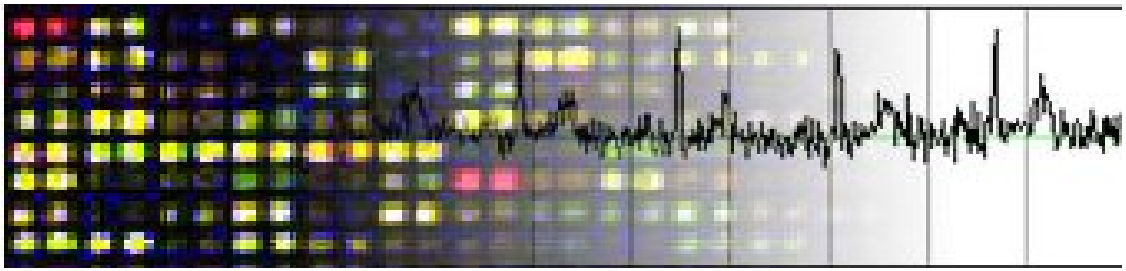
20.453J / 2.771J / HST.958J Biomedical Information Technology
Fall 2008



Biomedical Information Technology

2.771J 20.453J HST.958J SMA5304 Fall 2008

Lecture 21 October 2008

Creating Databases from Ontologies

C. Forbes Dewey, Jr.
Massachusetts Institute of Technology
Cambridge, MA USA

Here's our agenda

- Review of database technology
  - *History*
  - *What is their role?*
  - *What do they do well?*
  - *What do they do poorly?*
- Ontologies
  - *Why are they desirable?*
  - *What are their advantages?*
  - *What are their weaknesses?*
- Relationship between DB and Ontologies
- Exploiting the combination - OWLdb

Review of relational database technology

Started in mid-80's

- *Composed of tables with rows and columns*
- *Minimal language to construct and query them: SQL*
- *Generally fixed and unchanged relationships*

Java-based interface JDBC a real breakthrough; common interface for all flavors. Can write programs!!!

Databases don't talk to each other

Stove-piping, warehousing, federation

Updates, security, reliability, performance

OWL Ontologies

Describe meaning

- *OWL is a language (set of relationships) that can be extended with additional definitions*
- *OWL is written in RDF, is machine-readable, and can be parsed and combined*

Representing databases as ontologies

- *Can represent a database schema by an ontology*
- *Will discuss state-of-the art programs to do the conversion*
- *If we can convert, can have the best of both environments*

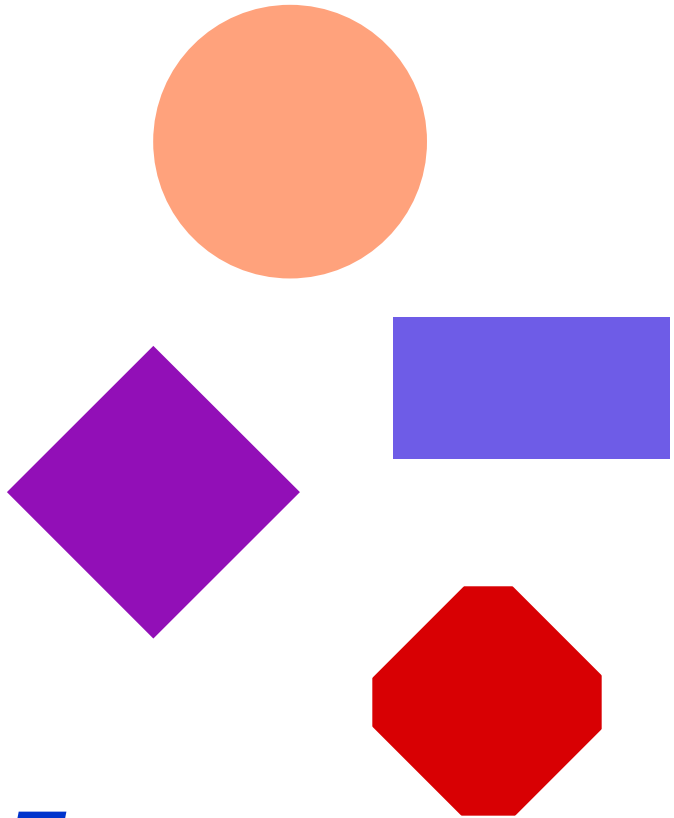
OWL – Web Ontology Language

- *An expressive and uniform way of defining meaning for terms used to transmit data and relationships*
- *Can be used for many key purposes*
  - *Guarantee that two definitions are the same*
  - *Discover that two terms are synonymous*
  - *Encode complete object descriptions*
  - *Define unambiguous database schema*
- *Comes in multiple flavors*
  - *OWL Lite – OWL DL – OWL Full*

http://www.w3.org/TR/2004/REC-owl-features-20040210/#s2.1

Owl Structure

- Classes
- Properties
- Types
- Meta-Data



Owl is written in RDF

OWL – Web Ontology Language (2)

RDF Schema Features:

Class (Thing, Nothing)
rdfs:subClassOf
rdf:Property
rdfs:subPropertyOf
rdfs:domain
rdfs:range
Individual

(In)Equalitiy:

equivalentClass
equivalentProperty
sameAs
differentFrom
AllDifferent
distinctMembers

Property Characteristics:
Property Restrictions:
Restricted Cardinality:
Header Information:
Class Axioms:
Arbitrary Cardinality

Class Intersection:
Versioning:
Annotation Properties:
Datatypes:
Boolean Combinations:
Filler Information

http://www.w3.org/TR/2004/REC-owl-features-20040210/#s2.1

OWL – Web Ontology Language (3)

RDF Schema Features:

rdfs:subPropertyOf

rdfs:subPropertyOf: Property hierarchies may be created by making one or more statements that a property is a subproperty of one or more other properties. For example, hasSibling may be stated to be a subproperty of hasRelative. From this a reasoner can deduce that if an individual is related to another by the hasSibling property, then it is also related to the other by the hasRelative property.

