

What is the Semantic Web?

Sir Tim Berners-Lee's vision of the Web as a universal medium for data, information, and knowledge exchange.

An evolving extension of the World Wide Web in which the semantics of information and services on the web is defined, making it possible for the web to understand and satisfy the requests of people and machines to use the web content.

...prospective future possibilities that are yet to be implemented or realized.

A set of design principles, collaborative working groups, and a variety of enabling technologies.

What is the Semantic Web? (continued)

“... supposed to make data located anywhere on the Web accessible and understandable, both to people and to machines.”

(Explorers Guide to the Semantic Web, p 3)

“... more a vision than a technology.”

(Explorers Guide to the Semantic Web, p 3)

“...a fluid, evolving, informally defined concept rather than an integrated, working system.”

(Explorers Guide to the Semantic Web, p 3)

Multiple Views

- The Machine Readable Data View
- The Intelligent Agents View
- The Distributed Database View
- The Automated Infrastructure View
- The Servant of Humanity View
- The Better Annotation View
- The Improved Searching View
- The Web Services View

The Machine Readable Data View

“The Semantic Web is a vision: the idea of having data on the Web defined and linked in a way that it can be used by machines not just for display purposes but for automation, integration and re-use of data across various applications.”

(Explorers Guide to the Semantic Web, p 3)

The Intelligent Agents View

“The aim of the Semantic Web is to make the present Web more machine-readable, in order to allow intelligent agents to retrieve and manipulate pertinent information.”

(Explorers Guide to the Semantic Web, p 3)

The Distributed Database View

“The Semantic Web concept is to do for data what HTML did for textual information systems: to provide sufficient flexibility to be able to represent all databases and logic rules to link them together to great added value.”

(Explorers Guide to the Semantic Web, p 4)

“A simple description of the Semantic Web is that it is an attempt to do for machine processable data what the World Wide Web did for human readable documents. Namely to transform information processing by providing a common way that data can be accessed, linked together and understood. To turn the Web from a large hyperlinked book into a large interlinked database.”

(Explorers Guide to the Semantic Web, p 4)

The Automated Infrastructure View

“In his recent Scientific American article Berners-Lee argues that the Semantic Web is infrastructure and not an application.”

(Explorers Guide to the Semantic Web, p 4)

The Servant of Humanity View

“The vision of the Semantic Web is to let computer software relieve us of much of the burden of locating resources on the Web that are relevant to our needs and extracting, integrating and indexing the information contained within.”

(Explorers Guide to the Semantic Web, p 4)

“The Semantic Web is a vision of the next generation Web, which enables web applications to automatically collect web documents, integrate and process information and interoperate with other applications in order to execute sophisticated tasks for humans.”

(Explorers Guide to the Semantic Web, p 4)

The Better Annotation View

“The idea of a 'Semantic Web' supplies the (informal) web as we know it with annotations expressed in a machine readable form and linked together.”

(Explorers Guide to the Semantic Web, p 4)

The Improved Searching View

“Soon it will be possible to access Web resources by content rather than just by keywords.”

(Explorers Guide to the Semantic Web, p 4)

The Web Services View

“Increasingly, the Semantic Web will be called upon to provide access not just to static documents that collect useful information, but also to services that provide useful behavior.”

(Explorers Guide to the Semantic Web, p 4)

“The Semantic Web promises to expand the services for the existing web by enabling software agents to automate procedures currently performed manually and by introducing new applications that are infeasible today.”

