

COM556

SEMANTIC WEB TECHNOLOGIES

Week 1

Semantic Web Vision and

Introduction

Assist. Prof. Dr. Melike Şah Direkoglu

Acknowledgements:

Dr. Myungjin Lee's lecture notes from Linked Data and Semantic Web Technology (Korea), Ivan Herman's tutorial from W3C, Marin Dimitrov's GATE tutorial slides and Declan O'sullivan's lecture slides from Trinity College Dublin were used in the preparation of these slides

Outline

- Semantic Web and Semantic Web Vision
- Semantic Web Technologies
- Semantic Web Case Studies

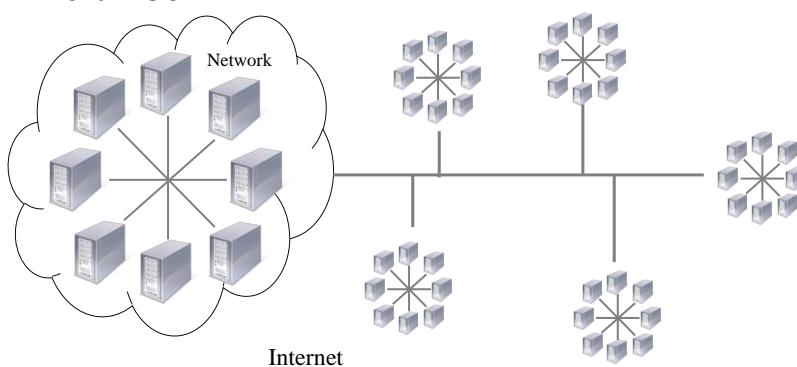
Outline

- **What is Semantic Web and its Vision?**
- Semantic Web Technologies
- Semantic Web Case Studies

3

Internet

- A global system of interconnected computer networks
- A network of networks
- Network
 - a collection of computers interconnected by communication channels



4
[Myungjin Lee]

Internet Services before the Web

- E-Mail Communication: SMTP, POP3
- File Transfer: FTP
- Remote Control: Telnet



```

C:\> netstat -an
Connection user: Enter your name and phone:
Name/phone:John Smith

C:\> netstat -an | more
Copyright www.n!l.com 2008 (C) x

      1. AN [bound] Show this help
      2. Open CDM port Open CDM port
      3. Open CDM port Open CDM port exclusively
      4. or lock 1. &h Open CDM port for reading only
      5. close CDM port Close CDM port
      6. echo Echoes the command
      7. exit Exit from the program
      8. or stat Show ports status
      9. connect Connect to a remote host
      10. or ... Show current opened port
      11. update Update ports information from file

connections:
ports listed:
  1. CDM is held by 127.0.0.1 (Local PC) For 03:10
  2. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  3. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  4. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  5. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  6. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  7. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  8. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  9. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  10. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  11. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)

converse:
  1. CDM is held by 127.0.0.1 (Local PC) For 03:10
  2. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  3. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  4. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  5. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  6. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  7. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  8. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  9. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  10. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)
  11. CDM, 127.0.0.1 (Local PC) For 03:10 (John Smith)

converse>

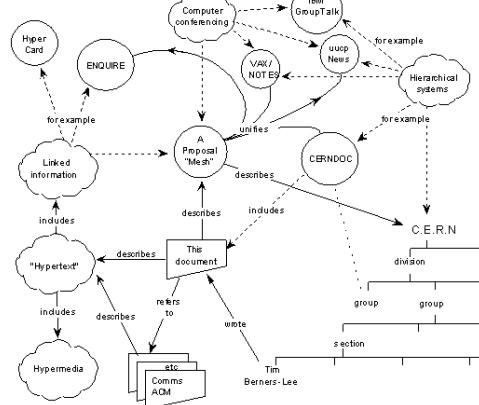
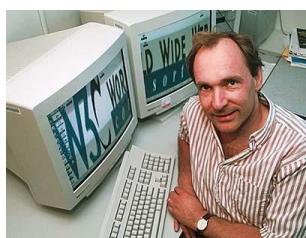
```

- Problem of these services:
 - Information access requires expert knowledge
 - Information access is expensive...
 - Information retrieval is very expensive...

5
[Myungjin Lee]

World Wide Web (WWW)

- A system of interlinked hypertext documents accessed via the Internet (invented by Sir Tim Berners-Lee in 1993)
- Berners-Lee also invented the first Web browser & Web server



Proposal of "Hypertext project" called "World Wide Web"

6
[Myungjin Lee]

Characteristics of Web

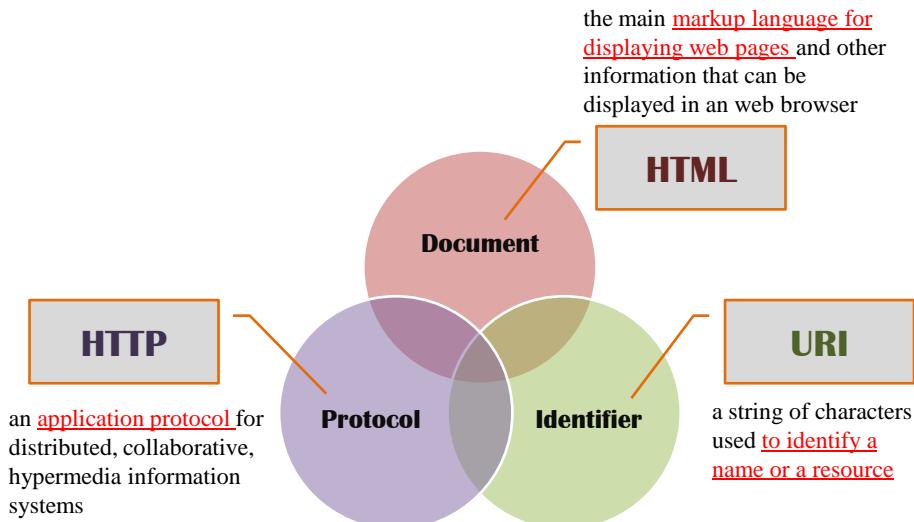
- Hyperlink and Multimedia



- Advantages:
 - No expert knowledge required
 - Simple information access
 - Information retrieval via search engines

⁷
[Myungjin Lee]

Web Architecture



[Myungjin Lee]⁸



Problem of HTML

- HTML describes
 - how information is presented, displayed, and linked for human readers
- There is no meaning of information.

AutoTrader.com

Primary information about this vehicle:	
AT Car ID:	AT-1209232A [?]
Price	United
MSRP	\$63,930
Mileage	28
Body Style	Sedan
Exterior Color	Black
Interior Color	Brown
Engine	6 Cylinder Supercharged
Transmission	8 Speed Shiftable Automatic
Drive Type	All wheel drive
Fuel Type	Gasoline
Doors	Four Door
Stock No.	A120027
VIN	View VIN

same
information,
but ...

Cars.com

About This A6	
Mileage:	40
Body Style:	Sedan
Exterior Color:	Black
Interior Color:	Black
Stock #:	CN131602
VIN:	WAUJGAFC7CN131602

Features:

- ABS (4-Wheel)
- Air Cond Rear
- AM/FM Stereo
- [Show Additional Features](#)

Problem of HTML

Google A6

Web Images Maps Shopping More + Search tools

About 283,000,000 results (0.32 seconds)

Audi A6 > Audi of America
models audiina.com/A6 - Cached
Given that the previous Audi A6 was in 100 winning comparison tests in its final year, its replacement had big shoes to fill. Yet the new model has proven every bit ...
Models - Features and ... Explore > Audi A6 > Audi of ... A6 specifications

Apple A6 - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Apple_A6 - Cached
The Apple A6 is a system on a chip (SoC) designed by Apple Inc. that drives the iPhone 5 which was introduced on September 12, 2012. Apple states that it is ... Design - Performance - Products that include the Apple A6 - See also

A6 road (England) - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/A6_road_(England) - Cached
The A6 is one of the main historic north-south roads in England. It currently runs from Luton in Bedfordshire to Carlisle in Cumbria, although it formerly started at ...

A6 > MotorWorld.de - Audi Worldwide
www.audiworld.de/contenut/a6?viewType=4&grid=...
A6 Models Audi Worldwide ... High-tech for the business class: the Audi A6 Limousine's assistance- and multimedia systems are full of functions, yet easy to ...

A6 - Sambandden.com
www.sambandden.com/contenut/a6?viewType=4&grid=...
A description for this result is not available because of this site's robots.txt – learn more.

A6 Center
alcenter.se/ - Cached - Translate this page
4 days ago – Ett av Sveriges största kopcentra. Information, fakta, annonser och nyheter.

Apple A6 News - Autoblog
www.autoblog.com / Automobiles / Audi - Cached
Read the latest Audi A6 news and reviews complete with photos, videos, and road tests of new Audi A6 vehicles.

