MediGRID

Grid-Computing für Medizin und Lebenswissenschaften

Abschlussbericht des Projektes MediGRID

Berlin, 07. August 2009

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Das diesem Projekt zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter den Förderkennzeichen 01AK803A bis H gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

Inhaltsverzeichnis

Mitarbeiterverzeichnis ..... **vii**

Abbildungsverzeichnis ..... **ix**

Kurzdarstellung des Verbundvorhabens **1**

1 Aufgabenstellung des Vorhabens 1
2 Voraussetzungen und Rahmenbedingungen 1
3 Planung und Ablauf des Vorhabens 1
4 Wissenschaftlicher und technischer Stand 1
5 Zusammenarbeit mit anderen Stellen 1

Projektbericht – Darstellung der Projektergebnisse **3**

1 Einführung 3
1.1 Projektziel und Projektinhalt 3
1.2 Projektlaufzeit und Projektorganisation 4
2 Ergebnisse des Verbundvorhabens 5
2.1 Medizinische Bildverarbeitung 5
2.2 Biomedizinische Informatik 16
2.3 Klinische Forschung 19
2.4 Middleware 27
2.5 Ressourcenfusion 54
2.6 Ontologiewerkzeuge 67
2.7 eScience – Enhanced Science 77
2.8 Geschäftsmodelle für Healthgrids 89
3 Weitere Aktivitäten im Verbundprojekt 103
3.1 Administrative Koordination 103
3.2 Öffentlichkeitsarbeit 103
3.3 Nationale und internationale Vernetzung 104
4 Schlussfolgerungen und Ausblick 107
5 Veröffentlichungen, Vorträge und Dissertationen 109
5.1 Veröffentlichungen 109
5.2 Vorträge 117
5.3 Dissertationen im Rahmen des Projektes 121

Anhang **123**

A1 Pressemitteilungen (Auswahl) 123
A2 Projektflyer 133

Mitarbeiterverzeichnis

Dr. A. Andrzejak	ZIB, Berlin
Prof. Dr. J. Bernarding	Universität Magdeburg
K. Beronov	LSTM, Erlangen
M. Blaurock	TMF e.V., Berlin
W.Cassel	Universität Marburg
Dr. S. Canisius	Universität Marburg
Dr. T. Deserno	RWTH, Aachen
H. Dietze	Technische Universität Dresden
O. Dzimova	LSTM, Erlangen
Dr. K.-H. Englmeier	GSI, München
T. Ernst	Fraunhofer FIRST, Berlin
J. Falkner	Fraunhofer IAO, Stuttgart
M. Freudigmann	TMF e.V., Berlin
Dr. J. Hampe	Universität Kiel
M. Hartung	Universität Leipzig
Dr. A. Herrmann	Universität Kiel
Prof. Dr. H. Herre	Universität Leipzig
A. Hoheisel	Fraunhofer FIRST, Berlin
Dr. T. Kirsten	Universität Leipzig
S. Kottha	Technische Universität Dresden
Assist. Prof. U.D. Dr. T. A. Knoch	Universität Heidelberg, DKFZ
Dr. D. Krefting	Charité, Berlin
Dr. T. Lingner	Universität Göttingen
F. Loebe	Universität Leipzig
K. Loebe	Universität Leipzig

Prof. Dr. M. Löffler	Universität Leipzig
Prof. Dr. P. Meinzer	Universität Heidelberg, DKFZ
Dr. Y. Mohammed	Universität Göttingen
P. Müller	Universität Heidelberg, DKFZ
R. Müller-Pfefferkorn	Technische Universität Dresden
Prof. Dr. T. Penzel	Charité, Berlin
K. Peter	ZIB, Berlin
T. Ploch	Universität Marburg
Prof. Dr. E. Rahm	Universität Leipzig
Prof. Dr. O. Rienhoff	Universität Göttingen
Prof. Dr. U. Sax	Universität Göttingen
S. Scholz	TMF e.V., Berlin
U. Schwardmann	GWDG, Göttingen
S. C. Semler	TMF e.V., Berlin
D. Sommerfeld	GWDG, Göttingen
Dr. T. Steinke	ZIB, Berlin
O. Strauß	Fraunhofer IAO, Stuttgart
A. Thiel	OFFIS, Oldenburg
Prof. Dr. T. Tolxdorff	Charité, Berlin
F. Viezens	Universität Göttingen
M. Vossberg	Charité, Berlin
PD Dr. A. Weisbecker	Fraunhofer IAO, Stuttgart
P. Welter	RWTH, Aachen

Abbildungsverzeichnis

Abbildung 1. D-Grid I Communities 3
Abbildung 2. MediGRID Verbundvorhaben - Projektpartner 4
Abbildung 3. Gridportlet mit Eingabe der Simulationsparameter, interaktiver 3D-Visualisierung und Statusanzeige des Workflows 6
Abbildung 4. Portlet zur Registrierung mit ITK. Für die Ultraschallbildgebung sind geeignete Parameterwerte vorgewählt. Beliebige 3D-Daten können registriert werden, wobei eine Vielzahl von Optionen möglich ist 8
Abbildung 5. Systeme und Datenfluss beim MediGRID-Anschluss an der Charité 10
Abbildung 6. PACS-Komponente für den Zugriff über das Gridportal, hier in der fMRI-Anwendung. Die Patientennamen wurden durch Pseudonyme ersetzt 11
Abbildung 7. Arbeitsumgebung des GLOBE 3D Genome Viewers 17
Abbildung 8. Beschreibung des Workflow 21
Abbildung 9. Screenshot des Portlets 22
Abbildung 10. Berechnung der Herzfrequenz aus dem EKG 23
Abbildung 11. Erkennung von schlafbezogenen Atmungsstörungen (SBAS) 23
Abbildung 12. Funktionsweise des EMG-Algorithmus 24
Abbildung 13. Pseudonymisierungskonzept 25
Abbildung 14. Portalbasierter Zugang zu MediGRID 32
Abbildung 15. Screenshot der im MediGRID-Portal integrierten grafischen Nutzerschnittstelle zur Ausführung und Monitoring von Workflows 34
Abbildung 16. Screenshot der Nutzerschnittstelle für die Verwaltung von MediGRID-Workflows 35
Abbildung 17. MediGRID Anwendungen a) „Ontologie-Anwendung“, b) „virtuelle Gefäßchirurgie“ 40
Abbildung 18. MediGRID Anwendungen a) „3D Ultraschall Prostatabiopsie“, b) „AUGUSTUS Gensequenzanalyse“ 40
Abbildung 19. MediGRID-Anwendungen a) „Funktionelle Hirnbilddaten“, b) QRS-Analyse von Schlaf-EKGs 40
Abbildung 20. Monitoring in MediGRID mittels Ganglia 41
Abbildung 21. MDS Ressourcen Monitoring im MediGRID Portal 42
Abbildung 22. Screenshot des Nagios-Überwachungssystems 43
Abbildung 23. Grid Workflow Execution Service - Jobdetails / Jobmonitoring / Jobmanagement 44
Abbildung 24. Monitoring von Workflows in MediGRID 44
Abbildung 25. SRB-Installation in MediGRID 60
Abbildung 26. Ressourcen der SRB-Installation in MediGRID 60

Abbildung 27. Datentransferzeiten (in Sekunden) zwischen SRB-Client und –Server für verschiedene Datengrössen (links) und beim Zugriff auf verschiedene Standorte (rechts) ... 61
Abbildung 28. Leserate (in kB/s) aus einer Datenbank mit OGSA-DAI mit verschiedenen Zugriffsmechanismen ... 63
Abbildung 29. Architektur der realisierten Ontologie-Middleware ... 70
Abbildung 30. Verwendung von Ontologien im MediGRID Portal ... 72
Abbildung 31. Entwicklungspfad von Grid-Typen ... 91
Abbildung 32. Referenzgeschäftsmodell für Healthgrids ... 92
Abbildung 33. Wertschöpfungskette Grid Computing ... 93
Abbildung 34. Typisierung von Geschäftsmodellen ... 94
Abbildung 35. Anwendungsfelder Grid Computing ... 95
Abbildung 36. Marktdurchdringung Healthgrids ... 96
Abbildung 37. Erlösmodelle für Healthgrids ... 97
Abbildung 38. Idealtypischer Entwicklungsprozess von MediGRID ... 98
Abbildung 39. Perspektive einer integrierten Grid-Infrastruktur für MediGRID ... 99

Kurzdarstellung des Verbundvorhabens

1 Aufgabenstellung des Vorhabens

Aufbau und Betrieb von Grid-Services und deren beispielhafte Nutzung durch Forscher aus der Community Medizin und Lebenswissenschaften in Zusammenarbeit mit anderen Communities und den Dienstleistungsrechenzentren.

2 Voraussetzungen und Rahmenbedingungen

Die Voraussetzungen, unter denen das Vorhaben durchgeführt wurde, ergaben sich durch die Vorarbeiten bis 2005 sowie durch die im Vorprojekt definierten Bedingungen und Optionen. Die im Projektantrag deutlich stärker vorgesehene Koordinierung seitens der TMF – in Kenntnis der Bedingungen in der Medizin und den Lebenswissenschaften wurde im Begutachtungsprozess unhilfreich stark reduziert.

3 Planung und Ablauf des Vorhabens

Das Projekt wurde administriert durch die Telematikplattform medizinischer Forschungsnetze (TMF), einer durch den BMBF seit 1999 aufgebauten Einrichtung zur Förderung von Forschungsinfrastruktur. Der Projektverlauf entsprach bis auf geringe Abweichungen den Planungen.

4 Wissenschaftlicher und technischer Stand

Der wissenschaftliche und technische Stand, an den angeknüpft wurde, ergab sich aus den Vorläuferprojekten in der EU und in den USA. Ferner aus der Tatsache, dass die IT-Dienstleisterinfrastruktur in Medizin und Lebenswissenschaften stark dezentralisiert ist und keiner Abstimmung, z.B. durch IT-intensive internationale Großprojekte, unterliegt. Die meisten Primärforscher arbeiten nicht direkt mit IT-Diensten sondern vermittelt über methodische Einrichtungen (oft Institute für Statistik, Epidemiologie oder Informatik). Die Ausgangslage 2005 war dadurch charaktersisiert, dass nur einzelne Personen und Projekte das Thema Grid-Services adressierten.

Demzufolge war auch die Verfügbarkeit von eingearbeitetem Personal auf Einzelpersonen beschränkt. Durch das Projekt MediGRID und die erheblichen finanziellen Anstrengungen des BMBF konnte die nationale Situation verbessert werden – jeoch nach wie vor weitab von einem Nutzungsdurchbruch.

5 Zusammenarbeit mit anderen Stellen

Die Zusammenarbeit mit D-Grid erfolgte über Workshops und die D-Grid internen Abstimmungsrunden. Die Zusammenarbeit mit den internationalen Grid-Vorhaben erfolgte über Studienbesuche und Workshops. Internationale Erfahrungen konnten auch über Forschungsbeziehungen der Modulleiter bzw. des Koordinators (DFG-Studienreisen) eingebracht werden.

Es zeigte sich, dass die komplexe Materie nicht leicht kommunizierbar ist. Zwischen den verschiedenen Ländern, Communities, Nutzern bzw. Anbietern ist nahezu jeweils alles anders. Da die Technologie jedoch bezüglich Akzeptanz und Ausgestaltung stark abhängig von den Rah-

menbedingungen ist, limitieren diese Rahmenaspekte den Austausch. Tatsächlich wäre erheblich mehr Geld für die nicht-technische Arbeit notwendig um tatsächlich Kooperationen zu etablieren.

Besonders hilfreich für den Community-übergreifenden Austausch erwiesen sich die All-Hands-Meetings, an deren Organisation sich MediGRID 2007 und 2009 beteiligte. MediGRID nahm ferner an den internationalen Tagungen HealthGRID 2006, 2007 und 2008 teil. Es warb die HealthGRID Konferenz 2009 nach Berlin ein.

Aus MediGRID ist das Grid-Forum der TMF hervor gegangen, dass inzwischen eine anerkannte Einrichtung aller Gridprojekte in Medizin und Lebenswissenschaft geworden ist und von der TMF getragen wird. Das Forum besteht über die Förderung des Projektes hinaus und erfreut sich einer wachsenden Zahl teilnehmender Projekte.

Projektbericht – Darstellung der Projektergebnisse

1 Einführung

1.1 Projektziel und Projektinhalt

Das Verbundvorhaben MediGRID ist ein Projekt der Initiative D-GRiD (www.d-grid.de) des Bundesministeriums für Bildung und Forschung – BMBF (siehe Abbildung 1). Es betreibt die Entwicklung einer IT-Integrationsplattform und darauf aufsetzender eScience-Dienste für die biomedizinische Wissenschaft. Dabei zeigt das Projekt am Beispiel biomedizinischer Forschung mit hochdimensionalen Daten und der Verknüpfung von vielfältigen, genotypischen und phänotypischen Daten die Anwendbarkeit und Relevanz von Grid-Computing-Diensten in der Medizin und in den Lebenswissenschaften.

Die vier methodischen Projekt-Module Middleware, Ontologie-Werkzeuge, Ressourcenfusion und eScience haben schrittweise eine Grid-Computing-Infrastruktur entwickelt, die auf den Anforderungen der Grid-Nutzer aus dem biomedizinischen Umfeld basieren. Auf dieser Infrastruktur setzen die Applikationen aus drei anwendungsorientierten Projektmodulen auf:

- Bildverarbeitung,
- biomedizinische Informatik,
- und klinische Forschung.

Die unterschiedlichen Anwendungen stehen den registrierten Nutzern im MediGRID Applikationsportal (portal.medigrid.de) zur Nutzung zur Verfügung. Unter ihnen befinden sich Anwendungen zur DNA-Analyse, aus dem Bereich der Schlafforschung und der Bilddiagnose.

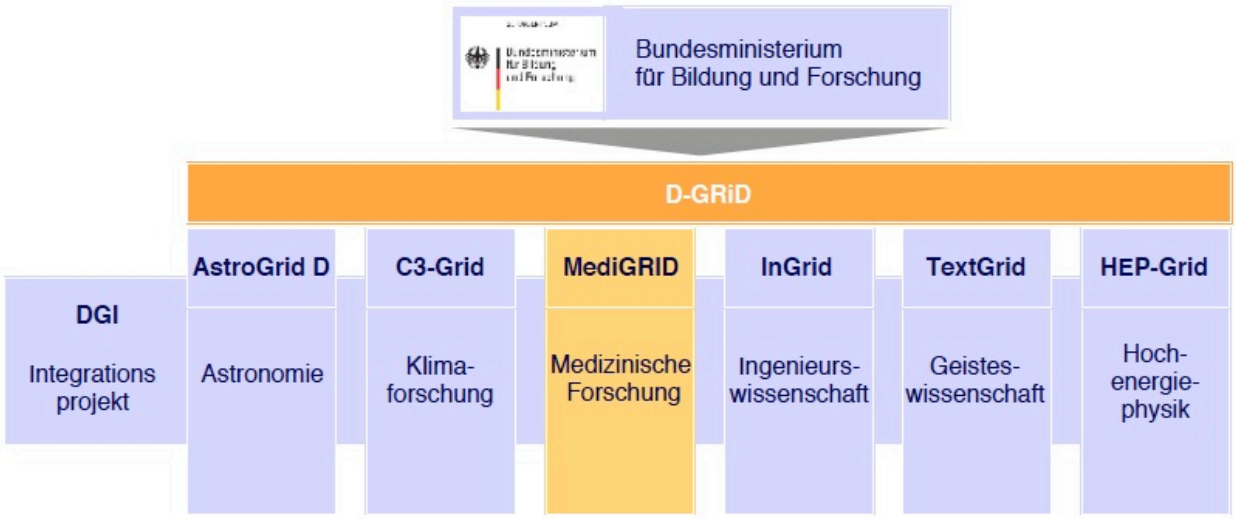


Abbildung 1. D-Grid I Communities

1.2 Projektlaufzeit und Projektorganisation

MediGRID wurde zwischen Dezember 2005 und Februar 2009 von dem BMBF umfassend gefördert.

Für das Projektvorhaben MediGRID wurde ein modularer Aufbau gewählt. In den vier methodischen Modulen Middleware, Ontologie-Werkzeuge, Ressourcenfusion und eScience wurde eine Grid-Infrastruktur erarbeitet. Für drei repräsentative Fachbereiche der biomedizinischen Forschung (Bildverarbeitung, Biomedizinische Informatik, Klinische Forschung) wurden exemplarische Anwendungen in diese Grid-Infrastruktur portiert. Die Gesamtstruktur des Projektes ist in Abbildung 2 dargestellt.

Die Einbeziehung zahlreicher Partner aus Industrie, Versorgungs- und Forschungseinrichtungen und die Kooperation mit weiteren internationalen Grid-Computing-Initiativen stellt das Projekt auf eine breite Interessensbasis.

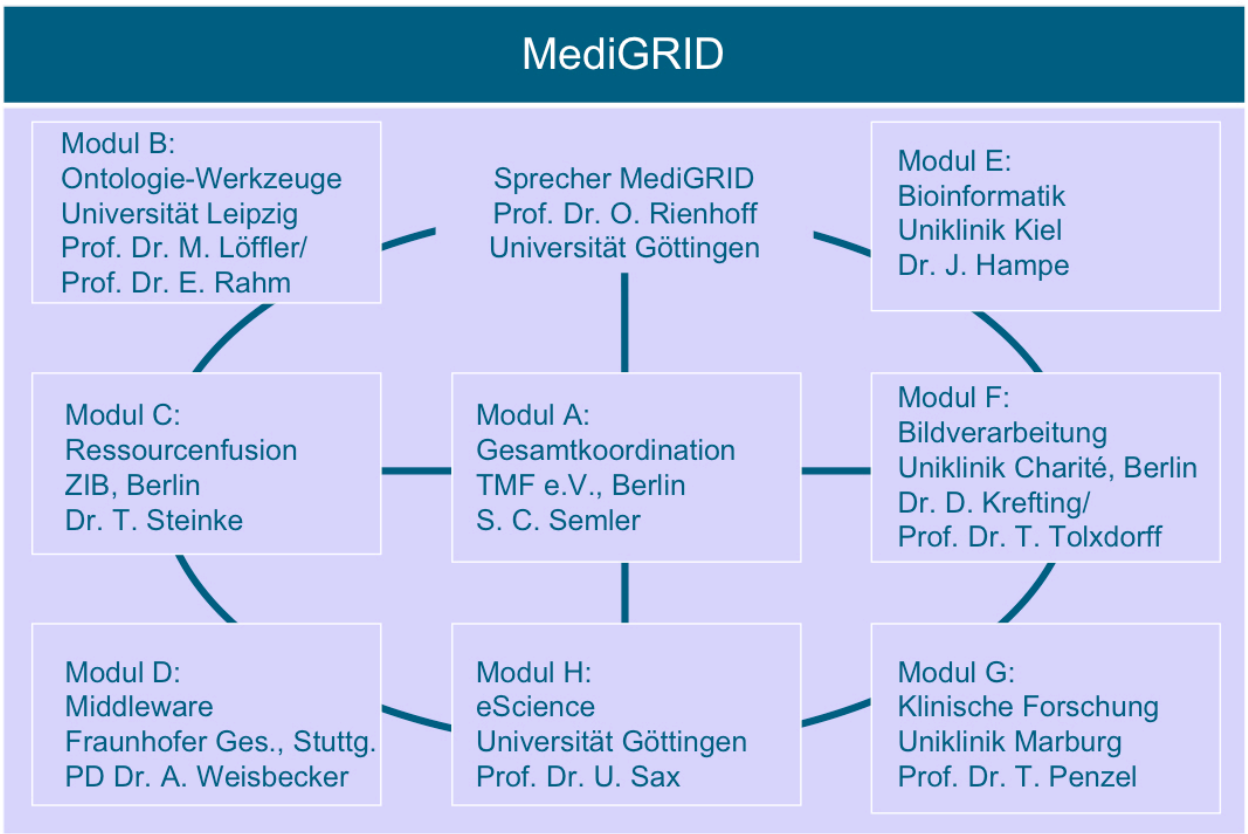


Abbildung 2. MediGRID Verbundvorhaben - Projektpartner

2 Ergebnisse des Verbundvorhabens

Die Ergebnisse des Verbundvorhabens werden im Folgenden für die einzelnen beteiligten Module separat dargestellt. Die verantwortlichen Projektinstitutionen ergeben sich aus Abbildung 2. Begleitmaßnahmen wie die Gesamtkoordination, die Öffentlichkeitsarbeit und die internationale Vernetzung werden in Kapitel 3 erläutert.

Für jedes Modul werden die Modulziele, die wissenschaftlich-technischen und sonstigen Modulergebnisse, der Nutzen und die Verwertbarkeit der Ergebnisse und die generelle Fortentwicklungen im betrachteten Themenbereich durch Dritte dargestellt und durch einen Ausblick abgeschlossen.

2.1 Medizinische Bildverarbeitung

2.1.1 Moduleinordnung und Modulziele

Ziel des Moduls Bildverarbeitung in MediGRID war es, medizinische Bildverarbeitungsprobleme mit Hilfe von Methoden und Werkzeugen des Grid Computing effizienter zu lösen, bzw. durch Einsatz der neuen Technologien neue Anwendungen zu erschließen. Dazu sollten die wesentlichen Methoden, Datenstrukturen und Datenmanagement der medizinischen Bildverarbeitung exemplarisch anhand von drei Anwendungsszenarien in das Grid implementiert werden. Die im Rahmen des Projektes entwickelten Grid-Anwendungen – mit besonderer Berücksichtigung der sicheren Bildübertragung und Speicherung – sollten über die wissenschaftliche Forschung hinaus eine solide Basis für spätere klinische Anwendungen darstellen.

2.1.2 Modulergebnisse

Drei Anwendungsszenarien sind erfolgreich in das MediGRID implementiert worden:

- Analyse funktioneller Hirnbilddaten,
- Virtuelle Gefäßchirurgie,
- 3D-Ultrasschall.

2.1.2.1 Analyse Funktioneller Hirnbilddaten

Für die Bestimmung von durch einen bestimmten sensorischen Reiz ausgelöste Aktivierungen von Hirnregionen sind bis zu 600 Einzelmessungen notwendig, die je eine Magnetresonanztomographie umfassen. Die Vorverarbeitung dieser Bilder ist sehr aufwändig und erfordert verschiedene Bildverarbeitungsschritte, kann jedoch parallelisiert erfolgen. Zur effizienten Verarbeitung der Magnetresonanztomographien wurde das freie, auf Matlab basierende Softwarepaket „SPM“ in das MediGRID implementiert und in das MediGRID-Portal integriert. Zu verarbeitende Datensätze können aus dem angeschlossenen Bildarchiv (PACS) und dem Datenmanagementsystem (SRB) ausgewählt, oder neu hochgeladen werden. Eingabeparameter wie die Halbwertsbreite des Glättungsfilters können über das Portlet übergeben werden. Die Applikation kann auf bis zu 8 Knoten gleichzeitig durchgeführt werden. Die Ergebnisse werden anschließend im Datenmanagementsystem gespeichert und können von dort zur lokalen Weiterverarbeitung heruntergeladen werden. Die Implementierung ist prototypisch für eine Gridimplementierung von auf Matlab basierendem Code. Dies schränkt jedoch auch die Skalierbarkeit ein, da es im Rahmen des D-Grid nicht gelungen ist, eine Lizenzvereinbarung mit dem Lizenzinhaber Mathworks zu treffen. Deshalb kann die Anwendung derzeit nur auf einer der MediGRID-Ressourcen ausgeführt werden.

2.1.2.2 Virtuelle Gefäßchirurgie

Die hämodynamische Simulation ermöglicht die patientenspezifische Analyse von Gefäßgeometrien bezüglich des Blutflusses, wie zum Beispiel die Bestimmung des Wanddruckes in der Nähe eines Aneurysmas. Die virtuelle Gefäßchirurgie ermöglicht nun die Bestimmung der hämodynamischen Größen nach Veränderung der Geometrie, wie zum Beispiel das Einsetzen eines Stents oder das Abklemmen eines Blutgefäßes. Verschiedene Szenarien können so vor der eigentlichen Operation durchgespielt und miteinander verglichen werden.

Während die Bestimmung und die Modifikation der Geometrie, also die virtuelle Operation, ein interaktiver Prozess ist, verläuft die hämodynamische Simulation automatisch, ist aber aufgrund der notwendigen Genauigkeit sehr rechenintensiv und übersteigt im Allgemeinen schnell die lokalen Rechenressourcen. Zudem erwartet das auf „fortran“ basierende Programm eine fest definierte Ordnerstruktur und Dateinamensbenennung, die wenig nutzerfreundlich ist und eine lokale Installation fehleranfällig macht. Eine Analyse der existierenden Tools zur Segmentierung der Blutgefäße aus den Angiographien ergab, dass eine interaktive Segmentierung im Grid zwar prinzipiell möglich ist, jedoch einen direkten Zugang zu den Ressourcen benötigt, der durch die Firewallbestimmungen der meisten klinischen IT-Netze nicht möglich ist. Zudem bieten die meisten Modalitäten, mit denen heute Angiographien durchgeführt werden, bereits entsprechende Segmentierungssoftware. Deshalb wurden die spezifischen, für die Durchführung und Bewertung der hämodynamischen Simulation notwendigen Schritte in das Grid implementiert. Diese umfassen die Auswahl des Simulationsgebietes auf Basis der vorsegmentierten Geometrie, die Eingabe der Simulationsparameter und die Visualisierung der Ergebnisse.

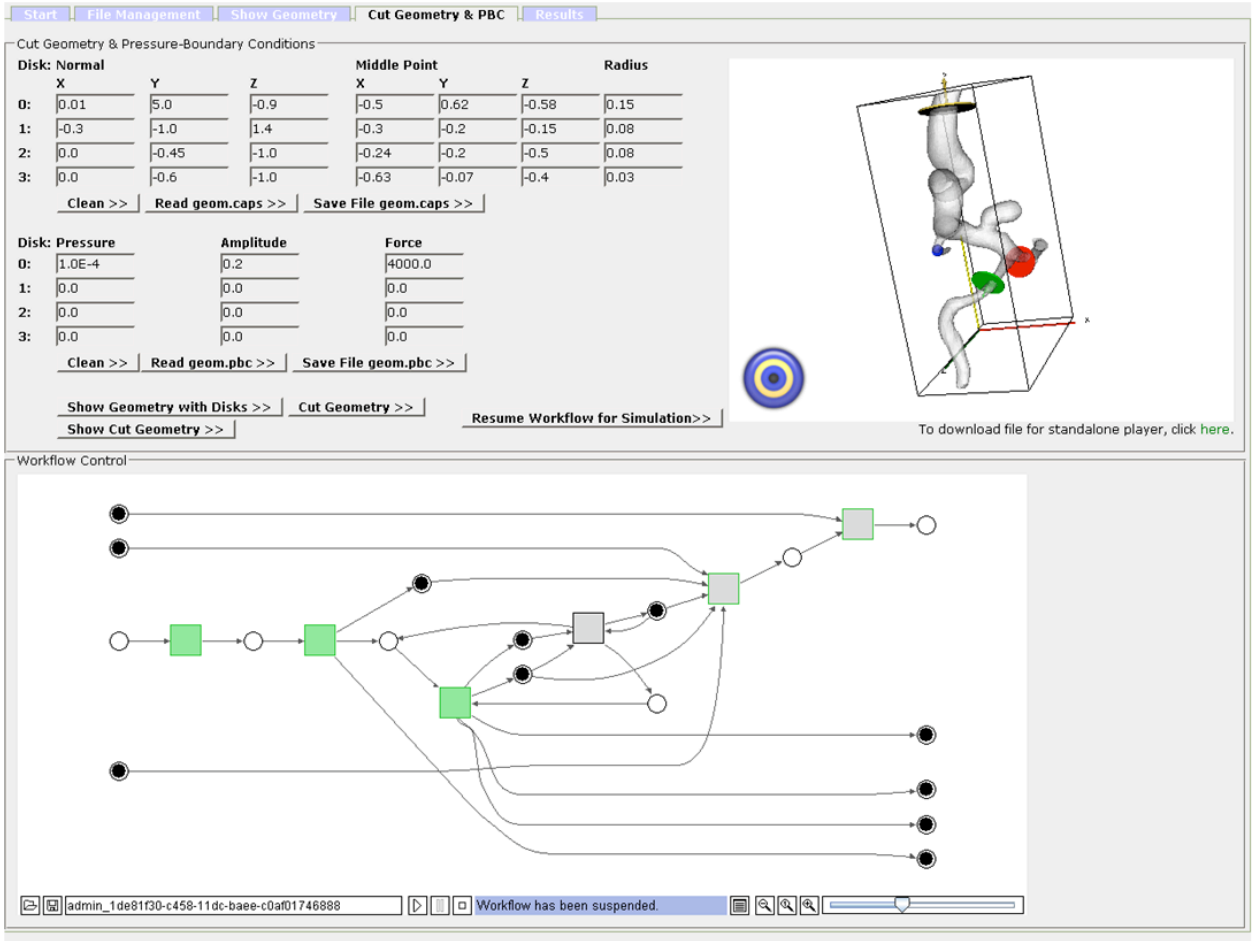


Abbildung 3. Gridportlet mit Eingabe der Simulationsparameter, interaktiver 3D-Visualisierung und Statusanzeige des Workflows

Da auch die Bestimmung des Simulationsgebietes ein interaktiver Schritt ist, der ggf. Nachkorrekturen verlangt, wurde ein mehrstufiger Workflow definiert, der die Wiederholung einzelner Schritte über die Nutzeroberfläche ermöglicht. Die Anwendung wurde ebenfalls in das MediGRID-Portal integriert. Die vorsegmentierten Daten können wieder aus dem Bildarchiv oder dem Datenmanagement ausgewählt und der automatische Workflow gestartet werden. Dieser berechnet aus der gegebenen Geometrie eine VRML-Repräsentation, die mithilfe eines entsprechenden Browserplugins eine interaktive passive 3D-Visualisierung ermöglicht, so dass die Geometrie von allen Seiten betrachtet werden kann. Dies ist ein notwendiger Schritt zur Festlegung der Simulationsgrenzen, welche im folgenden Schritt entweder aus einem Datenfile geladen oder manuell eingegeben werden können. Die gewählten Simulationsgrenzen werden ebenfalls in VRML dargestellt, so dass diese kontrolliert und gegebenenfalls verändert werden können. Dann können noch physiologische Größen wie Blutdruck und wirkende Kraft an den Simulationsgrenzen definiert werden, bevor die Simulation gestartet wird. Ist diese abgeschlossen, kann die Druck- und Geschwindigkeitsverteilung des Blutes, ebenfalls als VRML-Objekt, über das Portal visualisiert werden. Die Ergebnisse werden im Datenmanagement archiviert und können jederzeit herunter geladen werden. Die Implementierung ermöglicht Nutzern auch ohne weitere IT-Kenntnisse und ohne entsprechende Rechenressourcen die Durchführung von virtueller Operationsplanung in der Gefäßchirurgie und ist somit ein Prototyp für medizinische Grid-Dienste.

2.1.2.3 3D-Ultraschall

Ebenso wie kardiovaskuläre und neurologische Krankheiten nimmt auch der Prostatakrebs in der Entwicklung von alternden Gesellschaften stark zu. Die implementierte Anwendung umfasst verschiedene Teilaspekte der Prostatakrebsdiagnostik. Zum einen kann anhand von physiologischen Daten und Laborwerten das Prostatakrebsrisiko mithilfe einer Klassifikation abgeschätzt werden. Die Bestimmung des Risikos kann den Arzt bei der Entscheidung für den Einsatz weiterer Diagnoseverfahren, wie die Prostatabiopsie, unterstützen. Des weiteren kann ein Image Retrieval durchgeführt werden, mit dem ein Ultraschallbild der Prostata mit bereits in der Datenbank gespeicherten Aufnahmen verglichen wird und ähnliche Bilder zurückgegeben werden. Ist aufgrund der Datenlage eine Prostatabiopsie angezeigt, so können die bei der Biopsie zur Kontrolle aufgezeichneten Ultraschallbilder dazu verwendet werden, eine dreidimensionale Rekonstruktion aller im Rahmen der Untersuchung durchgeführter Gewebeproben zu erstellen und zu visualisieren. Dazu sind verschiedene Segmentierungs- und Registrierungsschritte notwendig, die von dem Nutzer ausgewählt werden können. Das Methodenspektrum dieser Applikation umfasst wesentliche Komponenten typischer Bildverarbeitungspipelines, wie Segmentierung, Registrierung, Klassifikation und Image Retrieval. Die Klassifikation und die Segmentierung der Biopsienadeln ist in Matlab entwickelt worden. Im Gegensatz zur funktionellen Hirnbildgebung ist der Code jedoch kompilierbar, so dass die Lizenzproblematik entfällt. Das Image-Retrieval, angeboten von der RWTH Aachen, wurde über einen Webservice angebunden. Die Registrierung, der rechenintensivste Schritt der Applikation, wurde auf Basis der frei erhältlichen c++ Softwarebibliothek ITK in das Grid implementiert. Die Nutzeroberfläche erlaubt eine Reihe von individuellen Anpassungen des Registrierungsalgorithmus, ebenso wie Parameterscans, die für die robuste Durchführung der Registrierung notwendig sind.

Segmentation | Registration | Classification | IRMA

Locates a 2D monitor image within a previously taken 3D Volume.

More info

p_RItn.dat
pipe.stl
prs_results_40000.wrl
prs_results_50000.wrl
seriesset_51123_1100_AL.tgz
seriesset_51123_1100_AR.tgz
seriesset_51123_1100_BL.tgz
seriesset_51123_1100_BR.tgz
seriesset_51123_1100_ML.tgz
seriesset_51123_1100_MR.tgz

StudySet — StudySet: studyset_51123_1100.tgz

SeriesSet — SeriesSet: seriesset_51123_1100_PR.tgz

Search

Config

Parameter

Advanced...

ImageSet:	Reg_51123_105200		
FirstImage:	Reg_51123_105200_sec_000		
VolumeImage:	3D_20051123_1100_shrink.ml		
VolumeMask:	3D_51123_1100_mask.mhd	SecondImage:	51123_105200_sec_000_top.r
FirstMask:	2dmask_bottom_new.mhd	SecondMask:	2dmask_top_new.mhd
SplineOrder:	3	Level:	1
StepLength:	0.5	Iteration:	1000
TranslateScale:		RotateScale:	
TranslateSample:		RotateSample:	
TranslateLimit:		RotateLimit:	
InitialTransform:	51123_105200_transform_bot	RatioTransform:	51123_105200_topvsbottom_r
Fixed Derivation		Moving Derivation	
SimplexSize		SampleNum	

Write Images: (•) All ( ) checkerboard ( ) diff ( ) substract ( ) none

Variation: ( ) On ( ) Off

Metric: (•) Norm Corr ( ) Mean Square ( ) MutualInfomr

Optimizer: (•) Amobea ( ) Gradient

Run

Start Workflow

Message: Workflow submitted, check the workflow tab for results

Status: RUNNING

Abbildung 4. Portlet zur Registrierung mit ITK. Für die Ultraschallbildgebung sind geeignete Parameterwerte vorgewählt. Beliebige 3D-Daten können registriert werden, wobei eine Vielzahl von Optionen möglich ist

Als wesentlich für die medizinische Bildverarbeitung hat sich ein hohes Maß an Datensicherheit und Nutzerfreundlichkeit herausgestellt. Auch die zunehmend rigideren Sicherheitsbestimmungen in klinischen IT-Netzen mussten berücksichtigt werden. Deshalb wurden alle Anwendungen in das webbasierte MediGRID-Portal integriert, so dass ein intuitiver und rein webbasierter Zugang zu allen drei Grid-Anwendungen existiert, die von allen MediGRID-Mitgliedern genutzt werden können. Die von der Grid-Middleware Globus unterstützte zertifikatsbasierte AAI-Infrastruktur wurde ebenfalls in allen drei Anwendungsszenarien durchgängig implementiert. Über die konkrete Anwendungsentwicklung hinaus, wurden im Rahmen des Moduls Bildverarbeitung verschiedene generische Methoden entwickelt und implementiert, die ohne oder mit geringen Anpassungen in vergleichbaren Bildverarbeitungsproblemen eingesetzt werden können

2.1.2.4 Sichere Übertragung medizinischer Bilder

Im Rahmen des Projektes wurden Methoden zur sicheren und fehlertoleranten DICOM-Übertragung entwickelt und implementiert, die sowohl die Pseudonymisierung und Speicherung im halböffentlichen Bereich des Kliniknetzes (demilitarisierte Zone) als auch die zertifikatsbasierte fehlertolerante DICOM-Übertragung umfasst.

Die Anbindung an das MediGRID erfolgt über ein speziell für diesen Zweck aufgesetztes PACS (dcm4chee), das in der demilitarisierten Zone der Charité steht und über ein speziell entwickeltes Protokoll (griddicom) zertifikatsbasierten gesicherten DICOM-Transfer zwischen dem Grid-PACS und dem MediGRID erlaubt. Innerhalb der Charité besteht eine Standard-DICOM Verbindung über einen Gateway-Rechner im Intranet.

Die für die Datenverarbeitung im Grid vorgesehenen DICOM-Daten werden mit einem vom Offis-Institut in Oldenburg entwickelten Programm verschlüsselt. Dabei werden die patientenbezogenen Metadaten im DICOM-Header durch DICOM-konforme Zeichenketten ausgetauscht, die keine patientenbezogenen Informationen enthalten. Die ursprünglichen patientenbezogenen Daten werden verschlüsselt in so genannten „Private Tags“ gespeichert. Im Gegensatz zu den anderen Pseudonymisierungskonzepten, die die Speicherung der verschlüsselten Daten in einem Block realisieren, ermöglicht das im Rahmen des MediGRID entwickelte Pseudonymisierungstool bei der Entschlüsselung die Zuordnung der abgeleiteten Bilder zu den Originalbildern, was für die Integrität der Bildverarbeitungskette von großer Bedeutung ist.

Die pseudonymisierten Daten sind in der DMZ der Charité auf einem PACS gespeichert.

Dieses Konzept ergänzt die von den Modulen „Klinische Forschung“ und „eScience“ entwickelten Sicherheitskonzepte für die Verarbeitung von klinischen Biosignaldaten und epidemiologischen Daten um medizinisches Bildmaterial.

Abbildung 5 stellt die Systeme und den Datenfluss beim MediGRID-Anschluss an der Charité dar: Das interne PACS (1) steht in der internen Netzwerkzone, das externe PACS (2) in der DMZ. Bilder können nur manuell vom Arbeitsplatzrechner (5) über den Router (6) verschoben werden. Der Zugriff auf die Anwendungen im Grid erfolgt mit dem Webbrowser über den Charité-Proxy (8). Gridrechner können auf das externe PACS nur über den GridDICOM-Router (7) zugreifen. Autorisierte Nutzer können über das Gridportal die Daten anwählen und weiterverarbeiten.

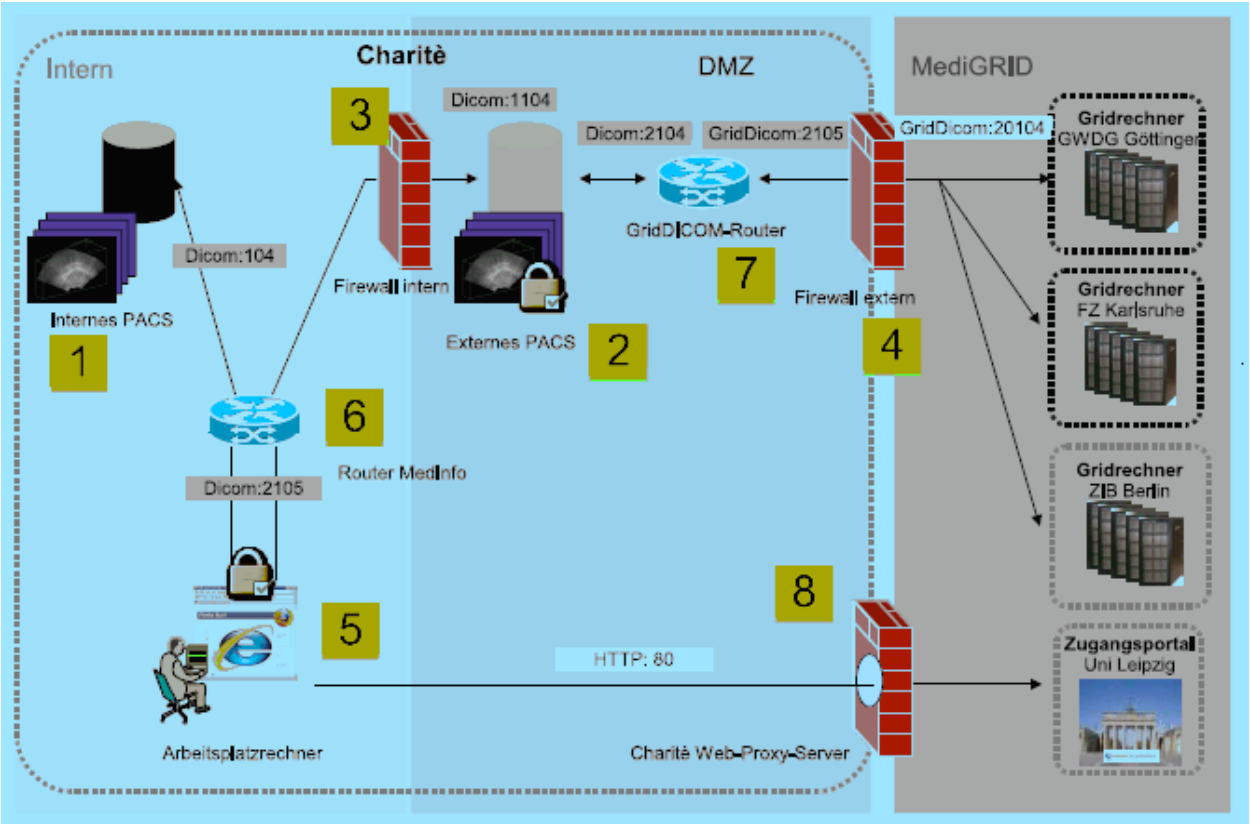


Abbildung 5. Systeme und Datenfluss beim MediGRID-Anschluss an der Charité

Im Falle der Weiterverarbeitung werden die Daten auf einen verfügbaren Grid-Rechner übertragen. Zur Anbindung des Datentransfers an die Middleware wurde analog zum Protokoll gridFTP ein grid-DICOM Protokoll entwickelt, welches das Globus Security Interface (GSI) verwendet. Durch Entwicklung dieser Gridvariante des DICOM-Standards (gDicom) können nun zum einen die Vorteile eines integrierten DICOM-Netzwerkes auch im Grid mit voller Unterstützung der Sicherheitsinfrastruktur genutzt werden. Das gridDICOM Protokoll ist als Erweiterung des DICOM-Standards formuliert und ermöglicht nun die Implementierung in DICOM-Geräte. Eine Referenzimplementierung ist öffentlich verfügbar. Zur Sicherung der Kompatibilität von DICOM-Geräten wie PACS (Picture Archive Systems) und Bildmodalitäten (CT, MRT), die das gridDICOM Protokoll nicht implementiert haben, wurde ein entsprechender Router entwickelt, der das Protokoll entsprechend umsetzt ohne die Daten zwischenzuspeichern. Diese Maßnahme soll Sicherheitsrisiken minimieren.

Da neben der reinen Datensicherheit auch Aspekte der Verfügbarkeit und Fehlertoleranz eine wichtige Rolle in der Gridimplementierung spielen, wurden fehlertolerante Transfermechanismen entwickelt und implementiert, die kurzfristige Netzausfälle abfangen können und bei der Übertragung von Bildserien einzelne, fehlerhaft übertragene Bilder erneut anfordern. Der in Analogie zum RFT (reliable File Transfer) benannte RDT-Dienst (reliable Dicom Transfer) basiert auf dem im MediGRID eingesetzten und vom Fraunhofer-Institut für Rechnerarchitektur und Softwaretechnik (FIRST) entwickelten Workflowmanager GWES und steht im MediGRID zur Verfügung.

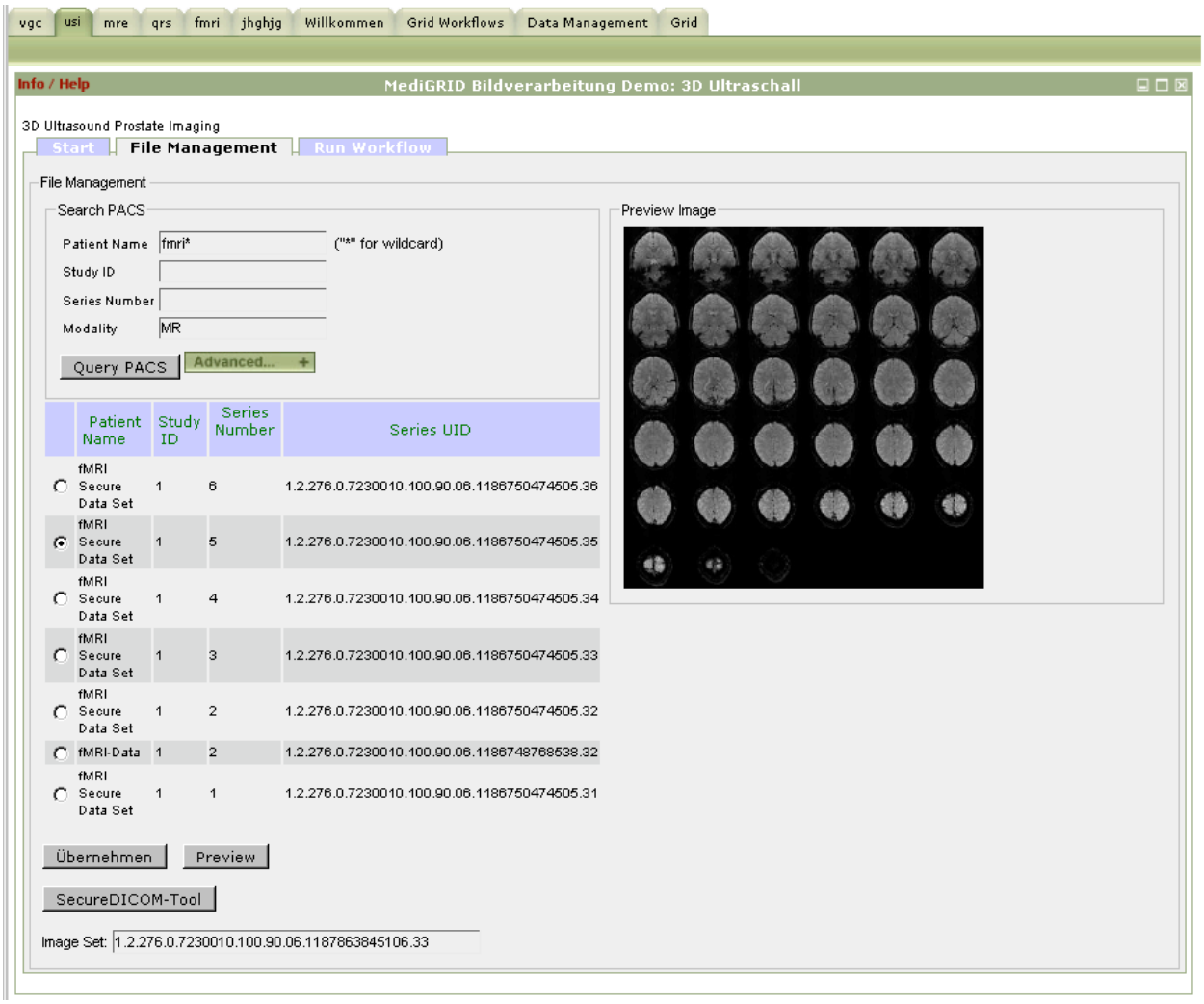


Abbildung 6. PACS-Komponente für den Zugriff über das Gridportal, hier in der fMRI-Anwendung. Die Patientennamen wurden durch Pseudonyme ersetzt

2.1.2.5 Portletkomponenten

Im Rahmen der Integration der Anwendungsszenarien in das MediGRID Portal wurden eine Reihe von generischen Portletkomponenten entwickelt, die die Entwicklung von anwendungsspezifischen Portlets stark vereinfachen. Die Entwicklungen umfassen zum einen die Datenselektion und die Übergabe als Inputparameter an den automatischen Workflow. Für das Hochladen in das Homeverzeichnis auf einem Gridknoten wurde das von Gridsphere bereitgestellte Portlet um die Funktion, mehrere Bilder mit Preview und verschiedene Daten mit einem Befehl hochzuladen, erweitert. Wichtiger ist jedoch die Funktionalitätserweiterung, die die URL des ausgewählten Datensatzes übergibt und global verfügbar macht. Damit kann die URL in anderen Portlets weiter verwendet werden. Ebenso wurde eine Komponente zur Auswahl von Daten aus dem SRB entwickelt. Es existiert zwar ein von den Entwicklern des Datenmanagementsystems bereitgestelltes SRB-Portlet, der aktuell ausgewählte Datensatz kann jedoch nicht abgefragt werden. Deshalb war hier eine Eigenentwicklung notwendig. Zur Anbindung des Bildarchivs wurde eine PACS-Komponente entwickelt, mit der Bilder nach bestimmten Kriterien gesucht, im Preview angesehen und ausgewählt werden können.

Zudem wurde ein vereinheitlichtes Framework mit vordefinierten Komponenten für Titelseite, Filemanagement, Workflowdefinition und Ergebnisvisualisierung entwickelt, so dass eine modu-

lare Zusammenstellung des anwendungsspezifischen Portlets möglich ist, und die Neuentwicklung von Portlets stark vereinfacht.

In diesem Zusammenhang ist auch der Entwurf und die Referenzimplementierung eines Prototypen zum webbasierten Upload von myproxy-Zertifikaten zu nennen. Das Hochladen der myproxy-Zertifikate ist die einzige notwendige Aktion, die bisher nur mit erweitertem Netzzugang möglich war. Klinikmitarbeiter mussten bisher entweder eine besondere Firewallkonfiguration beantragen, oder haben diesen Schritt, der etwa einmal die Woche erfolgen muss, von zuhause aus durchgeführt. Basierend auf dem von Fraunhofer Institut für Arbeitswirtschaft und Organisation (IAO) entwickelten MediGRID-Myproxy-Upload-Tools wurde eine Applet/Servlet Lösung entwickelt, die den webbasierten Upload ermöglicht.

2.1.2.6 Konzepte und Leitfaden zur Gridimplementierung

Im Rahmen des Projektes konnte breites technisches und konzeptionelle Know-how bezüglich der Gridimplementierung von medizinischen Bildverarbeitungsmethoden erworben werden, das bereits zur Projektlaufzeit zur Implementierung von weiteren Bildverarbeitungsanwendungen geführt hat. Für die Implementierung von Anwendungen in das MediGRID wurde ein Leitfaden entwickelt, mit dem mit geringem Aufwand (wenige Stunden) neue Anwendungen auf das Grid gebracht und sofort genutzt werden können. Dadurch können potentielle neue Nutzer ohne großen Entwicklungsaufwand mit einer Basisimplementierung die Vorteile des Grids gegenüber alternativen Implementierungen testen, und so an das Grid-Computing herangeführt werden. Dies war insbesondere möglich durch die enge Zusammenarbeit mit den Modulen Ressourcenfusion und Middleware, die entsprechende Middlwarekomponenten entwickelten und bereitstellten. Von besonderer Bedeutung ist der Workflowmanager GWES zu nennen, der ein hohes Maß an Virtualisierung erlaubt und damit die Gridimplementierung sehr stark vereinfacht.

2.1.3 Nutzen und Verwertbarkeit der Modulergebnisse

Betrachtet man existierende und abgeschlossene Healthgrid-Projekte, so findet man unter ihnen einen starken Anteil an Bildverarbeitungsprojekten, z.B. das europäische „Mammogrid“ oder die „Medical Imaging“ workbench des „caBIG“ Projektes. Auch der starke Anteil von Bildverarbeitungsaspekten in den MediGRID Nachfolgeprojekten „MedInfoGrid“, Services@Medigrid, sowie dem als förderungswürdig bewerteten „PneumoGRID“ zeigt, dass die Bildverarbeitung als eines der zukunftsträchtigen Bereiche der medizinischen Gridanwendungen gilt. Die rechen- und speicheraufwändige Verarbeitung von Volumendaten und dynamischer Bildgebung und die einfach zu implementierende Parallelisierung von zumindest Teilschritten auf Bildserien-, Bild- und Teilbildebene führt dazu, dass digitale Bildverarbeitung sehr stark von Grid-Computing profitieren kann. Auch das Datenmanagement von großen Datenmengen, wie sie in der medizinischen Bildverarbeitung vorhanden sind, wird durch das Grid vereinfacht. Die im Rahmen des Moduls Bildverarbeitung im MediGRID Ergebnisse ermöglichen zum einen die sichere, nutzerfreundliche und effiziente Nutzung der implementierten Bildverarbeitungsanwendungen. Zum anderen weisen sie über die konkreten Anwendungen hinaus und stellt der Community generische Methoden und Dienste zur Verfügung, die die Vorteile des Grid Computings auch für eine Vielzahl ähnlich gelagerter Forschungs- und Entwicklungsaufgaben im Bereich der medizinischen Bildverarbeitung nutzbar macht. Im Einzelnen sind es:

- Funktionelle Hirnbildgebung: Das implementierte Tool SPM ist eines der meistgenutzten Softwarepakete für die funktionelle Hirnbildgebung. Weitere, über die assoziierten Partner hinaus gehende, auch internationale Forschergruppen haben ein starkes Interesse an einer Gridnutzung zur Speicherung und Verarbeitung von funktionellen Hirnbilddaten bekundet. Zur Zeit ist im MediGRID jedoch aufgrund der Lizenzproblematik mit der darunterliegenden Software „matlab“ eine starke Skalierbarkeit nicht gegeben. Es bestehen jedoch Kontakte zu einer Arbeitsgruppe in Amsterdam, die eine kompilierte Version von

SPM erstellt. Des weiteren ist auch das oft alternativ verwendete Softwarepaket „FSL“ bereits im MediGRID implementiert, und erste Anwendungen der fMRI in Zusammenarbeit mit der niederländischen biomedizinischen Gridinitiative „VLEMED“ integriert worden. Beide Softwarepakete sind frei, benötigen jedoch eine spezielle Umgebung (Matlab-Lizenz, bzw. Linux-OS), haben ein hohes Datenaufkommen und hohen Rechenbedarf. Angesichts der schwierigen Finanzlage vieler universitärer Einrichtungen, die sich insbesondere auch auf die Investition in moderne Rechenanlagen auswirkt, wird für diese Anwendung, die sich im Wesentlichen auf die universitären Forschung bezieht, ein hohes Nutzungspotential gesehen. Da funktionelle Bildgebung auch zunehmend in der klinischen Praxis zur Erkennung und Differenzierung neurologischer Defekte eingesetzt wird, können die im Grid angebotenen Dienste auch für die Verwendung im Rahmen der Versorgung erweitert werden.

- Virtuelle Gefäßchirurgie: Auch die virtuelle Gefäßchirurgie ist ein zukunftsweisendes Gebiet, das neben Forschungsaspekten auch eine klinische Nutzung des entwickelten Dienstes anvisiert. Für die virtuelle Gefäßchirurgie wurde insbesondere der rechenintensive Teil – die hämodynamische Simulation – in das Grid implementiert. Des weiteren wurden neue Methoden zur interaktiven 3D-Visualisierung der Gefäßgeometrie sowie der Simulationsergebnisse über das MediGRID-Portal entwickelt. Die Veränderung der Gefäßgeometrie z.B. durch virtuelles Einsetzen von Stents erfordert eine sehr starke Nutzerinteraktion. Eine Gridimplementierung hat sich hier mit den bisherigen Ressourcen als nicht effizient erwiesen, weshalb diese Schritte lokal durchgeführt wurden. Die Erweiterung der D-Grid Ressourcen durch die Sonderinvestitionen 2008 um vier Visualisierungscluster ermöglicht aber nun eine voraussichtlich einfache Implementierung aller Schritte der Prozesspipeline in das MediGRID, die die Nutzerfreundlichkeit deutlich erhöht. Dadurch können dann Dienste zur virtuellen Gefäßchirurgie auf unterschiedlichen Qualitätsstufen angeboten werden, die von der gesamten Bearbeitung im Grid inklusive der Nutzung lizenzpflichtiger Software (AMIRA) bis zur ausschließlichen Nutzung der hämodynamischen Simulation reichen.
- 3D-Ultraschall: Die entwickelte Gridanwendung ermöglicht die Klassifikation des Prostatarisikos anhand verschiedener Kennzahlen, und Lokalisation von Gewebeproben im Prostatavolumen. Der Qualitätskontrolle und Qualitätssicherung von diagnostischen Untersuchungen, wie sie in dieser Anwendung für die Prostatabiopsie entwickelt wurde, kommt eine immer größere Bedeutung zu. Während die automatisierte Segmentierung ein spezialisiertes Verfahren voraussetzt, ist die Bildregistrierung weitgehend generisch, basierend auf den freien Programmbibliotheken ITK und VTK, implementiert worden und ist so sehr flexibel und erweiterbar einsetzbar. Ein Problem für die Skalierbarkeit der Anwendung sind die wenig standardisierten Ultraschallaufnahmen, die sich sowohl von Hersteller zu Hersteller, als auch von Schallkopf zu Schallkopf stark unterscheiden können. Im Moment kann das Verfahren deshalb nur auf die für das Forschungsprojekt verwendeten Geräte angewandt werden. Eine Erweiterung der Anwendung auf weitere gängige Schallwandler ist anvisiert. Auch für diese Anwendung bietet der Visualisierungscluster durch eine entsprechende Weiterführung ein hohes Potential an zusätzlicher Funktionalität wie virtuelles Training, Biopsie- und Operationsplanung, die alle ein hohes Maß an Interaktivität voraussetzen.
- secureDICOM: Das im Rahmen des MediGRID entwickelte Pseudonymisierungstool steht im MediGRID-Portal für den Datentransfer zwischen zwei DICOM-kompatiblen Geräten zur Verfügung. Diese müssen allerdings eine entsprechende Verbindung zulassen, so dass DICOM-Geräte im internen Kliniknetz nicht angesprochen werden können. Eine Java-Applikation ermöglicht das lokale Verschlüsseln von im Dateisystem abgelegten DICOM-Daten und die gleichzeitige Übertragung in ein PACS. Das Tool wird intern

bereits verwendet und liegt dem behördlichen Datenschutzbeauftragten zur Begutachtung vor. Das Tool wird im Rahmen des MedInfoGrid-Projektes weiterentwickelt, um die Authentifizierung über den Heilberufeausweis zu ermöglichen. Es soll auch im PneumoGRID eingesetzt werden.

- gridDICOM: Die für die sichere DICOM-Übertragung in Grids entwickelte Erweiterung des DICOM Standards ermöglicht nun die DICOM-Kommunikation innerhalb von Grids und – durch die Bereitstellung eines entsprechenden Routers – die nahtlose Anbindung von DICOM-kompatiblen Geräten, insbesondere Bildarchive (PACS). Zusammen mit dem Datenschutzkonzept für die Bereitstellung von Daten zur Verarbeitung im Grid sind die Entwicklungen die technische Voraussetzung für das sichere Datenmanagement von klinischen Bilddaten. gridDICOM soll im Rahmen des PneumoGRID um Auditing-Funktionalitäten erweitert werden, um es auch für klinische Studien einsetzen zu können. gridDICOM und die Referenzimplementierung ist unter GPL frei verfügbar. Die Implementierung von fehlertoleranter Datenübertragung basiert auf dem vom Projektpartner FIRST entwickelten Workflowmanager GWES, der für die universitäre Forschung frei verfügbar ist.
- Portletkomponenten: Der webbasierte Zugang zu den Gridressourcen hat sich als wesentlicher Faktor in der Nutzbarkeit von Healthgrids für medizinische Forschung und in noch stärkerem Maße für die klinische Forschung herausgestellt. Die entwickelten Portletkomponenten werden bereits für die Entwicklung weiterer nutzerfreundlicher Gridanwendungen aus dem Bereich der Biosignal- und Bildverarbeitung eingesetzt. Der Sourcecode der entwickelten Portlets wird in einem Versionsmanagement gepflegt und wird auf Anfrage auch anderen Arbeitsgruppen zur Verfügung gestellt. Die webbasierte Myproxy-Upload-Lösung soll in Zusammenarbeit mit dem Fraunhofer IAO um weitere Funktionalitäten erweitert und in das MediGRID-Portal integriert werden.
- Konzepte und Leitfaden: Das im Rahmen des Projektes erworbene Know-How ermöglicht die Beratung und Unterstützung von potentiellen Gridnutzern aus dem Bereich der digitalen Bild- und Signalverarbeitung. Durch die Implementierung der weitestverbreiteten Tools und Softwarepakete in diesem Bereich ist die Präsentation einer Referenzimplementierung möglich, und viele Schwierigkeiten und Besonderheiten können bereits in der Planungsphase berücksichtigt werden.

2.1.4 Fortschritte auf dem Markt für Grid Computing in der Medizin und der Bioinformatik aus Modulsicht - Darstellung Ergebnisse Dritter

Zunehmend werden Gridtechnologien für die biomedizinische Forschung eingesetzt. Dies äußert sich unter anderem in zahlreichen öffentlich geförderten Gridprojekten aus dem Life-Science-Bereich, z.B. caBIG oder medizinische Projekte in EGEE. Obwohl der potentielle Nutzen von Grids auch für die klinische Forschung und die Versorgung außer Frage steht und den so genannten Healthgrids eine wesentliche Bedeutung im zukünftigen Gesundheitswesen zugesprochen wird, hat keines der Projekte bisher das Stadium der Grundlagenforschung verlassen. Dies liegt vor allem an den erhöhten Ansprüchen von Healthgrids an Datensicherheit und Nutzerfreundlichkeit, die teilweise scheinbar im Widerspruch zu dem offenen Ansatz von Gridinfrastrukturen stehen. Mit der generellen Ausweitung von Grids werden aber zunehmend auch von anderen Nutzergruppen ähnliche Anforderungen an die Gridtechnologie gestellt, so dass aktuelle Entwicklungen in der Gridforschung die gestiegenen Anforderungen an Sicherheit, Stabilität und Fehlertoleranz berücksichtigen müssen. In diesen Zusammenhang können auch zahlreiche Entwicklungen im Rahmen des DGI/DGI-2-Projektes sowie einzelner Community Grids eingeordnet werden (s.u.). Der derzeitige Stand der Forschung erlaubt einerseits bereits die Verarbeitung

von pseudonymisierten Patientendaten im Grid; die fehlende lückenlose Dokumentation der Datenverarbeitung (Audit-Trail) verhindert aber andererseits noch die Nutzung des Grid für klinische Studien. Patientendaten können aufgrund von Datenschutzbestimmungen bisher nicht verarbeitet werden.

Abgesehen von den biomedizinischen Communityprojekten gibt es eine Reihe von generischen Entwicklungen in den anderen D-Grid-Projekten, die auch für die Gridnutzung für klinische Anwendungen von Bedeutung sind. So haben die Projekte mit Industriebeteiligung (z.B. InGRID, BauVOGrid) besondere – über den derzeitigen Standard hinausgehende – Sicherheitsinteressen und in einigen Aspekten ähnliche Nutzerprofile (z.B. Einsatz von Windows-Rechnern als Endgeräte). In Astrogrid werden Metadatendienste und Fehlertoleranzstrategien entwickelt. Im Integrationsprojekt DGI/DGI-2 werden Konzepte und Werkzeuge zum Accounting und Billing entwickelt und die aktuelle Förderphase (3. Call) berücksichtigt insbesondere Forschung im Bereich Service-Level-Agreements und Wissenstechnologien. Der Vergleich und Austausch mit internationalen Healthgrid-Projekten zeigt, dass trotz verschiedener rechtlicher Ausgangslage, anderer medizinischer Fachrichtung und divergierender Infrastruktur oft ähnliche Probleme bezüglich Sicherheit und Nutzerfreundlichkeit bestehen. Grade die Pionierprojekte in diesem Bereich (z.B. das französische MediGRID und Mammogrid) waren vor allem technischer Natur und auf Grundlagenforschung ausgelegt; aufgrund der Besonderheiten der medizinischen Nutzercommunity haben sie – wie bereits oben erwähnt – die Produktionsreife nicht erreicht. Auf der anderen Seite herrscht durch die Heterogenität der Projekte ein reicher Erfahrungsschatz über die Einsatzmöglichkeiten der existierenden Middlewarekomponenten im medizinischen Bereich. Am weitesten fortgeschritten scheint bezüglich der Integration von Kliniken das europäische Health-e-Child Projekt zu sein. Dieses entwickelt eine gridbasierte Informationsplattform für pädiatrische Krankheiten und hat in diesem Rahmen eine standardisierte, auf einem Gateway basierende Lösung zur Anbindung von medizinischen Einrichtungen an das Grid realisiert. Auch Businessmodelle spielen in der internationalen Healthgrid-Community eine zunehmende Rolle, wie zum Beispiel die Aufnahme eines medizinischen Projektes in die zweite Runde des BEinGRID zeigt. Eine weitere wichtige Entwicklung ist die zunehmende Berücksichtigung von Interoperabilität zwischen den verschiedenen Gridinfrastrukturen. Die gemeinsame Nutzung von Diensten und Ressourcen von zum Beispiel EGEE und MediGRID würden die Möglichkeiten der kollaborativen Forschung und Entwicklung stark erweitern.

2.1.5 Ausblick und Perspektiven im Bereich Medizinische Bildverarbeitung

Die derzeitigen Entwicklungen in der medizinischen Bildgebung hin zu dynamischer Volumenbildgebung (4D) und gleichzeitiger höherer räumlicher und zeitlicher Auflösung werden auch in Zukunft hohe Anforderungen an Datenspeicherung und –verarbeitung stellen, die mit Gridinfrastrukturen effizient gelöst werden können. Die Herausforderungen liegen dabei zur Zeit weniger in der Implementierung der zugrunde liegenden Verarbeitungsschritte, sondern in der sicheren und nutzerfreundlichen Bereitstellung der entsprechenden Dienste. Die laufenden und bewilligten Nachfolgeprojekte Services@Medigrid, MedInfoGrid und PneumoGRID weisen alle in diese Richtung. Sollte durch Integration des Heilberufeausweises in die Grid-Sicherheitsinfrastruktur der Gridzugang für Mediziner stark vereinfacht werden, könnte das Grid zu einer zentralen Plattform für die Bereitstellung und Nutzung von Diensten zur medizinischen Bild- und Signalverarbeitung werden.

2.2 Biomedizinische Informatik

2.2.1 Moduleinordnung und Modulziele

Aktuelle genomische, genetische und postgenomische Hochdurchsatzverfahren erzeugen eine vorher nicht gekannte Datenflut in den biologisch / medizinischen Wissenschaften. Die NCBI-Datenbanken enthalten beispielsweise über eine Billion Sequenztraces, die für öffentliche Analysen zur Verfügung stehen. Die in diesem Projekt dargestellten Anwendungen bildeten selbstverständlich nur einen kleinen Teil der bioinformatischen Aktivitäten in Deutschland ab. Es wurde jedoch eine offene Struktur geschaffen, die entscheidende Voraussetzung für die Akzeptanz des Grids in der bioinformatischen Nutzergemeinschaft ist und spätere spezifische Zugänge für Grid-Anwendungen erlaubt.

2.2.2 Modulergebnisse

Wichtigste Ergebnisse von Interesse für das Gesamtprojekt sind folgende:

- Nachweis der erfolgreichen Implementation von Anwendungen in die GRID-Umgebung (Göttingen, Kiel, Dresden)
- Praktische Anwendung der im MediGRID entwickelten Programme und Verfahren in vernetzten genomischen Projekten (als ein Beispiel siehe die Publikationen in ***Nature Genetics*** mit der Entdeckung eines neuen Krankheitsgenes für das Gallensteinleiden – die gemeinsam durch DGRID und das NGFN erfolgte)

Die Projekte werden unten in ihren Einzelprojekten in größerem Detail dargestellt.

