COM556
SEMANTIC WEB TECHNOLOGIES

Week 1
Semantic Web Vision and Introduction

Assist. Prof. Dr. Melike Şah Direkoğlu

Outline

• Semantic Web and Semantic Web Vision

• Semantic Web Technologies

• Semantic Web Case Studies

2

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

Network
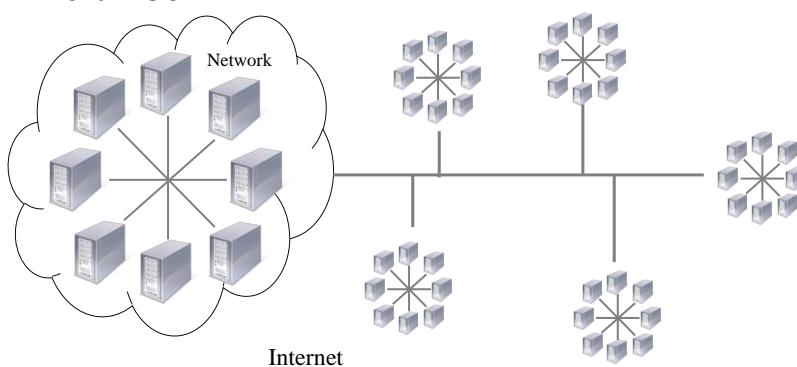
Internet

4
[Myungjin Lee]

Internet Services before the Web

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

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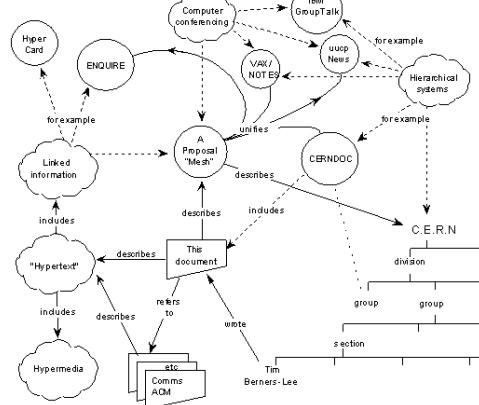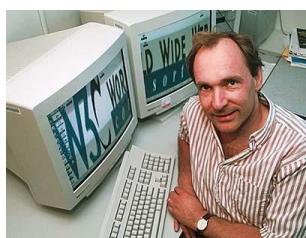
Proposal of "Hypertext project" called "World Wide Web"

6
[Myungjin Lee]

3

# Characteristics of Web

- Hyperlink and Multimedia

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

7
[Myungjin Lee]

Web Architecture

the main markup language for displaying web pages and other information that can be displayed in an web browser

HTML

Document

HTTP

an application protocol for distributed, collaborative, hypermedia information systems

Protocol

Identifier

URI

a string of characters used to identify a name or a resource
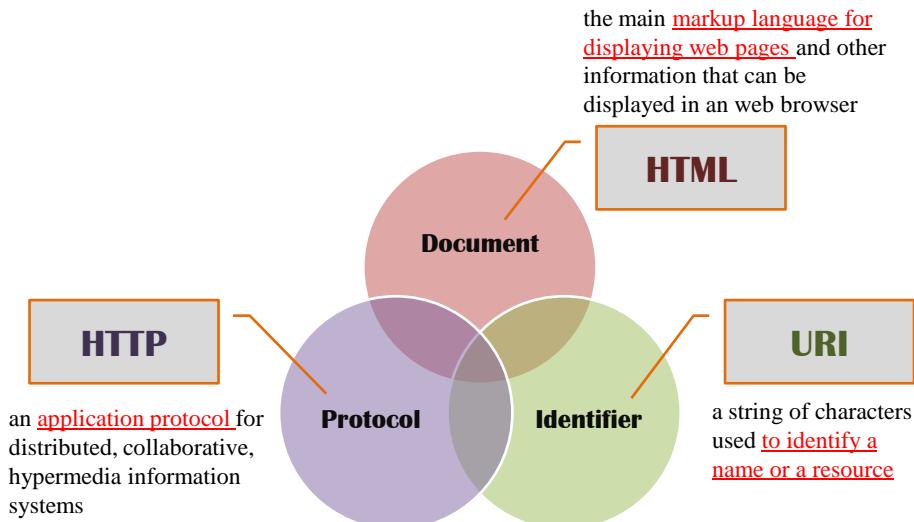
[Myungjin Lee] 8

Problem of HTML

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

AutoTrader.com

same information, but …

Cars.com

10
[Myungjin Lee]

Problem of HTML

Audi A6

Maserati A6

A6 Metrobus Lines

A6 Paper Size

Apple A6

[Myungjin Lee]

