

Module 13

Introduction to Semantic Technology, Ontologies and the Semantic Web

Module 13 Outline

10.30-12.30

- Introduction to the Semantic Web
- Ontologies
- Semantic Web related standards

12.30-14.00

Lunch break

14.00-16.00

- Semantic Web related standards (part II)
- Some Application of Semantic Technologies
- Tools

16.00-16.30

Coffee

About this tutorial

- The Web that we know → The Semantic Web
- Ontologies
- Semantic Web related standards
- Some Applications of Semantic Technologies
- Tools

Introduction to the Semantic Web

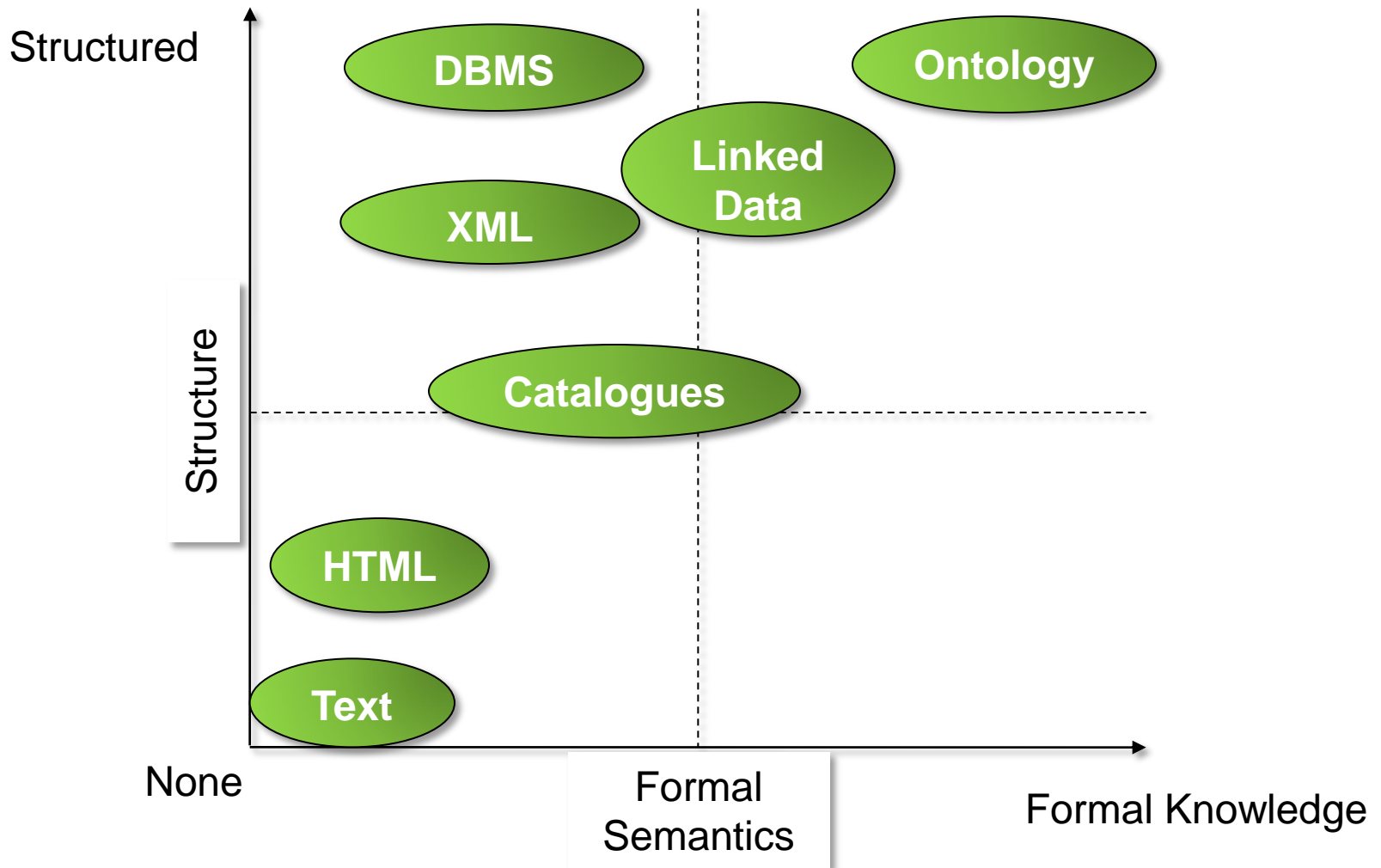
The Web as we know it

- *Target consumers*: humans
 - web 2.0 mashups provide *some* improvement
 - Rules about the *structure* and *visualisation* of information, but not about its *intended meaning*
 - Intelligent agents can't easily use the information
- *Granularity*: document
 - One giant distributed *filesystem* of documents
 - One *document* can link to other documents
- *Integration & reuse*: very limited
 - Cannot be easily automated
 - Web 2.0 mashups provide *some* improvement

Some problems with the current Web

- Finding information
- Data granularity
- Resource identification
- Data aggregation & reuse
- Data integration
- Inference of new information

Types of Data



The need for a smarter Web

- *"The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation."* (Tim Berners-Lee, 2001)

The need for a smarter Web (2)

- *“PricewaterhouseCoopers believes a Web of data will develop that fully augments the document Web of today. You’ll be able to find and take **pieces of data** sets from **different places**, aggregate them without warehousing, and **analyze them in a more straightforward, powerful way** than you can now.” (PWC, May 2009)*

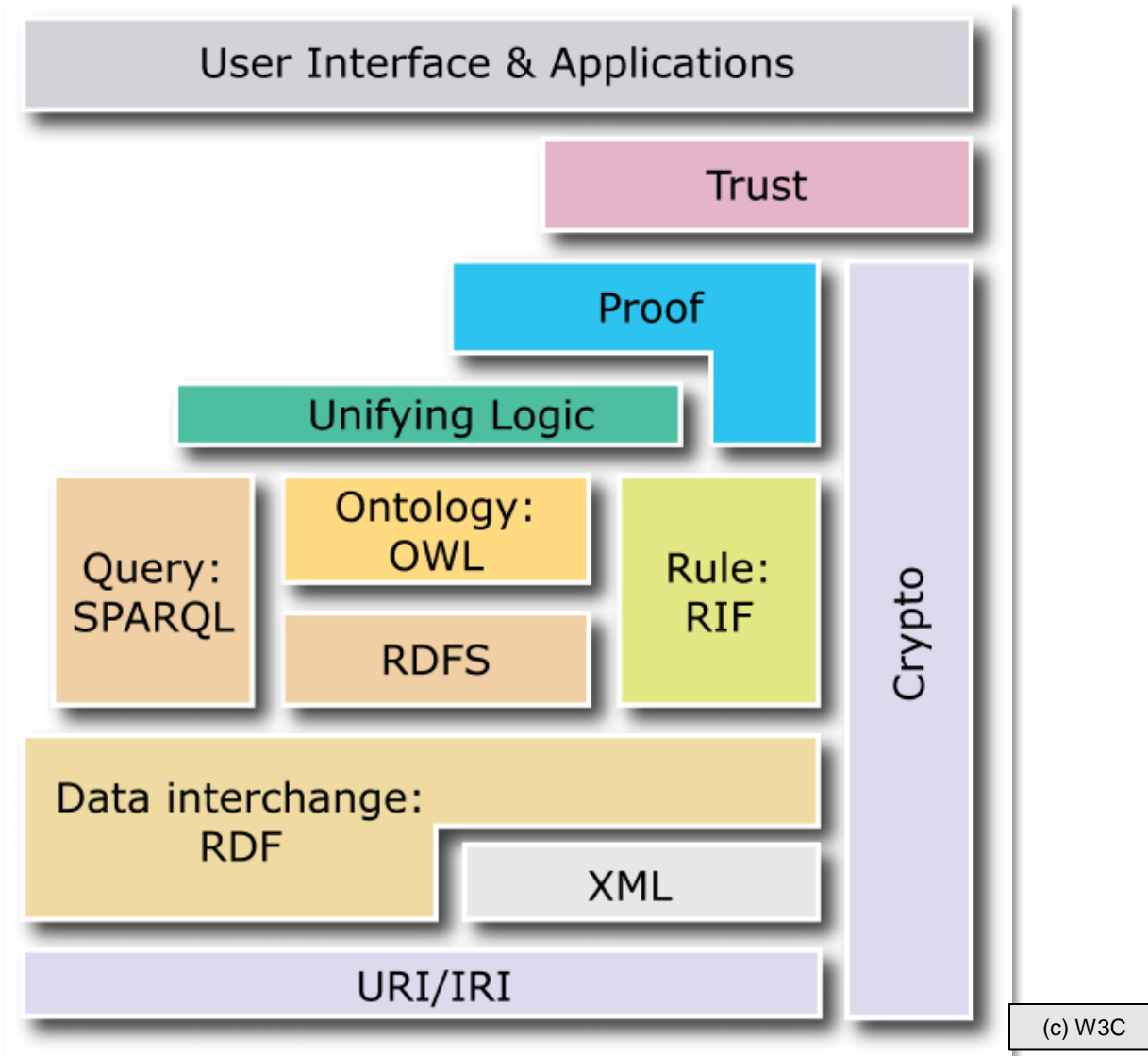
The Semantic Web

- *Target consumers*: intelligent agents
 - Explicit specification of the *intended meaning* information
 - Intelligent agents can make use the information
- *Granularity*: resource/fact
 - One giant distributed *database* of facts about resources
 - One *resource* can be linked (related) to other resources
- *Integration & reuse*: easier
 - Resources have unique identifiers
 - With explicit semantics transformation & integration can be automated

The Semantic Web vision (W3C)

- Extend principles of the Web from documents to data
- Data should be accessed using the general Web architecture (e.g., URI-s, protocols, ...)
- Data should be related to one another just as documents are already
- Creation of a common framework that allows
 - Data to be shared and reused across applications
 - Data to be processed automatically
 - New relationships between pieces of data to be inferred

The Semantic Web layer cake

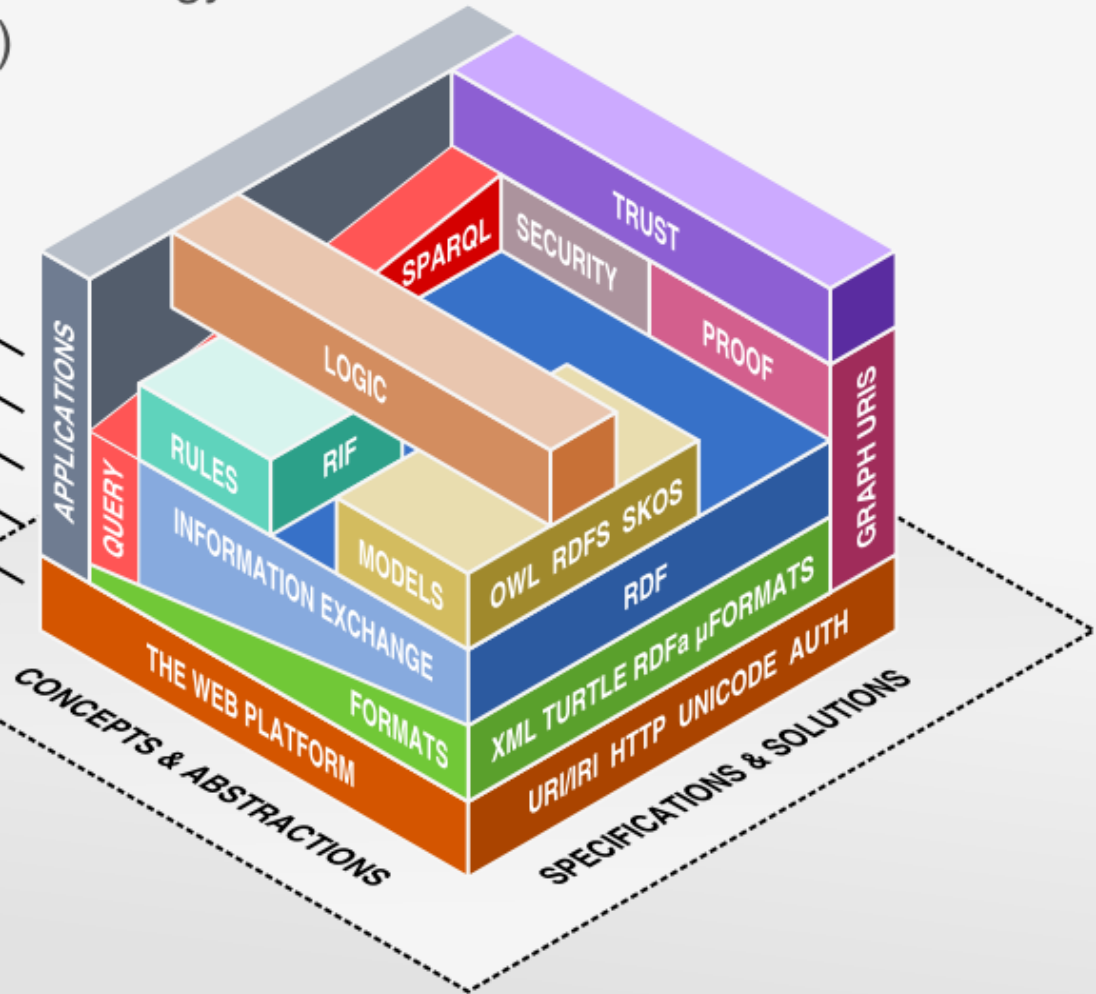
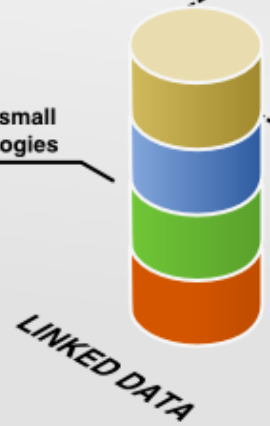


The Semantic Web layer cake (2)

