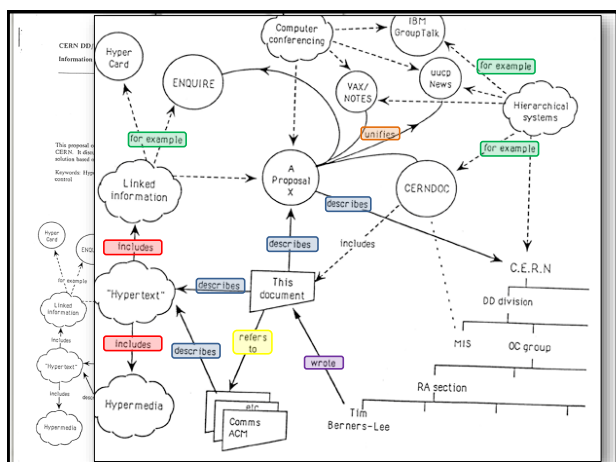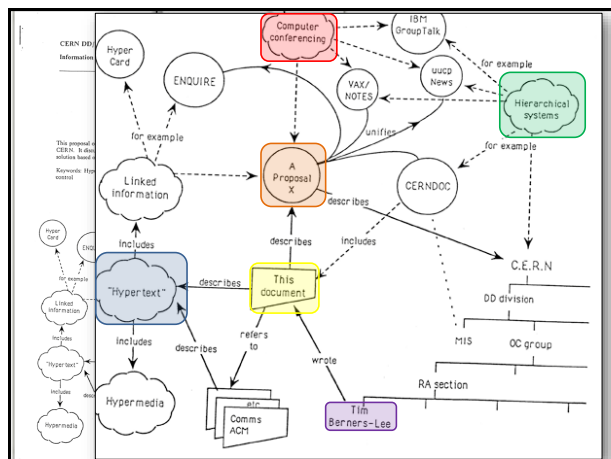INTRODUCTION TO
A WEB OF LINKED DATA
& A SEMANTIC WEB

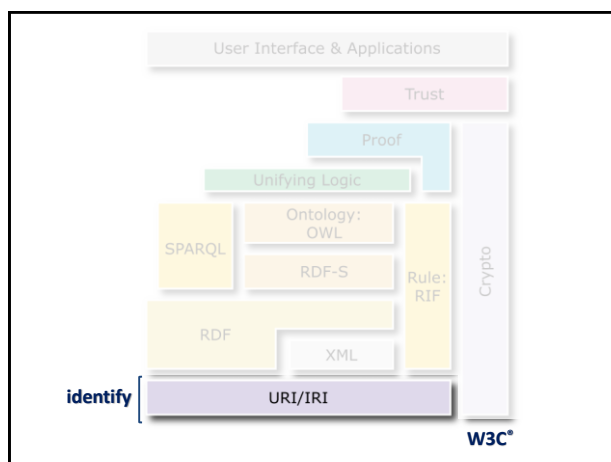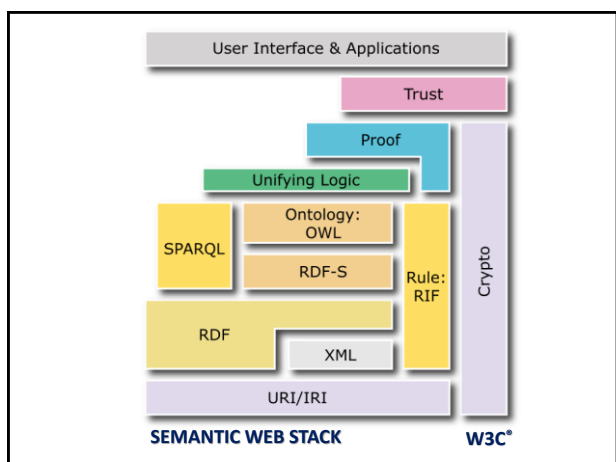Fabien GANDON, @fabien_gandon   http://fabien.info

semantic web
mentioned by Tim BL
in **1994** at WWW

[Tim Berners-Lee 1994, http://www.w3.org/Talks/WWW94Tim/]

