



eScience:
(digitally enhanced science)

Welche Vision soll da eigentlich Realität werden?

Achim Oßwald
Fachhochschule Köln
Institut für Informationswissenschaft
achim.osswald@fh-koeln.de

9. InetBib-Tagung, Münster 7.9.2006



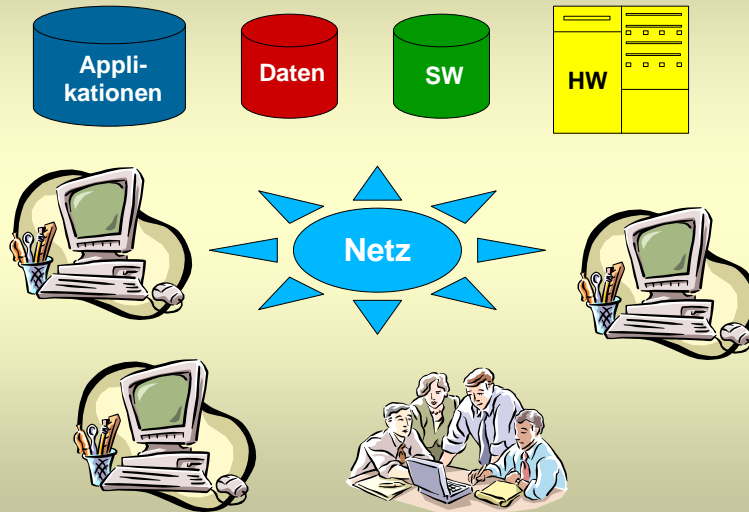
eScience: Thematischer Überblick





Wissenschaftliches Arbeiten heute

Gemeinsam genutzte Ressourcen



3 / 19



Einschränkungen der heutigen Arbeitsweise

- ↪ Ausgangspunkt ist Erfahrungswelt des **EIGEN**genutzen *Personal* Computer, d.h. jede/r arbeitet für sich und tauscht allenfalls Teil- oder Endergebnisse der Arbeit mit anderen aus
- ⇒ eher **kooperatives** als **kollaboratives** Arbeiten
- ↪ Software-Inkompatibilitäten, Versionsprobleme, mangelnde Programmkenntnisse, fehlende oder nicht berücksichtigte Standardisierungen
- ↪ nur z.T. gemeinsame Nutzung von Infrastruktur (HW, (Primär)Daten, Anwendungen, Rechenkapazität) in speziellen Wissenschaftsdisziplinen u.a. wegen Lizenzen, Abrechnung etc.

4 / 19



Wissenschaftliches Arbeiten in Zukunft



Vgl. die Animation „What is the Grid“ unter
<http://gridcafe.web.cern.ch/gridcafe/animations.html>

5 / 19



Ziele und Vorteile gridbasierten Arbeitens = Perspektive von eScience

- ↪ **eScience = digitally enhanced science**
- ↪ **gemeinsame Ressourcennutzung** bei Rechnerleistung, (Primär)Daten, Dienstleistungen etc.
⇒ **nahezu keine Begrenzung bei Verarbeitungskapazität und Speicherplatz**
- ↪ **die räumliche Verteilung** der ggf. heterogenen Daten und Anwendungen bleibt dem Anwender **verborgen**
- ↪ der Informationsaustausch über die Ressourcen erfolgt auf der Grundlage maschinenlesbarer Metadaten (**semantic web**-Konzept)
- ↪ der Regelfall ist **kollaboratives Arbeiten** mit Software, die diese Arbeitsweise unterstützt

6 / 19



Ziele und Vorteile gridbasierten Arbeitens = Perspektive von eScience

- ↙ **Virtuelle Organisationen / Arbeitsgruppen** bilden sich projektbezogen und ortsunabhängig; sie arbeiten ggf. zeitversetzt und mobil
- ↙ **Kooperationen** auf der nationalen wie internationalen Ebene werden **erleichtert**, in manchen Fällen erst ermöglicht
- ↙ **inhaltliche Synergien** und Beschleunigungen durch direkte Zusammenarbeit von Forschern und die **Transparenz der Ergebnisse**
- ↙ **Doppelforschung** aus Unwissen über parallele Aktivitäten wird eher **vermieden**



