

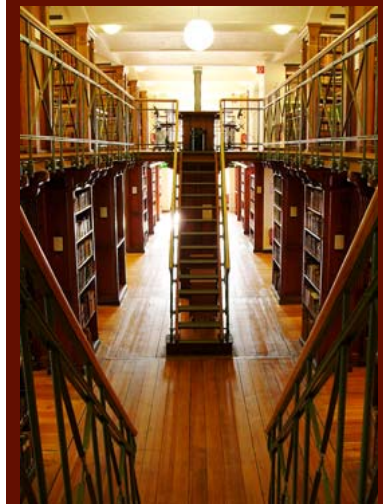
# Symposium FH Köln

29.01.2009

Von den Preußischen Instruktionen zu E-Science:  
80 Jahre Ausbildung und Studium für  
Bibliothekare und andere  
Informationsspezialisten in Köln

## eScience und Bibliotheken

Heike Neuroth  
MPDL Berlin / SUB Göttingen



# ToC

- Einführung
  - Wissensmanagement
- eScience
- Bibliotheken
- Aktivitäten
  - DFG Ausschreibungen
  - Allianz “Digitale Information”
- Ausblick



## **Aufgabe von Bibliotheken**

hier mit Fokus auf wiss. Bibliotheken!

- Management von Wissen = Wissensmanagement
- Unterstützung von Forschung und Lehre
- Unterstützung der Wissenschaftler in ihrem Prozess der Wissenserzeugung und -verarbeitung inklusive Publikation und Zugang

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick

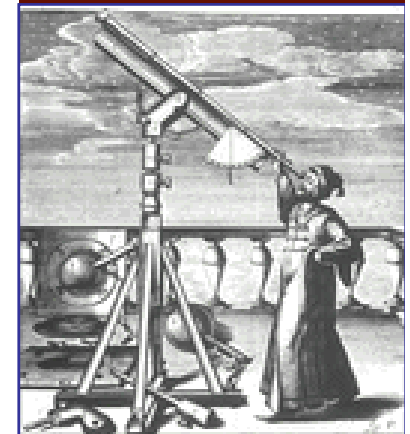


## Paradigmenwechsel in der Wissenschaft

- Vor Tausend Jahren - **Experimental Science**
  - Beobachtung/Beschreibung natürlicher Phänomene
- In den letzten Jahrhunderten - **Theoretical Science**
  - Newton's Gesetze, Maxwell's Gleichung ...
- In den letzten Jahrzehnten - **Computational Science**
  - Simulationen komplexer Phänomene
- Heute - **eScience or Data-centric Science**
  - Zusammenführen von Theorie, Experiment und Simulation
  - "data exploration", "data mining", ("text mining")

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick

[basierend auf Tony Hey Keynote, AHM im Sept. 07: [http://www.d-grid.de/fileadmin/user\\_upload/images/D\\_Grid\\_AHM\\_0907/presentations/D-Grid-AHM\\_1eroeffnung\\_01\\_Hey\\_Keynote\\_Paper.pdf](http://www.d-grid.de/fileadmin/user_upload/images/D_Grid_AHM_0907/presentations/D-Grid-AHM_1eroeffnung_01_Hey_Keynote_Paper.pdf)]



## Wissenschafts-Prozess

- Informations-Erzeugung  
Recording, Capturing
- Informations-Management  
Organizing, Storing, Preserving
- Informations-Auffindung  
Discovery, Mining, Browsing
- Informations-Verarbeitung  
Curation, Conversion, higher-order Processing
- Informations-Distribuiierung  
Communication, Publishing

[aus: Agenda AHD 2020 für eHumanities in Deutschland, Hinrichs, Lossau,  
Neuroth, Romary, Wittenburg]

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



# Wissensmanagement

- **Knowledge Management (KM)** comprises a range of practices used by organisations to
  - identify, [http://en.wikipedia.org/wiki/Knowledge\\_management](http://en.wikipedia.org/wiki/Knowledge_management)
  - create,
  - represent and distribute knowledge
- **Quellen für Wissen**
  - Bilder, Photographien
  - Manuskripte, Dokumente
  - Tonaufnahmen (gesprochen, gesungen)
  - Videos (Interviews, Folklore ...)
  - Artefakte
  - Forschungsdaten (inklusive Meßdaten, Rohdaten ..)
  - ...

