Knowledge Graphs

INFO216

Marc Gallofré Ocaña
marc.gallofre@uib.no

INFO216
Organization

Course information

- Course language: **English**
- Course materials and information:
  - mitt.uib.no
  - wiki.uib.no/info216
- 14 weeks course (tentative planning)
  - 12-14 lecture sessions
  - 12-14 lab sessions
- **You learn programming (mostly) through the lab exercises and project!**

Course recommended knowldege

- Good programming skills
  - **Java** is recommended
- **SQL** basic skills
- **Semantic Technologies course (INFO116)**
- Class or DBs modeling concepts (either UML, EER or others)

Course grading

- Group final project (**40%**)
  - Mandatory presentations
  - Graded on deliveries and progress
  - Follow-up meetings/presentations
- 3 hours individual written final exam (**60%**)
  - Exam requirements:
    - Submitted programming project
    - Participation in 80% of labs
  - Exam pass grade is 50 points

Follow-up meetings and presentations

- 3 project presentations during lab sessions
  - The last one is the final project presentation, during the last lab session.
- 1 elevator pitch (2min)
- 1 PechaKucha (15slidesx20’’)
- 1 Follow-up meeting

Course organization

- Lectures
  - **Mondays 14:15 to 16:00** from January 14th until April 8th (*last lecture date can change*)
- Labs (2 groups) ← Starting Tuesday 22nd
  - Group 1: **Tuesday 08:15 - 10:00**
  - Group 2: **Wednesday 14:15 - 16:00**
- Question sessions ← Starting Tuesday 22nd
  - **Tuesday 14:15 - 16:00**

Course organization

- Lectures: **BC Aud 128**
- Labs (2 groups): **Ulrikke Pihl PC 202** xor **UP PC 205**
- Question sessions: **BC Aud 128**

Labs organization

- Labs
  - 3 lab days used for project presentations and discussions.
  - The rest are practical assignments
  - **80% mandatory attendance, including all the presentation days**

Project organization

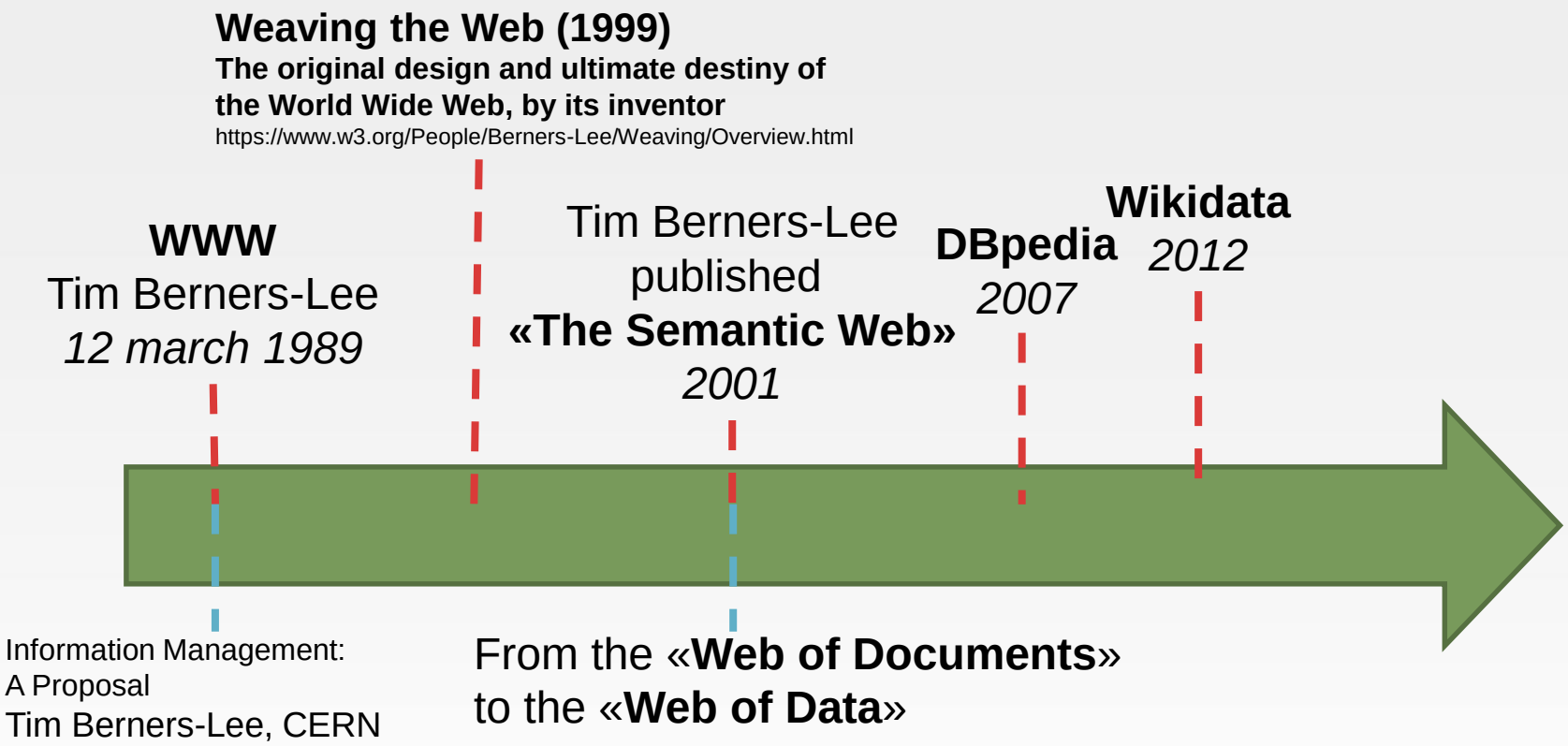
- Groups
  - Groups of 3 people.
  - In case of not multiple of 3, we will do a group of 2 people.
  - All group members should register themself into the same lab group
  - In case of other subjects overlaps with lab schedules, the last group registered will be changed allowing the attendance of the other group.