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!
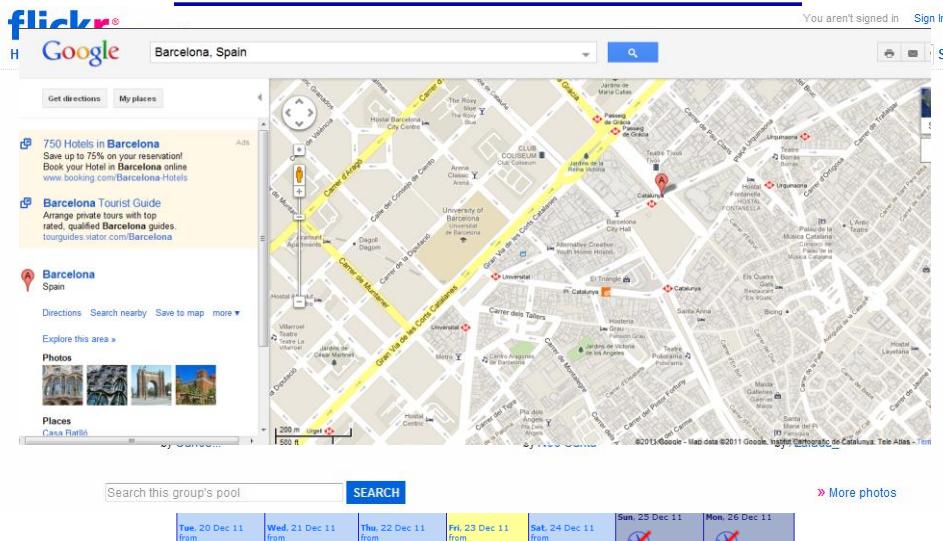
What happened?

- 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 <u>link data</u> (independent of their presentation) and use the data the way I want
- Agents, programs, scripts, etc. should be able to <u>interpret</u> part of that data
- But wait, representation of the data and access to that data should be <u>standardized</u> 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)

© Melike Sah

What we want on the Web?

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

> 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 <u>Web of Data</u> and <u>machine-processable Web</u>

- Web → **"links document to document", "documents to READ"**

- Semantic Web → **"links data to data", "data for all sorts of things"**

© Melike Sah

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 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 interoperation:
  - Understanding what the data means
  - Linking in insightful ways
  - Automated support for data integration
  - Smart applications

  - *Sharing data ⇒ 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
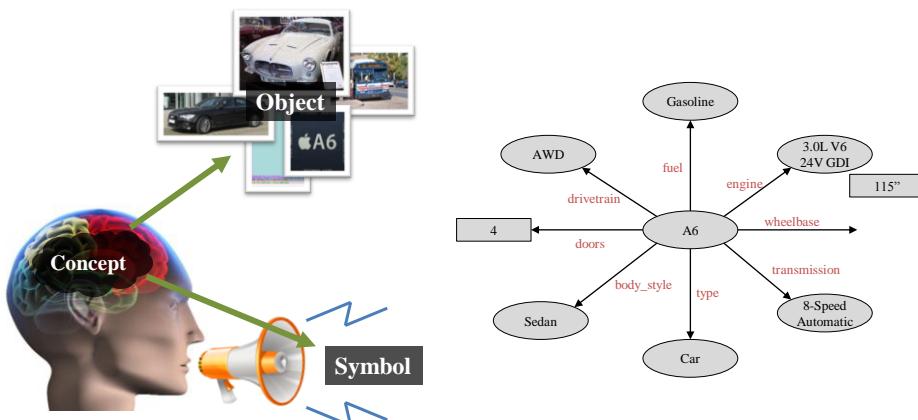

24
[Myungjin Lee]

Overview of the Semantic Web

- What is the Semantic Web?