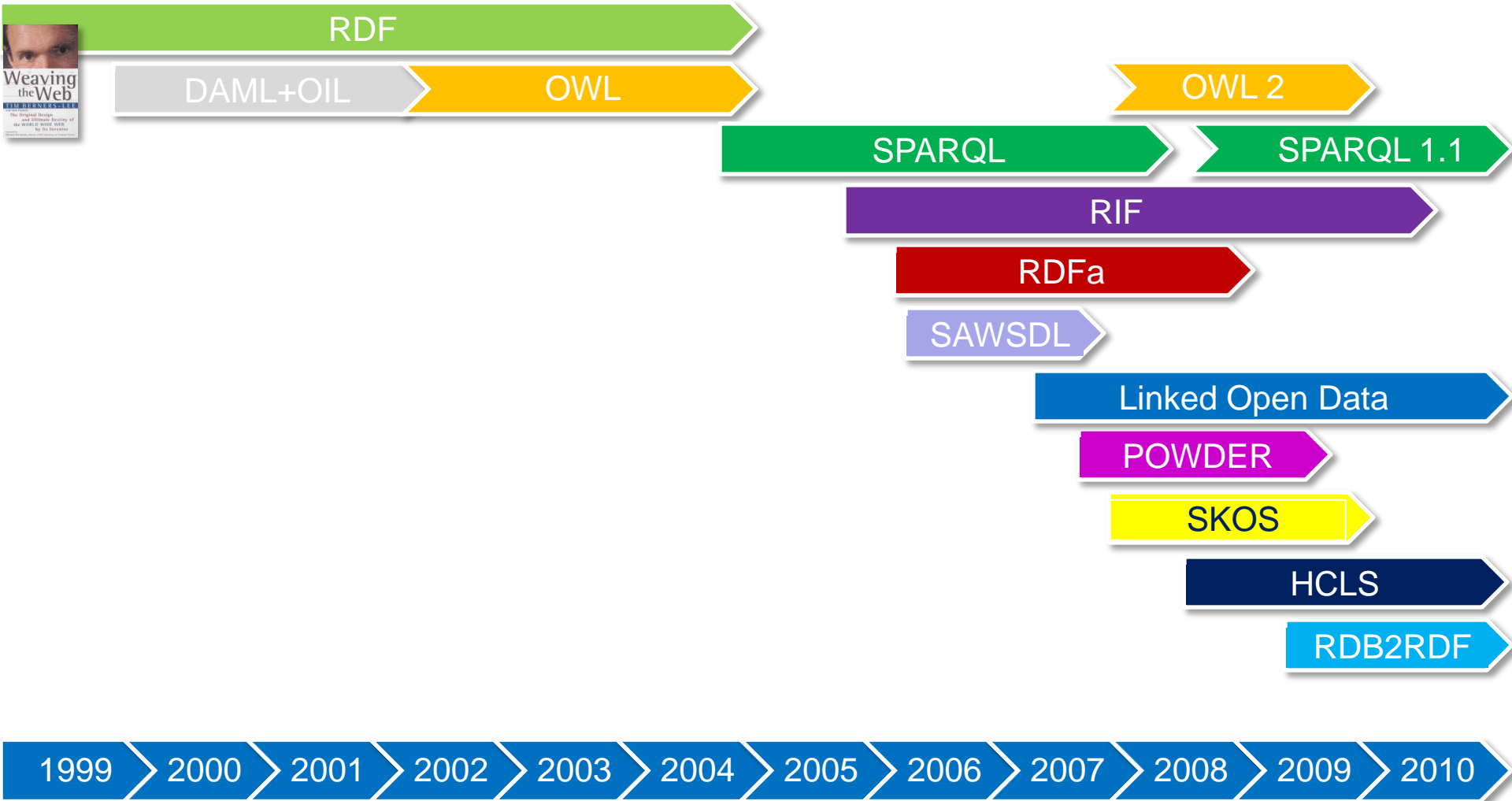
The Semantic Web Technology Stack
(not a piece of cake...)

- Most apps use only a subset of the stack
- Querying allows fine-grained data access
- Standardized information exchange is key
- Formats are necessary, but not too important
- The Semantic Web is based on the Web

Linked Data uses a small selection of technologies



The Semantic Web timeline



Ontologies

What is an ontology

- What is an ontology
 - A formal specification that provides sharable and reusable knowledge representation
 - Examples – taxonomies, thesauri, topic maps, E/R schemata*, formal ontologies
- Ontology specification includes
 - Description of the *concepts* in some domain and their properties
 - Description of the possible *relationships* between the concepts and the *constraints* on how the relationships can be used
 - Sometimes, the *individuals* (members of concepts) #16

Ontology dimensions (NIST, 2007)

	dimension	examples
Semantic	Degree of structure and formality	<ul style="list-style-type: none"> • Informal (specified in natural language) • taxonomy or topic hierarchy • very formal - unambiguous description of terms and axioms
	Expressiveness of the representation language	<ul style="list-style-type: none"> • different logic formalisms have different expressivity (and computational complexity)
	granularity	<ul style="list-style-type: none"> • simple taxonomies and hierarchies • detailed property descriptions, rules and restrictions
Pragmatic	Intended use	<ul style="list-style-type: none"> • data integration (of disparate datasources) • represent a natural language vocabulary (lexical ontology) • categorization and classification
	Role of automated reasoning	<ul style="list-style-type: none"> • is inference of new knowledge required? • simple reasoning (class/subclass transitivity inference) vs. complex reasoning (classification, theorem proving)
	Descriptive vs. prescriptive	<ul style="list-style-type: none"> • descriptive – less strict characterization, • prescriptive – strict characterization
	Design methodology	<ul style="list-style-type: none"> • bottom-up vs. top-down
	governance	<ul style="list-style-type: none"> • are there legal and regulatory implications • is provenance required?