2013 Audi A6



Audi A6



Maserati A6



A6 Metrobus Lines



[Myungjin Lee]



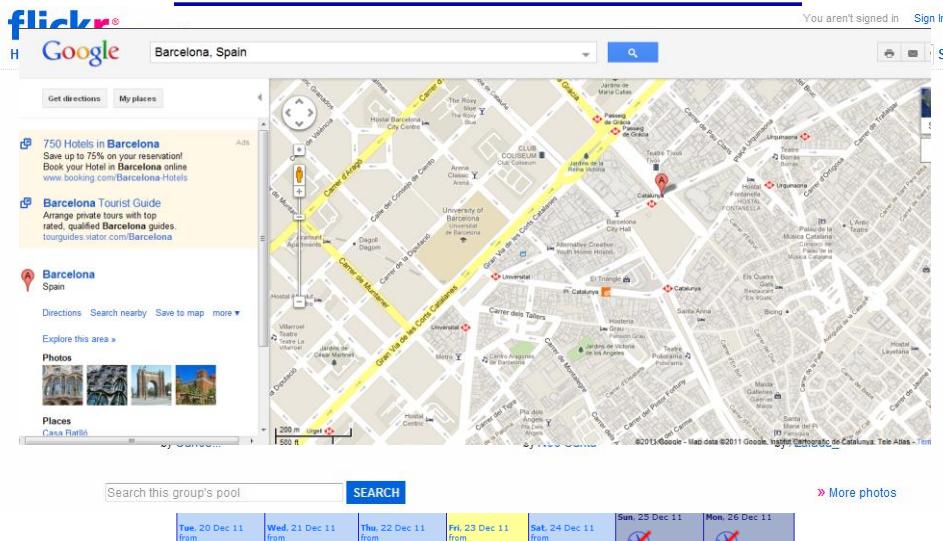
Apple A6

11

An Example to Illustrate the Problem of HTML

Let's organize a trip to Barcelona using the Web

Find a proper flight and accommodation!



What happened?

- You had to consult a large number of sites, all different in style, purpose and possibly in language
- You had to mentally integrate all these information to achieve your goals
- As you all know, sometimes it is long and tedious process
- In addition, what you see is the tip of the iceberg, the real data is hidden in databases, XML files, Excel sheets,...
- You can only access to what the Web page designer allows you to see

[Ivan Herman, Intro Semantic Web Technologies, 2010]

The Web

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

[Marin Dimitrov, 3rd GATE tutorial, 2010]

Limitations of the Current Web

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

[Marin Dimitrov, 3rd GATE tutorial, 2010]

What we would like to have?

- Able to link data (independent of their presentation) and use the data the way I want
- Agents, programs, scripts, etc. should be able to interpret part of that data
- But wait, representation of the data and access to that data should be standardized so that different applications, platforms, etc. can use it!

[Ivan Herman, Intro Semantic Web Technologies, 2010]

Semantic 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)

What we want on the Web?

- to process the meaning of information automatically
- to relate and integrate heterogeneous data
- to deduce implicit information from existing information in an automated way

The Web was designed as an information space, with the goal that it should be useful not only for human-human communication, but also that **machines would be able to participate and help.**



[Myungjin Lee]

So what is the Semantic Web?

- The Semantic Web is a collection of standard technologies to realize the Web of Data and machine-processable Web
- Web → “links document to document”, “documents to READ”
- Semantic Web → “links data to data”, “data for all sorts of things”

The Semantic Web

- Target consumers: intelligent agents
 - Explicit specification of the intended meaning information
 - Intelligent agents can make use of 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 and integration can be automated

[Marin Dimitrov, 3rd GATE tutorial, 2010]

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

[Marin Dimitrov, 3rd GATE tutorial, 2010]

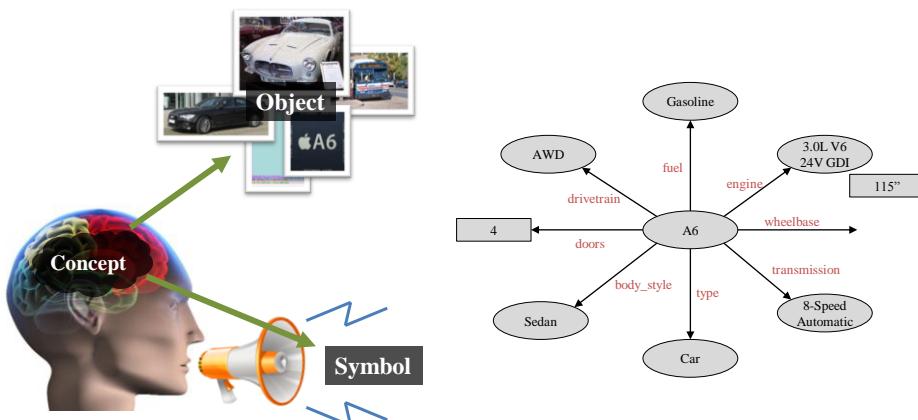
Next Steps on Web

- Next step is semantic interoperability:
 - Understanding what the data means
 - Linking in insightful ways
 - Automated support for data integration
 - Smart applications
- *Sharing data \Rightarrow Sharing meaning*

© Declan O'Sullivan

Approach of the Semantic Web

- Explicitly annotate metadata with its meaning that can be read and processed correctly by machines using Semantic Web technologies



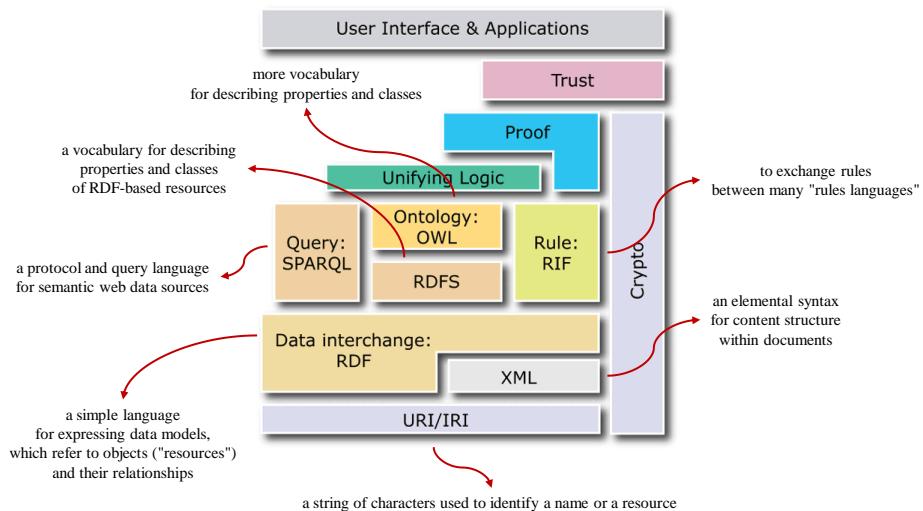
24
[Myungjin Lee]

Overview of the Semantic Web

- What is the Semantic Web?
- **Semantic Web Technologies**
- Semantic Web Case Studies

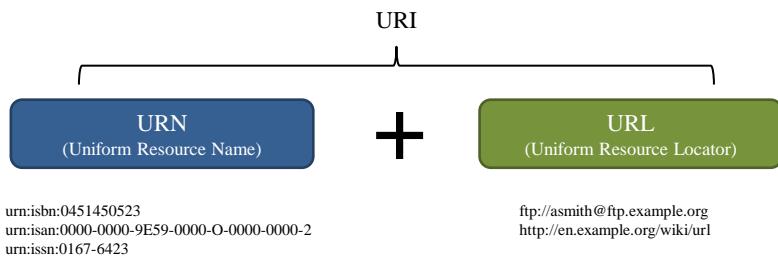
25

Semantic Web Layer Cake

[Myungjin Lee]²⁶

URI (Uniform Resource Identifier)

- a string of characters used to identify a name or a resource



²⁷
[Myungjin Lee]

XML (Extensible Markup Language)

- a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable

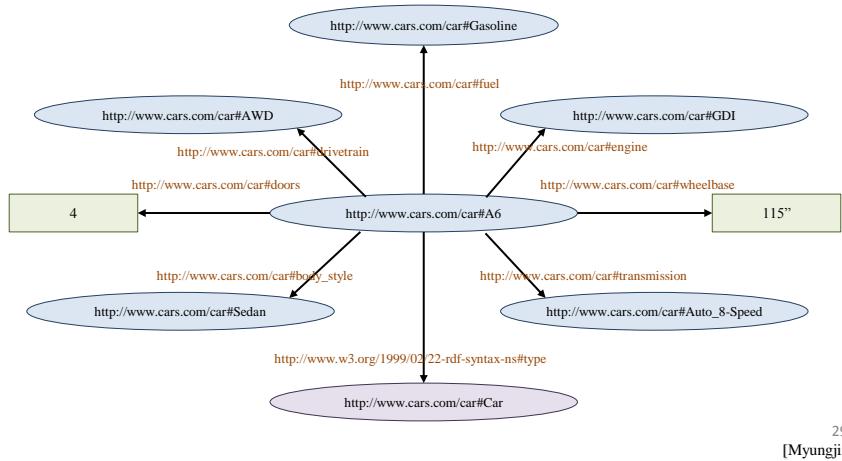
```

<?xml version="1.0" encoding="utf-8"?>
<note>
    <to>Tove</to>
    <from>Jani</from>
    <heading>Reminder</heading>
    <body>Don't forget me this weekend!</body>
</note>
    
```

