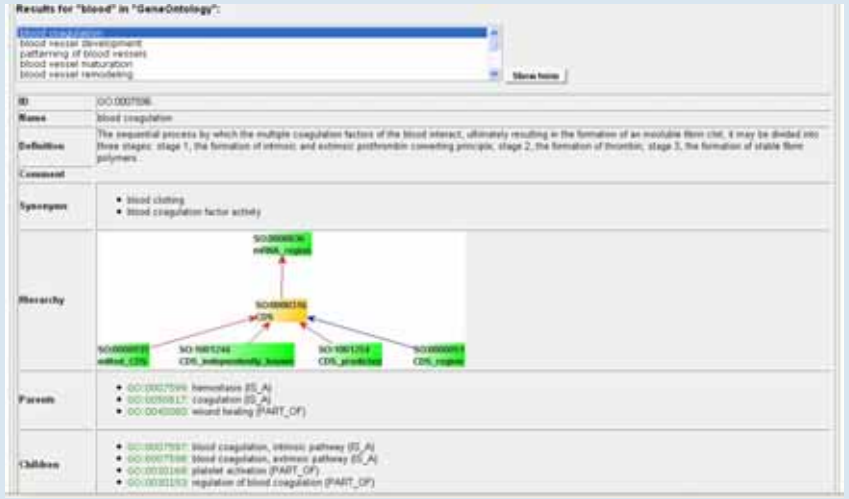
IZBI
Universität Leipzig

UNIVERSITÄT LEIPZIG

D-GRID



A Grid Middleware for Ontology Access

German e-Science Conference

Baden Baden, May 4, 2007

Michael Hartung, Erhard Rahm

Interdisciplinary Centre for Bioinformatics, Leipzig

MediGRID

Bundesministerium
für Bildung
und Forschung

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

Bundesministerium für Bildung und Forschung

Motivation

Key problem: Heterogeneity between existing ontologies

- **Yet no existing ontology access system in grids**
- **Base information**
  - Terms, concepts with name, definition, synonyms
  - Relations between concepts of an ontology
  - Cross references to other ontologies, data sources
- **Different source formats of ontologies**
  - Relational databases: GeneOntology
  - Structured flat files: NCI Thesaurus
  - Standardizations: OBO-Ontologies
  - Web Standards: OWL, RDF
- **Different design of ontology information**
  - Modeling of relationships
  - Presentation of synonyms, properties

OBO
open biomedical ontologies

National Cancer Institute

FMA
University of Washington
School of Medicine

Bundesministerium
für Bildung
und Forschung

Goals

Simple and transparent access middleware for ontologies in grids

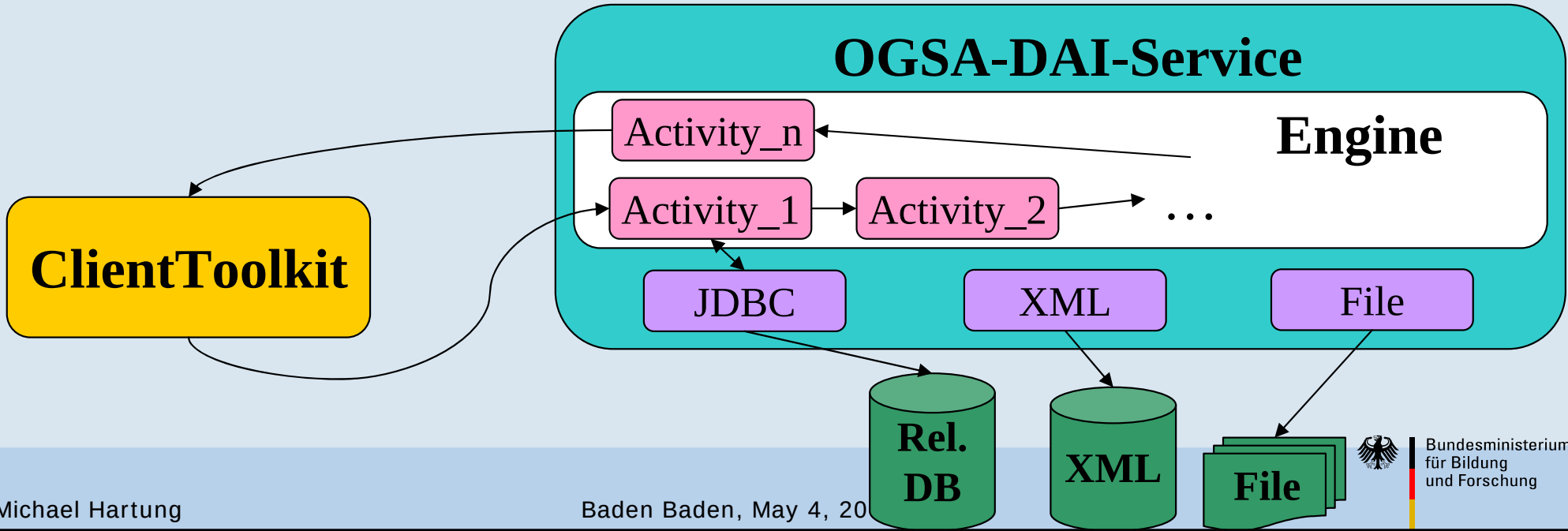
- **Functionalities**
  - Look Up of concepts/terms in available ontologies
  - Access to ontology information: definitions, synonyms, relations, cross references → uniform interface for all clients (API)
  - GUI: navigation in ontologies, graph representation of ontologies