User Interface & Applications
Trust
Proof
Unifying Logic
Ontology: OWL
SPARQL
RDF-S
Rule: RIF
Crypto
RDF
XML
URI/IRI
SEMANTIC WEB STACK
W3C°

User Interface & Applications
Trust
Proof
Unifying Logic
Ontology: OWL
SPARQL
RDF-S
Rule: RIF
Crypto
RDF
XML
identify URI/IRI
W3C°
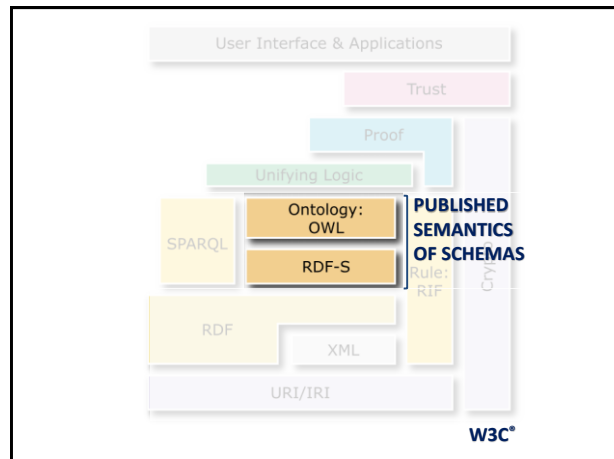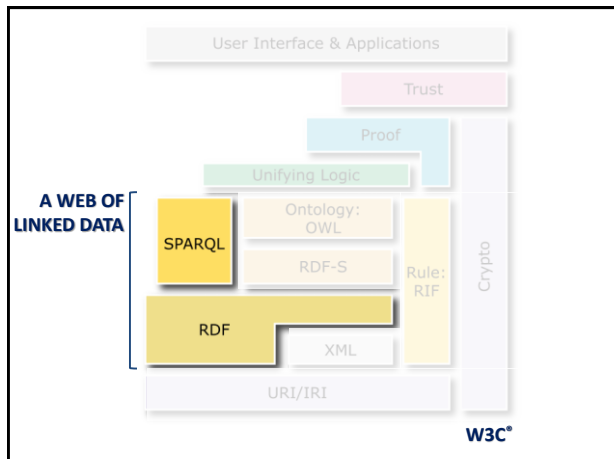
IDENTIFYING EVERYTHING ON THE WEB

What are the three keystones of the Web architecture?

?

propose your own languages (XML)

structure data

using tags

in a textual format

open standard family languages

composable languages

open non-proprietary

13

example of a name card

```
<card>
  <name>gandon</name>
  <tel type="office">+33492965170</tel>
  <page url="fabien.info"/>
</card>
```
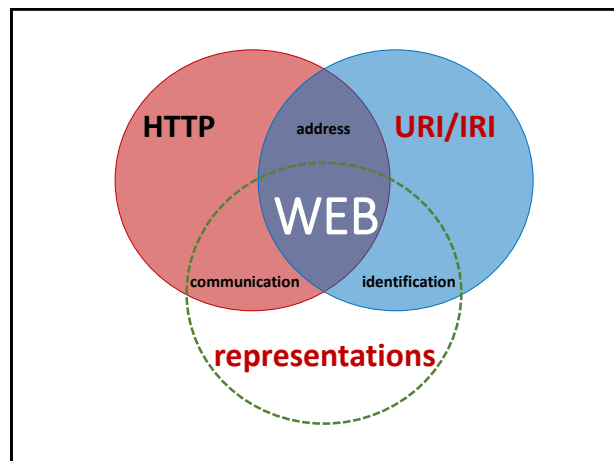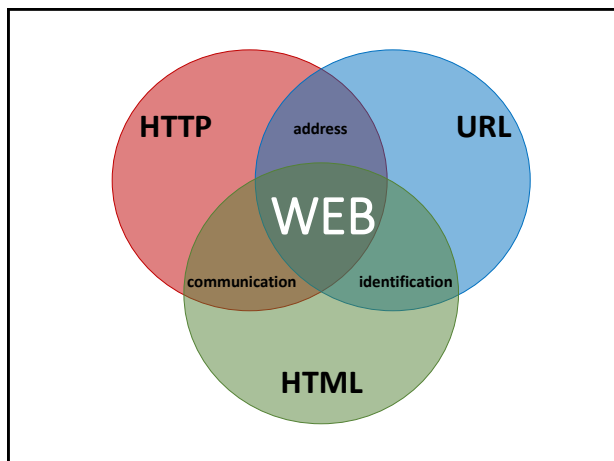
14