2.2.2.1 Krankheitsorientierte Genomdaten (Kiel)

Ziel des Teilprojekts war die Einbindung der in Kiel entwickelten Verfahren zur Analyse von krankheitsorientierten Genomdaten ins Grid. Dabei standen die Anwendungen SNPSELECTION zur Auswahl von Einzelbasenpolymorphismen für Krankheitskartierungen und GENOMIZER zur Analyse von genomweiten Assoziationsstudien im Mittelpunkt. GENOMIZER (als Standalone Anwendung verfügbar) war dabei die international erste „Public Domain“ Software für derartige Analysen. SNPSELECTION läuft als Webservice im Portal von MediGRID. Eine Publikation mit der Beschreibung der Methodologie ist zur Publikation eingereicht. Im Einzelnen wurden hier folgende Arbeitsschritte durchgeführt:

2.2.2.2 Sequenzanalyse (Göttingen)

Ziel des Teilprojekts war die Gridifizierung der in der Abteilung entwickelten Softwareprogramme AUGUSTUS und DIALIGN. Innerhalb der ersten Förderperiode ist der Schwerpunkt dabei geringfügig verschoben worden. Das Genvorhersage-Programm AUGUSTUS wird derzeit in zahlreichen Genomprojekten eingesetzt. Eine verbesserte Version des Programms liefert Ergebnisse von exzellenter Qualität, ist allerdings deutlich rechenintensiver als die ursprüngliche Version. Für die internationale Sichtbarkeit des Projekts erschien es daher vorteilhaft, den Schwerpunkt zunächst auf die Gridifizierung von AUGUSTUS zu setzen, und das Programm DIALIGN danach anzupassen, das dann als zweites implementiert wurde.

2.2.2.3 Proteininteraktionsanalyse (Dresden)

Die SCOPPI Protein interaction database wurde als Online-Service im Netz und als Web2.0 Service implementiert. Bei der Java-basierten Implementierung wurde besonders auf Portierbarkeit in die Grid-Umgebung geachtet. Anwendungen umfassen insbesondere die LIGSITE csc (ligand docking, BMC Structural biology, Huang et al. 2006) und SCOWLP Applikationen.

Prova ist eine Regel-basierte Java Skriptsprache um verteilte Rechenabläufe für Proteinmodellierung zu beschreiben. Anwendungen schlossen die Entwicklung von Datenintegrationsabläufen („Workflows“) und die Abfrage von Ontologien ein. Damit konnte ebenfalls ein semantischer Grid-Browser erstellt werden, der Proteindaten im Web zusammenführt. Die Java-Technologie ist gut auf die Grid-Umgebung portierbar. Prova ist ein Beitrag zu gemeinsamen Standardierungsbemühungen zum Beispiel mit CaBIG.

Die Ergebnisse des Projekts sind in den Publikationen dokumentiert.

2.2.2.4 Sequenzkorrelation (Heidelberg)

Hier wurden komplexe Interaktionen von Sequenzdaten und biologische Bilddaten aufgebaut – hier bestand eine enge Kooperation zur Bildverarbeitungsmodul des Projektes. Hier werden komplexe Interaktionen von Sequenzdaten und biologische Bilddaten aufgebaut – hier besteht eine enge Kooperation zur Bildverarbeitungsmodul des Projektes. Aufgrund der Komplexität der erwarteten Fragestellungen und der damit verbundenen Datenstrukturierungsmaßnahmen, der enormen Rechenanforderungen und der komplexen Interpretation der Analysedaten wurden vier Bereiche im Hinblick auf eine qualitativ hochwertige Korrelationsplattform implementiert; siehe Abbildung 7.

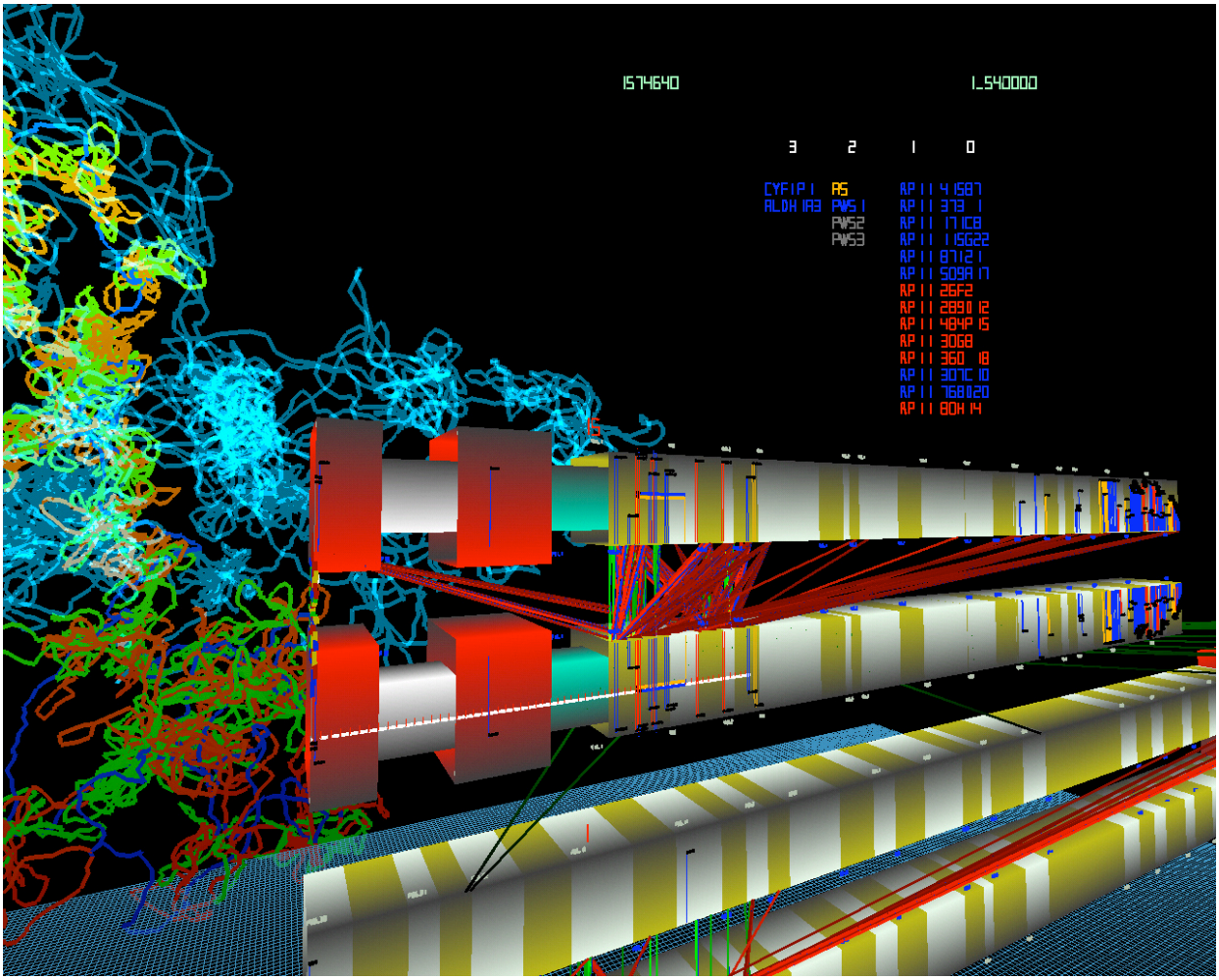


Abbildung 7. Arbeitsumgebung des GLOBE 3D Genome Viewers

Zu sehen sind Paare von Chromosome 15 (Mitte) und 1 (unten) sind in ihrer Metaphase-Repräsentation mit Ideogram Banden, sowie achrozentrischen und centromerischen Regionen

dargestellt. Auf die Chromosomen sind verschieden DNA Klonfragmente projiziert. Inter und intra-chromosomale Sequenzduplikationen sind durch Linien dargestellt (die Farben indizieren eine Positionsklassifizierung der Duplikate). Eine Interphase-Struktur-Simulation von Chromosome 15 basierend auf dem so genannten Multi-Loop-Subcompartment Model ist im Hintergrund dargestellt (die Farben repräsentieren verschieden Ideogramm Cluster). Die bedeutende Prader-Willi-Syndrom Region der Metaphase Repräsentation von Chromosome 15 (rechts neben der türgrünen Centromerregion) ist in der Simulation in hellem blau dargestellt.

2.2.2.5 Relevante Websites

Übergreifendes MediGRID-Portal mit Grid-Anwendungen: https://portal.medigrid.de

Projekt: Krankheitsbezogene Genomdaten (Kiel): www.snpselection.de

Projekt: Protein Interactions (Dresden): http://www.biotec.tu-dresden.de/schroeder

Projekt: Sequenz- und Promoteranalyse (Göttingen): http://gobics.de/department/software

2.2.3 Nutzen und Verwertbarkeit der Modulergebnisse

Die hier entwickelten Softwarelösungen werden weiter für die wissenschaftlichen Fragestellungen genutzt. Die Erfahrungen aus diesem Projekt fließen in die Arbeiten von Services@MediGrid und der Systembiologieprojekte der Antragsteller (QuantLiver) ein und haben hier die entscheidenden Grundlagen gelegt.

2.2.4 Fortschritte auf dem Markt für Grid Computing in der Medizin und der Bioinformatik aus Modulsicht - Darstellung Ergebnisse Dritter

In den hier bearbeiteten Anwendungsbereichen gibt es keine konkurrierenden Grid-Lösungen. Insofern hat das Projekt hier ein Alleinstellungsmerkmal.

2.2.5 Ausblick und Perspektiven im Bereich biomedizinische Informatik

Wir gehen davon aus, dass Grid-basierte Analyseverfahren für die neuen Herausforderungen der Bioinformatik von entscheidender Bedeutung sein werden.

2.3 Klinische Forschung

2.3.1 Moduleinordnung und Modulziele

Das Modul Klinische Forschung verfolgte das zentrale Ziel einer Analyse der prinzipiellen Gridfähigkeit der klinischen Forschung. Es wurden die Grundvoraussetzungen für eine Verbesserung der interdisziplinären Zusammenarbeit zwischen klinischer und grundlagenorientierter Forschung analysiert und formuliert sowie ein Datenschutzkonzept für Nutzung von klinischen Studiendaten im Grid entwickelt.

Das Modul unterhielt eine enge Zusammenarbeit mit allen weiteren Modulen von MediGRID, hier insbesondere dem Modul Bildverarbeitung und dem Modul eScience. Auch mit dem ebenfalls durch das BMBF geförderten Projekt InGRID fand eine Kooperation im Bereich der Algorithmenentwicklung statt.

Das Modulziel war die Definition und Implementierung eines Anwendungsszenarios „Schlafforschung“ zur Demonstration der Gridfähigkeit klinischer Forschung. Dies beinhaltete die Erstellung eines Pseudonymisierungskonzept für die Nutzung von Daten aus der klinischen Forschung sowie der Patientenversorgung im Grid. Dieses Konzept repräsentiert gleichzeitig das Deliverable 1 dieses Moduls. Der erste Meilenstein des Moduls ist die Realisierung der datenschutzrechtlichen Voraussetzungen (siehe Pseudonymisierungskonzept D1) im ersten konkreten Fall. Als zweiter Meilenstein ergibt sich daraus die Realisierung einer Gridanwendung für die klinische Forschung im ersten konkreten Fall. Beide Meilensteine wurden erreicht, das Deliverable 1 wurde fertig gestellt. Publikationen dazu wurden erarbeitet und erfolgreich veröffentlicht (siehe Liste der Veröffentlichungen).

2.3.2 Modulergebnisse

2.3.2.1 Wissenschaftlich-technische Ergebnisse

Innerhalb des Moduls wurde gemäß Meilenstein 2 eine Gridanwendung für die klinische Forschung im ersten konkreten Fall realisiert. Dafür wurden innerhalb des Grid Daten von schlafmedizinischen (polysomnographischen) Registrierungen zur Verfügung gestellt. Diese stammen aus einer Datenbank von nächtlichen Registrierungen, die von Hand ausgewertet und mit Daten zur Erkrankung und zur Anthropometrie der Patienten/Probanden versehen wurden (sog. SIESTA-Datenbank). Darüber hinaus wurden aktuelle Registrierungen für weitere Tests der konkreten Implementation hinzu gefügt. Somit wurde durch das Modul Klinische Forschung die Basis für eine Referenzdatenbank von polysomnographischen Registrierungen geschaffen, die unter Beachtung der Datenschutzrichtlinien erweitert werden kann.

Eine enge Kooperation auf Anwendungsebene fand mit dem Modul Bildverarbeitung statt. Durch einen intensiven Austausch der Erfahrungen aus den beiden Modulen konnten hier Projektziele in beiden Modulen besonders effektiv erreicht werden, zum Beispiel die Nutzung des Workflowsystems und die Implementierung der Algorithmen, die Berücksichtigung der in der Universität Marburg erarbeiten Richtlinien zum Datenschutz und das Datenmanagement. Alle Registrierungen wurden zusammen mit dem Modul Bildverarbeitung auf dem SRB erfasst und mit den entsprechenden Metadaten versehen um eine Selektion von Aufzeichnungen mit bestimmten Eigenschaften (z.B. Alter der Patienten, Geschlecht o.ä.) für eine automatische Analyse durchführen zu können. Durch Mitarbeit des Moduls „Bildverarbeitung“ und in Zusammenarbeit mit den Querschnittsmodulen „Ressourcenfusion“ und „Middleware“ wurde die im Rahmen der europäischen schlafmedizinischen Multicenterstudie „SIESTA“ erstellte umfassende und Referenzdatenbank in das SRB-Datenmanagementsystem des MediGRID implementiert. Die anonymisierten physiologischen Parameter wurden als nutzerdefinierte Metada-

ten in den SRB eingepflegt, so dass die Daten nun nach verschiedenen Kriterien, wie Alter oder Hauptdiagnose, durchsucht werden können.

Für das Anwendungsszenario „Schlafforschung“ wurde als Ziel definiert, eine Grid basierte Analyseplattform für nächtliche Biosignalaufzeichnungen (polysomnographische Aufzeichnungen) zu entwickeln. Zwar sind die anfallenden Datenmengen mit ca. 100 MByte bis 200 MByte je nach gewählter Abtastrate pro Aufzeichnung noch recht überschaubar, jedoch stellt die automatisierte Analyse komplexer Biosignale eine rechenintensive Aufgabe dar, die jedoch aufgrund der Datenstruktur der Aufzeichnungen (EDF-Format, und synchron aufgezeichnete unterschiedliche Biosignale wie Hirnströme, Augenbewegung, Muskeltätigkeit, Atmung, Herzschlag) leicht parallelisierbar ist. Somit stellt die automatische Analyse von Biosignalaufzeichnungen aus der schlafmedizinischen Forschung eine ideale Anwendung zur Beschreibung der prinzipiellen Gridfähigkeit der klinischen Forschung dar. Darüber hinaus kann die bereitgestellte Referenzdatenbank dazu genutzt werden, die Weiterentwicklung und Optimierung von automatischen Analyseverfahren zu beschleunigen und zu vereinfachen. Zum einen können die optimierten Analyseverfahren schnell und einfach eine große Zahl von Aufzeichnungen analysieren, zum anderen können die automatischen Ergebnisse leicht mit den hinterlegten Ergebnissen der visuellen Analyse (Goldstandard) verglichen werden. Auch ein Vergleich von zwei unterschiedlichen Analyseverfahren, die das gleiche Biosignal verwenden, kann so deutlich einfacher als bei einer rein lokalen Entwicklung durchgeführt werden.

Für polysomnographische Registrierungen werden unterschiedliche Biosignale im Laufe einer Nacht aufgezeichnet. Dies sind die Hirnströme EEG (Elektroenzephalogramm), Augenbewegungen EOG (Elektrookulogramm), Muskeltätigkeit EMG (Elektromyogramm), Herzschlag EKG (Elektrokardiogramm), Atmungsanstrengung (sog. respiratorische Induktionsplethysmographie) und transkutane Sauerstoffsättigung (Pulsoxymetrie). Eine visuelle Analyse der aufgezeichneten Signale durch schlafmedizinisch geschulte Experten erlaubt die Diagnostizierung von schlafmedizinischen Erkrankungen (z.B. schlafbezogene Atmungsstörungen oder Bewegungsstörungen). Die visuelle Analyse ist jedoch sehr aufwändig und zeitraubend (ca. 45-60 min pro Aufzeichnung). Somit wäre eine Vereinfachung der Diagnostik eine dringend benötigte und auch nachhaltige Entwicklung im Bereich der Schlafmedizin. Aufgrund der Häufigkeit der Erkrankung und auch der lokal vorhandenen Expertise beschränken sich die Entwicklungen im Modul Klinische Forschung aktuell auf die Vereinfachung und Optimierung der automatisierten Diagnostik schlafbezogener Atmungsstörungen (OSAS obstruktives Schlafapnoesyndrom). Darüber hinaus wurde in Kooperation mit der Klinik für Neurologie der Universitätsklinik Gießen und Marburg GmbH ein Analyseverfahren zur automatischen Analyse des EMG entwickelt.

Wie oben beschrieben werden zahlreiche Biosignale im Rahmen der polysomnographischen Untersuchung aufgezeichnet. Die automatisierte Analyse der Biosignale und speziell des EEG zur Bestimmung der Schlafstadien in der Nacht ist jedoch eine sehr komplexe Aufgabe, was sich vor allem darin widerspiegelt, dass es aktuell nur wenige klinisch validierte automatische Analyseverfahren für das EEG (zusammen mit EOG und EMG) gibt. Eines hiervon ist z.B. der Somnolyzer 24x7 der Siesta Group, der unter anderem an den in MediGRID bereitgestellten Referenzdaten entwickelt wurde (Siesta-Datenbank). Ein anderes Verfahren ist der Artisana Algorithmus der Firma MCC (Karlsruhe) und Weinmann GmbH (Hamburg) an dessen Entwicklung auch das Schlaflabor der Universitätsklinik Gießen und Marburg mit gearbeitet hat. Da der Fokus auf die automatische Erkennung von schlafbezogenen Atmungsstörungen gelegt wurde, konzentrierten sich die Entwicklungen von automatischen Analyseverfahren für ein Anwendungsszenario im Grid auf solchen, die eine zuverlässige Erkennung von Atmungsstörungen bieten. Aufgrund der vergleichsweise einfachen Signalstruktur und damit verbundenen Artefaktstabilität wurde somit für das Anwendungsszenario klinische Forschung das EKG-Signal gewählt. Zudem sind Analysen des EKG Signals mit einem sehr geringen Aufwand auf ganz

andere Projekte der klinischen Forschung zu übertragen. Heute wird bei jeder klinischen Prüfung eine Untersuchung des Herzschlages mit EKG verlangt.

Die open source Programmbibliothek „physio toolkit“ zur Analyse von Biosignaldaten wurde in das Grid implementiert und ein Grid-Dienst zur schlafstadienaufgelösten Berechnung der mittleren Herzfrequenz entwickelt. Dazu werden basierend auf den Suchergebnissen die Herzfrequenz aus den entsprechenden Polysomnographien ermittelt und mit den Hypnogrammen (die die Schlafphasen enthalten) verknüpft.

Die parallelisierbare Analyse der Polysomnographien und Hypnogramme wurde mithilfe eines einfachen Workflows modelliert, die Auswahl einzelner Datensätze und die anschließende EKG-Analyse konnte über ein Anwendungsportlet gesteuert werden.

Für eine statistische Analyse einer bestimmten Untergruppe der SIESTA-Datenbank musste die Metadatensuche allerdings manuell auf einem Gridknoten durchgeführt, die Workflowbeschreibung manuell angepasst und in das Portal hochgeladen werden; die anschließende statistische Analyse der Gesamtheit der gewählten Datensätze wurde lokal durchgeführt. Damit konnte zwar eine deutliche Reduktion der Laufzeit erreicht werden, das Programm ist aber nur für Nutzer mit direktem (gsissh) Zugriff auf die Gridressourcen und der entsprechenden lokalen Analyseprogramme und Softwarelizenzen verwendbar. Deshalb wurde anschließend die komplette Anwendung analog zu den Bildverarbeitungsanwendungen als Workflow modelliert und das existierende Portlet um die Metadatensuche und die Visualisierung der statistischen Analyse erweitert. Durch die Verwendung der bereits entwickelten Portletkomponenten und Workflowbeschreibungen konnte die Entwicklungsdauer dieser Komponenten auf wenige Wochen reduziert werden. Zur Zeit können über das Portlet Untergruppen nach verschiedenen Suchkriterien ausgewählt werden, außerdem stehen zur Analyse unterschiedliche Algorithmen zur Verfügung.

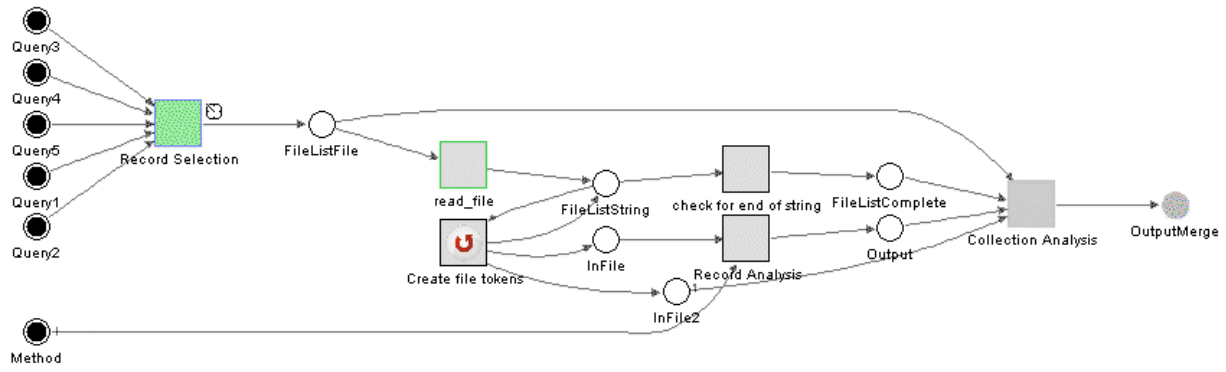


Abbildung 8. Beschreibung des Workflow

Das entwickelte Portal (siehe Abbildung 9) stellt somit den zweiten Meilenstein (Realisierung einer Gridanwendung für die klinische Forschung im ersten konkreten Fall) des Arbeitspaketes 1 (Anwendungsszenario „Schlafforschung“) dar.

Wie auch schon in der Publikation zum entwickelten Analyseverfahren beschrieben eignet sich das EKG-Signal deshalb ausgezeichnet hierfür, weil die nächtlichen Atmungsstörungen fast immer von einer Weckreaktion (sog. Arousal) beendet werden, wobei die Weckreaktionen selbst wiederum eine zyklische Variation der Herzfrequenz (CVHR cyclical variation in heart rate) verursachen. Diese Variationen in der Herzfrequenz können sehr gut automatisch erkannt werden. Ein erster Ansatz der Analyse des EKG zur Erkennung von CVHR wurde in Zusammenarbeit mit dem Projekt InGRID erarbeitet. Der Fachbereich Informatik der Philipps-Universität Marburg ist an diesem Projekt beteiligt, Mitarbeiter der Arbeitsgruppe entwickelten einen ersten

Analysealgorithmus auf JAVA-Basis. Dieser Algorithmus sowie das Konzept der Nutzung im Grid wurde im Rahmen der HealthGRID Tagung 2007 in Genf durch das Modul Klinische Forschung vorgestellt.

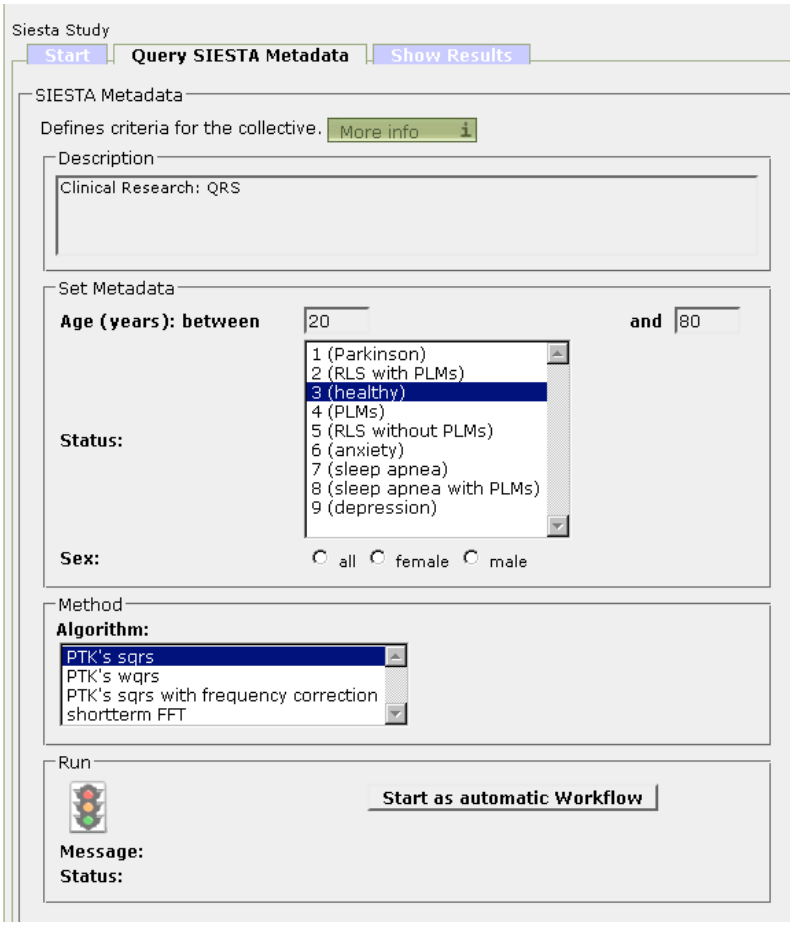


Abbildung 9. Screenshot des Portlets

Das schlafmedizinische Zentrum der Uniklinik Gießen und Marburg hat während der Laufzeit von MediGRID auf Basis des von den Kollegen des FB Informatik erstellten Verfahrens einen Algorithmus zur Erkennung der CVHR und somit Detektion von schlafbezogenen Atmungsstörungen aus dem EKG-Signal entwickelt. Der Algorithmus selbst wurde mit Hilfe von Referenzdaten des „Physionet“ erstellt, die letztendliche Validierung jedoch soll mit Hilfe der im Grid bereitgestellten polysomnographischen Daten (Siesta-Datenbank) erfolgen. In Abbildung 1 ist die grundsätzliche Funktionsweise erläutert. Prinzipiell besteht das Analyseverfahren aus zwei Schritten: 1. der Berechnung der Herzfrequenz aus dem EKG-Signal (siehe Abbildung 10), 2. Berechnung der „Frequency Power“ und des „Differential Power“ aus dem Herzfrequenzsignal (siehe Abbildung 11). Phasen, in denen das Differential Power Signal oberhalb einer bestimmten Schwelle liegt, werden als Phasen mit schlafbezogenen Atmungsstörungen erkannt.

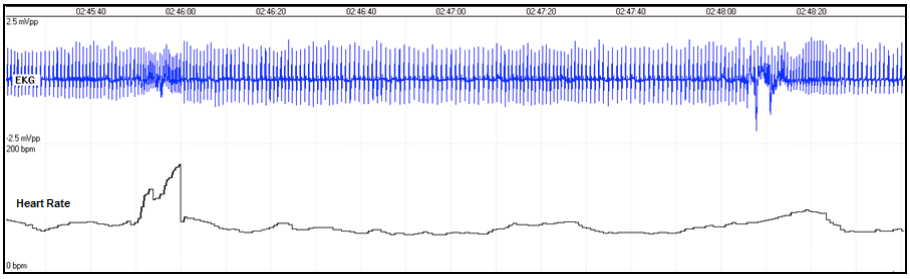


Abbildung 10. Berechnung der Herzfrequenz aus dem EKG

Oben ist das EKG mit jedem einzelnen Herzschlag zu sehen und unten ist die Herzfreqeunz zu sehen. Links sieht man bei einer Aufwachreaktion einen plötzlichen Anstieg der Herzfrequenz.

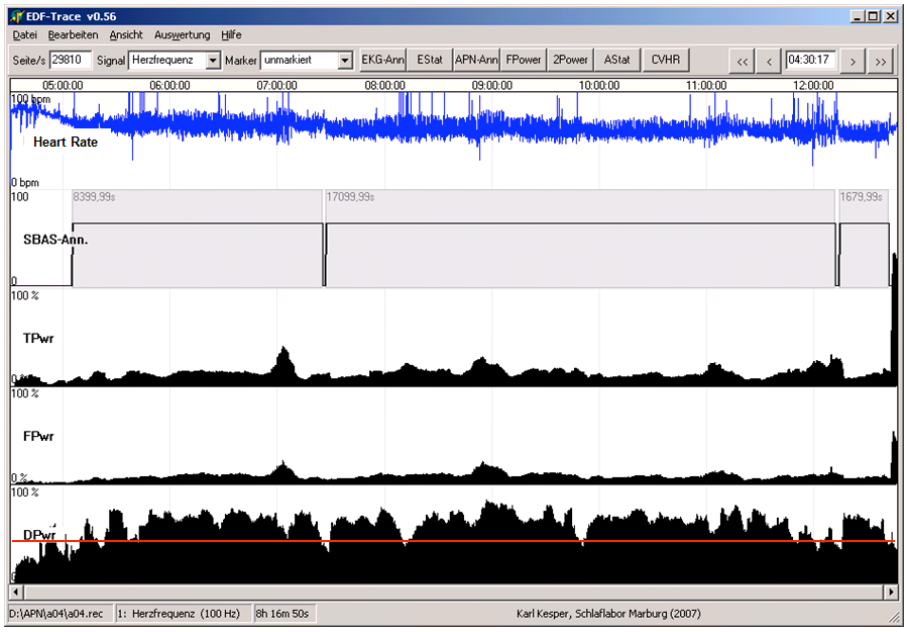


Abbildung 11. Erkennung von schlafbezogenen Atmungsstörungen (SBAS)

Erkennung von SBAS aus der Herzfrequenz. Aus der Herzfrequenz berechnete Parameter (TPwr, FPwr, DPwr) sind zusätzlich dargestellt.

Während der Laufzeit des Projektes MediGRID wurde ebenfalls ein Analyseverfahren für das EMG-Signal entwickelt, da die elektrische Muskelaktivität im Schlaf (EMG – Elektromyogramm) neben dem EKG klinisch ebenfalls sehr bedeutsam ist. So gehen einige schlafmedizinische Erkrankungen mit charakteristischen Veränderungen der Muskelaktivität im Schlaf einher. Solche Erkrankungen sind z.B. das Syndrom der unruhigen Beine (Restless Legs Syndrome – RLS) oder auch die REM-Schlaf-Verhaltensstörung (REM-Sleep-Behaviour Disorder – RBD). Bei letzterer bleibt die physiologische Entkopplung der muskulären Aktivität von der Hirnaktivität im Traumschlaf (REM-Schlaf) aus und es kommt zum ungewollten Ausagieren der Träume. Das RLS dagegen ist durch periodisch auftretende Muskelaktivierung vor allem in den Beinen während des nächtlichen Schlafes gekennzeichnet, was zu einem deutlichen Zucken der Beine und damit verbundenen Weckreaktionen führt. Eine automatische Analyse der elektrischen Muskelaktivität in der Nacht kann also die Diagnosestellung im schlafmedizinischen Bereich sehr erleichtern. Die nächtlichen Aufzeichnungen, die vom Modul Klinische Forschung zusammengetragen worden sind, enthalten auch Aufzeichnungen von Patienten mit diesen Erkrankungen.

Der in unserer Arbeitsgruppe entwickelte Analysealgorithmus der Muskelaktivität wurde validiert. Darüber hinaus wurden weitere Projekte zum Einsatz dieses Algorithmus geplant und deren Finanzierung durch öffentliche Mittel beantragt. Der Algorithmus wurde im Journal of Clinical Neurophysiology publiziert. Die prinzipielle Funktionsweise ist in Abbildung 12 dargestellt.

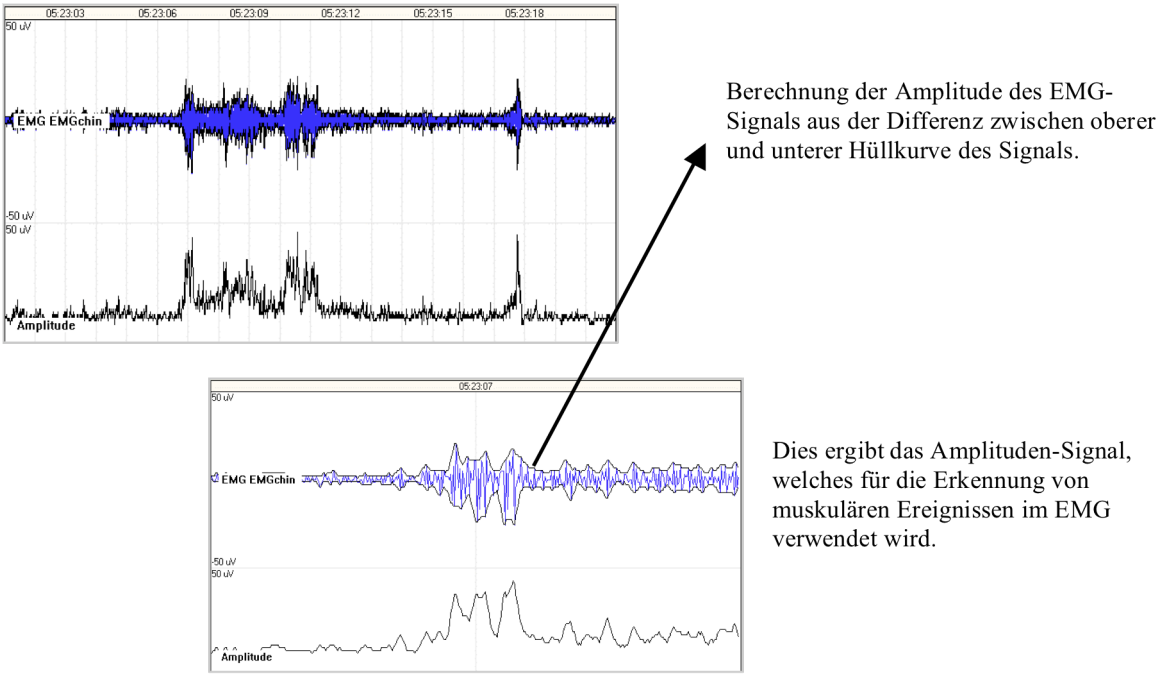


Abbildung 12. Funktionsweise des EMG-Algorithmus

Die im Rahmen der Siesta-Datenbank zur Verfügung stehenden polysomnographischen Aufzeichnungen (ca. 300) liegen zwar anonymisiert vor, jedoch soll das Anwendungsszenario der klinischen Forschung auch mit Daten aus dieser selbst nutzbar sein. Dies erfordert die Entwicklung eines neuen Datenschutzkonzepts für diese besondere Anwendung, da hierfür bisher keine existierenden Konzepte vorhanden sind. Einen großen Schwerpunkt des Moduls Klinische Forschung stellte daher die Entwicklung eines sog. Pseudonymisierungskonzepts dar, mit dessen Hilfe zum einen die Einhaltung von Datenschutzrichtlinien für die polysomnographischen Aufzeichnungen (medizinische Daten!) zum anderen die Nutzbarkeit einer Grid Umgebung außerhalb des Krankenhausnetzwerks erreicht werden soll. Da die durch die automatische Analyse im Grid gewonnenen Ergebnisse potentiell auch auf den Patienten, von dem die Aufzeichnung stammte, zuzuordnen sein sollen, kann hier nur eine Pseudonymisierung und keine vollständige Anonymisierung erfolgen. Dies erfordert aber besondere Beachtung der Datenschutzrichtlinien für medizinische Daten, wobei sowohl die Biosignalaufzeichnung selbst als auch die Ergebnisse der automatischen Analyse solche medizinischen Daten darstellen.

In Zusammenarbeit mit dem Modul eScience wurde somit ein Pseudonymisierungskonzept entwickelt, welches zum einen die technischen Gegebenheiten der Grid-Middleware zum anderen aber auch die Datenschutzrichtlinien zum Umgang mit medizinischen Daten berücksichtigt. Die Entwicklung des Konzepts erfolgte in Zusammenarbeit mit dem behördlichen Datenschutzbeauftragten der Universitätsklinik Gießen und Marburg GmbH, dieser hat das letztendliche Konzept auch hinsichtlich der Verwendbarkeit formal akzeptiert. In Abbildung 13 ist das Pseudonymisierungskonzept als Schema dargestellt. Es stellt das Deliverable 1 unseres Moduls dar.

Das präsentierte Konzept ist generisch und kann auch für klinische Studien außerhalb des Grid-Kontexts genutzt werden, da die zugrunde liegenden Datenschutzkonzepte nicht Grid-spezifisch sind.

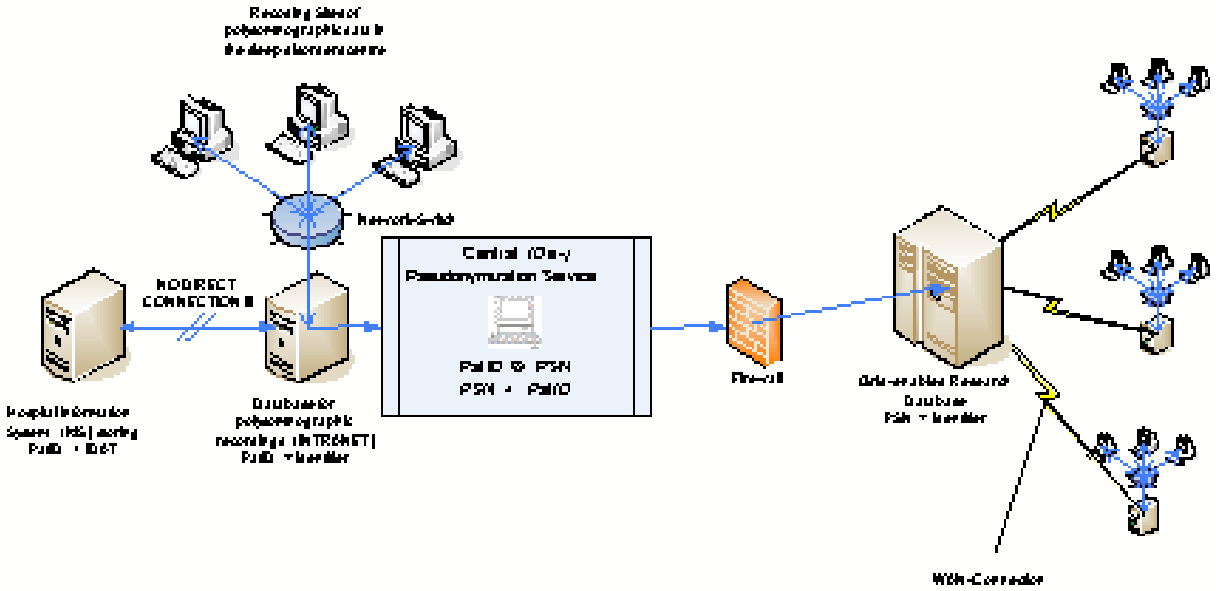


Abbildung 13. Pseudonymisierungskonzept

2.3.2.2 Sonstige Ergebnisse

- Zusammenarbeit mit InGRID (FB Informatik der Uni Marburg) im Bereich der Algorithmenentwicklung (erste Schritte zur EKG-Analyse) → Präsentation der vorläufigen Ergebnisse dieser Entwicklung im Rahmen der HealthGRID Tagung 2007
- Teilnahme an Security Workshops, All-Hands-Meeting, HealthGRID Tagung 2007, e-Health Tagung 2007, HealthGrid Tagung 2009, ccGrid Tagung 2009, TMF Workshop ID-Management 2008, DGSM-Tagungen mit Vorstellung der Ergebnisse, DGP-Tagungen mit Vorstellung der Ergebnisse, ERS-Tagungen mit Vorstellung der Ergebnisse.
- Publikationen: Es konnten einige Publikationen zum Konzept der Gridanwendung für die klinische Forschung sowie zu selbst entwickelten Analyseverfahren, die als Gridanwendung genutzt werden, veröffentlicht werden. Diese sind am Ende des Berichts (Kapitel 2.1) aufgelistet.
- Erarbeitung von Anträgen zur weiterführenden GRID-Projekten unter Einbeziehung der Anwender (PneumoGRID)

2.3.3 Nutzen und Verwertbarkeit der Modulergebnisse

Die erreichten Modulergebnisse sollen als Analyseplattform für polysomnographische Aufzeichnungen weitergenutzt werden. Da das MediGRID Portal einen Zugriff von jedem Internet-fähigen Rechner ermöglicht, kann das Portlet des Moduls klinische Forschung auch aus geschützten Netzwerken (hier insbesondere Krankenhausnetzwerke oder Netzwerke in Arztpraxen) genutzt werden. Mit Hilfe des entwickelten Pseudonymisierungskonzepts werden auch die datenschutzrechtlichen Bestimmungen für die Analyse von polysomnographischen Aufzeichnungen im Grid erfüllt.

Aufgrund der bereits vorhandenen Datenstruktur und der Möglichkeit, Algorithmen aus verschiedenen Programmiersprachen einzubinden, können mit Hilfe von MediGRID sehr leicht und effizient neue Analysealgorithmen für die schlafmedizinische Diagnostik entwickelt und an

Referenzdatensätzen validiert und getestet werden. Hier bietet insbesondere die große Menge an Referenzdaten aus der SIESTA Datenbank, die von unserem Modul bereitgestellt werden konnte, eine ideale Datenbasis. Somit können die Entwicklungen aus dem Modul Klinische Forschung nicht nur für die wissenschaftliche oder versorgungsorientierte Analyse von Polysomnographien genutzt werden, sondern auch die Entwicklung neuer Analyseverfahren beschleunigen und vereinfachen. Aktuell wird dies für eine Weiterentwicklung des Analysealgorithmus zur Detektion von schlafbezogenen Atmungsstörungen aus dem EKG genutzt, eine weitere Nutzung ist geplant um ein ähnliches Analyseverfahren zur Erkennung von Schlafstadien aus dem EKG zu testen und zu validieren.

Nicht zuletzt für das geplante Projekt „PneumoGRID“ (3. Call BMBF) stellt MediGRID eine Basisimplementierung da, auf die zukünftige Gridprojekte aufsetzen können. Konkret sollen viele Entwicklungen (Datenschutz, Zugriffsrechte usw.) weitergenutzt werden.

2.3.4 Ausblick und Perspektiven im Bereich Klinische Forschung

Das Modul Klinische Forschung konnte im Rahmen von MediGRID ein Portlet mit Analyseverfahren entwickeln, welches als Basis für die automatisierte Analyse von Polysomnographien und für die effizientere Weiterentwicklung von automatischen Analyseverfahren von nächtlichen Biosignalen dienen kann. Eine weitere Nutzung der erreichten Entwicklungen in zukünftigen Projekte (z.B. PneumoGRID) wird angestrebt.

2.4 Middleware

2.4.1 Moduleinordnung und Modulziele

Das Modul Middleware bildete die Klammer von den Portalen zu den Anwendungen. Ziel des Moduls Middleware war es, eine auf die Anforderungen der MediGRID Community abgestimmte Grid Plattform zu realisieren. Das Modul hatte die Aufgabe, vorhandene Middleware-Entwicklungen an die Anforderungen aus den Anwendungsszenarien anzupassen und die notwendigen Community-spezifischen Entwicklungen, insbesondere in den Bereichen Grid Portale und Grid Workflow Systeme zu leisten. Dies beinhaltete die Bereitstellung und Virtualisierung von Ressourcen, sowie die Erleichterung der Integration von Fachanwendungen aus dem Bereich der Lebenswissenschaften.

Die im MediGRID Middleware-Modul realisierte grid-basierte e-Science-Infrastruktur unterstützt die Zusammenarbeit von Wissenschaftlern im biomedizinischen Bereich und die Ausführung der dafür nötigen Anwendungen. Das Modul stellte auch die Verknüpfung zum D-Grid-Integrationsprojekt und zu internationalen Grid-Aktivitäten sicher.

Das Modul Middleware arbeitete eng mit dem Modul Ressourcenfusion zusammen, das die Koordination zwischen Anwendern, Middleware-Entwicklern und Ressourcenprovidern übernahm.

Die Realisierungen im Modul Middleware umfassen die folgenden Bereiche:

- Grid Workflow System
- Community Grid Portal
- Fachspezifische Erweiterung der Ressourcenbeschreibungssprache

Die Bereiche Datagrid und Grid-Basisplattform wurden im Modul Ressourcenfusion bearbeitet. Es erfolgte daher eine enge Abstimmung der Konzeptionen und Umsetzungen mit dem Modul Ressourcenfusion.

Dabei wurde in einem zyklischen Verfahren der Nutzen der MediGRID-spezifischen Middleware anhand der Anwendungen aus den anderen MediGRID Modulen überprüft. Die Erkenntnisse dieser Tests flossen direkt in den Entwicklungsprozess zurück.

Die Zielsetzung, die im Antrag für das Modul Middleware gesetzt wurde konnte somit vollständig realisiert werden.

Die Aufgabenverteilung war dabei folgendermaßen:

Fraunhofer IAO: Koordination des Moduls und Abstimmung mit den anderen MediGRID Modulen sowie mit dem D-Grid Integrationsprojekt, Portalentwicklung, Nutzermanagement und Nutzersupport, Entwicklerunterstützung im Portalbereich

Fraunhofer FIRST: Workflowentwicklung, Weiterentwicklung der D-Grid Ressourcenbeschreibungssprache anhand der Anforderungen aus MediGRID, Entwicklersupport

IAT, Universität Stuttgart: im Unterauftrag des Fraunhofer IAO wurden ab Juni 2006 der Portalbetrieb und die Wartung des Portals durchgeführt.

2.4.2 Modulergebnisse

Das Modul Middleware mit den Fraunhofer-Instituten IAO und FIRST sowie dem Unterauftragnehmer IAT, Universität Stuttgart konnte im MediGRID Projekt alle geplanten Projektergebnisse und Meilensteine fristgerecht erreichen und umsetzen. Die wesentlichen Projektergebnisse sind der Aufbau einer Middleware Infrastruktur – aufsetzend auf der MediGRID und D-Grid Hard-

ware-Infrastruktur – die es dem Endnutzer erlaubt, auf eine Reihe von Grid-Anwendungen aus den Bereichen Bioinformatik, medizinische Bildverarbeitung, klinische Forschung und biomedizinische Ontologien zuzugreifen und deren Potenziale zu nutzen, ohne dabei mit der Infrastruktur oder mit Grid Technologien vertraut zu sein. Dies ist eine wesentliche Voraussetzung, um diese Technologien einem breiten Nutzerkreis in der biomedizinischen Community zu eröffnen.

Ein wesentlicher Eckpfeiler ist dabei zum einen das MediGRID Portal, das über anwendungsspezifische Portlets einen webbasierten – und bzgl. der Grid-Nutzung völlig transparenten – Zugriff auf die MediGRID Anwendungen erlaubt. Ein weiterer Grundpfeiler der Architektur ist der Grid Workflow Execution Service (GWES), der mit seiner Funktionalität als Workflowsystem, Ressourcenbroker und Meta-Scheduler eine Schlüsselrolle bei der Virtualisierung der Grid Ressourcen einnimmt und es somit erst möglich macht, die Komplexität der Grid Infrastruktur vor dem Nutzer zu verbergen. Gleichzeitig kann diese Komplexität aber bei Bedarf und Interesse auch in all ihren Details nachvollziehbar gemacht werden. Letzteres wird insbesondere durch die Workflowmonitoring-Funktionen des GWES erreicht, die eine detaillierte Sicht auf Logging-Informationen zu einzelnen Jobs ermöglicht und somit auch ein wichtiges Mittel zur Fehlersuche und Fehlerbehebung beim Aufbau und Betrieb der Middleware-Infrastruktur war und ist.

Ein Kernelement der Arbeiten des Moduls Middleware lag auch in der Abstimmung mit den Arbeiten zum Datenmanagement via SRB und iRods im Modul Ressourcenfusion sowie in der Abstimmung zu Security-Fragestellungen mit dem Modul e-Science.

Weitere wichtige Ergebnisse des Moduls Middleware liegen in der Anbindung der MediGRID Infrastruktur an zentrale D-Grid Dienste wie das User und VO Management, die Zertifizierungsinfrastruktur oder das Monitoring. Ebenso von Bedeutung war die Kooperation mit dem DGI bei der (Weiter-) Entwicklung der D-Grid Ressourcenbeschreibungssprache „D-GRDL“. Ein großer Schritt in Richtung Nachhaltigkeit konnte gegen Projektende mit der Vorbereitung einer rechtlich abgesicherten Nutzung der zentralen D-Grid Accounting Infrastruktur in Zusammenarbeit mit dem entsprechenden Fachgebiet des DGI II und den Betreibern der Accounting Services erreicht werden.

Ein besonderer Erfolg der Arbeiten ist aber vor allem die nachhaltige Sicherung des Betriebs der MediGRID Infrastruktur, die zunächst durch den Weiterbetrieb und Erhalt der Infrastruktur im Projekt Services@MediGRID und längerfristig über das Forum Grid erreicht wird.

Im Folgenden sind die Projektergebnisse des Moduls Middleware im Detail dargestellt.

2.4.2.1 Anforderungsanalyse

Zu Beginn des Projekts wurde zunächst eine umfangreiche Anforderungsanalyse bei den Entwicklern und Nutzern der für MediGRID vorgesehenen Anwendungen durchgeführt. Zunächst wurden die Anwendungs-Entwickler gemeinsam mit dem Modul Ressourcenfusion per Fragebogen zu ihren Anwendungen, dem Einsatzzweck und den technischen Randbedingungen sowie zu ihren Anforderungen an die Middleware befragt. Nach Auswertung und Analyse der Umfrageergebnisse erfolgten mehrere Workshops vor Ort bei den Anwendungsentwicklern, um sich ein detailliertes Bild von den Anforderungen und Rahmenbedingungen machen zu können. Bei diesen Workshops wurden soweit möglich auch Endnutzer mit einbezogen. Begonnen wurde mit einem zentralen Workshop zur Vorstellung der Middlewareplanungen, bei dem alle Anwendungsbereiche und Ressourcenbetreiber anwesend waren. (Eintägiger MediGRID-Workshop mit 28 Projektteilnehmern am 22.11.2005 in Berlin: Verfügbare HW-Ressourcen, einzusetzende Anwendungen, geplante Grid-Middleware, Zertifizierungsverfahren). Anhand der Ergebnisse dieses Anwendungsbereich-übergreifenden Workshops erfolgte dann eine gezielte Abstimmung der Einzelworkshops mit den verschiedenen Anwendungen bzw. deren Entwicklern sowie mit den MediGRID Ressourcenbetreibern.

2.4.2.2 Spezifikation

Auf der Grundlage aller Workshop-Ergebnisse und der Anforderungen, die bei den Anwendern und Anwendungsentwicklern aufgenommen wurden, wurden anschließend ein Grobkonzept für die MediGRID-Middleware erstellt und die Spezifikationsaufgaben abgeleitet. Dazu zählte die Definition der Eckpunkte einer Middlewarestrategie für MediGRID. Im Fokus stand dabei die Nutzung des Globus Toolkits als Standard-Middleware in MediGRID, die Verwendung und Weiterentwicklung des Grid Workflow Execution Service (GWES), der portalbasierte Nutzerzugang zu den MediGRID Anwendungen und die Festlegung einer initialen Monitoringstrategie für die MediGRID Hardware-Ressourcen.

Im Rahmen der Ableitung der Spezifikationsaufgaben für die Definition der MediGRID Middleware wurden bereits erste Modellierungen von Anwendungsworkflows anhand der Workshop-Auswertungen vorgenommen. Darüber hinaus wurden Anforderungen für das D-Grid-Integrationsprojekt abgeleitet die im Folgenden mit dem D-Grid-Integrationsprojekt abgestimmt wurden. Dies betraf insbesondere die Themengebiete „Grid Portale“, „Ressourcenbeschreibung“, „Monitoring“, „Billing“ und „Security“. Zur Unterstützung biomedizinischer Ontologien erfolgte aber auch eine Koordination der Zusammenarbeit zwischen den Communties MediGRID und TextGrid im Rahmen des Workshops "Ontologien im Grid", 9.-10.3, Leipzig.

Über die Workshops hinaus erfolgte im Rahmen der Spezifikation der MediGRID Middleware einerseits eine permanente Aufnahme von Informationen für Ressourcenbetreiber und Anwendungsentwickler, und andererseits auch die Bereitstellung von Informationen für diese beiden Gruppen über ein MediGRID Wiki.

In Abstimmung mit dem Modul Ressourcenfusion wurden die Ergebnisse der Anforderungsanalyse und der Spezifikation der MediGRID Middlewareinfrastruktur in einem Dokuments zur Anforderungsaufnahme und deren Auswertung zusammengefasst.
Titel: „MediGRID - Analyse und Spezifikation der MediGRID Anwendungen“. Inhalt dieses Dokuments ist die Ableitung der Anforderungen an die MediGRID Middleware (insbesondere Portal und Workflowsystem) seitens der Anwender sowie die Auswahl der Middleware-Komponenten für MediGRID und der Entwurf einer MediGRID Middleware- und Datenmanagementarchitektur.

Die Ergebnisse der Middleware-Architektur-Konzeption wurden im Anschluss im Rahmen eines zweitägigen Workshops am 03./04.07.06 in Göttingen an die MediGRID-Anwender und Anwendungsentwickler weitergegeben.

2.4.2.3 Realisierung und Konsolidierung

MediGRID Portal

Parallel zur Aufnahme der Anforderungen wurde mit dem Aufbau der für das Projekt notwendigen Hardware- und Organisations-Infrastruktur begonnen. Neben der Einrichtung der Rechen- und Speicherressourcen sowie einer Abstimmung des Nutzermanagements mit dem DGI zählte hierzu beispielsweise auch die Einrichtung eines Wiki-Servers zur Unterstützung von Nutzern, Entwicklern und Ressourcenbetreibern in MediGRID. Als wesentliche Voraussetzung für den Betrieb der Infrastruktur wurde zunächst auch das Zertifizierungsverfahren in MediGRID mit den Zertifizierungsstellen in D-Grid abgestimmt und für die MediGRID Community definiert.

Die Grid-Portal-Infrastruktur wurde auf Basis des vom DGI unterstützten Gridsphere Portal Frameworks aufgebaut. Zunächst wurden hierfür Ressourcen der GWDG in Göttingen genutzt. Auch der Grid Workflow Execution Service (GWES) wurde zunächst auf dieser Ressource installiert. Mit der Beschaffung der Sonderinvestitionen 2006 wurden an der Universität Leipzig die im Rahmen der Sonderinvestition beschafften Ressourcen in Betrieb genommen, darunter ein

LoadBalancer zur Verteilung der Anfragen, zwei Server für den Betrieb des Portals und ein iSCSI-Storage als Backend zur Datensicherung.

Nach der Installation des Betriebsystems, der Basisdienste und der Anbindung des Backend wurde auf beiden Servern jeweils eine komplette Portal-Infrastruktur aufgebaut. Mit Hilfe des Loadbalancers konnten nun die Anfragen auf beide Portale aufgeteilt werden und so eine Lastverteilung erreicht werden. Im Falle eines Serverausfalles kann der noch funktionierende Server die Tätigkeit alleine weiterführen.

Auf beiden Servern läuft eine Gridsphere-Framework Instanz mit den Fach-Anwendungen. Die beiden Instanzen arbeiten auf den gleichen Datenbanken und geteilten Verzeichnissen und bleiben so synchron. Alle Tätigkeiten auf einer Instanz sind direkt auf der anderen verfügbar.

Im nächsten Schritt wurde zur Stabilitätssicherung eine gegenseitige Überwachung der Server eingeführt, damit bei Ausfällen die Dienste automatisch übernommen werden können. Der Loadbalancer reagiert auf Ausfälle und leitet Anfragen nur auf laufende Dienstinstanzen weiter.

Zu Testzwecken wurde zunächst eine Testphase mit einem Parallelbetrieb der Portale auf den Ressourcen in Göttingen und Leipzig durchgeführt bevor der Betrieb dann voll auf die neuen Ressourcen umgestellt wurde.

Neben Gridsphere wurden auf den Portalen auch Globus-Instanzen installiert, deren Zweck das Anstoßen von Rechnungen auf anderen Ressourcen ist. Das Abarbeiten von Rechenjobs auf den Portalressourcen ist nicht geplant.

Auch der GWES wurde im Zuge des Umstiegs auf die Sonderinvest-Ressourcen auf den Portalservern installiert.

Im weiteren Verlauf des Projekts wurde der redundante und ausfallsichere Betrieb der Portalressourcen aus Gründen des Lifecycle Managements zugunsten einer Konstellation bestehend aus Produktiv- und Entwicklungsportal geändert. Das Testportal ist dem Produktivportal technisch gleichgestellt, so dass die Preproduction-Tests zum einen für die Produktivumgebung aussagekräftig sind und zum anderen die getesteten Erweiterungen direkt übertragen werden können.

Die Einrichtung und Konfiguration des MediGRID Portals umfasste zunächst den Aufbau eines Systems zur Nutzerverwaltung am Portal, der im Verlauf des Projekts immer enger mit dem Virtual Organisation Management and Registration Service (VOMRS) verbunden wurde, welcher vom DGI betrieben und bereitgestellt wird. Darüber hinaus wurden zunächst die Gridportlets eingerichtet, die es erlauben, Globus-Ressourcen an das Portal anzubinden und über das Portal zu verwalten. Zum Funktionsumfang zählen dabei die Verwaltung der für die Globus-Nutzung notwendigen Credentials sowie ein Portlet zur Sumittierung von Jobs auf angeschlossene Globus-Ressourcen. Darüber hinaus steht ein File-Transfer Portlet zur Verfügung um auf Middleware-Ebene Daten zwischen den angeschlossenen Ressourcen transferieren zu können.

Zu diesem Zweck wurde eine Anbindung des Credential Management Portlets an den bei der GWDG im Rahmen des Moduls Ressourcenfusion betriebenen MyProxy Server angebunden. Des Weiteren wurden die für die MediGRID VO verfügbaren Globus-Ressourcen eingetragen und für die Grid-Portlet-Funktionen konfiguriert und getestet.

Weitere in MediGRID zur Verfügung gestellte Basiskomponenten des Portals sind die Monitoring-Portlets sowie Komponenten zur Erstellung und Verwaltung von Ressourcenmetadaten in der D-Grid Resource Description Language. Letztere wurde aus dem DGI Fachgebiet 1-5 übernommen. Die Monitoring Portlets bieten die Möglichkeit aktuelle Snapshots aus dem MediGRID Ganglia Clustermanagement sowie aus dem MediGRID MDS (später auch des gesamten D-Grid MDS) zu nutzen. Hierfür wurde ein Synchronisationsmechanismus entwickelt, da eine direkte Einbindung externer Webseiten in Gridsphere 2 noch nicht gegeben war.

Neben dem Aufbau und Betrieb der Portal-Infrastruktur und der Basisfunktionalität bzgl. Der Gridanbindung war vor allem die Integration Anwendungs- und Service-Interfaces in das MediGRID Portal ein wesentlicher Bestandteil der Arbeiten zur Realisierung und Konsolidierung. Im Laufe des Projekts wurde eine ganze Anzahl von anwendungsspezifischen Portlets in das MediGRID Portal integriert. Im Einzelnen waren dies folgende Portlets:

- Für den Zugriff auf biomedizinische Ontologien wurde eine Multi-Ontologie-Anwendung für den Zugriff auf diverse Ontologie-Datenbanken aus dem Bereich der Bioinformatik über OGSA-DAI integriert. Die Anwendung bereitet die Datenbankinhalte (auch graphisch) auf und erlaubt die einfache und intuitive Verknüpfung von Begriffen mittels der zugrunde liegenden Ontologien. Der OGSA-DAI-Zugriff wurde dabei zunächst als Proof-of-Concept durch die eine Authentifizierung und Autorisierung mittels gültiger D-Grid-Zertifikate geschützt. Aus Gründen der Usability wurde diese Zugriffskontrolle anschließend wieder entfernt, da der Zugriff auf die Datenbankinhalte nicht beschränkt ist. Darüber hinaus wurde eine Verknüpfung mit dem D-Grid Ontology Wiki hergestellt.
- Die Bioinformatik-Anwendung „AUGUSTUS“ zur Gensequenzanalyse wurde ebenso wie die Schwesteranwendung „DiAlign“ in das MediGRID Portal integriert. Da die Anwendung einer möglichst breiten Nutzerschaft – auch außerhalb der D-Grid Communities – zur Verfügung gestellt werden sollte wurden im Rahmen der Integration auch Konzepte entwickelt und getestet, wie man die Anwendung mit Service-Zertifikaten (bzw. Robot-Zertifikaten) betreiben kann, so dass auf Nutzerseite keine Verwendung von Zertifikaten erforderlich ist, man aber gleichzeitig die Globus-Rechner-Infrastruktur nutzen (die die Verwendung von Proxys voraussetzt) kann um die resultierenden Jobs im Grid zu verteilen.
- Eine weitere Anwendung aus dem Bereich der Bioinformatik war die SNP-Selection. Sie dient zur Auswahl von Einzelbasenpolymorphismen für Krankheitskartierungen. Mit „Single Nucleotide Polymorphism“ werden Variationen von einzelnen Basenpaaren in einem DNA-Strang bezeichnet. Die SNP-Selection wird genutzt um Gene zu identifizieren, die für die Auslösung bestimmter Krankheiten verantwortlich sind. Die Anwendung „SNPSelection“ ist gekennzeichnet durch große Datenmengen (ca. 1 Million Genotypen pro Patient), komplexe und verteilte Datenstrukturen, rechenintensive Analysen (Haplotyp- und Interaktionsanalysen) und eine verteilte und wenig technophile Nutzergruppe.
- Im Bereich der medizinischen Bildverarbeitung wurde eine Anwendung zur Erzeugung 3D-Ultraschallbildern von 2D-Ultraschallaufnahmen zur Unterstützung der Auffindung von Biopsienadeln im Rahmen einer Prostatabiopsie integriert. Diese mehrstufige Anwendung erfordert komplexe Workflows und eine enge Anbindung an die Datenmanagementsysteme. Im Fall der 3D – Ultraschallaufnahmen bei Prostatabiopsien ist es das Ziel, möglichst schnell eine genaue Lokalisation der Gewebeproben durch die Analyse mehrerer Bilder zu erhalten. Hier sind rechenintensive Bildverarbeitungsroutinen durchzuführen und der Zugriff auf heterogene, verteilte Daten zu gewährleisten.
- Die Strömungssimulation in Blutgefäßen (Virtuelle Gefäßchirurgie) dient der 3D Darstellung von Blutströmung in Gefäßen und deren Veränderung. Die Geometrie der Gefäße wird durch unterschiedliche Aufnahmemodalitäten erfasst. Anschließend wird die Blutströmung und die resultierende Belastung der Gefäßwände berechnet. Eine zu hohe Gefäßbelastung bildet die Gefahr einer Blutung, die bei Hirngefäßen zum Schlaganfall führt.
- Eine weitere Anwendung aus dem Bereich der medizinischen Bildverarbeitung ist die Analyse und Aufbereitung funktioneller Hirnbilddaten, wie sie in leistungsfähigen Magnetresonanztomographen erzeugt werden.

- Eine Anwendung zur Musteranalyse von EKGs in der Schlafforschung wurde aus dem Bereich „klinische Forschung“ zur Ermittlung von Krankheitsbildern bei Schlafstörungen in das MediGRID Portal und den Workflow Dienst integriert. Die Anwendung befasst sich mit der „QRS-Analyse von Schlaf-EKGs“. Die Integration umfasste die Entwicklung und das Deployment eines anwendungsspezifischen Portlets, sowie die Erstellung von Workflows für den Workflowservice und die Anpassung der Anwendung an das MediGRID Datenmanagementsystem (SRB).

Die Arbeiten zur Portletintegration umfassten vor allem die Konfiguration, Abstimmung und Fehlerbehebung der Anwendungsportlets. Es fanden über die gesamte Laufzeit permanent Erweiterungen und Updates vorhandener Anwendungsportlets statt.

Neben den Anwendungsportlets und den Grid-Basisfunktionen wurden zwei weitere wesentliche Bestandteile der MediGRID Middleware-Infrastruktur ins Portal integriert: ein Portlet zur Verwaltung und zum Monitoring von Grid Workflows und ein Portlet zum Zugriff auf das MediGRID Datenmanagementsystem SRB. Beide bieten einen großen Funktionsumfang sowie eine umfangreiche und nutzerfreundliche Möglichkeit zur Kontrolle der Workflows bzw. Daten auf den MediGRID Compute und Storage Ressourcen.

Abbildung 14 zeigt einen Überblick über die gesamte Portalinfrastruktur mit Zugängen zu Anwendungen aus verschiedenen Bereichen sowie zu Standard-Grid-Portlets, Infrastrukturdiensten und Administrationsdiensten. Das MediGRID Portal wurde somit zur zentralen Schnittstelle für End-Anwender, Anwendungsentwickler bzw. Servicebetreiber und Ressourcenbetreiber.

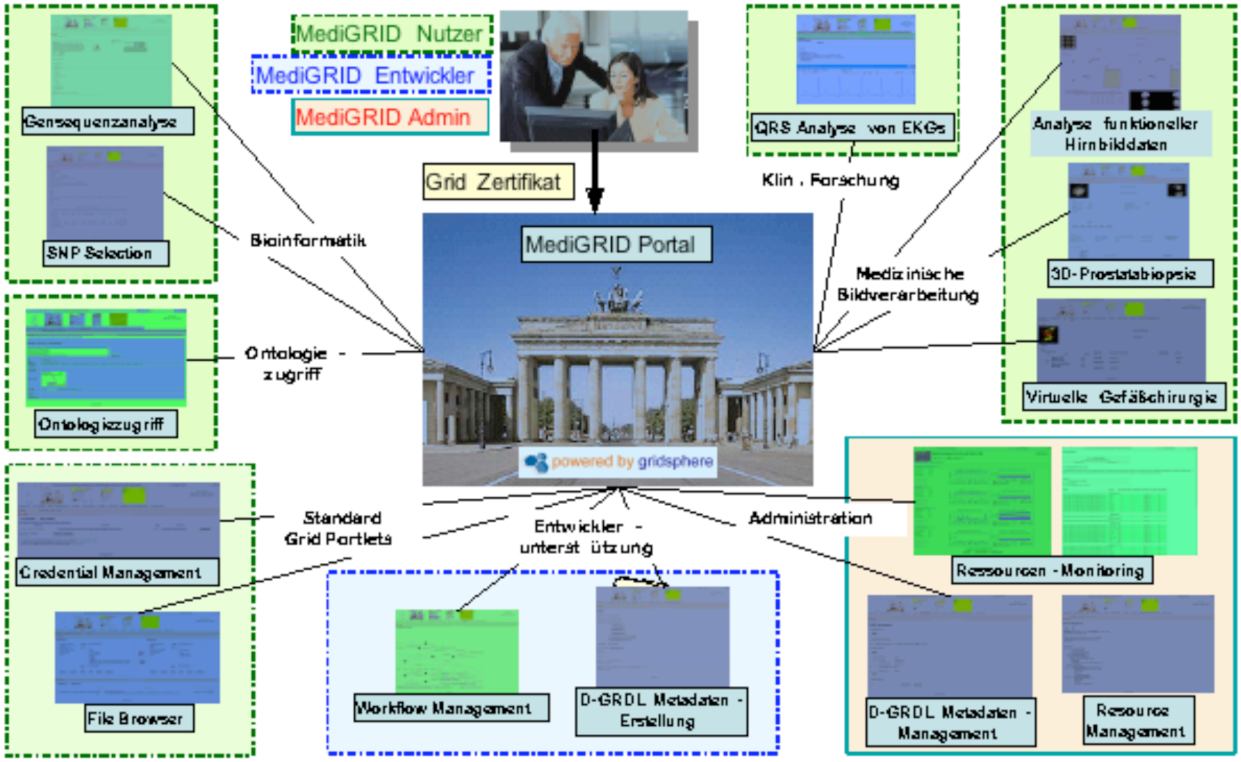


Abbildung 14. Portalbasierter Zugang zu MediGRID

Mit dem Release des Gridsphere Portal Frameworks in der Version 3.0 wurde auch in MediGRID ein Upgrade des Portals angestrebt. Die Gridportlets, die den wesentlichen für das Grid relevanten Funktionsumfang bereitstellen, waren aber für Gridsphere 3.0 nicht mehr verfügbar. Stattdessen sollten diese Funktionen über das Vine Toolkit bereitgestellt werden, dessen Fertig-

stellung sich aber zunächst lange verzögerte und das bis zum Projektende keinen Zustand erreicht hat, in dem es für Gridsphere 3 produktiv einsetzbar wäre. Aus diesem Grunde wurde schließlich auf eine Migration während der Projektlaufzeit verzichtet.

