

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

Marcia Zeng, 2018 DIS

4 Linked Data Principles

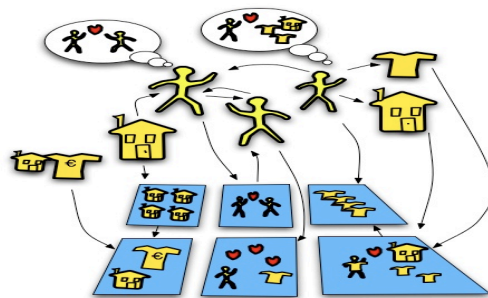
1. Use **URLs** as names for things.
2. Use **HTTP** URLs so that people can look up those names.
3. When someone looks up a URL, provide useful information, using the standards (RDF*, SPARQL).
4. Include links to other URLs so that they can discover more things.

Tim Berners-Lee: [Design Issues]

<http://www.w3.org/DesignIssues/LinkedData.html>

2

#1. Use URIs/IRIs as names for things



personA = [URI ...]
placeB = [URI ...]
eventC = [URI...]

Everything has a URI

Don't say "New York" say `<http://dbpedia.org/page/New_York>`

- URI= Uniform Resource Identifier (including URL and URN)
- In computing, a Uniform Resource Identifier (URI) is a string of characters used to identify a name or a resource on the Internet.

3

Marcia Zeng, 2018 DIS

Image credit: How it works: The Web <http://www.flickr.com/photos/danbri/2415237566>

Additional note: URIs / IRIs

- **URIs (Uniform Resource Identifiers)**

are limited to a subset of the ASCII character set.

<https://en.wikipedia.org/wiki/Himalayas>



- **IRIs (Internationalized Resource Identifier)**

allow non-ASCII characters such as non-English characters in the Universal Character Set (Unicode) to be used in the IRI character string.

<https://zh.wikipedia.org/wiki/喜马拉雅山脉>

https://my.wikipedia.org/wiki/တိမ္မော့_တောင်တန်း

4

Marcia Zeng, 2018 DIS

7

2. Use **HTTP URIs** so that people can look up those names

http://

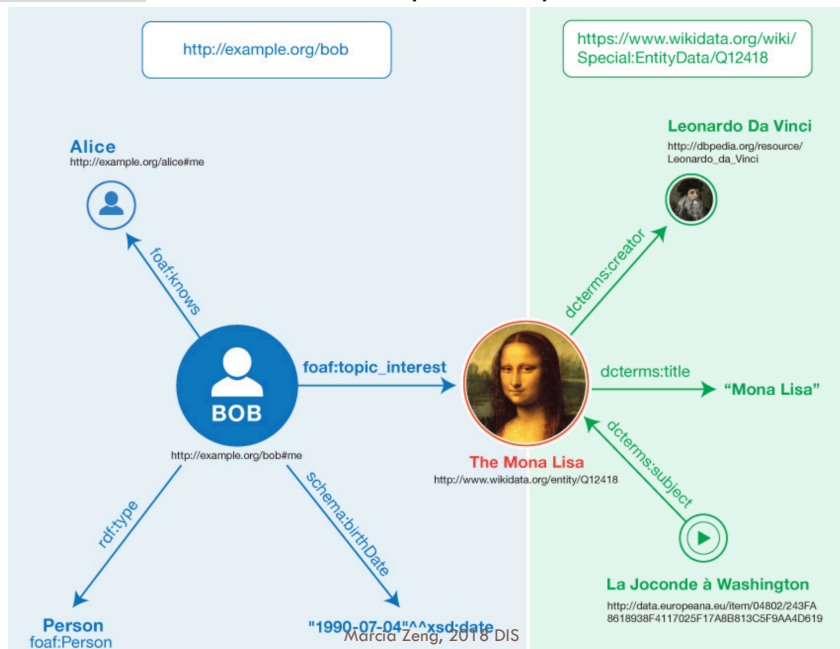
- The Hypertext Transfer Protocol (HTTP) is a networking protocol for distributed, collaborative, hypermedia information systems.
- HTTP is the foundation of data communication for the World Wide Web.

