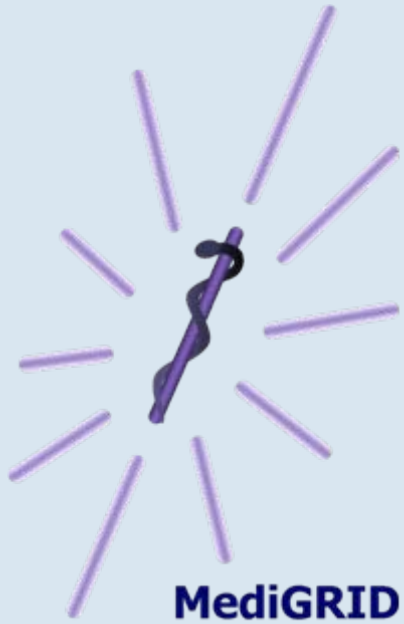
MediGRID / caBIG Workshop Demo

TMF, Berlin, 23.01.2008

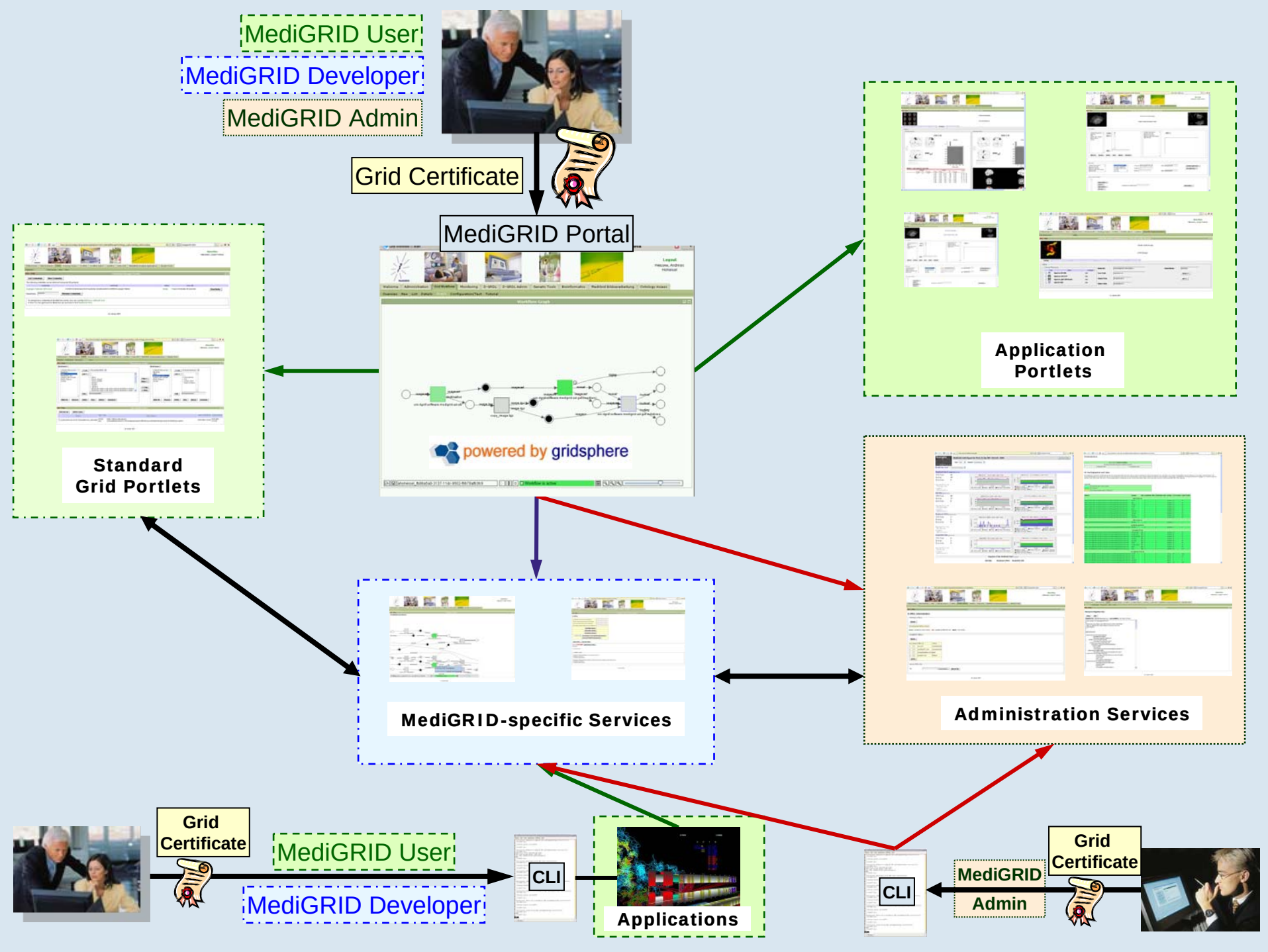
Sebastian Canisius
Jürgen Falkner
Dagmar Krefting
Thomas Lingner
Thomas Penzel
Erhard Rahm
Thomas Steinke



Bundesministerium für Bildung und Forschung

Demo Outline

- Technical Introduction
- Application demos (supported by explanatory slides):
  - Bioinformatics:
    Augustus genome sequence analysis
  - Ontologies:
    Biomedical ontology access + D-Grid ontology
  - Medical Image Processing:
    Virtual Vascular Surgery
  - Clinical Research:
    QRS-analyses of sleep-ECGs


MediGRID User
MediGRID Developer
MediGRID Admin
Grid Certificate
MediGRID Portal
powered by gridsphere
Standard Grid Portlets
Application Portlets
MediGRID-specific Services
Administration Services
Grid Certificate
MediGRID User
MediGRID Developer
CLI
Applications
CLI
MediGRID Admin
Grid Certificate

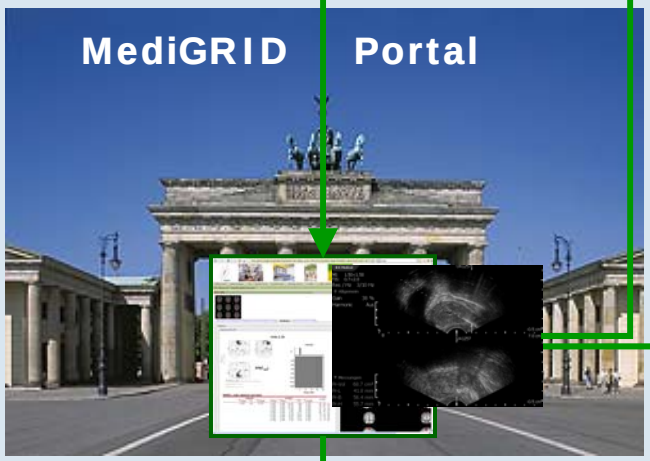
Grid Workflow Management and Resource Virtualization

