



Introduction

- **Module Ontology Tools in MediGRID**
 - Responsible for integration of ontologies in MediGRID
 - Tools that help to use and access ontology information in grids
- **Usage of ontologies depends on applications in MediGRID**
- **3 Pilot classes of applications**
 - Bioinformatics
 - Imaging
 - Clinical Research
- **Linking applications with currently existing ontologies**
 - Integration of further knowledge
 - Utilization of annotation databases, sources
 - First phase: No editing/construction of new ontologies



Example of an ontology – NCI Thesaurus

- **About the thesaurus**
 - Reference Terminology for NCI
 - Broad coverage of cancer domain
 - Findings and Disorders
 - Anatomy
 - Drugs, Chemicals
 - About 38000 concepts in 20 hierarchies
- **Why**
 - Uniform conceptualization in a domain
 - Standardization, interoperability, classification
 - Enable reuse of data and information
- **Usage**
 - Annotation of medical data (images, ...)
 - Better search capabilities (categories)

NCI_Thesaurus Taxonomy

- [-] Abnormal Cell
- [-] Activity
- [-] Anatomic Structure, System, or Substance
 - [+] Body Fluid or Substance
 - [+] Body Part
 - [+] Body Region
 - [+] Body Cavity
- [-] Embryologic Structure or System
- [-] Microanatomic Structure
 - [-] Organ
 - [.] Biliary Tract
 - [.] Bladder
 - [.] Bone Marrow
 - [.] Brain
 - [+] Breast
 - [.] Bronchial Tree
 - [.] Diaphragm
 - [+] Duct
 - [.] Epididymis
 - [.] Esophagus
 - [+] Fallopian Tube
 - [.] Gall Bladder
 - [+] Gland
 - [+] Gonad
 - [.] Heart



Goals

Simple and transparent access middleware for ontologies in grids

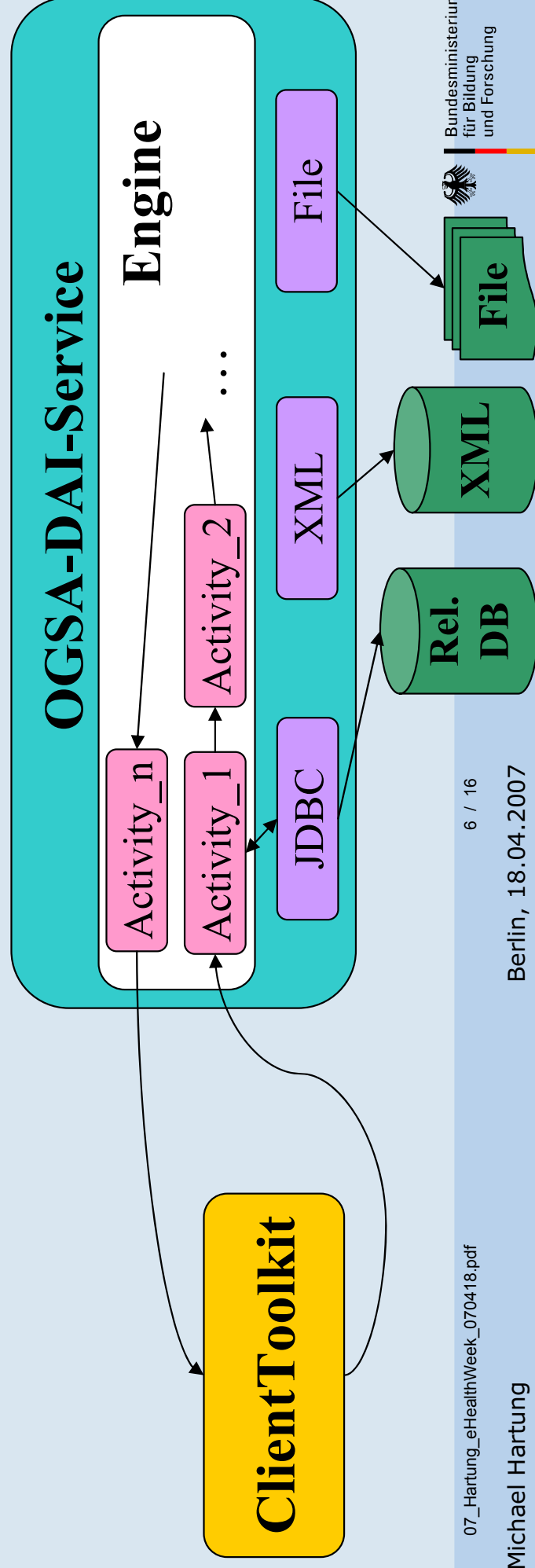
- **Functionalities**
 - Look Up of concepts/terms in available ontologies
 - Access to ontology information: definitions, synonyms, relations, cross references → uniform interface for all clients (API)
 - GUI: navigation in ontologies, graph representation of ontologies
- **Service based access to ontologies**
 - Ontology Services as middleware between applications and ontologies
 - Support for clients (MediGRID applications, other services)
 - Usage of grid: Distribution of ontologies → load balancing, failure safety
 - Extensibility: simple integration of new ontologies, extension in functionalities
 - Usage of D-Grid basis software, integration in D-Grid infrastructure



