

From Digital Libraries to e-Research in Europe: a Challenge for your Institutional Repository?

Matthias Razum
FIZ Karlsruhe

Library Fair
Yokohama, Japan
November 11, 2009

Agenda

- **Institutional Repositories, Digital Libraries, and their challenges**
- Data and Information Infrastructures in Europe
- Virtual Research Environments
- eSciDoc

Institutional Repositories and Digital Libraries

Characteristics of Institutional Repositories and Digital Libraries

- Content mostly traditional publications
- Maintained by librarians
- Selection and quality-ensurance processes in place
- Well-structured content
- High-quality metadata
- Data is fixed

Important IR Software

Usage of Institutional Repository software systems
(November 2009, based on ROAR):

IR System	World-wide	Germany
DSpace	32,1 %	6,3 %
EPrints	22,9 %	13,5 %
OPUS	2,4 %	30,6 %
Fedora	1,0 %	0,1 %



Registry of Open Access Repositories (ROAR), <http://roar.eprints.org/>

What is e-Research?

The invention and exploitation of advanced IT:

- to generate, curate and analyse **research data**;
- to develop and explore models and simulations;
- to enable dynamic distributed virtual organisations

Tony Hey (former Director of UK e-Science Core Programme, UK)

e-Science

+

e-Humanities

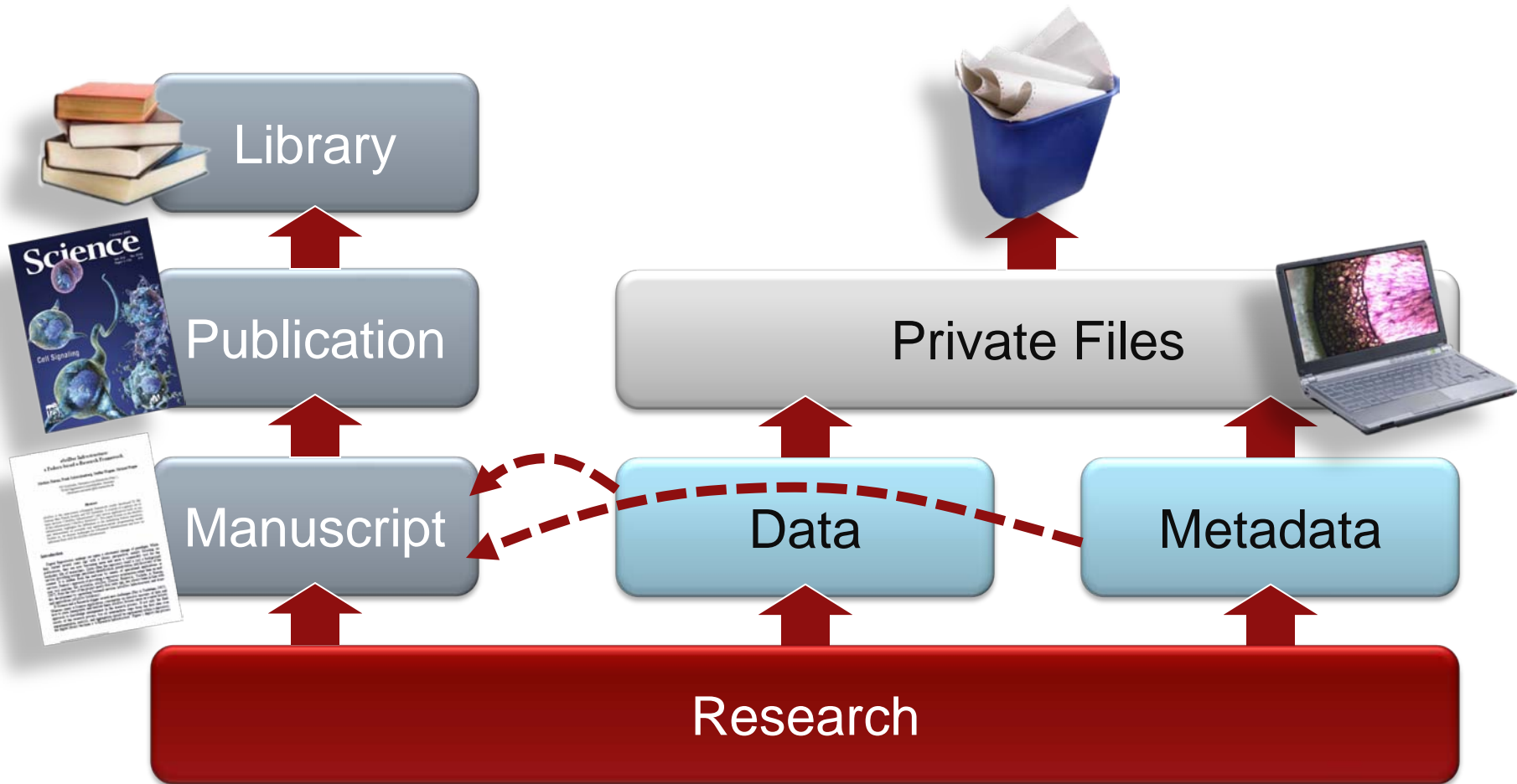
= e-Research

Data – The Next Challenge

- Most research today is data-driven, even in the humanities
- Reference data should be accessible independently of scientific papers in a citable form, allowing attribution
- Scholarly publication
 - remains essential for describing and contextualizing findings,
 - but is inadequate as the only document of research activity
- Databases
 - are essential for comprehensively archiving both published and unpublished data,
 - but are not yet fully integrated into the scientific process

Editorial „Sharing Data“, Nature Cell Biology, 11 (11), 2009

Data in the Publication Process Today



Helly, J., H. Staudigel, and A. Koppers, Scalable models of data sharing in Earth sciences, Geochem. Geophys. Geosyst., 4(1), 1010, doi:10.1029/2002GC000318, 2003

Agenda

- Institutional Repositories, Digital Libraries, and their challenges
- **Data and Information Infrastructures in Europe**
- Virtual Research Environments
- eSciDoc

Data and Information Infrastructures

“Equally, if not more important than its own data and information needs, today’s research community must also assume **responsibility** for building a **robust data and information infrastructure** for the **future**.”