5

Marcia Zeng, 2018 DIS

2. Use **HTTP URIs**

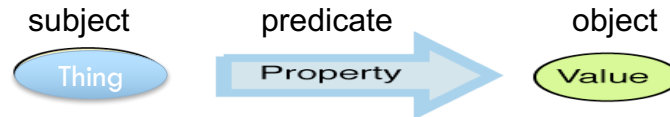
Back to the triple example



6

Marcia Zeng, 2018 DIS

#3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)



- *RDF (Resource Description Framework)
 - the language of the Semantic Web
- SPARQL (pronounced "sparkle")
 - an RDF query language

7

Marcia Zeng, 2018 DIS

http://dbpedia.org/page/University_of_San_Diego

thing About **University of San Diego**

An Entity of Type *Private university*, from Named Graph : <http://dbpedia.org>, within Data Space :

The University of San Diego (USD) is a private Roman Catholic university in San Diego, California. The university offers 42 baccalaureate degrees, and several degrees in law, nursing, (masters, PhD, and DNP), and other doctorate programs. The university comprises seven different academic colleges.

Property	Value
dbo:abstract	<ul style="list-style-type: none"> ▪ The University of San Diego (USD) is a private Roman Catholic university in San Diego, California. The university offers 42 baccalaureate degrees, and several degrees in law, nursing, (masters, PhD, and DNP), and other doctorate programs. The university comprises seven different academic colleges.
dbo:affiliation	<ul style="list-style-type: none"> ▪ dbr:Consortium_for_North_American_Higher_Education_Colleges ▪ dbr:Roman_Catholic ▪ dbr:Association_of_Catholic_Colleges_and_Universities ▪ dbr:National_Association_of_Independent_Colleges_and_Universities
dbo:athletics	<ul style="list-style-type: none"> ▪ dbr:Pioneer_Football_League ▪ dbr:NCAA_Division_I
dbo:campus	<ul style="list-style-type: none"> ▪ dbr:Urban_area
dbo:city	<ul style="list-style-type: none"> ▪ dbr:San_Diego
dbo:endowment	<ul style="list-style-type: none"> ▪ 3.956E8
dbo:facultySize	<ul style="list-style-type: none"> ▪ 875 (xsd:integer)

1. Use URIs as names for things
2. Use http URIs

3. When someone looks up a name, provide useful information, using the standards (RDF*, SPARQL)

Marcia Zeng, 2018 DIS

dbpedia.org/page/University_of_San_Diego

About: University of San Diego

An Entity of Type : Private university, from Named Graph : http://dbpedia.org, within Data Space : dbpedia.org

The University of San Diego (USD) is a private Roman Catholic university in San Diego, California. The university offers 42 baccalaureate degrees, and several degrees in law, nursing, (masters, PhD, and DNP), and other doctorate programs. The university comprises seven different academic colleges

Property	Value
dbo:abstract	<ul style="list-style-type: none"> The University of San Diego (USD) is a private Roman Catholic university in San Diego, California. The university offers 42 baccalaureate degrees, and several degrees in law, nursing, (masters, PhD, and DNP), and other doctorate programs. The university comprises seven different academic colleges
is dbp:workplaces of	<ul style="list-style-type: none"> dbr:Satyan_L dbr:Michael_Shulman_(mathematician)
is owl:differentFrom of	<ul style="list-style-type: none"> dbr:University_of_California,_San_Diego dbr:San_Diego_State_University
is foaf:primaryTopic of	<ul style="list-style-type: none"> wikipedia-en:University_of_San_Diego
dbo:enclowment	3.9886E8
dbo:facultySize	875 (xsd:integer)

9

http://dbpedia.org/page/University_of_San_Diego

Marcia Zeng, 2018 DIS

3. When someone looks up a name, provide useful information, using the standards (RDF*, SPARQL)

dbpedia.org/sparql

Virtuoso SPARQL Query Editor

Default Data Set Name (Graph IRI)
http://dbpedia.org

Query Text

```
select distinct ?Concept where {[] a ?Concept} LIMIT 10
```

Results Format: HTML (The XML output is disabled, see details)

Execution timeout: 0 milliseconds (values less than 1000 are ignored)

Options: Strict checking of void variables

(The result can only be sent back to browser, not saved on the server, see details)

Run Query Reset

We will use SPARQL query templates to get datasets during the hands-on session.