OGSA-DAI as base infrastructure

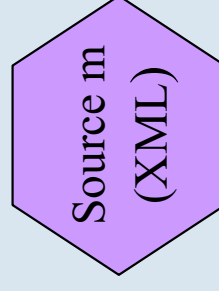
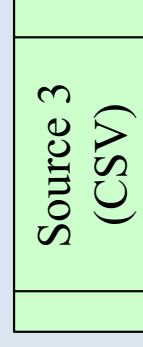
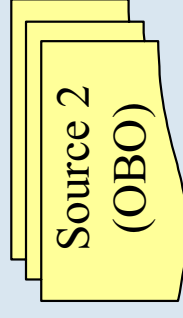
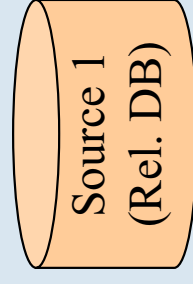
Access to data sources in grids

- **Extensible framework for “data access” and “integration”**
 - Access to distributed and heterogeneous data sources in grids
 - Web service based access system (WSRF, Globus Toolkit 4)
 - Extensibility as an important feature (activities, resources)
 - Client Toolkit API for service interaction

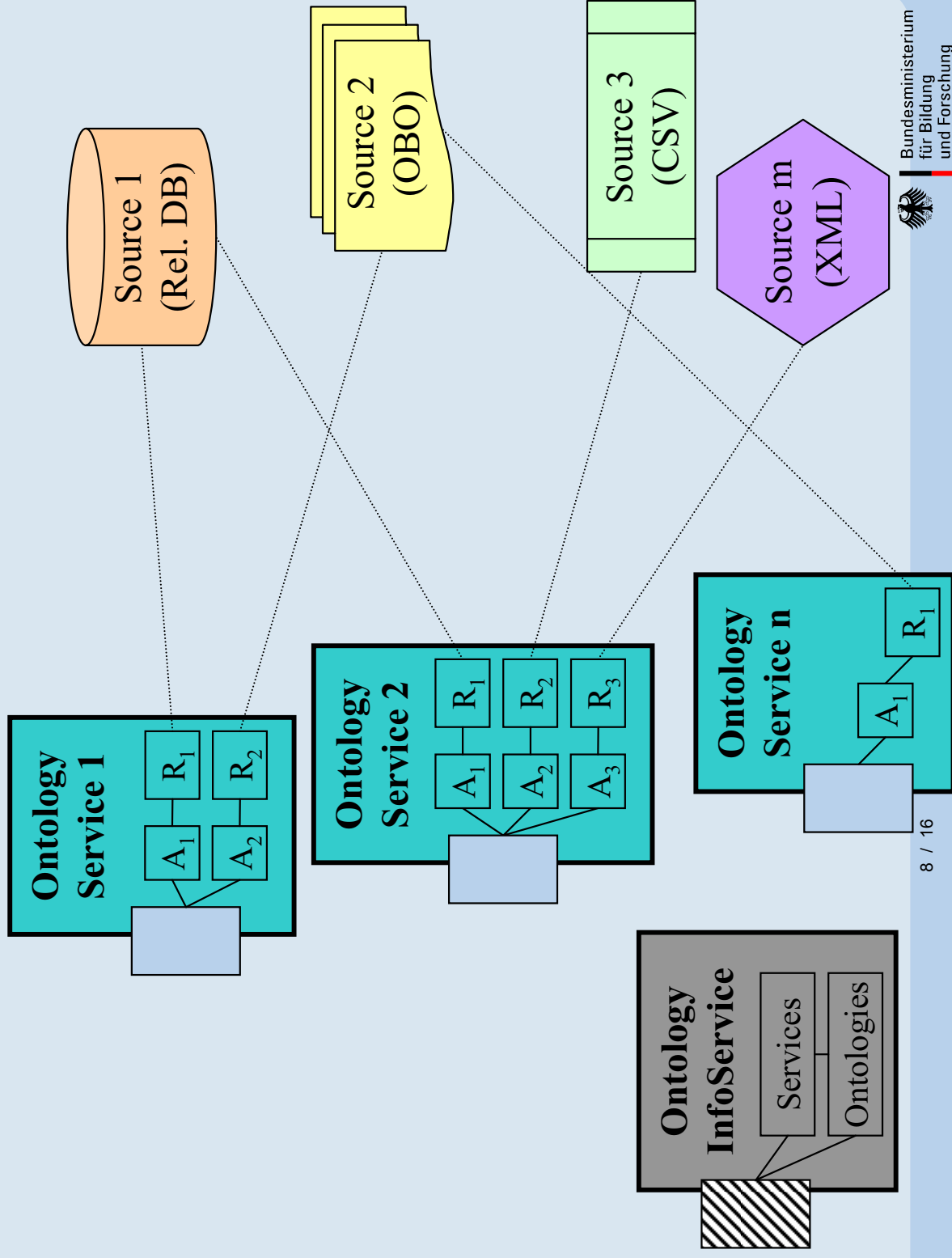


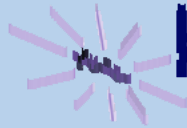


Ontology Access Middleware

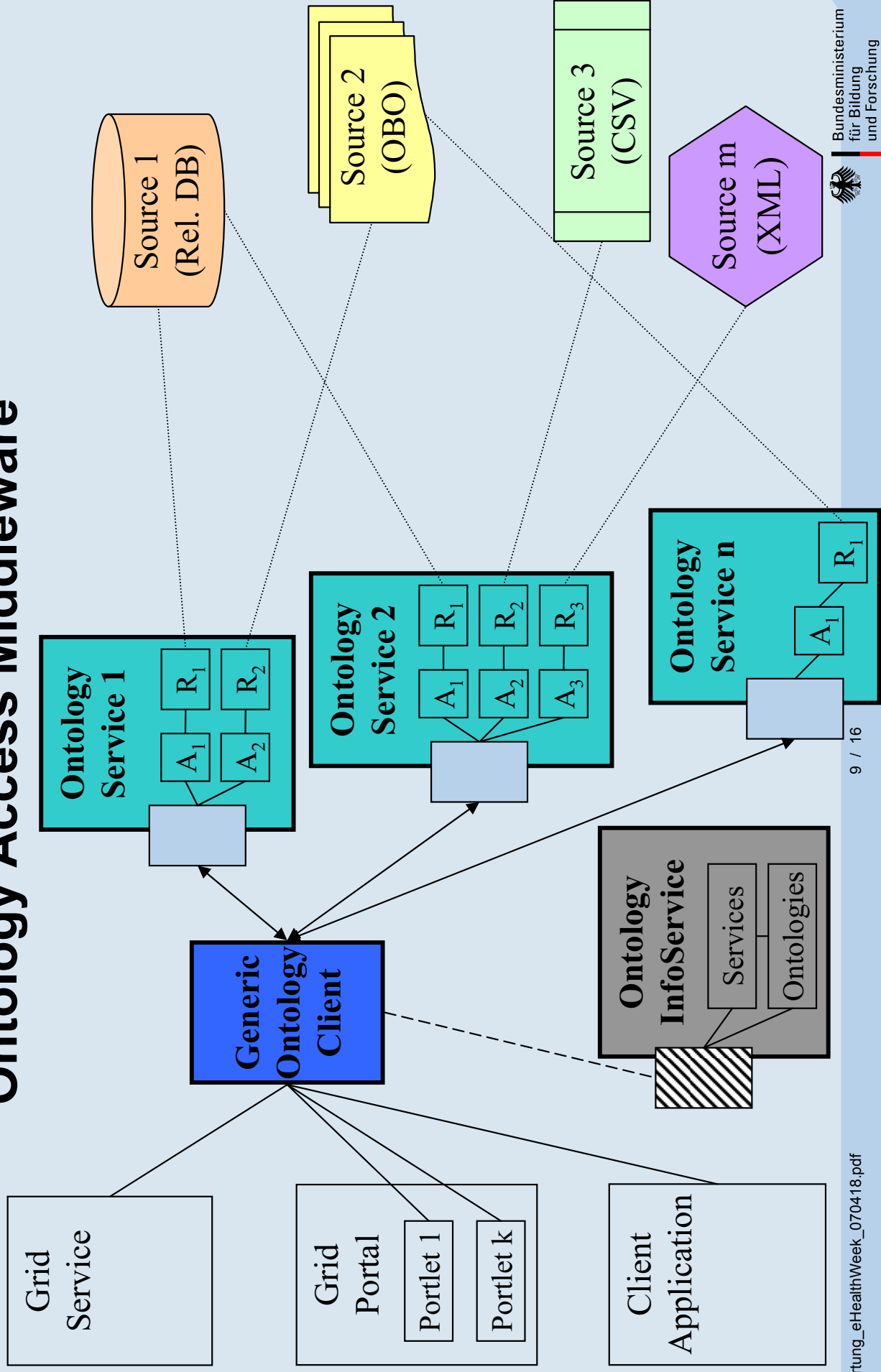


Ontology Access Middleware





Ontology Access Middleware

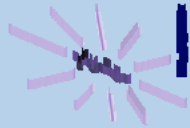




Applications using the middleware

MediGRID portal as central entry to grid applications

- **Advantages**
 - User friendly interaction with grid resources
 - Central access without knowledge about resource locations etc.
 - Portlets as re-usable software components
 - Graphical representation of results, workspace for users
 - **Ontology Access Middleware in the MediGRID portal**
 - Ontology Client interacts with distributed Ontology Services
 - Client is integrated in application portlets