International Council for Science, ICSU, 2004

Examples from Europe

- DRIVER
 - pan-European infrastructure for digital repositories
 - e-Infrastructure is used to store knowledge
- Enabling Grids for E-scienceE (EGEE)
 - Europe's leading grid computing project
 - providing a computing support infrastructure for over 10,000 researchers world-wide
- DILIGENT
 - advanced testbed for knowledge e-Infrastructure
 - integrating **grid** and **digital library** technology
- BRICKS
 - organizational and technological foundations of a digital library



Agenda

- Institutional Repositories, Digital Libraries, and their challenges
- Data and Information Infrastructures in Europe
- **Virtual Research Environments**
- eSciDoc

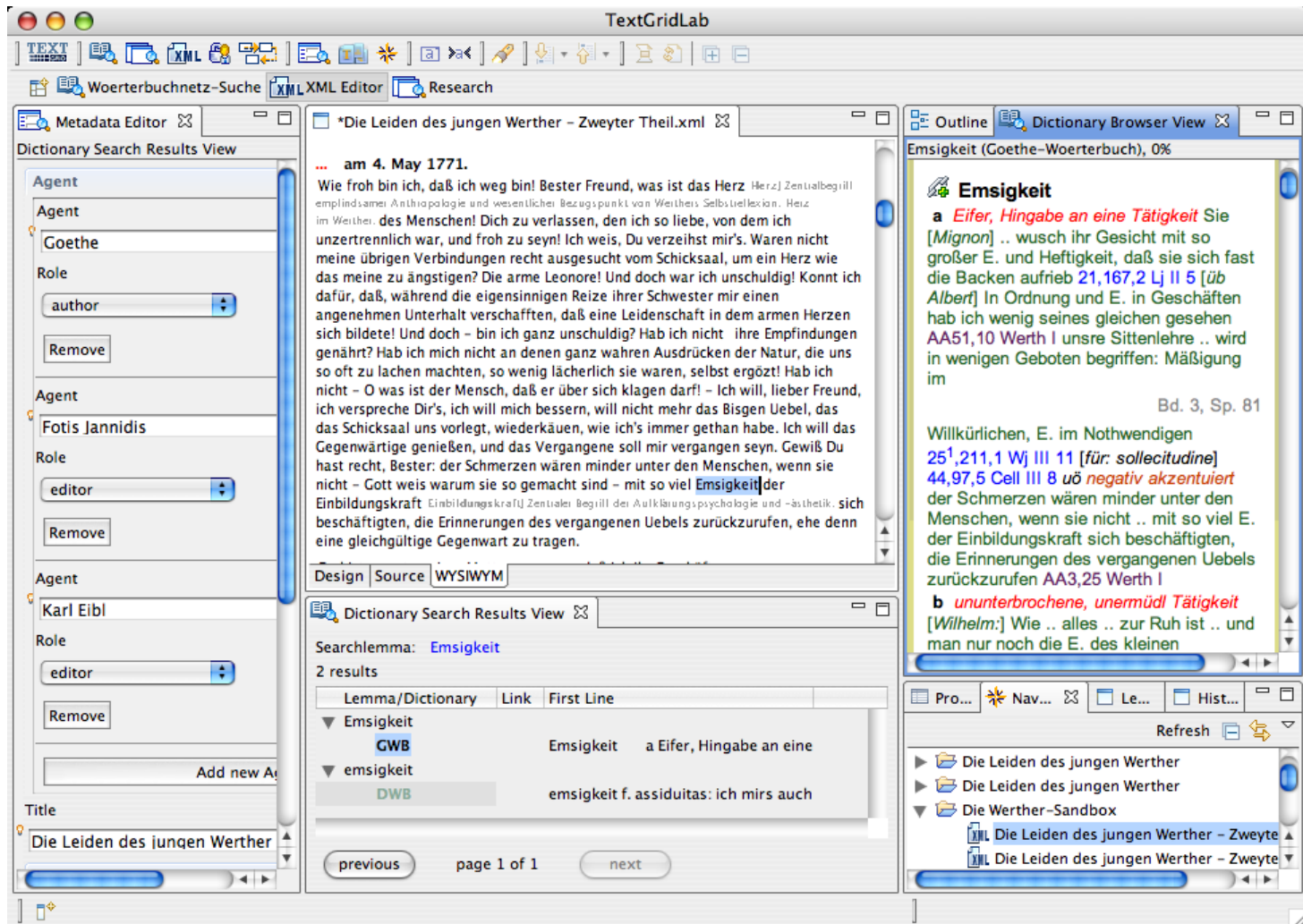
Virtual Research Environments

“The aim of a Virtual Research Environment (VRE) is to help researchers manage this complexity [of research] by providing an infrastructure specifically designed to support the activities carried out within research teams, on both small and large scales.”

JISC VRE Roadmap (2004)

- Originally planned to widen the reach of e-Science
- VRE's are not just about grand challenge science, but everyday research

TextGridLab



The screenshot displays the TextGridLab application window. The main interface is divided into several panes:

- Metadata Editor:** Shows a list of authors with their roles. The authors listed are Goethe (author), Fotis Jannidis (editor), and Karl Eibl (editor). Each entry has a 'Remove' button and an 'Add new Author' button at the bottom.
- Dictionary Search Results View:** Displays search results for the lemma 'Emsigkeit'. It shows 2 results:

Lemma/Dictionary	Link	First Line
GWB	Emsigkeit	a Eifer, Hingabe an eine
DWB	emsigkeit f.	assiduitas: ich mirs auch
- Main Text View:** Shows a snippet of text from 'Die Leiden des jungen Werther'. The text includes the date 'am 4. May 1771.' and a paragraph starting with 'Wie froh bin ich, daß ich weg bin! Bester Freund, was ist das Herz...'. The text is annotated with red and blue highlights.
- Dictionary Browser View:** Shows the full dictionary entry for 'Emsigkeit' from the Goethe-Wörterbuch. It includes the definition: 'a Eifer, Hingabe an eine Tätigkeit Sie [Mignon] .. wusch ihr Gesicht mit so großer E. und Heftigkeit, daß sie sich fast die Backen aufrieb 21,167,2 Lj II 5 [üb Alber] In Ordnung und E. in Geschäften hab ich wenig seines gleichen gesehen AA51,10 Werth I unsre Sittenlehre .. wird in wenigen Geboten begriffen: Mäßigung im'. It also mentions 'Bd. 3, Sp. 81' and provides further references like 'Willkürlichen, E. im Nothwendigen 25¹,211,1 Wj III 11 [für: sollicitudine] 44,97,5 Cell III 8 uö negativ akzentuiert der Schmerzen wären minder unter den Menschen, wenn sie nicht .. mit so viel E. der Einbildungskraft sich beschäftigten, die Erinnerungen des vergangenen Uebels zurückzurufen AA3,25 Werth I b ununterbrochene, unermüdt Tätigkeit [Wilhelm:] Wie .. alles .. zur Ruh ist .. und man nur noch die E. des kleinen'.
- Outline:** Shows a hierarchical view of the document structure, including 'Die Leiden des jungen Werther' and 'Die Werther-Sandbox'.

Microsoft Research Information Centre

YOUR 0901


[Contact Us](#) | [Feedback](#)

Research Information Centre

Welcome

 RIC Search

Home



Mike Chance
I'm researching into Cytokines
Interests
Microsoft SharePoint
[Edit Profile](#)

Projects Overview

Name	Role	Activities	Discussions
Cytokines		ARC_SEC2.doc Added by	No discussion started
Sharks		This is a document... Added by	No discussion started
Cancer		No activity performed	No discussion started
Projects		No activity performed	No discussion started

Feedback Alerts

Feedback Requests					
Name	Project	Requested By	Received On	Feedback	
ARC_c...proposal	Cytokines		6/25/2009 5:45:13 PM	Provide Feedback	<input type="button" value=""/>
ARTHRL_AMPAIGN	Cytokines		6/25/2009 5:16:59 PM	Provide Feedback	<input type="button" value=""/>
ARTLUDL_AMPAIGN	Cytokines		6/25/2009 5:16:59 PM	Provide Feedback	<input type="button" value=""/>

Feedback Received
There are no items to show in this view.

