

# Web Services and Web Data

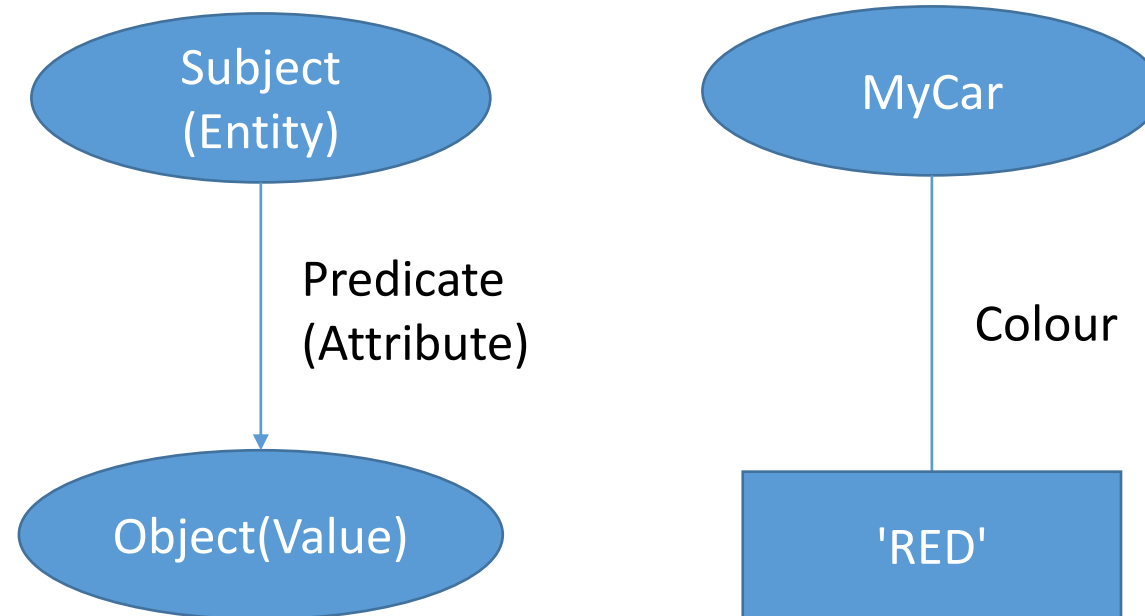
## XJCO3011



### *Session 15 - RDF: the Resource Description Framework*

# RDF Statements

- RDF descriptions are comprised of a number of statements, called **triples**.
- Each statement must have three parts: a subject, a predicate, and an object (hence the name triple)
- RDF is a World Wide Web Consortium (W3C) specification (<http://www.w3.org/standards/techs/rdf>)