My example here: Find from 'Artist' Class anyone who was influencedBy Pablo_Picasso and Henri_Matisse

Virtuoso SPARQL Query Editor

Default Data Set Name (Graph IRI)
http://dbpedia.org

Query Text

```
PREFIX dataset: <http://dbpedia.org/ontology/>
SELECT ?uri ?influencedBy
WHERE
{
  ?uri a dataset:Artist .
  ?uri dataset:influencedBy ?influencedBy .
  filter regex(?influencedBy, 'Pablo_Picasso', 'i') .
}
UNION
{
  ?uri a dataset:Artist .
  ?uri dataset:influencedBy ?influencedBy .
  filter regex(?influencedBy, 'Henri_Matisse', 'i') .
}
```

Results Format: Auto (The XML output is disabled, see details)

Execution timeout: 0 milliseconds (values less than 1000 are ignored)

Options: Strict checking of void variables

(The result can only be sent back to browser, not saved on the server, see details)

Run Query Reset

Marcia Zeng, 2018 DIS

3. When someone looks up a name, provide useful information, using the standards (RDF*, SPARQL)

dbpedia.org/sparql?default-graph-uri=http%3A%2F%2Fdbpedia.org&query=PRE

uri	influencedBy
http://dbpedia.org/resource/Ellsworth_Kelly	http://dbpedia.org/resource/Henri_Matisse
http://dbpedia.org/resource/Raoul_Dufy	http://dbpedia.org/resource/Henri_Matisse
http://dbpedia.org/resource/Karel_Appel	http://dbpedia.org/resource/Henri_Matisse
http://dbpedia.org/resource/Ben_Shahn	http://dbpedia.org/resource/Henri_Matisse
http://dbpedia.org/resource/Dick_Bruna	http://dbpedia.org/resource/Henri_Matisse
http://dbpedia.org/resource/Wifredo_Lam	http://dbpedia.org/resource/Henri_Matisse

We will use **SPARQL** query templates to get datasets during the hands-on session.

the results (portion)

Visualization of the results (portion) using GEPHI software

11 Marcia Zeng, 2018 DIS

Linked Data Principles

#4. Include links to other URIs so that they can discover more things

12 Marcia Zeng, 2018 DIS

dbpedia.org/page/University_of_San_Diego

DBpedia

is dbo:almaMater of

- dbr:Todd_Gloria
- dbr:Anna_Maria_Corazza_Bildt
- dbr:Thomas_J_Whelan_(judge)
- dbr:Miguel_Luis_Villafuerte
- dbr:Danny_Tarkanian
- dbr:Walid_Juffail
- dbr:Lorenzo_Fertitta
- dbr:Larry_Alán_Burns
- dbr:Judith_Won_Pat
- dbr:Ashley_Clements
- dbr:AJ_Agrawal
- dbr:Enrique_Morones
- dbr:Salvatore_Cordileone
- dbr:Fahed_Boodai
- dbr:Khaled_Juffail
- dbr:Ryan_Zinke
- dbr:Lowell_McAdam
- dbr:Frances_Townsend
- dbr:Heidi_Watney
- dbr:Gaétan_Barrette
- dbr:Dave_Cox
- dbr:Juan_Vargas
- dbr:Corey_Brown_(politician)
- dbr:Erik_Wahl
- dbr:Bill_Bavasi
- dbr:John_Carlos_Frey

http://dbpedia.org/page/University_of_San_Diego

Every person has a URI

dbpedia.org/page/John_Ca

About: John Carlos Frey

An Entry of Type : person, from Named Graph : http://dbpedia.org, within Data Space : dbpedia.org