Result

Input Parameters for Applications

Job-Initialization

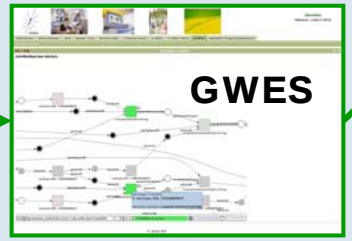
MediGRID Portal



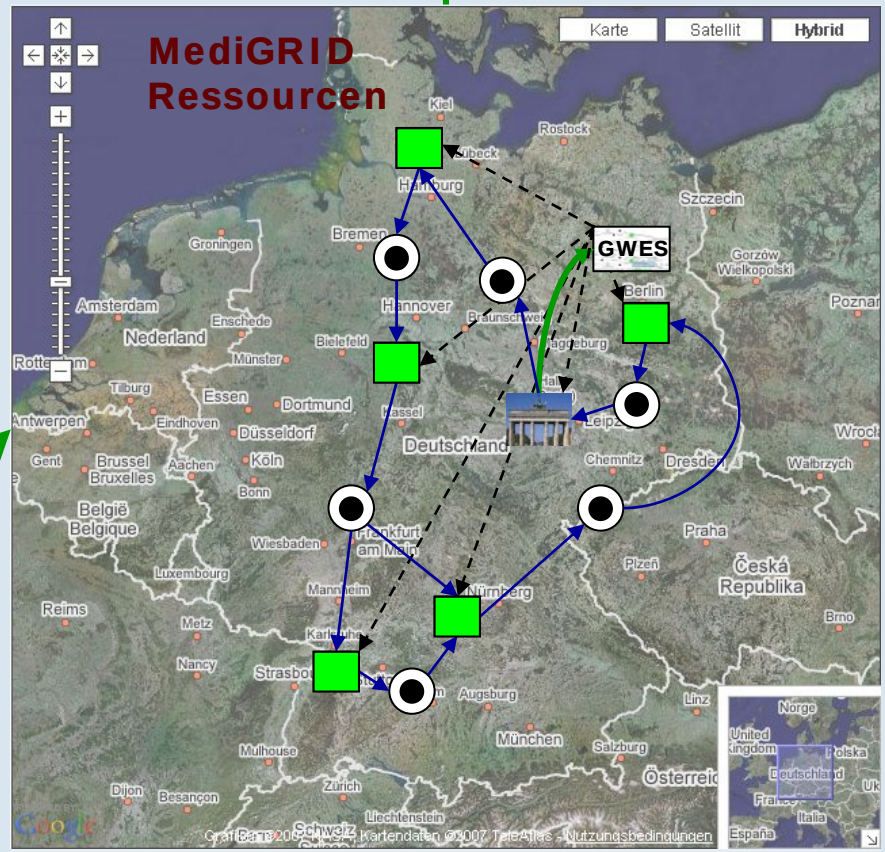
Automatic Resource Selection

Workflow-Initialization

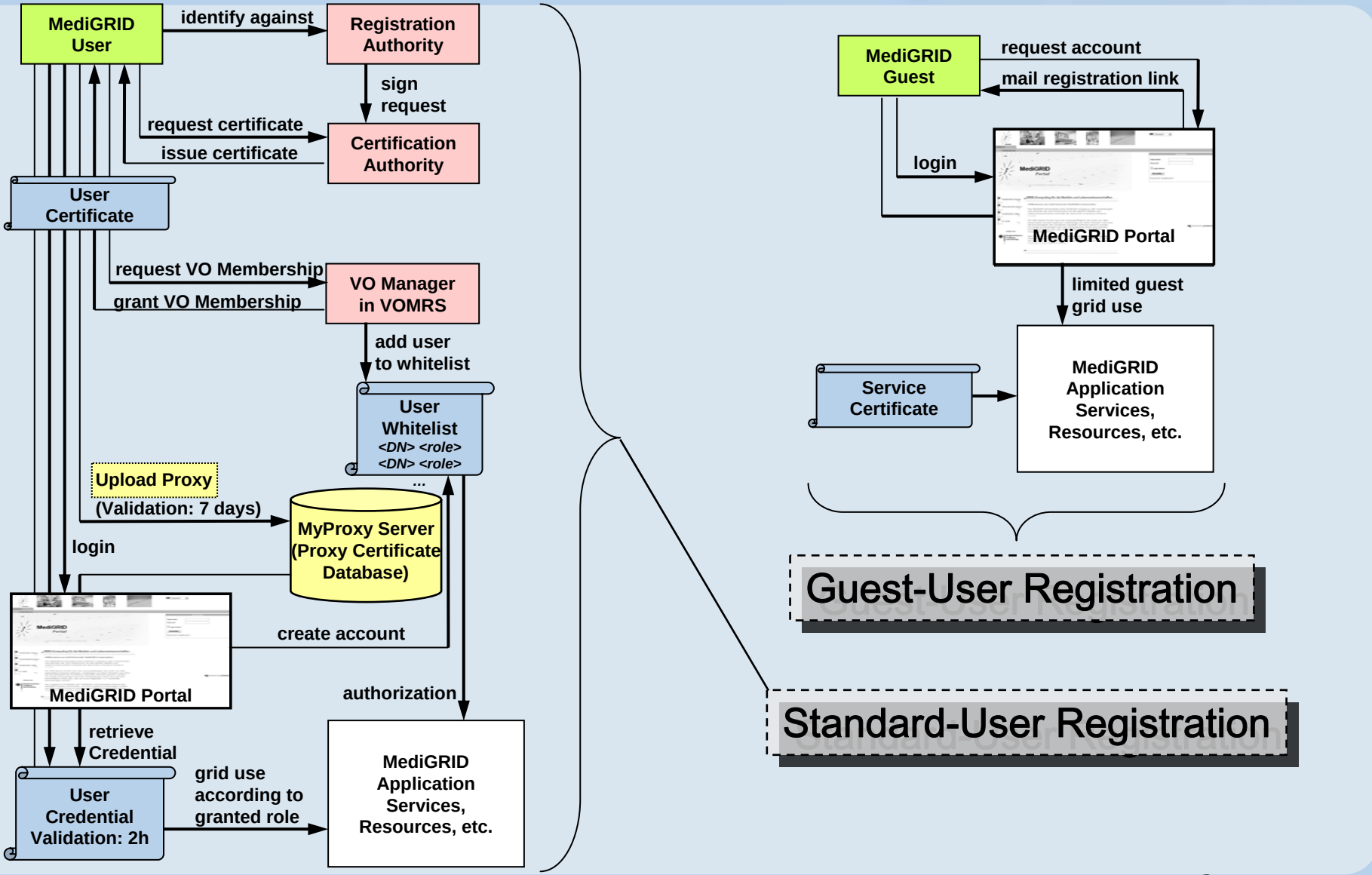
Job-Submission



Grid Workflow Execution Service



Secure Access to MediGRID



MediGRID and its Applications

Bioinformatics

- Augustus
- Agrippa
- SequCorr
- SNPselect

Image Analysis

- USI
- fMRI
- VirtSurgery

Clinical Research

- CR-QRS

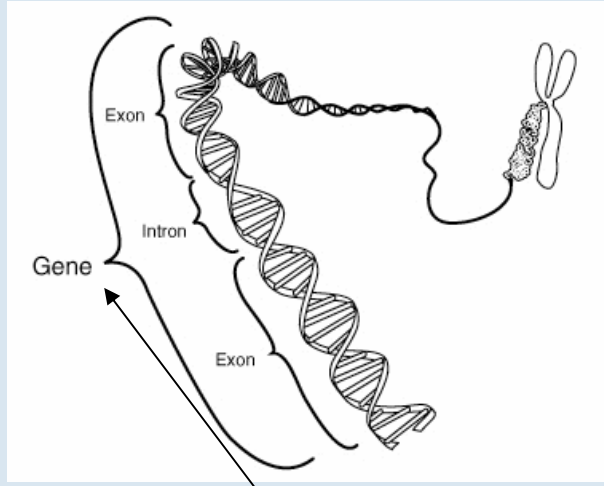
MediGRID core site being part of D-Grid

D-Grid core site

Gene prediction on DNA sequences

„in silico“ gene finding

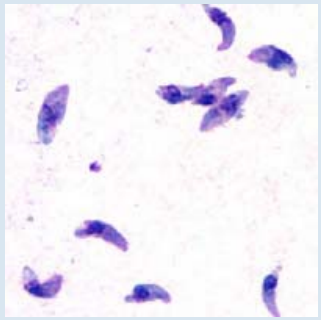
- find important functional regions in the genome (genes, proteins)
- “wet lab” gene finding is tedious and expensive



...ATAGCTAGCT AGCTGATCG...