My Space

[References ...](#)

[Bookmarks ...](#)

[Document Library ...](#)

- [ARC__cancer_proposal.docx](#)
- [This is a document.doc](#)

Upcoming Events [Manage](#)

BBC News | Science & Environment | UK Edition

- [Lift off for Nasa's lunar probes](#)
- [Flood protection 'needs doubling'](#)
- [Firms team up for ISS supply ship](#)
- [UK 'must plan' for warmer future](#)
- [Mobiles boost Africa climate data](#)

14

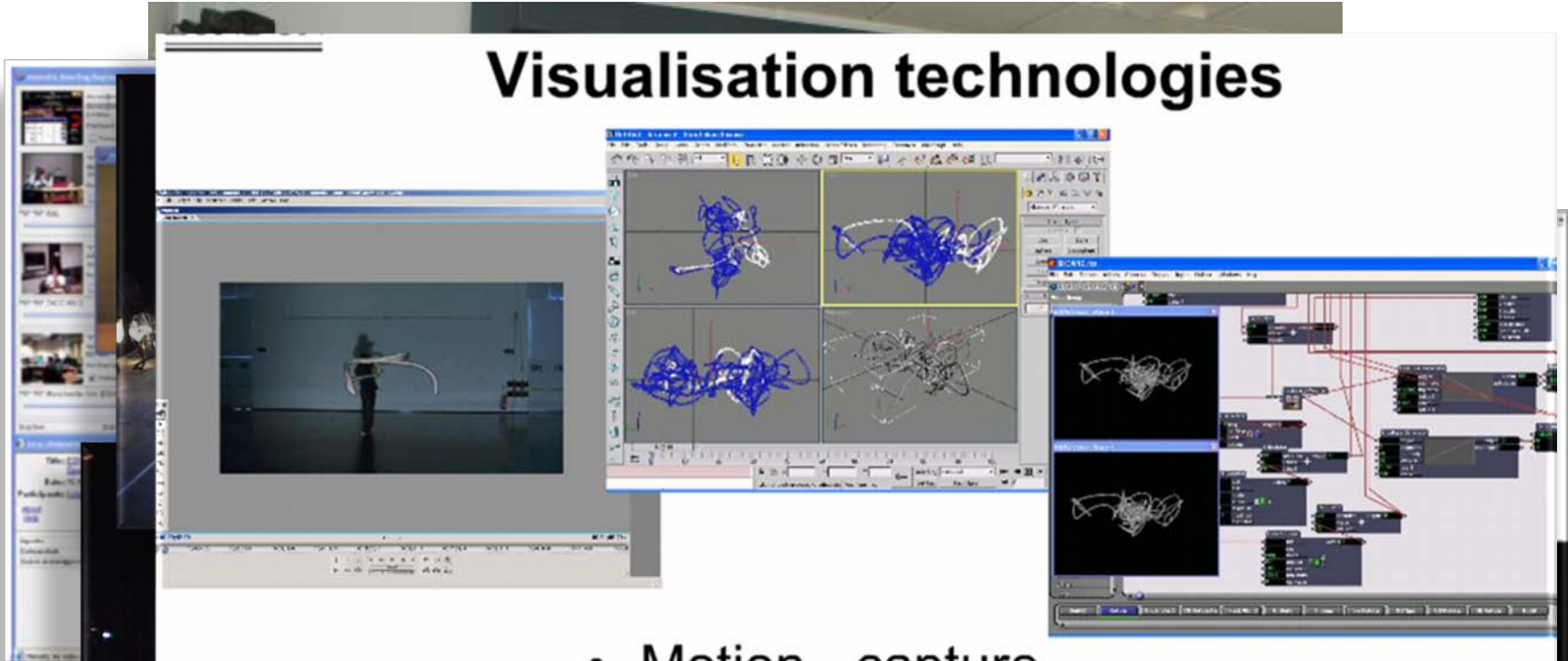
From DLs to e-Research: A Challenge to your Institutional Repository?

Matthias Razum, FIZ Karlsruhe

06.11.2009

From Meetings to Dancing


Visualisation technologies




- Motion - capture
- 3D animation
- Motion - tracking

http://laws1.reading.ac.uk - Project SILCH: 184 - Ces

SILCH: Cont
Fill of cess



Small fragment of possible amber ring SF3516




Small fragment of possible amber ring SF3516

http://laws1.reading.ac.uk - IADB Document Editor - Mozilla Firefox

Project SILCH: 124 - Timber Building 1

Document Metadata Discussion History [Print | Edit Document]

Early Roman Timber Building 1A



Timber Building 1

Occupying what became the angle between the main north-south street and the subsidiary east-west street our building consisted of two rooms. The larger, Room 1 ([Object 50043](#), [Plan](#), [Matrix 290](#)), measures c.10 by 7m, giving an internal surface area of 70 square metres, was bounded by shallow slots ([5393](#), [5304](#), [5375](#), [5357](#), [5356](#), [5346](#), [5396](#), [5371](#), [5323](#)) filled with yellow clay. This is assumed to be material associated with the infilling of the wall-structure, which, over time, replaced the voids created by the decayed horizontal sill-beams. At the centre of the room was a rectangular, tiled hearth ([1433](#)), built up of horizontally laid, broken tile fragments and associated with dark, charcoal-rich soils ([6038](#), [5919](#)). At least three separate phases of use of the hearth can be identified, each associated with a re-laying or re-patching of the floor. The associated floor surfaces of the room were composed

Design Principles for (Social) VREs

- Fit in, Don't Force Change
 - Act Local, think Global
 - Enable Users to Add Value
 - Design for Network Effects
 - Keep your Friends Close
 - Embed
-
- But: Researchers collaborate to compete!

De Roure, D. and Goble, C. "Software Design for Empowering Scientists," IEEE Software, vol. 26, no. 1, pp. 88-95, January/February 2009

MyExperiment



Home Us

Home » Tags » BLAST (Workflows only)

Tag Results

Workflows tagged with "BLAST" (24) [See results for all]

Note: some items may not be visible to you, due to viewing

Taverna 1 **BLAST using DDBJ service**

Original Uploader



Paul Fisher

Created: 03/10/07 @ 18:36:06 | Last updated: 13/11/07 @ 16:13:48

License: Creative Commons Attribution-No Derivs



Perform a sequence search through the DDBJ service

Rating: 4.5 / 5 (2 ratings) | Versions: 2 | Revisions: 1

Viewed: 282 times | Downloaded: 109 times

Tags (5):

bioinformatics | BLAST | ddbj | sequence | simulation

Taverna 1 **BLASTP with simplified results**

Original Uploader



Created: 03/10/07 @ 18:36:06 | Last updated: 13/11/07 @ 16:13:48

License: Creative Commons Attribution-No Derivs



Perform a blastp search based on the user input. The workflow does not function correctly as it is designed for use with NCBI blast

Version 2 (latest) (of 2)

View version: 2 (latest)

