TDT 44 – Semantic Web

Özlem Özgöbek

ozlem.ozgobek@ntnu.no

Let’s meet!

About the course

Goal: The students get an orientation of the theoretical, methodological and technological background and of the ongoing standardization work that support semantic Web.

You will gain:

- Background theoretical knowledge
- Insight to relevant problems within Web information resources and services
- Insight to how to use semantic web technology to solve problems in different domains

Mostly self-driven.

Learning Material

BOOKS

ACADEMIC PAPERS

PRESENTATIONS

EXTERNAL LINKS

Follow the course wiki page: https://www.ntnu.no/wiki/display/idiemner/TDT44+-+Semantic+Web

Schedule

22.09.2021 Introductory meeting
Presentation of semantic web concepts
14.10.2021 Mid-term follow-up meeting - 1 with presentations from students
21.10.2021 Mid-term follow-up meeting - 2 with presentations from students
04.11.2021 Mid-term follow-up meeting - 3 with presentations from students
XX.11.2021 Report deadline
XX.11.2021 (TBA) Oral exam – Presentation of reports

Mandatory activities: Presentations, project report, final presentation.

Final grade: Project report and final presentation.

The details will be available on the wiki page soon.

Brief Introduction to Semantic Web

Semantic Web

Machines understanding the meaning
or
Meaning is machine readable.

Why do we need it?
Too much information, too little structure.
Meaning (semantics) of data is only human readable.
Connected data is smarter data.

Semantic Web

Set of standards

W3C - World Wide Web Consortium
https://www.w3.org/

Semantic Data

Representation of data

Distributed and connected data

Application areas

- Information retrieval
- Knowledge management
- Natural language processing

“Smart systems of any type”

- Search engines
- E-learning
- E-commerce
- Recommender systems

Open World

The world is dynamic.

Information changes all the time.

The Web is an Open World, we treat it using the “Open World Assumption”.

It is very important when we want to draw conclusions based on distributed data.

Semantic Modeling

“to make sense of the world around us”

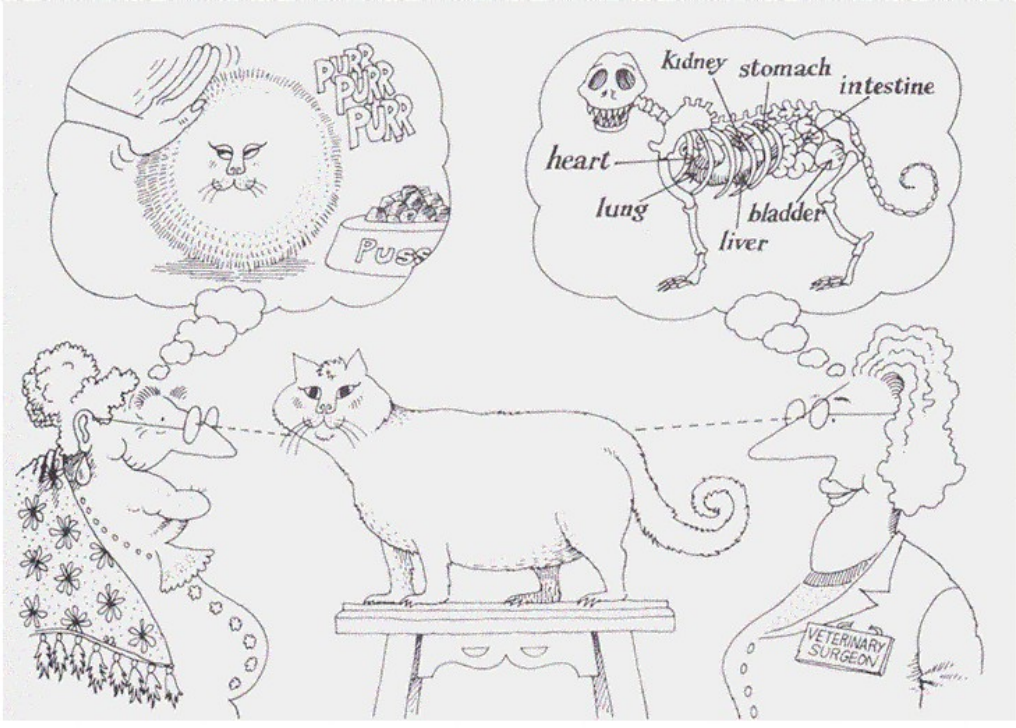
Models are abstract descriptions that hides certain details while illuminating others.