high-throughput shotgun sequencing

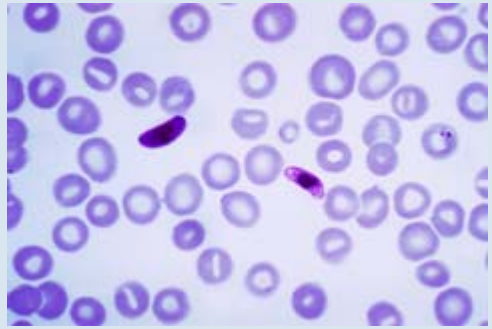
Medical Relevance



Toxoplasma gondii
(Toxoplasmosis)

can cause in enfants:

- *central nervous system disorders*
- *enlargement of the liver and spleen*
- *blindness*
- *mental retardation*



Plasmodium falciparum
(Malaria)

accounts for 80% of all human malarial infections and 90% of the deaths

AUGUSTUS is used in several international genome projects

Brugia malayi
(Filariasis)

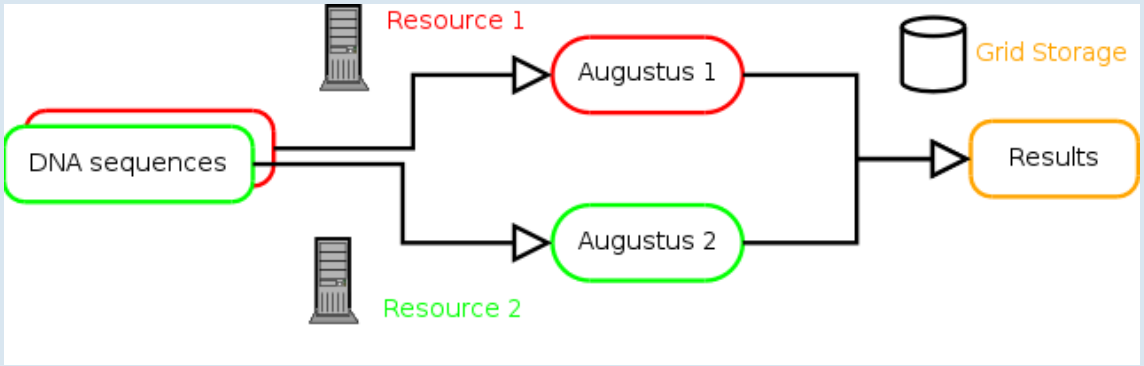
can cause Elephantiasis

Aedes aegypti
(transfers Yellow fever,
Dengue fever)

Schistosoma mansoni
(Schistosomiasis)

Precise prediction is computationally demanding!

- long running time (up to weeks for genomes), but easy to split into parallel processes
- wide-spread community without computer science background and low computing power
  - ➔ simple usage and accessibility
  - ➔ intuitive interpretation of results



Augustus Demo

- how to create guest user account
- easy access to application, easy job configuration
- automatic segmentation and workflow creation
- workflow management allows for nice extras
- intuitively interpretable presentation of prediction results
- links to Genome Browser, BLAST search of predicted genes and ontology component

Biomedical ontology access in grids
Motivation and realization

Motivation

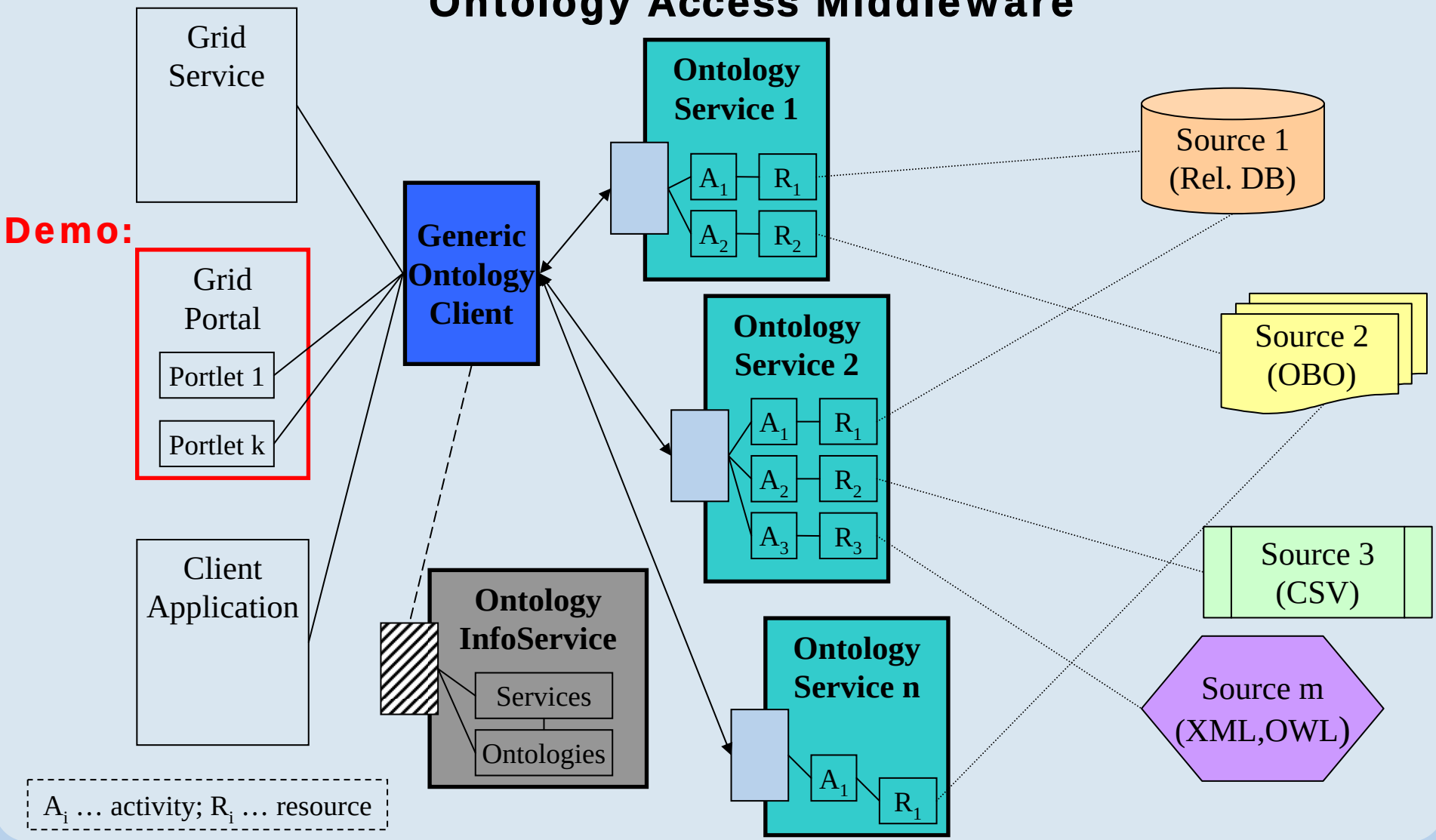
- **Key problem:** Heterogeneity between existing biomedical ontologies → different source types and formats
- Yet no existing, usable ontology access system in grids
- Integration of different ontologies for MediGRID applications

What has been realized?