URL

identify what exists on the web.

http://my-site.fr

URL → URI

identify what exists on the web.

identify,
on the web,
what exists.

http://my-site.fr

http://animals.org/zebra#this

URL → URI → IRI

identify what exists on the web.

identify,
on the web,
what exists.

identify
on the web,
in any language,
what exists.

http://my-site.fr

http://animals.org/zebra#this

http://الحيوانات.tn/斑馬#this

URL → URI → IRI

RESOURCE

Definition: a resource is anything that can be identified by a URI.

http://fabien.info/objects#mycar

e.g. a page, a person, a car, a dog, an idea, a country, a product, a service...

19

ratatouille.fr

datatouille.fr

Linking Open Data

May 2007   April 2008   September 2008

March 2009

400
300
200
100
0
10/10/2006 28/04/2007 14/11/2007 01/06/2008 18/12/2008 06/07/2009 22/01/2010 10/08/2010 26/02/2011 14/09/2011 01/04/2012

September 2011   September 2010

Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. http://lod-cloud.net/

thematic content

Domains	Number of datasets	Number of Triples	%	Out links	%
Media	25	1 841 852 061	5,82 %	50 440 705	10,01 %
Geography	31	6 145 532 484	19,43 %	35 812 328	7,11 %
Government	49	13 315 009 400	42,09 %	19 343 519	3,84 %
Publications	87	2 950 720 693	9,33 %	139 925 218	27,76 %
Inter-domain	41	4 184 635 715	13,23 %	63 183 065	12,54 %
Life Sciences	41	3 036 336 004	9,60 %	191 844 090	38,06 %
Users' content	20	134 127 413	0,42 %	3 449 143	0,68 %
	295	**31 634 213 770**		**503 998 829**	

0%
6%
9%
10%
13%
42%
20%

- Government
- Geography
- Inter-domain
- Life Sciences
- Publications
- Media
- Users' content

surf on the Web of data

Fichier Édition Affichage Historique Marque-pages Outils ?
Nouvel onglet
Saisir un terme à rechercher ou une adresse

Practice with BBC

Great White Shark

```
http://www.bbc.co.uk/nature/life/Great_white_shark
http://www.bbc.co.uk/nature/life/Great_white_shark.rdf
```

Reference: "Current and future uses of Semantic Web technologies at the BBC"
http://raimond.me/slides/isemantics-2013/

---

a Web approach to data publication

« http://fr.dbpedia.org/resource/Paris »

URI ???...

---

a Web approach to data publication

« **http://**fr.dbpedia.org/resource/Paris »

HTTP URI

---

a Web approach to data publication

GET

HTTP URI

---

a Web approach to data publication

GET

303 redirect

HTTP URI

---

a Web approach to data publication

GET

303 redirect

GET

HTTP URI

a Web approach to data publication

GET
303 redirect
GET
HTML, ...