1. Central Ontology Look Up Service
 2. AUGUSTUS gene prediction tool
 3. SNPSelection



Ontology Look Up Service (I)

Central ontology applications

The screenshot shows a web browser window with the URL www.grid.ac/portal/ontology. The page title is "MultiOntologyAccessPortlet". At the top, it says "Welcome, Michael Hartung" and "Logout". A navigation menu includes: Welcome, Grid, Bioinformatics, Genetic Tools, Ontology Admin, MedGrid Database Access, D-GRDL, D-GRDL Admin, Monitoring, MedGrid Imaging Applications, MultiOntologyAccess, GridWiki, Gene Ontology, NCI Thesaurus, MultiOntologyAccessPortlet, Info / Help, Search, Show Results, OntologyInfo, ResourceInfo. The main content area has a search bar with the text "ProteinProteinInteractionOntology" and a "Search" button. Below the search bar, a dropdown menu shows "Available ontologies:" with the following list: GeneOntology, NCI-Thesaurus, HumanDiseaseOntology, ProteinProteinInteractionOntology, CellTypeOntology. There is a "Refresh" button at the bottom left of the dropdown. The date "March 23, 2007" is visible in the bottom right corner of the page content.

Ontologies to Look Up





Ontology Look Up Service (II)

Results for "heart" in "GeneOntology":

[regulation of heart contraction](#)
 heart looping