Da sich diese Verzögerungen aber nicht vorhersehen ließen floss zwischenzeitlich ein nicht unerheblicher Aufwand in die Portierung der MediGRID Portlets auf die neue Version des Portalframeworks. Im Zusammenhang mit dem geplanten Wechsel auf das neue Gridshpere Framework 3.0 wurde auch eine automatisierte Einrichtung von Nutzeraccounts am Portal entwickelt – verbunden mit einer automatischen Zuweisung von Rollen und Nutzungsrechten. Das Portal wird sollte sich dazu automatisch mit dem D-Grid VOMRS abgleichen, ähnlich wie dies für die Einrichtung von Nutzeraccounts und Gridmappings auf den Gridressourcen geschieht. Für das Portal wurden hierfür im Berichtszeitraum Eingriffe in das Usermanagement des Gridsphere Frameworks und entsprechende Anpassungen vorgenommen. Diese Bestandteile des Portals wurden auch fertig gestellt und erprobt.

Im Zuge der geplanten Migration auf das Gridsphere 3 Portal Framework im Rahmen der Projektaufstockung erfolgte auch eine Erweiterung des Loginportlets die es erlaubt, sich am Portal per im Browser installierten Nutzerzertifikats anzumelden. Darüber hinaus wurde eine Selbsregistrierungsmöglichkeit am Portal hinzugefügt, die MediGRID VO-Mitgliedern mit gültigem Zertifikat die Selbstregistrierung am Portal ermöglicht. Weiterhin wurde die Möglichkeit eines Gastnutzerzugangs eingerichtet.

Grid Workflow Execution Service

Durch die Weiterentwicklung des *Grid Workflow Execution Services* von Fraunhofer FIRST zum generischen *Workflow-Management-System* können nun komplexe Prozesse, welche verschiedene IT-Anwendungen in MediGRID miteinander verknüpfen, schneller und einfacher auf heterogenen, verteilten D-Grid-Ressourcen ausgeführt werden. Damit Nutzer Workflows ohne spezielle Kenntnisse der Rechnerumgebung modellieren und ausführen können, liegen der Entwicklung folgende Anforderungen zugrunde:

- Einfache aber zugleich formale Modellierung von Prozessen
- Abstrakte Prozessbeschreibungen sollen in unterschiedlichen Infrastrukturen wiederverwendet werden können
- Automatische Abbildung der Prozesse auf die aktuell verfügbaren Dienste und Ressourcen
- Fehlertolerante und interaktive Ausführung und Kontrolle der Prozesse
- Berücksichtigung von dynamischen Prozessen, bei denen z.B. Entscheidungen erst während der Laufzeit gefällt werden oder sich die Struktur der Prozesse im Laufe der Abarbeitung ändert
- Geeignete Verknüpfung der Prozesse mit Daten, d.h. Modellierung von Kontroll- und Datenflüssen
- Einfache Ableitung von IT-Prozessen aus Geschäftsprozessen, sowohl konzeptionell, als auch technisch
- Unterstützung von serviceorientierten Architekturen und Grid-Computing-Umgebungen

Das Grid-Workflow-Management-System besteht aus mehreren Diensten, deren Kerndienst der „*Grid Workflow Execution Service*“ (GWES) ist, welcher Workflows automatisch, persistent und fehlertolerant auf verteilten IT-Ressourcen ausführt. Neben dem GWES besteht das Grid-Workflow-Management-System aus einer Ressourcen- und Workflow-Datenbank, einem *Re-*

sourceMatcher zur Abbildung von abstrakten Job-Anfragen auf verfügbare Hardware- und Software-Ressourcen, einem Monitoring-System, welches über verteilt installierte *ResourceUpdater* alle 10-20s Informationen über die aktuelle Auslastung der Zielsysteme erhält, sowie einem Scheduler, welcher die Auswahl der Zielsysteme optimiert. Das Grid-Workflow-Management-System ermöglicht somit neben einer automatischen Ausführung von Workflows ein Meta-Scheduling von Anwendungen, welches sich im Produktivbetrieb bereits vielfach bewährt hat und bei dem die Anwender keine besonderen Kenntnisse der Grid-Infrastruktur benötigen.

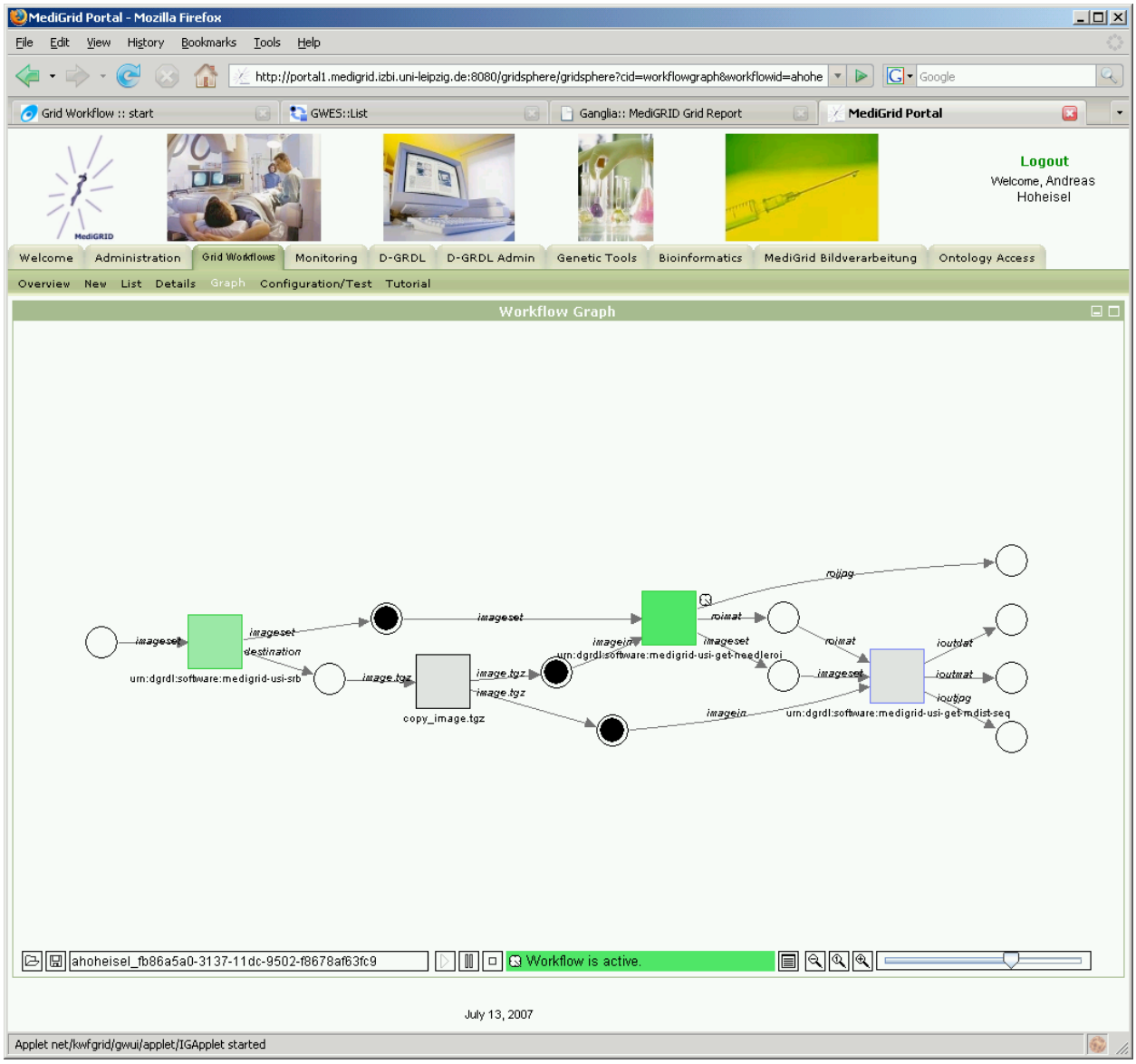


Abbildung 15. Screenshot der im MediGRID-Portal integrierten grafischen Nutzerschnittstelle zur Ausführung und Monitoring von Workflows

Die Prozessmodellierungssprache des Grid-Workflow-Management-Systems basiert auf Petrinetzen, welche eine formale Analyse, Verifizierung und Optimierung der Workflows ermöglichen. Hierbei werden Daten durch Marken (ausgefüllte Kreise) dargestellt, die auf Stellen (Kreise) liegen können. Workflow-Aktivitäten, wie zum Beispiel das Ausführen einer Anwendung oder der Aufruf eines Web-Services werden durch Transitionen dargestellt (Quadrate).

Durch die Pfeile, welche Stellen mit Transitionen bzw. Transitionen mit Stellen verbinden, wird der Kontroll- und Datenfluss festgelegt. An Transitionen können zusätzliche Bedingungen gestellt werden, die – neben der Existenz von Marken auf den Eingabestellen – für das Eintreten der entsprechenden Aktivität erfüllt sein müssen. Pfeilanschriften dienen der Zuordnung von Marken (Daten) auf die entsprechenden Parameter. Nach dem Eintreten einer Transition wird jeweils eine Marke von den Eingabestellen entfernt und jeweils eine neue Marke mit den entsprechenden Ausgabedaten auf die Ausgabestellen gelegt.

Weitere Dokumentation zum GWES ist auf der GWES-Homepage unter http://www.gridworkflow.org/gwes zu finden. Eine Anleitung zur Bereitstellung von Anwendungen für die Zielgruppe der Anwendungsentwickler wurde auf der MediGRID-Homepage unter http://www.medigrid.de/downloads/MediGRID_application_deployment_1-2.pdf bereitgestellt.

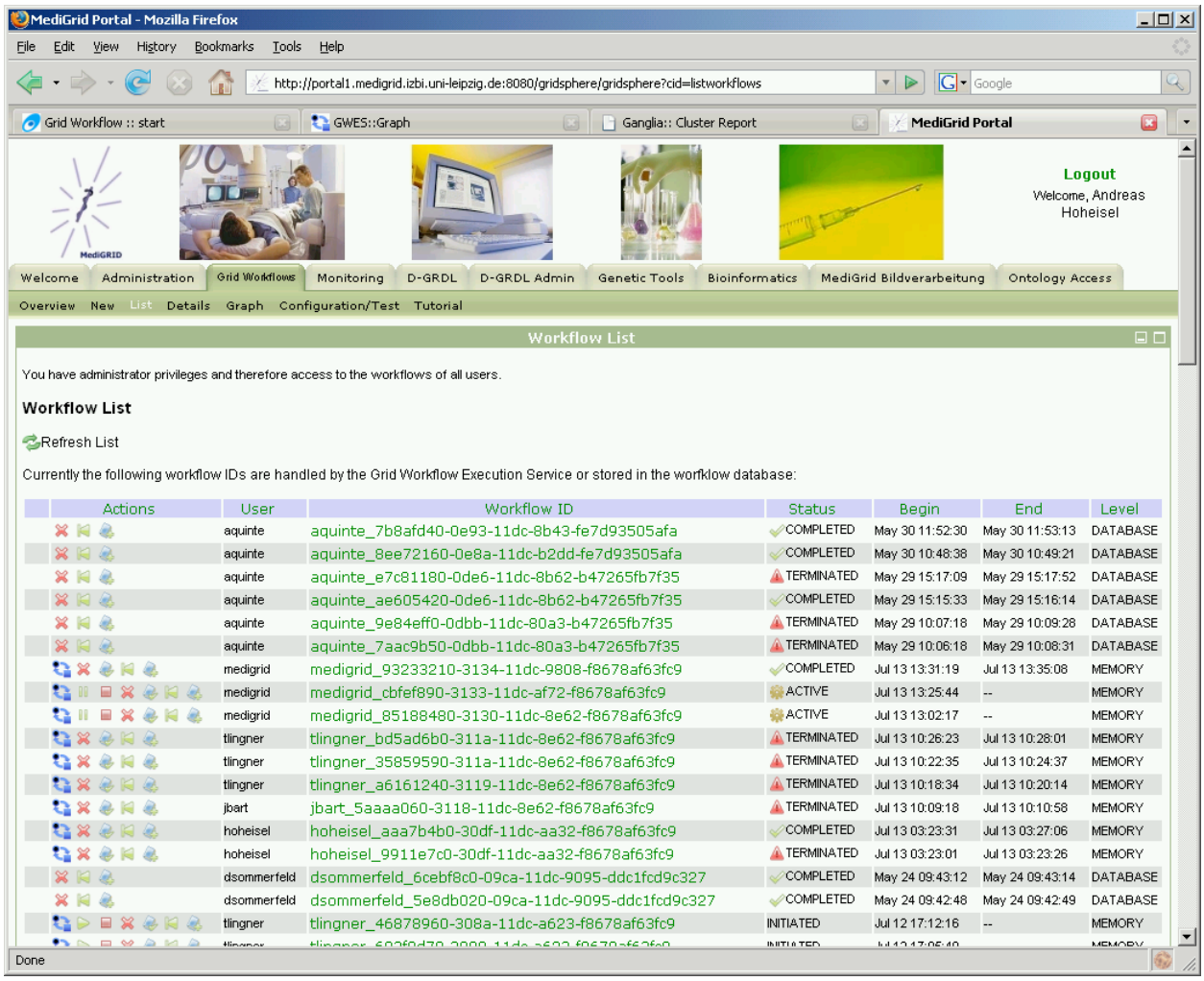


Abbildung 16. Screenshot der Nutzerschnittstelle für die Verwaltung von MediGRID-Workflows

Im Rahmen von MediGRID wurde durch Fraunhofer FIRST der Grid Workflow Execution Service GWES zusammen mit einigen Zusatzdiensten eingerichtet, an die Gegebenheiten von MediGRID und D-Grid erweitert bzw. angepasst sowie betrieben. Die eigentliche Kernsoftware wurde dabei von Fraunhofer FIRST als Eigenleistung zu Beginn in das Projekt eingebracht. Die freie Nutzung der Software durch die Projektpartner während der Projektlaufzeit wurde über den Konsortialvertrag geregelt.

Folgende wesentliche Anpassungsarbeiten wurden im Rahmen von MediGRID am GWES und den damit verbundenen Diensten durchgeführt:

- Anpassung des GWES an die D-Grid-Infrastruktur, Kopplung mit ResourceMatcher, Montoring-Werkzeugen (ResourceUpdater) und D-GRDL-Datenbank (siehe DGI-Projekt).
- Übernahme und Anpassung des im Projekt „Instant-Grid“ entwickelten einfachen Meta-Scheduling-Ansatzes. Während in Instant-Grid hingegen lediglich „Fork“-Ressourcen betrachtet wurden, wird in D-Grid überwiegend „PBS“ als lokales Ressourcen-Managementsystem verwendet. Durch die Anpassungen wurde für MediGRID ein robustes und schnelles Meta-Scheduling von Workflow-Aktivitäten auf viele in D-Grid verteilte Cluster ermöglicht.
- Ein solches Meta-Scheduling erfordert jedoch eine Überwachung der Grid-Knoten, welche über die standardmäßig in D-Grid installierte MDS4-Infrastruktur hinausgeht. Um insbesondere auch kurzfristige Änderungen von Eigenschaften wie zum Beispiel der Auslastung (Load oder Queuelänge) und dem freien Speicherplatz von Grid-Knoten zu erfassen, wurde eigens ein zur D-GRDL kompatibler „ResourceUpdater“ entwickelt und auf allen derzeit verfügbaren MediGrid-Knoten installiert. Dieser erfasst den lokalen Zustand des Grid-Knotens bzw. des Clusters und aktualisiert die entsprechenden D-GRDL-Einträge in der Ressourcendatenbank. Dabei werden sowohl langlebige (quasi-statische) als auch kurzlebige (dynamische) Eigenschaften überwacht, wie zum Beispiel die Anzahl von Jobs in Queues (wartend, aktiv, gesamt), Anzahl von Prozessoren, Load, Netzwerkverbindungen, Speicherauslastung, Festplattenauslastung, etc.
- Verbesserung der Skalierbarkeit des Scheduling-Ansatzes: Es wurden Anpassungen vorgenommen, damit der GWES in einem gewissen Zeitfenster die Jobs aller Transitionen aller Workflows sammelt und sie unter Berücksichtigung der Prioritäten gemeinsam auf die verfügbaren Ressourcen abbildet. Dadurch konnte zum einen die Skalierbarkeit des GWES für viele Workflows erhöht werden (zuvor hat jeder Workflow ein eigenes Scheduling durchgeführt), zum anderen werden jetzt auch Prioritäten zwischen den Workflows berücksichtigt.
- Zudem wurde der GWES um die fehlertolerante Ausführung von Workflows, sowie um die Fähigkeit zur Erstellung und Wiederaufnahme von Workflow-Checkpoints erweitert, sodass ein zuverlässiger Betrieb auch für lang laufende Workflows gegeben ist.
- Des Weiteren wurde eine Methodik zum automatischen Installieren und Ausführen von Programmen im Rahmen von Workflows konzipiert und realisiert, wodurch der Aufwand bei der Verwendung insbesondere kleinerer Programme in Workflows merklich reduziert wurde.
- Kopplung von WS-GRAM-Jobs und Web Services: Es wurden Hilfsoperationen implementiert und als Web-Service zur Verfügung gestellt, die es ermöglichen, in einem Workflow WS-GRAM-Jobs mit normalen SOAP-Methodenaufrufen zu verknüpfen.
- Weiterentwicklung der „Grid Workflow Description Language“ (GWorkflowDL). Wesentliche Änderungen:
    - Einführung von Lesekanten (readPlaces - Marken werden nach dem Schalten NICHT entfernt)
    - Neue Syntax für „operations“ (gemeinsame „operationClass“ für WS-Methodenaufrufe und WS-GRAM-Kommandozeilenprogramme)
    - Annotationen für Marken

- Erstellung eines Konverters, der Workflow-Dokumente im alten Format automatisch in das neue Format überführt. Diese Funktion sichert die Abwärtskompatibilität des GWES zu älteren GWES-Versionen.
- Verbesserung der Ausführung von Programmen auf den Grid-Knoten: Kommandozeilenprogramme auf Grid-Knoten werden mittelbar durch ein zwischengeschaltetes Skript (gwes-command-line-operation.sh) gestartet. Somit entfallen einige Hilfs-Gram-Jobs (z.B. anlegen temporärer Verzeichnisse) und der Overhead (Queue-Wartezeit + WS-GRAM-Overhead) konnte nahezu halbiert werden.
- Räumliche Co-Allocation: Über spezielle Annotationen der Workflows kann jetzt bei Bedarf sichergestellt werden, dass mittel- oder unmittelbar aufeinanderfolgende Aktivitäten auf derselben Ressource ausgeführt werden.
- Gruppieren von Daten bei Parameterstudien: Über Annotationen von Marken können jetzt Daten bzw. Marken, die auf unterschiedlichen Eingabe-Stellen liegen, bestimmten Datensätzen zugeordnet werden, sodass auch die Daten zusammen verarbeitet werden, die zusammen gehören. Dies vereinfacht unter Anderem die Modellierung von komplexen Parameterstudien erheblich.

Neben den Anpassungsarbeiten wurden zudem Arbeiten im Bereich Beratung, Betrieb und Test durchgeführt:

- Entwicklung eines Vorgehensmodells zur Bereitstellung von Anwendungen, auf die per Workflow-Management zugegriffen werden soll.
- Der GWES wurde redundant auf den beiden Portalrechnern zusammen mit einigen zusätzlichen Diensten (ResourceMatcher, Ressourcendatenbank, Graph-Layout, etc) zunächst in Göttingen beim GWDG und später auf den Sonderinvestitions-Rechnern der Universität Leipzig installiert und über eine für MediGRID überarbeitete graphische Benutzeroberfläche im GridSphere-Portal zur Verfügung gestellt und überwacht.
- Neben den Tests von MediGRID-spezifischen Anwendungen (siehe Arbeitspaket 5) wurden umfangreiche technische Workflow-Tests zur Qualitätssicherung definiert und mindestens einmal pro Tag automatisiert ausgeführt. Zusätzlich wurden Workflow-Lasttests zur Messung des Speedups realisiert. (siehe auch Kapitel „Test der MediGRID-spezifischen Anwendungen“)
- Weiterhin wurde Support für Anwendungsentwickler in Fragen bezüglich Grid-Workflows geleistet, und eine Vielzahl von Beispiel-Workflows bereitgestellt.

User- und VO-Management

Zu den Arbeiten des Moduls Middleware gehörte die Nutzerverwaltung am MediGRID Portal, das Management der MediGRID VO sowie der MediGRID Gruppen innerhalb der VO und im späteren Projektverlauf auch das Anlegen von Nutzeraccounts im Grid-Datenmanagementsystem SRB. Letzteres wurde aus Effizienzgründen mit übernommen um eine vollständige Abstimmung und Koordination der verschiedenen Nutzermanagement-Aufgaben in den unterschiedlichen Teilsystemen zu erreichen. In MediGRID erfolgt inzwischen bei Anmeldung an der VO eine automatische Registrierung am Portal und am Storage Resource Broker. Dabei wird die Zugehörigkeit zu Gruppen innerhalb der VO in diesen Systemen berücksichtigt.

In der MediGRID VO existieren im Moment folgende Gruppen:

- bi: Bioinformatik
- bv: Bildverarbeitung
- kf: Klinische Forschung

- on: Ontologien
- rp: Resource Provider (Ressourcenfusion)
- sv: Services@MediGRID

Darüber hinaus wurde in Abstimmung mit Fachgebiet 1-10 des DGI und mit dem Modul e-Science in MediGRID eine Umsetzung der MediGRID Security Policies umgesetzt indem auf die VO zugeschnittene MediGRID Security Policies entwickelt wurden. Die MediGRID Nutzer stimmen diesen Nutzungsbedingungen bei der VO-Registrierung explizit zu. Die Nutzungsbedingungen unterscheiden sich von den Standard-Nutzungsbedingungen, was aufgrund der besonderen Datenschutzthematik im biomedizinischen Umfeld erforderlich ist.

MediGRID Proxy Upload Tool

Die MediGRID Community ist aufgrund ihres Nutzerprofils darauf angewiesen, einerseits einen möglichst einfachen und intuitiven Zugang zu den Grid-Ressourcen und vor allem den Anwendungsinterfaces zu bieten und andererseits besonders hohe Sicherheitsanforderungen zu erfüllen. Aus diesem Grunde wurde der Zugang zu fast allen MediGRID Anwendungen portalbasiert umgesetzt. Das heißt, dass der Nutzer sich nicht direkt auf einer Globus-Ressource einloggen kann und insofern das Problem hat, sein Proxy – das für die Nutzung von Globus-Ressourcen erforderlich ist – in das Grid hochzuladen. Eine Lösung wurde zunächst am CCLRC in Großbritannien gefunden – das MyProxy Upload Tool. Diese Java Webstart Anwendung bietet im Wesentlichen die Möglichkeit vom lokal auf dem Nutzerrechner vorhandenen D-Grid Schlüsselpaar ein Proxy zu erstellen und dieses auf den MediGRID MyProxy Server hochzuladen.
In der Praxis war die Bedienung des Tools aufgrund der großen Vielzahl an Konfigurationsmöglichkeiten und –notwendigkeiten allerdings so kompliziert, dass die Endnutzer nicht damit zurecht kamen. Da das Tool nicht quelloffen zur Verfügung stand und eine entsprechende Anpassung von Seiten der Entwickler nicht möglich war wurde vom Modul Middleware ein ähnliches Tool – das MediGRID Proxy Upload Tool – entwickelt, bei dem die Konfiguration bereits komplett für die MediGRID Infrastruktur voreingestellt ist und den Nutzer somit von der komplizierten Konfiguration entlastet. Dieses Tool konnte die Praxistests schließlich erfolgreich bestehen und ist im produktiven Einsatz in MediGRID. Es ermöglicht zudem den Einsatz des im Browser üblichen Zertifikatsformats PKCS12, so dass die sonst für die Grid-Nutzung erforderliche Konvertierung von Zertifikaten ins PEM-Format entfallen kann. Dies führte zu einer weiteren Entlastung der Endnutzer.

Developer- und User-Support

Zur Unterstützung der Ressourcen- und Service-Provider sowie der Anwender in MediGRID fand eine Erweiterung des elektronischen Informationsangebots für die MediGRID Anwender im MediGRID Wiki statt. Darüber hinaus wurden zur Verbesserung der Nutzerführung im MediGRID Portal Bedienungsanleitungen zu allen eingebundenen Anwendungen integriert.

Im Verlauf des Projekts erfolgte in Zusammenarbeit mit der Support-Koordinierungsgruppe (SuKo) des DGI der Aufbau von nachhaltigen Support-Strukturen für MediGRID. Dazu wurden zunächst die Verantwortlichkeiten für Anwendungen, Dienste, Infrastrukturdienste und Ressourcen geklärt und eindeutig zugewiesen.

Auf dieser Basis wurden Mailinglisten zu allen Anwendungen, Middlewarediensten und Infrastrukturkomponenten in MediGRID eingerichtet um die Kommunikationswege im Support von Personalwechseln zu entkoppeln. Es erfolgte eine Anbindung der MediGRID Supportstrukturen an die D-Grid User Support Dienste (DGUS) und eine Erweiterung des DGUS-Systems durch das DGI auf Basis der Anforderungen in MediGRID.

Die Supportstrukturen wurden auch mit dem Projekt Services@MediGRID abgestimmt, welches die MediGRID Infrastruktur mit nutzt.

2.4.2.4 Test der MediGRID-spezifischen Anwendungen

Die Integration der MediGRID-spezifischen Anwendungen betraf die MediGRID Middleware während des Projekts in mehrerlei Hinsicht. Zum einen mussten die Anwendungsinterfaces zu allen Anwendungen zunächst ins MediGRID Portal integriert und anschließend bis zum Erreichen des fehlerfreien und stabilen Betriebs getestet werden. Darüber hinaus musste eine Abstimmung mit den über das Portal bereitgestellten Grid-Basisfunktionen wie Credential Management, Globus-Anbindung und SRB-Anbindung erfolgen.

Ein weiterer wesentlicher Punkt bei der Integration war die Erstellung von Workflowbeschreibungen und die anschließenden Tests und Fehlerbehebungszyklen beim Deployment und der Integration der Anwendungen ins MediGRID. Im Verlauf des Projekts wurden acht Anwendungen aus den Bereichen Biomedizinische Informatik, Medizinische Informatik, Klinische Forschung und Biomedizinische Ontologien in das Portal integriert und somit weltweit für registrierte Benutzer zugänglich gemacht.

Aufgrund der permanenten Weiterentwicklung der Anwendungen, der Anwendungsinterfaces (Portlets) und der damit verbundenen steigenden Komplexität der Workflows wurden laufend Testzyklen und Debuggingzyklen durchlaufen. Die daraus resultierenden Anforderungen an die Portalinfrastruktur, die Authentifizierungsinfrastruktur (z.B. Proxy Upload Tool) und den Grid Workflow Execution Service flossen zurück ins Arbeitspaket „Realisierung und Konsolidierung“ und triggerten dort die kontinuierliche Weiterentwicklung und Optimierung der MediGRID Middleware Tools.

Zusätzlich zu den acht bereits erwähnten Anwendungen wurde im Verlauf des Projekts eine weitere Anwendung im MediGRID verfügbar gemacht. Es handelt sich dabei um die Anwendung „SequCorr“ zur Korrelationsanalyse biologischer Daten aus dem Modul Bioinformatik. Bei der Anwendung zur Korrelationsanalyse von Gensequenzdaten handelt es sich um die bisher einzige MediGRID Anwendung, die ohne Portal auskommt und bei der die Nutzer über einen direkten Zugriff auf Shell-Ebene arbeiten. Die Jobs dieser Anwendung sind hochparallelisiert und feingranular. Um die Virtualisierungsmechanismen im Zusammenhang mit dem MediGRID Workflow Execution Service zu nutzen wurde für die entsprechenden Anwender ein gesonderter Zugang zum MediGRID Portal per gsissh eingerichtet. Darüber hinaus wurden spezielle Skripte angefertigt mit denen Jobs an den Workflow-Service übergeben werden können, so dass im Anschluss die Vorteile der Ressourcen-Virtualisierung ebenso wie bei den bereits bestehenden Anwendungen genutzt werden können.

Die MediGRID Anwendungen wurden in Zusammenarbeit mit den jeweiligen Anwendungsmodulen umfangreichen Tests unterzogen. Auf Basis dieser Tests werden die MediGRID Anwendungen iterativ weiter entwickelt und verbessert. Wie aus den Integrationstests in AP6 auch hervorgeht (siehe auch Abbildung 4) wechseln sich Test und Weiterentwicklungsphasen in kurzen Abständen ab. Die Integrationstests für die Anwendungen umfassen neben Tests der ins Grid-Portal integrierten Interfaces auch die reibungslose Integration ins Workflowsystem sowie Tests der Basismiddlewarekomponenten.

Die Logging- und Monitoring-Funktionen des Workflowservices boten dabei wertvolle Möglichkeiten, die Ausführung komplexer Workflows im MediGRID nachzuvollziehen und Fehlerquellen in allen Bereichen der Ressourcen- und Middlewareinfrastruktur zu identifizieren. Davon wurde im Rahmen der Integrationstests ausführlich Gebrauch gemacht.

Die Graphiken in Abbildung 17 bis Abbildung 19 zeigen exemplarische Screenshots der ins MediGRID Portal integrierten Anwendungen.

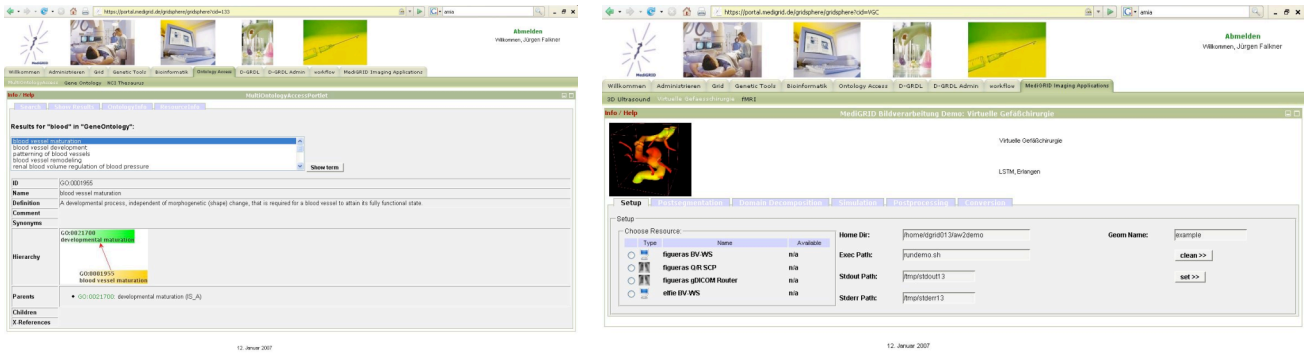


Abbildung 17. MediGRID Anwendungen a) „Ontologie-Anwendung“, b) „virtuelle Gefäßchirurgie“

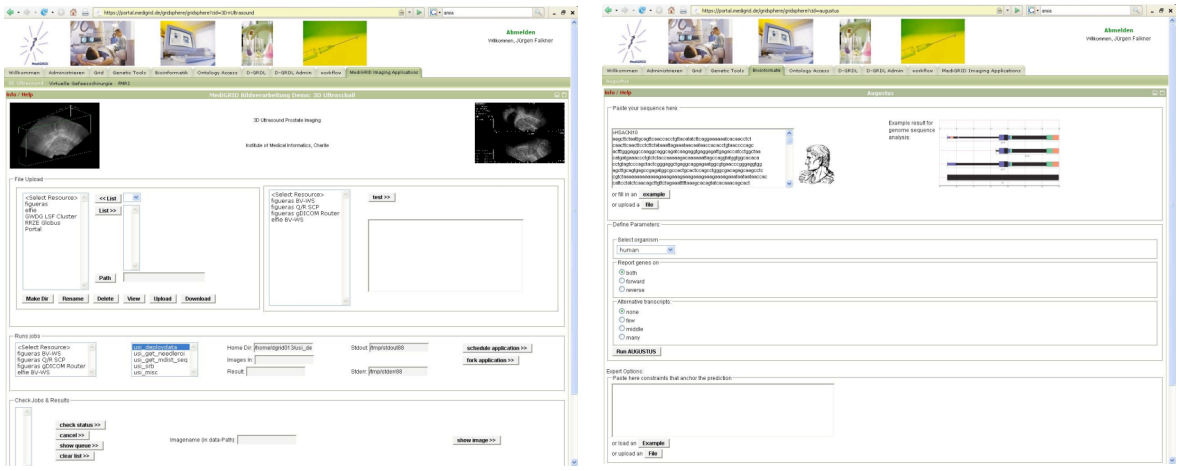


Abbildung 18. MediGRID Anwendungen a) „3D Ultraschall Prostatabiopsie“, b) „AUGUSTUS Gensequenzanalyse“

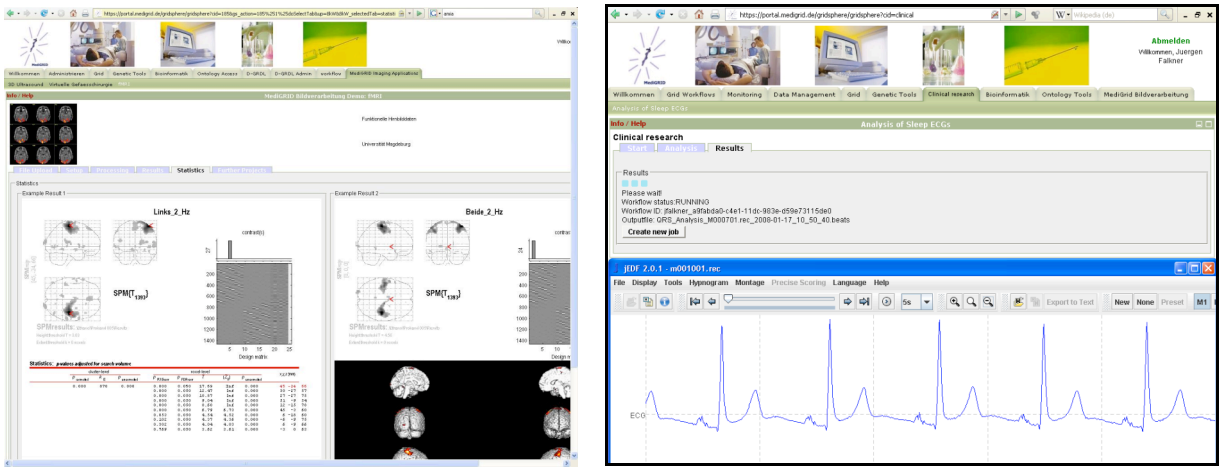


Abbildung 19. MediGRID-Anwendungen a) „Funktionelle Hirnbilddaten“, b) QRS-Analyse von Schlaf-EKGs

2.4.2.5 Integrationstests

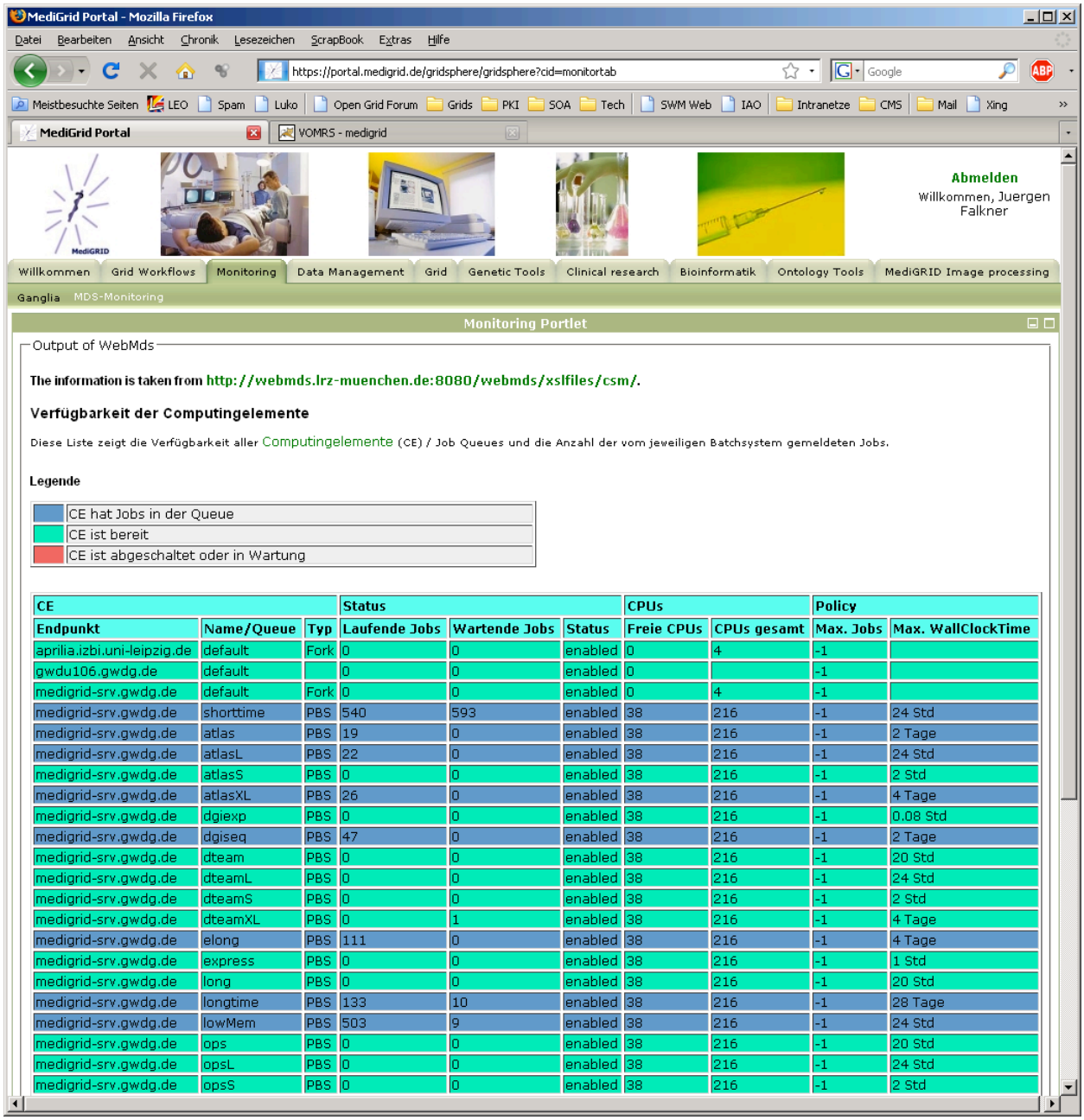
Als Basis der Integrationstest wurden zunächst im Zuge der Realisierung (AP4) diverse Überwachungs- und Monitoring-Systeme eingerichtet. Neben dem D-Grid-weiten Monitoring auf MDS-Basis kamen in MediGRID zusätzlich das Clustermanagement Ganglia, welches auch eine graphische Auswertung der Ressourcennutzungsstatistiken ermöglicht und darüber hinaus NAGIOS zur Diensteüberwachung sowie das GWES-Monitoring für Grid-Workflows zum Einsatz.

Die Monitoringsysteme MDS und Ganglia dienten zur Überwachung der MediGRID Hardware-Ressourcen bzgl. Auslastung und Verfügbarkeit (siehe Abbildung 20 und Abbildung 21).

Abbildung 20. Monitoring in MediGRID mittels Ganglia

Während das Ganglia Monitoring für die Überwachung des Zustands von einzelnen Ressourcen und zur schnellen Erfassung von Zustandsstatistiken genutzt wurde, bietet das MDS Monitoring detaillierte Informationen zur Hardwareausstattung einzelner Ressourcen sowie zu den Batch-queues und deren Füllständen.

Zum Test und Monitoring der Workflow-Dienste sowie der Verfügbarkeit der Workflow-Anwendungen wurde bei Fraunhofer FIRST ein Überwachungssystem auf Basis von „Nagios“ eingerichtet, welches in regelmäßigen Abständen die jeweiligen Dienste testet. Zum Test von Workflow-Anwendungen in MediGRID wurden für die einzelnen Anwendungen jeweils Referenz-Workflows spezifiziert, die in das Testsystem eingepflegt wurden.



CE			Status			CPUs		Policy	
Endpunkt	Name/Queue	Typ	Laufende Jobs	Wartende Jobs	Status	Freie CPUs	CPUs gesamt	Max. Jobs	Max. WallClockTime
aprilia.izbi.uni-leipzig.de	default	Fork	0	0	enabled	0	4	-1	
gwdu106.gwdg.de	default		0	0	enabled	0		-1	
medigrid-srv.gwdg.de	default	Fork	0	0	enabled	0	4	-1	
medigrid-srv.gwdg.de	shorttime	PBS	540	593	enabled	38	216	-1	24 Std
medigrid-srv.gwdg.de	atlas	PBS	19	0	enabled	38	216	-1	2 Tage
medigrid-srv.gwdg.de	atlasL	PBS	22	0	enabled	38	216	-1	24 Std
medigrid-srv.gwdg.de	atlasS	PBS	0	0	enabled	38	216	-1	2 Std
medigrid-srv.gwdg.de	atlasXL	PBS	26	0	enabled	38	216	-1	4 Tage
medigrid-srv.gwdg.de	dgiexp	PBS	0	0	enabled	38	216	-1	0.08 Std
medigrid-srv.gwdg.de	dgiseq	PBS	47	0	enabled	38	216	-1	2 Tage
medigrid-srv.gwdg.de	dteam	PBS	0	0	enabled	38	216	-1	20 Std
medigrid-srv.gwdg.de	dteamL	PBS	0	0	enabled	38	216	-1	24 Std
medigrid-srv.gwdg.de	dteamS	PBS	0	0	enabled	38	216	-1	2 Std
medigrid-srv.gwdg.de	dteamXL	PBS	0	1	enabled	38	216	-1	4 Tage
medigrid-srv.gwdg.de	elong	PBS	111	0	enabled	38	216	-1	4 Tage
medigrid-srv.gwdg.de	express	PBS	0	0	enabled	38	216	-1	1 Std
medigrid-srv.gwdg.de	long	PBS	0	0	enabled	38	216	-1	20 Std
medigrid-srv.gwdg.de	longtime	PBS	133	10	enabled	38	216	-1	28 Tage
medigrid-srv.gwdg.de	lowMem	PBS	503	9	enabled	38	216	-1	24 Std
medigrid-srv.gwdg.de	ops	PBS	0	0	enabled	38	216	-1	20 Std
medigrid-srv.gwdg.de	opsL	PBS	0	0	enabled	38	216	-1	24 Std
medigrid-srv.gwdg.de	opsS	PBS	0	0	enabled	38	216	-1	2 Std

Abbildung 21. MDS Ressourcen Monitoring im MediGRID Portal

Überwachte Dienste waren hierbei:

- MediGRID-Portal (http, https)
- Ressourcen- und Workflow-Datenbank (eXist-DB, siehe Abbildung)
- GWES: Augustus-Workflow der Uni-Göttingen
- GWES: Workflow zur generischen Ausführung von Programmen in MediGRID
- GWES: Workflow mit Lasttest
- GWES: Bildverarbeitungs-Workflow der Charité
- GWES: Workflow mit Filetransfer auf alle Grid-Knoten von MediGRID (Umfassender RFT-Test)

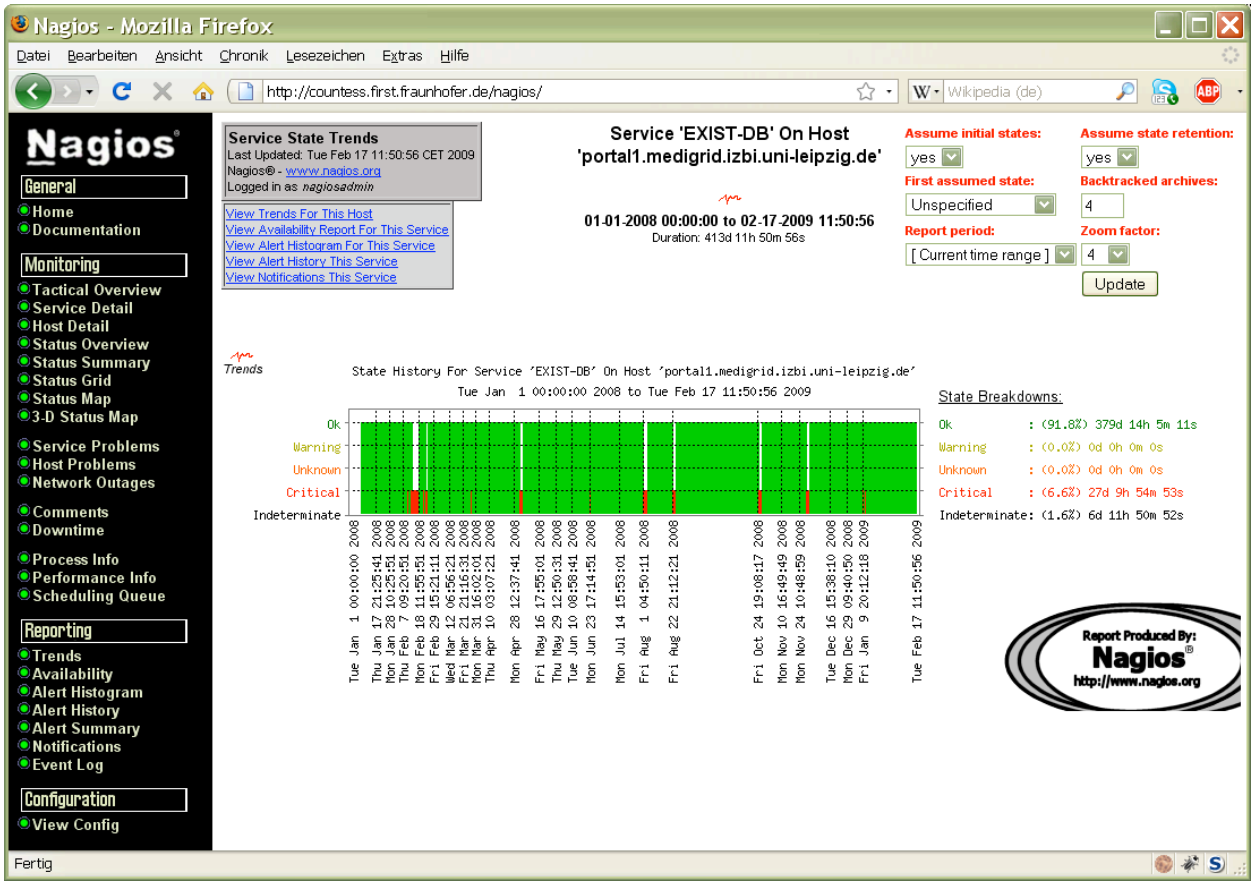


Abbildung 22. Screenshot des Nagios-Überwachungssystems

Abbildung 22 zeigt einen Screenshot aus Nagios, der Verfügbarkeitsstatistiken für das MediGRID Portal darstellt. Wie auf dem Bild zu sehen ist, sind die Wartungsintervalle aufgrund der permanenten Erweiterungen deutlich sichtbar.

Die Workflowübersicht des Grid Workflow Execution Service bietet ein weiteres zentrales Mittel zur Überwachung von Jobs und Jobszuständen. Wie in Abbildung 23und Abbildung 24 zu sehen ist, stehen über den GWES detaillierte Informationen zu den Jobs, zu Jobszuständen, Fehlermeldungen, Ausführungsorten usw. zur Verfügung. Anhand dieser Daten können bei der Integration von Anwendungen und Diensten, aber auch bei der Integration neuer Ressourcen, der Einrichtung neuer Batchqueues etc. wertvolle Informationen auf einfach Weise eingesehen werden. Auf diese Weise lassen sich Integrationstests sehr effizient durchführen und Fehlerquellen können vergleichsweise einfach identifiziert und beseitigt werden.

Weiterhin kann mittels des Grid Workflow Execution Service verfolgt werden in welchem Zustand sich ein einzelner Workflowschritt befindet (siehe Abbildung 15).

ie Integrationstests wurden für alle Ressourcen und Middlewarekomponenten (Portal, GWES, Datenmanagementsystem SRB, Monitoringkomponenten, User- und VO-Management, Credential Management und Proxy Upload, etc.) sowie für die Anwendungen und Anwendungsinterfaces (siehe auch Abschnitt 2.4.2.4) durchgeführt. In einem iterativen Verfahren wurden aufgrund der Erkenntnisse aus der Ressourcen- und Dienste-Überwachung ständige Verbesserungen an allen Bestandteilen durchgeführt und die Quality of Service der Middlewaredienste sowie der Anwendungen und Ressourcen ständig erhöht. Die QoS-Statistiken fließen inzwischen sogar als Auswahlkriterium für Grid-Ressourcen in die Metriken der Broker-Komponente des GWES mit ein.

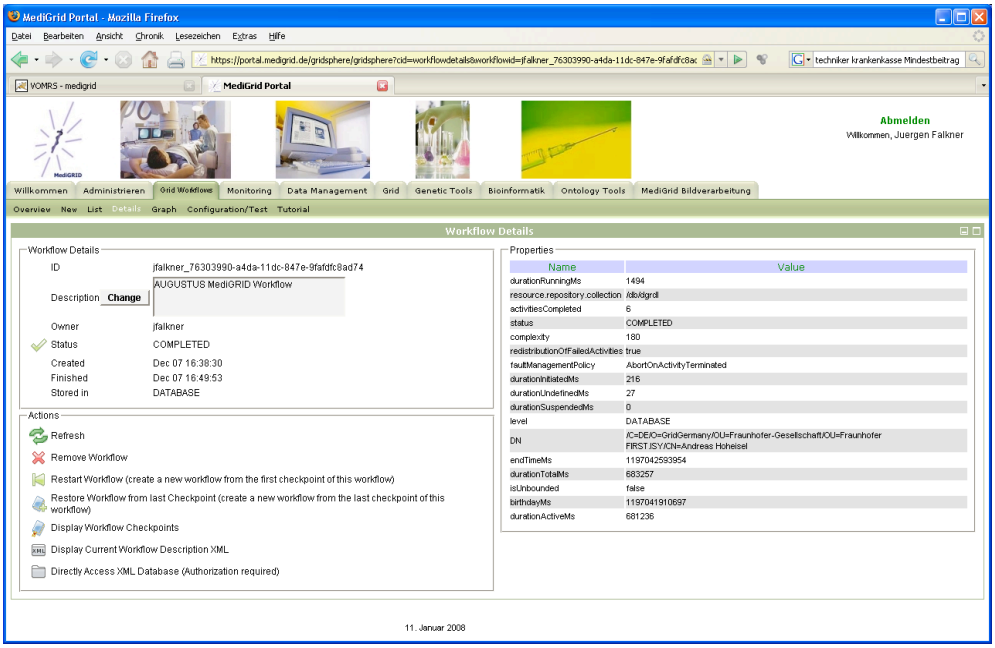


Abbildung 23. Grid Workflow Execution Service - Jobdetails / Jobmonitoring / Jobmanagement

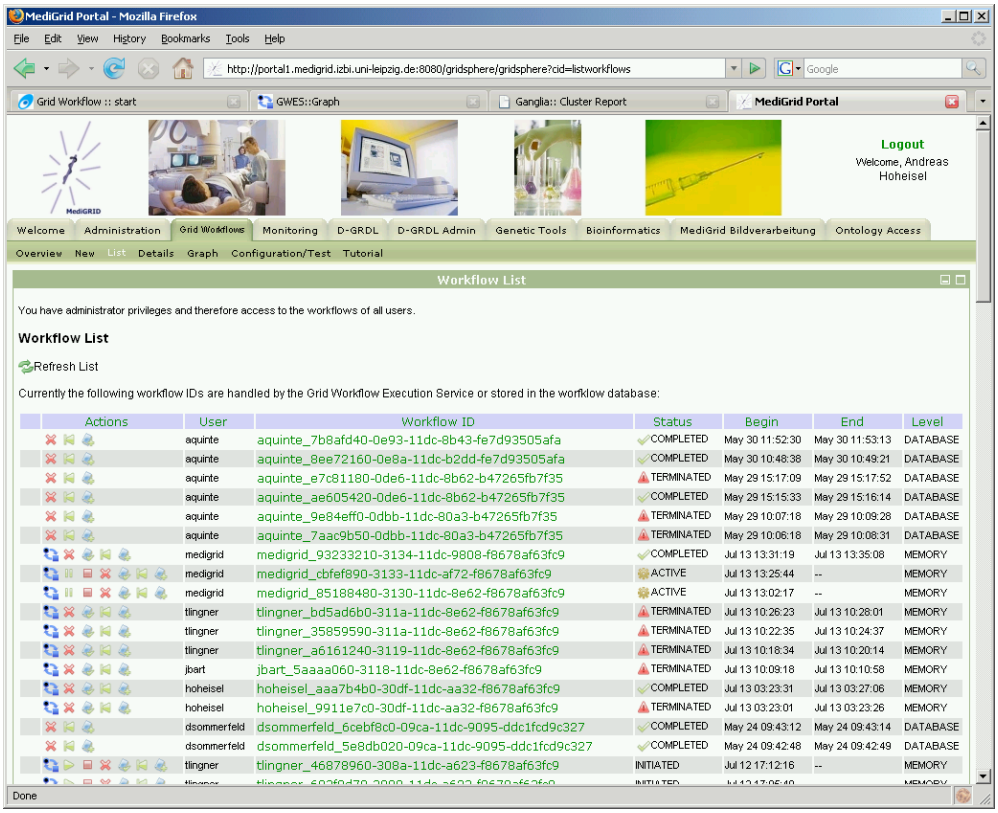


Abbildung 24. Monitoring von Workflows in MediGRID

DEntwicklungen wie die fehlertolerante Ausführung und automatische Re-Submittierung von Workflows im GWES sind direkt auf die umfangreichen Erkenntnisse aus den Integrationstests zurückzuführen.

Ausgehend von den Erkenntnissen der Integrationstests wurden beispielsweise auch die Betriebsstrategien auf den Compute-Ressourcen hinsichtlich der Einführung von Express-Queues auf den MediGRID Ressourcen überarbeitet. Dies diente dazu, kurzlaufende Kleinstjobs (z.B. zur Ausführung von Filetransfers oder zum Ändern von Zugriffsrechten) zeitnah ausführen zu können, da diese ansonsten hinter lang laufenden Jobs „gefangen“ wären, was zunächst im Testbetrieb der Fall war.

Zur Vorbereitung der Umstellung des MediGRID Portals auf das Gridsphere Framework 3.0 wurden zusätzliche Integrationstests auf Basis einer Testinstallation gemacht. Die weitgehend bereits ans neue Framework angepassten Portlets und Dienste sowie Erweiterungen des Frameworks wurden intensiven Tests unterzogen um die produktive Einführung vorzubereiten und einen reibungslosen Umstieg zu ermöglichen.

2.4.2.6 Weitere Ergebnisse

- Projektinterne Abstimmung

  Die Abstimmung innerhalb des MediGRID Projektes erfolgte über zahlreiche persönliche Meetings, sowie über regelmäßige Telefon- und Videokonferenzen sowohl im MediGRID Vorstand als auch auf Ebene der MediGRID Entwickler, wo sich Vertreter der Module Middleware, Ressourcenfusion und eScience regelmäßig mit den Betreibern der Anwendungen in MediGRID austauschten. Darüber hinaus wurde über die gesamte Laufzeit umfangreicher telefonischer und E-Mail-Support für die MediGRID Nutzer und Entwickler geleistet. Im Folgenden sind die persönlichen projektinternen Treffen aufgelistet an denen Vertreter des Moduls Middleware teilgenommen haben.

  - Sitzung des MediGRID-Vorstands (Modulleiter) am 21./22.11.2005 in Berlin
  - MediGRID Vorstandssitzung am 02.08.2006 in Berlin
  - Vorstandssitzung im Rahmen des AMIA Annual Symposium 2006 in Washington DC, 11.-15.11.2006
  - Treffen zur Abstimmung der Projektkoordination zwischen dem Modul Middleware und dem Projektkoordinator am 26.09.2006 in Stuttgart
  - 10.01.2007 Vorstandssitzung MediGRID, TMF, Berlin
  - 14.02.2007 Projekttreffen MediGRID, ZIB, Berlin
  - 28.02.2007 Projekttreffen MediGRID, ZIB, Berlin
  - 01.-02.03.2007 Zwischenbegutachtung D-Grid, Präsentation MediGRID, ZIB, Berlin
  - 05.06.2007 Vorstandssitzung MediGRID, Universität Göttingen, Göttingen
  - 14.08.2007 Meeting zur Ressourcenplanung in MediGRID, TMF, Berlin
  - 28.09.2007 Vorstandssitzung, Berlin
  - 18.10.2007 Arbeitsmeeting zwischen den Modulen Middleware, Bildverarbeitung und Klinische Forschung, Charité, Berlin
  - 05.12.2007 Vorstandssitzung, Berlin
  - 23.01.2008 Vorstandssitzung, TMF, Berlin
  - 20.02.2008 Vorstandssitzung, Göttingen
  - 15.04.2008 Vorstandssitzung, Göttingen
  - ??? Konsortialmeeting zum Projektabschluss

- Projektbegutachtung (Review)
    - Am 01. und 02.03.2007 wurde das MediGRID Projekt den D-Grid Gutachtern zum Review präsentiert
- Abstimmung mit anderen Grid-Communities
    - Über die gesamte Laufzeit erfolgte eine Zusammenarbeit mit InGrid hinsichtlich Security Fragestellungen
    - 2008 wurde TextGrid bei der Integration des Grid Workflow Execution Service in die TextGrid Infrastruktur unterstützt
    - Abstimmung mit caBIG™, caGRID, I2B2 (USA) -> siehe Veranstaltungen und Vorträge
    - Abstimmung mit dem Europäischen HealthGrid (EU) -> siehe Veranstaltungen und Vorträge
- Durchführung von Veranstaltungen
    - Eintägiger MediGrid-Workshop am 22.11.2005 in Berlin zur Information der Projektteilnehmer über verfügbare HW-Ressourcen, einzusetzende Anwendungen, die geplante Grid-Middlewares sowie das D-Grid-Zertifizierungsverfahren
    - Workshop zu den Anwendungen Sequenzanalyse und Genomdatenanalyse bei GOBICS in Göttingen, 09.02.06
    - Workshop zu den Anwendungen Medizinische Bildverarbeitung in der Prostatabiopsie und Virtuelle Gefäßchirurgie bei Charité in Berlin, 21.02.06
    - Workshop zu den Anwendungen Protein Structure Prediction Pipeline und RNAi bei ZIB in Berlin, 22.02.06
    - Workshop zur Anwendung Bildverarbeitung von funktionellen Hirnbilddaten an der Uniklinik Magdeburg, 02.03.06
    - Vermittlung der Ergebnisse der Middleware-Architektur-Konzeption an die MediGRID-Anwender im Rahmen eines Workshops am 03./04.07.06
    - MediGRID-Workshop in Rahmen „AMIA Annual Symposium 2006 - Biomedical and Health Informatics“ mit Austausch mit Kollegen aus den US-Projekten caBIG™, caGRID, I2B2 und dem Europäischen HealthGrid
- Teilnahme an Veranstaltungen
    - D-Grid Workshops
    - Koordination der Zusammenarbeit zwischen den Communties MediGRID und TextGrid anlässlich des Workshops "Ontologien im Grid", 9.-10.3, Leipzig (ZIB, ZIH)
    - D-Grid Security Workshop, UKG, Göttingen, 27./28.03.2007
    - D-Grid All-Hands-Meeting, Göttingen, 10.-12.09.2007
    - D-Grid Nachhaltigkeitsworkshop, Berlin, 09.-10.10.2007
    - D-Grid Scheduling-Workshop, Dortmund, 27.11.2007
    - D-Grid IVOM Workshop, Hannover, 19.02.2008
    - D-Grid Billing Workshop, Hannover, 13-14.03.2008
    - 3. D-Grid Security Workshop, Göttingen, 01./02.04.2008

- D-Grid Support-Koordinierungs-Kick-Off-Workshop, Karlsruhe, 03.-04.07.2008
- 2. D-Grid Support-Koordinierungs-Workshop, Karlsruhe, 15.-16.01.2009

Nationale und internationale Veranstaltungen

- Präsentation des MediGRID Projekts bei einem Industrieworkshop auf der ITeG in Frankfurt am Main 01.06.2006
- Paper und Vortrag zu „Integration of Applications in MediGRID“ im Rahmen des Cracow Grid Workshop 2006
- Teilnahme am „AMIA Annual Symposium 2006 - Biomedical and Health Informatics“ in Washington DC, USA, 11.-15.11.2006
- Vorträge „IT-Virtualisierung und Grid Computing“ „Grid Computing - Standards, Architekturen und Anwendungsbeispiele“ im Rahmen des Stuttgarter Softwaretechnik Forums 2006
- Treffen mit Industrievertretern, Frankfurt, 15.03.2007
- International Symposium on Grid Computing (ISGC) 2007, Academia Sinica, Taipei, Taiwan, 26.-29.03.2007
- eHealth Week Berlin, Special Interest Session „HealthGrid”, 18.04.2007, Berlin
- Teilnahme an der German E-Science Conference 2007 02.-04.05.2007 Baden-Baden
- MediGRID Messestand auf der International Supercomputing Conference 2007, 26.-29.06.2007, Dresden
- Open Grid Forum (OGF) 20 , 07.-11.05.2007, Manchester
- MEDNET 2007, Leipzig, 07.-10.10.2007
- MediGRID/caBig Workshop, Berlin, 23.01.2008
- Conhit Messe, Berlin, 10.04.2008
- International Symposium on Grid Computing, Taipei, 11.04.2008
- Forum Virtualisierung und Grid Computing, Stuttgart, 27.05.2008
- Forum Grid, Berlin, 02.07.2008
- MEDICA Messe, Düsseldorf, 19.-22.11.2008

- Publikationen (siehe Berichtsende)
- Deliverables / Dokumentationen / Projektberichte
  - MediGRID - Analyse und Spezifikation der MediGRID Anwendungen
  - MediGRID Resource Usage Policy phase one – development
  - MediGRID-developer/user/guest agreement
  - Guidelines for the Deployment of Applications in MediGRID (Version 1.0 bis 1.3)

2.4.3 Nutzen und Verwertbarkeit der Modulergebnisse

2.4.3.1 Erfindungen und Schutzrechte

Die MediGRID-Middleware wurde zu großen Teilen durch die Integration und Anpassung bereits vorhandener Open-Source-Technologien realisiert, sodass hier die direkte Verwertung von Schutzrechten nicht anwendbar ist. Eine Ausnahme ist der „Grid Workflow Execution Service“ der von Fraunhofer FIRST entwickelt wurde und für den ggf. Lizenzeinnahmen durch die zukünftige kommerzielle Nutzung der Software durch Dritte zu erwarten sind. Die Nutzungslizenz des GWES erlaubt den kostenlosen Einsatz für die eigene wissenschaftliche Forschung und Lehre. Die darüber hinausgehende Nutzung, z.B. für kommerzielle Dienste oder der Bereitstellung von Diensten für Dritte, bedarf hingegen einer gesonderten Lizenzvereinbarung (siehe http://www.gridworkflow.org/kwfgrid/gwes/docs/license.html). Konkrete Verhandlungen diesbezüglich sind z.B. bezüglich der kommerziellen Nutzung des GWES im Rahmen von Services@MediGRID geplant.

2.4.3.2 Wirtschaftliche Erfolgsaussichten nach Projektende

MediGRID Infrastruktur arbeitet zurzeit noch nicht kostendeckend ist aber auf einem guten Weg, die nötigen organisatorischen Strukturen sowie die Technologien dafür aufzubauen. Im Anschlussprojekt Services@MediGRID werden für eine Anzahl von Projekten Geschäftsmodelle entwickelt, die dann auf andere Grid-Anwendungen übertragen werden können. Im Förderungsfall ist durch das Projekt WissGrid für weitere drei Jahre die Weitergabe der Erfahrungen aus dem MediGRID Projekt an neue Anwender und Anwendungsentwickler sichergestellt. In Services@MediGRID werden gegenwärtig die rechtlichen Grundlagen für die Durchführung sowohl eines technischen als auch fachlichen Accountings gelegt indem SLAs zwischen der MediGRID VO einerseits und den Grid Resource Providern und Grid Service Providern andererseits konzipiert und ausgestaltet werden. In der Projektlaufzeit von Services@MediGRID wird voraussichtlich eine prototypische Umsetzung eines durchgehenden Accountings auf allen Ebenen erreicht werden. Auf Grundlage der dabei entstehenden Accountingdaten wird dann grundsätzlich eine Leistungsverrechnung möglich sein. Diese Möglichkeit zur Leistungsverrechnung und Rechnungstellung wiederum wird eine Grundlage bilden um einerseits die direkte Förderung der Nutzung von Anwendungen und Dienstleistungen im MediGRID – und daraus folgende eine auf tatsächlicher Nutzung basierende Bezahlung der Ressourcenbetreiber ermöglichen. Andererseits wird auch der kommerzielle Betrieb von Anwendungen und Dienstleistungen in MediGRID ermöglicht werden. Auf diese Weise kann dann ein nachhaltiger Betrieb der MediGRID Infrastruktur sowie der darin vorhandenen Anwendungen und Dienste gewährleistet werden. Zurzeit ist jedoch nicht abzuschätzen inwiefern zur Produktivstellung noch eine Übergangsfinanzierung über die im Moment absehbaren Folgeförderungen hinaus erforderlich sein wird.

Für den Grid Workflow Execution Service ergeben sich neben den oben genannten Lizenzeinnahmen zusätzliche Verwertungsmöglichkeiten sowohl innerhalb der Medizinbranche als auch in weiteren Anwendungsbereichen. Dies wird durch die weitgehend generisch gehaltene Methodik der Prozessmodellierung und Prozessausführung im GWES unterstützt sodass der Transfer in andere Domänen mit relativ wenig Aufwand möglich ist. So wurde der GWES bereits neben der Medizin zudem in vielen anderen Bereichen, wie zum Beispiel in der Umweltsimulation, Katastrophenmanagement, Ingenieurwesen, Medien, Verkehrsmanagement sowie der Warenwirtschaft eingesetzt. Dadurch erbeben sich zusätzliche Verwertungsmöglichkeiten:

- Einsatz der Methodik (Modellierung, Analyse und Ausführung von nebenläufigen Prozessen in verteilten Umgebungen) in Forschungs- und Beratungsaufträgen für die Industrie.
- Verwendung der Software für verwandte Anwendungsgebiete.

- Gemeinsame Vermarktung mit Partnern (z.B. Fraunhofer IAO, Charité Berlin, Daasi GmbH)
- Einbringen und Verwertung der Technologie und Kompetenz in weitere Projekte. Derzeit beantragt sind z.B. die D-Grid-Call-3-Projekte PneumoGRID, WissGRID und DGSI.

2.4.3.3 Wissenschaftliche und technische Erfolgsaussichten nach Projektende

Die MediGRID Middleware und Infrastruktur konnte im Verlauf des Projekts nicht nur eingerichtet und zu einem Proof-of-Concept geführt werden, sondern hat auch einen beachtlichen Reifegrad erreicht. Dieser äußert sich vor allem im Grad der Virtualisierung und der damit verbundenen einfachen Nutzbarkeit der MediGRID Anwendungen. MediGRID ist darüber hinaus sehr gut mit den zentralen D-Grid Diensten wie User- und VO-Management, Monitoring, Accounting und den Supporteinrichtungen vernetzt.

