Learning, Understanding, and Using Linked Data

Linked Data

1 2 3 4 5

1 – One goal
2 – Two types of questions
3 – RDF triples
4 – Four principles
5 – Five star LOD

Learn by Understanding

Learn by Analyzing

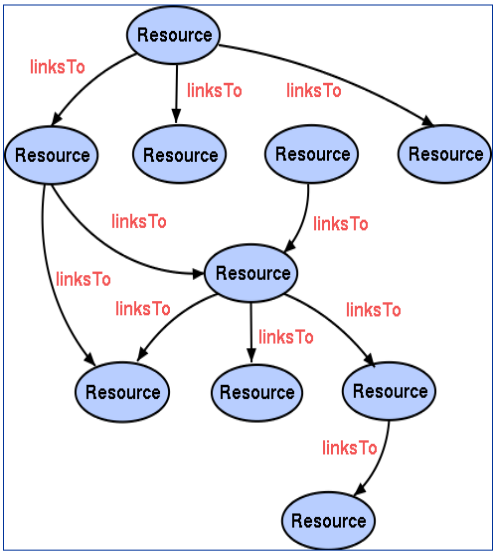
1

Why Linked Data?

One Goal: Breaking the silos, linking data

Image: Slide from T.B.Lee at TED 2009 conference, "The Great Unveiling" in Long Beach, CA. USA, 4, Feb 2009

From "Web of Documents" →to → "Web of Data"



"web of documents"

Resources:
identified by URI's; un-typed

Links:
href, src, ... limited, non-descriptive

Machine:
Very little information available - significance of the links only evident from the context around the anchor.

Source: W3C Semantic Web Activity - http://www.w3.org/2001/sw/

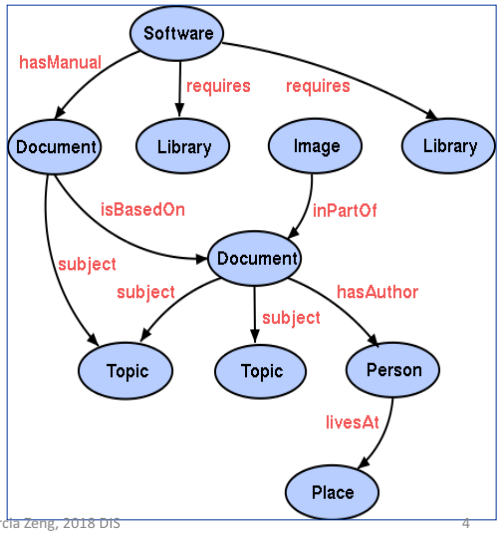
The Semantic Web

"web of data"

Resources:
Globally Identified by URI's; typed

Links:
Identified by URI's; descriptive

Machine:
More processable information is available



Linked Data

-- is about:

- using the Web to connect related data that was not previously linked,
- using the Web to lower the barriers to linking data currently linked using other methods. [1]

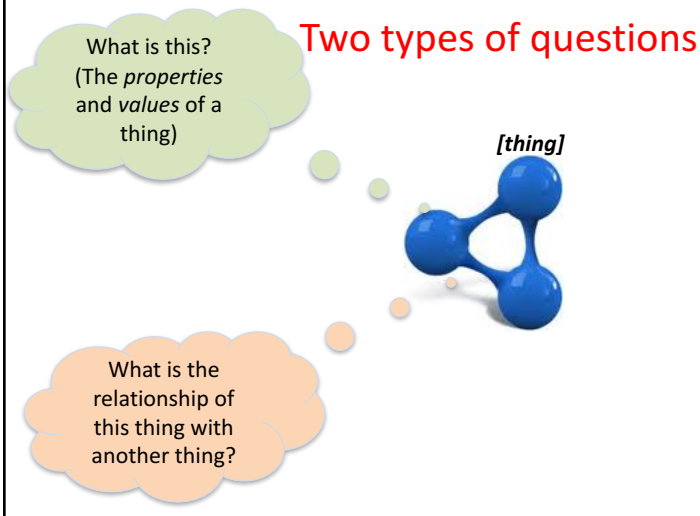
-- is a term used to describe a method of exposing, sharing, and connecting data on the Web using URIs and RDF

HOW?

[1] http://linkeddata.org/

2

Two types of questions



3

Use triples to express answers

What is this? (The *properties* and *values* of a thing)

[thing]	*[property]*	*[value]*
	name:	Steve Paul Jobs
	name:	Steve Jobs
	dateOfBirth:	1955－02－24
	Spouse:	Laurene Powell (m. 1991-2011)
	Education:	Reed College
	placeOfBirth:	San Francisco, CA

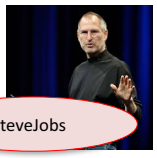
What are independent things here? (A thing would have its own properties）

[thing]	*[property]*	*[value]*
	name:	Steve Paul Jobs
	name:	Steve Jobs
	dateOfBirth:	1955－02－24
	Spouse:	Laurene Powell
	Education:	Reed College
	placeOfBirth:	San Francisco, CA

What is the relationship of this thing with another thing?

