LOD-LAM

Jenel Farrell - SAA 2011
Metadata & Digital Object Roundtable

LINKED OPEN DATA
The current web is a web of documents intended for human interpretation.
Data is locked in closed silos

The **linked data vision** is to...
• Move away from the repository-centric mind-set and open data silos
• Publish data applications computers can access & interpret
• Use “standards” in the form of common Web technologies

Making it possible for applications to look up (meta)data...
...about things
books concepts people places
by looking up their names
http://dbpedia.org/resource/The_Metamorphosis
on the web

But, to allow applications to look up metadata in your database, we need some technology
Uniform Resource Identifier (URI)

- URI is a string of characters used to identify (UI) a name or resource on the internet
- URIs are more specific than URLs, which name the location of something. URIs are the ‘thing’

Resource Description Framework (RDF)

- In addition to having a unique identifier for your ‘thing’, you need RDF, which is a model for representing metadata on the web
- This representation comes in the form of statements, which are called ‘triples’
- Triples define the relationships between things
- Language for describing vocabularies (there are others)
- Way to represent info on the web
Simple Knowledge Organization System (SKOS)

• A language (built on RDF) for describing controlled vocabularies
  • thesauri
  • classification schemes
  • taxonomies
  • subject-heading systems

SPARQL

• And, of course, a language to query the data
• SPARQL is a query language built on RDF
DBPedia...

The main objective of DBpedia is to extract structured information from Wikipedia, convert it into RDF, and make it freely available on the Web. In a nutshell, DBpedia is the Semantic Web mirror of Wikipedia.

In a nutshell, the basics of

LINKED OPEN DATA
Make your stuff available on the web (whatever format) under an open license

Make your data available as structured data (ex. CSV table instead of an image scan of a table)
Use **non-proprietary formats** (CSV instead of Excel)

Use **URIs to identify things**, so that people can point at your stuff
Link your data to other data to provide context.

Linked data is essential to connect the semantic web.
By linking up data, we gain access to many more databases & discover new connections