Models:
Help people communicate
Explain and make predictions
Mediate among multiple viewpoints

The Semantic Web standards enable people to collaborate on models that organizes the information. Common collection of knowledge.

A semantic web model must be able to organize **variations** in a meaningful and manageable way.

In semantic web context, modeling is an ongoing process.



How to do it?

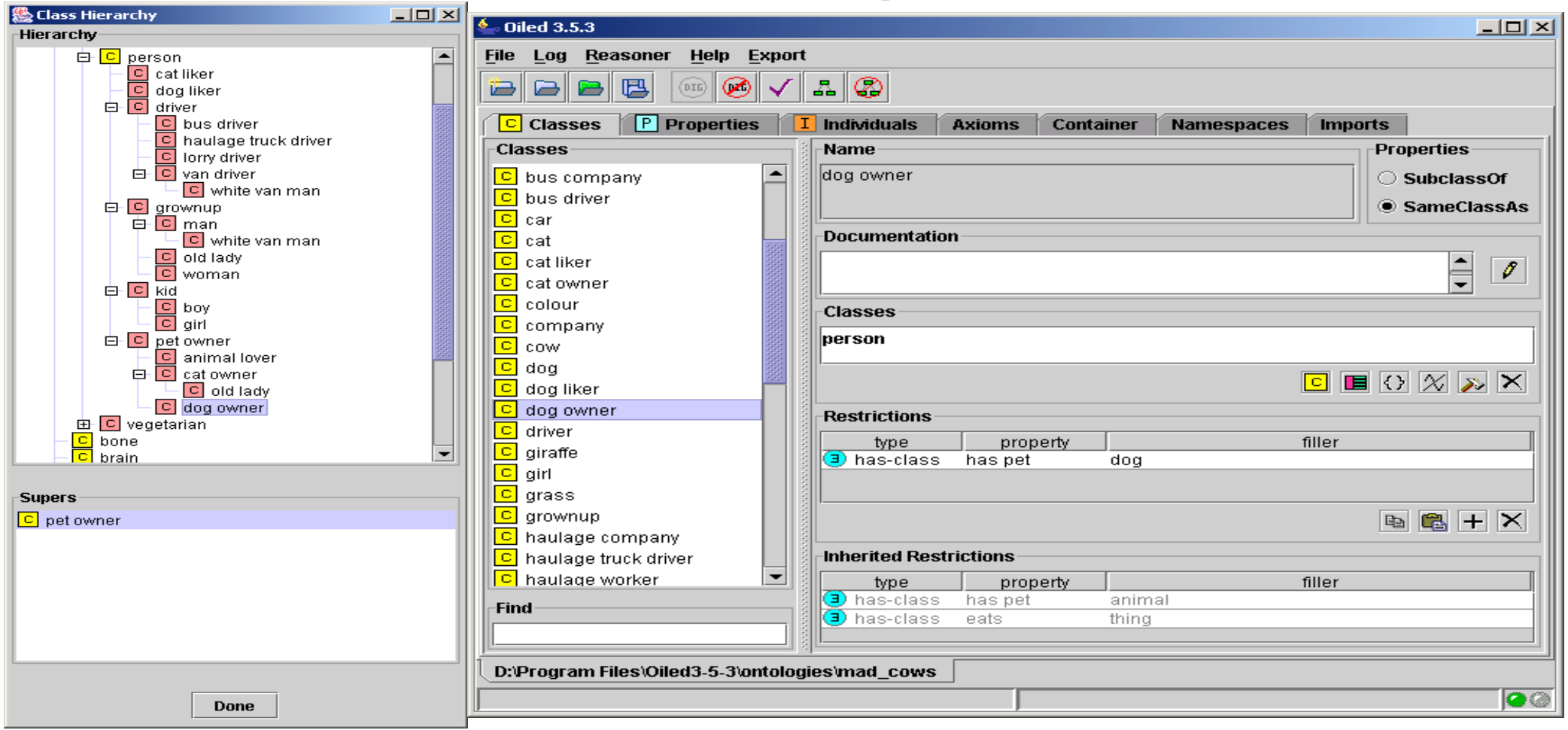
Frameworks, descriptions, structured representation of data, links, inference etc.