(Explorers Guide to the Semantic Web, p 4)

Enhanced Web Search

- More precise searching, using metadata fields. Search for content published before or after a certain date, content written by a certain author, etc.
- Search based on assertions as validity, trustworthiness, etc.
- Caveat: trust and rating mechanisms for the SW are a whole separate ball of wax.
- Search for content by license; let's say I want pages that are licensed under the GFDL or a certain Creative Commons license.
- Platform for Internet Content Selection – PICS <http://www.w3.org/PICS/>
- Content Labels – successor to PICS
<http://www.w3.org/2005/Incubator/wcl/XGR-wcl-20070220/>

Assisting With Retail Purchases

- The current web is a valuable tool if you want to do something like, say, buy flowers.
- But...
- A lot of manual steps are still involved. Google for “Flower Shops,” Raleigh. Visit 2 or 3 or 4 of the websites you find. Check the location, hours, prices and availability for each, possibly requiring one or more phone calls as well.
- Why can't you, for example, ask your computer to find you a flower shop, within 10 miles of your office, that's open until 7:00pm, which stocks the exact roses you want, with an acceptable price?

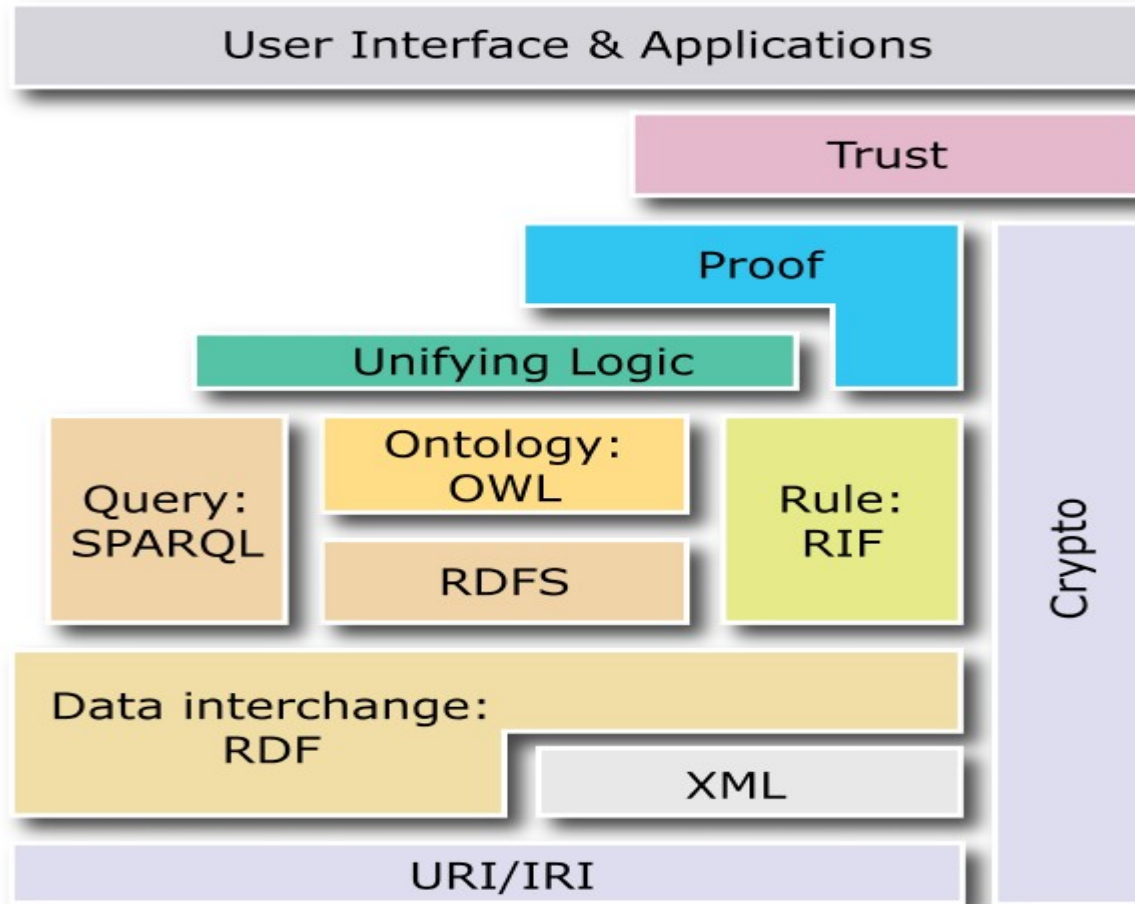
Hiring & Recruitment