One of 54 base language constructs

http://www.w3.org/TR/2004/REC-owl-features-20040210/#s2.1

OWL – Web Ontology Language (2)

RDF Schema Features:
Class (Thing, Nothing)
rdfs:subClassOf
rdf:Property
rdfs:subPropertyOf
rdfs:domain
rdfs:range
Individual

(In)Equalitiy:
equivalentClass
equivalentProperty
sameAs
differentFrom
AllDifferent
distinctMembers

Property Characteristics:
Property Restrictions:
Restricted Cardinality:
Header Information:
Class Axioms:
Arbitrary Cardinality

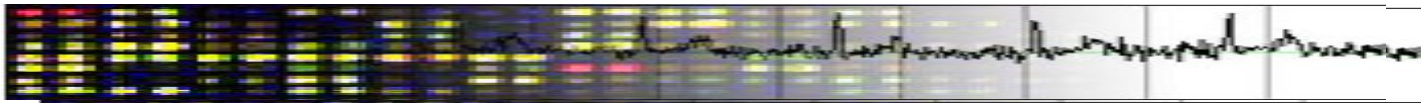
Class Intersection:
Versioning:
Annotation Properties:
Datatypes:
Boolean Combinations:
Filler Information

http://www.w3.org/TR/2004/REC-owl-features-20040210/#s2.1

XML Schema Datatypes used in OWL

http://www.w3.org/2001/XMLSchema#name

xsd:string, xsd:boolean, xsd:decimal, xsd:float, xsd:double, xsd:dateTime, xsd:time, xsd:date, xsd:gYearMonth, xsd:gYear, xsd:gMonthDay, xsd:gDay, xsd:gMonth, xsd:hexBinary, xsd:base64Binary, xsd:anyURI, xsd:normalizedString, xsd:token, xsd:language, xsd:NMTOKEN, xsd:Name, xsd:NCName, xsd:integer, xsd:nonPositiveInteger, xsd:negativeInteger, xsd:long, xsd:int, xsd:short, xsd:byte, xsd:nonNegativeInteger, xsd:unsignedLong, xsd:unsignedInt, xsd:unsignedShort, xsd:unsignedByte and xsd:positiveInteger

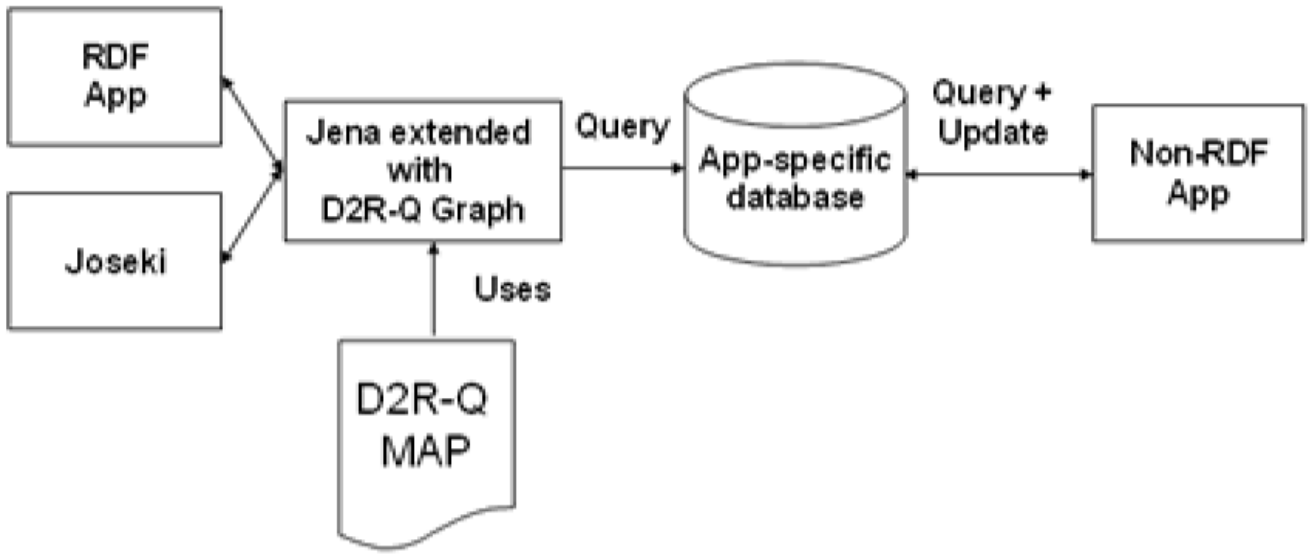


Our approach to creating and merging databases

- *Create ontologies from database schema*
- *Edit and maintain the ontologies*
- *Utilize merged & aligned ontologies*
- *Create on-the-fly databases from ontologies*
- *Model scientific processes - use cases*
- *Allow for interoperability*