- **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

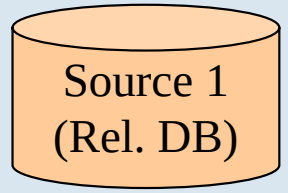
Access to data sources in grids

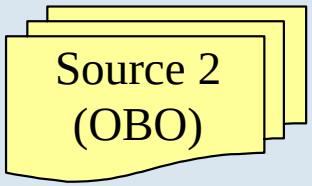
- **Extensible framework for “data access” and “integration”**
  - Access to distributed and heterogeneous data sources in grids
  - Web service based access system (WSRF, Globus Toolkit 4)
  - Extensibility as an important feature (activities, resources)
  - Client Toolkit API for service interaction

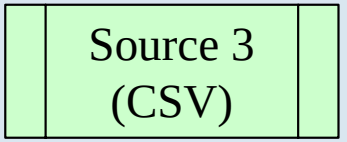


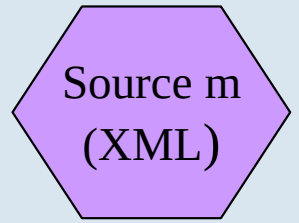
Bundesministerium für Bildung und Forschung

Ontology Access Middleware



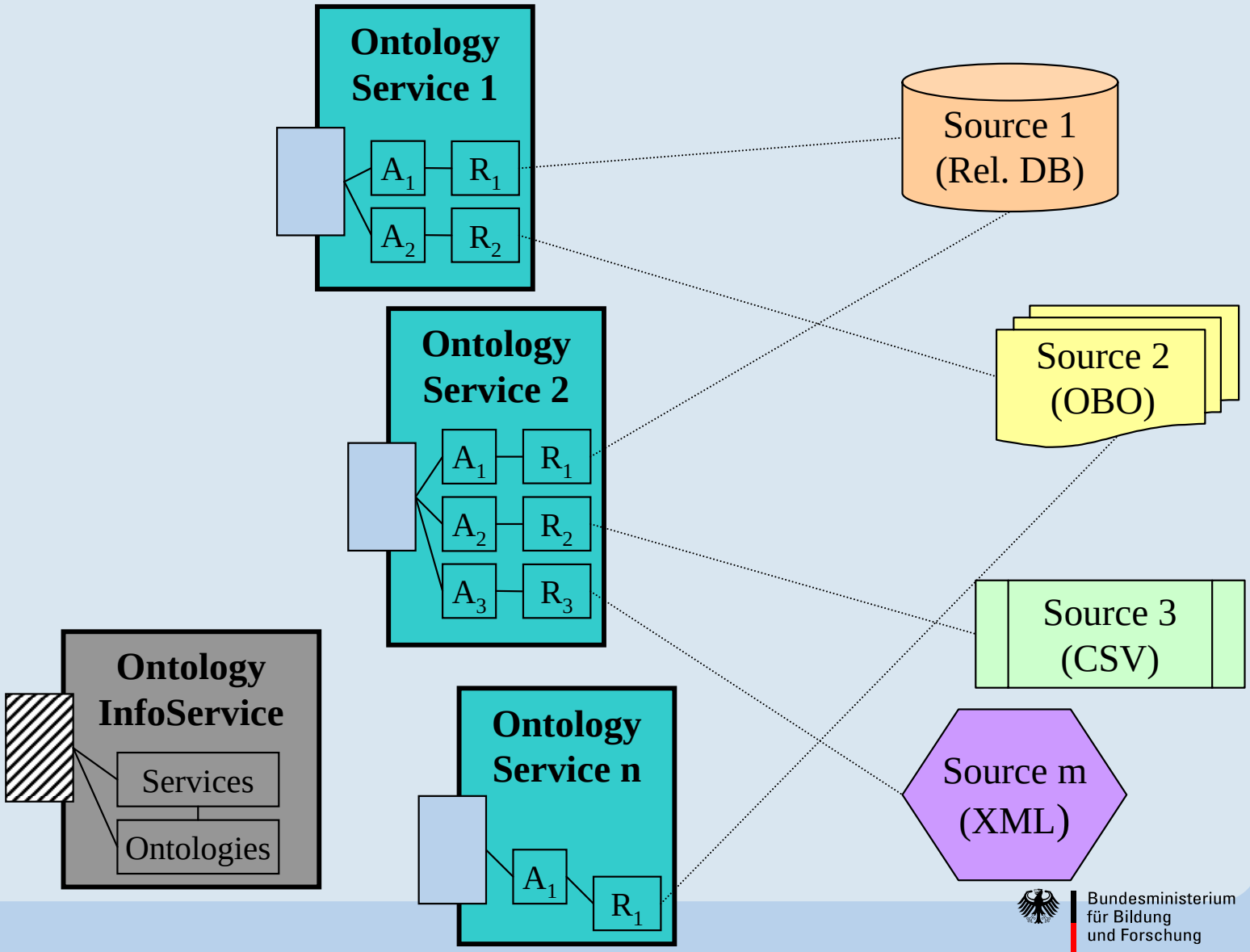






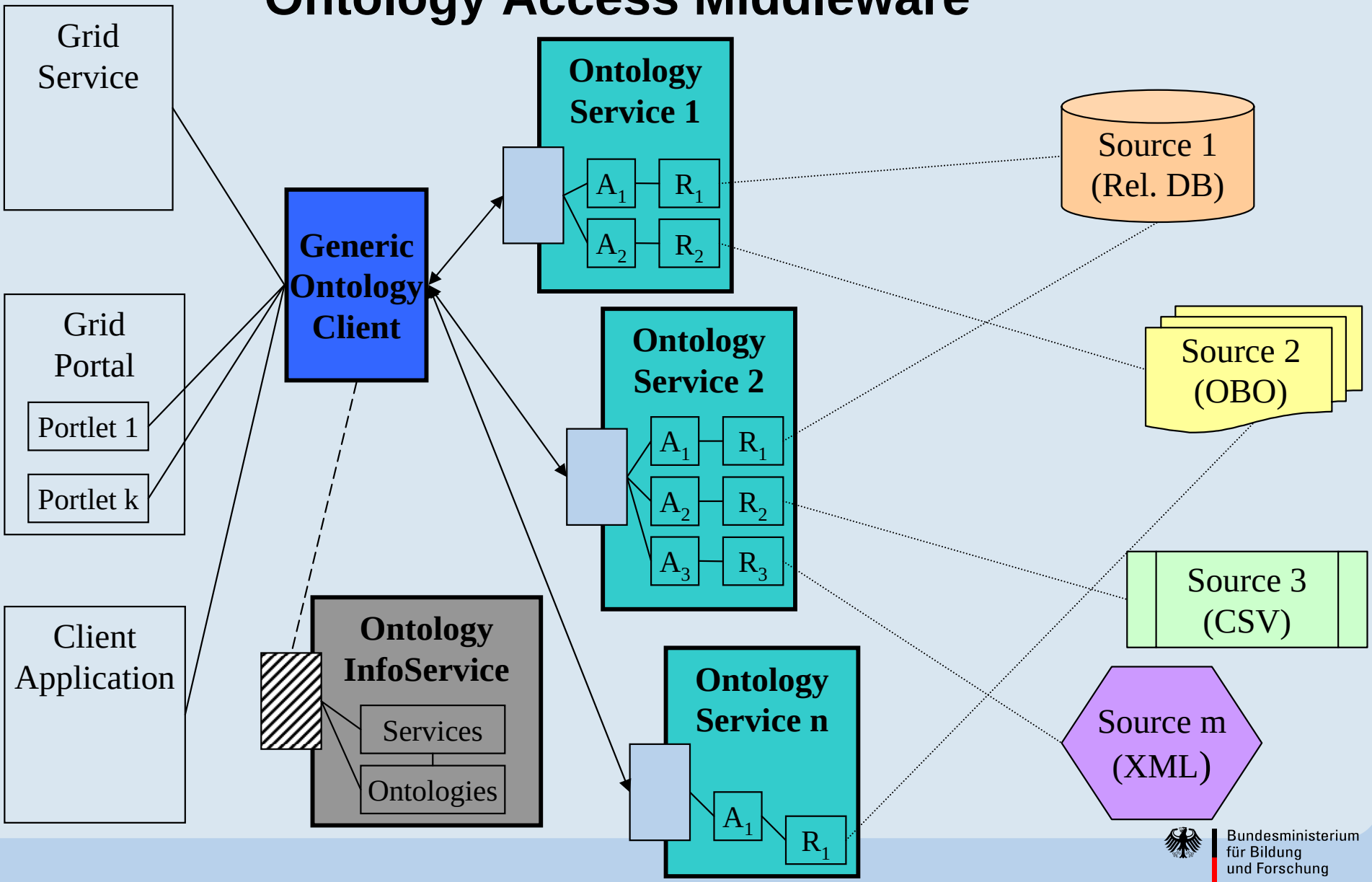
Bundesministerium
für Bildung
und Forschung

