



Services for the eHumanities

Heike Neuroth & Laurent Romary

Max Planck Digital Library

3 - 5 February 2009 in Bielefeld, Germany

9th International Bielefeld Conference

ToC

- **Introduction**
- **Research Infrastructure (RI)**
- **Services for the eHumanities**
- **Final Words**

eScience

- Services for Science nowadays = eScience
- eScience = enhanced Sciences
- John Taylor (UK): „*eScience is about global collaboration in key areas of science and the next generation of infrastructure that will enable it.*”
 - virtual **integration** of a wide variety of services that offer various types of resources and **resource sharing** (compute power, storage, data, tools, knowledge etc)
 - **interoperability** at all levels – structural and semantic
 - enabling new forms of **virtual and interdisciplinary collaboration**
 - all based on **persistent infrastructures**

Examples of eScience Services

- Particle Physics – Large Hadron Collider
 - The largest machine in the world...
 - The fastest racetrack on the planet...
 - The biggest and most sophisticated detectors ever built...
 - The most powerful supercomputer system in the world...
(<http://public.web.cern.ch/public/en/LHC/Facts-en.html>)
 - funded by and built in collaboration with over 10,000 scientists and engineers from over 100 countries as well as hundreds of universities and laboratories
 - The total cost of the project is expected to be **€3.2 – 6.4 billion**
(http://en.wikipedia.org/wiki/Large_Hadron_Collider)



Examples of eScience Services

- Astronomy – European Southern Observatory (ESO)



Spiral Galaxy NGC 1232

builds and operates a suite of the world's most advanced ground-based astronomical telescopes

- Atacama Large Millimeter/submillimeter Array (ALMA):
The telescope is expected to revolutionize modern astronomy by providing an insight on star formation in the early universe
- With a cost in excess of **1 billion US dollars**, it is the most ambitious ground-based telescope currently under construction
(http://en.wikipedia.org/wiki/Atacama_Large_Millimeter_Array)



ALMA antenna

New Science Paradigm

- Thousand years ago – **Experimental Science**
 - Description of natural phenomena
 - Last few hundred years – **Theoretical Science**
 - Newton's Laws, Maxwell's Equations...
 - Last few decades – **Computational Science**
 - Simulation of complex phenomena
 - Today – **eScience or Data-centric Science**
 - Unify theory, experiment, and simulation
 - Using data exploration and data mining
 - Data captured by instruments
 - Data generated by simulations
 - Data generated by sensor networks
- Scientist analyzes databases/files
- Commercial and academic applications

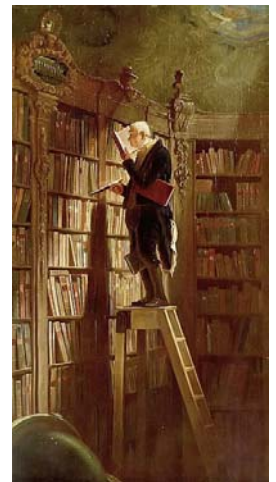


Summary given by Tony Hey at Grid AHM 2007:

http://www.d-grid.de/fileadmin/user_upload/images/D_Grid_AHM_0907/presentations/D-Grid-AHM_1eroeffnung_01_Hey_Keynote_Paper.pdf

eHumanities in General

- What does this mean for eHumanities?
 - fragmentation in general (subject, format, data ..) is extreme
 - landscape will always be heavily distributed
 - large variation of resource types and structures
 - semantic relations between resources are complex
 - semantics within resources are complex
 - problems in general not compute intensive (lack of algorithms)
 - but research is increasingly data driven as well
 - *** funds are in general very small ***
 - often no expert ICT support in the institution
 - AND: awareness of researchers about new methods is poor
 - resource sharing not always wanted and possible



