

A Hands-On Overview of the Semantic Web



Shamod Lacoul
Mike Hewett

League of Extraordinary Modelers

Outline

- What is the Semantic Web?
- Why use the Semantic Web?
- RDF for data representation
- RDFS - data schemas (models)
- SPARQL - RDF query language
- RDFa
- Linked Open Data
- OWL - Web Ontology Language
- SWRL - Semantic Web Rule Language
- Useful links

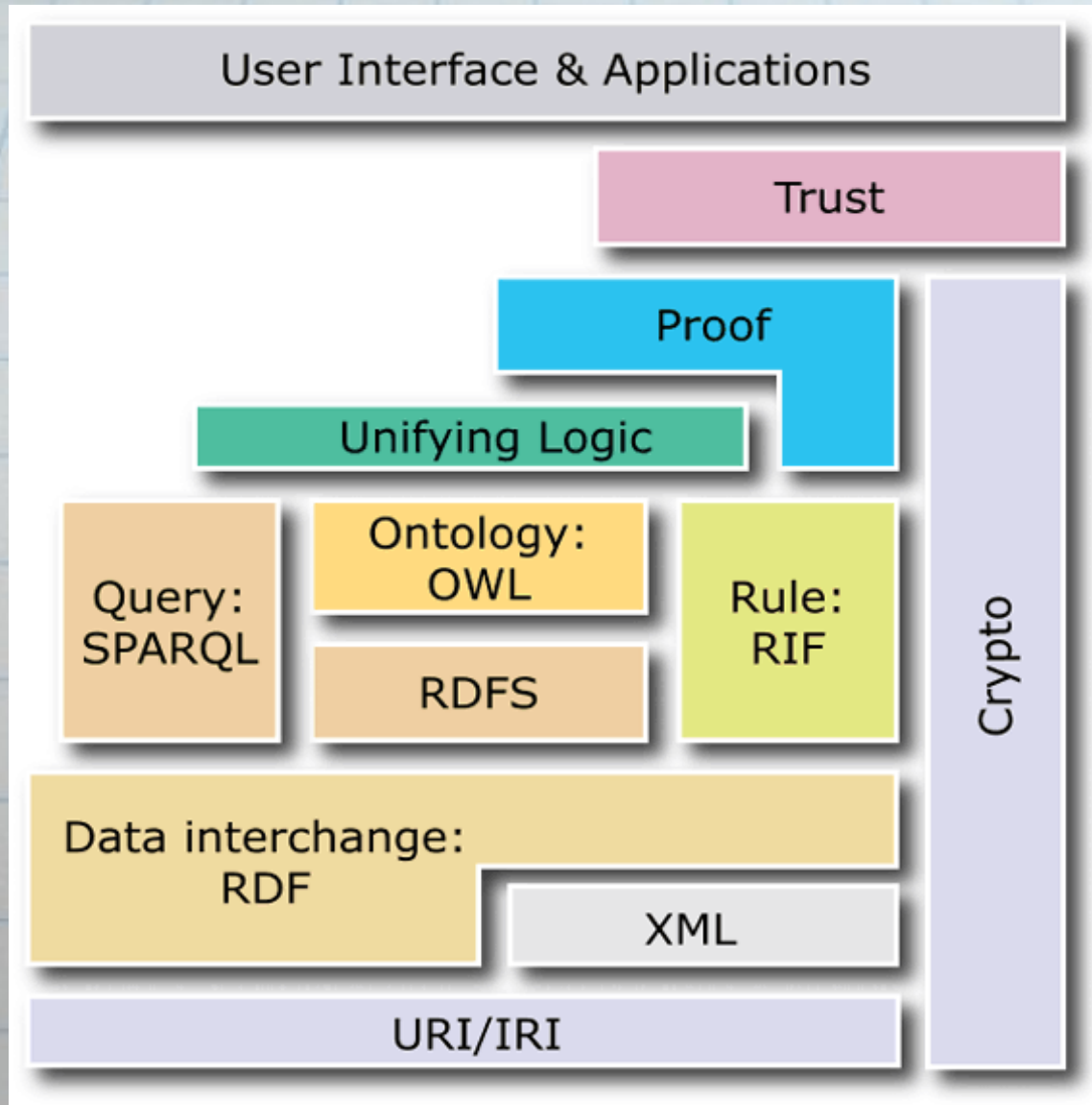
... interspersed with demos

See <http://www.hewettresearch.com/svcc2009/>
for related materials.

Who is this person?



Who has seen this diagram before?



Web vs. Semantic Web

Web --> links documents to documents

Semantic Web --> links data to data.

aka. Data Web (Web of Data), Structured Web or LINKED DATA.