 negative regulation of heart contraction rate in baroreceptor response to increased blood pressure
 decreased strength of heart contraction during baroreceptor response to increased blood pressure
 positive regulation of heart contraction rate in baroreceptor response to decreased blood pressure

Show term

Search Results

ID	GO:0008016
Name	regulation of heart contraction
Definition	Any process that modulates the frequency, rate or extent of heart contraction.
Comment	
Synonyms	<ul style="list-style-type: none"> regulation of heart rate
Hierarchy	<p>GO:0051239 regulation of organismal physiological process</p> <p>GO:0008016 regulation of heart contraction</p> <p>GO:0002026 cardiac inotropy</p> <p>GO:0002027 cardiac chronotropy</p> <p>GO:0045822 negative regulation of heart contraction</p> <p>GO:0045823 positive regulation of heart contraction</p>
Parents	<ul style="list-style-type: none"> GO:0051239; regulation of organismal physiological process (IS_A)
Children	<ul style="list-style-type: none"> GO:0002026; cardiac inotropy (IS_A) GO:0002027; cardiac chronotropy (IS_A) GO:0045822; negative regulation of heart contraction (IS_A) GO:0045823; positive regulation of heart contraction (IS_A)
X-References	<ul style="list-style-type: none"> GOC:go_curators TIGR_TIGRFAMS:TIGR01294_phospholamban