- **Semantic Web Technologies**

- Semantic Web Case Studies

25

Semantic Web Layer Cake
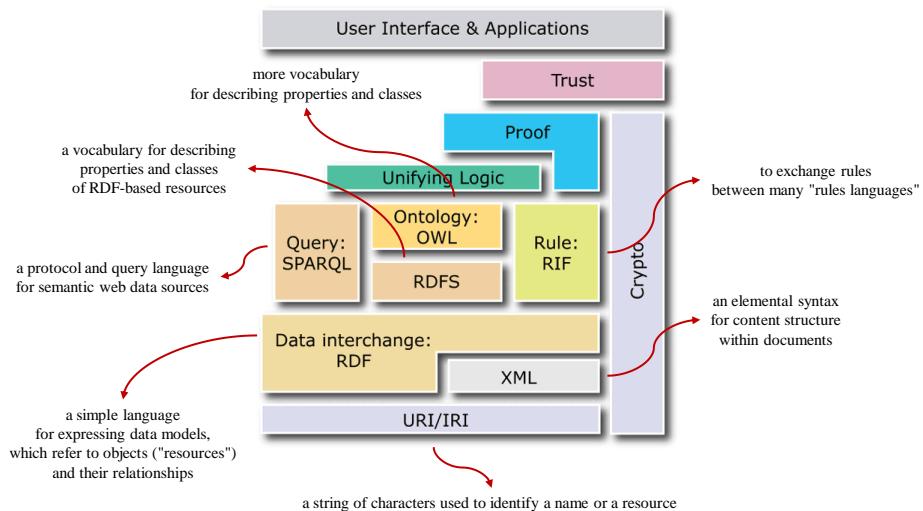
[Myungjin Lee]

26

URI (Uniform Resource Identifier)

- a string of characters used <u>to identify a name or a resource</u>

URI

URN (Uniform Resource Name)	**+**	URL (Uniform Resource Locator)

urn:isbn:0451450523
urn:isan:0000-0000-9E59-0000-O-0000-0000-2
urn:issn:0167-6423

ftp://asmith@ftp.example.org
http://en.example.org/wiki/url
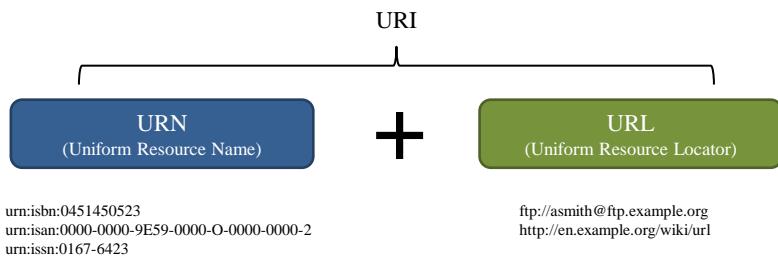
27
[Myungjin Lee]

XML (Extensible Markup Language)

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

```
<?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.)
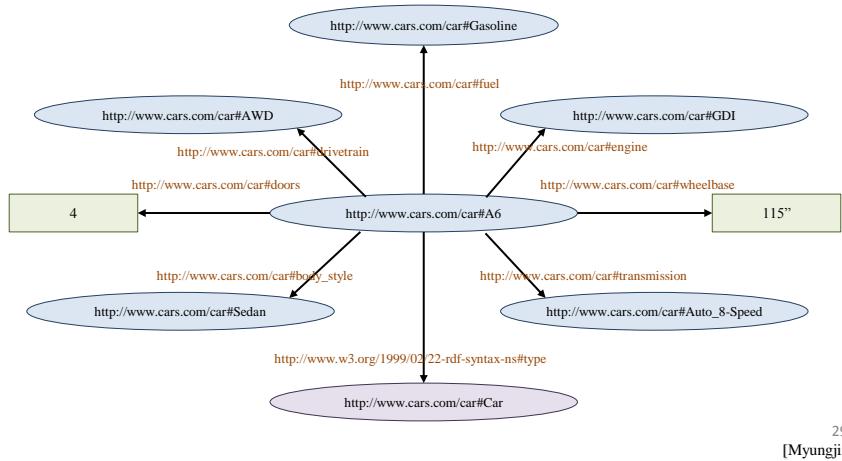
29
[Myungjin Lee]

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
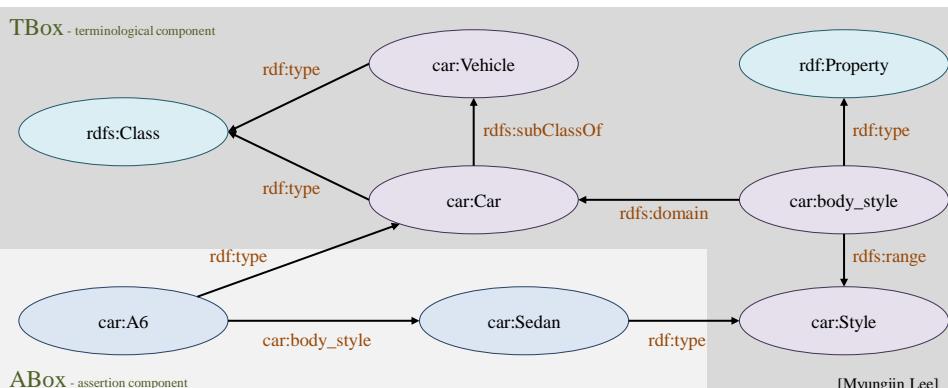
[Myungjin Lee]

Ontology

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

> "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 <u>knowledge representation languages</u> 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

- <u>SWRL (Semantic Web Rule Language)</u> is a proposal for a Semantic Web <u>rules-language</u>, 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 <u>derive new data from data</u> that you already know

hasWife

hasParent   hasParent