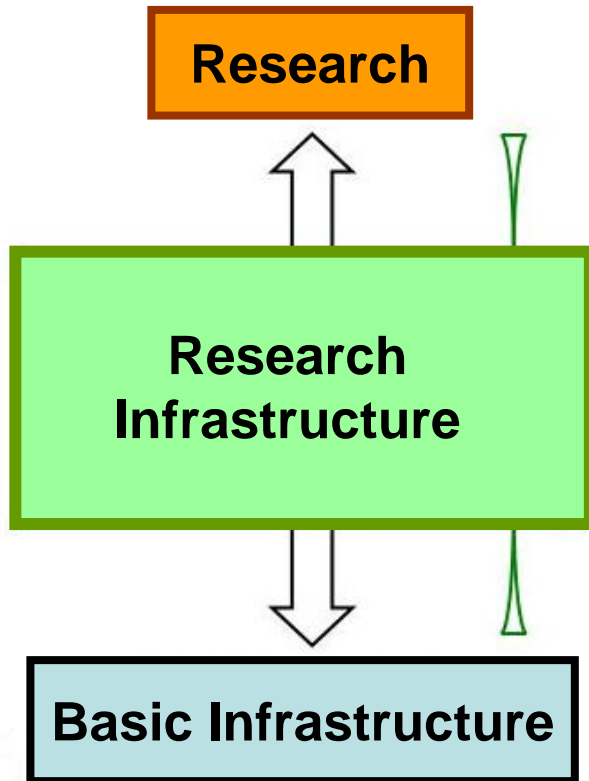
ToC

- Introduction
- **Research Infrastructure (RI)**
- Services for the eHumanities
- Final Words

Definition of Research Infrastructure (RI)

- Research infrastructure is defined as “the physical, informational and human resources essential for researchers to conduct high-quality research.” It includes:
 - tools, equipment, instrumentation, computer platforms and facilities;
 - software and information resources, including enabling (i.e., used for more than one project) computer systems, databases, data analysis and interpretation systems, and communication networks;
 - technical support (human or automated) and services needed to operate infrastructure and keep it working effectively; and
 - special environments and installations (e.g., buildings and research space) necessary to effectively create, deploy, access and use research tools.

Research Infrastructures: Services



R	<p>e.g. new research questions based on new methods like text mining, mass-analyses</p> <p>...</p>
RI	<p>Services, tools standards, training</p> <p>...</p>
BI	<p>e.g. Grid infrastructure (compute, storage), repository infrastructure</p> <p>...</p>

Examples of RI in eHumanities

- **Europe**

ESFRI – European Strategy Forum on Research Infrastructures (<http://cordis.europa.eu/esfri/>)

- **CLARIN**: offers scholars the tools to allow computer-aided language processing (<http://www.clarin.eu/>)
- **Dariah**: Digital Research Infrastructure for the Arts and Humanities (<http://www.dariah.eu/>)

- **Germany**

- **eAqua**: Extraction of structured knowledge from ancient sources (<http://www.eaqua.net/>)
- **eSciDoc**: eResearch Environment (<http://www.escidoc.org/>)
- **TextGrid**: Virtual Research Environment for Humanities (<http://www.textgrid.de/>)
- ...

ToC

- Introduction
- Research Infrastructure (RI)
- **Services for the eHumanities**
- Final Words

Examples eSciDoc

- Based on RI **eSciDoc** - The open source eResearch Environment
 - **ViRR**: Virtueller Raum Reichsrecht (MPI for European History of Law): provides a digital compilation and working environment for various artefacts of the period of the German Holy Empire. The collection will be indexed, edited, and enlarged cooperatively within the discipline http://192.129.1.123:8080/virr_presentation/
 - **FACES** (MPI for Human Development): set of images of naturalistic faces of 171 young, middle-aged, and older women and men displaying each of six facial expressions: neutrality, sadness, disgust, fear, anger, and happiness <http://faces.mpdl.mpg.de/faces/>

ViRR service

- Functionalities
 - browsing within the works of ViRR
 - search within all bibliographic and descriptive metadata
 - online editing interface for privileged users to gather structural metadata about the digitized works
 - compilation and export of private subsets of pictures (as jpg or within one pdf), based on special research questions
 - export of structural metadata in METS
 - persistent storage of all data ...