Ontology Access Middleware



Bundesministerium für Bildung und Forschung

Ontology Access Middleware

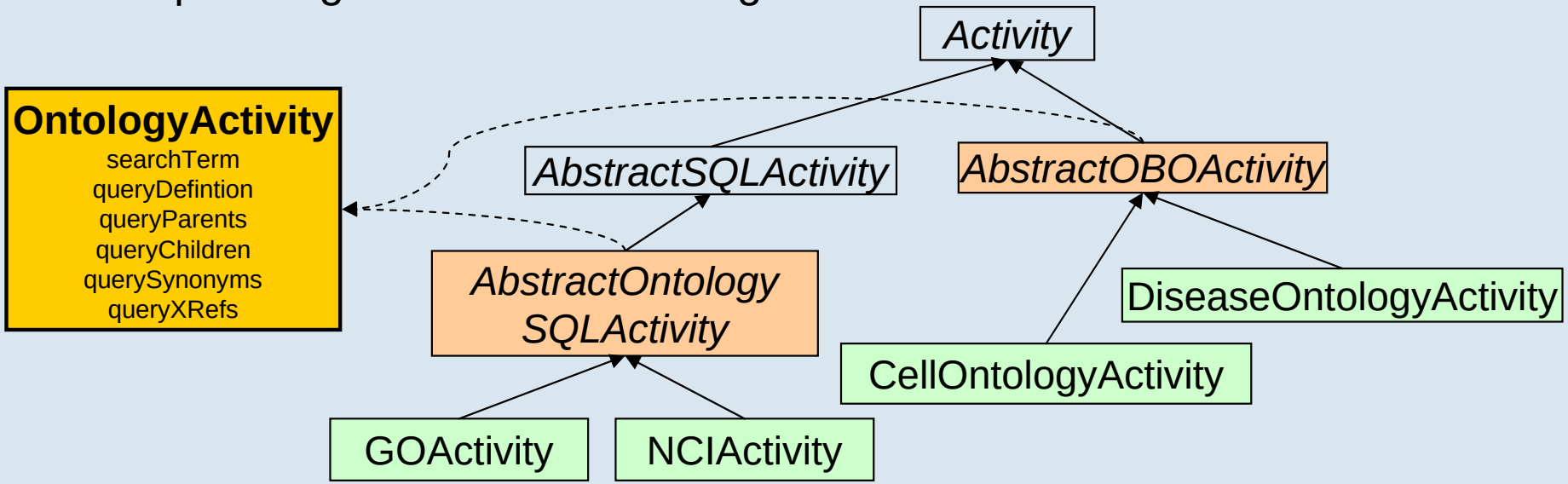


Bundesministerium für Bildung und Forschung

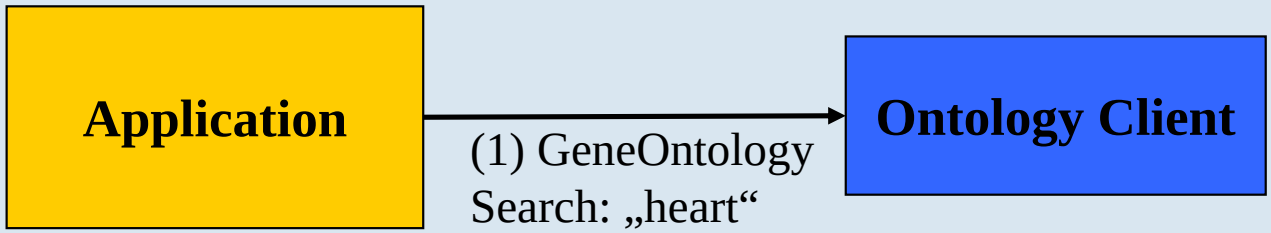
Extensions to OGSA-DAI

Ontology services as extension of OGSA-DAI data services

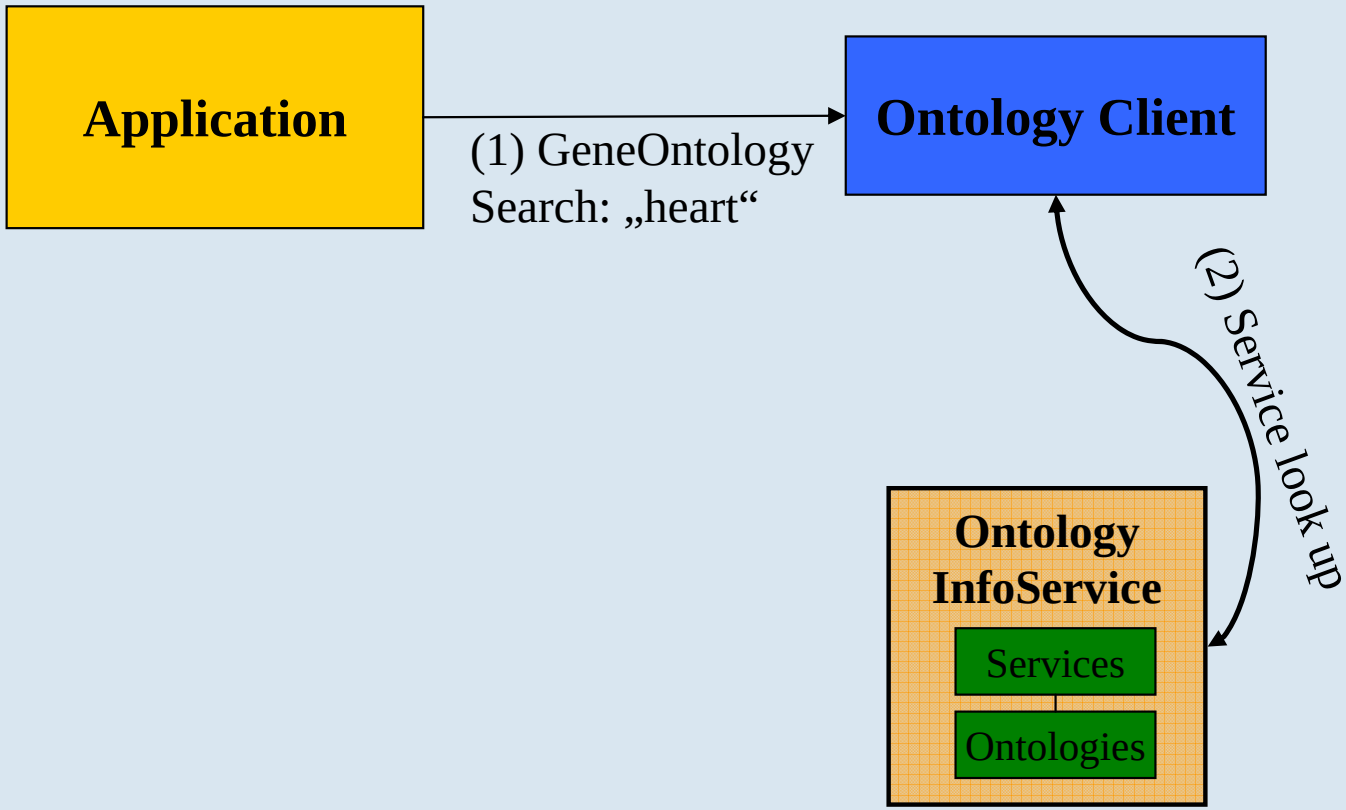
- **Activities**
  - Special ontology activities as wrapper for ontology sources
  - Central ontology interface: OntologyActivity
  - Hierarchy of ontology activities
  - Simple integration of new ontologies