If we identify every **thing** with an ID …

*[**ex**: = our example namespace]*

[thing]	*[property]*	*[value]*
 ex:SteveJobs	name:	Steve Paul Jobs
	name :	Steve Jobs
	dateOfBirth:	1955—02—24
	Spouse :	Laurer ex:LaurenePowell
	Education :	Reed (ex:ReedCollege
	placeOfBirth:	San Fr ex:SanFrancisco

If we identify every **thing** with an ID …

we will be able to produce triples

ex:SteveJobs

name: Steve Jobs

ex:SteveJobs

spouse: ex:LaurenePowell

ex:SteveJobs

birthPlace: ex:SanFrancisco

[thing]

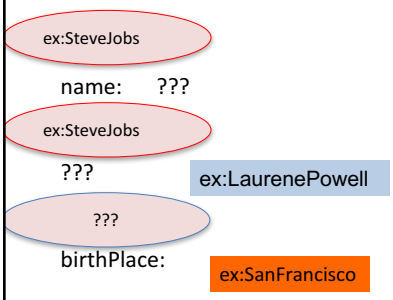
ex:SteveJobs

ex:LaurenePowell

ex:ReedCollege

ex:SanFrancisco

What questions can a triple answer?



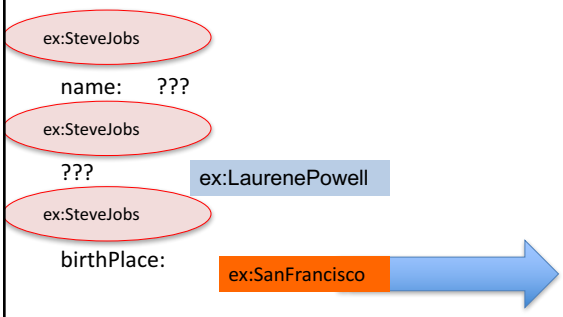
What is the name of this person ?

What was his relationship with Laurene Powell?

Who was born in San Francisco, California?

Could an "object" also become a "subject"?

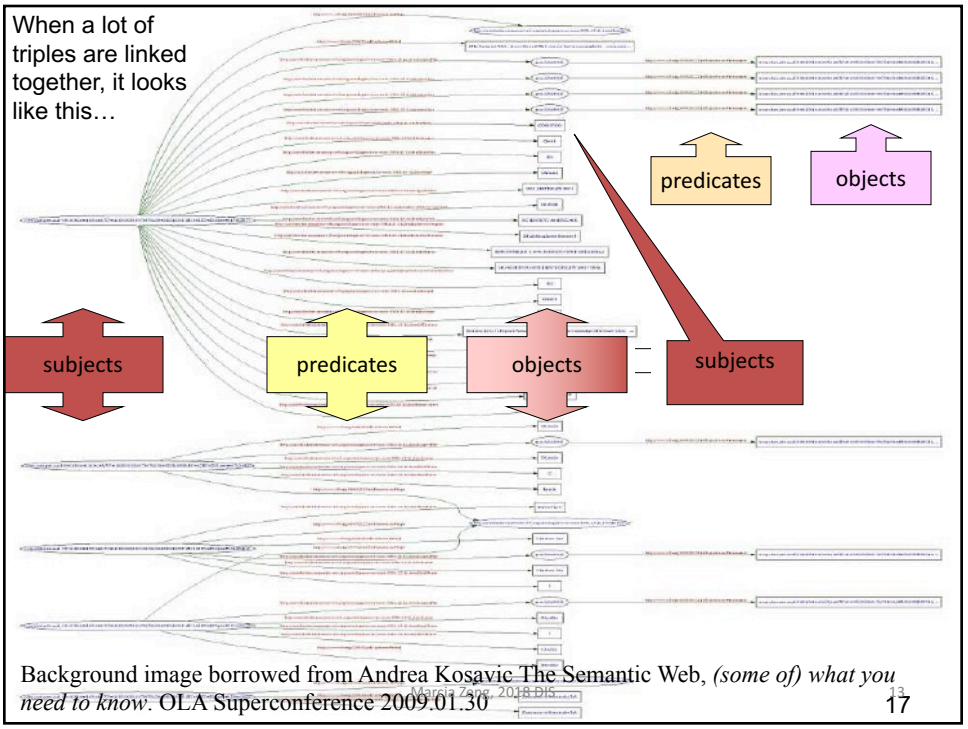
Can it reveal more triples?



Let's think about San Francisco, what other 'properties' can you think of?

Among these, what are identifiable 'things'?

