

Primers
Tutorial T05
November 11, 2006



Semantic Web for Health Care and Life Sciences



Olivier Bodenreider
Vipul Kashyap
Eric Neumann



TERANODE®

Overview

- Overview of the Semantic Web
- Semantic Web technologies
- Semantic Web applications in biomedicine
- W3C Semantic Web *Health Care and Life Sciences Interest Group*
- Current trends and future directions

Part 1

Overview of the Semantic Web

A web of information

- Many biomedical resources available on the Web
- Information retrieval paradigm
- For humans to read
 - Human “in the middle”
 - No automated integration

Web technologies

- Uniform naming scheme for locating resources
 - Unified Resource Identifier (URI) <http://www.w3.org/Addressing/>
- Protocols to access named resources
 - HyperText Transfer Protocol (HTTP) <http://www.w3.org/Protocols/>
- Hypertext navigation among resources
 - HyperText Markup Language (HTML) <http://www.w3.org/MarkUp/>
 - To link documents to other documents
 - Document structure (and presentation)
 - Hypertext/hypermedia links
 - No semantics

HTML today

- HTML 4 (1997)
 - Internationalization
 - Scripting
 - Stylesheet (decouple structure and presentation)
- XHTML (2000)
 - Extensible HyperText Markup Language
 - HTML in XML
 - Can be processed by XML tools

Limitations of the current Web

- Heterogeneity
 - Syntax (e.g., document format)
 - Semantics (e.g., values for schemas)
- Based on natural language, not represented in a controlled way
 - Not interoperable
- Web for humans
 - Not easily processable by agents
 - Information retrieval paradigm

Semantic Web approach

- From information designed for human consumption
 - Web of documents
 - Discovered and read by humans
- To information expressed in a machine processable form
 - Web of data and information
 - Accessed and processed by agents

Semantic Web The vision

The entertainment system was belting out the Beatles' "We Can Work It Out" when the phone rang. When Pete answered, his phone turned the sound down by sending a message to all the other *local* devices that had a *volume control*. His sister, Lucy, was on the line from the doctor's office: "Mom needs to see a specialist and then has to have a series of physical therapy sessions. Biweekly or something. I'm going to have my agent set up the appointments." Pete immediately agreed to share the chauffeuring. At the doctor's office, Lucy instructed her Semantic Web agent through her handheld Web browser. The agent promptly retrieved information about Mom's *prescribed treatment* from the doctor's agent, looked up several lists of *providers*, and checked for the ones *in-plan* for Mom's insurance within a *20-mile radius* of her *home* and with a *rating* of *excellent* or *very good* on trusted rating services. It then began trying to find a match between available *appointment times* (supplied by the agents of individual providers through their Web sites) and Pete's and Lucy's busy schedules. (The emphasized keywords indicate terms whose semantics, or meaning, were defined for the agent through the Semantic Web.)

[...]

[Berners-Lee et al., Scientific American, 2001]

Semantic Web In practice

- Extension of the current Web
- For both humans and agents
- Seamless integration of resources
 - Common format
 - Common, harmonized data model
 - Shared meaning
- Semantic bus
- Support for reasoning and decision making

Semantic Web In practice

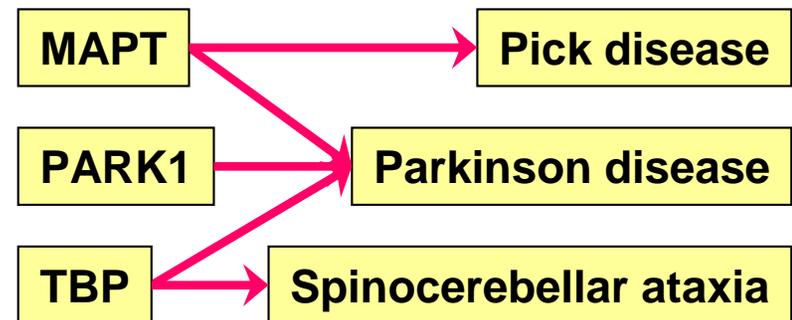
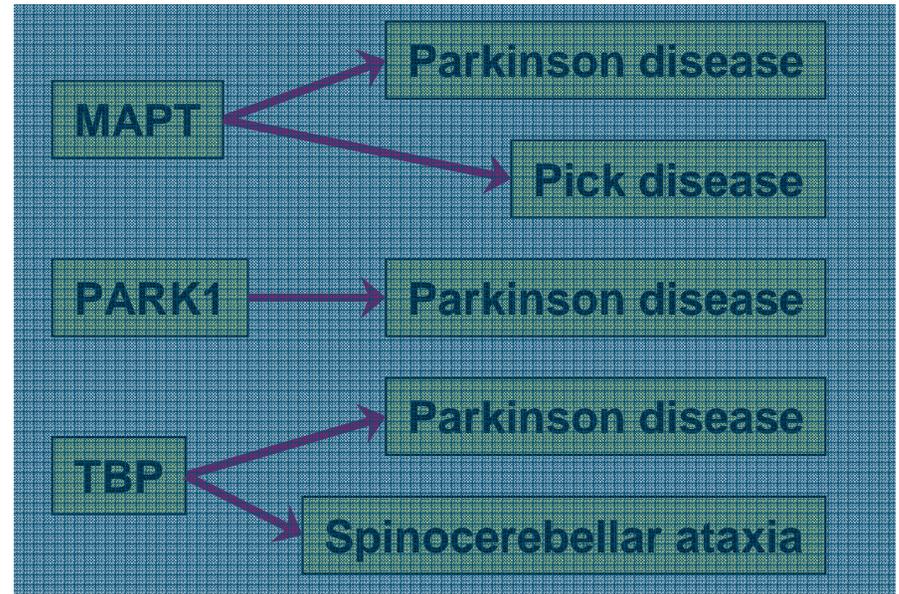
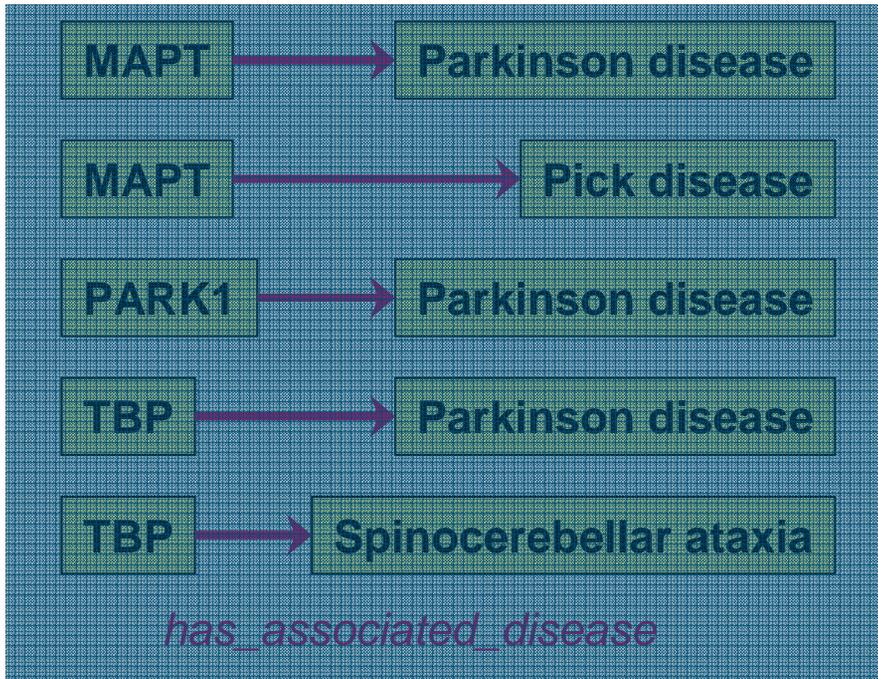
- Representing facts (assertions) on the Web
- Querying facts
- Reasoning about them

- Examples of facts
 - *PARK1* *has associated disease* Parkinson disease
 - Parkinson disease *isa* Neurodegenerative disease

Facts as triples

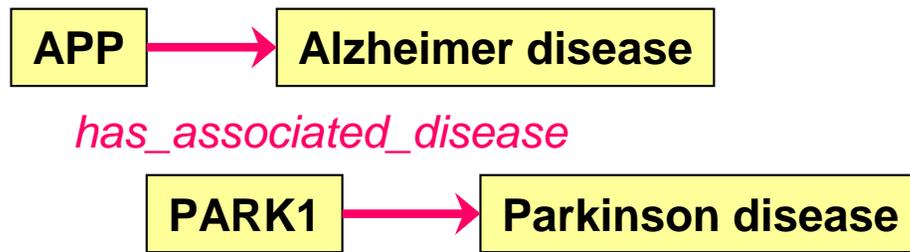
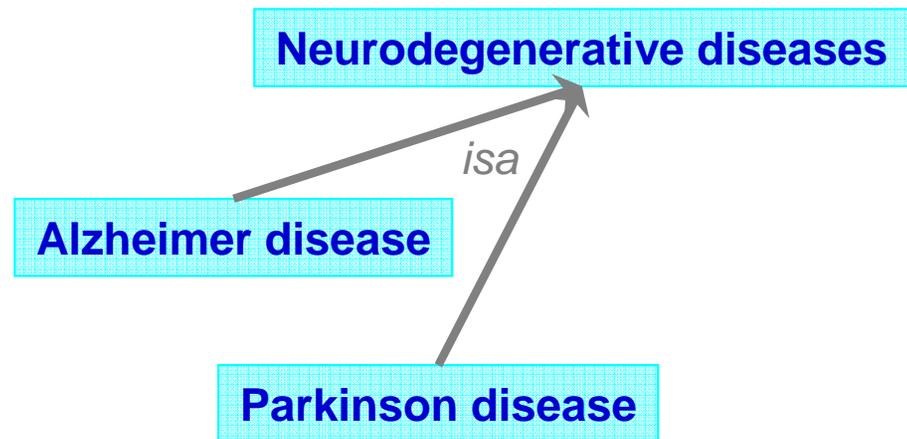


From triples to a graph



Connecting graphs

- Integrate graphs from multiple resources
- Query across resources



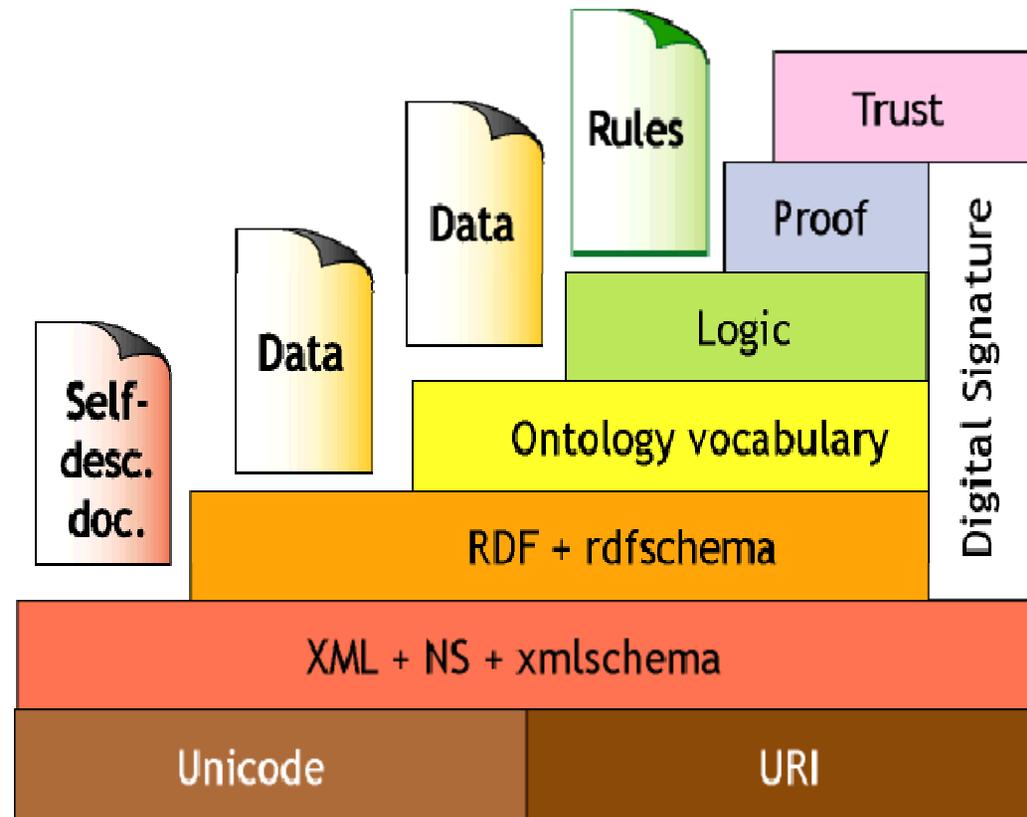
Needed to realize the SW vision

- A standard way of identifying things
- A standard way of describing things
- A standard way of linking things
- Standard vocabularies for talking about things

Semantic Web Technologies

- Richer structure for resources
 - eXtensible Markup Language (XML)
- Exposed semantics
 - Resource Description Framework (RDF)
- Explicit semantics
 - Ontologies
 - Web Ontology Language (OWL)

The “layer cake” slide



[Tim Berners Lee, XML-2000 Conference]

Part 2

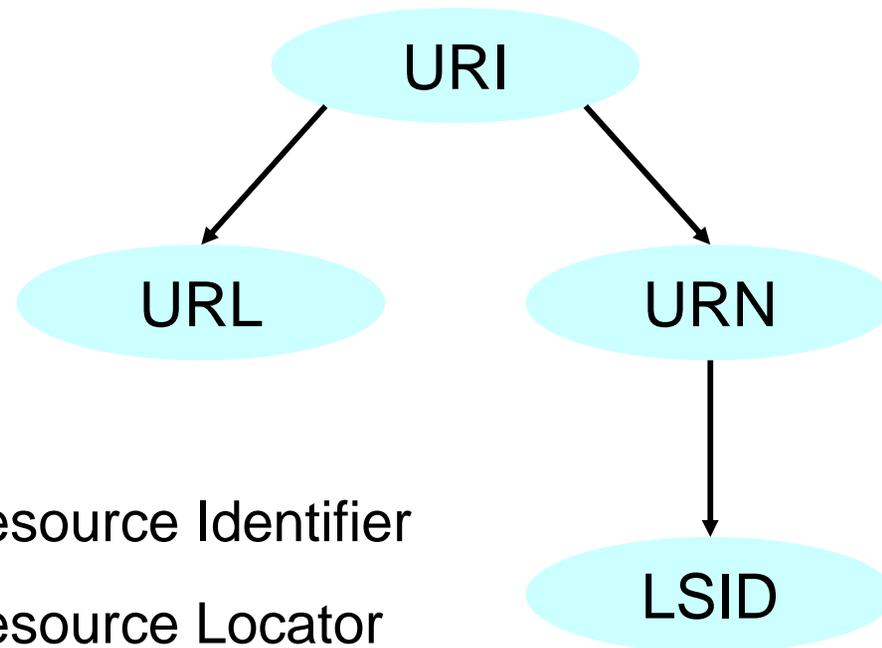
Semantic Web technologies

Overview

- Resource identification schemes
- Extensible Markup Language (XML)
- Resource Description Framework (RDF)
- Ontologies
 - Web Ontology Language (OWL)
- Logic
 - OWL-based automated reasoning
 - Rule languages (RuleML, SWRL)
- Web services

Resource identification schemes

A Family of Identifiers



URI = Uniform Resource Identifier

URL = Uniform Resource Locator

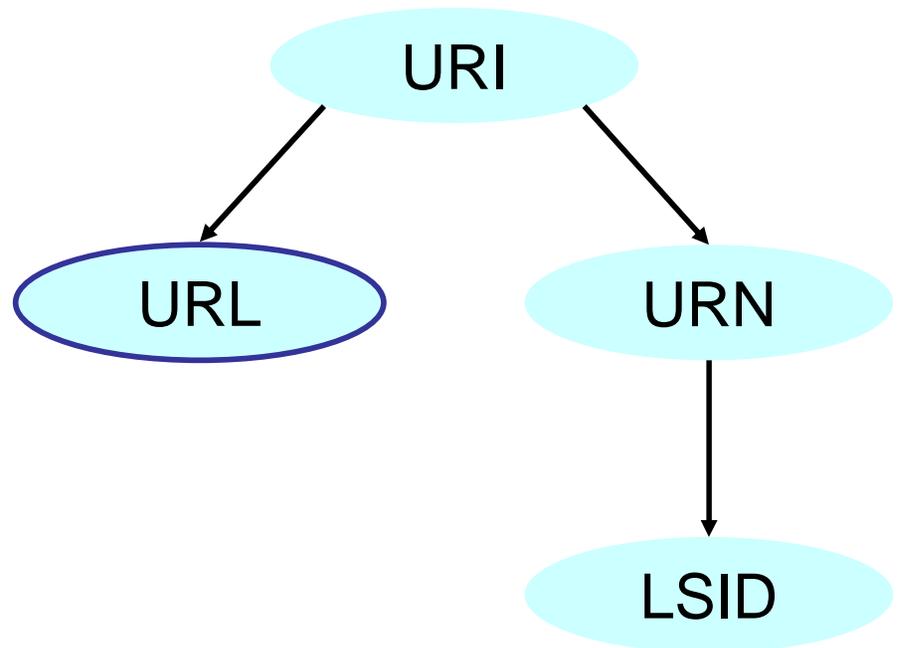
URN = Uniform Resource Name

LSID = Life Science Identifier

<http://www.w3.org/Addressing/>

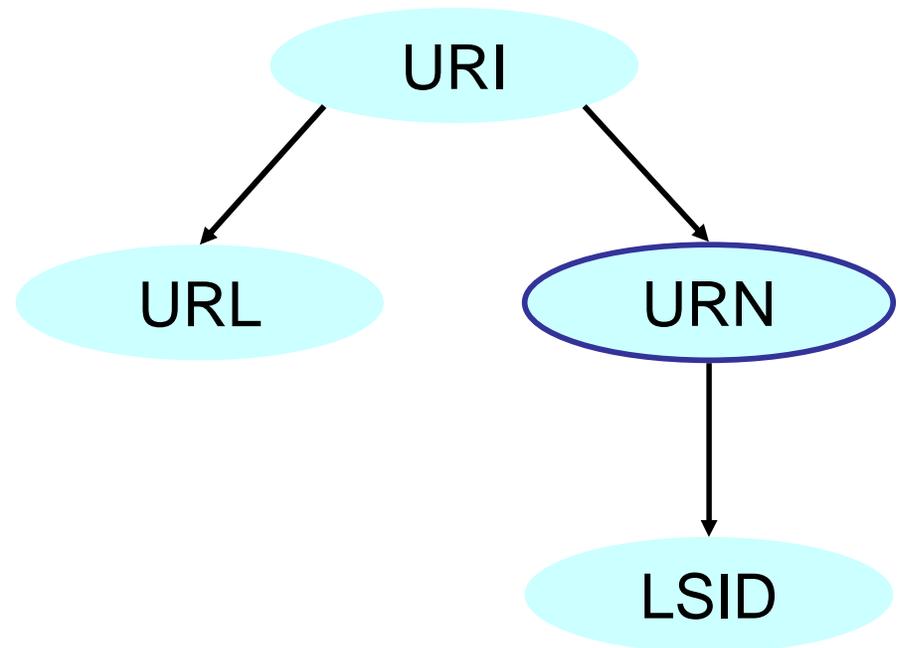
Uniform Resource Locator

- A type or resource identifier
- Identifies the **location** of a resource (or part thereof)
- Specifies a protocol to access the resource
 - http, ftp, mailto
- E.g.,
 - <http://www.nlm.nih.gov/>



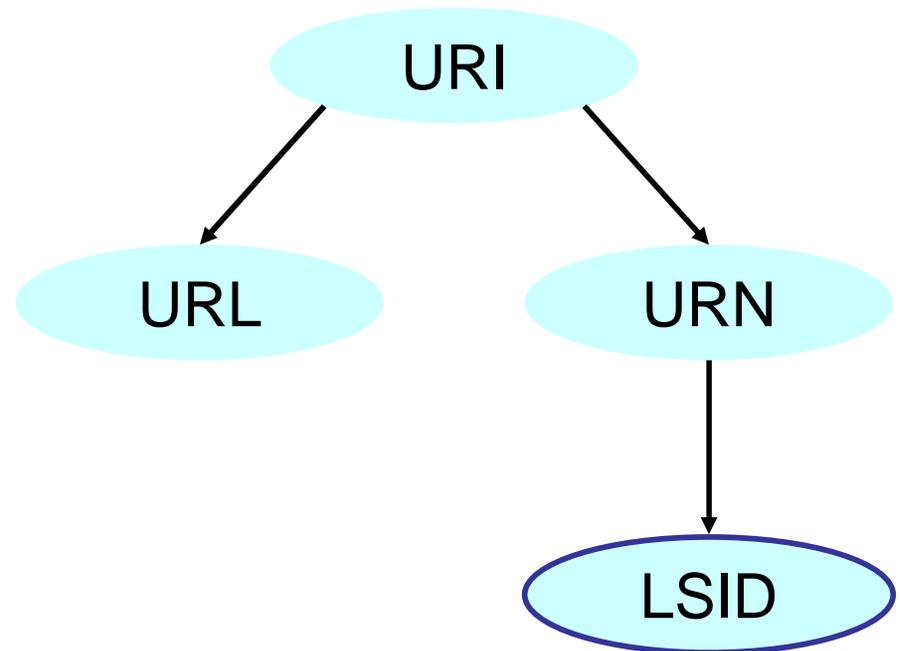
Uniform Resource Name

- A type or resource identifier
- Identifies the **name** of a resource
- Location independent
- Defines a namespace
- E.g.,
 - urn:isbn:0-262-02591-4
 - urn:umls:C0001403



Life Science Identifier

- A type or resource identifier
- A type of URN
- For biological entities
- Specific properties
 - Versioned
 - Resolvable
 - Immutable



- E.g.,

`urn:lsid:ncbi.nlm.nih.gov:pubmed:12571434`

 DNS name namespace unique ID

<http://lsid.sourceforge.net/>

Extensible Markup Language (XML)

XML Introduction

- Derived from the Standard Generalized Markup Language (SGML)
 - designed to enable the sharing of machine-readable documents
- A kind of markup language
- Designed to describe data
- No predefined tags

