

# Semantics - I

Matthew J. Graham  
CACR

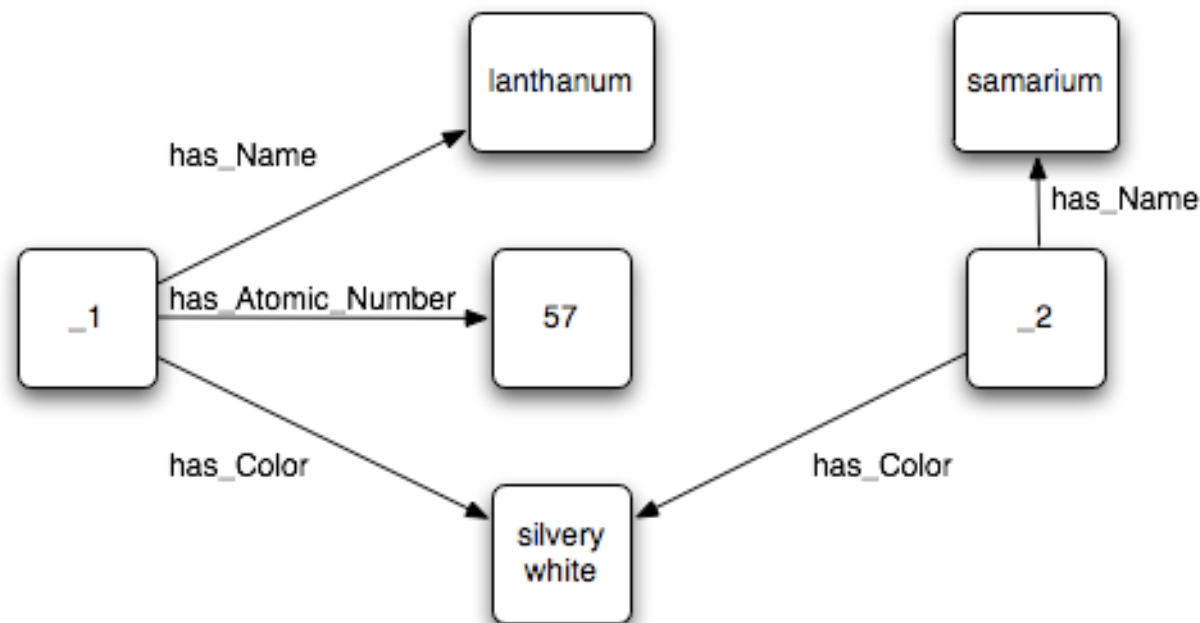
Methods of Computational Science  
Caltech, 2009 April 2

*matthew graham*

- The future of the Internet (Web 3.0)
- Decentralized platform for distributed knowledge
- A web of databases
- Logical pieces of meaning that can be mechanically manipulated by a machine
- Vocabularies for making assertions about things
- Data consistency and inferencing

# knowledge as a graph

Information is best expressed as a labelled, directed graph (entity-attribute-value data model)



# resource description framework

---

- W3C standard for encoding knowledge

- A fact is expressed as a *subject-predicate-object* triple or statement

- Subjects, predicates and objects are given as names for entities

- Names are URIs

- Objects can also be given as text values, called literal values, which may or may not be typed using XML Schema datatypes

- URIs are expressed as `<...>` or abbreviated with namespaces

- Statement is just written as subject URI, predicate URI and object URI/literal value with a period:

```
@prefix pt: <http://www.example.org/PeriodicTable#> .  
<http://www.example.org/PeriodicTable#La> pt:name "lanthanum" .
```

- Multiple statements with the same subject (and predicate) can be grouped together with a semicolon (commas)

```
<http://www.example.org/PeriodicTable#La> pt:name "lanthanum" ;  
    pt:atomicNumber 57 ; pt:color "silvery white" .
```

## XML-based representation

### Resource nodes:

- Statement subjects and objects
- Usually have `rdf:about` attribute
- Can only contain property nodes

### Property nodes:

- Literal values
- Reference to an object resource using `rdf:resource` attribute
- Resource node

## rdf/xml example

---

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:pt="http://www.example.org/PeriodicTable#">
  <rdf:Description rdf:about="http://www.example.org/PeriodicTable#La">
    <pt:name> lanthanum </pt:name>
    <pt:atomicNumber> 57 </pt:atomicNumber>
    <pt:color> silvery white </pt:color>
  </rdf:Description>
</rdf:RDF>
```

For comparison:

```
@prefix pt: <http://www.example.org/PeriodicTable#> .
<http://www.example.org/PeriodicTable#La> pt:name "lanthanum" ;
  pt:atomicNumber 57 ; pt:color "silvery white" .
```

*matthew graham*

---

Embed knowledge in an (X)HTML document

Uses attributes applicable to all elements:

- **about**
- **rel** and **rev**
- **href**, **src** and **resource**
- **property**
- **content**
- **datatype**
- **typeof**



# rdfa example

---