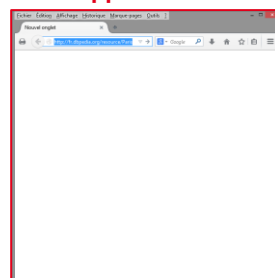HTTP URI

a Web approach to data publication

GET
303 redirect
GET
HTML, XML,...
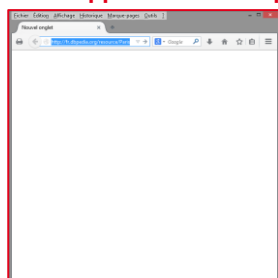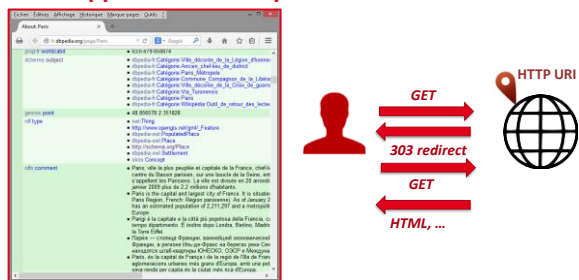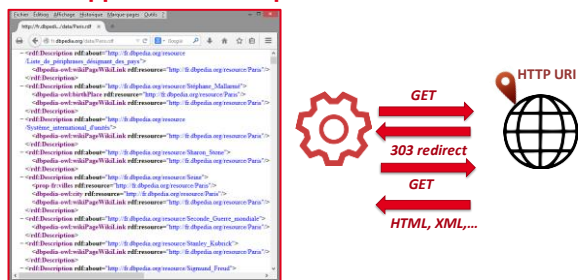
HTTP URI

linked data

linked data principles
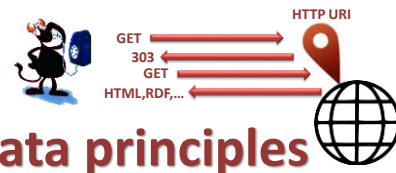
HTTP URI
GET
303
GET
HTML,RDF,...

- Use **RDF** as data format
- Use **HTTP URI**s as names for things so that people can look up those names
- When someone looks up a URI, provide useful information (RDF, HTML, etc.) using **content negotiation**
- Include **links to other URIs** so that related things can be discovered

content negotiation

HTTP URI
GET
303
GET
HTML,RDF,...

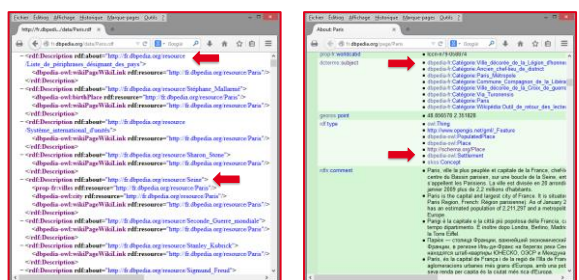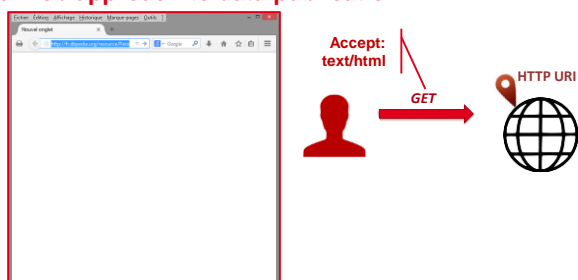- mechanism defined in the HTTP protocol specification
- serve different representation of a resource at the same URI
- user agents inform the servers of media types preferences (format, language, etc.)

```
Accept-Language: fr; q=1.0, en; q=0.5
Accept: text/html; q=1.0, text/*; q=0.8, image/gif; q=0.7,
image/jpeg; q=0.6, image/*; q=0.5, */*; q=0.1
```