```
if      hasParent(?x, ?y)
        hasParent(?x, ?z)
        Man(?y)
        Woman(?z)


then    hasWife(?y, ?z)
```
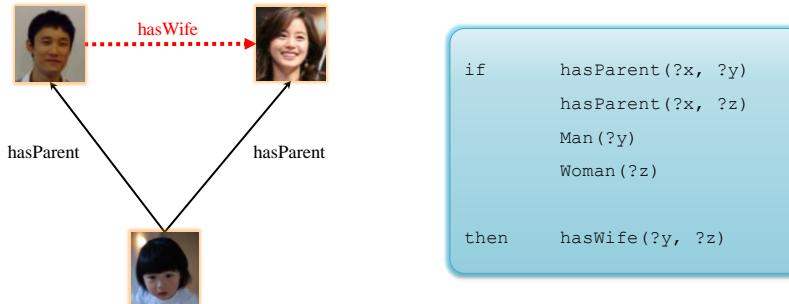
34
[Myungjin Lee]

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

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?name ?email
WHERE {
    ?person a foaf:Person.
    ?person foaf:name ?name.
    ?person foaf:mbox ?email.
}
```

RDF Knowledge Base

?name	?email
Myungjin Lee	mjlee@li-st.com
Gildong Hong	gildong@daum.net
Grace Byun	grace@naver.com

35
[Myungjin Lee]
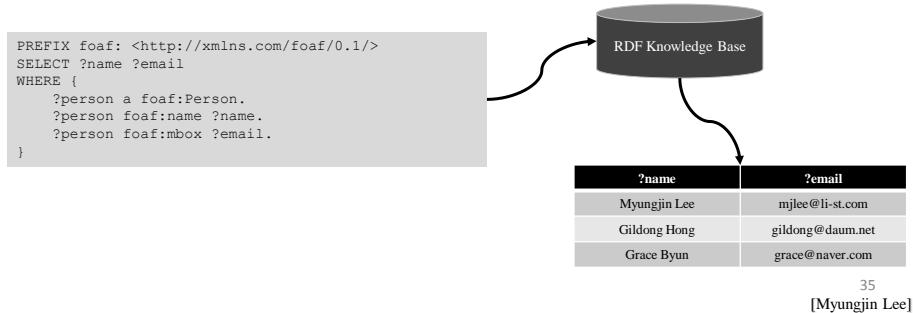
Ontologies

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

Weak Semantics

Can be arbitrary
- Taxonomy
  - "is subclassification of"
Relational Model

Term Semantic relations used: Synonyms, homonyms, narrower meaning, broader meaning
- Thesaurus
  - "x is a homonym of y, e.g. tank"
ER

Concepts, Properties, Relationships, Rules
- Conceptual Model
  - "Is subclass of"
RDF/S, XTM

Axioms (range of statements asserted to be true) +
- Logical Theory
  - "Is disjoint subclass of with transitivity property"
UML

Inference Rules (rules that given assumptions provide valid conclusions)
OWL
Description Logic

First Order Logic

Strong Semantics

[Daconta 2003]

© Declan O`Sullivan

19

The cost of semantic clarity
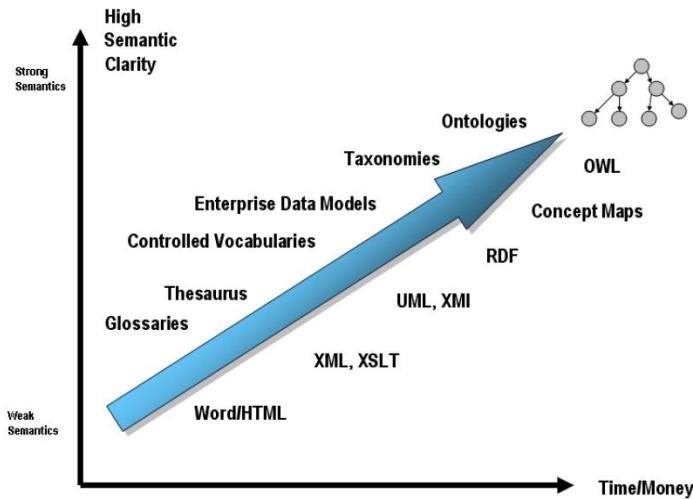
[Marin Dimitrov, 3rd GATE tutorial, 2010]

