MAX PLANCK
digital library

eSciDoc Data Repository Infrastructure in Science Trends and Practice in Europe

Malte Dreyer

This work is licensed under a Creative Commons Attribution 2.0 Germany License
http://creativecommons.org/licenses/by/2.0/de/

The Max Planck Society (MPG – Max-Planck Gesellschaft)

- 80 institutes
  - 265 directors, 4700 Scientists, ~ 12.000 Young Academics from allover the world
- 30 locations
- Intense cooperation with universities

distributed organization

cooperation between different locations is very common

locations outside Germany are not shown (Nijmegen, Florence, Rome, …)

Fields

- diverse set of fields
  - Chemistry, Physics, Technology
  - Biology, Medicine, Brain Science
  - Law, Art, History, Cognition
- not a broad as many universities
- but more interdisciplinary

Max Planck Digital Library MPDL

- Founded in 2007
- Departments
  - Information Provision
  - Research and Development
  - Open Access
- Complemented by activities in many institutes
- Close cooperation with Institutes

eSciDoc

- eSciDoc
  - Initially Joint Project with FIZ Karlsruhe (JST is their partner in Japan)
  - Initially Funded by German Ministry
    - eResearch Environment
    - eScience Infrastructure
    - Solutions on top
  - Now Community Open Source

- Strategic Platform for Max Planck Society

Verkehr
Mehr...
Karte
Baffin Sea
Greenland
Iceland
Hudson Bay
Sverige
Sweden
Suomi
Finland
Norge
Norway
Россия
Russia
United Kingdom
Ireland
Danmark
Polska
Poland
Беларусь
Belarus
Україна
Ukraine
Österreich
Austria
France
Romania
Italia
Italy
España
Spain
Portugal
Ελλάς
Greece
Türkiye
Turkey
Kazakhstan
Mongolia
Uzbekistan
Kyrgyzstan
Turkmenistan
Syria
Iraq
Iran
Afghanistan
Pakistan
中国
China
대한민국
S Korea
日本
Japan
Morocco
Tunisia
Algeria
Libya
مصر
Egypt
Saudi Arabia
Oman
Yemen
India
Nepal
Myanmar (Burma)
Thailand
Vietnam
Malaysia
Indonesia
Papua New Guinea
Western Sahara
Mauritania
Mali
Niger
Chad
Sudan
Burkina Faso
Guinea
Ghana
Nigeria
Ethiopia
Somalia
Kenya
Gabon
DR Congo
Tanzania
Angola
Zambia
Mozambique
Namibia
Zimbabwe
Botswana
Madagascar
South Africa
Indian Ocean
Australia
NT
QLD
WA
SA
NSW
VIC
South Atlantic Ocean
North Atlantic Ocean
Gulf of Mexico
México
Cuba
PR
Guatemala
Nicaragua
Venezuela
Guyana
Suriname
Colombia
Ecuador
Brasil
Brazil
Perú
Bolivia
Paraguay
Chile
Uruguay
Argentina
NU
MB
SK
NL
ON
QC
ND
SD
WY
NE
CO
KS
NM
OK
TX
LA
MN
WI
MI
IA
IL
IN
OH
PA
NY
VT
ME
NB
NS
PE
NH
MA
CT
RI
NJ
DE
MD
VA
WV
KY
TN
NC
SC
GA
AL
MS
FL
AM
PA
MA
CE
RN
PB
PE
AL
SE
PI
AC
RO
TO
BA
MT
GO
MG
MS
SP
ES
RJ
PR
SC
RS
©2009 Google - Kartendaten ©2009 Tele Atlas, ZENRIN, MapLink/Tele Atlas, AND, Europa Technologies -