²⁸
[Myungjin Lee]

RDF (Resource Description Framework)

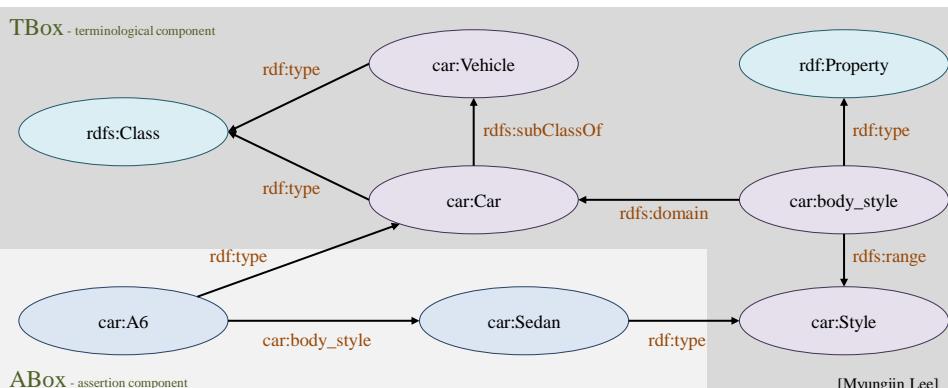
- A general method for conceptual description or modeling of information in web resources,
- There are variety of syntax formats (RDF/XML, n3, turtle, etc.)



29
[Myungjin Lee]

RDFS (RDF Schema)

- RDFS is a semantic extension of RDF
- Intends to structure RDF resources using classes and properties
 - describing groups of related resources and the relationships between these resources



[Myungjin Lee]

Ontology

- Knowledge representation as a set of concepts within a domain, and the relationships between those concepts
 - More vocabulary for describing classes and properties
- Formal, explicit specification of a shared conceptualisation

"Ontologies are often equated with taxonomic hierarchies of classes, class definitions, and the subsumption relation, but ontologies need not be limited to these forms. Ontologies are also not limited to conservative definitions — that is, definitions in the traditional logic sense that only introduce terminology and do not add any knowledge about the world. To specify a conceptualization, one needs to state axioms that do constrain the possible interpretations for the defined terms."

³¹
[Myungjin Lee]

OWL (Web Ontology Language)

- A family of knowledge representation languages for authoring ontologies on the Semantic Web

RDF Schema Features:

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

(In)Equality:

- equivalentClass
- equivalentProperty
- sameAs
- differentFrom
- AllDifferent
- distinctMembers

Property Characteristics:

- ObjectProperty
- DatatypeProperty
- inverseOf
- TransitiveProperty
- SymmetricProperty
- FunctionalProperty
- InverseFunctionalProperty

Property Restrictions:

- Restriction
- onProperty
- allValuesFrom
- someValuesFrom

Restricted Cardinality:

- minCardinality (only 0 or 1)
- maxCardinality (only 0 or 1)
- cardinality (only 0 or 1)

Header Information:

- Ontology
- imports

Class Intersection:

- intersectionOf

Versioning:

- versionInfo
- priorVersion
- backwardCompatibleWith
- incompatibleWith
- DeprecatedClass
- DeprecatedProperty

Annotation Properties:

- rdfs:label
- rdfs:comment
- rdfs:seeAlso
- rdfs:isDefinedBy
- AnnotationProperty
- OntologyProperty

Datatypes

- xsd:datatypes

³²
[Myungjin Lee]

Language for the Rule Description

- [SWRL \(Semantic Web Rule Language\)](#) is a proposal for a Semantic Web rules-language, combining sublanguages of the OWL Web Ontology Language (OWL DL and Lite) with those of the Rule Markup Language (Unary/Binary Datalog).

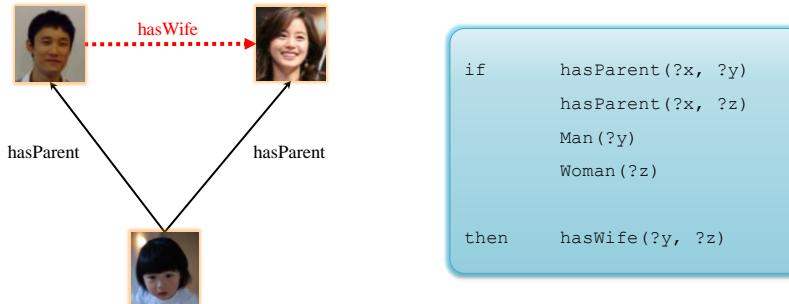
```
hasParent(?x1,?x2) ∧ hasBrother(?x2,?x3) ⇒ hasUncle(?x1,?x3)
```

```
<ruleml:imp>
  <ruleml:_rlab ruleml:href="#example1"/>
  <ruleml:_body>
    <swrlx:individualPropertyAtom swrlx:property="hasParent">
      <ruleml:var>x1</ruleml:var> <ruleml:var>x2</ruleml:var>
    </swrlx:individualPropertyAtom>
    <swrlx:individualPropertyAtom swrlx:property="hasBrother">
      <ruleml:var>x2</ruleml:var> <ruleml:var>x3</ruleml:var>
    </swrlx:individualPropertyAtom>
  </ruleml:_body>
  <ruleml:_head>
    <swrlx:individualPropertyAtom swrlx:property="hasUncle">
      <ruleml:var>x1</ruleml:var> <ruleml:var>x3</ruleml:var>
    </swrlx:individualPropertyAtom>
  </ruleml:_head>
</ruleml:imp>
```

33
[Myungjin Lee]

Inference

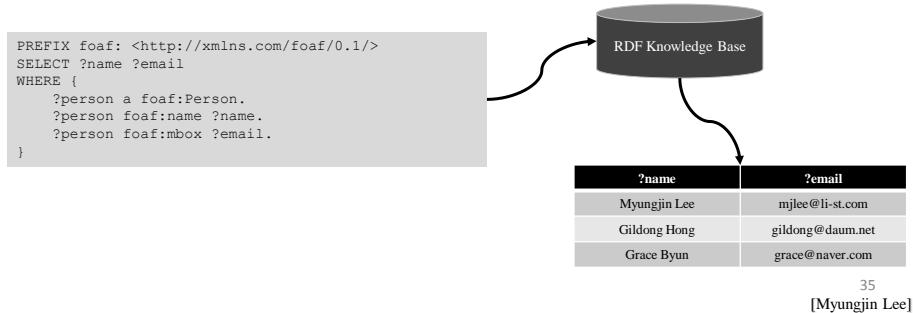
- Being able to derive new data from data that you already know



34
[Myungjin Lee]

SPARQL

- Why do we need a query language for RDF?
 - to get to the knowledge from RDF
- SPARQL Protocol and RDF Query Language
 - to retrieve and manipulate data stored in RDF format
 - to use SPARQL via HTTP



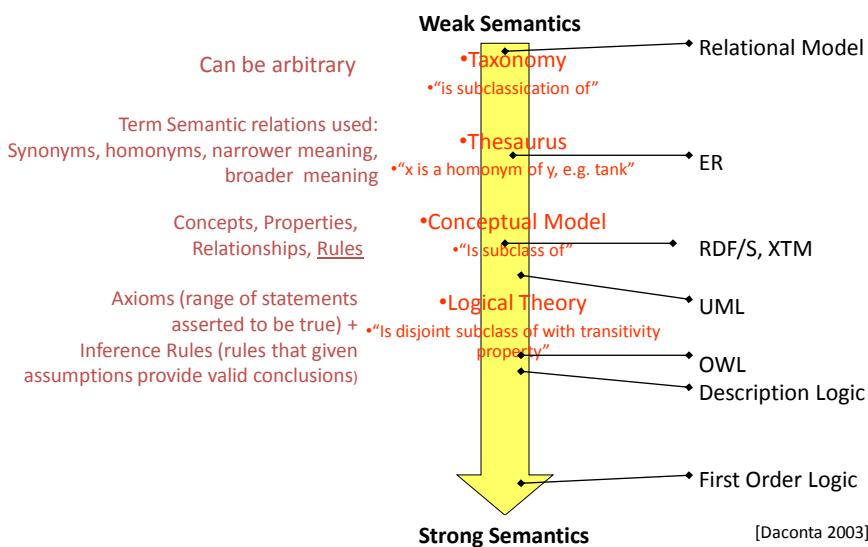
Ontologies

What is an ontology?

- An "**ontology**" describes the common words, concepts and relationships between concepts used to describe and represent an area of knowledge
- An ontology can range from a
 - **Taxonomy** (knowledge with minimal hierarchy or a parent/child structure)
 - **Thesaurus** (words and synonyms)
 - **Conceptual Model** (with classes, relationships, constraints)
 - **Logical Theory** (with very rich, complex, consistent and meaningful knowledge).