Version created on: 03/10/07 @ 18:36:06 by: Paul Fisher | Revision comments

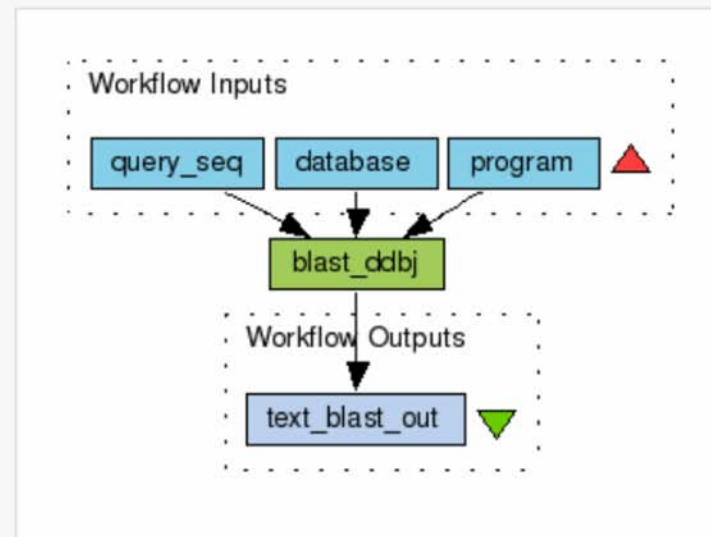
Last edited on: 13/11/07 @ 16:13:48 by: Paul Fisher

Title: **BLAST using DDBJ service**

Type: Taverna 1

Preview

(Click on the image to get the full size)



Download Scalable Diagram (SVG)

Social VREs, Collaboration, and Policies

- Researchers use informal and trusted sources of advice from colleagues, rather than institutional service teams, to help identify information sources and resources
- The use of social networking tools for scientific research purposes is far more limited than expected
- Data and information sharing activities are mainly driven by needs and benefits perceived as most important by life scientists rather than ‘top-down’ policies and strategies

Williams, R., Pryor, G. et al. "Patterns of information use and exchange: case studies of researchers in the life sciences", Report by the Research Information Network and the British Library, November 2009

Lessons Learned: Data

- Lots of new digital-born objects (tags, relations, annotations)
- Clash of Digital Library models:
Curated by experts (librarians) \leftrightarrow self-organizing systems
- New expertise is required to curate research data
- Librarians need to manage trust, while sharing control

Lessons Learned: VREs

- Standardization is difficult to achieve, even within a discipline
 - We need to avoid constraining researchers and their methodologies
 - Enable researchers to create their own solutions
- Talk to researchers: they understand their processes, data entities, and requirements best
 - Show them early prototypes to avoid misunderstandings
- Researchers are reluctant to change processes and tools that work for them
 - Try to re-use and integrate existing tools instead of re-implementing functionality

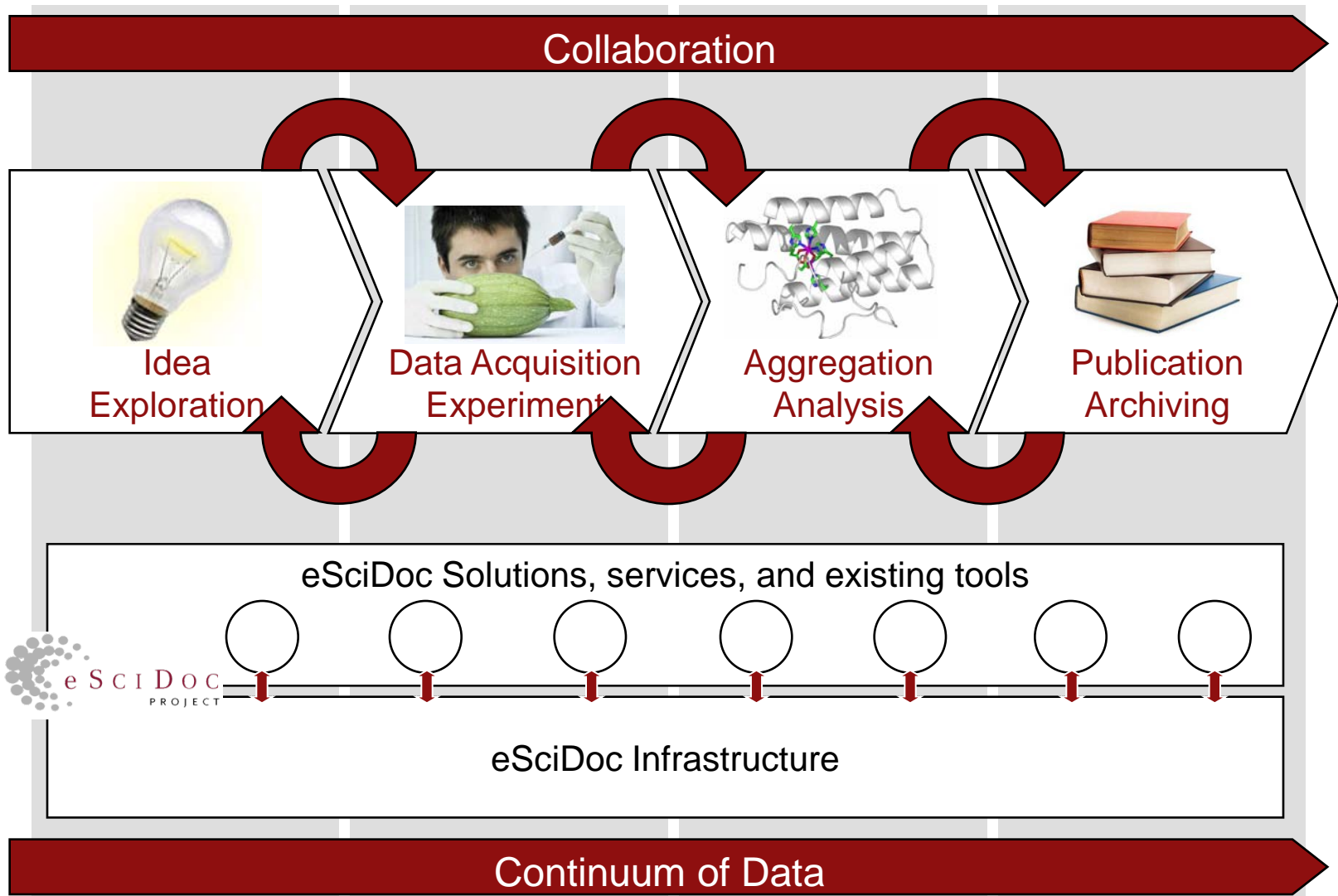
Agenda

- Institutional Repositories, Digital Libraries, and their challenges
- Data and Information Infrastructures in Europe
- Virtual Research Environments
- **eSciDoc**