Creating ontologies from database schema

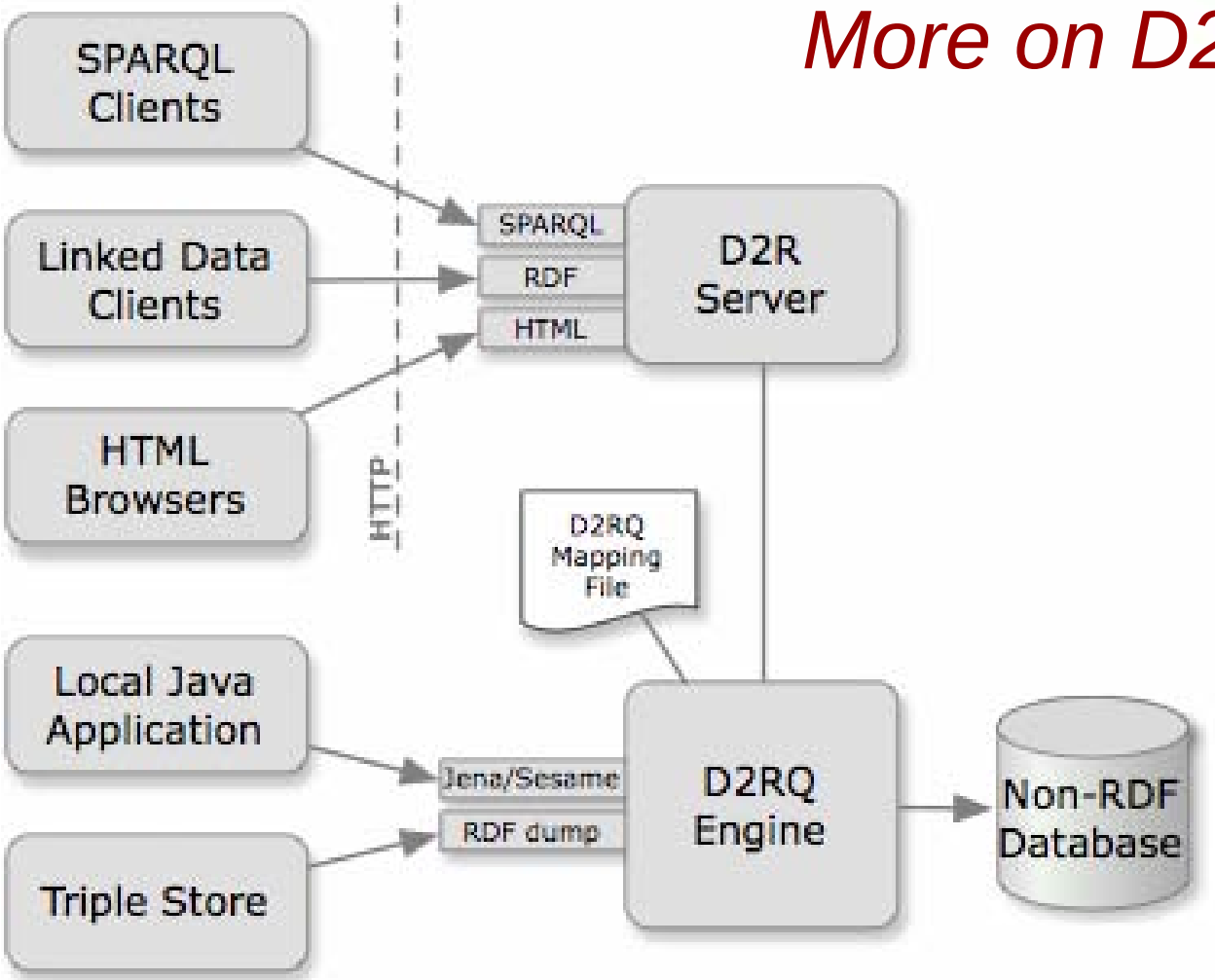
D2RQ is a declarative language to describe mappings between relational database schemata and OWL/RDFS ontologies.



Courtesy of Prof. Dr. Christian Bizer. Used with permission.