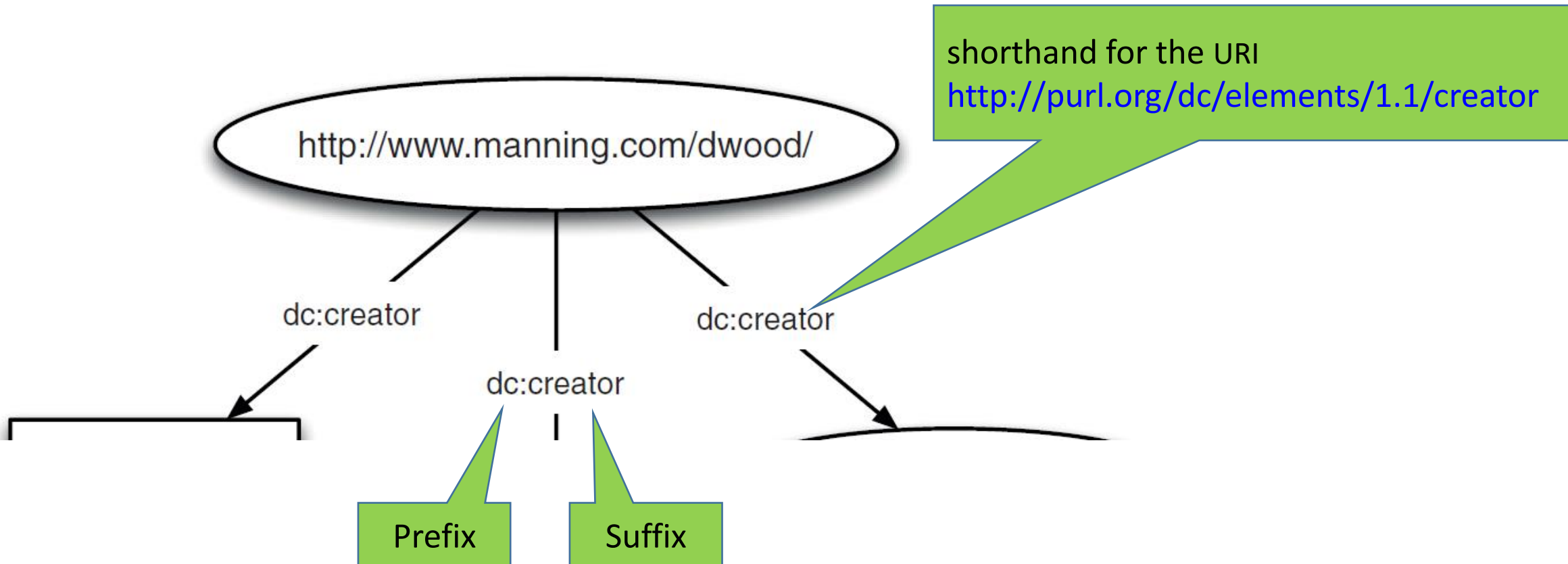
# *The Subject*

- Represents the entity that is being described
- Must be a resolvable URI (Web URL)
- Requesting the URL should return information about the subject (entity)