<http://www.w3.org/XML/>

XML Example (MeSH 2007)

```
<DescriptorRecordSet>
  <DescriptorRecord DescriptorClass = "1">
    <DescriptorUI>D000224</DescriptorUI>
    <DescriptorName>
      <String>Addison Disease</String>
    </DescriptorName>
    <TreeNumberList>
      <TreeNumber>C19.053.500.263</TreeNumber>
      <TreeNumber>C20.111.163</TreeNumber>
    </TreeNumberList>
    <ConceptList>
      <Concept PreferredConceptYN="Y">
        <ConceptUI>M0000346</ConceptUI>
        <ConceptName>
          <String>Addison Disease</String>
        </ConceptName>
        <ConceptUMLSUI>C0001403</ConceptUMLSUI>
        [...]
      </Concept>
      [...]
    </ConceptList>
  </DescriptorRecord>
</DescriptorRecordSet>
```

To download MeSH in XML, see
<http://www.nlm.nih.gov/mesh/filelist.html>

The MeSH browser

National Library of Medicine - Medical Subject Headings

2007 MeSH

MeSH Descriptor Data

[Return to Entry Page](#)

MeSH Heading	Addison Disease
Tree Number	C19.053.500.263
Tree Number	C20.111.163
Scope Note	An adrenal disease characterized by the progressive destruction of the ADRENAL CORTEX , resulting in insufficient production of ALDOSTERONE and HYDROCORTISONE . Clinical symptoms include ANOREXIA ; NAUSEA ; WEIGHT LOSS ; MUSCLE WEAKNESS ; and HYPERPIGMENTATION of the SKIN due to increase in circulating levels of ACTH precursor hormone which stimulates MELANOCYTES .
Entry Term	Addison's Disease
Entry Term	Primary Adrenal Insufficiency
Entry Term	Primary Adrenocortical Insufficiency
Entry Term	Primary Hypoadrenalism
Allowable Qualifiers	BL CF CI CL CN CO DH DI DT EC EH EM EN EP ET GE HI IM ME MI MO NU PA PC PP PS PX RA RH RI RT SU TH UR US VE VI
Entry Version	ADDISON DIS
History Note	2005 (1963)
Unique ID	D000224

<http://www.nlm.nih.gov/mesh/MBrowser.html>

XML vs. HTML

- HTML

- Main focus:
Display information
- How data looks

```
[...]  
<TABLE border>  
<TITLE>Addison Disease</TITLE>  
<TR><TH align=left>MeSH Heading</TH>  
<TD>Addison Disease</TD></TR>  
<TR><TH align=left>Tree Number</TH>  
<TD><A HREF="#TreeC19.053.500.263">  
C19.053.500.263</A></TD></TR>  
<TR><TH align=left>Tree Number</TH>  
<TD><A HREF="#TreeC20.111.163">  
C20.111.163</A></TD></TR>  
[...]  
<TR><TH align=left>Unique ID</TH>  
<TD>D000224</TD></TR>  
[...]
```

- XML

- Main focus:
Describe information
- What data is

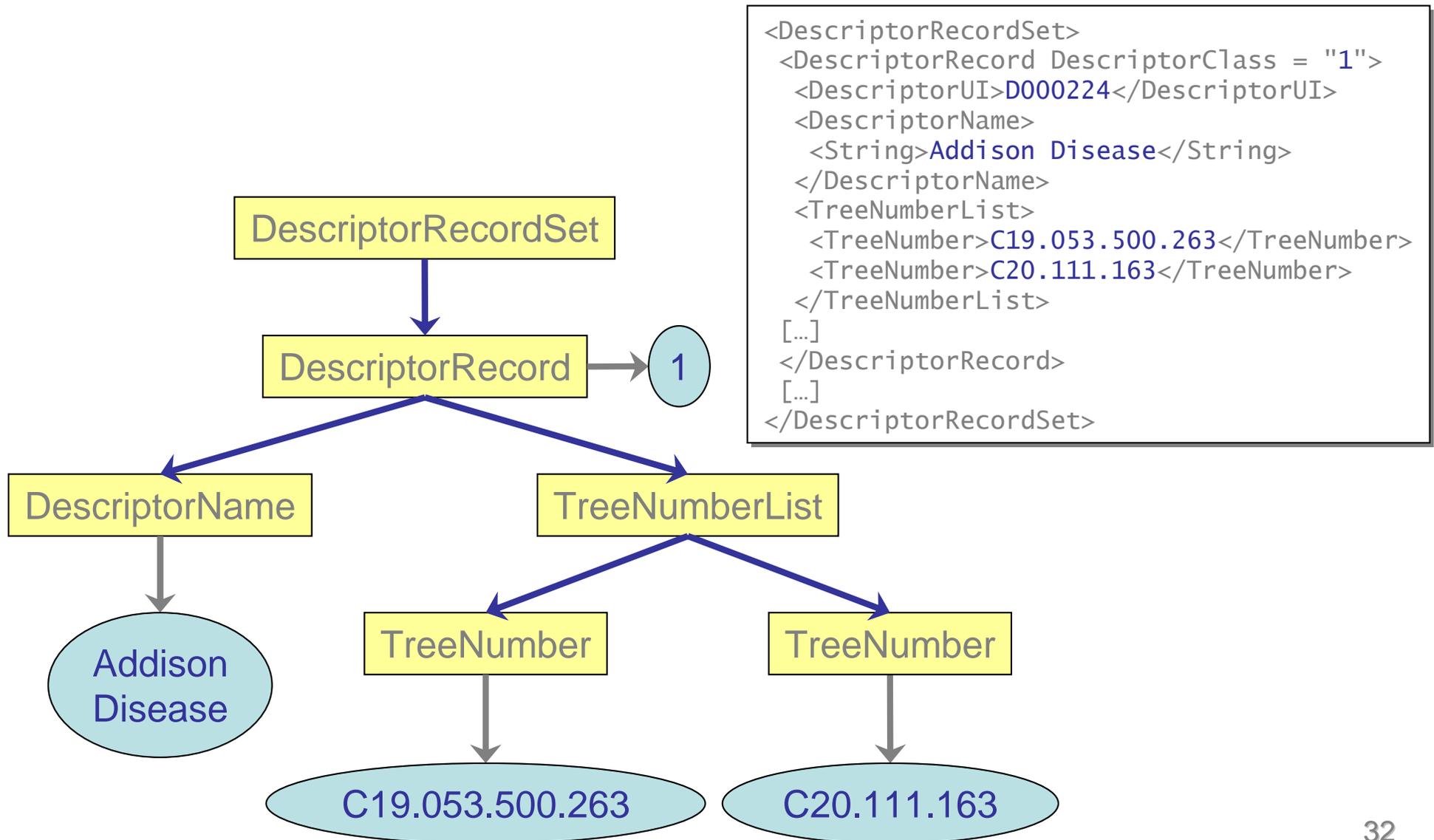
```
<DescriptorRecordSet>  
  <DescriptorRecord DescriptorClass = "1">  
    <DescriptorUI>D000224</DescriptorUI>  
    <DescriptorName>  
      <String>Addison Disease</String>  
    </DescriptorName>  
    <TreeNumberList>  
      <TreeNumber>C19.053.500.263</TreeNumber>  
      <TreeNumber>C20.111.163</TreeNumber>  
    </TreeNumberList>  
    [...]  
  </DescriptorRecord>  
  [...]  
</DescriptorRecordSet>
```

XML syntax

- Mandatory document root
- XML tags
 - Not predefined
 - Mandatory closing tag
 - Case-sensitive
 - Can be nested
 - Order matters
- XML attributes
 - Values must be quoted
 - Order does not matter

```
<DescriptorRecordSet>
  <DescriptorRecord DescriptorClass = "1">
    <DescriptorUI>D000224</DescriptorUI>
    <DescriptorName>
      <String>Addison Disease</String>
    </DescriptorName>
    <TreeNumberList>
      <TreeNumber>C19.053.500.263</TreeNumber>
      <TreeNumber>C20.111.163</TreeNumber>
    </TreeNumberList>
    [...]
  </DescriptorRecord>
  [...]
</DescriptorRecordSet>
```

XML seen as a tree (partial)



Document definition

- Two mechanisms
 - DTD (Document Type Definition)
 - XSD (XML schema)
- DTDs can be converted to XSDs
- Serve for validating the structure of XML documents

Processing XML documents

- Parsing
 - Xerces, Expat, ...
 - SAX (API to XML parsers)
 - DOM (parser and representation)
- Transformation
 - Stylesheets
 - XSLT – Extensible Stylesheet Language Transformation

Resources available in XML

- Literature
 - MEDLINE (citations)
 - PubMedCentral (full-text)
- Terminology
 - Medical Subject Headings (MeSH)
- Knowledge bases
 - Entrez databases (e.g., Gene, GenBank,...)
- ...

Resource Description Framework (RDF)

RDF Introduction

- Based on XML and URI
- Major differences with XML
 - Explicit semantics
 - Many-to-many relationships
 - Limited support for inference
 - Open-world assumption

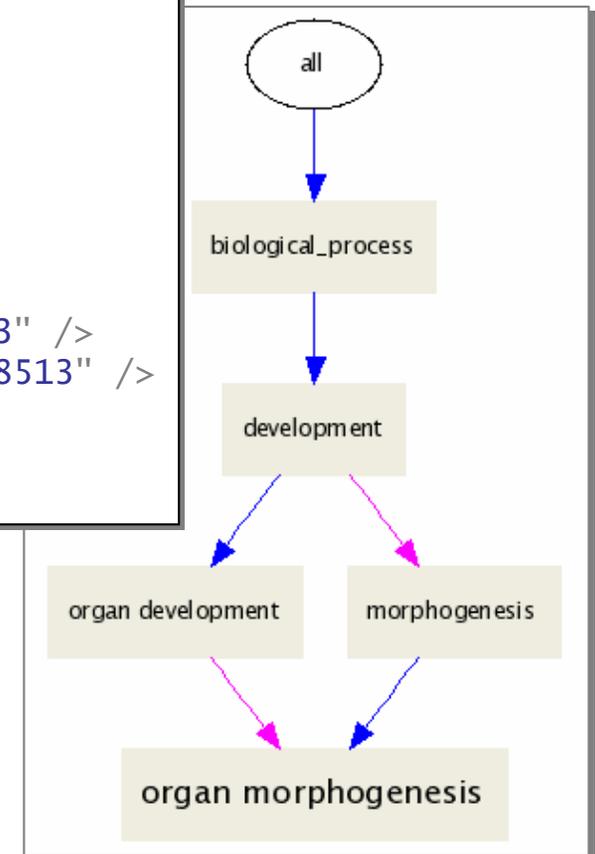
<http://www.w3.org/RDF/>

RDF Example

```
<go:go xmlns:go="http://www.geneontology.org/dtds/go.dtd#"
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:RDF>
    [...]
    <go:term rdf:about="http://www.geneontology.org/go#GO:0009887" n_associations="0">
      <go:accession>GO:0009887</go:accession>
      <go:name>organ morphogenesis</go:name>
      <go:synonym>histogenesis and organogenesis</go:synonym>
      <go:definition>Morphogenesis of an organ. An organ is defined as a tissue or set
of tissues that work together to perform a specific function or functions.
Morphogenesis is the process by which anatomical structures are generated and
organized. Organs are commonly observed as visibly distinct structures, but may also
exist as loosely associated clusters of cells that work together to perform a
specific function or functions.</go:definition>
      <go:is_a rdf:resource="http://www.geneontology.org/go#GO:0009653" />
      <go:part_of rdf:resource="http://www.geneontology.org/go#GO:0048513" />
    </go:term>
    [...]
  </rdf:RDF>
</go:go>
```

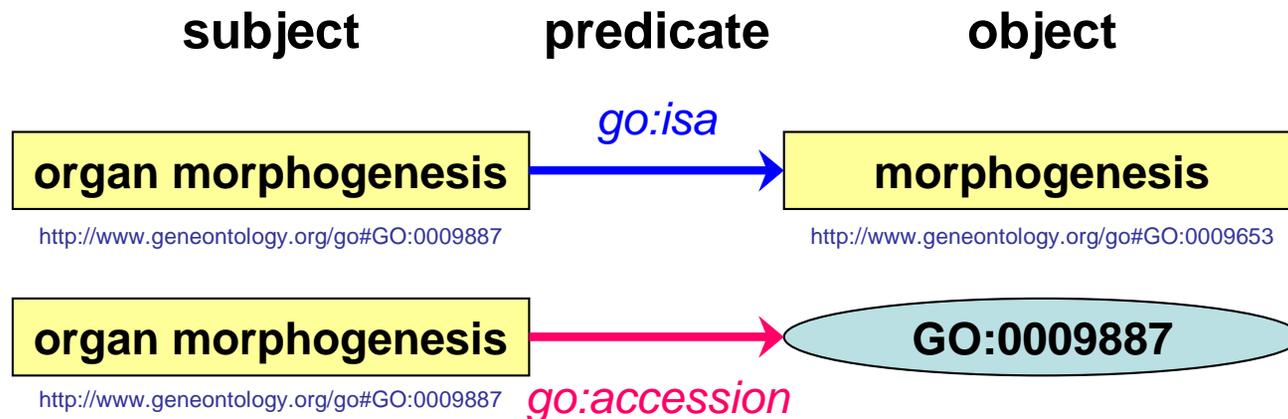
RDF Example

```
<go:go xmlns:go="http://www.geneontology.org/dtds/go.dtd#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:RDF>
    <go:term rdf:about="http://www.geneontology.org/go#GO:0009887">
      <go:accession>GO:0009887</go:accession>
      <go:name>organ morphogenesis</go:name>
      <go:synonym>histogenesis and organogenesis</go:synonym>
      [...]
      <go:is_a rdf:resource="http://www.geneontology.org/go#GO:0009653" />
      <go:part_of rdf:resource="http://www.geneontology.org/go#GO:0048513" />
    </go:term>
  </rdf:RDF>
</go:go>
```

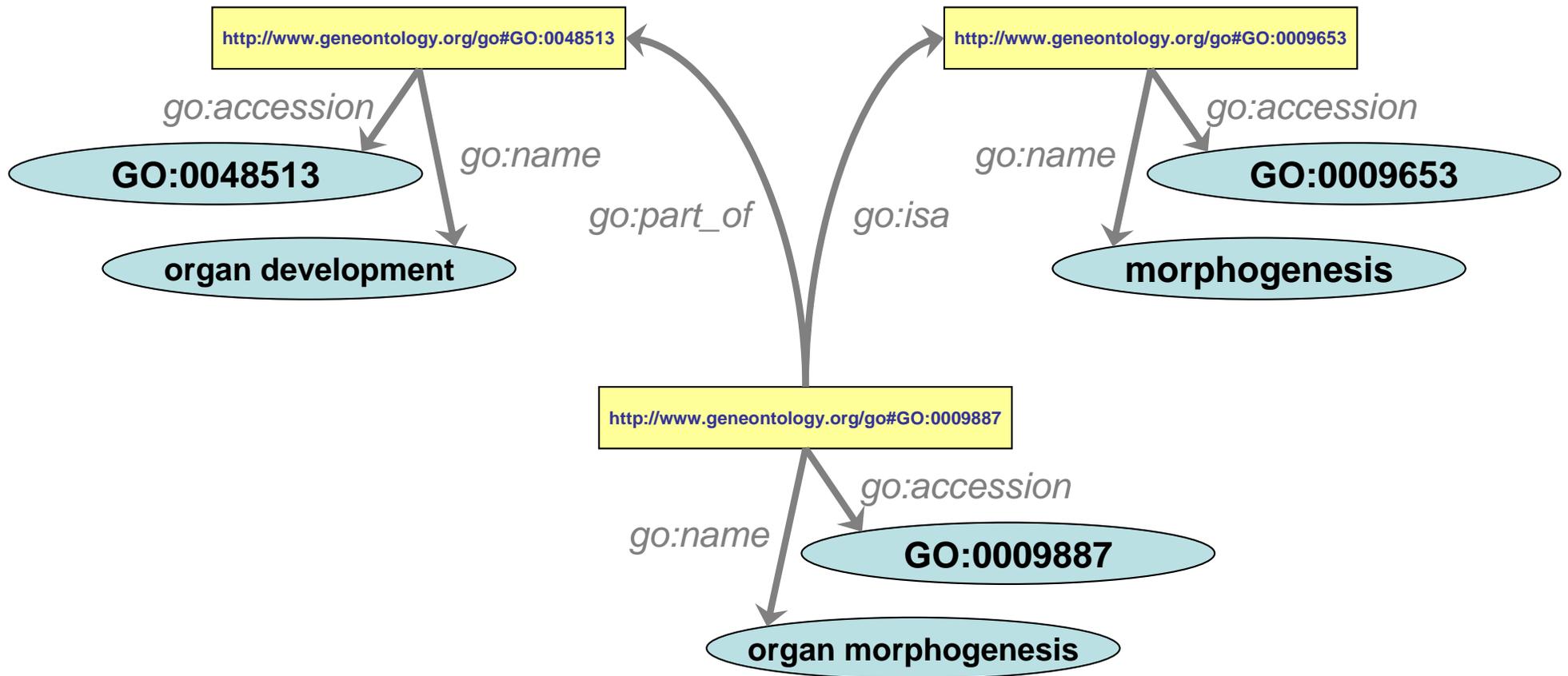


RDF Triples

```
<go:go xmlns:go="http://www.geneontology.org/dtds/go.dtd#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:RDF>
    <go:term rdf:about="http://www.geneontology.org/go#GO:0009887">
      <go:accession>GO:0009887</go:accession>
      <go:name>organ morphogenesis</go:name>
      <go:synonym>histogenesis and organogenesis</go:synonym>
      [...]
      <go:is_a rdf:resource="http://www.geneontology.org/go#GO:0009653" />
      <go:part_of rdf:resource="http://www.geneontology.org/go#GO:0048513" />
    </go:term>
  </rdf:RDF>
</go:go>
```



RDF Graph



RDF syntax

- DTD
- Properties
 - e.g., `go:name`
- Attributes
 - e.g., `rdf:resource`
- Values
 - Literals:
`organ morphogenesis`
 - Resources:
`http://www.geneontology.org/go#GO:0009653`

```
<go:go xmlns:go="http://www.geneontology.org/dtds/go.dtd#"
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:RDF>
    <go:term rdf:about="http://www.geneontology.org/go#GO:0009887">
      <go:accession>GO:0009887</go:accession>
      <go:name>organ morphogenesis</go:name>
      <go:synonym>histogenesis and organogenesis</go:synonym>
      [...]
      <go:is_a rdf:resource="http://www.geneontology.org/go#GO:0009653" />
      <go:part_of rdf:resource="http://www.geneontology.org/go#GO:0048513" />
    </go:term>
  </rdf:RDF>
</go:go>
```

RDF Container elements

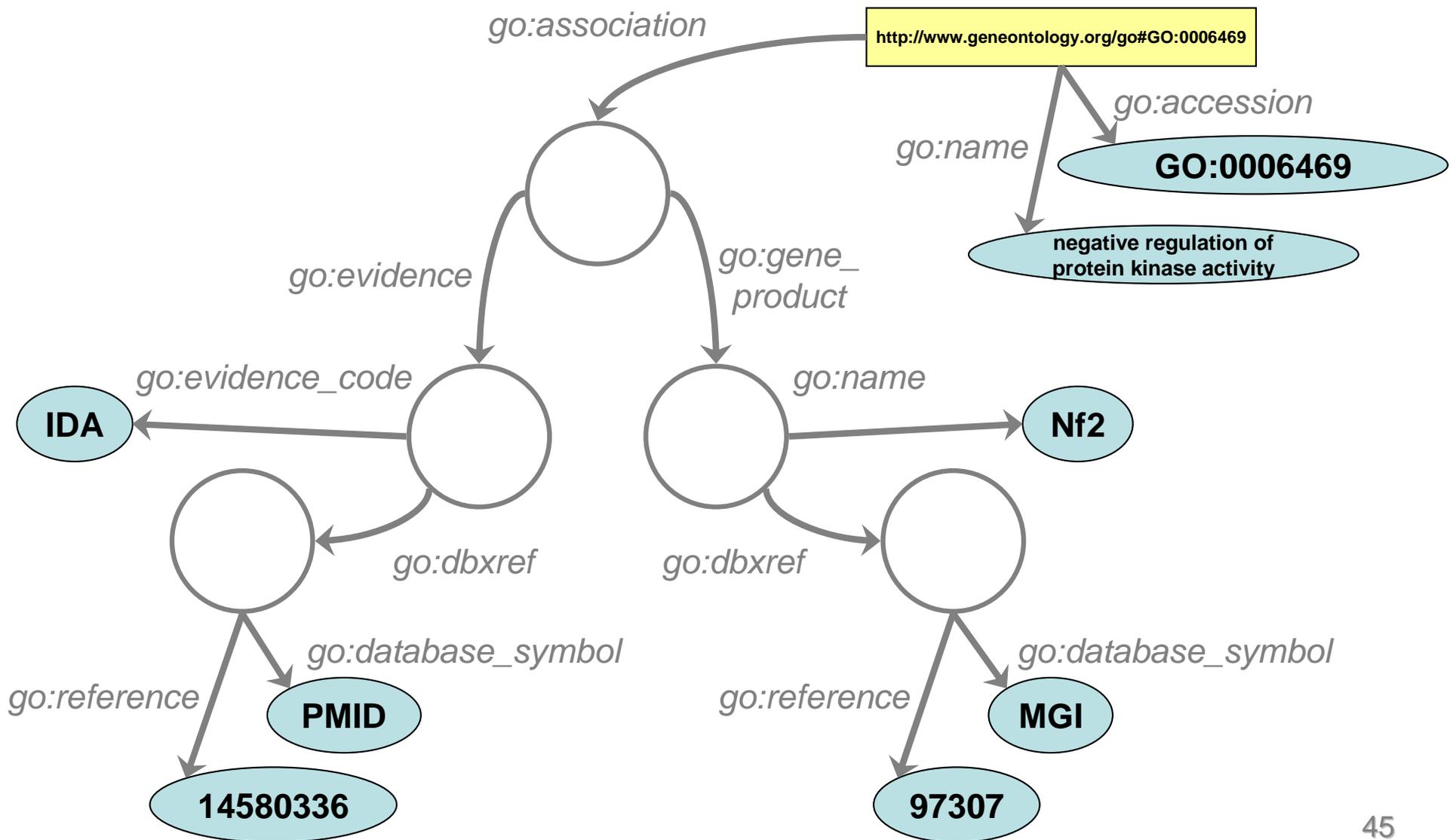
- **rdf:Bag**
 - Unordered container
- **rdf:Seq**
 - Ordered container
- **rdf:Alt**
 - Set of alternatives

