2002   other tools 74%   eprints 26%

Aug 2003   other tools 61%   eprints 39%
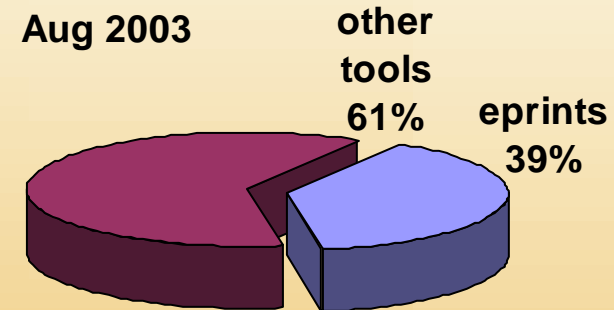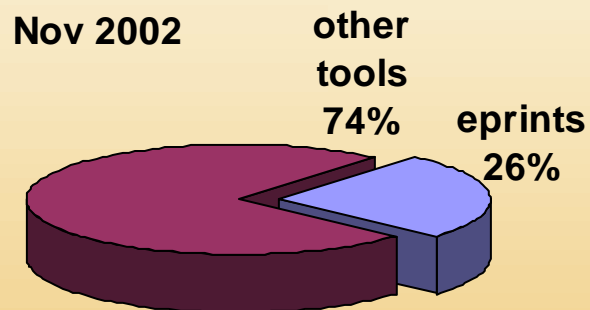
Tools: Eprints - DSpace

➢ Open source developments for archiving

➢ Nearly identically in their functionality

  ➢ search functions, document archiving, online interfaces for self archiving, integration of the OAI PMH, …

➢ Systems base on different technologies

  ➢ Eprints: traditional technologies, runs on pure open source systems: mySQL and Apache, programmed by using the script language "Perl"

  ➢ Dspace: operates with new technologies such as the Postgres database and Tomcat for jsp/java web application, higher performance

Questionnaire: Metadata Formats

Multiple answers possible

Format	already OAI compatible	not yet OAI compatible
Others (single mentioned)	12	6
METS		3
TEI	1	2
MAB	1	2
EAD	1	3
UNIMARC	3	2
MARC 21	4	2
Dublin Core qualified	10	5
Dublic Core simple	17	7

□ already OAI compatible ■ not yet OAI compatible

➢ **Problem: Standardisation**

"heterogeneity of the content of the metadata records requires the service provider to expend a lot of effort in normalizing the data in order to make it more comparable and usable"

  ✖ could be done at lesser cost by the individual DP

  ✖ or by the development of middleware tools that service providers could use for data normalisation

➢ **Lacking interoperability**

  ✖ different metadata standards, terminology, languages, access strategies, interfaces / transfer protocols, copyright regulations

➢ **Difficulty to establish joint services based on open archives**

Creating Recommendations

Example of a possible solution:

German Initiative for Networked Information:

Recommendations for usage of OAI-PMH

➢ created by DINI-OAI working group (http://www.dini.de/)

➢ target: agreement on syntax and semantics of OAI set definitions for German data and service providers

➢ enhance retrieval quality and support subject gateways (e.g. Physnet, Dissertation search engine, ...)

➢ definition of three classification types

- subjects (according to DNB)
- formal publication types (e.g. dissertation)
- formal document types (e.g. text, audio)

➢ example service provider based on recommended sets: http://edoc.hu-berlin.de/e_suche/oai.php

OAI advantages ?

➢ **Importance of OAI**

- provide additional services to existing services

- replace existing services through OAI interface

- better retrieval, make Metadata exchange available

➢ **Advantages of OAI**

- share scientific knowledge, harvest other knowledge databases, cross-search in institutional assets

- major dissemination of researchers' results

- simple and cheap in implementation

→ „provide access to all of human knowledge"

→ „nothing other than political expediency"

http://www.oaforum.org/oaf_db/

Resources:

Information Resource | Glossary | Talks and Presentations |

Articles | Other Resources | Technical Validation

Home

News

Overview

Resources

Workshops

Project Documents

Contact us

Information Resource Database

In order to bring interested parties together, to build a community of interest, and to enable exchange of information OAForum established this web-based European information source for open archives.