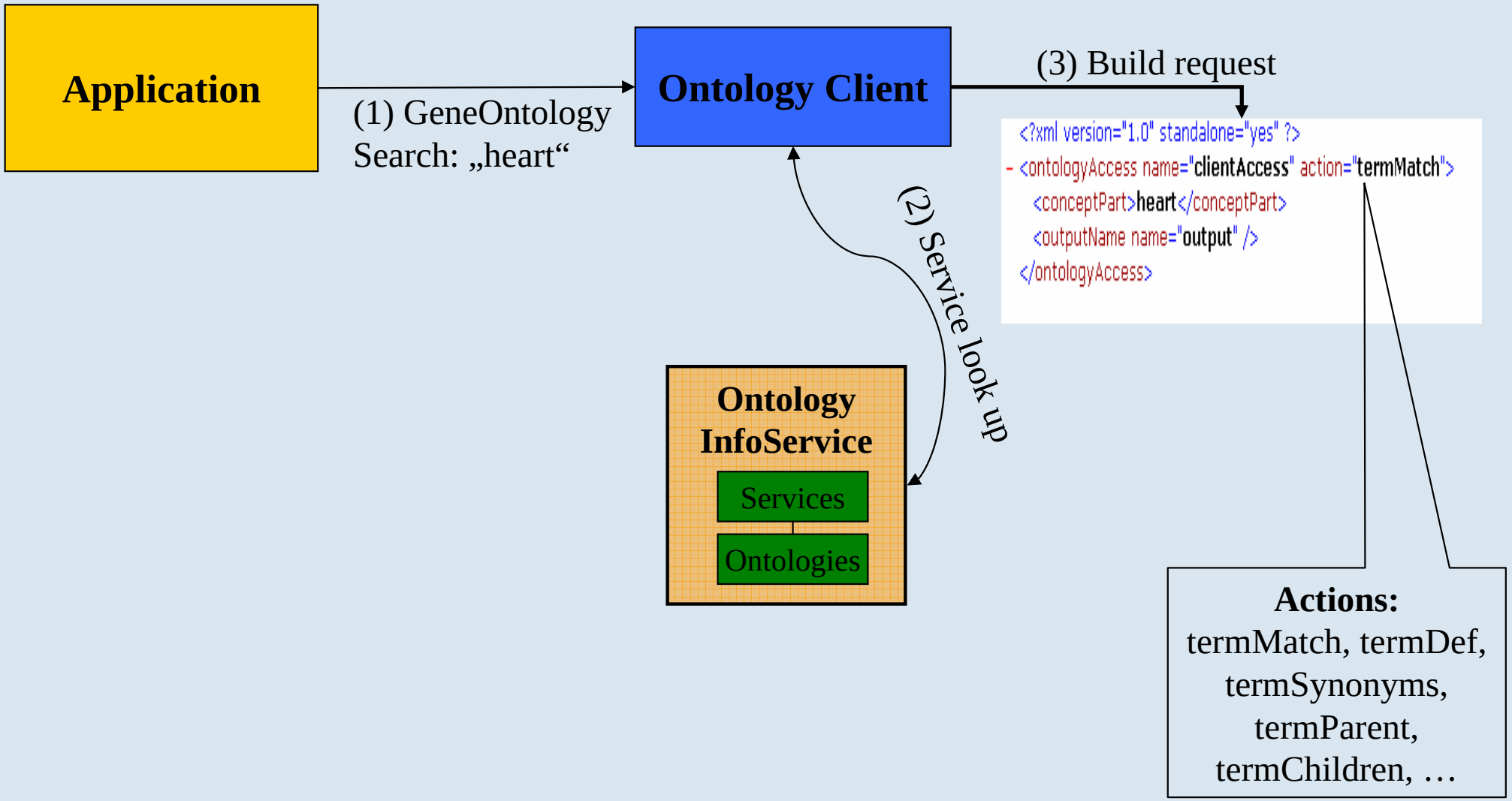
Bundesministerium für Bildung und Forschung