RDF Blank nodes

- No URI associated
- Local scope
- n-ary relations

```
[...]
<go:association rdf:parseType="Resource">
  <go:evidence evidence_code="IDA">
    <go:dbxref rdf:parseType="Resource">
      <go:database_symbol>PMID</go:database_symbol>
      <go:reference>14580336</go:reference>
    </go:dbxref>
  </go:evidence>
  <go:gene_product rdf:parseType="Resource">
    <go:name>Nf2</go:name>
    <go:dbxref rdf:parseType="Resource">
      <go:database_symbol>MGI</go:database_symbol>
      <go:reference>MGI:97307</go:reference>
    </go:dbxref>
  </go:gene_product>
</go:association>
[...]
```

RDF Blank nodes



RDF schema (RDFS)

- Classes and Subclasses
 - #Glycoprotein subclass of #Protein
- Properties and Subproperties
 - has_update_date subproperty of has_date
- Domain and range (for properties)
 - encodes
 - Domain: gene
 - Range: protein

Querying RDF

- Multiple query languages
 - RQL, RDQL, SPARQL, ...

e.g., <http://www.w3.org/TR/rdf-sparql-query/>

- Example of SPARQL query

```
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix : <http://example.org/book/> .
@prefix ns: <http://example.org/ns#> .

:book1 dc:title "SPARQL Tutorial" .
:book1 ns:price 42 .
:book2 dc:title "The Semantic Web" .
:book2 ns:price 23 .
```

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX ns: <http://example.org/ns#>
SELECT ?title ?price
WHERE
  { ?x ns:price ?price .
    FILTER ( ?price < 30 ) .
    ?x dc:title ?title .
  }
```

<http://www.w3.org/TR/rdf-sparql-XMLres/>

RDF Inference rules

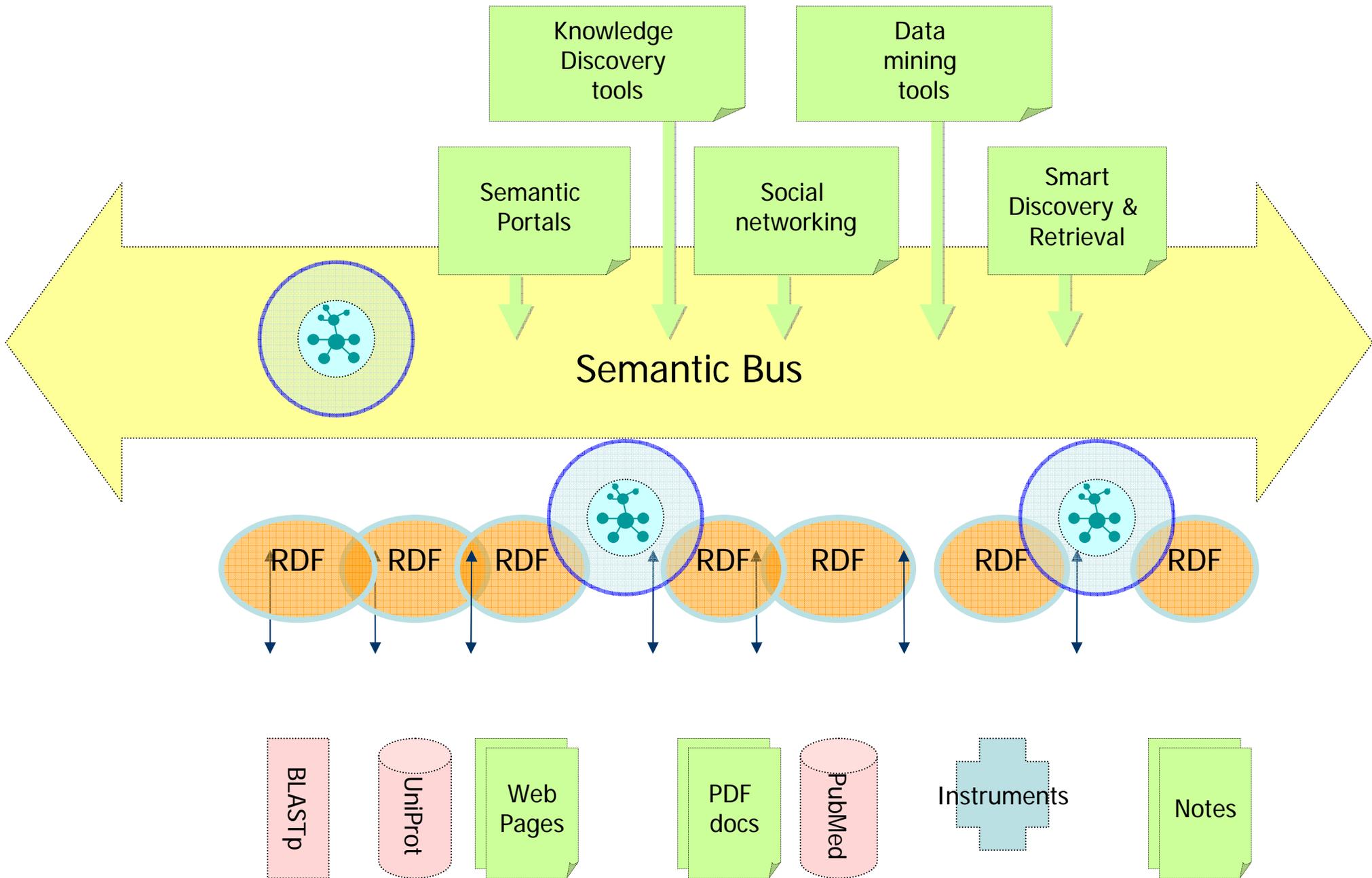
- Transitivity of isa
 - Infer $x \text{ isa } z$ from $x \text{ isa } y$ and $y \text{ isa } z$
- Transitivity of part_of
 - Infer $x \text{ part_of } z$ from $x \text{ part_of } y$ and $y \text{ part_of } z$
- isa and part_of combinations
 - Infer $x \text{ part_of } z$ from $x \text{ part_of } y$ and $y \text{ isa } z$
 - Infer $x \text{ part_of } z$ from $x \text{ isa } y$ and $y \text{ part_of } z$

Storing RDF

- RDF data management system
 - RDF store
 - Triple store
- Several implementations available
 - Publicly available
 - e.g., Sesame <http://www.openrdf.org/>
 - Proprietary
 - e.g., Oracle (from 10g) <http://www.oracle.com/>

Resources available in RDF

- Few resources currently available in RDF
- Many databases being converted to RDF
 - **UniProt** <http://expasy3.isb-sib.ch/~ejain//rdf/>
 - **NCBI Entrez Gene** <http://mor.nlm.nih.gov/pubs/alum/2006-sahoo.pdf>
- Examples of projects using RDF
 - **SWAN** (Semantic Web Application in Neuromedicine) [\[Gao et al, JWS 2006\]](#)
 - **SenseLab** (integrating neuroscience databases) <http://senselab.med.yale.edu/>



Courtesy of
R. Stevens

An RDF world

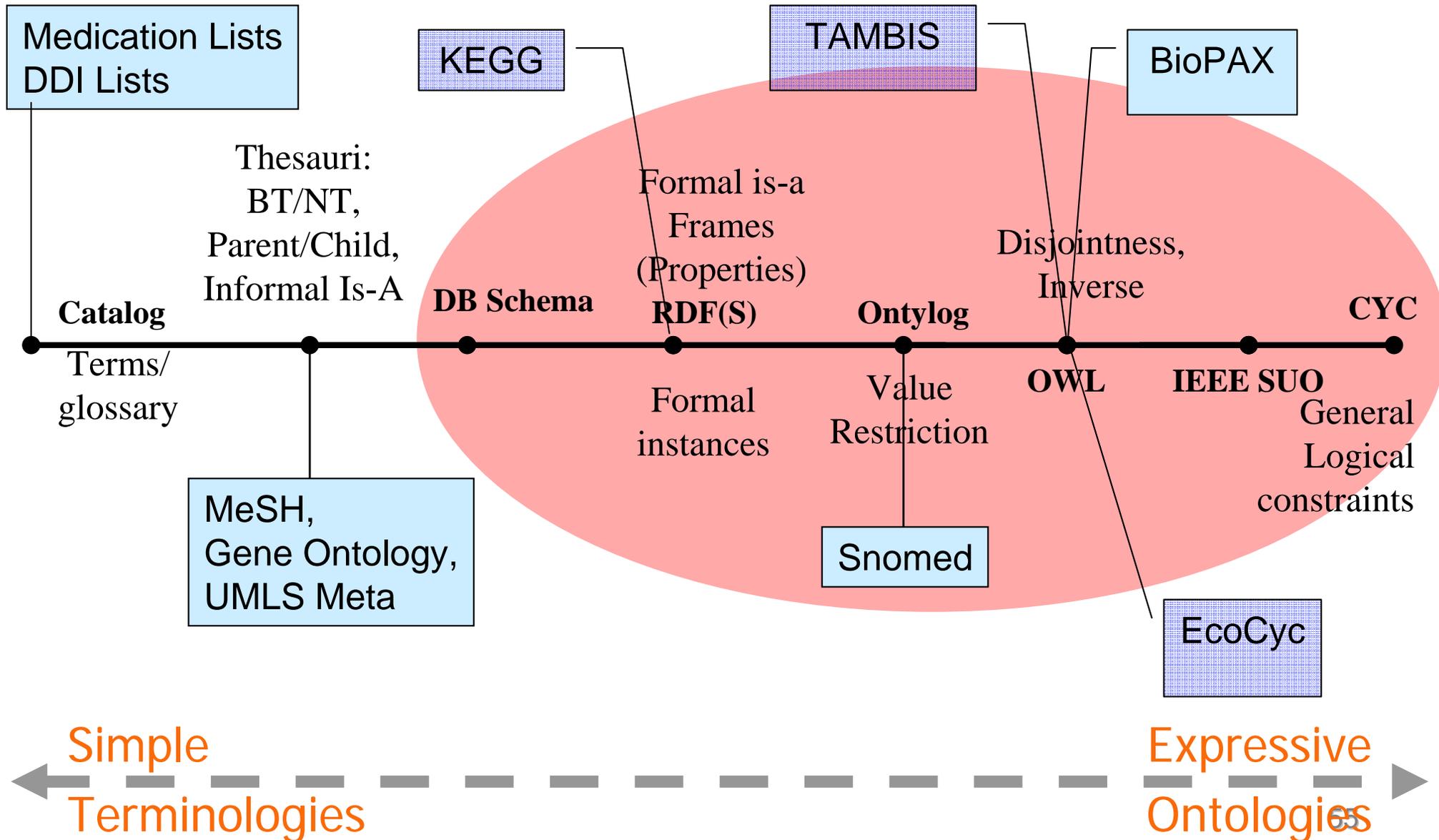
- Distributed heterogeneous resources present their data as RDF
- A common data model for a sea of data
- A “bus” into which resources can plug
- Common, syntax, common data model
- But no common vocabulary for values on the bus
- Also need vocabularies from ontologies
- Build ontology is the Web Ontology Language (OWL) and use via RDF Schema

Ontologies and Web Ontology Language (OWL)

OWL Introduction

- History: DAML + OIL = OWL (2001)
 - DAML – DARPA Agent Markup Language (1999)
 - OIL – Ontology Inference Layer (1997)
- Based on RDF(S)
- Added features, mostly related to identity
 - Restrictions
- Three flavors of increasing expressiveness, but decreasing tractability
 - OWL Lite
 - OWL DL (used for most applications)
 - OWL Full

The Knowledge Semantics Continuum



Ontology Dimensions based on McGuinness and Finin

OWL DL Example

- Class: *Benign intracranial meningioma*
in the NCI Thesaurus <http://cancer.gov/cancerinfo/terminologyresources/>

