

Digitale Langzeitarchivierung und Grid: Gemeinsam sind wir stärker?

Dr. Heike Neuroth
SUB Göttingen
neuroth@sub.uni-goettingen.de

Leipzig, 21. März 2007

Überblick

Was ist Grid / eScience?

Was hat das mit LZA zu tun?

Was macht nestor dazu?

Was ist Grid?



Was ist Grid?

Old World

Static

Silo

Physical

Manual

Application



New World

Dynamic

Shared

Virtual

Automated

Service

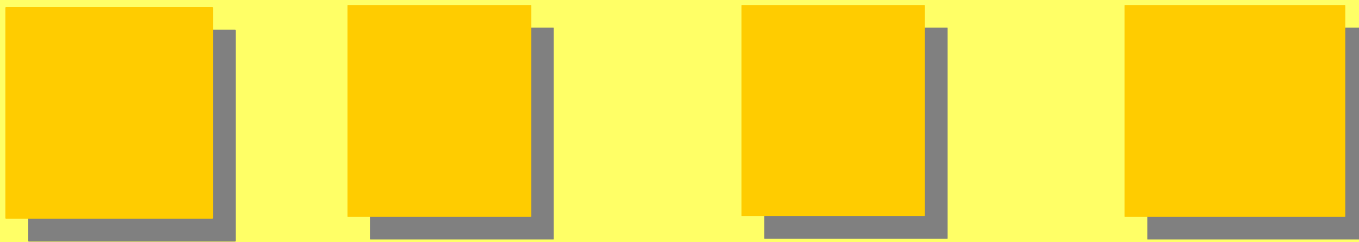
Schichten des Grid

Wissenschaftliche Communities, Institutionen

Disziplinspezifische Werkzeuge und Dienste

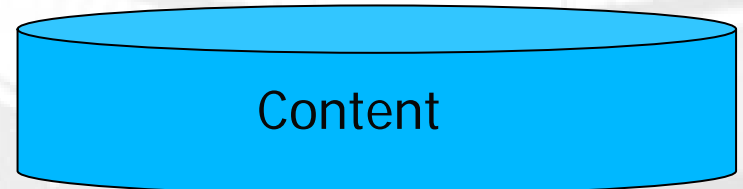


Disziplinübergreifende Werkzeuge und Infrastruktur

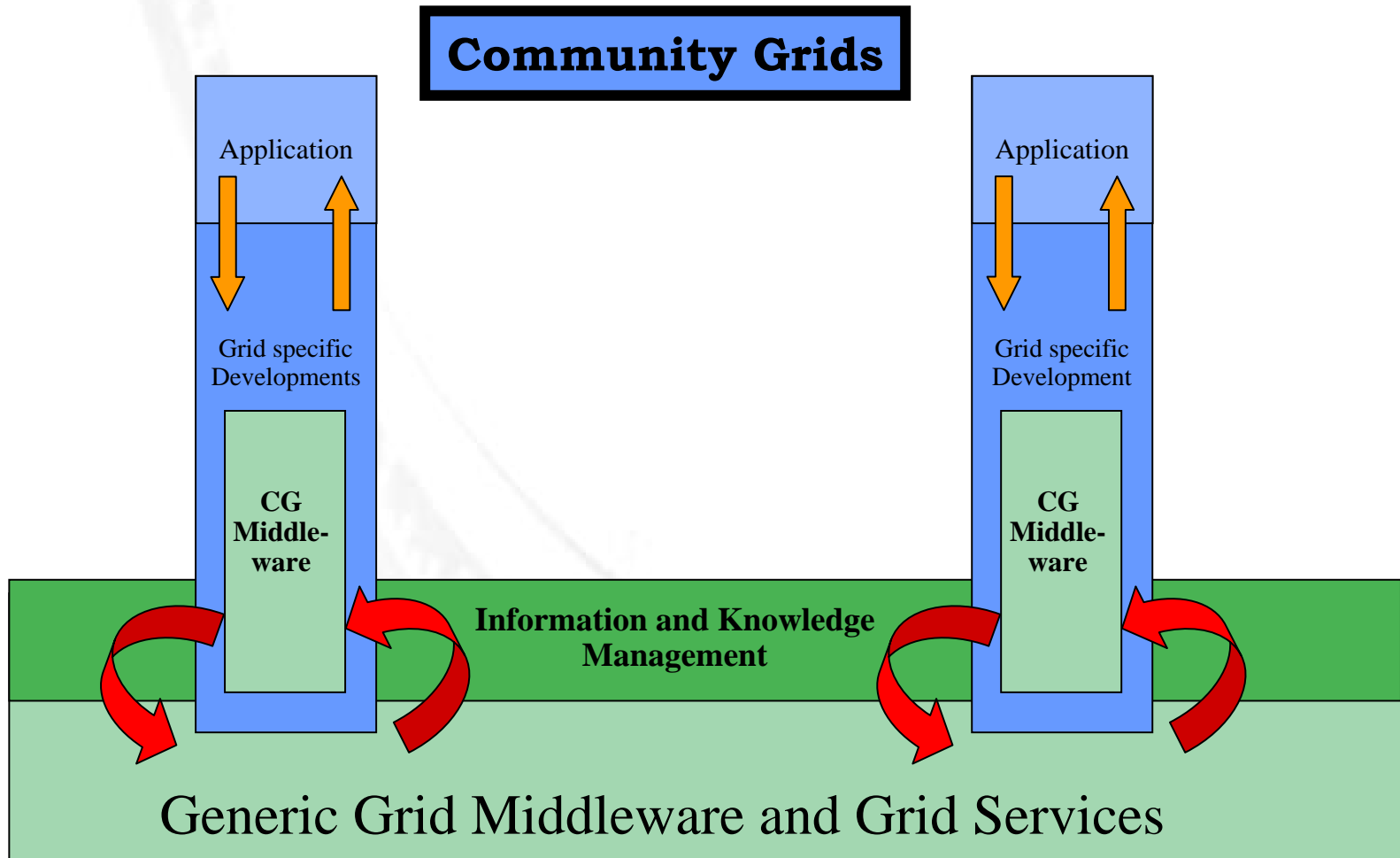


Virtualisierte Hardwareressourcen

Content



Technische Darstellung



eScience

"e-Science is about global collaboration in key areas of science and the next generation of infrastructure that will enable it."

(John Taylor, Director General of the UK Research Council)

- **Infrastruktur** für die Wissenschaft
- Technische **Integration**
- **Vernetzung** von Wissenschaftlern

eScience