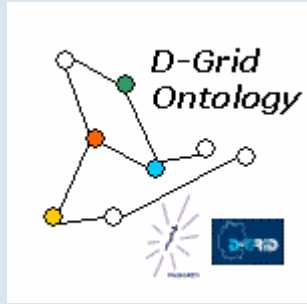
- Middleware for ontology access in grids
- Based on D-Grid standards (Globus, OGSA-DAI)
- Extensible / flexible: easy integration of new ontologies
- Different usage scenarios
  1. Browsing / Look Up Service (interactive use) **(Demo)**
  2. Utilization in MediGRID applications **(AUGUSTUS Demo)**

Ontologies Demo

Ontology Access Middleware



D-Grid Ontology (DGO)
Motivation, goals and realization



Motivation

- D-Grid – nationwide and growing project (D-Grid I/II)
  - Hundreds of participants, organizations and resources
- Meta information about D-Grid is distributed and heterogeneous
- Missing semantic linkage between different information objects

Goals

- Development of a common D-Grid Ontology by ontological management of *content types* (Person, Project, Resources, ...)
- Provision of service interfaces for third-party grid applications

What has been realized? (Demo)

- Wiki-like system for collaborative development of DGO
- Usage of Web 2.0 techniques
  - Easy addition or change of content, different search capabilities
  - Visualization of semantic content (maps, trees)

http://buell.izbi.uni-leipzig.de/dgo

Future challenges and topics

Ontology based data integration

- Computation of similarity between different biomedical objects
- Utilization of ontology mappings (→ ontology matching)

Matching of biomedical ontologies

- Computation of matchings/mappings between different biomedical ontologies
- Validation of match results by experts
- Metadata vs. instance-based matching techniques
- Management and processing of generated or publicly available ontology matchings/mappings

Analysis of ontology evolution

- Semantic changes between versions of an ontology
- Effects of ontology evolution on ontology matching results

Virtual Vascular Surgery

Hemodynamic simulations based on a patient's vascular geometry allows for virtual surgery of cardiovascular deseases

Segmentation of vascular geometry from CT images

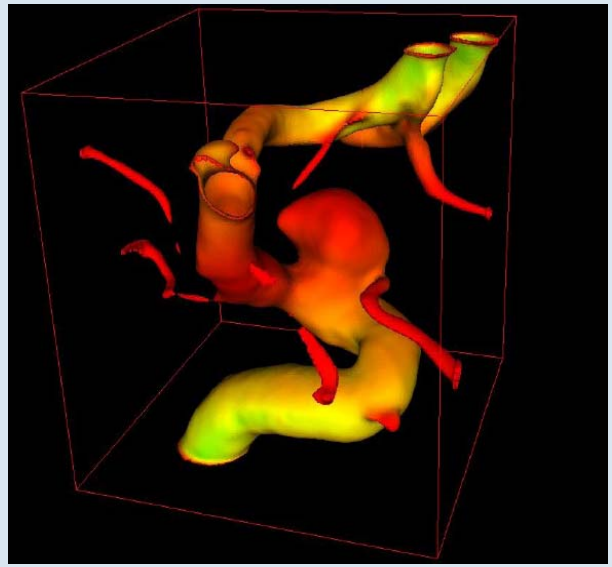
- interactive segmentation and virtual surgery

Numerical simulation of blood flow

- time consuming processing step
- initial parameters/geometry

Visualization of results

- Blood flow, pressure field



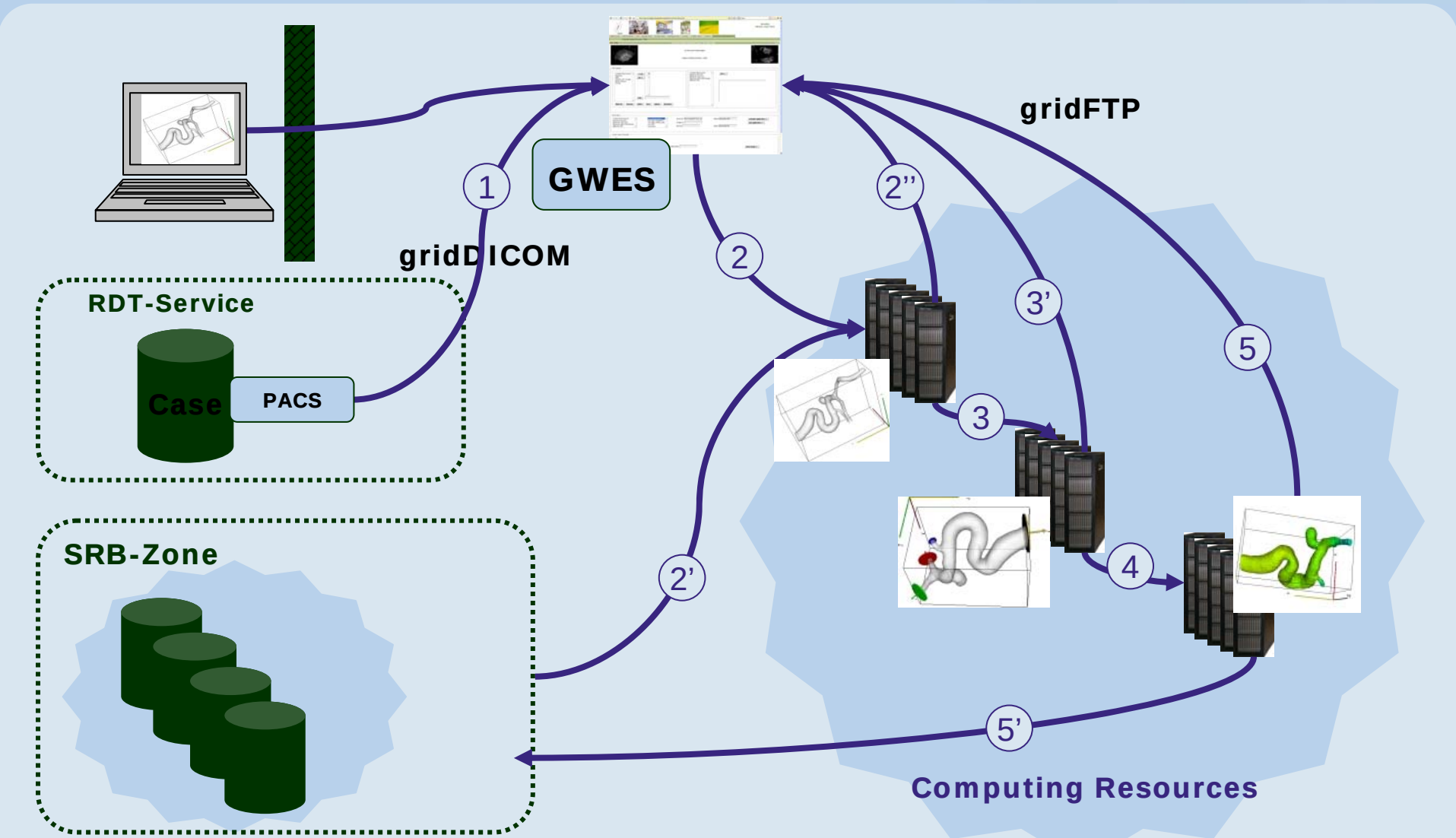
Preprocessing locally (not shown):

Interactive Segmentation with licensed software

Live Demonstration:

1. Access to PACS resource, data selection
2. 3D View on chosen case
3. Setting of cut positions and pressure boundary conditions
4. Starting of domain decomposition and simulation
5. 3D View on results

Middleware Components



Rationale: GRID and Sleep Medicine

- Roughly 20% of the population suffer from sleep disorders
- New investigational approaches for sleep disorders with high prevalence and/or limitations in daytime performance
- Current diagnostics include sophisticated overnight examinations (polysomnography)

Analysis Features suitable for GRID

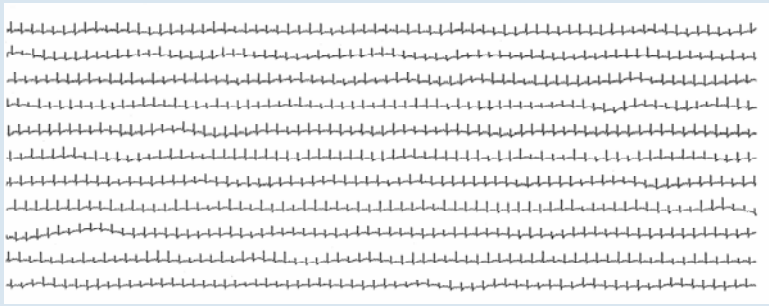
- Analysis of ECG (Electrocardiogram) for the determination of heart rate variations (HRV) during sleep.
  → Development of algorithms finished, implementation in Grid environment in progress.