```
<html xmlns="http://www.w3.org/1999/xhtml"
      xmlns:pt="http://www.example.org/PeriodicTable#">
  <head>
    <title> Lanthanum </title>
    <base href="http://www.example.org/PeriodicTable#La" />
  </head>
  <body>
    <p>
      The element <span property="pt:name"> lanthanum </span> has an atomic
      weight of <span property="pt:atomicWeight"> 57 </span> and is
      <span property="pt:color"> silvery white </span> in color.
    </p>
  </body>
</html>
```

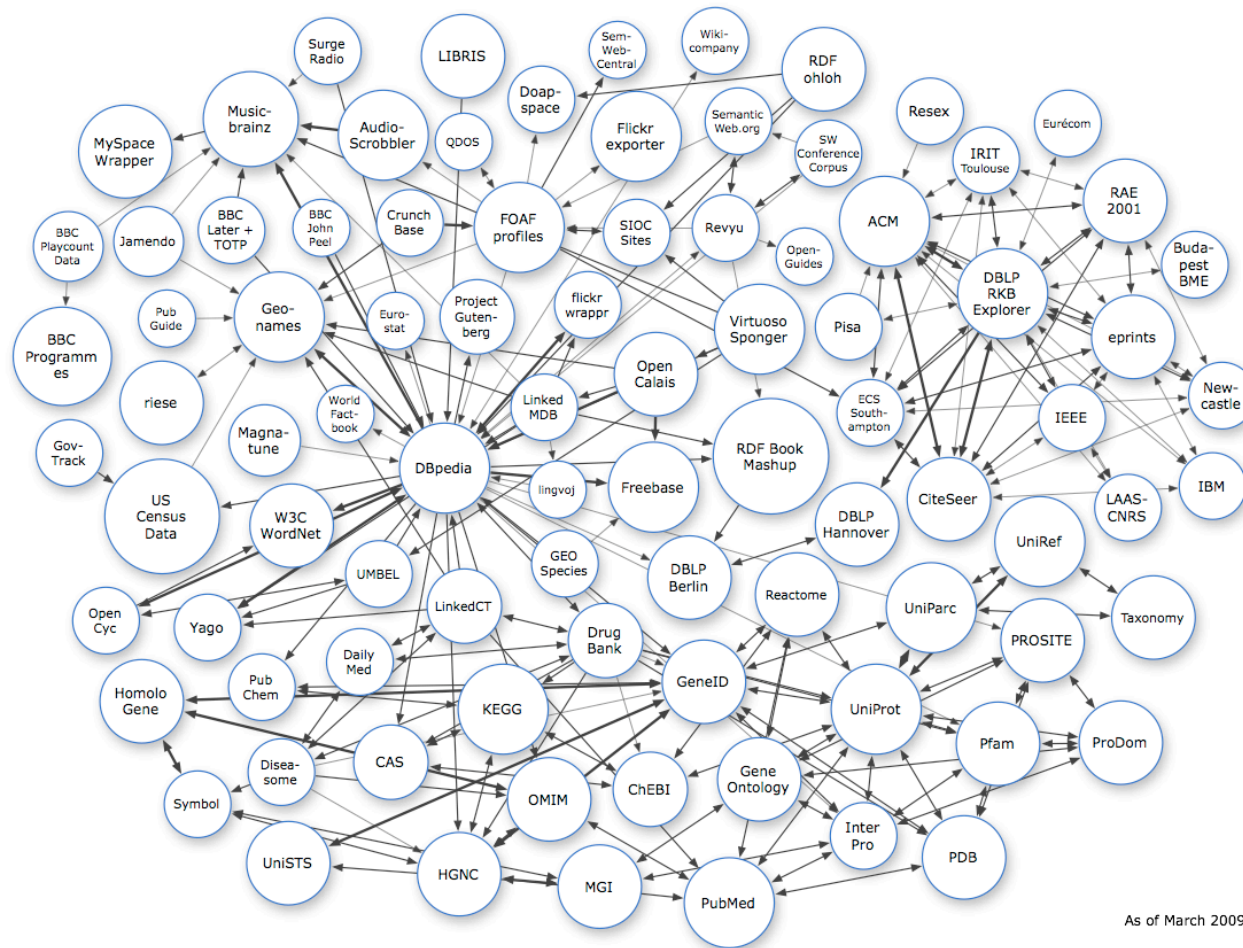
*matthew graham*

---

A term coined by Tim Berners-Lee who outlined four principles:

- Use URIs to identify things that you expose to the Web as resources
- Use HTTP URIs so that people can locate and look up (dereference) these things
- Provide useful information about the resource when its URI is dereferenced
- Include links to other, related URIs in the exposed data

# linked data web



# linked data example

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:dbpprop="http://dbpedia.org/property/"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:owl="http://www.w3.org/2002/07/owl#" >
  <rdf:Description rdf:about="http://dbpedia.org/resource/Lanthanum">
    <rdfs:label xml:lang="en">Lanthanum</rdfs:label>
    <dbpprop:abstract xml:lang="en">Lanthanum is a chemical element with the symbol La and atomic number
57.</dbpprop:abstract>
    <foaf:page rdf:resource="http://en.wikipedia.org/wiki/Lanthanum"/>
    <skos:subject rdf:resource="http://dbpedia.org/resource/Category:Chemical_elements"/>
    <skos:subject rdf:resource="http://dbpedia.org/resource/Category:Lanthanides"/>
    <rdf:type rdf:resource="http://sw.opencyc.org/2008/06/10/concept/Mx4rIGTaIPAIQdaffLzGWDo0Zw"/>
    <owl:sameAs rdf:resource="http://rdf.freebase.com/ns/guid.9202a8c04000641f80000000045bfbf8"/>
    <dbpprop:hasPhotoCollection
rdf:resource="http://www4.wiwiss.fu-berlin.de/flickrwrappr/photos//Lanthanum"/>
    <foaf:img rdf:resource="http://upload.wikimedia.org/wikipedia/commons/5/5c/Lanthan_1.jpg"/>
  </rdf:Description>
</rdf:RDF>
```