Semantic Web Advantages

Universal data representation (using RDF)

Reusable data models (using RDF, RDFS, and OWL)

W3C Standard query language (SPARQL)

Information validation and classification (reasoners)

Rule-based inferencing (SWRL)

Use Cases for Semantic Data

Business Data Exchange

- RDF is a universal data format
- RDF data can be mapped to and from relational, XML and object models
- Even better, execute SPARQL queries remotely to retrieve just the subset of data you need

Use Cases for Semantic Data

Biomedical modeling and processing

- Hundreds of OWL biomedical knowledge bases available
- Reuse knowledge in different applications

Other domains:

Business

Engineering

Scientific

E-Commerce

...

Uses of XML in RDF

- RDF/XML is one type of RDF serialization
- XSD datatypes
- Namespaces



Vocabulary: Namespace

Namespace - a concept borrowed from XML

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

prefix -> **rdf:**

namespace URI -> **http://www.w3.org/1999/02/22-rdf-syntax-ns#**

<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>

is the same as

rdf:type

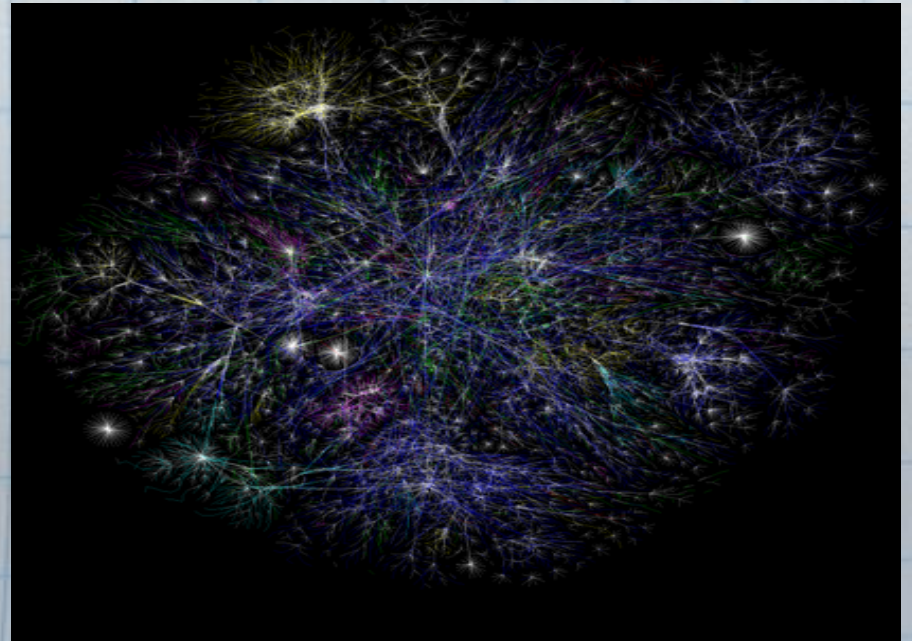
Vocabulary: URI

Universal Resource Identifier - a unique ID

<http://www.siliconvalley-codecamp.com/Sessions#SemWeb>

Namespace - <http://www.siliconvalley-codecamp.com/Sessions#>

Fragment Identifier - [SemWeb](#)



Any questions so far?

See <http://www.hewettresearch.com/svcc2009/>
for related materials.

Coming next:

- RDF - data model and examples
- RDFS - data schemas (models)
- SPARQL - RDF query language
- SPARQL demos
- OWL - Web Ontology Language
- SWRL - Semantic Web Rule Language

What is RDF?

A Universal Data Model consisting of statements:

subject - predicate - object

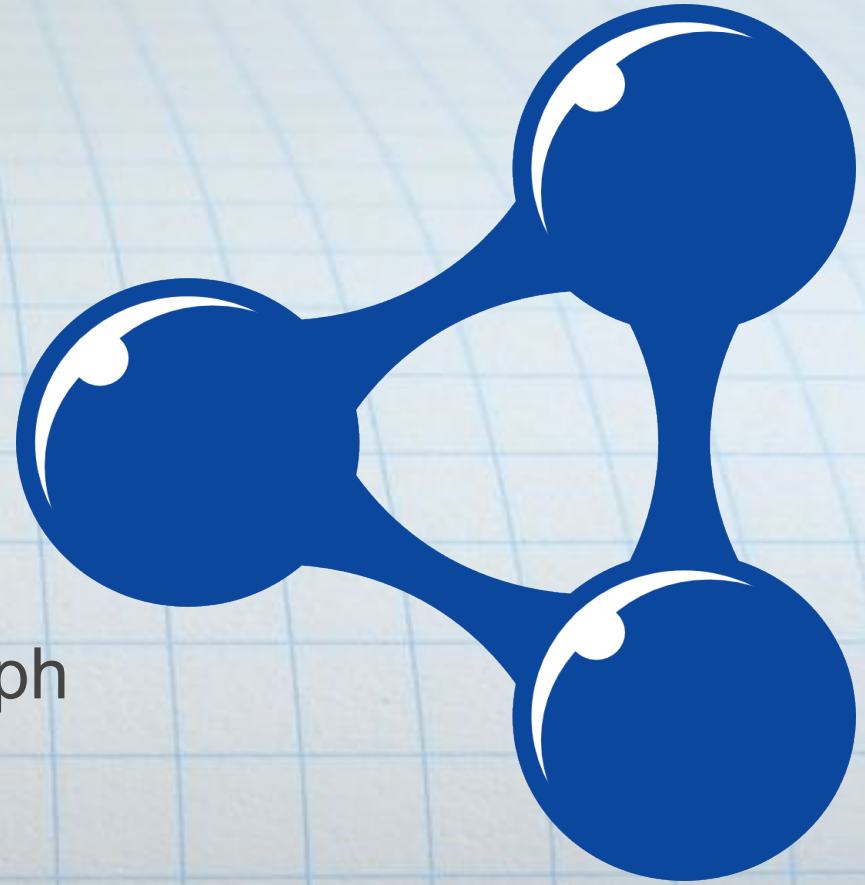
the set of RDF statements form a graph

Informal example:

ns1:myHouse ns2:hasColor pantone:chartreuse

pantone:chartreuse pantone:redComponent "13"^^xsd:int

pantone:chartreuse pantone:greenComponent "214"^^xsd:int



An RDF Example

This particular SVCC session (let's call it "SemWeb")
has the title "A Hands-On Introduction to the Semantic Web"
the description "The Semantic Web is in its infancy ..."

<http://www.siliconvalley-codecamp.com/sessions#SemWeb> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>

<http://www.siliconvalley-codecamp.com/rdfs#Sessions>

has the

<http://purl.org/dc/elements/1.1/title>

of

"A Hands-On Introduction to the Semantic Web"^^xsd:string

and the

<http://purl.org/dc/elements/1.1/description>

of

"The Semantic Web is in its infancy ..."^^xsd:string

RDF Graph

<http://www.siliconvalley-codecamp.com/rdfs#Sessions>

"The Semantic Web is in its infancy..."

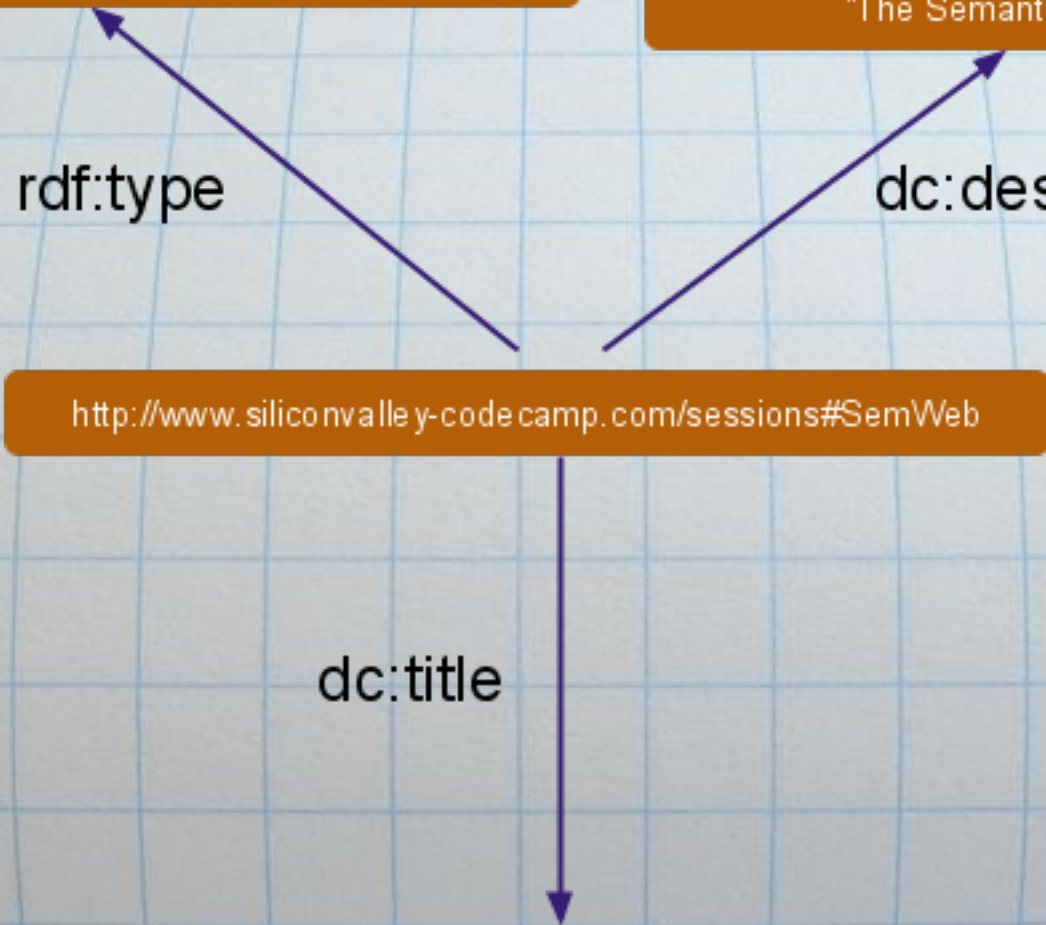
rdf:type

dc:description

<http://www.siliconvalley-codecamp.com/sessions#SemWeb>

dc:title

"A Hands-On Introduction to the Semantic Web"