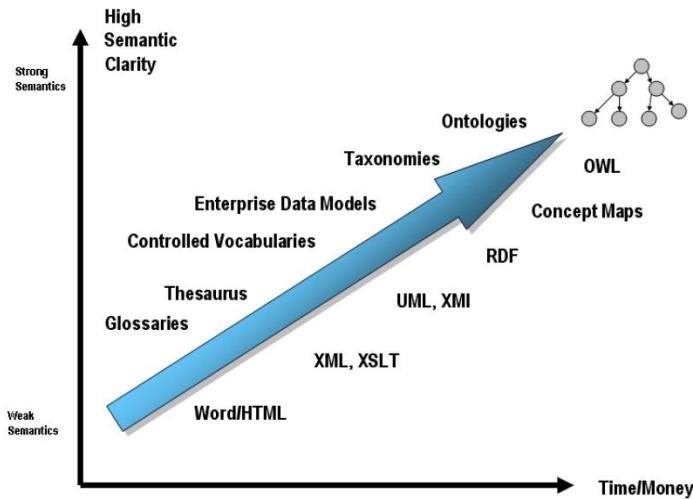
© Declan O'Sullivan

Ontology Spectrum



© Declan O'Sullivan

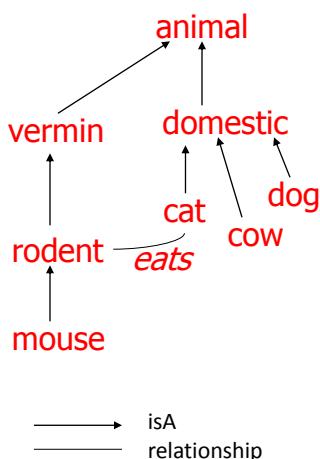
The cost of semantic clarity



[Marin Dimitrov, 3rd GATE tutorial, 2010]

Ontology Modeling An explicit description of a domain

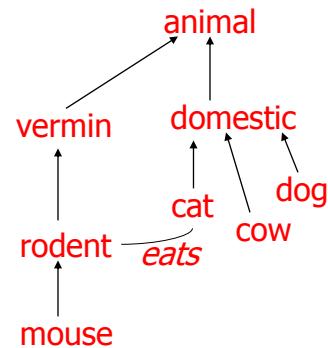
- Concepts (class, set, type)
 - animal, domestic, cat, dog,...
- Properties of concepts and relationships between them (slot, attribute)
 - *Taxonomy*: generalisation ordering among concepts *isA*, *partOf*, *subProcess*
 - *Relationship, Role or Attribute*: *functionOf*, *hasActivity*, *location*, *eats*, *size*



[Carole Goble, Nigel Shadbolt, Ontologies and the Grid Tutorial]

An explicit description of a domain

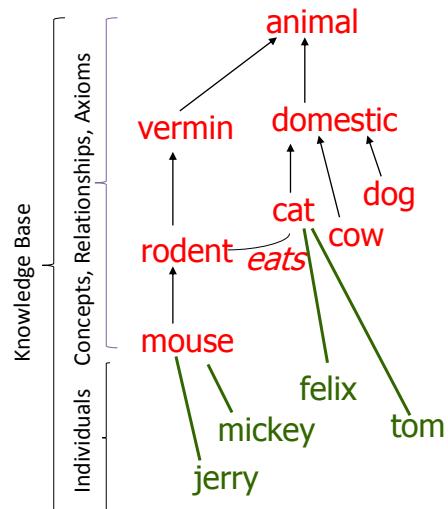
- Constraints or axioms on properties and concepts:
 - value: integer
 - domain: cat
 - cardinality: at most 1
 - range: $0 \leq X \leq 100$
 - cows are larger than dogs
 - cats cannot eat only vegetation
 - cats and dogs are disjoint
- Values or concrete domains
 - integer, strings
 - 20, mouse



© Declan O'Sullivan [Carole Goble, Nigel Shadbolt, Ontologies and the Grid Tutorial]

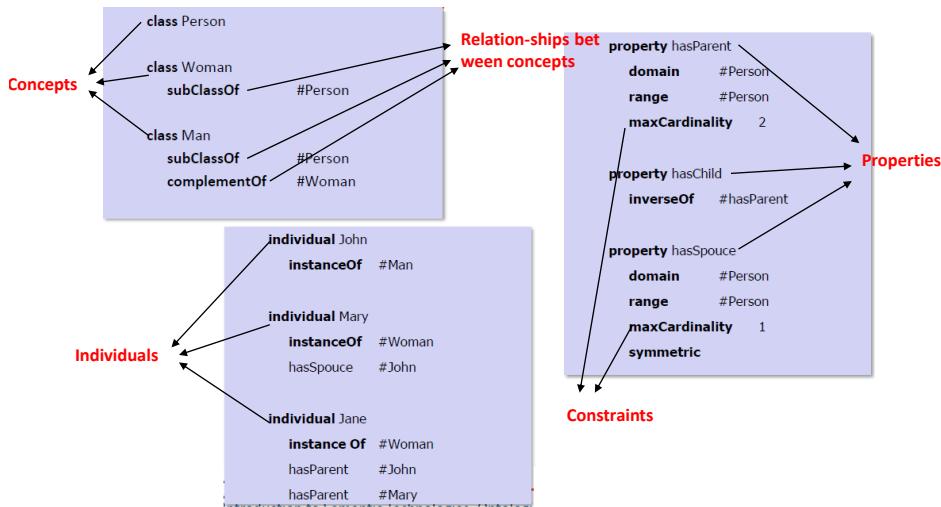
An explicit description of a domain

- Individuals or Instances
 - jerry, mickey, felix, tom
- Ontology versus Knowledge Base
 - An *ontology* = concepts+properties+axioms +values
 - A *knowledge base* = ontology+instances



© Declan O'Sullivan [Carole Goble, Nigel Shadbolt, Ontologies and the Grid Tutorial]

Ontology Specification Example



© Melike Sah

Semantic Web Standards

Resource Description Framework (RDF) - data model

RDF Schema (RDFS) - vocabulary

Web Ontology Language (OWL)

RDF Query Language (SPARQL)

Resource Description Framework (RDF)

Resource Description Framework (RDF)

- A simple data model for
 - Formally describing the semantics of information in a machine accessible way
 - Representing meta-data (data about data)
- Semantics = a way of encoding meaning (link between term and a model of the world) → Good for building applications
- Syntax = a way of encoding terms so that they can be distinguished, structured, grouped and related to each other in a grammar → Good for building parsers
- Note! We need syntaxes for expressing a machine-readable semantics
- Meta-data = data about data
 - Describe the information content of the underlying data independent of representational details
 - Describe the domain knowledge about the information domain, which allows inferences about the underlying data to be made
 - Examples: modification date of document, textual annotations describing an image, etc.

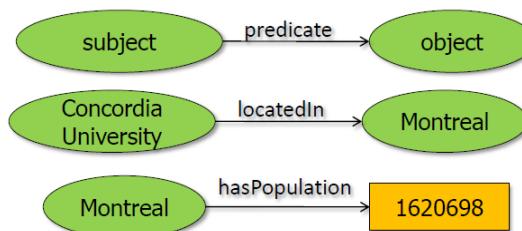
RDF (Cont.)

- A set of representation syntaxes
 - XML (standard) but also N3, Turtle, ...
- Building blocks
 - Resources (with unique identifiers – URI) as a global namespace of identifiers of things)
 - Unique across entire WWW
 - Literals
 - Named relations between pairs of resources (or a resource and a literal)

[Marin Dimitrov, 3rd GATE tutorial, 2010]

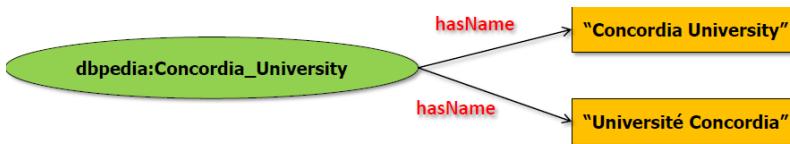
RDF (Cont.)