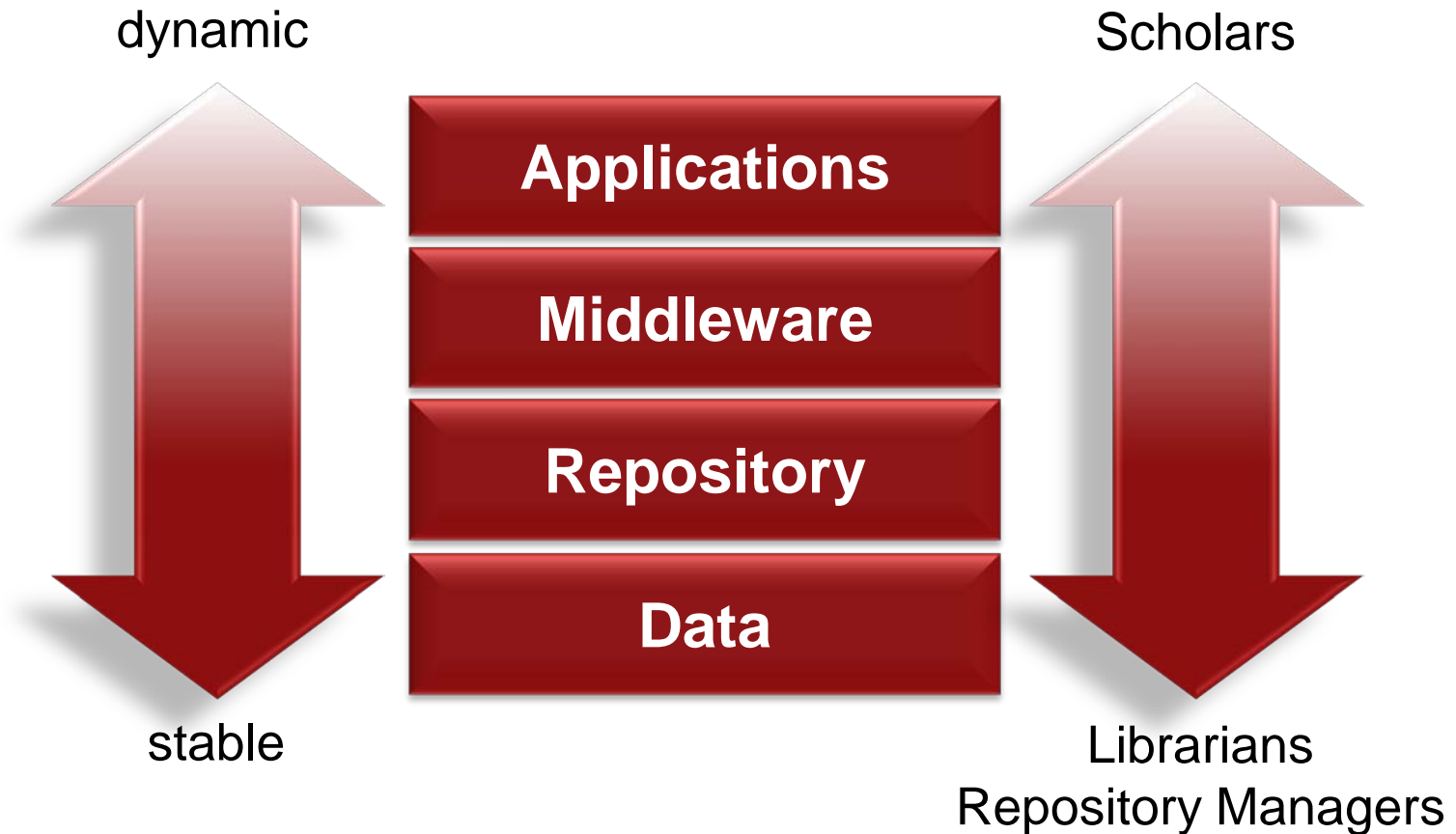
eSciDoc



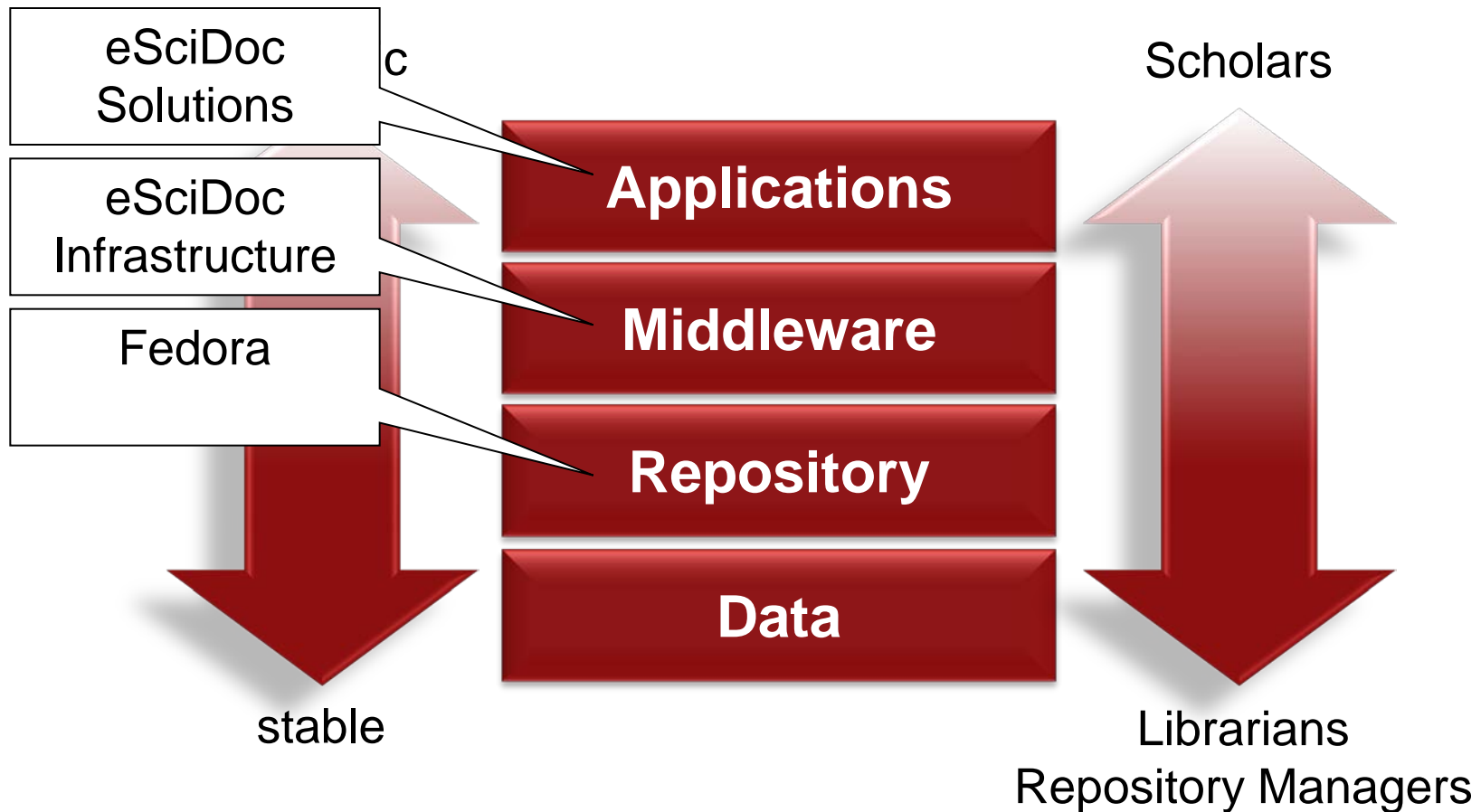
- eSciDoc is as a joint project of the Max Planck Society and FIZ Karlsruhe, funded by the Federal Ministry of Education and Research (BMBWF), aimed at building an e-Science platform for multi-disciplinary research organizations.
- eSciDoc
 - integrates research results and materials in an emerging e-research network,
 - provides effective and comprehensive access to data and information
 - supports collaboration and interdisciplinary research in future e-Science scenarios
 - increases the accountability of research
 - improve the visibility of research institutions and organizations