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



# Wissensmanagement

- **Knowledge Management (KM)** comprises a range of practices used by organisations to

• identify [http://en.wikipedia.org/wiki/Knowledge\\_management](http://en.wikipedia.org/wiki/Knowledge_management)

**d.h. heterogene Quellen,  
unterschiedliche Standards,  
unterschiedliche technische Umgebungen,  
und unterschiedlich tief erschlossen**

...

- Videos (Interviews, Folklore ...)
- Artefakte
- Forschungsdaten (inklusive Meßdaten, Rohdaten ..)
- ...

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



# Wissensmanagement

- **Vision:**

Semantische Vernetzung von Wissen und Inhalten/Daten mit Hilfe geeigneter Technologien und Diensten/Tools

- **Schlüsseltechnologien:**

- Services, Dienste, Tools (disziplinär, interdisziplinär)
- Metadaten und Ontologien (fachspezifisch, fachübergreifend)
- Informationsextraktion, -annotation etc.
- Wissensvernetzung
- Semantische Suche
- ...

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



- **Vision:**

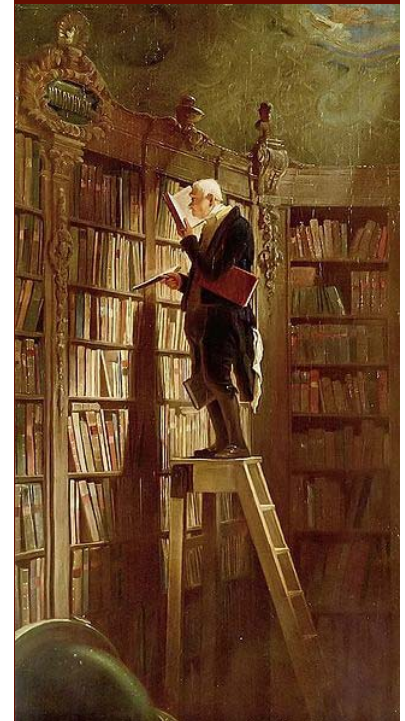
**Suche nach und *resource sharing* von:  
Daten/Inhalten, Wissen,  
Software, Diensten/Tools,  
Daten- und Wissensquellen**

...

**= eScience**

- Wissensvernetzung
- Semantische Suche
- ...

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



John Taylor (UK):

*„eScience is about global collaboration in key areas of science and the next generation of infrastructure that will enable it.“*

- **Virtuelle Integration** verschiedener Dienste, Services und Tools
- **Zusammenbringen** unterschiedlichster Ressourcen wie (Forschungs)Daten, Digitalisate, Inhalte ...
- **Resource Sharing**: Rechenleistung, Speicherplatz, Dienste/Services/Tools, Wissen ...
- **Interoperabilität**: auf allen Ebenen (strukturell, technisch, semantisch ...)
- **Kollaboration und Kooperation**: neue Formen, virtuell und interdisziplinär
- basierend auf persistenten und nachhaltigen **Infrastrukturen**

Einführung

eScience

Bibliotheken

Aktivitäten

Ausblick



- Services für die Wissenschaft und Forschung (und Lehre) heutzutage = eScience
  - "enhanced Science" (zu deutsch etwa erweiterte Wissenschaft)
  - e-Science zielt auf eine Forschung, die auf der Basis einer umfassend **digitalen Infrastruktur** vollzogen wird. Diese Infrastruktur **integriert** dabei alle **relevanten Ressourcen** für einen **Forschungsbereich** und stellt zugleich **Werkzeuge** zu ihrer Verarbeitung bereit. In **rechenintensiven** Forschungsszenarien organisiert sie die Verteilung der Rechenkapazitäten. Bei der Erarbeitung von Forschungsergebnissen unterstützt sie **kollaborative Prozesse** und bietet **eine Plattform** für die Publikation der Resultate - die **dadurch selbst wieder** zu Ressourcen für die Weiterforschung werden.

Einführung

eScience

Bibliotheken

Aktivitäten

Ausblick



<http://de.wikipedia.org/wiki/E-Science>

# eScience ... aber:

- e-Science zielt auf eine Forschung, die auf der Basis einer umfassend **digitalen Infrastruktur** vollzogen wird. Diese Infrastruktur **integriert** dabei alle **relevanten Ressourcen** für **einen Forschungsbereich** und stellt zugleich **Werkzeuge** zu ihrer Verarbeitung bereit. In **rechenintensiven** Forschungsszenarien organisiert sie die Verteilung der Rechenkapazitäten. Bei der Erarbeitung von Forschungsergebnissen unterstützt sie **kollaborative Prozesse** und bietet **eine Plattform** für die Publikation der Resultate - die **dadurch selbst wieder** zu Ressourcen für die Weiterforschung werden.