Ontology

An ontology describes a formal specification of a certain domain. The types of entities and relationships are defined in ontologies.

There are different ontology/semantic web modeling languages: XML, RDF, OWL ...

An example:


Class Hierarchy
Hierarchy
person
cat liker
dog liker
driver
bus driver
haulage truck driver
lorry driver
van driver
white van man
grownup
man
white van man
old lady
woman
kid
boy
girl
pet owner
animal lover
cat owner
old lady
dog owner
vegetarian
bone
brain
Supers
pet owner
Done
Oiled 3.5.3
File
Log
Reasoner
Help
Export
Classes
Properties
Individuals
Axioms
Container
Namespaces
Imports
Classes
bus company
bus driver
car
cat
cat liker
cat owner
colour
company
cow
dog
dog liker
dog owner
driver
giraffe
girl
grass
grownup
haulage company
haulage truck driver
haulage worker
Find
Name
dog owner
Properties
SubclassOf
SameClassAs
Documentation
Classes
person
Restrictions
type
property
filler
has-class
has pet
dog
Inherited Restrictions
type
property
filler
has-class
has pet
animal
has-class
eats
thing
D:\Program Files\Oiled3-5-3\ontologies\mad_cows

What is the usefulness of an ontology?

- To make domain assumptions explicit
  - Ontological analysis
    - clarifies the structure of knowledge
    - allow domain knowledge to be explicitly defined and described
- Enrich software applications with the additional semantics
- To facilitate communications among systems with out semantic ambiguity. i,e to
- achieve inter-operability
- Thus, practically, improving: computer-computer, computer-human, and human-
- human communication
- To provide foundations to build other ontologies (reuse)
- To save time and effort in building similar knowledge systems (sharing)

Challenges

- Scalability
- Logical structure

“The challenge of the Semantic Web is to find a representation language powerful enough to support automated reasoning but simple enough to be usable” [AKT 2003]

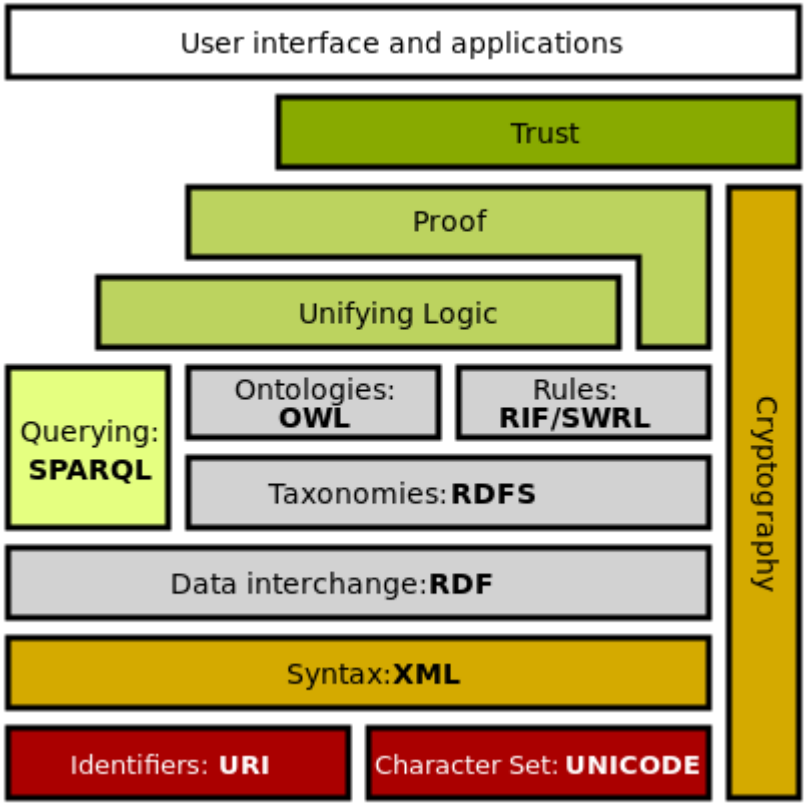
However...

“The Semantic Web is not about getting everyone agree on a single ontology, but rather about coping in a world where not everyone will agree, achieving some degree of interoperability.”