Insofern bestehen gute Aussichten mit dem MediGRID ein Erfolgsmodell und Beispiel für andere Communities zu etablieren. Technologien wie z.B. der Grid Workflow Execution Service bieten einen bisher nur selten erreichten Leistungsumfang bzgl. der automatisierten Auswahl von Ressourcen, der Abarbeitung von Workflows sowie bzgl. des Fail-over-Managements von Grid Jobs. Im Bereich der Authentifizierung konnte bereits ein sehr hohes Maß an Nutzerfreundlichkeit erreicht werden, das die Grundlage für eine weitere Verbesserung im D-Grid 3 Projekt „Gap-SLC“ bilden wird.

Die Übernahme des Grid Workflow Execution Service durch die TextGrid Community und das allgemein gute Interesse an den Authentifizierungslösungen in MediGRID sind starke Indikatoren für die guten technischen Erfolgsaussichten der MediGRID Middleware Infrastruktur.

Die produktive Bereitstellung einer signifikanten Anzahl an für die biomedizinische Community interessanten Anwendungen und die Sicherstellung des Weiterbetriebs der MediGRID Infrastruktur im Projekt Services@MediGRID sowie über das TMF Forum Grid stellen darüber hinaus die wissenschaftliche Anschlussfähigkeit nach Projektende sicher.

2.4.3.4 Wissenschaftliche und wirtschaftliche Anschlussfähigkeit

Um die wissenschaftliche und wirtschaftlichen Anschlussfähigkeit des MediGRID Projekts und insbesondere der Aktivitäten im Modul Middleware sicherzustellen erfolgten im Laufe des Projekts eine große Anzahl von Aktivitäten um MediGRID und die im Projekt erzielten Inhalte sowohl im wissenschaftlichen Bereich als auch gegenüber Unternehmen zu präsentieren.

Zur Verbesserung der Wahrnehmung in der Öffentlichkeit und zum Aufbau von Kontakten mit Vertretern der Wirtschaft im medizinischen Bereich wurde das MediGRID Projekt bei einem Industrieworkshop auf der Fachmesse ITeG in Frankfurt am Main 01.06.2006 präsentiert.

Im Rahmen der Öffentlichkeitsarbeit wurde MediGRID auf dem Grid-Tag „IT-Virtualisierung und Grid-Computing“ des Stuttgarter Softwaretechnik Forums am 09.11.2006 vorgestellt.

International wurde MediGRID auf dem „International Symposium on Grid Computing“ vom 26. bis zum 29.03.2007 in Taipei (Taiwan) durch einen Vortrag des Moduls Middleware vorgestellt. Auf der German e-Science Conference vom 02. bis zum 04.05.2007 wurde in Baden-Baden die Nutzung von Grid Computing in der Medizin anhand der Anwendungen aus der medizinischen Bildverarbeitung in dem Beitrag „Medical Image Processing in MediGRID“ vorgestellt.

Im Rahmen der „eHealthweek” vom 17. bis zum 19.04.2007 in Berlin wurde MediGRID durch einen Beitrag auf der Telemed 2007 vorgestellt. Der Beitrag „Auf dem Weg zur individualisierten Medizin - Grid-basierte Services für die EPA der Zukunft“ zeigt auf, wie Grid-Services für die elektronische Patientenakte genutzt werden können. Der Beitrag wurde vom Berufsverband

Medizinischer Informatiker e.V. und der Deutschen Gesellschaft für Gesundheitstelematik e.V. als bester wissenschaftlicher Vortrag ausgezeichnet.

Auf der eHealth Conference wurde die Special Interest Session „HealthGrid" am 18.04.2007 vom Projekt MediGRID gemeinsam mit den Organisationen des europäischen HealthGrid organisiert. In der Veranstaltung wurden die Möglichkeiten von Grid Services in der medizinischen Forschung und der Versorgung diskutiert. Das Modul Middleware ist in seinem Beitrag „Interoperability in Grid Computing" auf die Notwendigkeit der Interoperabilität von Daten und Anwendungen in Grid Umgebungen eingegangen.

Auf der Internationalen Supercomputing Conference 2007 vom 26. bis zum 29.07.2007 in Dresden präsentierte das Modul Middleware MediGRID auf dem D-Grid-Stand.

National wurde der Beitrag zu MediGRID „Benutzerfreundliche und sichere Zugänge zu Grid Ressourcen" auf der 21. DFN-Arbeitstagung über Kommunikationsnetze am 01.07.2007 in Kaiserslautern präsentiert. Darüber hinaus erfolgte die Teilnahme am D-Grid Security Workshop vom 27. bis zum 28. 03.2007 an der Universität Göttingen sowie am D-Grid All-Hands Meeting vom 10. bis zum 12.09.2007 in Göttingen und am D-Grid Nachhaltigkeitsworkshop vom 09. bis zum 10.10.2007 in Berlin.

Es wurden in 2007 Treffen mit verschiedenen Industrieunternehmen durchgeführt, auf denen die Nutzung von Grids in der Medizin diskutiert wurde.

Darüber hinaus fanden die bereits vorliegenden Ergebnisse von MediGRID Eingang in die Vorlesung „Softwaretechnik und -management" im Sommersemester 2007 und 2008 und in die Vorlesung „Electronic Business" im Wintersemester 2007/2008, 2008/2009 an der Universität Stuttgart.

International wurde MediGRID auf der MEDNET 2007 vom 07.-10.10.2007 in Leipzig durch das Modul Middleware mit dem Vortrag "MediGRID - Grid Computing for cooperative work in e-Health" vorgestellt.

International wurde MediGRID auf dem International Symposium on Grid Computing in Taipei vom 07. bis 11.04.2008 durch das Modul Middleware mit dem Vortrag "Secure Grid Services for Cooperative Work in Medicine and Life Science" vorgestellt. Darüber hinaus wurden die Arbeiten aus MediGRID internationalen Besuchern präsentiert, wie beispielsweise einer chinesischen Delegation am 11.07.2008 am Fraunhofer IAO und der European Medicines Agency (EMEA) am Fraunhofer IAO am 21.11.2008 in Stuttgart mit dem Vortrag „Short presentation on Grid computing activities".

National wurde MediGRID in 2008 im Rahmen diverser D-Grid Workshops, am Forum »Virtualisierung und Grid Computing« am 27.05.2008 und beim jährlich stattfindenden Stuttgarter Softwaretechnik Forum 25.-28.11.2008 präsentiert.

Darüber hinaus war das Modul Middleware auf den Fachmessen „conhit" in Berlin und „Medica" in Düsseldorf vertreten.

Im Rahmen verschiedener Veranstaltungen wurde MediGRID auch in 2008 Unternehmen und Wissenschaftlern vorgestellt.

Mit Hilfe dieser zahlreichen Aktivitäten erfolgte zum einen der Wissenstransfer aus dem Modul Middleware des MediGRID Projekts direkt in die Lehre an der Universität Stuttgart. Zum anderen wurde über die zahlreichen Kontakte zu Forschungs- und Industriepartnern eine sehr gute Vernetzung erreicht, die neben einer inhaltlichen Abstimmung – beispielsweise mit Europäischen und US-Amerikanischen Aktivitäten – auch den Grundstein für eine stärkere Einbindung von und Kooperation mit Industriepartnern legte. Erste greifbare Auswirkungen dieser Vernetzung sind beispielsweise die Industriebeteiligungen am Projekt Services@MediGRID.

2.4.4 Fortschritte auf dem Markt für Grid Computing in der Medizin und der Bioinformatik aus Modulsicht - Darstellung Ergebnisse Dritter

Im Verlauf des Projekts wurde zunächst außerhalb und später auch innerhalb der D-Grid-Projekte im Security-Bereich eine verstärkte Verbreitung und Akzeptanz von Shibboleth-basierten A&A Infrastrukturen bemerkt. Diese Entwicklungen wurden ernsthaft in Erwägung gezogen und beobachtet, haben jedoch in der Projektlaufzeit keinen ausreichend produktiven Status erreicht um für MediGRID in Frage zu kommen.

Für MediGRID äußerst interessant sind auch die Entwicklungen im Bereich der Authentifizierungs- & Autorisierungs-Infrastrukturentwicklungen im Telematikbereich und im Zusammenhang mit der Einführung der elektronischen Gesundheitskarte und des Heilberufeausweises. Diese wurden und werden aufmerksam verfolgt, konnten jedoch auch zur Projektlaufzeit noch keinen produktiven Zustand erreichen.

Für das VO- und Usermanagement war die Umsetzung des in Fachgebiet 1-10 im D-Grid Integrationsprojekt entwickelten VOMRS Dienstes von großer Bedeutung.

Außerhalb der D-Grid-Projekte wurde im Security-Bereich mit dem MyProxy Upload Tool aus dem Uk e-Science Programm eine Anwendung zum Upload von Proxy-Zertifikaten gefunden, die es ermöglicht, Zertifikate auf einem MyProxy Server zu hinterlegen ohne selbst direkten Zugang zu einem Grid-Rechner haben zu müssen. Dies ist für die rein portalbasierte Nutzung von Grid-Diensten sehr hilfreich. Die Anwendung wurde auch direkt an die Kollegen des Fachgebiets 3 im D-Grid Integrationsprojekt weitergegeben, mit der Bitte, diese auf ihre Sicherheit zu überprüfen. Da der Quellcode der Anwendung von den Entwicklern nicht offen gelegt wurde und neben den Sicherheitsfragestellungen auch erheblicher Bedarf an einer Verbesserung der Usability bestand wurde eine Anwendung mit ähnlicher Funktionalität im Projekt entwickelt. Das MediGRID Proxy Upload Tool ist inzwischen produktiv im Einsatz.

Im Portalbereich wurde die Entwicklung des Gridsphere Portal Frameworks aufmerksam verfolgt. Mit dem Release der Version 3.0 wurden einige Verbesserungen am Basisframework eingeführt. Leider wurden mit der Version 3.0 die Grid Portlets allerdings inkompatibel. Die Verzögerung in der Entwicklung des Vine Toolkits, welches die Grid Portlets ablösen sollte und die Tatsache, dass diese auch zu Projektende noch nicht einwandfrei funktionieren machten allerdings die mehrfach geplante Migration des Portals auf die Gridsphere Version 3.0 unmöglich.

Gleichzeitig erfreut sich im Bereich der Grid Portal Entwicklung das Liferay Opensource Portal Framework (www.liferay.com) mittlerweile großer Beliebtheit. Aufgrund der großen Anzahl an verfügbaren Portlets für dieses Framework und auch aufgrund des zu Liferay erhältlichen kommerziellen Supports durch die Firma Liferay, Inc. ist das Framework eine ernstzunehmende Alternative zum Gridsphere Portal Framework. Ein unmittelbarer Umstieg kam aber aufgrund des dafür notwendigen Aufwands zur Projektlaufzeit nicht mehr in Frage.

Darüber hinaus wurden die Entwicklungen in den Fachgebieten zu Accounting und Billing im D-Grid Integrationsprojekt mit Interesse verfolgt. Gegenwärtig befindet sich die MediGRID VO in Verhandlungen mit dem Betreiber des zentralen D-Grid Accountings (RRZN) zum Abschluss eines Service-Level-Agreements, um die Weitergabe von MediGRID Accountingdaten ans RRZN zu regeln und somit die Voraussetzung für eine Leistungsverrechnung innerhalb der MediGRID VO zu schaffen.

Nicht zuletzt hatte die Beschaffung von Hardware im Rahmen der D-Grid Sonderinvestitionen einen positiven Einfluss auf die Leistungsfähigkeit und Verfügbarkeit der MediGRID Infrastruktur. Trotz des zusätzlichen Integrationsaufwandes waren die Auswirkungen auf das Projekt hierbei ausschließlich positiv.

2.4.5 Ausblick und Perspektiven im Bereich Middleware

Die Entwicklungen, die im Verlauf des Projekts im Modul Middleware sowohl auf technischer als auch auf organisatorischer und prozeduraler Ebene erzielt werden konnten sind sehr positiv. Im MediGRID Projekt wurde eine weitgehend produktive Infrastruktur aus Ressourcen, zentralen Diensten und Anwendungen aufgebaut, die über ein sehr nutzerfreundliches Portal von einer breiten Nutzerbasis im Bereich der Medizin und Bioinformatik wird. Es konnte eine sehr gute Integration mit zentralen D-Grid Diensten wie dem VO- und Nutzer-Management, dem Accounting und Monitoring und dem D-Grid User-Support erreicht werden.

Die Entwicklungen hin zu einer nachhaltigen und langfristig tragfähigen Infrastruktur sind auf einem sehr guten Weg, wenngleich sie innerhalb dieses Pionierprojekts selbstverständlich nicht abgeschlossen werden konnten. Weiterentwicklungspotenziale und –notwendigkeiten bestehen beispielsweise im Bereich der Autorisierung – dort insbesondere im feingranularen Management von Rechten und Rollen sowie bei der entsprechenden Umsetzung an den Ressourcen und Diensten. Auch im Bereich des Single-Sign-On sind mittel- und langfristig noch deutliche Verbesserungen erforderlich, die aber im D-Grid III Projekt „Gap-SLC“ unter Beteiligung von Vertretern des MediGRID Moduls Middleware in Angriff genommen werden.

Im Portalbereich gibt es mittlerweile ernstzunehmende Konkurrenz für das in D-Grid weit verbreitete Gridsphere Portal Framework, die auch unter dem Gesichtspunkt eines nachhaltigen Betriebs und Supports ernsthaft in Erwägung gezogen werden muss. Auch im Bereich der Unterstützung der Leistungsverrechnung besteht noch deutlicher Handlungsbedarf, der allerdings bereits jetzt im Projekt Services@MediGRID und im DGI II in Angriff genommen wird.

Im MediGRID Projekt konnten aber bereits sehr gute Erfolge im Bereich der Virtualisierung von Grid Diensten und Ressourcen und vor allem beim Verbergen der hohen Komplexität dieser Infrastruktur vor dem Endnutzer erzielt werden. Mit insgesamt 9 leistungsfähigen Anwendungen aus den Bereichen Bioinformatik, medizinische Bildverarbeitung, klinische Forschung und Biomedizinische Ontologien konnten in allen angestrebten Themenbereichen wichtige Pilotanwendungen technisch umgesetzt werden und den Mehrwert für die Endnutzer unter Beweis stellen.

Auf dieser Basis werden im Folgeprojekt Services@MediGRID fünf weitere Anwendungen technisch in die MediGRID Infrastruktur integriert. Darüber hinaus erfolgt in diesem Projekt eine zielgerichtete Entwicklung von nachhaltigen Dienstleistungen für die neuen Anwendungen, die auch die detaillierte Ausgestaltung von Geschäftsmodellen einschließt.

Die Erfolge beim Aufbau des MediGRIDs, bei der Integration der Anwendungen und beim Betrieb der Infrastruktur und Dienste haben auch bereits zu positiven Rückmeldungen von Europäischen und US-Amerikanischen Partnern aus dem HealthGrid und aus caBIG geführt.

Darüber hinaus hat beispielsweise der Grid Workflow Execution Service einen Funktionsumfang und Reifegrad erreicht, der dazu geführt hat, dass die D-Grid Community TextGrid in 2008 die Nutzung ihres bisherigen Workflowsystems eingestellt und auf die Nutzung des GWES umgestellt hat (siehe auch nächster Abschnitt).

Wie in Abschnitt 2.4.3.1 beschrieben eröffnet diese Entwicklung auch die Perspektive eines kommerziellen Vertriebs des GWES.

Die Erfolge des Projekts und auch der Arbeiten im Modul Middleware spiegeln sich darüber hinaus in der Zahl bereits genehmigten oder in Begutachtung befindlichen Folgeprojekte wieder. Hierzu zählen die Projekte Services@MediGRID, Gap-SLC, PneumoGRID, WissGrid und DGSI.

Die MediGRID Infrastruktur und ihre Middlewaredienste und Organisationsstrukturen sind auf einem sehr guten Weg. Mit den Folgeaktivitäten ist – im Falle positiver Begutachtung der noch

in der Antragsphase befindlichen Projektanträge – auch für die unmittelbare Zukunft ein Fortbetrieb gesichert. Es bleibt abzuwarten ob bereits im Rahmen dieser Folgeaktivitäten ein nachhaltiger Betrieb ohne die Notwendigkeit weiterer Förderung für Forschung und Entwicklung möglich sein wird. Dies wird letztlich von den Geschäftsmodellen und den Möglichkeiten zur Leistungsverrechnung ebenso abhängen wie von den zukünftigen Fördermodellen für die Endnutzer solcher Infrastrukturen. Wenn die im Betrieb einer Grid Infrastruktur entstehenden Sach-, Betriebs- und Personalkosten von den zum großen Teil wissenschaftlichen Endnutzern getragen werden sollen, muss bis zum Ende der aktuellen Folgeprojekte eine Klärung der zukünftigen Förderung dieser Endnutzer erfolgen.

2.5 Ressourcenfusion

2.5.1 Moduleinordnung und Modulziele

Das Modul Ressourcenfusion ordnet sich im MediGRID-Projekt zwischen den Modulen mit Schwerpunkt IT-Infrastruktur (Module „Middleware“ und „eScience“) und den Modulen der Anwendungsentwickler (Module „Bioinformatik“ und „Medizinische Bildverarbeitung“) ein. Die Tätigkeiten im Modul Ressourcenfusion umfassten sowohl koordinierende und organisatorische Arbeiten zusammen mit den MediGRID-Ressourcenprovidern als auch die Bearbeitung von Forschungs- und Entwicklungsaufgaben im Bereich des verteilen Datenmanagements und Job-Scheduling.

Das Modul Ressourcenfusion hat zwischen Anwendern, Middleware-Entwicklern und Ressourcenprovidern moderiert und generelle Implementierungsfragen der MediGRID-Infrastruktur koordiniert. Damit konnte eine effiziente Bereitstellung der MediGRID-spezifischen Grid-Infrastruktur garantiert werden. Das Modul war zuständig für den praktischen Betrieb der MediGRID-Middleware-Infrastruktur. Als Vertreter der Ressourcenprovider hat das Modul Ressourcenfusion in kooperativer Weise mit dem Middleware-Modul und den Anwendern den Prozess der Auswahl, Definition und Installation von Middleware-Komponenten unter den Randbedingungen der D-Grid-Infrastruktur bei den Ressourcenprovidern begleitet. Gemeinsam mit dem Modul Middleware wurde eine dafür notwendige technische Strategie entwickelt und implementiert sowie allgemeine Aspekte der MediGRID-Anwendungsszenarien herausgearbeitet und in der MediGRID-Plattform umgesetzt.

Auch von Seiten der Forschungs- und Entwicklungsaufgaben im Bereich „Datenmanagement“, „Job-Scheduling“, und „Portal-Implementierung“ gab es naturgemäß eine starke Vernetzung mit dem Modul Middleware.

Das Modul Ressourcenfusion war gegenüber dem DGI Ansprechpartner für technisch-betriebliche Belange.

Um den Koordinations- und Moderatorprozess in Punkt a) möglichst effektiv umsetzen zu können, sind im Modul Ressourcenfusion mit dem ZIB (Berlin), ZIH (TU Dresden) und der GWDG (Göttingen) diejenigen MediGRID-Partner konzentriert worden, die auch hauptsächlich die Rechen- und Speicherressourcen für MediGRID-Anwendungen bereitstellen. Diese MediGRID-Infrastruktur wird zusätzlich durch die Bereitstellung weiterer Ressourcen durch den assoziierten Partner Forschungszentrum Karlsruhe erweitert.

Die ursprünglichen Ziele im Modul Ressourcenfusion wurden bis Projektende verfolgt. Lediglich in der Phase der Erfassung und Systematisierung der Nutzeranforderungen zu Beginn und zur Halbzeit des Projektes stellte sich heraus, dass es keinen Bedarf für die Nutzung des Grid Application Toolkit (GAT) gibt, so dass dieser Arbeitspunkt nicht weiterverfolgt wurde.

Zusätzlich zu den ursprünglichen Projektzielen ergab sich der Bedarf, ein für das MediGRID adäquates Accounting zu implementieren. Die systematische Erfassung der Ressourcennutzung erlangte insbesondere nach der Installation von zusätzlichen Rechen- und Speicherressourcen aus Sondermitteln des BMBF ab 2006 ihre Bedeutung. Da es zu diesem Zeitpunkt keine allgemeine D-Grid-Lösung für ein strukturiertes Accounting gab, ist ein Prototyp für MediGRID im Modul Ressourcenfusion implementiert worden.

Die Arbeiten im Modul Ressourcenfusion fokussierten sich auf die folgenden Aufgabenschwerpunkte:

a) die Koordination, Abstimmung von betrieblichen Aspekte zwischen Grid-Ressourcenprovidern und Entwicklern/Nutzern sowie der Betrieb des MediGRID,

b) die Konzeption, Implementierung und Betrieb des verteilten Datenmanagements im MediGRID,

c) die Optimierung des Job-Scheduling für datenintensive Workflows in MediGRID, und

d) der Betrieb des MediGRID-Portals.

Die unter Punkt a) zusammengefassten Aufgaben mit ihrem organisatorisch-administrativen Schwerpunkt wurden bereits oben beschrieben.

Der Hauptschwerpunkt der Forschungs- und Entwicklungsaufgaben war die Implementierung eines „Data Grids“ in MediGRID (Punkt b)), um einen effizienten Zugriff auf die im Grid verteilten Daten zu ermöglichen. Die Nutzerinformationen können dabei in lokalen relationalen Datenbanksystemen oder als Dateien in hierarchischen Dateisystemen organisiert sein. Dabei erfordern die verschiedenen MediGRID-Anwenderszenarien je nach ihrer Charakteristik bezüglich Zahl und Größe der Dateien, ihrem temporären oder permanenten Charakter verschiedene Umsetzungsvarianten für das verteilte Datenmanagement. Die notwendige Interoperabilität mit den im Middleware-Modul implementierten Grid-Diensten wie zum Beispiel dem Grid Workflow Execution Service (GWES) bzw. den D-Grid-Basisdiensten musste garantiert werden. Die wesentlichen Arbeiten zum Aufbau des „Data Grid“ im MediGRID wurden gemeinsam vom ZIB und ZIH bearbeitet.

Der Arbeitsschwerpunkt c) befasste sich mit der Analyse und Optimierung des Job-Scheduling im MediGRID. Um aufwendige Datentransferoperationen zu vermeiden oder Einschränkungen bei der Auswahl von Datenquellen und -senken in den Workflows der medizinischen Anwendungen zu berücksichtigen, wurden Kriterien für eine Optimierung durch ein datenabhängiges Job-Co-Scheduling erarbeitet. Hierbei flossen auch Erfahrungen und Lösungsansätze aus den anderen Community-Projekten „C3-Grid“ und „Astro-Grid“ ein. Die Arbeiten zum Co-Scheduling wurden von der GWDG bearbeitet.

Für den einheitlichen einfachen Zugang zu den MediGRID Anwendungen wird ein GridSphere-basiertes Portal betrieben. Der Software-Stack des Portals wurde weitestgehend durch das Modul Middleware definiert. Die betriebliche Umgebung wurde zunächst an der GWDG und später beim MediGRID-Projektpartner an der Universität Leipzig realisiert. Hier stehen eine Produktions- und eine Entwicklungsumgebung aus BMBF-Sondermitteln zur Verfügung, um neue Anwendungen in das MediGRID Portal integrieren und testen zu können.

2.5.2 Modulergebnisse

2.5.2.1 Koordination, Policies und Betrieb

Koordination und Policies

Zur Regelung des Betriebs der MediGRID-Ressourcen und deren Nutzung wurden folgende MediGRID-Policy-Dokumente erstellt:

- Mohammed Y, Viezens F, Sax U, Falkner J, Steinke Th (2006). Resource usage policy phase one – development. MediGRID-Deliverable, 27.10.2006
- Mohammed Y, Viezens F, Sax U, Falkner J, Steinke Th (2006). MediGRID-developer/user/guest agreement phase one – development phase. MediGRID-Deliverable, 27.10.2006

In diesen Dokumenten werden Nutzerrollen definiert und die Verantwortung zur Administration der Module festgelegt. Dieser Vereinbarung müssen Ressourcenprovider und Anwender zustimmen, bevor sie an MediGRID teilnehmen können.

Die Koordinierung und Absprache mit dem D-Grid Kerngrid erfolgte direkt mit den Ansprechpartnern der Ressourcenprovider oder über die Managementstrukturen im D-Grid und MediGRID.

Betrieb des MediGRIDs

Mit Beginn von MediGRID wurde ein erstes Testbed durch die Ressourcenprovider ZIB, ZIH und GWDG auf den aus Eigenmittel bereitgestellten Clustersystemen implementiert, damit den MediGRID-Nutzern ein frühzeitiger Zugang zu einer Grid-Umgebung ermöglicht werden konnte. In diesem Testbed wurden die für MediGRID nötigen Middleware-Komponenten wie Globus erstmalig installiert, eine lokale Grid-RA zur Ausstellung von X.509-Zertifikaten für Maschinen und Nutzer eingerichtet und die Datenmanagementsysteme SRB und OGSA-DAI betrieben. Mit dem Testbed wurden die notwendigen Erfahrungen gesammelt, um die Konfiguration der mit Sondermitteln des BMBF beschafften neuen Ressourcen zu vereinfachen. Das MediGRID-Testbed wurde fristgerecht an die Nutzer übergeben (AP1, Deliverable D2).

Auf der ZIB-Plattform wurde SRB für das Datenmanagement verwendet und mit OGSA-DAI der Zugriff auf Datenbanken im MediGRID eingerichtet. Hierfür wurden die Datenbanken „Geneontology“ (MySQL), „NCI Thesaurus“ und die Bilddatenbank der Charité (PostgreSQL) installiert.

Sonderinvestition 2006 - 2008

Das Modul Ressourcenfusion war federführend an der Beschaffung der Sonderinvestitionen 2006, 2007 und 2008 für die Standorte Berlin, Dresden und Göttingen beteiligt.

Am ZIB wird zum Zeitpunkt Februar 2009 ein D-Grid-Cluster mit 480 Cores (80 Knoten) betrieben. Auf diesem Cluster befinden sich neben dem durch die DGI-Referenzinstallation vorgegebenen Grid-Middleware-Stack auch die MediGRID spezifischen Dienste wie das Datenmanagementsystem SRB (Storage Resource Broker), GWES (Grid Workflow Execution Service) sowie Ganglia zum Ressourcenmonitoring. Alle registrierten D-Grid Nutzer haben Zugang zu den Ressourcen und können mit Hilfe der installierten Middleware-Komponenten Jobs starten und den Speicherplatz für Daten bzw. Datenbanken nutzen.

Bei der GWDG wurden während der Projektlaufzeit Ressourcen aus den Sonderinvestition 2006, 2007 und 2008 konfiguriert und in Betrieb genommen. Alle Ressourcen am Standort Göttingen sind heute im Grid-Ressourcen-Zentrum **GoeGrid** zusammenfasst. Der GoeGRID-Cluster mit mehr als 1600 Cores ist von mehreren D-Grid-Communities beschafft worden und wird gemeinsam von der GWDG und dem II. Physikalischen Institut der Universität Göttingen betrieben. Diese Ressourcen stehen über ein Fairshare-Scheduling dem MediGRID und seinen Nachfolgeprojekten wie Services@MediGRID zu ca. 25% zur Verfügung.

Am ZIH wurden 2006 die Sondermittel benutzt, um die Compute-Ressourcen substantiell zu erweitern, die vor allem den Communities von MediGRID und HEP-Grid zur Verfügung stehen.

Monitoring

Als Monitoringsystem wird an den Standorten Göttingen, Dresden und Berlin Nagios benutzt. Nagios ermöglicht die flexible Definition von zusätzlichen Prüfungen, wie sie im Grid-Kontext notwendig sind. Mithilfe von Nagios werden Grid-Services wie z. B. von Globus (globus jobsubmit) und die Gültigkeit von Zertifikaten (Maschinenzertifikaten) geprüft sowie der noch freie Speicherplatz überwacht. Mit Nagios kann die Funktionsfähigkeit der Ressourcen und die Verfügbarkeit von Grid-Diensten fortlaufend überwacht werden. Durch das Nagios-Monitoring

konnten Probleme zeitnah festgestellt und in vielen Fällen behoben werden, bevor Nutzer durch DGUS auf sie hinweisen mussten. Weiterhin konnte durch die bessere Überwachung auch die Zuverlässigkeit der Rechenressourcen insgesamt verbessert werden.

Von der GWDG werden die Ressourcen darüber hinaus mit dem Monitoring-Dienst Jawari überwacht. Dabei handelt es sich um eine Entwicklung von Fraunhofer ITWM innerhalb des DGI. Es wurde eine Jawari-Konfiguration vorgenommen, welche die Antwortzeiten der MediGRID-Rechenressourcen und die Datentransferraten zwischen den Standorten überwacht.

Mit Beginn des MediGRID-Projektes wurde an der GWDG das übergreifende Monitoring der MediGRID-Ressourcen realisiert. Die Aufgaben des Systems umfassen die zentrale Erfassung der Ganglia-Informationen und ihre Darstellung im Ganglia-Webinterface, die Aggregation der MDS-Informationen aller MediGRID-Ressourcen, die Zugriffsmöglichkeit über WebMDS sowie die Bereitstellung eines MyProxy-Servers, in dem die Credentials der Grid-Nutzer abgelegt werden. Mit der beträchtlichen Vergrößerung der MediGRID-Ressourcen durch Mittel der Sonderinvestition 2008 seit Mitte 2008 zeigte die alte Serverplattform Funktionseinschränkungen. Deshalb wurde mit der Planung der neuen Plattform gridmon begonnen. Die Konfigurationsarbeiten dafür sind mittlerweile abgeschlossen und in die Produktion überführt.

Accounting

Für das Accounting der benutzen Ressourcen (Rechenzeit) wurde das Tool "pacct" evaluiert und weiter verbessert. Neben Korrekturen bei der Erfassung paralleler Jobs wurden Erweiterungen hinzugefügt, die eine direkte Aufschlüsselung der Nutzung nach Communities erlauben. Die in MediGRID implementierte gemeinsame Accounting-Infrastruktur erlaubt es, basierend auf den Accounting-Rohdaten eine hierarchische, der MediGRID-Community adäquate Präsentation entsprechend typischer Nutzungsprofile wie z. B. Anwendungsklassen zu erstellen.

2.5.2.2 Anforderungsanalyse und Lösungsstrategien

Die Anforderungsanalyse und Spezifikation der Anforderungen an die Funktionalität der Grid-Middleware wurde fristgerecht abgeschlossen.

Für die Ableitung der Spezifikation der Anforderungen an die Funktionalität der Grid-Middleware im MediGRID wurden die beiden Dokumente

- AP2, Deliverable 1: „MediGRID - Analyse der MediGRID-Anwendungen“ (29.05.2006)
- AP2, Deliverable 2: „Analyse und Spezifikation der MediGRID-Anwendungen“ (31.07.2006)

gemeinsam mit dem Modul Middleware erarbeitet. Die Aufteilung in zwei Dokumente erfolgte mit dem Ziel einer klaren Trennung der Analyse von der nachfolgenden Ableitung der Spezifikation und der möglichst zeitnahen projektinternen Dokumentation.

Das Dokument „MediGRID - Analyse der MediGRID-Anwendungen“ ist das Ergebnis mehrerer fokussierter Treffen von Projektpartnern der Module Middleware und Ressourcenfusion mit den jeweiligen Anwendergruppen. Es beschreibt die Ergebnisse der informatischen Analyse der MediGRID-Anwendungen aus den Bereichen der Bioinformatik und der Bildverarbeitung. In ihm werden verbindlich für die entsprechenden Projektpartner die Anwendungen der beiden oben genannten Module knapp und systematisch aus informatischer Sicht dokumentiert und soweit möglich ein umfassender Anforderungskatalog für jedes Anwendungszenario zusammengestellt. Speziell wird für jedes der vier Anwendungsszenarien des Moduls „Bioinformatik“ und bzw. der drei Anwendungsszenarien des Moduls „Medizinische Bildverarbeitung“ die Art und Weise der Nutzung der Anwendung in einer Grid-Umgebung und der sich daraus erwartende Vorteil festgehalten. Für jede der Anwendungen wird ein Workflow unter Angabe spezifischer

Details (z. B. Ein- und Ausgabeobjekte) und Randbedingungen dokumentiert. Soweit möglich werden anhand des Skalierungsverhaltens der Algorithmen produktionstechnische Charakteristika wie Rechenzeitaufwand und Bedarf an permanenter und temporärer Plattenkapazität abgeleitet.

Basierend auf der Analyse der MediGRID-Anwendungen wurden im nachfolgenden Dokument „Analyse und Spezifikation der MediGRID-Anwendungen“ die abgeleiteten Anforderungen und die geplante technische Umsetzung festgehalten. Für jede der Anwendungen werden Anforderungen in den Kategorien

- Nutzerschnittstelle (Portal),
- Datenmanagement,
- Workflow-Management und
- Sicherheit

unterschieden. Für jede dieser Kategorien werden bestimmte Grid-Middleware-Komponenten ausgewählt und evaluiert. Aufbauend auf diesen Erkenntnissen wurde der Entwurf der MediGRID-Architektur vorgestellt.

Workflow- und DataGrid Coaching

Für die Ausführung der MediGRID-Anwendungen mit dem Grid Workflow Execution Service (GWES) müssen Workflows und Ressourcen in maschinenlesbaren Format (XML) spezifiziert werden. Zusammen mit den Entwicklern der Bioinformatik-Anwendung AUGUSTUS wurden Beispiel-Workflows für die Sequenzanalyse erstellt und passende Software-Ressourcenbeschreibungen formuliert.

Im Zuge der Arbeiten waren mehrfach Änderungen notwendig, die sich aufgrund der ebenfalls noch in Arbeit befindlichen Anpassung des GWES an die Globus Toolkit 4 Middleware ergaben. Weiterhin erfolgte eine Anpassung an die neue D-Grid Ressource Description Language (D-GRDL), die im gleichen Zeitraum vom DGI spezifiziert wurde.

Die Tätigkeiten erfolgten in enger Zusammenarbeit und Abstimmung mit den Anwendungs-Entwicklern, die Portlets zur automatischen Erzeugung angepasster Workflows erstellten. Die Erfahrungen bei der Erstellung von Workflows für GWES wurden in einem technischen Bericht festgehalten.

Die Partner im Modul Ressourcenfusion (GWDG, ZIH, ZIB) unterstützten während der gesamten Projektlaufzeit die Anwendungsentwickler und Nutzer bei Fragen der Datenspeicherung im Grid, speziell bei der Nutzung der Werkzeuge SRB und iRODS. Im ersten Quartal 2008 wurden die Coaching-Aktivitäten für die Entwickler der Anwendungen der Klinischen Forschung bei der Nutzung des Datenmanagementsystems SRB in Kooperation mit der Charite (MediGRID-Modul Bildverarbeitung) weiter fortgeführt.

2.5.2.3 DataGrid in MediGRID

Die Anforderungsanalyse aus AP 2 ergab, dass die meisten Anwendungen folgende Arten von Datentransfers bzw. langfristiger Datenspeicherung benötigen:

- Dauerhafte Speicherung von (größeren) Datenmengen der Anwendungs-Workflows. Eine gemeinsame Datennutzung (innerhalb einer Nutzergruppe) sollte möglich sein.
- Transfer von Eingabedaten zu den Ressourcen, auf denen der Job ausgeführt werden soll (Data Stage-In) .

- Transfer von temporären Zwischenergebnissen, wenn der nächste Workflow-Schritt nicht auf der gleichen Ressource ausgeführt werden kann.
- Transparente Speicherung der Ergebnisse mit der Möglichkeit zur künftigen Weiterverarbeitung.

In MediGRID werden deshalb zur grid-weiten Datenspeicherung Services betrieben, die auf dem SDSC SRB (Storage Resource Broker, San Diego Supercomputing Center - SDSC), Nirvana SRB und iRODS (integrated Rule-Oriented Data System, SDSC) basieren.

Der **SRB** (Storage Resource Broker)[1] als existierendes System (erstes Release 1997) wird bereits in verschiedenen internationalen Projekten verwendet und stellte sich für die MediGRID Anforderungen als am besten geeignet dar. Der SRB ist ein verteiltes Datenmanagementsystem und basiert auf einer Client-Server Architektur, welches zur Datenvirtualisierung eingesetzt wird. Als bereits existierendes Werkzeug konnte der SRB bereits kurz nach Projektstart den Entwicklern für die Gridifizierung ihrer Anwendungen zur Verfügung gestellt werden. Die aktuelle Installation umfasst Speicherressourcen in Berlin, Dresden und Göttingen.

In Abbildung 25 ist beispielhaft für die MediGRID SRB Zone die verteilte Client-Server Architektur des SRB mit seiner Metadatenbank MCAT dargestellt.

Der Administrator des SRB besitzt eine grafische Schnittstelle zur Konfiguration der Zone. In Abbildung 26 werden die in der MediGRID Zone integrierten Ressourcen gezeigt.

Die wichtigsten Funktionalitäten des SRB für MediGRID sind:

- Datenabstraktion: global eindeutige Identifier für Daten bieten Datenvirtualisierung mit Abstraktion vom physikalischen Speicherort. Das heißt der Benutzer kann Daten verteilt über mehrere Ressourcen speichern und hat auf alle Daten innerhalb seines SRB Accounts eine Dateisystem-ähnliche Sicht. Daten lassen sich transparent über die Ressourcen replizieren.
- Metadaten: Zu Datenobjekten können zusätzlich Metadaten in Form von Attribut-Value-Unit Tripeln abgespeichert werden. Diese können auch zur Suche nach Daten verwendet werden. Die Metadaten werden in der SRB Metadatenbank MCAT gespeichert.
- Vielfalt an Schnittstellen: APIs (Java, C) und Clients (Scommands, SRB Portlet)
- Erweiterbarkeit: Neue Ressourcen können in eine existierende Zone integriert werden. Zwischen Zonen ist eine Zonenföderation möglich, um Nutzern gegenseitigen Zugriff auf die jeweiligen Ressourcen der Zonen zu ermöglichen.

1 http://www.sdsc.edu/srb

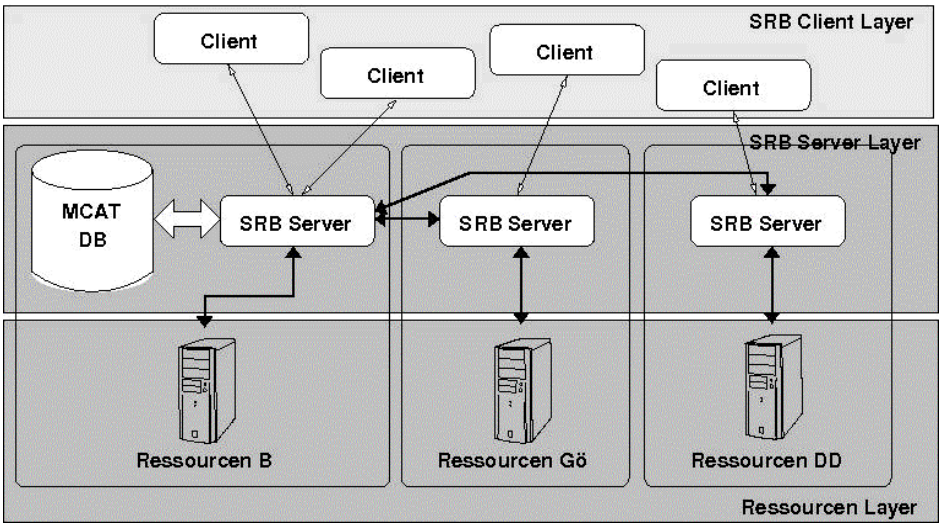


Abbildung 25. SRB-Installation in MediGRID

Integration des Datenmanagementsystems in die MediGRID Infrastruktur

Mit dem **Datenmanagement-Portlet** im MediGRID-Portal wird dem Anwender eine einfache Möglichkeit gegeben, um auf Daten im SRB zuzugreifen. Dazu wurde das GridSphere-Portlet aus dem BIRN Projekt[2] an die Software-Umgebung im MediGRID-Portal angepasst.

Eine weitere Zugriffsmöglichkeit lässt der **SRB Fuse-Client** zu. Für einen einzelnen Nutzer kann dessen SRB Verzeichnis in die Dateisystemstruktur auf dem (Linux-) Rechner des Nutzers eingebunden werden (Linux Fuse Mount). Somit sind die Daten im SRB für Anwendungen über die normale Dateisystemschnittstelle des lokalen Betriebssystems zugreifbar und benötigen keine speziellen Client-Befehle zum Datentransfer. In Abbildung X3 ist die Sicht auf eine SRB-Nutzerkennung aus Portalsicht und mit Fuse dargestellt. Allerdings steht mit der Fuse-Schnittstelle nicht der volle Funktionsumfang des SRB zur Verfügung.

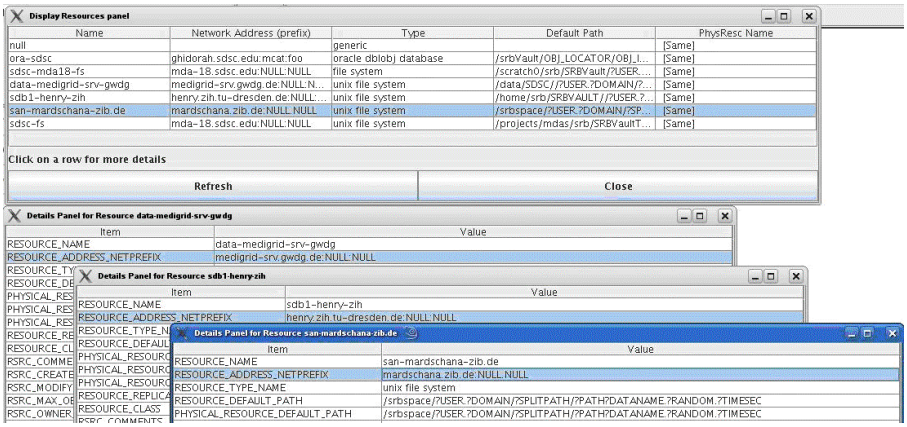


Abbildung 26. Ressourcen der SRB-Installation in MediGRID

Im Gegensatz dazu exponiert der **Scommand-Client** die volle SRB-Funktionalität dem Nutzer. Der Scommand-Client wird in MediGRID hauptsächlich zum Datenzugriff a) zwischen dem

2 http://www.nbirn.net/index.shtm

lokalen Rechner beim Anwender und SRB, b) auf den Rechenressourcen zum Zugriff auf Eingabedaten vor dem Start des Anwendungs-Workflows, und c) zur Speicherung von Ergebnisdaten verwendet. Alle Schritte a) bis c) können als Datentransferschritte Bestandteil des Anwender-Workflows zur aktuellen Laufzeit sein, d. h. die Datentransferschritte werden genau wie die eigentlichen Rechensschritte im Workflow modelliert. Die Anbindung von SRB an GWES ist allerdings noch verbesserungswürdig. Die Verwendung des Fuse-Clients zum Datenzugriff direkt aus der Anwendung konnte auf den Grid-Ressourcen wegen 32/64-Bit-Inkompatibilität nicht umgesetzt werden. Eine alternative Lösung wäre der Datenzugriff direkt über separate Grid-Services.

In Abbildung 27 sind die Datentransferzeiten beim Zugriff auf Daten im SRB mit dem Scommands-Client dargestellt. In Abbildung 27, links werden die Transferzeiten für verschiedene Datengrößen mit und ohne Optimierung durch die Bulk-Transfer Option gezeigt (kleine Datei: 500 KB, größere Datei: 100 MB, viele kleine Dateien: 100 x 500 KB und viele große Dateien: 100 x 100 MB). Man erkennt, dass der Gewinn durch die Optimierung mit wachsender Datengröße leicht abnimmt. Abbildung 27, rechts zeigt die Transferzeit für eine einzelne 500 MB große Datei mit einem Scommands-Client in Berlin und den Servern in Berlin, Göttingen und Dresden. Die Transferzeiten sind bis auf kleinere Schwankungen gleich und zeigen, dass bei ausreichender Netzwerkbandbreite kein signifikanter Unterschied bei den Zugriffszeiten besteht. Die durchschnittlich erreichte Bandbreite betrug ca. 10 MB/s.

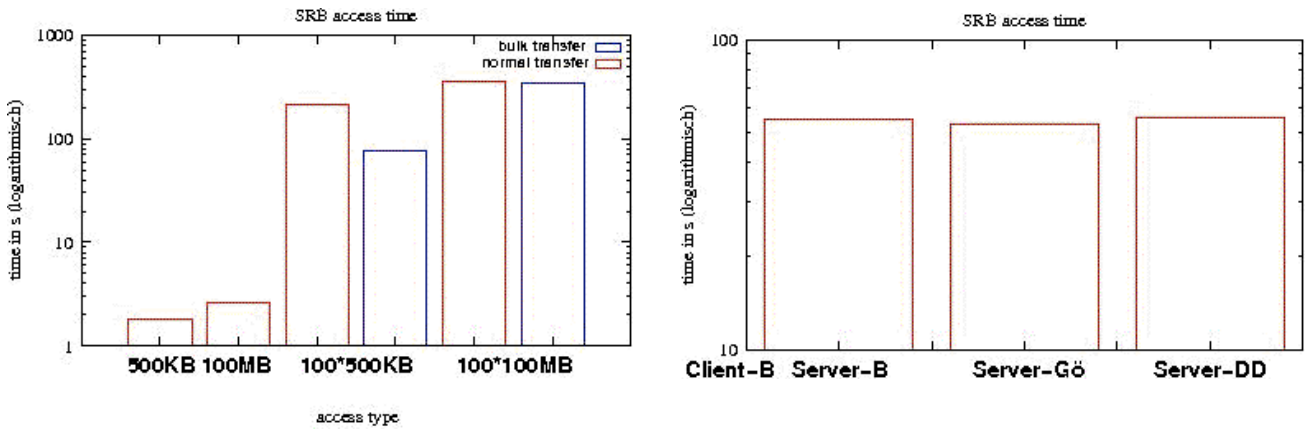


Abbildung 27. Datentransferzeiten (in Sekunden) zwischen SRB-Client und –Server für verschiedene Datengrössen (links) und beim Zugriff auf verschiedene Standorte (rechts)

Weitere Arbeiten zur Integration des Datenmanagements mit SRB in die MediGRID-Infrastruktur umfassten folgenden Aktivitäten:

- Das SRB-Nutzermanagement wurde in das MediGRID Benutzermanagement integriert. Nutzerkennungen im SRB werden automatisch angelegt, sobald neuer Benutzer im MediGRID VOMRS registriert wird.
- Am ZIB wird das SRB Nachfolgesystem iRODS[3] evaluiert und die Migration von SRB zu iRODS begleitet.
- Es wurde ein Konzept für eine fehlertolerante, weitverteilte Datenspeicherung entwickelt

3 https://www.irods.org/index.php

(s. Veröffentlichung).

- Die SRB-Server werden mit Nagios überwacht; automatische Neustart werden im Fehlerfall initiiert und im Problemfall der Administrator benachrichtigt.
- Die kommerzielle Version Nirvana SRB wurde mit einer kostenlosen Testinstallation evaluiert. Nirvana SRB ist nicht kompatibel zum SDSC SRB. Da die Nutzer sich auf die Nutzung vom SDSC beschränkten, wird der Nirvana SRB nicht weiterbetrieben. Die für den SRB einschränkende Nutzumgslizenz wird im Nachfolgewerkzeug iRODS aufgelöst (iRODS ist auch für kommerzielle Nutzer frei)
- Die MediGRID-Entwickler wurden bei der Umstellung ihrer Anwendungen auf die Datenzugriffe im Grid intensiv unterstützt. Dazu dienten die Durchführung von Nutzer-Workshops oder individuelle Unterstützungsleistungen.

Für die Bereitstellung der DataGrid-Dienste fielen periodisch Arbeiten zur Wartung und Aktualisierung der installierten Dienste, d.h. der Instanzen von OGSA-DAI, des SRB und der iRODS-Testinstallation, an.

Am ZIH (Dresden) wurden die Arbeiten zum Aufsetzen der Data-Grid-Infrastruktur mit der Installation und Bereitstellung der Biodatenbanken sowie der Durchführung von Leistungstests im Berichtszeitraum den Planungen gemäß abgeschlossen (AP3, Deliverable D3). Die Leistungsmessungen wurden bereits vorfristig im Jahr 2007 beendet und deren Ergebnisse veröffentlicht (AP3, Deliverable D4 und D5).

Transparenter und performanter Zugriff auf Datenbanksysteme im Grid

Das Hauptziel war die Bereitstellung eines grid-weiten, transparenten und performanten Zugriffs auf relationale Datenbanken aus den MediGRID-Anwendungen. Die dafür erforderlichen Funktionalitäten bieten sowohl die Grid-Middleware-Komponenten **OGSA-DAI** als auch der Storage Resource Broker (SRB). Dazu wurden zunächst Installationen am ZIH und ZIB aufgesetzt und biomedizinische und bioinformatische Datenbanken wie z.B. MSD, CATH, PIBASE, PFAM, GOA oder SCOP installiert.

Schwerpunkt der Arbeiten am ZIH war eine tief greifende Analyse der Leistungscharakteristik von OGSA-DAI und SRB (Performanceanalyse) und die Ableitung von Empfehlungen für den praktischen Betrieb. Dafür wurden Leistungsdaten für ein breites Spektrum von Zugriffsmethoden und -charakteristika wie

- lokaler oder. entfernter Zugriff,
- Umfang der Datenmengen (klein/groß),
- Speichern von Bilddaten,
- Kopieren von Daten,
- Anwendung verschiedener OGSA-DAI Transfer-Methoden,
- Verschlüsselung

ermittelt. Weiterhin wurde die Option des Distributed Query Processing (DQP) evaluiert. Die Ergebnisse sind in der Veröffentlichung „Accessing Bio-Databases with OGSA-DAI - A Performance Analysis“ zusammengestellt.

Abbildung 28 zeigt beispielhaft das Ergebnis der Messung beim Lesen aus einer Datenbank in OGSA-DAI unter Nutzung der verschiedenen Transfer-Mechanismen, die OGSA-DAI bietet.

Deutlich ist der Leistungseinbruch in OGSA-DAI gegenüber dem nativen Zugriff (JDBC) zu sehen, wobei sich die unterschiedlichen OGSA-DAI-Transfervarianten sich nicht qualitativ untereinander unterscheiden. Der synchrone OGSA-DAI-Transfermode liefert dabei die beste Datenzugriffsraten, der asynchrone Transfermode die schlechteste.

Das anfängliche MediGRID-DataGrid-Testbed wurde erfolgreich in den Produktionsbetrieb überführt und in die MediGRID-Anwender-Workflows integriert.

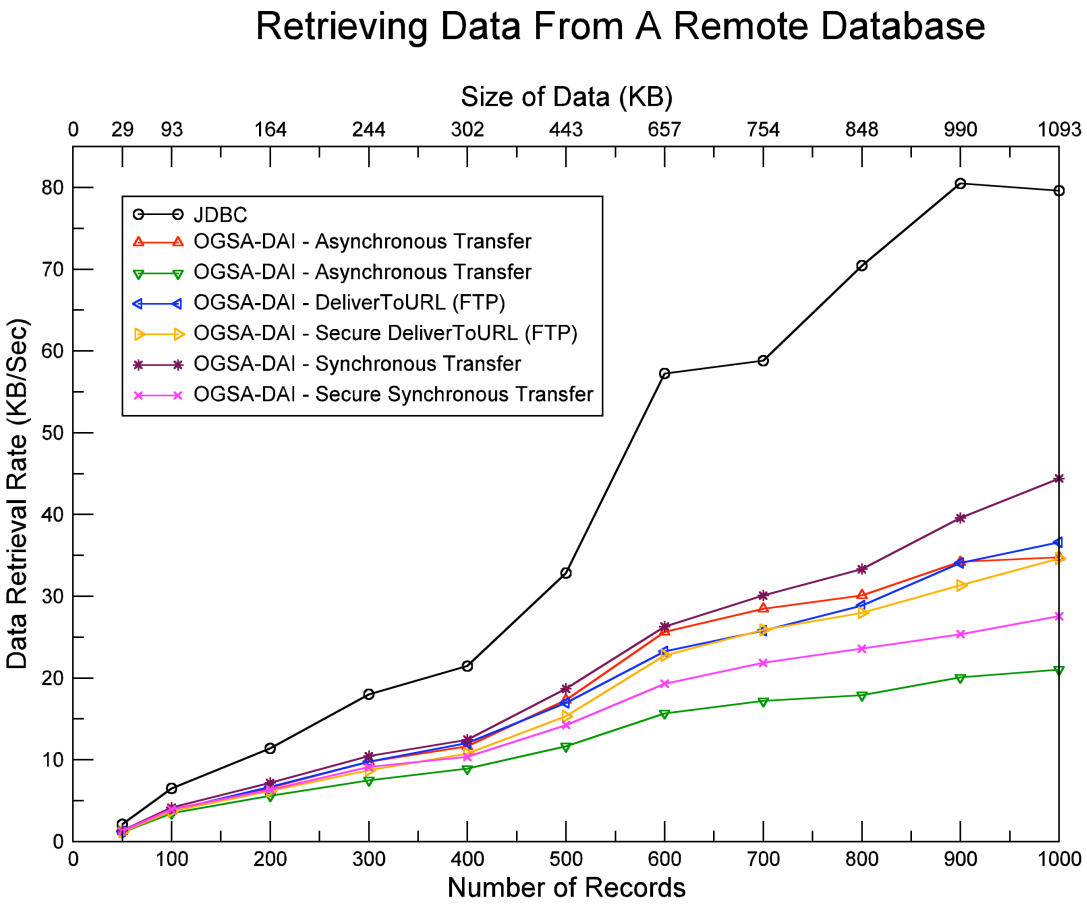


Abbildung 28. Leserate (in kB/s) aus einer Datenbank mit OGSA-DAI mit verschiedenen Zugriffsmechanismen

2.5.2.4 Ressourcenverteilung für datenintensive Anwendungen im MediGRID

Die Umsetzung dieses Arbeitspakets erfolgte in Kooperation mit dem Modul Middleware (A. Hoheisel, FIRST), welches das GWES Workflow-System weiterentwickelte, mit dem das Scheduling der MediGRID-Anwendungen realisiert wird.

Das AP4-Deliverable D7 "Definition der Leistungsparameter" wurde zum 01.05.2007 in der Version 1.0 abgeschlossen. Dazu wurde eine Abstimmung mit den D-Grid-Communities von

AstroGrid-D und C3-Grid vorgenommen, indem eine Untersuchung der Informationsdienste in beiden Communities durchgeführt wurde, um potentielle Synergien mit MediGRID zu nutzen. Neben dem GWES Workflow-System zur Verteilung der Anwendungen im Grid verwendet MediGRID die D-GRDL Ressourcenbeschreibungssprache. Auf dieser Ebene bestand keine Übereinstimmung mit beiden Projekten, die eigene Komponenten entwickelten.

Ein gemeinsames Community-übergreifendes Interesse gab es bei den Monitoring-Systemen. In der Folge wurden mehrere Monitoring-Workshops veranstaltet, bei denen die Monitoring-Entwicklungen präsentiert und Anforderungen diskutiert wurden. Ansätze für das Scheduling und Datenmanagement in datenintensiven Anwendungen in MediGRID wurden beim Scheduling-Workshop während des D-Grid All-Hands-Meetings 2007 vorgestellt.

Mit der Bereitstellung der Version 1.1.0 des Grid Workflow Execution Service (GWES) konnten die Anforderungen für ein Daten- und Job-Co-Scheduling für MediGRID fristgerecht erfüllt werden (AP4, Deliverable D9). Diese Version des GWES bietet Möglichkeiten zum datenabhängigen Scheduling von Aktivitäten. Über eine räumliche Co-Allokation kann erreicht werden, dass Aktivitäten mit mittel oder unmittelbaren Datenabhängigkeiten auf derselben Ressource ausgeführt werden, um unnötige Datentransfers über das WAN zu vermeiden. Darüberhinaus werden von dieser GWES-Version auch erweiterte Ablaufstatistiken von Aktivitäten und Workflows zur Verfügung gestellt. Mit Hilfe einer angepassten Ressourcenbewertung für große Cluster wurde die Verteilung weiter verbessert.

Zuletzt erfolgte eine Analyse der Job-Wartezeiten auf den MediGRID-Ressourcen. Die gewonnenen Daten sollen der Entwicklung von Modellen zur stochastischen und statistischen Vorhersage von Queue-Wartezeiten mit dem Ziel der Laufzeitreduzierung von Workflows dienen. Diese Modelle können in späteren GWES-Versionen berücksichtigt werden.

Gemeinsam mit FIRST wurden von der GWDG Ideenskizzen für eine Interoperabilitätsschicht für Community-Scheduler entwickelt, die im Projekt "D-Grid Scheduler Interoperability (DGSI)" im 3. D-Grid-Call realisiert werden sollen.

2.5.2.5 MediGRID-Portal

Von 2006 bis Mitte 2007 wurde bei der GWDG die erste Version des MediGRID-Portals betrieben. Das Portal benutzt GridSphere als Basisplattform.

Zur Aufnahme des Portalbetriebes waren eine Reihe von technischen Maßnahmen an den beteiligten Grid-Standorten umzusetzen:

- Konfiguration der Firewalls aller beteiligten Hardware-Ressourcen, um den Zugriff auf Globus-Dienste zu ermöglichen.
- Test der Basisdienste von Globus wie GridFTP, RFT, GSI mit X.509- Zertifikaten und Job-Submission mit GRAM und WS-GRAM.
- Implementierung einer initiale Prozedur zum Mapping von Nutzerzertifikaten auf lokalen Systemkennungen.
- Verwaltung der X.509 Grid-Nutzerzertifikate der Portalbenutzer über einen auf dem Portalserver installierten MyProxy-Server

Das Deliverable D11 „Demo – MediGRID-Anwendung im Portal” wurde durch die Demonstrationen von MediGRID-Anwendungen auf Konferenzen AMIA-Konferenz (November 2006), HealthCare (Januar 2007, Genf) und nicht zuletzt durch die Anwendungsdemonstration für die Zwischenbegutachtung am 2. März 2007 erfüllt.

Mit der Inbetriebnahme der Ressourcen aus Sondermitteln 2006 am MediGRID-Standort Leipzig wurde wie geplant das MediGRID-Portal auf die dortigen redundanten Portalserver migriert. In Leipzig läuft je eine Instanz im Produktionsbetrieb und eine für Test- und Entwicklungsarbeiten auf den redundanten Portalservern.

Die Arbeiten zum AP5 sind fristgerecht abgeschlossen worden.

2.5.2.6 Abstimmung mit anderen Grid-Communities

Bezüglich des Scheduling in MediGRID erfolgte die vorgesehene Abstimmung mit den D-Grid-Communities von AstroGrid-D und C3-Grid. Es zeigte sich, dass beide Projekte aufgrund der spezifischen Anforderungen eigene Scheduling-Komponenten entwickelten, die nicht mit dem in MediGRID verwendeten GWES Workflow System und der D-GRDL Ressourcenbeschreibungssprache interoperabel sind.

Partner des Moduls Ressourcenfusion haben an mehreren nationalen und internationalen Grid-Konferenzen und –Workshops aktiv teilgenommen, um über den Stand der Entwicklung im MediGRID zu berichten und weitere Entwicklungen bzw. Kooperationen abzustimmen.

2.5.3 Nutzen und Verwertbarkeit der Modulergebnisse

Mit dem im Modul Ressourcenfusion entwickelten Migrationsweg vom SRB zum iRODS kann die Nachhaltigkeit der MediGRID-Infrastruktur für das verteilte Datenmanagement ermöglicht werden.

Eine Integration des im XtreemOS-Projektes entwickelten verteilten Dateisystems XtreemFS als Alternative zu iRODS hängt von verfügbaren Personalressourcen ab. Dieser Weg wird im Projektantrag **MoSGrid** des 3- D-Grid-Calls verfolgt.

Im Modul Ressourcenfusion wurden keine Erfindungen oder Schutzrechte angemeldet.

Im MediGRID konnte gezeigt werden, dass grid-verteilte Rechen- und Speicherressourcen für die Lösungen von Fragestellungen der Medizin und Bioinformatik nutzbringend erschlossen werden können. Vorrangiges geeignetes Rechenmodell ist das High-Throughput-Computing (HTC) im Grid, mit dem temporärer Spitzenbedarf ökonomisch abgedeckt werden kann.

Die sich im MediGRID erstmalig zusammengeschlossenen Partner haben aufgrund ihrer positiven Erfahrungen im MediGRID erfolgreich weitere D-Grid-Projekte einwerben können. Im Bereich Biomedizin/Lebenswissenschaften zählen dazu **Services@MediGRID** und **MedInfoGrid** im 2. D-Grid-Call und **PneumoGrid** im 3. D-Grid-Call.

2.5.4 Fortschritte auf dem Markt für Grid Computing in der Medizin und der Bioinformatik aus Modulsicht - Darstellung Ergebnisse Dritter

Im Bereich „Verteiltes Datenmanagement“ haben sich in den vergangenen Jahren Entwicklungen ergeben, die neben einer prinzipiellen Leistungsverbesserung insbesondere auf die Erhöhung der Transparenz für den Endnutzer und die Berücksichtung von etablierten Standardschnittstellen abzielen.

Im Rahmen des EU-Projektes „XtreemOS“ wird **XtreemFS** entwickelt, das ein geographisch verteiltes Dateisystem realisiert. Die POSIX-konforme Schnittstelle für I/O-Operationen, wie sie auch in Betriebssystemen auf Serverplattformen zu finden ist, wird beibehalten, so dass eine Transparenz beim Zugriff auf verteilte Daten für den Endnutzer garantiert werden kann. Für einen effizienten Zugriff auf die verteilten Daten werden Replikamechanismen in XtreemFS implementiert.

Das SRB Nachfolgesystem **iRODS** bietet neben seinem offenen Lizenzmodell auch eine prinzipielle höhere Flexibilität bei der Integration feingranularer Dienste und ermöglicht somit eine verbesserte Unterstützung datenintensiver Anwendungsszenarien. iRODS ist auch im Tahmen von MediGRID evaluiert worden und mögliche Migrationsstrategien vom SRB zum iRODs in Zusammenarbeit mit den Anwendungsentwicklern erarbeitet worden

Eines der weltweiten Leuchtturmprojekte im Bereich Grid-Computing für die Lebenswissenschaften stellt das vom US National Institute of Health geförderte **caBIG**-Projekt dar. Durch die Verfügbarkeit des TeraGrids musste man sich hier weniger auf die Entwicklung und den Betrieb einer Infrastruktur für das Grid-Computing orientieren. Vielmehr wurde der Schwerpunkt auf die Definition und Implementierung von standardisierten Datenmodellen (Ontologien), Datenstrukturen und Zugriffsmechanismen gelegt, so dass caBIG-zertifizierte Werkzeuge konsistent auf die Vielzahl von verteilten Datenbanken der caBIG-Partner zugreifen können. Die Herausforderung lag hierbei nicht nur in der Entwicklung von Datenstandards als vielmehr auch deren konsequente Durchsetzung bei den Datenprovidern speziell im onkologischen Bereich. Eine praktikable Migration wird über ein mehrschichtiges Zertifizierungssystem erreicht.

2.5.5 Ausblick und Perspektiven im Bereich Ressourcenfusion

Die vergangenen Jahre haben gezeigt, dass die Erwartungen an das Grid nur teilweise erfüllt werden konnten. Hauptsächliche Schwächen liegen in a) dem hohen administrativen Aufwand bei der Implementierung der Grid-Infrastruktur, b) der unzureichenden Stabilität der Grid-Middleware, c) der fehlenden Transparenz für den Endnutzer, d) der unzureichend umgesetzten Standardisierung bei den Nutzerschnittstellen oder e) der inoperablen Weiterentwicklung strategischer Software-Komponenten („Dilemma inkompatibler Globus Varianten“).

Cloud-Computing restrukturiert die Ressourcenverteilung auf wenige aber mächtige Daten- und Rechenzentren, um teilweise Nachteile des Grids aufzuheben. Die Virtualisierung der Anwendungen stellt ein Fortschreiben der mit dem Grid-Computing begonnenen Bestrebung nach Virtualisierung der Ressourcen dar. Problematisch verbleibt der Sachverhalt einer unzureichenden Standardisierung für verschiedenen Rollen (Provider, Entwickler, Nutzer).

Die in MediGRID benutzen Methoden und Werkzeuge für den Zugriff und die Verwaltung von verteilten Daten sind bzw. werden durch externe Gruppen (SDSC, UK eScience) entwickelt. Im Modul Ressourcenfusion wurden die Werkzeuge SRB, Nirvana, OGSA-DAI sowie jüngst auch iRODS evaluiert und im Zusammenarbeit mit dem DGI/DGI-II für die D-Grid-Community Erfahrungen gesammelt. Die genannten Werkzeuge haben sich überwiegend als brauchbar in der angestrebten MediGRID-Produktionsumgebung gezeigt. Teilweise muss der vereinfachte Zugang mit substantiellen Leistungseinbußen erkauft werden (OGSA-DAI).

Die in MediGRID gesammelten Erfahrungen beim Datenmanagement in Grid-Umgebungen und beim Betrieb von Grid-Ressourcen flossen in die D-Grid-Projekte **Services@MediGRID, MedInfoGrid** und **PneumoGrid** ein.

Die im AP4 „Job-Co-Scheduling“ gesammelten Erfahrung sind von der GWDG gemeinsam mit dem FhG-Institut FIRST zu Ideenskizzen für eine Interoperabilitätsschicht für Community-Scheduler weiterentwickelt und im Projekt **D-Grid Scheduler Interoperability** (DGSI) im Rahmen des 3. D-Grid-Call vorgeschlagen worden.

Im D-Grid-Projektantrag **MoSGrid** werden Technologien, die in anderen EU-Projekten entwickelt wurden (ChemMomentum, XtreemOS), auf ihre Möglichkeiten und Grenzen in der D-Grid-Umgebung für den Anwendungsbereich molekularer Simulationen evaluiert.

2.6 Ontologiewerkzeuge

2.6.1 Moduleinordnung und Modulziele

2.6.1.1 Modulaufbau und Zusammensetzung

Am Modul Ontologiewerkzeuge sind drei Institute der Universität Leipzig beteiligt: Institut für Statistik, Epidemiologie und Medizinische Informatik (IMISE), Institut für Informatik (IfI) und das Interdisziplinäre Zentrum für Bioinformatik (IZBI). Die Leitung des Moduls erfolgt über die Modulleiter Prof. Markus Löffler (IMISE) und Prof. Erhard Rahm (IfI / Abteilung Datenbanken). Die durch das Projekt eingeworbene Personalstelle wurde im November 2005 mit Herrn Michael Hartung (IZBI) besetzt. Weitere beteiligte Personen am Modul sind Herr Dr. Toralf Kirsten (IZBI), Frank Loebe (IfI), Prof. Heinrich Herre (IMISE) und Ronald Speer (IMISE). Interne Abstimmungen in Form regelmäßiger Treffen während der Projektlaufzeit garantierten einen kontrollierten Fortgang der Arbeiten innerhalb des Moduls. MediGRID-weit wurden über regelmäßige Besuche von Vorstandssitzungen, speziellen Workshops / Treffen oder durch Teilnahme an Telefonkonferenzen Informationen ausgetauscht und der Fortgang der Arbeiten diskutiert.