This application allows you to

... browse current lists of

- **Services**
- **Repositories**
- **Projects**
- **Software**
- **Organisations**
- **Protocols**
- **Metadata schemes**

... search in the database

Make sure you have only used blank (" ") to seperate different search terms.
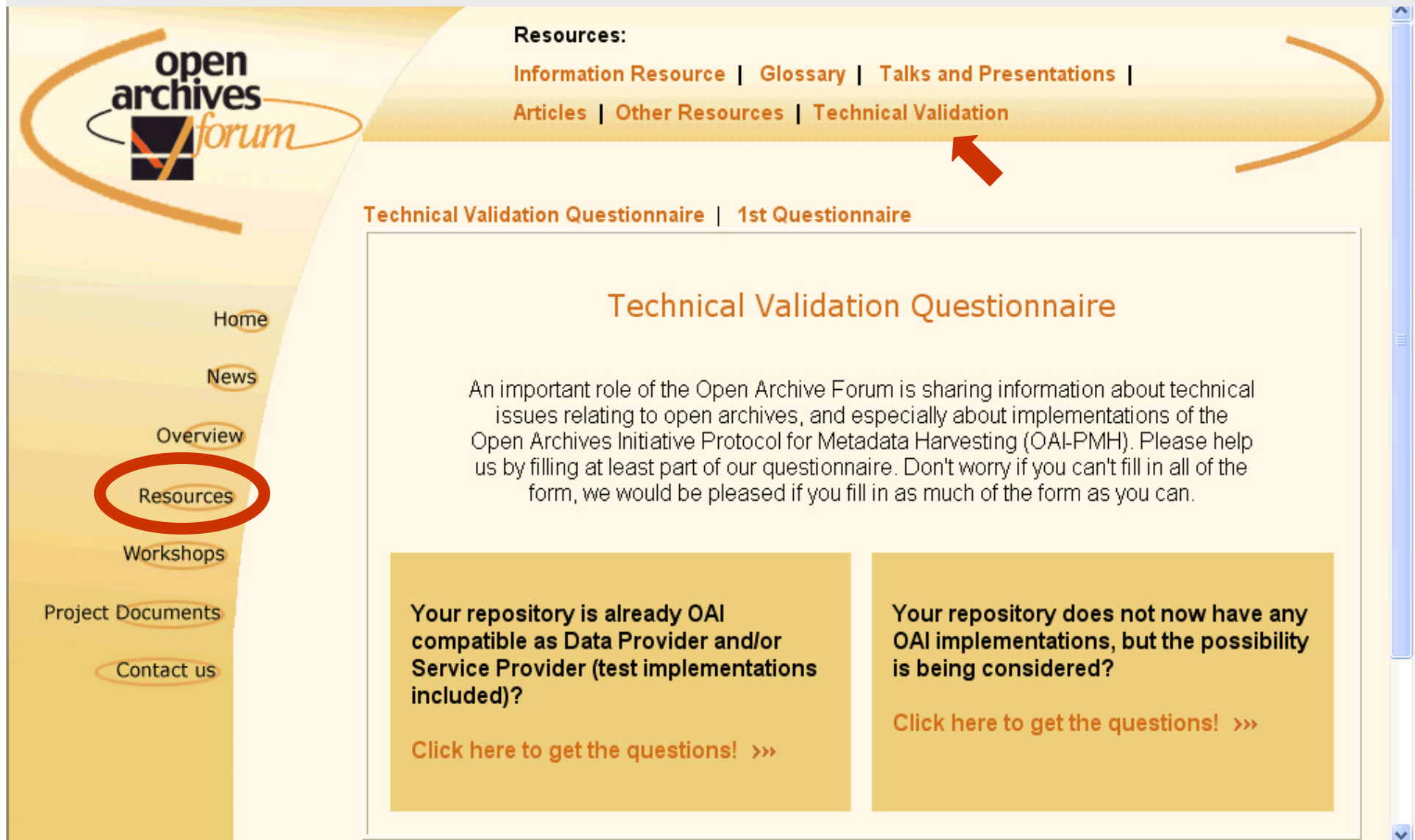
Search for: [_____] [Go]

... be part of the OAForum open archives information source?

›» **Register here**

All European services, repositories, projects and organisations that use an open archives approach, have started metadata initiatives or have developed appropriate software can register themselves using our web

http://www.oaforum.org/resources/tecvalq2.php

Resources:

Information Resource | **Glossary** | **Talks and Presentations** |

Articles | **Other Resources** | **Technical Validation**

Home

News

Overview

Resources

Workshops

Project Documents

Contact us

Technical Validation Questionnaire | **1st Questionnaire**

Technical Validation Questionnaire

An important role of the Open Archive Forum is sharing information about technical issues relating to open archives, and especially about implementations of the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). Please help us by filling at least part of our questionnaire. Don't worry if you can't fill in all of the form, we would be pleased if you fill in as much of the form as you can.
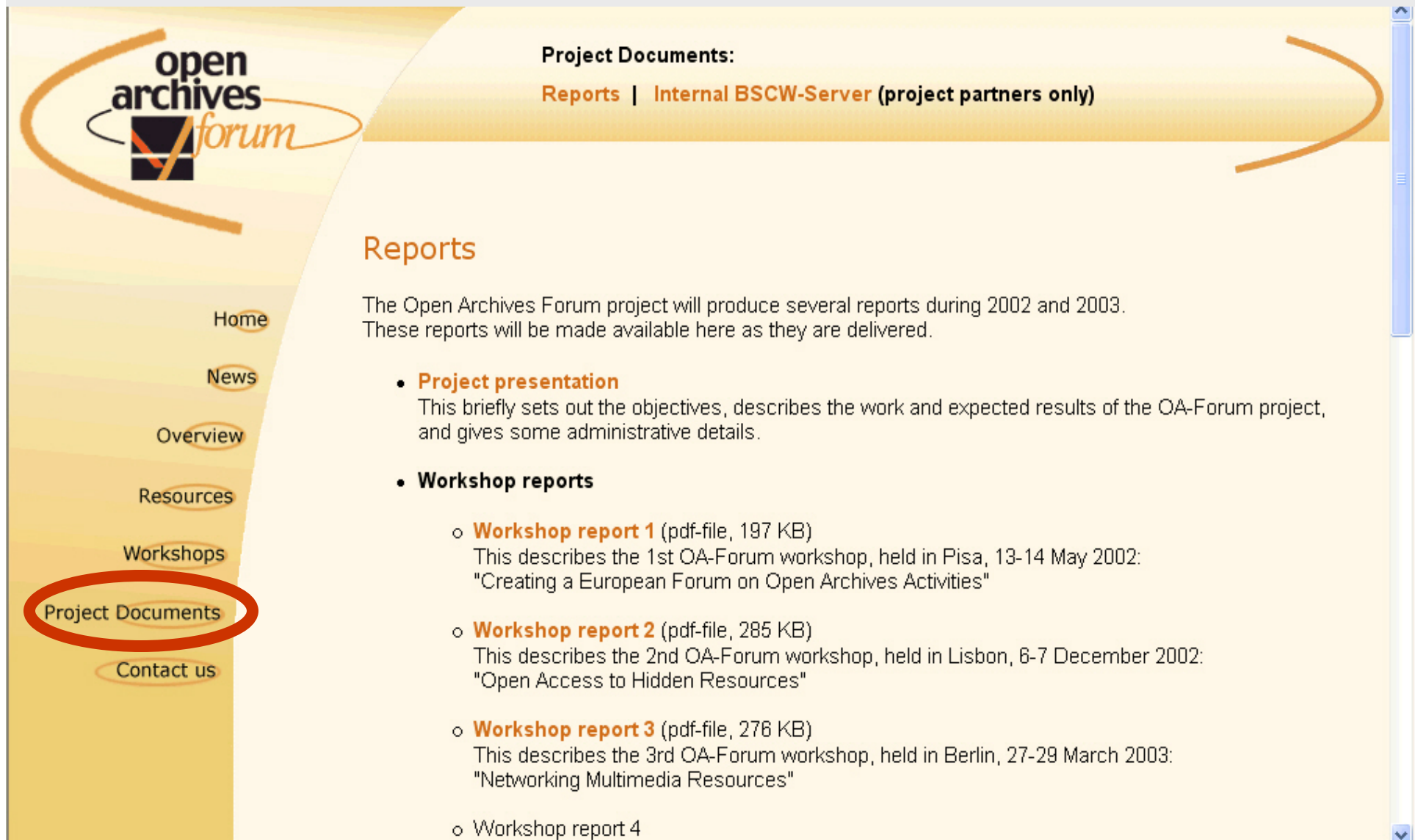
Your repository is already OAI compatible as Data Provider and/or Service Provider (test implementations included)?

Click here to get the questions! ⋙

Your repository does not now have any OAI implementations, but the possibility is being considered?

Click here to get the questions! ⋙

http://www.oaforum.org/documents/

Project Documents:

Reports | **Internal BSCW-Server** (project partners only)

Reports

The Open Archives Forum project will produce several reports during 2002 and 2003.
These reports will be made available here as they are delivered.

- **Project presentation**
  This briefly sets out the objectives, describes the work and expected results of the OA-Forum project, and gives some administrative details.

- **Workshop reports**

  - **Workshop report 1** (pdf-file, 197 KB)
    This describes the 1st OA-Forum workshop, held in Pisa, 13-14 May 2002:
    "Creating a European Forum on Open Archives Activities"

  - **Workshop report 2** (pdf-file, 285 KB)
    This describes the 2nd OA-Forum workshop, held in Lisbon, 6-7 December 2002:
    "Open Access to Hidden Resources"

  - **Workshop report 3** (pdf-file, 276 KB)
    This describes the 3rd OA-Forum workshop, held in Berlin, 27-29 March 2003:
    "Networking Multimedia Resources"

  - Workshop report 4

Navigation

Home

News

Overview

Resources

Workshops

Project Documents

Contact us

Thank you!

Birgit Matthaei

Humboldt University Berlin, Germany

Computer and Media Service, Electronic Publishing Group

birgit.matthaei@cms.hu-berlin.de