- Everything is a triple
 - Subject (resource), Predicate (relation), Object (resource or literal)
 - An RDF subject is always a resource => always a URI
 - An RDF object can be a resource or a literal value
 - What about predicates?
- The RDF graph is a collection of triples



[Marin Dimitrov, 3rd GATE tutorial, 2010]

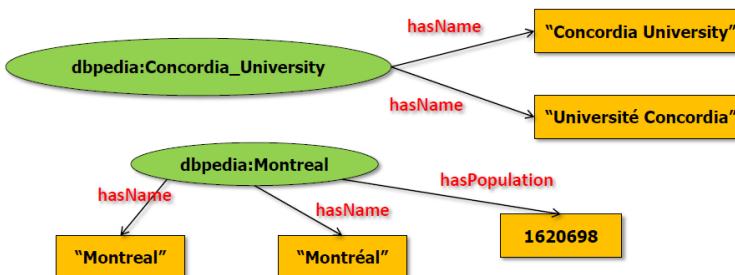
RDF Graph Example



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

[Marin Dimitrov, 3rd GATE tutorial, 2010]

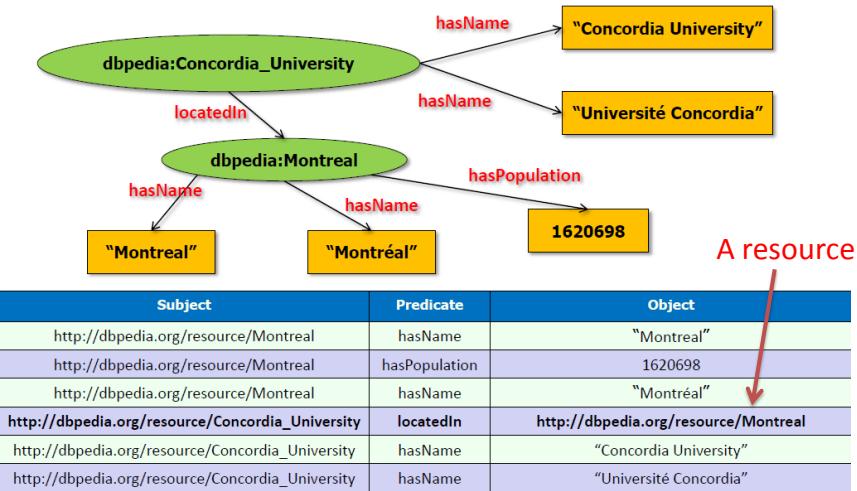
RDF Example (2)



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	hasName	"Concordia University"
http://dbpedia.org/resource/Concordia_University	hasName	"Université Concordia"

[Marin Dimitrov, 3rd GATE tutorial, 2010]

RDF Example (3)



[Marin Dimitrov, 3rd GATE tutorial, 2010]

RDF Advantages

- Simple but expressive data model
- Global identifiers of all resources (URIs)
- 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

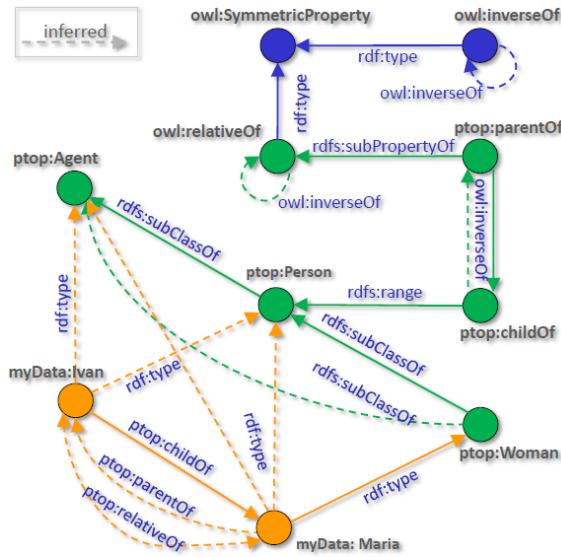
[Marin Dimitrov, 3rd GATE tutorial, 2010]

Resource Description Framework Schema (RDFS)

RDF Schema (RDFS)

- RDFS is a semantic extension of RDF
- RDFS provides mechanisms for describing groups of related resources and the relationships between these resources
- RDFS provides means for:
 - Defining Classes and Properties – rdfs:Class, rdfs:Property
 - Defining hierarchies (of classes and properties) – rdf:type, rdfs:subClassOf, rdfs:subPropertyOf
 - Restrictions – rdfs:domain, rdfs:range
- Using relationships between resources, new triples can be inferred from existing ones (RDFS axioms)

RDFS (Cont.)



[Marin Dimitrov, 3rd GATE tutorial, 2010]

Web Ontology Language (OWL)

OWL

- More expressive than RDFS
 - Identity equivalence/difference
 - owl:sameAs, owl:differentFrom, owl: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
- 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?

[Marin Dimitrov, 3rd GATE tutorial, 2010]

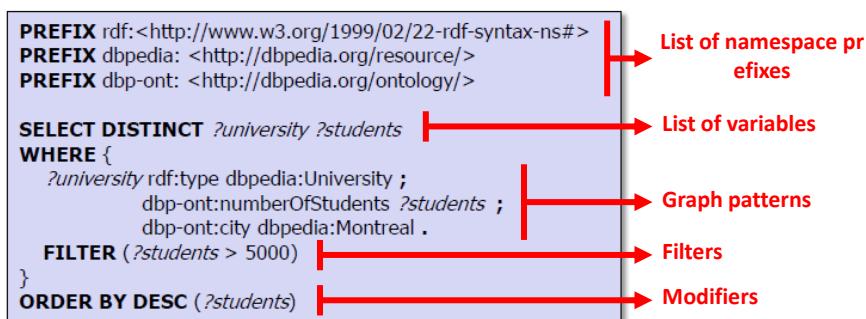
SPARQL Protocol and RDF Query Language for RDF

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

[Marin Dimitrov, 3rd GATE tutorial, 2010]

Anatomy of a SPARQL query



[Marin Dimitrov, 3rd GATE tutorial, 2010]

Linked Data

- Currently data is sitting in databases, pages, etc. out of reach, not useful...
- **Unlock the 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 and consumed

© Melike Sah

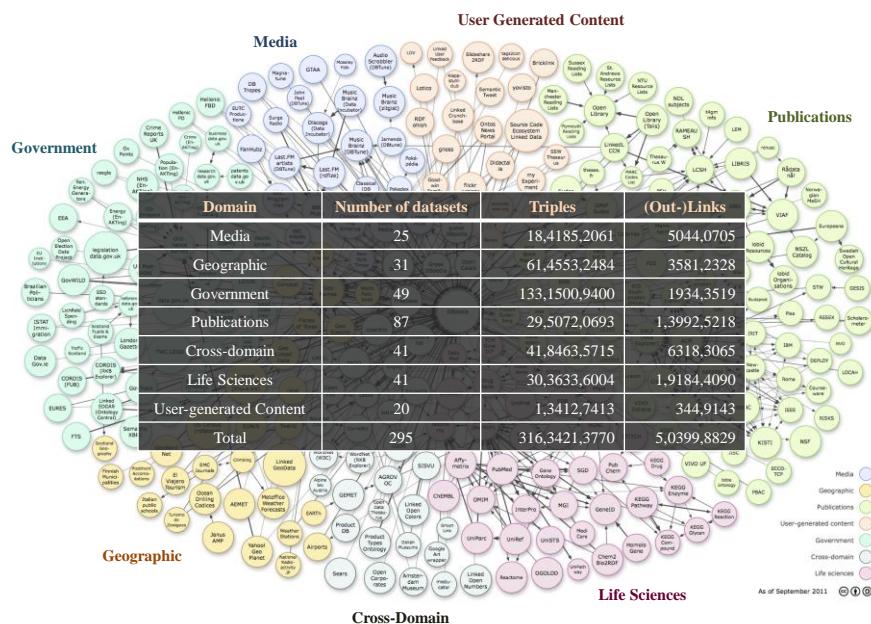
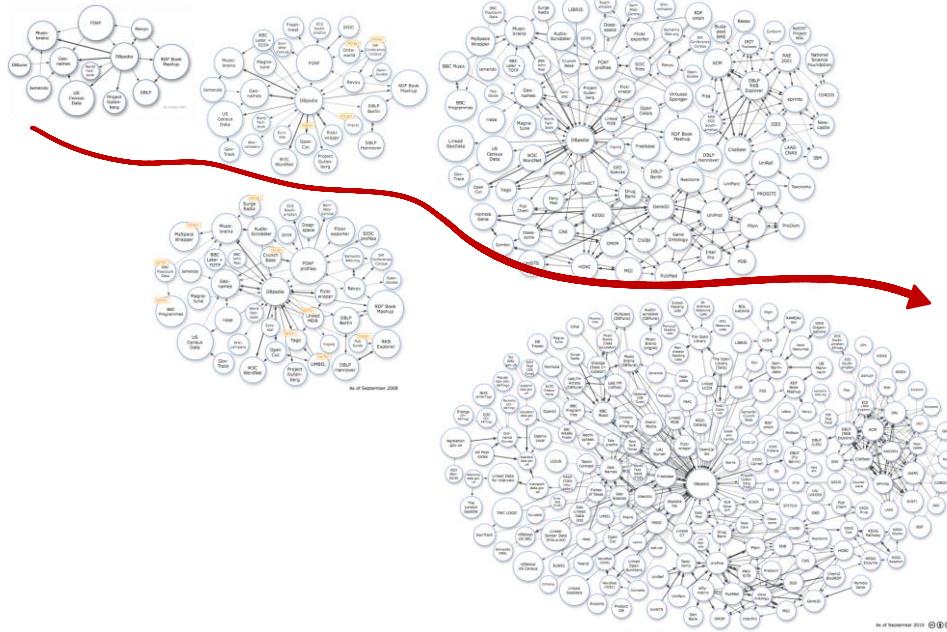
[Marin Dimitrov, 3rd GATE tutorial, 2010]