## linked data queries

---

- | List all episodes of the HBO television series "The Sopranos" ordered by their air-date

- | Find the official websites of companies with more than 50000 employees

- | Find me things close to the Eiffel Tower

- | Discover new drugs to treat Alzheimer's:

- What proteins are involved in signal transduction **and** are related to pyramidal neurons?

- \* Google: 223000 hits, 0 results

- \* Linked healthcare data: 32 hits, 32 results

## Browsers:

- Tabulator (also has Firefox plugin)
- Disco
- Openlink Data Browser

## Libraries

- Semantic Web Client Library (Java)

## Exposing data

- d2r for exposing relational data in a db

---

## concept schemes

- | controlled vocabulary: a closed list of terms that can be used for classification
- | taxonomy: a controlled vocabulary with hierarchy (term - sub-term)
- | thesaurus: a taxonomy with broader/narrower term, synonymous term, top term, scope note and related term
- | subject heading list, terminology, glossary, faceted classification

# \_\_\_\_\_simple knowledge organization system

■ W3C standard for expressing knowledge concept schemes in a machine-understandable way

■ Set of RDF properties and RDFS classes to express the content and structure of a concept scheme as an RDF graph:

- Collection
- ConceptScheme
- prefLabel
- definition
- broader
- CollectableProperty
- Concept
- altLabel
- example
- narrower
- OrderedCollection
- hiddenLabel
- scopeNote
- related



# skos example

---

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:skos="http://www.w3.org/2004/02/skos/core#"
        xml:base="http://www.astro.physik.uni-goettingen.de/~hessman/rdf/IVOAT">
  <skos:ConceptScheme rdf:about="">
    <skos:hasTopConcept rdf:resource="#temperatureScales"/>
  </skos:ConceptScheme>
  <skos:Concept rdf:about="#absoluteTemperatureScale">
    <skos:prefLabel>absolute temperature scale</skos:prefLabel>
    <skos:definition>absolute temperature scale</skos:definition>
    <skos:broader rdf:resource="#temperatureScales"/>
    <skos:narrower rdf:resource="#Kelvin"/>
    <skos:related rdf:resource="#celsiusTemperatureScale"/>
    <skos:related rdf:resource="#temperature"/>
  </skos:Concept>
</rdf:RDF>
```

*matthew graham*

---

# \_\_\_\_\_sparql protocol and rdf query language

- W3C standard for RDF query language and data access protocol
- Variables indicated by "?"
- Query forms: SELECT, CONSTRUCT, ASK, DESCRIBE
- Patterns: WHERE, FILTER, OPTIONAL, UNION
- Modifiers: ORDER BY, DISTINCT, REDUCED, LIMIT, OFFSET

# sparql examples

---

```
PREFIX table: <http://www.example.org/PeriodicTable#>
SELECT ?name ?symbol ?weight ?number ?color
FROM <http://www.example.org/PeriodicTable.owl>
WHERE
{
    ?uranium table:name "uranium".
    ?uranium table:atomicWeight ?uraniumWeight.
    ?element table:name ?name.
    ?element table:symbol ?symbol.
    ?element table:atomicWeight ?weight.
    ?element table:atomicNumber ?number.
    OPTIONAL { ?element table:color ?color }.
    FILTER ( ?weight > ?uraniumWeight )
}
ORDER BY ASC[?weight]
```

*matthew graham*

---

# sparql examples

---

```
PREFIX table: <http://www.example.org/PeriodicTable#>
CONSTRUCT { ?element table:name ?name ; table:symbol ?symbol ;
  table:atomicWeight ?weight ; table:color ?color }
FROM <http://www.example.org/PeriodicTable.owl>
WHERE {
  ?uranium table:name "uranium".
  ?uranium table:atomicWeight ?uraniumWeight.
  ?element table:name ?name.
  ?element table:symbol ?symbol.
  ?element table:atomicWeight ?weight.
  ?element table:atomicNumber ?number.
  OPTIONAL { ?element table:color ?color }.
  FILTER ( ?weight > ?uraniumWeight ) }
ORDER BY ASC[?weight]
```

*matthew graham*

---