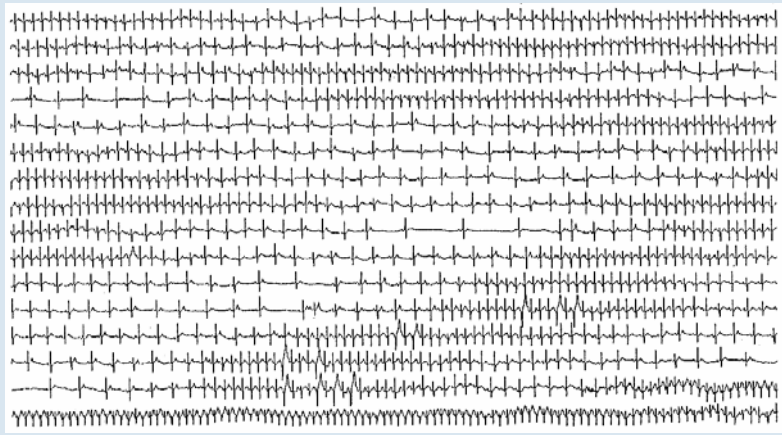
- Analysis of respiratory signals (airflow) for the diagnosis of sleep related breathing disorders (e.g. insp. flow limitation).
  → Development of algorithm started, implementation in Grid in development.

- Analysis of EMG (Electromyogram) for the diagnosis of movement disorders during sleep (e.g. RLS).
  → Development of algorithm finished – validation pending.

Sub-Project: Clinical Research

Analysis of ECG - Rationale





Healthy Subject

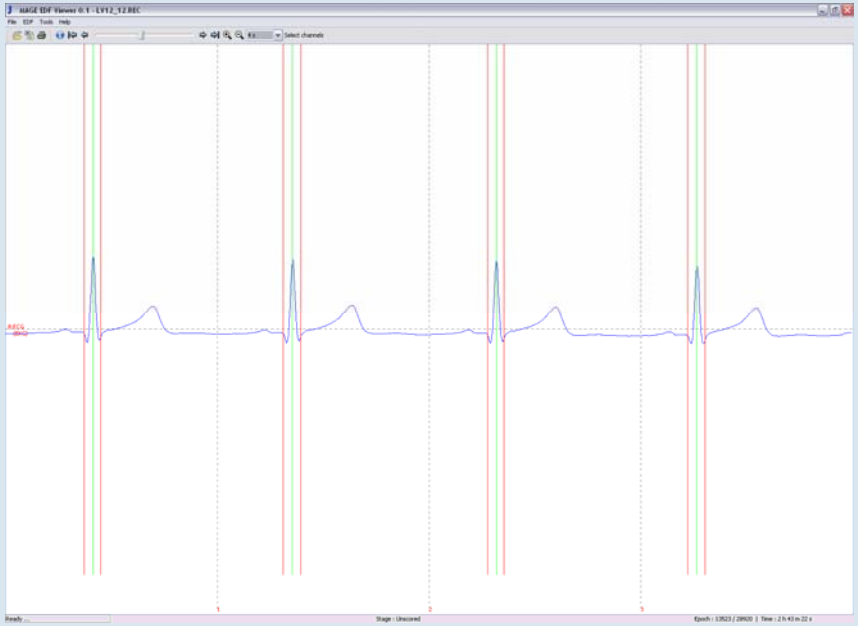
Patient with OSAS

The heart rate shows characteristic cyclic variations in patients with OSAS

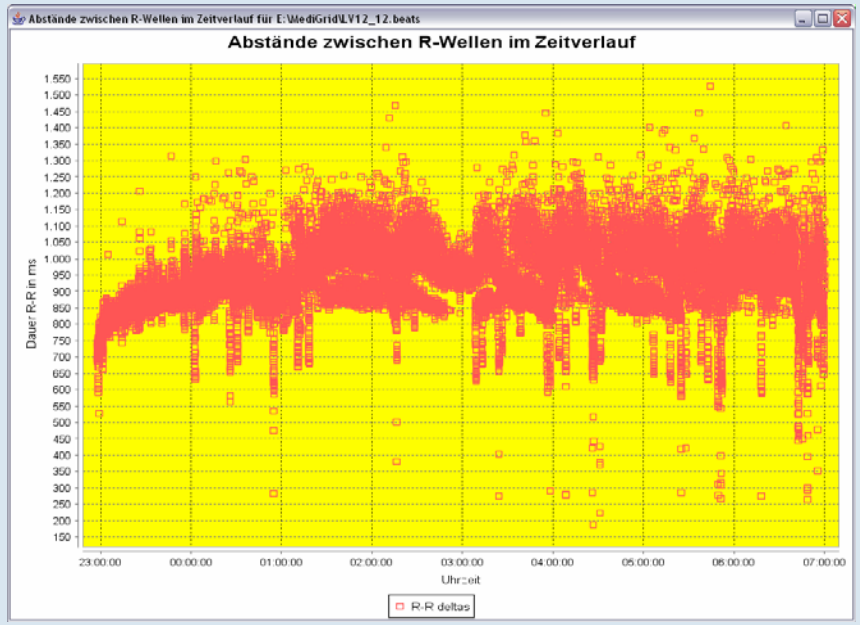
Penzel T et al. IEEE Trans. Biomed. Eng. 50: 1143-1151 (2003)
Stein PK et al. J. Cardiovasc. Electrophysiol. 14: 467-473 (2003)
DeChazal P et al. Physiol. Meas. 25: 967-983 (2004)

Sub-Project: Clinical Research

Grid based analysis of ECG



Automatic detection of QRS complex within ECG signal

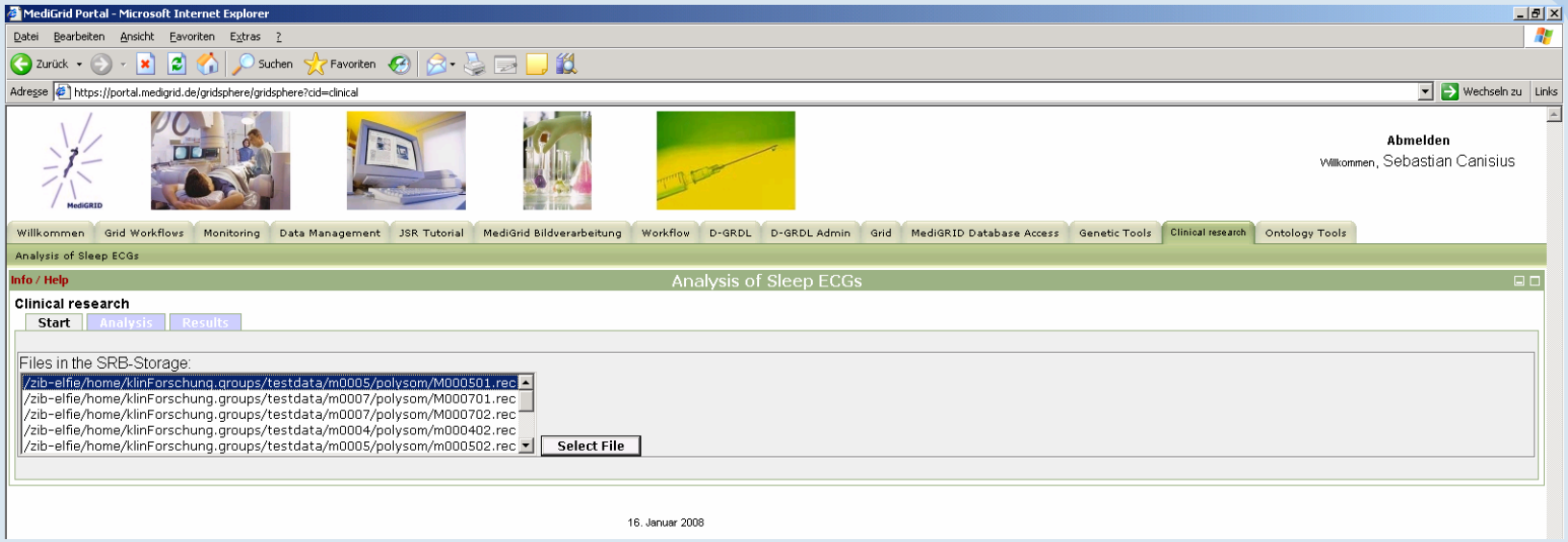


Duration of R-R-Intervals throughout the night recording (y-axis: duration of R-R, x-axis: time)