http://www4.wiwiss.fu-berlin.de/bizer/D2RQ/

More on D2RQ

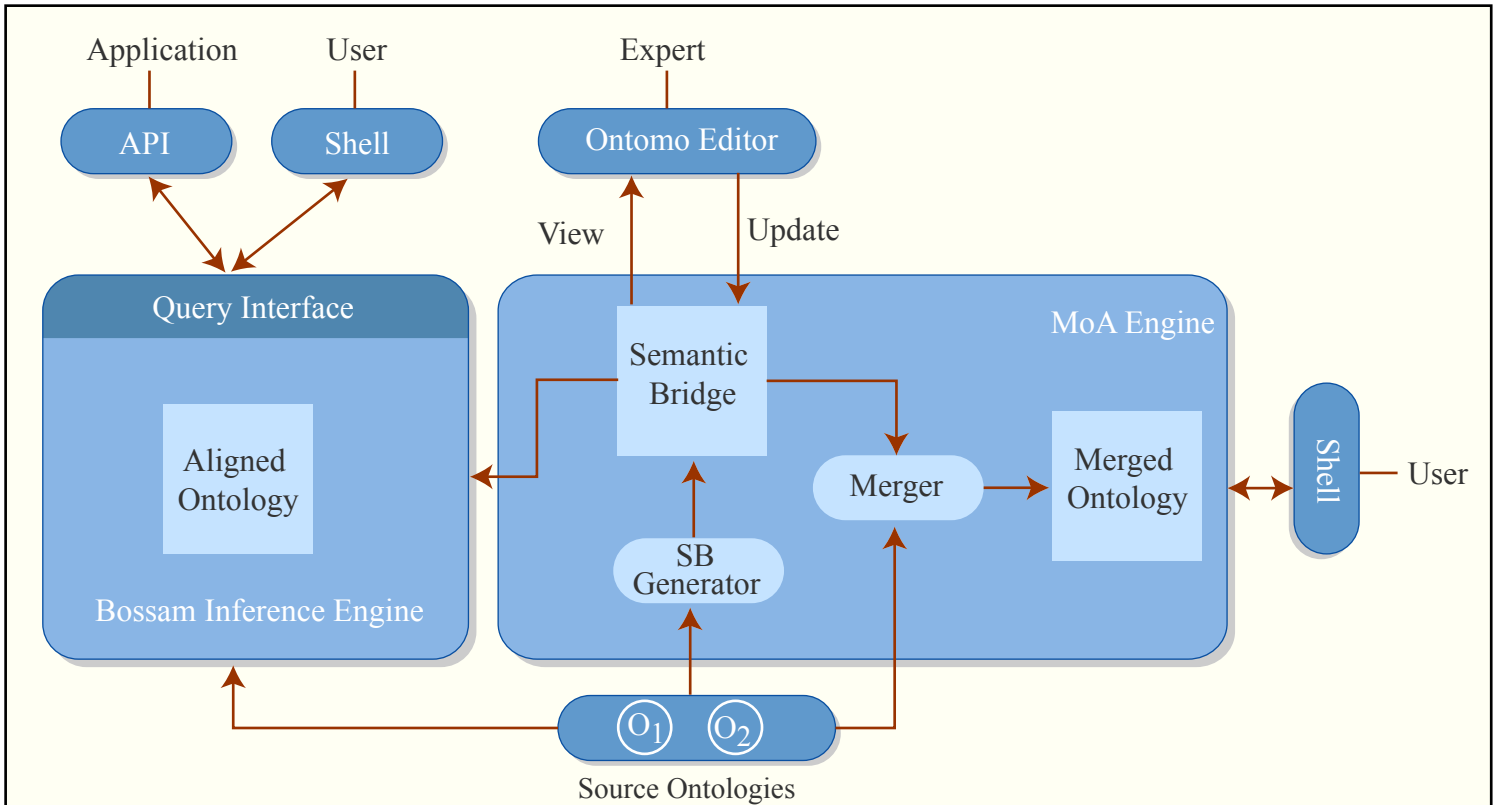


Courtesy of Prof. Dr. Christian Bizer. Used with permission.

http://www4.wiwiss.fu-berlin.de/bizer/D2RQ/

Options for Ontology Merging & Aligning

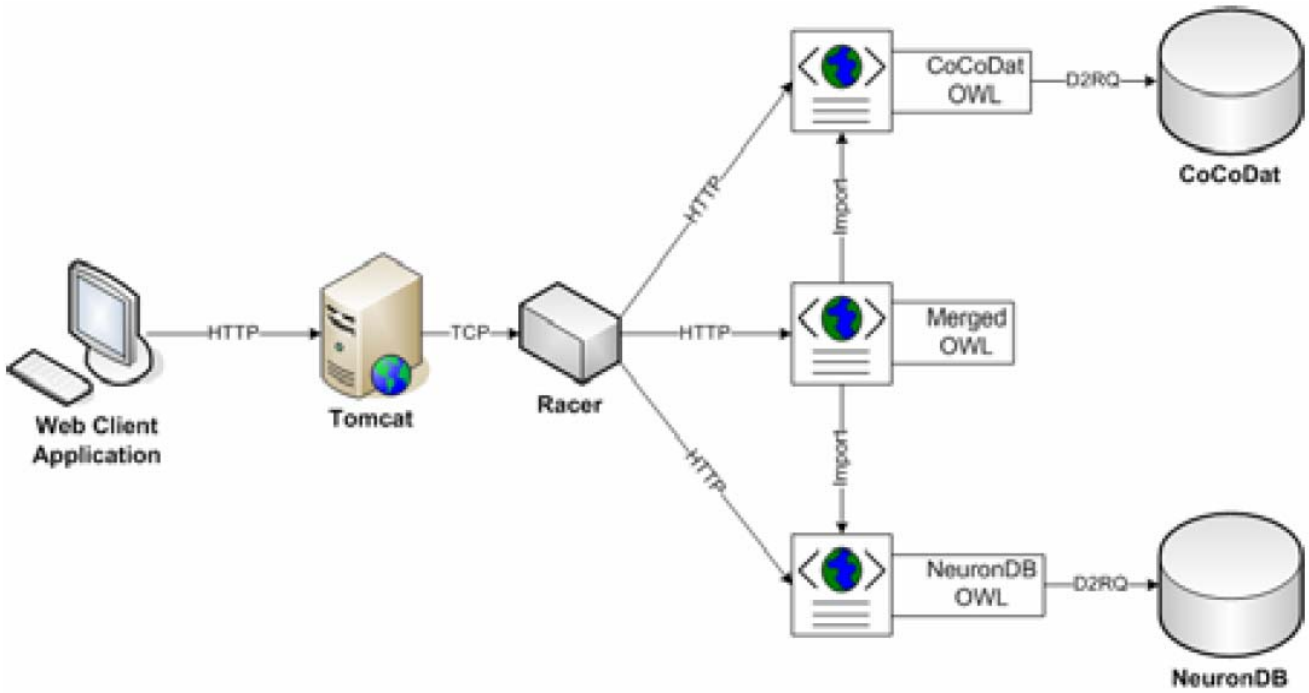
Kim, Jaehong, Minsu Jang, Young-Guk Ha, Joo-Chan Sohn, and Sang Jo Lee. "MoA: OWL Ontology Merging and Alignment Tool for the Semantic Web." 18th Conference on IAAI, Bari Italy, 2005.



Architecture of MoA system.

Figure by MIT OpenCourseWare, after Kim et al, 2005.

Options for Ontology Merging & Aligning (2)



Lam, H.Y.K. et al. "Using Web Ontology Language to Integrate Heterogeneous Databases in the Neurosciences." *AMIA Annual Symposium Proc*. 2006, 464-468.
[PubMed Central OpenAccess article.]