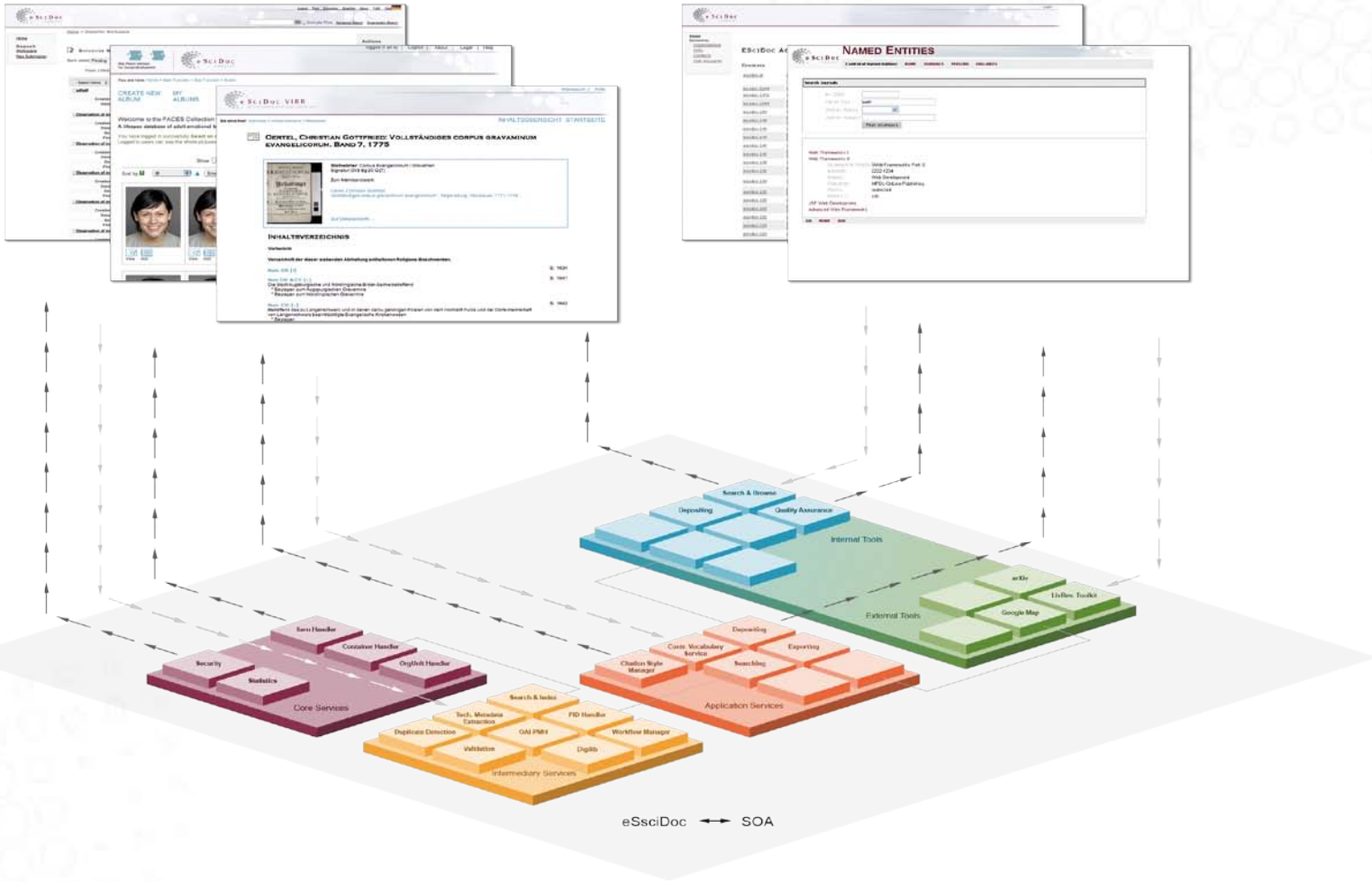
Infrastructure /1: Fixed Purpose

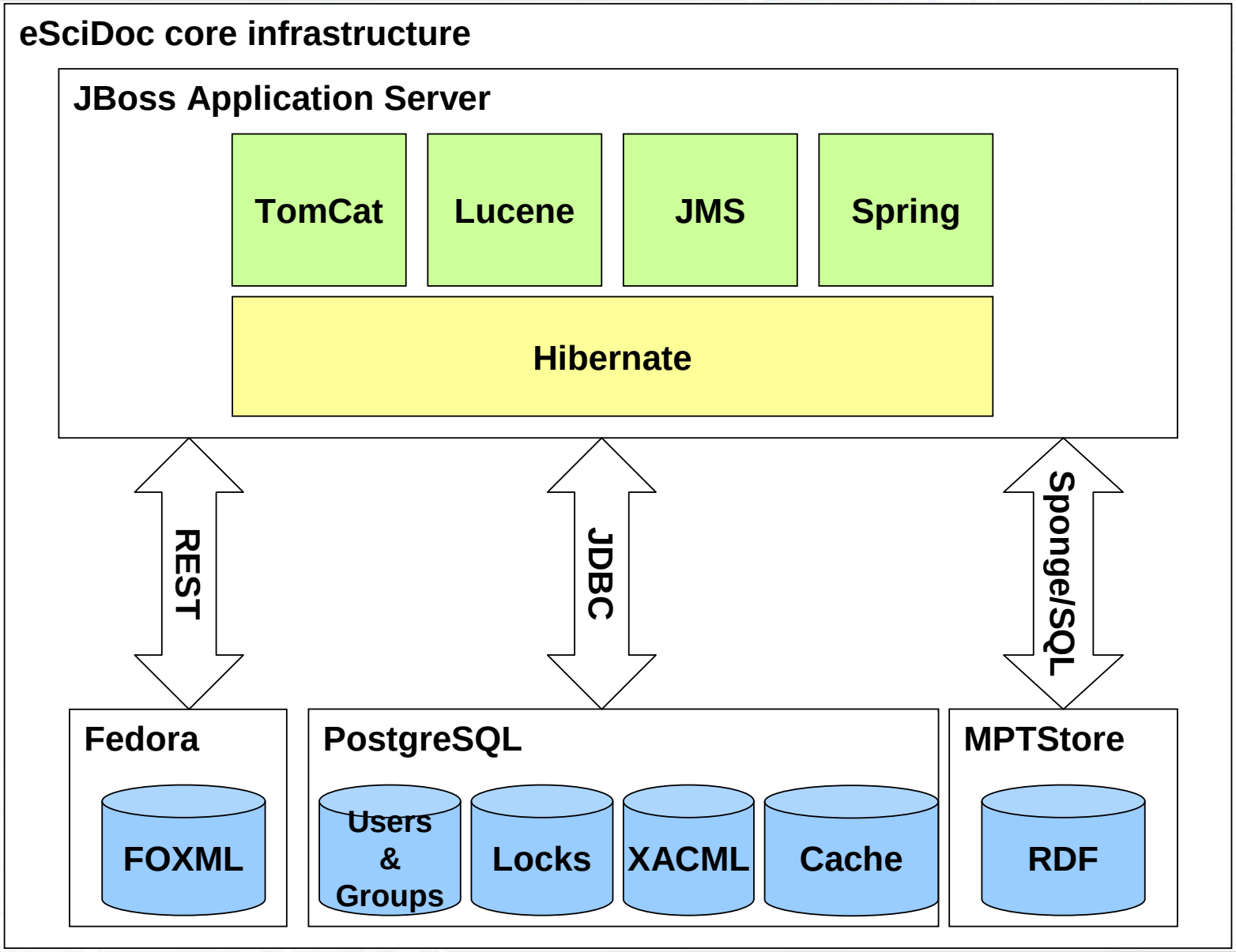
http://commons.wikimedia.org/wiki/File:DA_ph_dreveny_most01.jpg (CC-By-SA 3.0)

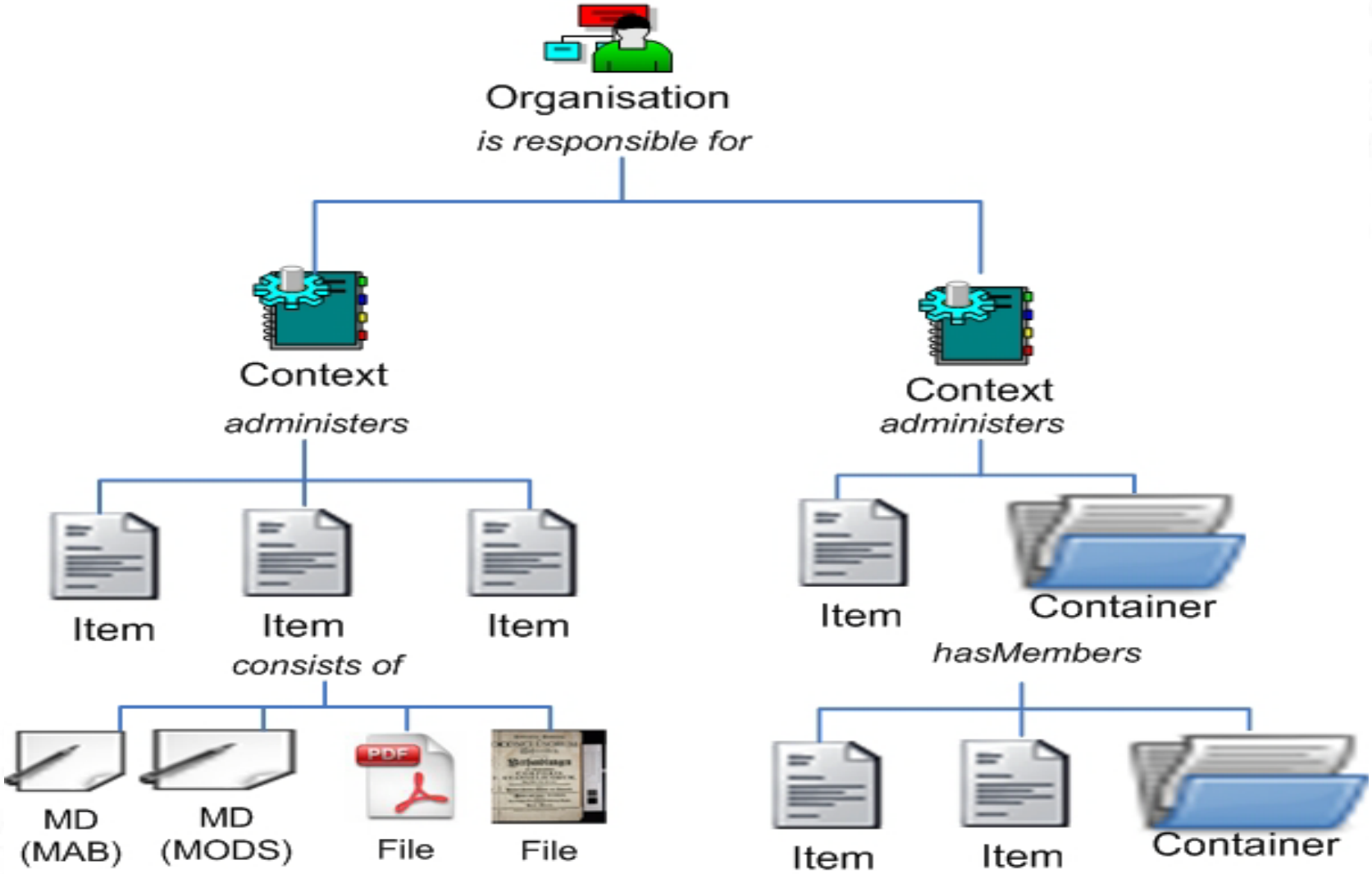
Infrastructure /2: More flexible

http://commons.wikimedia.org/wiki/File:Wellington_infrastructure.jpg (CC-By 2.0)