- digitale Infrastruktur: ***Forschungsinfrastruktur (RI)***
- relevante Ressourcen: ***Daten, Inhalte, Dienste ...***
- einen Forschungsbereich: ***nicht nur!***
- rechenintensiv: ***nicht nur!***
- kollaborative Prozesse: ***auch interdisziplinär!***
- eine Plattform: ***nun ja ... stimmt das?***
- dadurch selbst wieder: ***sehr wichtig!***

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick

<http://de.wikipedia.org/wiki/E-Science>



# Beispiele für eScience

- 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...
- 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://public.web.cern.ch/public/en/LHC/Facts-en.html>)

([http://en.wikipedia.org/wiki/Large\\_Hadron\\_Collider](http://en.wikipedia.org/wiki/Large_Hadron_Collider))

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



# Beispiele für eScience

- **Astronomy – European Southern Observatory (ESO)**

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](http://en.wikipedia.org/wiki/Atacama_Large_Millimeter_Array))

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



# Beispiele für eScience

- **TextGrid – a Community Grid for the Humanities**

community grid for the collaborative editing, annotation, analysis and publication of specialist texts

- im Rahmen der deutschen Grid Initiative  
([www.d-grid.de](http://www.d-grid.de))
- bisher das einzige Infrastrukturprojekt in den Geisteswissenschaften in Deutschland, weltweit das einzige Projekt mit Grid-Technologie in GW
- in Astronomie auch AstroGrid, in Teilchenphysik auch HepGrid ...
- Kosten insgesamt bisher ca. **2.5 Mio Euro** inklusive BMBF Sonderinvestitionen für Hardware (u.a. 250 TB Speicher für Langzeitarchivierung)
- Kosten D-Grid Initiative bisher: ca. 120 Mio Euro, über 100 Partner aus Wissenschaft, Forschung und Industrie

Einführung

eScience

Bibliotheken

Aktivitäten

Ausblick



# Beispiele für eScience

Tools, Dienste in TextGrid:

- **TextGrid Rep**

- **Repository** für Forschungsdaten und Inhalte (z.B. Digitalisate, Text Corpora etc.) mit Langzeitarchivierungsanspruch

- **TextGrid Lab**

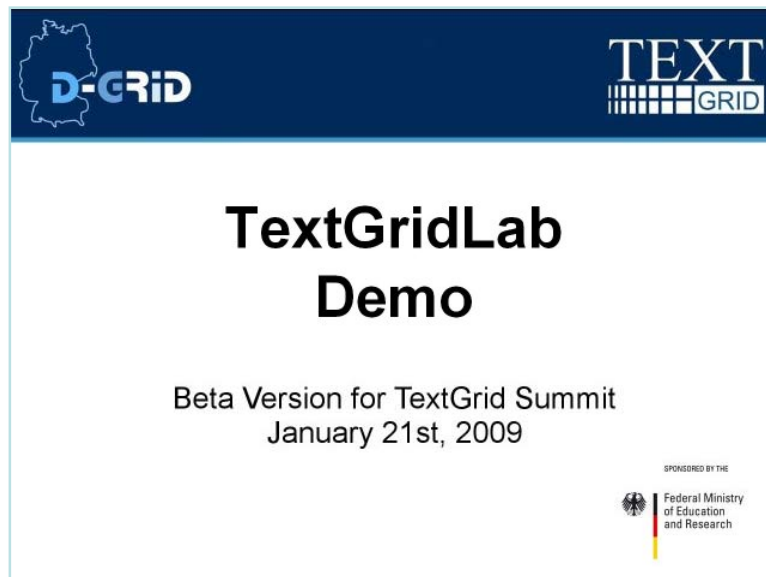
- **Laboratory** für textwissenschaftliche Tools: XML Editor, grafischer Linkeditor, Metadatentool, Annotationstool ...
- Zusammen mit Wissenschaftler entwickelt, aus unterschiedlichen GW Disziplinen

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



# Demoverision TextGrid

- **Beta Version zum Download**  
<http://www.TextGrid.de/beta.html>
- **Video Demo – Highlights**



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

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



# Rolle der (wiss.) Bibliotheken

D-Lib Magazine

September/October 2007

Volume 13 Number 9/10

ISSN 1082-9873

## Library Support for "Networked Science"

„Bench-top“ science  
„Big Science“  
„Team Science“  
„Networked Science“

An interesting article appeared in the September 7, 2007, issue of *The Chronicle for Higher Education*: "The Dawn of Networked Science" by Diana Rhoten.<sup>1</sup> The article particularly caught my eye because of its relevance to the two-part commentary by Anna Gold that appears in this issue of *D-Lib Magazine* in which Gold discusses cyberinfrastructure and possible new roles for librarians to support scientific research and data curation.