- Using eSciDoc as infrastructure – can be used for other research groups as well, e.g. MPI for Art History (Italy)

Service for working with text

ViRR service

- Close Frames
- Startseite
 - Log_In
 - Browse_&_Display
 - Inhaltsübersicht
 - Mehrbandwerk
 - Bandansicht**
 - Detailansicht
 - Editor
 - Inhaltsübersicht_Edit
 - Bibl_MD_ändern
 - Bibl_MD_ändern_mel
 - Paginierung_&_Struktur
 - Strukturelement_anle
 - Seiten_hinzufügen
 - Weiteres_Strukturelei
 - Weitere_Seiten_hinzu
 - Metadaten_eingeben
 - Details_zu_Metada
 - Bestätigung
 - Hilfsseiten_für_Fenster_
 - Seitenliste_Paginieru
 - Struktur_neu_anleger
 - Strukturelement_anle
 - Baum_Seiten_hinzufü

INHALTSÜBER



Oertel, Christian Gottfried: Vollständiges corpus gravaminum evangelicorum. Band 7, 1775.

► Struktur Ausblenden



[Zu den digitalen Faksimile](#)
[Zu den digitalen Faksimile im DFG Viewer](#)

Multivolume | **Oertel, Christian Gottfried:
Vollständiges corpus gravaminum eva
Heubauer, 1771-1775.**

Zeige

Vorbericht

Num. CIII.

Num CIV. & CV.

Die Stadt Augsburgische und Nördlingische Bilder-Sac

▫ [Beylagen zum Augspurgischen Gravamine](#)

▫ [Beylagen zum Nördlingischen Gravamine](#)

Num. CVI.

**Betreffend das zu Langenschwarz und in denen darzu
Herrschaft von Langenschwarz beeinträchtigte Evan**

▫ [Beylagen](#)

Num. CVII.

**Die aus der Nassau-Hadamarischen Succesßion und so
Durchlaucht zu Nassau-Siegen Evangelisch-Reformirte**

Signatur Dt 9 Bg 20 Q [7]
Autor Oertel, Christian Gottfried
Title Vollständiges corpus gravaminum evangelicorum
Band 7
Erscheinungsjahr 1775
Anmerkung
Sprache
Katalognummer 66237
Schlagwörter Corpus Evangelicorum / Gravamen
Genre Bandsatz
Anzahl Scans 2345
Barcode 1111111111111

Bandansicht

FACES service

- Functionalities
 - picture search based on the available attributes (currently ID, age group, gender, facial expression and picture group)
 - detailed view of single pictures (and navigation within the pictures) and their attributes
 - creation and export of public and private subsets for the analysis of special research questions
 - management of an unlimited quantity of attributes per picture (e.g. adding new attributes based on the output of rating studies)
 - user and rights management to realize different access rights
 - possibility to create a new, personalized album

Service for working with images

FACES service

Close Frames

- Home
- Login
- About
- Registration
- Help
- Browse_not_logged_in
- Browse_logged_in
- Actions_selecting_a**
- Detailed_View
- Detailed_View_Pictur
- Browse_Admin
- Statistic_Visits
- Search
- Search_Results
- Search_Results_0_Hi
- All_pictures_of_one_p
- Create_Album
- Edit_Album
- Confirm_Publish_or_L
- Confirm_Withdrawn
- My_Albums
- My_Albums_Search
- List_View_Private_Alt