Book: Semantic Web for the Working Ontologist

The Semantic Web Stack

- XML
  Surface syntax, no semantics
- XML Schema
  Describes structure of XML documents
- RDF
  Datamodel for “relations" between “things"
- RDF Schema
  RDF Vocabulary Definition Language
- OWL
  A more expressive Vocabulary Definition Language

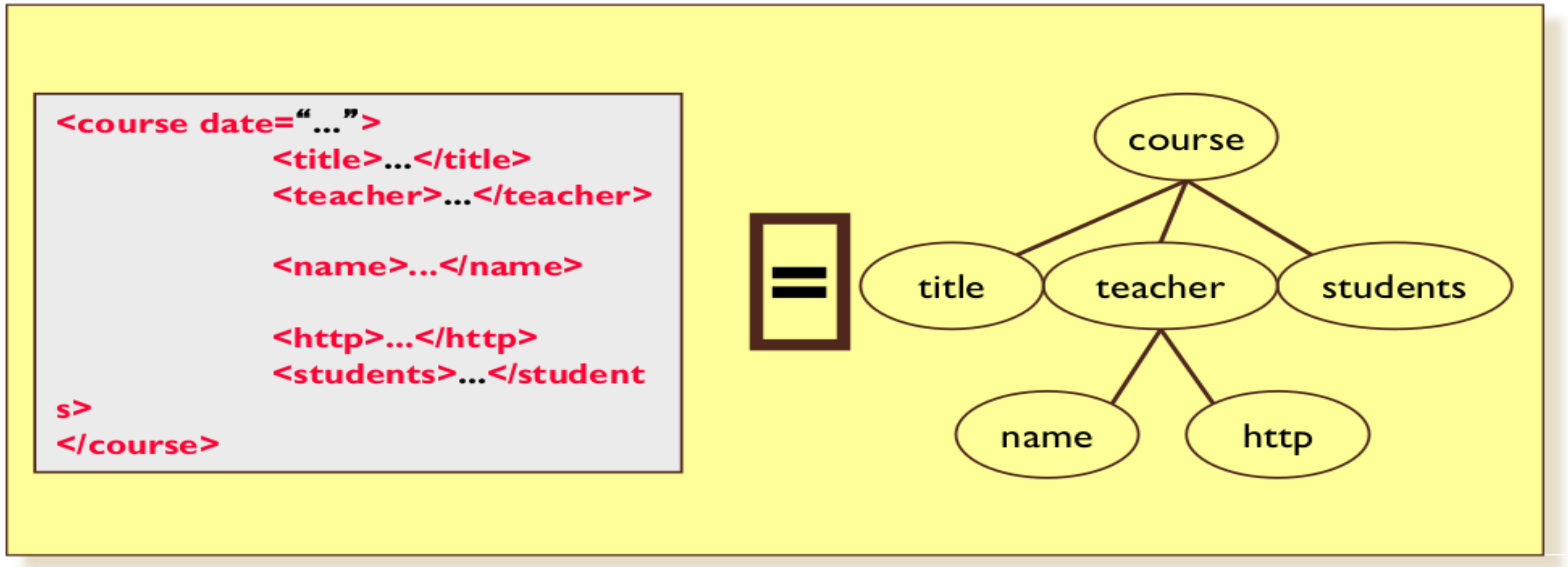


https://en.wikipedia.org/wiki/Semantic_Web

XML

Document is a labeled tree. XML is a markup language that enables creation of documents composed of semi-structured data.

- node = label + attr/values + contents



- XML Schema: grammars for describing legal trees and datatypes

RDF → the standard for representing knowledge graphs

RDFS and OWL → two standards for defining ontologies

SPARQL → the standard for querying knowledge graphs

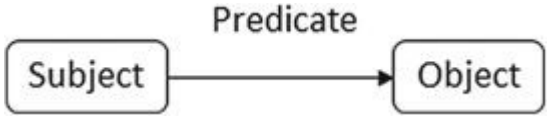
RDF (Resource Description Framework)

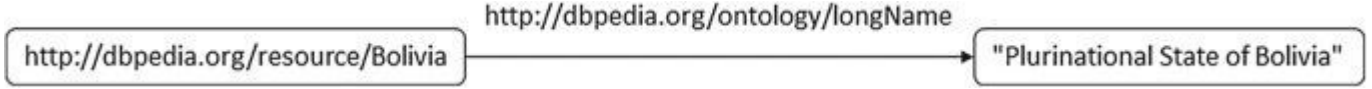
RDF is a standard to describe entities/resources. A resource can be anything we can identify, such as a person, a homepage or great dragons in the Game of Thrones.