Rhoten outlines how scientific research has moved forward through various phases, beginning with "bench-top" science in which individual scientists conducted their research alone; to "Big Science" involving research by many scientists employed at one facility (for example, the Manhattan Project); to "Team Science" in which many researchers employed by multiple organizations conduct research on one project that is "anchored in a large cooperative research center" (Rhoten's example here is the Human Genome Project); and most recently to what Rhoten calls "Networked Science" in which scientists from multiple organizations collaborate or cooperate to solve a particular problem, but do so as individual researchers in a distributed research environment (one of her examples of this is the Biomedical Informatics Research Network<sup>2</sup>).

Libraries have traditionally been the custodians of the scholarly scientific record. Advances in digital technologies have introduced both new opportunities and challenges for achieving their mission in that regard. As cyberinfrastructure creates the setting in which "Team Science" and "Networked Science" can be accomplished, information professionals will need to learn new skills and take on new roles to support scientific research and data curation. They will also need to deal with complex issues, such as ensuring privacy and protecting intellectual property within a distributed information environment.

Bonita Wilson

Editor

# Rolle der (wiss.) Bibliotheken

- **Entwicklungen in der wissenschaftlichen Kooperation, Kollaboration**
  - „**Bench-top**“ **Science**: Einzelne Wissenschaftler
  - „**Big Science**“: Mehrere Wissenschaftler an einer Fakultät
  - „**Team Science**“: Mehrere Wissenschaftler aus verschiedenen Fakultäten, Organisationen arbeiten an einer bestimmten Fragestellung zusammen (z.B. Human Genome Project), verankert in einem Forschungszentrum
  - „**Networked Science**“: Wissenschaftler aus unterschiedlichen Gebieten arbeiten an einer bestimmten Fragestellung zusammen, **verteilt, kooperativ, kollaborativ, vernetzt**

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



# Rolle der (wiss.) Bibliotheken

- **Unterstützung**
  - des komplexen (nicht mehr linearen) wissenschaftlichen Prozesses (Daten, Inhalte, Dienste ...)
  - der wissenschaftlichen Kommunikation, Kooperation, Kollaboration, Austausch generell
- **Bereitstellung von Diensten/Tools, Services:**
  - Speicherung / Archivierung
  - Tool-Entwicklung
  - ...
- **Virtuelle Forschungsumgebung / Forschungsinfrastruktur**

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



# DFG Ausschreibungen I

## Ausschreibung „Repositorien – Ausbau und Entwicklung vernetzter Informationsdienstleistungen für die Wissenschaft“ (30.04.2009)

- Aufbereitung der Metadaten für die übergreifende Suche nach fachlichen und formalen Kriterien;
- Optimierung der Interoperabilität von Repositorien und ihrer Kompatibilität mit webbasierten Diensten;
- Entwicklung von Harvesting-Routinen;
- Aggregation von Daten zur Nutzungsanalyse, Zugriffsstatistik und Zitatanalyse;
- Aufbereitung der Publikationen für eine übergreifende Volltextindexierung;
- Verknüpfung von Ergebnissen aus einer maschinellen Aufbereitung von Volltexten mit Systemen zur Sacherschließung;
- Aufbau von Repositorien für multimediale Inhalte;
- wechselseitige Verzahnung institutioneller und fachbasierter Repositorien;
- Vernetzung von Repositorien über Ländergrenzen hinweg;
- aggregierende Zusammenführung und Übermittlung von Publikationen an Systeme zur Langzeitarchivierung;
- Gestaltung von Funktionen, die sich in Lehr- und Forschungsumgebungen transparent integrieren lassen;
- ...

<http://de.wikipedia.org/wiki/E-Science>

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



# DFG Ausschreibungen I

Ausschreibung Digitalisierung der Forschung

V  
V

Führung

a

ken

n

**Fokus liegt auf Publikationen,  
nicht auf Forschungsdaten -  
aber auch Betonung auf Geschäftsmodelle**

[http://www.dfg.de/forschungsfoerderung/wissenschaftliche\\_infrastruktur/lis/aktuelles/download/ausschreibung\\_massnahme\\_repositorien\\_081212.pdf](http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/aktuelles/download/ausschreibung_massnahme_repositorien_081212.pdf)

– ...  
transparent integrieren lassen;

– ...