Technische Implikationen von eScience

Hardware-Voraussetzung für eScience ist eine **leistungsfähige IT-Infrastruktur**

Deshalb wird investiert u.a. in

- ↙ **Hochleistungsnetze** mit garantierten Service-Levels
- ↙ **Supercomputer** sowie Rechner-Cluster



Jülicher Supercomputer JUMP (<http://www.fz-juelich.de>);
Teil von DEISA, einem virtuellen europäischen Supercomputer



Technische Implikationen von eScience

Ziel = **Software** so gestalten, dass sie die mit eScience angestrebten Formen der Zusammenarbeit ermöglicht und unterstützt

Entwickelt hierfür wird u.a.

- ↪ **Anwendungssoftware**,
z.B. für das
 - kollaborative Arbeiten
 - Management sehr großer Datenmengen
 - interaktives Arbeiten mit lokal nicht verfügbaren Speichersystemen
- ↪ sog. **Middleware** = Software, die in der Lage ist, anwendungsunabhängig Dienstleistungen / Ressourcen zwischen Programmen zu vermitteln, ohne dass dies dem Nutzer transparent wird



Vernetzte Informationsdienstleistungen

Für das Funktionieren und den Erfolg von eScience sind informationstechnische und klassische Informationsdienstleistungen Voraussetzung, z.B.:

Authentifizierung
= Identitätsprüfung



Autorisierung
= Berechtigungsprüfung



Standardisierte Beschreibung aller prinzipiell verfügbaren Ressourcen

Bereitstellung und Präsentation von Diensten und Ressourcen zur Bearbeitung und Analyse

Nachweis aller aktuell im Grid verfügbaren Ressourcen



Vernetzte Informationsdienstleistungen

Personalisierung und Einbettung von Daten bzw. Informationen **in Anwendungsumgebungen** entsprechend den Bedürfnissen einzelner Wissenschaftler oder von Gemeinschaften von Wissenschaftlern,
z.B.

- Auswahl + Zusammenführung von Quellen unter Bezugnahme auf deren Metadaten (semantic web)
- Anpassung von Retrieval-Parametern
- Bereitstellungs- und Ausgabeoptionen angepasst an die jeweils aktuelle Arbeitsumgebung
- proaktive Informationsbereitstellung
- Bewertungs- und Empfehlungsdienste
- Dokumentations- und Datensicherungsdienste



Vernetzte Informationsdienstleistungen

Aufbau von **publikationsunterstützender Infrastruktur** wie z.B.

- Verknüpfung von Publikationen und Daten, auf die sie Bezug nehmen
- Reference linking
- Verlinkung mit domänenspezifischen bzw. domäneübergreifenden Normdateien zur einheitlichen Benennung von Personen und Objekten
- halbautomatische Generierung von Metadaten
- Generierung von inhaltsbeschreibenden Wörtern zur Anmeldung bei Suchmaschinen
- Vergabe von dauerhaften Identifikatoren
- Einbindung in Prozeduren der Langzeitarchivierung



Anforderungen an Wissenschaftler/innen durch eScience

- ↪ offline arbeiten ist out, da offline = ohne Grid
- ↪ wissenschaftliche Aktivitäten werden transparenter, aber auch kontrollierbarer
- ↪ wiss. Arbeiten wird von lokaler IT-Kompetenz unabhängiger; gleichzeitig wird die Kenntnis der IT-Potenziale für den einzelnen Wissenschaftler notwendig(er)
- ↪ Offenheit + Bereitschaft zur Zusammenarbeit incl. kollegialer Ergebnisteilnahme wird eingefordert
- ↪ Teamfähigkeit von Spezialisten gewinnt an Bedeutung



Anforderungen an Wissenschaftler/innen durch eScience

- ↪ neue Anerkennungssysteme, die die Zusammenarbeit bewerten und Eigenanteile im Gesamten erkennen lassen, werden relevant
- ↪ Open Access und die Nutzung von Open Source Software gewinnen im Gefolge von eScience weiter an Bedeutung