example

class Person

class Woman

subClassOf #Person

class Man

subClassOf #Person

complementOf #Woman

individual John

instanceOf #Man

individual Mary

instanceOf #Woman

hasSpouce #John

individual Jane

instance Of #Woman

hasParent #John

hasParent #Mary

property hasParent

domain #Person

range #Person

maxCardinality 2

property hasChild

inverseOf #hasParent

property hasSpouce

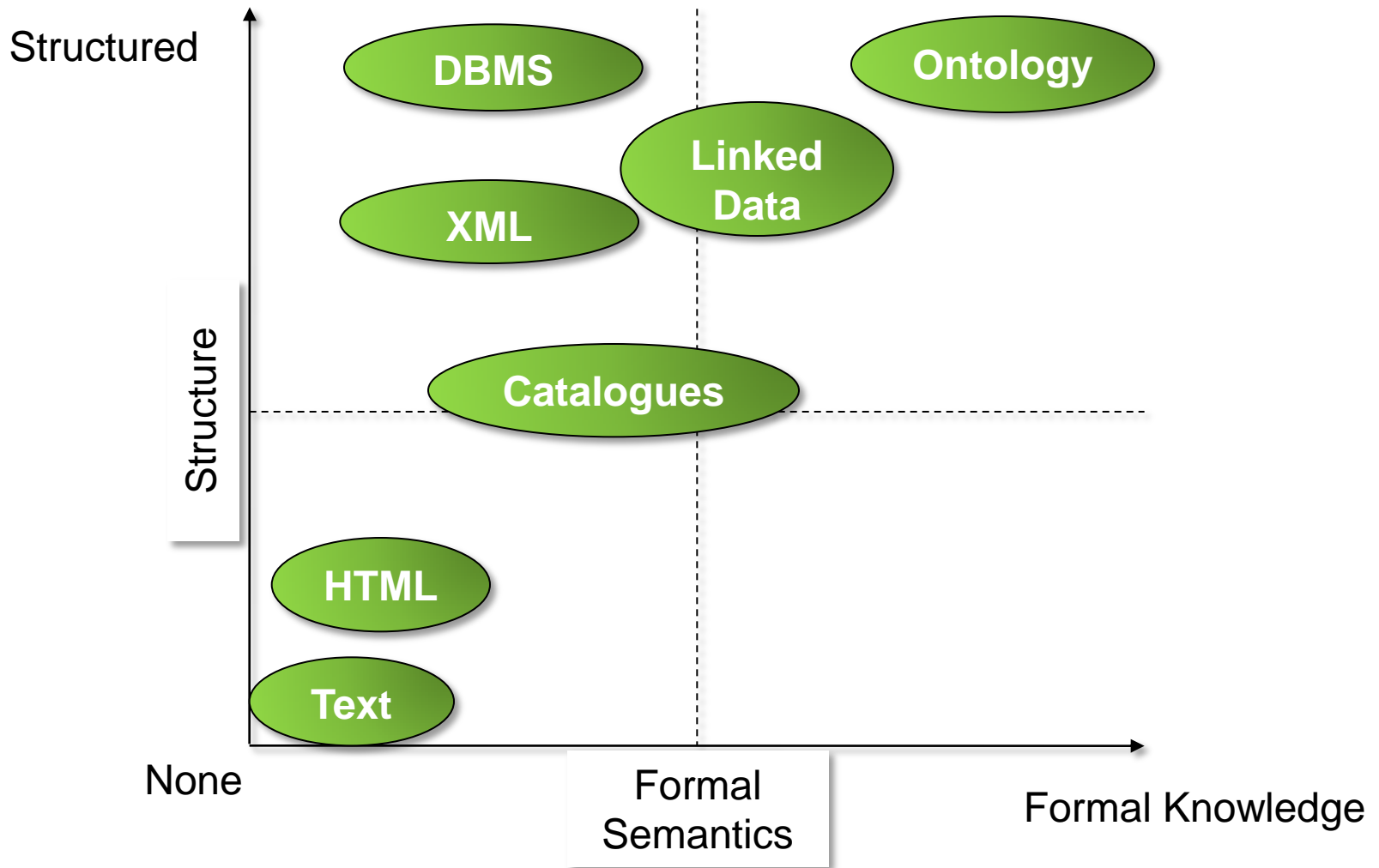
domain #Person

range #Person

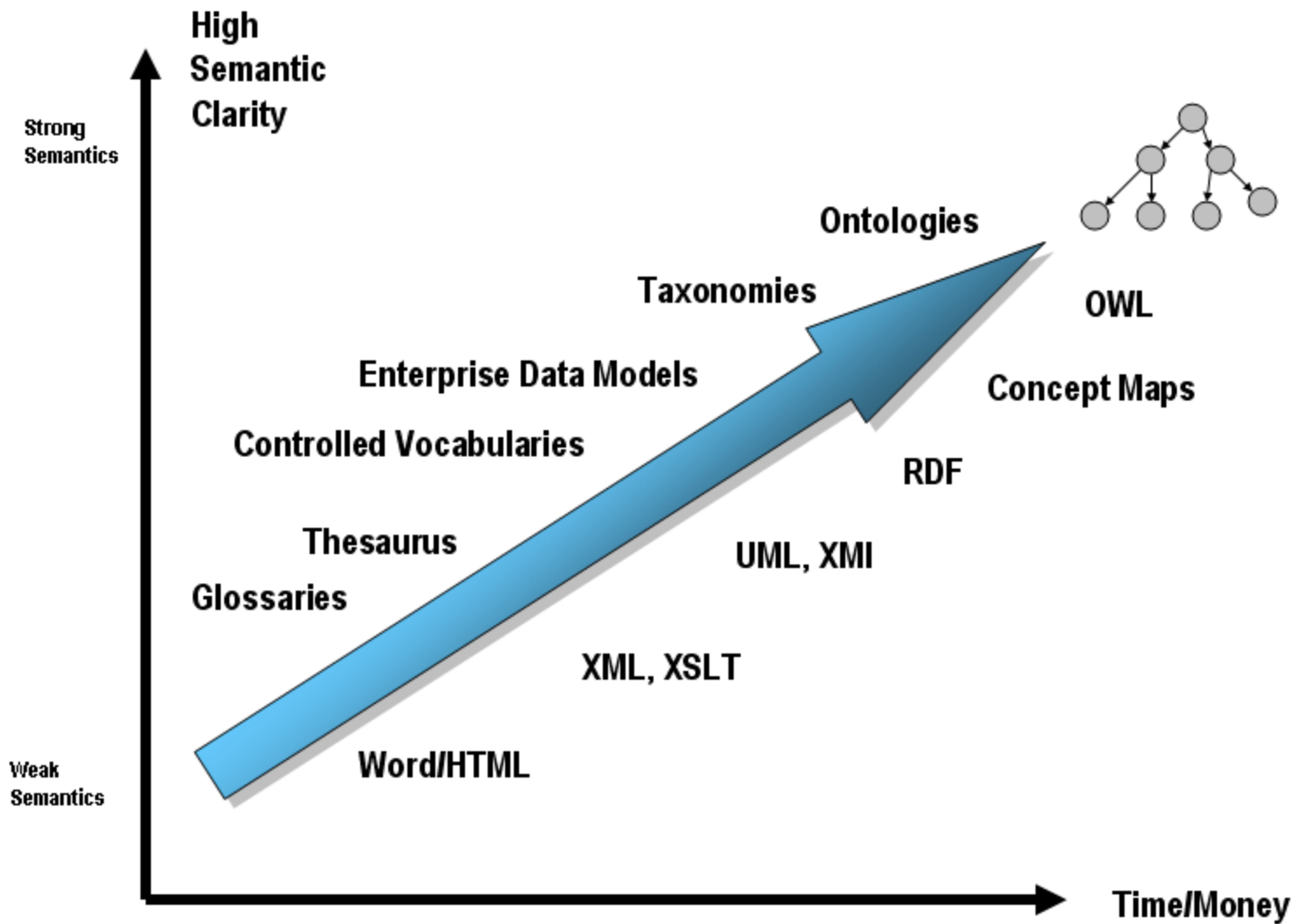
maxCardinality 1

symmetric

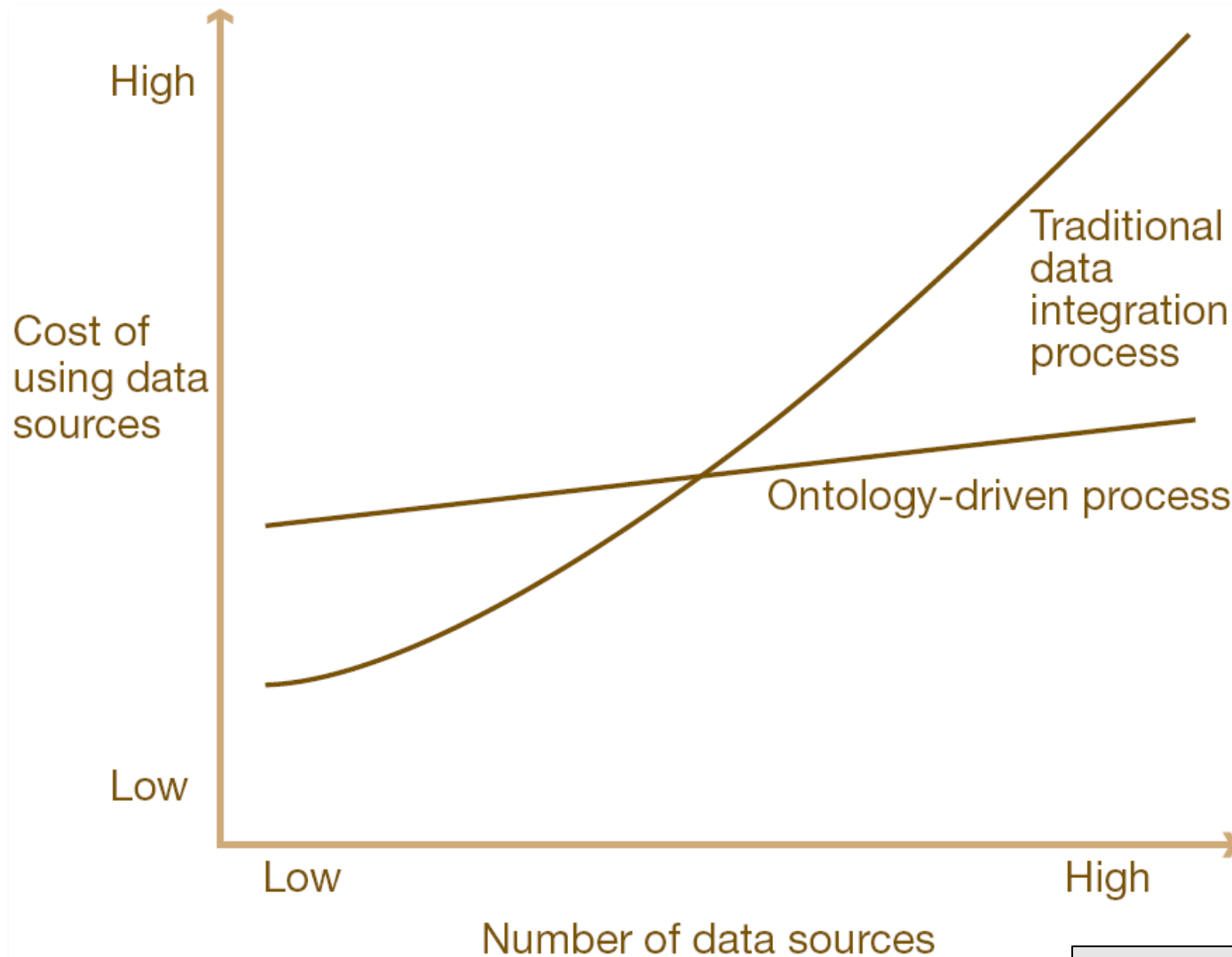
Types of Data



The cost of semantic clarity



Data integration cost



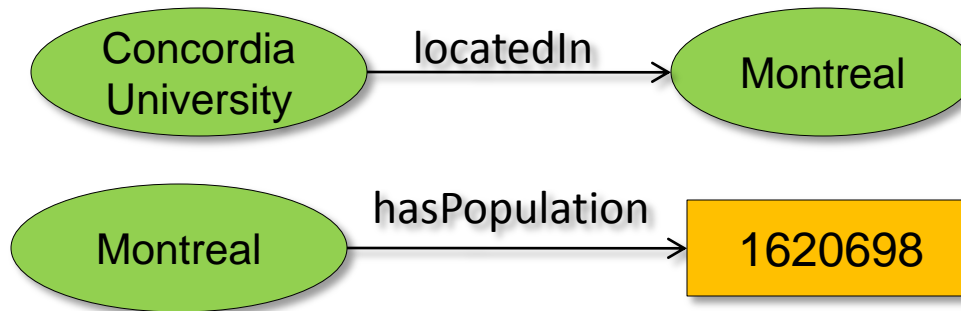
Semantic Web related standards

Resource Description Framework (RDF)

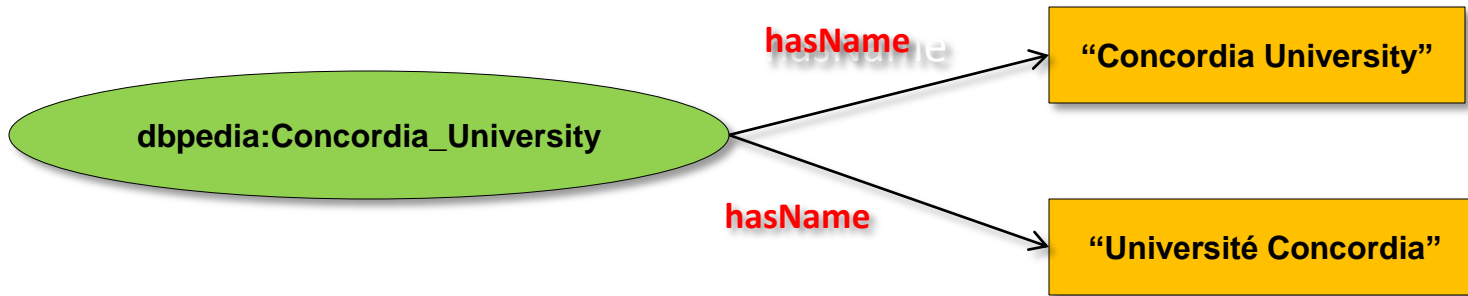
- A simple data model for
 - describing the *semantics* of information in a machine accessible way
 - representing meta-data (data about data)
- A set of representation syntaxes
 - XML (standard) but also N3, Turtle, ...
- Building blocks
 - *Resources* (with unique identifiers)
 - *Literals*
 - Named *relations* between pairs of resources (or a resource and a literal)

RDF (2)

- Everything is a triple
 - **Subject** (resource), **Predicate** (relation), **Object** (resource or literal)
- The RDF graph is a collection of triples

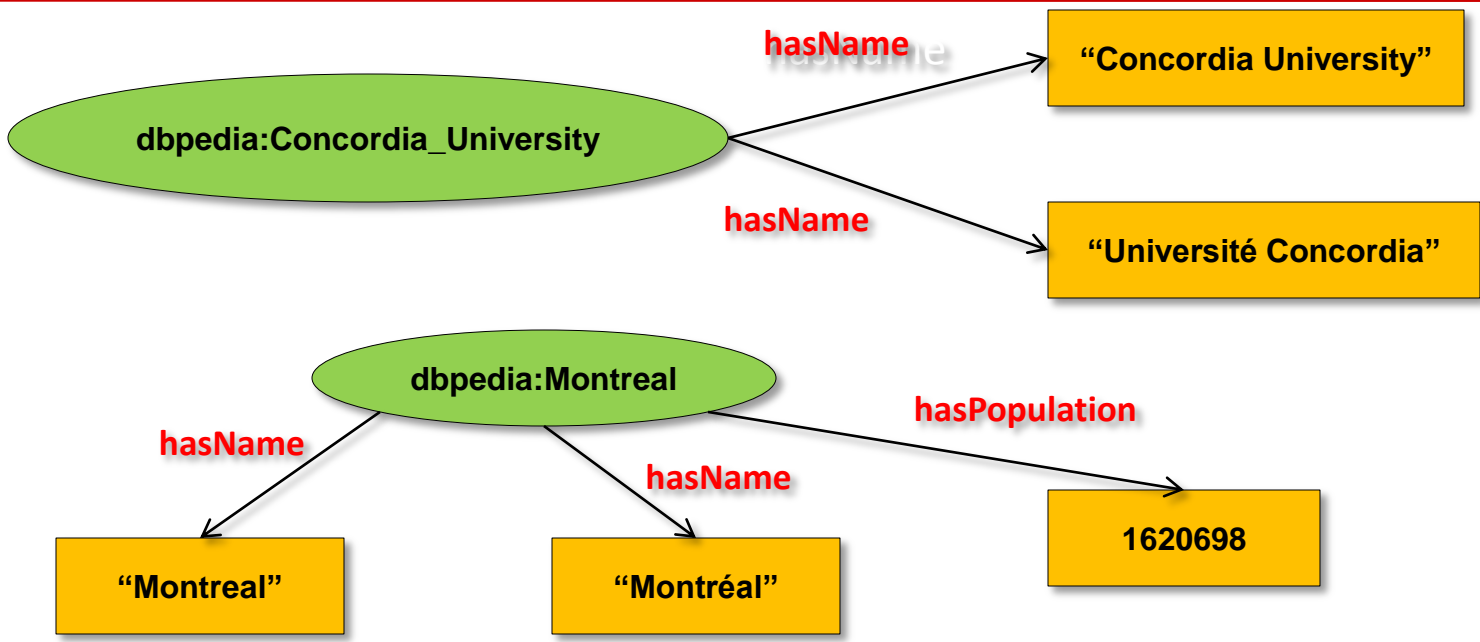


RDF (3)



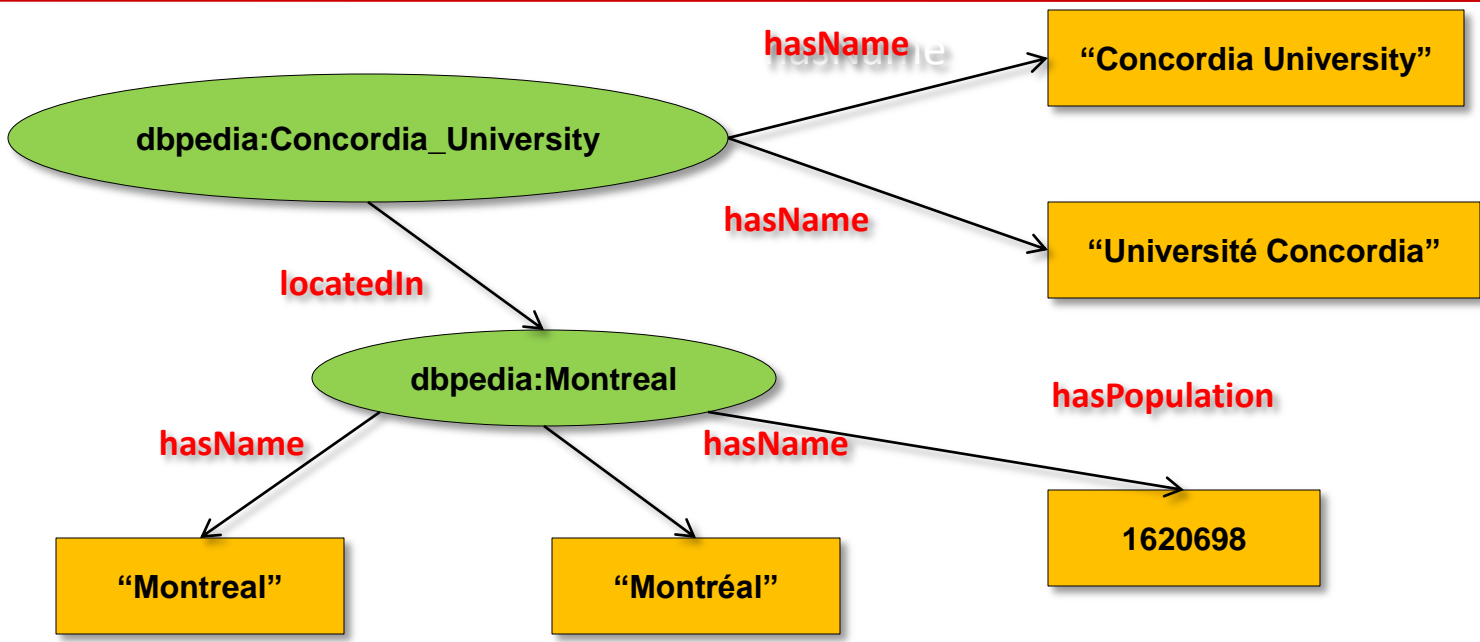
Subject	Predicate	Object
<code>http://dbpedia.org/resource/Concordia_University</code>	<code>hasName</code>	<code>"Concordia University"</code>
<code>http://dbpedia.org/resource/Concordia_University</code>	<code>hasName</code>	<code>"Université Concordia"</code>

RDF (4)



Subject	Predicate	Object
<code>http://dbpedia.org/resource/Montreal</code>	<code>hasName</code>	<code>"Montreal"</code>
<code>http://dbpedia.org/resource/Montreal</code>	<code>hasPopulation</code>	<code>1620698</code>
<code>http://dbpedia.org/resource/Montreal</code>	<code>hasName</code>	<code>"Montréal"</code>
<code>http://dbpedia.org/resource/Concordia_University</code>	<code>hasName</code>	<code>"Concordia University"</code>
<code>http://dbpedia.org/resource/Concordia_University</code>	<code>hasName</code>	<code>"Université Concordia"</code>

RDF (5)



Subject	Predicate	Object
http://dbpedia.org/resource/Montreal	hasName	"Montreal"
http://dbpedia.org/resource/Montreal	hasPopulation	1620698
http://dbpedia.org/resource/Montreal	hasName	"Montréal"
http://dbpedia.org/resource/Concordia_University	locatedIn	http://dbpedia.org/resource/Montreal
http://dbpedia.org/resource/Concordia_University	hasName	"Concordia University"
http://dbpedia.org/resource/Concordia_University	hasName	"Université Concordia"

RDF (6)

- RDF advantages
 - Simple but expressive data model
 - Global identifiers of all resources
 - Remove ambiguity
 - Easier & incremental data integration
 - Can handle incomplete information
 - Open world assumption
 - Schema agility
 - Graph structure
 - Suitable for a large class of tasks
 - Data merging is easier

SPARQL Protocol and RDF Query Language (SPARQL)

- SQL-like query language for RDF data
- Simple protocol for querying remote databases over HTTP
- Query types
 - *select* – projections of variables and expressions
 - *construct* – create triples (or graphs)
 - *ask* – whether a query returns results (result is true/false)
 - *describe* – describe resources in the graph

SPARQL (2)

- Anatomy of a SPARQL query
 - List of namespace prefix shortcuts
 - Query result definition (variables)
 - List of datasets
 - Graph patterns (restrictions)
 - Conjunctions, disjunctions, negation
 - Modifiers
 - Sort order, grouping

SPARQL (3)

```

PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX dbpedia: <http://dbpedia.org/resource/>
PREFIX dbp-ont: <http://dbpedia.org/ontology/>

SELECT DISTINCT ?university ?students
WHERE {
    ?university rdf:type dbpedia:Academic_institution .
    ?university dbp-ont:numberOfStudents ?students .
    ?university dbp-ont:city dbpedia:Montreal .
    FILTER (?students > 5000)
}
ORDER BY DESC (?students)
    
```

FactForge

SPARQL Query

Results for [PREFIX rdf:<http://www.w3... \(6\)](#) View as [Exhibit](#) Downlo

university	students
dbpedia:Con_U	43944
dbpedia:HEC_Montreal	12000
dbpedia:Collège_Ahuntsic	10100
dbpedia:John_Molson_School_of_Business	8026
dbpedia:CEGEP_Vanier	6100
dbpedia:Cégep_du_Vieux_Montréal	6100

RDF Schema (RDFS)