```
<owl:Class rdf:ID="Benign_Intracranial_Meningioma">
  <rdfs:label>Benign Intracranial Meningioma</rdfs:label>
  <code>C5133</code>
  <owl:equivalentClass>
    <owl:Class>
      <owl:intersectionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Benign_Intracranial_Neoplasm"/>
        <owl:Class rdf:about="#Benign_Meningioma"/>
        <owl:Class rdf:about="#Intracranial_Meningioma"/>
      </owl:intersectionOf>
    </owl:Class>
  </owl:equivalentClass>
  <Preferred_Name>Benign Intracranial Meningioma</Preferred_Name>
  <Semantic_Type>Neoplastic Process</Semantic_Type>
  <dSynonym>Benign Intracranial Meningioma</dSynonym>
  [...]
  <NCI_META_CUI>CL006955</NCI_META_CUI>
</owl:Class>
```

OWL Class Constructors

Constructor	DL Syntax	Example	Modal Syntax
intersectionOf	$C_1 \sqcap \dots \sqcap C_n$	Human \sqcap Male	$C_1 \wedge \dots \wedge C_n$
unionOf	$C_1 \sqcup \dots \sqcup C_n$	Doctor \sqcup Lawyer	$C_1 \vee \dots \vee C_n$
complementOf	$\neg C$	\neg Male	$\neg C$
oneOf	$\{x_1\} \sqcup \dots \sqcup \{x_n\}$	{john} \sqcup {mary}	$x_1 \vee \dots \vee x_n$
allValuesFrom	$\forall P.C$	\forall hasChild.Doctor	$[P]C$
someValuesFrom	$\exists P.C$	\exists hasChild.Lawyer	$\langle P \rangle C$
maxCardinality	$\leq_n P$	≤ 1 hasChild	$[P]_{n+1}$
minCardinality	$\geq_n P$	≥ 2 hasChild	$\langle P \rangle_n$

OWL Axioms

Axiom	DL Syntax	Example
subClassOf	$C_1 \sqsubseteq C_2$	Human \sqsubseteq Animal \sqcap Biped
equivalentClass	$C_1 \equiv C_2$	Man \equiv Human \sqcap Male
disjointWith	$C_1 \sqsubseteq \neg C_2$	Male $\sqsubseteq \neg$ Female
sameIndividualAs	$\{x_1\} \equiv \{x_2\}$	{President_Bush} \equiv {G_W_Bush}
differentFrom	$\{x_1\} \sqsubseteq \neg\{x_2\}$	{john} $\sqsubseteq \neg$ {peter}
subPropertyOf	$P_1 \sqsubseteq P_2$	hasDaughter \sqsubseteq hasChild
equivalentProperty	$P_1 \equiv P_2$	cost \equiv price
inverseOf	$P_1 \equiv P_2^-$	hasChild \equiv hasParent ⁻
transitiveProperty	$P^+ \sqsubseteq P$	ancestor ⁺ \sqsubseteq ancestor
functionalProperty	$\top \sqsubseteq \leq 1P$	$\top \sqsubseteq \leq 1$ hasMother
inverseFunctionalProperty	$\top \sqsubseteq \leq 1P^-$	$\top \sqsubseteq \leq 1$ hasSSN ⁻

Borrowed from Tutorial on OWL by Bechhofer, Horrocks and Patel-Schneider
<http://www.cs.man.ac.uk/~horrocks/ISWC2003/Tutorial/>

Existential vs. universal quantification

- Existential quantification
 - owl:someValuesFrom
 - Necessary condition
 - E.g., migraine = headache & has_symptom throbbing pain [only if one-sided]
- Universal quantification
 - owl:allValuesFrom
 - Necessary and sufficient condition
 - E.g., heart disease = disease & located_to heart

OWL reasoners

- For OWL DL, not OWL Full
- Reasoners
 - Fact++ <http://owl.man.ac.uk/factplusplus/>
 - Pellet <http://www.mindswap.org/2003/pellet/>
 - RacerPro <http://www.racer-systems.com/>
- Functions
 - Consistency checking
 - Automatic classification

OWL Reasoners Details

- **CEL**
 - Polynomial time classifier for the description logic EL+
 - EL+ is specially geared towards biomedical ontologies
- **Cerebra**
 - Commercial C++ reasoner, Support for OWL-API
 - Tableaux based reasoning for TBoxes and ABoxes
- **Fact++**
 - Free open source reasoner for DL reasoning
 - Support for Lisp API and OWL API
- **KAON2**
 - Free Java based DL reasoner with support for SWRL fragment
 - Support for DIG API
- **MSPASS**
 - A generalized theorem prover for numerous logics, also works for DLs
- **Pellet**
 - Free open source Java based reasoner for DLs
 - Support for OWL, DIG APIs and Jena Interface
- **RacerPro**
 - Commercial lisp based reasoner for DLs
 - Support for OWL APIs and DIG APIs

Editing OWL ontologies

The screenshot displays the Protégé 3.2 beta interface for editing an OWL ontology. The main window is titled "Thesaurus Protégé 3.2 beta" and shows the "CLASS EDITOR" for the class "Benign_Intracranial_Meningioma".

Subclass Explorer: Shows the asserted hierarchy for the project "Thesaurus". The class "Benign_Intracranial_Meningioma" is highlighted under the "Intracranial_Meningioma" class.

Class Editor: Shows the class "Benign_Intracranial_Meningioma" (instance of owl:Class). The "Annotations" tab is active, displaying a table of properties and their values:

Property	Value	Lang
ccode	C5133	
d\$synonym	Benign Intracranial Meningioma	
FULL_SYN	<term-name>Benign Intracranial Meningioma</term-name><term-group>PT</term-group><term-source>NCI</term-source>	
NCI_META_CUI	CL006955	
Preferred_Name	Benign Intracranial Meningioma	
rdfs:label	Benign Intracranial Meningioma	
Semantic_Type	Neoplastic Process	

Asserted Conditions: Shows the class hierarchy and associated conditions for "Benign_Intracranial_Meningioma". The conditions are listed as follows:

- NECESSARY & SUFFICIENT: Benign_Intracranial_Neoplasm, Benign_Meningioma, Intracranial_Meningioma
- NECESSARY: Disease_Excludes_Abnormal_Cell **only** Malignant_Cell [from Benign_Neoplasms_of_the_Meninges], Disease_Excludes_Primary_Anatomic_Site **only** Spinal_Cord [from Intracranial_Neoplasm], Disease_Has_Abnormal_Cell **only** Neoplastic_Cell [from Neoplasm], Disease_Has_Abnormal_Cell **only** Neoplastic_Meningothelial_Cell [from Meningothelial_Cell_Neoplasm], Disease_Has_Associated_Anatomic_Site **only** Central_Nervous_System [from Central_Nervous_System_Disorder], Disease_Has_Associated_Anatomic_Site **only** Nervous_System [from Nervous_System_Disorder], Disease_Has_Associated_Anatomic_Site **only** Meninges [from Meningeal_Neoplasm], Disease_Has_Finding **only** Slow_Growing_Mass [from Benign_Meningioma]
- DISJOINTS: Disjoint

The interface also includes a menu bar (File, Edit, Project, OWL, Code, Tools, Window, Help), a toolbar, and a status bar at the bottom with "Log c View" and "Properties View" options.

Resources available in OWL

- Many resources currently available in OWL
 - Gene Ontology <http://www.geneontology.org/>
 - NCI Thesaurus <http://cancer.gov/cancerinfo/terminologyresources/>
- Many projects using OWL
 - e.g., BioPax <http://www.biopax.org/>

OBO format

http://www.godatabase.org/dev/doc/obo_format_spec.html

- Used to represent many ontologies in the OBO family (Open Biological Ontologies)

<http://obo.sourceforge.net/>

- Essentially a subset of OWL DL

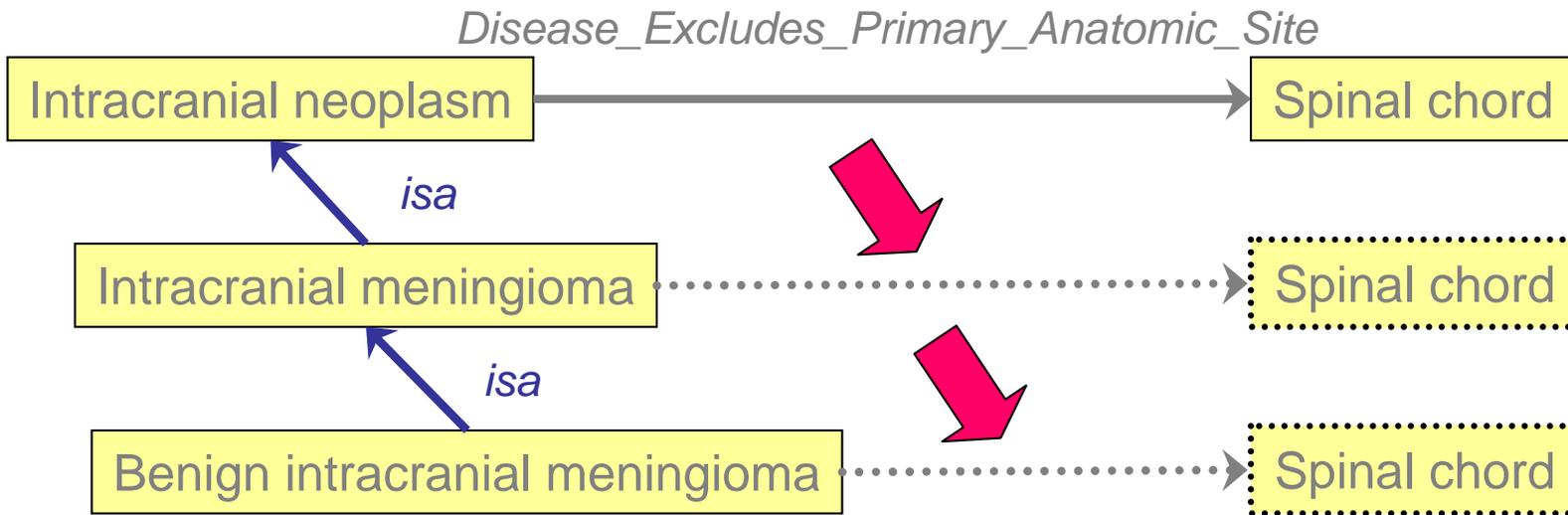
```
[Term]
id: GO:0019563
name: glycerol catabolism
namespace: biological_process
def: "The chemical reactions and pathways resulting in the breakdown of glycerol ..."
subset: gosubset_prok
exact_synonym: "glycerol breakdown" []
exact_synonym: "glycerol degradation" []
xref_analog: MetaCyc:PWY0-381
is_a: GO:0006071 ! glycerol metabolism
is_a: GO:0046174 ! polyol catabolism
```

Logic and Rule languages

Introduction

- Ontologies represent knowledge
- Automated reasoners infer conclusions from the given knowledge
 - Make implicit knowledge explicit
 - Help validate the ontology (e.g., consistency checking and automatic classification in DL)
- Need for more expressive logic
 - Inference rules

Simple inference



The screenshot shows a software interface with a class hierarchy on the left and a list of asserted conditions on the right.

Class Hierarchy (Left):

- Neoplasm
 - Hemangioblastoma_of_the_Central_Nervous_System
 - Hemangiopericytoma_of_the_Central_Nervous_System
 - Intracranial_Neoplasm
 - Adult_Intracranial_Neoplasm
 - Benign_Intracranial_Neoplasm
 - Brain_Neoplasm
 - Childhood_Intracranial_Neoplasm
 - Intracranial_Meningioma
 - Adult_Brain_Meningioma
 - Anaplastic_Malignant_Intracranial_Meningioma
 - Benign_Intracranial_Meningioma
 - Cerebello_Papillary_Meningioma
 - Childhood_Brain_Meningioma
 - Choroid_Plexus_Meningioma

Asserted Conditions (Right):

Condition	Source	Assertion
Benign_Intracranial_Neoplasm		NECESSARY & SUFFICIENT
Benign_Meningioma		=
Intracranial_Meningioma		NECESSARY
Disease_Excludes_Abnormal_Cell only Malignant_Cell	[from Benign_Neoplasms_of_the_Meninges]	INHERITED
Disease_Excludes_Primary_Anatomic_Site only Spinal_Cord	[from Intracranial_Neoplasm]	C
Disease_Has_Abnormal_Cell only Neoplastic_Cell	[from Neoplasm]	C
Disease_Has_Abnormal_Cell only Neoplastic_Meningothelial_Cell	[from Meningothelial_Cell_Neoplasm]	C
Disease_Has_Associated_Anatomic_Site only Central_Nervous_System	[from Central_Nervous_System_Disorder]	C
Disease_Has_Associated_Anatomic_Site only Nervous_System	[from Nervous_System_Disorder]	C
Disease_Has_Associated_Anatomic_Site only Meninges	[from Meningeal_Neoplasm]	C
Disease_Has_Finding only Slow_Growing_Mass	[from Benign_Meningioma]	C

Complex inference

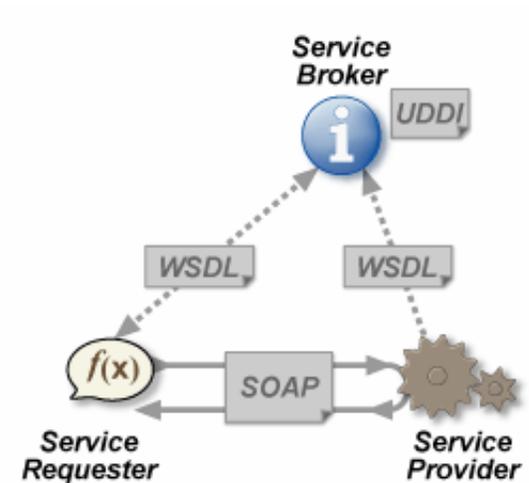
- Clinical decision support
 - If patient is treated by aminoglycosides and patient has impaired renal function then reduce dose (or frequency of administration) of aminoglycosides
- Not directly supported by DL reasoners
- Require rule languages
 - RuleML <http://www.ruleml.org/>
 - SWRL (Semantic Web Rule Language)

<http://www.w3.org/Submission/SWRL/>

Web services

Introduction

- Software system designed to support interoperable machine-to-machine interaction over a network
 - Services can be discovered
 - Service description (WSDL)
 - Standard communication mechanism
 - Protocol (http, SOAP, UDDI)
 - Data format (XML)



Web Services Benefits

- Standardized protocols for I/O
- Cross-platform, language independent
- Automated architecture for high throughput analyses, querying
- Automated service discovery
- Integration with Semantic Web technologies

Web Services Technology

- **WSDL – Web Service Description Language**
 - XML format for describing service interfaces
<http://www.w3.org/TR/wsdl>
- **SOAP – Simple Object Access Protocol**
 - Protocol for exchanging XML-based messages over the Internet (http)
 - Client/Server protocol: Remote Procedure Call (RPC)
<http://www.w3.org/TR/soap/>
- **UDDI – Universal Description, Discovery, and Integration**
 - Service registry
<http://www.uddi.org/>

Web Services and ontology

- **WSDL-S – Web Service Semantics**

<http://www.w3.org/Submission/WSDL-S/>

- **OWL-S (OWL-based Web service ontology)**

<http://www.daml.org/services/owl-s/1.0/>

- **WSMO – Web Service Modeling Ontology**

<http://www.wsmo.org/>

Examples of Web Services in biomedicine

- **BioMOBY** <http://biomoby.open-bio.org/>
 - Biomedical Web Services registry
 - Ontology-based messaging standard
 - Discover and interact with task-appropriate biological data and analytical service providers
- **Entrez Utilities Web Service**
 - Access NCBI's Entrez Utilities via SOAP

http://www.ncbi.nlm.nih.gov/entrez/query/static/esoap_help.html

Part 3

Semantic Web applications in biomedicine

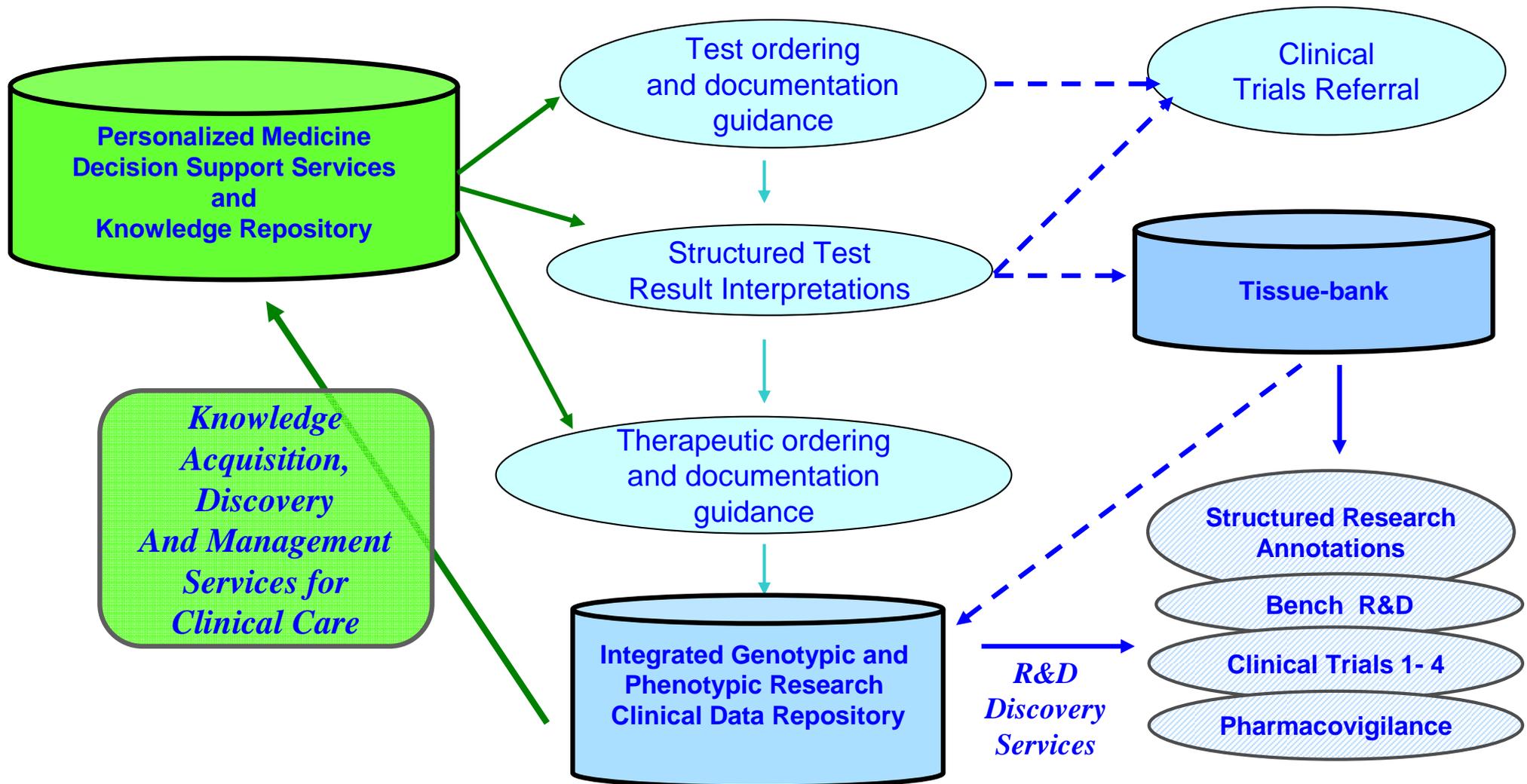
Overview

- Use Case Flow
 - Need for Shared Semantics in the context of Personalized Medicine
- Functional Requirements
 - Metadata-based Annotation
 - Semantic Data Integration
 - Ontology Driven Clinical Decision Support
 - Knowledge Change and Provenance

Use Case Flow based on Shared Semantics

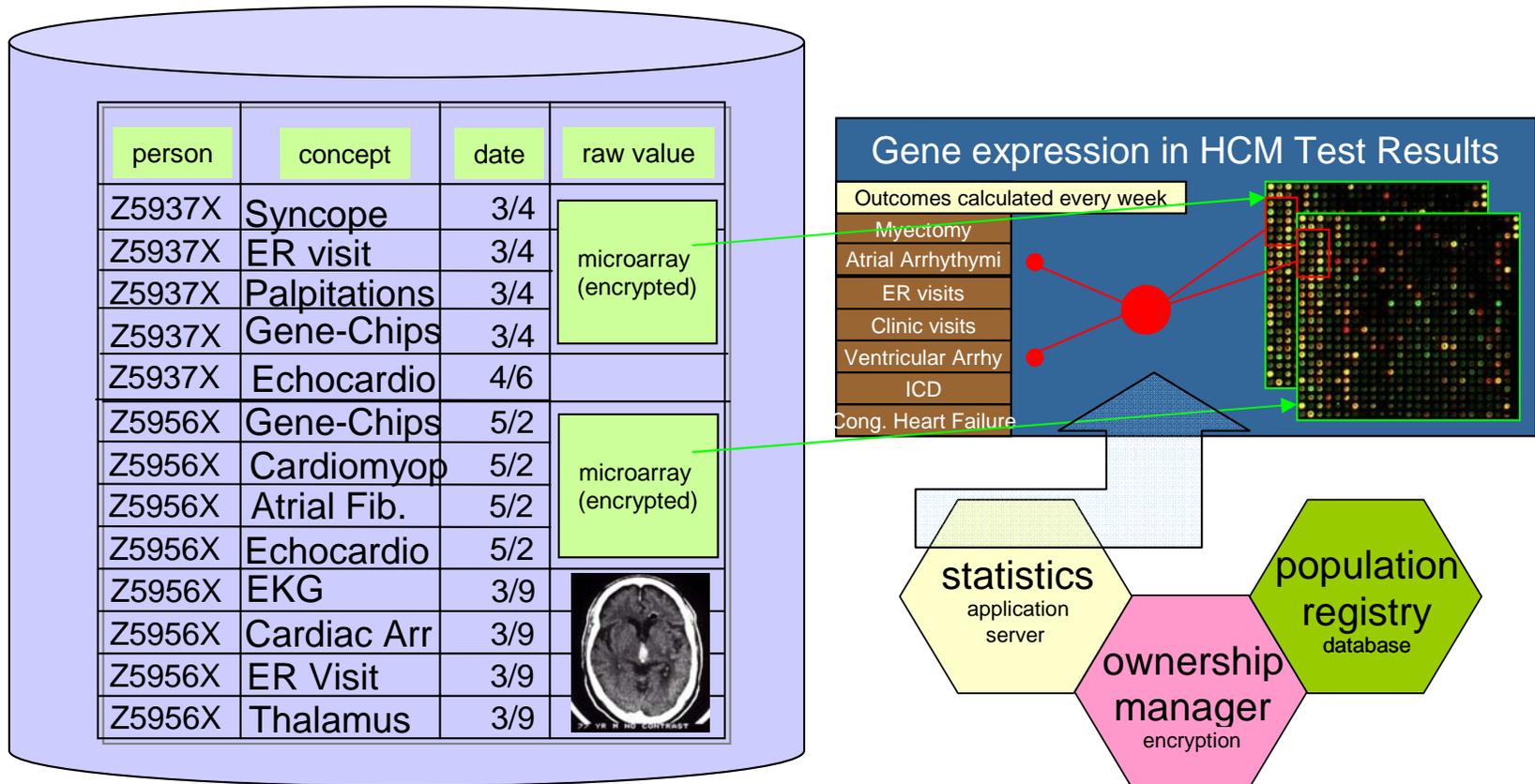


Patient Encounter



Need for Metadata Annotations

Connecting Dx, Rx, Outcomes and Prognosis Data to Genotypic Data for Cardiomyopathy



Need for Data Integration

- Need for integrating Clinical and Molecular Diagnostics Data
- Integrated Genotypic-Phenotypic Research Clinical Data Repository

Need for Clinical Decision Support

LMR Smart Form - Microsoft Internet Explorer provided by Partners HealthCare System

Address: http://is.partners.org/prototype/dtaylor/marchdmsf/smartform3.htm

Patient: GUBERNATH, JANETTA
MRN: 0000001 (MGH) DOB: 08/25/1947 Age: 58 Sex: F Tel: (H) 617-555-1212

Conditions to include: DM CAD ARI

Summary Graphs Note Preview Patient-friendly

Carry forward all note content from: Most recent note 3/18/05

	Order	12/20/04	11/10/04	9/14/04	6/4/03