Linked Data Principles

- **Very Simple Three Rules**
- **1.** Use HTTP URIs for things (objects/resources) so that people can look up the names (using HTTP protocol)
- **2.** Provide useful information about that object (resource)
- **3.** Link the object (resource) to related objects – include links to other HTTP URIs – **data is relationships**

© Melike Sah

The Linking Open Data cloud diagram



Overview of the Semantic Web

- What is the Semantic Web?
- Semantic Web Technologies
- Semantic Web Case Studies

65

Sig.ma

SIGMA
SEMANTIC INFORMATION MASHUP

Tim Berners Lee [Add More Info](#) [Start New](#) [Options](#) [Use it!](#) [Order](#)

Tim Berners-Lee

picture:  [\[1\]](#) [\[7\]](#)  [\[13,14\]](#)

given name: Tim [\[2\]](#)

family name: Berners-Lee [\[2\]](#)

comment: Sir Timothy John "Tim" Berners-Lee, OM, KBE, FRS, FREng, FRS (born 8 June 1955), is a British engineer and computer scientist and MIT professor credited with inventing the World Wide Web, making the first proposal for it in March 1989. On 25 December 1990, with the help of Robert Cailliau and a young student at CERN, he implemented the first successful communication between an HTTP client and server via the Internet. [\[1\]](#)

is creator of: [Tabulator](#) [\[1\]](#)

is alternate of: [Tim Berners-Lee - semanticweb.org](#) [\[15\]](#)

alternate: http://www.squidoo.com/xml/syndicate_lens/tim-berners-lee [\[7\]](#)

[Essay metadata](#) [\[17\]](#)

birth year: 1955-01-01T00:00:00-05:00 [\[1\]](#)

Sources (20) Approved (0) Rejected (0) [Filter](#)

2 Tim Berners-Lee : Inform... 12 facts 2010-11-23 http://encyclopedia.vbxml.net/Tim_Berners-Lee
3 Smeerch.it :: Tim Be... 3 facts 2010-10-02 http://www.smeerch.it/tag/tim-berners-lee/
4 » tim berners lee NYLON 3 facts 2010-11-19 http://www.nylon.gr/tag/tim-berners-lee/
5 Tim Berners Lee Webz 3 facts 2009-10-22 http://www.webz.gr/tag/tim-berners-lee/
6 Untitled document 2 facts 2010-11-19 http://e-bergi.com/2008/Mayis/Tim-Berners-Lee
7 Tim Berners-Lee 12 facts 2010-10-02 http://www.squidoo.com/tim-berners-lee
8 Tim Berners Lee 4 facts 2010-10-02 http://plumbot.com/Tim_Berners-Lee.html
9 » tim berners lee 4 facts 2010-11-19 http://www.thewavingcat.com/tag/tim-berners-lee/
10 Tim Berners-Lee - Wikipedia 2 facts 2010-10-02 http://hi.wikipedia.org/wiki/Tim_Berners-Lee
11 tim berners lee 2 facts 2010-11-19 http://trakin/Tags/Business/tim-berners-lee/
12 Tim Berners Lee 3 facts 2010-11-23 http://blog.joeadrieu.com/tag/tim-berners-lee/
13 Tim Berners-Lee Biograph... 13 facts 2010-10-02 http://en.wikipedia.org/wiki/Tim_Berners-Lee

66

Naver Semantic Movie Search

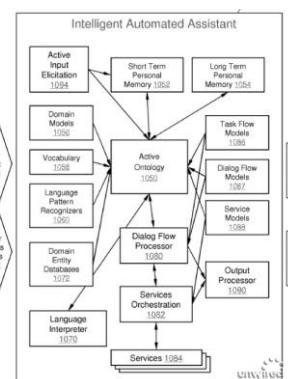
The screenshot shows the Naver LAB Semantic Movie Search interface. At the top, there's a search bar with the query "현빈 출연 영화" (Movies starring Hyunbin) and a "검색" (Search) button. Below the search bar, there are two sections: "추천 확장카워드" (Recommended expansion keywords) and "추천 연관관계" (Recommended relationships). The "추천 확장카워드" section lists terms like "현빈 조연영화 영화 감독", "현빈 조연영화 영화 조연배우", and "현빈 조연영화 영화 제작진정보". The "추천 연관관계" section lists terms like "현빈 조연영화 영화 주연배우", "현빈 조연영화 영화 단역배우", and "현빈 조연영화 영화 수상정보". Below these sections, there are two main movie lists. The first list is titled "▣ 현빈 주연 영화 5" and includes five movies: "만추" (Director: Kim Taewoo, 2010, Rating: 5.5), "나는 행복합니다" (Director: Yoon Jong-chang, 2008, Rating: 6.4), "백만장자의 첫사랑" (Director: Kim Tae-gon, 2006, Rating: 7.5), "돌려차기" (Director: Nam Sang-uk, 2004, Rating: 7.2), and "사위" (Director: Kim Jin-hyuk, 2002, Rating: 10). The second list is titled "▣ 현빈 조연 영화 1" and includes one movie: "기다리 아저씨" (Director: Gong Sung-sik, 2005, Rating: 7.2). Each movie entry includes a thumbnail, title, director, year, rating, and a "화장검색 + 더보기" (Search + View more) link.

67

Apple's Siri

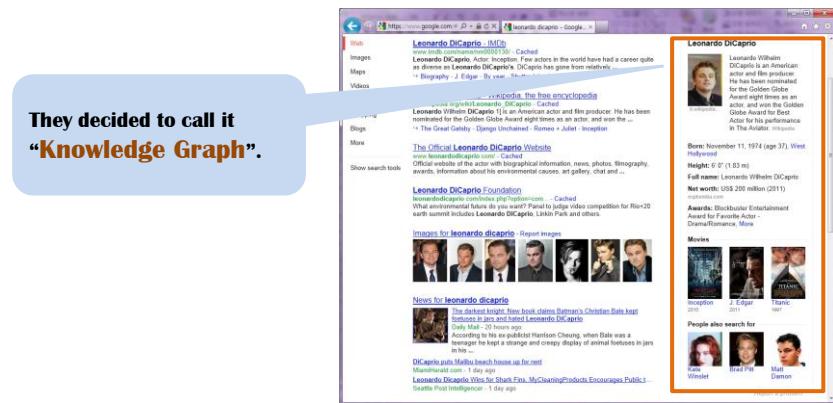
- an intelligent personal assistant and knowledge navigator which works as an application for Apple's iOS
- a natural language user interface to answer questions, make recommendations, and perform actions by delegating requests to a set of Web services

Siri's knowledge is represented in a unified modeling system that combines ontologies, inference networks, pattern matching agents, dictionaries, and dialog models. ... Siri isn't a source of data, so it doesn't expose data using Semantic Web standards.



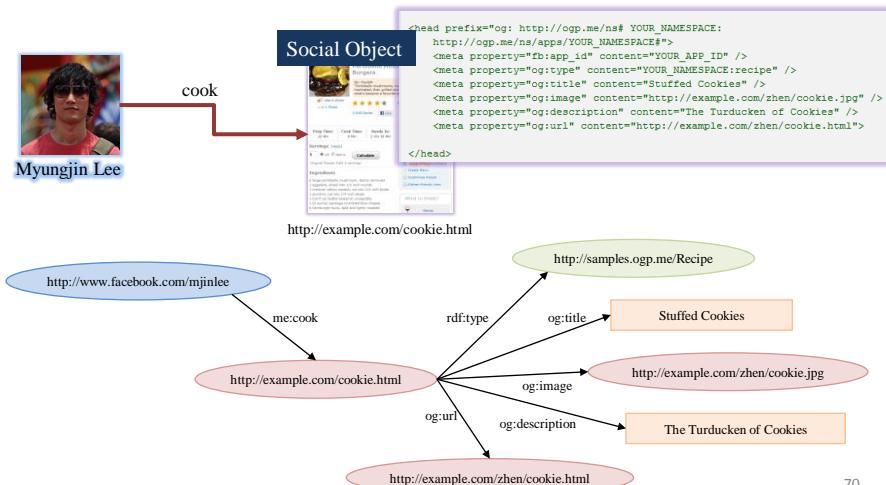
Google's Knowledge Graph

- A knowledge base used by Google to enhance its search engine's search results with semantic-search information gathered from a wide variety of sources (schema.org)
- over 570 million objects and more than 18 billion facts about and relationships between these different objects



Facebook's Open Graph Protocol

- simple protocol for enabling any web page to become a rich object in a social graph



Twitter Annotations

- Add one or more annotations that represent structured metadata about the tweet

First element is a type.

Every Annotations has a type.

Type maps to attribute and value pair.

Second element is one or more attribute names with values.