- servers select the most suited representation

a Web approach to data publication

Accept:
text/html

GET

HTTP URI

a Web approach to data publication

Accept: text/html

GET

303 redirect

URL of HTML

HTTP URI

a Web approach to data publication

Accept: application/ rdf+xml

GET

303 redirect

URL of RDF/XML

HTTP URI
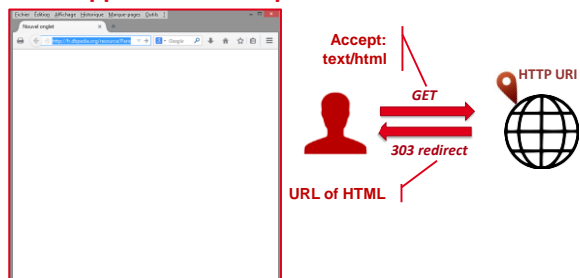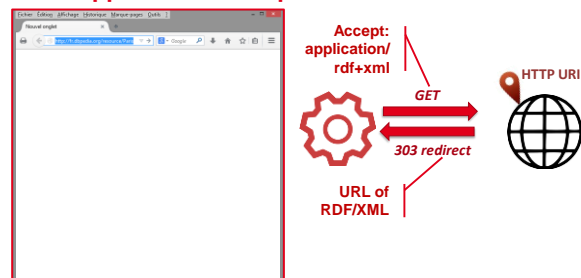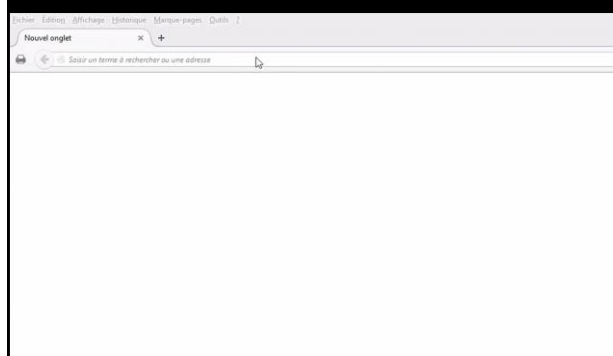
DBpedia demo

Practical Session

1. Find "London" on DBpedia.org
   e.g. Google: "london site:dbpedia.org"
2. Find `dbp:populationDemonym`
3. Find `rdf:type`
4. Find value `yago:CapitalsInEurope`
5. Find "Vienna"
6. Find its URI

use CURL to get data

C:\Users\gandon\Documents\mooc>
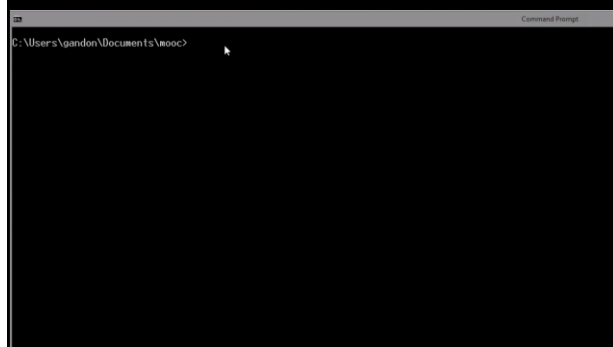
Practical Session

Do you have CURL? (windows=no, mac= yes, linux=?)
CURL : http://curl.haxx.se/
Installation wizard: http://curl.haxx.se/dlwiz/?type=bin

curl -o Paris.html -L http://dbpedia.org/resource/Paris

curl -o Paris-rdf-xml.txt -L -H "Accept: application/rdf+xml"
http://dbpedia.org/resource/Paris
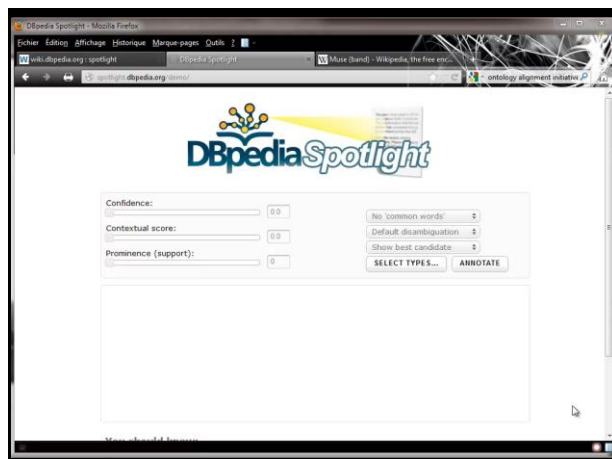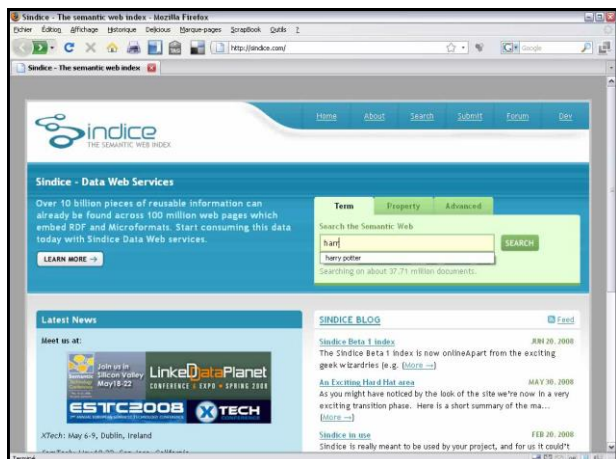
Free book !!!