Important dates

- Final project submission
  - **14:00 on May 24th** through Inspera
- Final exam
  - **09:00 on June 13th** (date and time can change)

Semantic web history

Semantic web and WWW history



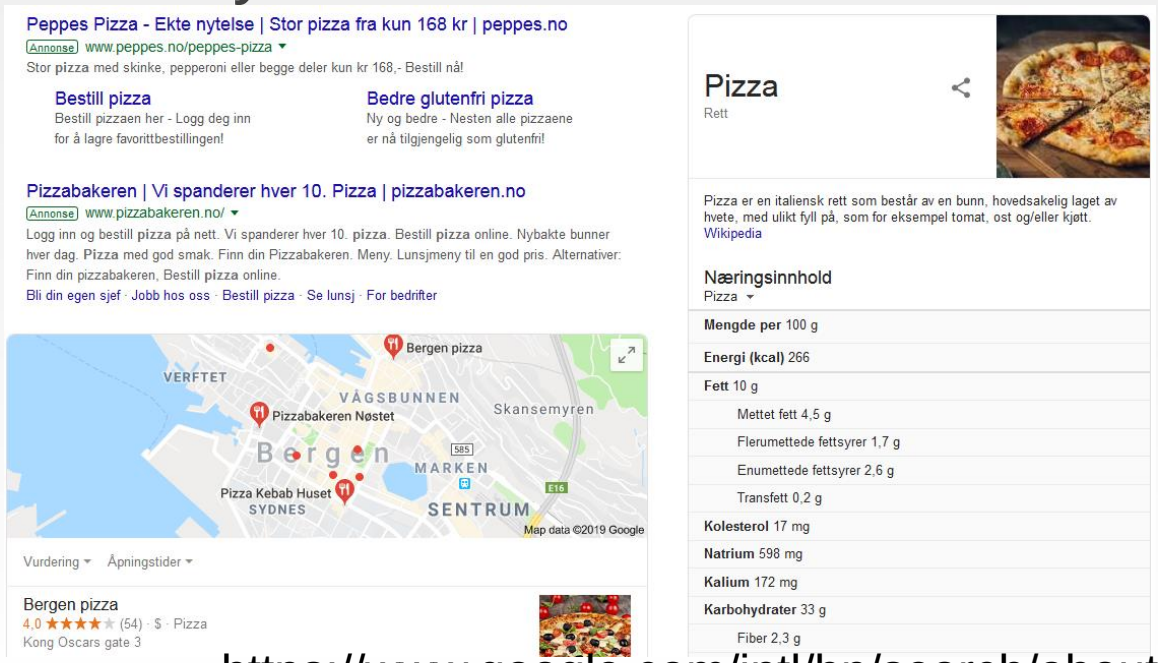
Tim Berners-Lee: http://www.youtube.com/watch?v=HeUrEh-nqtU
Information Management: A Proposal: https://cds.cern.ch/record/369245/files/dd-89-001.pdf