- In 2004 almost 28 % of 14-64 years old people used the web to research the employment market and look for open job positions.
- No way (much less easy way) for the job seeker to gain an overview of all the offers published on the internet.
- More and more companies post their vacancies only on their company websites.
- Most of these sites cannot be found by the popular search engines due to the lack of “meaning” (semantics), i.e. they are not recognised as “job postings.”
- If postings are published using a known ontology using RDF, RDFS and OWL, we can picture a “personal agent” that could crawl the websites of companies you are particularly interested in, alerting you to potential matches.
- Or a portal site could try to crawl the entire “job web” and then provide a front-end UI for searching for available jobs.
- Better semantic matching could mean less time digging through results you are not really interested in.

Enabling Technologies

- XML, XML-Schema
- RDF (RDF/XML, N3, TURTLE), RDFS, RDFa
- OWL (OWL-Lite, OWL-DS, OWL-FULL)
- SPARQL
- RULEML
- GRDDL
- FOAF, Dublin Core, SKOS, SIOC, DOAC, DOAP, etc.

The “Semantic Web Layer Cake”



RDF – Resource Description Framework

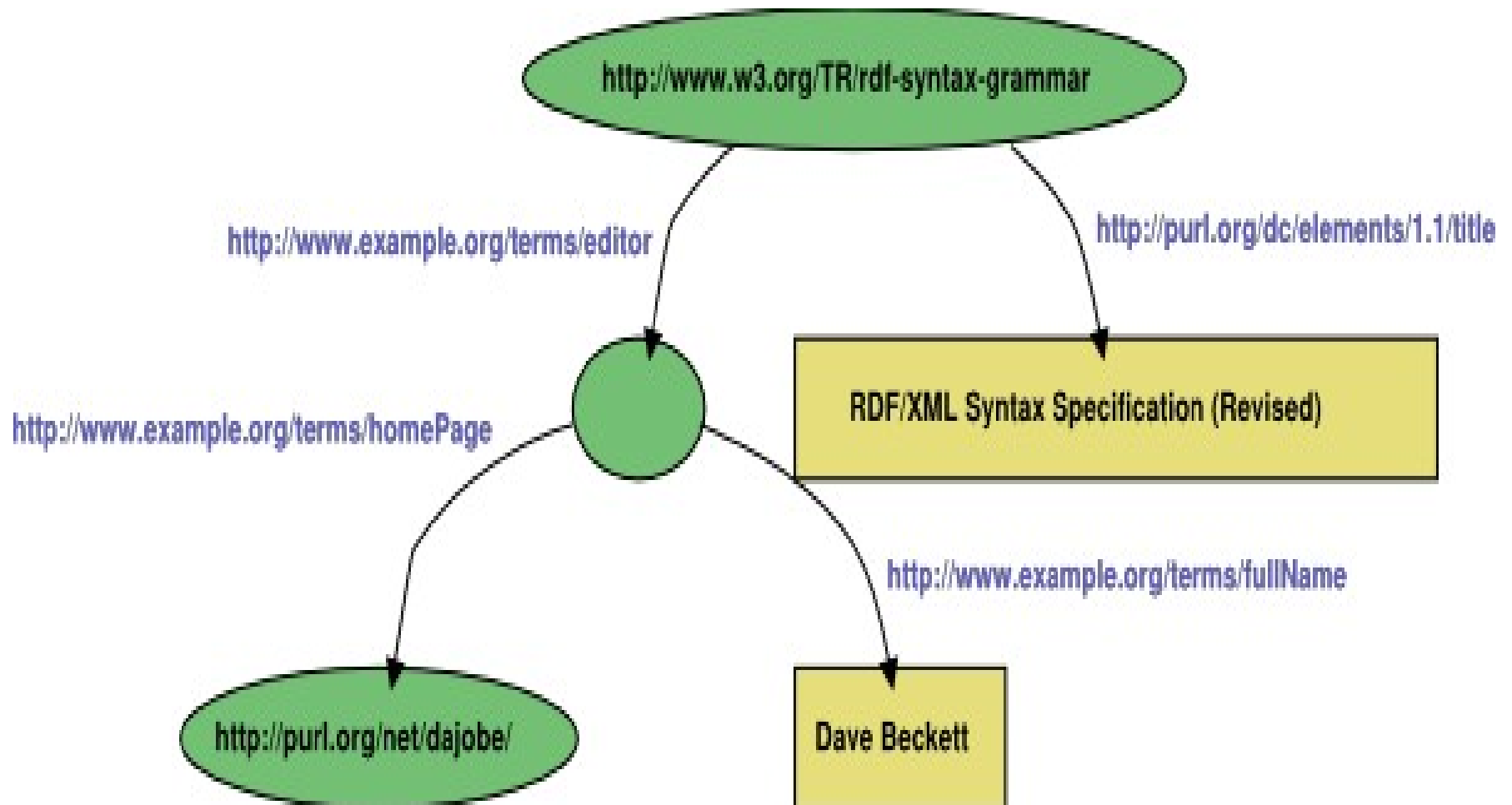
Models statements as “triples”