John Carlos Frey (born November 3, 1969 in Tijuana, Mexico), is a Mexican-American freelance investigative journalist and documentary filmmaker based in Los Angeles. His investigative work has been featured on 60 Minutes, PBS, NBC News, CBS News, The Weather Channel, Dan Rather Reports, Fusion TV, Current TV, Univision and Telemundo to name a few. John Carlos Frey has also written articles for the Los Angeles Times, the Huffington Post, Salon, Need to Know online, the Washington Monthly, and El Diario (in Spanish).

Property	Value
dbo:abstract	<ul style="list-style-type: none"> ■ John Carlos Frey (born November 3, 1969 in Tijuana, Mexico), is a Mexican-American freelance investigative journalist and documentary filmmaker based in Los Angeles. His investigative work has been featured on 60 Minutes, PBS, NBC News, CBS News, The Weather Channel, Dan Rather Reports, Fusion TV, Current TV, Univision and Telemundo to name a few. John Carlos Frey has also written articles for the Los Angeles Times, the Huffington Post, Salon, Need to Know online, the Washington Monthly, and El Diario (in Spanish). Frey is currently a Reporting Fellow at the Investigative Fund and a Special Correspondent for The NewsHour on PBS. (en)
dbo:activeYears	
dbo:almaMater	
dbo:birthName	<ul style="list-style-type: none"> ■ Juan Carlos Frey (en)
dbo:birthPlace	<ul style="list-style-type: none"> ■ dbr:Tijuana ■ dbr:Mexico ■ dbr:Baja_California

4. Include links to other URIs so that they can discover more things

Marcia Zeng, 2018 DIS

http://dbpedia.org/page/John_Carlos_Frey

dbpedia.org/page/University_...

DBpedia

About: University of San Diego

An Entry of Type : Private university, from Named Graph : http://dbpedia.org, within Data Space : dbpedia.org

The University of San Diego (USD) is a private Roman Catholic university in San Diego, California. The university offers 42 baccalaureate degrees, and several degrees in law, nursing, (masters, PhD, and DNP), and other doctorate programs. The university comprises seven different academic colleges.

Property	Value
dbo:abstract	<ul style="list-style-type: none"> ■ The University of San Diego (USD) is a private Roman Catholic university in San Diego, California. The university offers 42 baccalaureate degrees, and several degrees in law, nursing, (masters, PhD, and DNP), and other doctorate programs. The university comprises seven different academic colleges. (en)
dbo:affiliation	<ul style="list-style-type: none"> ■ dbr:Consortium_for_North_American_Higher_Education_Collaboration ■ dbr:Roman_Catholic ■ dbr:Association_of_Catholic_Colleges_and_Universities ■ dbr:National_Association_of_Independent_Colleges_and_Universities
dbo:athletics	<ul style="list-style-type: none"> ■ dbr:Pioneer_Football_League ■ dbr:NCAA_Division_I
dbo:campus	<ul style="list-style-type: none"> ■ dbr:Urban_area
dbo:city	<ul style="list-style-type: none"> ■ dbr:San_Diego
dbo:endowment	<ul style="list-style-type: none"> ■ 3.956E8
dbo:facultySize	<ul style="list-style-type: none"> ■ 875 (xsd:integer)

4. Include links to other URIs so that they can discover more things

Marcia Zeng, 2018 DIS

http://dbpedia.org/page/University_of_San_Diego

dbpedia.org/page/San_Diego

DBpedia

1 About: San Diego

An Entity of Type: City (California), from Named Graph : <http://dbpedia.org>, within Data Space : dbp:

San Diego /sæn diˈeɪɡoʊ/ (Spanish for "Saint Didacus") is a major city in California on the Pacific Ocean in Southern California, approximately 120 miles (190 km) from the border with Mexico.