Options for Ontology Merging & Aligning

Ontology Merging for Federated Ontologies on the Semantic Web

Gerd Stumme
Institute for Applied Computer Science and
Formal Description Methods (AIFB)
University of Karlsruhe
D-76128 Karlsruhe, Germany
www.aifb.uni-karlsruhe.de/WBS/gst

Alexander Maedche
FZI Research Center
for Information Technologies
Haid-und-Neu-Strasse 10-14
D–76131 Karlsruhe, Germany
www.fzi.de/wim

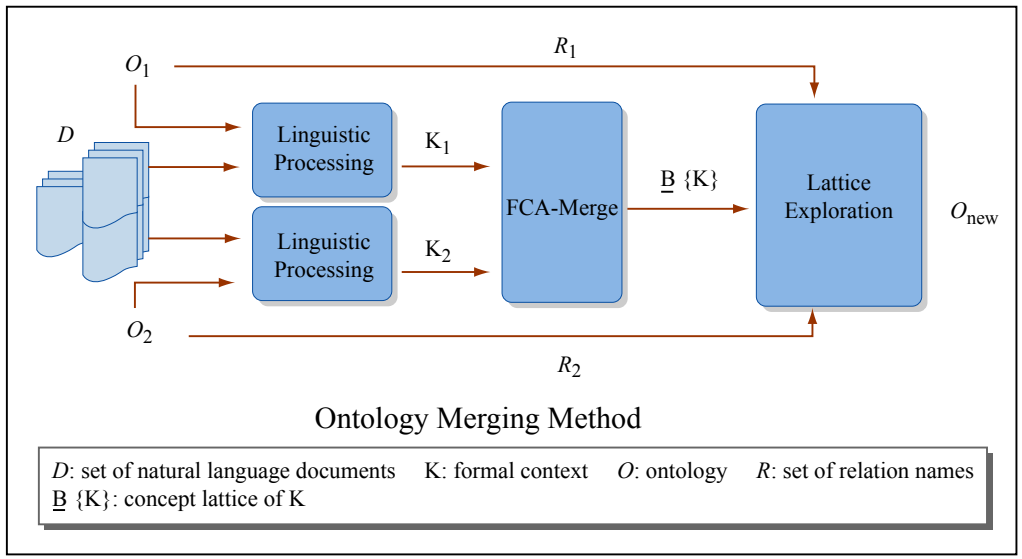
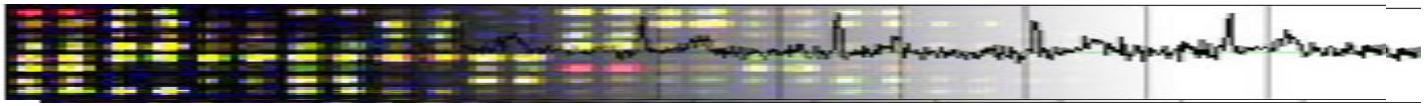


Figure by MIT OpenCourseWare.

Proceedings of the International Workshop for Foundations of Models for Information Integration (FMII-2001), Viterbo, Italy, September 2001.