Sub-Project: Clinical Research
Current Portlet Application

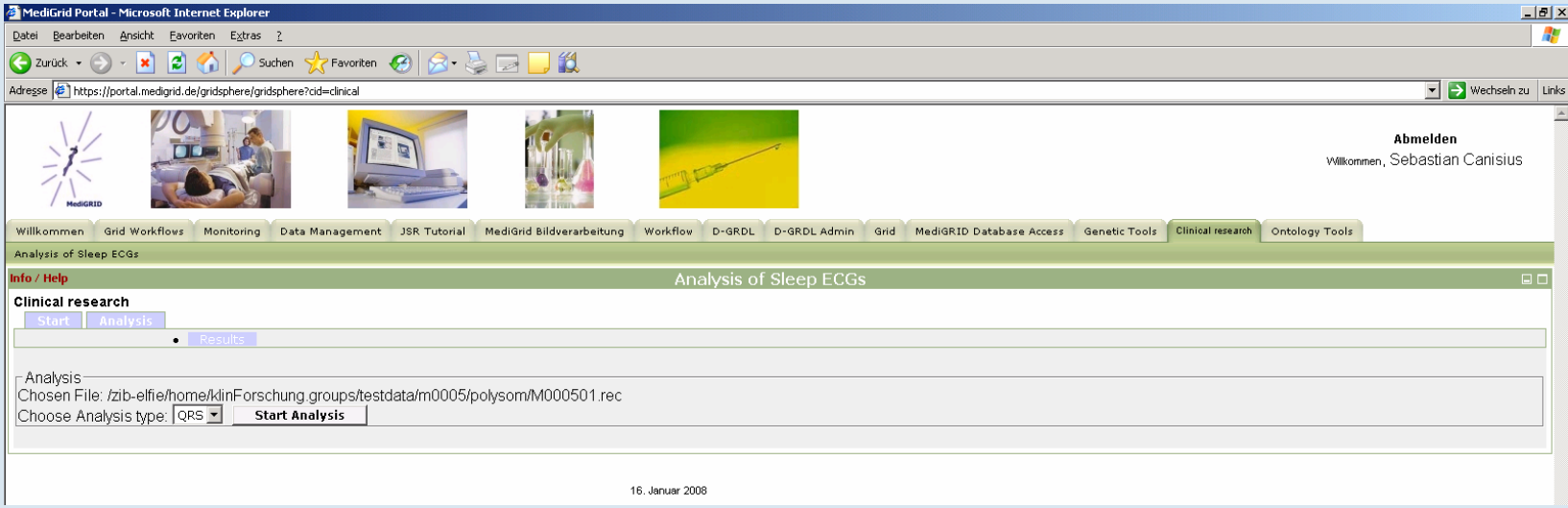
1)

Select File from SRB to be processed



2)