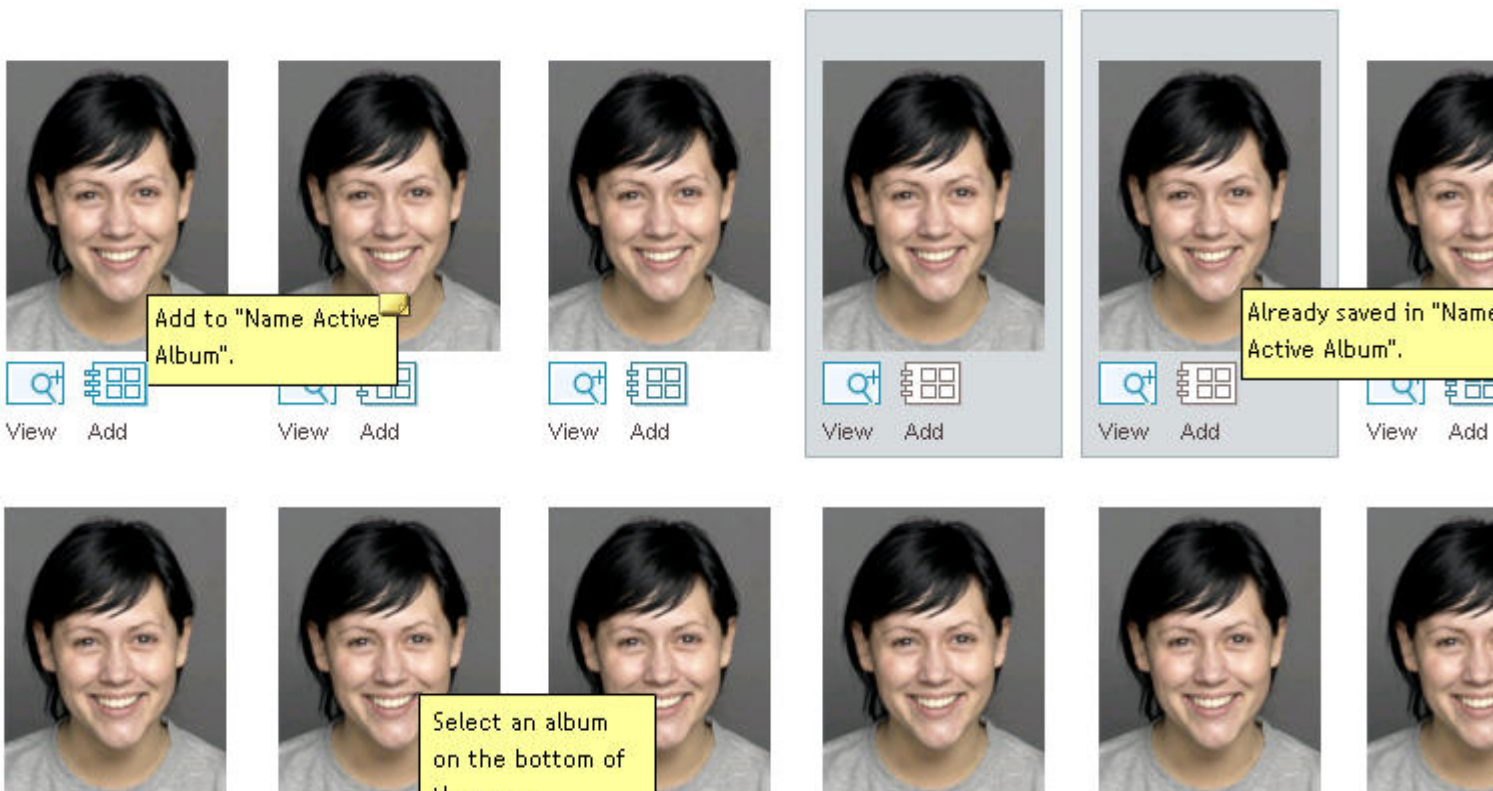
2 Hits per Page

Previous | 123 | 123 | 123 | 123 | 123 | 123 | 123 | Next

Go to Page

Add all Pictures

Add



Add to "Name Active Album".

Already saved in "Name Active Album".