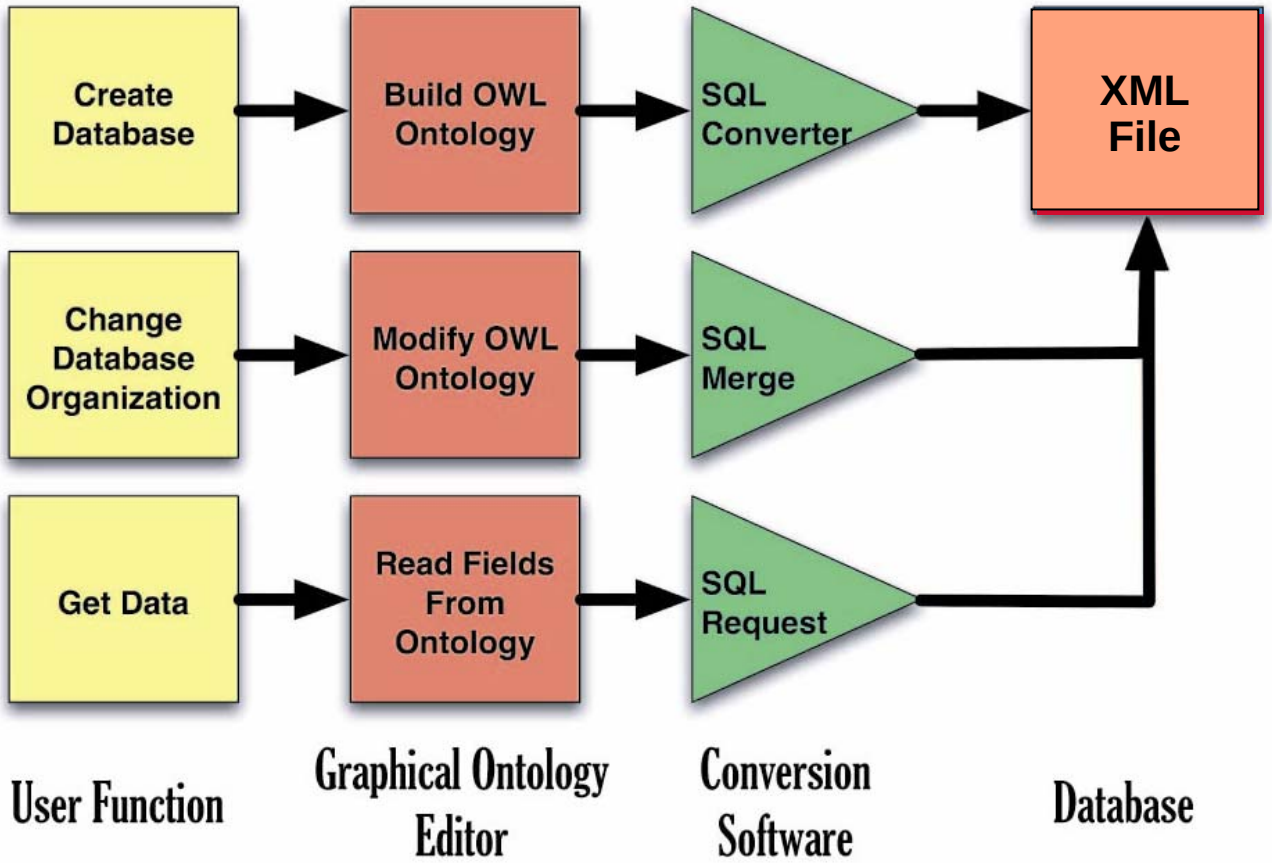
Options for Ontology Merging & Aligning

An Algorithm for Merging and Aligning Ontologies: Automation and Tool Support

Natalya Fridman Noy and Mark A. Musen

Stanford Medical Informatics
Stanford University
Stanford, CA 94305-5479
{noy, musen}@smi.stanford.edu

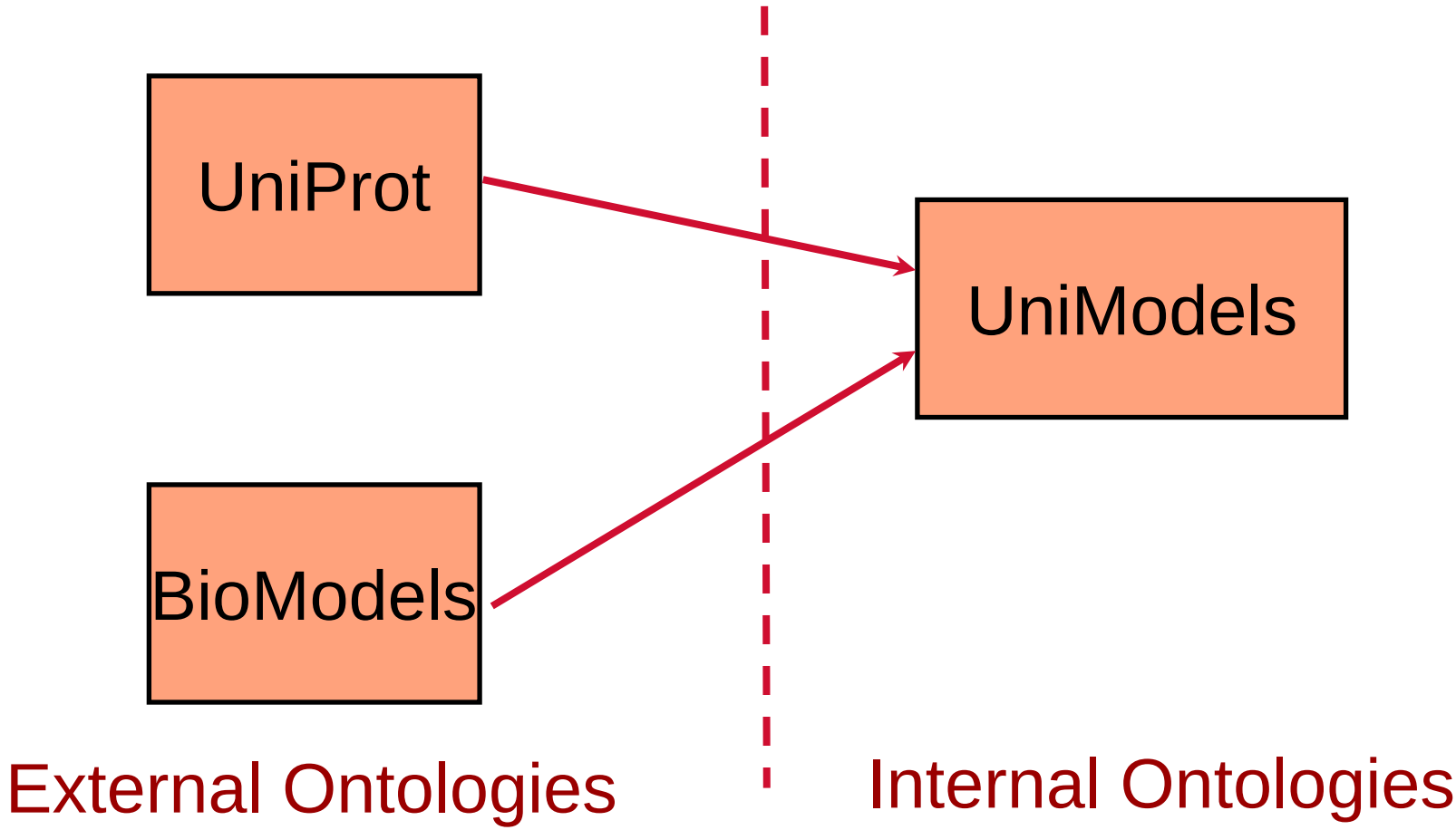
OWLdb: A new paradigm for creating databases