- **Rechner- und Speicherkapazitäten**

→ virtuelle Wissensumgebungen mit umfassenden Datenbeständen, Visualisierungen und wissenschaftlichen Informationen

- **Technische Integration**

→ Tiefgreifende Verknüpfung vielfältiger Akteure, Dienste, Ressourcen und Methoden

- **Vernetzung von Menschen**

→ Unterstützung von Arbeitsabläufen, Kommunikation und Wissensaustausch

D-Grid

- **DGI:** Projektleitung, D-Grid-Basis-Infrastruktur und Software
- **Astro-Grid:** Nutzung und Management von Astronomie-Ressourcen (Messinstrumente, Rechner, Datenarchive)
- **C3-Grid:** gemeinsame und verteilte Verwaltung und Nutzung von Modell- und Beobachtungsdaten
- **HEP-Grid:** Datenauswertung in der Hochenergie-/Teilchenphysik, extremer Speicher- und Rechnerbedarf
- **InGrid:** Grid für zentrale, rechenintensive Probleme der Ingenieurwissenschaften mit verschiedenen Anwendungen
- **MediGrid:** Nutzung verteilter, heterogener Daten unter Wahrung des Datenschutzes
- **TextGrid:** Workbench und Tools für die gemeinschaftliche philologische Bearbeitung, Analyse, Annotation, Edition und Publikation von Textdaten