<http://r.github.com/annotationsformatter/>

The screenshot shows a web-based tool for generating annotations from a tweet. At the top, there's a text input field containing a tweet: "Check out @twitterapi's new developer portal!". Below it, there's a section for defining annotations. A dropdown menu labeled "Annotation type" is set to "website". Underneath, there are two input fields: "Attribute" with the value "url" and "Value" with the value "http://dev.twitter.com". A large teal arrow points from this annotation definition down to a code block below. The code block contains a single JSON object:

```
[{"website": {"url": "http://dev.twitter.com"}}]
```

71

Linking Open Data Applications

72

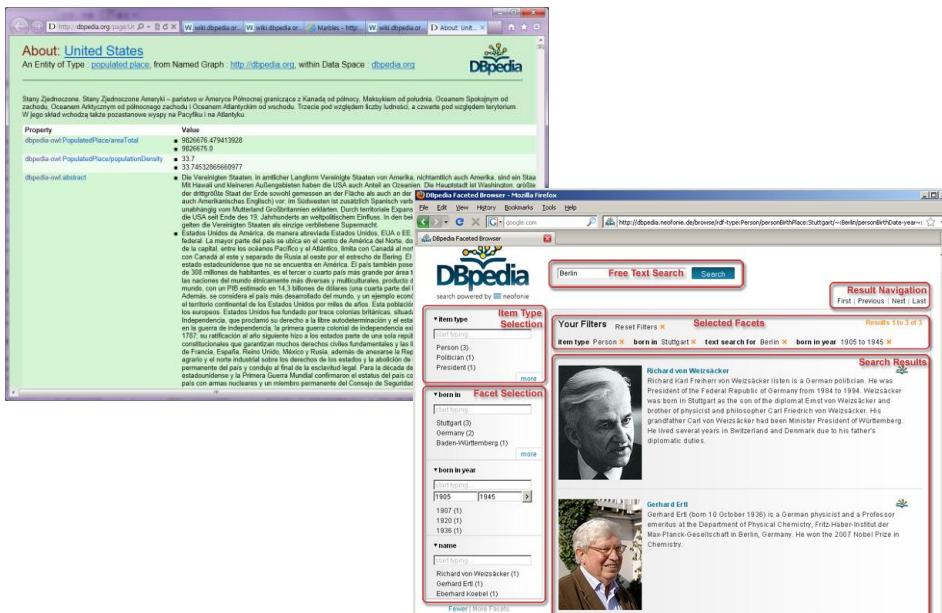


DBpedia

- A project aiming to extract structured content from Wikipedia using the Resource Description Framework (RDF) to represent the extracted information
- More than 3.64 million things, out of which 1.83 million are classified in a consistent ontology
- 2,724,000 links to images and 6,300,000 links to external web pages
- Over 1 billion pieces of information (RDF triples)

73

DBpedia



The screenshot displays two separate search results pages from the DBpedia website.

Search Result for 'United States':

- About:** United States
- An Entity of Type: populated place, from Named Graph: http://dbpedia.org, within Data Space: dbpedia.org
- Property:**
 - dbpedia-owl:PopulatedPlace/areaTotal: Value: 986978.479413928
 - dbpedia-owl:PopulatedPlace/populationDensity: Value: 33.7
 - dbpedia-owl:PopulatedPlace/populationDensity: Value: 1743285660077
- Description:** Die Vereinigten Staaten, in amtlicher Langform Vereinigte Staaten von Amerika, räumlich auch Amerika, sind ein Staat Mit Hauptverwaltungssitz Washington, D.C., und Hauptstadt Washington, geografisch der größte Staat Nordamerikas. Sie sind ein Mischstaat aus einer Union von 50 Bundesstaaten und einer unabhängigen Hauptstadt. Auch die USA sind eine Mischung aus englischer Sprache (auch Amerikanisches Englisch) vor. In Südosten ist zusätzlich Spanisch weit verbreitet. Der Name United States ist eine Übersetzung des englischen Namens United States of America. Der Name United States ist Ende des 19. Jahrhunderts ein wettbewerblicher Einfluss. In den bei

Search Result for 'Berlin':

- Item Type Selection:** Item type: Person (3)
- Selected Facets:** Selected Facets: born in Berlin (3), text search for Berlin (3), born in year 1905 to 1945 (3)
- Result Navigation:** First | Previous | Next | Last
- Search Results:**
 - Richard von Weizsäcker:** Richard von Weizsäcker ist ein deutscher Politiker. Er war Präsident der Federal Republic of Germany from 1984 to 1988. Weizsäcker was born in Stuttgart as the son of the diplomat Ernst von Weizsäcker and brother of physicist and philosopher Carl Friedrich von Weizsäcker. His grandfather Carl von Weizsäcker had been Minister President of Württemberg. He lived his last years in Switzerland and Denmark due to his father's diplomatic duties.
 - Gerd Ehr:** Gerd Ehr (born 10 October 1938) is a German physicist and a Professor emeritus at the Department of Physical Chemistry, Fritz-Haber-Institut der Max-Planck-Gesellschaft in Berlin, Germany. He won the 2007 Nobel Prize in Chemistry.

Linked Data on BBC

Data from Wikipedia

Data from MusicBrainz

Data in RDF

```

<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://purl.org/ontology/stuff/#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:mo="http://purl.org/ontology/mo/"
  xmlns:core="http://www.w3.org/2002/07/owl#relationship"
  xmlns:geo="http://opendatacommons.org/terms/"; xmlns:mo="http://purl.org/ontology/mo/"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
<mo:Artist rdf:about="http://musicbrainz.org/artist/650e7db6-b795-4eb5-a702-5ea2fc46c48d#artist">
  <rdf:type rdf:resource="http://purl.org/ontology/mo/SoloMusicArtist"/>
  <dc:subject>Lady Gaga</dc:subject>
  <dc:sortLabel>Lady Gaga</dc:sortLabel>
  <bio:events>
    <bio:birth>
      <dc:date datatype="http://www.w3.org/2001/XMLSchema#dateTime">1986-03-28</dc:date>
    </bio:birth>
  </bio:events>
  <owl:sameAs rdf:resource="http://dbpedia.org/resource/Lady_gaga"/>
  <mo:musicbrainz rdf:resource="http://musicbrainz.org/artist/650e7db6-b795-4eb5-a702-5ea2fc46c48d.html"/>
  <mo:resource rdf:resource="http://static.bbco.uk/music/images/artists/234x132/650e7db6-b795-4eb5-a702-5ea2fc46c48d.jpg"/>
  <rdf:homepage rdf:resource="http://www.ladygaga.com/"/>
  <mo:fanpage rdf:resource="http://ladygaga.wikia.com/"/>
  <mo:wikidata rdf:resource="http://en.wikipedia.org/wiki/Lady_Gaga"/>

```

Best Buy with GoodRelations

Best Buy - The Colony

STORE HOURS

SUN	MON	TUE	WED	THU	FRI	SAT
11:00a	10:00a	10:00a	10:00a	10:00a	10:00p	10:00a
8:00p	9:00p	9:00p	9:00p	9:00p	10:00p	10:00p

STORE SERVICES

- Geek Squad
- Best Buy Mobile
- Best Buy For Business
- Electronics Recycling
- Hablamos Español
- Hardware Trade-In
- Video Game Software Trade-In
- Best Buy Credit Card
- Car & GPS Installation Services

SPECIAL PRODUCTS

- Open Box Items
- Clearance Items
- TRADE-IN
- Special Products

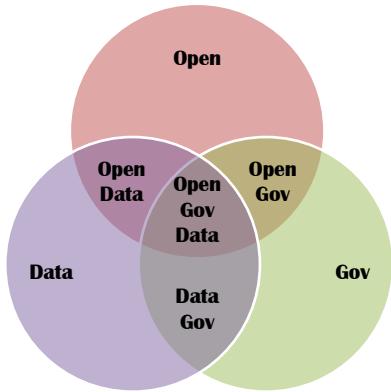
```

<div class="vcard" typeof="gr:LocationOfSalesOrServiceProvisioning" about="#store_1796">
<div class="hours" rel="gr:hasOpeningHoursSpecification">
<li class="day0" typeof="gr:OpeningHoursSpecification" about="#storehours_sun">
<span rel="gr:hasOpeningHoursDayOfWeek" resource="http://purl.org/goodrelations/v1#Sunday" class="day">
<span property="gr:opens" datatype="xsd:time" content="11:00:00" class="open">
...

```

Open Government Data

- By “open”, “open” data is free for anyone to use, re-use and re-distribute.
- By “**government data**” we mean data and information produced or commissioned by government or government controlled entities.



77

Data.gov (the United States Government)

An Official Website of the United States Government

Tuesday, June 12, 2012 | Text: A+ A- A | Share □

DATA.GOV
EMPOWERING PEOPLE

HOME DATA An Official Web Site of the United States Government

HOME INTERACTIVE DATASETS RAW DATA APPS GEO DATA COMMUNITY METRICS OPEN DATA SITES GALLERY WHAT'S NEW

Central Contractor Registration (CCR) FOIA Extract
This dataset lists all government contractors previously available under FOIA.