RDF/XML

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:svcc="http://www.siliconvalley-codecamp.com/rdfs#"
  xml:base="http://www.siliconvalley-codecamp.com/sesssions">
  <rdf:Description
    rdf:ID="http://www.siliconvalley-codecamp.com/sesssions#SemWeb">
    <rdf:type rdf:resource="svcc:Session" />
    <dc:title>A Hands-On Introduction to the Semantic Web</dc:title>
    <dc:description>The Semantic Web is in its infancy ...</dc:description>
  </rdf:Description>
</rdf:RDF>
```

Other forms of RDF Serialization

N3, TURTLE, N-Triple, etc.

@prefix dc: <http://purl.org/dc/elements/1.1/> .

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

@prefix svcc: <http://www.siliconvalley-codecamp.com/rdfs#> .

<http://www.siliconvalley-codecamp.com/sesssions#SemWeb>

a svcc:Session ;

dc:description "The Semantic Web is in its infancy ..." ;

dc:title "A Hands-On Introduction to the Semantic Web" ;

RDF vs. XML

RDF (Graph-based) != XML (Tree-based).

RDF graphs represent information

XML trees represent data

RDF/XML serializes RDF

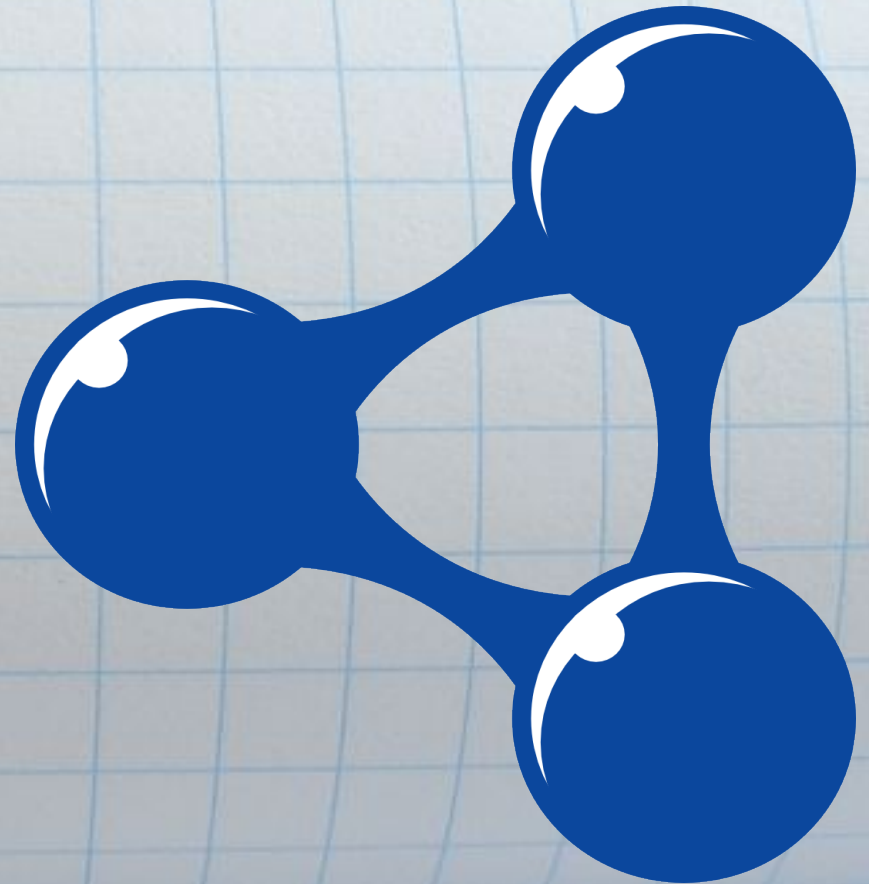


RDF storage

- RDF statements are stored in RDF Repositories
 - Also called triple stores
- Data can be accessed via:
 - SPARQL queries
 - API calls
 - Text searches

What is RDFS?

a representation for defining schemas for RDF



RDF Schema

```
<?xml version="1.0"?>
<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://www.siliconvalley-codecamp.com/rdfs">