Subject, Predicate, Object

- Resources unambiguously named using URIs
- Everything is a triple... ex: “the shoe is red” would be the triple with subject = “shoe”, predicate (or property) = “color”, and object (or value = “red”
- Standard serialization format is XML, known as RDF/XML.
- Developer friendly serialization formats include N3 and Turtle

RDF Example



Plain old RDF/XML

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:ex="http://example.org/stuff/1.0/">
<rdf:Description rdf:about="http://www.w3.org/TR/rdf-syntax-grammar">
  <ex:editor>
    <rdf:Description>
      <ex:homePage>
        <rdf:Description rdf:about="http://purl.org/net/dajobe/">
          </rdf:Description>
        </ex:homePage>
      </rdf:Description>
    </ex:editor>
  </rdf:Description>

<rdf:Description rdf:about="http://www.w3.org/TR/rdf-syntax-grammar">
  <ex:editor>
    <rdf:Description>
      <ex:fullName>Dave Beckett</ex:fullName>
    </rdf:Description>
  </ex:editor>
</rdf:Description>

<rdf:Description rdf:about="http://www.w3.org/TR/rdf-syntax-grammar">
  <dc:title>RDF/XML Syntax Specification (Revised)</dc:title>
</rdf:Description>
</rdf:RDF>
```

Simplified RDF

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:ex="http://example.org/stuff/1.0/">
  <rdf:Description rdf:about="http://www.w3.org/TR/rdf-syntax-grammar">
    <ex:editor>
      <rdf:Description>
        <ex:homePage>
          <rdf:Description rdf:about="http://purl.org/net/dajobe/">
            </rdf:Description>
          </ex:homePage>
          <ex:fullName>Dave Beckett</ex:fullName>
        </rdf:Description>
      </ex:editor>
      <dc:title>RDF/XML Syntax Specification (Revised)</dc:title>
    </rdf:Description>
  </rdf:RDF>
```

The same example, using Turtle

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix ex: <http://example.org/stuff/1.0/> .