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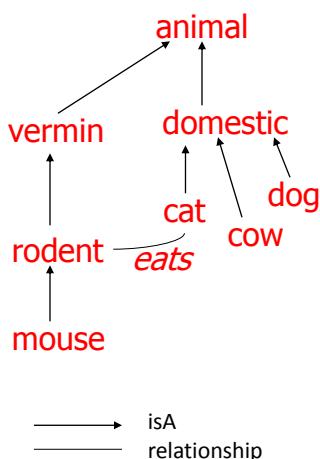
[Carole Goble, Nigel Shadbolt, Ontologies and the Grid Tutorial]

An explicit description of a domain

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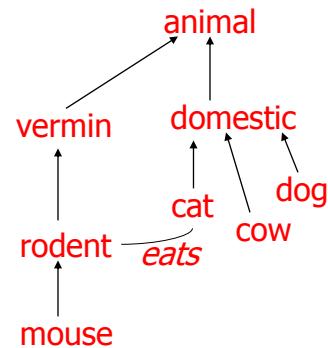
animal

vermin   domestic

rodent   cat   dog
        eats   cow

mouse

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

An explicit description of a domain

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

Knowledge Base

Concepts, Relationships, Axioms

Individuals

animal

vermin   domestic

rodent   cat   dog
        eats   cow

mouse

felix   tom

mickey

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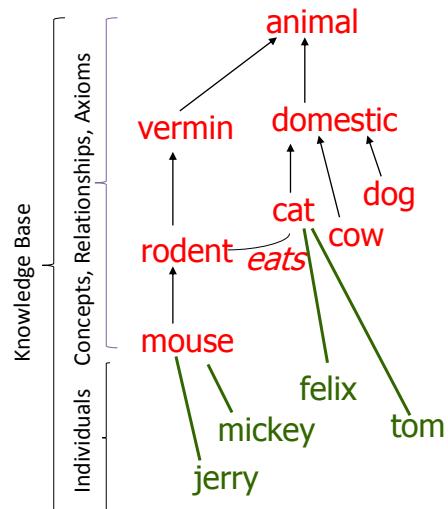
© Declan O'Sullivan   [Carole Goble, Nigel Shadbolt, Ontologies and the Grid Tutorial]

Ontology Specification Example
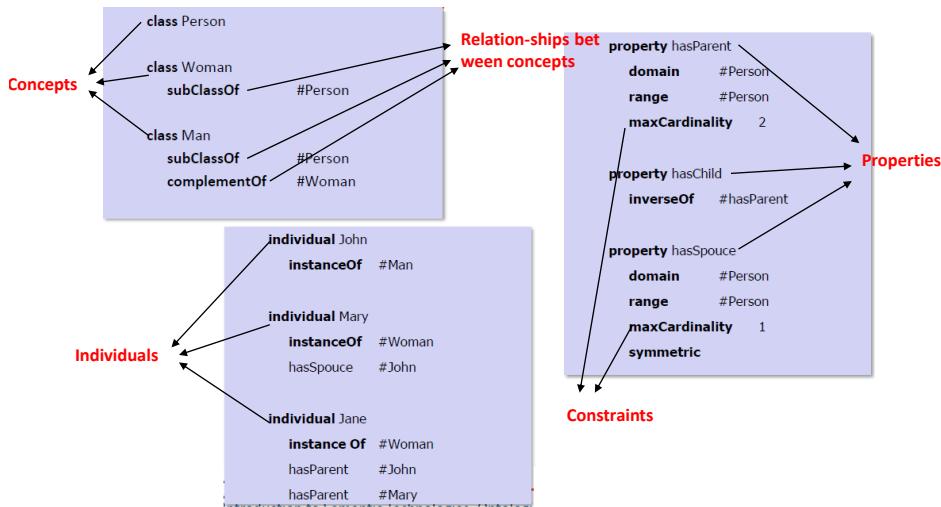
Semantic Web Standards

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 <u>data model</u> for
  - Formally describing the <u>semantics</u> of information in a machine accessible way
  - Representing meta-data (data about data)