- **links4science:** Analyse des Integrations- und F&E-Bedarfs im Bereich eScience-Wissensvernetzung
- **eSciDoc:** Workbench und Repository in web-basierter Umgebung, bindet externe Dienste ein
- **FRESO:** Workbench und verteilte Informationsinfrastruktur, bindet externe Dienste ein
- **HyperImage:** Bildrepository und -editor, insbesondere zur Annotation und Verlinkung
- **Im Wissensnetz:** Workbench mit Diensten zur Forschervernetzung durch Data Mining und Data Sharing
- **Ontoverse:** kooperative und semi-automatische Erstellung von Ontologien im Bereich Life Sciences
- ...

Was hat das mit LZA zu tun?

Warum Grid für LZA?

LZA braucht GRID:

1. Grid-Technologie ist die neuste Informationstechnologie, die entwickelt wird.
2. Mit Grid-Technologie werden Datenaufkommen in einer enormen Größordnung verarbeitet.
3. Tools, Services etc. werden zunehmend in Grid Umgebung integriert (z.B. Fedora, Editions-Tools ...)

Grid braucht LZA:

1. Die wissenschaftlichen Datenaufkommen müssen auch in der Grid Umgebung nachhaltig gesichert werden
2. Erfahrungen in diesem Bereich (Workflows, Ingest ...) kommen aus der LZA Community

Grid für LZA

Grid abstrahiert von der konkreten Hardware. Dies entbindet LZA-Dienste von vielen Problemen, z.B.:

- beim Speicher (große Mengen on-demand, verteilte und redundante Speicherung, ...)
- bei der Rechenkapazität (für Migrationen, Verwaltung großer Mengen, beim Zugriff, ...)

Und es gibt weitere nützliche und nachnutzbare Dienste im Grid (z.B. verteilte Authentifizierung).

Was macht nestor dazu?

Was macht nestor dazu?

Arbeitsziel: Kontextualisierung von
Langzeitarchivierung, Grid-Computing und e-Science

- Expertisen
- AG LZA/Grid

Expertisen

- 1. Rohdaten:** "Langzeitarchivierung meets e-Science im Bereich wissenschaftlicher Rohdaten",
GeoForschungszentrum Potsdam, Dr. Jens Klump
- 2. Standards:** "Standards und Standardisierung im Kontext von Grid-Technologien und Langzeitarchivierung",
Universität der Bundeswehr München, Prof. Dr. Uwe M. Borghoff
- 3. Synergien:** "Synergiepotenziale zwischen GRID- und e-Science-Technologien für die Langzeitarchivierung",
Fernuniversität Hagen, Prof. Dr. Wolfram Schiffmann

Koordinierung der Expertisen, Zusammenführung der Ergebnisse und Formulierung weiterer Bedürfnisse

Definition von Testbeds

Durchführung von Veranstaltungen für und von beiden Communities (z.B. Workshop auf GES)

Beobachtung internationaler Entwicklungen

Erarbeitung einer **Roadmap** zur Zusammenführung von Grid und LZA bzw. Gridifizierung von Archivsystemen

Inhaltliche Fragen:

Was ist der LZA-Bedarf und was sind die Anforderungen im Grid-Bereich?

Was ist das Potential der eScience-/Grid-Technologie für die LZA?

Was bedeutet Vertrauenswürdigkeit eines Archivs im Grid-Kontext?
(dezentrales Archiv, Virtuelle Organisation ...)

Mission Statement der AG

e-Science steht für kollaborativ und gleichzeitig verteilt betriebene Wissenschaft, ermöglicht durch modernste Infrastruktur.

e-Science in Verbindung mit Grid-Technologie erfordert aufgrund eines hohen Datenaufkommens und großer technischer Veränderungen besondere Lösungen im Bereich der digitalen Langzeitarchivierung. Umgekehrt könnte diese Technologie ein großes Potential für die Implementierung von Langzeitarchivsystemen haben.