2.6.1.2 Modulinhalte

Das Modul fokussiert vorrangig auf die Nutzbarmachung vorhandener Ontologien (z.B. Gene Ontology, NCI Thesaurus) im Bereich der Lebenswissenschaften für die am Projekt beteiligten Anwendungsmodule Bioinformatik, Bildverarbeitung und Klinische Forschung. Hierzu sind entsprechende Werkzeuge, im Speziellen Grid-Services, welche einen Zugang und die Integration von Ontologien in Grids ermöglichen, zu konzipieren und umzusetzen. Des weiteren werden semantische Korrespondenzen zwischen vers. Ontologien in Form von Ontologiemappings aufgrund der wachsenden Anzahl von Ontologien immer wichtiger. Auf diesem Gebiet gilt es semi-automatische Match-Verfahren zur Erzeugung solcher Ontologiemappings zu untersuchen und zu entwickeln, um Ergebnisse Anwendern des Projekts über Grid-Schnittstellen zur Verfügung zu stellen. Neben der Integration existierender Ontologien und zugehöriger Mappings wird gleichermaßen die Erstellung eigener, projekt-spezifischer Ontologien angestrebt. Hierzu stellen kollaborative Werkzeuge in Form von Web-Applikationen ein mögliches Mittel zum Aufbau ontologisch fundierter Begriffssammlungen bzw. Standardisierungen innerhalb einer Domäne dar. Alle im Modul entwickelten Werkzeuge / Services sollen Anwendern des Projekts über einfache, intuitive Web-Schnittstellen zugänglich sein, was eine enge Zusammenarbeit mit dem Modul Middleware voraussetzt.

2.6.1.3 Ziele des Moduls

Das Modul setzte sich laut Antrag zum Ziel die bisher entwickelten Werkzeuge gridfähig zu machen, um eine Basisunterstützung zur Erstellung und Verwendung vorhandener biomedizinischer Ontologien in MediGRID-Applikationen zu garantieren. Aufgrund der Vorgaben des D-Grid möglichst auf Standard-Software aufzubauen (z.B. Globus Toolkit, OGSA-DAI, ...) wurde der angestrebte Fusionsdienst zur Integration von Ontologien auf Basis der OGSA-DAI Framework umgesetzt (siehe AP 1). Dies erleichterte die Integration in das MediGRID, da Routinen/Aufgaben wie Zertifikatsbehandlung oder Ressourcenbeschreibungen bereits umgesetzt waren und direkt verwendet werden konnten. Zusätzlich zum ursprünglichen Antrag wurde die Erstellung projekt-eigener Ontologien mittels eines kollaborativen Werkzeugs auf Basis von Wiki-Technologie betrachtet. Im Speziellen wurde der Aufbau einer Projekt-Ontologie für MediGRID und anschließend für das gesamte D-Grid (D-Grid Ontology DGO) durchgeführt. Aufgrund der Anforderungen aus MediGRID wurde in AP 2 auf die Verfügbarmachung berechneter Ontologiemappings aus Match-Verfahren über Grid-Schnittstellen fokussiert. Auf Seiten der Hardware konnten durch die Sonderinvestitionsmaßnahmen 2006 drei Server beschafft werden,

was laut Projektplan nicht einplant war. Die Administration der Server erfolgt im Zusammenarbeit mit dem Modul Middleware.

2.6.1.4 Umgesetzte Aufgaben des Moduls

Die aus den Zielen des Moduls abgeleiteten Aufgaben gliedern sich in zwei Arbeitspakete. Während Arbeitspaket 1 (AP 1) die Integration existierender sowie dem Aufbau eigener projektspezifischer Ontologien umfasst, beschäftigt sich Arbeitspakt 2 (AP 2) mit der Erstellung und Integration von Ontologiemappings in das MediGRID.

- *Arbeitspaket 1 (AP 1) – Integration vorhandener Ontologien, Aufbau projekt-spezifischer Ontologien*
    - AP 1.1: Anforderungsanalyse
    - AP 1.2: Konzeption und Realisierung einer Middleware für die Integration von Ontologien in Grids
    - AP 1.3: Integration neuer Ontologien
    - AP 1.4: Integration mit Grid-Middleware
- *Arbeitspaket 2 (AP 2) – Ontologiematching*
    - AP 2.1: Anforderungsanalyse
    - AP 2.2: Integration von Ontologiemappings

Ergebnisse eines jeden Arbeitspakets sind in entsprechenden Deliverables (D1 – D6) dokumentiert. Realisierte Softwarebausteine, Services, Schnittstellen und Web-Oberflächen wurden in Kooperation mit den Modulen Middleware und Ressourcenfusion in das MediGRID und dessen Infrastruktur integriert.

2.6.2 Modulergebnisse

2.6.2.1 Wissenschaftlich-technische Ergebnisse

Nachfolgend werden die wissenschaftlich-technischen Ergebnisse des Moduls für jedes der Arbeitspakete dargestellt.

Für die Anforderungsanalyse in AP 1 wurde ein Fragebogen erstellt, um den gegenwärtigen Stand und geplante Entwicklungen zu Ontologien und relevanten Datenquellen zu ermitteln. Dieser Fragebogen wurde zusammen mit Fragen der Module Middleware und Ressourcenfusion an die Anwendungsmodule versandt. Die Befragung zeigte, dass die Projektpartner bisher kaum Erfahrungen im Bereich Ontologien haben und existierende Ontologien nur gering nutzen. Die Notwendigkeit der Vereinheitlichung von Begriffen wird jedoch prinzipiell erkannt, z.B. für Begriffe wie „Genotyp“, „Chromosom“, etc. Aufgrund der Ergebnisse wurde ein Ontologie-Workshop organisiert (siehe eigene Veranstaltungen), um den Projektpartnern die Grundlagen und Möglichkeiten von Ontologien sowie Integration von Datenquellen aufzuzeigen. Hierzu wurde eine Kooperation mit dem ebenfalls vom BMBF geförderten Projekt „OntoVerse“ (Prof. Arndt v. Haeseler) ins Leben gerufen. Während des Workshops wurden Arbeiten des eigenen Moduls wie auch von OntoVerse vorgestellt und diskutiert. Im Anschluss an den Workshop wurden zwei konkrete Anwendungen des Moduls Bioinformatik (AUGUSTUS, SNPSelection) als Anwender von Ontologieinformationen identifiziert. Während AUGUSTUS auf Genvorhersagen fokussiert und für die Ergebnisrepräsentation Terme / Konzepte der SequenceOntology benötigt, werden bei der Markerselektion in SNPSelection zusätzliche Informationen aus der GeneOntology als hilfreich erachtet. Da neben den beiden aufgeführten Anwendungen künftig weitere MediGRID-Anwendungen Ontologieressourcen nutzen, erschien die Entwicklung einer generischen und transparenten Middleware für den Zugang zu Ontologien im Grid sinnvoll.

Aufbauend auf dem bereits existierenden und von D-Grid für Datenbankzugriffe favorisierten Framework OGSA-DAI wurde mit der Konzeption und Entwicklung einer Middleware für die Integration von Ontologien in Grids begonnen (AP 1.2). Ergebnisse der Anforderungsanalyse bspw. detaillierte Angaben zu den betrachteten Anwendungen, Inhalte der Fragebögen bzw. Informationen zum durchgeführten Ontologie-Workshop sind in Deliverable „D1 – Anforderungsanalyse“ dokumentiert.

Das Problem beim Zugriff auf mehrere verteilte und bereits existierende Ontologien im Grid ist u.a. deren Heterogenität. So existieren im biomedizinischen Bereich zahlreiche Formate und Quelltypen für Ontologien. Hierzu zählen relationale Datenbanken, XML-Daten, standardisierte Dateiformate wie OBO oder Standardisierungen aus dem Web-Bereich, z.B. OWL und RDF. Die Überbrückung der Heterogenität mit dem Ziel eines einfachen, transparenten und flexiblen Zugriffs auf Ontologien in Grids floss in die Realisierung der Middleware ein. Die Erweiterbarkeit von OGSA-DAI bildete die Basis für die Entwicklung der Architektur. Da bei der Interaktion mit Ontologien grundlegende Funktionen/Anfrageoperatoren verwendet werden, z.B. Suche von Begriffen/Termen, Ermittlung der Definition eines Konzepts oder Abfrage von Beziehungen eines Konzepts, müssen diese durch spezielle Zugriffsroutinen (Activities) für bestimmte Ontologieressourcen abgebildet werden (Wrapperprinzip).

Die in Abbildung 29 dargestellte Architektur der Ontologie-Middleware besteht aus *Ontologieservices* (OGSA-DAI-Services mit Ontologiezugriffsfunktionalität), einem zentralen *Infoservice*, *Ontologiesourcen* (Ontologiequellen) und *Clients*. Zentrale Komponenten bilden die Ontologieservices mit ihren speziellen Activities für den transparenten Zugriff auf Ontologien. Jeder Ontologieressource ist eine spezielle Activity für die Integration von Ontologieinformationen mittels vers. Funktionen zugeordnet. Ein Ontologieservice kann Zugriff auf mehrere Ontologiequellen haben. Andererseits kann eine Ontologiequelle mehreren Ontologieservices zugeordnet werden. Durch diese Art der Verteilung und Bereitstellung der Ontologiequellen werden eine erhöhte Ausfallsicherheit sowie Lastbalancierung ermöglicht. Der zentrale Infoservice dient als Registry für die verteilten Ontologien und deren Ressourcen. Über ihm sind Informationen zu Ontologien und den zugehörigen Ontologieservices abfragbar. Es wird registriert, welcher Ontologieservice welche Ontologien anbietet bzw. umgekehrt welche Ontologie ist über welchen Ontologieservice erreichbar. Diese Informationen sind über Serviceschnittstellen abrufbar und dienen als Ausgangspunkt für die Interaktion mit Ontologieservices und somit den Ontologien im Grid. Der Service wurde nach dem WSRF-Standard implementiert und in die Middleware integriert. Möchte ein Client auf eine bestimmte Ontologie zugreifen, so wendet er sich mit Hilfe eines generischen Clients an den Infoservice und erhält eine Liste möglicher Services. Durch die Wahl eines Services in Verbindung mit seiner Ontologieressource ist ein Zugang zur gewünschten Ontologie möglich. Clients spezifizieren Ontologiefunktionen, z.B. Suche nach Konzepten bzw. Abfrage von Informationen zu Konzepten (Name, ID, Definition, Beziehungen, Synonyme, Referenzen), welche anschließend an den gewählten Service gesendet und durch die speziellen OntologyActivities verarbeitet werden. Ergebnisse werden in den generischen Client transferiert und der aufrufenden Anwendung zur Verfügung gestellt. Die Nutzung des Clients kann sowohl in Portalapplikationen anderen Grid-Services oder eigenständigen Applikationen erfolgen.

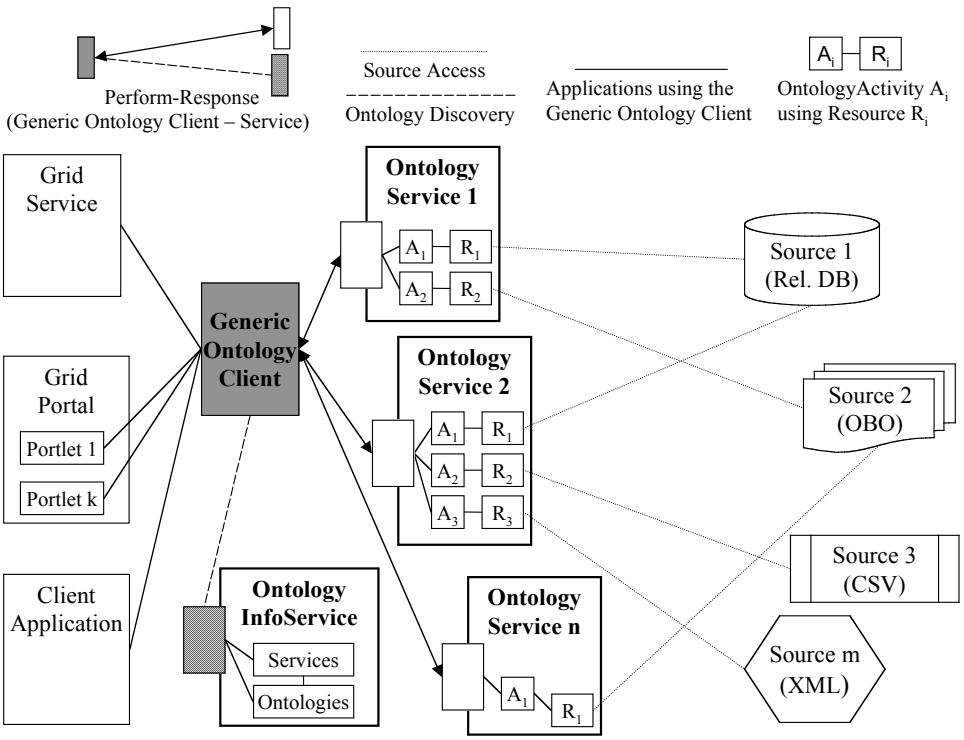


Abbildung 29. Architektur der realisierten Ontologie-Middleware

In einer ersten Testumgebung am IZBI in Leipzig wurden neben GeneOntology und dem NCIThesaurus zahlreiche biomedizinische Ontologien des OBO-Projektes, z.B. SequenceOntology, HumanDiseaseOntology oder MESH über die Architektur integriert. Im weiteren Verlauf des Projekts wurde das lokale Testbed am IZBI nach und nach in die MediGRID Infrastruktur integriert und mit Anwendungen des MediGRID Portals verknüpft (siehe AP 1.4). Die Middleware selbst wird im Deliverable „D2 - Konzeption und Realisierung einer Middleware für die Integration von Ontologien in Grids“ ausführlicher dargestellt und diskutiert. Veröffentlichungen erfolgten auf der German e-Science Conference (GES) 2007 sowie dem Workshop über die Grundlagen von Datenbanken (GvDB) 2008. Gleichermaßen wurde die Middleware auf dem vom DGI durchgeführten Metadaten-Workshop vorgestellt (2007).

Die zuvor beschriebene Middleware für die Integration von Ontologien in Grids wurde während der Projektlaufzeit verwendet, um erforderliche biomedizinische Ontologien in das MediGRID einzubinden. Neben Ontologien der Bioinformatik (z.B. GeneOntology, SequenceOntology) wurden gleichermaßen Ontologien auf dem Gebiet der Bildverarbeitung (z.B. RadLex für den Bereich radiologischer Systeme) oder Ontologien kooperierender Projekte (z.B. NCI-Thesaurs aus dem caBIG Projekt) nutzbar gemacht. Bis zum aktuellen Zeitpunkt wurden 16 Ontologien über die Middleware in das MediGRID integriert und sind somit über die Grid-Schnittstellen Anwendern und Applikationen zugänglich. Die integrierten Ontologien werden aufgrund ihrer ständigen Weiterentwicklung (Evolution) stets angepasst und aktualisiert. Automatisierte Routinen aktualisieren die Ontologien in periodischen Abständen, so dass ein hoher Grad an Aktualität für Nutzer gewährleistet werden kann. Ergebnisse des Arbeitspakets sind im zugehörigen Deliverable „D3 - Integration neuer Ontologien“ vermerkt.

Neben der Integration existierender Ontologien fokussierte das Modul andererseits auf das kollaborative Entwickeln/Editieren von domänenspezifischen Ontologien. Für die Erstellung solcher Ontologien boten sich zum Zeitpunkt der Bearbeitung insbesondere web-basierte Applikationen

aus dem Bereich der semantischen Wikis (Wikis mit einem semantischen Metamodell) an. Auf Basis der hohen Akzeptanz, als auch einer internen Kooperation mit dem Max Planck Institut für Evolutionäre Anthropologie (MPI-EVA) in Leipzig, hat sich das Semantic MediaWiki (http://www.ontoloworld.org) als erweiterbares Basissystem herausgestellt. Neben der Möglichkeit Artikelseiten über semantische Links miteinander zu verknüpfen bzw. Attribute einzelnen Seiten zuweisen zu können, bietet das System zahlreiche Basis-Features zur semantischen Suche und zur Versionierung an. Aufgrund der eher mangelnden Nutzbarkeit (z.B. komplexe Syntaxen zur Beschreibung von semantischen Beziehungen), wurde das Basissystem innerhalb des Moduls weiterentwickelt, um den Anforderungen eines einfachen, intuitiv bedienbaren, webbasierten Editors gerecht zu werden. Dies betraf insbesondere das Metamodell. So wurden einerseits so genannte *Content-Typen* (z.B. Person, Projekt, Organisation, Publikation) in das existierende Basissystem integriert. Content-Typen werden verwendet, um neue Inhalte strukturiert erfassen zu können. Über automatisiert erzeugte Eingabeformen können Nutzer mit Hilfe von Web 2.0 Techniken (z.B. Autocomplete) Eingaben tätigen bzw. existierenden Inhalt verändern. Relationen (semantische Links zwischen Content-Typen) werden bidirektional gepflegt und erlauben somit eine verbesserte Navigierbarkeit und Auswertung erfassten Wissens. Die Erfassung neuer Content-Typen bzw. Änderungen an existierenden Content-Typen werden über eine Evolutionskomponente unterstützt. Neben den Content-Typen werden andererseits *Kategorien* zur Klassifikation von Inhalten nach bestimmten Eigenschaften eingesetzt. So können bspw. Projekte nach Projekttyp bzw. Projektstatus in entsprechende Kategorien eingeordnet werden.

Als erstes Anwendungsgebiet für die Erstellung einer domänenspezifischen Ontologie diente das Thema Grid, im Speziellen die semantische Erfassung der laufenden D-Grid-Initiative. So wurden typische Content-Typen im Bereich Grid wie bspw. *Personen*, *Ressourcen* (*Hard-* und *Software*), *Projekte* und *Organisationen* in die Ontologie aufgenommen. Zahlreiche visuelle Features (z.B. dynamisch erzeugte Karten über beteiligte Organisationen oder registrierte Gridressourcen) bzw. Navigationshilfen (z.B. Projekthierarchien), aber auch verbesserte Suchmöglichkeiten runden die Aufbereitung und Darstellung von bereits erfassten Wissen ab. Das System ist unter http://aprilia.izbi.uni-leipzig.de/dgo verfügbar und wurde nach Evaluierung in der MediGRID-Community D-Grid-weit bekannt gemacht (z.B. Demo auf dem All-Hands-Meeting 2007). Ebenfalls wurde das System inkl. des konzipierten Metamodells auf einem internationalen Grid-Symposium vorgestellt. Eine erweiterte Beschreibung für ein Special Issue des Information Sciences Journal ist derzeit noch unter Begutachtung.

Im letzten Teilpaket von AP 1 wurde die in AP 1.2 erstellte Ontologie-Middleware mit der MediGRID Infrastruktur verknüpft. Einerseits wurde die Middleware technisch mit den Komponenten des MediGRID verbunden. Konkret werden Ontologien auf den durch die Sonderinvestitionen beschafften Servern am IZBI in Leipzig verwaltet. So stellt der Ontologieserver alle derzeit in MediGRID benötigten Ontologien über einen OntologyService zur Verfügung. Der OntologyService sowie der zugehörige InfoService laufen innerhalb der Umgebung des Globus Toolkit GT 4 und sind über die Adressen http://buell.izbi.uni-leipzig.de:8081/wsrf/services/ogsadai/DataService bzw. http://buell.izbi.uni-leipzig.de:8081/wsrf /services/izbi/services/OntologyInfoService erreichbar. Des weiteren wurde der Zugriff auf die Middleware und somit auf verfügbare Ontologien über Anwendungen (Portlets) im zentralen MediGRID Portal sichergestellt. Für einen ersten Zugang (Browsing, Suche, Navigation) wurde ein zentraler „Ontology LookUp Service“ als Portlet im MediGRID Portal freigeschaltet (siehe Abbildung 30). Gleichermaßen wurden anwendungsspezifisch Portlets von MediGRID-Applikationen erweitert, um Ontologien mit anwendungsrelevanten Daten / Ergebnissen zu verknüpfen. So wurde die Applikation AUGUSTUS dahingehend erweitert, dass berechnete Ergebnisse mit Termen der SequenceOntology verknüpft werden, d.h. Anwender können sich mit Hilfe der verfügbarer Ontologie-Informationen detaillierter über Ergebnisteile informieren. In gleicher Art und Weise nutzt die Applikation SNPSelection die Schnittstellen der Middleware

aus, um Annotationen zu Gene Ontology Konzepten darzustellen. Details über die Integration der Middleware und Realisierung der Web-Schnittstellen im MediGRID Portal können dem Deliverable „D4 - Integration mit Grid Middleware“ entnommen werden. Die Verwendung der Ontologien im MediGRID Portal wurde in einem gemeinsamen Artikel der MediGRID Community im Future Generation Computer Systems Journal veröffentlicht (2009).

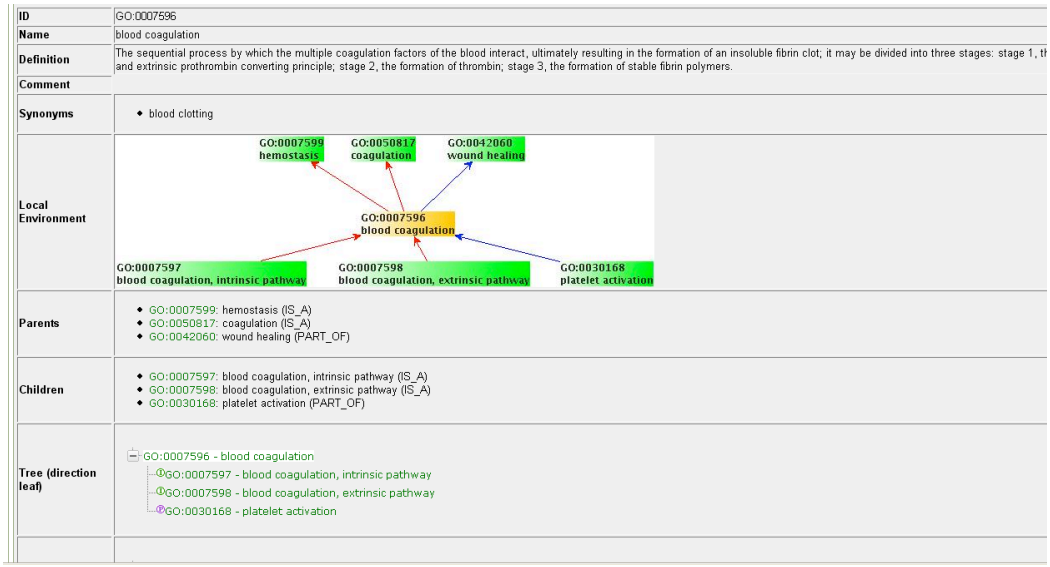


Abbildung 30. Verwendung von Ontologien im MediGRID Portal

Die Anforderungsanalyse für das Arbeitspaket 2 Ontologiematching umfasste die Befragung möglicher Nutzer innerhalb der MediGRID-Applikationen sowie eine Aufstellung existierender Match-Verfahren zur semi-automatischen Erzeugung von Ontologiemappings. Die Ergebnisse der Analyse sind im Deliverable „D5 - Anforderungsanalyse Ontologiematching“ dokumentiert. Aufgrund der vorhandenen Vorarbeiten zum Thema Ontologiematching im Modul wurde das Ziel der Integration vorhandener/berechneter Ontologiemappings (z.B. zwischen den Sub-Ontologien der Gene Ontology) in das MediGRID verfolgt. Dazu ist ein Repository für das Management und die Verwaltung berechneter Ontologiemappings zu konzipieren. Anschließend sind Mappings in das Repository einzuspielen, über Grid-Schnittstellen wird das Repository mit der MediGRID Middleware verknüpft.

Die Konzeption und Implementierungsarbeiten für das Repository fanden hauptsächlich in AP 2.2 statt. So wurde in einem ersten Schritt ein konzeptionelles Schema entwickelt. Das Schema umfasst Angaben zu den beteiligten Ontologien zwischen den Ontologiemappings erstellt werden. Zudem wird dokumentiert wie das entsprechende Mapping generiert wurde bspw. durch eine metadatenbasiertes, instanzbasiertes oder kombiniertes Match-Verfahren. Zu jeder Korrespondenz (Verknüpfung eines Konzepts der Quellontologie mit einem Konzept der Zielontologie) wird der Typ der semantischen Beziehung wie auch die Stärke (confidence) der Beziehung festgehalten. Anschließend wurde das konzeptionell entwickelte Schema in einem relationalen DBMS (MySQL) umgesetzt. Darauf aufbauend wurde eine Kernelsoftware mit Schnittstellen für den Import von Datenquellen (Ontologien) wie auch Mappings entwickelt und Tests unterzogen. Als erste prototypische Daten wurden die Ontologiemappings zwischen den drei Sub-Ontologien der GeneOntology in das Repository eingespielt. Die Daten stammen aus Ergebnissen, welche 2007 während der Evaluierung von instanz-basierten Matchverfahren produziert worden sind (siehe Publikation: „Instanced-based matching of large life science ontologies“).

Für die weitere Einbindung des Repository in das MediGRID und dessen Infrastruktur wurde auf dem Ontologieserver am IZBI in Leipzig ein entsprechender OGSA-DAI Datenservice eingerichtet. Der Datenservice http://buell.izbi.uni-leipzig.de:8081/wsrf/services/ogsadai/ DataService bietet unter dem Ressourcennamen Mappings@aprilia.izbi.uni-leipzig.de eine Grid-Schnittstelle zu verfügbaren Ontologiemappings in MediGRID an. Um die nun neu verfügbaren Informationen MediGRID-weit nutzen zu können, fand im letzten Schritt eine Einbindung in das zentrale MediGRID Portal statt. Dazu wurde das Ontologieportlet „Ontology LookUp Service“ erweitert und ergänzt. Neben den Informationen zu einem Ontologiekonzept, u.a. dessen Name, Accession, Definition und seine semantische Umgebung, können nun auch Korrespondenzen des Konzepts zu Konzepten in anderen Ontologien eingesehen werden. Dem Anwender wird eine Tabelle mit allen Korrespondenzen absteigend sortiert nach Stärke (confidence) präsentiert. Durch eine Verlinkung ist es möglich direkt zu den Konzepten in den anderen Ontologien zu navigieren, um bspw. Informationen über ähnliche Konzepte zu erhalten. Bei Anforderungen aus anderen bzw. neuen MediGRID Anwendungen können weitere Ontologiemappings in das Repository eingespielt werden. Eine ebenfalls direkte Verwendung der Ontologiemappings, bspw. für Integrationszwecke innerhalb einer MediGRID Anwendung, ist ebenfalls denkbar. Die Konzeption und Umsetzung des Repository zur Verwaltung von Ontologiemappings ist in Deliverable „D6 - Integration von Ontologiemappings“ beschrieben.

Weitere Arbeiten im Modul umfassten die Installation und Wartung der durch die Sonderinvestitionen 2006 beschafften Server. Zwei der drei beschafften Server dienen als redundante Portal-Lösung für das zentrale MediGRID Portal. In Zusammenarbeit mit dem Modul Middleware wurden die Server eingerichtet und das anfänglich in Göttingen betriebene Portal nach Leipzig migriert. Der dritte beschaffte Server dient als Ontologieserver und wird vorrangig für die Verwaltung benötigter Ontologien, des D-Grid Ontologie Systems sowie weiterer relevanter Datenbanken aus dem Bereich der Lebenswissenschaften verwendet. Nach der Teilnahme am Sonderinvestitionsworkshop Anfang 2007 wurde auf Basis der Referenzinstallation eine Globus-Umgebung sowie das zugehörige OGSA-DAI Framework installiert. Des weiteren wurde der Server in die Monitoring-Systeme von MediGRID und D-Grid (MDS, Ganglia) integriert. Das Nutzermanagement greift auf den in Jülich verfügbaren VORMS Service zurück.

2.6.2.2 Sonstige Ergebnisse

Die Abstimmung mit anderen Grid-Communities fand vorzugsweise auf Konferenzen oder Workshops statt. So wurden innerhalb von D-Grid der Sonderinvestitionsworkshop, das All-Hands-Meeting sowie der DGI Metadatenworkshop genutzt, um sich mit anderen Communities abzustimmen und Informationen auszutauschen. Ebenfalls wurde über gegenseitige Workshopbesuche (eigener Ontologie-Workshop sowie OntoVerse-Workshops) mit dem zeitgleich geförderten Projekt OntoVerse kooperiert. Abstimmungen mit internationalen Grid-Projekten insbesondere dem Cancer Biomedical Informatics Grid (caBIG) wurden durch einem Besuch des AMIA Meetings 2006 sowie einem gemeinsamen caBIG/MediGRID-Workshop Anfang 2008 in Berlin durchgeführt.

Der vom Modul und dem Projekt OntoVerse durchgeführte Workshop „Ontologien im Grid“ fand am 9./10. März 2006 in Leipzig statt. Neben Einführungsvorträgen zu Ontologien, wurde das am IMISE entwickelte Tool OntoBuilder und die Fusions- bzw. Datenintegrationsstrategien des Fachgebietes Datenbanken der Universität Leipzig vorgestellt. OntoVerse präsentierte in eigenen Vorträgen Inhalte und Ziele ihres Projektes. Der Workshop wurde von Mitarbeitern des MediGRID als von auch Interessenten anderer Projekten bzw. Institute besucht. Über die Website des Workshops (http://buell.izbi.uni-leipzig.de/gridworkshop) können nähere Informationen zum Programm und Ablauf abgerufen werden.

Weiterhin wurden während der Projektlaufzeit zahlreiche relevante internationale und nationale Veranstaltungen besucht. Oft wurden eigene Arbeiten vorgestellt und Ergebnisse mit teilneh-

menden Wissenschaftlern diskutiert. Zu den Veranstaltungen zählen u.a. Data Integration in the Life Sciences (2006, 2007, 2008), German e-Science Conference (2007), Annual AMIA Meeting (2006), GI-Fachtagung Datenbanksysteme für Business, Technologie und Web (2007, 2009), eHealth-Week (2007) oder Intl. Symposium on Intelligent Distributed Computing (2008).

Ergebnisse des Moduls wurden in relevanten Journals bzw. auf nationalen oder internationalen Konferenzen / Workshops veröffentlicht. So wurde die beschriebene Middleware auf der German e-Science Conference 2007 und am Workshop über die Grundlagen von Datenbanken 2008 vorgestellt. Die Arbeiten zur D-Grid Ontologie wurden auf einem internationalen Grid Symposium (IDC 2008) vorgetragen. Eine erweiterte Fassung für ein Special Issue des Information Sciences Journal ist noch unter Begutachtung. Arbeiten zum Matching bzw. zur Evolution von Ontologien wurden im Rahmen der internationalen Workshop-Reihe Data Integration in the Life Sciences (2007, 2008) publiziert. Gemeinsame Artikel mit anderen Modulen betrafen Veröffentlichungen in Future Generation Computer Systems sowie dem E-HEALTH-COM Magazin.

2.6.3 Nutzen und Verwertbarkeit der Modulergebnisse

- *Erfindungen/Schutzrechtsanmeldungen und erteilte Schutzrechte, die vom Zuwendungsempfänger oder von am Vorhaben Beteiligten gemacht oder in Anspruch genommen wurden, sowie deren standortbezogene Verwertung (Lizenzen u. a.) und erkennbare weitere Verwertungsmöglichkeiten:*

  Im Rahmen des Projektes sind weder Erfindungen/Schutzrechtanmeldungen erfolgt noch Schutzrechte angemeldet worden.

- *Wirtschaftliche Erfolgsaussichten nach Projektende (mit Zeithorizont) – z. B. auch funktionale/ wirtschaftliche Vorteile gegenüber Konkurrenzlösungen, Nutzen für verschiedene Anwendergruppen/-industrien am Standort Deutschland, Umsetzungs- und Transferstrategien (Angaben, soweit die Art des Vorhabens dies zulässt):*

  Wirtschaftliche Erfolge sind nicht absehbar.

- *Wissenschaftliche und/ oder technische Erfolgsaussichten nach Projektende (mit Zeithorizont) – u.a. wie die geplanten Ergebnisse in anderer Weise (z.B. für öffentliche Aufgaben, Datenbanken, Netzwerke, Transferstellen etc.) genutzt werden können. Dabei ist auch eine etwaige Zusammenarbeit mit anderen Einrichtungen, Firmen, Netzwerken, Forschungsstellen u.a. einzubeziehen:*

  Die Integration der verschiedenen Ontologien und die Bereitstellung im Rahmen des MediGRIDs ermöglichen im Bereich der Biomedizin neue Forschungsansätze. Aus diesem Grund sind die Erfolgsaussichten als gut zu bewerten. Insbesondere die Zusammenarbeit mit den Projekten der vernetzten Forschung (Kompetenznetze in der Medizin, Koordinierungszentren Klinische Studien KKS) kann dabei die Dissemination der Projektergebnisse unterstützen. Weiterhin können Ergebnisse direkt in Folgeprojekten wie Services@MediGRID oder MedInfoGrid weiter verwendet werden. Bspw. stellt die realisierte Ontologie-Middleware Schnittstellen zur Verfügung um weitere projektrelevante Ontologien ins Grid zu integrieren und in Anwendungen zu nutzen.

  Der kollaborative Aufbau einer Ontologie für die D-Grid Initiative ist für das gesamte Vorhaben inklusive seiner Teil- bzw. Folgeprojekte von Bedeutung. Das über die Communities bzw. DGI einfließende Wissen wird semantisch (ontologisch) fundiert gesichert und ermöglicht eine effizientere und somit nachhaltigere Nutzung gegenüber einer rein textuellen Erfassung.

- *Wissenschaftliche und wirtschaftliche Anschlussfähigkeit für eine mögliche notwendige nächste Phase bzw. die nächsten innovatorischen Schritte zur erfolgreichen Umsetzung der Ergebnisse:*

  Die noch unzureichende ontologische Unterstützung im Bereichen der Biomedizin (z.B. Klinische Studien) verhindert eine vollständige Integration aller Bereiche der Lebenswissenschaften. In diesen Feldern sind künftig noch entsprechende Entwicklungsarbeiten zu leisten. Dies betrifft vorrangig die Integration und Nutzung bereits existierender Ontologien sowie dem Aufbau entsprechender Ontologien / Standards insbesondere zum Datenaustausch und Verarbeitung im Grid. Gleichermaßen muss Ontologieentwicklungen /-veränderungen (Evolution) Rechnung getragen werden, da diese möglicherweise Inkonsistenzen / Fehlverhalten in den ontologie-nutzenden Applikationen bzw. Daten hervorrufen.

  Der entwickelte Prototyp zum Aufbau einer D-Grid Ontologie steht teilweise noch unter Entwicklung, ist aber bereits lauffähig und einsetzbar. Neue Anforderungen wie Mehrsprachigkeit oder Evolutionsaspekte können iterativ in das System eingebracht werden. Um das System jedoch professionell und längerfristig in allen Communities innerhalb des D-Grid betreiben so können, bedarf es breiterer Entwicklungen am System und sowie Abstimmungen mit den Communities.

  Das entwickelte Repository zum Management von Mappings zwischen unterschiedlichen Ontologien sowie die Möglichkeiten zur effizienten Speicherung von Ontologieversionen können projektübergreifend verwendet werden. Bspw. wurde eine webbasierte Applikation (OnEX – Ontology Evolution Explorer) zur Analyse der Evolution von Ontologien in den Lebenswissenschaften geschaffen. Das frei zugängliche System erlaubt Ontologiedesignern und Wissenschaftlern der jeweiligen Ontologiedomänen die Entwicklung (Evolution) ausgewählter Ontologien nachzuvollziehen und zu analysieren. Gewonnene Erkenntnisse können bspw. zur Annotation von Instanzdaten (z.B. Proteine) bzw. im Rahmen des Ontologiematchings verwendet werden.

2.6.4 Fortschritte auf dem Markt für Grid Computing in der Medizin und der Bioinformatik aus Modulsicht - Darstellung Ergebnisse Dritter

Für die Integration von Daten insbesondere strukturierte Datenquellen wie relationale Datenbanken, XML oder Ontologien wurden die Arbeiten / Entwicklungen zum OGSA-DAI Framework (http://www.ogsadai.org.uk/) beobachtet. OGSA-DAI stellt eine auf Globus basierende Middleware für die Integration von Datenquellen in Grids zur Verfügung. So genannte Datenservices gewähren einen Zugriff auf integrierte Datenquellen in einem Grid. Nach anfänglichen Performance-Tests wurde das Framework erweitert (siehe AP 1), um eine transparente Integration vers. Ontologien in das MediGRID zu gewährleisten. Gleichermaßen konnten über OGSA-DAI weitere relevante Datenbanken bzw. Repositories in MediGRID verfügbar gemacht werden, z.B. das Repository für Ontologiemappings aus AP 2.

Weiterhin Beobachtet wurden insbesondere die Arbeiten am National Cancer Instiute (NCI) innerhalb der Projekte caBIG (http://cabig.nci.nih.gov/) und caGRID (http://cagrid.org/). Durch ein zentrales Vokabular (NCI-Thesaurus), welches ebenfalls in MediGRID integriert wurde, und die Nutzung eines zentralen Repository für gemeinsam genutzte Datenobjekte (caDSR) ist ein ontologiebasiertes Metadatenmodell für den Austausch von Daten und die Kommunikation in Grids geschaffen worden. Der im Januar 2008 in Berlin durchgeführte caBIG / MediGRID Workshop bot MediGRID und somit auch dem Modul Ontologiewerkzeuge die Gelegenheit eines Informationsaustauschs mit caBIG. Es bleibt zu eruieren, inwieweit weitere Vorarbeiten / Werkzeuge aus obig genannten Projekten in das MediGRID bzw. dessen Folgeprojekte einfließen können.

Durch das rasante und dynamische Wachstum sowie die stark ansteigende Anzahl von Ontologien (insbesondere in Lebenswissenschaften) kommt dem Ontologie-Management eine immer größere Bedeutung zu. Bisherige Arbeiten konzentrierten sich vorrangig auf die Erstellung und Verarbeitung einzelner Ontologien. Hingegen wurden Probleme der Evolution von Ontologien insbesondere mögliche Einflüsse auf nutzende Anwendungen bzw. Daten kaum adressiert. Auch Match-Verfahren zur semi-automatischen Erstellung von Ontologiemappings bspw. für Datenintegrationszwecke bzw. Analysen wurden erst in den vergangenen Jahren aufgrund der wachsenden Einsatzgebiete von Ontologien näher untersucht, befinden sich jedoch vorrangig noch in prototypischer Entwicklung. Besonders das effiziente Matching großer Ontologien wie in den Lebenswissenschaften häufig anzutreffen, erfordern weitere Forschungsarbeit. Für einen über Überblick über relevante Arbeiten im Bereich Ontologie-Management wurde im Wintersemester 2008/2009 ein Problemseminar an der Abteilung Datenbanken der Universität Leipzig durchgeführt (http://dbs.uni-leipzig.de/stud/ws0809 /omsem).

Auf dem Gebiet der kollaborativen Ontologieentwicklung wurden relevante Entwicklungen insbesondere zu so genannten Semantischen Wikis (Wikis mit einem semantischen Metadatenmodell) beachtet. Die bisherigen Arbeiten konzentrierten sich vorrangig auf die Erweiterung bereits bestehender Wiki-Systeme um ein semantisches Metadatenmodell sowie die Einführung erweiterter Sprachkonstrukte zur semantische Erfassung von Wissen/Inhalt. So können Wiki-Artikel semantisch miteinander verlinkt werden oder erhalten über zugewiesene Attribute semantische Beschreibungen, welche bspw. bei Auswertungen (Anfragen) oder zur Navigation verwendet werden. In den Arbeiten zur D-Grid-Ontologie wurde insbesondere auf die bisher noch mangelnde Unterstützung von Eingabe-/Editiermöglichkeiten (Nutzbarkeit) Wert gelegt. Des weiteren erlauben Metamodell-Veränderungen (z.B. an Content-Typen oder Kategorien) die Umsetzung veränderter Anforderungen, z.B. aus interessierten Communities.

2.6.5 Ausblick und Perspektiven im Bereich Ontologiewerkzeuge

Während der Projektlaufzeit fokussierte das Modul Ontologiewerkzeuge innerhalb von MediGRID auf die Konzeption und Entwicklung von Werkzeugen für die Erstellung, Verfügbarmachung und das Matching von Ontologien in den Lebenswissenschaften. Die Integration existierender Ontologien in das MediGRID wurde mit einer transparenten und generischen Ontologie-Middleware begegnet (AP 1). Die Erstellung projekt-spezifischer Vokabulare / Ontologien wird über eine kollaborative Webumgebung ermöglicht, welche am Bsp. der D-Grid Ontologie evaluiert wurde (AP 1). Ontologiemappings werden über entsprechende Match-Verfahren erzeugt und können mittels eines Repositories in MediGRID verwendet werden (AP 2).

Die am Projekt im Bereich Ontologiewerkzeuge durchgeführten Arbeiten stellen einen ersten Schritt in Richtung semantischer Grids insbesondere in den Lebenswissenschaften dar. Wie aus dem caBIG-Vorhaben ersichtlich ist, wird der Einsatz von Ontologien bzw. Vokabularen im Bereich der Grid-Technologie verstärkt vorangetrieben. Insbesondere im Bereich der Medizin ist eine Standardisierung medizinischer Items und die Bereitstellung und Verwendung von Begriffssammlungen und Definitionen ein wichtiges Hilfsmittel zur Herstellung von Interoperabilität zwischen den beteiligten Organisationen. Die Interaktion verschiedener Services und das Zusammenwirken von Grid-Diensten erfordert neben der technischen Interoperabilität auch eine semantische Interoperabilität der Dienste und Daten. Die ontologiebasierte Infrastruktur des caBIG-Projekts beschreibt die unterschiedlichen Ressourcen des Grids in semantisch eindeutiger Art und Weise und verbessert somit deren Nutzbarkeit. Ontologieentwicklungen wie in caBIG bilden die Voraussetzung für organisationsübergreifende Arbeit in Gridumgebungen und sind ein wichtiges Element für die Qualitätssicherung in Grids.

Erzielte Ergebnisse können in direkte/indirekte Folgeprojekte des MediGRID einfließen. Hierzu zählen die beiden direkten Folgeprojekte Services@MediGRID und MedInfoGrid. Gleichermaßen ist das Modul im beantragten Projekt „Grid für die Wissenschaft - Wiss-GRID“ involviert.

In der Abteilung Datenbanken wurde synchron dazu ein Antrag auf Sachbeihilfe zum Thema „Evolution von Ontologien und Mappings“ bei der Deutschen Forschungsgemeinschaft (DFG) gestellt. Bei positiver Begutachtung können ebenfalls Erkenntnisse aus dem MediGRID in das neue Projekt einmünden bzw. begonnene Arbeiten / Forschungsrichtungen weiter vertieft und ausgebaut werden. Ein Verwendung bzw. der Einsatz von Grid-Technologie im Forschungsprogramm LIFE der Universität Leipzig wird derzeit gerade eruiert.

2.7 eScience – Enhanced Science

2.7.1 Moduleinordnung und Modulziele

Im Modul eScience / Rahmenarchitektur wurden einige der wichtigen community-spezifischen Eckpunkte behandelt und in die erforderliche Rahmenarchitektur einer eScience-Plattform implementiert. Neben einer wissenschaftlichen Integration flankierte das Modul eScience die anderen thematischen Module, in den drei Aufgaben, die im Umgang mit medizinischen Daten und biomedizinischen Informatik bedeutend sind: Datenschutz, Biometrie hochdimensionaler Daten und Übernahme international verfügbarer Lösungen für die biomedizinische GRID-Community in Deutschland. Ohne die flankierende Bearbeitung wären GRID-Anwendungen in diesem Kontext rechtlich und wissenschaftlich riskant.

Das Modul leistete die wissenschaftliche Integration der sechs Produktionsmodule durch Dokumentation und Analyse der Ergebnisse und Zusammenführung der komplexen Details in den verschiedenen Modulen zu einer Gesamtsicht. Diese Gesamtsicht wurde durch zwei Teilprojekte erarbeitet: die Entwicklung datenschutzrechtlicher Regeln und die Ableitung von biometrischen Standard Operating Prozeduren (SOPs) für hochdimensionale Daten im Grid. Die Ergebnisse wurden mit Erkenntnissen, die aus Fact-Finding Misssions (USA) sowie den europäischen Partnerprojekten (HealthGrid, EGEE) und Großbritannien (eScience) gewonnen werden konnten, korreliert.

Diese Arbeiten und seine Ergebnisse garantieren eine wissenschaftlich gesicherte Vorbereitung für die weitere Etablierung einer GRID-Infrastruktur in der Biomedizin. Wegen ähnliche Anforderungen in Punkto Datenschutz und Datensicherheit entwickelte sich im Verlauf eine Kooperation mit dem ingenieurswissenschaftlichen Community, die in der Gestaltung eines gemeinsamen Arbeitspaketes mündet, das im Folgenden beschrieben wird.

2.7.1.1 Enhanced Security and Trust

In diesem Arbeitspaket wurden Anforderungen bezüglich Sicherheit und Vertrauen definiert, die als Grundlage für die Arbeit in späteren Phasen von D-Grid dienen sollen. Die Anforderungen gehen teilweise weit über die Anforderungen anderer Wissenschaftsbereiche hinaus, haben aber andererseits einen hohen Grad an Überlappung. Dies liegt daran, dass man es in beiden Gebieten in der Regel mit Kollaborationsmodellen zu tun hat, die sich als Private Public Partnerships (PPP) beschreiben lassen. Auf der öffentlichen Seite sind die Partner Universitäten, Kliniken und Großforschungseinrichtungen, auf der privaten Seite im Bereich der Ingenieurwissenschaften die Industrie und im Falle der Medizin die Patienten. Aus inhaltlichen und ökonomischen Gründen war es nahe liegend dieses Teilprojekt in enger Abstimmung mit DGI durchzuführen.

Es wurden die Begriffe Safety, Security, Confidentiality und Trust behandelt. Unter Safety wird die physikalische Absicherung der Daten gegen Verlust und Veränderung verstanden. Unter Security versteht man den Schutz von Daten und Systemen vor unbefugtem Zugriff. Bei der Abbildung von Anwenderprozessen muss die Entscheidung darüber, welchen Ressourcen-Providern und Services vertraut werden kann, aber auch welchen Benutzern Zugriff auf einen Dienst oder eine Ressource gewährt werden kann, teilweise an Softwareinstanzen delegiert werden. Für die Delegation dieser Entscheidung ist das Vorhandensein von Trust-Metriken und -Infrastrukturen zur Vertrauensbewertung eine zwingende Voraussetzung, für die Anforderungen

und Spezifikationen erarbeitet, aber auch eigene auf die ingenieursspezifischen und medizinischen Gegebenheiten erweiterte Mechanismen entwickelt werden. Die Vertraulichkeit von Information und Wissen (Confidentiality) ist eine elementare Grundlage für die Zusammenarbeit gerade in Private Public Partnerships. Dies setzt voraus, dass der Besitzer von Information und Wissen die uneingeschränkte Kontrolle über sein intellektuelles Eigentum behält und dieses nicht an eine Zentralinstanz abgeben muss. Insbesondere im Bereich des Datenmanagements hat dies einen entscheidenden Einfluss auf die Akzeptanz und Einsetzbarkeit in sensitiven Anwendungen.

Ein Ziel war die Analyse der Eignung existierender Systeme und Technologien für das Management von Sicherheit und Vertrauen. Diese Analyse erfolgte insbesondere im Hinblick auf die angebotenen Lösungen und Konzepte in service-orientierten Infrastrukturen bzw. service-orientierten Grid-Lösungen. Ferner wurden die besonderen Anforderungen der ingenieurwissenschaftlichen und medizinischen Anwendungen an Arbeitsabläufe für die integrierte Grid-gestützte Problemlösung in die Betrachtung einbezogen. Es wurde eine Spezifikation der entsprechenden Anforderungen erarbeitet, die auf den Basisdiensten des D-Grid Integrationsprojektes aufsetzen. Die ausgewählten und teilweise durch das Integrationsprojekt zur Verfügung gestellten Basismechanismen wurden in die prototypischen Anwendungen von InGrid und MediGRID integriert. Die nicht statische Natur insbesondere des ingenieurwissenschaftlichen Grids führte zur Notwendigkeit, Methoden zu entwickeln, die eine Automatisierung der Interaktionsrechte von Anwendungskomponenten auf Basis von Sicherheits- und Vertrauensmetriken regelt.

Als weiteres Ziel dieses Arbeitspaketes wurden die hier entwickelten Sicherheits- und Vertrauensmechanismen in MediGRID und InGrid getestet. Hierbei wurde besonderes Augenmerk auf die Usability der Methoden in einer dynamischen Ad-hoc-Grid-Umgebung gelegt. Ein Bericht fasst die zur Verfügung stehenden Sicherheits- und Vertrauensmechanismen sowie eine Bewertung ihrer Anwendbarkeit für Beispielanwendungen zusammen.

Das Arbeitspaket wurde gemeinsam von InGrid und MediGRID bearbeitet. Die Arbeitsaufteilung innerhalb dieses gemeinsamen Arbeitspaketes war so gewählt, dass beide Projekte einzeln bestehen können, jedoch die Ergebnisse der Partnerprojekte in Kombination eine höhere Ergebnisqualität in beiden Projekten ermöglicht.

2.7.1.2 Biometrie vieldimensionaler Daten

Im Mittelpunkt der medizinischen GRID- Anwendungen stehen immer mehr hochdimensionale Daten, d.h. zu einer. Person liegen tausende von Merkmalen vor – im Vergleich zur Probandenzahl in den statistischen Modellen kehrt sich das Merkmal/Fall-Verhältnis um. Im Vergleich zu „klassischen“ Grid Communities stellt dies ein besonderes Problem dar. In der Projektlaufzeit wurden erst Vorgespräche mit den medizinischen Projektbeteiligten geführt. Der Projektplan sah die Erstellung von initialen Standard Operating Procedures vor. Dieses Arbeitspaket konnte erst nach der Fertigstellung der grundlegenden Infrastruktur begonnen werden, zusammen mit dem Modul Bioinformatik. Die Verarbeitung/ Aufbereitung und Interpretation von multidimensionalen Bilddaten stand im Mittelpunkt aller Betrachtungen. Somit waren die Arbeiten dynamisch an den zu liefernden Daten der Anwendungen gekoppelt. Der Beginn leitete sich aus dem zeitlichen Fortschritt der Module ab. Nach Fertigstellung der Grid-Infrastruktur und der Einarbeitung der beeinflussenden Faktoren wurde eine initiale SOP entwickelt. Die Datenerhebung ist für genetische Sequenzen im Modul Bioinformatik, für Bilddaten im Modul Bildverarbeitung realisiert worden. Im initialen Report zu einer Standard Operating Procedure (SOP) für hochdimensionale Daten in MediGRID sind drei Szenarien klassifiziert und untersucht worden. Es handelt sich dabei um die Bereiche klinische Studien, Registrierung und Meta-Analyse. Die dabei gewonnen Kenntnisse flossen in die beiden Revisionen ein. Die Problematik der Homogenisierung von biomedizinischen Daten in Metadaten ist aus den drei Anwendungsszenarien an der Struktur des gesamten MediGRID-Projektes angelehnt und die Prozesse wurden damit begleitet. Durch die

verstärkte Überschneidung bei biomedizinischen und bioinformatischen Fragestellungen war eine Erkenntnis für weitere Forschungen, dass Abläufe übergreifend in nachvollziehbaren hochwertigen Standard Operating Procedures (SOPs) zu dokumentieren bzw. die Systeme auf denen sie laufen weiter zu validieren sind.

2.7.2 Modulergebnisse

Die gewonnenen Erkenntnisse und wesentliche Ergebnisse werden in den folgenden Unterpunkten in den jeweiligen Arbeitspaketen beschrieben.

2.7.2.1 Ergebnisse des AP Enhanced Security and Trust (mit InGrid und DGI)

Eine ausführliche wissenschaftliche Analyse für Sicherheit in medizinischen Gridanwendungen führte zum Ergebnis, dass die meisten Sicherheitsanforderungen von „klassischen“ Grid-Infrastrukturen nur teilweise erfüllt werden. Den Bedarf zu eruieren und Lösungswege aufzuzeichnen war der Hauptpunkt dieses Arbeitspaketes. Im Projektzeitraum erfolgte eine detaillierte Abstimmung mit der ingenieurs-wissenschaftlichenCommunity und dem DGI. Dazu wurde ein Workshop „Enhanced Security für MediGrid und InGrid“ mit dem DGI vorbereitet. Ziel war eine Evaluierung der erfolgten Sicherheitsanalyse. Eine enge Abstimmung mit InGrid ist initiiert worden und wurde weiter intensiv betrieben. Dazu dienten insbesondere interaktive Telefonkonferenzen. Eine Vorarbeit zum Deliverable D2 war der Report zur Anforderungsanalyse existierender Sicherheitsmechanismen und Erarbeitung von Spezifikationen. Innerhalb dieses Arbeitspaketes fanden die D-Grid Security Workshops in Zusammenarbeit mit dem DGI und InGrid statt. In den Workshops wurden die existierenden Sicherheitsmechanismen gemeinsam mit den allen beteiligten Communities diskutiert und analysiert. Dabei wurden wichtigste sicherheitsrelevante Erkenntnisse erzielt und strategische Absprachen für das weitere gemeinsame Vorgehen getroffen.

Deliverable 2: Anforderungsanalyse

Die Analyse zeigt Sicherheitsmängel an den verwendeten Grid-Sicherheitstechniken auf. Die Notwendigkeit für einen erweiterten Sicherheitsservice mit Techniken wie Audit und feingranularen Zugriffsmechanismen wurde festgestellt.

Deliverable 3: D-Grid Security Workshop

Es wurden Themen aus den Bereichen PKI, Grid-CA, AAI, VOMS, GridShib, Firewalls, CERT-Dienste, juristische Rahmenbedingungen und die „Enhanced Security“ diskutiert und bearbeitet. Ergebnis war die Einigung auf Erweiterung der benutzten Sicherheitstechniken, um existierende Mängel zu beseitigen. Nachdem sich alle Beteiligten auf eine einheitliche Strategie für die Sicherheitskonzeption geeinigt haben, sind die Grid-Anwendungen entsprechend dieser Konzepte in den Testbetrieb gegangen und wurden dort einer umfassenden Evaluierung und Testung unterzogen. Erfahrungen und Vergleiche auf internationaler Ebene waren Gegenstand eines weiteren D-Grid Security Treffens im November 2006.

Deliverable 4: Initiale Datenschutzempfehlung (usage policy phase 1)

Als erster Schritt wurde eine Policy für die Nutzung von MediGRID-Ressourcen erstellt. In diesem Zusammenhang sind Szenarien und Use Cases für die Phase der Entwicklung von Applikationen festgelegt und Rahmenbedingungen geschaffen worden. Die Policy ist nach einem Review in die Beschlussphase gegangen. Letzte Änderungen wurden eingearbeitet und sind bindend für alle biomedizinischen Provider und Portaladministratoren bzw. - nutzer. Es wurden die Leitlinien für die Nutzung von Ressourcen während der Applikationsentwicklung (gridifing process) beschrieben.

Deliverable 5: Software und Beschreibung: Pilotumgebung „Enhanced Security“

Die Beschreibung der Pilotumgebung umfasst neben den Schritten für eine Nutzung der verfügbaren Gridanwendungen, auch das technische Konzept und die Lösungsansätze für die spezifischen Probleme in der medizinischen Community. Die Pilotumgebung wurde in zwei Abschnitten beschrieben. Der erste Abschnitt beschäftigt sich mit der Nutzersicht und den Anforderungen, was die Klassifizierung von Anwendungsklassen betrifft. Im zweiten Teil werden die medizinisch relevanten Sicherheitsaspekte beleuchtet. Lösungen werden für zwei Gruppen angeboten (mit und ohne Personendatenbezug).

Deliverable 6: D-Grid Security Workshop

Mit einem gleichbleibenden Teilnehmerkreis hat sich diese Veranstaltung zu einem festen Termin innerhalb des D-Grid etabliert. Eine enge Verzahnung von Lösungen konnte verzeichnet werden. Die erzielten Ergebnisse im Sicherheitsbereich mit Schwerpunkt Authentifizierung wurden mit dem zukünftigen Kernthema Autorisierung und den strategischen Aufgaben gegenübergestellt. Eine Aussage konnte herausgestellt werden, die enge Verknüpfung von AAI (Authentication and Authorization Infrastructure) und VO (Virtuelle Organisationen) –Management im Zusammenspiel mit Firewalls.

Deliverable 7: Generische Datenschutzempfehlung

Nach der erstellten Übersicht der Anforderungen der einzelnen Module hat sich aus Sicht der Strategie und eines standardisierten Vorgehens eine Überarbeitung notwendig gemacht. Im Rahmen eines MediGRID All-Hands-Meetings im Juli 2007 wurde eine Vorgehensweise analog dem Arbeitspaket erstellt und veröffentlicht. Die Überarbeitung konnte mit allen neuen Anforderungen und Ergebnissen finalisiert werden und wurde der MediGRID-Community zur Verfügung gestellt. Als neue Randbedingung ist das Service-Zertifikat dazugekommen, das mit der DGI intensiv diskutiert wurde.

Deliverable 8: Evaluation, Erfahrungen und Sicherheitsmechanismen in den Bereichen Engineering und Medizin

Dieser Report wurde zusammen mit InGrid, DGI und MediGRID im Vorfeld des vierten D-Grid Security-Workshops erstellt und vorgestellt. Dort enthalten ist der zum Zeitpunkt verfügbare State of the Art im D-Grid. Es ist eine Bestandsaufnahme bezogen auf die D-Grid I-Projekte.

Deliverable 9: Vorstellung der Ergebnisse aller Communities

Die Ergebnisse aus dem MediGRID All-Hands-Meeting mit den Informationen aus dem Security-Workshop beim D-Grid All-Hands-Meeting und die aufkommenden Service-Zertifikaten von Grid-Applikationen sind zusätzlich zu den bestehenden Punkten von AAI bis Trackability mit feingranularen Zugriffsrechten hinzugefügt worden. Diese Themen nahmen den Großteil des vierten D-Grid Security-Workshops ein, um auch den D-Grid II-Communities Lösungen zu ihren Anforderungen zu präsentieren. Das Service-Zertifikat hat sich als Notwendigkeit zur Identifizierung und Sicherheit gebildet. Dabei ist in Abstimmung mit DGI eine Spezifikation erforderlich, bei der die Abläufe und Integration in bestehende Strukturen beschrieben sein müssen. Als Ergebnis wurde eine thematische Veranstaltung zum „Ein Tag im Leben eines Grid-Jobs“ der breiten D-Grid I und II-Community durchgeführt. Eine Aufzeichnung im SuGI-Portal bereit.

Deliverable 10: Dissemination: Papier für internationale Grid Community

Die im Projekt erzielten Ergebnisse fanden internationale Beachtung, so dass eine global übergreifende Veröffentlichung im biomedizinischen Kontext mit gestaltet werden konnte. Das Proposal wurde Ende 2007 eingereicht und akzeptiert. Daraufhin begannen die Arbeiten an diesem Papier als Kapitel in dieser Veröffentlichung unter dem Thema „Data Protection and Data Security regarding Grid Computing in Biomedical Research“. Diese Arbeit wurde erfolgreich reviewed und wird im Mai 2009 in „Handbook of Research on Computational Grid Tech-

nologies for Life Sciences, Biomedicine andHealthcare“ im amerikanische IGI-Verlag erscheinen.

2.7.2.2 Ergebnisse des AP Biometrie vieldimensionaler Daten

Detaillierte Dokumente und Vorgehensweisen bei der Verarbeitung hochdimensionaler Daten auf dem Grid liegen in Form von generalisierten Working Procedures vor. Diese bilden die Grundlage für alle weiteren biomedizinischen Grid-Vorhaben und werden durch die Community fortgeschrieben.

Deliverable 11: Initiale SOPs für vieldimensionale Daten (Draft)

In Zusammenarbeit mit dem KKS Düsseldorf sind vorhandene Daten und Methoden für Grid-Anwendungen evaluiert und validiert worden. Ergebnis ist eine erste SOP für die Verarbeitung von biomedizinischen Daten im Grid. Diese SOP ist in allen Anwendungsszenarien evaluiert worden. Es wurde sichergestellt das aktuelle europäische Entwicklungen Beachtung finden.

Deliverable 12: Einführung und Revision der SOPs für hochdimensionale Daten

In einer ersten Phase wurden die SOPs vorgestellt und am realen Ablauf getestet (Anwendungsszenarien aus D11). Die SOPs (Deliverable 11) wurden auf Vorstandsebene eingeführt. Es wurde über die Inhalte und rechtlichen Rahmenbedingungen innerhalb der methodischen und klinischen Module diskutiert. Ergebnis sind die redigierten Papiere in der Phase der ersten Revision. Die Dokumente waren Vorlage der Data Protection Working Group innerhalb des caBIG/ MediGRID-Workshops.

Deliverable 13: 2. Revision der SOPs für hochdimensionale Daten

In der Revision wurden die in Deliverable 12 gesammelten Reaktionen und offenen Fragen bearbeitet. Fokussiert wurde auf die Punkte Haftungsrecht, Betriebskonzept und Musterverträge. Dieses Ergebnis und der Meilenstein M1 (Verabschiedung/ Inkraftsetzung) wurden in der Produktivphase eingesetzt und werden Grundlage für das Biometrie-AP im service@MediGRID sein.

Mit dem Verabschieden der zweiten Revision liegt eine Fassung vor, die jene Prozesse im klinischen Umfeld bei der Nutzung von Grid Techniken beschreibt und regelt. Darauf aufbauend sind erste weiterführende Aufgabenstellungen mit dem service@MediGRID-Projektes beschrieben und begonnen worden, der inhaltliche Übergang für Umsetzungen und Anwendungen ist damit realisiert worden.

2.7.2.3 Ergebnisse des AP Internationale Kooperationen

Der Bereich internationale Kooperation wurde auf zwei Bereiche aufgegliedert, um die Ergebnisse besser darzustellen. Die Arbeit als zusammenführende und integrierende Komponente wird damit herausgestellt. Es wurde nicht übernommen oder adaptiert, sondern es ist an einem generellen für alle Grid-Community gültigen Ansatz erfolgreich gearbeitet worden.

Europäische Union

Eine Kontaktaufnahme mit dem französischen Grid-Projekt MediGRID wurde im Projektzeitraum intensiv versucht. Durch intensive Recherche ist eine Kontaktaufnahme mit einem Vertreter gelungen. Dadurch konnten wichtige sicherheitsrelevante Ergebnisse aus den MediGRID-Projekten ausgetauscht werden. Mit den Grid-Kollegen aus Österreich (UMIT, Prof. Ammenwerth) fanden Abstimmungsgespräche statt, eine Intensivierung wurde vorbereitet. Innerhalb der eHealth-Konferenz 2007 in Berlin wurde ein Track Medical Grid Computing von MediGRID und HealthGrid im EU-Kontext durchgeführt. Eingebettet darin war die Skizzierung der direkten Beeinflussung der Versorgung auf Ebene deutsches und europäisches Gesundheitswesen. Klinische Aspekte spielten zusammen mit kostenkalkulatorischen Eckpunkten eine wichtige Rolle im

Bezug auf Infrastruktur und Nachhaltigkeit in den vorhandenen Strukturen. Die Zusammenarbeit auf europäischer Ebene mit der HealthGrid-Initiative verläuft sehr erfolgreich. Ein Beispiel ist die gemeinsame Veröffentlichung im Nano-Verlag mit der gemeinsamen europäischen Sicht der Entwicklungen.

International/ USA

Durch die Initiierung bei der Jahrestagung der AMIA wurde ein MediGRID- Workshop in Boston 2006 mit Vertretern der HealthGrid-Initiative der EU, der Universitäten von Pennsylvania und Stanford (US HealthGrid), NCI, NCBi, Fox Chase Cancer Center, caBIG, caGRID, dem MIT, dem General Hospital Boston, I2B2, sowie der Harvard Medical School in Washington DC vorbereitet und durchgeführt. Dies war eine Folge der Abstimmung mit US Medizin- Grids zur Intensivierung der Zusammenarbeit und der Partizipierung an Forschungsergebnissen. Die Zielsetzung war Informationen über den Stand der Grid-Aktivitäten im Sektor HealthCare/ Life Sciences in den USA (und international) zu erhalten, Kontakte mit führenden Experten in diesem Bereich zu knüpfen, eine konkrete Klärung vorzunehmen, ob und welche gemeinsamen Projekte und Aktivitäten mit internationalen Projekten möglich und sinnvoll sind. Es konnte resümiert werden, dass das Thema „Grid“ in USA und EU viel weiter aufgefasst und in Projekten breiter angelegt wird, das technology-driven vs. use-case-driven vorherrscht, eine stärkere Förderung der semantischen Interoperabilität angestrebt wird, die Verzahnung von Versorgungs- und Forschungsdaten bereits begonnen hat, eine strenge OpenSource-Philosophie verfolgt wird und keine Verwertungsauflagen (caBIG) existieren. Es wurde auf einem Treffen mit caBIG-Verantwortlichen aus den USA die Zusammenarbeit auf die Punkte Applikationen und Kooperation festgelegt, und ist 2008 in Berlin konkretisiert worden. Eine erfolgreiche internationale Kollaboration mit gemeinsamen Vorhaben wurde gestartet. Die vereinbarte Zusammenarbeit mit den europäischen und amerikanischen Kollegen wurde in einem Fachtreffen weiter konkretisiert und intensiviert. In diesem Fachgremium sind Vertreter der HealthGrid-Initiative, dem US HealthGrid, dem caBIG/ caGRID und den universitären Einrichtungen im Bostoner Umfeld zusammengekommen, um als Zielsetzung gemeinsame Aktivitäten und Projekte zu beraten. In den Vereinbahrungen integriert ist die Invitation von MediGRID zu wichtigen Tagungen der caBIG-Gremien. Besonderes Interesse hat die Nachhaltigkeit Die unterschiedlichen Begrifflichkeiten und der breite Ansatz von Grid-Aktivitäten wurden zusammengefasst. Es wird an einem grundsätzlichen strategischen Papier für die Grid-Community gearbeitet. Erste Informationen sind der biomedizinischen Community in der Fachpresse (Szolovits 2007) angekündigt worden, was sukzessive ausgebaut wird. Mit den assoziierten Partnern in Boston fand und findet ein regelmäßiger Austausch statt, insbesondere zur Einschätzung aktueller Entwicklungen im Umgang mit personenbezogenen Daten (Personal Health Record, genomische Daten, Datenschutz). Das Dach Forum Grid bei der TMF e. V. hat dabei eine besondere Rolle, da der Austausch auf eine möglichst breite Basis gestellt worden ist.