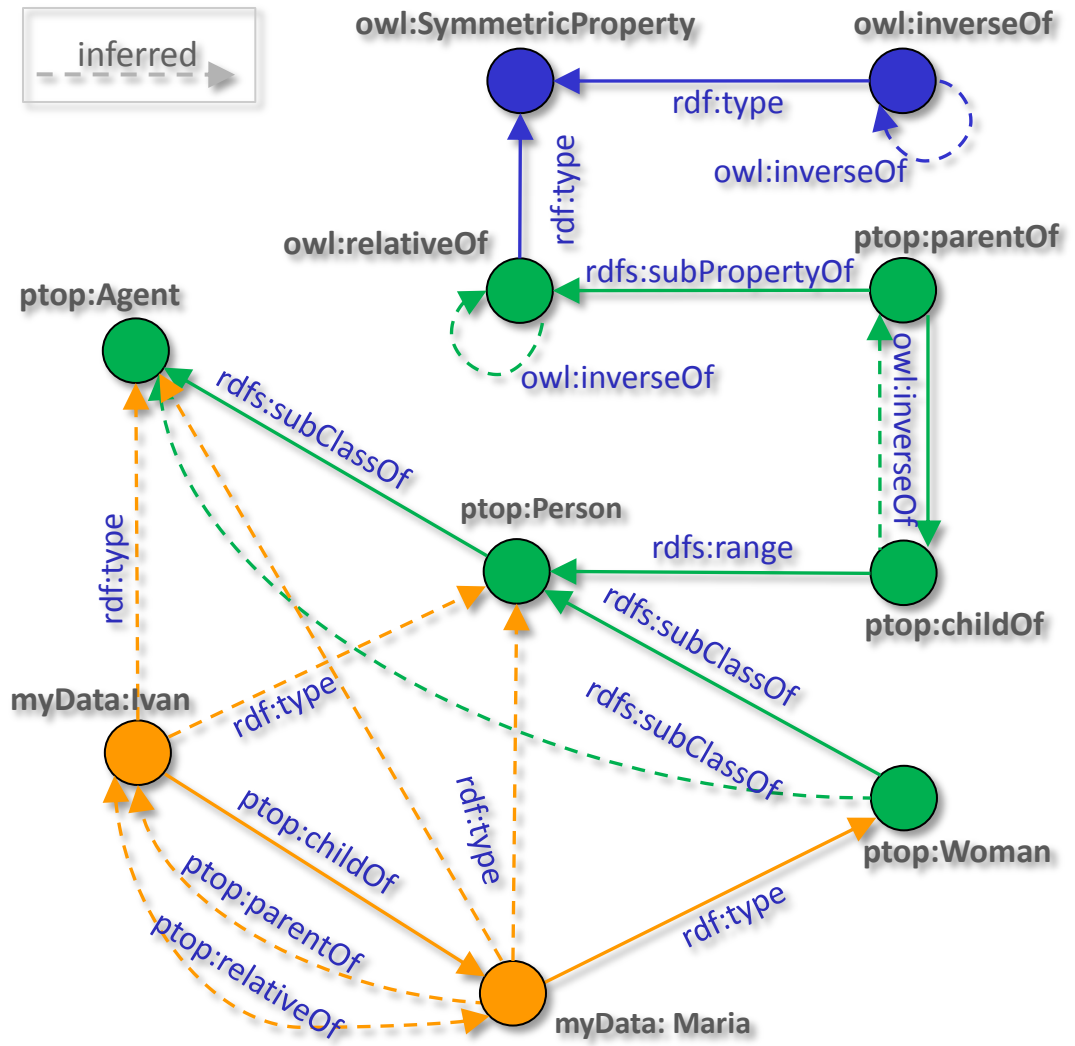
- RDFS provides means for
 - Defining *Classes* and *Properties*
 - Defining hierarchies (of classes and properties)
- RDFS differs from XML Schema (XSD)
 - Open World Assumption
 - RDFS is about describing resources, not about validation
- Entailment rules (axioms)
 - Infer new triples

RDFS (2)

- Entailment rules

1: $s p o$ (if o is a literal)	\Rightarrow $\neg \exists n \text{ rdf:type rdfs:Literal}$
2: $p \text{ rdfs:domain } x$ $\& s p o$	$\Rightarrow s \text{ rdf:type } x$
3: $p \text{ rdfs:range } x$ $\& s p o$	$\Rightarrow o \text{ rdf:type } x$
4a: $s p o$	$\Rightarrow s \text{ rdf:type rdfs:Resource}$
4b: $s p o$	$\Rightarrow o \text{ rdf:type rdfs:Resource}$
5: $p \text{ rdfs:subPropertyOf } q$ $\& q \text{ rdfs:subPropertyOf } r$	$\Rightarrow p \text{ rdfs:subPropertyOf } r$
6: $p \text{ rdf:type rdf:Property}$	$\Rightarrow p \text{ rdfs:subPropertyOf } p$
7: $s p o$ $\& p \text{ rdfs:subPropertyOf } q$	$\Rightarrow s q o$
8: $s \text{ rdf:type rdfs:Class}$	$\Rightarrow s \text{ rdfs:subClassOf rdfs:Resource}$
9: $s \text{ rdf:type } x$ $\& x \text{ rdfs:subClassOf } y$	$\Rightarrow s \text{ rdf:type } y$
10: $s \text{ rdf:type rdfs:Class}$	$\Rightarrow s \text{ rdfs:subClassOf } s$
11: $x \text{ rdfs:subClassOf } y$ $\& y \text{ rdfs:subClassOf } z$	$\Rightarrow x \text{ rdfs:subClassOf } z$
12: $p \text{ rdf:type rdfs:ContainerMembershipProperty}$	$\Rightarrow p \text{ rdfs:subPropertyOf rdfs:member}$
13: $o \text{ rdf:type rdfs:Datatype}$	$\Rightarrow o \text{ rdfs:subClassOf rdfs:Literal}$

RDFS (3)



Web Ontology Language (OWL)

- More expressive than RDFS
 - Identity equivalence/difference
 - *sameAs, differentFrom, equivalentClass/Property*
 - More expressive class definitions
 - Class intersection, union, complement, disjointness
 - Cardinality restrictions
 - More expressive property definitions
 - Object/Datatype properties
 - Transitive, functional, symmetric, inverse properties
 - Value restrictions

OWL (2)

- What can be done with OWL?
 - *Consistency checks* - Are there contradictions in the logical model
 - *Satisfiability checks* - Are there classes that cannot have any instances?
 - *Classification* - What is the type of a particular instance?
- OWL sublanguages
 - **OWL Lite** – low expressiveness / low computational complexity
 - **OWL DL** – high expressiveness / decidable & complete
 - **OWL Full** – max expressiveness / no guarantees

Simple Knowledge Organization System (SKOS)



- Represent simple knowledge bases
 - Taxonomies, thesauri, classifications, vocabularies, etc
 - RDF based
- SKOS essentials
 - *Concepts* – describe entities
 - *Labels* – lexical means to refer to a concept
 - *Relationships* – hierarchy (*skos:broader*, *skos:narrower*) or relatedness (*skos:related*)
 - *Notes* – human readable documentation
 - *Schemes* – compilation of concepts

Rule Interchange Format (RIF)

- Goals
 - Define a framework for rule languages for the Semantic Web
 - If *<condition>* then *<conclusion>*
 - Define a standard format/syntax for interchanging rules
- Several dialects defined so far
 - different expressivity & complexity
 - **RIF BLD** (Basic Logic Dialect)
 - Rule condition/conclusion are *monotonic* (like in OWL/RDF)
 - **RIF PRD** (Production Rule dialect)
 - Condition/conclusion are *non-monotonic* (retraction)

RDFa

- A set of XHTML attributes
 - Embed RDF annotations in web pages
 - The DC and FOAF vocabularies can be easily used for most simple annotations
 - Creator, title, contact info, ...

Original HTML

```

<div >
  <ul>
    <li>
      <a href="http://example.com/bob/">Adam</a>
    </li>
    <li>
      <a href="http://example.com/eve/">Eve</a>
    </li>
  </ul>
</div>
    
```

Annotated XHTML

```

<div xmlns:foaf="http://xmlns.com/foaf/0.1/">
  <ul>
    <li typeof="foaf:Person">
      <a property="foaf:name"
        rel="foaf:homepage"
        href="http://example.com/bob/">Adam</a>
    </li>
    <li typeof="foaf:Person">
      <a property="foaf:name"
        rel="foaf:homepage"
        href="http://example.com/eve/">Eve</a>
    </li>
  </ul>
</div>
    
```

RDFa (2)

- Less than 5% of web pages have RDFa annotations (Google, 2010)
- But many organizations already publish or consume RDFa:
 - Google, Yahoo
 - Facebook, MySpace, LinkedIn
 - Best Buy, Tesco, O'Reilly,
 - SlideShare, Digg
 - WhiteHouse.gov, Library of Congress, UK government
 - Newsweek, BBC

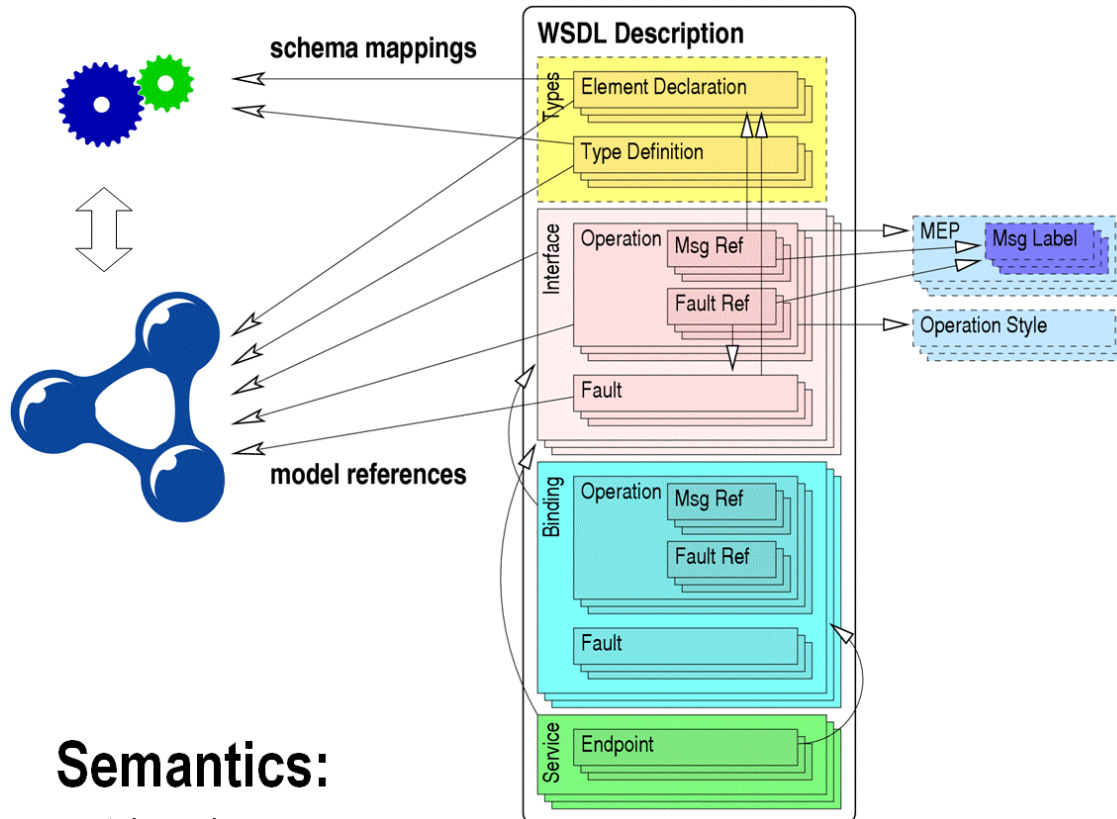
Semantic Annotations for WSDL (SAWSDL)



- Semantic annotations for Web Services
 - Embed semantic annotations within WSDL 2.0 service descriptions
 - Annotations are agnostic of the ontology language
 - Based on the extensibility mechanism of WSDL 2.0
 - Existing tools that do not understand the semantics will just ignore the annotations
- Elements
 - *Model reference* – associate a WSDL element with a concept in some ontology
 - *Lifting/lowering schema* – mappings between the XML and the ontology data

Web Services and Business Process Management (2)

- SAWSDL at a glance



Semantics:

- ontology classes
 - discovery, composition
 - filtering, ranking
- lifting/lowering mappings
 - mediation, invocation
- functionality categories
 - publishing, discovery, composition
- anything, really

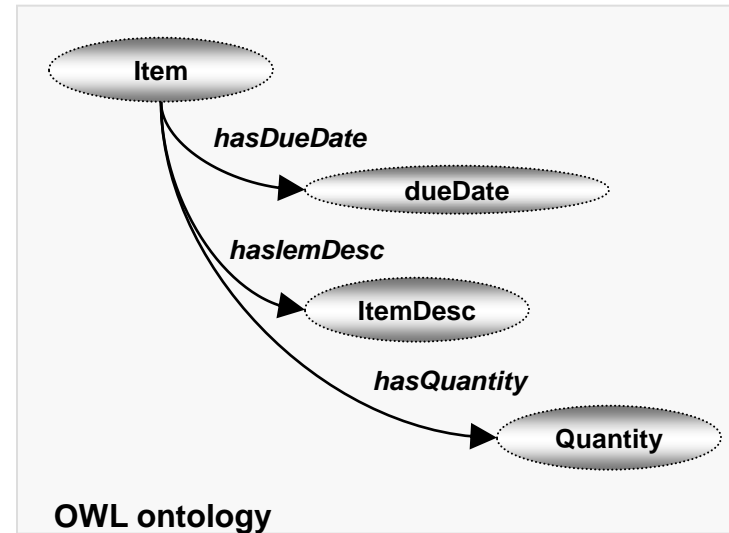
Web Services and Business Process Management (3)

- SAWSDL example

```

<complexType name="POItem" >
  <all>
    <element name="dueDate" nillable="true" type="dateTime"
      sawsdl:modelReference="
        http://www.w3.org/2002/ws/sawsdl/spec/ontology/purchaseorder#DueDate"/>
    <element name="qty" type="float"
      sawsdl:modelReference="
        http://www.w3.org/2002/ws/sawsdl/spec/ontology/purchaseorder#Quantity"/>
    <element name="EANCode" nillable="true" type="string"
      sawsdl:modelReference="
        http://www.w3.org/2002/ws/sawsdl/spec/ontology/purchaseorder#ItemCode"/>
    <element name="itemDesc" nillable="true" type="string"
      sawsdl:modelReference="
        http://www.w3.org/2002/ws/sawsdl/spec/ontology/purchaseorder#ItemDesc" />
  </all>
</complexType>
  
