Learning Outcomes

• After learning this chapter, you should be able to
  ▪ Understand the basics of Semantic Web and its techniques
  ▪ Be able to do simple SPARQL SELECT queries
The Semantic Web

• Syntax VS Semantics
• Web 3.0: Make the web easier to understand by computers
  ▪ Web 2.0: user generated web contents, usability, and interoperability. Examples are blog, wiki, social network sites and mashup.
• Techniques/standards for Semantic Web
  ▪ Resource description framework (RDF) - describing data with a global naming scheme (URIs)
  ▪ RDF Schema (RDFS) - A standard way of describing the properties of that data
  ▪ Web Ontology Language (OWL) - A standard way of describing relationships between data items.
  ▪ RDF Stores and SPARQL - A standard way of storing and querying graphs.
From Current Web to Semantic Web

a) Current Web

b) Semantic Web
Resource Description Framework (RDF)

• **RDF** is a graph model built by triples: `<resource/subject, property/predicate, value/object>`
• RDF describes data with a global naming scheme (URIs)
• RDF is independent of any serialization
• Common RDF serialization notion/language are **N3** and XML
• Commonly used RDF vocabularies include **Dublin Core**, **FOAF**
  ▪ You can define your own vocabularies
• **W3C's RDF Validation Service** for validation and graph representation
  ▪ **One graph representation example**
N3 and XML Serialization for RDF

@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema> .

<http://userpages.umbc.edu/~jianwu/foaf.rdf>
foaf:name "Jianwu Wang"^^xsd:string .

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/">
  <rdf:Description
    rdf:about="http://userpages.umbc.edu/~jianwu/foaf.rdf">
    <foaf:name>Jianwu Wang</foaf:name>
  </rdf:Description>
</rdf:RDF>
RDFSchemas

- RDF Schema is an extension of the basic RDF vocabulary
- It provides a standard way of describing the properties of that data
- It adds the concept of class (sub-class), and associated properties (sub-property) to RDF triples
RDFSchema (2)

• The RDF Schema class and property system is similar to the type systems of object-oriented programming languages such as the Java language.

• A difference is RDF Schema defines a property separately/independently by its range class and domain class.

• A specific resource is a type of a class
  ▪ UMBC is a type of university class

• You can find RDFSchema example in N3 at book chapter.
OWL (Web Ontology Language)

• **OWL** is a language for defining and instantiating Web ontologies

• Ontology is a term borrowed from philosophy that refers to the science of describing the kinds of entities in the world and how they are related

• An OWL ontology may include descriptions of classes, properties and their instances

• Given such an ontology, the OWL formal semantics specifies how to derive its logical consequences, i.e. facts not literally present in the ontology, but *entailed* by the semantics
SPARQL

- **SPARQL** is a query language for RDF
  - Recursive acronym: SPARQL Protocol and RDF Query Language
- Four types of SPARQL queries
  - `SELECT`
  - `CONSTRUCT`
  - `ASK`
  - `DESCRIBE`
- You can specify one or more SPARQL endpoints in SPARQL query
  - Endpoint list: [https://www.w3.org/wiki/SparqlEndpoints](https://www.w3.org/wiki/SparqlEndpoints)
SPARQL Demo

- **SPARQL Explorer**
- An SELECT query example

```sparql
PREFIX dbo: <http://dbpedia.org/property/>
SELECT ?name ?birth
{
?person foaf:name ?name .
FILTER (?birth < "1900-01-01"^^xsd:date) .}
ORDER BY ?name
#People who were born in Berlin before 1900
```
RDF (Triple) Stores

• A RDF triple store is a purpose-built **graph** database for the storage and retrieval of RDF metadata

• RDF Store Software: [Jena](https://jena.apache.org/), [RDF4J/Sesame](https://www.openrdf.org/), [Neo4j](https://neo4j.com/), etc.

create RDF files from other formats
Linked Data

• A method of publishing and interlinking structured data so it can become more useful through semantic queries
• It builds upon HTTP, RDF and URIs
• This enables data from different sources to be connected and queried
• Famous linked data sets
  ▪ DBpedia for extracted data from Wikipedia
  ▪ FOAF for persons, their properties and relationships
Semantic Web Criticism

• Web is too vast and vague
  ▪ But what about specific areas?
• Too computer-friendly