# The Predicate

- Determines an attribute of the subject
- Must have its own resolvable URI



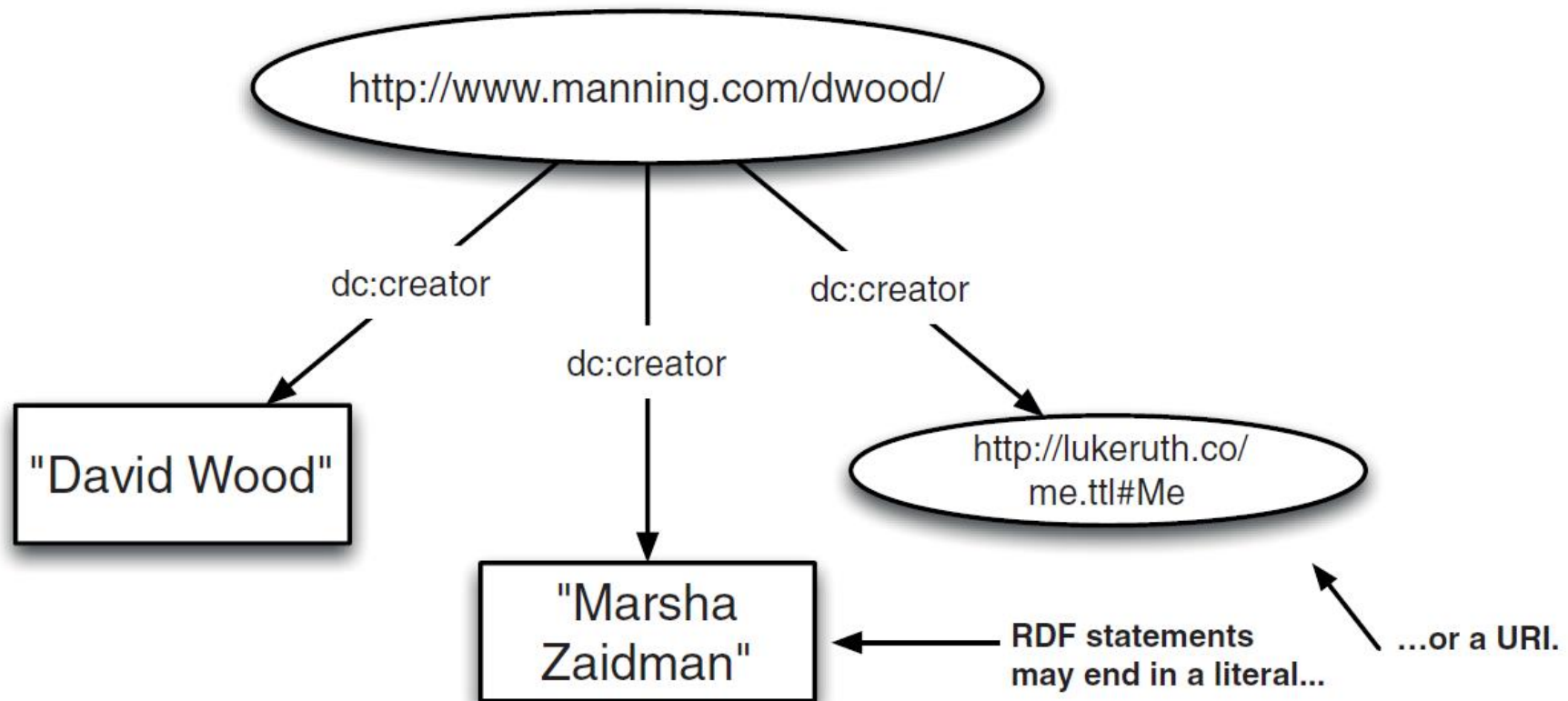
# Resolving Predicates

- If you need to know what the predicate means and find other information about it, just type the full URL of the predicate into your browser.
- For example, Here is the result of resolving the URL <http://purl.org/dc/elements/1.1/creator>.
- You can find that the predicate itself defined as an RDF statement

```
<http://purl.org/dc/elements/1.1/creator>
  dcterms:description "Examples of a Creator include a person, an organization, or a service. Typically,
  dcterms:hasVersion <http://dublincore.org/usage/terms/history/#creator-006> ;
  dcterms:issued "1999-07-02"^^<http://www.w3.org/2001/XMLSchema#date> ;
  dcterms:modified "2008-01-14"^^<http://www.w3.org/2001/XMLSchema#date> ;
  a rdf:Property ;
  rdfs:comment "An entity primarily responsible for making the resource."@en ;
  rdfs:isDefinedBy <http://purl.org/dc/elements/1.1/> ;
  rdfs:label "Creator"@en ;
  skos:note "A second property with the same name as this property has been declared in the dcterms: name
document \"DCMI Metadata Terms\" (http://dublincore.org/documents/dcmi-terms/) for an explanation."@en .
```

# The Object

- Determines the value of the attribute, or the object of the subject