```

WSDL complex type element



OWL ontology

Some Applications of Semantic Technologies

Linked Data

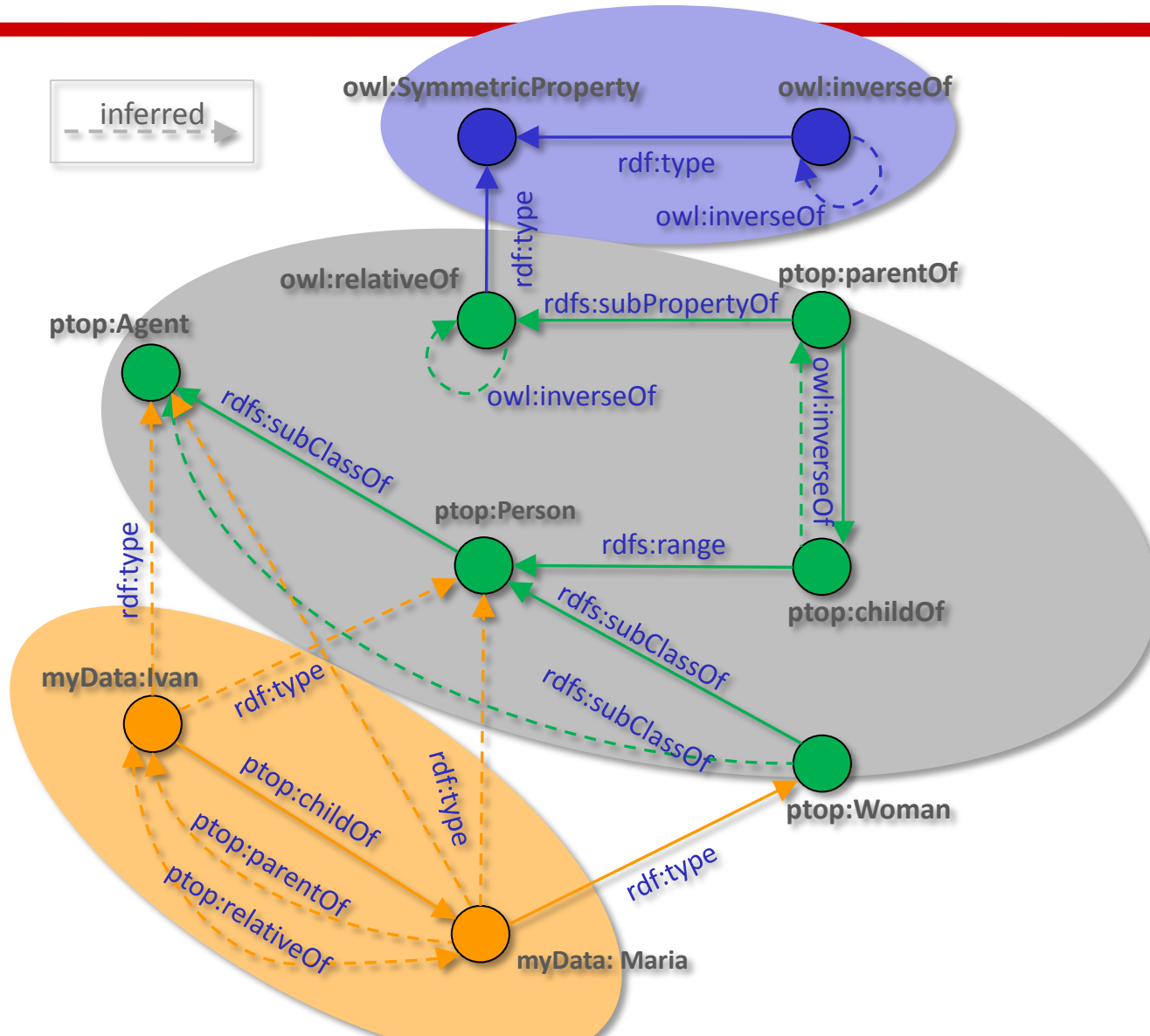
*“To make the Semantic Web a reality, it is necessary to have a large volume of data available on the Web in a standard, reachable and manageable format. In addition the relationships among data also need to be made available. This collection of interrelated data on the Web can also be referred to as **Linked Data**. Linked Data lies at the heart of the Semantic Web: large scale integration of, and reasoning on, data on the Web.” (W3C)*

- *Linked Data* is a set of principles that allows publishing, querying and browsing of RDF data, distributed across different servers
 - similar to the way HTML is currently published & consumed

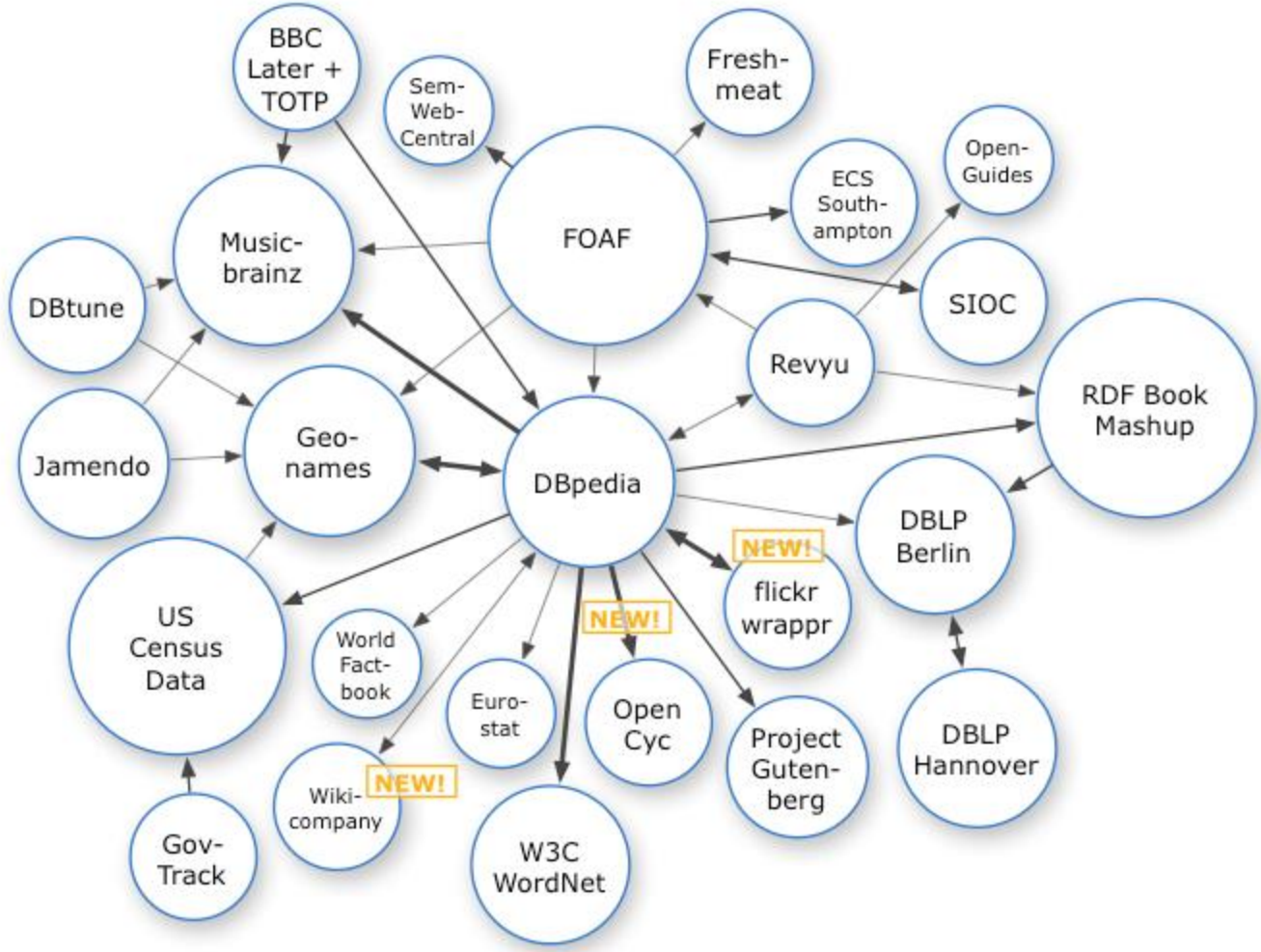
Linked Data design principles

1. Unambiguous identifiers for objects (resources)
 - Use URIs as names for things
2. Use the structure of the web
 - Use HTTP URIs so that people can look up the names
3. Make it easy to discover information about an object (resource)
 - When someone looks up a URI, provide useful information
4. Link the object (resource) to related objects
 - Include links to other URIs

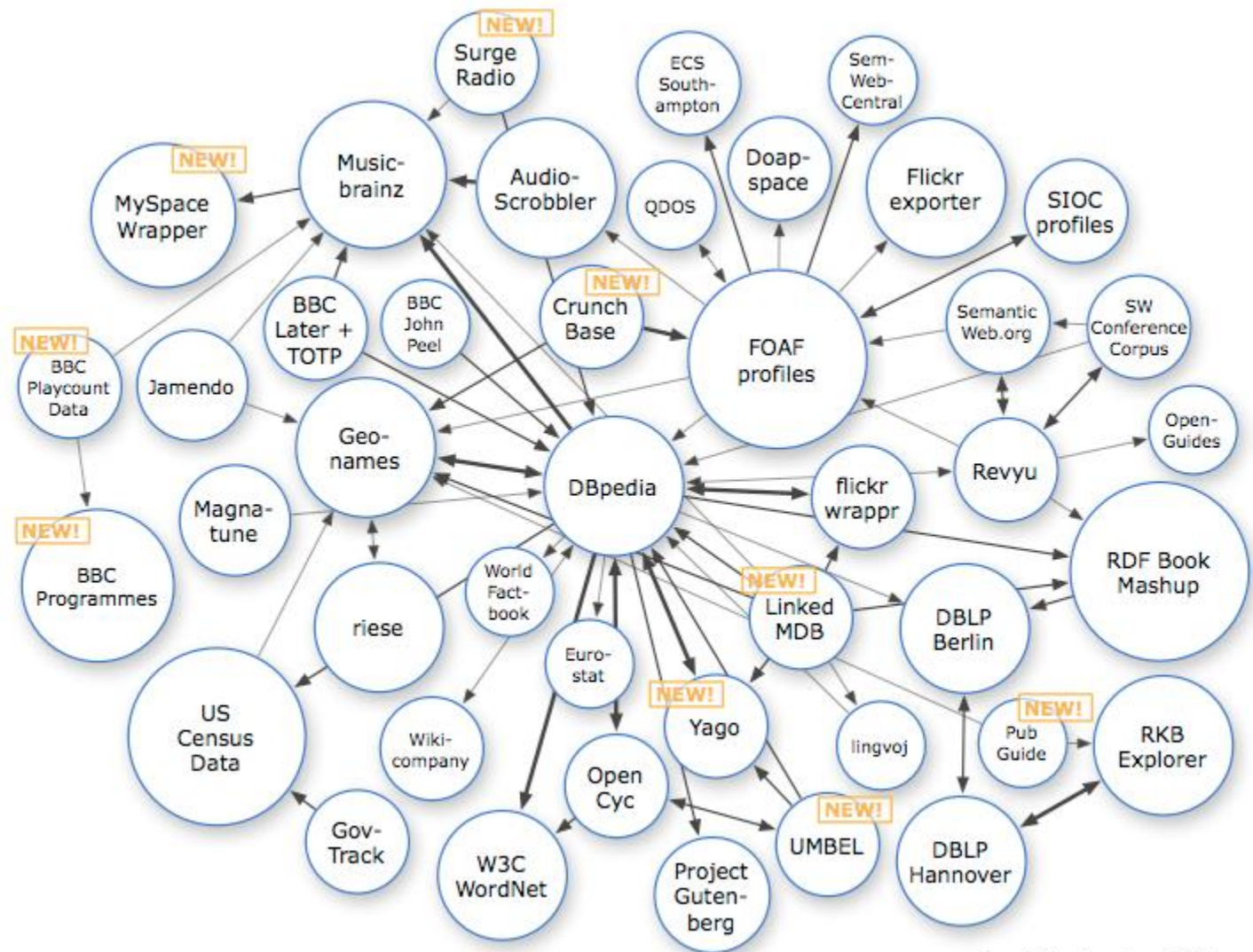
Linked Data (2)



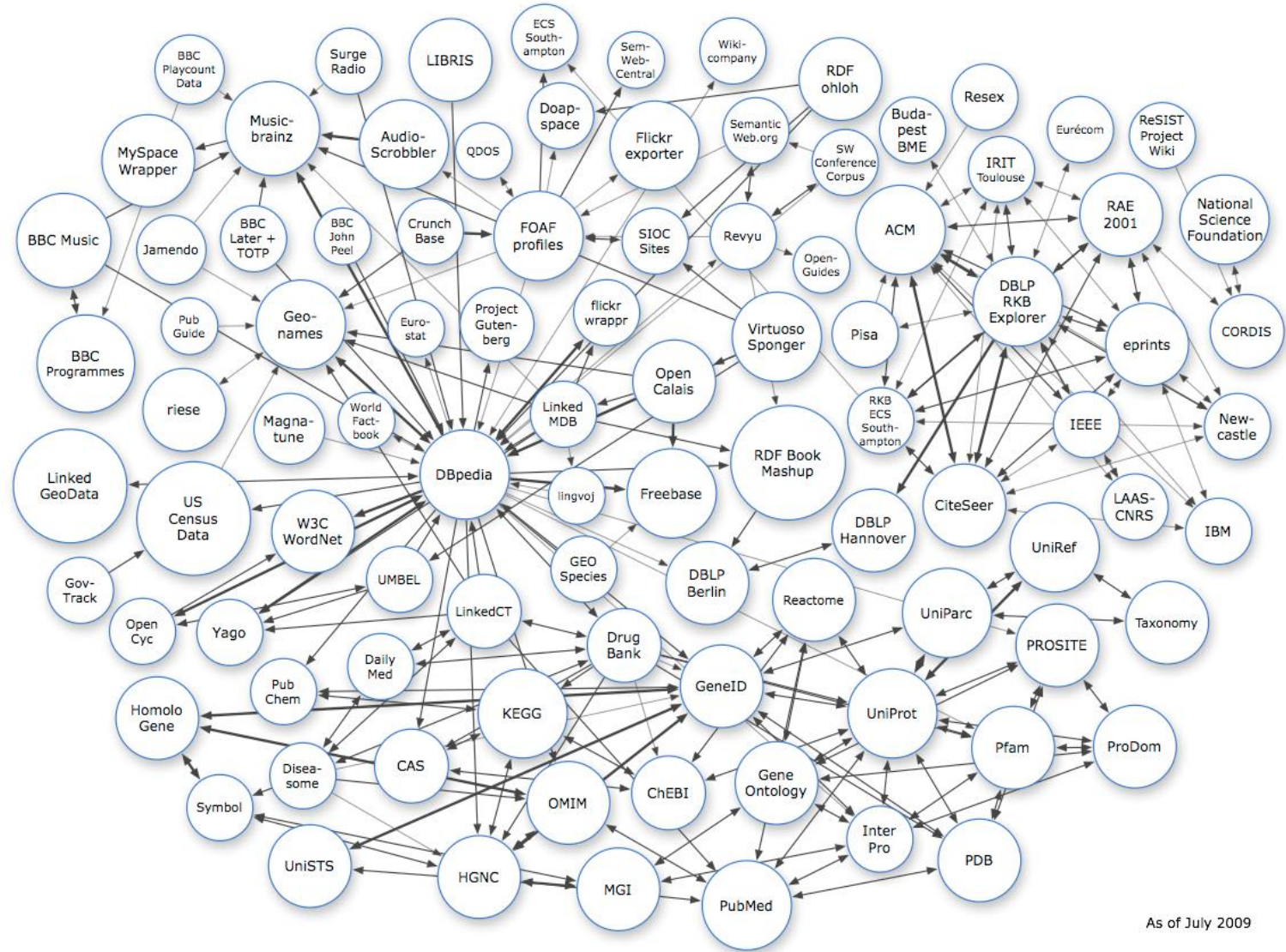
Linked Data evolution – Oct 2007



Linked Data evolution – Sep 2008



Linked Data evolution – Jul 2009



As of July 2009

Open Government Data



An Official Web Site of the United States Government Thursday, August 19, 2010 Text: A+ A- A Share

DATA.GOV
EMPOWERING PEOPLE

HOME DATA TOOLS COMMUNITY METRICS DIALOGUE GALLERY WHAT'S NEW

DEEPWATER HORIZON RESPONSE

VIEW MORE ▶



Most Popular Datasets

1. Food and Drug Administration--Recalls
2. Worldwide M1+ Earthquakes, Past 7 Days
3. AVAILABLE TECHNOLOGIES
4. TSCA Inventory
5. 2000 Federal Register in XML

SEARCH OUR CATALOGS

Search our catalogs.. SEARCH ▶

APPS



With so much government data to work with, developers are creating a wide variety of

COMMUNITY

Data.gov is leading the way in democratizing public sector data and driving innovation. The data is being surfaced from many locations making the Government data stores available to researchers to perform their own analysis. Developers are finding good uses for the datasets, providing interesting and useful applications that allow for new views and public analysis. This is a work in progress, but this movement is spreading to cities, states, and

SEMANTIC WEB

As the Web of linked documents evolves to include the Web of linked data, we're working to maximize the potential of Semantic Web technologies to realize the promise of Linked Open Government Data.