Linked Data: Evolving the Web into a Global Data Space,

Tom Heath and Christian Bizer, Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool (2011)

http://linkeddatabook.com/

A WEB OF LINKED DATA
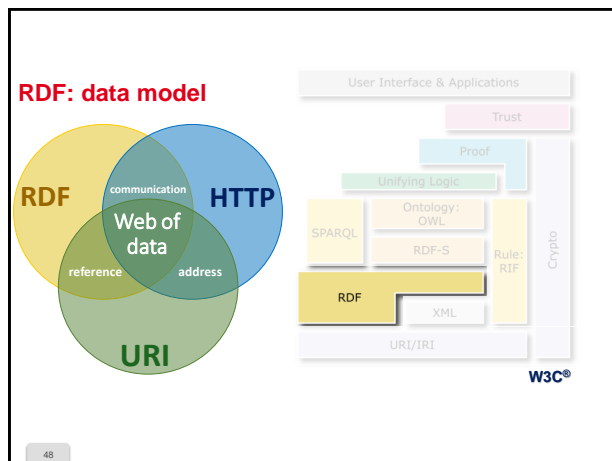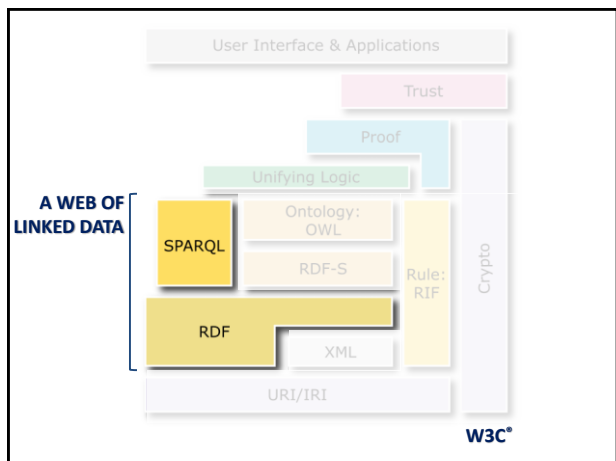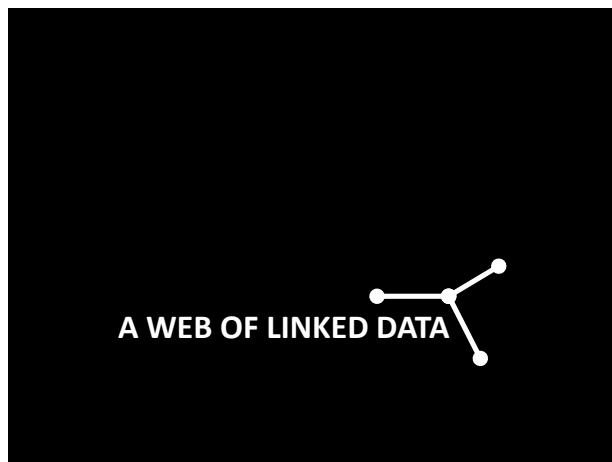
RDF stands for
Resource: *pages, dogs, ideas...*
everything that can have a URI
Description: *attributes, features, and*
relations of the resources
Framework: *model, languages and*
syntaxes for these descriptions

RDF is a triple model *i.e.* every
piece of knowledge is broken down into
(subject , predicate , object)