- Can be a resolvable URI or a literal



# URI Abbreviations

- A URL can be abbreviated into a prefix:suffix, such as dc:Creator
- Here are some commonly used abbreviations in RDF files, but you can always define your own abbreviations if needed.
- Each one of these domains contains definitions for standard RDF vocabulary that are commonly used in Linked Data.

Prefix	Namespace URI
dc:	<a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/</a>
foaf:	<a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/</a>
rdf:	<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>
rdfs:	<a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#</a>
vcard:	<a href="http://www.w3.org/2006/vcard/ns#">http://www.w3.org/2006/vcard/ns#</a>

## *Guidelines for Minting URIs*

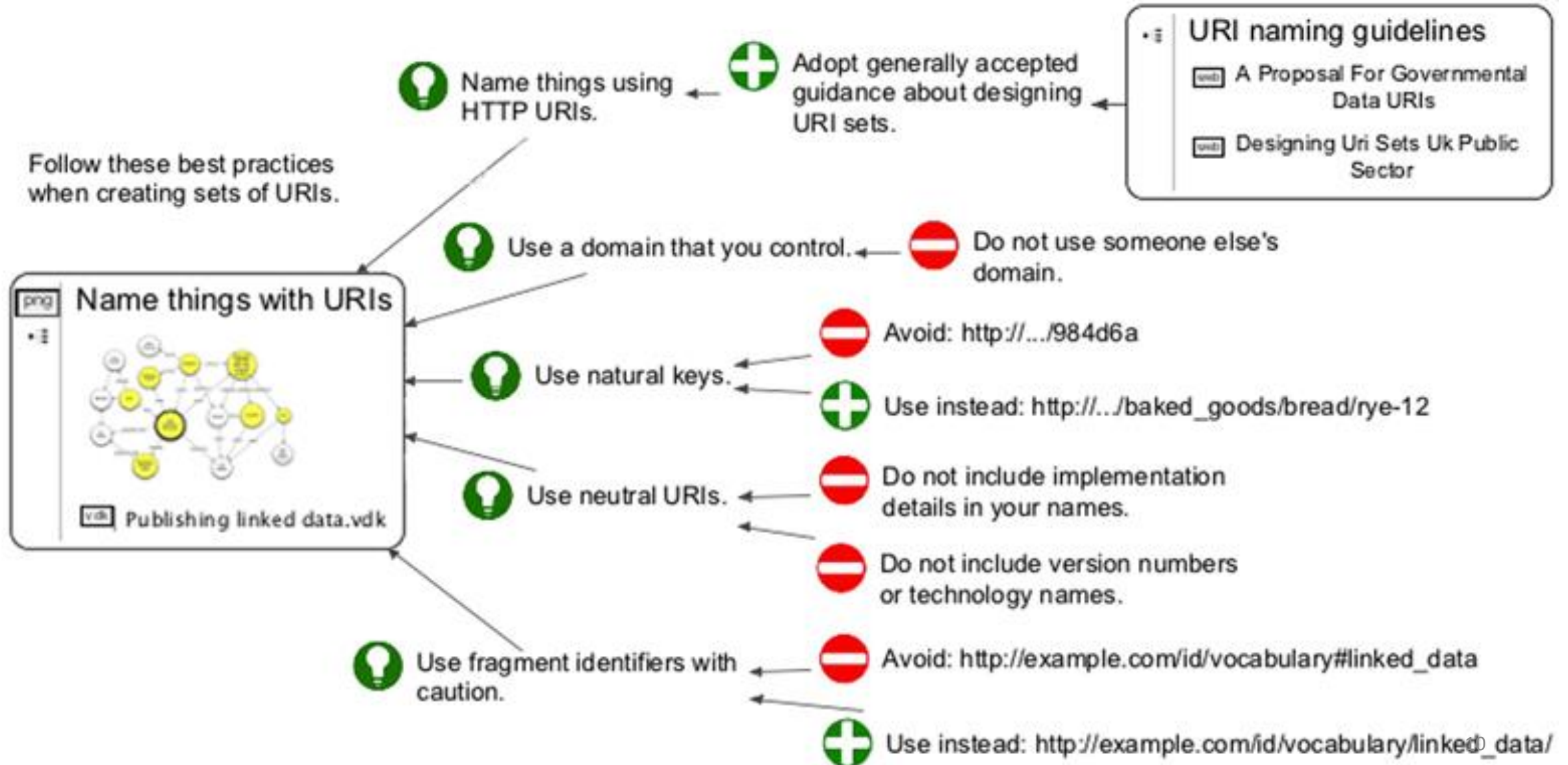
- Name things with URIs most of the time.
- Use a DNS domain that you control.
- Use natural keys that can be easily understood by humans. For example, do not use  
(<http://paulsbakery.example.com/984d6a>)  
but use:  
([http://paulsbakery.example.com/baked\\_goods/bread/rye-12](http://paulsbakery.example.com/baked_goods/bread/rye-12)).
- Make your URIs neutral to implementation details. For example, do not use  
<http://example.com/index.aspx>