Open Government Data (2)

The screenshot shows the data.gov.uk website. At the top, it features the HM Government logo and the text 'data.gov.uk BETA'. A search bar is present with the placeholder text 'What are you looking for?' and a 'Search' button. Below this is a navigation menu with links for Data, Apps, Ideas, Forum, Wiki, Blog, Transparency, Linked Data, and Resources. A 'Log in or Sign up' button is also visible. The main content area includes a message about transparency, a blue molecular structure icon, and a 'Latest Blog Post' section for 'Data.gov.uk - site redesign' dated 19 AUG. Below this are sections for 'Data', 'Apps', and 'Ideas', each with a 'view all' link and a 'share your' button. The 'Data' section lists datasets like 'Combined Online Information...' and 'Higher Education Statistics...'. The 'Apps' section lists 'Where Can I Live?' and 'Post Box Finder'. The 'Ideas' section lists 'Link Post Code to Go...'. A 'Popular tags' section on the right lists tags like 'health (725)', 'care (435)', 'health-and-social-care (427)', 'population (388)', and 'health-well-being-and-care (327)'.

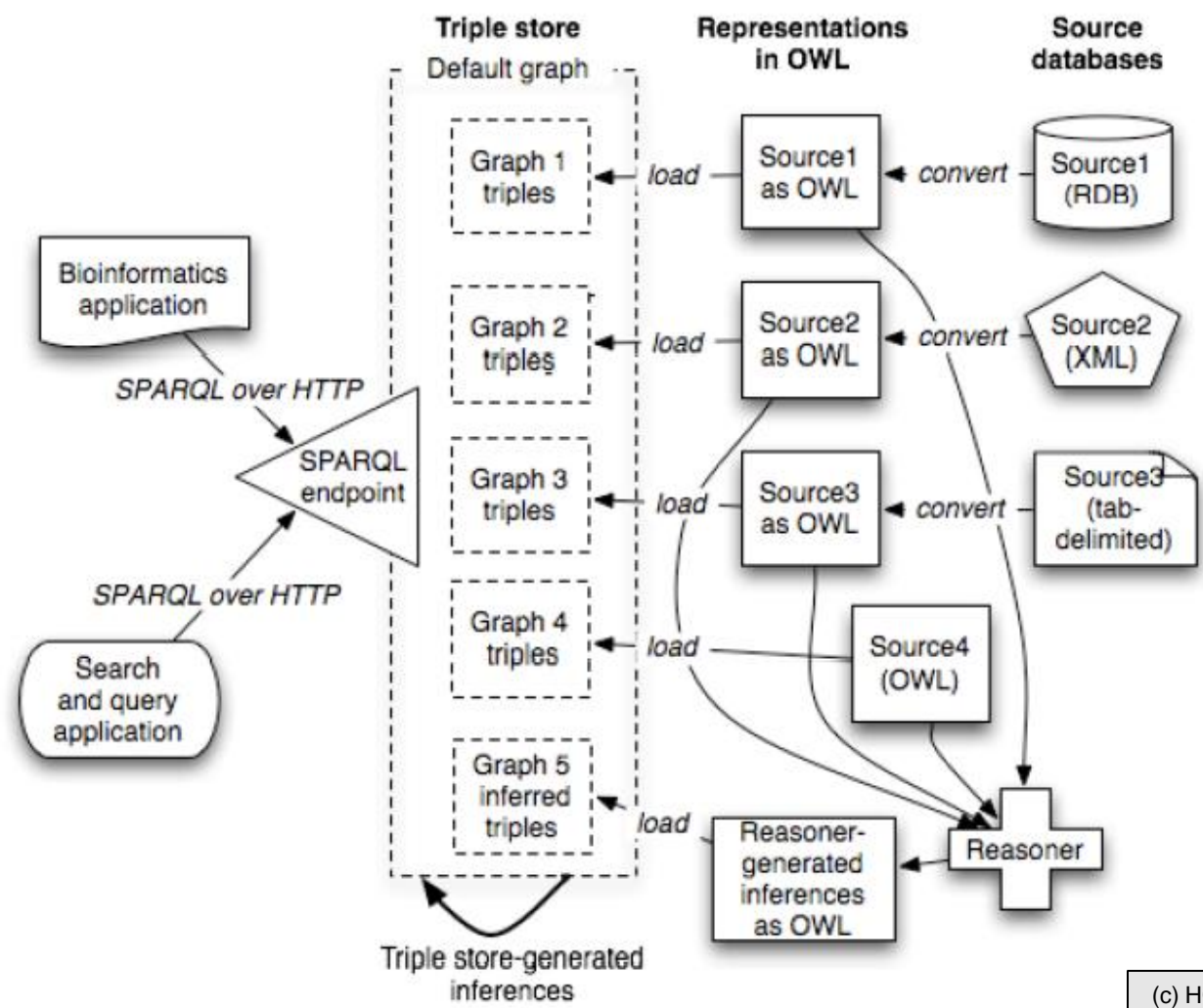
Linked Data success stories

- BBC Music
 - Integrates information from MusicBrainz and Wikipedia for artist/band infopages
 - Information also available in RDF (in addition to web pages)
 - 3rd party applications built on top of the BBC data
 - BBC also contributes data back to the MusicBrainz
- NY times
 - Maps its thesaurus of 1 million entity descriptions (people, organisations, places, etc) to DBpedia and Freebase

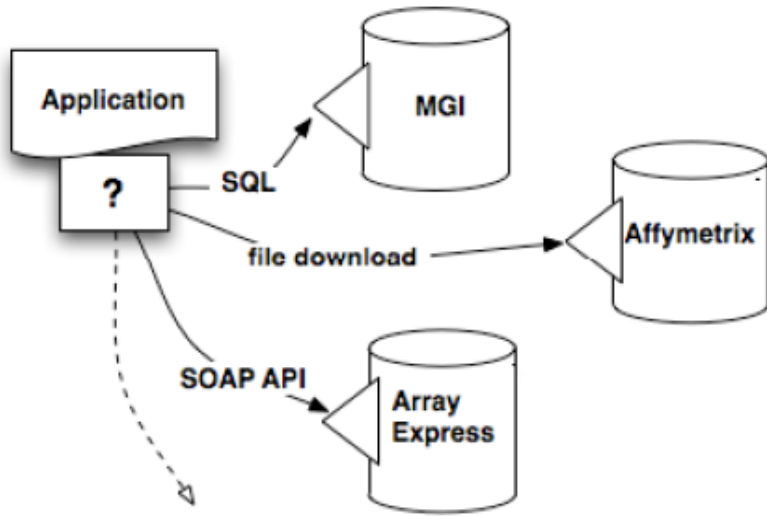
Life Sciences

- Semantic Web Health Care and Life Sciences (HCLS) @ W3C
 - <http://www.w3.org/blog/hcls/>
- Potential benefits of ST
 - Interoperability – re-using common models
 - Data integration – too many disparate datasources
 - Efficient query answering

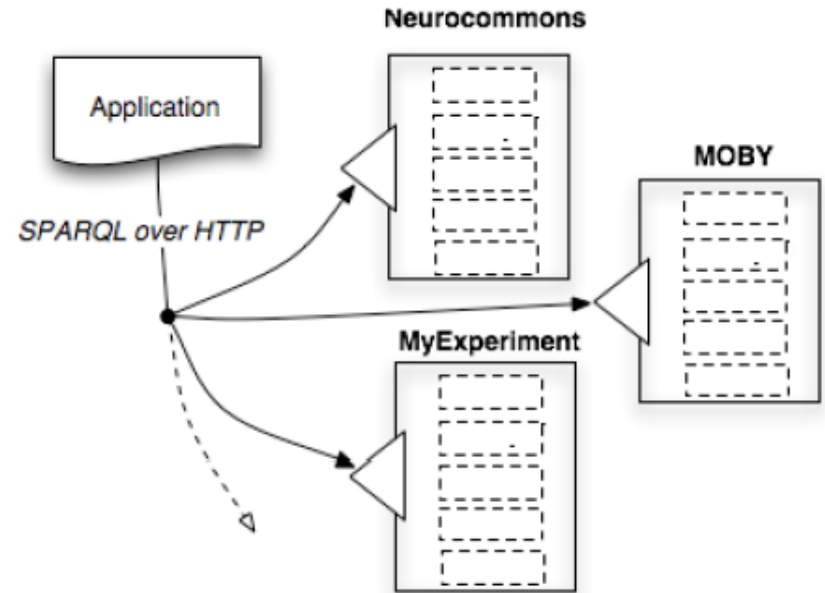
Life Sciences (2)



Life Sciences (3)

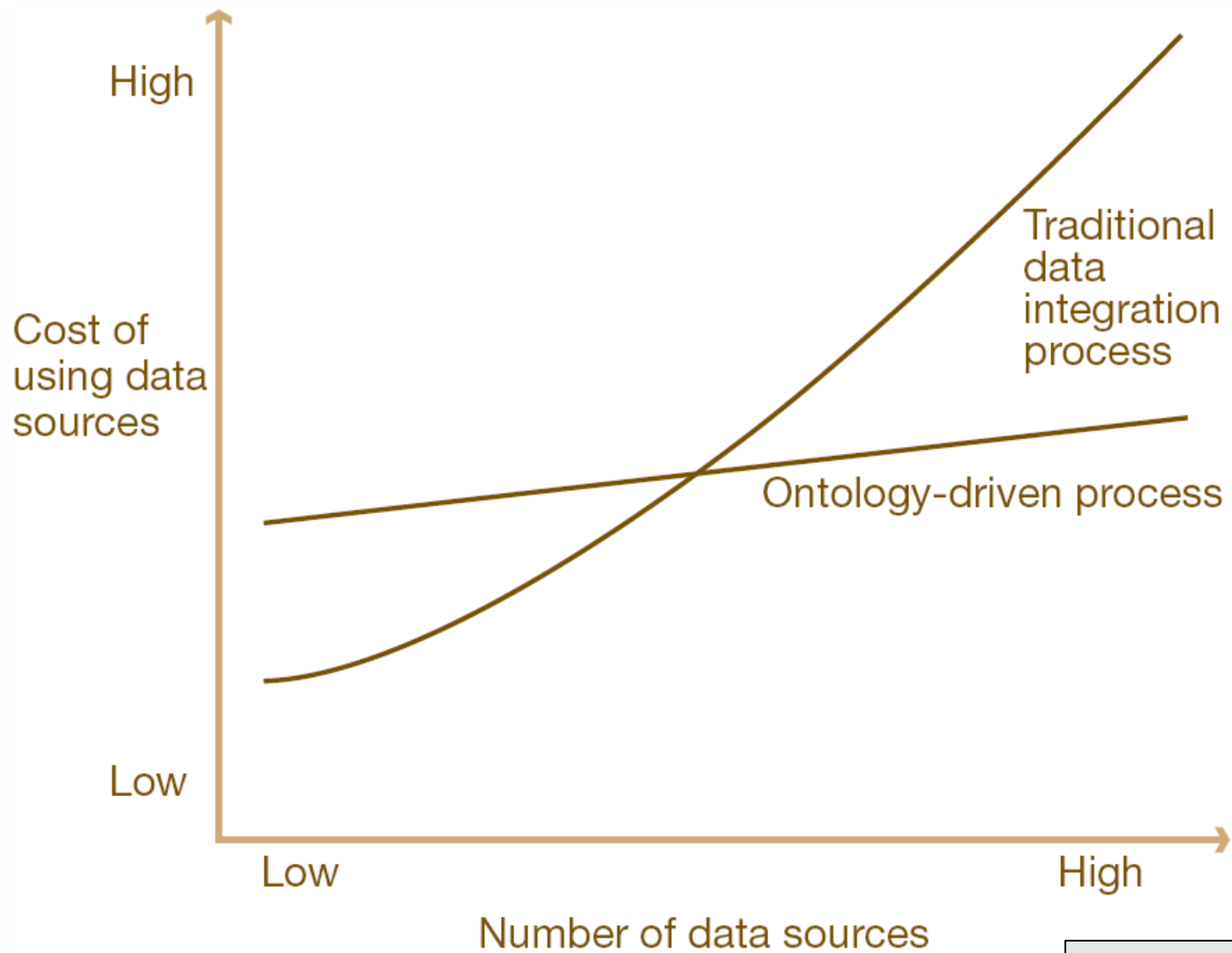


distributed querying at present



distributed querying with RDF and SPARQL

Data integration cost



Semantic Life Science Models

OWL	OBO
Generic or top-down	Specific or bottom-up
Describe any domain (in theory)	Focus on supporting existing users and applications
Background in AI	Background in genome annotations
Ontology (strict semantics)	Vocabularies (relaxed semantics)

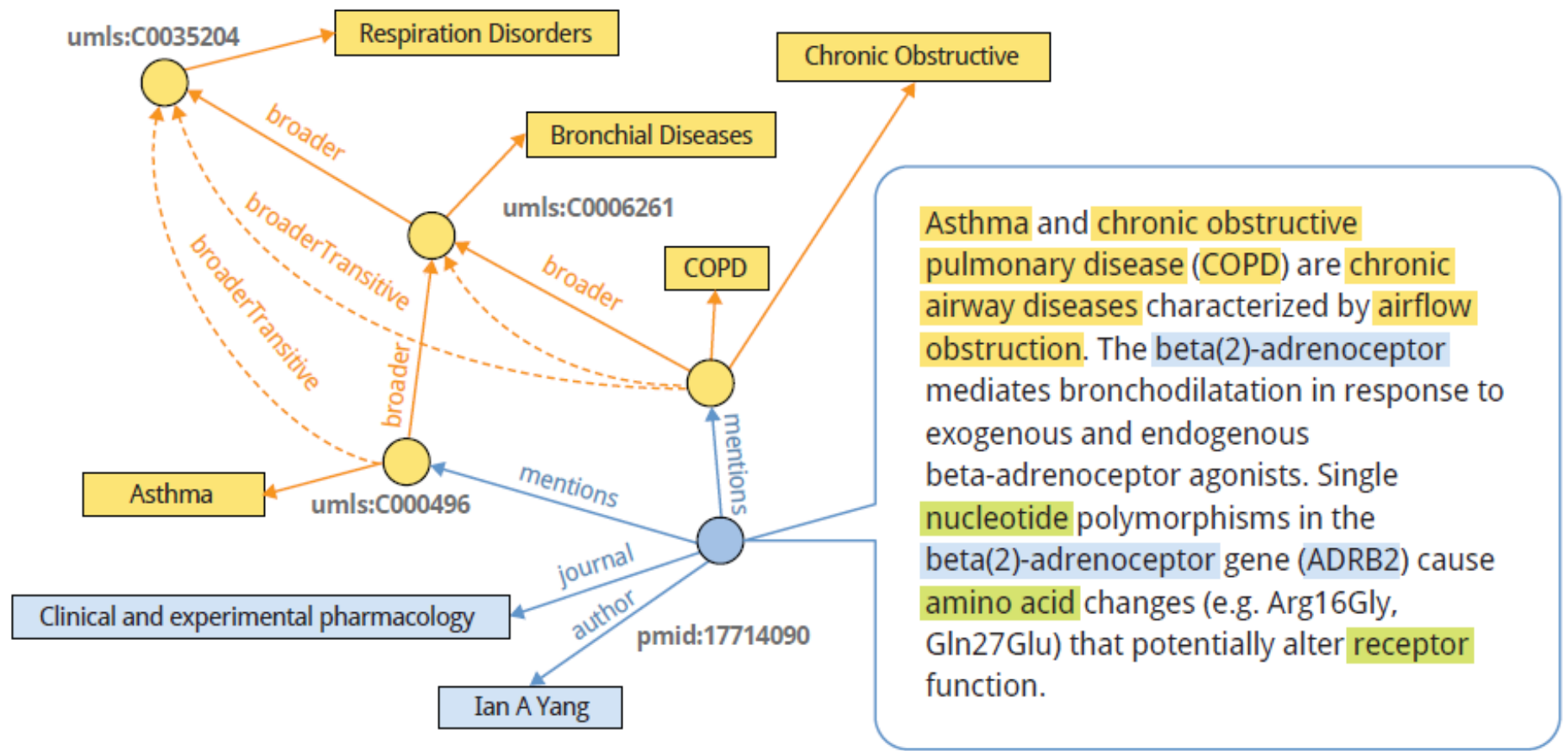
Semantic Annotation

- Semantic annotation (of text)
 - The process of linking text fragments to structured information
 - Organisations, Places, Products, Human Genes, Diseases, Drugs, etc.
 - Combines Text Mining (Information Extraction) with Semantic Technologies
- Benefits of semantic annotations
 - *Improves the text analysis process*
 - by employing Ontologies and knowledge from external Knowledge Bases / structured data sources