			12142001	10202011	
1	1ZB52	A	7,112,002	9,152,011	DWIGHT MATTHEWS AIA AF
2	4TEZ2	A	11,072,005	11,102,011	SPINELLI RAVIOLI MFG CO
3	1VV19	A	4,122,002	5,072,011	J C R CORP
4	3X4W6	A	7,072,004	6,292,011	H LOEB CORPORATION
5	33801	A	12,051,999	1,262,012	PRECISION SYSTEMS INC
6	1D046	A	5,212,001	1,252,012	BOMAS MACHINERY SPECIAL
7	1XBK5	A	4,292,002	10,012,011	OPTOMETRICS CORPORAT
8	94689	A	6,192,001	12,072,011	DAMPNEY COMPANY, INC.
9	9P408	A	3,152,002	2,022,012	NORTH ATLANTIC FISH CO.
10	1BK05	A	2,222,002	2,022,012	TRIANGLE ENGINEERING, I
11	535D7	A	5,192,008	10,062,011	METHOW VALLEY INDUSTR
12	5UY37	A	1,142,010	1,192,012	MUELLER CORPORATION
13	91480	A	1,142,002	5,262,011	AMES SAFETY ENVELOPE C
14	1CT45	A	1,082,002	11,162,011	STERLINGWEAR OF BOSTC
15	4HTW1	A	8,222,006	5,302,011	TRAIL MFG INCORPORATE
16	94633	A	3,152,002	3,012,012	FRANKLIN SPORTS, INC.
17	01456	A	3,082,002	11,172,011	MICROWAVE DEVELOPMEN

Export

API
Print
Download

Download a copy of this dataset in a static format

Download As:

CSV
JSON
PDF
RDF
RSS
XLS
XLSX
XML

Done

DATA.GOV

Home | About | FAQ | Contact Info | Data Policy | Accessibility | Privacy Policy | Sitemap

78

Data.gov.uk (HM Government)

The screenshot shows the Data.gov.uk homepage for HM Government. At the top, there's a navigation bar with links for Home, Data, Apps, Consultation, Forum, Blogs, Ideas, Linked Data, Resources, and About. A search bar is located at the top right. Below the navigation, there's a section titled "Opening up government" featuring several cards:

- Count of Gypsy and Traveller Caravans** [Dataset] by Department for Communities and Local Government. Description: The bi-annual Count of Gypsy and Traveller Caravans takes place twice a year and records the number of caravans on both authorised and unauthorised ...
- Communities and Local Government group spending data** [Dataset] by Department for Communities and Local Government. Description: Department for Communities and Local Government Department for Communities and Local Government Department for Communities and Local ...
- Spend over £500 in Homes and Communities Agency** [Dataset] by Department for Communities and Local Government. Description: Department for Communities and Local Government Homes and Communities Agency The Homes and Communities Agency is fully committed to the ...
- Code for Sustainable Homes and Energy Performance of Buildings** [Dataset] by Department for Communities and Local Government. Description: Department for Communities and Local Government Statistics on the number of dwellings that have been certified to the standards set out in the Code Technical Guide. In which local authority area, at ...
- Expenses data in Audit Commission** [Dataset] by Department for Communities and Local Government. Description: Department for Communities and Local Government Audit Commission The Audit Commission's Board and management team are fully committed to the ...

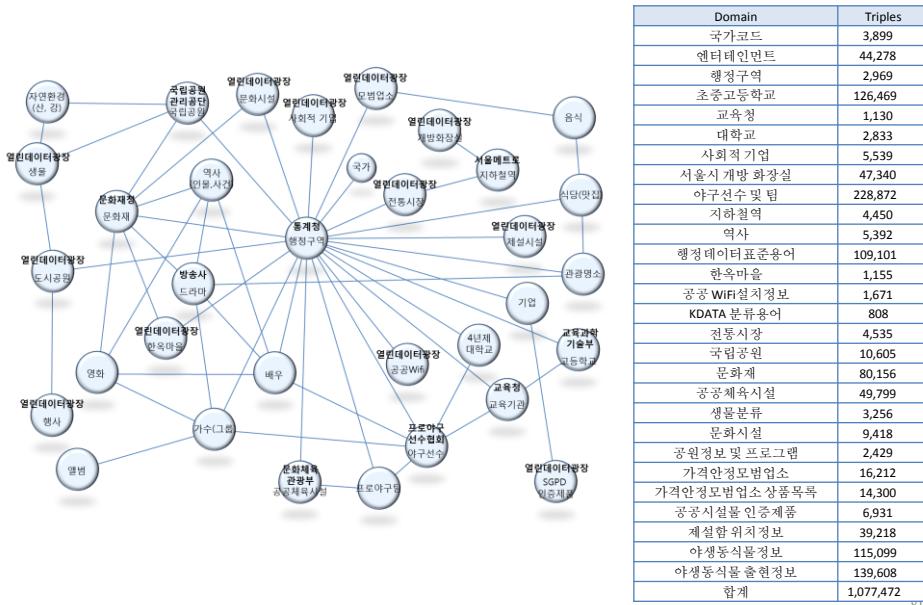
On the right side, there are three filter panels:

- Map Based Search**: Options include "Conduct Map Based Search".
- Filter by type**: Options include "Dataset (768)".
- Filter by resource format**: Options include "XLS (460)", "CSV (46)", "PDF (8)", "RDF (7)", "Csv (6)", "XML (5)", "Shapefile (4)", "HTML (3)", "TIFF LZW (3)", and "DXF (2)".

Data-Gov Wiki

- A project for investigating open government datasets using semantic web technologies
 - 417 RDFized datasets covering the content of 703 out of 5762 datasets with 6.46 billion RDF triples.
 - Additional RDF-ized datasets including 35 Non-Data.gov Datasets with 0.9 billion more RDF triples.
- http://data-gov.tw.rpi.edu/wiki/The_Data-gov_Wiki

KDATA (Linked Data for Korea)



81

References

- <http://en.wikipedia.org/wiki/Internet>
- http://en.wikipedia.org/wiki/Computer_network
- http://en.wikipedia.org/wiki/World_Wide_Web
- <http://www.slideshare.net/lysander07/openhpi-11>
- <http://en.wikipedia.org/wiki/HTML>
- <http://www.google.com/insidesearch/howsearchworks/thetheory/>
- <http://www.go-gulf.com/blog/60-seconds/>
- <http://www.slideshare.net/lysander07/openhpi-15>
- <http://www.w3.org/DesignIssues/Semantic.html>
- http://en.wikipedia.org/wiki/Semantic_web
- <http://www.slideshare.net/lysander07/openhpi-13>
- <http://www.w3.org/2001/sw/>
- Tim Berners-Lee, James Hendler, and Ora Lassila, "The Semantic Web", Scientific American Magazine, 2001.
- [http://www.w3.org/2007/Talks/0130-sb-W3CTechSemWeb/#\(24\)](http://www.w3.org/2007/Talks/0130-sb-W3CTechSemWeb/#(24))
- <http://www.slideshare.net/onlyjiny/semantic-web-13288556>
- <http://www.slideshare.net/onlyjiny/linked-open-government-data-15708234>
- <http://www.slideshare.net/onlyjiny/linkeddata>
- <http://www.slideshare.net/sonagi/ss-16734202>
- <http://www.slideshare.net/lysander07/13-semantic-web-technologies-linked-data-semantic-search>
- <http://kdata.kr/index.jsp>
- <http://linkeddata.org/>
- <http://lod-cloud.net/>

82

Individual Task 1: Literature Review

1. Read “initial Papers” on the Semantic Web:
 - (i) [The Semantic Web](#) by Tim Berners-Lee, Ora Lassila and James Hendler, Scientific American
 - http://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf
 - (ii) [The Semantic Web Revisited](#) by Nigel Shadbolt, Wendy Hall and Tim Berners-Lee
 - http://eprints.ecs.soton.ac.uk/12614/1/Semantic_Web_Revisited.pdf
 - (iii) [Linked Data](#) by Tim Berners-Lee
 - <http://www.w3.org/DesignIssues/LinkedData.html>
- **2. Bring 5 bullet points (even if just questions) about each paper to the lecture on Thursday and be prepared to discuss with your peers!**

©

Declan O'Sullivan

Individual Task 2: Selecting a Project Title

1. Search for Semantic Web applications and read/research topics that you like to work on:
 - Semantic Search
 - Semantic Mobile Web Applications
 - Social media analysis and visualization
 - Intelligent User interfaces in a domain
 - Knowledge extraction
 - Contributing to linked data
 - Linked data applications that use existing knowledge
 -
 - While selecting a topic, think if you can contribute to the field (add something [new/original](#)), which improves the state of the art in the field).
 - **AA or BA will be guaranteed for those who perform a project that is publishable in an international conference.**
 - **Write one page proposal about your project and send it to melike.sah@neu.edu.tr by 19 March 2015 for approval!!!**

©

Declan O'Sullivan