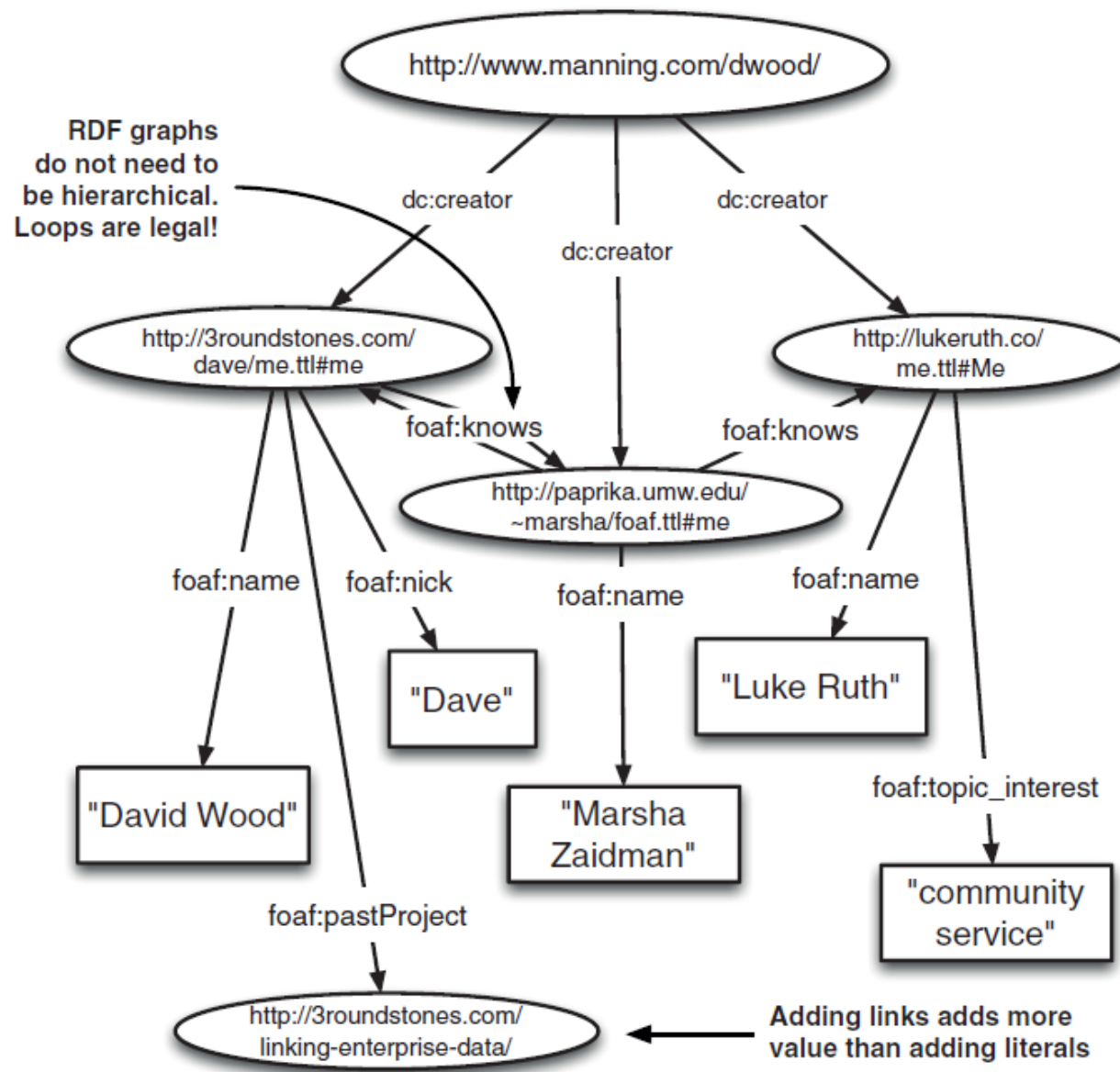
## *Fragment Identifiers (the #)*

- Fragment identifiers are the part of the URL that follows the hash symbol (#)
- Fragment identifiers are often used in HTML pages to point to a particular section in the page
- Fragment identifiers are not passed to web servers by web clients
- This means that if you try to resolve a URI such as [http://example.com/id/vocabulary#linked\\_data](http://example.com/id/vocabulary#linked_data), your browser will just send <http://example.com/id/vocabulary> to the server.
- Fragment identifiers are used by many Linked Data vocabularies because **the vocabulary is often served as a document and the fragment is used to address a particular term within that document.**