Information about ontology concepts





Ontology Look Up Service (III)

The screenshot shows a web browser window with the URL `MultiOntologyAccessPortlet`. The page title is "Überblick Ontologie- und DB-Ressourcen". A search bar contains "engbackta". Below the search bar, a list of ontology services is displayed, each with a green checkmark indicating it is available. A red location pin is placed on a map of Germany, pointing to Halle (Saale). A green callout box points to the list of services, and another green callout box points to the map.

Monitoring of Ontology Services and corresponding ontologies

Locations

Host: <http://buell.izbi.uni-leipzig.de:8081>
 Service: wsrf/services/ogsadai DataService

- GeneOntology - GO@ducati.izbi.uni-leipzig.de - ✓
- NCI-Thesaurus - NCIThesaurusResource - ✓
- GeneOntology - GO@buell.izbi.uni-leipzig.de - ✓
- PathwayOntology - PathwayOntologyResource - ✓
- HumanDiseaseOntology - DiseaseOntologyResource - ✓
- ProteinProteinInteractionOntology - ProteinProteinInteractionOntologyResource - ✓
- CellTypeOntology - CellOntologyResource - ✓
- MultipleAlignmentOntology - MultipleAlignmentOntologyResource - ✓
- ProteinModificationOntology - ProteinModificationOntologyResource - ✓
- FlyBaseControlledVocabulary - FlyBaseControlledVocabularyResource - ✓
- MammalianPhenotypeOntology - MammalianPhenotypeOntologyResource - ✓
- MeSH - MeSHResource - ✓
- SequenceOntology - SequenceOntologyResource - ✓
- MobyObjectsOntology - MobyObjectsOntologyResource - ✓





AUGUSTUS

The screenshot shows the MedIGRID Augustus web interface. The top navigation bar includes links for Welcome, Grid, Bioinformatics, Genetic Tools, Ontology Access, MedIGRID Database Access, D-GRDL, D-GRDL Admin, Monitoring, and MedIGRID Imaging Applications. The main content area displays the following information:

back
 Executed command:
 augustus --species=human --strand=both --singlestrand=true --genemodel=partial --codingseq=on --alternatives=false sequence.file

Result file:
[Download Result file](#)

Transformed result files:
[Graph of the result: png-File](#)
[PDF of the result: pdf-File](#)
[Postscript of the result: ps-File](#)

Output:
 # This output was generated with AUGUSTUS (version 1.8.2).
 # AUGUSTUS is a gene prediction tool for eukaryotes written by Mario Stanke (mstanke@gmdg.de).
 # Please cite: Mario Stanke and Stephan Waack (2003) "Gene prediction with a hidden Markov model and a new intron submodel", Bioinformatics, Vol. 19 Suppl. 2, ii215-ii225
 # No extrinsic information on sequences given.
 # Initialising the parameters ...
 # Looks like sequence.file is in fasta format.
 # We have hints for 0 sequence and for 0 of the sequences in the input set.
 # ----- prediction on sequence number 1 (length = 6483, name = HSACKI10) -----
 # Constraints/Hints:
 # (none)
 # Predicted genes for sequence number 1 on both strands
 # (none)
 ### gene gl
 HSACKI10 AUGUSTUS 1674 5499 1 + . gl
 HSACKI10 AUGUSTUS 1674 5499 0.35 + . gl.tl
 HSACKI10 AUGUSTUS 1674 1676 . + 0 transcript_id "gl.tl"; gene_id "gl";
 HSACKI10 AUGUSTUS 2300 0.55 + 0 transcript_id "gl.tl"; gene_id "gl";
 HSACKI10 AUGUSTUS 1674 2300 0.55 + 0 transcript_id "gl.tl"; gene_id "gl";

GFF result file

Information about GFF region names

- Linkage between result file and SequenceOntology (SO)