- <u>Semantics</u> = a way of encoding meaning (link between term and a model of the world) → Good for building applications
- <u>Syntax</u> = 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
- <u>Note!</u> 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
  - <u>Resources (with unique identifiers – URI</u> as a global namespace of identifiers of things)
    - Unique across entire WWW
  - <u>Literals</u>
  - Named <u>relations</u> 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 <u>or</u> a literal value
  - What about predicates?
- The RDF graph is a collection of triples
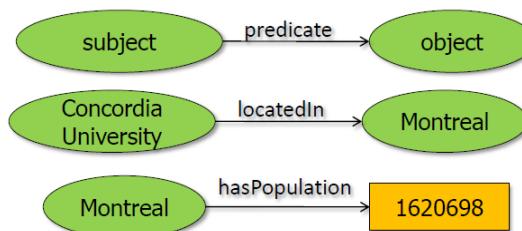
[Marin Dimitrov, 3rd GATE tutorial, 2010]

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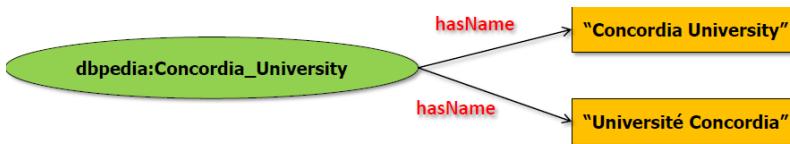
Subject	Predicate	Object
http://dbpedia.org/resource/Concordia_University	hasName	"Concordia University"
http://dbpedia.org/resource/Concordia_University	hasName	"Université Concordia"

[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)

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"

[Marin Dimitrov, 3rd GATE tutorial, 2010]

RDF Advantages

- <u>Simple but expressive</u> 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 <u>extension</u> of RDF
- RDFS provides mechanisms for describing groups of related resources and the <u>relationships</u> between these resources
- RDFS provides means for:
  - Defining <u>Classes and Properties</u> – rdfs:Class, rdfs:Property
  - Defining <u>hierarchies</u> (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.)
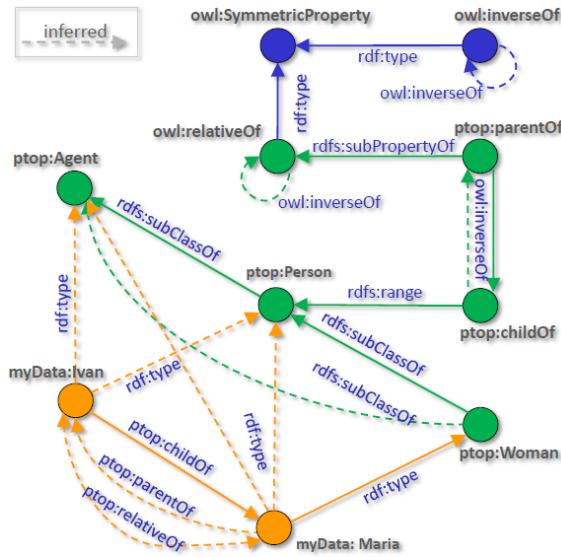
[Marin Dimitrov, 3rd GATE tutorial, 2010]

Web Ontology Language (OWL)

OWL

- <u>More expressive than RDFS</u>
  - Identity equivalence/difference
    - owl:sameAs, owl:differentFrom, owl:equivalentClass/Property
- <u>More expressive class definitions</u>
  - Class intersection, union, complement, disjointness , Cardinality restrictions
- <u>More expressive property definitions</u>
  - 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

- <u>SQL-like</u> query language for RDF data
- Simple protocol for querying remote databases over <u>HTTP</u>
- 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

```
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:University ;
            dbp-ont:numberOfStudents ?students ;
            dbp-ont:city dbpedia:Montreal .
  FILTER (?students > 5000)
}
ORDER BY DESC (?students)
```

- List of namespace prefixes
- List of variables
- Graph patterns
- Filters
- Modifiers
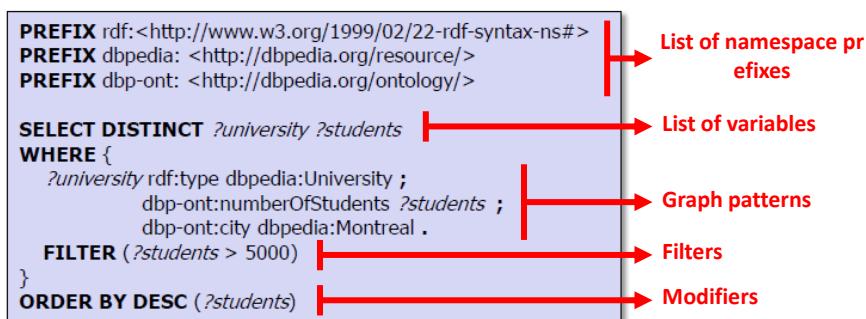
[Marin Dimitrov, 3rd GATE tutorial, 2010]