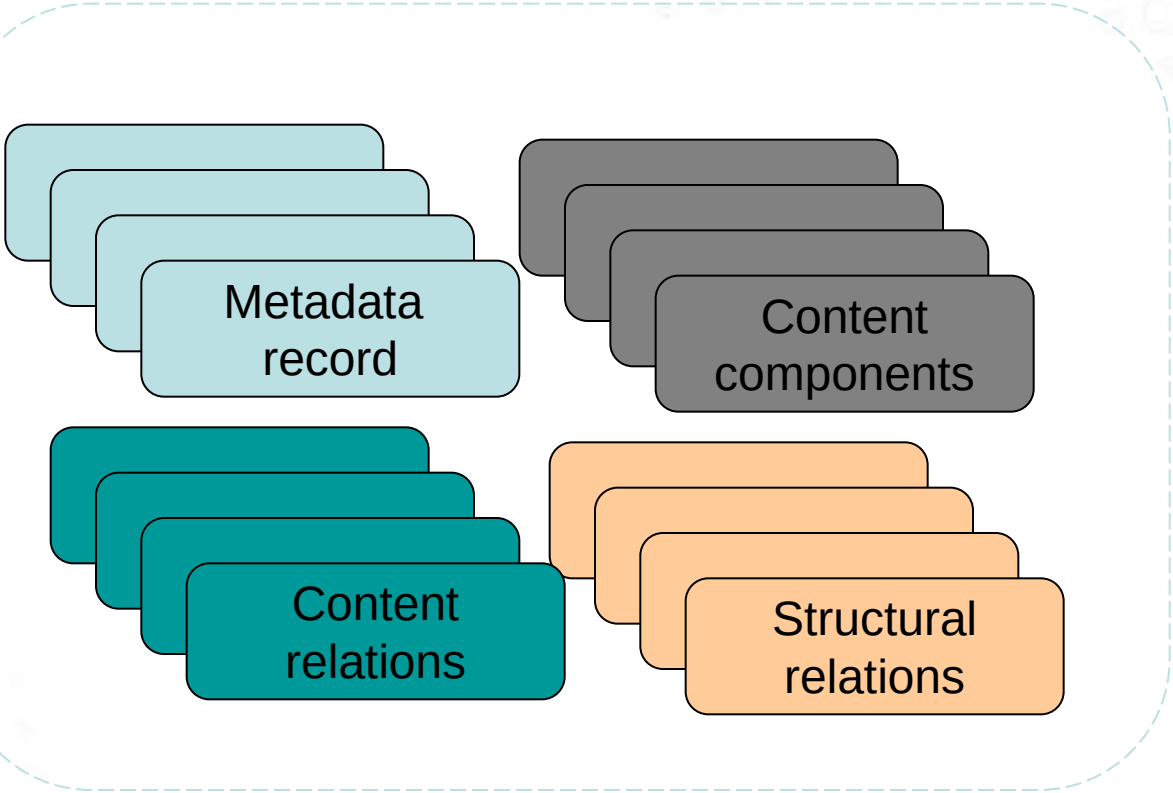
eSciDoc Architecture




eSciDoc core infrastructure
JBoss Application Server
TomCat
Lucene
JMS
Spring
Hibernate
REST
JDBC
Sponge/SQL
Fedora
FOXML
PostgreSQL
Users & Groups
Locks
XACML
Cache
MPTStore
RDF

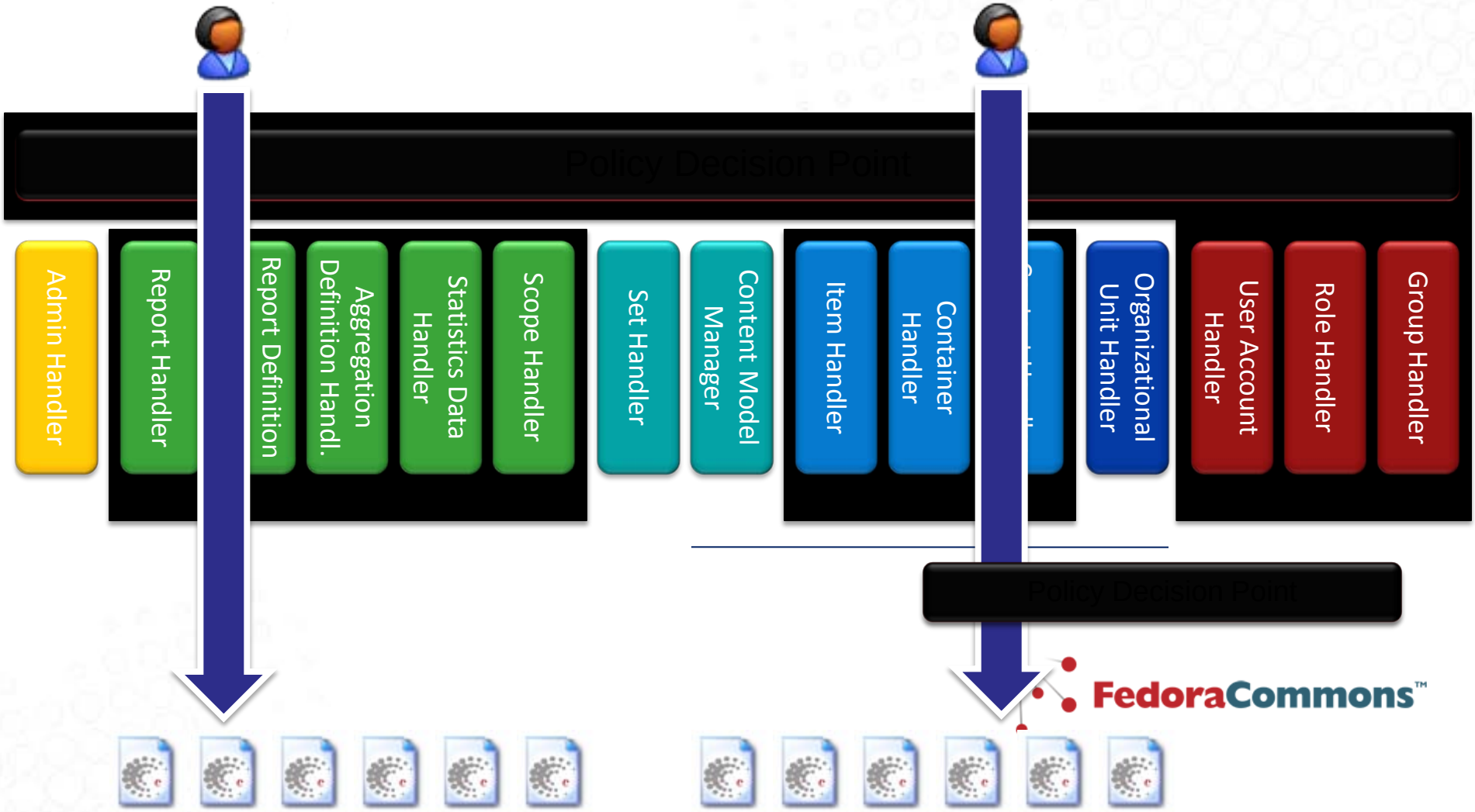

Organisation
is responsible for
Context
administers
Item
Item
Item
consists of
MD
(MAB)
MD
(MODS)
PDF
File
File
Context
administers
Item
Container
hasMembers
Item
Item
Container

Items



- Plus: Audit Trail, Properties, Techn. MD, Context Information,…….