In der gesamten Projektlaufzeit konnten folgende Veranstaltungen vorbereitet, durchgeführt und besucht werden:

2.7.2.4 Durchführung von eigenen Veranstaltungen

- Kickoff-Meeting am 1.9.2005
- Heidelberger Informatik-Kolloquium, 09.01.2006 (1 Vortrag)
- Gründertreffen GoeGrid-Runde, Göttingen, 10.03.2006 (Organisation: eScience-Modul)
- D-Grid Security Workshop, Göttingen, 21-22.03.2006 (Organisation: eScience-Modul)
- Fact Finding Mission, Boston, 16-25.09.2006
- D-Grid Security Treffen mit InGrid, DGI, TextGrid und MediGRID, Göttingen, 06.11.06

- USA-Workshop mit caBIG, Healthgrid und caGRID, Washington, 11-15.11.06
- D-Grid Security Workshop, UMG, 27.03.-28.03.2007 (2 Vorträge),
- Telemed/Special Interest Session (eHealth-week), Berlin, 16.04.-19.04.2007 (5 Vorträge)
- Teststellung CertiWare (eGK, HBA, Gematik-Infrastruktur), UMG, 23.04.-25.04.2007
- Workshop Generisches Sicherheitskonzept (All-Hands-Meeting), Berlin, 13.07.2007
- Meeting zur Ressourcenplanung, Berlin, 14.08.2007
- D-Grid All-Hands-Meeting, Göttingen, 10.-12.09.2007 (1 Vortrag)
- Security-Workshop mit DGI und InGrid, Göttingen (Session), 11.09.2007
- US Workshop mit Vertretern von caBIG (Disseminierung), Philadelphia, 01.10.2007
- Datenschutzkonzept für die klinische Forschung, Marburg, 25.02.2008
- D-Grid Security Workshop, Göttingen, 01.-02.04.2008
- Einweihung des Göttinger Grid Ressourcenzentrums, Göttingen, 13.05.2008
- Konstituierende Sitzung Forum Grid, Berlin, 02.07.2008
- Gründung Executive Board GoeGrid, GWDG10.07.2008
- Sitzung Technical Board GoeGrid, Physik-Campus, 14.07.2008

2.7.2.5 Teilnahme an Veranstaltungen

- Stand auf der MEDICA 2005 in Düsseldorf (Präsentation in MedicaMedia-Reihe)
- Sitzung des MediGRID-Vorstands, Berlin, am 21./22.11.2005
- MediGrid-Workshop, Berlin, 22.11.2005
- Präsentation des Projektes auf der MEDICA in Düsseldorf, BMBF- Stand, 16.-19.11.2005
- D-Grid GAT-Workshop, Potsdam (ZIB), 19.01.2006
- Globus-Workshop, München, 24.01.2006
- Anwendertreffen Bioinformatik, Göttingen, 10.02.2006
- Gridsphere-Workshop, Göttingen, 01.03.2006
- Ontologie-Workshop, Leipzig, 09-10.03.2006
- BMBF Kick Off, Bonn, 06.04.2006
- „Rechtliche Aspekte der Telemedizin“, Berlin, 06.04.2006
- Telemed, Berlin, 07-08.04.2006 (Vortrag)
- Datamanagement Workshop, Berlin, 18-19.05.2006
- Industrieworkshop auf der ITeG, Frankfurt, 01.06.2006
- Healthgrid-Workshop, Valencia, 07.-09.06.2006
- ICT, Brüssel, 29-30.06.2006.(1 Poster)
- Gridsphere-Workshop für MediGRID-Anwender, Göttingen, 03-04.07.2006
- Sitzung der AG Datenschutz der TMF e.V., Berlin, 07.09.2006
- GMDS-Jahrestagung, Leipzig, 10-14.09.2006 (1 Poster, 1 Vortrag)
- GridKA, Karlsruhe, 11-15.09.2006

- EGEE, Genf, 25-29.09.2006 (1 Poster)
- Nachhaltigkeitsworkshop, Karlsruhe, 09-10.10.2006 (1 Vortrag)
- Jahrestagung der HL7 Benutzergruppe, Göttingen, 25-27.10.2006
- Grid-Seminar, Paulinerkirche, Göttingen, 23.11.06
- Internationale Datenschutzkonzepte Workshop der TMF e.V., Berlin, 08.12.2006,
- Grid@Asia, Seoul, 11-14.12.2006 (1 Vortrag)
- Sicherheit in medizinischer Verbundforschung (TMF), Berlin, 12.12.2006, (1 Vortrag)
- Workshop für Förderempfänger der D-Grid-Sonderinvestitionen, Karlsruhe, 15.01.2007
- Workshop zu rechtlichen Fragestellungen bei Biomaterialbanken, UMG, 24.01.2007
- Zwischenbegutachtung MediGRID, Berlin, 01.03.-02.03.2007
- D-Grid- und links4science-Workshop "Wissensschicht für Grid", Göttingen, 29.03.2007
- HealthGrid 2007 Conference, Genf, 24.04.-27.04.2007 (1 Poster)
- German eScience Conference, Baden-Baden, 02.05.-04.05.2007 (1 Vortrag, 1 Workshop)
- Nachhaltigkeitsworkshop, DFN Berlin, 11.06.-12.06.2007
- MediGRID Exponat auf ISC2007, Dresden, 26.06.-29.06.2007
- MedInfo 2007, Brisbane, 20.-24.08.2007 (1 Vortrag)
- 2. Göttinger Grid-Seminar, Göttingen, 13.09.2007 (1 Vortrag)
- GMDS-Jahrestagung, Augsburg, 17.-19.09.2007 (1 Postervortrag)
- 3. MediGRID Vorstandssitzung, Berlin, 28.09.2007
- D-Grid Nachhaltigkeitsworkshop, Berlin, 09.-10.10.2007
- International Conference CeHR, Regensburg, 02.-05.12.2007 (1 Vortrag)
- 4. MediGRID Vorstandssitzung, Berlin, 05.12.2007
- caBIG Workshop, Berlin, 22.-23.01.2008
- D-Grid Accounting/Billing-Workshop, Hannover, 13./14.03.2008
- Healthgrid-Tagung, Chicago (USA), 02.-04.06.2008
- 2. D-Grid Accounting/Billing-Workshop, Hannover, 07.08.2008
- Sitzung Technical Board GoeGrid, Physik-Campus, 11.08.2008
- Sitzung Technical Board GoeGrid, Physik-Campus, 08.09.2008
- GMDS-Jahrestagung, Stuttgart, 15.-18.09.2008
- Sitzung Technical Board GoeGrid, Physik-Campus, 09.12.2008

Publikationen, die in der Projektlaufzeit entstanden, sind im nachfolgenden Punkt drei aufgelistet, diese sind modulübergreifend als Gesamtwerke und als Ergebnis von Arbeitspaketen entstanden und zu verstehen.

2.7.3 Nutzen und Verwertbarkeit der Modulergebnisse

Alle im Modul erzielten Ergebnisse können in der biomedizinischen Community und darüber hinaus weiter verwendet werden. Der Nutzen liegt im Bereich Coaching, der Ressourcennutzung und deren Ausbau, sowie der Disseminierung von Erkenntnissen.

- Erfindungen/Schutzrechtsanmeldungen und erteilte Schutzrechte, die vom Zuwendungsempfänger oder von am Vorhaben Beteiligten gemacht oder in Anspruch genommen wurden, sowie deren standortbezogene Verwertung (Lizenzen u. a.) und erkennbare weitere Verwertungsmöglichkeiten,

Im Rahmen des Projektes sind weder Erfindungen noch Schutzrechtanmeldungen erfolgt, noch Schutzrechte angemeldet worden. In Bezug auf eine Nachhaltigkeit sind Diskussionen seit der Zwischenbegutachtung erfolgt, die weiter getrieben worden und eine konkrete Umsetzung im Rahmen von Sonderinvestitionen und den weiter geförderten Projekten erleben. Dort sind Konzepte und Pläne vorgelegt worden, die maßgebend vom Modul und seinen integrativen Wirken gezeichnet sind.

- Wirtschaftliche Erfolgsaussichten nach Projektende (mit Zeithorizont) – z. B. auch funktionale/ wirtschaftliche Vorteile gegenüber Konkurrenzlösungen, Nutzen für verschiedene Anwendergruppen/-industrien am Standort Deutschland, Umsetzungs- und Transferstrategien (Angaben, soweit die Art des Vorhabens dies zulässt),

Die wirtschaftlichen Erfolgsaussichten konnten zum Anfang des Projektes noch nicht abgeschätzt oder verbindlich getroffen werden. Im Laufe des Projektes sind nachhaltige Betriebsmodelle in enger Kooperation mit den Communities erarbeitet worden. Eine Nachhaltigkeit darf aber nicht nur unter finanziellen Gesichtspunkten gesehen werden. Eine nachhaltige Infrastruktur hat primär wissenschaftlichen Nutzen. Neue Arten der wissenschaftlichen Zusammenarbeit wurden gebildet. Ziel ist eine langfristige Unabhängigkeit von Fördermitteln. Dazu musste die Anzahl der Benutzer und Applikationen erweitert werden, Anforderungen aus dem Bereich Kunden, Ressourcen und Support wurden eruiert und betrachtet. Die Beteiligung der Industrie ist notwendig und gewünscht. KMUs mit kurzfristigem, hohem Bedarf an Grid-Ressourcen haben großes Interesse an einer dauerhaft verfügbaren Grid-Infrastruktur. Größere Unternehmen sehen ihre Rolle in erster Linie als Anbieter von Diensten. Insbesondere wird von der Sonderinvestition ein Impuls zur Verbreitung von Grid-Ideen und damit eine vergrößerte Nutzerbasis aus Forschung und Wirtschaft erwartet. Durch die getätigten Sonderinvestitionen mit diesem Hintergrund wurde eine größere Anzahl von Nutzern darauf vorbereitet und der Einsatz getestet im Bereich Coaching. Es wurde seit Projektbeginn an diesem Problem gearbeitet und im Rahmen der Sonderinvestitionen wurde community-übergreifend eine Rahmenvereinbahrung geschlossen. Die Punkte Betrieb und die Pflege, sowie Coaching wurden herausgestellt. Eine fair share-Regelung wird weit reichende Bedeutung für alle D-Grid-Bereiche haben.

- Wissenschaftliche und/ oder technische Erfolgsaussichten nach Projektende (mit Zeithorizont) – u.a. wie die geplanten Ergebnisse in anderer Weise (z.B. für öffentliche Aufgaben, Datenbanken, Netzwerke, Transferstellen etc.) genutzt werden können. Dabei ist auch eine etwaige Zusammenarbeit mit anderen Einrichtungen, Firmen, Netzwerken, Forschungsstellen u.a. einzubeziehen,

Die gewonnenen Ergebnisse sind sehr interessant für medizinische Forschungsnetze in Deutschland (TMF). Im Rahmen der Mitarbeit in den Arbeitsgruppen Datenschutz und Biomaterialbanken der TMF wurde eine Fortschreibung der generischen Datenschutzkonzepte in Bezug auf Versorgungsaspekte, Internationalität und verteilter Datenhaltung (Grid) begonnen und weiter vorangetrieben. Besonders durch die neuen Aspekte von Biomaterialbanken und genomweiten Assoziationsstudien, Aspekte der Bildverarbeitung, klinischen Studien und Datenschutz aus dem Grid-Kontext sind in dem generischen Konzept der TMF weiter modelliert und fortgeschrieben worden. Zusätzlich sind generelle SOPs für Grid-Anwendungen in die klinischen Bereiche aktuell eingebracht worden, was direkten Einfluss auf Arbeiten im Bereich Management Klinischer Studien hat. Ein Dachverband innerhalb der TMF ist gebildet worden, der alle nationalen biomedizinischen Grid-Projekte bündelt und koordiniert. Mit dem

Forum Grid ist die enge Abstimmung national und international der biomedizinischen Grid-Projekte sichergestellt.

Die assoziierten Partner in den USA zeigten sich bei Arbeitstreffen in Washington und Boston interessiert an der Homogenisierung von heterogenen Datenquellen, an Ontologie und auch an Datenschutzthemen. Diese Aspekte waren bei gemeinsamen Workshops mit der Harvard Medical School und dem MIT von besonderem Interesse. Zusammen wurde die Herausforderung von genomweitern Assoziationsstudien angegangen. Harvard Medical bringt die Erfahrung im Umgang mit genomischen Daten ein, MediGRID/eScience Erfahrung in Datenschutzfragen und in der Strukturierung medizinischer Daten. Mit caBIG-Vertretern wurde ein Workshop erfolgreich durchgeführt, wo die Zusammenarbeit langfristig vereinbart wurde. Mit einer eigenen Session bei der HealthGrid-Tagung in Chicago 2008 überzeugte die MediGRID-Anwendung die Fachkollegen, auch da global die Verfügbarkeit und Realisierung erfolgreich demonstriert wurde und im direkten Vergleich überzeugte.

Weiterhin wurde das Thema „Grid computing“ in die Lehrinhalte des Studienganges Medizinische Informatik in Göttingen aufgenommen. Somit wurden wichtige Grundlagen gelegt, um den Nachwuchs zu fördern. In studentischen Arbeiten ist eine Strategie für die Universitätsmedizin Göttingen mit Beteiligung von MediGRID erarbeitet worden und eine virtualisierte Testumgebung ausgetestet worden, um erste Sicherheitskonzepte zu validieren, die dann auf medizinische Anwendungen übertragbar sind. Diese Aktivitäten sind über die Medizinische Fakultät hinaus bekannt geworden, was dazu führte das die Anregung eines Grid-Seminars in der universitären Ausbildung aufgenommen wurde. Die ansässigen Communities (HEP, InstantGrid, MediGRID und TextGrid) führten zum Wintersemester 2007/8 ein interdisziplinäres Seminar "Grid Computing in der Wissenschaft" - Konzepte, Methoden und Anwendungen – zusammen mit dem Institut für Informatik durch. Zielgruppe waren Studierende im Hauptstudium der Naturwissenschaften und Forscher. Es erfolgte ein Wissenstransfer über die Bandbreite von Middleware , Monitoring, Data Storage, Transfer, Performance, Quality Assurance und Parallelisierungskonzepte. Das Seminar stellte „die“ fakultätsübergreifende Veranstaltung in Göttingen dar und hat daher Vorbildcharakter. Als Konsequenz wurde im Sommersemester 2008 ein Praktikum "Programmierung von Grid-Anwendungen“ am Zentrum für Informatik unter Beteiligung der Grid-Communities gestartet. Dieses Praktikum bewegte das Institut für Informatik die Grid Programmierung tiefer in den Lehrinhalten zu verankern mit eigenen Angeboten im Bereich Ausbildung der Kerninformatik. Inhalte des Praktikums bezogen sich auf Testung der MediGRID-Komponenten, der VO education und den Ausbau von Coaching-Tools. Hier wurde demonstriert, das Ergebnisse von und in die Forschung fließen und kommerzielle Partner davon profitieren können und ein Anreiz für Grid-Nutzung geschaffen wird.

- .Wissenschaftliche und wirtschaftliche Anschlussfähigkeit für eine mögliche notwendige nächste Phase bzw. die nächsten innovatorischen Schritte zur erfolgreichen Umsetzung der Ergebnisse.

Durch die intensiven Vorbereitungsarbeiten und Abstimmungen für den D-Grid II-Call wurden wichtige Kontakte in die biomedizinische Industrie geknüpft, wodurch insbesondere bei High Throughput-Providern ein großes Interesse an MediGRID und D-Grid entstanden ist. Es war eine feste Vorgehensweise, das Resultate aus MediGRID in D-Grid II einflossen und weitergeführt bzw. ausgebaut wurden. Gewonnene Ergebnisse sind und werden in aktuelle und zukünftige Anträge im Grid-Kontext einfließen. Das services@MediGRID - Projekt partizipiert direkt von vorhandenen Ergebnissen. Damit wird ein erster Teil der Nachhaltigkeit der D-Grid I-Projekte in diesem D-Grid II-Projekt gewährleistet.

2.7.4 Fortschritte auf dem Markt für Grid Computing in der Medizin und der Bioinformatik aus Modulsicht - Darstellung Ergebnisse Dritter

Die Bereiche Monitoring, Accounting und Billing sowie GridSphere Portalentwicklung und vor allem Authentication und Authorization-Infrastruktur (AAI) wurden und werden auch in Zukunft (D-Grid II) aufmerksam mitverfolgt um eine enge Abstimmung mit dem DGI sicherzustellen. Einfluss werden betrachtete Entwicklungen auf dem Bereich der Ident- Dokumente für den Bereich der Authentifizierung haben (Digitaler Personalausweis kontaktlos), elektronische Gesundheitskarte und der Heilberufe- Ausweis (kontaktbasiert). Herausgestellt wurde die Bedeutung einer homogenen AAI- Struktur. Im Security-Bereich wurde eine verstärkte Verbreitung und Akzeptanz von Shibboleth-basierten AA Infrastrukturen bemerkt und verfolgt, was eine Anpassung des Middlewarekonzepts zur Folge hat.

Die aktuelle Entwicklung von Biomaterialbanken innerhalb medizinischer Forschungsverbünde hat auch Auswirkung auf MediGRID. Einerseits werden die Datenschutzanforderungen durch das Einbringen von Material- bzw. Analysedaten noch komplexer, andererseits können Lösungsansätze für die Verarbeitung im Grid mit eingebracht werden. Die Gesetzgebung im BMB-Bereich (Gewebegesetz- Entwurf) schafft Rahmenbedingungen für die Umsetzung im Grid. Durch die Mitarbeit in der AG der TMF wird dies in die Netze getragen und gleichzeitig die Modelle und Szenarien angepasst.

Die Etablierung der in MediGRID entwickelten Sicherheitskonzepte und –lösungen ist im nationalen - und internationalen Gebiet weiter vorangeschritten. Ob im D-Grid, HealthGrid oder EGEE die Policies und die nun die neu benötigten Service-Zertifikate setzen neue Maßstäbe im Vergleich zu anderen Organisationen. Das amerikanische caBIG ist dem D-Grid/ MediGRID keineswegs voraus. Eine Nutzung und Übertragung aus caBIG nach Deutschland ist möglich und kann geleistet werden.

Perspektivisch soll im Rahmen von Services@MediGRID anwendungsnah Grid Computing an Sequenzierungs- und Microarrayexperimenten herangetragen werden. Hierbei sind Querbezüge zu weiteren Projekten in Göttingen hergestellt worden. Die Analyse der Haplotypen ist in der Versuchsphase, wo Einflüsse auf Datenstrom, Verarbeitung und Prozesssteuerung getestet werden. Die Ergebnisse haben unmittelbaren Einfluss auf die nachgelagerten stochastischen Verfahren. In einen Workshop zu Anwendungsszenarien ist die Umsetzung von R-Algorithmen auf Grid-Ressourcen skizziert und durchgespielt worden. Hier zeigt sich der Vorteil der OpenSource-Philosophie. Die statistischen Methoden können dort auf den Nodes problemlos installiert werden. Das Zuladen der Daten geschieht über die bekannten Speicherkonzepte von OGSA und SRB. Über eine interdisziplinäre Veranstaltung im Fachbereich Medizin unter dem Titel „Personalisierte Medizin“ konnten weitere Abteilungen, wie z.B. die Molekulare Medizin für den Bereich biomedizinisches Grid Computing gewonnen werden.

2.7.5 Ausblick und Perspektiven im Bereich eScience

In der gesamten Laufzeit ist die Produktivphase des Grid-Portals durch Anwender/ Entwickler und Nutzer (auch für die Lehre) evaluiert und ausgiebig getestet werden. Die Applikationen und die Infrastruktur überzeugten und wurden sicher für den Dauereinsatz konzipiert. Die eingesetzte PKI erfüllte die im Rahmen des Moduls eScience definierten Bedingungen einer sicheren medizinischen Grid-Umgebung für die Anwenderklassen. Die Ergebnisse waren die Grundlage für die Aktivitäten im zweiten Call mit Beteiligung von Unternehmen. Die Erfahrungen mit Daten im Grid konnten biometrische Methoden hinzugestellt werden, was Einfluss auf komplexe systembiologische Strukturen hat. Zusätzlich wurde über einen Workshop die Integrationsfähigkeit der gematik-Infrastruktur beleuchtet.

Die strategischen Sonderinvestitionen des BMBF unterstützten die Communities bei der Vorbereitung des Forschungs-Grid für den industriellen Einsatz und eines Wissenschafts-Grid. In den

Nutzungsberichten belegten das zentrale Controling und die Zahlen der Ressourcenprovidern den kalkulierten und eingetretenen Nutzen und die richtige Auslegung der Investitionen. Dabei sind die standortförderlichen vorhandenen Synergien, die durch GöGrid existieren, auf die HEP-Community ausgeweitet worden. Das Ergebnis ist ein zukunftsträchtiges Grid-Ressourcen-Zentrum für vier Communities. In der Veranstaltung zur Einweihung des Göttinger Grid-Ressourcenzentrums überzeugte dies neben nationalen auch internationale Fachleute. Die Auslastung der Ressourcen konnte 2008 stabil auf über 80% ausgebaut und gehalten werden. Ziel war und bleibt eine höhere Effizienz bei der Auslastung.

2.8 Geschäftsmodelle für Healthgrids

2.8.1 Moduleinordnung und Ziele

Neben der Ausgestaltung und Stabilisierung der MediGRID-Infrastruktur ist ein nachhaltiger Betrieb der Grid-Infrastruktur über das Projektende hinaus ein wesentliches Ziel des Konsortiums.

Um eine strategische Ausrichtung nachhaltiger Folgeaktivitäten zu ermöglichen, wurde im Rahmen des Moduls Koordination (TMF e.V.), primär zuständig für die Gesamtkoordination und Vernetzung der Projektpartner, wissenschaftlich untersucht, wie nachhaltige Geschäftsmodelle für Grid Computing innerhalb der Medizin und der Life Sciences ausgestaltet sein sollen. Neben konkreten projektbezogenen Untersuchungen wurden die Erkenntnisse einer themenbegleitenden Promotion und die Ergebnisse einer beauftragten Marktuntersuchung in die Analyse einbezogen.

Die Untersuchung gliedert sich in folgende Abschnitte:

Methodische Module:

- Marktstudie anhand 23 weltweiter Grid-Projekte aus dem Medizin- und Life Sciences Bereich, einschließlich der Wirtschaftlichkeitsuntersuchung am Beispiel von Fallstudien ausgewählter Grid-Projekten
- Expertenbefragung innerhalb des MediGRID-Projektes sowie umfassende internationale Expertenbefragung von 33 Grid-Experten

Inhaltliche Module:

- Entwicklung eines Geschäftsreferenzmodells als Modellrahmen für nachhaltige Grid-Computing-Aktivitäten
- Erfolg versprechende Geschäftsmodelltypen für Grid Computing in der Medizin und den Life Sciences
- Analyse der Wertkette und Identifizierung der beteiligten Akteure des Wertschöpfungsnetzwerks und ihrer Beziehungen
- Entwicklung generischer Erlösmodellen für unterschiedliche Anwendungsszenarien
- Betriebs- und Kostenanalyse des MediGRID-Projektes für einen nachhaltigen Betrieb

2.8.2 Modulergebnisse

2.8.2.1 Beobachtung des Marktes für Grid Computing

Der Markt für Informationssysteme wächst und ist in den kommenden Jahren dynamischen Veränderungen unterworfen. Allein innerhalb der letzten zehn Jahre ist die Nachfrage nach Rechen- und Netzwerkleistung sprunghaft gestiegen (Buyya und Venugopal 2005). Dieser Trend ist auch im Gesundheitswesen zu beobachten. Der Digitalisierungsgrad in der Medizin steigt zunehmend. Zusätzlich nimmt gerade im wissenschaftlichen Umfeld die Zahl an Projekten zu, deren Mitglieder organisationsübergreifend miteinander kooperieren und komplexe Daten und Informationen austauschen. Neue Erkenntnisse wie die Entschlüsselung des menschlichen Genoms beflügeln IT-affine Bereiche wie die Biomedizin zusätzlich. Die Entschlüsselung bestimmter Krankheitserreger ist z.T. allein von der zur Verfügung stehenden Rechenkapazität abhängig. Diese Entwicklungen verlangen nach adäquaten technologischen Lösungen, die durch Grid-Computing-Infrastrukturen bereit gestellt werden können.

Der Markt für Grid Computing lässt sich aus Anbietersicht in unternehmensinterne Grids (Department Grids und Enterprise Grids), in Partner Grids und in Open Grids aufteilen (vgl. Plaszczak und Wellner Jr. 2006) (siehe Abbildung 31). Enterprise Grids haben sich in verschiedenen Branchen bereits etabliert, so auch im Gesundheitssektor; u.a. betreiben die Pharmaunternehmen GlaxoSmithKline und Novartis eigenen Grid-Systeme zur in-silico-Forschung. Open Grids sind freie dynamische Grids mit wechselnden Partnern, die perspektivisch angestrebt werden, aber momentan noch nicht am Markt etabliert sind. Partner Grids, zu denen auch MediGRID zählt, formen eine Virtuelle Organisation fixierter voneinander unabhängiger Partner. Sie befinden sich in einer frühen Entwicklungsphase und kommen derzeit primär im öffentlich geförderten Forschungsbereich zum Einsatz. Die hohe technologische und organisationsspezifische Komplexität von Partner-Grids, u.a. bezüglich der Frage der technischen Vernetzung heterogener Systeme und die komplexe vertragliche Gestaltung im Rahmen einer Virtuellen Organisation bestehend aus diversen unabhängigen Partnern, bremsen eine schnelle Marktdurchdringung. Auf kommerzieller Seite werden homogene Hardwareplattformen mit wenigen Partnern auf der Anbieterseite bevorzugt.

In den vergangenen Jahren wurden Marktgröße und Marktpotenzial von Grid Computing in unterschiedlichen Studien bewertet (vgl. u.a. Cohen 2004; Insight Corp. 2005, 2006; Longbottom und Tarzey 2006). Grundsätzlich erschweren neue Technologien einer seriösen Abschätzung des Marktvolumens, da die Abgrenzung des Marktes häufig nicht eindeutig und schwierig ist. Allen vorliegenden Studien ist jedoch gemein, allgemein als Grid Computing verstandenen Infrastrukturen ein überdurchschnittliches positives Wachstum in den kommenden Jahren zu bescheinigen.

Insight Research Corp. beziffert das Marktvolumen für Grid Computing in allen Branchen im Jahre 2005 auf 715 Mio. US$ und prognostiziert ein Wachstum auf über 24 Mrd. US$ im Jahre 2011 (vgl. Insight Corp. 2006). Das prognostizierte Marktwachstum wird zu einem Großteil von unternehmensweiten Grids (Enterprise Grids) getragen, signifikante Marktanteile für Partner-Grids und Open Grids sind erst ab dem Jahre 2007 zu erwarten.

Konkrete Zahlen für die Entwicklung bei Healthgrids liegen nicht vor. Dennoch ist erkennbar, dass in den vergangenen Jahren eine Vielzahl von Healthgrid-Projekten, primär Partner Grids, initiiert wurden. Während frühere Grid-Projekte den Fokus auf hochrechenintensive Problemstellungen legten, steht gerade in jüngeren Projekten mehr und mehr die Vernetzung von patientenrelevanten Daten im Vordergrund (vgl. Breton et al. 2005).

Im Vergleich zu anderen Branchen, wie der Hochenergiephysik oder den Finanzdienstleistungen, kann aufgrund komplexerer Rahmenbedingungen, insbesondere rechtlicher Beschränkungen, in der Biomedizin von einer verzögerten Adaption und Verbreitung von Grid Computing Infrastrukturen ausgegangen werden. Aktuelle Projekte können als *Early Adopters* verstanden werden, während mit einer Stabilisierung der Nutzung auf breiter Basis zwischen 2011 und 2017 gerechnet werden kann (*Early Majority*). Eine vollständige Marktdurchdringung ist erst ab dem Jahre 2017 denkbar (*Late Majority*) (vgl. Andoulsi et al. 2008).

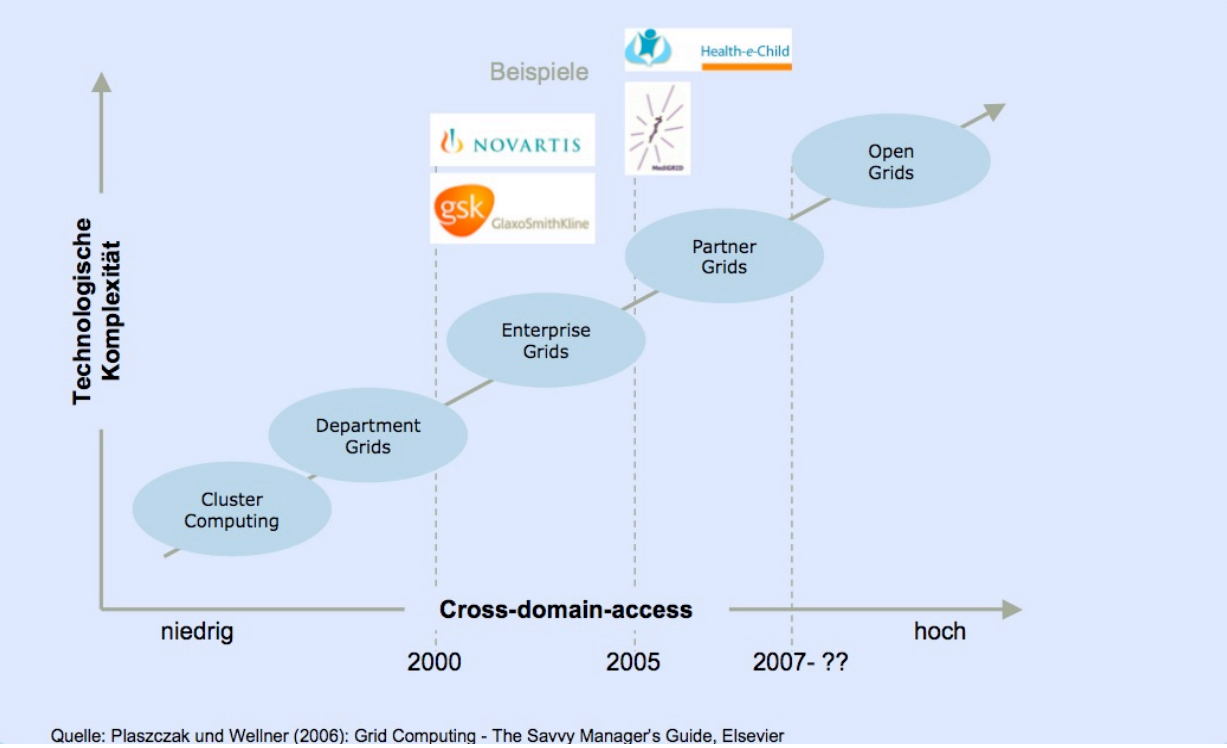


Abbildung 31. Entwicklungspfad von Grid-Typen

2.8.3 Geschäftsmodelle für Grid Computing

Forschungsanstrengungen, die insbesondere unter dem Oberbegriff „Grid Economics“ zu finden sind, konzentrieren sich hauptsächlich auf eine effiziente preisbasierte Allokation von Ressourcen innerhalb von Grid-Infrastrukturen. Dies ist im Rahmen interner Kostenverrechnung und einer effizienten Werterstellung von Bedeutung, allerdings beschreiben diese Modelle nicht die übergeordnete Fragestellung, inwieweit Grid-Aktivitäten selbst marktgerecht und kundenorientiert sind. Geschäftsmodelle leisten hier einen entscheidenden Beitrag.

Innerhalb des Electronic Business haben sich verschiedene Geschäftsmodelltheorien herausgebildet, die in ihrer Grundsystematik auf Grid-Computing-Strukturen übertragbar sind. Die wirtschaftswissenschaftliche Forschung untersuchte im Rahmen der Entwicklung des Internets hin zu einer Breitenanwendung tiefgehend die aufkommenden Geschäftsideen und ihre Erfolgsfaktoren. Darauf aufbauend entwickelte sich eine umfassende Literaturbasis zur Theorie der Geschäftsmodelle im Electronic Business (Timmers 1998; Scheer et al. 2003; Osterwalder et al. 2005).

Unter einem Geschäftsmodell kann ein Handlungsrahmen verstanden werden, der eine (neue) Technologie und andere unternehmensinhärente Ressourcen als Inputvariablen versteht, und diese über Markterfolge und Kunden in wirtschaftliche Leistung wandelt. Ein Geschäftsmodell dient insofern auch als fokussierendes Medium zwischen technologischer Innovation und der Schaffung neuer Werte.

Vom Geschäftsmodelle abzugrenzen ist das Konzept der Nachhaltigkeit. Während Geschäftsmodelle eine konkrete Architektur zur Umsetzung eines Geschäftszieles bereit stellen, beschreibt der Nachhaltigkeitsbegriff eine übergeordnete Zielstellung oder Mission eines Unternehmens und seiner Geschäftstätigkeit. Geschäftsmodelle, die sich an der Zielsetzung der Nachhaltigkeit orientieren, verfolgen insofern kein kurzfristiges Gewinnziel, sondern streben eine dauerhafte und stabile Teilnahme am Marktgeschehen an.

Eine sinnvolle Geschäftsmodellarchitektur sollte verschiedene Teilmodelle berücksichtigen. Wichte Teilmodelle eines Geschäftsmodells sind: Nutzensicht, Wertschöpfungssicht und Kapitalsicht. Eine ausführliche Erläuterung der Ebenen findet sich in Scholz et al. (2008).

Neben diesen drei Grundsäulen sollte ein Geschäftsmodell gerade für Healthgrids limitierende und fördernde Rahmenbedingungen berücksichtigen. Hierzu zählen die Grid-Technologie selbst, die sowohl neue Geschäftsfelder erschließen lässt als auch gleichzeitig aufgrund ihrer Natur bestimmte Geschäftsfelder nicht ermöglicht, rechtliche Rahmenbedingungen, insbesondere die Gesetzgebung in datenschutzrelevanten und ethischen Bereichen. Ein mögliches Referenzgeschäftsmodell für Healthgrids, welches diesen Vorgaben folgt, wird in Abbildung 32 vorgeschlagen.

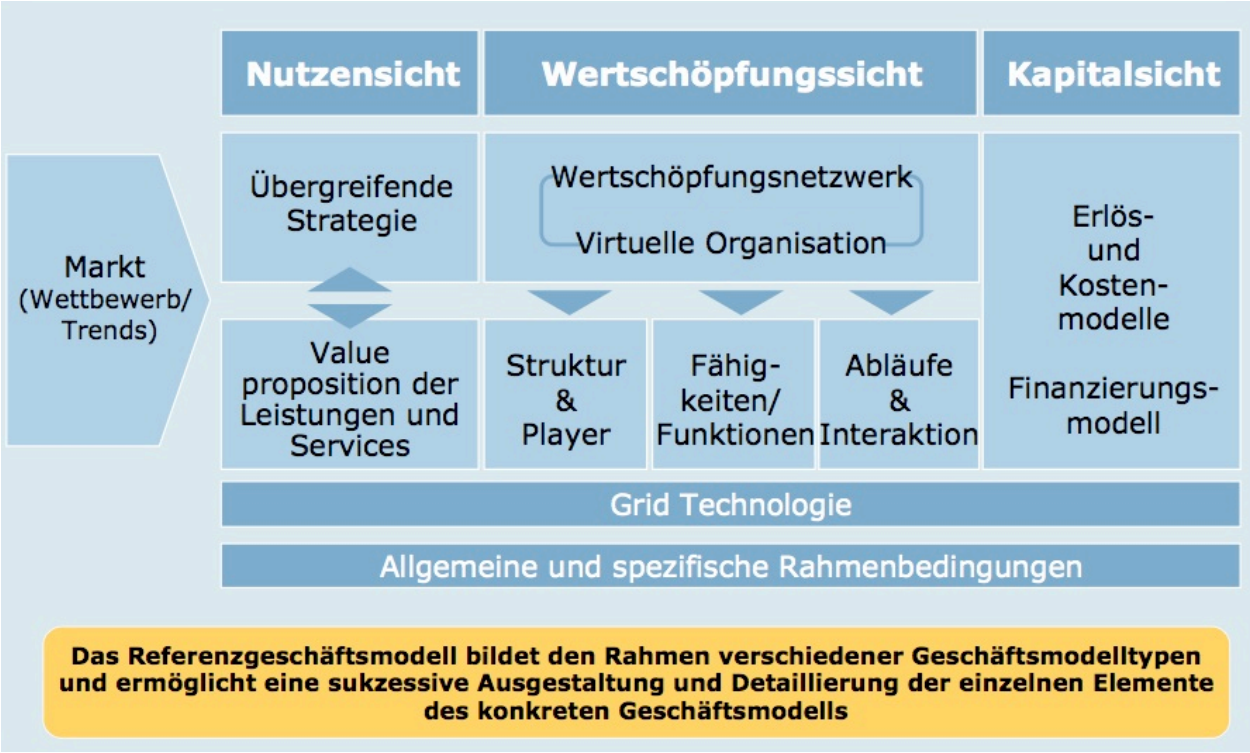


Abbildung 32. Referenzgeschäftsmodell für Healthgrids (Scholz et al. 2008)

Die konkrete Ausgestaltung eines Geschäftsmodells anhand eines Referenzmodells ist abhängig vom konkreten Typ des Geschäftsmodells im untersuchten Marktumfeld. Typisierungen[4], also die Kategorisierung verschiedener Geschäfte anhand definierter Kriterien, wurden in der wissenschaftlichen Literatur anhand der Wertschöpfungsstruktur, der angebotenen Leistung, des Erlösmodells und anderer Klassifizierungsrahmen vorgenommen. In vielen Arbeiten wurde insbesondere auf die Position und Interaktion in der Wertschöpfungskette verwiesen.

Die Wertschöpfungskette für Grid Computing (siehe Abbildung 33) lässt sich in sechs Rollengruppen gliedern (Scholz 2008):

- **Hardware Provider**: Bereitsteller der physischen IT-Infrastruktur (Netze, Rechen- und Speicherressourcen)

[4] In der Literatur findet sich häufig der Begriff der Taxonomie zur Beschreibung einer Klassifikation von (Geschäftsmodell-)typen

- **Support Service Provider**: Plattformsoftware und funktionsrelevante Softwarekomponenten (Middleware, Allokations- und Controllingsysteme, Billing & Accounting)
- **Ressourcen Service Provider**: Unterstützende i.d.R. „nichtdigitaler" Dienstleistungen (Vertikale Dienste für mehrere Wertschöpfungsstufen, z.B. rechtliche Fragen)
- **Content Provider**: Zur Nutzung bereitgestellte Inhalte (Anwendungen, Datenbanken, etc.; einschließlich Services zur Gridifizierung von Anwendungen)
- **Content Service Provider**: Vermittlungsleistung, Bereitstellungsleistungen, Anpassungsleistungen für Zielkunden
- **Endnutzer**: Endnutzer der Dienste - für Healthgrids: Forschung (gewerbl. und akadem.), Klinken, niedergelassene Ärzte, (Patienten, öffentliche Hand)

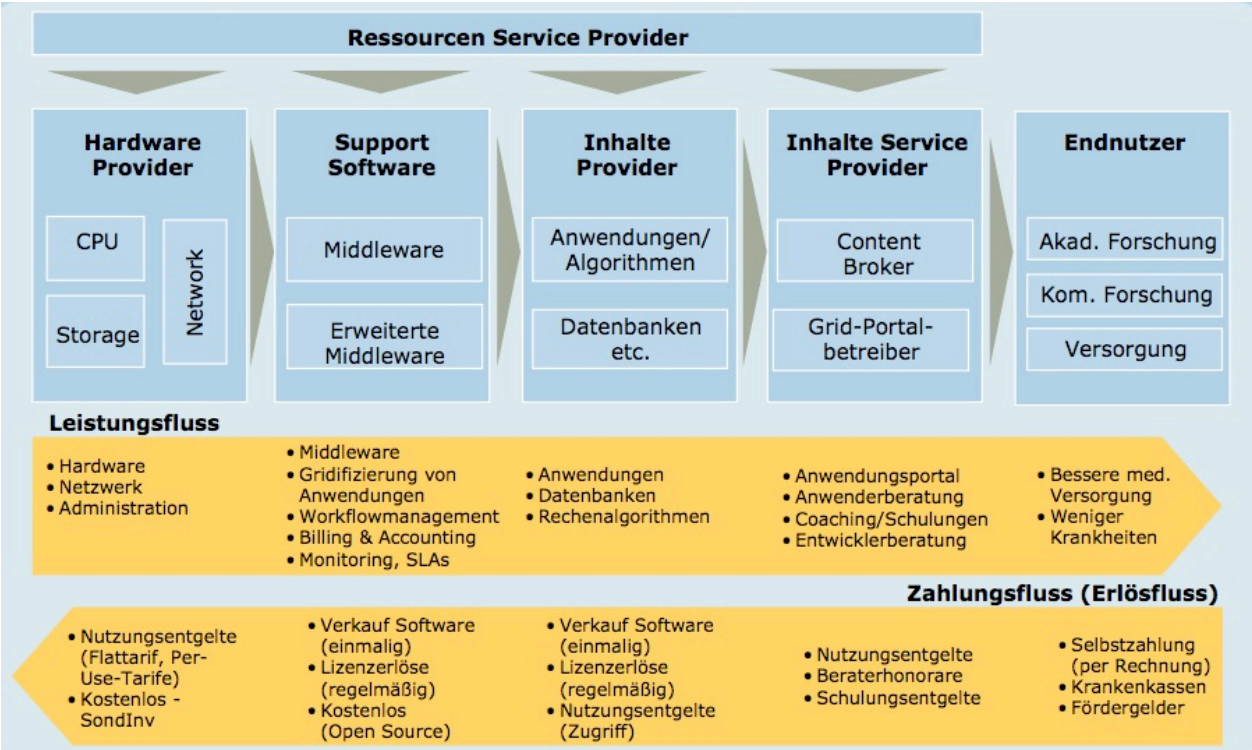


Abbildung 33. Wertschöpfungskette Grid Computing (Scholz 2008)

Die unterschiedlichen Rollenmodelle am Markt für Grid Computing bieten die Möglichkeit zur Herausbildung spezifisch konfigurierter Wertschöpfungsnetzwerke.

Auf Basis der besonderen Konfiguration der Wertschöpfungskette und der organisatorischen Struktur von Grid-Computing-Infrastrukturen lassen sich insbesondere sechs unterschiedlich zusammengesetzte Wertschöpfungsnetzwerke als Geschäftsmodelltypen für Grid-Computing-Infrastrukturen identifizieren (Scholz 2009). Alle sechs Grundtypen unterscheiden sich hinsichtlich der organisatorischen Integration und des Grades der Wertintegration, d.h. des Einbezugs unterschiedlicher Wertschöpfungsstufen (siehe Abbildung 34). Einige dieser Geschäftsmodelltypen sind bereits am Markt etabliert (insbes. Typ I, Typ III, Typ IV und Typ V), andere werden sich mit der Weiterentwicklung des Grid Marktes herausbilden (insbes. Typ II und Typ VI). Eine Beschreibung der einzelnen Typen gibt Tabelle 1.

Tabelle 1 Geschäftsmodelltaxonomie für Grid Computing

Geschäftsmodelltyp	Kurzbeschreibung
Inhouse Gridbox (Typ I)	Grundressourcenpaket aus Hard- und Middleware, Plug&Play, standardisierte Schnittstellen, insbes. für KMU
Grid-ASP (Typ II)	Vertragliche Zusammenarbeit mit Grids, die Anwendungen bereit stellen, umfassendes Anwendungsportfolio
Inhouse Grid (Typ III)	Grundressourcenpaket aus Hard- und Middleware, umfangreicher Implementierungsservice, Einbindung eigener Anbieterdienste, insbes. für große Unternehmen/ Einrichtungen
Utility Grid (Typ IV)	Bereitstellung von Rechen- und Speicherressourcen extern, Plug&Play, definierte Schnittstellen, häufig zusätzliche Services
Statisches SaaS-Grid (Typ V)	Umfassende Grid-Infrastruktur, Bereitstellung definierter Anwendungen, def. Kundenkreis, Kunden i.d.R. Teil der VO
Dynamisches SaaS-Grid (Typ VI)	Umfassende Grid-Infrastruktur, dynamisches Anwendungsportfolio, dynamisches Kundenportfolio

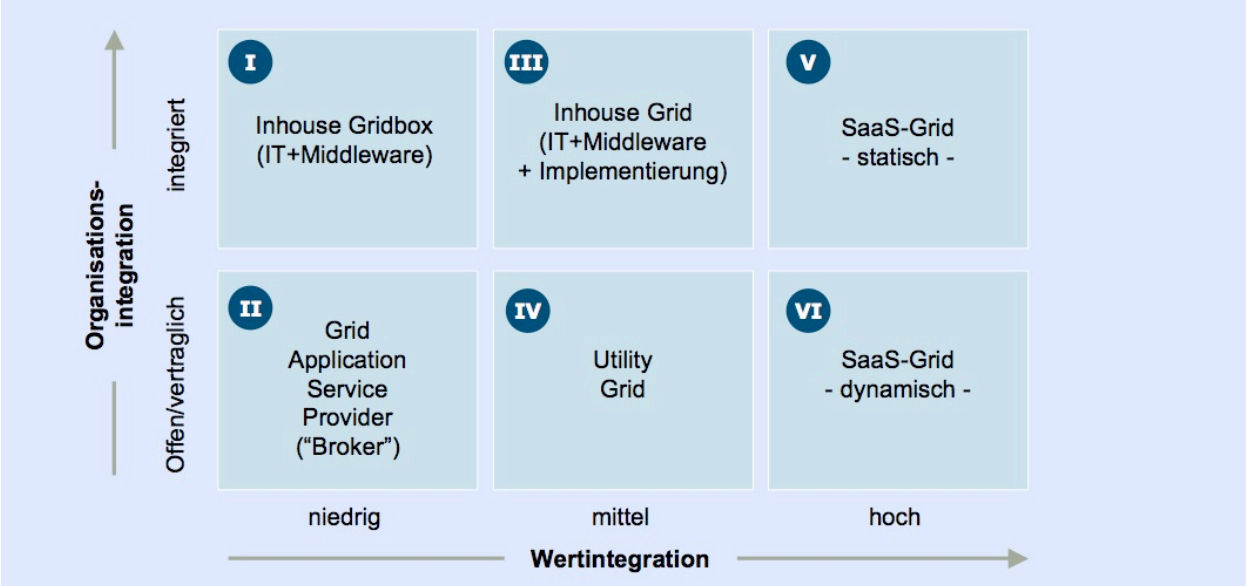


Abbildung 34. Typisierung von Geschäftsmodellen (Scholz, 2009)

2.8.3.1 Zielgruppenmärkte für Healthgrids – Anwendungsfelder und Kunden

Grid Computing wird in den Gebieten der Bildverarbeitung, der biomedizinischen Simulationen und Kalkulationen, der Visualisierung/Modellierung, der Verarbeitung von Biosignalen, der Verwaltung von Patientendaten und der Klinischen Forschung eingesetzt. Viele der am Markt aktiven Projekte dienen als Plattform für verschiedene Anwendungsbereiche (z.B. MediGRID – Biomedizinische Grundlagenforschung, Bildverarbeitung und Biosignalauswertung), selten sind Projekte auf ein einzelnes Anwendungsfeld spezialisiert (z.B. MammoGrid - Mammographien). In einigen Fällen dient ein breites Anwendungsspektrum der Untersuchung eines einzelnen Krankheitsbildes oder eines medizinischen Fachgebietes (z.B. @NeurIST – cerebrale Aneurysmen, Health-e-Child – Kinderheilkunde). Einen Überblick über Anwendungsfelder des Grid Computing gibt Abbildung 35.

Bildarchivierung und die Bildverarbeitung bilden den am häufigsten zum Einsatz kommenden Anwendungstyp. Neben der Vernetzung von Bilddaten stellen verschiedene Projekte gleichzeitig

Diagnose- und Analysetools zur Bildverarbeitung bereit. Der ausgeprägte Einsatz ist insbesondere auf die Grid Computing begünstigenden Rahmenbedingungen zurückzuführen (hohes Datenvolumen, etablierte Standards).

Biomedizinische Simulationen und Kalkulationen sind ein weiteres häufig zum Einsatz kommendes Anwendungsgebiet. Hierunter fallen Analysen aus der Genetik als auch der Proteomik sowie daraus abgeleitete Simulationen.

An dritter Stelle stehen Projekte, die ihre Plattformen explizit auf die Unterstützung *Klinischer Studien* ausrichten. Im Untersuchungsfeld fallen in diesen Bereich epidemiologische und genetische Studien, aber auch Studien zu spezifischen Krankheitsbildern. Projekte, die sich auf dieses Anwendungsgebiet konzentrieren, kombinieren dies häufig mit komplementären Anwendungen, wie z.B. Bildverarbeitung oder biomedizinischen Simulationen.

Eine kleine Anzahl der am Markt aktiven Projekte integriert Anwendungen, die explizit der *Entscheidungsunterstützung* im Rahmen der Patientenbehandlung dienen. Dies sind zum Teil interaktive Systeme zum Abruf vorausgewerteter Patienteninformationen, zum Teil Systeme, die ex-post eine Beurteilung durchgeführter Maßnahmen erlauben. Des weiteren ermöglichen verschiedene Projekte die *Modellierungen oder Visualisierungen* biologischer Vorgänge.

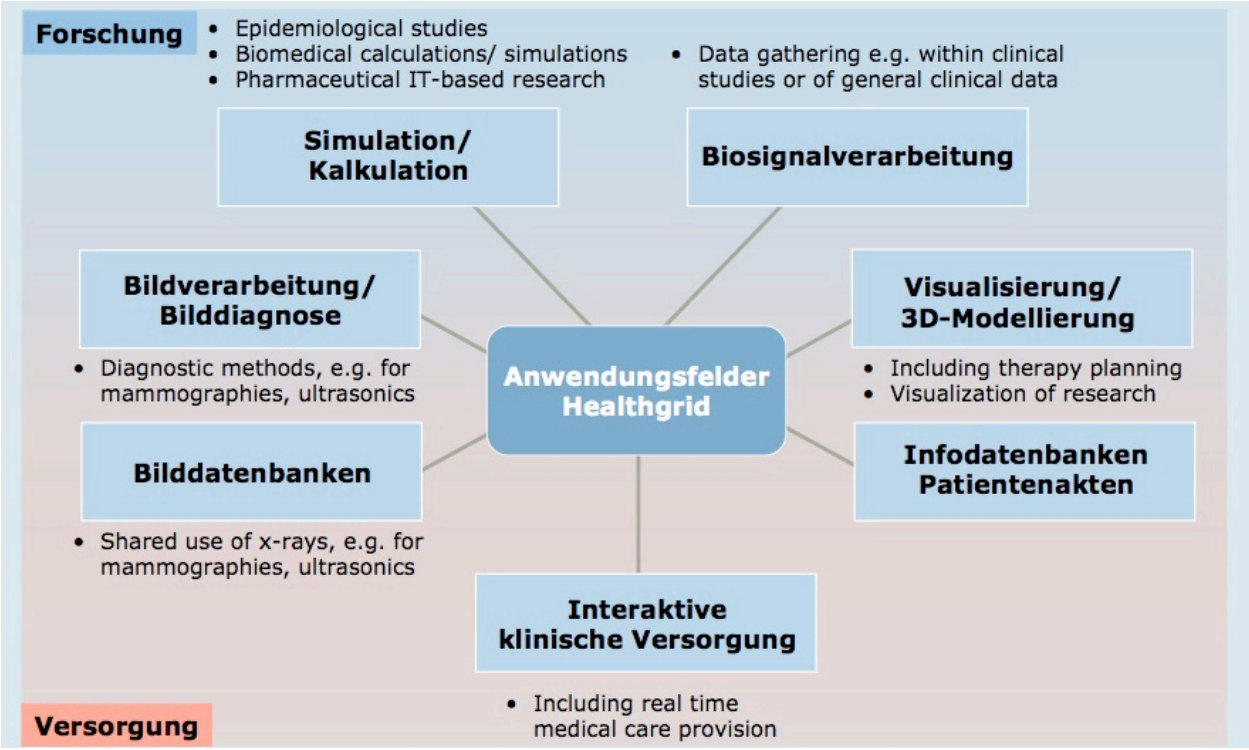


Abbildung 35. Anwendungsfelder Grid Computing

Grundsätzlich konzentrieren sich gerade neuere Projekte explizit auf die Entwicklung und Bereitstellung von Anwendungen zur *Integration und Vernetzung von Datenbeständen*.

Aus Sicht *medizinischer Krankheitsbilder* sind die im Rahmen des MediGRID-Projektes untersuchten Grid-Projekte insbesondere in der Onkologie und der Neurologie aktiv – hier insbesondere in den bildverarbeitenden Anwendungsbereichen mit ihrer ressourcenintensive Datenbasis, insbesondere von Mammographien und Gehirnbilddaten.

Grid-Computing-Projekte im Umfeld der Partner-Grids sind primär auf die medizinische/klinische oder biomedizinische Forschung ausgerichtet. Viele Projekte befassen sich mit biomedizinischer Grundlagenforschung, die eine entsprechende Kundengruppe anspricht. Ebenfalls eine große Anzahl an Projekten konzentrieren sich auf die medizinische und klinische Forschung und bieten Anwendungen, die perspektivisch Ärzte in Krankenhäusern und

niedergelassene Ärzte ansprechen sollen. Allerdings ist gerade in diesem Segment mit einer Durchdringung frühestens in einem längerfristigen Zeitraum zu rechnen (siehe Abbildung 36).

Potenzielle Anwendergruppen	Potenzial	Aufwand	Etablierung /Finanzierung 3-5 J.	5-10 J.	>10 J.
Forschung					
öffentlich			Förderung	Förderung*	Förderung*
kommerziell			Teilförderung	Teilförderung	Rechnung
Sekundäre Versorgung (Facharzt)					
öffentlich			X	Teilförderung	Kassen
kommerziell			X	Teilförderung	Rechnung
Primäre Versorgung (Allgemeinmedizin)					
öffentlich			X	X	Teilfö. Kas.
kommerziell			X	X	Teilfö. Re.

* Öffentliche Förderung verstärkt mit Wirtschaftlichkeitsanreizen (Kostenstellenrechnung, Bedarfsgerechte Kostenerstattung)

Abbildung 36. Marktdurchdringung Healthgrids

2.8.3.2 Erlösmodelle für Healthgrids

Ein wesentlicher Faktor für die Ausgestaltung von Geschäftsmodellen ist die Frage, mit welcher Methodik langfristig kostendeckende Erlöse erzielt werden. Hierzu dienen Erlösmodelle. Erlösmodelle lassen sich als Kombination aus Erlösquelle (d.h. die angebotene Leistung), Erlöspartner (i.d.R. der Kunde oder Leistungsempfänger) und eines sich daraus ableitenden Preisschemas beschreiben. Letztendlich bepreist ein Erlösmodell den zusätzlichen Nutzen, den die Erlösquelle für den Erlöspartner darstellt.

Für Healthgrids ergeben sich auf Ebene der letzten Wertschöpfungsstufe Preisschemata aus der Kombination der angebotenen Softwareanwendungsbereiche und der verschiedenen Endkundengruppen (Scholz et al. 2007). Im Rahmen der für MediGRID vorgenommenen Untersuchungen wurden vier verschiedene Preismechanismen (dynamische Bepreisung entsprechend der genutzten Leistung, ein fest vereinbarter Pauschalpreis, ein Preis-Stufen-Modell und eine wiederkehrende Gebühr (d.h. Mitgliedschaftsgebühr)) berücksichtigt und unterschiedlichen Kombinationen von Anwendungsfeldern und Anwendergruppen zugeordnet (siehe Abbildung 37).

Die Analyse zeigt, dass es kein fixes einheitliches Preisschema für Anwendungsbereiche oder Kundengruppen gibt, sondern Preisschemata immer spezifisch auf die zu erbringende Leistung und die jeweilige Kundengruppe abzustimmen sind. Kundengruppen, die häufig wiederkehrende Nutzungsmuster zeigen, werden eine regelmäßige wiederkehrende gleichmäßige Zahlung einer fallabhängigen Bepreisung vorziehen (z.B. Radiologen, die regelmäßig Bilddaten abfragen), während Nutzer mit einem unregelmäßigen Nutzungsprofil eine Fallpauschale vorziehen (z.B. Allgemeinmediziner, die eine einmalige Abfrage in einer Datenbank vornehmen). Anwendungsfelder, die einen im Vorfeld schwer kalkulierbaren Leistungsbedarf aufzeigen (z.B. Simulatio-

nen) sollten in der Regel mit einem der tatsächlichen Leistung entsprechenden Bepreisungsschema belegt werden.

	Industry	Researchers	Specialists	Physician
Simulations/calc.	Dynamic pricing	Dynamic pricing		
Visualization/ m.		Level pricing / membership	Level pricing / membership	Level pricing
Image processing	Level pricing	Level pricing / membership	Level pricing / membership	Level pricing
Image handling	Fixed pricing	Membership fee	Membership fee	Level pricing / membership
Data gathering/ h.	Fixed pricing	Membership fee	Membership fee	Fixed pricing
Databases			Membership fee	Fixed pricing
Medical care prov.			Level pricing / membership	

Abbildung 37. Erlösmodelle für Healthgrids (Scholz et al. 2007)

2.8.3.3 Erfolgsfaktoren für Healthgrids

Die Durchsetzung von Healthgrids am Markt beim Kunden ist von der Erfüllung und Berücksichtigung verschiedener Rahmenbedingungen des Marktes abhängig. Der Soll-Zustand der wesentlichen rahmengebenden Faktoren kann als Erfolgsfaktor bezeichnet werden.

Im Rahmen von MediGRID wurden internationale Experten befragt, welche Erfolgsfaktoren bei Healthgrids eine besondere Rolle spielen (Scholz et al. 2009). Im Verlauf der Untersuchung wurden drei Kategorien von Erfolgsfaktoren identifiziert, deren Bedeutung der hier wiedergegebenen Reihenfolge entspricht:

- Qualität und Leistung,
- Wirtschaftliche Faktoren,
- Rahmengebende Marktbedingungen.

In Bezug auf die *Qualität und die Leistung* müssen die angebotenen Produkte zwingend überzeugende Lösungen für die Datensicherheit und den Datenschutz beinhalten. Fast ebenso wichtig ist den Experten eine nutzerfreundliche Schnittstelle und nicht zuletzt eine stabile und vertrauenswürdige Infrastruktur.

Innerhalb der zweitwichtigsten Gruppe der wirtschaftlichen Faktoren sahen die Experten stand die Bereitstellung einer lauffähigen in ein überzeugendes Betreibermodell eingebetete „Killer-Application“ ganz oben auf der Prioritätenliste. Weitere wichtige Erfolgsfaktoren sind belastbare Erlös- und Kostenmodelle.

Die rahmengebenden Marktbedingungen konzentrieren sich auf die Forderung nach marktgerechten, der technologischen Entwicklung entsprechenden Rechtsgebungen, insbesondere unter Berücksichtigung der länderübergreifenden Zusammenarbeit vieler Partner-Grids und der jeweils

unterschiedlichen Rechtsgebung in Hinblick auf Datenschutz und Haftungsfragen. Ein weiterer wichtiger Erfolgsfaktor ist aus Sicht der internationalen Experten die optimale Adaption der Lösungen auf die spezifische Kultur im Gesundheitssektor, insbesondere die Zeitknappheit und geringe IuK-Affinität von medizinischem Personal.

2.8.3.4 Entwicklungspotenziale auf Basis der MediGRID-Infrastruktur

Neben der Beleuchtung des Healthgrid-Marktes im Allgemeinen wurde MediGRID auf eine nachhaltige Entwicklung hin untersucht. Entsprechend der oben beschriebenen Gesamtmarktsituation kann MediGRID als Basisprojekt zum Aufbau einer Grundinfrastruktur und bezüglich der zum Einsatz gekommenen Anwendungen als Proof-of-Concept-Projekt eingestuft werden. MediGRID ordnet sich damit in den Gesamtmarkt ein. Der Entwicklungsstand des Projektes entspricht dem international typischen Stand für Partner-Grids.

MediGRID verfolgt perspektivisch das Ziel der Etablierung und des störungsfreien, nachhaltigen Betriebs der wichtigsten Grid-Infrastruktur für die Medizin und die Lebenswissenschaften in Deutschland.

Um dauerhaft eine solche nachhaltige Struktur aufzubauen, wäre es erforderlich, dass MediGRID einen dreistufigen Entwicklungsprozess durchläuft (siehe Abbildung 38).

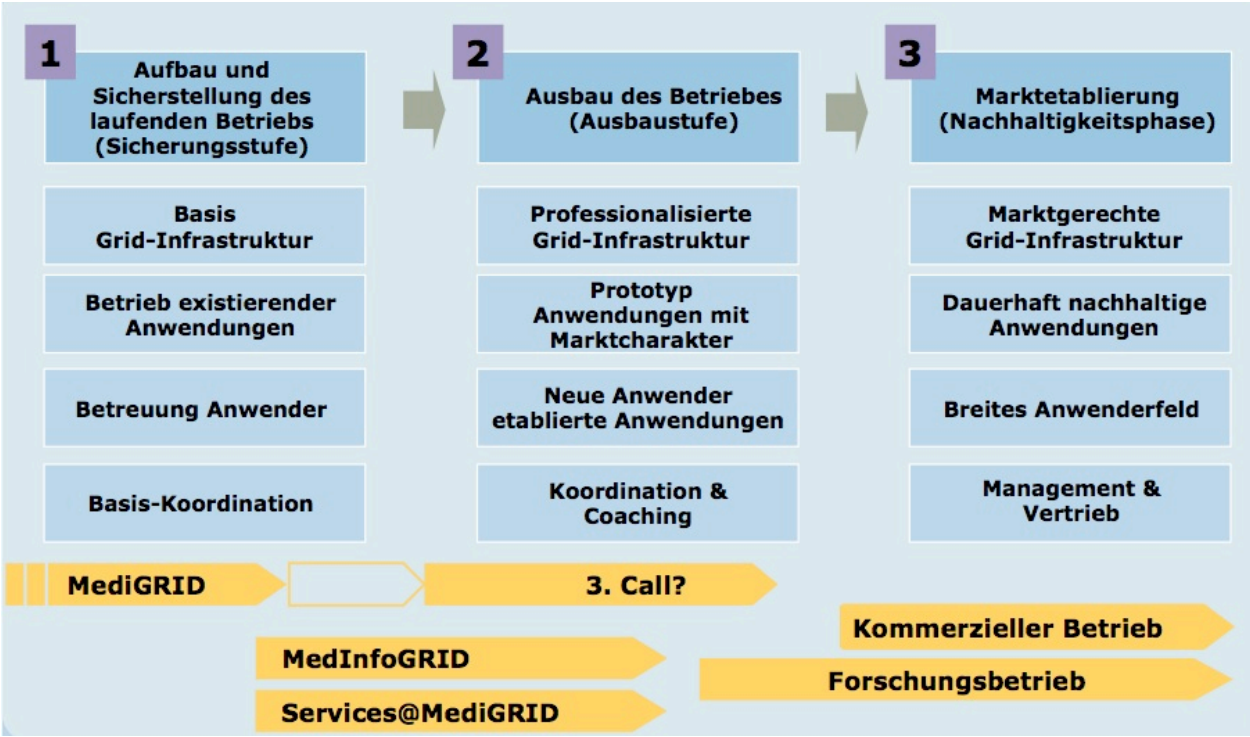


Abbildung 38. Idealtypischer Entwicklungsprozess von MediGRID

Prinzipiell wäre innerhalb dieses Prozesses die Entwicklung von drei Kernbereichen voranzutreiben: Sicherung und Ausbau der Basis-Grid-Infrastruktur, Sicherstellung und Ausbau eines nutzerorientierten Anwendungsportfolios und Sicherstellung und Ausbau des Wissentransfers über verschiedene Entwicklungsstufen hinweg innerhalb des Projektes und nach Außen Coaching potenzieller Dienstleister und Anwender.

Als Ergebnis stünde für MediGRID ein Geschäftsmodell als Gesamtszenario zur Diskussion, dass die Etablierung einer dezentralen, flexiblen Organisation mit One-Stop-Agency für Kunden

(Entwickler und Anwender) als zentrale Anlauf- und Koordinierungsstelle als Marktplatz für Angebot und Nachfrage zum Ziel hat.

Eine mögliche Gesamtstruktur unter Berücksichtigung verschiedener Ausbaustufen zeigt Abbildung 39.

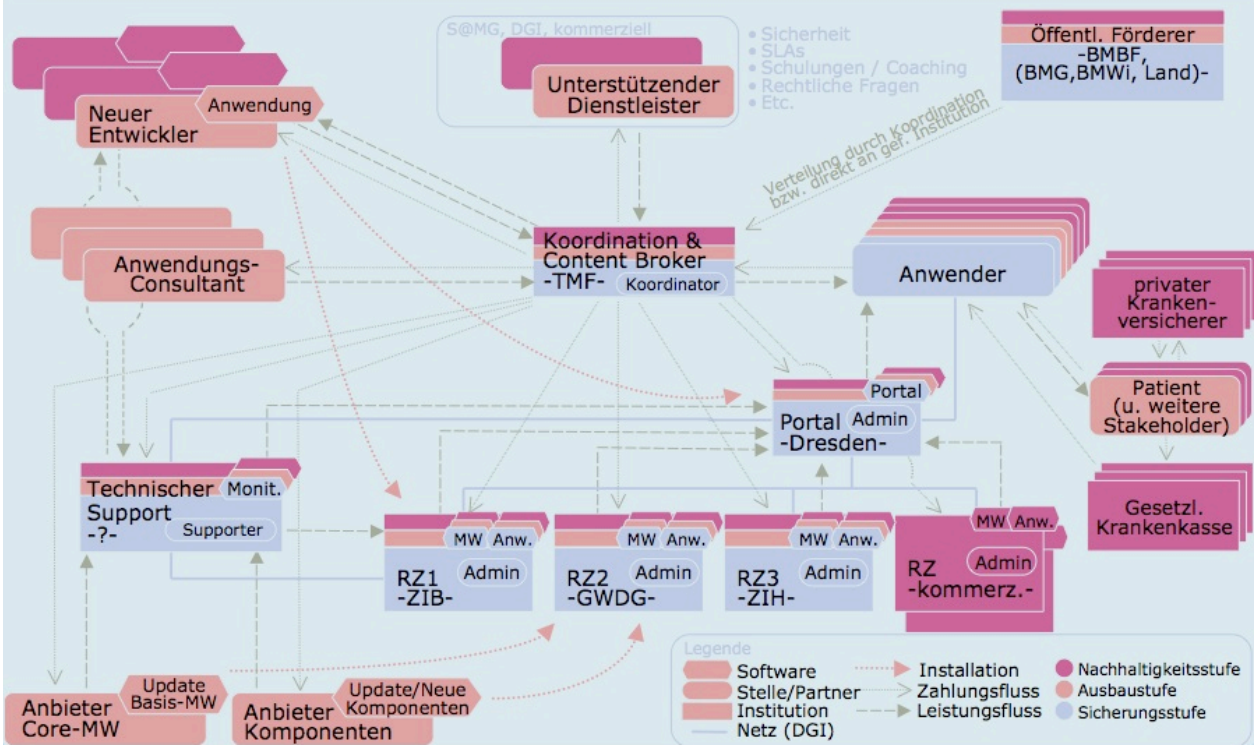


Abbildung 39. Perspektive einer integrierten Grid-Infrastruktur für MediGRID

2.8.3.5 Sonstige Ergebnisse und Aktivitäten

Ergebnisse der Geschäftsmodellbetrachtungen zum Healthgrid-Markt und zu MediGRID wurden regelmäßig auf Veranstaltungen innerhalb der D-Grid-Initiative (z.B. Sugi-Workshop 2008, DGI-Nachhaltigkeitstreffen) und darüber hinaus kommuniziert (Mednet 2007, MKWI 2008, Healthgrid 2008 und 2009; siehe Publikationen und Vorträge im Anhang). Dies schließt verschiedene Beiträge zu Konferenztagungsbänden ein.

Die Ergebnisse zu Geschäftsmodellen und zur Nachhaltigkeit waren insbesondere das Ergebnis einer Zusammenarbeit mit dem Institut für Wirtschaftsinformatik von Prof. Michael H. Breitner der Leibniz-Universität Hannover im Rahmen einer Promotion zum Thema „Geschäftsmodelle für Grid Computing in der Medizin und den Lebenswissenschaften“ (S. Scholz, TMF, Berlin).

In die Marktuntersuchungen flossen Ergebnisse einer von der TMF e.V. in einer frühen Projektphase in Auftrag gegebenen Studie zur Untersuchung des aktuellen Marktes für Healthgrid-Projekte ein. Durchgeführt wurde diese Studie von der Unternehmensberatung empirica, Bonn.

Zur Verbesserung der Vernetzung der Beteiligten der einzelnen Wertschöpfungsstufen des Healthgrid-Marktes hat die TMF e.V. im Jahr 2008 die Netzwerkplattform ForumGrid gegründet (siehe Öffentlichkeitsarbeit).

2.8.4 Nutzen und Verwertbarkeit der Modulergebnisse

Erkenntnisse aus den wirtschaftlichen Betrachtungen sind konkret in die Projektanträge Services@MediGRID, MedInfoGrid, WissGrid und PneumoGRID eingeflossen. Das in dem Projekt PneumoGRID zum Einsatz kommende Betreibermodell wurde anhand im MediGRID-Projekt hergeleiteten Erkenntnisse aufgesetzt und beurteilt.

2.8.5 Fortschritte auf dem Markt für Grid Computing

Während Geschäftsmodelle für elektronische Märkte bezogen auf das Internet intensiv von der wissenschaftlichen Literatur beleuchtet wurden, ist eine ähnlich umfangreiche Beschreibung von Geschäftsmodellen für Grid Computing nicht verfügbar (vgl. Veit und Gentzsch 2008, S. 215). Zum Teil wurden insbesondere in der Europäischen Union öffentlich geförderte Projekte ins Leben gerufen, gezielt ökonomische Aspekte von Grid-Computing-Infrastrukturen zu beleuchten, z.B. GridEcon und BIG(vgl. Schikuta et al. 2005; Altmann et al. 2007a). Die wenigen Arbeiten, die ökonomisch relevante Aspekte beleuchten sind zu einem signifikanten Teil im Rahmen dieser Projekte entstanden.