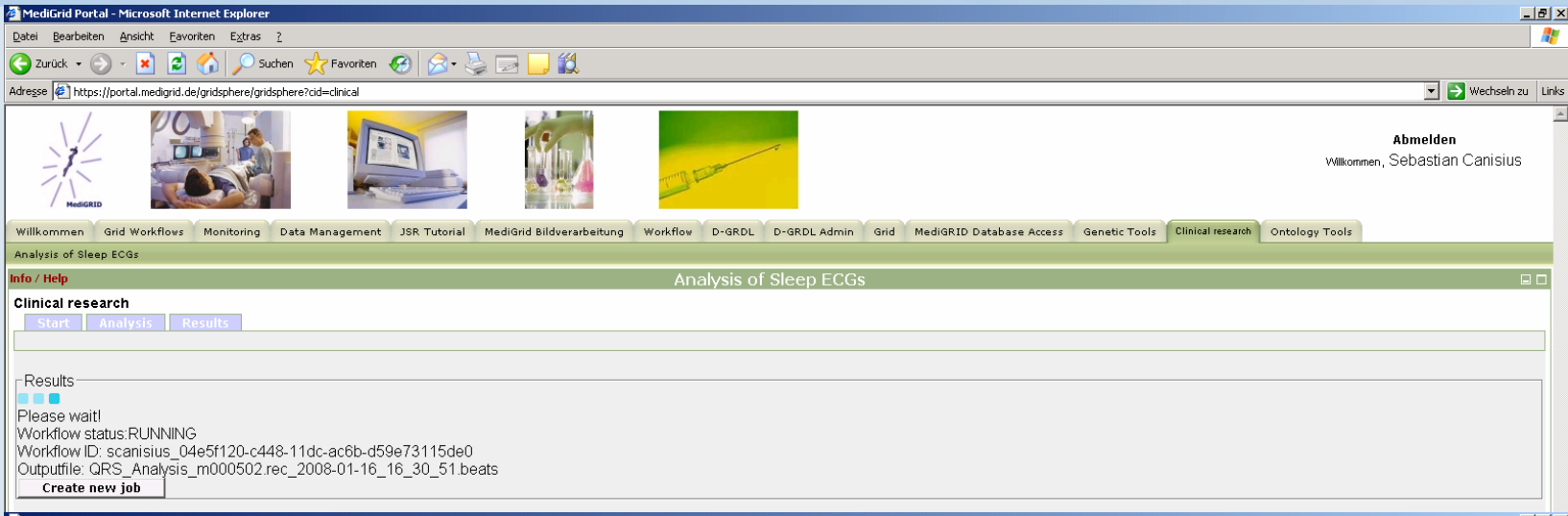
Select Analysis to be performed



Sub-Project: Clinical Research
Current Portlet Application

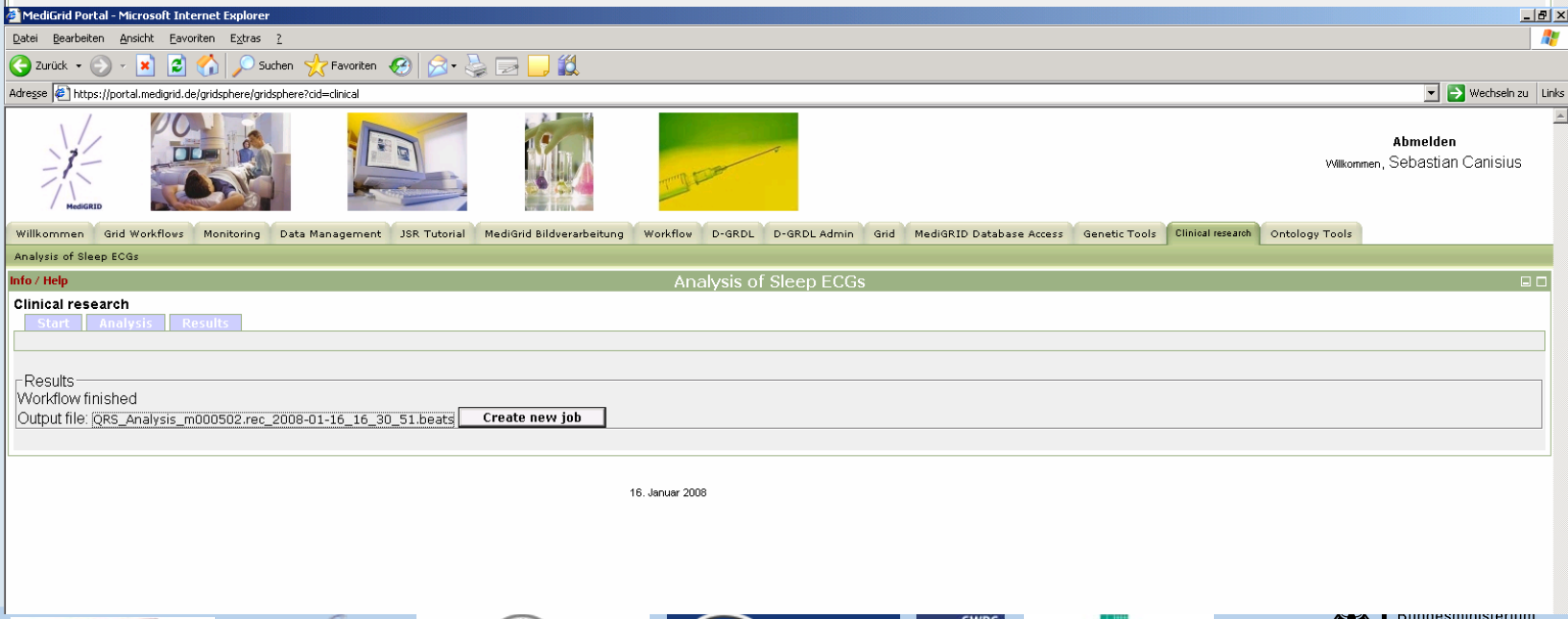
3)

Analysis is running



4)

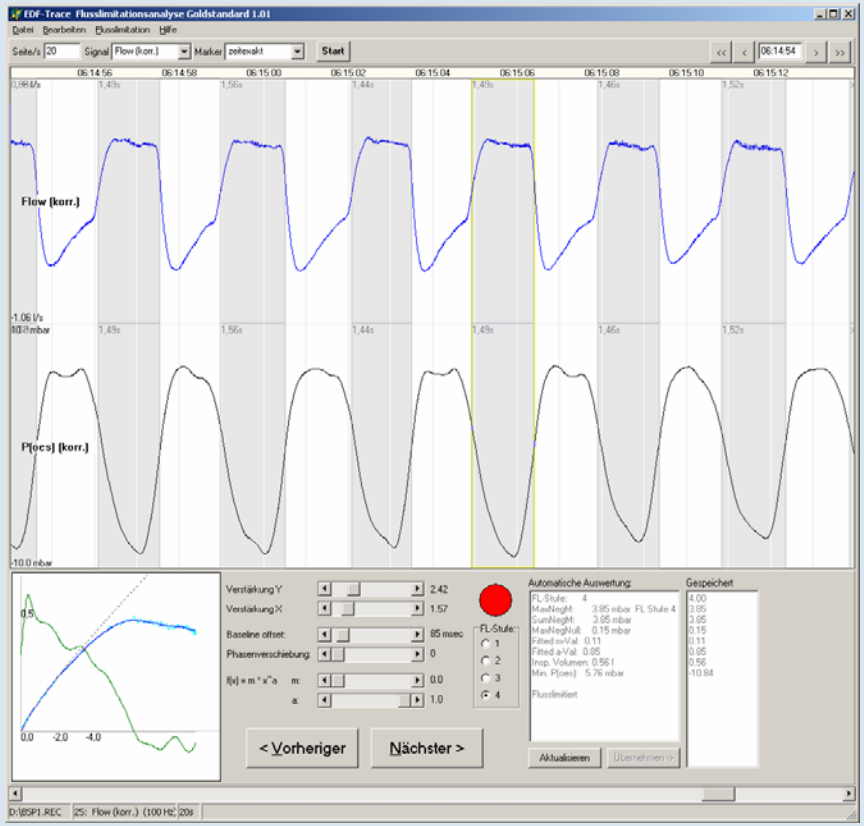
Analysis finished, results file for download



Grid based analysis of Respiratory Signals

Detection of Inspiratory Flow Limitation (IFL)

Clark SA, et al. Am J Respir Crit Care Med. 1998 Sep;158(3):713-22.



Computer based recognition of inspiration (grey background) and expiration (white background), recognition and classification of flow limitation based on respiratory airflow.

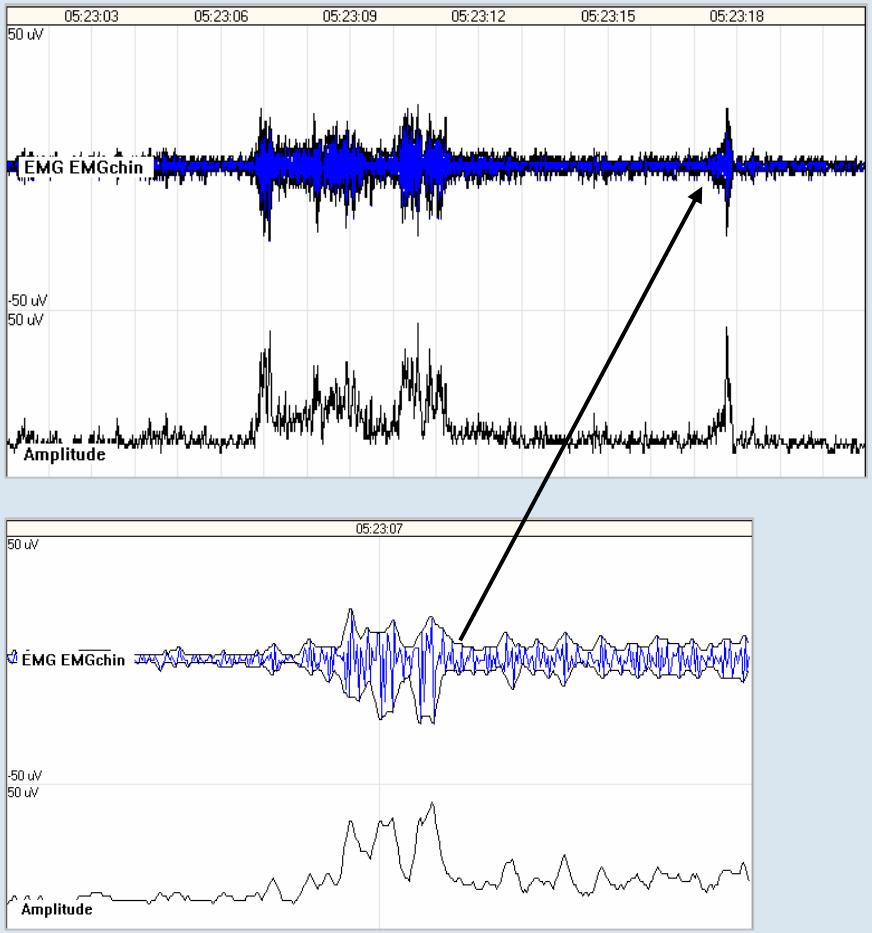
Grid based analysis of EMG

Detection of sleep related movement disorders

- Neurologic disorders RLS and narcolepsy are often accompanied by motor events in the EMG.
  - Events between 0.5 and 5 sec → RLS
  - Short twitches < 1 sec → narcolepsy
- Analysis of EMG is time consuming (app. 1-2 hours for one 8-hour recording) and requires a lot of experience.
- Motor events are counted for severity rating of RLS, detection of twitches can alleviate diagnosing narcolepsy.