Core Services / Security

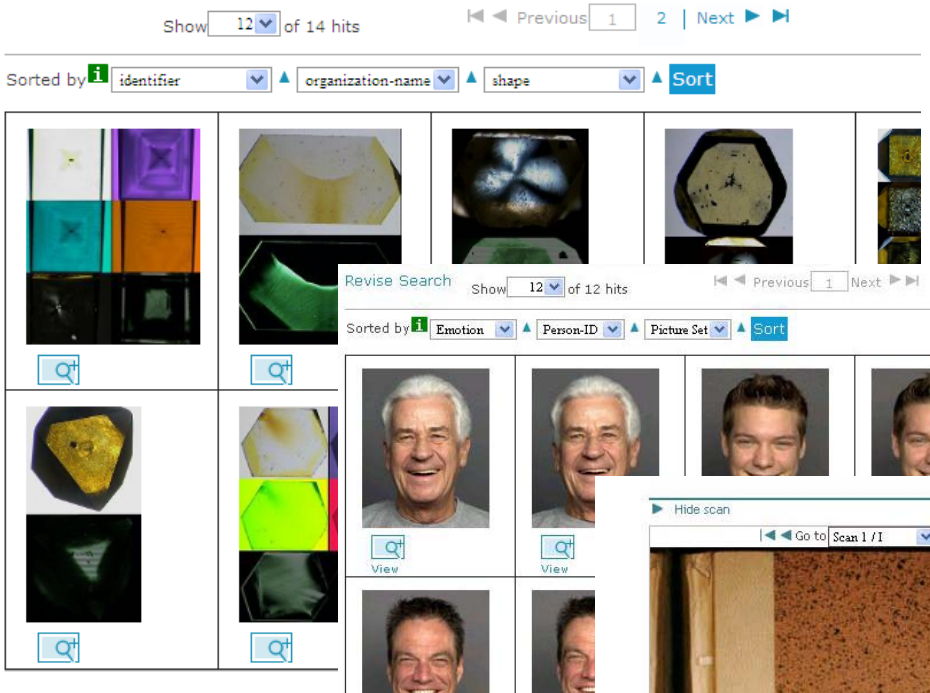


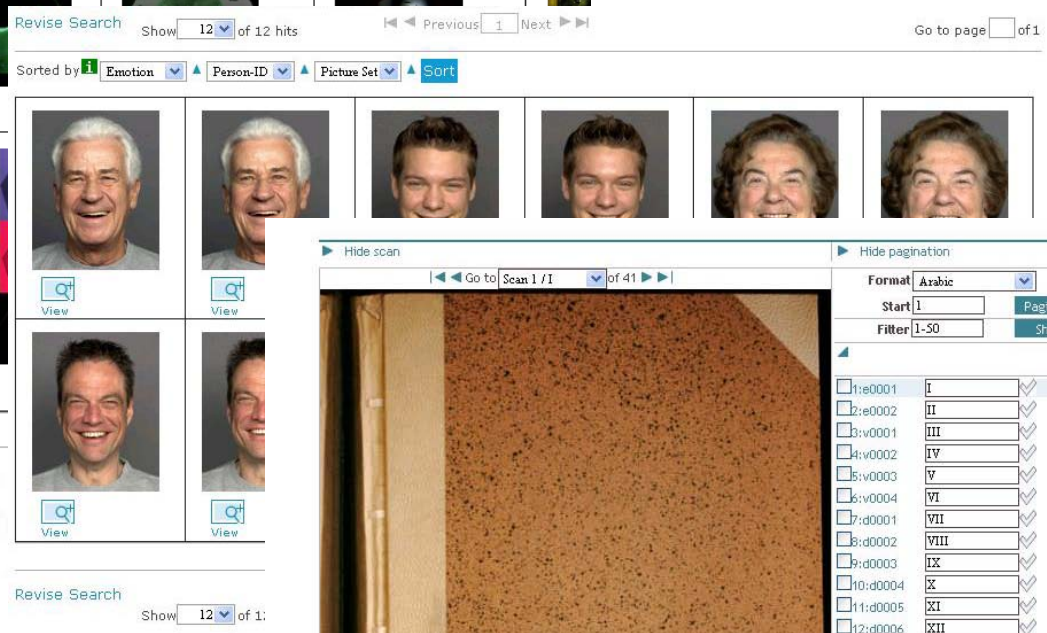
eSciDoc Scenarios Cloud

Max Planck digital library

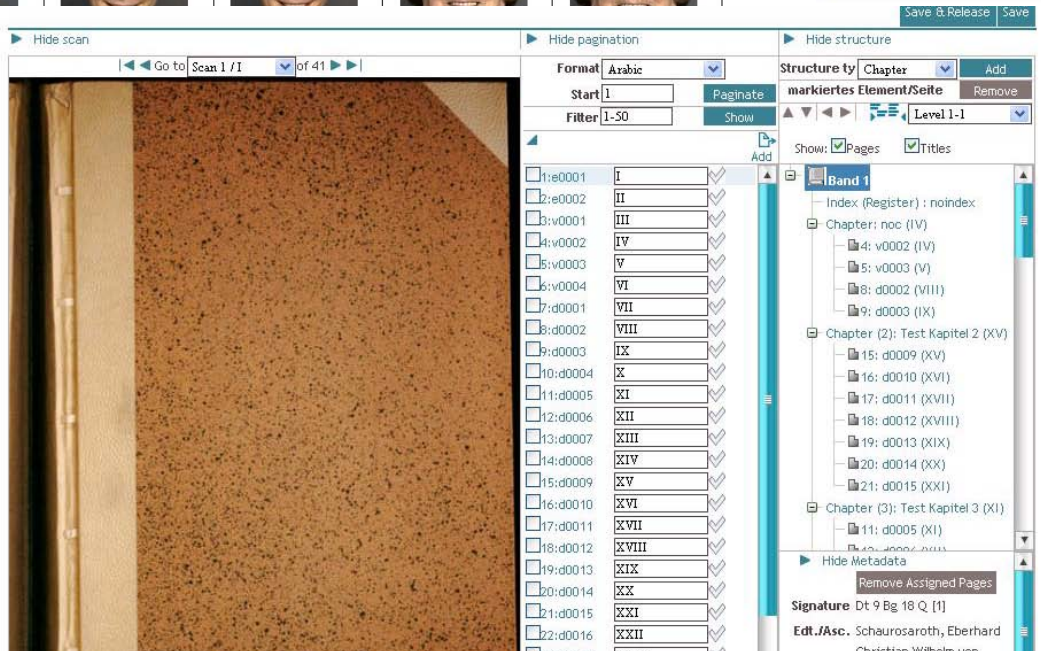
Solutions Online /1

Solutions /2

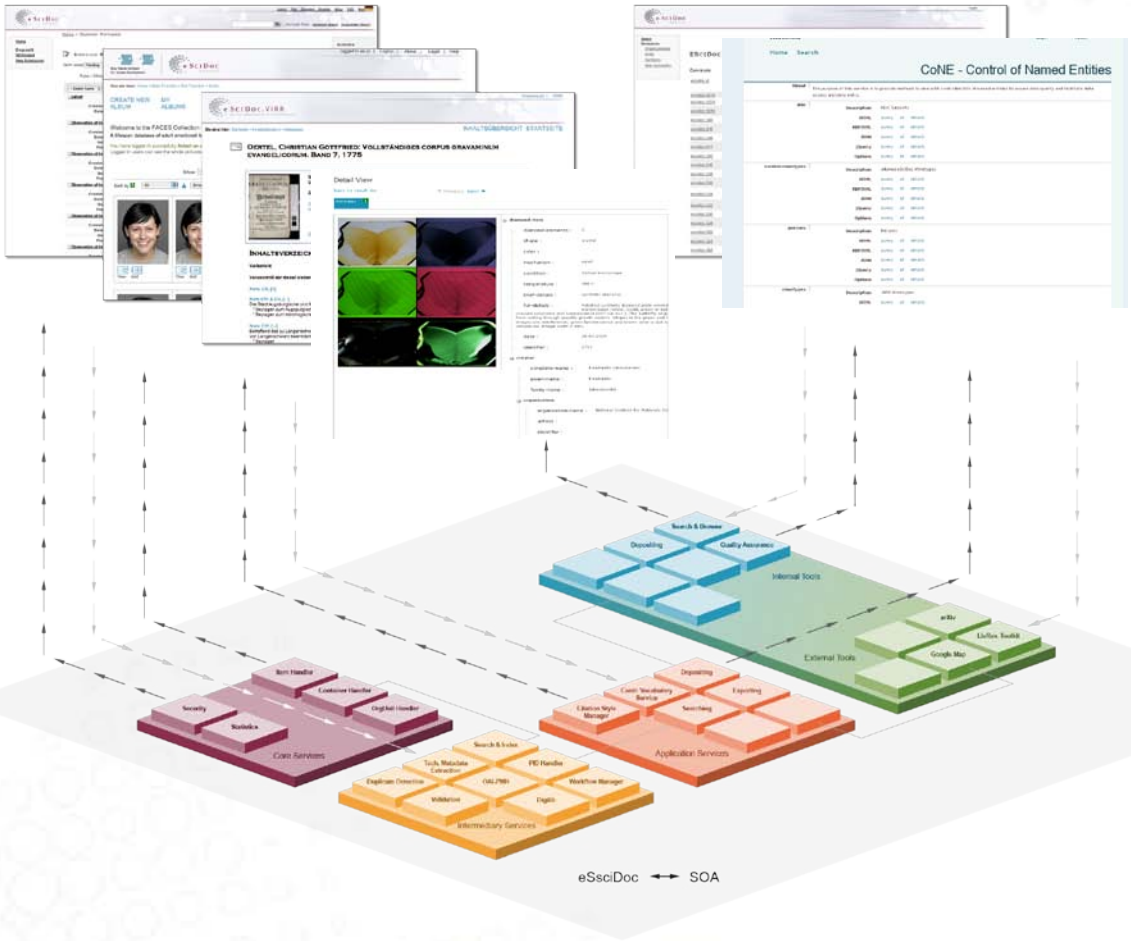

Show 12 of 14 hits
Previous 1 2 | Next
Sorted by identifier organization-name shape Sort