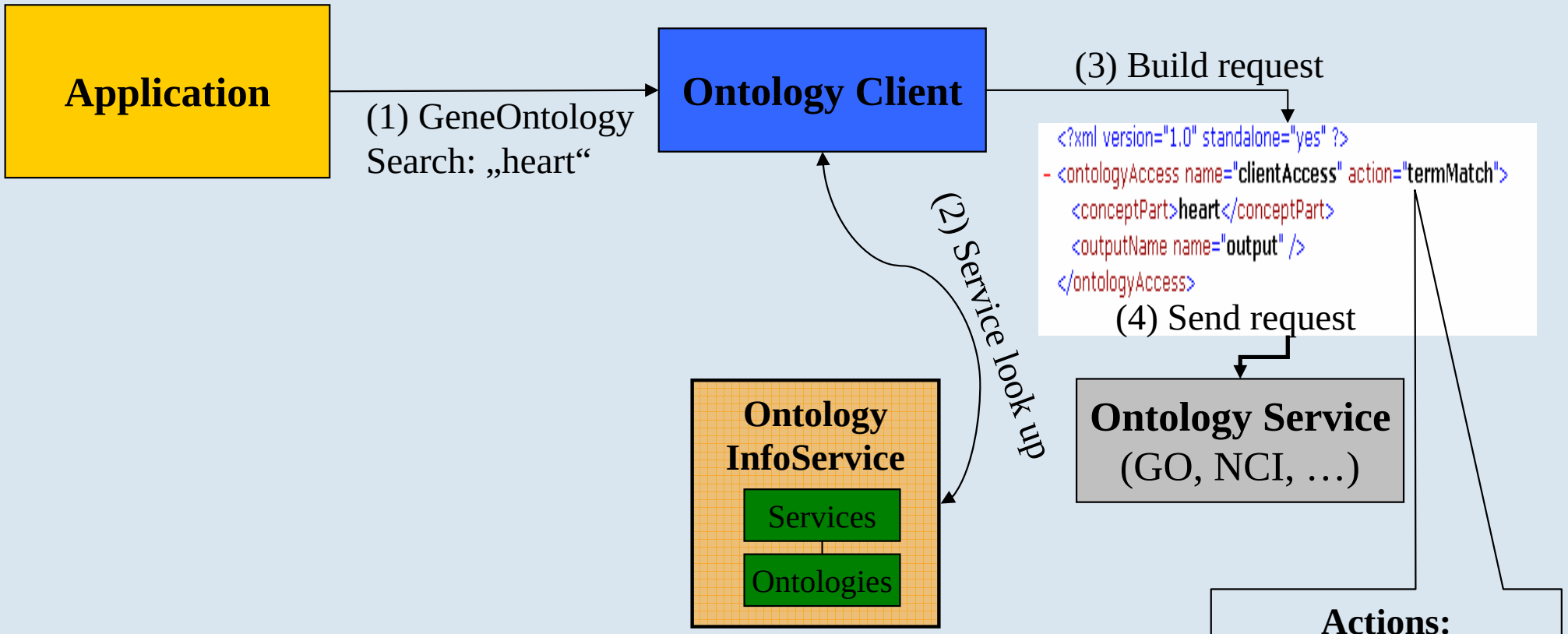


Bundesministerium
für Bildung
und Forschung

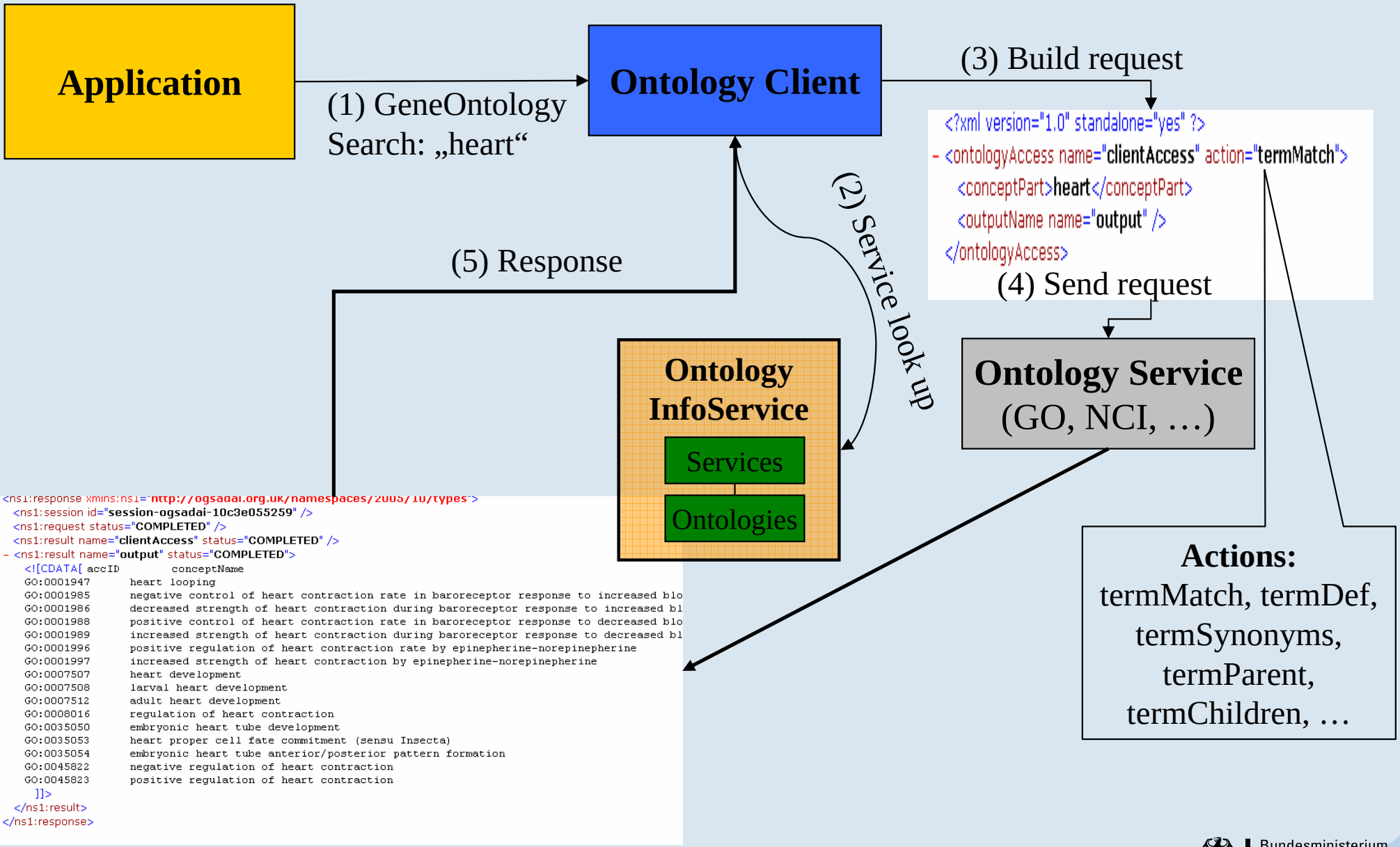


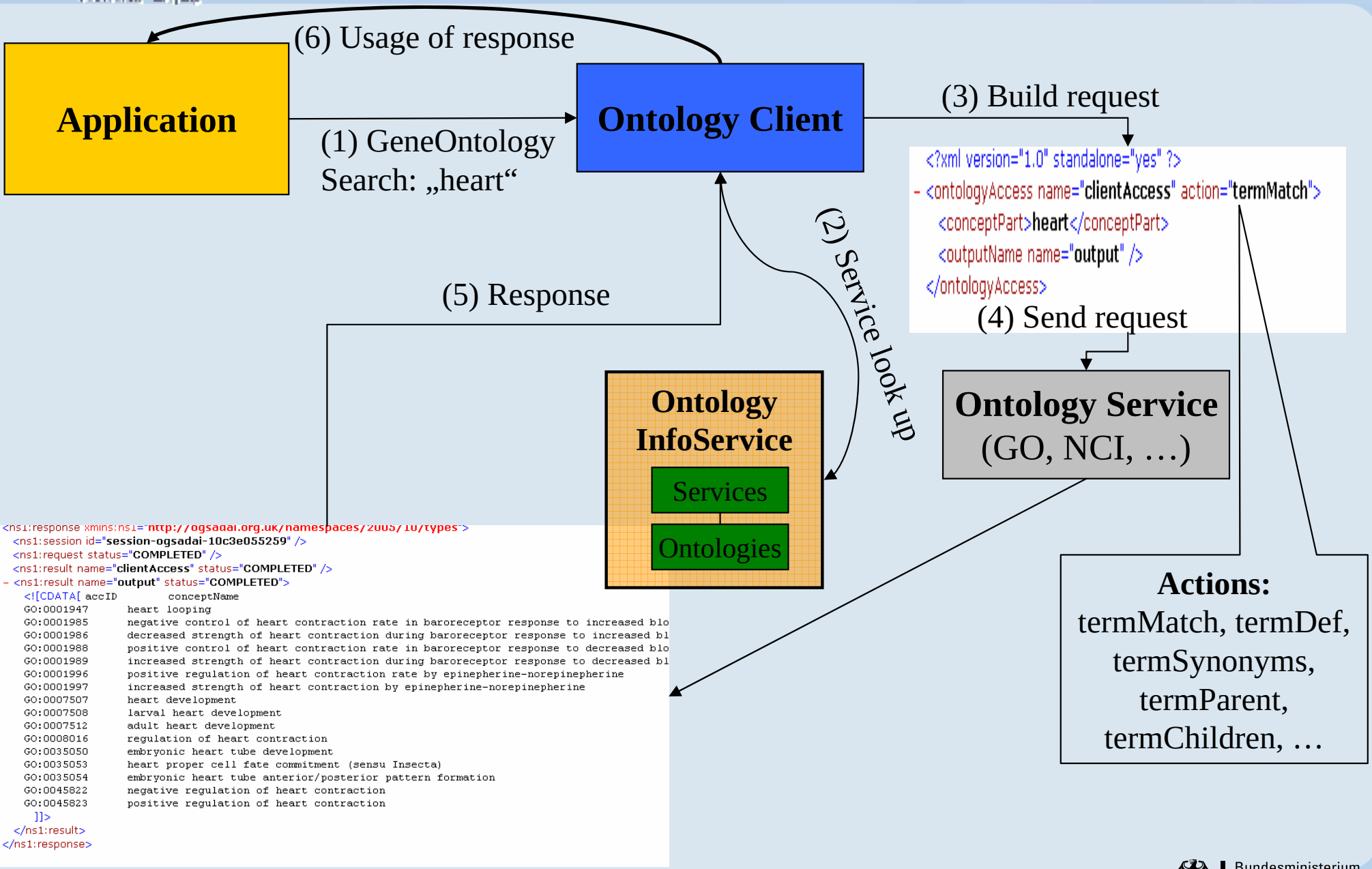
Bundesministerium
für Bildung
und Forschung


Application
(1) GeneOntology Search: „heart“
Ontology Client
(3) Build request
<?xml version="1.0" standalone="yes" ?>
- <ontologyAccess name="clientAccess" action="termMatch">
<conceptPart>heart</conceptPart>
<outputName name="output" />
</ontologyAccess>
(2) Service look up
Ontology InfoService
Services
Ontologies
Actions:
termMatch, termDef, termSynonyms, termParent, termChildren, …



Bundesministerium für Bildung und Forschung


Application
(1) GeneOntology Search: „heart“
Ontology Client
(3) Build request
<?xml version="1.0" standalone="yes" ?>
- <ontologyAccess name="clientAccess" action="termMatch">
<conceptPart>heart</conceptPart>
<outputName name="output" />
</ontologyAccess>
(4) Send request
(2) Service look up
(5) Response
Ontology InfoService
Services
Ontologies
Ontology Service (GO, NCI, …)
- <ns1:response xmlns:ns1="http://ogsadai.org.uk/namespaces/2005/10/types">
<ns1:session id="session-ogsadai-10c3e055259" />
<ns1:request status="COMPLETED" />
<ns1:result name="clientAccess" status="COMPLETED" />
- <ns1:result name="output" status="COMPLETED">
<![CDATA[accID conceptName
GO:0001947 heart looping
GO:0001985 negative control of heart contraction rate in baroreceptor response to increased blo
GO:0001986 decreased strength of heart contraction during baroreceptor response to increased bl
GO:0001988 positive control of heart contraction rate in baroreceptor response to decreased blo
GO:0001989 increased strength of heart contraction during baroreceptor response to decreased bl
GO:0001996 positive regulation of heart contraction rate by epinepherine-norepinepherine
GO:0001997 increased strength of heart contraction by epinepherine-norepinepherine
GO:0007507 heart development
GO:0007508 larval heart development
GO:0007512 adult heart development
GO:0008016 regulation of heart contraction
GO:0035050 embryonic heart tube development
GO:0035053 heart proper cell fate commitment (sensu Insecta)
GO:0035054 embryonic heart tube anterior/posterior pattern formation
GO:0045822 negative regulation of heart contraction
GO:0045823 positive regulation of heart contraction
]]>
</ns1:result>
</ns1:response>
Actions:
termMatch, termDef, termSynonyms, termParent, termChildren, …