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

  <rdfs:Class rdf:ID="Speaker">
    <rdfs:subClassOf rdf:resource="http://xmlns.com/foaf/0.1/Person" />
  </rdfs:Class>
  <rdf:Property rdf:ID="presenter">
    <rdfs:domain rdf:resource="#Session" />
    <rdfs:range rdf:resource="#Speaker" />
    <rdfs:label>Presenter</rdfs:label>
  </rdf:Property>
</rdf:RDF>
```

RDFS: Classes and Properties

Class Name
rdf:List
rdfs:ContainerMembershipProperty
rdfs:Container
rdf:Alt
rdf:Seq
rdf:Bag
rdf:Statement
rdfs:Datatype
rdf:Property
rdfs:Class
rdf:XMLLiteral
rdfs:Resource

Property Name
rdf:type
rdfs:subClassOf
rdfs:subPropertyOf
rdfs:domain
rdfs:range
rdfs:label
rdfs:comment
rdfs:member
rdf:first
rdf:rest
rdfs:seeAlso
rdfs:isDefinedBy
rdf:value
rdf:subject
rdf:object
rdf:predicate

RDF or RDFS questions?

Any questions on RDFS?

See <http://www.hewettresearch.com/svcc2009/>
for related materials.

Coming next:

- SPARQL - Semantic Web Query Language
- RDFa
- Linked Open Data
- OWL - Web Ontology Language
- SWRL - Semantic Web Rule Language

SPARQL

- a W3C standard query language to fetch data from distributed Semantic Web data models (mainly, RDF and OWL)
- a concept similar to SQL for Database
- can query a triple-store (local RDF repository) or data on the Web (at a URL)

SPARQL - a query language for RDF

```
PREFIX svcc:<http://www.siliconvalley-codecamp.com/rdfs#>
```

```
PREFIX dc:<http://purl.org/dc/elements/1.1/>
```

```
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
```

```
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
```

```
SELECT DISTINCT ?title ?presenter ?description
```

```
WHERE
```

```
{
```

```
  ?session rdf:type svcc:Session .
```

```
  ?session dc:title ?title .
```

```
  ?session svcc:presenter ?presenter .
```

```
  ?session dc:description ?description .
```

```
}
```

SPARQL - a query language for RDF

PREFIX svcc:<<http://www.siliconvalley-codecamp.com/rdfs#>>

PREFIX dc:<<http://purl.org/dc/elements/1.1/>>

PREFIX rdfs:<<http://www.w3.org/2000/01/rdf-schema#>>

PREFIX rdf:<<http://www.w3.org/1999/02/22-rdf-syntax-ns#>>

PREFIX fn: <<http://www.w3.org/2005/xpath-functions#>>

SELECT ?title ?presenter ?description

WHERE

{

?session dc:title ?title .

?session svcc:presenter ?presenter .

?session dc:description ?description .

FILTER (fn:string-length(?description) < 100)

}

SPARQL - a query language for RDF

```
PREFIX svcc:<http://www.siliconvalley-codecamp.com/rdfs#>
```

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
```

```
CONSTRUCT { ?session svcc:abstract ?description }
```

```
WHERE
```

```
{
```

```
  ?session dc:description ?description .
```

```
}
```

```
LIMIT 2
```

SPARQL - a query language for RDF

PREFIX svcc: <<http://www.siliconvalley-codecamp.com/rdfs#>>

PREFIX speaker: <<http://www.siliconvalley-codecamp.com/speakers#>>

ASK

{

?x svcc:presenter speaker:Shamod_Lacoul

}

Other SPARQL operations

Operation

DESCRIBE

WHERE operations

OPTIONAL

UNION

ORDER BY

DISTINCT

REDUCED

OFFSET

LIMIT

Reference Clause

FROM

FROM NAMED

SPARQL/UPDATE - a query language to fulfill CRUD operations
INSERT & DELETE

SPARQL results

- Many different formats:
 - SPARQL/XML
 - JSON
 - CSV
 - RDF/XML
- You will need to parse the results for your needs

SPARQL endpoints

- A SPARQL endpoint is a web service that accepts SPARQL queries and returns results
- Example: <http://www.govtrack.us/developers/rdf.xpd>