Einen signifikanten Mehrwert leisten insbesondere jüngere Arbeiten zur Wertschöpfungsstruktur in Grid-Infrastrukturen (vgl. u.a. Forge und Blackmann 2006; Altmann et al. 2007b; Stanoevska-Slabeva et al. 2007).

Neben der ökonomisch-theoretischen Analyse von Grid Computing Geschäftsmodellen auf der einen Seite verfolgt ein paralleler Forschungszweig die Modellierung und Entwicklung geschäftsunterstützender Services (im Sinne von Software-Lösungen) zur Abbildung und Unterstützung ökonomischer Fragestellungen im Grid. Übergreifend wird dabei die Frage beantwortet, welche Services im Rahmen einer marktorientierten Grid Infrastruktur notwendig sind (vgl. Altmann und Routzounis 2006). Im Zentrum der Untersuchungen der vergangen Jahren standen die Frage der effizienten Allokation von Ressourcen und die und damit zusammenhängende nutzungsadäquate Bepreisung und Verrechnung der in einem Grid geteilten Ressourcen (vgl. Stuer et al. 2007). Arbeiten dieses Forschungszweiges werden allgemein unter dem Oberbegriff *Grid Economics* zusammengefasst (vgl. Neumann et al. 2006). Dabei werden unterschiedliche Ansätze verfolgt (vgl. Buyya et al. 2005; Eymann et al. 2005; Yeo und Buyya 2006).

2.8.6 Ausblick und Perspektiven von Healthgrids

Firmeninternen Grids haben sich bereits im Gesundheitssektor etabliert. Partner-Grids stehen hier noch zurück, gewinnen jedoch aufgrund einer zunehmenden Vernetzung und Kollaboration insbesondere im Forschungsumfeld eine immer größere Rolle.

Häufig sind die momentanen Healthgrid-Projekte jedoch nicht langfristig ausgelegt, sondern verfolgen ein singuläres Projektziel ohne übergreifende langfristige strategische Zielsetzung. Diese Definition und Anpassung der langfristigen Zielsetzung ist für Healthgrids, und im konkreten Fall für MediGRID, aber unbedingt erforderlich.

Um dieses Ziel zu erreichen, sind aus öffentlicher Sicht zwei Dinge erforderlich. Zum einen müssen geförderte Einzelprojekte viel stärker als bisher in einer übergreifenden langfristigen Initiative integriert sein und Projekten eine stetige Entwicklung über einen drei Jahreszeitraum hinaus ermöglicht werden. Die derzeitige Planung auf Dreijahresbasis ist aus betriebswirtschaftlicher Sicht ungenügend. Zum anderen sind akademische Einrichtungen, die mehrheitlich an öffentlichen Projekten beteiligt sind, insbesondere in Bezug auf die Darstellung indirekter Kosten transparenter zu gestalten. Dies ermöglicht transparentere Verrechnung von Kosten innerhalb des Projektes und der Abrechnung mit dritten Partnern als Basis für eine wirtschaftlich nachhaltige Lösung.

Auf der Seite ist potenziellen akademischen Nutzern, die gerade in der Etablierungsphase primär das Kundenportfolio stellen, die Möglichkeit zu geben, öffentliche Fördergelder stärker als bisher in Verbrauchsmittel zu investieren. Der Einsatz einer dezentralen vernetzten IT-Infrastruktur, wie Grid-Computing, basiert auf der verbrauchsbezogenen Abrechnung von IuK-Leistungen.

2.8.6.1 Verwendete Literatur

Altmann, J., et al. 2007a. GridEcon – The Economic-Enhanced Next-Generation Internet. Veit, D., et al., eds. GECON 2007, LNCS 4685. Springer, Heidelberg, 188-193.

Altmann, J., et al. 2007b. Taxonomy of Grid Business Models. Veit, D., et al., eds. GECON 2007, LNCS 4685. Springer, Heidelberg, 29-43.

Altmann, J., Routzounis, S. 2006. Economic Modeling of Grid Services, eChallenges 2006. Online: http://www.gridecon.eu/downloads/Economic_Modeling.pdf, Stand: 23.03.2009.

Andoulsi, I., et al. 2008. The SHARE Road Map: Healthgrids for Biomedical Research and Healthcare. Studies in health technology and informatics 138 238-278.

Breton, V., et al. 2005. The HealthGrid White Paper, Proceedings of Third HealthGrid Conference. Studies in health technology and informatics 112 249-321.

Buyya, R., et al. 2005. The Grid Economy. Proceedings of the IEEE 93(3) 698-714.

Buyya, R., Venugopal, S. 2005. A Gentle Introduction to Grid Computing and Technologies. Computer Science of India 9.

Cohen, R. 2004. The Impact of Grid Computing: The Results of the ESI Grid Study. GT'04 VIP Summit. Online: http://www.rhsmith.umd.edu/dingman/events/summaries/presentations/Cohen2005.ppt, Stand: 05.11.2008.

Eymann, T., et al. 2005. Catallaxy-based grid markets. Multiagent Grid Systems 1(4) 297– 307.

Forge, S., Blackmann, C. 2006. Commercial Exploitation of Grid Technologies and Services - Drivers and Barriers, Business Models and Impact of Using Free and Open Source Licensing Schemes. SCF Associates Ltd. for European Commission / DG Information Society and Media, ed.

Insight Corp. 2005. Grid Computing - A Vertical Market Perspective 2005-2010 (Executive Summary). The INSIGHT Research Corporation, Boonton, New Jersey, USA.

Insight Corp. 2006. Grid Computing - A Vertical Market Perspective 2006-2011 (Executive Summary). The INSIGHT Research Corporation, Boonton, New Jersey, USA.

Longbottom, C., Tarzey, B. 2006. Grid Computing Update - Pilots move to Mainstream. Quocirca Inc., ed. Quocirca Insight Report.

Neumann, D., et al. 2006. Grid-Economics. Wirtschaftsinformatik 48(3) 206-209.

Osterwalder, A., et al. 2005. Clarifying Business Models: Origins, Present and Future of the Concept. Communications of the Association for Information Systems (CAIS) 16(1) 1–25.

Plaszczak, P., Wellner Jr., R. 2006. Grid Computing: The Savvy Manager's Guide. Elsevier, Amsterdam.

Scheer, C., et al. 2003. Geschäftsmodelle und internetbasierte Geschäftsmodelle-Begriffsbestimmung und Teilnehmermodell. ISYM Working Paper. Online unter: http://isym.bwl.uni-mainz.de/publikationen/isym012.pdf, last reviewed 23.09.07.

Schikuta, E., et al. 2005. Business In the Grid: Project Results, 1st Austrian grid symposium. OCG Verlag, Hagenberg. Online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.90.1176&rep=rep1&type=pdf, Stand: 03.04.2007.

Scholz, S. 2008. Geschäftsmodellansätze und Wertschöpfung in Healthgrid-Infrastrukturen - Präsentation. Institut für Wirtschaftsinformatik (IWI), Leibniz-Universität Hannover.

Scholz, S. 2009. Forschungsseminar 2009: Geschäftsmodelle für Grid Computing in der Medizin und den Life Sciences - Präsentation. Institut für Wirtschaftsinformatik (IWI), Leibniz-Universität Hannover.

Scholz, S., et al. 2008. A Sustainable Business Model Approach for Grid Computing – And a Life Sciences Example. Bichler, M., et al., eds. Multikonferenz Wirtschaftsinformatik 2008. GITO-Verlag, Berlin, Volltext (auf CD), S. 1-12.

Scholz, S., et al. 2007. Business Models for Grid Computing in Life Science: An Approach to Identifying Health Grid Revenue, Mednet 2007, Leipzig. Online: http://www.mednet2007.com/pdfs/fullpaper/0708MednetProceedingTMFUniHan.pdf, Stand: 4.11.2008.

Scholz S, Semler SC, Breitner MH (2009). Business Aspects and Sustainability for Healthgrids – an Expert Survey. Healthgrid 2009. Studies in Health Technology and Informatics 147, 163-172.

Stanoevska-Slabeva, K., et al. 2007. Development of a Generic Value Chain for the Grid Industry Grid Economics and Business Models. Springer, Berlin, Heidelberg, 29-43.

Stuer, G., et al. 2007. A commodity market algorithm for pricing substitutable Grid resources. Future Generation Computer Systems 23 688–701.

Szolovits, P. 2007. What is a grid? J Am Med Inform Assoc 14(3) 386.

Timmers, P. 1998. Business Models for Electronic Markets. Electronic Markets (EM) 8(2) 3-8.

Veit, D., Gentzsch, W. 2008. Grid Economics and Business Models. Journal of Grid Computing 6(3) 215-217.

Yeo, C.S., Buyya, R. 2006. A Taxonomy of Market-Based Resource Management Systems for Utiltiy-driven Cluster Computing. Software Practice and Experience 36(13) 1381-1419.

3 Weitere Aktivitäten im Verbundprojekt

Neben den wissenschaftlich-technischen Modulen war ein Modul zuständig für die administrative und fachliche Gesamtkoordination des Vorhabens, die unterstützende Öffentlichkeitsarbeit sowie die nationale und internationale Vernetzung innerhalb der Community.

3.1 Administrative Koordination

Die administrative und fachliche Koordinierung der acht Konsortialpartner wurde von der Telematikplattform für medizinische Forschungsnetze - TMF e.V. übernommen. Hierzu zählten das übergreifende Berichtswesen, das Projektmanagement und die fachliche Koordination des Gesamtprojektes.

Beim Start des Projektes übernahm die TMF in der Projektinitiierungsphase die Ausgestaltung und Abstimmung des MediGRID-Konsortialvertrages (Deliverable 1).

Zur Koordination gehörte die Organisation und Durchführung von projektbezogenen Sitzungen und Konferenzen. Insbesondere fielen hierunter:

- Vorstandssitzungen des Konsortiums (alle Modulleiter zzgl. Sprecher, 4 x p.a.)
- Web- und Telefonkonferenzen (ergänzende regelmäßige Projektabstimmung)
- MediGRID All-Hands-Meeting (07/2007, Berlin)

Zur Sicherstellung einer gezielten projektinternen Kommunikation verwaltete und koordinierte die TMF gruppenbezogene E-Mail-Verteiler, u.a. für die Modulleiter und Projekt-Entwickler.

Zusätzlich koordinierte die TMF das Berichtswesen des Konsortiums. Hierzu zählte die Erstellung der Statusberichte und des Abschlussberichtes (Deliverables 2-4). Die Projektaktivitäten wurden konkret über folgende Dokumentationen sichergestellt:

- D-Grid-Monatsberichte (09/2005 – 08/2008, Koordination und Erstellung durch TMF)
- Halbjährliche Statusberichte (12/2005 – 12/2007, Erstellung durch TMF)
- Halbjährliche Zwischenberichte (12/2005 – 12/2008, Koordination TMF)
- Erweiterte Berichtspflichten zur Zwischenbegutachtung (03/2007, Koordination TMF)
- Abschlussbericht (1. Hj. 2009, Koordination und Gesamterstellung durch TMF)

Alle projektrelevanten Dokumente des MediGRID Projektes wurden auf einem projektinternen Sharepoint-Server abgelegt. Neben Protokollen, Berichten und Präsentationen wurde ein Kalender mit allen MediGRID internen und externen Veranstaltungen sowie nationale und internationale Veranstaltungen bereitgestellt, die für MediGRID Mitarbeiter interessant sind. Im Juli 2007 wurde ein separater Bereich für die Gutachter von MediGRID angelegt, in dem regelmäßig die Modul- und Statusberichte sowie der Status der Deliverables abgelegt wurden.

3.2 Öffentlichkeitsarbeit

Die Öffentlichkeitsarbeit des Projektes war Aufgabe der TMF e.V.. Zu den Kernelementen der öffentlichkeitswirksamen Maßnahmen des Projektes zählten der Aufbau einer Projekt-Website, die Bereitstellung von Werbe- und Informationsmitteln, die Durchführung von Community-Workshops, die Platzierung von projektbezogenen Artikeln in der Fachpresse, die Präsenz auf Veranstaltungen und Messen sowie die Vernetzung mit internationalen Einrichtungen.

Bereits zu Projektbeginn war MediGRID mit einer eigenen Internetpräsenz vertreten (www.medigrid.de), die über die Projektlaufzeit ständig aktualisiert wurde. Zum Projektab-

schluss wurde der Internetauftritt noch einmal überarbeitet und wichtige Projektdokumente für die weitere Nutzung bereit gestellt.

Ebenfalls in einer frühen Projektphase erstellte die TMF einen Projektflyer in Form eines doppelseitig bedruckten Faltblattes in englischer und deutscher Sprache zur Verbreitung auf Konferenzen und Tagungen zur Verfügung (siehe Anhang).

Gemeinsam mit der Pressestelle der TMF e.V. wurden in der Projektlaufzeit verschiedene Artikel in der Fachpresse publiziert (u.a. in den Gesundheitsnachrichten, eHealh.com) und Presseanfragen zu den Aktivitäten von MediGRID betreut.

Zur stärkeren Verbreitung des Grid-Ansatzes in der Fachcommunity wurden zwei Industrieworkshops durchgeführt - bereits im Startjahr des Projektes ein Workshop für Ressourcenprovider in Berlin und im Jahr 2006 ein weiterer Industrieworkshop im Rahmen der Industriemesse ITeG (Deliverable Workshop).

MediGRID wurde zusätzlich durch die Präsenz auf verschiedenen Fachtagungen und Messen durch die TMF vertreten. Neben der Beteiligung und Präsenz auf Veranstaltungen von D-Grid teilweise mit Präsentationen (D-Grid All-Hands-Meetings in Göttingen, Nachhaltigkeitsworkshops in Karlsruhe, Berlin und Hannover (MediGrid-Koordination und –Vorbereitung) präsentierte die TMF das MediGRID-Projekt auch gegenüber anderen Grid-Communities (u.a. Workshop „Gridbasierte Wertschöpfung und Nachhaltigkeit“ im Rahmen des SuGI-Projektes, Demonstration auf der HealthGrid 2008 in Genf) und Fachpublikum außerhalb der Grid-Community (u.a. MEDICA-Fachmesse).

Die Vernetzung und Kooperation der Healthgrid-Community im europäischen und internationalen Rahmen ist ein weiteres Element zur nachhaltigen Bekanntmachung der Healthgridaktivitäten im deutschen Raum. So wurde die TMF während der Projektlaufzeit aktives Mitglied von HealthGrid (www.healthgrid.org), der führenden europäischen Vereinigung für Grid Computing im Bereich Medizin und Lebenswissenschaften sowie Veranstalter der gleichnamigen internationalen Tagung. Für das 2009 ist es der TMF gelungen, die Konferenz als begleitender Koordinator in Berlin auszurichten.

Aufgrund gemeinsamer Fragestellungen besteht mit dem Grid-Infrastrukturprojekt der amerikanischen Krebsforschungszentren caBIG (cabig.cancer.gov) ein intensiver Austausch, der in zwei Workshops gefestigt wurde. Der erste gemeinsame Workshop fand bereits im November 2007 während der AMIA-Konferenz in Washington statt, der zweite Workshop wurde im Januar 2008 in Berlin abgehalten.

Die internationale Vernetzung wurde zusätzlich durch die Initiierung des TMF „Forum Grid“ auf eine weitere Säule gestellt. Das Forum dient der Zusammenarbeit der deutschen Grid-Projekte in der Medizin und den Lebenswissenschaften und der Vertiefung nationaler und internationaler Kooperationen. Die Auftaktsitzung fand im Juli 2008 in Berlin statt.

3.3 Nationale und internationale Vernetzung

Im Rahmen von Anwenderworkshops und Koordinierungstreffen fand eine enge Abstimmung mit den Communities und dem Integrationsprojekt in D-Grid, sowie der eScience-Initiative statt. Aktuelle Ergebnisse waren Gegenstand der Beiträge und Demonstrationen bei besuchten Veranstaltungen und bildeten die Grundlagen des schnellen und übergreifenden Informationsaustausches zwischen den Communities. Der zusätzliche Bedarf an Abstimmung wurde über interaktive Konferenzen abgedeckt. Als Vorbereitung für den 2. und 3.Call im D-Grid fanden massive Abstimmungen innerhalb von MediGRID und weiteren D-Grid- und Nicht-D-Grid Communites statt. Das Modul eScience war und ist mit den D-Grid I -und auch D-Grid II-Projekten in allen Bereichen gut abgestimmt. Das schlägt sich in der intensiven (assoziierten) Mitarbeit in den neuen Communities nieder. Die Zusammenarbeit mit dem DGI auf dem Gebiet

von VO-Management, Autorisierung/ Authentifizierung und der Bildung nachhaltiger Grid-Infrastrukturen für laufende und zukünftige Grid-Vorhaben in Deutschland ist weiter eng abgestimmt und erfolgreich.

Als erstes Ergebnis hat das Modul e-Science mit HealthGrid-Partnern und den methodischen Modulen in MediGRID einen Leitfaden veröffentlicht. Unter dem Titel „Grid-Computing in der Biomedizinischen Forschung - Datenschutz und Datensicherheit“ eruieren die beteiligten Autoren Anforderungen, Aspekte, Realisierungs-Möglichkeiten bzw. -Schwierigkeiten, Visionen und zukünftige Arbeiten von Grid-Umgebungen für die Community Biomedizin.

In einer gemeinsamen Initiative von HealthGrid-Partnern und dem Modul eScience ist eine Präsentation von Lösungen aus verschiedenen Projekten und ein Diskussionspanel auf der e-Health-week 2007 organisiert worden. Dort sind die verschiedenen Anforderungen, Ansätze, Methoden und Standardisierungsinitiativen vorgestellt worden. In einem Round-Table-Gespräch haben die Fachleute erste Phasen skizziert und eine Time-Table aufgezeichnet.

Die intensive Arbeit an Sicherheitslösungen und Konzepten im Grid-Umfeld mit medizinischem Background führte zu Einladungen für nationale, europäische und internationale Veranstaltungen (GMDS, CeHR, caBIG, HealthGrid, EGEE und MedInfo). Das Fachpublikum wurde hierbei über die Machbarkeit der Umsetzung von medizinischen Applikationen im Grid informiert, wobei Data Protection und Data Security kein Hinderungsgrund sein müssen. Die Methodiken und Ansätze sind auf die jeweilige Plattform portierbar, Beispiel dafür ist die Integration entwickelter De-Identifizierungsalgorithmen als Grid-Service.

Das in MediGRID verwendete Informations-Portal zur kollaborativen Zusammenarbeit dient neben einem Wiki zum modulübergreifenden Informationsaustausch und wird zur Kommunikation des MediGRID-Projektes und weiterer Projektvorhaben gepflegt und zur Verfügung gestellt. Während und nach der Projektlaufzeit sind die Beteiligten in der biomedizinischen Grid-Community, in der D-Grid und internationalen Grid-Comnmunity weiter gut abgestimmt, ein Zeichen dafür sind die Arbeiten im Projekt services@MediGRID, MedInfoGrid und den Beteiligungen an Projekten der dritten Phase (PneumoGrid und dem gemeinsamen Antrag der D-Grid Projekte).

Im Berichtszeitraum erfolgte eine intensive Kooperation und Abstimmung mit TextGrid und Instant-Grid. Die Vorteile liegen in der Standortnähe dieser Initiative und die Möglichkeit der gemeinsamen Koordinierung. Am Standort Göttingen wurde eine lokale Grid-Community (GoeGrid-Runde) aufgebaut, in der neben MediGRID, TextGrid und InstantGrid (GWDG) beteiligt sind. Die gemeinsamen Zusammenkünfte in der GoeGrid-Runde haben zu einer intensiveren Verzahnung der gemeinsamen Vorhaben und Resultate geführt. Die Communities führten ein gut besuchtes Grid-Symposium durch. Die Informationen fließen über dieses Gremium bidirektional zu den Communities und dem DGI. 2007 wurde GoeGrid durch die Community HEP erweitert. GoeGrid co-organisierte das erste D-Grid-All-Hands-Meeting mit Vertretern der alten und neuen Grid-Communities. Dort eingebettet war das 2. Grid Symposium. Dort partizipierten besonders die neuen Communities von den gesammelten Erfahrungen.

Weiter kann auf eine gute Zusammenarbeit während der Beantragung von Sonderinvestitionsmitteln verwiesen werden. Alle Vertreter waren sich über den Ausbau vorhandener Ressourcen und die Einrichtung eines Grid-Ressourcen-Zentrums einig und machten sogleich Überlegungen zur nachhaltigen Nutzung und Betrieb. Die GoeGrid-Community bildete im Sommer ein executive-Board und ein technical-Board, die letzten Punkt übergreifend begleiten.

Weiterhin konnte im Wintersemester 2007/2008 zusammen mit TextGrid, dem II. Physikalischen Institut (HEP), der GWDG als Grid-Ressourcenzentrum und dem Zentrum für Informatik der Universität Göttingen ein interdisziplinäres Grid-Seminar aufgelegt werden, das fakultäts-

übergreifend großes Interesse bei Dozenten und Studierenden fand. Diese Erfolgsgeschichte schrieb sich auch im gemeinsamen Grid-Praktikum im Sommersemester 2008 weiter.

Durch Teilnahme an der Jahrestagung der American Medical Informatics Association (AMIA) 2005 in Washington kam es zu einer erfolgreichen Abstimmung mit US Medizin-Grids. Weiterhin wurde der Kontakt zu den assoziierten Partnern in den USA durch ein Arbeitstreffen in Boston nach der AMIA-Tagung intensiviert. Diese Informationen und Kontakte waren Grundlage für den MediGRID-US-Workshops 2006 in Boston.

Das rege Interesse auf der GMDS-Jahrestagung in Leipzig, der EGEE-Konferenz in Genf, dem Grid-Symposium in Göttingen, sowie der USA-Workshop in Washington am Rande der AMIA-Jahrestagung und die Präsentation von D-Grid auf der Grid@Asia-Konferenz in Korea zeigten, dass eine Durchdringung der Community erreicht wurde. Auf der HealthGrid-Konferenz 2007 in Genf war das eScience-Modul mit einem Beitrag auf Einladung der europäischen Community vertreten. Dort wurde die enge Verknüpfung von MediGRID mit europäischen und internationalen Partnern deutlich. In der German eScience Conference wurden die sicherheitsrelevanten Ergebnisse am Beispiel einer Anwendung der Bildverarbeitung vorgestellt. Zudem hat das Modul eScience maßgeblich am Accounting und Billing Workshop im Zusammenhang mit DGI-Aktivitäten beigetragen (Weiterführung in services@MediGRID, MedInfoGrid). Bei der Medinfo 2007 in Brisbane (Australien) wurde der Vortrag aus MediGRID sehr aufmerksam verfolgt und war auf Einladung mit einem Besuch des Gridlabs in Melbourne und anschließenden Studienaufenthalt verbunden.

Zu einem Treffen zur Vertiefung der Zusammenarbeit mit caBIG kam es im Rahmen eines caBIG/ MediGRID-Meeting 2008 in Berlin. Beide Seiten erklärten die gemeinsame Zielsetzung von einander zu partizipieren und eine internationale biomedizinische Community mit Infrastrukturen und Applikationen zu versorgen. Die Ergebnisse sind über das Forum Gird der TMF vorgestellt worden. Eine weitere Demonstration der Ergebnisse aus MediGRID und die Fortschritte in services@MediGRID auf internationaler Ebene war die Session bei der HealthGrid 2008 in Chicago. Dort wurden der aktuelle Stand und die Verzahnung mit den amerikanischen Grid-Aktivitäten präsentiert.

4 Schlussfolgerungen und Ausblick

Die Arbeit des Projektes MediGRID steht zwischen zwei wichtigen Abschnitten der deutschen Förderung von Forschungsvorhaben mit großem Bedarf an IT-Infrastruktur:

1. Der Motivationsphase 2002-2005, in der eine multidisziplinäre Arbeitsgruppe von Wissenschaftlern BMBF davon überzeugen konnte, den bereits fortgeschrittenen Förderprogrammen der USA, der EU sowie verschiedener EU-Mitgliedsstaaten ein deutsches Programm folgen zu lassen.
2. Der Reifungsphase seit 2007, in der die Ergebnisse der Projekte der ersten Förderphase in nachhaltige Infrastrukturlösungen fortentwickelt werden sollen.

Das Projekt durchlief – wie alle Vorhaben der ersten Runde und auch alle nachfolgenden – einen aufwändigen Begutachtungs- und Verhandlungsprozess, der zeigte, wie sehr Antragsteller, Gutachter und Förderer sowohl bezüglich der Zielerwartungen wie des methodischen Vorgehens auseinander lagen. Diese Diskrepanzen belasteten die Arbeiten genauso wie die offenen Dispute zwischen den Rechenzentrumsbetreibern und den Nutzercommunities.

Im Falle von MediGRID kam noch hinzu, dass die Community außerordentlich vielgestaltig ist und den meisten Gutachtern nicht vertraut war. Die aus dem Gesundheitsforschungsreferat des BMBF seit 1999 aufgebaute TMF, die MediGRID Projektkoordinatorin, war den Gutachtern nicht bekannt. Dies führte zu erheblichen Fehleinschätzungen.

In diesem Kontext kann die ausgleichende und hochkompetente Leistung des Projektträgers nicht hoch genug eingeschätzt werden. Unser Dank gilt insbesondere den Projektbetreuern Dr. Torsten Asselmeyer-Maluga, Dr. Dirk Franke, Dr. Rüdiger Krahl, Matthias Kuom und Ramona Schulz.

Trotz dieser Rahmenprobleme kann das Projekt als sehr erfolgreich eingestuft werden. Dies liegt daran, dass es gelang, eine freundschaftlich produktive Arbeitsstimmung aufzubauen, in der gemeinsames lernen und verstehen kultiviert werden konnte. Die Verantwortlichkeit und der Nutzen des Projektes wurde primär bei den Teilprojekten belassen. Die Konsortialführung blieb eher im Hintergrund.

Es gelang in allen Anwendungsszenarien beispielhafte Forschungsprozesse auf Basis einer gemeinsamen Grid-Middleware, die über ein „MediGRID-Portal“ zugreifbar ist, zu realisieren. Auch die Themen Sicherheit und Ontologie konnten erfolgreich bearbeitet werden. Die internationale Kollaboration konnte eingeleitet werden. Insofern ist MediGRID ein großer Fortschritt seit den Tagen der Motivationsphase gelungen. Allerdings zeigten die Arbeiten auch, wie weit die Nutzung der Grid-Technologie noch von einer verbreiteten Nutzung in Medizin und Lebenswissenschaften entfernt ist:

1. Die Services sind immer noch zu schwierig zu erreichen, benötigen zu viel Spezialkenntnisse und sind zu instabil, um von den Forschern angenommen zu werden.
2. Die Sicherheit der Grid-Services liegt deutlich unter den dem heute üblichen Level der wissenschaftlichen Rechenzentren und unakzeptabel tief im Hinblick auf verwertungsrechtlich hoch brisante Forschung.
3. Die methodischen Einrichtungen der Community verfügen in der Regel noch nicht über das Know How, ihren Bedarf an IT-Kapazität aus ihren Forschungen gegen die Angebote verschiedener Rechenzentren zu skalieren.
4. In der Community fehlt das Know How, geeignete Rechenvorgänge für die Grid-Technologie zu erkennen und anzupassen.

5. Die nachhaltige Finanzierung aller Fragen in Zusammenhang mit der Nutzung der Grid-Technologie ist völlig offen.
6. Die Defizite der deutschen medizinischen Forschung im Bereich der Ontologien wirkt sich auch bei der Nutzung von Grids aus.

Aus den genannten Punkten ergibt sich, dass noch erhebliche Anstrengungen notwendig sind – hier erweisen sich die nachfolgenden Förderungen des BMBF als richtig und zielführend. Die Nachfolgeprojekte von MediGRID geben der Grid-Technologie eine Chance, doch noch für geeignete Forschungsprojekte Verwendung zu finden.

Allerdings werden auch die Nachfolgeprojekte von der Entwicklung der Rahmenbedingungen abhängen:

1. Wird es gelingen, die Grid-Services nutzerfreundlicher zu machen – dies müsste gezielt untersucht werden.
2. Wird es gelingen, Grid-Services bezüglich Sicherheit, Verlässlichkeit und Transparenz auf das Niveau der Forschung in Medizin und Lebenswissenschaften anzuheben?
3. Wird es möglich, eine Community-bezogene Governance-Struktur zu etablieren (etwa das aus MediGRID hervorgegangene Forum Grid der TMF) und zu finanzieren? Welche Macht/Marktbasis wird sie haben können?
4. Werden die methodischen Zentren in Medizin und Lebenswissenschaften die Services annehmen und genügend Anwender Know-How aufbauen?
5. Werden die Forscher in Medizin und Lebenswissenschaften erkennen, dass Grid-Technologie ihnen erlaubt, einige Fragen effizienter zu bearbeiten als bisher?
6. Wird es gelingen, Auftrags-, Abrechnungs- und Finanzierungsprozesse zu etableren, die die nachhaltige Sicherung von Grid-Services ermöglich
7. Werden andere Serviceangebote, den Bedürfnissen der Community gerechter werden als die Grid-Dienste?

Zusammenfassend kann man feststellen, dass es noch keineswegs erwiesen ist, ob und wenn ja wie die Grid-Technologie in Medizin und Lebenswissenschaften Fuß fassen wird. Etwa 10 Jahre währen bis jetzt die Bemühungen. Möglicherweise wird Grid-Computing eine Teilkomponente der Serviceangebote der Rechenzentren für ganz spezielle Anwendungen. Es besteht jedoch auch noch die Möglichkeit, dass sie retrospektiv betrachtet nur ein überschaubarer Lernprozess in der Entwicklung verteilter Dienstleistungsangebote der Informationstechnologie war. Am Ende der momentan laufenden Nachfolgeprojekte von MediGRID werden diese Fragen beantwortet werden können.

5 Veröffentlichungen, Vorträge und Dissertationen

5.1 Veröffentlichungen

Andraschek M (2007). Datenmanagement- / Workflow-Integration für MediGRID, Diplomarbeit, Universität Potsdam, Institut für Informatik, 2007.

Bartsch R, Kantelhardt JW, Penzel T, Havlin S (2007). Experimental evidence for phase synchronization transitions in the human cardiorespiratory system. Phys.Rev.Lett. 98(5).

Beronov K, Dzhimova O, Delgado A, Vossberg M, Krefting D, et al. (2008). Virtual endovascular correction and hemodynamic analysis over the MediGRID-portal, eMBEC 2008 IFMBE Proceedings ECIFMBE 2008, 4th European Conference of the International Federation for Medical and Biological Engineering.

Beronov K, Dzhimova O, Tolxdorff T, Vossberg M, Krefting D (2009). Grid computing for detailed hemodynamics-simulation-based planning of endovascular interventions Healthgrid 2009.

Buch S, Schafmayer C, Völzke H, Becker C, Franke A, von Eller-Eberstein H, Kluck C, Bäßmann I, Brosch M, Lammert F, Miquel JF, Nervi F, Wittig M, Rosskopf D, Timm B, Höll C, Seeger M, ElSharawy A, Fändrich F, Fölsch UR, Krawczak M, Schreiber S, Nürnberg P, Tepel J, Hampe J (2007). A genome-wide association scan identifies the hepatic cholesterol transporter ABCG5/ABCG8 as a susceptibility factor for human gallstone disease. Nat Genet 39(8), S. 995-999.

Canisius S, Mayer G, Ploch T, Koehler U, Kesper K (2007). Automatische Analyse der Muskelaktivitätt zur Untersuchung der REM-Schlaf-Verhaltensstîrung (RBD). Somnologie Suppl. 1, S. 180.

Canisius S, Penzel T, Kesper K, Krefting D (2009). Application of Grid technology for automated detection of sleep disordered breathing Healthgrid 2009.

Canisius S, Ploch T, Gross V, Jerrentrup A, Penzel T, et al. (2008). Detection of sleep disordered breathing by automated ECG analysis. Conf.Proc.IEEE Eng Med.Biol.Soc. 2008, S. 2602-2605.

Canisius S, Ploch T, Kesper K, Smith M, Freisleben B, et al. (2007). Sleep medicine as a scenario for medical grid application. Stud.Health Technol.Inform. 126, S. 37-46.

Canisius S, Viezens F, Hornung B, Ploch T, Krefting D, et al. (2008). Datenschutz für die vernetzte Phänotypisierung in der Schlafmedizin. Somnology - Sleep Research and Sleep Medicine 12(Supplement 1), S. 22-23.

Dickmann F, Viezens F, Sax U (2008). Grid Computing und die Zukunft der IT im Gesundheitswesen. German Medical Science, S. 101-103.

Dobrev A, Scholz S, Zegners D, Stroetmann KA, Semler SC (2009). Economic Performance and Sustainability of HealthGrids: Evidence from Two Case Studies, Healthgrid 2009. Studies in Health Technology and Informatics 147, 151-162.

Drepper J, Semler SC, Mohammed Y, Sax U (2006). Aktuelle Themen des Datenschutzes und der Datensicherheit in der biomedizinischen Forschung. In: Sax U, Mohammed Y, et al., (Eds.) Grid-Computing in der Biomedizinischen Forschung – Datenschutz und Datensicherheit. Urban&Voge, München, S. 25-36.

El Sharawy A, Manaster C, Teuber M, Rosenstiel P, Kwiatkowski R, Huse K, Platzer M, Becker A, Nürnberg P, Schreiber S, Hampe J (2006). SNPSplicer – Systematic analysis of SNP-dependent splicing in genotyped cDNAs. Human Mutation.

Falkner J, Steinke T, Weisbecker A (2006). Grid-Computing. In: Sax U, Mohammed Y Viezens F, Rienhoff, O. (Hrsg.). Grid-Computing in der biomedizinischen Forschung: Datenschutz und Datensicherheit. München: Urban und Vogel, S. 44-55.

Falkner J, Weisbecker A (2006). Integration of ASlications in MediGRID In: Bubak, M., Turala, M., Wiatr, K.: Cracow 2006 Grid Workshop ProceedingS Krakow: Academic Computer Centre Cyfronet AGH, S. 511-518.

Falkner J, Weisbecker A (2008). Medizin – Unterstützung von Prostatabiopsien durch Ultraschall Bildbearbeitung. In: Weisbecker A, Pfreundt, F.-J, Linden, J, Unger S (Hrsg.). Fraunhofer Enterprise Grids: Business CaseS Stuttgart: Fraunhofer IRB, S 31-33.

Falkner J, Weisbecker A, Spath D (2008). Virtualisierung und Grid Computing: Aktuelle Trends und Erfahrungsberichte zum Einsatz von Virtualisierung – und Gridtechnologien. Spath, D, Weisbecker A, Falkner J (Hrsg.). (2008) Virtualisierung und Grid Computing. Stuttgart: Fraunhofer IRB, S. 1-13.

Franke A, Wollstein A, Teuber M, Wittig M, Lu T, Hoffmann K, Nürnberg P, Krawczak M, Schreiber S, Hampe J (2006). GENOMIZER - An Integrated Analysis System for Genome-Wide Association Data. Human Mutation, 27(6), S. 583-588.

Groß A, Hartung M, Kirsten T, Rahm E (2008). Evolution-based analysis of functional protein annotation. In: Thiery et al. (Eds.), Leipzig Research Festival for Life Sciences 2008. S. 50.

Groß A, Hartung M, Kirsten T, Rahm, E (2009). Estimating the quality of ontology-based annotations by considering evolutionary changes. In: Paton et al. (Eds.), Data Integration in the Life Sciences 2009. Lecture Notes in Computer Science (LNBI). Springer, Berlin / Heidelberg 2009, To appear.

Hampe J, Franke F, Rosenstiel P, Till A, Teuber M, Huse K, Albrecht M, Häsler R, Sipos B, Mayr G, De La Vega FM, Briggs J, Günther S, Prescott NJ, Onnie CM, Fölsch UR, Lengauer T, Platzer M, Mathew CG, Krawczak M, Schreiber S (2009). A genome-wide association scan of non-synonymous SNPs identifies a susceptibility variant for Crohn disease in the autophagy-related 16-like (ATG16L1) gene. Nat. Genet. in press.

Hartung M (2008). Management von Ontologien in den Lebenswissenschaften. In: Höpfner, Klan (Eds.), Proceedings of the 20. GI-Workshop on Foundations of Databases (Grundlagen von Datenbanken). Technical Report 01/2008 School of Information Technology, International University in Germany 2008, S. 111-115.

Hartung M, Kirsten T, Groß A, Rahm E (2008). Exploring changes in life science ontologies with OnEX. In: Thiery et al. (Eds.), Leipzig Research Festival for Life Sciences 2008. S. 49.

Hartung M, Kirsten T, Groß A, Rahm E (2009). OnEX: Exploring Changes in Life Science Ontologies. Submitted to BMC Bioinformatics (in review).

Hartung M, Kirsten T, Rahm E (2008). Analyzing the Evolution of Life Science Ontologies and Mappings. In: Bairoch et al. (Eds.), Data Integration in the Life Sciences 2008. Lecture Notes in Computer Science (LNBI) 5109. Springer, Berlin / Heidelberg 2008, S. 11-27.

Hartung M, Loebe F, Herre H, Rahm E (2008). A platform for collaborative management of semantic grid metadata. In: Badica et al. (Eds.), Intelligent Distributed Computing Systems & Applications. Studies in Computational Intelligence 162. Springer, Berlin / Heidelberg 2008, S. 115-125.

Hartung M, Loebe F, Herre H, Rahm E (2009). Management of Evolving Semantic Grid Metadata Within a Collaborative Platform. Submitted to Information Sciences – Special Issue on "Intelligent Distributed Information Systems" (in review).

Hartung M, Rahm E (2007). A Grid Middleware for Ontology Access. In: Proc. 1st German eScience Conference (GES). Conference Paper, ID: 315463.0, Max Planck Digital Library, 2007.

Henschel A, Kim WK, Schroeder M (2006). Equivalent binding sites reveal convergently evolved interaction motifs. In Bioinformatics, 22(5), S. 550-556.

Hoheisel A (2008). Grid-Workflow-Management. In: Weisbecker A, Pfreundt, F.-J, Linden, J, Unger S (Hrsg.). Fraunhofer Enterprise Grids - Software. Fraunhofer IRB Verlag, Stuttgart.

Hoheisel A (2008). Yagsi - Sicherheitsinfrastruktur für serviceorientierte Architekturen. In: Weisbecker A, Pfreundt, F.-J, Linden, J, Unger S (Hrsg.). Fraunhofer Enterprise Grids - Software. Fraunhofer IRB Verlag, Stuttgart.

Hoheisel A, Sommerfeld, D, Peter K (2008). Guidelines for the Deployment of ASlications in MediGRID Technical report, MediGRID, 2008..

Hoheisel A, Unger S (2008). Pharma - Screening Anwendung aus dem Drug-Design. In: Weisbecker A, Pfreundt, F.-J, Linden, J, Unger S (Hrsg.). Fraunhofer Enterprise Grids: Business CaseS Stuttgart: Fraunhofer IRB Verlag, S. 34-35.

Huang B und Schroeder M (2006). LIGSITEcsc: predicting ligund binding sites using the Connolly surface und degree of conservation. BMC Structural Biology.

Kesper K, Canisius S, Jerrentrup A, Penzel T, Koehler U (2007). Entwicklung einer computergestützten EKG-Analyse zur Bestimmung respiratorischer Parameter. Somnologie Suppl. 1.

Kim WK, Henschel A, Winter C, Schroeder M (2006). The many faces of protein-protein interactions: A compendium of interface geometry. In Plos Computational Biology.

Kirsten T, Rahm E (2006). BioFuice: Mapping-based data integration in bioinformatics. In: Leser et al. (Eds.), Data Integration in the Life Sciences 2006. Lecture Notes in Computer Science (LNBI) 4075. Springer 2006, S. 124-135.

Kirsten T, Thor A, Rahm E (2007). Instance-based matching of large life science ontologies. In: Cohen-Boulakia, Tannen (Eds.), Data Integration in the Life Sciences 2007. Lecture Notes in Computer Science (LNBI) 4544. Springer 2007, S. 172-187.

Kottha S, Peter K, Steinke Th, Bart J, Falkner J, Weisbecker A, Viezens F, Mohammed Y, Sax U, Hoheisel A, Ernst T, Sommerfeld D, Krefting D, Vossberg M (2007). Medical Image Processing in MediGRID. German eScience Conference (GES 2007), Baden-Baden, Online: http://edoc.mpg.de/display.epl?mode=doc&id=316613&col=100&grp=1414.

Kottha S., Abhinav K., Müller-Pfefferkorn R., Mix H (2007). Accessing Bio-Databases with OGSA-DAI - A Performance Analysis. In: LNCS Vol. 4360: Proceedings of Distributed, High-Performance and Grid Computing in Computational Biology - International Workshop GCCB 2007, S 141-156, Springer, 2007.

Krawczak M, Thomas NST, Hundrieser B, Mort M, Wittig M, Hampe J, Cooper DN (2006). Single Base-Pair Substitutions in Exon–Intron Junctions of Human Genes: Nature, Distribution, und Consequences for mRNA Splicing. Human Mutation.

Krefting D, Bart J, Beronov K, Dzhimova O, Falkner J, Hartung M, Hoheisel A, Knoch T, Lingner T, Mohammed Y, Peter K, Rahm E, Rienhof O, Sax U, Sommerfeld D, Steinke T, Tolxdorff T, Vossberg M, Viezens F, Weisbecker A (2009). MediGRID: Towards a user friendly secured grid infrastructure. Future Generation Computer Systems, 25(3), S. 326-336.

Krefting D, Canisius S, Hoheisel A, Tolxdorff T, Penzel T (2009). Gridbased Sleep Research - Analysis of Polysomnographies using a Grid Infrastructure 9th IEEE International Symposium on Cluster Computing and the Grid, Shanghai.

Krefting D, Canisius S, Penzel T (2008). Automatisierte Signalanalyse von Polysomnographien unter Verwendung eines Griddatenmanagementsystems 16. Jahrestagung der Deutschen Gesellschaft für Schlafforschung und Schlafmedizin.

Krefting D, Canisius S, Peter K, Tolxdorff T, Penzel T (2008). Signalanalyse von Polysomnographien unter Verwendung einer Gridinfrastruktur 53. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (gmds), Stuttgart.

Krefting D, Falkner J, Hoheisel A, Weisbecker A et al. (2008). MediGRID: Towards a user friendly secured grid infrastructure, Future Generation Computer Systems 25(3), March 2009, S. 326-336.

Krefting D, Vossberg M, Beronov K, Tolxdorff T (2008). Standortunabhängiger Zugang und interaktive Steuerung von gridbasierten Diensten zur medizinischen Bildverarbeutung über ein Gridportal 53. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (gmds).

Krefting D, Vossberg M, Tolxdorff T (2008). Medizinische Bildverarbeitung im MediGRID Poster Presentation, 22. Treffpunkt Medizintechnik, Berlin, Charité - Universitätsmedizin Berlin, Campus Virchow, Berlin.

Krefting D, Vossberg M, Tolxdorff T (2008). Simplified Grid Implementation of Medical Image Processing Algorithms using a Workflow Managment System. Proc. of the Workshop on Medical Imaging on Grids (MICCAI-Grid), S. 23-32.

Krefting D, Vossberg, M, Hoheisel A, Tolxdorff T (2008). Simplified Grid Implementation of Medical Image Processing Algorithms using a Workflow Management System. In: Future Generation Computer Systems, Elsevier (submitted).

Legré Y, Mohammed Y, Viezens F, Rienhoff O, Sax U (2008). HealthGRID/ SHARE: Grids for Health - International Interoperability eHealth Conference 2007 - Experts' Special Interest Sessions. nanos Verlag oHG, Bonn, S. 81-99.

Lingner T und Meinicke P (2006). Remote homology detection based on oligomer distances. Bioinformatics 22, S. 2224-2231.

Luchtmann M, Baecke S, Bernarding J (2007). Neuroimaging: SPM als verteilte Komponente in Grid- und Cluster-Architekturen BVM2007 - Bildverarbeitung für die Medizin.

Luchtmann M, Baecke S, Lätzkendorf R, Bernarding J (2007). MediGRID: Distributed Computing to Accelerate fMRI Analysis 52. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (gmds).

Lützkendorf R, Bernarding J, Hertel F, Viezens F, Thiel A, et al. (2009). Enabling of Grid based Diffusion Tensor Imaging Using a Workflow Implementation of FSL Healthgrid 2009, Berlin.

Manaster C, Zheng W, Teuber M, Wächter S, Döring F, Schreiber S, Hampe J (2005). InSNP – a tool for automated detection und visualization of SNPs und InDels, Human Mutation 26, S. 1-9.

May P, Ehrlich HC, Steinke Th (2006). ZIB Structure Prediction Pipeline: Composing a Complex Biological Workflow through Web Services. In: Nagel et al. (Eds.), Euro-Par 2006 Parallel Processing. Lecture Notes in Computer Science. Springer, Berlin u.a., S. 1148-1158.

Mayer G, Kesper K, Ploch T, Canisius S, Penzel T, et al. (2008). Quantification of tonic and phasic muscle activity in REM sleep behavior disorder. J.Clin.Neurophysiol. 25(1), S. 48-55.

Meinicke P, T. Brodag, Fricke WF, Waack S (2006). P-value based visualization of codon usage data. Algorithms for Molecular Biology1,10.

Merelli E et al. (2006). Agents in bioinformatics, computational und systems biology. Briefings in Bioinformatics.

Mohammed Y (2006). Erweiterte Sicherheit und Datenschutz Techniken. In: Sax U, Mohammed Y, et al., (Eds.) Grid-Computing in der Biomedizinischen Forschung – Datenschutz und Datensicherheit. Urban&Vogel, München, S. 70-77.

Mohammed Y, Rey S, Rakebrandt F, Sax U (2007). The Integration of Grid Resources into a Portal for Research Collaboratories. Studies in Health Technology and Informatics Technol Inform 129, S. 335-339.

Mohammed Y, Sax U, Viezens F, Rienhoff O (2007). Shortcomings of Current Grid Middlewares Regarding Privacy in HealthGrids. Studies in Health Technology and Informatics Technol Inform 126, S. 322-329.

Mohammed Y, Viezens, F., Sax, U., Rienhoff, O. (2006). Rechtliche Aspekte bei Grid-Computing in der Medizin. In: Niederlag W, Dierks C, et al., (Eds.) Rechtliche Aspekte der Telemedizin, Dresden, S. 235-245.

Mohammed Y, Viezens, F., Sax, U.. (2006). Privacy in Grid Computing Projects for Biomedicine – The MediGRID Project. In: Sax U, Mohammed Y, et al., (Eds.) Grid-Computing in der biomedizinischen Forschung - Datenschutz und Datensicherheit. Urban&Vogel, München, S. 9-15.

Morgenstern B, Prohaska SJ, Pöhler D, Stadler PF (2006). Multiple sequence alignment with user-defined anchor points. Algorithms for Molecular Biology 1(1), S. 6.

Müller S (2007). Realisierung von freingranularem Rechtemanagement in einer service-orientierten Grid-Architektur. Diplomarbeit, Universität Potsdam, Institut für Informatik, 2007.

Olabarriaga SD, Glatard T, Hoheisel A, Nederveen AJ, Krefting D (2009). Crossing HealthGrids Borders: Early Results in Medical Imaging. Proceedings of the HealthGrid 2009, Berlin (submitted).

Pellegrini, S, Hoheisel A, Giacomini F, Ghiselli A (2008). Using GWorkflowDL for Middleware-Independent Modeling and Enactment of WorkflowS In: Proceedings of the CoreGRID Integration Workshop 2008, Crete, Greece, 2008.

Penzel T, Hirshkowitz M, Harsh J, Chervin RD, Butkov N, et al. (2007). Digital analysis and technical specifications. J.Clin.Sleep Med. 3(2), S. 109-120.

Peter K (2007). Erasure-tolerant Codes for Rule-based Grid Storage Systems. In: 1st Baltic Conference on Advanced Topics of Telecommunications. Oktober, 2007

Plantikow S, Peter K, Högqvist M, Grimme C, Papaspyrou A (2009). Generalizing the Data Management of Three Community Grids. Future Generation Computer Systems 25 (2009), S. 281-289.

Platzer M, Hiller M, Szafranskzy K, Jahn N, Hampe J, Schreiber S, Backofen R, Huse K (2006). Splice acceptor SNPs - no hypervariability but sequencing errors. Nat. Biotechnology. 24(9), S. 1068-70.

Pommerening K, Sax U, Müller T, Speer R, Ganslandt T, et al. (2008). Integrating eHealth and Medical Research: The TMF Data Protection Scheme. In: Blobel B, Pharow P, et al., (Eds.) eHealth: Combining Health Telematics, Telemedicine Biomedical Engineering and Bioinformatics to the Edge. Akademische Verlagsgesellschaft Aka GmbH, Berlin, S. 5-10.

Rienhoff O (2006). Lösungen für sichere Grid-Anwendungen in der medizinischen Forschung. In: Sax U, Mohammed Y, et al., (Eds.) Grid-Computing in der Biomedizinischen Forschung – Datenschutz und Datensicherheit. Urban&Vogel, München, S. 86-90.

Rienhoff O, Kleinoeder T, Lorberg K (2006). Archivierung von Telemedizin-Sessionen. In: Niederlag W, Dierks C, et al., (Eds.) Rechtliche Aspekte der Telemedizin, Dresden, S. 219-225.

Royer L, Linse B, Wachter T, Bry F, Schroeder M (2006). Querying the semantic web: A case study. In Christopher Baker und Kei-Hoi Cheung, editors, Semantic Web: Revolutionizing Knowledge Discovery in the Life Sciences. Springer.

Sax U (2006). Stand der generischen Datenschutz-Konzepte sowie deren technischen Realisierung in biomedizinischen Grids. In: Sax U, Mohammed Y, et al., (Eds.) Grid-Computing in der Biomedizinischen Forschung – Datenschutz und Datensicherheit. Urban&Vogel, München, S. 38-43.

Sax U, Beckmann M, Müller S, Rienhoff O (2006). Von der Forschung in die Versorgung - Eine Lösung für das Kompetenznetz Angeborene Herzfehler. In: Tolxdorff T, Steyer G, (Eds.), Telemed. Akademische Verlagsgesellschaft Aka GmbH, Berlin, Berlin, 84-88. Online.

Sax U, Mohammed Y, Viezens F, Rienhoff O (2006). Grid-Computing in der Biomedizinischen Forschung – Datenschutz und Datensicherheit. Urban&Vogel, München.

Sax U, Schmidt S (2005). Integration of genomic data in Electronic Health Records--opportunities and dilemmas. Methods Inf Med 44(4), S. 546-550.

Sax U, Weisbecker A, Falkner J, Viezens F, Mohammed Y, et al. (2007). Auf dem Weg zur individualisierten Medizin - Grid-basierte Services für die Elektronische Patientenakte der Zukunft. In: Jäckel A, (Ed.) Telemedizinführer Deutschland 2008, Ausgabe 2008 ed. Medizin Forum AG, Bad Nauheim, S. 47-51.

Sax U, Weisbecker A, Falkner J, Viezens F, Mohammed Y, Hartung M, Bart J, Krefting D, Knoch, TA, Semler, SC (2007). Grid-basierte Services für die elektronische Patientenakte der Zukunft, E-HEALTH-COM 2007(4), KomPart Verlagsgesellschaft mbH & Co. KG, S. 61-63.

Sax U., Weisbecker A, Falkner J, Viezens F, Mohammed Y, Hartung M., Bart J, Krefting D, Knoch, T., Semler SC. (2007). Auf dem Weg zur individualisierten Medizin - Grid-basierte Services für die EPA der Zukunft. Telemed 2007, Berlin. Ausgezeichnet als bester wissenschaftlicher Beitrag.

Schroeder M, Burger A, Kostkova P, Stevens R, Habermann B, Dieng-Kuntz R (2006). Sealife: A Semantic Grid Browser for the Life Sciences Applied to the Study of Infectious Diseases. Proceedings of HealthGrid 2006, Studies in Health Technology and Informatics 120, S. 167-78.

Scholz S, Breitner MH, Blaurock M (2008). A Sustainable Business Model Approach for Grid Computing – And a Life Sciences Example. In: Bichler M, Hess T, et al., (Eds.) Multikonferenz Wirtschaftsinformatik 2008. GITO-Verlag, Berlin, Volltext auf CD, S. 1-12.

Scholz S, Breitner MH, Semler SC, Blaurock M (2007). Business Models for Grid Computing in Life Science: An Approach to Identifying Health Grid Revenue, Mednet 2007, Leipzig. Online: http://www.mednet2007.com/pdfs/fullpaper/0708MednetProceedingTMFUniHan.pdf.

Scholz S, Dobrev A, Breitner MH, Semler SC (2009). Grid Computing in Medizin und Life Sciences - Analyse zu Nachhaltigkeitskonzepten und Business Cases. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie 2009 (accepted).

Scholz S, Dobrev A, Semler SC, Breitner MH (2009). Nachhaltigkeitskonzepte für Grid Computing in der Medizin und den Life Sciences. Telemedizinführer 2009 (accepted).

Scholz S, Semler SC, Breitner MH (2009). Business Aspects and Sustainability for Healthgrids – an Expert Survey. Healthgrid 2009. Studies in Health Technology and Informatics 147, 163-172.

Schultz AK, Zhang M, Leitner T, Kuiken C, Korber B, Morgenstern B, Stanke M (2006). A jumping profile Hidden Markov Model und applications to recombination sites in HIV und HCV genomes. BMC Bioinformatics 7, S. 265.

Schulz M (2007). Realisierung fehlertoleranter Workflows in einer Service-orientierten Grid-Architektur. Diplomarbeit, Universität Potsdam, Institut für Informatik, 2007.

Sommerfeld D, Lingner T, Heilgeist J, Richter H (2008). Gridification and virtualization: enabling e-Science in the life sciences; MG '08: Proceedings of the 15th ACM Mardi Gras conference, Baton Rouge, Louisiana, 2008, S. 1.

Sommerfeld D, Lingner T, Richter H (2007). Stepwise Enabling of AUGUSTUS for MediGRID. Technischer Report TU Clausthal, IfI-07-12 (2007), ISSN 1860-8477, http://www.in.tu-clausthal.de/fileadmin/homes/techreports/ifi0712sommerfeld.pdf.

Sommerfeld D, Lingner T, Stanke M, Morgenstern B, Richter H (2007). AUGUSTUS at MediGRID: Adaption of a Bioinformatics Application to Grid Computing for Efficient Genome Analysis. Future Generation Computer Systems, 25 (2009), S. 337 – 345.

Sommerfeld D, Richter H (2009). A two-tier approach to efficient workflow scheduling in MediGRID. Grid-Technologie in Göttingen - Beiträge zum Grid-Ressourcen-Zentrum GoeGrid, GWDG-Bericht Nr. 74 (2009), ISSN 0176-2516, S. 39-51.

Spath, D, Weisbecker A, Falkner J (Hrsg.) (2008). Virtualisierung und Grid Computing. Stuttgart: Fraunhofer IRB Verlag.

Spath, D, Weisbecker A, Höß, O (Hrsg.) (2006). IT-Virtualisierung und Grid-Computing. Tagungsband Stuttgarter Softwaretechnik Forum 2006. Stuttgart: Fraunhofer IRB Verlag.

Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B (2006). AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res.34, S. 435-439.

Tech M, Morgenstern B, Meinicke P (2006). TICO: a tool for postprocessing the predictions of prokaryotic translation initiation sites. Nucleic Acids Res.34, S. 588-590.

Teyra J, Doms A, Schroeder M, Pisabarro M (2006). SCOWLP: a web-based database afr detailed characterization und visualization of protien interfaces. In BMC Bioinformatics.

Thor A, Hartung M, Groß A, Kirsten T, Rahm E (2009). An Evolution-based Approach for Assessing Ontology Mappings - A Case Study in the Life Sciences. In: Freytag et al. (Eds.), BTW 2009: Datenbanksysteme in Business, Technologie und Web. GI-Edition - Lecture Notes in Informatics (LNI) 144. Bonner Köllen Verlag 2009, S. 277-286.

Tolosana-Calasanz R, Bañares JA, Rana OF, Álvarez P, Ezpeleta J, Hoheisel A (2008). Adaptive Exception Handling for Scientific WorkflowS Concurrency and Computation: Practice and Experience, Wiley (submitted).

Viezens F (2006). Grid-Computing in der Biomedizin. In: Sax U, Mohammed Y, et al., (Eds.) Grid-Computing in der Biomedizinischen Forschung – Datenschutz und Datensicherheit. Urban&Vogel, München, S. 56-62.

Viezens F, Lorberg K, Barz A, Dickmann F, Sax U (2008). Integration und Verarbeitung von Patientendaten aus Krankenhausinformationssystemen im Kontext interdisziplinärer Forschung unter Nutzung von Grid-Technologie. German Medical Science, S. 643-646.

Viezens F, Sax U (2008). Collaboration at the Transition of Personalized Medicine, Electronic Health Record and HealthGRIDs – GRID Services and Applications for the Future. In: Blobel B, Pharow P, et al., (Eds.) eHealth: Combining Health Telematics, Telemedicine Biomedical Engineering and Bioinformatics to the Edge. Akademische Verlagsgesellschaft Aka GmbH, Berlin, S. 49-57.

Vossberg M, Hoheisel A, Tolxdorff T, Krefting D (2008). A Reliable DICOM Transfer Grid Service Based on Petri Net Workflows, CCGRID, 441-448. Online: http://doi.ieeecomputersociety.org/10.1109/CCGRID.2008.122.

Vossberg M, Krefting D, Tolxdorff T (2007). Gridcomputing in der medizinischen Bildverarbeitung: Das MediGRID Projekt BVM2007 - Bildverarbeitung für die Medizin, S. 429-433.

Vossberg M, Krefting D, Tolxdorff T (2007). Using DICOM in Medical Grids. Secure Image Communication and Integration of External DICOM Devices in Globus Grids Proceedings of 5th IASTED BIOMED.

Vossberg M, Tolxdorff T, Krefting D (2008). DICOM image communication in globus-based medical grids. IEEE Trans. Inf. Technol. Biomed. 12, S. 145-153.

Vossberg, M, Hoheisel A, Tolxdorff T, Krefting D (2008). A Workflow-based ASroach for Fault-tolerant Medical Image Transfer in Health GridS In: Future Generation Computer Systems, Elsevier (submitted).

Vossberg, M., Hoheisel A, Tolxdorff, T., Krefting D (2008). A Reliable DICOM Transfer Grid Service Based on Petri Net Workflows In: ccgrid, S. 441-448, 2008 Eigth IEEE International Symposium on Cluster Computing and the Grid (ccgrid), 2008.

Weisbecker A (2006). Grid Computing für Medizin und Lebenswissenschaften. Erscheint in: Wegweiser (Hrsg.). Jahrbuch eHealth & Gesundheitswirtschaft 2006/2007: Modernisierung, Prozessoptimierung und Vernetzung im deutschen Gesundheitswesen.

Weisbecker A (2008). SaaS & Web 2.0 – Aktuelle Trends und Potenziale für Unternehmen. In: Spath, D, Weisbecker A, Höß, O. (Hrsg.). SaaS und Web 2.0. Tagungsband Stuttgarter Softwaretechnik Forum 2008. Stuttgart: Fraunhofer IRB Verlag, S. 15-33.

Weisbecker A, Bart J, Falkner J (2007). MediGRID – Grid Computing for cooperative work in e-health. In: Elsner, C., Häckl, D, van der Slikke, H., Della Mea, V., Arveanitis, T.N., Wiemeth, H. (EdS). Technology and Health Care. International Journal of Health Care Engineering, Volume 15, Number 5, 2007, S. 369-370. Amsterdam: IOS press.

Weisbecker A, Falkner J (2007). Benutzerfreundliche und sichere Zugänge zu Grid Ressourcen. 21. DFN-Arbeitstagung, 30.06.-01.07.2007, Kaiserslautern.

Weisbecker A, Falkner J, (2009). Secure Grid Services for Cooperative Work in Medicine and Life Science. To aSear in: Lin, SC., Yen, E. (EdS). Grid Computing – International Symposium on Grid Computing (ISGC 2008). New York: Springer.

Weisbecker A, Falkner J, Bart J (2006). Vom Internet zum Utility Computing – durch die optimierte Nutzung von Unternehmensressourcen zur vernetzten Arbeitswelt der Zukunft. In: Spath, D: Technologiemanagement in der PraxiS Fraunhofer IRB, S. 179-184.

Weisbecker A, Falkner J, Höß, O, Spath, D (2006). IT-Virtualisierung und Grid-Computing. In: Spath, D, Weisbecker A, Höß, O. (Hrsg.). IT-Virtualisierung und Grid-Computing. Tagungsband Stuttgarter Softwaretechnik Forum 2006. Stuttgart: Fraunhofer IRB, S 3-6.

Weisbecker A, Falkner J, Rienhoff O (2008). MediGRID – Grid Computing For Medicine and Life ScienceS In: Lin, SC., Yen, E. (EdS). Grid Computing – International Symposium on Grid Computing (ISGC 2007). New York: Springer, S. 57-65.

Wieczorek M, Hoheisel A, Prodan R (2008). Towards a General Model of the Multi-criteria Workflow Scheduling on the GriD In: Future Generation Computer Systems, Elsevier (accepted).

Winter C, Henschel A, Kim WK, Schroeder M (2006). SCOPPI: a structural classification of protein-protein interfaces. Nucleic Acid Research, 34(Database issue), S. 310-314.

Zhang M, Schultz AK, Calef C, Kuiken C, Leitner T, Korber B, Morgenstern B, Stanke M (2006). jpHMM at GOBICS: a web server to detect genomic recombinations in HIV-1. Nucleic Acids Res. 34, S. 463-465.

5.2 Vorträge

Bart J (2007). Anwendungsintegration und Load-Balancing in MediGRID, D-Grid AHM, Portal-Workshop, 11.09.2007.

Blaurock M, Scholz S (2008). Geschäftsmodelle in MediGRID. SUGI - Gridbasierte Wertschöpfung und Nachhaltigkeit (Workshop), Siegen.

Canisius S, Ploch T, Kesper K, Smith M, Freisleben B, et al. (2007). Sleep medicine as a scenario for medical grid application. Healthgrid Conference 2007, Chicago.

Canisius S, Viezens F, Hornung B, Ploch T, Krefting D, et al. (2008). Datenschutz für die vernetzte Phänotypisierung in der Schlafmedizin. Somnologie.

Falkner J (2006). Grid Computing – Standards, Architekturen und Anwendungsbeispiele, Stuttgarter Softwaretechnik Forum 2006, 09.11.2006, Stuttgart.

Falkner J (2006). Integration of Applications in MediGRID, Cracow Grid Workshop 18.10.2006.

Falkner J (2007). Benutzerfreundliche und sichere Zugänge zu Grid Ressourcen. 21. DFN-Arbeitstagung, 01.07.2007, Kaiserslautern.

Falkner J (2008). Service Zertifikate und Portal Security, Langfristige Anforderungen an die Sicherheitsmechanismen, 3. D-Grid Security Workshop, Göttingen, 02.04.2008.

Gentzsch W, Sax U (2006). D-Grid The German Grid Initiative in International Context, The 3rd Grid@Asia & GFK 2006 International Joint Workshop. Korea Institute of Science and Technology Information, The Westin Chosun Hotel, Seoul, Korea, 65-79. Online: http://www.gridatasia.net/images/seoul/d-grid.ppt.

Hartung, M (2006). Ontologie- und Datenbankzugriff im Grid-Umfeld. Workshop Ontologien im Grid, 09.03.2006, Leipzig.

Hartung, M (2006). Portlet-basierter Zugang zu Ontologien im Grid. Anwenderworkshop zur Anwendungs- und Portletprogrammierung, 04.07.2006, Göttingen.

Hartung, M (2007). A Grid Middleware for Ontology Access. German eScience Conference, 04.05.2007, Baden-Baden.

Hartung, M (2007). Integration und Nutzung von Ontologien in MediGRID. D-Grid Metadaten Workshop, 26.03.2007, Göttingen.

Hartung, M (2008). A platform for collaborative management of semantic grid metadata. 2nd International Symposium on Intelligent Distributed Computing (IDC), 18.09.2008, Catania (Italien).

Hartung, M (2008). Analyzing the Evolution of Life Science Ontologies and Mappings. 5th International Workshop on Data Integration in the Life Sciences (DILS), 26.06.2008, Evry (Frankreich).

Hartung, M (2008). Management von Ontologien in den Lebenswissenschaften. 20. Workshop über die Grundlagen von Datenbanken (GvDB), 15.06.2008, Apolda.

Hartung, M, Rahm, E (2007). Ontological Applications at MediGRID. eHealth Conference – Experts' Special Interest Session HealthGRID / SHARE Grids for Health – International Interoperability, 18.04.2007, Berlin.

Herre, H (2007). Development, Semantic Integration, and Applications of Ontologies: A Principled Top-Level Approach. 2nd OntoVerse Workshop, 20.11.2007, Düsseldorf.

Hoheisel A (2007). JSDL – Related Work: D-GRDL + XQuery, OGF20, 09.05.2007, Manchester.

Hoheisel A (2007). Meta-Scheduling aus Sicht des MediGrid-Workflow-Managements, D-Grid AHM, Scheduling-Workshop, 11.09.2007.

Hoheisel A (2007). Simple Workflow Scheduling, CoreGRID Workshop, 11.05.2007, Manchester.

Hoheisel A (2008). Using GWorkflowDL for Middleware-Independent Modeling and Enactment of Workflows, Integration Workshop 2008, Crete, Greece, 03.04.2008.

Hoheisel A (2008). Workflows in MediGRID und BauVOGrid, Grid Workflow Workshop, Oldenburg, 03.04.3008.

Kottha S (2007). Accessing Bio-Databases with OGSA-DAI - A Performance Analysis. 21.01.2007, Eilat (Israel).

Krefting D (2007). Medical Image Processing (Forschungsprojekte zur Bildanalyse). Oral Presentation, eHealth2007, Messezentrum, Berlin.

Krefting D (2008). MediGRID : Using Workflows for Medical image and signal processing. Oral Presentation, University of Amsterdam, Academic Medical Center Amsterdam.

Krefting D (2008). MediGRID. Oral Presentation, Universität Potsdam, Institut für Informatik, Potsdam.

Krefting D (2008). MediGRID: Medical image and signal processing in D-Grid. Oral Presentation, BE-Grid Seminar, Brüssel.

Lingner T, Viezens F (2007). Bioinformatic and Ontology in MediGRID. D-GRiD All-Hands-Meeting, Göttingen.

Mohammed Y (2006). Juristische, organisatorische und Management Fragestellungen - Datenschutz, D-Grid Security Workshop, Göttingen. Online: http://www.ingrid-info.de/modules.php?name=AddFile&file=addfile&datei=SW_presentations.

Penzel T (2007). Technical and Digital Requirements for Sleep Recording. APSS Conference 2007.

Penzel T, Nottrott M, Canisius S, Greulich T, Becker HF, et al. (2006). Morphology Changes Improves Detection of Obstructive Sleep Apnea. World Congress for Sleep Apnea, Montreal 2006.

Penzel T, Viezens F (2007). Research Projects to support clinical Medicine. eHealth Conference 2007, Berlin.

Peter K (2006). Verwaltung von Dateien und Datenbanken im Grid mit SRB und OGSADAI. D-Grid Datenmanagement-Workshop, 18.05.2006.

Peter K (2006). Datenmanagement und SRB Anwendungsinteraktion. 04.07.2006 Göttingen

Peter, K (2008). Datenmanagement mit Storage Resource Broker und iRODS. 13.06.2008, Dresden.

Rahm, E (2006). Datenintegration in den Lebenswissenschaften. Workshop Ontologien im Grid, 10.03.2006, Leipzig.

Rahm, E (2007). Data Integration in Bioinformatics and Life Sciences. EDBT Summer School, 06.09.2007, Bolzano (Italien).

Rienhoff O (2007). MediGRID: Aufbau einer Grid-Infrastruktur für Medizin und Lebenswissenschaften, D-GRiD All-Hands-Meeting, Göttingen, 10. Online: http://www.d-grid.de/.