# Minting URIs Summary

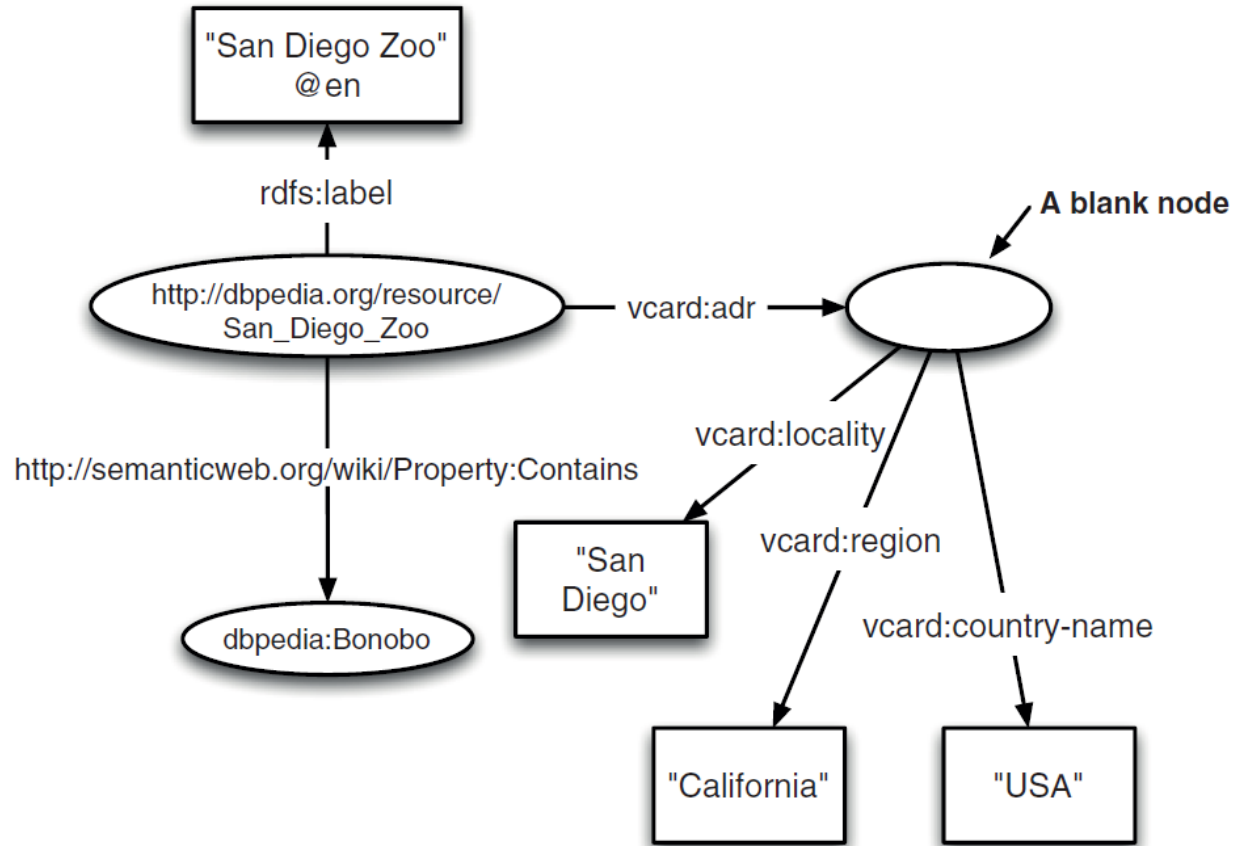


# More Comprehensive RDF example



# Blank Nodes

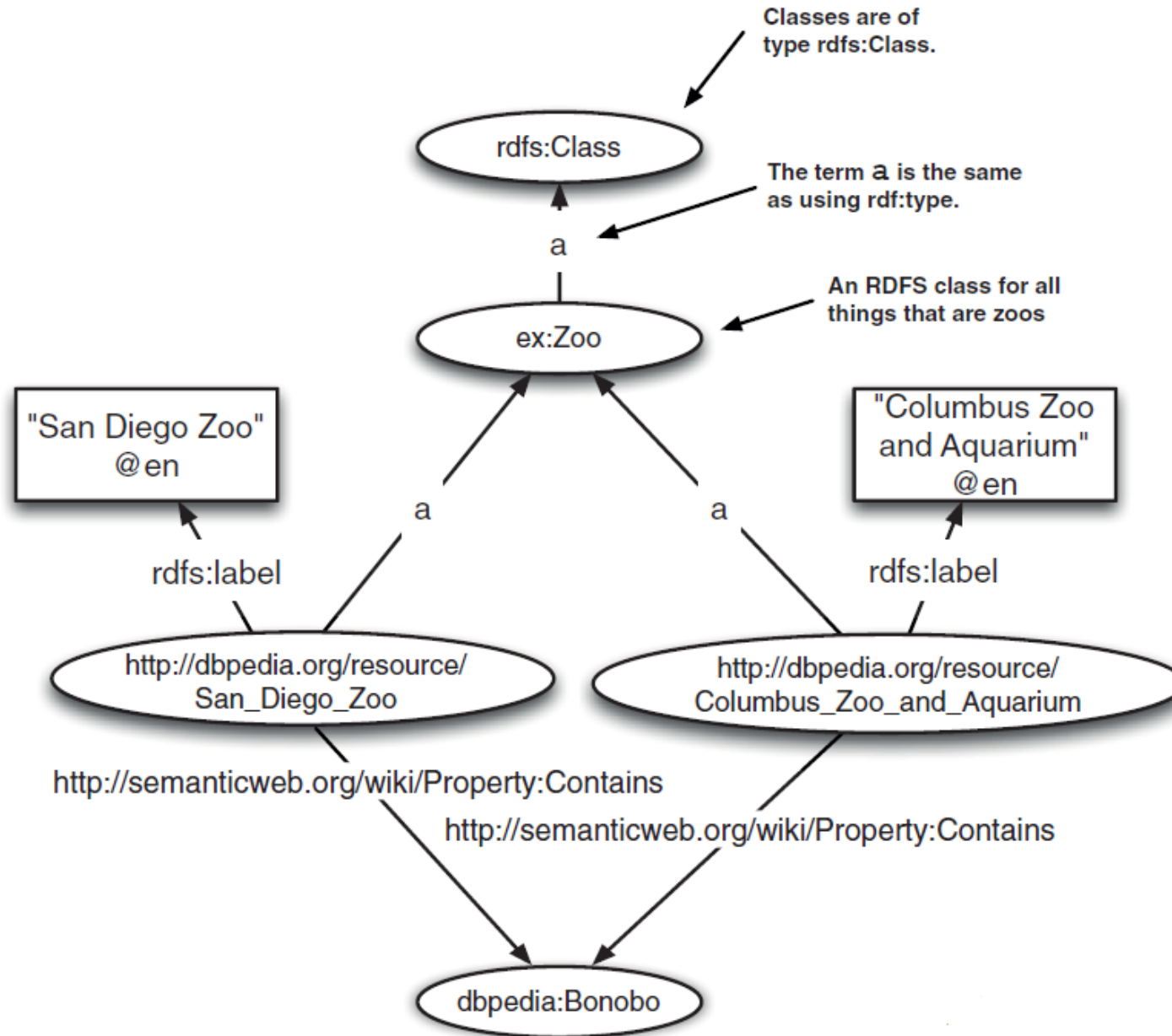
- Useful when you need to link to a collection of items but don't want to bother making up a URI for it
- Some RDF databases automatically assign URIs to blank nodes so they may be more easily operated upon. This process is known as *skolemization*
- In general, avoid using blank nodes as much as possible



# *Classes*

- RDF resources may be divided into groups called classes using the property `rdf:type` in the RDF Schema (RDFS) standard
- The members of a class are known as instances of the class, just as they are in object-oriented programming
- RDFS classes are themselves RDF resources and are of type `rdfs:Class`.
- The `rdfs:subClassOf` property may be used to state that one class is a subclass of another.

# Class Example



# ***RDF Vocabulary***

- RDF vocabularies provide definitions of the terms used to make relationships between data elements
- RDF vocabularies are distributed over the Web, are developed by people all over the world, and only come into common use in Linked Data if a lot of people choose to use them.
- We have already seen a number of terms used in Linked Data such as foaf:name, rdfs:label, and vcard:locality.
- These terms are grouped together to form RDF vocabularies, examples:
  - The terms rdfs:label, rdfs:comment, and rdfs:seeAlso are all defined in the RDF Schema vocabulary (<http://www.w3.org/2000/01/rdf-schema#>)
  - The terms vcard:locality, vcard:region, and vcard:country-name all come from the vCard vocabulary associated with the vcard prefix (<http://www.w3.org/2006/vcard/ns#>)