Select an album on the bottom of the page.

Examples TextGrid

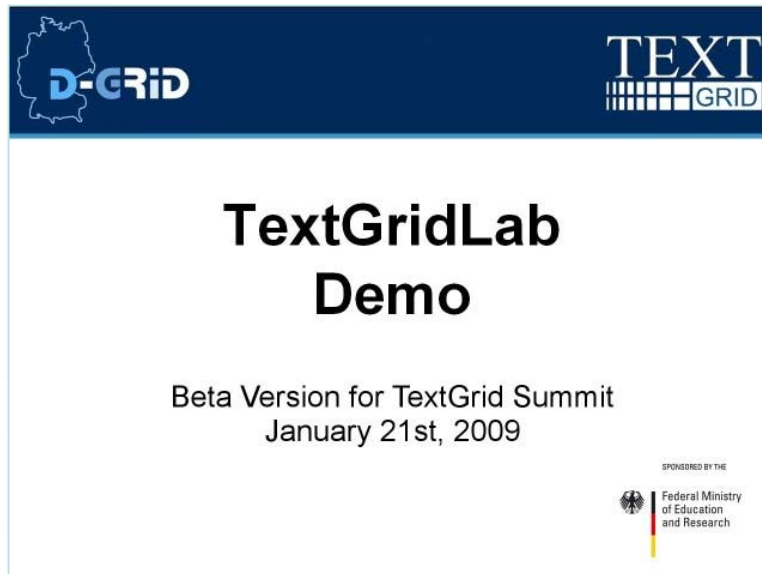
- *Data Grid* - Virtual archive(s) for research data curation
TextGrid Rep (Repository)

- *Service Grid* - Toolbox for collaborative work
TextGrid Lab (Laboratory)
 - Interactive tools (Eclipse-based)
 - XML-Editor
 - Research
 - Graphical Editor
 - Interfaces to dictionaries and reference material
 - Streaming tools (Web Service-based)
 - Collation
 - Lemmatise
 - Sort, tokenise, etc



Demo TextGrid Lab

- Beta version available for download
<http://www.textgrid.de/beta.html>
- Video demo – Highlights



[play swf](#)
[play wmv](#)

ToC

- Introduction
- Research Infrastructure (RI)
- Services for the eHumanities
- **Final Words**

eHumanities

- scholars are not that IT affine. Closer cooperation between scholars & IT experts is needed
- maybe not a data deluge (yet) but much more heterogeneous data, which need special services
- difficult to encourage cooperation/collaboration across disciplines
change of culture/tradition?
- sustainability? Who is paying after project end? For infrastructure (RI) and services?
- roles, responsibilities (tool developer, IT experts e.g. for Grid, scholars, data curator, data host ...) still unclear

DFG Announcement

- DFG announcement “Virtual Research Environment”

http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/ausschreibung_virtuelle_forschungsumgebungen.pdf

- Die Projekte müssen im Antrag die **konkrete Zusammenarbeit** zwischen **Informationseinrichtungen** (Bibliotheken, Rechenzentren, Medienzentren u.ä.) und **Forschern** (...) skizzieren.

Dabei sollen die jeweiligen Vertreter der Infrastruktur sowie der Wissenschaft **gleichberechtigt** an der Erarbeitung der Konzepte beteiligt sein.

- Close cooperation between service infrastructure organizations (like libraries, computer centers) **AND** researchers, scientists - each one with equal rights

Services

- embedded in research practice and needs
- must provide evidence of value for research
- services needed with a wider scope to increase the acceptance within humanities (balance between specific and more generic services)
- collaborations across domains and sectors
- need for training, education
- international in scope
- standards: scholars might be encouraged to develop standards like TEI

Thank you!
Questions, comments?

Acknowledgment:
Malte Dreyer, Peter Wittenburg

3 - 5 February 2009 in Bielefeld, Germany

9th International Bielefeld Conference