*doc.html has for author Fabien
and has for theme Music*

**doc.html has for author Fabien
doc.html has for theme Music**

**(doc.html , author , Fabien)
(doc.html , theme , Music)**

(subject , predicate , object)

Predicate

Object
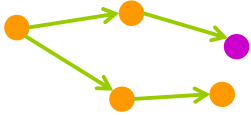
Subject

a triple
the RDF atom

RDF is also a graph model
to link the descriptions of resources

RDF triples can be seen as arcs
of a graph (vertex,edge,vertex)

(doc.html , author , Fabien)
(doc.html , theme , Music)

Fabien

author

doc.html

theme

Music

RDF is a model for **directed** labeled multigraphs

Fabien

author

doc.html

theme

Music
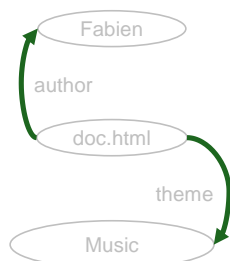
edges have a direction:
starting/head node (subject)
arrival/tail node(object)

59

RDF is a model for directed **labeled** multigraphs

Fabien

author

doc.html

theme

Music

edges and nodes have labels

60

10

RDF is a model for directed labeled **multigraphs**

Fabien

author

several edges/arcs between nodes/vertices

doc.html

contains   theme

Music

61

---

several points of views on a graph

r(a,b)
r(a,c)
r(c,a)

r	s	o
1	a	b
2	a	c
3	a	a

62
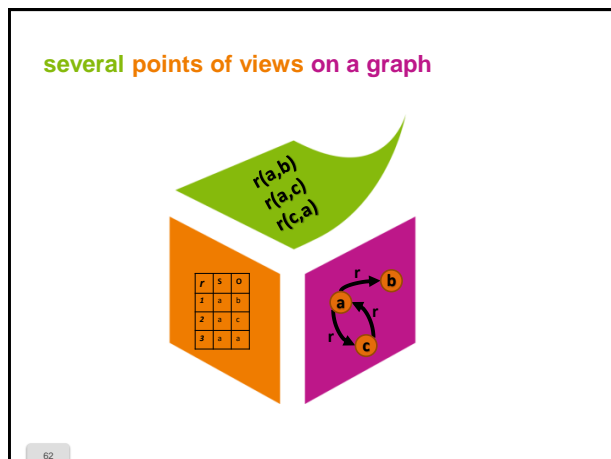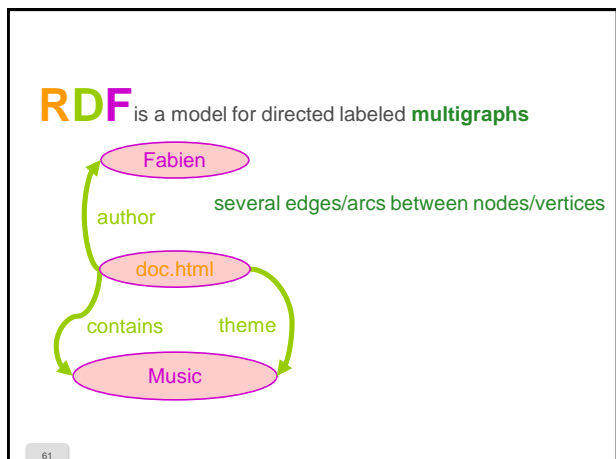
What is the mathematical structure built by the RDF triples?
(give the type of structure and its definition/explanation)

?