Property	Value
dbo:PopulatedPlace/areaTotal	<ul style="list-style-type: none"> 964.51 964.5115722891264
dbo:PopulatedPlace/populationDensity	1545.4
dbo:abstract	San Diego /sæn diˈeɪɡoʊ/ is a major city in California on the Pacific Ocean in Southern California, approximately 120 miles (190 km) from the border with Mexico.

is dbo:City of

- dbo:San_Diego_State_University_College_of_Education
- dbo:Point_Loma_High_School
- dbo:Point_Loma_Nazarene_University
- dbo:San_Diego_City_College
- dbo:San_Diego_Mesa_College
- dbo:University_of_San_Diego_High_School
- dbo:San_Diego_Miramar_College
- dbo:University_of_California,_San_Diego
- dbo:San_Diego_State_University
- dbo:University_of_San_Diego
- dbo:Brown_Field_Municipal_Airport
- dbo:KGB-FM
- dbo:The_Academy_of_Our_Lady_of_Peace
- dbo:Lincoln_High_School_(San_Diego)

4. Include links to other URIs so that they can discover more things

3. When someone looks up a name, provide useful information, using the standards

1. Use URIs as names for things

2. Use http URIs

Every place has its URI

Marcia Zeng, 2018 DIS

http://dbpedia.org/page/San_Diego

Is it machine understandable?

Let's have a look from another example. Go to: <http://rdfa.info/play/>

Human-readable webpage.

Machine-processable raw data

Machine-understand-able codes for this part of the webpage.

The dataset tells us:

```

<div vocab="http://xmlns.com/foaf/0.1/">
  <div resource="#manu" typeof="Person">
    <span property="name">Manu Sporny</span> knows
    <a property="knows" href="#alex">Alex</a> and
    <a property="knows" href="#brian">Brian</a>.
  </div>
  <div resource="#alex" typeof="Person">
    <span property="name">Alex Milowski</span> wrote the
    <div resource="#brian" typeof="Person">
      <span property="name">Brian Sletten</span> wrote the
    </div>
  </div>
</div>

```

Manu Sporny knows Alex and Brian.
Alex Milowski wrote the RDFa processor for this page.
Brian Sletten wrote the syntax highlighting for the raw data.

```

@prefix rdfa: <http://www.w3.org/ns/rdfa#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix rdfs: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

<http://rdfa.info/play/>
  rdfa:usesVocabulary foaf: .
<http://rdfa.info/play/#manu>
  rdf:type foaf:Person ;
  foaf:name "Manu Sporny" ;
  foaf:knows <http://rdfa.info/play/#alex> ;
  foaf:knows <http://rdfa.info/play/#brian> .
<http://rdfa.info/play/#alex>
  rdf:type foaf:Person ;
  foaf:name "Alex Milowski" .
<http://rdfa.info/play/#brian>
  rdf:type foaf:Person ;
  foaf:name "Brian Sletten" .

```

Visualization Raw Data

Web Page #manu

- type: Person
- name: Manu Sporny
- knows: #alex
- knows: #brian
- type: Person
- name: Alex Milowski
- type: Person
- name: Brian Sletten

Marcia Zeng, 2018 DIS

16

References

- Tim Berners-Lee. (2006) Linked Data. Last change: 2009/06/18. <http://www.w3.org/DesignIssues/LinkedData.html>
- The Linking Open Data cloud diagram. <http://lod-cloud.net/>
- Linking Open Data (LOD) Project Webpage. <http://linkeddata.org/>
- *Tim Berners-Lee speaks on Linked Data, (2009)*. TED 2009 conference, "The Great Unveiling" in Long Beach, CA. USA, 4, Feb 2009. Video: https://www.ted.com/talks/tim_berniers_lee_on_the_next_web
- *RDF 1.1 Primer*. (2014) W3C Working Group Note. <https://www.w3.org/TR/rdf11-primer/>
- LODLAM - Linked Open Data in Libraries, Archives, and Museums <http://lodlam.net/>
- LOV – Linked Open Vocabularies <http://lov.okfn.org/dataset/lov>