

Metadata and Semantics

RDF and Dublin Core

Resource Description Framework

- RDF is about adding semantics to metadata.
- It can be written as XML, but is not, *per se*, an XML vocabulary.
- Why on earth would we want *semantics*, and what do I mean by that anyway?

Semantics

- Broadly speaking, *semantics* involves questions of *meaning*.
- But what is meaning to a machine?
- Machines are simple creatures, really. Understanding, in this context, means that a machine has a useful response to a given piece of data.

RDF Schemas

- RDF has a schema language too.
- This one isn't concerned with validity, though, but modeling relationships between concepts.
- In RDF Schemas, everything is fundamentally a Class or a Property.

RDF Syntax

- In RDF, everything can be described with *triples* (or statements), a subject, a predicate, and an object. Each of these can be associated with a URI.
- These can be written as XML, but they don't have to be, they could be rows in a database table, for example.
- The RDF XML vocabulary is frequently

A word about ontologies

- There are things.
- There are things about things that make them different from other things. These are called properties.
- Things can be sub-classes of other things, like a Chihuahua is a type of dog.

What's the point?

- Why should we care about ontologies?
- The main use-case here is a program looking at a document of a type it's never encountered and isn't programmed to handle.
- Imagine we have an indexing program that's set up to save information about documents it encounters into a database....

Dublin Core

- We've already heard about DC.
- 15 elements. Title, description, creator, etc.
- + a set of qualifiers and types.
- And it has both XML and RDF Schemas.
Why?

XML for Data

- We're now on the other side of the coin.
What's different?
- Mixed content rare.
- Data typing more useful (and important).
- Combining vocabularies more common.
Lots of namespaces.

Why XML?

- Is XML a good fit for these technologies?
- When would you use it as opposed to a database or some other mechanism?