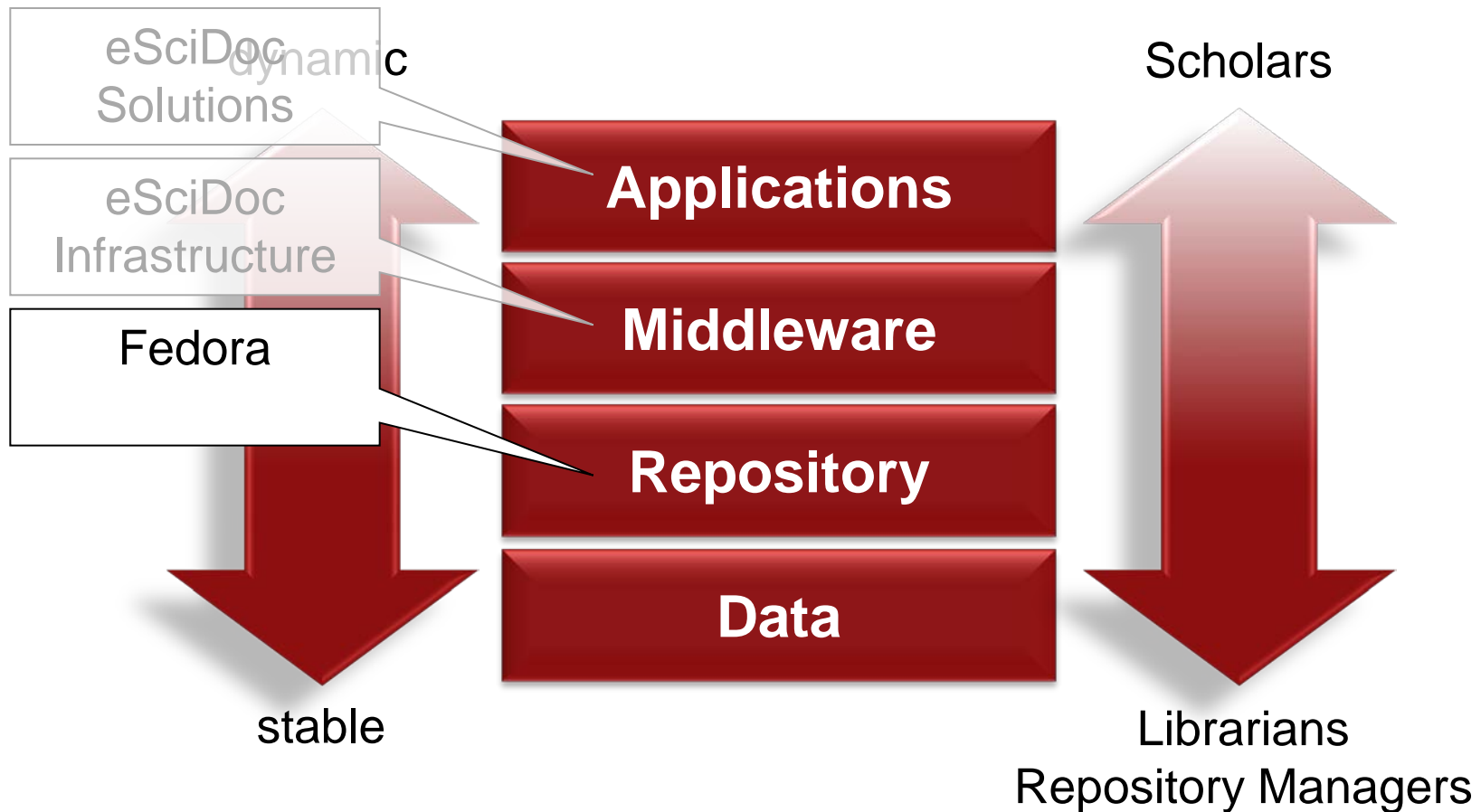
Hierarchy of Persistence



Hierarchy of Persistence



Hierarchy of Persistence



Fedora is an Acronym for

- **F**lexible
- **E**xtensible
- **D**igital
- **O**bject
- **R**epository
- **A**rchitecture



What means „Repository Architecture“?