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
SELECT * WHERE {
<http://www.rdfabout.com/rdf/usgov/congress/109/bills/h867> ?p ?o
}
```

Setting up a SPARQL endpoint

- Use Sesame - <http://www.openrdf.org/>
- Includes a web service that runs under Tomcat
- Configure it to use your favorite backend RDF store or use the built-in RDF store
- Some good, scalable RDF stores that support Sesame:
 - SwiftOWLIM - <http://ontotext.com/owlim/>
 - AllegroGraph - <http://www.franz.com/agraph/allegrograph/>

SPARQL demo

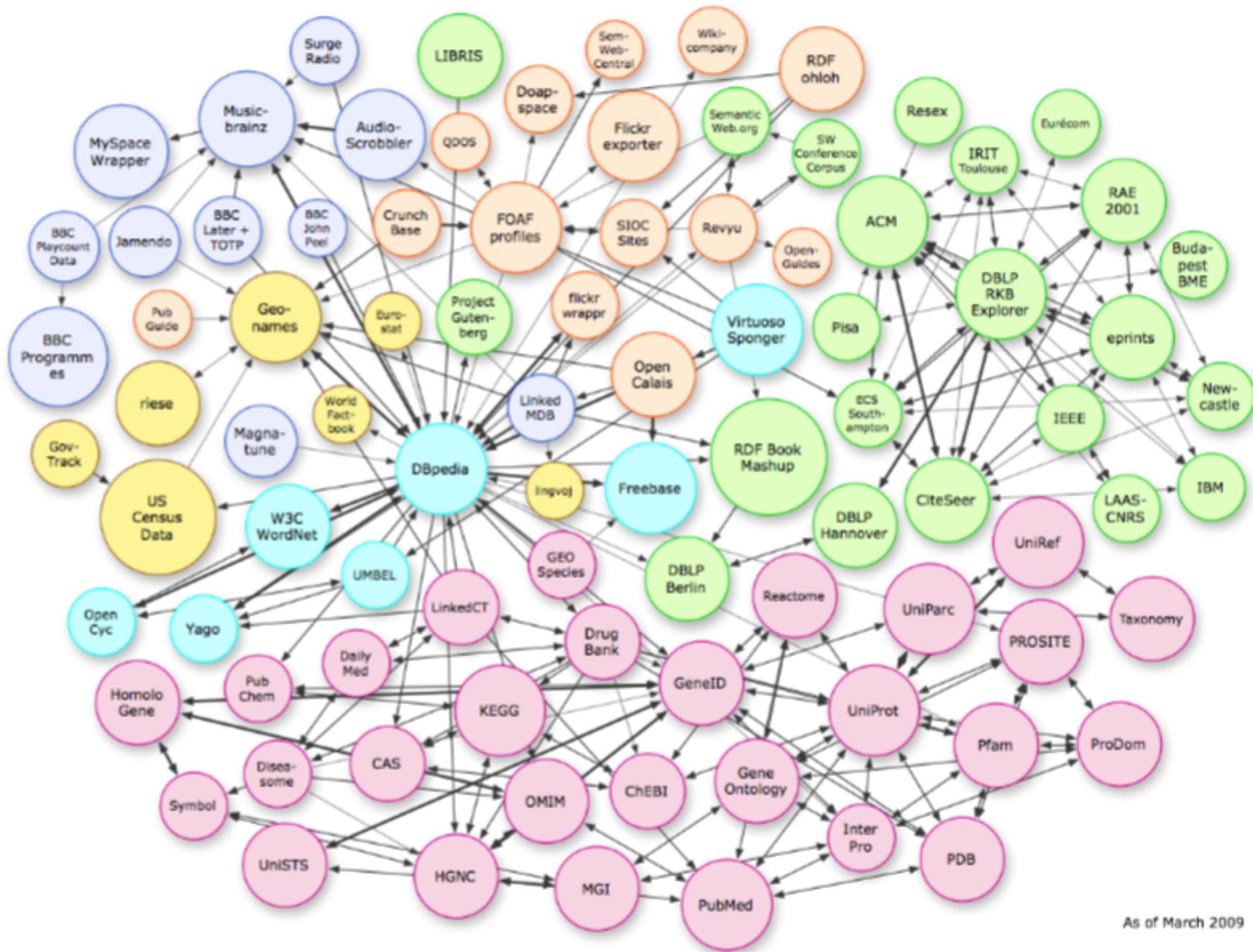
See <http://www.hewettresearch.com/svcc2009/>
for related materials.

RDFa example

```
<?xml version="1.0" encoding="UTF-8" ?>
<html xmlns="http://www.w3.org/1999/xhtml"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  version="XHTML+RDFa 1.0" xml:lang="en">
<head><title>Silicon Valley Code Camp Session</title></head>
<body>
  <h3>Silicon Valley Code Camp Session Registration</h3>
  <div about="http://www.silicon-valley.com/sessions#SemWeb">
    <div>
      <b>Title: </b>
      <span id="title" property="dc:title">A Hands-On Intro To the Semantic Web</span>
    </div><br /><br />
    <div>
      <b>Description: </b>
      <span id="desc" property="dc:description">The Semantic Web is in its infancy ...
    </span>
    </div><br /><br />
  </div>
</body>
</html>
```

Linking Open Data (RDF)

- 4.7 billion RDF triples
- 142 million RDF links (as of May 2009)
- LOD examples
 - DBPedia, Freebase
 - BBC, MusicBrainz, Flickr
 - SIOC, FOAF



As of March 2009

How to access SW information?

- Semantic Web Search Engines
 - Indexing RDF information by keywords
 - Zemanta, Sindice, SWSE, Swoogle
- Query SPARQL endpoints
 - Apply SPARQL queries directly to the RDF data
- Query SWRL endpoints (when they exist)
 - Run rules on OWL ontologies

Questions about SPARQL or RDFa?

See <http://www.hewettresearch.com/svcc2009/>
for related materials.

Coming next:

- OWL - Web Ontology Language
- SWRL - Semantic Web Rule Language

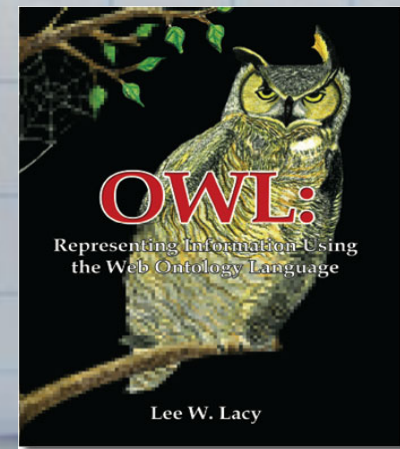
Web Ontology Language (OWL)

An extension to RDF/RDFS to enable complex knowledge representations

a language for defining and instantiating *ontologies*

An *OWL ontology* may include descriptions of *classes*, *properties* and their instances.

Based on Open-World Assumption - what is not known is not "untrue", just "unknown".



Flavors of OWL ontologies

OWL Lite - supports classification in hierarchies and simple constraints

OWL DL - correspondence with Description Logics

OWL Full - maximum expressiveness

OWL 2 - (W3C Candidate Recommendation, June 2009)

- OWL 2 EL - has computational guarantees
- OWL 2 QL - maps to database technologies
- OWL 2 RL - computationally efficient

OWL