Grid based analysis of EMG

Detection of sleep related movement disorders

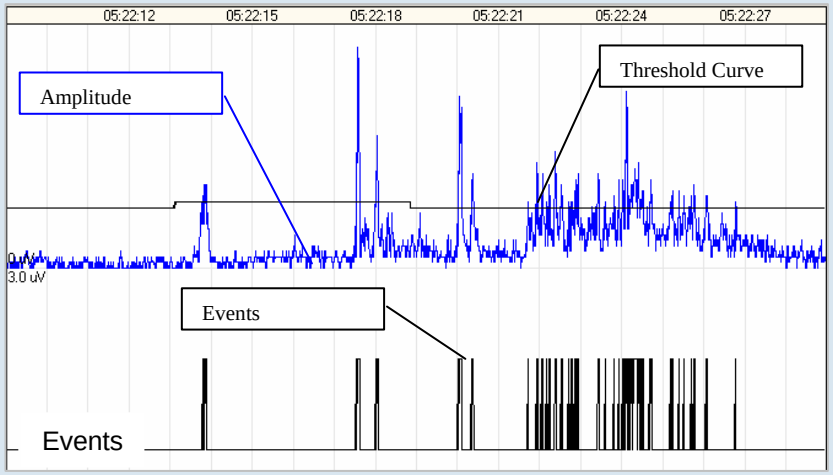


Calculation of the amplitude of the EMG signal from the difference between the upper and lower envelopes of the signal.

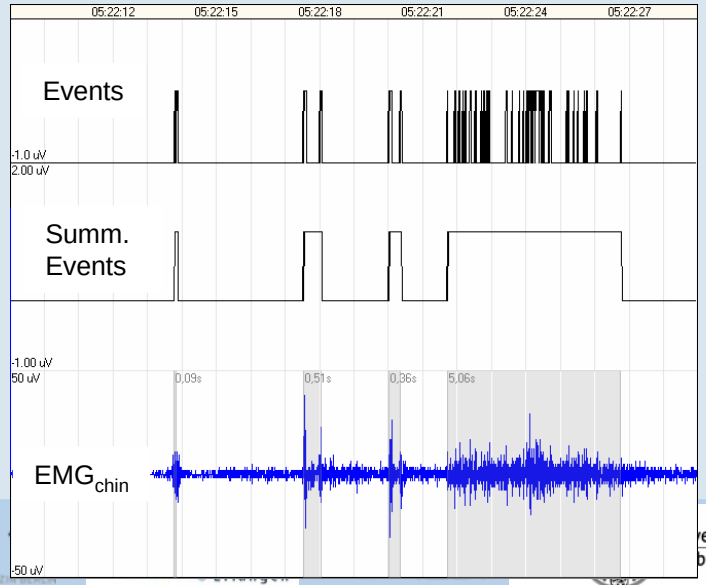
This results in the amplitude signal, which is used for the detection of events in the EMG signal.

Grid based analysis of EMG

Detection of sleep related movement disorders



Detection of muscle events by calculation of threshold curve over 200 sec. Peaks of the amplitude that are higher than the threshold curve are considered as events.



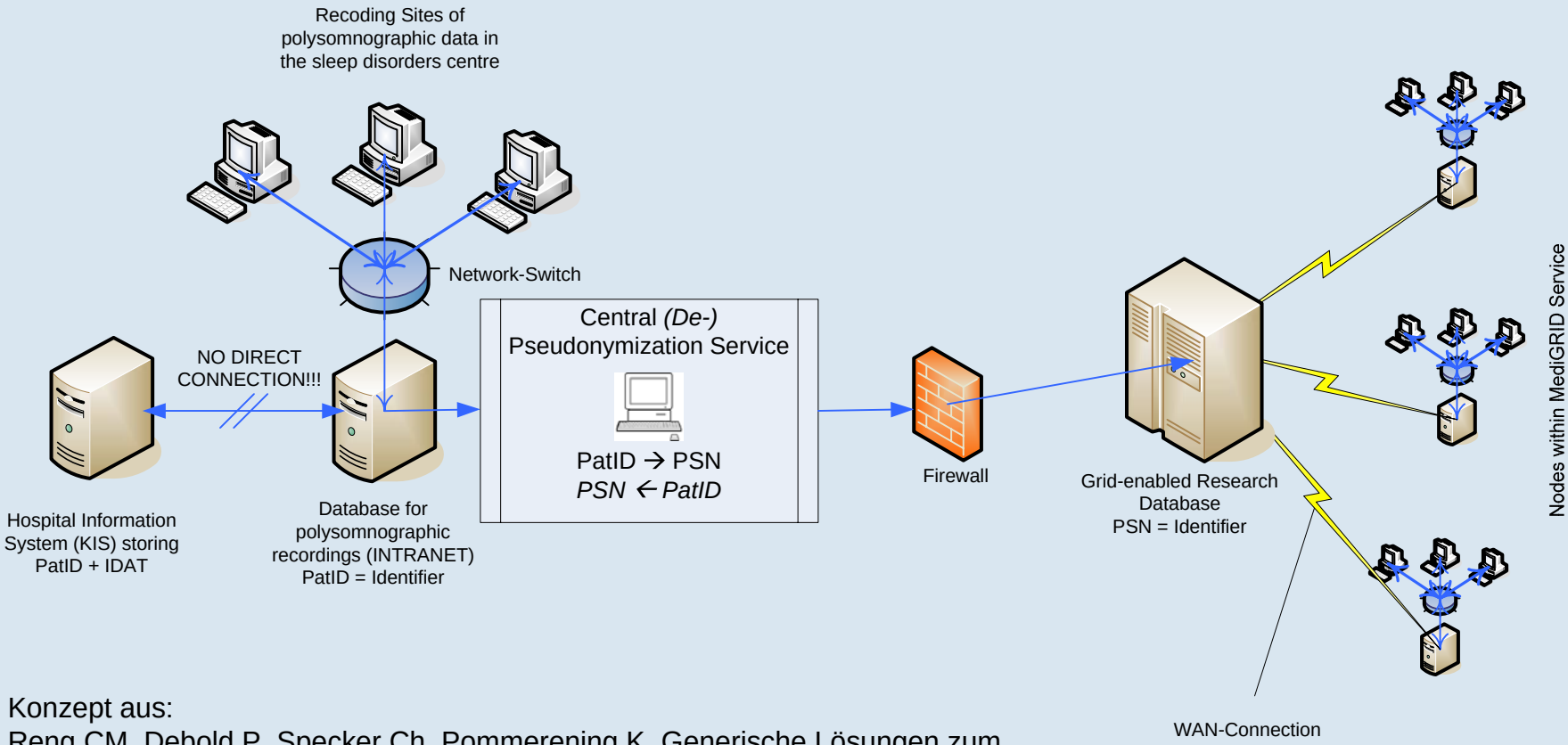
Summarization of coherent EMG events with an interval of < 1 sec between activations. This allows the differentiation between short twitches (occurring e.g. in Narcolepsy) and longer muscle activations (occurring e.g. in RLS)

Data Protection and Data Security

- Strict separation between IDAT (identification data) and MDAT (medical data) by means of separate databases
- Multi-Level Pseudonymization
- Central Pseudonymization Service
- Re-Allocation only possible for clinical sites providing data for analysis

Drepper J, Semler SC, Mohammed Y, Sax U. Aktuelle Themen des Datenschutzes und der Datensicherheit in der biomedizinischen Forschung. In: Sax U, Mohammed Y, Viezens F, Rienhoff O, editors. Grid-Computing in der biomedizinischen Forschung - Datenschutz und Datensicherheit. München: Urban&Vogel, 2006: 25-36.

Data Protection and Data Security



Konzept aus:
Reng CM, Debold P, Specker Ch, Pommerening K. Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin. Medizinisch Wissenschaftliche Verlagsgesellschaft, 2006.