Semantic Annotation (2)

- Benefits of semantic annotations (cont.)
 - Provides *unambiguous (global) references for entities* discovered in text
 - Different from tagging
 - Provide the means for *semantic search*
 - Together or independently of the original text
 - Improved *data integration*
 - Documents from different data sources can share the same semantic concepts

Semantic Annotation (2)



Semantic Annotation (3)

The screenshot shows the Guardian website interface with a KIM Explorer window overlaid. The Explorer window displays semantic annotations for the word "Greece":

- Greece is a Country , Trusted^{tip!}
- Located in Europe , Southern Europe
- has Alias Greece , Hellenic Republic , Ellas , Eliniki Dhimokratia , Grecia ... (5 more)
- has Main Alias Greece
- Part of Europe , Southern Europe
- Subregion of Europe , Southern Europe
- has Capital Athens
- hasAdjective Greek

A yellow arrow points from the Explorer window to the main article on the page, which is titled "Greece protesters clash with police".

Semantic Search

- Semantic Search
 - In addition to the terms/keywords, explore the entity descriptions found in text
 - Make use of the semantic relations that exist between these entities
- Example
 - Query – “*Documents about a telecom companies in Europe related to John Smith from Q1 or Q2/2010*”
 - Document containing “*At its meeting on the 10th of May, the board of Vodafone appointed John G. Smith as CTO*” will **not match**

Semantic Search (2)

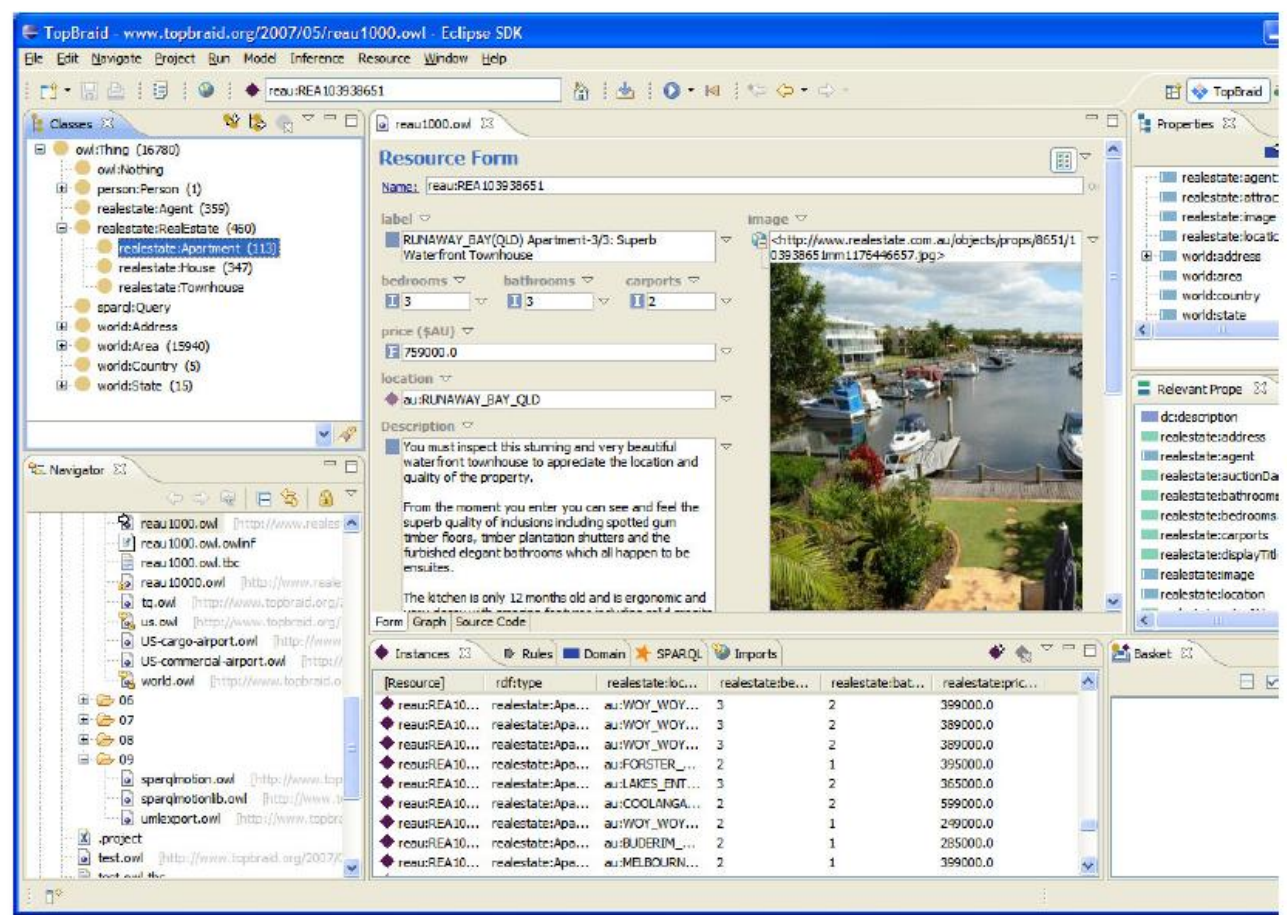
- Classical IR will fail to recognise that
 - Vodafone is a mobile operator, and mobile operator is a type of telecom
 - Vodafone is in the UK , which is part of Europe
 - => Vodafone is a “telecom company in Europe”
 - 5th of May is in Q2
 - John G. Smith may be the same as John Smith

- Potential benefits of using ST
 - Automated mediation & interoperability
 - Capability based discovery
 - Automated composition of complex workflows
- Relevant activities & standards
 - SAWSDL – www.w3.org/2002/ws/sawSDL
 - Embed semantic annotations within WSDL descriptions
 - WSMO / Conceptual Models for Services (CMS) - <http://cms-wg.sti2.org>
 - Ontology and methodology for semantic annotation of services and business processes

Tools

Ontology editors

- TopBraid Composer
- <http://www.topquadrant.com>



Ontology editors (2)

- Altova SemanticWorks
- <http://www.altova.com/semanticworks.html>

The screenshot shows the Altova SemanticWorks application window titled "Altova SemanticWorks - [AltovaProducts *]". The interface includes a menu bar (File, Edit, View, RDF/OWL, Tools, Window, Help), a toolbar with various icons, and a main workspace displaying an ontology diagram. The diagram features several classes and properties:

- prod:hasEdition** (Property): A dashed yellow box highlights this property and its domain **prod:Enterprise**. The property is shown as a box with a cardinality of "1" and a range of "1".
- prod:Enterprise** (Class): A dashed yellow box highlights this class and its domain **prod:hasEdition**. It is shown as a box with a cardinality of "1" and a range of "1".
- prod:XMLSpyEnterprise** (Class): A box representing this class, with a cardinality of "1" and a range of "1".
- prod:version** (Property): A box representing this property, with a cardinality of "1" and a range of "1".
- prod:XMLSpy** (Class): A dashed yellow box highlights this class and its domain **prod:XMLSpyEnterprise**. It is shown as a box with a cardinality of "1" and a range of "1".
- prod:Product** (Class): A box representing this class, with a cardinality of "1" and a range of "1".
- prod:XMLSpy** (Property): A box representing this property, with a cardinality of "1" and a range of "1".

On the right side, the "Details" panel shows the selected element's information:

Property	Value
prod:hasEdition	prod:Ent...
prod:hasEdition	edit to a...
prod:version	2006
prod:version	edit to ad...

At the bottom, the "Overview" panel shows a small thumbnail of the ontology diagram, and the "Errors" panel displays the message: "This ontology is well-formed."

Ontology editors (3)

- Protégé

- <http://protege.stanford.edu>
- <http://webprotege.stanford.edu>

The screenshot shows the Protégé 3.1 interface with a project named 'travel'. The main workspace displays a class hierarchy for 'Destination'. The hierarchy includes 'Destination' at the root, which has subclasses: 'RuralArea', 'UrbanArea', 'BudgetHotelDestination', 'RetireeDestination', 'Beach', 'FamilyDestination', 'QuietDestination', 'BackpackersDestination', 'AccommodationRating', and 'Contact'. 'RuralArea' and 'UrbanArea' are further divided into 'Farmland', 'NationalPark', 'Town', and 'City'. 'City' is a subclass of 'Capital'. 'Surfing' is associated with 'BackpackersDestination'. A legend at the bottom right explains the symbols: a yellow circle for 'Destination', a circle with a vertical line for 'hasAccommodation BudgetAccommodation', and a circle with a horizontal line for 'hasActivity (Sports Adventure)'. The interface also shows a 'CLASS BROWSER' on the left and a toolbar at the top.

The screenshot shows the Protégé interface for the 'Hematopoietic_Growth_Factor' class. The 'Properties for Hematopoietic_Growth_Factor' panel is active, showing a table with the following data:

Property	Value	Language
NCI_META_CUI	C1026822	
DEFINITION	NCIEndogenous growth factors which promote the development of blood cells.	
Preferred_Name	Hematopoietic Growth Factor	

Below the properties, there are sections for 'Axioms for Hematopoietic_Growth_Factor', 'Equivalent classes (Necessary and Sufficient conditions)', and 'Notes for Hematopoietic_Growth_Factor'. The 'Notes' section contains a table with columns for 'Text', 'Author', and 'Date', and a 'Comment' section with text and dates.

RDF-izers

- Triplify

- <http://triplify.org>

- Transform relational data into RDF / Linked Data

- D2RQ platform

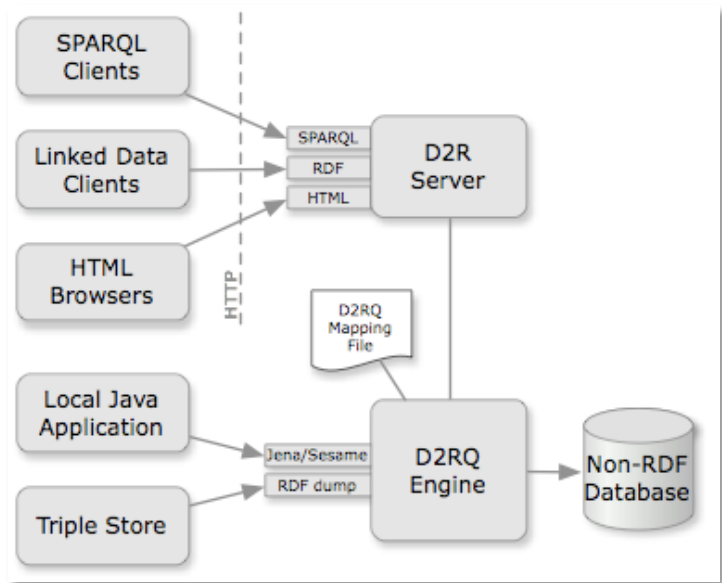
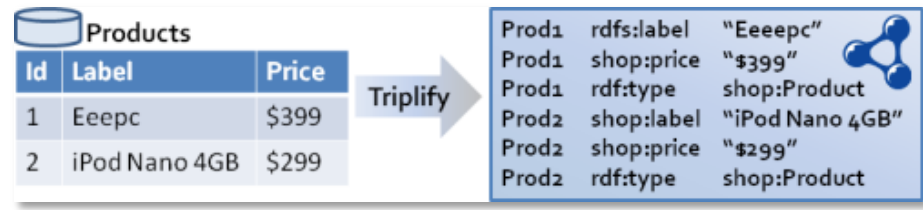
- <http://www4.wiwiss.fu-berlin.de/bizer/d2rq/index.htm>

- D2RQ mapping language

- D2RQ plugin for Sesame/Jena

- D2R server

- Linked Data & SPARQL endpoint



RDF APIs

- Jena
 - <http://jena.sourceforge.net>
 - RDF/OWL API (Java)
 - In-memory or persistent storage
 - SPARQL query engine
- OpenRDF (Sesame)
 - <http://www.openrdf.org>
 - RDF API (Java), high performance parser
 - Persistent storage
 - SPARQL query engine

Triplestores

- OWLIM
 - <http://www.ontotext.com/owlim/>
 - Unmatched materialisation & query performance, replication cluster, RDFRank, “sameAs” optimisation
 - 10-20 billion triples
- Virtuoso
 - <http://virtuoso.openlinksw.com>
 - RDBMS integration, geo-spatial extensions, federated / virtual databases
 - 10-20 billion triples

Triplestores (2)

- Oracle
 - RDBMS integration, parallel inference/query, built-in compression
- AllegroGraph RDFStore
 - <http://www.franz.com/agraph/allegrograph>
 - Geo-spatial extensions, built-in SNA functionality, federated / virtual databases, built-in compression
- 4Store
 - <http://4store.org>
 - distributed cluster

Reasoners

- Pellet
 - <http://clarkparsia.com/pellet>
 - OWL DL, OWL2 Profiles, datatypes, SWRL extensions, ontology analysis & repair, incremental reasoning, integration with Oracle DB
- Fact++
 - <http://code.google.com/p/factplusplus>
 - OWL DL, partial OWL2 and datatype support
- Racer
 - <http://www.racer-systems.com>