Die Aufgabe der AG ist es, dieses neue Gebiet mit seinen Chancen und Risiken zu skizzieren und eine Roadmap zu seiner Erschließung sowie zur Gridifizierung von Archivsystemen aufzustellen.

GES-Workshop

German e-Science Conference 2007,
bietet Überblick über eScience/Grid in Deutschland, auch
mit internationale Experten

stellt nestor der Grid-Community vor und den Zwischenstand
der Expertisen

ist ein Schritt zur Zusammenführung von Grid und
Langzeitarchivierung zur Schaffungen eines gemeinsamen
Verständnisses der Anforderungen und Begrifflichkeiten der
bisher getrennten Communities

Informationsplattform

[➤ Informationsdatenbank](#)[➤ Wer Wo Was](#)[➤ Aktuelles](#)[➤ News Archiv](#)[➤ News Suche](#)[➤ Hilfe](#)[➤ Feedback](#)[➤ Presse](#)[➤ Kalender](#)[➤ Qualifizierung](#)[➤ Projekte](#)[➤ Newsletter](#)[➤ Annotation](#)

Kompetenznetzwerk

[➤ Home](#)

Kommunikationsplattform

[➤ Mailinglisten](#)[➤ Arbeitsbereich](#)

Links

[➤ Padi](#)[➤ Forum Bestandserhaltung](#)[➤ Netzwerk Mediatheken](#)

gefördert vom



Bundesministerium
für Bildung
und Forschung

[Übersicht](#) | [Ablauf](#) | [Anmeldung](#) | [Teilnehmerliste](#) | [Kontakt](#)

Workshop "Digitale Langzeitarchivierung und Grid: Gemeinsam sind wir stärker?"

Termin: Mittwoch, 2. Mai 2007, 11:00 bis 13:00

Veranstalter: nestor im Rahmen der **German e-Science Conference 2007**

Moderation: Heike Neuroth

Ort: Raum 1, Kongresshaus Baden-Baden

Die Veranstaltungssprache ist Deutsch.

Grid- und e-Science-Technologien zeichnen sich oft durch ein hohes Datenaufkommen und schnelle technische Veränderungen aus, was einen besonderen Bedarf nach Lösungen im Bereich der digitalen Langzeitarchivierung nahelegt. Umgekehrt könnten die Grid-Technologien ein großes Potential für die Implementierung von Langzeitarchivierungsdiensten und -werkzeugen haben. Was sind die Chancen und Herausforderungen bei der Kooperation der bisher getrennten Grid- und Langzeitarchivierung-Communities? Gibt es ein gemeinsames Verständnis der Anforderungen und Begrifflichkeiten?

Das Thema Grid ist ein neuer Schwerpunkt des Kompetenznetzwerks Langzeitarchivierung "nestor". Als Teil dieses Schwerpunktes werden drei Expertisen erstellt, die das Thema unter den Gesichtspunkten wissenschaftlicher Rohdaten, Standards und möglicher Synergien untersuchen, und es wurde eine Arbeitsgruppe mit Vertretern aus der D-Grid Initiative und den Bereichen e-Science und Langzeitarchivierung gegründet. Der Workshop stellt die Expertisen vor, diskutiert die bisherigen Ergebnisse und versucht, eine Grundlage für die Zusammenführung von Grid und Langzeitarchivierung zu schaffen.



AG Mitglieder

- ***Bibliotheken*** (z.B. SUB, DNB, BSB)
- ***Fachinformationszentrum*** (z.B. FIZ Karlsruhe)
- ***Universitäten*** (z.B. FU Hagen, Uni BW)
- ***Rechenzentrum*** (z.B. GWDG)
- ***Forschungszentren*** (z.B. GFZ Potsdam)
- ...

**Vielen Dank für
Ihre Aufmerksamkeit !**

Fragen, Anmerkungen ... ?



Heike Neuroth, SUB Göttingen
neuroth@sub.uni-goettingen.de