Glucose (mg/dL)	<input type="checkbox"/>	175	185	-	145
A1C (4.4%-6.4%)	<input type="checkbox"/>	8.1	8.2	-	8.3
Total Chol (<200 mg/dL)	<input type="checkbox"/>	210	-	240	190
HDL (>45 mg/dL)	<input type="checkbox"/>	59	-	60	53
LDL (<100 mg/dL)	<input type="checkbox"/>	112	-	115	118
Trig (<200 mg/dL)	<input type="checkbox"/>	125	-	125	125
SGOT (U/L)	<input type="checkbox"/>	10	7	-	-
BUN (mg/dL)	<input type="checkbox"/>	12	11	-	-
Cr (mg/dL)	<input type="checkbox"/>	1.2	1.1	-	-
ualb / Cr ratio (mg Ald/g Cr)	<input type="checkbox"/>	-	-	-	-

Needs Attention

- A1C high (8.1 on 12/20/04)
- LDL high (112 on 12/20/04)
- Total Chol high (210 on 12/20/04)
- ualb / Cr overdue (Last 11/9/03)
- Foot exam overdue (Last 4/14/03)
- Eye exam overdue (Last 10/6/03)
- Need current BP
- BMI high (34.0 today)
- Home glucose monitoring not documented

A1C

A1C high (8.1 on 12/20/04)

- ◆ Adjust glycemic therapy
- Refer to CDE
- Have patient report AM FBG after first 3-5 days
- Patient ed: *What is insulin?*
- Patient ed: *Giving an insulin injection*

Cancel Complete

Echo triggers guidance to screen for possible mutations:

- MYH7, MYBPC3, TNN2, TNNI3, TPM1, ACTC, MYL2, MYL3

Need for Knowledge Maintenance

- Need for a Knowledge Repository to support Clinical and Genomic Decision Support
- However:
 - Clinical Knowledge changes over time, e.g., value ranges for Clinical Normality
 - New knowledge is created, e.g., new molecular diagnostic test hits the market.
- Need for rapid knowledge change and maintenance

*The first step of any biomedical activity
(research, practice, knowledge gathering)
should be on the computer!*

Overview

- Use Case Flow
 - Need for Shared Semantics in the context of Personalized Medicine
- Functional Requirements
 - Metadata-based Annotation
 - Data Integration
 - Clinical Decision Support
 - Knowledge Change and Provenance

Metadata-based Semantic Annotations: Connotea

Connotea: Bookmarks with search terms cox-2 and inhibitors - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Refresh Print Mail Print Preview Print Setup Print... Search Favorites Refresh Print Mail Print Preview Print Setup Print... PageRank 7524 blocked Check AutoLink AutoFill Options Connotea

Address http://www.connotea.org/search?q=cox-2+inhibitors Go Links

Google Connotea Search PageRank 7524 blocked Check AutoLink AutoFill Options Connotea

Connotea logged in as vipul_kashyap Logout

Search All cox2 inhibitors Search My Library | Registration

Home | Latest News | About This Site | Site Guide | FAQ | Community Pages | Popular Links | Recent Activity | Contact Us

Tags used on these bookmarks:

- therapy
- Medical Sciences
- Pharmacology
- COX-2 inhibitors
- viral
- figure
- Pregnancy
- HIV
- network
- COX inhibitors
- Amino Acids - Protei...
- Sonya's Database Que...
- rheumatoid arthritis
- mammalian protein ki...
- US GOV
- oligohydrannios
- Gene Bank -Genomics ...
- Drug resistance
- active rheumatoid ar...
- immunology
- molecular research
- fetus
- Capsid Proteins
- Zoonoses-vector jump...
- Economics and or Tra...
- Ecological Sciences
- ulcers
- 2005
- Diagnosis and screen...
- review

Bookmarks with search terms cox-2 and inhibitors EXPORT LIST RSS ?

Note: Your search term matches the global tag [COX-2 inhibitors](#).

Number of bookmarks per page: 10 | 25 | 50 | 100

[copy](#)

[Replication-selective virotherapy for cancer: Biological principles, risk management and future directions \(info\)](#)
David Kirn, Robert Martuza, and James Zwiebel
Nat Med **7** (7), 781-7 (Jul 2001)
[doi:10.1038/89901](#)
Posted by [madhu](#) to [virotherapy](#) on [Mon Jan 30 2006](#) at 15:23 UTC

[copy](#)

[COX-2 inhibitors and metabolism of essential fatty acids. \(info\)](#)
Undurti N Das
Med Sci Monit **11** (7), RA233-7 (Jul 2005)
[PMID: 15990700](#)
Posted by [ingerida](#) to [review 2005 metabolism PRINT-D F: Das U N COX-2 inhibitors pathway essential fatty acids network figure](#) on [Thu Sep 01 2005](#) at 09:25 UTC

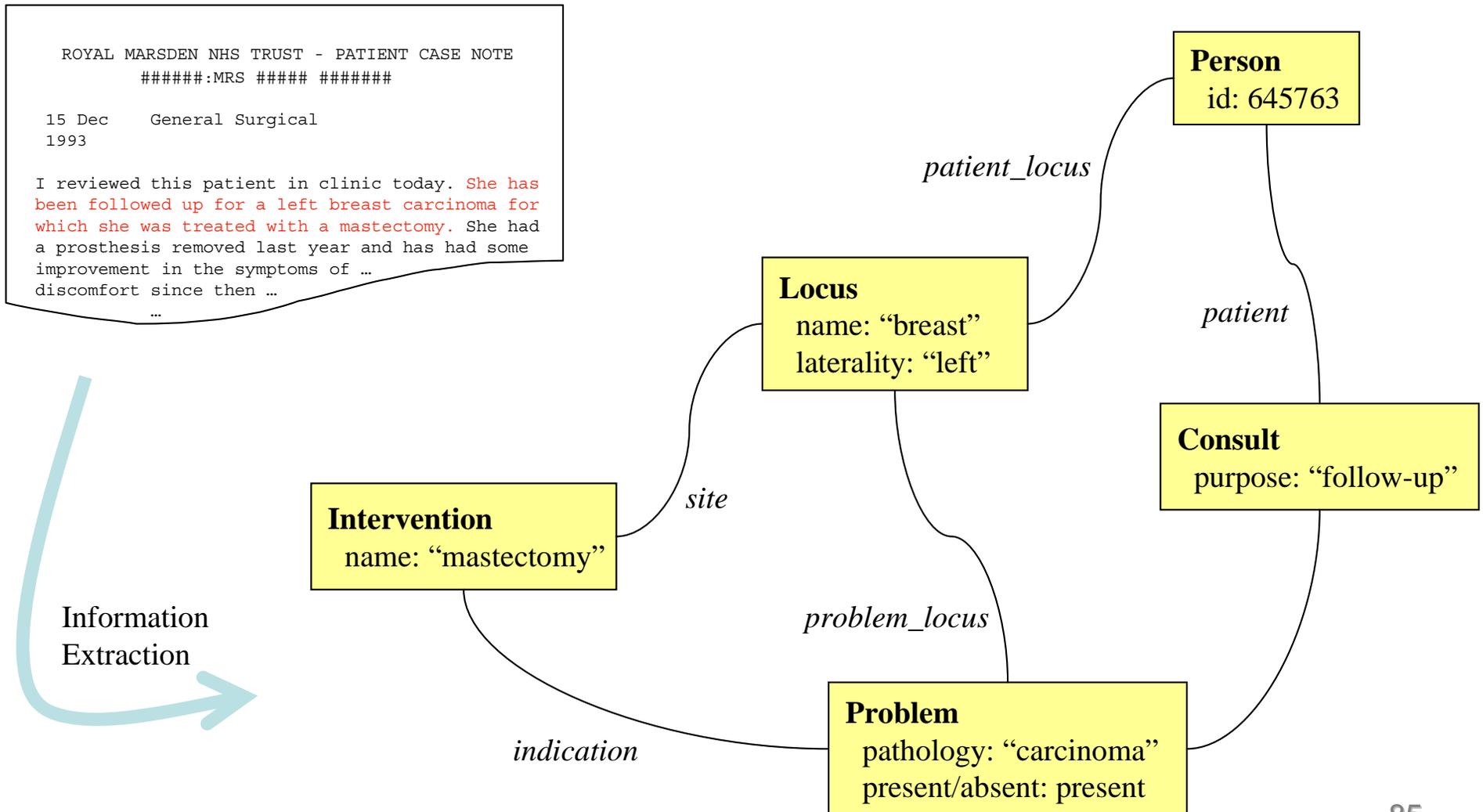
[copy](#)

[Is the use of COX-2 inhibitors in gastroenterology cost-effective? \(info\)](#)
David Graham and Francis KL Chan
Nat Clin Pract Gastroenterol Hepatol **1** (2), 60-1 (Dec 2004)
[doi:10.1038/ncpgasthep0043](#)
Posted by [NatureClinicalPractice](#) to [ulcers therapy](#) on [Thu Dec 08 2005](#) at 18:08 UTC

[copy](#)

[Genome Biology | Full text | Attacking pathogens through their hosts \(info\)](#)
Genome Biology **7** (1), 201 (2006)
[doi:10.1186/gb-2006-7-1-201](#)
.."This is partly because the disease burden of any one pathogen is unlikely to reach sufficient levels for pharmaceutical companies to justify the enormous cost of developing a new drug, which (although hotly debated) is estimated to be between \$0.5 billion and \$1.7 billion [4-6]. Although legislation over the past 20 years in the USA, especially the 'Orphan Drug Act', is designed to reduce such barriers, the difficulty of developing new antimicrobial drugs remains, and it is compounded by the fact that many infectious diseases requiring treatment occur in developing countries, which cannot cope with the costs of new drugs.."The concept of attacking the microbe by altering or augmenting a host-cell function or process is not new. The use of interferon alpha (IFNalpha) in combination with ribavirin in the treatment of hepatitis C virus infection is successful in 50% of infected

Metadata-based Semantic Annotations: Clinical E-Science Framework



Metadata-based Semantic Annotations HubMed

HubMed: cox-2 inhibitors - Microsoft Internet Explorer

Address: <http://www.hubmed.org/search.cgi?q=cox-2+inhibitors&sort=relevance>

Search: **cox-2 inhibitors** Results: 1-20 of 3293

sort by: date/relevance 0 Next 20

All Display Checked Abstracts Show Extracts

1 **Cox-2 inhibitors.**
Brown E
Physician Exec. 1999 Jan-Feb ; 25(1): 74-6
[Abstract](#) [FullText](#) [SFX](#) [Clip](#) [Export](#) [Related](#) [Cites](#) [Tag](#)

tags (space-separated)
 annotation

2 **COX-2 inhibitors.**
Brooks PM, Day RO
Med J Aust. 2000 Oct 16; 173(8): 433-6
[Abstract](#) [FullText](#) [SFX](#) [Clip](#) [Export](#) [Related](#) [Cites](#) [Tag](#)

3 **COX-2 inhibitors.**
Becker RC
Tex Heart Inst J. 2005; 32(3): 380-3
[Abstract](#) [FullText](#) [SFX](#) [Clip](#) [Export](#) [Related](#) [Cites](#) [Tag](#) [Review](#)

4 **COX-2 inhibitors.**
Hawkey CJ
Lancet. 1999 Jan 23; 353(9149): 307-14
[Abstract](#) [FullText](#) [SFX](#) [Clip](#) [Export](#) [Related](#) [Cites](#) [Tag](#)

5 **4,5-Diaryloxazole inhibitors of cyclooxygenase-2 (COX-2).**
Talley JJ, Bertenshaw SR, Brown DL, Carter JS, Graneto MJ, Koboldt CM, Masferrer JL, Norman BH, Rogier DJ, Zwwifel BS, Seibert K
Med Res Rev. 1999 May ; 19(3): 199-208

<http://www.hubmed.org/tags/edit/10387277> Internet

Metadata-based Semantic Annotations: HubMed

HubMed Tag Storage: demo - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address <http://www.hubmed.org/tags/users/demo>

HUBMED TAGS (BETA!)

Rat CD8+ FOXP3+ T suppressor cells mediate tolerance to allogeneic heart transplants, inducing PIR-B in APC and rendering the graft invulnerable to rejection. 2004
demo: [pir-b](#)

Tolerization of dendritic cells by T(S) cells: the crucial role of inhibitory receptors ILT3 and ILT4. 2002
demo: [ilt4](#)

Generation and function of antigen-specific suppressor and regulatory T cells.
demo: [ilt4](#) [specificity](#) [treg](#)

Manipulation of immune regulation in systemic lupus erythematosus. 2005
demo: [test](#)

Induction of allopeptide-specific human CD4+CD25+ regulatory T cells ex vivo. 2003
demo: [test](#)

Targeted CTLA-4 engagement induces CD4+CD25+CTLA-4high T regulatory cells with target (allo)antigen specificity. 2004
demo: [test](#)

The role of TCR specificity in naturally arising CD25+ CD4+ regulatory T cell biology. 2005
demo: [test](#)

CD4+ regulatory T cell responses induced by T cell vaccination in patients with multiple sclerosis. 2006
demo: [test](#)

Identification of a CD4+CD25+ T cell subset committed in vivo to suppress antigen-specific T cell responses without additional stimulation. 2004
demo: [cd134](#) [treg](#)

- [home](#)
- [search](#)
- [login](#)
- [register](#)
- [tag cloud](#)
- [recommendations](#)
- [→hubmed](#)

Internet

Metadata-based Semantic Annotations: Active Semantic EMR

AHC Acct No: 222222 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address C:\lex\final.xml

30606 Phone: 706-208-9700 Fax: 706-208-0806

Athens Heart Center
Shyam "Sham" Prabhakar
333Dogs Drive, Apt 604, ATHEHS, GA 30606
SSN: 222-22-2222 MR #: 222222 Sex: M DOB: 05/07/1970 Age: 33

Referred doctor from Practice Ontology

Visit on 07/29/2005

Office Visit Note - Complete History & Physical

Other Physicians: David Almand, M.D. E, Timothy Gibson, M.D. E, Alfredo Alarcon, E
Emergency Medicine (770) 922-3023 Family Practice 706-227-2027 (404) 256-5212

Problem List:

3. Backache unspecified E
4. Hypercholesterolemia E
5. Chest Pain E
 - A. Hypertension E
 - B. Shortness of Breath.
6. Dilated cardiomyopathy gthghjg
7. Abnormal ECG
9. Acute Glomerulonephritis with other specified Pathological Lesion in Kidney E
10. Something
11. Chest Pain E

Chief Complaint: Follow up of abdominal aortic aneurysm, angina, aortic stenosis, aortic valve replacement, dental clearance, and atrial fibrillation status post abnormal stress test. Cardiac clearance for aneurysm removal. Follow up of recent hospitalization at BJC - Commerce for atrial fibrillation.

History of Present Illness: Mr. Prabhakar is a 35 year old patient of Dr David Almand, Dr Timothy Gibson, and Dr Alfredo Alarcon. He was admitted to Ty Cobb Memorial by Dr. Alfredo Alarcon for bradycardia. He was found to have complete heart block. He was treated with mitral valve replacement. And he did not respond well. The patient was then transferred to Emory. He was admitted to Emory by Dr. Timothy Gibson for angina. He was found to have atrial fibrillation and complete heart block. He was treated with cholecystectomy and he responded well. The patient was then transferred to St. Mary's Hospital. He is here today for follow up management of arrhythmia, atrial fibrillation, and ICD function. Since his last visit new problems have developed. He is taking his medications as prescribed. There appear to be possibly some side effects related to the medications. Overall, he believes that his arrhythmia, atrial fibrillation, and ICD function is poorly controlled. He is here today for follow up management of cardiomegaly and coronary artery disease. Since his last visit no new problems have developed. He is taking his medications as prescribed. There appear to be possibly some side effects related to the medications. Overall, he believes that his cardiomegaly and coronary artery disease is stable. The chest pain is associated with itchy. He reports that his chest pain is aggravated by bending. The chest pain is relieved by belching. The patient reports this morning with her spell. Dizziness is associated with palpitations. He states that this symptom is aggravated by activity, bending, and high blood pressure. The dizziness is relieved by sitting down. He states that the palpitations are aggravated by position changes.

ICD9 codes from Diagnosis Procedure Ontology

Lexical annotation

Current Medications Medications After Visit

Done My Computer

Metadata-based Semantic Annotations: Active Semantic EMR

AHC Acct No: 222222 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address C:\lex\final.xml

Current Medications	Medication
Intropin injection 40mg/ml, 1 inj qd I E	Intropin injection 40mg/ml, 1 inj qd I E
Tasmar tablets 200mg, 1 tab qd I E	Advil 100mg/5ml, 1 susp qd F E
	Tasmar tablets 200mg, 1 tab qd I E
	Tylenol extended release 300mg, 1 tab er qd A F E