When a lot of triples are linked together, it looks like this…



Background image borrowed from Andrea Kosavic The Semantic Web, *(some of) what you need to know*. OLA Superconference 2009.01.30

3 (cont.)

RDF Triples

- RDF = **R**esource **D**escription **F**ramework
  - RDF is a language for representing information about resources in the World Wide Web.
    - RDF is to data as HTML is to documents.
      If you want to create a webpage, you use HTML.
- The RDF model is based on the principle of making logical statements about resources in the form of ***subject-predicate-object*** expressions (called ***triples*** in RDF terminology).
- The greatest use of RDF is
  - not restricted to encoding information about *web resources*;
  - used to provide information about, and relations between, ***things*** in the real world: people, places, concepts, and the like.

3 (cont.)

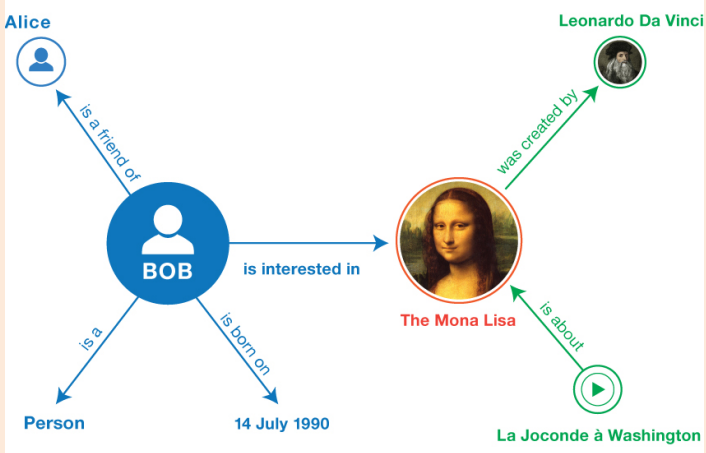
RDF Triples

http://www.example.org/index.html has a creator whose value is John Smith

The RDF terms for the various parts of the statement are:

- the *subject* is the URL **http://www.example.org/index.html**
- the *predicate* is the word "**creator**"
- the *object* is the phrase "**John Smith**"

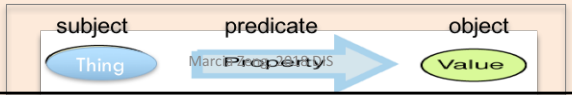
15



<Bob> <is a> <person>.
<Bob> <is a friend of> <Alice>.
<Bob> <is born on> <the 4th of July 1990>.
<Bob> <is interested in> <the Mona Lisa>.

<Leonardo da Vinci>
<is the creator of> <the Mona Lisa>.
<The video 'La Joconde à Washington'>
<is about> <the Mona Lisa>.

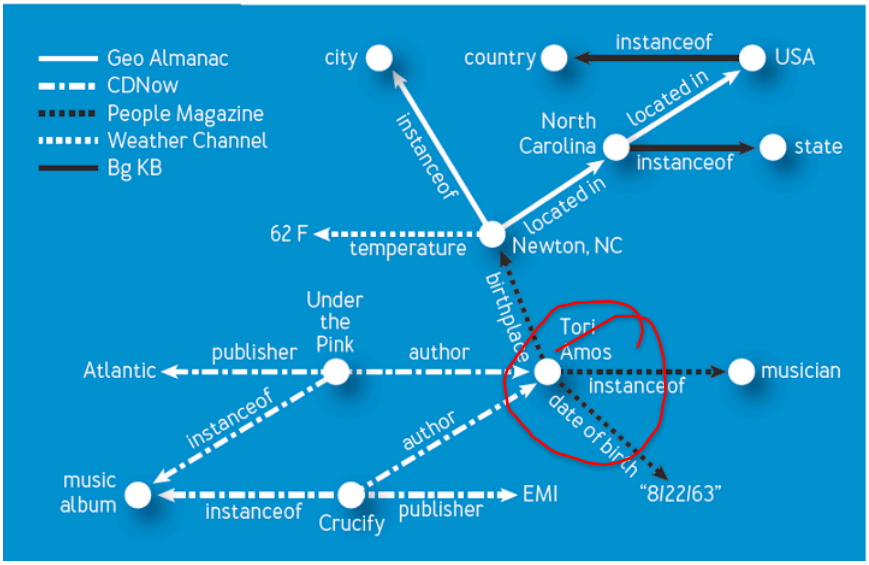
https://www.w3.org/TR/rdf11-primer/



16

That is how the triples in knowledge bases from silos can be connected.

FIGURE 1: **EXAMPLE OF A KNOWLEDGE BASE SOURCED FROM MULTIPLE SITES**



Guha, Brickley, and Macbeth, 2015. Schema.org: Evolution of Structured Data on the Web
http://queue.acm.org/detail.cfm?id=2857276