<http://www.w3.org/TR/rdf-syntax-grammar>
  dc:title "RDF/XML Syntax Specification (Revised)" ;
  ex:editor [
    ex:fullname "Dave Beckett";
    ex:homePage <http://purl.org/net/dajobe/>
  ] .
```

And again, in Notation 3 (N3)

```
@prefix ex: <http://example.org/stuff/1.0/> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
_:bnode0 ex:homePage <http://purl.org/net/dajobe/> ;
          ex:fullName "Dave Beckett" .
<http://www.w3.org/TR/rdf-syntax-grammar> ex:editor _:bnode0 ;
          dc:title "RDF/XML Syntax Specification (Revised)" .
```

And finally, in N-Triples

```
_:bnode0 <http://example.org/stuff/1.0/homePage> <http://purl.org/net/dajobe/> .  
<http://www.w3.org/TR/rdf-syntax-grammar> <http://example.org/stuff/1.0/editor> _:bnode0 .  
_:bnode1 <http://example.org/stuff/1.0/fullName> "Dave Beckett" .  
<http://www.w3.org/TR/rdf-syntax-grammar> <http://example.org/stuff/1.0/editor> _:bnode1 .  
<http://www.w3.org/TR/rdf-syntax-grammar> <http://purl.org/dc/elements/1.1/title> "RDF/XML  
Syntax Specification (Revised)" .
```

RDF/S – RDF Schema

- Specification: <http://www.w3.org/TR/rdf-schema/>
- For a good intro to RDF/S see the RDF-Schema chapter of the RDF Primer:
<http://www.w3.org/TR/2004/REC-rdf-primer-20040210/#rdfschema>
- Provides a type system for RDF
- Define classes, properties, relationships between classes, etc.

RDFS Example Schema

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [<!ENTITY xsd "http://www.w3.org/
2001/XMLSchema#">]>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-
ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xml:base="http://example.org/schemas/vehicles">

<rdfs:Class rdf:ID="MotorVehicle"/>

<rdfs:Class rdf:ID="PassengerVehicle">
  <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
</rdfs:Class>

<rdfs:Class rdf:ID="Truck">
  <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
</rdfs:Class>

<rdfs:Class rdf:ID="Van">
  <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
</rdfs:Class>

<rdfs:Class rdf:ID="MiniVan">
  <rdfs:subClassOf rdf:resource="#Van"/>
  <rdfs:subClassOf rdf:resource="#PassengerVehicle"/>
</rdfs:Class>
```

```
<rdfs:Class rdf:ID="Person"/>

<rdfs:Datatype rdf:about="&xsd;integer"/>

<rdf:Property rdf:ID="registeredTo">
  <rdfs:domain rdf:resource="#MotorVehicle"/>
  <rdfs:range rdf:resource="#Person"/>
</rdf:Property>

<rdf:Property rdf:ID="rearSeatLegRoom">
  <rdfs:domain
rdf:resource="#PassengerVehicle"/>
  <rdfs:range rdf:resource="&xsd;integer"/>
</rdf:Property>

<rdf:Property rdf:ID="driver">
  <rdfs:domain rdf:resource="#MotorVehicle"/>
</rdf:Property>

<rdf:Property rdf:ID="primaryDriver">
  <rdfs:subPropertyOf rdf:resource="#driver"/>
</rdf:Property>

</rdf:RDF>
```

Using the RDFS Schema

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [<!ENTITY xsd
"http://www.w3.org/2001/XMLSchema#">]>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:ex="http://example.org/schemas/vehicles#"
        xml:base="http://example.org/things">