Tim Berners-Lee
Semantic Web of Data

Industrial Solutions

Google Knowledge Graph (2012)

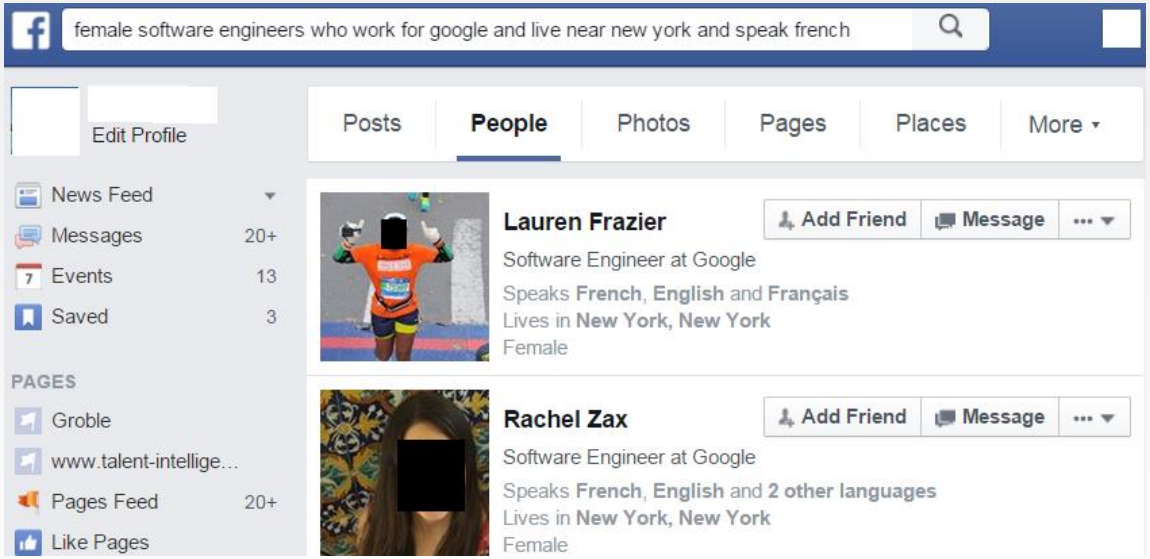
Google uses Knowledge Graph to enhance its search engine's results with information gathered from a variety of sources.



https://www.google.com/intl/bn/search/about

Facebook Graph Search (2013)

Facebook Graph Search was a **semantic search** designed to give answers to user natural language queries rather than a list of links.



https://www.facebook.com

Amazon Neptune (2017)

Amazon Neptune is a graph database service that supports **W3C's RDF** and **SPARQL** among others.

https://aws.amazon.com/neptune/

BBC Ontologies

BBC is using ontologies and its own vocabulary to support their news.

https://www.bbc.co.uk/ontologies

UiB-SSIS Research

UBIMOB-Ubiquitous Data-Driven Urban Mobility

To analyse and interlink data for being utilized to understand, optimise and manage mobility and make it more efficient, sustainable and smart.