Resources in RDF are described as triplets:

[subject, predicate, object]

[lemon, is, yellow]
[Ann, knows, Barbara]
[Mary, likes, pizza]





RDF

RDF models the relations between things.

- It has more interoperability,
- Provides better understanding of the data.

RDFS

Resource Description Framework Schema

RDF defines the structure of the data
RDFS defines semantic relationships

```
org:Organization rdf:type rdfs:Class .
org:Start-up a rdfs:Class .
org:hasHomePage rdf:type rdfs:Property .
org:Start-up rdfs:subClassOf org:Organization .
```

OWL

Web Ontology Language

https://www.w3.org/2001/sw/wiki/OWL

“It is a Semantic Web language designed to represent rich and complex knowledge about things, groups of things, and relations between things.”

SPARQL

Standard query language for RDF and OWL

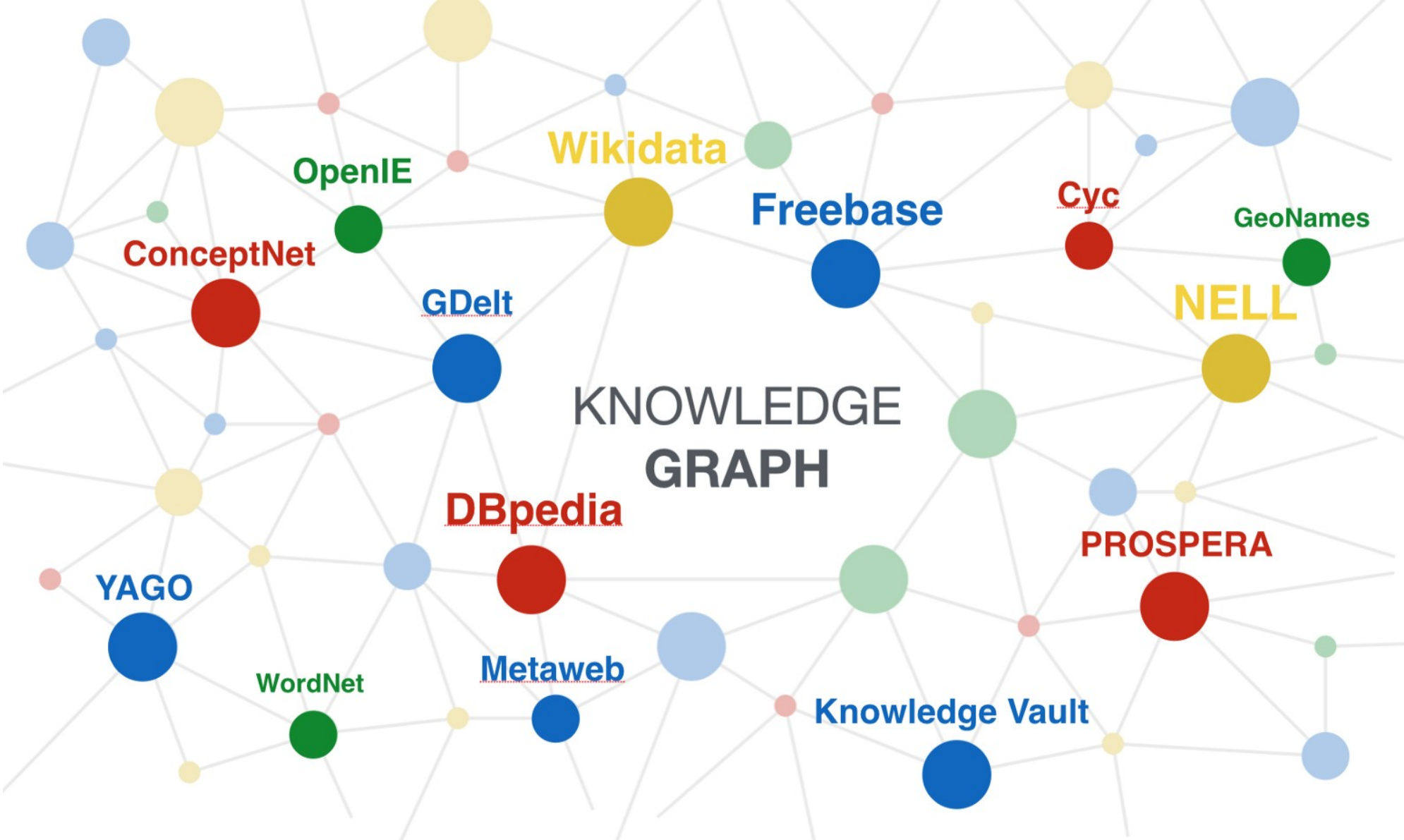
- Turtle syntax
- SQL like
- Based on pattern-matching mechanism

Knowledge Graphs

“A knowledge graph consists of a set of interconnected typed entities and their attributes.”