  <ex:PassengerVehicle rdf:ID="johnSmithsCar">
    <ex:registeredTo rdf:resource="http://www.example.org/staffid/85740"/>
    <ex:rearSeatLegRoom
      rdf:datatype="&xsd;integer">127</ex:rearSeatLegRoom>
    <ex:primaryDriver rdf:resource="http://www.example.org/staffid/85740"/>
  </ex:PassengerVehicle>
</rdf:RDF>
```

OWL – Web Ontology Language

- Originates from research into exploring how the idea of knowledge representation (KR) from AI could be made useful on the World Wide Web.
- a revision of the DAML+OIL web ontology language.
- 3 different versions of OWL, that make various tradeoffs between expressiveness and decidability.
- OWL DL and OWL Lite semantics are based on Description Logics, which have attractive and well-understood computational properties.
- Description Logics are a family of logics that are decidable fragments of first-order logic.
- OWL Full uses a novel semantic model intended to provide compatibility with RDF Schema.
- OWL Lite was originally intended to support those users primarily needing a classification hierarchy and simple constraints. For example, while it supports cardinality constraints, it only permits cardinality values of 0 or 1.
- OWL DL was designed to provide the maximum expressiveness possible while retaining computational completeness (all conclusions are guaranteed to be computed), decidability (all computations will finish in finite time), and the availability of practical reasoning algorithms.
- OWL Full is based on a different semantics from OWL Lite or OWL DL, and was designed to preserve some compatibility with RDF Schema. For example, in OWL Full a class can be treated simultaneously as a collection of individuals and as an individual in its own right; this is not permitted in OWL DL. OWL Full allows an ontology to augment the meaning of the pre-defined (RDF or OWL) vocabulary. It is unlikely that any reasoning software will be able to support complete reasoning for OWL Full.

OWL Ontologies

The data described by an OWL ontology is interpreted as a set of "individuals" and a set of "property assertions" which relate these individuals to each other. An OWL ontology consists of a set of axioms which place constraints on sets of individuals (called "classes") and the types of relationships permitted between them. These axioms provide semantics by allowing systems to infer additional information based on the data explicitly provided.

For example, an ontology describing families might include axioms stating that a "hasMother" property is only present between two individuals when "hasParent" is also present, and individuals of class "HasTypeOBlood" are never related via "hasParent" to members of the "HasTypeABBlood" class. If it is stated that the individual Harriet is related via "hasMother" to the individual Sue, and that Harriet is a member of the "HasTypeOBlood" class, then it can be inferred that Sue is not a member of "HasTypeABBlood".

For existing schemas and ontologies, see:

<http://www.schemaweb.info/>

and

http://protegewiki.stanford.edu/index.php/Protege_Ontology_Library

OWL Sample

```
<owl:Class rdf:ID="Interval">
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >An interval is a composition of multiple periods of time.</rdfs:comment>
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Temporal_Occurrence"/>
  </rdfs:subClassOf>
  <owl:disjointWith>
    <owl:Class rdf:ID="Moment"/>
  </owl:disjointWith>
  <rdfs:subClassOf rdf:resource="http://www.estrellaproject.org/lkif-core/mereology.owl#Composition"/>
</owl:Class>
<owl:Class rdf:about="#Temporal_Occurrence">
  <rdfs:subClassOf rdf:resource="http://www.estrellaproject.org/lkif-core/lkif-top.owl#Spatio_Temporal_Occurrence"/>
  <rdfs:subClassOf>
    <owl:Class>
      <owl:intersectionOf rdf:parseType="Collection">
        <owl:Restriction>
          <owl:someValuesFrom rdf:resource="#Temporal_Occurrence"/>
          <owl:onProperty>
            <owl:ObjectProperty rdf:ID="immediately_after"/>
          </owl:onProperty>
        </owl:Restriction>
        <owl:Restriction>
          <owl:someValuesFrom rdf:resource="#Temporal_Occurrence"/>
          <owl:onProperty>
            <owl:ObjectProperty rdf:ID="immediately_before"/>
          </owl:onProperty>
        </owl:Restriction>
      </owl:intersectionOf>
    </owl:Class>
  </rdfs:subClassOf>
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >A period of time, has a duration</rdfs:comment>
</owl:Class>
```

Gödel's incompleteness theorems & decidability

- http://en.wikipedia.org/wiki/Incompleteness_theorem
- “Any effectively generated theory capable of expressing elementary arithmetic cannot be both consistent and complete. In particular, for any consistent, effectively generated formal theory that proves certain basic arithmetic truths, there is an arithmetical statement that is true, but not provable in the theory.”
- “Gödel's first incompleteness theorem shows that any formal system that includes enough of the theory of the natural numbers is incomplete: it contains statements that are neither provably true nor provably false. Or one might say, no formal system which aims to define the natural numbers can actually do so, as there will be true number-theoretical statements which that system cannot prove.”
- [http://en.wikipedia.org/wiki/Independence_\(mathematical_logic\)](http://en.wikipedia.org/wiki/Independence_(mathematical_logic))
- http://en.wikipedia.org/wiki/Zermelo-Fraenkel_set_theory
- http://en.wikipedia.org/wiki/Axiomatic_set_theory
- http://en.wikipedia.org/wiki/First-order_logic

SPARQL

- is an RDF query language
- **Simple Protocol and RDF Query Language**
- Supported by many popular semantic reasoners and semantic web toolkits
- Including Jena, Pellet and Sesame
- http://en.wikipedia.org/wiki/Semantic_reasoner
- <http://pellet.owldl.com/>
- <http://jena.sourceforge.net/>
- <http://sparql.sourceforge.net/>
- <http://www.openrdf.org/about.jsp>

SPARQL Sample