https://www.uib.no/fg/ssis/111998/ubimob-ubiquitous-data-driven-urban-mobility

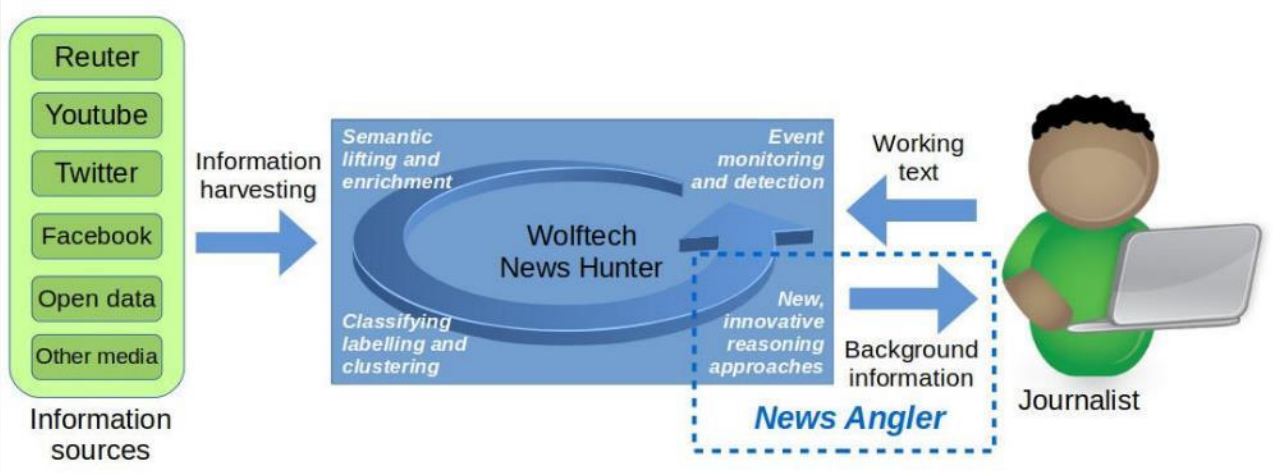
BDEM-Big Data for Emergency Management

Aims to leverage open and big data to assist strategic and operational deciding and acting during all phases of emergency management, from preparation, detection, rescue, relief, and rebuilding.

https://www.uib.no/fg/ssis/111999/bdem-big-data-emergency-management

NewsAngler

To develop tools that can help journalists finding new and unexpected background information, angles and stories, from news, big data and social media sources.



https://www.uib.no/en/rg/ssis/114224/discovering-unexpected-connections-news

Semantic web conecepts

Basic conecepts

- **AAA: a**nyone can say **a**nything about **a**ny topic.
  - Two theories: Flat earth (https://theflatearthsociety.org) and Spherical Earth (https://en.wikipedia.org/wiki/Spherical_Earth)
- **Nonunique Naming Assumption**: we have to assume that some Web resource might be referred to using different names by different people.
  - A star can have diferent names: Sirius or HD 48915 or HR 2491

Basic conecepts

Open world assumption: an open world in this sense is one in which we must assume that at any time any new information could come to light, and we may draw no conclusions that rely on assuming that the information available at any one point is all the information available.

URI and IRI

Uniform Resource Identifier and *Internationalized Resource Identifier.* Is a string of characters that **unambiguously** identifies a particular resource.

URI works only with English alphabet (US-ASCII) and IRI uses Unicode/ISO 10646.

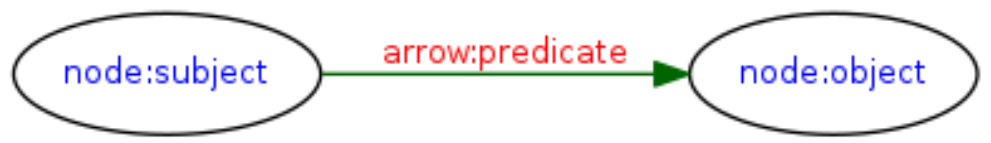
foo://example.com:8042/over/there?name=ferret#nose

https://www.ietf.org/rfc/rfc3986.txt
https://www.ietf.org/rfc/rfc3987.txt

RDF

Resource Description Framework. Basic framework that the Semantic Web is based on.

- Treated as a set of **triples**
- The relations form a directed graph
  - "Nodes": represent resources
  - "Leaf nodes": can contain literal values
  - "Arrows": relations between resources or literals



Resource

Resources may be physical phenomena (including people and artefacts), information resources, concepts, constructs, and information about all of them.

- Can be the **subject** or **object** in a statement but only rdf:Property can be predicate
- Can be named by an **IRI** or **anonymous** (blank node)
- Resources can have one or more rdf:type

Basic conecepts

Semantic Graph: is a collection of triples/statements (possibly zero).