Sax U (2006). Data-related Challenges of Genotype - Phenotype studies, 51. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie. German Medical Science; Doc 06gmds392, Leipzig. Online: http://www.egms.de/en/meetings/gmds2006/06gmds054.shtml.

Sax U (2006). eScience & Grid, Lange Nacht der Informatik, Göttingen. Online: http://www.uni-goettingen.de/de/sh/40244.html.

Sax U (2006). Grid-Computing für die Medizin und Lebenswissenschaften, e-Science und Grid-Aktivitäten in Göttingen, Göttingen. Online: http://www.textgrid.de/fileadmin/TextGrid/konferenzen_vortraege/2006_10_Programm_Goettingen_GridSeminar.pdf.

Sax U (2006). Bedarf für qualifizierte / akkreditierte Zertifikate, D-Grid Security Workshop, Göttingen. Online: http://www.ingrid-info.de/modules.php?name=AddFile&file=addfile&datei=SW_presentations.

Sax U (2007). eScience and Grid, Grid-Seminar, Göttingen, 23. Online: http://textgrid.gwdg.de.

Sax U, Mohammed Y, Viezens F, Lingner T, Morgenstern B, et al. (2006). Medical Grid Computing, International Supercomputer Conference, Dresden. Online: http://www.supercomp.de/show_exhibitor.php?id=325.

Sax U, Mohammed Y, Viezens F, Lingner T, Morgenstern B, et al. (2006). MediGRID - Medical Grid Computing, European Conference on Parallel Computing, Dresden. Online: http://www.europar2006.de/Download/Europar-GV-DGrid.pdf.

Sax U, Mohammed Y, Viezens F, Lingner T, Morgenstern B, et al. (2006). MediGRID - Medical Grid Computing, ICT for BIO-Medical Sciences, Brussels. Online: http://europa.eu.int/information_society/events/ict_bio_2006/images/pictures_conference/ict bio-demos.jpg.

Sax U, Viezens F, Mohammed Y, Lingner T, Morgenstern B, et al. (2006). MediGRID - Medical Grid Computing, EGEE'06 - Capitalising on e-infrastructures, Geneva. Online: http://indico.cern.ch/conferenceOtherViews.py?confId=1504&view=egee_meeting&showDate=all&showSession=all&detailLevel=contribution.

Sax U, Weisbecker A, Falkner J, Viezens F, Mohammed Y, et al. (2007). Auf dem Weg zur individualisierten Medizin - Grid-basierte Services für die EPA der Zukunft, Telemed. Berufsverband Medizinischer Informatiker e. V. (BVMI) und Telematikplattform für medizinische Forschungsnetze e. V. (TMF), Berlin, 9. Online: http://www.e-health-com.de/newsletter/index_11191.html.

Sax U, Weisbecker A, Falkner J, Viezens F, Mohammed Y, et al. (2007). Grid-basierte Services für die elektronische Patientenakte der Zukunft. E-HEALTH-COM 2(4), S. 61-63.

Scholz S (2006). Nachhaltig profitable Geschäftsmodelle für verteilte Ressourcen in der Medizin. Konzeptpräsentation, Doktorandenkolloquium, IWI, Universität Hannover.

Scholz S (2007). A Sustainable Business Model Approach for Grid Computing – And a Life Sciences Example. Multikonferenz Wirtschaftsinformatik 2008, München.

Scholz S (2007). Business Models for Grid Computing in Life Science: An Approach to Identifying Health Grid Revenue. Mednet 2007, Leipzig.

Scholz S (2007). Cost Structures and Cost Analysis in Grid Infrastructures - MediGRID as a Use Case. eHealth Conference 2007, Berlin.

Scholz S (2007). Grid Computing in der Medizin - Anwendungsszenarien und Geschäftsmodellsystematik. Doktorandenkolloquium, IWI, Universität Hannover.

Scholz S (2008). Der Markt für Grid Computing aus Expertensicht - Erste Ergebnisse einer Experteninterviewreihe. Doktorandenkolloquium, IWI, Universität Hannover.

Scholz S (2008). Geschäftsmodellansätze für Grid Computing in der Medizin und den Life Sciences. Doktorandenkolloquium, Wirtschaftsinformatik, Universitäten Niedersachsen, Sankt Andreasberg.

Scholz S (2008). Geschäftsmodellansätze und Wertschöpfung in Healthgrid-Infrastrukturen. Doktorandenkolloquium, IWI, Universität Hannover.

Scholz S (2008). Geschäftsmodelle für Grid Computing in der Medizin und den Life Sciences. Forschungsseminar, Universität Hannover.

Scholz S (2009). Business Aspects and Sustainability for Healthgrids – An Expert Survey, Healthgrid 2009, Berlin.

Steinke T, Sax U (2007). Usability of Solutions from caBIGProject in MediGRID, eHealth Conference, Berlin. Online: http://ehealth.gvg-koeln.de/.

Steinke, T, Krefting, D, Lingner, T, Rahm, E, Penzel, T (2008). MediGRID / caBIG Workshop Demos. MediGRID / caBIG Workshop, 22.01.2008, Berlin.

Steinke, Th (2006). Feingranulare Zugriffskontrolle. D-Grid Security Workshop, 21.03.2006, Göttingen.

Steinke, Th (2006). ZIB Structure Prediction Pipeline: Composing a Complex Biological Workflow Through Web Services. EuroPar 2006, 29.08.2006, Dresden.

Steinke, Th (2007). Medical Image Processing in MediGRID. German e-Science Conference 2007, 03.05.2007, Baden-Baden.

Steinke, Th (2007). On the Usability of Solutions from the caBIG Project in MediGRID. eHealth Conference 2007, 18.04.2007, Berlin.

Steinke, Th (2007). Workflow Systems for Distributed Computing: an User's Perspective, Invited. Distributed Computing in Computational Science (DCCS), 21.06.2007, Wien.

Steinke, Th (2008). OpenAccelerator Portable Accelerated Medical Application Libraries Project. Invited. Bio-Health Applications and Multi-Core Reconfigurable Supercomputing Workshop, 17.07.2008, Manchester

Strauß O (2008). Short presentation on Grid computing activities, EMEA/Fraunhofer IAO, 21.11. 2008.

Viezens F (2006). Trust und Trust-Delegation, D-Grid Security Workshop, Göttingen. Online: http://www.ingrid-info.de/modules.php?name=AddFile&file=addfile&datei=SW_presentations.

Viezens F, Barz A, Lorberg K (2007). Pseudonymisierung von Daten in einem Grid, 52. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie. German Medical Science, Rheinware Verlag, Augsburg, 370-371. Online: http://www2.egms.de/en/meetings/gmds2007/07gmds736.shtml.

Viezens F, Sax U (2006). New concept for processing medical data in distributed environments, 51. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie. German Medical Science, Leipzig, 324-325. Online: http://www.egms.de/de/meetings/gmds2006/06gmds344.shtml.

Weisbecker A (2006). Präsentation des MediGRID Projekts bei einem Industrieworkshop auf der ITeG in Frankfurt am Main 01.06.2006.

Weisbecker A (2007). "MediGRID - Grid Computing for cooperative work in e-Health", MEDNET 2007, Leipzig, 07.-10.10.2007.

Weisbecker A (2007). Interoperability in Grid Computing. eHealth Week Berlin, Special Interest Session „HealthGrid", 18.04.2007, Berlin.

Weisbecker A (2007). MediGRID – Grid Computing for Medicine and Life Sciences. International Symposium on Grid Computing, 28.03.2007, Taipei.

Weisbecker A (2008). Grid Computing. Delegation aus China, Fraunhofer IAO, 11.07.2008.

Weisbecker A (2008). SaaS & Web 2.0 – Aktuelle Trends und Potenziale für Unternehmen, Stuttgarter Softwaretechnik Forum 2008, Fraunhofer IAO, Stuttgart, 26.11.2008.

Weisbecker A (2008). Secure Grid Services for Cooperative Work in Medicine and Life Science, International Symposium on Grid Computing, Taipei, 11.04.2008.

Weisbecker, A (2006). IT-Virtualisierung und Grid Computing, Keynote. Stuttgarter Softwaretechnik Forum 2006, 09.11.2006, Stuttgart.

5.3 Dissertationen im Rahmen des Projektes

Hartung M. Evolution von Ontologien in den Lebenswissenschaften. Prof. Dr. Erhard Rahm, Institut für Informatik. Universität Leipzig (in Arbeit).

Lingner T (2008). Alignmentfreie Analyse von Proteinsequenzen mit Verfahren des maschinellen Lernens. Prof. Dr. Burkhard Morgenstern. Georg-August-Universität Göttingen.

Mohammed Y (2008). Protection and Data Security Concept for Medical Applications in a Grid Computing Environment. Prof. Dr. Otto Rienhoff. Abteilung Medizinische Informatik. Universitätsmedizin Göttingen.

Müller P. Fluoreszenzmikroskopische Darstellung und Analyse genomrelevanter Daten am Beispiel der Her2/neu(Erbb2)-Komplexe. Prof. Dr. Michael Hausmann. Technische Informatik. Kirchhoff-Institut für Physik. Universität Heidelberg (in Arbeit).

Scholz S. Geschäftsmodelle für Grid Computing in der Medizin und den Life Sciences. Prof. Dr. Michael H. Breitner, Institut für Wirtschaftsinformatik. Leibniz Universität Hannover (eingereicht).

Anhang

A1 Pressemitteilungen (Auswahl)

- Spiegel-Online vom 16.07.2007
- Göttinger Tageblatt vom 17.09.2007
- Göttinger Tageblatt vom 23.01.2008
- Göttinger Tageblatt vom 15.04.2008
- Pressemitteilung Universität Göttingen vom 08.05.2008

Spiegel-Online vom 16.07.2007

SPIEGEL ONLINE

16. Juli 2007, 16:49 Uhr

CHOLESTERIN-PUMPE

Risiko-Gen für Gallenstein entdeckt

Wer zu dick oder gar fettleibig ist, bekommt eher Gallensteine als andere Menschen. Doch auch im Erbgut steckt ein Risikofaktor: Deutsche Forscher haben eine Gen-Mutation gefunden, die offenbar die Bildung von Gallensteinen verursachen kann.

Kleine Klümpchen aus Kristallen und Ablagerungen verursachen stechende Schmerzen - bei 170.000 Patienten. So viele Gallenblasen-Operationen werden jedes Jahr in deutschen Krankenhäusern vorgenommen. Übergewicht und Fettleibigkeit sind die Hauptrisikofaktoren, doch Forscher suchen nach weiteren Merkmalen: Wen werden einmal die Stiche im Unterleib piesacken und wen nicht? Konkrete Hinweise auf die betroffenen Erbanlagen gab es aber bislang kaum.

DDP / Medizinhistorisches Museum Charite Berlin

Kleiner Keim, großer Klumpen: Gallensteine im medizin-historischen Museum der Berliner Charité

Nun hat ein Wissenschaftlerteam aus Greifswald, Köln, Berlin, Bonn und Santiago de Chile eine genetische Disposition für Gallensteine entdeckt. Eine bestimmte Mutation des Gens ABCG8 verdoppelt die Wahrscheinlichkeit für Ablagerung in der Gallenblase, berichten die Forscher um Jochen Hampe von der Universität Kiel in einer Online-Vorabveröffentlichung der Fachzeitschrift "Nature Genetics".

Für ihre Studie verglichen Hampe und seine Kollegen zunächst die Erbanlagen von 280 deutschen Gallenstein-Patienten mit denen von 360 gesunden Kontrollpersonen. Dabei ergaben sich an 235 Stellen des Erbgutes besonders deutliche Unterschiede zwischen diesen beiden Gruppen, schreiben die Wissenschaftler. Diese Positionen analysierten die Forscher im zweiten Schritt im Erbgut von insgesamt 1105 Patienten und 873 Kontrollpersonen genauer. Ein Variante des Gens ABCG8 kam dabei als Risikofaktor zum Vorschein.

Offenbar wird vielen den Trägern dieser Mutation eine Überleistung zum Verhängnis, für die das veränderte Gen verantwortlich ist. Denn es ist für den Transport des Blutfetts Cholesterin aus der Leber in die Gallenwege verantwortlich. Dies hatten Forscher um Frank Lammert von der Universität Bonn unlängst in der Fachzeitschrift "Hepatology" (Nr. 46) berichtet.

Die Gallenblase konzentriert Lebersekrete zum Gallensaft, der bei der Verdauung hilft. Wenn die Gallenflüssigkeit zu viel Cholesterin oder Bilirubin (ein Abbauprodukt roter Blutkörperchen) enthält, kristallisieren diese Substanzen in mehr oder weniger großen Körnern aus und verstopfen das Organ - ein schmerzhafter Gallenstein plagt den Patienten.

Wie eine Pumpe auf Hochtouren

Hampe und seine Kollegen vermuten nun, dass durch die Veränderung des Gens ABCG8 mehr Cholesterin in die Gallenflüssigkeit gelangt. Die Mutation sorge gleichsam für einen Pumpe, die permanent "auf Hochtouren" laufe.

Bereits in ihrer "Hepatology"-Studie hatten Hampe und seine Kollegen ABCG8 als Risikofaktor identifizieren können, wenngleich bei weniger Untersuchungspersonen: Demnach waren in 21,4 Prozent der Fälle Gallenstein-Patienten Träger der Genmutation, hingegen nur deutlich weniger als die Hälfte der gesunden Kontrollpersonen. Lammert betonte aber, er rechne noch mit mindestens drei oder vier weiteren Genvarianten, die allesamt das Gallenstein-Risiko erhöhten.

Nach Angaben der Universität Bonn haben 15 bis 20 Prozent aller Deutschen Gallensteine - meist ohne es zu merken. Bei einem Viertel von ihnen melden sich die Steine aber irgendwann mit schmerzhaften

Koliken zu Wort. Für solche Patienten könnte die neue Erkenntnis der Mediziner einst von ganz praktischem Nutzen sein: Einerseits um neben den Risikofaktoren Übergewicht und Fettleibigkeit einen weiteren Risikofaktor auszumachen, andererseits um diesem höheren Risiko dann vorzubeugen. Wenn Träger der Mutation vorbeugende Medikamente gegen die schmerzhaften Klümpchen nehmen würden, ließen sich möglicherweise eines Tages viele der 170.000 Operationen jährlich vermeiden.

stx/dpa

URL:

http://www.spiegel.de/wissenschaft/mensch/0,1518,494630,00.html

© SPIEGEL ONLINE 2007
Alle Rechte vorbehalten
Vervielfältigung nur mit Genehmigung der SPIEGELnet GmbH

Göttinger Tageblatt vom 17.09.2007

Georg August II. (1683-1760)

Vor 270 Jahren

Universität gegründet

Mit der Inauguration der Geogia Augusta wurde am 17. September 1737 die erste deutsche Universität mit voller Forschungs- und Lehrfreiheit eröffnet. Zwar nach Georg August II., König von England und Kurfürst von Hannover, benannt, so gilt doch der hannoversche Premierminister Gerlach Adolph von Münchhausen (1688-1770) als der eigentliche Gründer der Universität in Göttingen. Diese wurde im Sinne der Aufklärung konzipiert und bot somit eine Gleichberechtigung der Fakultäten Philosophie, Medizin, Jura und Theologie.

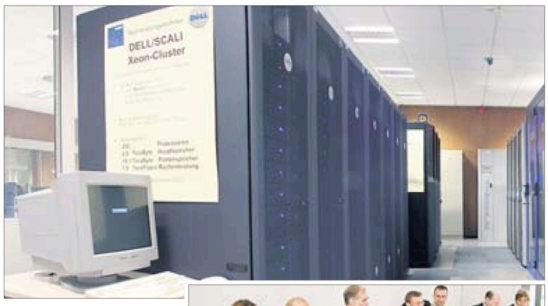
Seit 1832 gab es Pläne zur Gründung einer Universität im Kurfürstentum Hannover: Lüneburg, Celle, Göttingen und Hannover standen zur Debatte. Mit der Entscheidung für Göttingen und dem entsprechenden kaiserlichen Privileg von 1733 begann die Vorbereitung der Universitätsgründung. Das Pauliner Kloster am Papendiek bildete den baulichen Grundstock. *jes*

Die Wissenschaftsredaktion ist per E-Mail erreichbar: hochschule@goettinger-tageblatt.de

Mit dem Netzwerk ändert sich das Denken

Gemeinsames Computer-Netz „D-Grid“ für Forschung und Industrie

Ein offenes Computer-Netzwerk für ganz Deutschland, in dem gemeinsam geforscht und Ressourcen geteilt werden ist Ziel der „D-Grid“-Initiative. In Göttingen trafen sich rund 250 Experten aus allen Sparten der an D-Grid beteiligten Firmen und Einrichtungen zu einem „All Hands Meeting“.



Labor der Zukunft: Hochleistungs-Rechenzentrum. Nolte

Arbeiten am D-Grid: Dr. Sabine Roller, Dr. Alfred Geiger, Prof. Otto Rienhoff, Prof. Wolfgang Gentzsch, Dr. Norbert Lossau, Kalus-Peter Mickel, Prof. Gabriele von Voigt (von links). Windisch

D-Grid sei ein „Verbundprojekt mit Verbundprojekten“, sagt Prof. Wolfgang Gentzsch, Koordinator des durch das Bundeforschungsministerium geförderten Projekt: Derzeit ließen sich dort 150 Einzelprojekte in 20 Verbundprojekten finden. Dabei sei das Vorhaben, Forschungsarbeit ganz praktisch virtuell zu vernetzen, noch gar nicht so alt: Erste Pläne habe es 2003 gegeben, begonnen hätten die Arbeiten im März 2004 mit dem Ziel innerhalb der kommenden fünf Jahren 100 Millionen Euro in die Bereiche Grid-Computing (verteiltes Rechnen) e-Learning (elektronisches Lernen) und Wissensvernetzung zu investieren. So ist es über das Netz möglich Experimente auf „fremden“ Rechner auszuführen, gemeinsame Datenbanken zu nutzen, oder bei der Erstellung von Forschungssoftware zu kollaborieren.

Erste konkrete Maßnahmen seien dann Ende 2005 die Schaffung einer gemeinsamen Ressourcenplattform gewesen. Bislang seien in die unterschiedlichen Ausbaustufen des (D-Grid 1/2) rund 60 Millionen Euro an Bundesmittel geflossen. Mit einem Ziel: Die Deutsche Wissenschaft und Industrie im internationalen Wettbewerb konkurrenzfähig zu halten. Dabei sei die Ausgangssituation für Deutschland sehr gut – im Vergleich der Länder-Netze (das erste sei 2001 in den USA etabliert worden) sehe er Deutschland in einer Spitzengruppe aus sechs Nationen. D-Grid biete ein „sehr mächtiges Kern-Grid“ für Wissenschaft und Wirtschaft. Im Bereich der Ingenieurswissenschaften sie Deutschland mit dem Stuttgarter IN-Grid allerdings führend. Derzeit kann D-Grid eine Kapazität von 2200 Prozessoren, 800 Terabyte (1 Terabyte = 1000 Gigabyte) Platten- und 1400 Terabyte Bandspeicher vorweisen, berichtet D-Grid-Koordinator Klaus-Peter Mickel vom Forschungszentrum Karlsruhe zusammen. Allerdings nicht mehr lange: Binnen vier Monaten soll die Zahlen jeweils vervierfacht worden sein.

Bedarf an Rechenleistung

Ein Projekt der Initiative ist das Geodateninfrastruktur-Grid „GDI-Grid“. Damit sei es etwa Möglich Überflutungszenarien höchst präzise darzustellen und etwa Zeit- und Umgebungsaktuelle Fluchtwege zu berechnen, sagt Prof. Gabriele von Voigt vom Regionalen Rechenzentrum für Niedersachsen, Hannover. Damit solche wichtigen Informationen von den Einsatzkräften schnell online abgefragt werden könnten, müsste natürlich kurzfristig eine immense Rechkapazität zur Verfügung stehen.

Probleme nicht technischer Natur entstehen hingegen bei Projekten an Forschungseinrichtung oder Beteiligung der Wirtschaft, berichten Dr. Norbert Lossau, Leiter der Göttinger Universitätsbibliothek, die die D-Grid-Tagung ausgerichtet hat, und Prof. Otto Rienhoff (Medi-Grid) von der Abteilung Medizinische Informatik der Georgia Augusta. Insbesondere Geisteswissenschaftler müssten lernen, ihre Arbeitsweisen grundlegend umzustellen, im Verbund zu arbeiten, Werkzeug und Quellen zu teilen. Es gebe noch viele, die Projekt wie das Göttinger Text-Grid mit Scheu betrachten, dabei böten diese gerade in Zeiten in denen die Mittel für geisteswissenschaftliche Forschung gering seien, effektive Möglichkeiten verteilten Arbeitens.

Insbesondere im Bereich der Lebenswissenschaften sei im Hinblick darauf schon etwas weiter. Ende der 90er Jahre hätten Wissenschaftler begonnen, aus ihren „Forschungsilos“ herauszutreten, in der Forschungs zu kooperieren, nicht nur im Labor zu forschen, sondern auch Computersimulationen zu nutzen. Doch brauche es Jahrzehnte, bis dieser Prozess vollzogen sein werde, gibt Prof. Rienhoff zu bedenken. Es sei ja auch nicht nicht allzu lange her, dass von international hochrangigen Forschungsmagazinen Arbeiten mit mehreren Autoren akzeptiert würden. *fra*

www.d-grid.de, text-grid.de, medigrid.de, gdi-grid.de.

Bewertung von Futtermitteln

Die Untersuchung und Bewertung von Futtermitteln als Grundlage für die Lebensmittelqualität ist Thema eines Kongresses vom 18. bis 21. September an der Universität Göttingen. Der Verband Deutscher Landwirtschaftlicher Untersuchungs- und Forschungsanstalten (VDLUFA) veranstaltet ihn mit dem Institut für Tierphysiologie und Tierernährung sowie dem Department für Nutzpflanzenwissenschaften der Fakultät für Agrarwissenschaften.

Im Rahmen der Tagung findet am Mittwoch, 19. September, eine Akademische Feier der Fakultät für Agrarwissenschaften statt. Dabei werden mit dem Henneberg-Lehmann-Preis und mit der Verleihung einer Ehrendoktorwürde herausragende Forschungsleistungen auf dem Gebiet der Tierernährung ausgezeichnet. *eb*

Mathematik und E-Learning

Projekte und Anwendungen interaktiver Lernplattformen (E-Learning) sowie Möglichkeiten des Einsatzes von Lernmodulen in der Mathematikausbildung sind Themen einer Tagung am Mathematischen Institut der Universität Göttingen. Zu den SourceTalk-Tagen 2007 vom 17. bis 19. September erwarten die Initiatoren ein internationales Fachpublikum sowie am Dienstag Niedersachsens Wissenschaftsminister Lutz Stratmann.

„Mathematik und E-Learning“ lautet das Thema am heutigen, ersten Veranstaltungstag. Die große deutsche Open-Source-Lernplattform Stud.IP steht am Dienstag und Mittwoch im Mittelpunkt. *jes*

www.sourcetalk.de

Nach freundlichem Beginn ziehen dichte Wolken auf, die Regen mit sich bringen

Wetterlage

Anfangs herrscht im Binnenland Hochdruckeinfluss, von der Nordsee greift eine Kaltfront auf Norddeutschland über.

Vorhersage

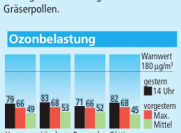
Zunächst wolkig, im Tagesverlauf dichter bewölkt mit nachfolgendem Regen. Werte bei mäßigem Südwestwind bis 22 Grad. Nachts Regen, um 10 Grad. Morgen fällt noch Regen, gebietsweise lockern die Wolken im Tagesverlauf mehr und mehr auf. Am Mittwoch ist es wechselnd wolkig mit Regenschauern.

Biowetter und Pollenflug

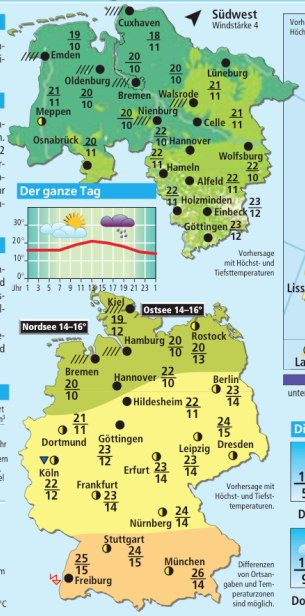
Der Wettereinfluss ist heute eher ungünstig. Bei zu niedrigem Blutdruck sind Kreislaufbeschwerden möglich. Kreislaufanregende Maßnahmen können helfen.

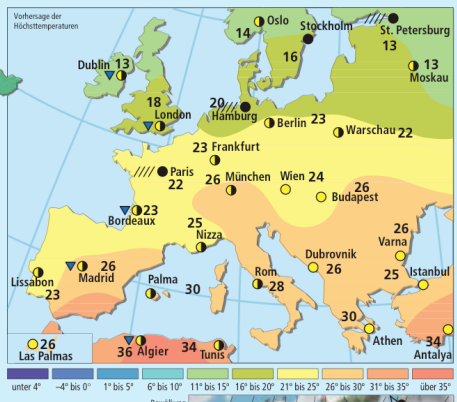
Es gibt nur eine schwache allergene Belastung durch den Flug von Beifuß- und Gräserpollen.

Ozonbelastung

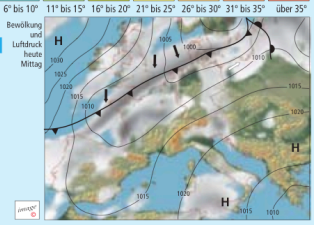


heiter, bewölkt, bedeckt, Nebel, Regen, Schauer, Gewitter, Schnee, Warmfront, Okklusion, Kaltfront, Warmluft, Kaltluft, Kaltluft i. d. Höhe, H Hochdruckzentrum, T Tiefdruckzentrum, 24 Isobaren in hPa, Temperaturen in °C





Die nächsten Tage



Reisewetter

Nord- und Ostseeküste: Meist bedeckt, örtlich Regen, 18 bis 21 Grad.

Harz: Wolkig, gelegentlich auch heiter, trocken, 21 bis 23 Grad.

Schwarzwald und Bodensee: Gebietsweise leichter, schauerartiger Regen, 22 bis 25 Grad.

Südbayern: Erst sonnig, später wolkiger bei angenehmen Temperaturen bis 26 Grad.

Österreich, Schweiz: Meist sonnig, örtlich wolkig mit gewittrigen Schauern, 22 bis 25 Grad.

Südskandinavien: Wechselnd bewölkt mit Regen oder Schauern, Tageshöchstwerte 11 bis 18 Grad.

Großbritannien, Irland: In Westirland freundlich, sonst Schauer, Temperaturen 12 bis 18 Grad.

Italien: Meist sonnig, im Norden wolkig, 28 bis 30 Grad.

Spanien, Portugal: Meist sonnig, örtlich wolkig, Temperaturen 23 bis 29, im Westen 21 Grad.

Griechenland, Türkei, Zypern: Viel Sonne mit nur wenigen Wolken, trocken, Tageshöchstwerte 26 bis 34, in der Ägäis 24 Grad.

Benelux, Nordfrankreich: Verbreitet stark bewölkt mit Schauern, Temperaturen 19 bis 24 Grad.

Südfrankreich: Überwiegend freundlich, 23 bis 27 Grad.

Sonne und Mond

	auf	unter	auf	unter
17. 9. 07	6.59	19.34	13.54	21.01
18. 9. 07	7.01	19.32	15.04	21.35

Mond: erstes Viertel 19. 9., Vollmond 26. 9., letztes Viertel 3. 10. , Neumond 11. 10.

Göttinger Tageblatt vom 23.01.2008

Rechenleistung soll wie Strom aus der Steckdose fließen

Göttinger Zentrum will innovative Netzwerktechnologien für Grid Computing entwickeln

Mit rund 660 000 Euro finanziert das Bundesministerium für Bildung und Forschung (BMBF) am Wissenschaftsstandort Göttingen die Realisierung virtueller Forschungsnetzwerke mit Hilfe innovativer Netzwerktechnologien. Konkret geht es um den Aufbau eines Grid-Ressourcenzentrums.

Ziel des Grid Computing ist es, vorhandene Hardware wie Recheneinheiten und Datenspeicher über das Internet mit minimalem Administrationsaufwand und maximaler Leistung zu vernetzen. Damit soll es möglich sein, Rechenleistung wie Strom aus der Steckdose zu beziehen, ohne sich mit technischen Einzelheiten beschäftigen zu müssen.

Mit Hilfe der Grid Computing-Technologie können Nutzer ihre Aufgaben in einem virtuellen Netzwerk erledigen lassen, wobei das Grid selbst die Zuteilung eines freien Prozessors sowie das Auffinden und Abspeichern der genutzten Daten übernimmt. Von der Realisierung dieser Netzwerke versprechen sich die Experten einen vergleichbaren Effekt wie von der Entwicklung des World Wide Web, das die Internetnutzung für wissenschaftliche, kommerzielle und private Zwecke revolutioniert hat.

Prof. Arnulf Quadt, Teilchenphysiker an der Georg-August-Universität und einer der Initiatoren der Göttinger Grid-Aktivitäten: „Diese Technologie ist sowohl für die Naturwissenschaften als auch für die Geisteswissenschaften von besonderem Interesse, da sie kooperatives Arbeiten in immer komplexer werdenden Forschungsumgebungen unterstützt." Die Realisierung des Göttinger Grid-Ressourcenzentrums basiert auf der interdisziplinären Zusammenarbeit verschiedener Forscherteams in vier Verbundprojekten.

Die Initiatoren von GöGrid: Oswald Haan, Ulrich Sax, Arnulf Quadt, Heike Neuroth, Ulrich Schwardmann (von links). EF

Weltweit einmalig ist das von Göttingen aus koordinierte TextGrid Projekt, mit dem geisteswissenschaftliche Texte digitalisiert und in einer Art riesiger Datenbank gespeichert werden. Forscher können darin schnell und gezielt nach Inhalten und Formulierungen suchen, diese verknüpfen und damit Textstudien durchführen, die bislang nicht möglich waren. Im MediGrid Verbundvorhaben wird das Verhalten von biologischen Molekülen und Wirkstoffen simuliert und mit experimentellen Daten verglichen. Außerdem arbeiten die Wissenschaftler an leistungsfähigen Algorithmen zur Rekonstruktion großer Datenmengen in der modernen medizinischen Diagnostik.

Grid-Computing: Dezentrale Recheneinheiten (CPUs) und Datenspeicher werden sicher über das Internet zu einem großen und extrem leistungsfähigen Ressourcen-Netz verbunden. pug

Gigabyte an Messdaten

In der physikalischen Forschung zur Struktur der Materie werden Großexperimente im Teilchenbeschleuniger „Large Hadron Collider" in Genf in den kommenden Jahren mehrere Millionen Gigabyte an Messdaten erzeugen, die mit herkömmlichen Methoden weder gespeichert noch ausgewertet werden können. Diese Aufgabe wird das HEP Community Grid übernehmen. Im Mittelpunkt des Instant-Grid-Projekts steht der Aufbau vorkonfigurierter Grid-Umgebungen, die von einer CD aus gestartet werden.

Die Göttinger Forscher setzen parallel dazu auf die Ausbildung der Studierenden. In dem interdisziplinären Seminar „Grid-Technologien in der Wissenschaft" werden in diesem Wintersemester erstmals Konzepte, Umsetzung und Anwendung vermittelt.

Die Initiatoren der Grid-Aktivitäten in Göttingen sind – neben Prof. Quadt in der Physik (HEP Community Grid) – Dr. Heike Neuroth von der Staats-und Universitätsbibliothek (TextGrid), Prof. Ulrich Sax von der Universitätsmedizin (MediGrid) sowie Prof. Oswald Haan und Dr. Ulrich Schwardmann von der Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen (InstantGrid). Die aus BMBF-Sondermitteln finanzierten Ressourcen sollen für den Betrieb vorbereitet werden. Dazu wird auch ein bundesweites Expertentreffen stattfinden.

Die Realisierung des Grid-Computing ist Teil der bundesweiten D-Grid-Initiative, die bereits eine Kern-Grid-Infrastruktur für deutsche Wissenschaftler aufgebaut hat. *pug*

www.d-grid.de und die Einzelprojekte unter www.instantgrid.de, www.textgrid.de, www.medigrid.de, www. hepcg.de

Kurz notiert

Mittwoch, 23. Januar

Indologie: „Jainistisches Tempelritual: Praxis und Hintergründe": Prof. Robert Zydenbos, München. Waldweg 26, H1, um 9.15 Uhr.
Nordamerikastudien: „Streitbares Christentum: Evangelikalismus und Fundamentalismus in der Geschichte der USA": Prof. Michael Hochgeschwender, München. Zentrales Hörsaalgebäude (ZHG), Platz der Göttinger Sieben 5, Raum 002, um 12.15 Uhr.
Medizin: „Medikamente und EKG": Kinderärztlicher Nachmittag mit Dr. Heike Schneider, Dr. Thomas Kriebel und Prof. Thomas Paul. Klinikum, Robert-Koch-Straße 40, Hörsaal 01/02, um 16 Uhr.
Wirtschaftswissenschaften: „Investment Timing, Liquidity, and Agency Cost of Debt": Prof. Marliese Uhrig-Homburg, Karlsruhe. Oeconomicum, Platz der Göttinger Sieben 3, Raum 1.122, 16 Uhr.
Agrarwissenschaften: „Moderne Pflanzenschutzforschung – Herausforderung durch neue Technologien und Indikationen": Dr. Klaus Stenzel, Monheim. Pflanzenbau, Von-Siebold-Straße 8, Hörsaal L01, um 16 Uhr.
Mikrobiologie: „RNA Controls B Cell Maturation": Prof. Hans-Martin Jäck, Erlangen. Hygiene-Institut, Kreuzbergring 57, Forum, um 16.30 Uhr.
Medizin: „Neues zur Genetik des Restless-Legs-Syndroms": Dr. Juliane Winkelmann, München. Klinikum, Robert-Koch-Straße 40, Hörsaal 55, um 17 Uhr.
Geowissenschaften: „Architektur und Entwicklung von Riftzonen auf Ozeaninseln": Andreas Klügel, Bremen. Geowissenschaftliches Zentrum, Goldschmidtstr., MN 14, 17.15 Uhr.
Philologie: „Konsolationskonzepte bei Seneca und Plinius": Prof. Claudia Wiener, München. Goßlerstraße 10, AP26, 18.15 Uhr.
Umweltgeschichte: „Das Erdbeben von Lissabon 1755: Die Rezeption der Erschütterung": Prof. Gerhard Lauer. Zoologie und Anthropologie, Bürgerstraße 50, Hörsaal, um 18.15 Uhr.
Jura: „Völkerrecht in der Praxis": Dr. Susanne Wasum-Rainer, Auswärtiges Amt, Berlin. ZHG, Raum 101, um 18.15 Uhr.
Philosophie: „Geist und praktisches Selbstverständnis": Prof. Holmer Steinfath. Auditorium, Weender Landstr. 2, Hörsaal 11, um 18.15 Uhr.
Biochemie: „Structural and Functional Analyses of Viral Attachment Mechanismus": Prof. Thilo Stehle, Tübingen. Humboldtallee 23, Gr. Hörsaal, um 18.15 Uhr.

Kein Vortrag von Ackermann

Die für den heutigen Mittwoch an der Universität Göttingen geplante Veranstaltung von Prof. Lutz Ackermann muss entfallen. Die Fakultät für Chemie teilte mit, die Antrittsvorlesung werde im Sommersemester nachgeholt. *jes*

Weisbrod in Bielefeld

Historiker forscht als Gastwissenschaftler

Der Göttinger Historiker Prof. Bernd Weisbrod wird das erste Halbjahr 2008 als Gastwissenschaftler am Zentrum für interdisziplinäre Forschung (ZiF) der Universität Bielefeld arbeiten. Er befasst sich dort mit den Mitgliedern einer internationalen Arbeitsgruppe mit der „Kontrolle der Gewalt" im 20. Jahrhundert. Im Mittelpunkt seiner historischen Untersuchungen steht politische Gewalt als „Kommunikationsstrategie". Die ZiF-Forschungsgruppe unter Leitung von Prof. Wilhelm Heitmeyer, Bielefeld, und Prof. Heinz-Gerhard Haupt, Florenz, wird neue Erkenntnisse zur Rolle der Gewalt zusammentragen. Dabei geht es um Amok, Terrorismus und Gewaltmärkte.

Weisbrod lehrt und forscht am Seminar für Mittlere und Neuere Geschichte der Universität Göttingen. Zu seinen Arbeitsschwerpunkten gehören die englische Sozialgeschichte des 19. Jahrhunderts und die deutsche Geschichte des 20. Jahrhunderts. *jes*

Überwiegend stark bewölkt und zeitweise Regen bei Tageshöchstwerten um 6 Grad

Wetterlage

Nach kurzem Zwischenhocheinfluss überquert uns von der Nordsee her eine Warmfront. Dabei kommt milde Luft.

Vorhersage

Meist ist es stark bewölkt, und es regnet zeitweise etwas. Um 6 Grad, nachts kaum kälter mit Werten um 5 Grad. Schwacher bis mäßiger Südwestwind. Morgen herrschen dichte Wolken vor. Es regnet zunächst nur wenig, erst später zieht von Nordwesten her ein Regengebiet heran. Auch Freitag wieder Regen.

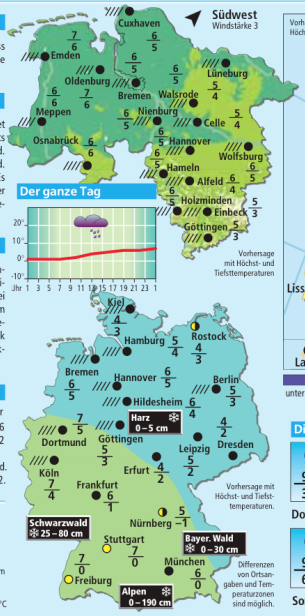
Biowetter

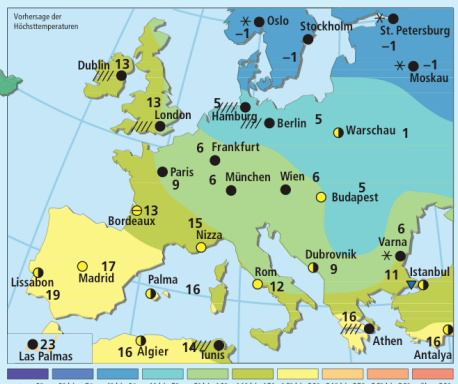
Heute ist der Wettereinfluss leicht ungünstig. Die Kopfschmerz- und die Migräneanfälligkeit ist erhöht. Sowohl bei zu hohem als auch bei zu niedrigem Blutdruck können Herz-Kreislauf-Beschwerden auftreten. Bei Bluthochdruck am Morgen und Abend ungewohnte Aktivitäten im Freien vermeiden!

Sonne und Mond

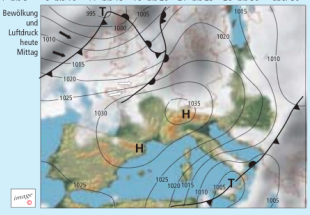
	auf	unter	auf	unter
23. 1. 08	8.16	16.52	18.07	8.56
24. 1. 08	8.15	16.53	19.31	9.12

Mond: letztes Viertel 30. 1. , Neumond. 7. 2. , erstes Viertel 14. 2. , Vollmond 21. 2.





Die nächsten Tage



Reisewetter

Nord- und Ostseeküste: Regen oder Schneefall, Werte 1 bis 7 Grad.
Harz: Wenig Sonnenschein, zeitweise leichter Regen, 3 bis 5 Grad.
Schwarzwald und Bodensee: Vielerorts freundliches Wetter, meist trocken, 3 bis 7 Grad.
Südbayern: Anfangs heiter bis wolkig, später Wolkenaufzug bei Temperaturen von 0 bis 5 Grad.
Österreich, Schweiz: Im Osten Österreichs einige Wolken, sonst viel Sonne, 1 bis 7 Grad.
Südskandinavien: Meist bedeckt, in weiten Teilen Regen, –1 bis 6 Grad.
Großbritannien, Irland: Ergiebige Regenfälle oder Schauer, 10 bis 13 Grad.
Italien: Sonniges Wetter, in Süditalien zeitweise Regen, Tageshöchstwerte 9 bis 14 Grad.
Spanien, Portugal: Heiteres Wetter mit viel Sonnenschein, 15 bis 20 Grad.
Griechenland, Türkei, Zypern: Regen, an der türkischen Riviera freundlich, 8 bis 18 Grad.
Benelux, Nordfrankreich: Sonne und Wolken, im Nordosten etwas Regen, 7 bis 12 Grad.
Südfrankreich: Weitgehend sonnig bei 6 bis 14 Grad.
Israel, Ägypten: In weiten Teilen sonnig, in Israel aber leichter Regen, 16 bis 21 Grad.
Mallorca, Ibiza: Heiteres Wetter mit viel Sonnenschein, niederschlagsfrei, 14 bis 17 Grad.
Tunesien, Marokko: Sonne und Wolken, meist trocken, 14 bis 20, in Südmarokko 23 Grad.

Göttinger Tageblatt vom 15.04.2008

Göttingen erprobt Nachfolger des Internet

Wissenschaftsminister eröffnet Grid-Ressourcen-Zentrum Am Faßberg / Vorteile für Forscher

Wissenschaftler und Politiker versprechen sich vom Grid-Computing die Nachfolge des Internet – nur wesentlich leistungsfähiger. Am Dienstag ist das Göttinger Grid-Ressourcen-Zentrum (GoeGrid) Am Faßberg in Betrieb genommen worden.

VON JENS WUCHERPFENNIG

Heutzutage sind Forscher, die für ihre Arbeit große Datenmengen bewältigen müssen, auf Superrechner an ihrem Forschungsstandort angewiesen. Mit dem Datennetzwerk Grid soll das anders werden: Weltweit sollen Wissenschaftler künftig Zugriff auf Supercomputer, Datenspeicher und Messinstrumente erhalten.

Vertreter von Wissenschaft und Politik versprechen sich von GoeGrid, das Hochleistungsrechner mit mehr als 800 Rechenkernen enthält, viel: „Die Georgia-Augusta hat damit wiederum einen Beitrag zu einer elementaren Zukunftsfrage geleistet", sagte Niedersachsens Wissenschaftsminister Lutz Stratmann (CDU). „Göttingen ist damit zum Motor des Grid-Computing geworden", so der Minister. Prof. Uwe Schwiegelshohn, Geschäftsführer von D-Grid, einer vom Bund geförderten Initiative zur Weiterentwicklung der Grid-Technologie, sagte: „Grid ist die Zukunft." Wenn es in Göttingen gelänge, ein anwenderfreundliches Grid-Netzwerk auzubauen, „ist auch die Wirtschaft bereit aufzusprigen".

Texte und Dunkle Materie

Den Festvortrag zum Thema „Grid Computing als Werkzeug zur Erschließung neuer Welten" hielt Prof. Rolf-Dieter Heuer, Forschungsdirektor des Deutschen Elektronen-Synchrotons (DESY) und zukünftiger Generaldirektor des CERN, der Europäischen Organisation für Kernforschung in Genf (Schweiz). Mit dem Festakt wurde zugleich ein regionales Rechencluster der Göttinger Teilchenphysik im weltweiten Grid-Netzwerk für den neuen Teilchenbeschleuniger Large Hadron Collider des CERN eingeweiht. „Mit dem Teilchenbeschleuniger soll der Ursprung der Masse und der Dunklen Materie geklärt werden", so Heuer. Dafür müssten gewaltige Datenmengen verarbeitet werden. „Das geht nur, wenn wir Grid-Computing haben", so Heuer. Auch für andere Wissenschaften wie die Textanalyse umfangreicher Werke oder für die Medizin ergeben sich neue Möglichkeiten, die mit GoeGrid getestet werden sollen. Initiatoren des Zentrums sind Experten der Niedersächsischen Staats- und Universitätsbibliothek Göttingen und der Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen (GWDG), in deren Räumen GoeGrid angesiedelt ist, Physiker und Bioinformatiker der Georg-August-Universität sowie Vertreter der Medizinischen Informatik an der Universitätsmedizin Göttingen. Das Bundesministerium für Bildung und Forschung (BMBF) hat den Aufbau von GoeGrid mit rund 930 000 Euro unterstützt. Weitere 200 000 Euro stammen aus einem Berufungsverfahren der Theoretischen Physik.

Von Pionierrolle Göttingens überzeugt: Uni-Vizepräsident Markus Hoppe, GWDG-Geschäftsführer Prof. Bernhard Neumair, Schwiegelshohn, Stratmann, Heuer, Ministerialdirektor Wolf-Dieter Lukas und Prof. Thomas Purschke (von links). Mischke

Georg-August-Universität Göttingen

Der Präsident
Presse, Kommunikation & Marketing

Pressemitteilung
Nr. 97/2008 – 8. Mai 2008

Göttinger Grid-Ressourcen-Zentrum wird offiziell in Betrieb genommen

Festakt mit Minister Lutz Stratmann – Vortrag über Einsatzmöglichkeiten des Grid Computing

(pug) Zur offiziellen Eröffnung des Göttinger Grid-Ressourcen-Zentrums (GoeGrid) findet am 13. Mai 2008 ein Festakt statt. Im Rahmen der Veranstaltung soll die internationale Bedeutung des Grid Computing – die Realisierung virtueller Forschungsnetzwerke mit Hilfe innovativer Netzwerktechnologien – der Öffentlichkeit vorgestellt werden. Eingeladen haben dazu die Initiatoren des Zentrums: Dazu gehören Experten der Niedersächsischen Staats- und Universitätsbibliothek Göttingen und der Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG), Physiker und Bioinformatiker der Georg-August-Universität sowie Vertreter der Medizinischen Informatik an der Universitätsmedizin Göttingen. Das Bundesministerium für Bildung und Forschung (BMBF) hat den Aufbau von GoeGrid mit rund 930.000 Euro unterstützt. Weitere 200.000 Euro stammen aus einem Berufungsverfahren der Theoretischen Physik. An der Eröffnungsveranstaltung wird der Niedersächsische Minister für Wissenschaft und Kultur, Lutz Stratmann, teilnehmen.

Zur Begrüßung sprechen Minister Stratmann, Universitäts-Vizepräsident Markus Hoppe und als Vertreter des BMBF Dr. Wolf-Dieter Lukas sowie Prof. Dr. Uwe Schwiegelshohn, Geschäftsführer der bundesweiten D-Grid-Initiative. Den Festvortrag zum Thema „Grid Computing als Werkzeug zur Erschließung neuer Welten“ hält Prof. Dr. Rolf-Dieter Heuer, Forschungsdirektor des Deutschen Elektronen-Synchrotons (DESY) und zukünftiger Generaldirektor des CERN, der Europäischen Organisation für Kernforschung in Genf (Schweiz). Im Anschluss informieren mehrere Vorträge über die vielfältigen Grid-Einsatzmöglichkeiten in verschiedenen Wissenschaftsdisziplinen. Mit dem Festakt wird zugleich ein regionales Rechencluster der Göttinger Teilchenphysik im weltweiten Grid-Netzwerk für den neuen Teilchenbeschleuniger Large Hadron Collider des CERN eingeweiht.

Ziel der Grid-Technologie ist es, vorhandene Rechen- und Speicherressourcen über das Internet mit minimalem Administrationsaufwand und maximaler Leistung zu vernetzen. Dabei können Nutzer in einem virtuellen Netzwerk arbeiten, wobei das Grid selbst die Zuteilung eines freien Prozessors sowie das Auffinden und Abspeichern der genutzten Daten übernimmt. GoeGrid ist in den Räumen der GWDG angesiedelt und bündelt die Ressourcen der am Wissenschaftsstandort Göttingen ansässigen Grid-Projekte. Die Bereitstellung von Hochleistungs-Rechenclustern mit umfangreichen Speicherkapazitäten bildet die zentrale technische Basis für die künftige digitale Wissenschaftsinfrastruktur und -kommunikation. Informationen sind im Internet unter www.gwdg.de/goegrid abrufbar.

Hinweis an die Redaktionen – Einladung zum Fototermin:
Der Festakt zur Eröffnung des Göttinger Grid-Ressourcen-Zentrums am Dienstag, 13. Mai 2008, findet im Max-Planck-Institut für biophysikalische Chemie, Am Faßberg 11, Hörsaal, statt und beginnt um 13.45 Uhr. Zuvor ist ein **Fototermin** (13.30 Uhr) bei der Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen,

Verantwortlich:
Marietta Fuhrmann-Koch

Wilhelmsplatz 1 · 37073 Göttingen
Tel. + 49 551/39-4342 · Fax + 49 551/39-4251

pressestelle@uni-goettingen.de
www.uni-goettingen.de

Am Faßberg 11, Turm 6, geplant. Treffpunkt ist die Eingangshalle. Dabei besteht die Möglichkeit, Aufnahmen der Grid-Ressourcen im Maschinenraum der GWDG zu machen.

An dem **Fototermin** nehmen teil:
Lutz Stratmann (Niedersächsischer Minister für Wissenschaft und Kultur)
Prof. Dr. Rolf-Dieter Heuer (Deutsches Elektronen-Synchrotron DESY)
Markus Hoppe (Vizepräsident der Georg-August-Universität Göttingen)
Dr. Wolf-Dieter Lukas (Bundesministerium für Bildung und Forschung)
Prof. Dr. Bernhard Neumair (Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen)
Dr. Heike Neuroth (Niedersächsische Staats- und Universitätsbibliothek Göttingen)
Prof. Dr. Thomas Pruschke (Institut für Theoretische Physik der Universität Göttingen)
Prof. Dr. Arnulf Quadt (II. Physikalisches Institut der Universität Göttingen)
Prof. Dr. Ulrich Sax (Abteilung Medizinische Informatik der Universitätsmedizin Göttingen)
Prof. Dr. Uwe Schwiegelshohn (D-Grid)

A2 Projektflyer

- MediGRID Flyer, deutsche Version
- MediGRID Flyer, englische Version

Was heißt eigentlich Grid-Computing?

Fasst man die Rechenleistung von Computern innerhalb eines Netzwerks zusammen, um nicht nur Daten auszutauschen, sondern zudem die parallele Lösung von rechenintensiven Problemen zu ermöglichen (verteiltes Rechnen), so spricht man von **Grid-Computing**. Grids sind gut skalierbar: durch Einbindung weiterer Rechner oder gar Verknüpfung mehrerer Grids kann die Leistung um Größenordnungen verändert werden. Ein solches Grid kann die Rechenkapazität und -geschwindigkeit von Supercomputern weit übertreffen.

Werden auch Datenbestände, auf welche die Grid-Nutzer zugreifen können, vernetzt bzw. für Grid-Anwendungen erschlossen, entsteht ein **Data Grid**, in Unterscheidung zum **Computing Grid**, in welchem die Verarbeitung dieser Daten stattfindet. Schließlich können auch gezielt unterschiedliche Funktionen oder Algorithmen, die jeweils in verschiedenen Systemen verfügbar sind, miteinander verknüpft und für komplexe Aufgaben simultan nutzbar gemacht werden.

Bestimmte Dienste oder Anwendungen sind oftmals integraler Bestandteil von Grids und werden durch diese den Nutzern zur Verfügung gestellt (**Application Sharing**).

Grid-Computing zielt also insbesondere auf die Zusammenführung, Auswertung und Darstellung großer Datenmengen ab, wie sie beispielsweise in der medizinischen Forschung, der Meteorologie oder Astronomie anfallen.

Einige Charakteristika von Grids erinnern an das herkömmliche **Stromnetz**. Das Grid stellt durch eine intelligente dynamische Verwaltung seiner verteilten Ressourcen Rechenleistung und Datenmaterial zur Verfügung. Dem Nutzer eines Grids ist letztlich egal, wo seine Daten genau verarbeitet werden. Er will sich möglichst einfach mit dem Grid verbinden und Rechenleistung beziehen.
Aus der Analogie zum Stromnetz drängen sich auch unmittelbar die Fragen nach geeigneten Betreibermodellen, Sicherheitsaspekten sowie Verfahren zur Messung des „Verbrauchs" von Leistung und dessen Abrechnung auf. Insofern muss ein Grid als Gesamtkonstrukt gedacht werden: seiner „Inbetriebnahme" muss die Beantwortung dieser Fragen ebenso vorausgehen wie die Lösung der hard- und softwaretechnischen Herausforderungen.

Das MediGRID-Konsortium

Das MediGRID-Konsortium besteht aus acht Partnern.

Sprecher: Prof. Otto Rienhoff
Universität Göttingen, Abteilung Medizinische Informatik
haegar@med.uni-goettingen.de

Modul A: Koordination
Telematikplattform für Medizinische Forschungsnetze e.V.
medigrid@tmf-ev.de
Sebastian C. Semler, Mathias Blaurock

Modul B: Ontologie-Werkzeuge
Universität Leipzig
rahm@informatik.uni-leipzig.de
Prof. Markus Löffler, Prof. Erhard Rahm, Toralf Kirsten,
Michael Hartung, Katrin Loebe

Modul C: Ressourcenfusion
Konrad Zuse Institut, Berlin
steinke@zib.de
Dr. Thomas Steinke, Ralph Müller-Pfefferkorn, Ulrich Schwardmann

Modul D: Middleware
Fraunhofer Gesellschaft, Stuttgart
anette.weisbecker@iao.fraunhofer.de
PD Dr. Anette Weisbecker, Jürgen Falkner, Thilo Ernst

Modul E: Bioinformatik
Uniklinikum Schleswig-Holstein
j.hampe@mucosa.de
PD Dr. Jochen Hampe, Dr. Tobias A. Knoch,
Prof. Burkhard Morgenstern, Prof. M. Schroeder

Modul F: Bildverarbeitung
Charité, Universitätsmedizin Berlin
thomas.tolxdorff@ukbf.fu-berlin.de
Prof. Thomas Tolxdorff, Dr. Dagmar Krefting, Michael Vossberg

Modul G: Klinische Forschung
Universität Marburg
canisius@staff.uni-marburg.de
Prof. Thomas Penzel, PD Dr. Richard Dodel, Prof. Wolfgang Oertel,
Sebastian Canisius

Modul H: eScience
Universität Göttingen
usax@med.uni-goettingen.de
Prof. Ulrich Sax, Yassene Mohammed

Internetseite: **http://www.medigrid.de**
Applikationsportal: **https://portal.medigrid.de**

02_2007 v09

MediGRID

Vernetzte Rechnerleistung für die Biomedizinische Forschung

– ein Projekt im D-GRiD-Verbund –

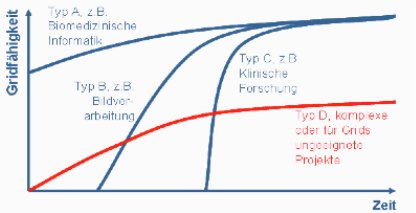
Gefördert vom
Bundesministerium
für Bildung
und Forschung

Das MediGRID-Projekt

Grid-Computing gilt inzwischen auch in der biomedizinischen Forschung als wichtiger Fortschrittsfaktor. Im Verbundprojekt MediGRID haben sich renommierte Forschungseinrichtungen in den Bereichen Medizin, Bioinformatik und Gesundheitswissenschaften als Konsortialpartner zusammengeschlossen und haben eine **Grid-Middleware-Integrationsplattform** und darauf aufsetzende eScience-Dienste für die biomedizinische Wissenschaft entwickelt. Die weitere Einbeziehung zahlreicher assoziierter Partner aus Industrie, Versorgungs- und Forschungseinrichtungen stellt das Projekt auf eine breite Interessensbasis.

Für das Projekt wurde eine modulare Aufgabenverteilung gewählt: In den vier methodischen Projekt-Modulen (Middleware, Ontologie-Werkzeuge, Ressourcenfusion und eScience) entwickelten die entsprechenden Konsortialpartner schrittweise eine Grid-Infrastruktur. Sie berücksichtigten dabei insbesondere die Anforderungen der Grid-Nutzer aus dem biomedizinischen Umfeld, welche exemplarisch in den drei anwendungsorientierten Projektmodulen (Biomedizinische Informatik, Bildverarbeitung, Klinische Forschung) erarbeitet wurden.

MediGRID soll anhand ausgewählter Anwendungsbeispiele zeigen, ob, wie und in welchen Fällen Grid-Technologie in der biomedizinischen Forschung gewinnbringend eingesetzt werden kann. Die Untersuchung der Anwendungsbeispiele wird dabei nicht parallel gestartet (Pipelining):



Typ-A-Projekte (s. Grafik) sind auf Grid-Computing vorbereitet und wurden sofort gestartet, andere können aufgrund der Erfahrungen aus den A-Projekten dann rasch gridfähig gemacht werden (Typen B und C). Projekte vom Typ D sind von vornherein zu vermeiden.

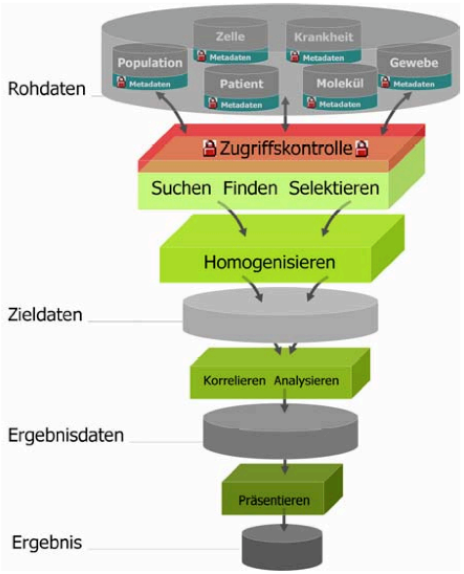
Die ersten Applikationen im Bereich Biomedizinische Informatik und Bildverarbeitung sowie Ontologie-Werkzeuge können Sie im MediGRID Applikationsportal selbst aufrufen:

https://portal.medigrid.de

Zur internationalen Anbindung des Projektes findet ein intensiver Austausch mit Partnern in Frankreich, UK und den USA statt. Deshalb hat MediGRID auch internationale Workshops durchgeführt, z.B. auf dem AMIA-Kongress 2006 in Washington, DC (USA).

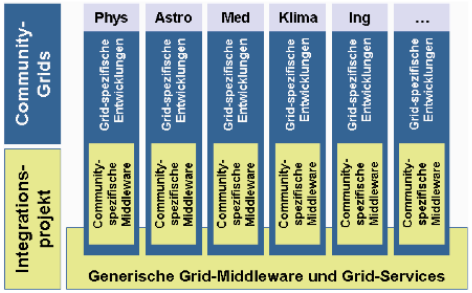
Von den Rohdaten zum Forschungsergebnis

Ein Grid, das in erster Linie für Forschungs-Communities designt wurde, muss zunächst aus **heterogenen Datenquellen**, beispielsweise aus dem biomedizinischen Bereich, über entsprechende **Zugriffskontrollmechanismen** Datensätze **suchen und selektieren**. Wegen der Heterogenität der Daten ist dann ein **Homogenisierungsschritt** notwendig, bei dem mit Hilfe von Ontologien semantische Interoperabilität erreicht wird. Die homogenisierten Zieldaten können dann mit biometrischen Methoden **korreliert und ausgewertet**, schließlich verdichtet und als Ergebnisdaten **präsentiert** werden.



MediGRID im D-GRiD-Verbund

Unter dem Dach der D-GRiD-Initiative werden vom Bundesministerium für Bildung und Forschung (BMBF) weitere Grid-Projekte gefördert, unter anderem für die Bereiche Astronomie, Klimaforschung, Hochenergiephysik und Ingenieurwissenschaften. Durch das D-GRiD-Integrationsprojekt werden die Aktivitäten der einzelnen Grid-Projekte koordiniert und Synergien genutzt, mit dem Ziel, letztlich eine gemeinsame Grid-Plattform für die deutsche Wissenschaftsgemeinde zu schaffen.



Weitere Informationen zu D-GRiD und den verschiedenen Community-Projekten:

www.d-grid.de

What is Grid Computing?

Grid computing uses the resources of many separate computers connected by a network (usually the internet) to solve large-scale computation problems.
Grid computing offers a model for solving massive computational problems by making use of the distributed resources (computing power, storage capacity or data algorithms) of large numbers of disparate, often desktop, computers treated as a virtual cluster embedded in a distributed telecommunication infrastructure.
Grid computing focuses on the ability to support computation across administrative domains, which sets it apart from traditional computer clusters or traditional distributed computing.

Grids offer a way to solve grand challenge problems like protein folding, financial modeling, earthquake simulation, climate/weather modeling an others found in Biomedical Informatics. They also provide means for offering information technology as a utility bureau for commercial and non-commercial clients, with those clients paying only for the resources they actually use, as with electricity or water. Grid computing has the design goal of solving problems too big for any single supercomputer, whilst retaining the flexibility to work on multiple smaller problems. Thus Grid computing provides a multi-user environment. Its secondary aims are: better exploitation of the available computing power, and catering for the intermittent demands of large computational exercises.
This implies the use of secure authorization techniques to allow remote users to access computing resources.
Grid computing involves sharing heterogeneous resources (based on different platforms, hardware/software architectures, and computer languages), located in different places belonging to different administrative domains over a network using open standards. In short, it involves virtualizing computing resources.
Grid computing is often confused with cluster computing. The key difference is the resources which comprise the grid are not all within the same administrative domain.

Functionally, one can classify grids into several types:
Computational Grids (including CPU scavenging grids), which focuses primarily on computationally-intensive operations.
Data grids, or the controlled sharing and management of large amounts of distributed data.
Equipment Grids which have a primary piece of equipment e.g. a telescope, and where the surrounding Grid is used to control the equipment remotely and to analyze the data produced.
Semantic Grids which contain complex, heterogeneous data types.

Source: partially Wikipedia.org

The MediGRID Consortium

MediGRID is realized in a consortium with eight partners.

Spokesperson: Prof. Otto Rienhoff
University Göttingen, Head of Department of Medical Informatics
haegar@med.uni-goettingen.de

Module A: Coordination
Telematikplattform for Medical Research Networks (TMF e.V.)
medigrid@tmf-ev.de
Sebastian C. Semler, Mathias Blaurock

Module B: Ontology tools
University Leipzig
markus.loeffler@imise.uni-leipzig.de
Prof. Markus Löffler, Prof. Erhard Rahm, Toralf Kirsten, Michael Hartung, Katrin Loebe

Module C: Resource fusion
Konrad-Zuse-Institute, Berlin
steinke@zib.de
Dr. Thomas Steinke, Ralph Müller-Pfefferkorn, Ulrich Schwardmann

Module D: Middleware
Fraunhofer-Gesellschaft IAO / FIRST, Stuttgart
anette.weisbecker@iao.fraunhofer.de
PD Dr. Anette Weisbecker, Jürgen Falkner, Thilo Ernst

Module E: Bioinformatics
Uniklinikum Schleswig-Holstein, Kiel
j.hampe@mucosa.de
PD Dr. Jochen Hampe, Dr. Tobias A. Knoch, Prof. Burkhard Morgenstern, Prof. M. Schroeder

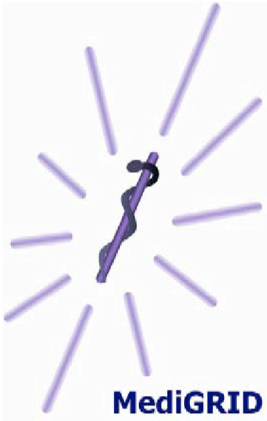
Module F: Image Processing
Charité, University Medicine Berlin
thomas.tolxdorff@ukbf.fu-berlin.de
Prof. Thomas Tolxdorff, Dr. Dagmar Krefting, Michael Vossberg

Module G: Clinical Research
University Marburg
penzel@med.uni-marburg.de
Prof. Thomas Penzel, PD Dr. Richard Dodel, Prof. Wolfgang Oertel, Sebastian Canisius

Module H: eScience
University Göttingen, Department of Medical Informatics
usax@med.uni-goettingen.de
Prof. Ulrich Sax, Yassene Mohammed

Website: **http://www.medigrid.de**
Applicationsportal: **https://portal.medigrid.de**

06_2007 v09



Development of a Grid Infrastructure for the Medical and Life Sciences

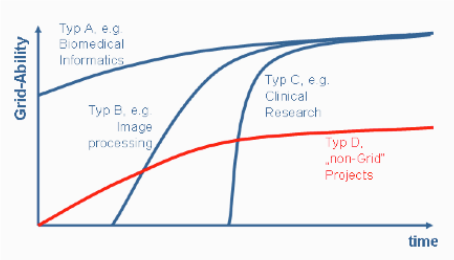
– A Project of the D-GRiD Group –

Sponsored by

Grant ID: 01AK803A - F

The MediGRID-Project

Grid computing is regarded as one of the critical success factors in biomedical Research. The joint project MediGRID unifies well known research institutes in the area of medicine, biomedical informatics and life sciences into a consortium. Numerous associated partners from industry, healthcare and research facilities ensure a broad representation of these communities.
The main goal of MediGRID is the Development of a **Grid middleware integration platform** enabling eScience services for biomedical life science. Therefore the consortium allocated the tasks in different modules. The four methodological modules (middleware, ontology, resource fusion and eScience) plan to incrementally develop and provide a Grid infrastructure while taking into account the need of the biomedical users. The user communities are represented in three research modules for biomedical informatics, image processing and clinical research.
MediGRID is initially funded until 2008. Within this time period our goal is to demonstrate the benefit of Grid computing in use cases covering different areas of biomedical informatics.



Type-A projects (s. figure) can start immediately; they are prepared for Grid computing. Other projects can be "gridified" using the experience of the first projects (type B and C). Projects of the type D should be avoided. The idea behind is to increase the slope of the learning curve (pipelining).

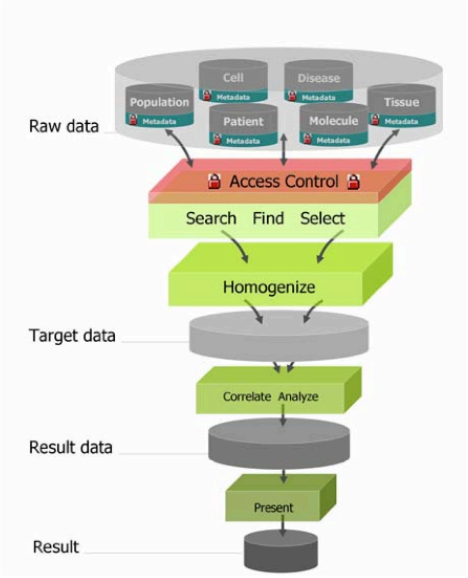
The first applications in image processing, ontology tools and bio-medical informatics are incorporated in our portal and can be called up:

https://portal.medigrid.de

We are member of HealthGrid.Org and organized a workshop during the AMIA 2006 in Washington, D.C. (USA). We hope to improve our services through international cooperation with project like caBIG.

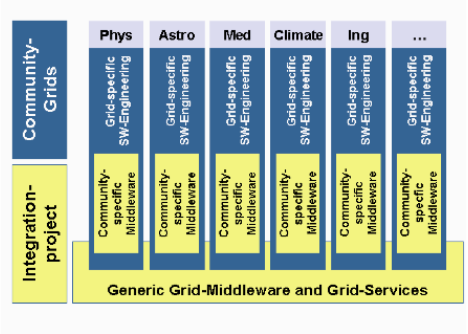
From Raw Data to Research Results

Whereas Grid projects are pretty similar in the lower layers, the biomedical community has to face particular challenges in the upper layers. The top layer represents the heterogeneous biomedical data sources. Beyond the problem to find the relevant data sets via metadata description, access control to the data is of paramount interest, as the owner of the data are foremost patients. Due to the heterogeneity of the data we need an additional ontology layer to homogenize the data. Given semantic interoperability the researcher can correlate and analyze the data with biomedical informatics methods. Finally the result data can be presented.



MediGRID is Part of the D-GRiD-Group

Under the roof of the D-GRiD initiative the German Federal Ministry of Education and Research (BMBF) is funding additional Grid projects e.g. for astronomy, climate research, high-energy physics and engineering science.
The D-GRiD integration project coordinates the activities of the individual community projects using synergies to create and consolidate a Grid platform for the German research community.



You can find further information about D-GRiD and the related community projects on

www.d-grid.de