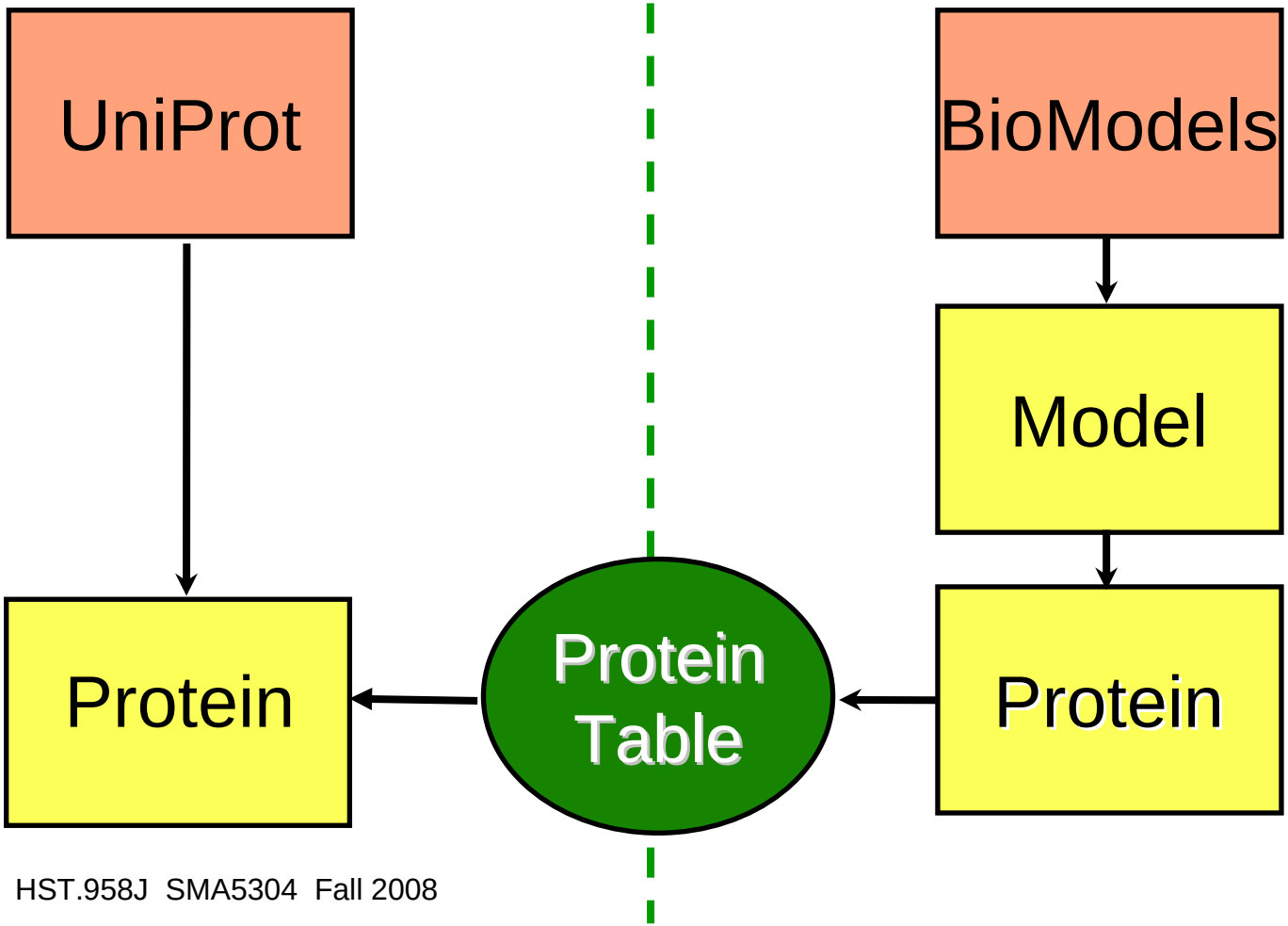
. . . . Courtesy of Kurt Stiehl

From Stiehl, Kurt R. "Development of Dynamic Database Structures Using OWL Ontologies." MIT BSME Thesis, June 2007.