```
<?xml version="1.0" ?>
<rdf:RDF
  xml:base = "http://www.siliconvalley-codecamp.com/tags#"
  xmlns:dc = "http://purl.org/dc/elements/1.1/"
  xmlns:owl = "http://www.w3.org/2002/07/owl#"
  xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs = "http://www.w3.org/2000/01/rdf-schema#"
  xmlns:tags = "http://www.holygoat.co.uk/owl/redwood/0.1/tags/">

  <owl:Ontology rdf:about="">
    <rdfs:comment>An ontology to define tags of Silicon Valley Code Camp</rdfs:
comment>
    <rdfs:label>A tag ontology</rdfs:label>
    <owl:imports rdf:resource="http://www.holygoat.co.uk/owl/redwood/0.1/tags/" />
    <dc:creator>Shamod Lacoul</dc:creator>
  </owl:Ontology>

  <tags:Tag rdf:ID="OWL">
    <tags:name>Web Ontology Language</tags:name>
    <tags:isTagOf rdf:resource="http://www.siliconvalley-codecamp.
com/sessions#SemWeb" />
    <tags:relatedTag rdf:resource="http://www.siliconvalley-codecamp.com/tags#RDF"
/>
  </tags:Tag>

</rdf:RDF>
```

OWL: Classes and Properties

Classes

owl:Class

rdfs:subClassOf

Property Restrictions

owl:allValuesFrom

owl:someValuesFrom

owl:cardinality

owl:hasValue

Properties

owl:ObjectProperty

owl:DatatypeProperty

rdfs:subPropertyOf

rdfs:domain

rdfs:range

Property Characteristic

owl:TransitiveProperty

owl:FunctionalProperty

owl:inverseOf

owl:InverseFunctionalProperty

Equivalence

owl:equivalentClass

owl:equivalentProperty

owl:sameAs

Complex Classes

owl:intersectionOf

owl:unionOf

owl:complementOf

NOTE: A subset of OWL Classes and Properties

Source: <http://www.w3.org/TR/owl-guide/>

Compare other models to SemWeb Models

	Database	XML	RDF	OWL
Expressiveness	medium	low	medium	high
Accessibility	low	medium	high	high
Inferencing	low	low	medium	high
Flexibility	low	medium	medium	medium

Advantages of RDF and OWL

Expressiveness (explicit metadata)

Inherently Distributed

Easier data interchange

Reasoning and Inferencing capabilities.

Reasoning / Inferencing

Both refer to ways of inferring more information than is in the asserted data model

- **Reasoning** does two things:
 - Validates information:
 - domain & range of properties
 - valid subclasses & instances
 - Classifies instances based on their properties
 - Uses logic - usually first-order predicate logic
- **Inferencing** uses explicit rules
 - typically domain-specific rules
 - SWRL and RIF are the W3C standard rule languages

OWL example / demo

See <http://www.hewettresearch.com/svcc2009/>
for related materials.

Coming next:

- SWRL - Semantic Web Rule Language
- Useful tools, companies and links

SWRL

- The proposed W3C standard rule language
- Part of the RuleML family of languages
- Use it to reason with data in OWL knowledge bases
- Support for OWL is still fairly primitive