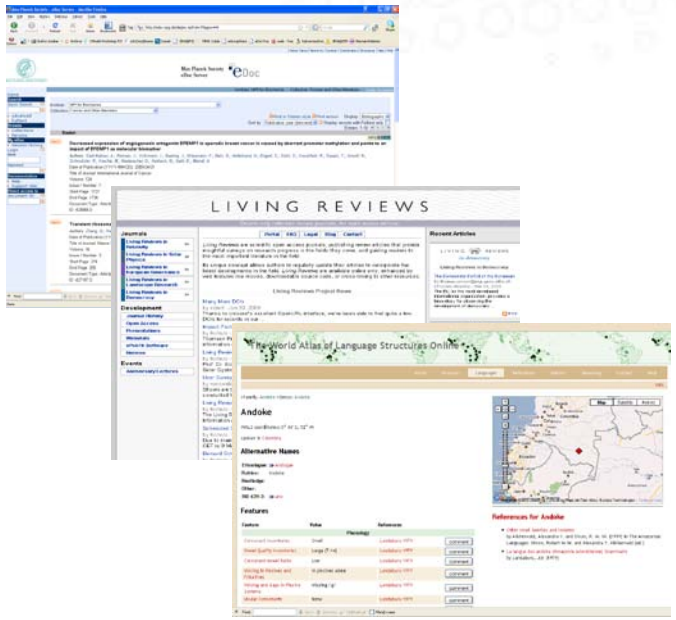

Revise Search
Show 12 of 12 hits
Previous 1 Next
Go to page 1 of 1
Sorted by Emotion Person-ID Picture Set Sort
View

Project Home
Subprojects
People
Publications
Tools
News
Events
Information structure in language acquisition - Publications
Journal article
Max Planck Institute for Psycholinguistics
Reinhold von Sengbusch
About

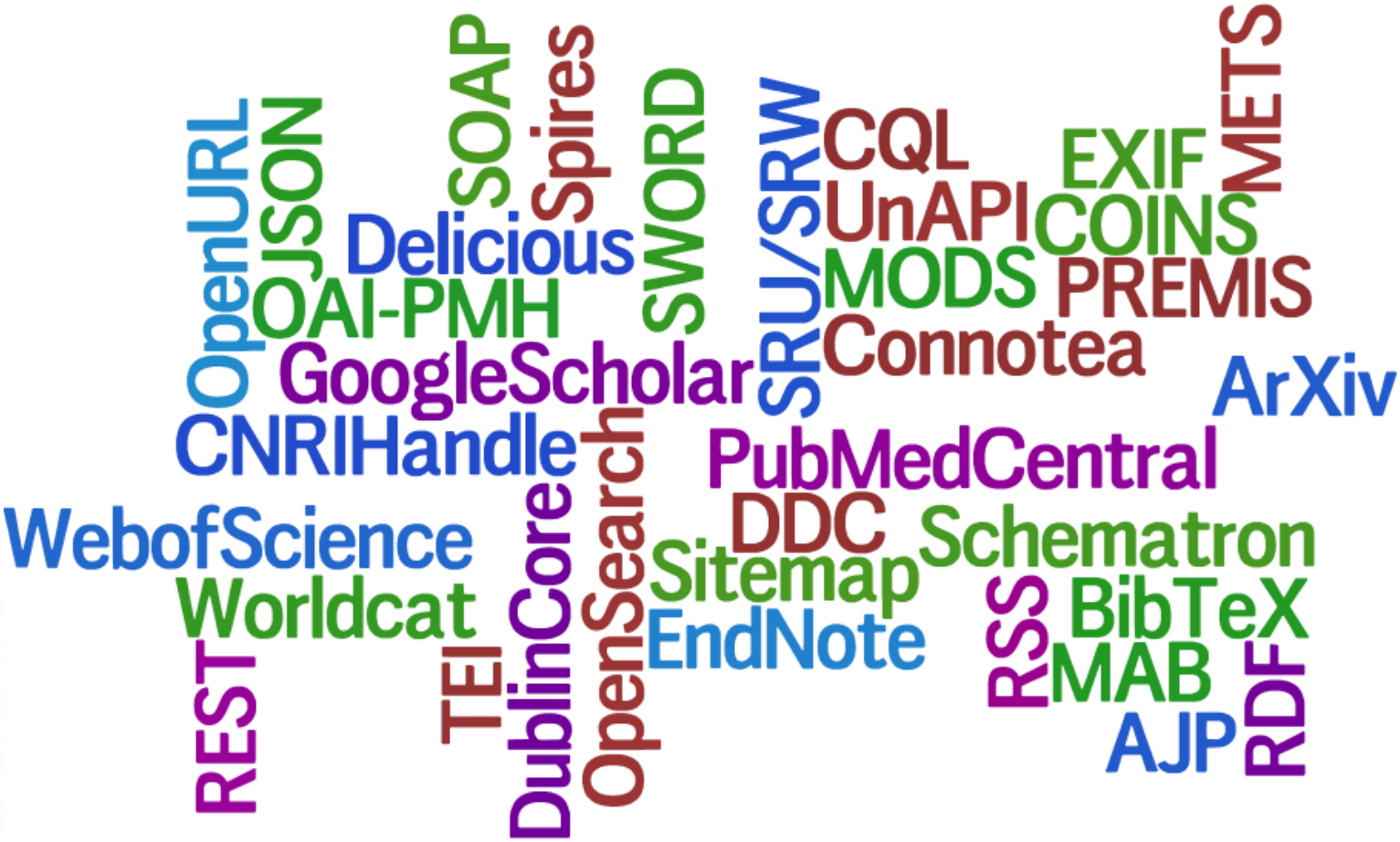

Hide scan
Go to Scan 1 / I of 41
Hide pagination
Format Arabic
Start 1
Filter 1-50
Paginate
Show
Hide structure
Structure ty Chapter
Add
markiertes Element/Seite
Remove
Level 1-1
Show: Pages Titles
Band 1
Index (Register) : noindex
Chapter: noc (IV)
Chapter (2): Test Kapitel 2 (XV)
Chapter (3): Test Kapitel 3 (XI)
Hide Metadata
Remove Assigned Pages
Signature Dt 9 Bg 18 Q [1]
Edt./Asc. Schaurosaroth, Eberhard