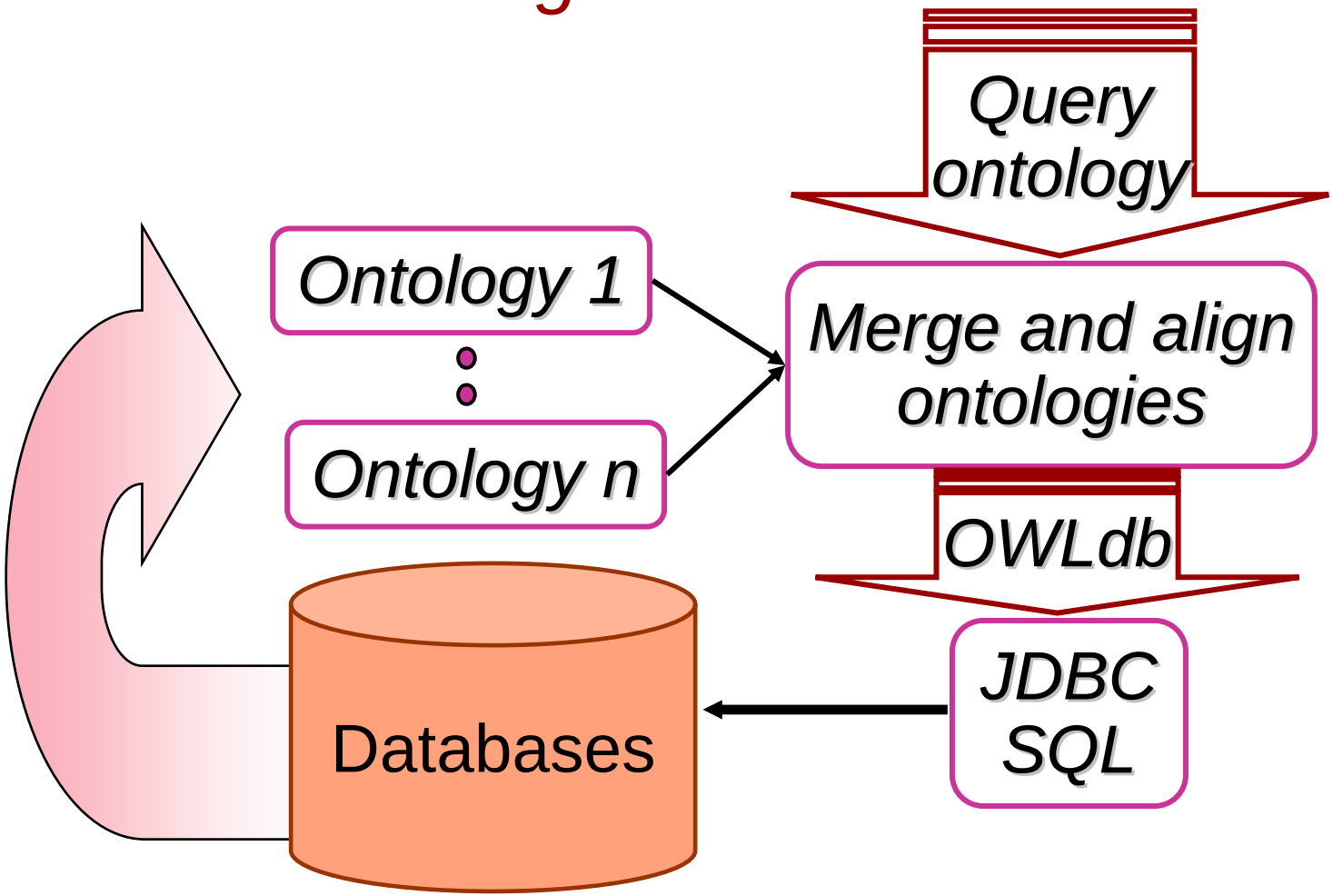
Integration - Ontology



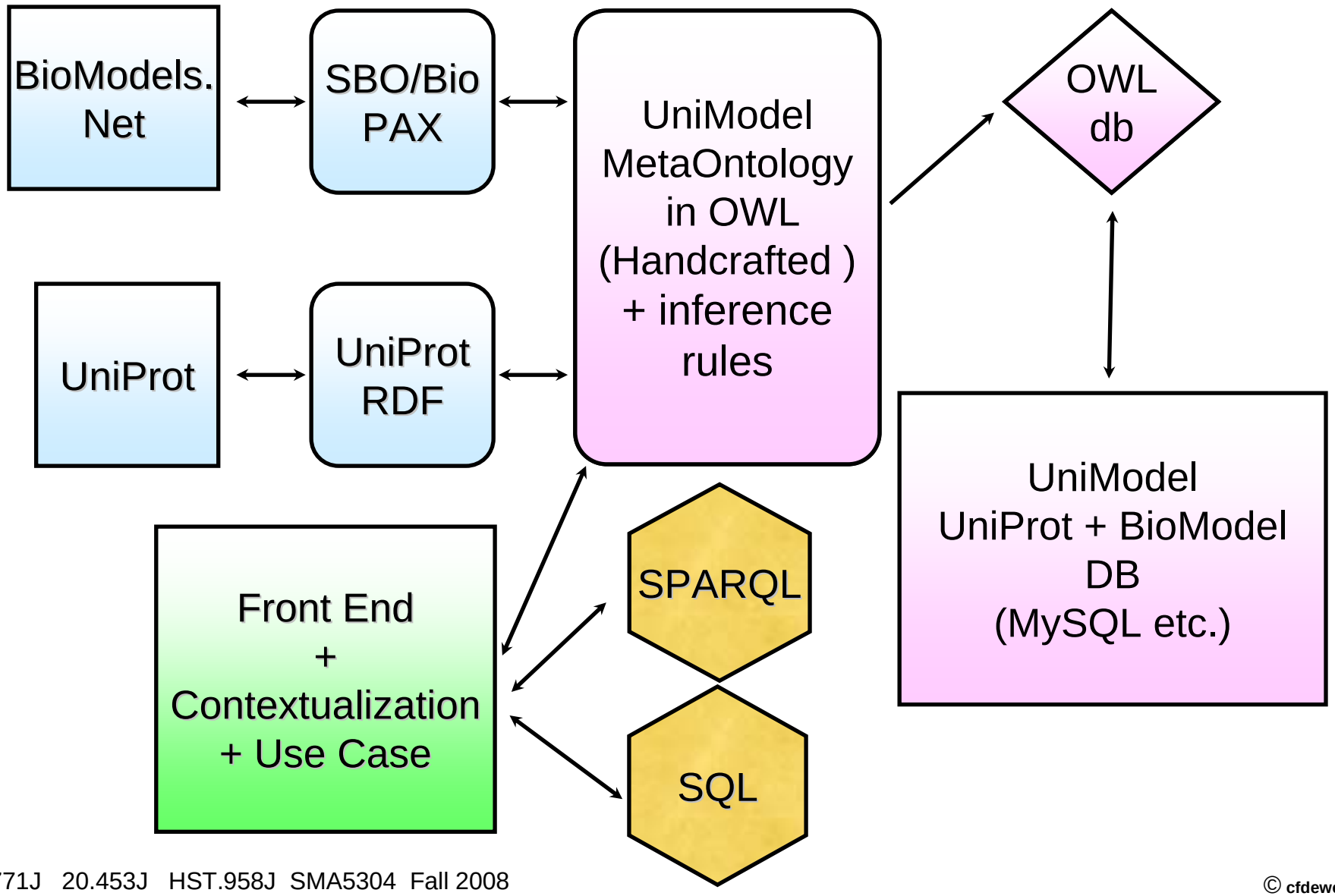
The Species Link



The OWLdb design



Architecture



Key Components

	UniModels
Registry	LSID
Interface	UniModels Ontology
Nomenclature	UniModels / SBO / UniPROT RDF
Editor	Protege/SBMLEditor
Parser	Jena
Reasoner	Jena/OWLdb
Storage	Oracle / MySQL / PostgresSQL / Sesame
Query	SPARQL / SQL / D2RQ / SeRQL

Computational and Systems Biology