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
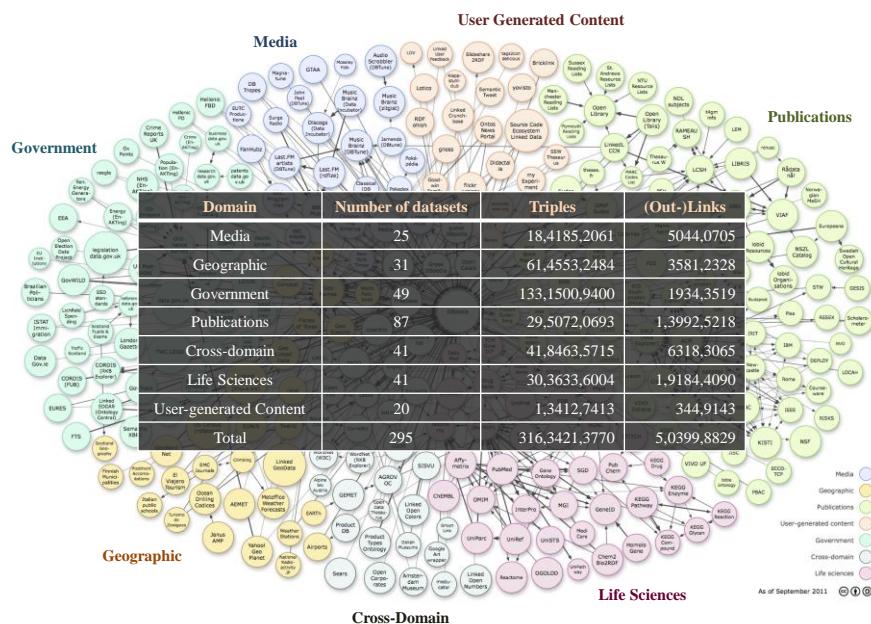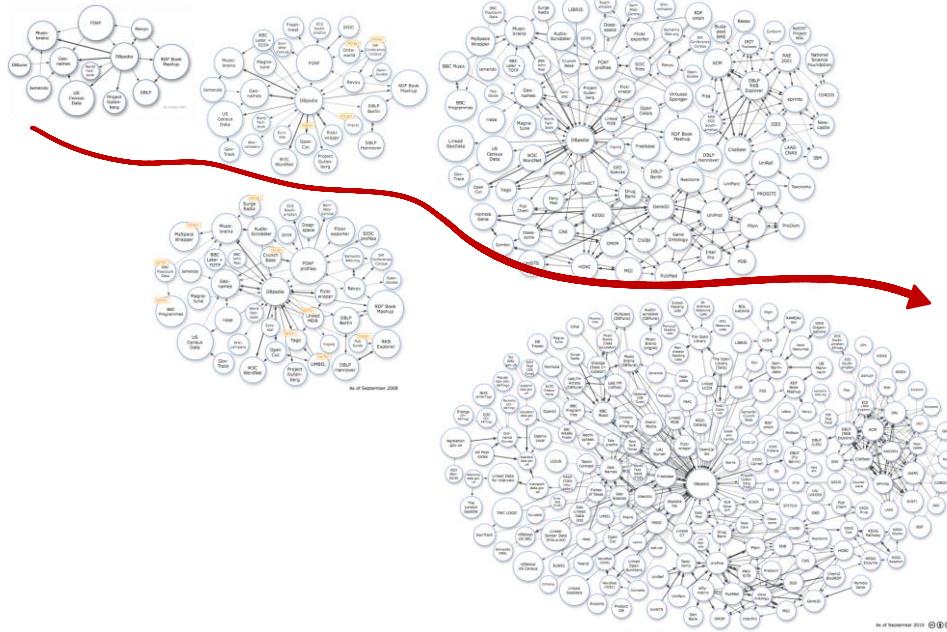
Domain	Number of datasets	Triples	(Out-)Links
Media	25	18,4185,2061	5044,0705
Geographic	31	61,4553,2484	3581,2328
Government	49	133,1500,9400	1934,3519
Publications	87	29,5072,0693	1,3992,5218
Cross-domain	41	41,8463,5715	6318,3065
Life Sciences	41	30,3633,6004	1,9184,4090
User-generated Content	20	1,3412,7413	344,9143
Total	295	316,3421,3770	5,0399,8829

Overview of the Semantic Web

- What is the Semantic Web?