Yokohama National University

簡易検索

検索

詳細検索

ホーム

一覧

- 学部・研究科等
- タイトル
- 著者
- 日付
- 高頻度閲覧文献
- アクセス統計
- リポジトリ説明サイト
- お問い合わせ

お知らせ

- 工学部歴史資料を登録
- 『エコノミア』第58巻を
- 富蔵コレクション(57冊
- 横浜国立大学で博士号

お知らせのバックナンバーは

お問い合わせ先: 横浜国立大学

検索

検索語を入力して、検索ボタ

学部・研究科等

1 教育人間科学部: 教育学研

Yokohama National University

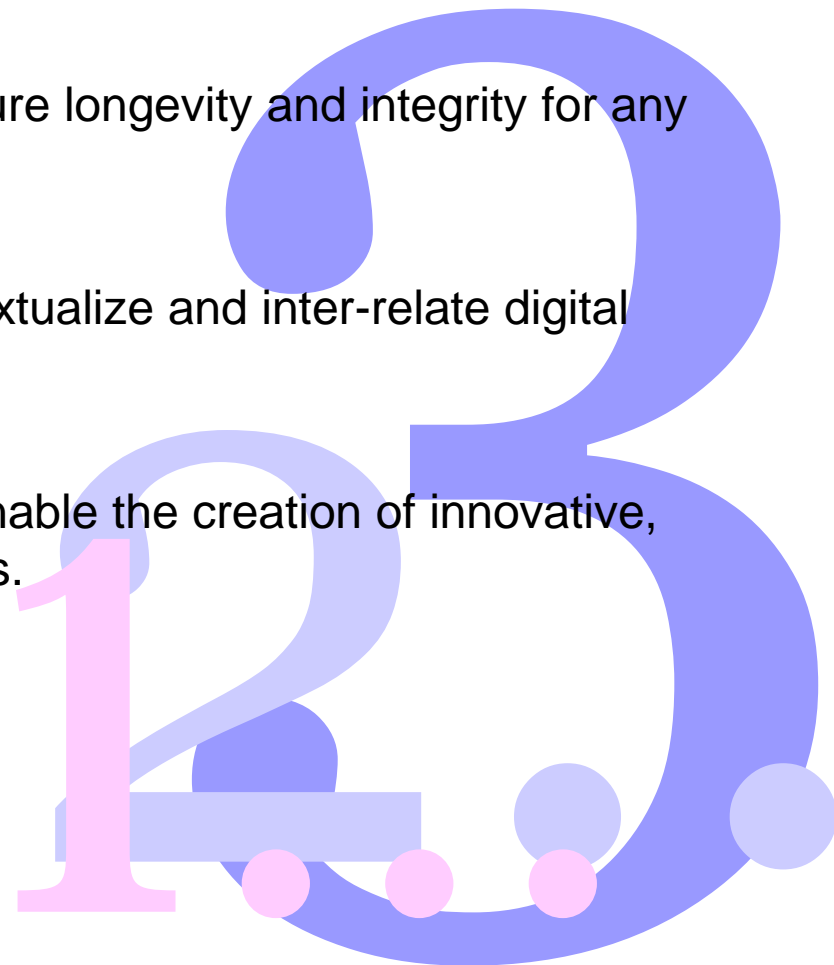
```
<?xml version="1.0" encoding="UTF-8"?>
<foxml:digitalObject VERSION="1.1" PID="escidoc:ex1"
xmlns:foxml="info:fedora/fedora-system:def/foxml#"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="info:fedora/fedora-system:def/foxml# http://www.fedora
<foxml:objectProperties>
<foxml:property NAME="info:fedora/fedora-system:def/model#state" VALUE="Ac
<foxml:property NAME="info:fedora/fedora-system:def/model#label" VALUE="Co
<foxml:property NAME="info:fedora/fedora-system:def/model#createdDate" VAL
<foxml:property NAME="info:fedora/fedora-system:def/view#lastModifiedDate"
</foxml:objectProperties>
<foxml:datastream ID="AUDIT" STATE="A" CONTROL_GROUP="X" VERSIONABLE="fals
<foxml:datastreamVersion ID="AUDIT.0" LABEL="Audit Trail for this object"
<foxml:xmlContent>
<audit:auditTrail xmlns:audit="info:fedora/fedora-system:def/audit#">
<audit:record ID="AUDREC1">
<audit:process type="Fedora API-M"/>
<audit:action>ingest</audit:action>
<audit:componentID></audit:componentID>
<audit:responsibility>fedoraAdmin</audit:responsibility>
<audit:date>2009-09-16T15:55:51.790Z</audit:date>
<audit:justification>eSciDoc object created</audit:justification>
</audit:record>
</audit:auditTrail>
</foxml:xmlContent>
</foxml:datastreamVersion>
```

DSpace

Fedora

The Fedora Triple Play

- Storage
 - Repository technologies to ensure longevity and integrity for any kind of digital content.
- Semantics
 - Semantic technologies to contextualize and inter-relate digital content from many sources.
- Services
 - Collaborative technologies to enable the creation of innovative, collaborative information spaces.



Fedora's Key Features

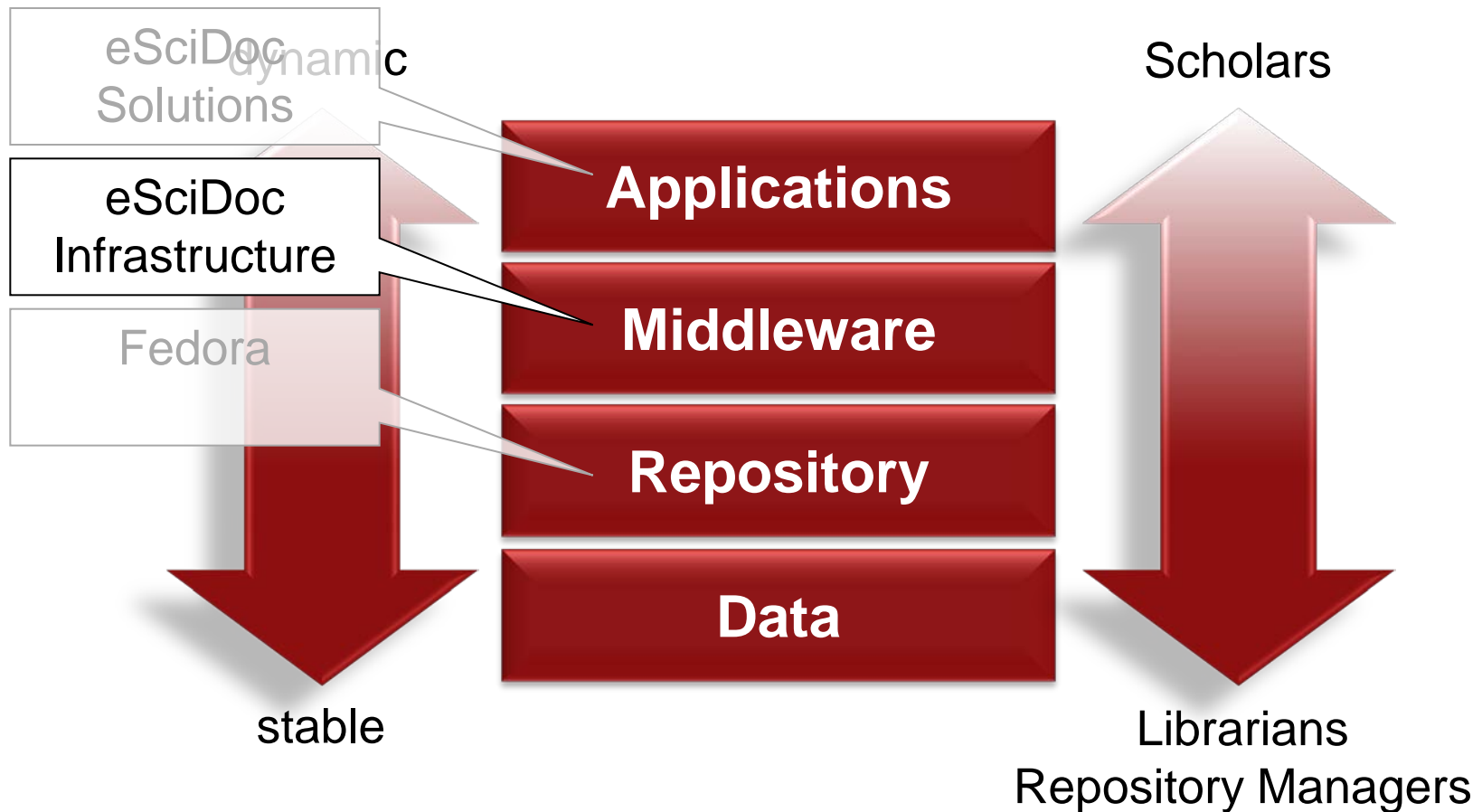
- Store whatever you want
- Easy access
- Express relationships
- Free open source software
- Enables permanence
- A supportive community
- Extensibility
- Maturity
- Scalability, both small and large
- An engine for innovation

Fedora, Duraspace, and eSciDoc



- Fedora Commons and Dspace Foundation are joining their organizations to pursue a common mission.
- The new organization called Duraspace will sustain and grow both repository systems: Dspace and Fedora.
- eSciDoc benefits from the strengthened Duraspace foundation and the increased development and innovation power.
- We are Fedora committers, thus influencing the further development of the Fedora software.