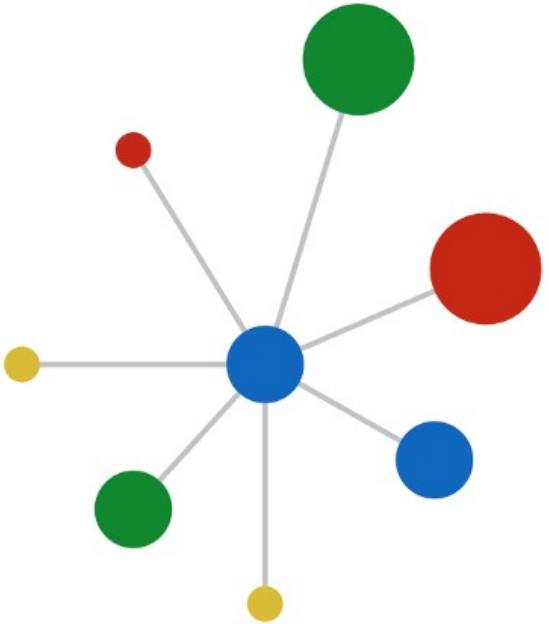
Popularized by Google in 2012.

Now also used by Facebook,
LinkedIn and Microsoft.

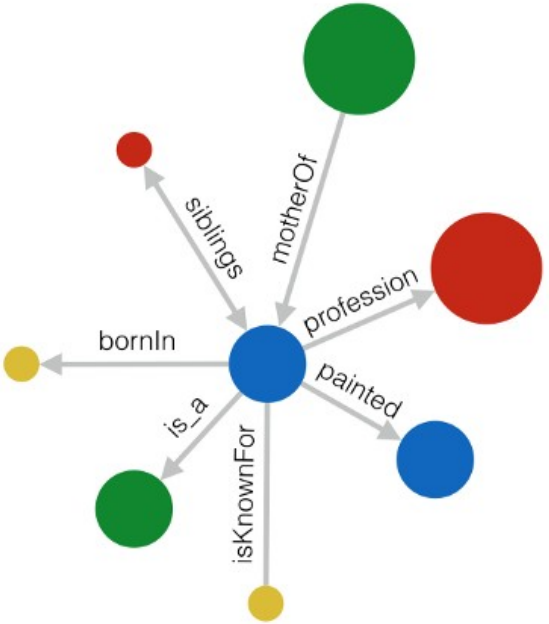

KNOWLEDGE
GRAPH
Wikidata
OpenIE
Freebase
Cyc
GeoNames
ConceptNet
GDelt
NELL
DBpedia
PROSPERA
YAGO
Metaweb
WordNet
Knowledge Vault

“leonardo da vinci”
String

Leonardo da Vinci
Recognized entity



Leonardo da Vinci
Recognized entity
Related entities



Leonardo da Vinci
Recognized entity
Related entities
Named Relationship

https://medium.com/@sderymail/challenges-of-knowledge-graph-part-1-d9ffe9e35214

Table 2.3 RDF Datasets versus Linked Data versus Knowledge Graph

Features	Pure RDF datasets	Linked Data	Knowledge Graph
Machine readability	Y	Y	Y
Human readability	NN	NN	Y
Data distribution	N	Y	NN
Inter-dataset linkage	L	Y	Y
Data integration	NN	NN	Y
Data consistency	NN	NN	Y
Reliability	NN	NN	Y
High quality	NN	NN	Y

Y= Yes, NN = Not Necessarily, N = No, L= Limited

Some Standard Ontologies

- Friend-of-a-Friend (FOAF)
  https://en.wikipedia.org/wiki/FOAF_(ontology)
- schema.org

- Read and study relevant learning materials.
- Check last year’s assignment.
- Follow the course wiki page.
- Contact me for further questions.