Pharmacy: Carson's Commerce Drug Company
Phone: 706-754-4128
Phone: 706-335-3111

Allergies: AMPICILLIN, IVP DYE, PENICILLIN, TYLENOL

Past Medical History: No past trauma.

Family History: Mr. Prabhakar has a positive family history of coronary artery disease. The patient's **crap2** is deceased.

Psychosocial: Mr. Prabhakar resides in an apartment home. He lives with her daughter to hkk1. He has excellent social support.

Life History:

Review of Systems:

General Patient reports daily chills associated with chest pain. These symptoms have not been worked up by his primary care physician. Patient reports recent unintentional weight gain. This problem has not been worked up by his primary care physician. Patient reports recent onset of severe trouble falling asleep. This problem has been working up by his primary care physician.

HEENT The patient reports migraine headache.

Genitourinary Patient denies dysuria. Patient complains of recent onset of hematuria. Patient complains of chronic presence of hesitancy. Patient denies dribbling. Patient complains of recent onset of burning with urination.

Hematologic Patient reports history of blood transfusion as a result of anemia. Patient reports he did experience a transfusion reaction.

Skin Patient reports frequent of mild pruritis associated with weakness. This problem has been worked up by the patient's dermatologist. Patient reports a walnut-sized keratosis that is located over the entire body.

Psychiatric Patient admits to a history of panic attacks that is currently managed by common mental health. His symptoms are felt to be not under control.

Vital Signs: Height: 170 Weight: 280 lbs BP: 120/80 Pulse: 80 Respirations: 20

Physical Examination:

General The patient appears the stated age.

Formulation Recommendation Using Insurance ontology

Drug Interaction using Drug Ontology

Drug Allergy

Done My Computer

Metadata-based Semantic Annotations: Active Semantic EMR

Explore: Drug *Tasmar*

AHC Acct No: 222222 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites

Address C:\lex\final.xml

Current Medications	Medications After Visit
Intropin injection 40mg/ml, 1 inj qd I E Tasmar tablets 200mg, 1 tab qd I E	Intropin injection 40mg/ml, 1 inj qd I E Advil 100mg/5ml, 1 susp qd F E Tasmar tablets 200mg, 1 tab qd I E Tylenol extended release 650mg, 1 tab er qd A F E

Pharmacy: Carson's Commerce Drug Company
Phone: 706-754-4128
Phone: 706-335-3111

Allergies: AMPICILLIN, IVP DYE, PENICILLIN, TYLENOL

Past Medical History: No past trauma.

Family History: Mr. Prabhakar has a positive family history of coronary artery disease. The patient's **crap2** is deceased.

Psychosocial: Mr. Prabhakar resides in a apartment home. He is lives with her daughter to hkk. He has excellent social support.

Life History:

Review of Systems:

General Patient reports daily chills associated with chest pain. These symptoms have not been worked up by his primary care physician. Patient reports recent unintentional weight gain. This problem has not been worked up by his primary care physician. Patient reports recent onset of severe trouble falling asleep. This problem has been working up by his primary care physician.

HEENT The patient reports migraine headache.

Genitourinary Patient denies dysuria. Patient complains of recent onset of hematuria. Patient complains of chronic presence of hesitancy. Patient denies dribbling. Patient complains of recent onset of burning with urination.

Hematologic Patient reports history of blood transfusion as a result of anemia. Patient reports he did experience a transfusion reaction.

Skin Patient reports frequent of mild pruritis associated with weakness. This problem has been worked up by the patient's dermatologist. Patient reports a walnut-sized keratosis that is located over the entire body.

Psychiatric Patient admits to a history of panic attacks that is currently managed by common mental health. His symptoms are felt to be not under control.

Vital Signs: Height: 170 Weight: 280 lbs BP: 120/80 Pulse: 80 Respirations: 20

Physical Examination:

General The patient appears the stated age.

Done My Computer

Metadata-based Semantic Annotations: Active Semantic EMR

Query View - DrugOnt_test_alt

antiparkinsonian_agents
neurological_agents
cpnum_group_2118
cont_inhibitors
Tolcapone
formulary_1498
Tasmar
interaction_with_cpnum_group_206_1
cpnum_group_206
interaction_with_cpnum_group_2118

classification
classification
classification
belongs to group
belongs to group
brand / generic

From: Tasmar
To: cpnum_group_2118
Type: has_cpnum_group

Semantic browsing and querying-- perform decision support (how many patients are using this class of drug, ...)

Nodes of Interest
 Use Selected (none selected)
 Search By Name
Name: cpnum_group_2118
 case sensitive
 regular expression
 Restrict Node Types
Defined Class
Enumeration
Individual
Logical Operation
Primitive Class
RDFS Class
Restriction

Arcs of Interest
 Restrict Arc Types
equivalent to
has instance
has subclass
has_brandname_equivalent
has_brandname_equivalent (Domain>Range)
has_brandname_equivalent (Necessary and Sufficient)
has_brandname_equivalent (Necessary)
has_brandname_prescription_drug
has_brandname_prescription_drug (Domain>Range)
has_brandname_prescription_drug (Necessary and Sufficient)
has_brandname_prescription_drug (Necessary)
has_cpnum_group
has_cpnum_group (Domain>Range)

Neighbours
 Show Neighbours
Arc Direction
 Outgoing and Incoming Arcs
 Outgoing Arcs
 Incoming Arcs
 Restrict Levels
Levels: 2

Results
Query

Overview

- Use Case Flow
 - Need for Shared Semantics in the context of Personalized Medicine
- Functional Requirements
 - Metadata-based Annotation
 - Data Integration
 - Clinical Decision Support
 - Knowledge Change and Provenance

Semantic Data Integration

- Create a focused ontology based on a collection of well defined use cases
- Develop and deploy “wrappers” that give an RDF-view of the underlying data and map it to ontological concepts.
- Specify mapping rules that associate data items across multiple RDF graphs

Semantic Data Integration: Ontology

OWL ontologies that blend knowledge from the Clinical and Genomic Domains

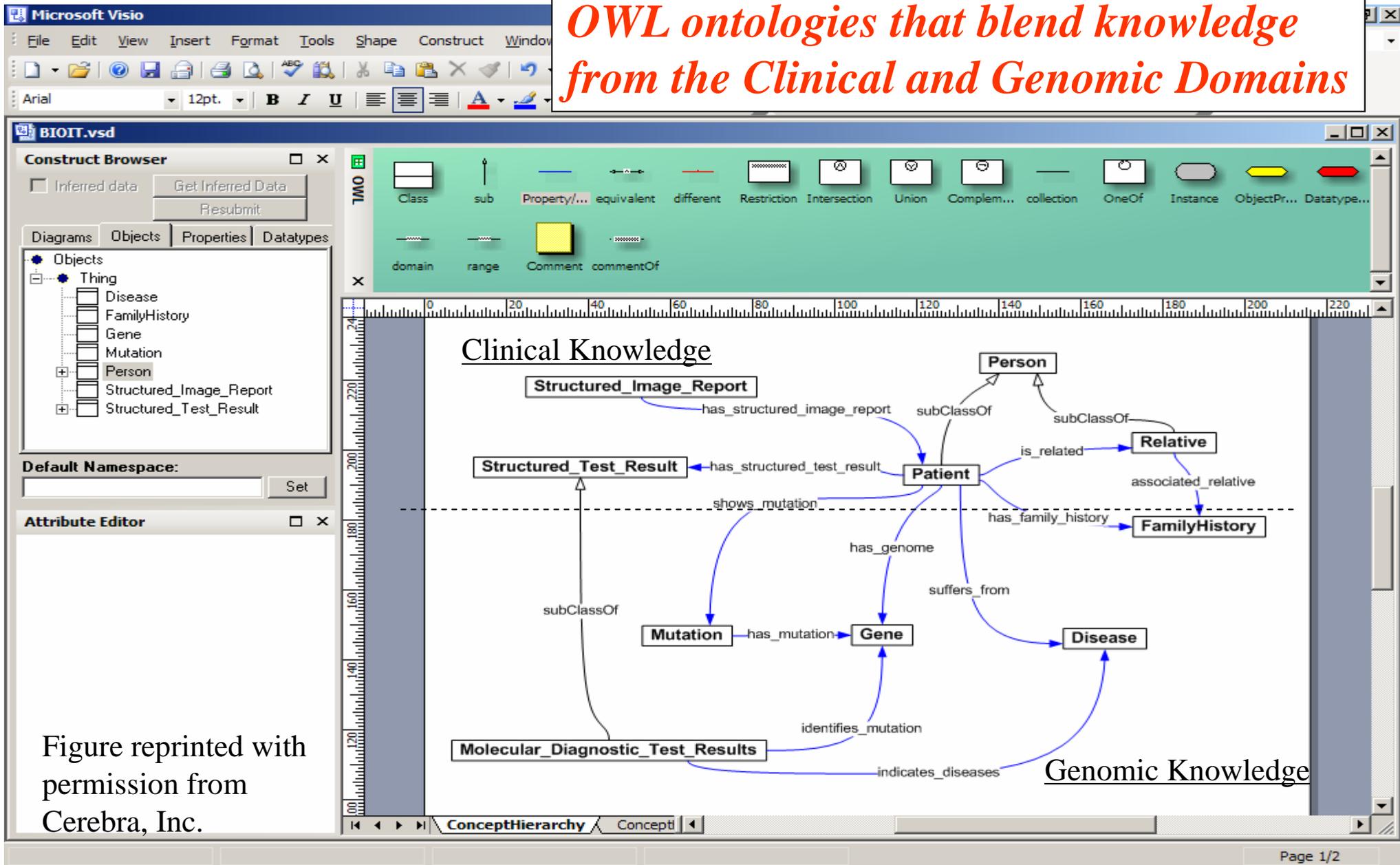
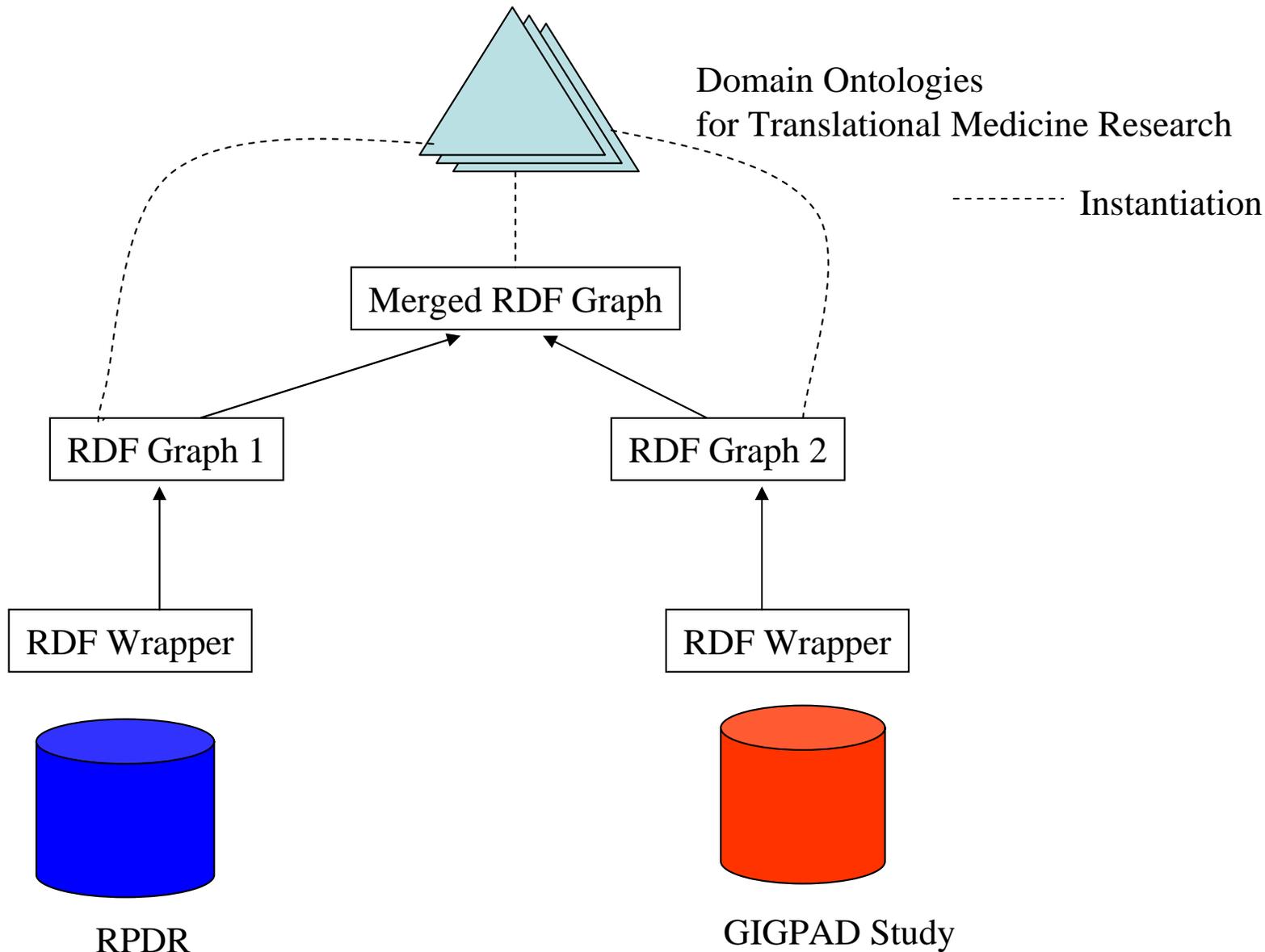


Figure reprinted with permission from Cerebra, Inc.

Semantic Data Integration: Architecture

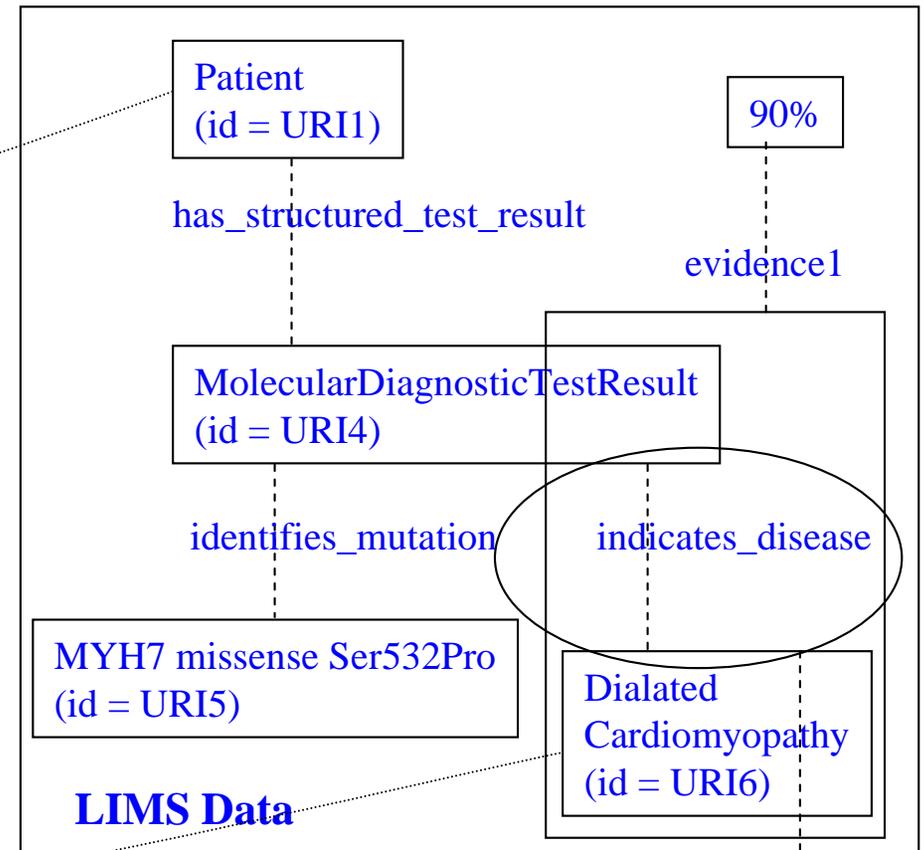
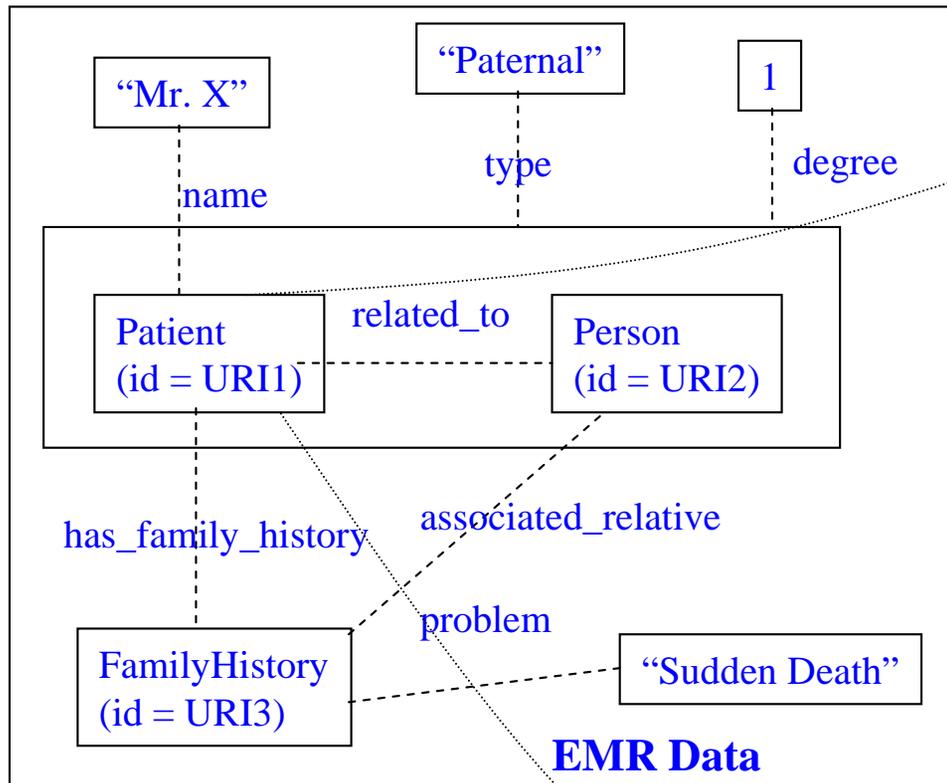


Semantic Data Integration: Incremental Roadmap

- Data assets remain as they are!
They do not need to be modified
- The wrapper abstracts out details related to location, access and data structure
- Integration happens at the information level
- Highly configurable and incremental process
- Ability to specify declarative rules and mappings for further hypothesis generation

Semantic Data Integration

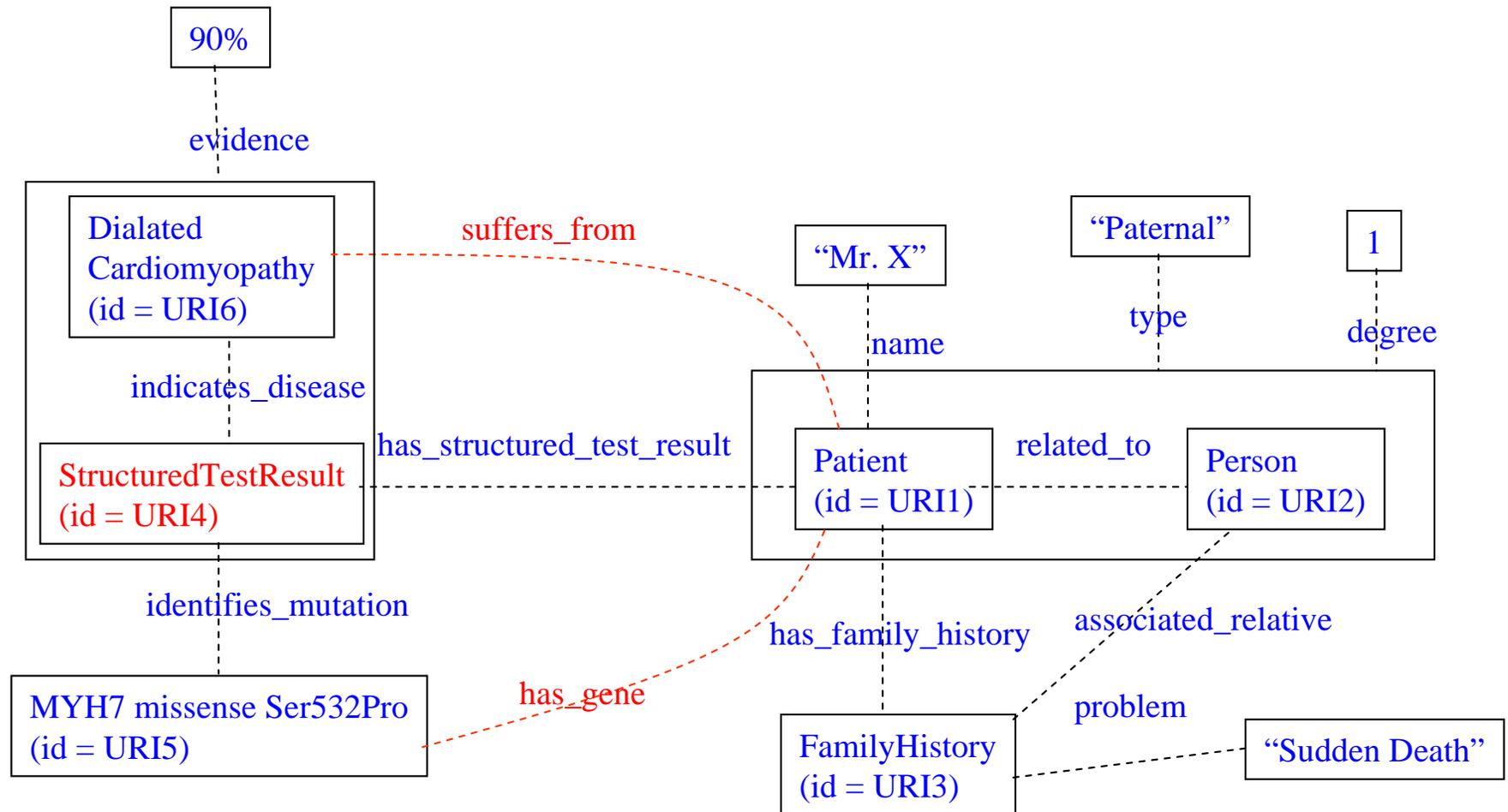
Bridging Clinical and Genomic Information



Rule/Semantics-based Integration:

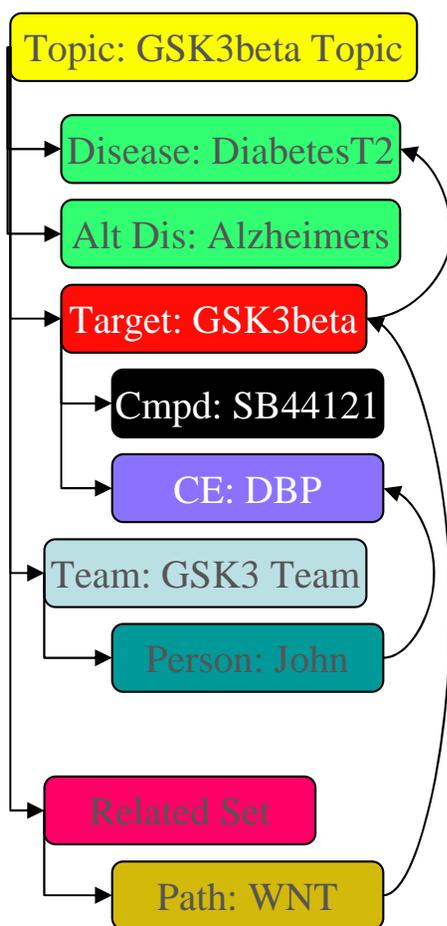
- Match Nodes with same Ids
- Create new links: IF a patient's structured test result indicates a disease THEN add a "suffers from link" to that disease

Semantic Data Integration: Bridging Clinical and Genomic Information



RDF Graphs provide a semantics-rich substrate for decision support. Can be exploited by SWRL Rules

Semantic Data Integration: Drug Discovery



Project Haystack Run Window Help

GSK3beta Topic

Target overview

DBP Lead, SB..., A..., CHI..., AKAPaulic NCE, CHI...

targets targets targets targets targets targets

GSK3beta

Group members

Title	role	Department	E-mail
John Tegler	Medicinal Chemist	Chemistry	john.tegle
Steve Smith	Synthetic Chemist	Chemistry	steve.smit
Tim Gross	Molecular Modeler	Cheminformatics	tim.gross

Primary disease

Type 2 Diabetes

#125853 [Links](#)

DIABETES MELLITUS, NONINSULIN-DEPENDENT; NIDDM

Alternative titles; symbols

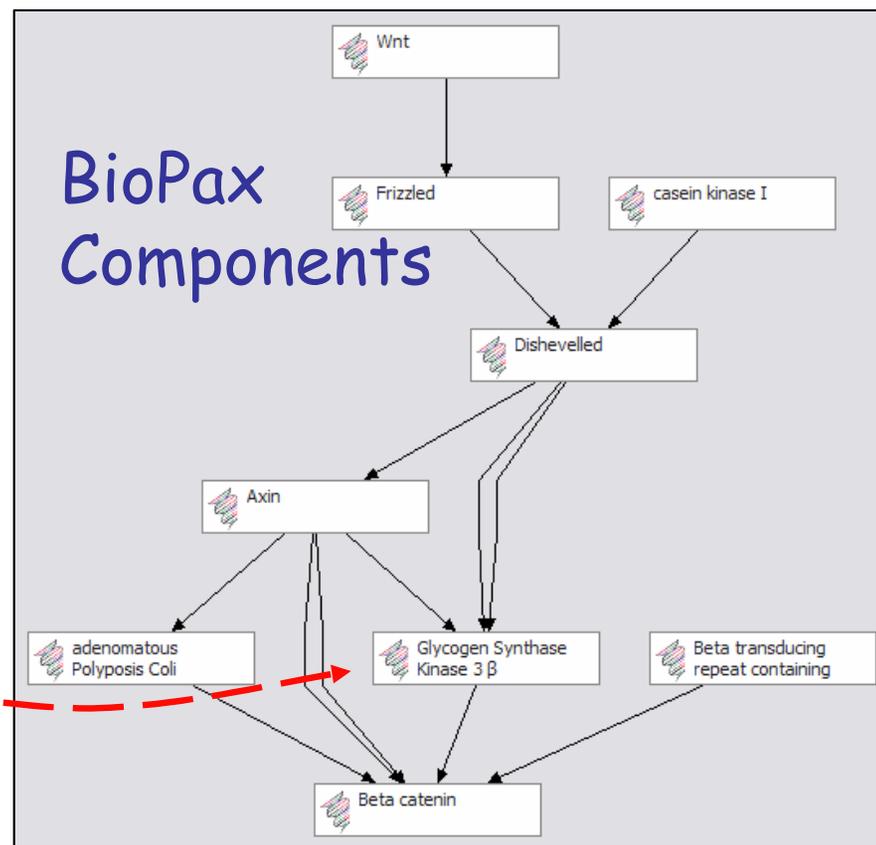
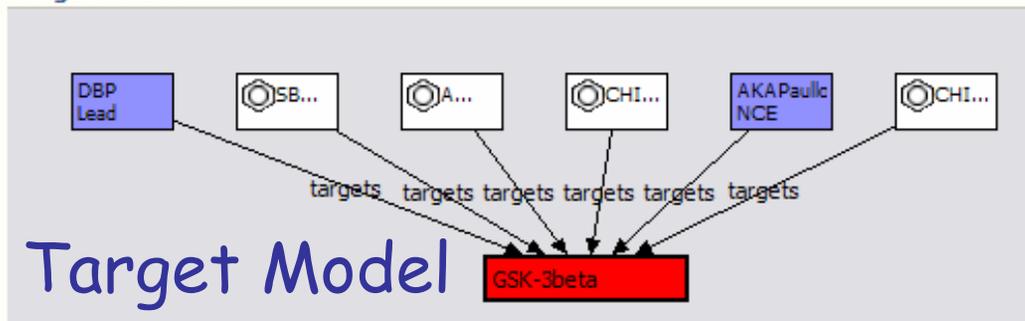
DIABETES MELLITUS, TYPE II
NONINSULIN-DEPENDENT DIABETES MELLITUS
MATURITY-ONSET DIABETES
INSULIN RESISTANCE, SUSCEPTIBILITY TO, INCLUDED

Gene map locus [20q12-q13.1, 20q12-q13.1](#)

Semantic Data Integration: Bridging Chemistry and Molecular Biology

Semantic Lenses: Different Views of the same data

Target overview



[urn:lsid:uniprot.org:uniprot:P49841](https://www.uniprot.org/uniprot/P49841)

Apply Correspondence Rule:

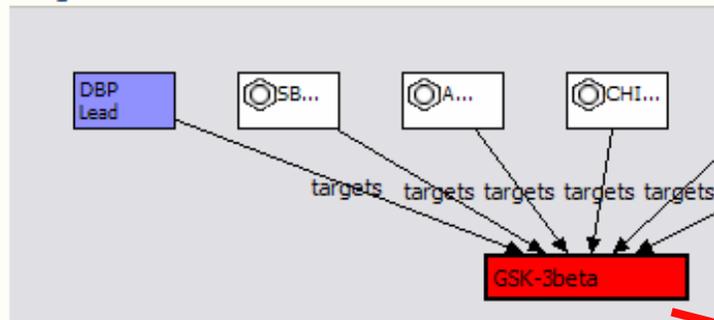
if ?target.xref.lsid == ?bpx:prot.xref.lsid
then ?target.correspondsTo.?bpx:prot

Semantic Data Integration

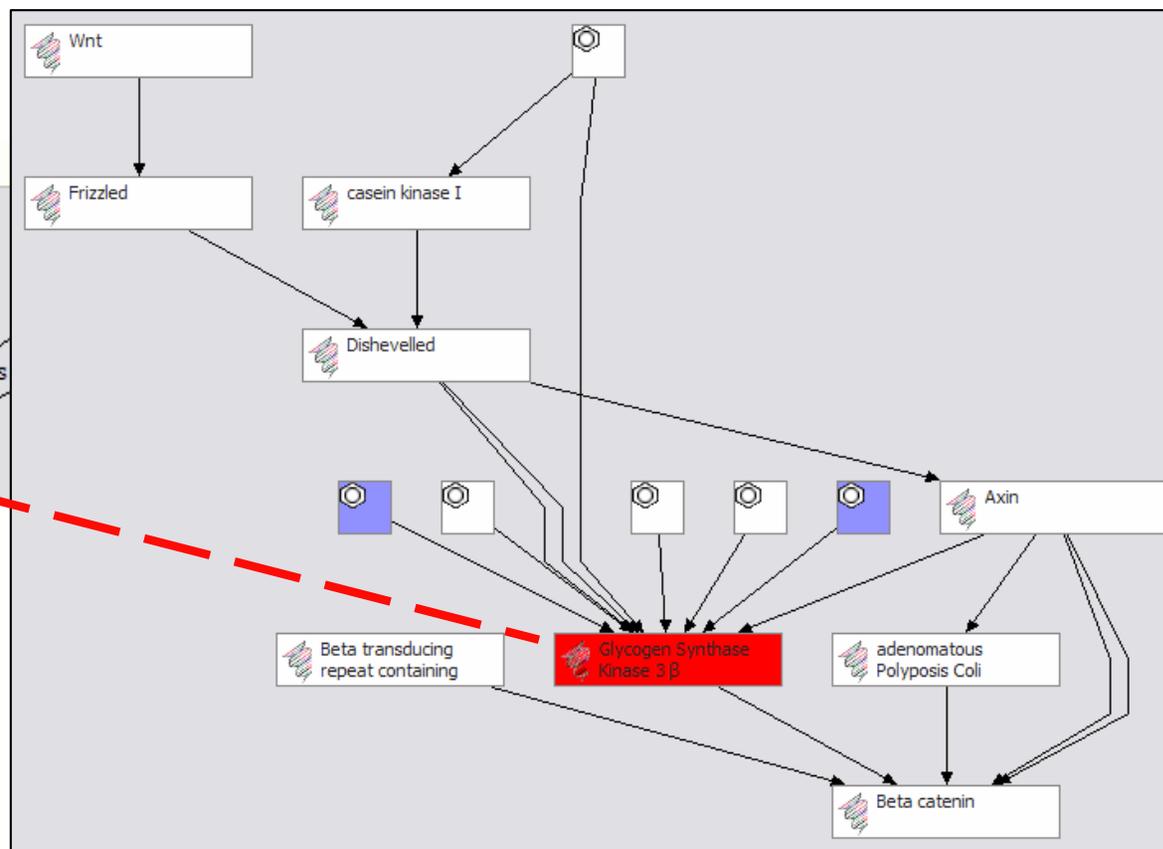
Bridging Chemistry and Molecular Biology

- Lenses can aggregate, accentuate, or even analyze new result sets
- Behind the lens, the data can be persistently stored as RDF-OWL

Target overview



- Correspondence does not need to mean "same descriptive object", but may mean objects with identical references



Semantic Data Integration: Advantages

- RDF: Graph based data model
 - More expressive than the tree based XML Schema Model
- RDF: Reification
 - Same piece of information can be given different values of belief by different clinical genomic researchers
- Potential for “Schema-less” Data Integration
 - Hypothesis driven approach to defining mapping rules
 - Can define mapping rules on the fly
- Incremental approach for Data Integration
 - Ability to introduce new data sources into the mix incrementally at low cost
- Use of Ontology to disallow meaningless mapping rules?
 - For e.g., mapping a gene to a protein...

Semantic Data Integration

“Schema-free” data integration

- Low cost approach for data integration
- No need for maintenance of costly schema mappings
- Ability to “merge” RDF graphs based on simple declarative rules that specify:
 - Equality of URIs
 - Connecting nodes of same type
 - Connecting two nodes associated by a “path”
- Disadvantage: Potential for specifying spurious non-sensical rules

Semantic Data Integration

Use of Reification

- Level of accuracy of test result.
 - Sensitivity and Specificity of lab result
 - Level of confidence in genotyping or gene sequencing
- Probabilistic relationships
 - Likelihood that a particular test result or condition is indicative of a disease or other medical condition
- Level of trust in a resource
 - Results from a lab may be trusted more than result from another
 - Results from well known health sites (NLM) may be trusted more than others
- Belief attribution
 - Scientific hypotheses may be attributed to appropriate researchers

Overview

- Use Case Flow
 - Need for Shared Semantics in the context of Personalized Medicine
- Functional Requirements
 - Metadata-based Annotation
 - Data Integration
 - **Clinical Decision Support**
 - Knowledge Change and Provenance

Clinical Decision Support

- Create the Business Object Model
- Specify Rules to encode Decision Support Logic
- Delineate definitions of Patient States
 - Represent them in an ontology

Clinical Decision Support

IF the patient's LDL test result is greater than 120
AND the patient has a contraindication to Fibric Acid
THEN

Prescribe Zetia Lipid Management Protocol

Contraindication to Fibric Acid: Clinical Definition (Old)

The patient is contraindicated for Fibric Acid if he has an allergy to Fibric Acid or has elevated Liver Panel

Contraindication to Fibric Acid: Clinical+Genomic Definition (New)

The patient is contraindicated for Fibric Acid if he has an allergy to Fibric Acid or has elevated Liver Panel or has a genetic mutation
Missense: XYZ3:Ser@ \$#Pro

Please note: Hypothetical – assume a genetic variant is a biomarker for patients contraindicated to Fibric Acid.

Clinical and Genomic Decision Support: Business Object Model

Class Patient: Person