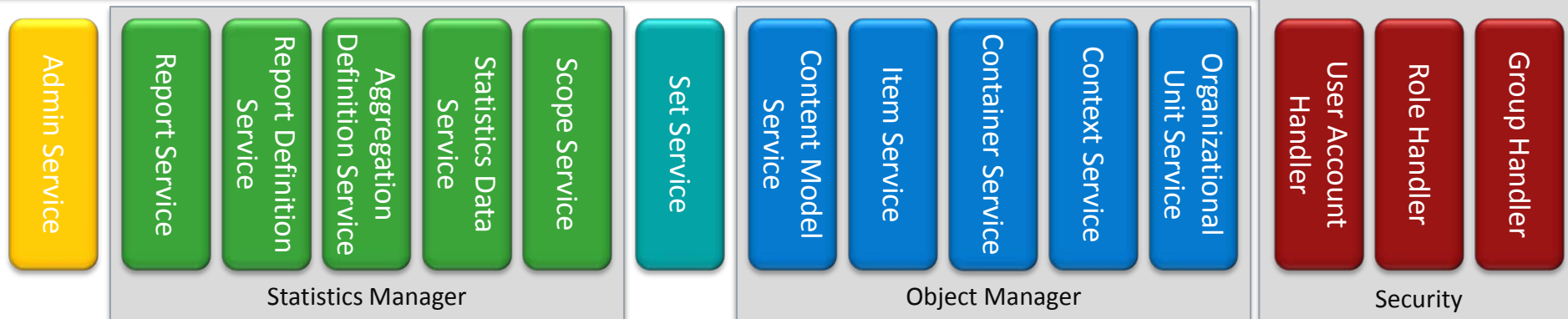
Hierarchy of Persistence



eSciDoc Services



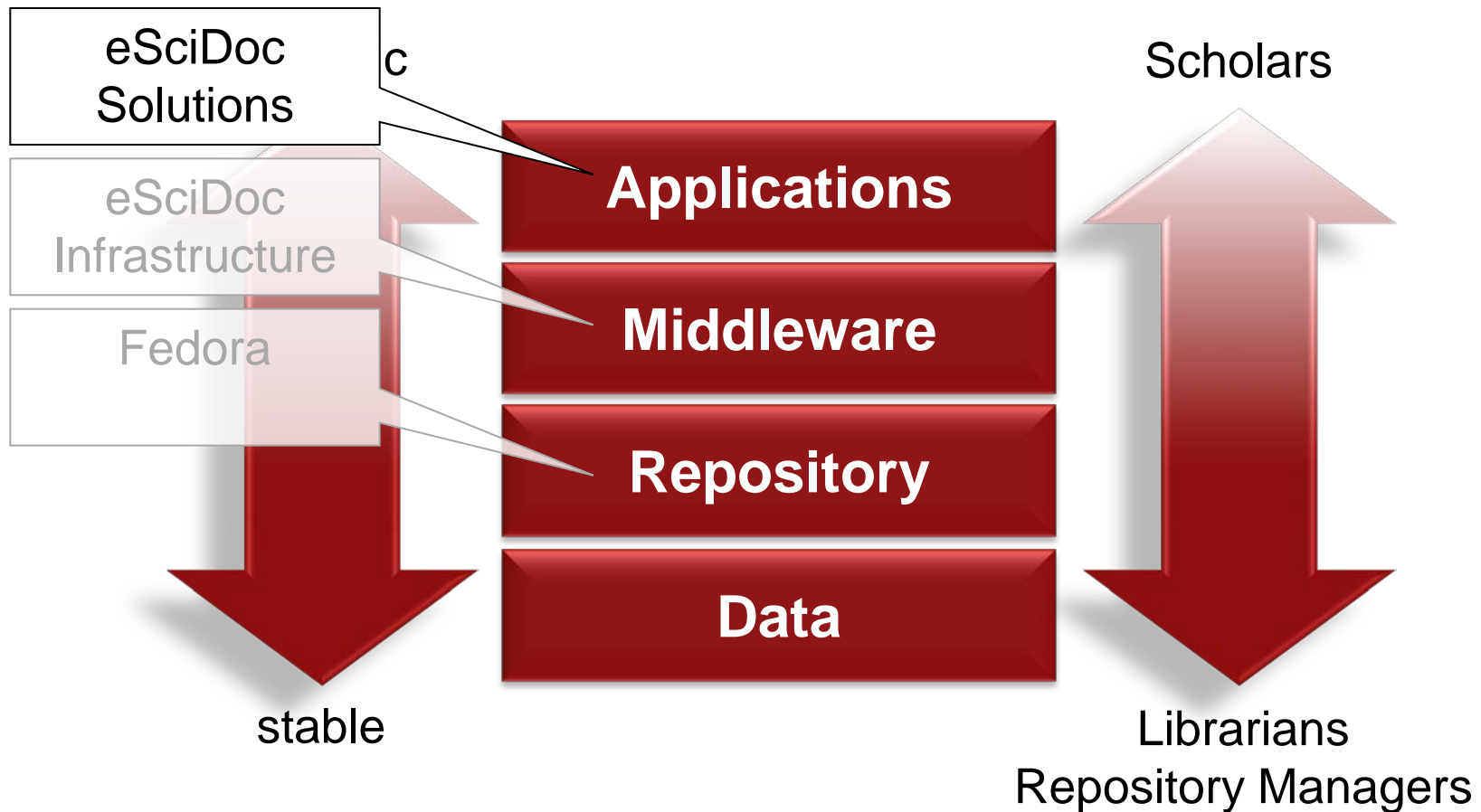
Policy Decision Point



Key Features of the eSciDoc Infrastructure

- Service-oriented architecture
- Stable Core Services
- Flexible content models
- Arbitrary metadata profiles
- Application-independent design
- Support for object relations and multiple ontologies
- Search (OpenSearch, SRW/SRU)
- Distributed Authentication/Authorization (Shibboleth)

Hierarchy of Persistence



eSciDoc Solutions

Christine Hasemann Logout Contact us Policy&Disclaimer Help English

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

Sorted by diamond-elements diamond-elements diamond-elements Sort Add all items of this page Add

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

The screenshot displays a search results interface for diamonds. At the top, there is a navigation bar with the user's name 'Christine Hasemann', a 'Logout' button, and links for 'Contact us', 'Policy&Disclaimer', and 'Help'. A language dropdown menu is set to 'English'. Below this, the search results are shown as a grid of 12 items (2 rows by 6 columns). Each item consists of a diamond image and a small icon for search and grid view. A sorting dropdown menu is open, showing a list of sorting criteria: 'diamond-elements', 'shape', 'color', 'mecharism', 'condition', 'temperature', 'brief-details', 'full-details', 'date', 'identifier', 'complete-name', 'given-name', 'family-name', 'identifier', 'organization-name', 'address', and 'identifier'. The current sort order is 'diamond-elements'. The page also includes pagination controls showing '1' of '2' pages and a 'Go to page' field.

Thank you!

Questions?

Matthias Razum
matthias.razum@fiz-karlsruhe.de

<http://www.escidoc.org/>