Data set: a collection of at least one graph

- One of the graphs is **default/unnamed**
- the others are **named**

Blank node (anonymous resource)

- Some nodes have no IRIs:
  - cannot be referenced by other data sets
  - can be referenced by graphs in same data set
- Can have a (non-IRI) identifier, but local identifier only meaningful inside the data set
- **Cannot be used to merge nodes from different data sets**
- Uses:
  - you are not sure of the IRI
  - you do not want to provide an IRI
  - logical grouping of related properties
  - not supported by all RDF technologies

Vocabulary

Semantic web use vocabulary to express the meaning of concepts:

- IRIs for *resource types*
- IRIs for *properties*
- Standard types for *literals*
- *Rules* about how they combine

Other *open semantic datasets* define standard IRIs for *individual resources*

Vocabulary

- **RDFS** (RDF Schema) is a general-purpose language for representing simple RDF vocabularies on the Web
- **OWL** (Web Ontology Language) is a Semantic Web language designed to represent knowledge about things, groups of things, and relations between things.
- **SKOS** (Simple Knowledge Organization System) is a common data model for sharing and linking knowledge organization systems via the Web.

Qnames

A **QName** is a simplified version of a IRI abbreviation scheme. A QName concisely associates the IRI of a resource namespace with the local name in that namespace.

owl: <http://www.w3.org/2002/07/owl#> .
rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
rdfs: http://www.w3.org/2000/01/rdf-schema#
skos: <http://www.w3.org/2004/02/skos/core#> .

Qnames

A **QName** has two parts: a namespace and a identifier, written with a colon between (namespace **:** identifier).

owl:sameAs → http://www.w3.org/2002/07/owl#sameAs
rdf:type → http://www.w3.org/1999/02/22-rdf-syntax-ns#type

TBox and ABox

- **TBox** ("*T from term*") are statements that describe a domain of interest by defining classes and properties as a domain vocabulary.

  All students are Persons

- **ABox** ("*A from assertion*") are facts associated with a conceptual model or ontologies within a knowledge base.

  John is a Student

  John is a Person

RDF

RDF triples

- The basic form of a RDF block is called **triple.**
- A **triple** has a **subject, predicate** and **object** (as a grammatical sentence).

Jonh bought apple

| | |

subject predicate object

RDF triples

- **Subject**: represents what the statement is about
  - the IRI of a semantic resource
- **Predicate**: represents a property of the subject resource
  - the IRI of a semantic property
- **Object**: represents the value of a property for a subject
  - the IRI of a semantic resource
  - a literal (number, string, boolean…)

RDF Triples

- Triples of **subject predicate object** or **subject predicate literal**
  - The subject:
    - must be a **resource** (→rdfs:Resource)
    - named by an IRI or anonymous (blank node)
  - The predicate:
    - must be a property (→rdf:Property)
    - properties are **resources** too!
  - The object:
    - either a **resource** (named or anonymous/blank) or a constant **value** (→rdfs:Literal)

A resource can be either a **subject** or **object** or **predicate**

RDF Triples

:John :bought :Apple

:Apple :from :Norway

:Apple :brand “NorgeEpler”

Triples

Representing data as triples

Author	Title
Mary Shelley	Frankenstein
Shakespeare	Macbeth

Subject	Predicate	Object
Mary Shelley	wrote	Frankenstein
Shakespeare	wrote	Macbeth

Triples exercise

Individual (5 minutes): Transform the given table into a triples table (subject – predicate – object)

Author	Book	Year
Mary Shelley	Frankenstein	1818
Isaac Asimov	Foundation	1942

Triples solution

Individual (5 minutes): Transform the given table into a triples table (subject – predicate – object)

Subject	Predicate	Object
Mary Shelley	wrote	Frankenstein
Isaac Asimov	wrote	Foundation
Frankenstein	publishedIn	1818
Foundation	publishedIn	1942

RDF triples

Based on the previous exercise using RDF triples and Qnames.

lit: → referencing to literature schema resources
http://www.INFO216.org/foo/literature

Subject	Predicate	Object
lit:Mary Shelley	lit:wrote	lit:Frankenstein
lit:Isaac Asimov	lit:wrote	lit:Foundation
lit:Mary Shelley	rdf:type	lit:Writer
lit:Foundation	lit:author	lit:Isaac Asimov