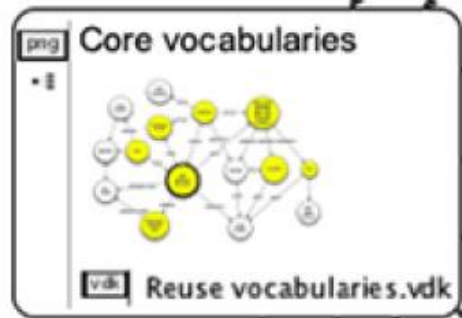
# *Who Mints RDF Vocabularies*

- Anyone can make an RDF vocabulary, and many people do.
- This would seem to be a recipe for disaster. How can anyone reuse Linked Data if it contains terms that you've never seen before?
- There are two ways to make this problem tractable:
  - Make certain that the URIs defining Linked Data vocabularies themselves follow the Linked Data principles
  - Reuse existing vocabularies whenever possible.



# Core RDF Vocabularies

Use terms from these core vocabularies to describe commonly understood data.



? Naming things? ← Use `rdfs:label`, `foaf:name`, `skos:prefLabel`.

? Describing people? ← Use FOAF, vCard.

FOAF: Friend of a Friend

? Describing addresses? ← Use vCard.

vCard: Virtual Contact File

? Describing projects? ← Use Description of a Project (DOAP).

? Describing web pages and other publications? ← Use `dc:creator` and `dc:description`.

dc: Dublin Core

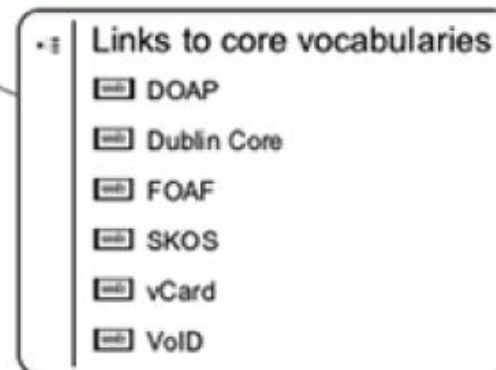
? Describing an RDF vocabulary? ← Use a VoID description.

? Describing existing taxonomies? ← Use SKOS.