Linked Data browsers – Marbles

- <http://marbles.sourceforge.net>
- XHTML views of RDF data (SPARQL endpoint), caching, predicate traversal


http://dbpedia.org/resource/Montreal Open marbles

<http://dbpedia.org/resource/Montreal>

seeAlso

- http://dbpedia.org/data/1924_International_Law_Tennis_Challenge
- http://dbpedia.org/data/%C3%89cole_de_technologie_sup%C3%A9rieure
- http://dbpedia.org/data/1928_International_Law_Tennis_Challenge
- http://dbpedia.org/data/10.5_Apocalypse
- http://dbpedia.org/data/1931_International_Law_Tennis_Challenge
- http://dbpedia.org/data/%C3%89cole_Polytechnique_de_Montr%C3%A9al
- http://dbpedia.org/data/1927_International_Law_Tennis_Challenge
- http://dbpedia.org/data/1923_International_Law_Tennis_Challenge
- http://dbpedia.org/data/1929_International_Law_Tennis_Challenge
- <http://dbpedia.org/data/%C3%89quiterre>

depiction



is <http://sindice.com/vocab/search#link> of

- Montreal, 蒙特利尔, Montréal, Ville de Montréal, Moupeans, モントリオール, Montréal (Canada), City of Montreal

is primary topic of

- [Photos for DBpedia.org resource Montreal](#)

Sources

- http://dbpedia.org/data/1929_International_Law_Tennis_Challenge failed, retrieved Tue, 17 Aug 2010 14:29:19 GMT (clear)
- <http://dbpedia.org/resource/Montreal> failed, retrieved Tue, 17 Aug 2010 14:29:16 GMT (clear)
- http://www4.wiwis.fu-berlin.de/flickrwrappr/photos/Montreal%2C_Quebec success (200), retrieved Tue, 17 Aug 2010 14:29:22 GMT (clear)
- http://www4.wiwis.fu-berlin.de/flickrwrappr/photos/Montreal_Canadiens success (200), retrieved Tue, 17 Aug 2010 14:29:52 GMT (clear)
- http://dbpedia.org/data/10.5_Apocalypse failed, retrieved Tue, 17 Aug 2010 14:29:18 GMT (clear)
- http://dbpedia.org/data/1928_International_Law_Tennis_Challenge failed, retrieved Tue, 17 Aug 2010 14:29:17 GMT (clear)

Linked Data browsers – ReFinder

- <http://refinder.dbpedia.org>
- Explore & navigate relationships in a RDF graph

The screenshot shows the ReFinder interface on the left and a large RDF graph on the right. The interface includes a search area with two input fields containing 'Concordia University' and 'McGill University', and buttons for 'add', 'clear', and 'Find Relations'. Below this is a filter section with a table showing the number of objects for different filters.

number of objects	num	vi
1	5/5	
2	12/12	

The RDF graph on the right shows a network of entities and their relationships. The central node is 'Montreal', highlighted in green. Other nodes include 'Concordia University', 'McGill University', 'Sir George Williams University', 'Howard Alper', 'Urban area', 'Public university', 'Macdonald Campus', 'Concordia Stingers', 'Loyola College', 'Canada', 'Quebec', and 'Association of American Universities'. Relationships are labeled with terms like 'affiliation', 'campus', 'type', 'city', 'country', 'state', 'location', 'redirect', and 'almaMater'. Red arrows highlight the path from the search results to the Montreal node.

Linked Data browsers – OpenLink RDF Browser

- <http://demo.openlinksw.com/DAV/JS/rdfbrowser/index.html>
- Explore & navigate relationships in a RDF graph

OpenLink RDF Browser

Data Source (URL):






<http://dbpedia.org/resource/Montreal> - 1936 triples - [Remove from storage](#) - [permalink](#)

TOTAL: 1936 triples - [permalink](#)

Filters

No filters are selected. Create some by clicking on values in Categories you want to view.

This module is used to navigate through locally cached data, one resource at a time. Note that filters are not applied here.

Ville de Montréal

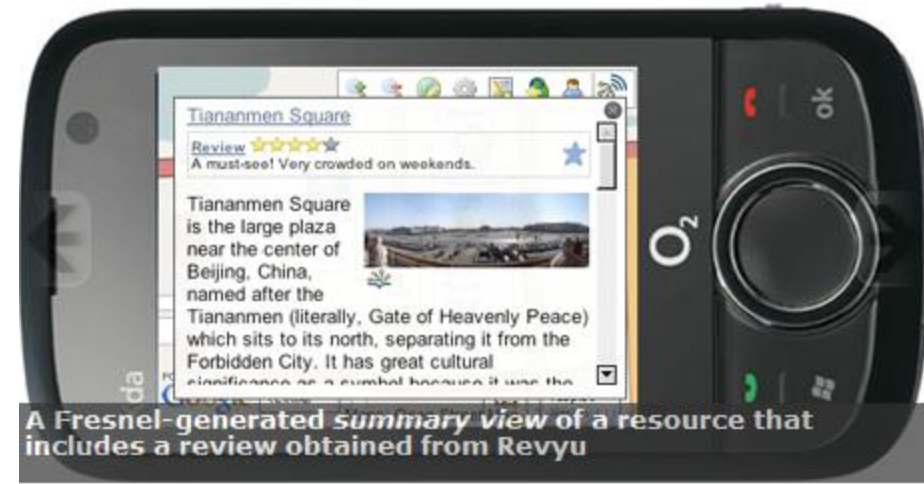
- ▼ **type**
 - [City](#)
 - [PortCitiesInCanada](#)
 - [HostCitiesOfTheSummerOlympicGames](#)
 - [Mx4rwPflcJwpEbGdrcN5Y29ycA](#)
 - [Settlement](#)
 - 8 more...
- ▼ **name**
 - [Ville de Montréal](#)
 - [City of Montreal](#)
- ▼ **sameAs**
 - [N59179828586486930801](#)
 - [id](#)
 - [CityOfMontrealCanada](#)
 - [CityOfMontrealCanada](#)
 - [Mx4rvViBU5wpEbGdrcN5Y29ycA](#)
 - 44 more...
- ▼ **maximumElevation**
 - [888](#)

DBpedia Mobile

- <http://wiki.dbpedia.org/DBpediaMobile>
- Based on user's GPS position, renders a map with nearby places of interest (from DBpedia)



DBpedia Mobile's *map view* of resources in the user's proximity



A Fresnel-generated *summary view* of a resource that includes a review obtained from Revyu

FactForge and LinkedLifeData

- *FactForge*
 - Integrates some of the most central LOD datasets
 - General-purpose information (not specific to a domain)
 - 1.2B explicit plus 1B inferred statements (10B retrievable)
 - The largest upper-level knowledge base
 - <http://www.FactForge.net/>
- *Linked Life Data*
 - 25 of the most popular life-science datasets
 - 2.7B explicit and 1.4B inferred triples
 - <http://www.LinkedListLifeData.com>

FactForge and LinkedLifeData (2)



a semantic data integration platform for the biomedical domain

[SPARQL Query](#)

Search and explore over 4 billion RDF statements from various sources including [UniProt](#), [PubMed](#), [EntrezGene](#) and [20 more...](#)
Perform complex SPARQL queries and retrieve more than 500 million RDF resources.

Disclaimer: Part of the information in the Linked Life Data knowledge base is from copyrighted data. Linked Life Data is a prototype demonstration service and its users are solely responsible for compliance with any copyright.

Linked Life Data is partly funded by the EU IST project [LarkC \(FP7-215535\)](#).
© 2009-2010 Ontotext AD. All rights reserved.



[RDF Search and Explore](#) | [SPARQL Query](#) | [Relfinder](#) | [About](#) | [Contact](#)

FactForge represents a [reason-able view](#) to the [web of data](#). It aims to allow users to find resources and facts based on the semantics of the data, like web search engines index WWW pages and facilitate their usage.

RDF Search and Explore

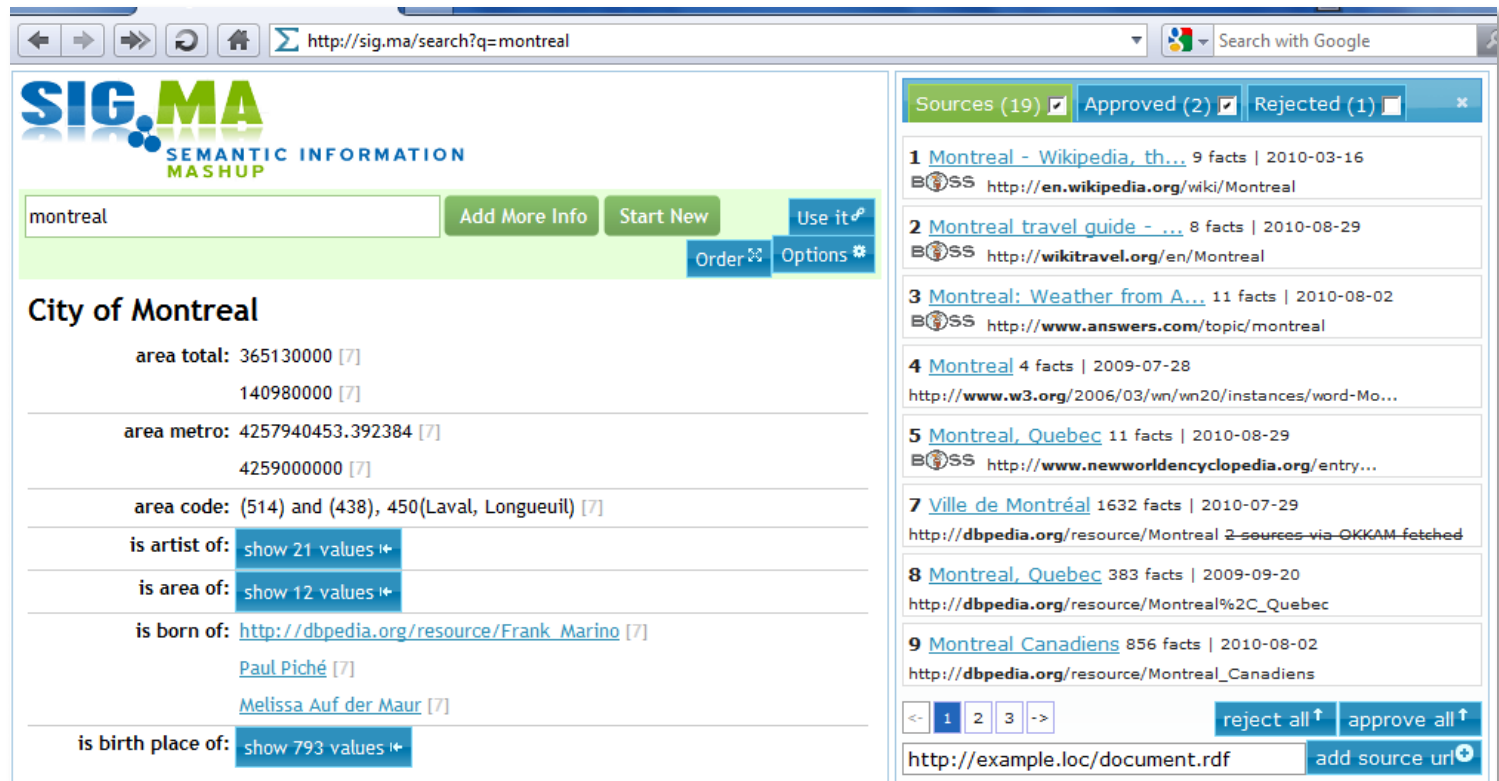
Keyword search retrieves a ranked list of RDF molecules. The automatic suggestions allow direct exploration of URIs.
Structured queries can be defined using a [SPARQL Form](#).

Repository overview

Engine: Big-OWLIM-3.3	Information: http://ontotext.com/factforge
Inference ruleset: factforge.pie (more)	Number of statements: 2,237,550,383 (Dataset statistics)
Number of expl. statements: 1,357,013,225	Number of retr. statements: 9,818,667,408
Number of entities: 404,796,665	Number of URI: 145,218,491
Number of Literals: 259,578,031	Number of Bnodes: 143

Sig.ma – the semantic mashup platform

- Aggregate RDF/RDFa data sources for „live views“ of the Web of Data
 - <http://sig.ma>



The screenshot shows the Sig.ma Semantic Information Mashup interface. The search bar contains 'montreal'. The main content area displays information for the 'City of Montreal', including area total, area metro, area code, and birth place. The right sidebar shows a list of sources used for the mashup, such as Wikipedia, Wikitravel, and Answers.com. The interface includes navigation buttons like 'Add More Info', 'Start New', 'Use it', 'Order', and 'Options'.

SIG.MA SEMANTIC INFORMATION MASHUP

montreal

City of Montreal

area total: 365130000 [7]
140980000 [7]

area metro: 4257940453.392384 [7]
4259000000 [7]

area code: (514) and (438), 450(Laval, Longueuil) [7]

is artist of:

is area of:

is born of: http://dbpedia.org/resource/Frank_Marino [7]
[Paul Piché](#) [7]
[Melissa Auf der Maur](#) [7]

is birth place of:

Sources (19) **Approved (2)** **Rejected (1)**

- 1 Montreal - Wikipedia, th...** 9 facts | 2010-03-16
BIBSS <http://en.wikipedia.org/wiki/Montreal>
- 2 Montreal travel guide - ...** 8 facts | 2010-08-29
BIBSS <http://wikitravel.org/en/Montreal>
- 3 Montreal: Weather from A...** 11 facts | 2010-08-02
BIBSS <http://www.answers.com/topic/montreal>
- 4 Montreal** 4 facts | 2009-07-28
<http://www.w3.org/2006/03/vn/vn20/instances/word-Mo...>
- 5 Montreal, Quebec** 11 facts | 2010-08-29
BIBSS <http://www.newworldencyclopedia.org/entry...>
- 7 Ville de Montréal** 1632 facts | 2010-07-29
http://dbpedia.org/resource/Montreal_2-sources-via-OKKAM-fetched
- 8 Montreal, Quebec** 383 facts | 2009-09-20
http://dbpedia.org/resource/Montreal%2C_Quebec
- 9 Montreal Canadiens** 856 facts | 2010-08-02
http://dbpedia.org/resource/Montreal_Canadiens

< 1 2 3 >

<http://example.loc/document.rdf>

Sindice – the Semantic Web index

- Crawls RDF/RDFa pages on the web and provides a consolidated index
 - <http://sindice.com>
 - 120+ million indexed pages



Summary of this module

- The Semantic Web is an extension of the current web, where information has well-defined meaning, so that it is usable by intelligent agents
- Ontologies provide means to share and reuse knowledge, by formal modelling of the concepts and relationships between them in a domain
- Since 1999 W3C has provided a rich family of Semantic Web related standards for modelling knowledge and rules, querying knowledge bases, embedding semantic annotations in web pages, etc

Summary of this module (2)

- Semantic Technologies are applicable for a wide set of domains where interoperability, data integration and knowledge reuse is crucial
- There is a rich tool support for Semantic Technologies

Useful links

- RDF
 - Resource Description Framework (RDF): Concepts and Abstract Syntax – <http://www.w3.org/TR/rdf-concepts/>
- Linked Data
 - <http://linkeddata.org/>