Building Infrastructure and Solutions





Selected Supported Protocols/Standards/Specifications

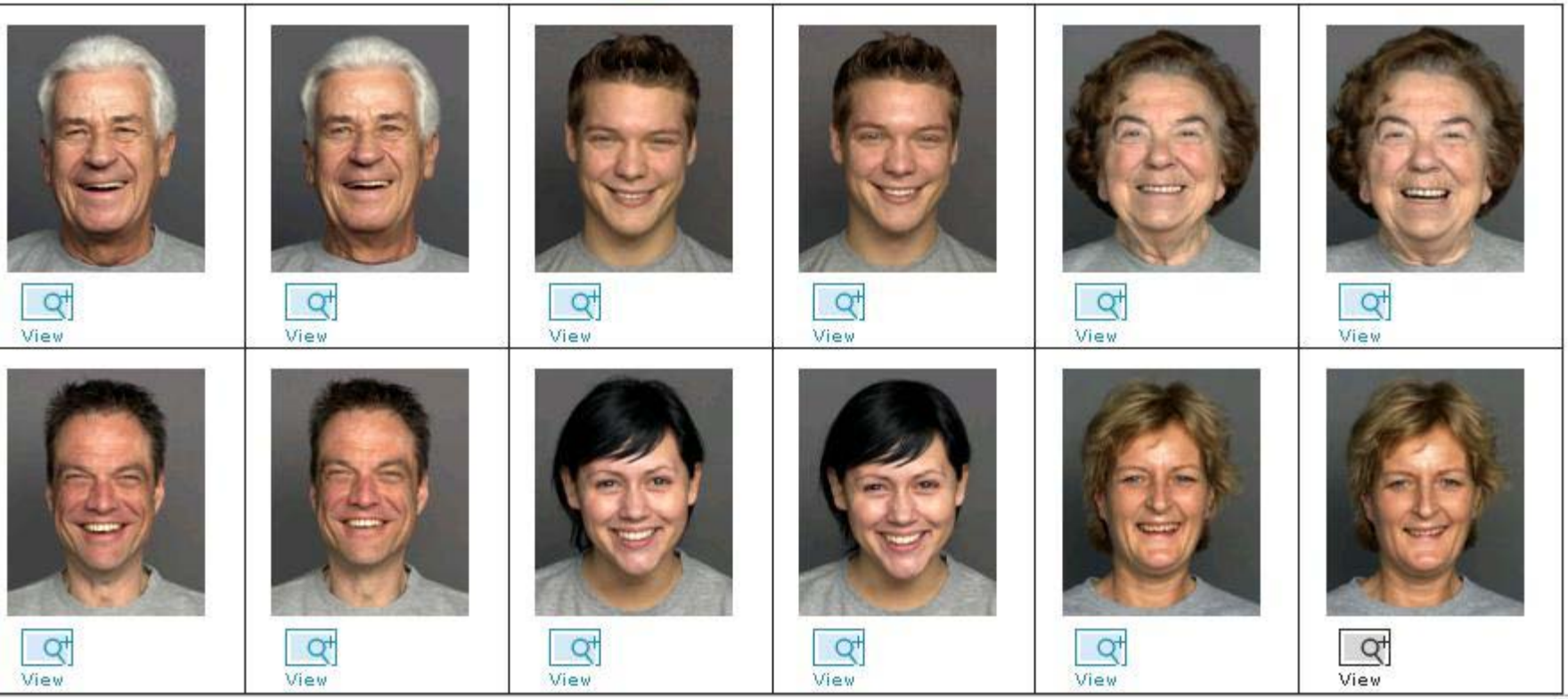


Search Result

You searched in the the whole Faces Collection for (happiness)

Revise Search Show 12 of 12 hits Previous 1 Next Go to page of 1

Sorted by Emotion ▲ Person-ID ▲ Picture Set ▲ Sort



Revise Search Show 12 of 12 hits Previous 1 Next Go to page of 1

Published Albums

Search in album list Search

Show 10 of 2 hits Previous 1 Next Go to page of 1

Sorted by Date published

Name	Number of pictures	Description	Owner	Affiliation	Date published
Expression Research on Middle Aged Men	8	This album contains a selection of middle aged men with different facial expressions.	Kleinfercher, Friederike	Max Planck Digital Library	2009-03-27 at 13:49:12 CET
DEMO FACES Collection	72	This is the FACES Collection of test-faces.mpdl.mpg.de	Saquet, Bastien	MPDL	2009-03-27 at 13:14:17 CET

Show 10 of 2 hits Previous 1 Next Go to page of 1

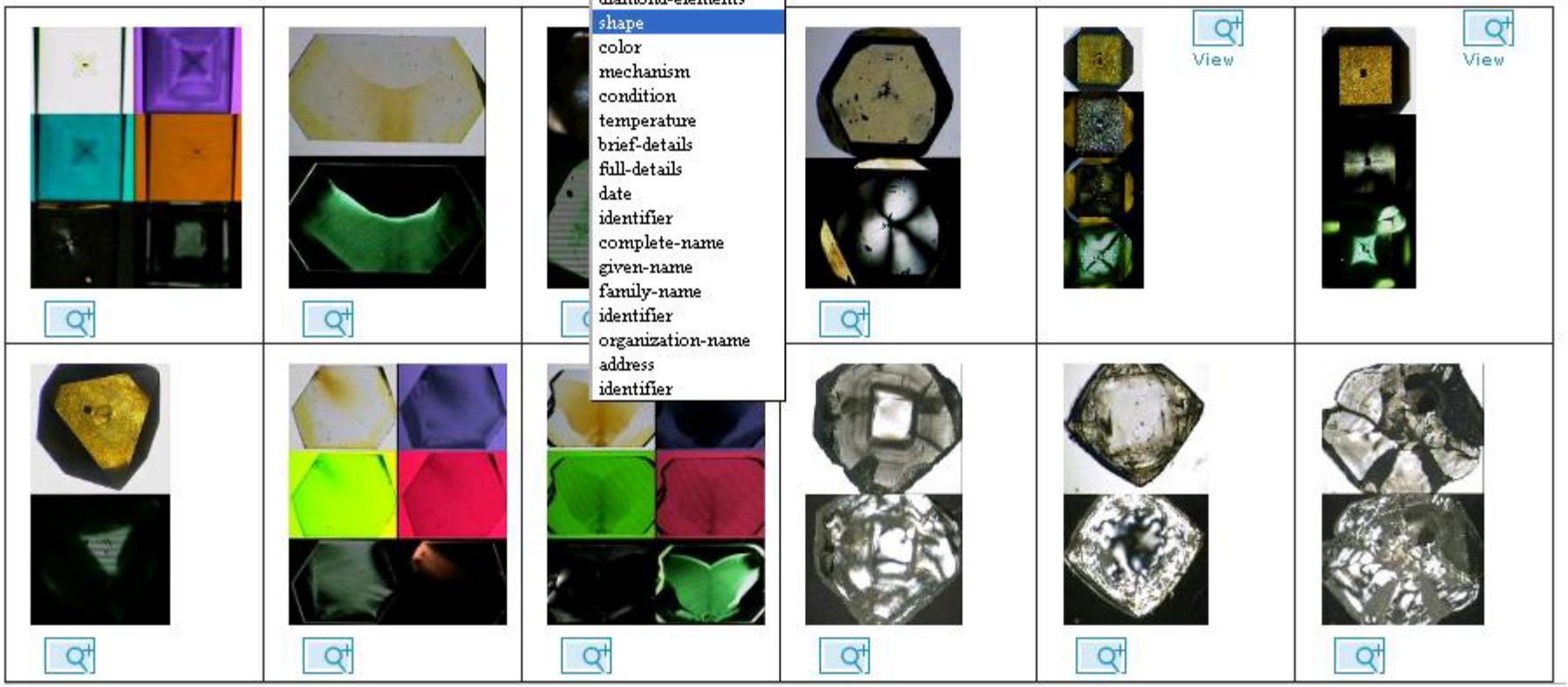
Browse

14 images to browse

Show 12 of 14 hits Previous 1 2 | Next Go to page of 2

Sorted by identifier ▲ organization-name ▲ shape ▲ Sort

- diamond-elements
- shape
- color
- mechanism
- condition
- temperature
- brief-details
- full-details
- date
- identifier
- complete-name
- given-name
- family-name
- identifier
- organization-name
- address
- identifier