<http://de.wikipedia.org/wiki/E-Science>



# DFG Ausschreibungen II

Ausschreibung „Virtuelle Forschungsumgebungen.  
Infrastruktur und Demonstrationsprojekte“ (31. Mai 2009)

## Ziel:

Die Förderung zielt auf die Verbreitung virtueller Forschungsumgebungen zur Unterstützung von Forschungsverbänden auf nationaler als auch internationaler Ebene ab. Dabei sollen die Entwicklung neuer oder die Zusammensetzung verschiedener bereits bewährter interoperabler Komponenten datenbezogener Infrastruktur (z.B. Wikis, Blogs, Projekt Management, Virtuelle Organisationen, Primärdatenrepositorien, Authentifizierungsstrukturen, Data-Tagging usw.) gefördert werden.

Weiteres wesentliches Ziel der Förderung ist die **verbesserte Zusammenarbeit zwischen den Forschungsverbänden und Informations-einrichtungen, vor allem Bibliotheken.**

<http://de.wikipedia.org/wiki/E-Science>

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



# DFG Ausschreibungen II

Aus

Die Projekte müssen im Antrag die konkrete Zusammenarbeit zwischen Informationseinrichtungen und Forschern skizzieren.

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

Forschungsvorhaben können in dieser Aktionslinie nicht gefördert werden!

[http://www.dfg.de/forschungsfoerderung/wissenschaftliche\\_infrastruktur/lis/download/ausschreibung\\_virtuelle\\_forschungsumgebungen.pdf](http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/ausschreibung_virtuelle_forschungsumgebungen.pdf)



# Allianz Schwerpunktinitiative „Digitale Information“

## Allianz-Partnerorganisationen:

- Alexander von Humboldt-Stiftung
- Deutscher Akademischer Austauschdienst
- Deutsche Forschungsgemeinschaft
- Fraunhofer-Gesellschaft
- Helmholtz-Gemeinschaft deutscher Forschungszentren
- Hochschulrektorenkonferenz
- Leibniz-Gemeinschaft
- Max-Planck-Gesellschaft
- Wissenschaftsrat

## Handlungsfelder

1. Nationale Lizenzierungen
2. Open Access
3. Nationale Hosting-Strategie
4. Forschungsprimärdaten
5. Virtuelle Forschungsumgebungen
6. Rechtliche Rahmenbedingungen

[http://www.dfg.de/aktuelles\\_presse/das\\_neueste/download/pm\\_allianz\\_digitale\\_information\\_details\\_080612.pdf](http://www.dfg.de/aktuelles_presse/das_neueste/download/pm_allianz_digitale_information_details_080612.pdf)

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



## Noch vieles ungelöst, z.B.:

- Fragmentierung in den Wissenschaften ist extrem und Forschung, Ressourcen sind sehr verteilt
- Große Bandbreite an unterschiedlichen Ressourcen-Typen, Standards, Formaten und Strukturen
- Semantische Beziehungen zwischen Ressourcen ist weitest gehend unerforscht und komplex
- Herausforderungen/Probleme haben wenig mit Rechenleistung oder Speicherplatz zu tun, eher mit dem Fehlen geeigneter Algorithmen
- Forschung wird aber mehr und mehr “data driven” sein
- Nur in vereinzelten Disziplinen ist wirklich IT Expertise vorhanden
- Wissenschaft nimmt neue (Forschungs)Methoden auch aus anderen Disziplinen nur sehr verzögert wahr
- “resource sharing” ist nicht immer gewünscht und möglich
- ...

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



## Aber das Aufgabenspektrum der (wiss.) Bibliotheken muß erweitert werden:

- Was früher und heute (noch) die Bücher sind, sind “(research) data” heute und in Zukunft
- Digitale Dienstleistungen und eServices für die Wissenschaftler wichtig
- Bibliotheken sollten eine Rolle spielen im veränderten Wissenschafts- und Forschungsprozess
- Aufbau IT Kompetenz notwendig
- ..

➤ **Entsprechende Ausbildung,  
Weiterqualifizierung und Lehre wichtig ☺**

Einführung  
eScience  
Bibliotheken  
Aktivitäten  
Ausblick



# Symposium FH Köln

29.01.2009

Vielen Dank für die Aufmerksamkeit!  
Anregungen, Fragen?

Heike Neuroth  
MPDL Berlin / SUB Göttingen