```
method get_name(): string;  
method has_genetic_test_result(): StructuredTestResult;  
method has_liver_panel_result(): LiverPanelResult;  
method has_ldl_result(): real;  
method has_contraindication(): set of string;  
method has_mutation(): string;  
method has_therapy(): set of string;  
method set_therapy(string): void;  
method has_allergy(): set of string;  
Method get_category(): set of string;
```

Class StructuredTestResult

```
method get_patient(): Patient;  
method indicates_disease(): Disease;  
method identifies_mutation(): set of string;  
method evidence_of_mutation(string): real;
```

Class LiverPanelResult

```
method get_patient(): Patient;  
method get_ALP(): real;  
method get_ALT(): real;  
method get_AST(): real;  
method get_Total_Bilirubin(): real;  
method get_Creatinine(): real;
```

Clinical and Genomic Decision Support: A Rules-based Implementation

```
IF the_patient.has_ldl_result() > 120
AND ((the_patient.has_liver_panel_result().get_ALP() ≥ <NormalRange>
      AND the_patient.has_liver_panel_result().get_ALT() ≥ <NormalRange>
      AND the_patient.has_liver_panel_result().get_AST() ≥ <NormalRange>
      AND the_patient.has_liver_panel_result().get_Total_Bilirubin() ≥ <NormalRange>
      AND the_patient.has_liver_panel_result().get_Creatinine() ≥ <NormalRange>)
OR "Fibric Acid Allergy" ∈ the_patient.has_allergy()
OR "Missense: XYZ3:Ser@ $#Pro" ∈ the_patient.has_mutation())
THEN
  the_patient.set_therapy("Zetia Lipid Management Protocol")
```

Definition of “Fibric Acid Contraindication”

Clinical Decision Support: Definitions vs. Decisions

Commonly occurring design pattern:

- The **definition** of a “Fibric Acid Contraindication” is represented using rules.
- The **decision** related to therapeutic intervention is also represented using rules.

Currently, both these inferences are performed by the rules engine.

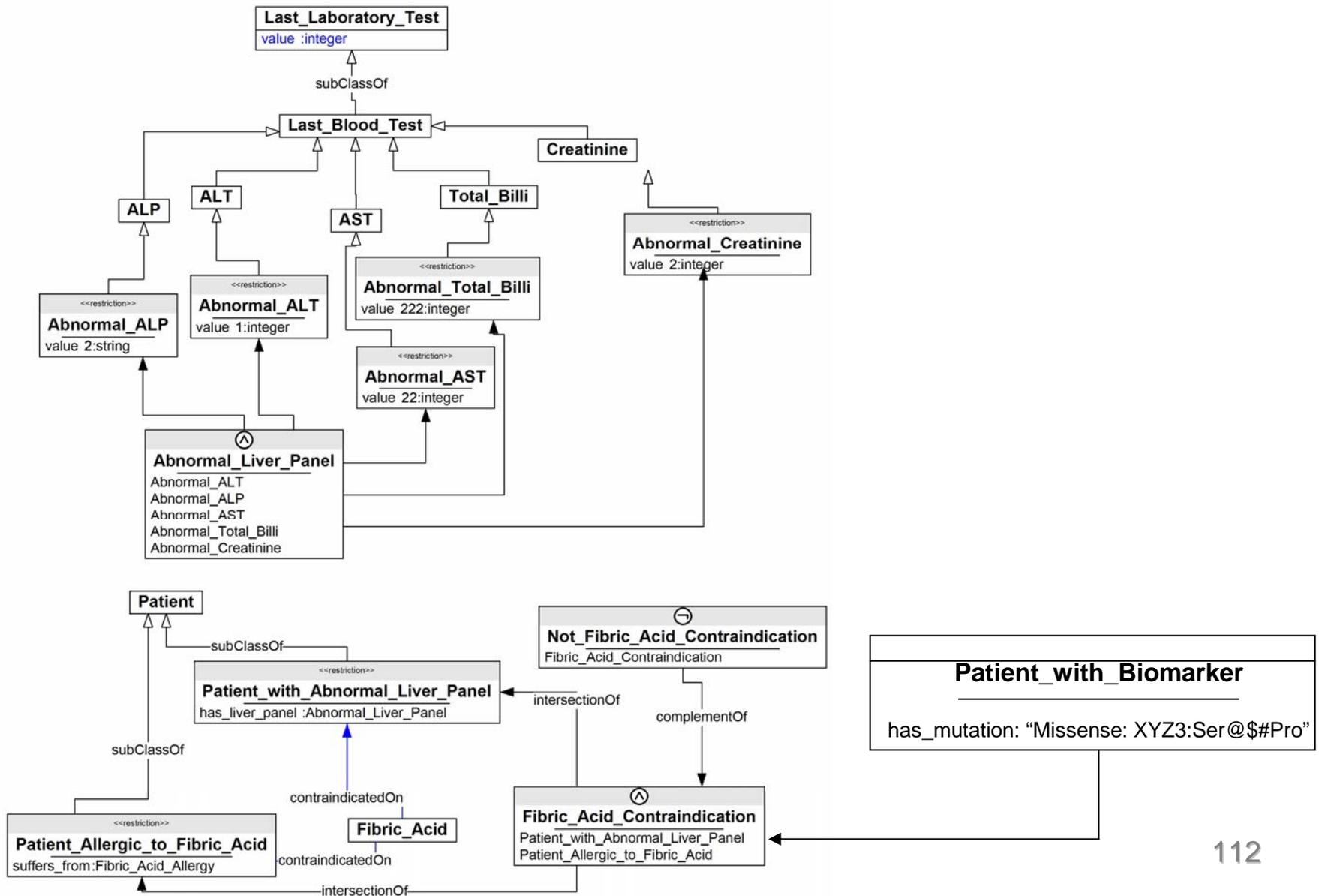
Clinical Decision Support

Delineating patient states

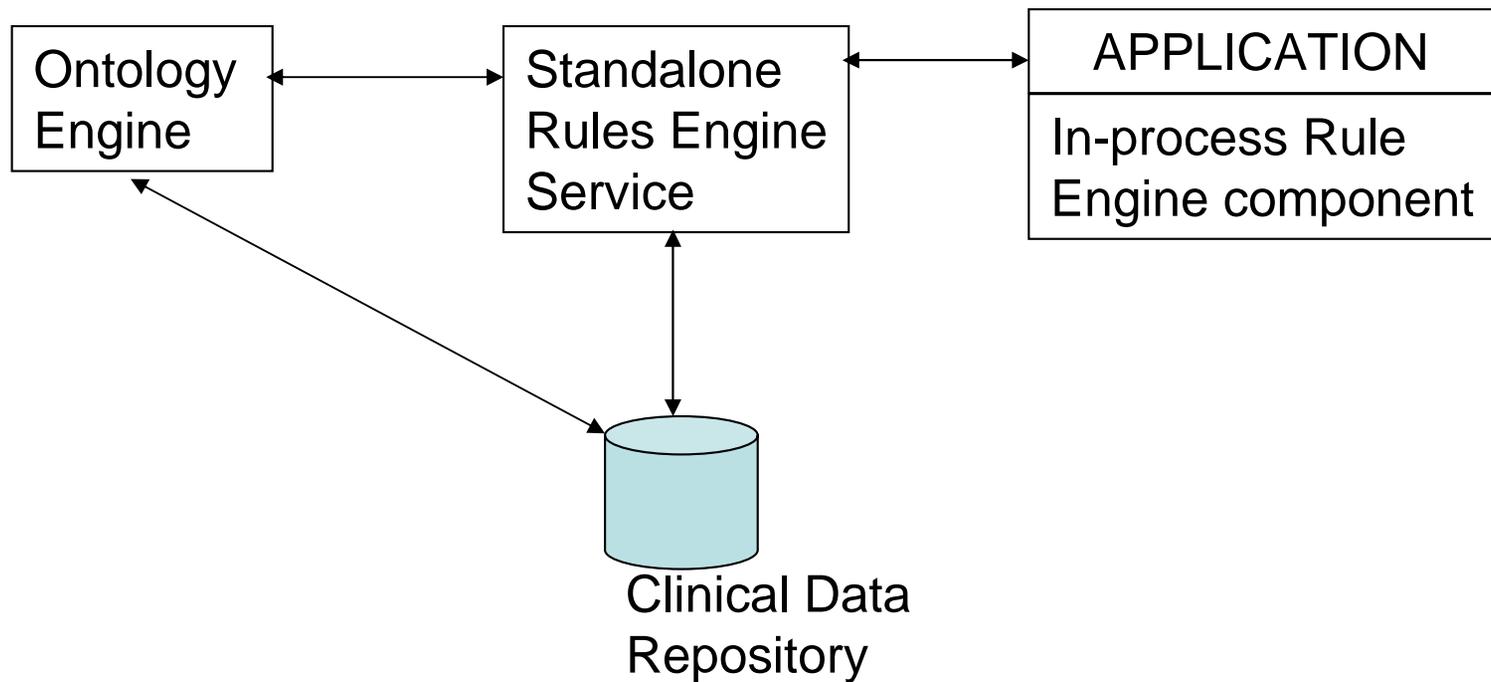
```
IF the_patient.has_ldl_result() > 120  
AND the_patient.get_category() = PatientWithFibricAcidContraindication  
  