- Semantic Web Technologies

- **Semantic Web Case Studies**

65

Sig.ma

66

Naver Semantic Movie Search

67

Apple's Siri

- an <u>intelligent personal assistant and knowledge navigator</u> 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.
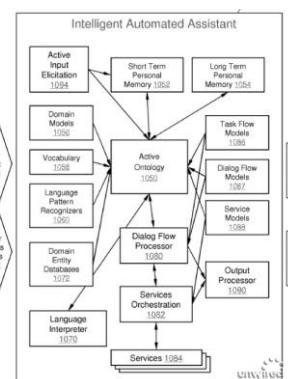
Google's Knowledge Graph

- A knowledge base used by Google to enhance its search engine's search results with <u>semantic-search information</u> 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

They decided to call it "**Knowledge Graph**".
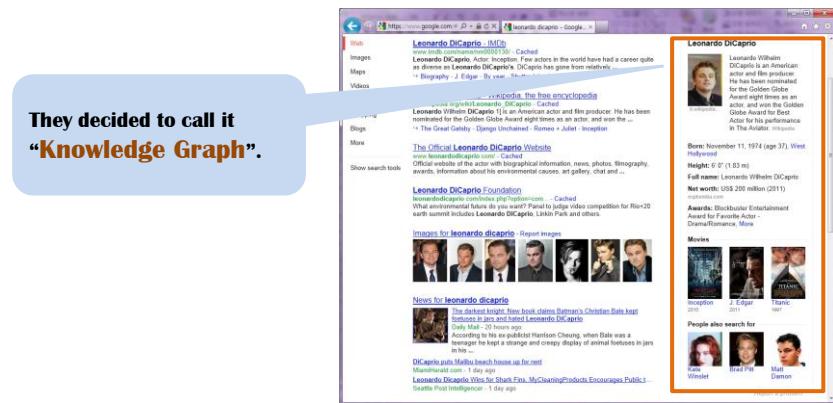
Facebook's Open Graph Protocol

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


```
<head prefix="og: http://ogp.me/ns# YOUR_NAMESPACE:
   http://ogp.me/ns/apps/YOUR_NAMESPACE#">
   <meta property="fb:app_id" content="YOUR_APP_ID" />
   <meta property="og:type" content="YOUR_NAMESPACE:recipe" />
   <meta property="og:title" content="Stuffed Cookies" />
   <meta property="og:image" content="http://example.com/zhen/cookie.jpg" />
   <meta property="og:description" content="The Turducken of Cookies" />
   <meta property="og:url" content="http://example.com/zhen/cookie.html">

</head>
```

Social Object

cook

Myungjin Lee

http://example.com/cookie.html

http://www.facebook.com/mjinlee

me:cook

rdf:type

og:title

og:image

og:description

og:url

http://example.com/cookie.html

http://samples.ogp.me/Recipe

Stuffed Cookies

http://example.com/zhen/cookie.jpg

The Turducken of Cookies

http://example.com/zhen/cookie.html

70

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/

Tweet text
Check out @twitterapi's new developer portal!
    Annotation type website
       Attribute url   **Value** http://dev.twitter.com

[{"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

Linked Data on BBC

Best Buy with GoodRalations


```
<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">
...
```

76

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.
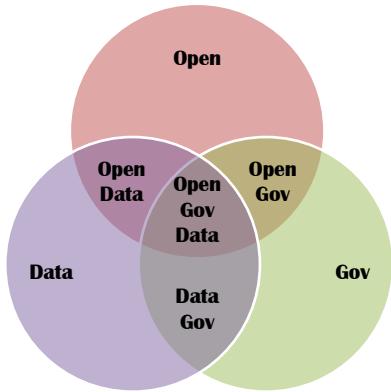
Data.gov (the United States Government)

Data.gov.uk (HM Government)

Data-Gov Wiki

- A project for investigating <u>open government datasets using semantic web</u> technologies
  - 417 RDFlized 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

80

KDATA (Linked Data for Korea)

Domain	Triples
국가코드	3,899
엔터테인먼트	44,278
행정구역	2,969
초중고등학교	126,469
교육청	1,130
대학교	2,833
사회적 기업	5,539
서울시 개방 화장실	47,340
야구선수 및 팀	228,872
지하철역	4,450
역사	5,392
행정 데이터표준용어	109,101
한옥마을	1,155
공공 WiFi설치정보	1,671
KDATA 분류용어	808
전통시장	4,535
국립공원	10,605
문화재	80,156
공공체육시설	49,799
생물분류	3,256
문화시설	9,418
공원정보 및 프로그램	2,429
가격안정모범업소	16,212
가격안정모범업소 상품목록	14,300
공공시설물 인증제품	6,931
제설함위치정보	39,218
야생동식물정보	115,099
야생동식물 출현정보	139,608
합계	1,077,472

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/thestory/
- 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.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_Revisted.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 vizualization
- 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 contibute 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