RDF/XML description

Qname

```
1: <?xml version="1.0"?>
2: <rdf:RDF
3:              xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:              xmlns:lit="http://www.INFO216.org/foo/literature#">
5:    <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Mary_Shelly">
6:       <lit:wrote>
7:          <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Frankenstein"/>
8:       </lit:wrote>
9:       <rdf:type rdf:resource="http://www.INFO216.org/foo/literature#Writer"/>
10:   </rdf:Description>
11:   <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Isaac_Asimov">
12:      <lit:wrote>
13:         <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Fundation">
14:            <lit:author>
15:               <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Isaac_Asimov"/>
16:            </lit:author>
17:         </rdf:Description>
18:      </lit:wrote>
19:      <rdf:type rdf:resource="http://www.INFO216.org/foo/literature#Writer"/>
20:   </rdf:Description>
21: </rdf:RDF>
```

RDF/XML description

Subject

```
1: <?xml version="1.0"?>
2: <rdf:RDF
3:              xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:              xmlns:lit="http://www.INFO216.org/foo/literature#">
5:    <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Mary_Shelly">
6:       <lit:wrote>
7:          <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Frankenstein"/>
8:       </lit:wrote>
9:       <rdf:type rdf:resource="http://www.INFO216.org/foo/literature#Writer"/>
10:   </rdf:Description>
11:   <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Isaac_Asimov">
12:      <lit:wrote>
13:         <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Fundation">
14:            <lit:author>
15:               <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Isaac_Asimov"/>
16:            </lit:author>
17:         </rdf:Description>
18:      </lit:wrote>
19:      <rdf:type rdf:resource="http://www.INFO216.org/foo/literature#Writer"/>
20:   </rdf:Description>
21: </rdf:RDF>
```

RDF/XML description

```
1: <?xml version="1.0"?>
2: <rdf:RDF
3:              xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:              xmlns:lit="http://www.INFO216.org/foo/literature#">
5:    <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Mary_Shelly">
6:       <lit:wrote>
7:          <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Frankenstein"/>
8:       </lit:wrote>
9:       <rdf:type rdf:resource="http://www.INFO216.org/foo/literature#Writer"/>
10:   </rdf:Description>
11:   <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Isaac_Asimov">
12:      <lit:wrote>
13:         <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Fundation">
14:            <lit:author>
15:               <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Isaac_Asimov"/>
16:            </lit:author>
17:         </rdf:Description>
18:      </lit:wrote>
19:      <rdf:type rdf:resource="http://www.INFO216.org/foo/literature#Writer"/>
20:   </rdf:Description>
21: </rdf:RDF>
```

Predicate

RDF/XML description

```
1: <?xml version="1.0"?>
2: <rdf:RDF
3:              xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:              xmlns:lit="http://www.INFO216.org/foo/literature#">
5:    <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Mary_Shelly">
6:       <lit:wrote>
7:         <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Frankenstein"/>
8:       </lit:wrote>
9:       <rdf:type rdf:resource="http://www.INFO216.org/foo/literature#Writer"/>
10:   </rdf:Description>
11:   <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Isaac_Asimov">
12:      <lit:wrote>
13:         <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Fundation">
14:            <lit:author>
15:               <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Isaac_Asimov"/>
16:            </lit:author>
17:         </rdf:Description>
18:      </lit:wrote>
19:      <rdf:type rdf:resource="http://www.INFO216.org/foo/literature#Writer"/>
20:   </rdf:Description>
21: </rdf:RDF>
```