THEN  
set the_patient.has_therapy("Zetia Lipid Management Protocol")
```

implemented in an OWL-based
ontology engine

OWL representation of Fibric Acid Contraindication



Ontology Driven Clinical Decision Support: Architecture



Clinical Decision Support:

Decoupling definitions vs. decisions

- Classification inferences (does patient have a fibrin acid contraindication?) can be evaluated by an ontology engine.
- Reduces overhead on Rule Engine
- Opens up the possibility of plugging-in other specialized inference engines (e.g., spatio-temporal conditions)
- Makes knowledge maintenance easier
 - Each definition may be referred to in 100s of rules..

Decision Support: Statistical vs. Symbolic Approaches

- Symbolic:
 - Knowledge Driven: Needs input of Subject Matter Experts
 - Not scaleable: Knowledge Bases can get huge in case of interacting conditions
 - Example:
 - Set of Rules for “CAD”
 - Set of Rules for “Diabetes”
 - What about rules for “Diabetes” and “CAD”
 - In general for N conditions, the Knowledge base size can be of the order of 2^N .
- Statistical:
 - Data Driven: Models can be “learned” from the data
 - More scaleable
 - Probabilistic conclusions, Thresholding required
 - **Blackbox: No explanations possible!**
- Hybrid: Need some combination of the two...

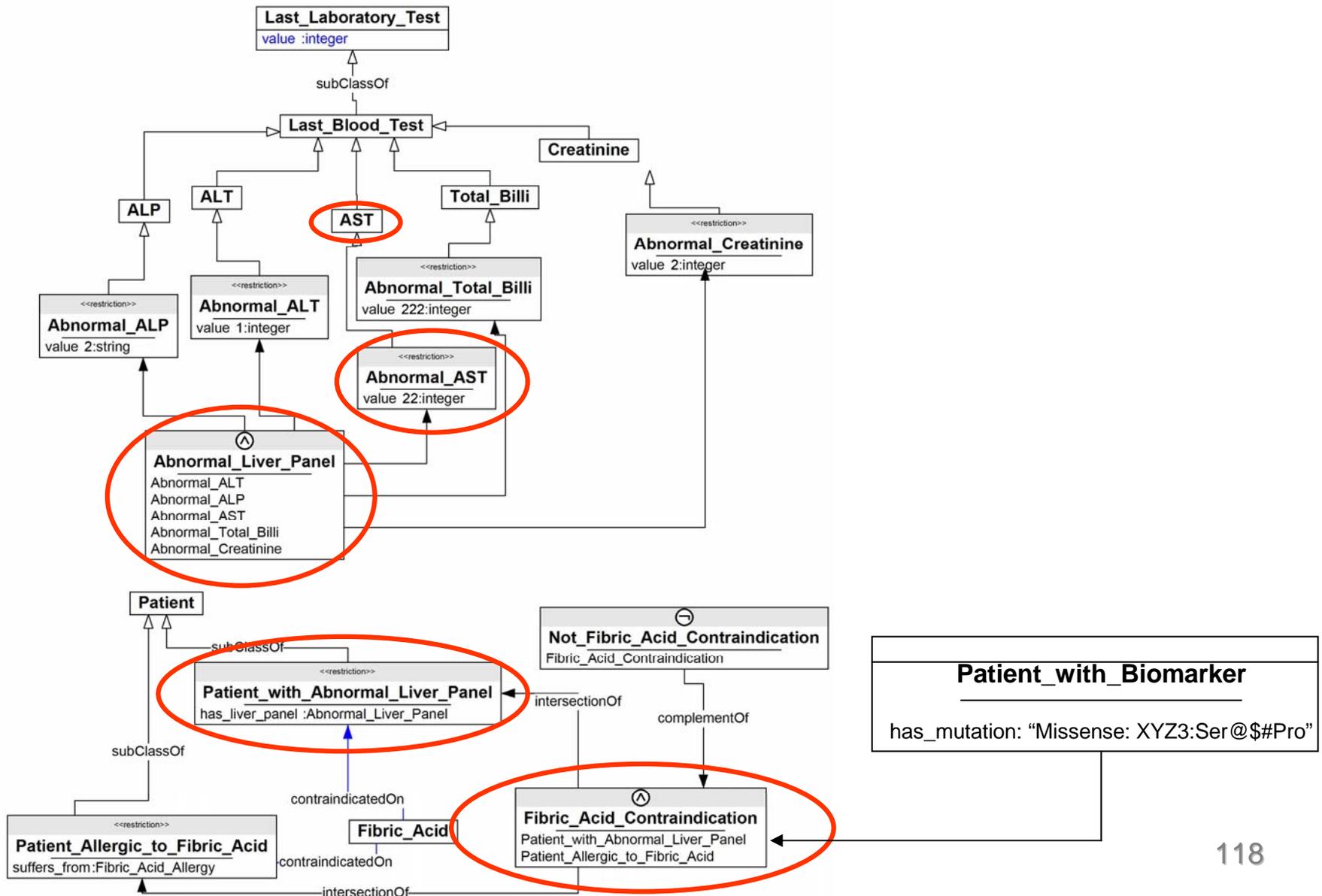
Overview

- Use Case Flow
 - Need for Shared Semantics in the context of Personalized Medicine
- Functional Requirements
 - Metadata-based Annotation
 - Data Integration
 - Clinical Decision Support
 - Knowledge Change and Provenance

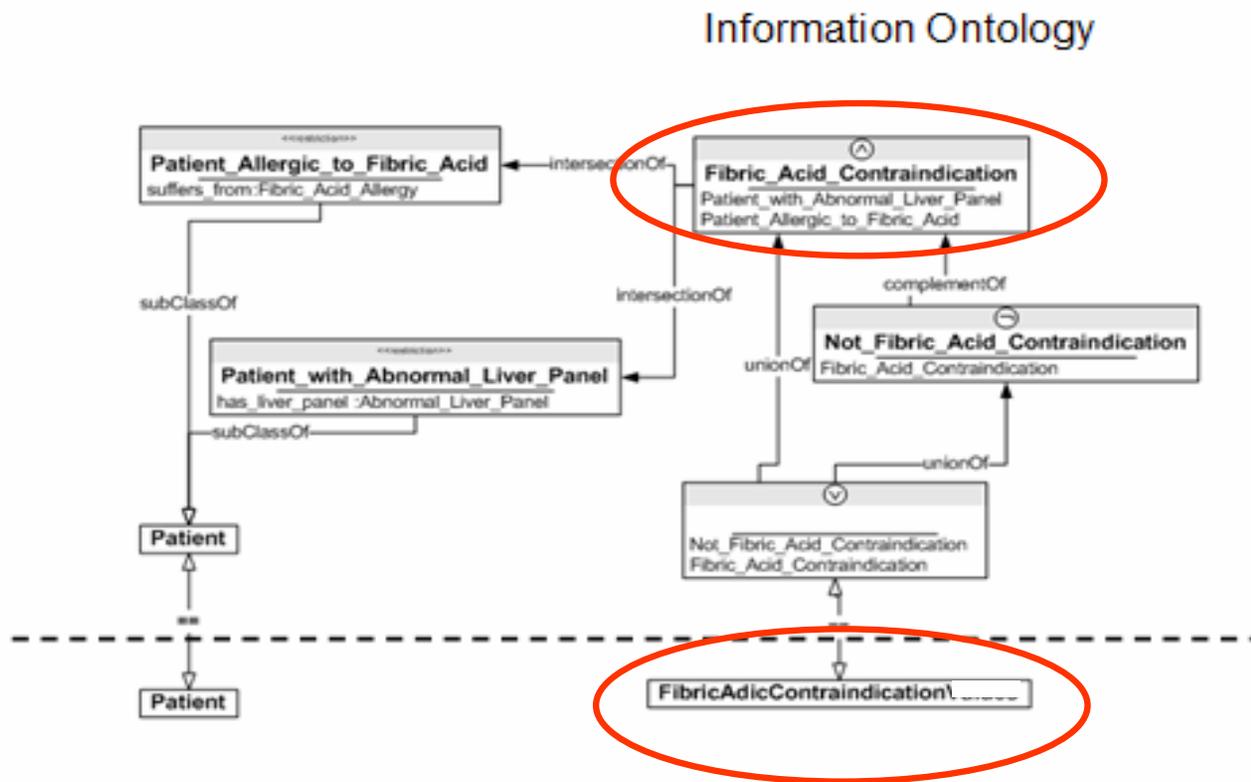
Knowledge Change and Provenance

- There is rapid knowledge discovery and evolution in the Healthcare and Life Sciences
- Provenance is an important aspect of maintaining knowledge consistence
- There is a close interrelationship between knowledge change and provenance
 - What has changed? – Change
 - Why did it change? – Provenance
 - Did someone change it? – Provenance
 - Did its components change? – Change
 - Who changed it? – Provenance

Knowledge Change and Propagation



Knowledge Change and Propagation



Rule base

Knowledge Change and Provenance

- At each stage, Knowledge Engineer gets notified of:
 - What has changed?
 - The definition of Fibric Acid Contraindication
 - Why did it change?
 - Fibric Acid Contraindication ← Patient with Abnormal Liver Panel ← Abnormal Liver Panel ← Abnormal AST ← Change in AST Values
 - Fibric Acid Contraindication ← Patient with Biomarker
 - Who was responsible for the change?
 - Knowledge Engineer who entered the changed AST values?
 - Change in a Clinical Guideline?
 - New Molecular Diagnostic Test appears in the market?

Knowledge Update and Maintenance

- Knowledge Dependency Propagation
 - If the definition of a concept changes,
 - What other concepts does it impact?
 - What other clinical decision support rules does it impact?
- Assertion Dependency Propagation
 - If a clinical decision support rule is changed, how does this impact potential decisions made for a patient?
 - How and when should such decisions be updated?
 - What if the decision involved a drug which has already been administered?

Part 4

W3C Semantic Web *Health Care and Life Sciences* Interest Group

The image features three computer keyboard keys arranged diagonally from top-left to bottom-right. The top key is white with a black 'w' character. The middle key is white with a black '3' character. The bottom key is white with a black 'c' character. The keys are set against a light blue, textured background that resembles a carpet or fabric. The lighting is soft, creating subtle shadows and highlights on the keys.

Part 4

W3C Semantic Web
Health Care and Life Sciences
Interest Group

Enabling Bench-to-Bedside: W3C Semantic Web for HealthCare and Life Sciences Interest Group

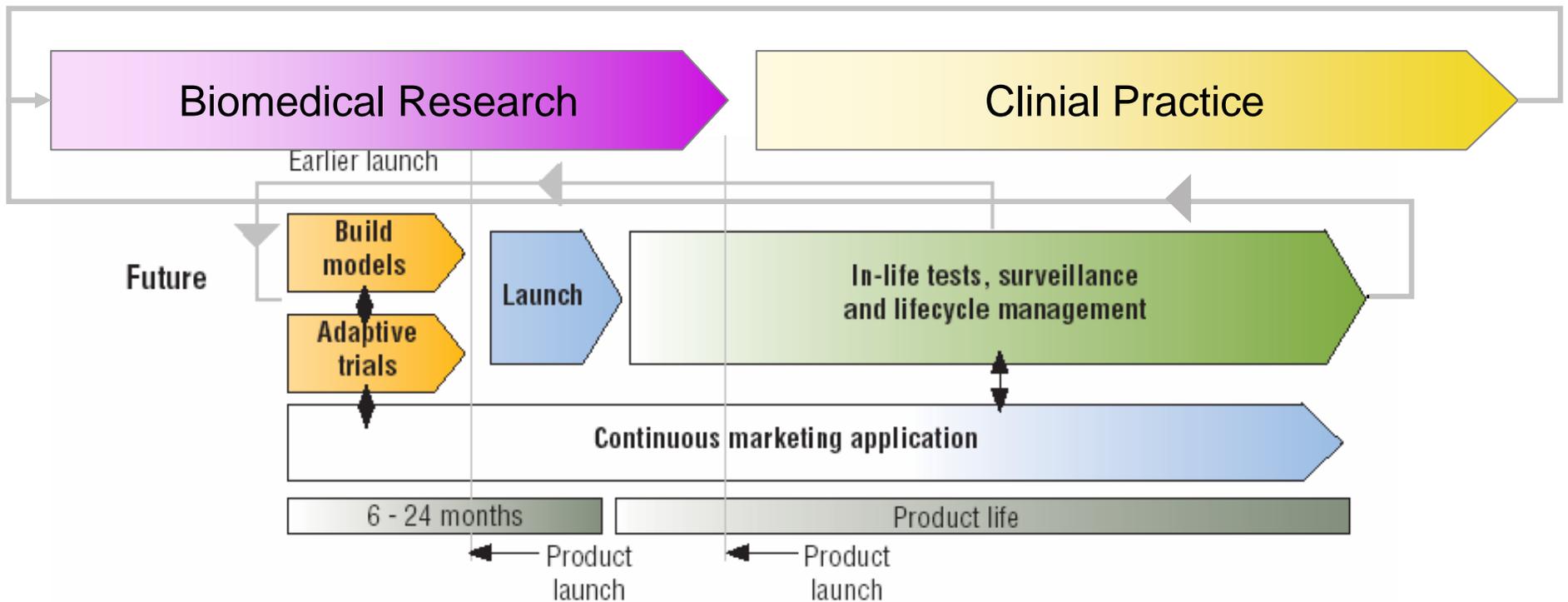
- **Launched Nov 2005:** <http://www.w3.org/2001/sw/hcls>
 - Co-chairs: Dr. Tonya Hongsermeier (Partners HealthCare); Eric Neumann (Teranode)
- **Membership: 37, 3 Pharmaceuticals, 4 Healthcare groups**
- **Chartered to develop and support the use of SW technologies and practices to improve collaboration, research and development, and innovation adoption in the of Health Care and Life Science domains**
- **Based on a foundation of semantically rich specifications that support process and information interoperability**
- **HCLS Objectives:**
 - Core vocabularies and ontologies to support cross-community data integration and collaborative efforts
 - Guidelines and Best Practices for Resource Identification to support integrity and version control
 - Better integration of Scientific Publication with people, data, software, publications, and clinical trials

Other W3C standards/techs

- SPARQL - querying the SW
- GRDDL - transforming data for SW
- Rules - Applying policies and inferencing on the SW
- RDFa - Embedding RDF into XHTML
- FRESNEL - Visualizing the SW

Ecosystem: Goal State

/* Need to expand this with Biomedical Research + Clinical Practice */



Overview

- A Forum for Scientists and Clinicians to
 - Share use-cases and experiences on “how-to”
 - Drive vendor adoption in HC and LS applications
 - Create vision demonstrations
 - Translation Research Informatics
- Expose collections of public resources as RDF
 - NCBI, Uniprot, EBI
- Develop new core vocabularies for data integration, and migrate existing ontologies to RDFS/OWL/SKOS
 - SNOMED, BioPAX, OBO, MESH, FMA

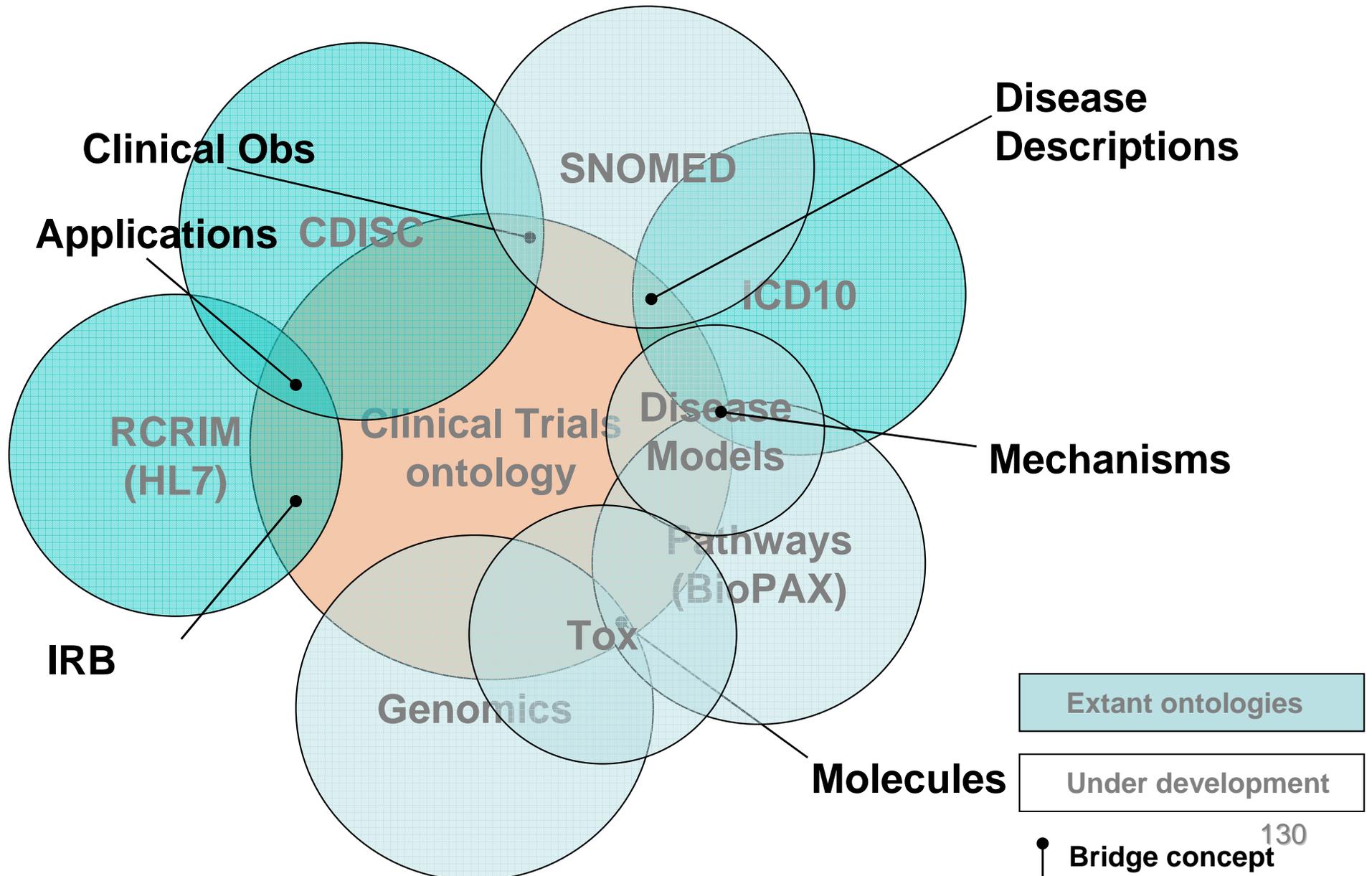
HCLS Tasks Overview

- **BioRDF**
 - exposing data as RDF
- **BioONT**
 - ontology guidelines and coordination
- **Adaptive Clinical Protocols and Pathways**
 - flexible healthcare management
- **Drug Safety and Efficacy**
 - pharmaceuticals
- **Scientific Publishing**
 - Supporting Knowledge through Text and Data

HCLS Activities

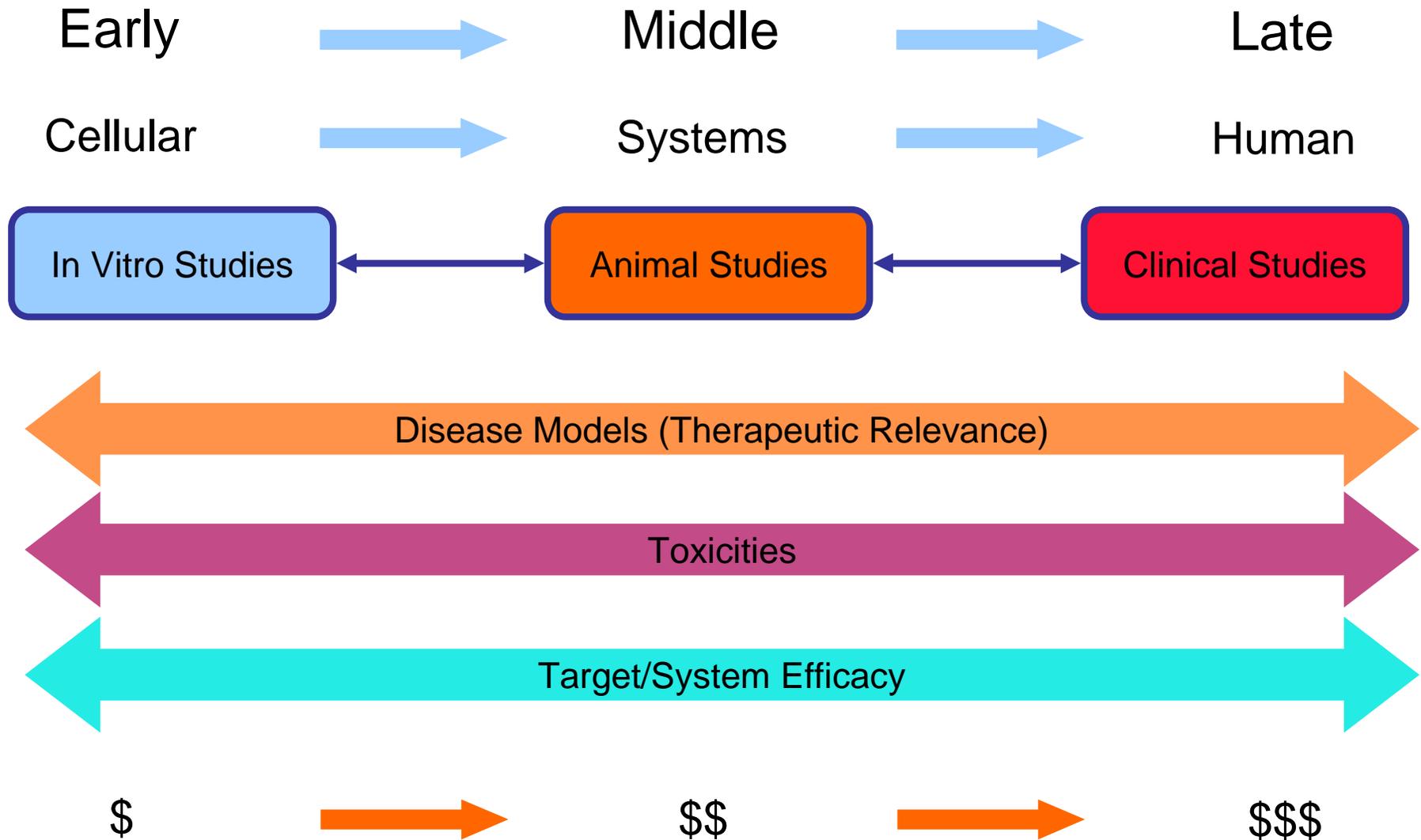
- BioRDF+BioONT: Parkinson Disease use-cases
 - Exposed MolBio Data
 - Parkinson's Ontology
- Adaptive Protocols
 - Ontology Development (with consideration of RIM v3)
 - Temporal Reasoning
- Drug Safety and Efficacy
 - Semantically enable CDISC SDTM Model
 - Adding SW annotations and hypotheses to a JANUS-style DB
 - Provenance and trust (non-reputability) and security

Linking Clinical Ontologies with the Semantic Web

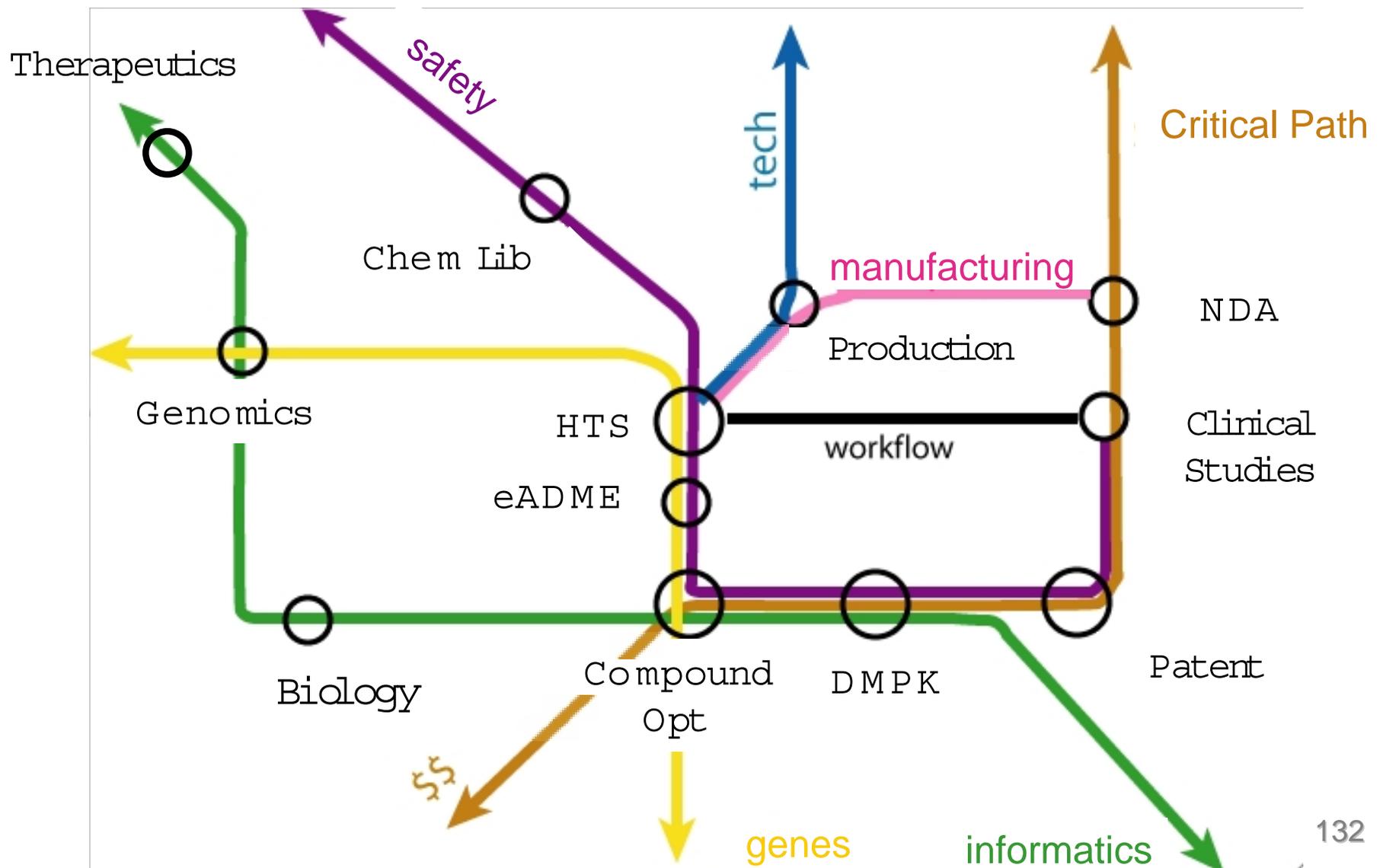


Translational Medicine in Drug

R&D



Application Space : Semantic Web Drug DD

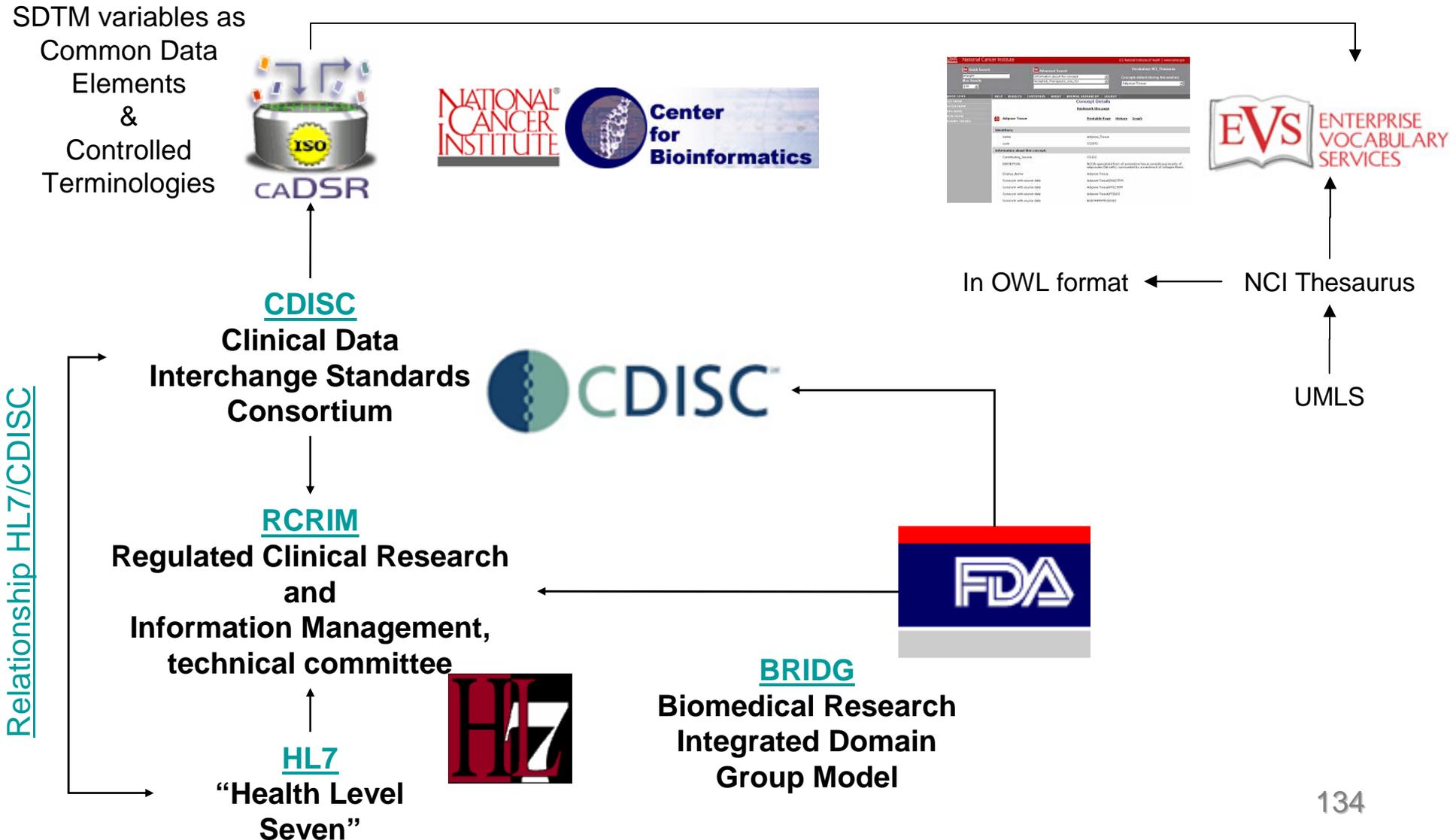


Domain Semantics in Clinical Trials

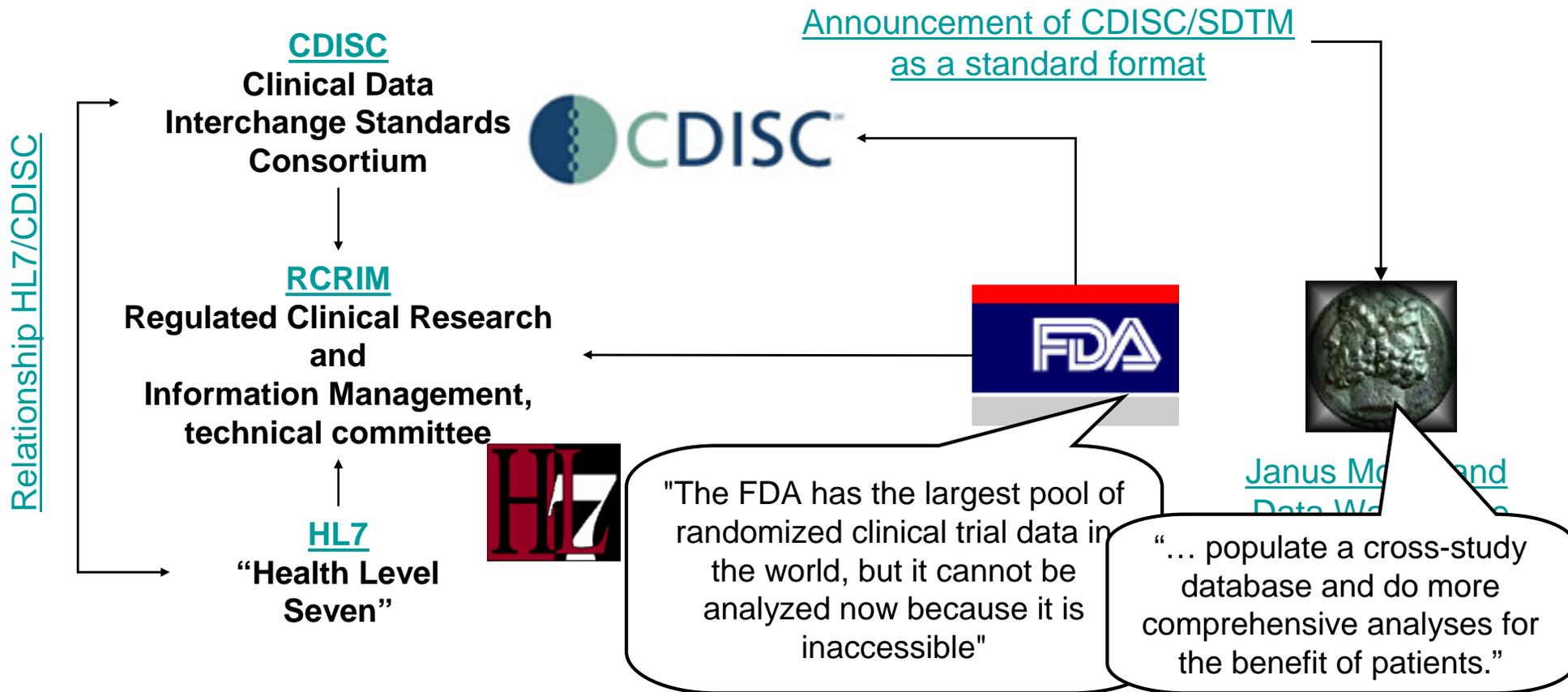
Clinical Semantics

- Patient/Subject → Disease/Health state
- Diagnostics → Findings
- Findings → Inferred (proposed) Disease state
- Disease state → Patient Classification / Segmentation
- Design → Trial arms / treatments
- Observation → POC, safety, mechanisms

During 2006-2007



Ongoing work at FDA



Dr. Janet Woodcock, Deputy Commissioner for 135
Operations and Chief Operating Officer, FDA
27 January 2006

Part 5

Current trends and future directions

Key Semantic Web Principles

- **Plan for change**
- **Free data from the application that created it**
- **Lower reliance on overly complex Middleware**
- **The value in "as needed" data integration**
- **Big wins come from many little ones**
- **The power of links - network effect**
- **Open-world, open solutions are cost effective**
- **Importance of "Partial Understanding"**

References

Books

- G. Antoniou and F. van Harmelen. *A Semantic Web primer*. 2004. MIT Press.
- K. Baclawski and T. Niu. *Ontologies for bioinformatics*. 2006. MIT Press.
- S. Staab and R. Studer (Eds). *Handbook on ontologies*. 2004. Springer.