SKOS: Simple Knowledge Organization System

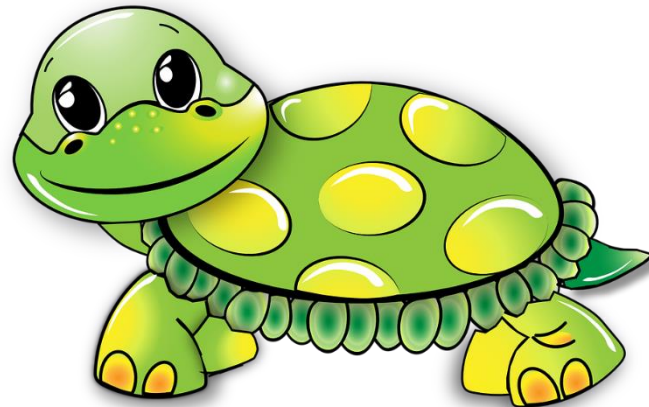
See also

Authoritative vocabularies.vdk

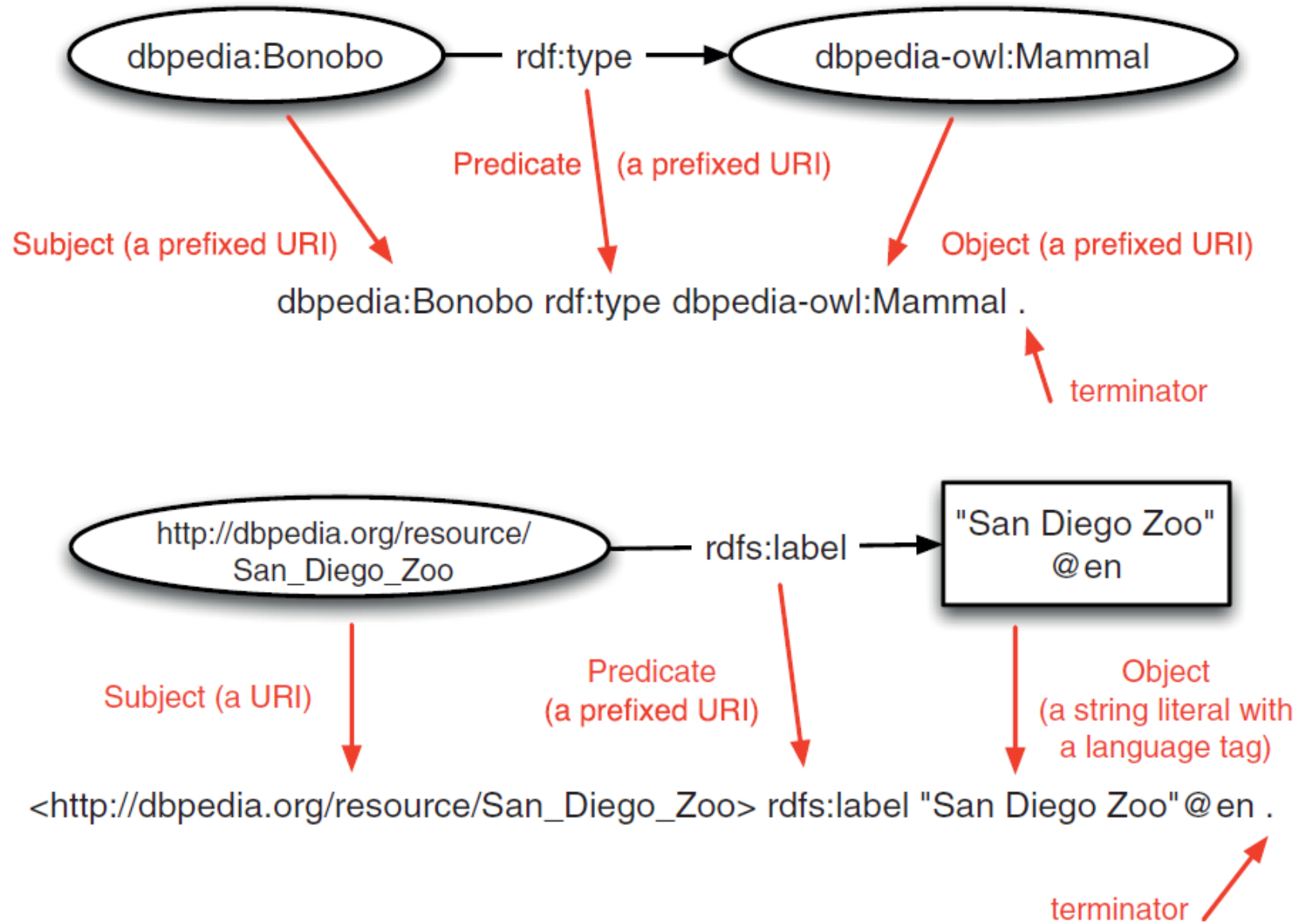


# ***RDF formats for Linked Data***

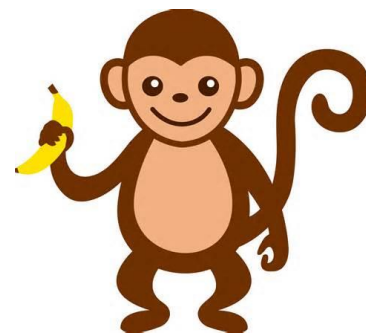
- RDF is a data model, not a format.
- Different formats can be used to serialize (express) RDF data
- Some of the most commonly used formats are:
  - Turtle—A simple, human-readable format
  - RDF/XML—The original RDF format in XML
  - RDFa—RDF embedded in HTML attributes
  - JSON-LD—A newer format aimed at web developers



# Turtle: Terse RDF Triple Language



## Bonobo example data in Turtle format



```
@prefix dbpedia:      <http://dbpedia.org/resource/> .
@prefix dbpedia-owl:  <http://dbpedia.org/ontology/> .
@prefix foaf:         <http://xmlns.com/foaf/0.1/> .
@prefix ex:          <http://example.com/> .
@prefix rdf:          <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs:         <http://www.w3.org/2000/01/rdf-schema#> .
@prefix vcard:        <http://www.w3.org/2006/vcard/ns#> .
@prefix xsd:          <http://www.w3.org/2001/XMLSchema#> .
```

Prefixes are  
defined first

Every statement must  
end with a period

**dbpedia:Bonobo**

```
    rdf:type      dbpedia-owl:Eukaryote , dbpedia-owl:Mammal ,
dbpedia-owl:Animal ;
    rdfs:comment  "The bonobo, Pan paniscus, previously called the pygmy
    chimpanzee and less often, the dwarf or gracile chimpanzee, is a great
    ape and one of the two species making up the genus Pan; the other is Pan
    troglodytes, or the common chimpanzee. Although the name \"chimpanzee\"
    is sometimes used to refer to both species together, it is usually
    understood as referring to the common chimpanzee, while Pan paniscus is
    usually referred to as the bonobo."@en ;
    foaf:depiction <http://upload.wikimedia.org/wikipedia/commons/a/a6/
    Bonobo-04.jpg> ;
    foaf:name      "Bonobo"@en ;
    rdfs:seeAlso   <http://eol.org/pages/326448/overview>
```

Commas separate  
multiple objects of the  
same predicate

Semicolons separate  
multiple predicates of  
the same subject

# *That's it Folks*



Further Reading

Chapter 2: Linked Data Structured Data on the Web, David Wood et al.