RDF/XML description

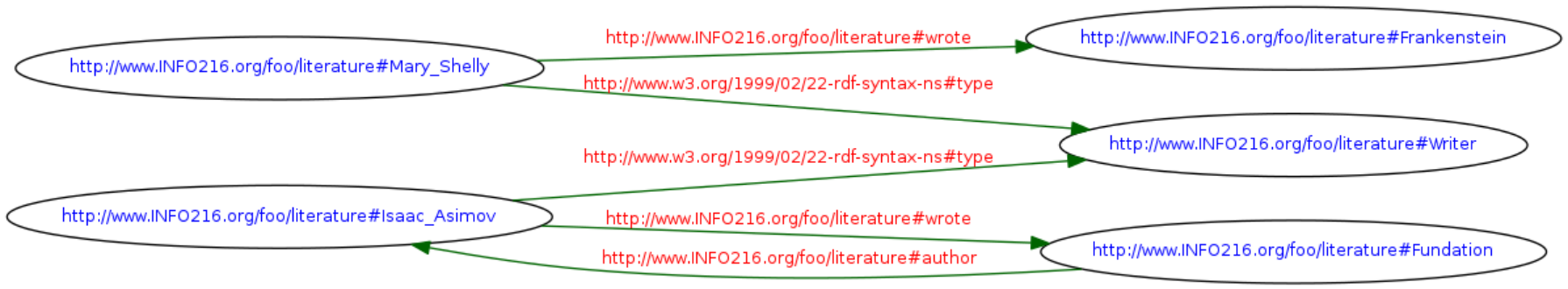
```
1: <?xml version="1.0"?>
2: <rdf:RDF
3:              xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:              xmlns:lit="http://www.INFO216.org/foo/literature#">
5:    <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Mary_Shelly">
6:       <lit:wrote>
7:         <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Frankenstein"/>
8:       </lit:wrote>
9:       <rdf:type rdf:resource="http://www.INFO216.org/foo/literature#Writer"/>
10:   </rdf:Description>
11:   <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Isaac_Asimov">
12:      <lit:wrote>
13:         <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Fundation">
14:            <lit:author>
15:               <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Isaac_Asimov"/>
16:            </lit:author>
17:         </rdf:Description>
18:      </lit:wrote>
19:      <rdf:type rdf:resource="http://www.INFO216.org/foo/literature#Writer"/>
20:   </rdf:Description>
21: </rdf:RDF>
```


RDF/XML description

```
1: <?xml version="1.0"?>
2: <rdf:RDF
3:              xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:              xmlns:lit="http://www.INFO216.org/foo/literature#">
5:    <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Mary_Shelly">
6:       <lit:wrote>
7:         <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Frankenstein"/>
8:       </lit:wrote>
9:       <rdf:type rdf:resource="http://www.INFO216.org/foo/literature#Writer"/>
10:   </rdf:Description>
11:   <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Isaac_Asimov">
12:      <lit:wrote>
13:         <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Fundation">
14:            <lit:author>
15:               <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Isaac_Asimov"/>
16:            </lit:author>
17:         </rdf:Description>
18:      </lit:wrote>
19:      <rdf:type rdf:resource="http://www.INFO216.org/foo/literature#Writer"/>
20:   </rdf:Description>
21: </rdf:RDF>
```

Nested description

RDF graph



https://www.w3.org/TR/rdf-syntax-grammar/
https://www.w3.org/RDF/Validator/rdfval

RDF/XML exercise

Groups of 3 (5 minutes): Write the RDF/XML statement which represents the following graph.

```
1: <?xml version="1.0"?>
2: <rdf:RDF
3:          xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:          xmlns:lit="http://www.INFO216.org/foo/literature#">
```

RDF/XML solution I

Groups of 3 (5 minutes): Write the RDF/XML statement which represents the following graph.

```
1: <?xml version="1.0"?>
2: <rdf:RDF
3:            xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:            xmlns:lit="http://www.INFO216.org/foo/literature#">
5:   <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Mary_Shelly">
6:     <lit:wrote>
7:       <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#Frankenstein">
8:           <lit:character>
9:                <rdf:Description
rdf:about="http://www.INFO216.org/foo/literature#Victor_Frankenstein"/>
10           </lit:character>
11:       </rdf:Description>
12:     </lit:wrote>
13:   </rdf:Description>
14: </rdf:RDF>
```

RDF/XML solution II

Groups of 3 (5 minutes): Write the RDF/XML statement which represents the following graph.

```
1: <?xml version="1.0"?>
2: <rdf:RDF
3:            xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4:            xmlns:lit="http://www.INFO216.org/foo/literature#">
5:   <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#:Mary_Shelly">
6:     <lit:wrote>
7:       <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#:Frankenstein"/>
8:     </lit:wrote>
9:   </rdf:Description>
10: <rdf:Description rdf:about="http://www.INFO216.org/foo/literature#:Frankenstein">
11:    <lit:character>
12:      <rdf:Description
rdf:about="http://www.INFO216.org/foo/literature#:Victor_Frankenstein"/>
13:    </lit:character>
14:  </rdf:Description>
15: </rdf:RDF>
```

UNIVERSITY OF BERGEN