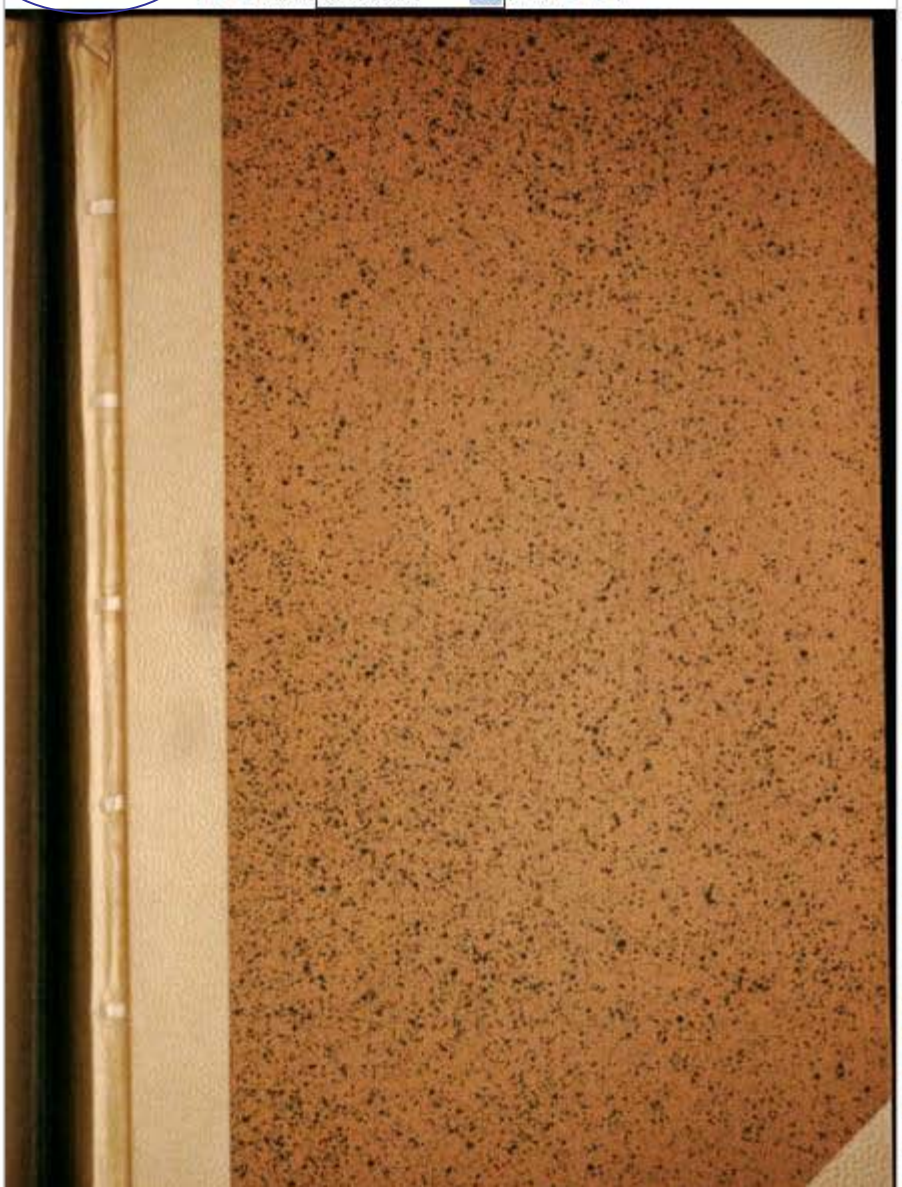


Show 12 of 14 hits Previous 1 2 | Next Go to page of 2

Vollständige Sammlung Aller Conclusorum, Schreiben Und anderer übrigen Verhandlungen Des Hochpreißlichen Corporis Evangelicorum nnnnnn. Vol. 1, 1751

Save & Release | Save

Hide scan | Hide pagination | Hide structure



Format: Arabic | Start: 1 | Paginate | Filter: 1-50 | Show | Add

1:e0001	I
2:e0002	II
3:v0001	III
4:v0002	IV
5:v0003	V
6:v0004	VI
7:d0001	VII
8:d0002	VIII
9:d0003	IX
10:d0004	X
11:d0005	XI
12:d0006	XII
13:d0007	XIII
14:d0008	XIV
15:d0009	XV
16:d0010	XVI
17:d0011	XVII
18:d0012	XVIII
19:d0013	XIX
20:d0014	XX
21:d0015	XXI
22:d0016	XXII
23:d0017	XXIII
24:d0018	XXIV
25:d0019	XXV
26:d0020	XXVI
27:d0021	XXVII

Structure ty: Chapter | Add | markiertes Element/Seite | Remove | Level 1-1

Show: Pages Titles

- Band 1
  - Index (Register) : noindex
  - Chapter: noc (IV)
    - 4: v0002 (IV)
    - 5: v0003 (V)
    - 8: d0002 (VIII)
    - 9: d0003 (IX)
  - Chapter (2): Test Kapitel 2 (XV)
    - 15: d0009 (XV)
    - 16: d0010 (XVI)
    - 17: d0011 (XVII)
    - 18: d0012 (XVIII)
    - 19: d0013 (XIX)
    - 20: d0014 (XX)
    - 21: d0015 (XXI)
  - Chapter (3): Test Kapitel 3 (XI)
    - 11: d0005 (XI)