```
PREFIX abc: <http://example.com/exampleOntology#>
SELECT ?capital ?country
WHERE {
  ?x abc:cityname ?capital ;
     abc:isCapitalOf ?y .
  ?y abc:countryname ?country ;
     abc:isInContinent abc:Africa .
}
```

Will return all country capitals in Africa

Collaborative Working Groups

- Mostly part of the W3C “Semantic Web Activity”
- Most work done more or less openly, via mailing lists, IRC, etc.
- Membership doesn't require you be a W3C member, see “Invited Expert” status.
- Currently Active WG's
 - Rules Interchange Format WG:
http://www.w3.org/2005/rules/wiki/RIF_Working_Group
 - OWL WG:
http://www.w3.org/2007/OWL/wiki/OWL_Working_Group
- Past WG's (work completed)
 - Best Practices & Deployment WG:
<http://www.w3.org/2001/sw/BestPractices/>
 - RDF Core WG:
<http://www.w3.org/2001/sw/RDFCore/>
 - Web Ontology WG:
<http://www.w3.org/2001/sw/WebOnt/>

Non W3C Working Groups / Projects

- FOAF – Friend of a Friend
<http://www.foaf-project.org/>
- SIOC – Semantically Interlinked Online Communities
<http://sioc-project.org/>
- DOAP – Description of a Project
<http://trac.usefulinc.com/doap>
- Dublin Core
<http://dublincore.org/>

Future Possibilities

- FOAF: Data portability / social networking
- SIOC: Distributed conversations
- Enhanced Web Search
- Negotiating Purchases
- Personal assistant helps with hiring / recruitment
- For more use-cases & scenarios, see:
<http://www.w3.org/2001/sw/sweo/public/UseCases/>

FOAF – Friend of a Friend

- Federation of “Social Graph” information across social-networks and other repositories of data.
- Search my network to level N for people who meet criteria X
- Combined with other SW data sources to create more interesting queries: search my graph for people who live in places where the average temperature is < 50 degrees?
- <http://weather.noaa.gov/weather/current/KRDU.html>
- <http://www.weather.gov/xml/>
- Combine with DOAC (Description of a Career) data for something like “Search my network for people I've worked with who have at least 7 years of Java experience,” etc.
- <http://ramonantonio.net/doac/>
- Note that none of this is anything new, but the Semantic Web makes it easier for computers to help us do things that would otherwise be tedious, manual operations.

SIOC – Semantically Interlinked Online Communities

- Enable “distributed conversations” by linking threads across blogs, forums, Usenet groups, etc.
- Use as an enhanced export/import format, with access to either the entire content or summaries
- Enable publishing and subscribing to decentralised discussion channels and communities
- Plays nicely with FOAF
- Possible use case: subscribe to a “virtual” feed of posts made to selected blogs, by selected members of my personal “social graph.”
- <http://sioc-project.org/>