⇒ **Mentalitätswandel nötig**



Ziele deutscher wiss.-politischer Förderung

- ↪ **eScience wird sowohl auf nationaler wie auch internationaler Ebene gefördert**
vgl. <http://www.e-science-forum.de/de/73.php>
- ↪ **Entwicklung eines eScience-Kompetenznetzwerks als Kristallisationspunkt von Kompetenzen im Bereich Grid-Computing mit dem Ziel, Grid-Ressourcen und Grid-Know-how bereit zu stellen**
- ⇒ **Anschluss an internationale Entwicklung finden**
- ↪ **Sammlung von Erfahrungen**
 - mit **virtuellen Organisationen** und deren Nutzung von Ressourcen und Dienstleistungen
 - Entwicklung von **domäneübergreifenden Anwendungsszenarien** in Wissenschaft und Industrie



Wissenschaftspolitische Förderung

- ↪ **BMBF fördert durch die *e-Science-Initiative* Verbundvorhaben auf der Grundlage eines Konzeptes der “D-Grid-Initiative” vom Juli 2004**
Ziel: Bereitstellung einer deutschlandweiten Grid-Infrastruktur für den Aufbau von Dienstleistungen und deren breite Nutzung durch die Wissenschaft
- ↪ **Verbundvorhaben in dieser Förderlinie:**
 - DGI - D-Grid Integrationsprojekt
 - AstroGrid-D - im Bereich Astronomie
 - C3-Grid - im Bereich Klimaforschung
 - HEP-Grid - im Bereich Hochenergiephysik
 - InGrid - im Bereich Ingenieurwissenschaften
 - MediGrid - im Bereich medizinische Grundlagenforschung
 - **TextGrid** - im Bereich der Geisteswissenschaften



Wissenschaftspolitische Förderung

SYNERGIE

Demonstrator zur personalisierten Informationsversorgung wiss. Communities
492.712 €

Im Wissensnetz

Vernetzte Informationsprozesse in Forschungsverbänden
1.959.159 €

WIKINGER

Wiki Next Generation Enhanced Repository
1.844.079 €

STEMNET

Wissensmanagement systems für Sicherheit im Bereich der Stammzellbiologie
1.357.702 €

e-Science und vernetztes

Wissensmanagement

Ausschreibung des BMBF 2004/5 für ausgewählte Anwendungsbereiche; Volumen ca. 17 Mio €

Ontoverse

Kooperatives vernetztes Wissensmanagement im Bereich der Life Sciences
1.662.304 €

HyperImage

Bildorientierte e-Science-Netzwerke
1.266.779 €

WISENT

Wissensnetz Energiemeteorologie
2.019.180 €

FRESCO

Fraunhofer e-Science Cockpit
400.067 €

eSciDoc

wissenschaftliche Informations-, Kommunikations- und Publikationsplattform für die Forschung
6.036.623 €



Zusammenfassung + Schlussfolgerungen

- ↙ eScience = digitally enhanced science im Grid ist die **wissenschaftliche Arbeitsweise der nahen Zukunft** – und Wissenschaftler wie Informationsdienstleister müssen darin Kompetenz erwerben
- ↙ konkrete Software-Anwendungen sind im dt.en wiss. Kontext in der Projektentwicklung
- ↙ als primäre Dienstleistungen für eScience dienen **IT-basierte Erschließungs- und Mehrwertdienste**
=> Herausforderung und Chance für Bibliotheken
- ↙ Die Transparenz und Qualität der Erstellungsprozesse dieser Informationsdienstleistungen bestimmt das **Vertrauen in die Dienstleister**



3 von diversen offenen Fragen

1. ***Ungleichzeitigkeit:***

Für manche Wissenschaftsdisziplinen ist eScience heute schon ein Muss – für andere eine Entwicklung, die in der Zukunft liegt.
Wie kann dies ausgeglichen werden?

2. ***Chancengleichheit:***

eScience könnte ein 2-Klassen-System auslösen: Wissenschaft im Grid und solche außerhalb.
Wie kann das verhindert werden?

3. ***Widerstrebende Entwicklungen:***

Wie können wir Studierende für teamorientierte eScience vorbereiten, während gleichzeitig föderale und partikulare Strukturen in der Hochschulwelt an Bedeutung gewinnen?

Lassen Sie uns darüber diskutieren! 😊