SWRL example

- $\text{hasParent}(\text{Adam}, ?\text{parent}) \wedge \text{hasParent}(?parent, ?g\text{-parent}) \wedge \text{hasGender}(?g\text{-parent}, \text{Male})$
- $\text{hasParent}(\text{Adam}, ?parent)$; Clause 1
- $\text{hasParent}(?parent, ?g\text{-parent})$; Clause 2
- $\text{hasGender}(?g\text{-parent}, \text{Male})$; Clause 3

- Each clause may have multiple bindings for its variable(s)
- Bindings carry forward to the next clause
- At the end, all binding sets are returned

Results

- `hasParent(Adam, ?parent) ^ hasParent(?parent, ?g-parent) ^ hasGender(?g-parent, Male)`
- Result 1:
 - `?parent = Bob`
 - `?g-parent = Cassius`
- Result 2:
 - `?parent = Betty`
 - `?g-parent = Charles`

SWRL Summary

- More complex queries and actions than SPARQL
- Support is weak at this point
- JESS – now supports SWRL
- Pellet – free reasoner that partially supports SWRL

SWRL questions?

See <http://www.hewettresearch.com/svcc2009/>
for related materials.

Coming next:

- Useful tools, companies and links

Semantic Web Tools

Java libraries for RDF/OWL

Jena (O)



Sesame (O)



RDF Repositories (Triplestores)

AllegroGraph (OC)



Virtuoso (C)



OWLIM (OC)

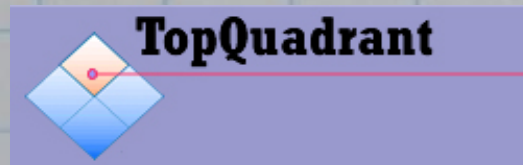


Semantic Web Editor

Protege (O)



TopQuadrant (OC)



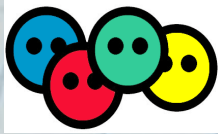
O - Open Source

C - Closed Source

OC - Open & Closed

Popular Public Ontologies

FOAF



NCBO



Dublin Core



GoodRelations



MusicBrainz

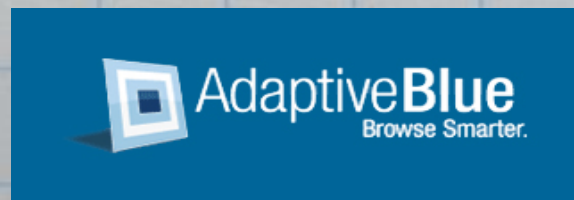
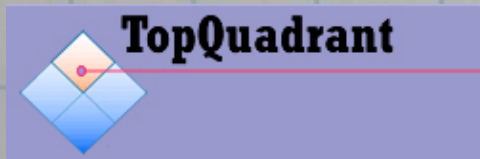


NCBO BioPortal

- OWL Ontologies
- <http://bioportal.bioontology.org/>
- 161 ontologies
- 723,806 concepts
- From ATMO (African Traditional Medicine) to ZFA (Zebrafish Anatomy and Development)
- Government funded, freely available

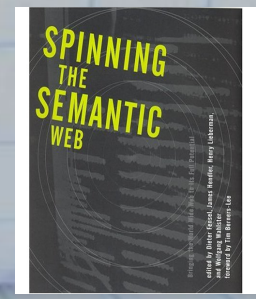
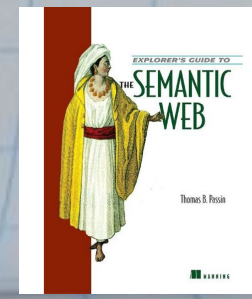
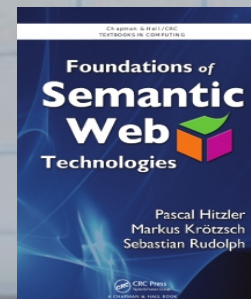
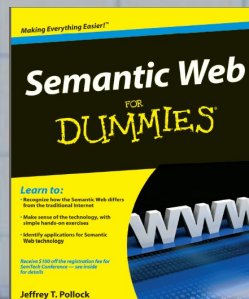
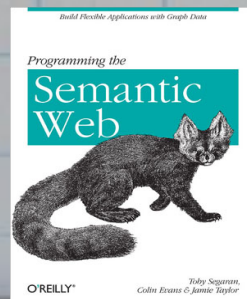
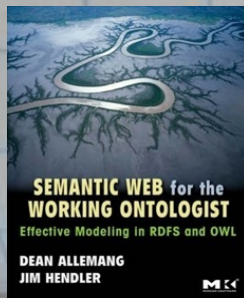
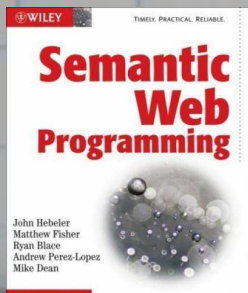


Prominent SemWeb Companies



Helpful References and Books

- <http://www.w3.org/2001/sw/BestPractices/Tutorials>
- <http://www.w3.org/TR/rdfa-syntax/>
- <http://jena.sourceforge.net/documentation.html>
- http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html
- <http://www.devx.com/semantic/Door/34578>
- <http://semanticweb.com>
- <http://semanticuniverse.com/>
- <http://www.mkbergman.com/>



Final Remarks

Semantic Web is not that hard. It comes with a learning curve, but so does everything else in life.

"If Yan can cook, so can you!"



See <http://www.hewettresearch.com/svcc2009/>
for related materials.



shamod@gmail.com

mike@hewettresearch.com

The END