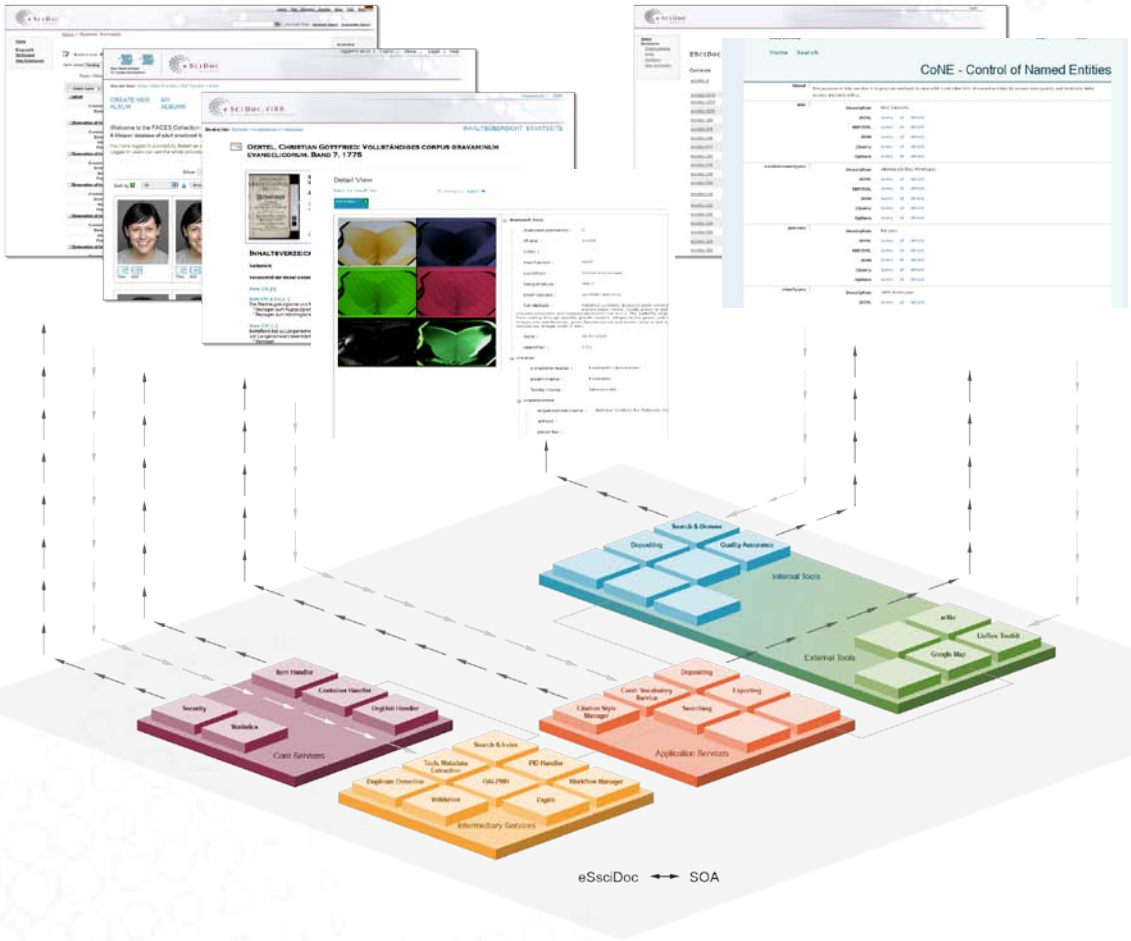
Hide Metadata | Remove Assigned Pages

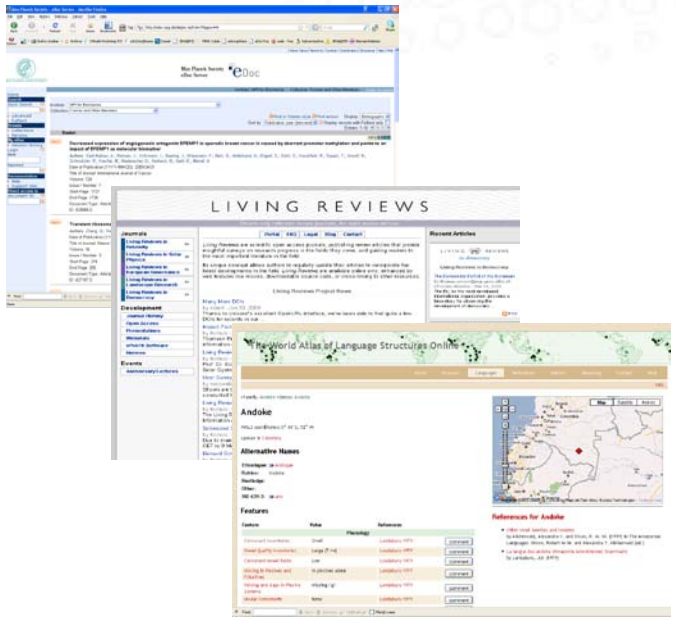
Signature Dt 9 Bg 18 Q [1]

Edt./Asc. Schaurosaroth, Eberhard Christian Wilhelm von

Title Vollständige Sammlung Aller Conclusorum, Schreiben Und anderer übrigen Verhandlungen

Building Infrastructure and Solutions

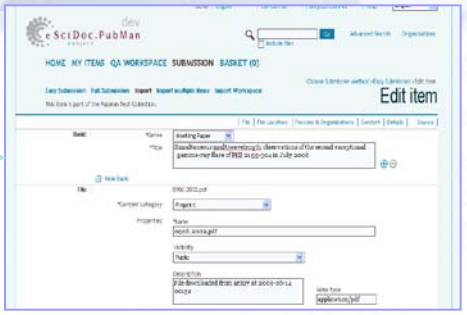




MAX PLANCK digital library

Services in action

3: display fetched resource



1: fetch resources

unAPI

Data Acquisition Service

*

4: Deposit

arXiv.org

PubMed Central

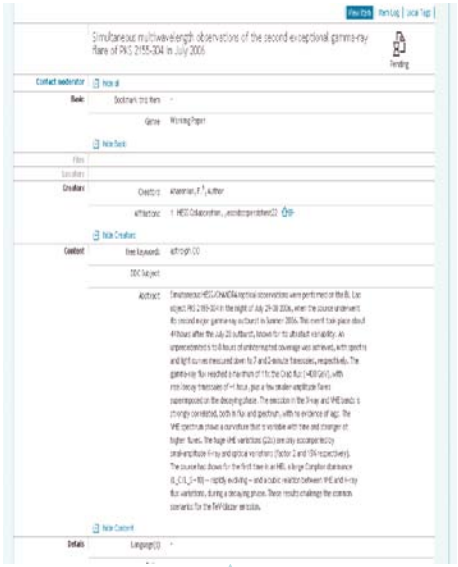
SPIRES

eSciDoc.PubMan PROJECT

2: transform

Fetches (metadata and files)
and deposits to eSciDoc repository

Depositing Service



Transformation Service

5: Validate

Validation Service

** Optionally, acquisition result may be forwarded to Depositing service directly*

6: Item deposited

Services in action

I need my (enriched) data back in *XML with APA Snippet* format

1:Query and ask format

Search&Export Service

2:Search and retrieve results

Search Service <core>

3:Results (raw)

3:Results (formatted)

Citation Style Manager

Search&Export: Process a (CQL) search query and delivers results in desired format

Citation style Manager: delivers bibliographic metadata formatted in a citation style (e.g. APA, AJP)

Results can be further reused in external applications, e.g. Researcher portfolio Wordpress Blog

Astronomer's Workbench

- Within the astronomy community, the need for a user-centred community platform to work collaboratively with distributed research data has been identified. The project will provide a discipline-specific variant of a "scholarly workbench scenario".

- Partners:
  - Max Planck Institute for Extraterrestrial Physics (MPE)
  - Max Planck Institute for Astrophysics (MPA)
  - Max Planck Digital Library

BW-eLabs

- BW-eLabs gives access to virtual and remote experiments in the field of nano technology. Key concepts of the project include the reproducability of experiments, discoverability of and access to primary data, and the storage and curation of all artifacts that emerge throughout the research process.

- Project Partners
  - University Stuttgart
  - HDM Stuttgart
  - University Freiburg
  - FIZ Karlsruhe

DARIAH

- Just like astronomers require a virtual observatory to study the stars and other distant objects in the galaxy, researchers in the arts and humanities need a digital infrastructure to get access to and join together the information and the knowledge that is embedded in digital content. The Digital Research Infrastructure for the Arts and Humanities – DARIAH – will be such an infrastructure with a European dimension. Results from eSciDoc serve for building one of the demonstrators

- Partners:
  - 14 Partners (UK, GR, DE, DK, FR, NL, CY, SI, HR, IE)

E-Kinematix

- E-Kinematix focuses on mechanical engineering, especially gearbox design. It will be based on an existing online library of gearbox design concepts. It will provice a library of calculation algorithms, tools for documenting the design process in a structured way, and a set of services supporting the collaborative aspects of the design process, which typically includes several stakeholders across institutions.

- Project Partners
  - RWTH Aachen
  - Technical University of Ilmenau
  - FIZ Karlsruhe

Micropublications

- To facilitate the collection of published description of the world's languages, the project aims to provide an integral online sourceindex of language descriptions and an online workspace to collaboratively store, access and publish interpretations of these sources.

- Partners:
  - University of Jena
  - University Leipzig
  - Max Planck Institute for Evolutionary Anthropology
  - Max Planck Digital Library

MATHEON – a research center for mathematics

- MATHEON - a hierarchical infrastructure for research (different levels)
- Top level: center (organizing and control)
- Basis level of MATHEON: projects
- The projects are organized in clusters corresponding to their
  - mathematical character
  - application area

- MATHEON – a distributed research center (TUB, HU, FU, WIAS, ZIB)

- MATHEON – a dynamic infrastructure: the projects have a lifecycle

Virtual research environments in MATHEON – What is needed?

- analyzing research processes in projects
- requirements to a virtual communication platform on project-level (MATHEON-level)
- typical Information flows

different aims:

- **building a virtual envoronment for research, e.g., a standardized desktop (internal)**
- organizing and nmanaging MATHEON (internal)
- information about MATHEON for the public (the existing MATHEON Web server provides a lot of information) (external)

One important aspect: The lifecycle of a project

- typically, at the beginning of a project in MATHEON there is a problem and an idea to solve it
- State of the art analysis of the problem: studying of existing results (proposals, ideas, publications)
- critical review of existing results (discussion, comments, meetings and conferences)
- Working plans, decisions
- Realization, implementation of ideas and control of the project

Knowledge base

Knowledge base of VirtMATHEON

- **archiving**: for existing literature (each scientist can store interesting publications here)
- **annotating** (automatically, manually): bibliographic and semantic metadata, e.g., keyphrases, classification, clustering
- **linking** (also with other types of information, e.g. , the profiles of scientists, events, software, etc.)
- **forum**: discussion, comments, conferences
- **publishing**: creating own publications (publishing function)

eSciDoc – a candidate for the knowledge base (and more)

Thanks for your attention!