(6) Usage of response
Application
(1) GeneOntology Search: „heart“
Ontology Client
(3) Build request
<?xml version="1.0" standalone="yes" ?>
- <ontologyAccess name="clientAccess" action="termMatch">
<conceptPart>heart</conceptPart>
<outputName name="output" />
</ontologyAccess>
(4) Send request
(2) Service look up
(5) Response
Ontology InfoService
Services
Ontologies
Ontology Service (GO, NCI, …)
- <ns1:response xmlns:ns1="http://ogsadai.org.uk/namespaces/2005/10/types">
<ns1:session id="session-ogsadai-10c3e055259" />
<ns1:request status="COMPLETED" />
<ns1:result name="clientAccess" status="COMPLETED" />
- <ns1:result name="output" status="COMPLETED">
<![CDATA[accID conceptName
GO:0001947 heart looping
GO:0001985 negative control of heart contraction rate in baroreceptor response to increased blo
GO:0001986 decreased strength of heart contraction during baroreceptor response to increased bl
GO:0001988 positive control of heart contraction rate in baroreceptor response to decreased blo
GO:0001989 increased strength of heart contraction during baroreceptor response to decreased bl
GO:0001996 positive regulation of heart contraction rate by epinepherine-norepinepherine
GO:0001997 increased strength of heart contraction by epinepherine-norepinepherine
GO:0007507 heart development
GO:0007508 larval heart development
GO:0007512 adult heart development
GO:0008016 regulation of heart contraction
GO:0035050 embryonic heart tube development
GO:0035053 heart proper cell fate commitment (sensu Insecta)
GO:0035054 embryonic heart tube anterior/posterior pattern formation
GO:0045822 negative regulation of heart contraction
GO:0045823 positive regulation of heart contraction
]]>
</ns1:result>
</ns1:response>
Actions:
termMatch, termDef, termSynonyms, termParent, termChildren, …

Applications using the middleware

MediGRID portal as central entry to grid applications

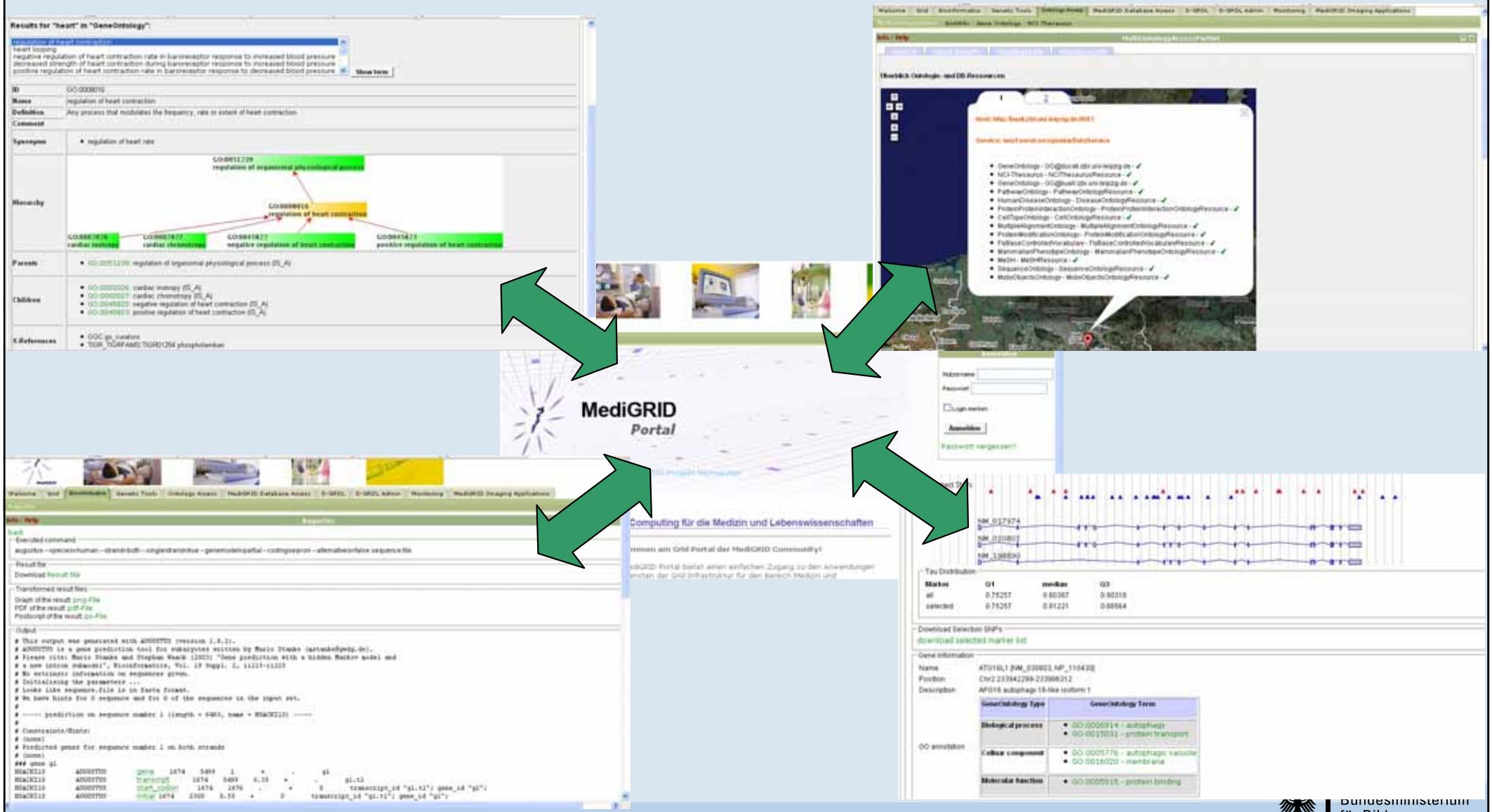
- **Advantages**
  - User friendly interaction with grid resources
  - Central access without knowledge about resource locations etc.
  - Portlets as re-usable software components
  - Graphical representation of results, workspace for users

- **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

Bundesministerium für Bildung und Forschung

Applications using the middleware (II)



Bundesministerium
für Bildung
und Forschung

Summary and Future

- **Service-based middleware to integrate ontologies in grids**
  - Based von grid standards (GT4, OGSA-DAI)
  - Distribution of ontologies, simple and uniform access to ontologies in grids
  - Integration in the central portal of MediGRID
  - Usage in different MediGRID applications

- **Future topics**
  - Integration of further ontologies for upcoming MediGRID applications
  - Matching of ontologies
  - Wiki-like system for collaborative editing and development of domain-specific ontologies (e.g. D-Grid initiative)