---

références au web

URL → IRI

identify what exists on the web

identify, on the web, what exists

http://my-site.fr

http://animals.org/this-zebra

---

http://ns.inria.fr/fabien.gandon#me

http://inria.fr/schema#author

http://inria.fr/rr/doc.html

http://inria.fr/schema#theme

Music

open and link data in a
global giant graph
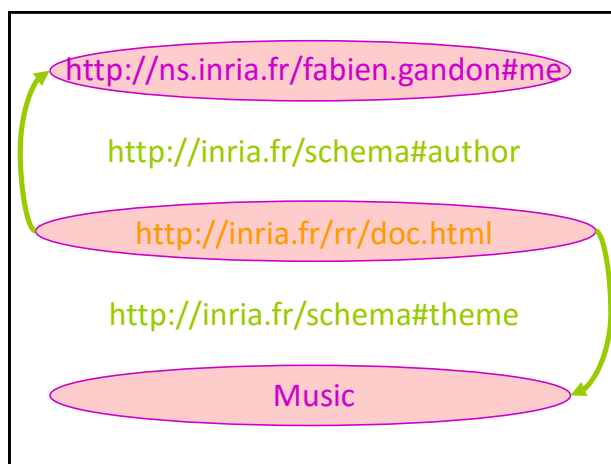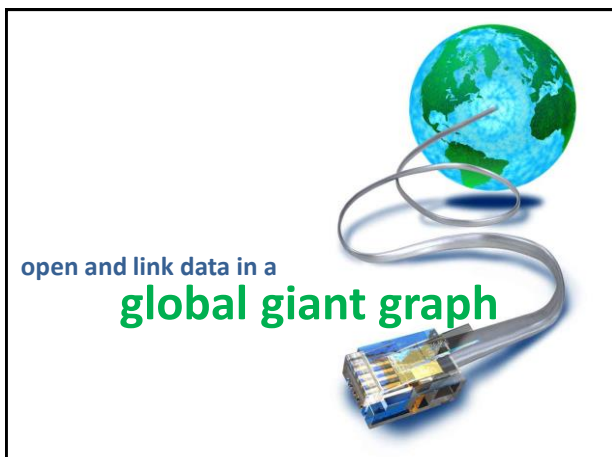
in **RDF** values of properties can also be literals i.e. strings of characters

(doc.html , author , Fabien)
(doc.html , theme , "Music")

http://ns.inria.fr/fabien.gandon#me

http://inria.fr/schema#author

http://inria.fr/rr/doc.html

http://inria.fr/schema#theme

"Music"

http://ns.inria.fr/fabien.gandon#me

http://inria.fr/schema#author

http://inria.fr/rr/doc.html

http://inria.fr/rr/doc.html

http://inria.fr/schema#theme

Music

see the graph

Slide 73
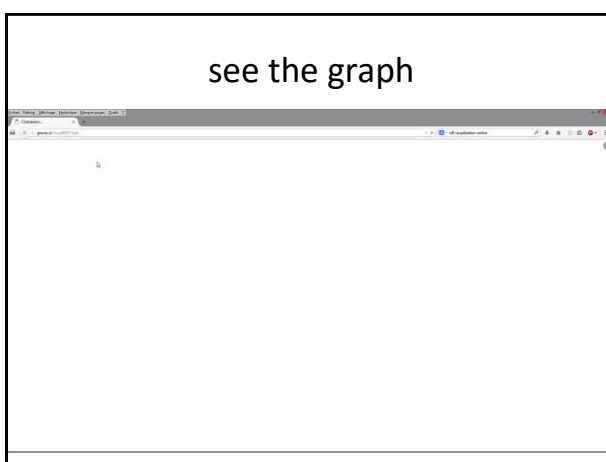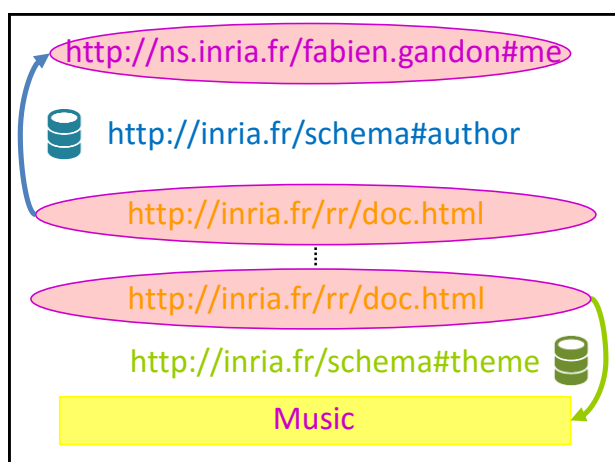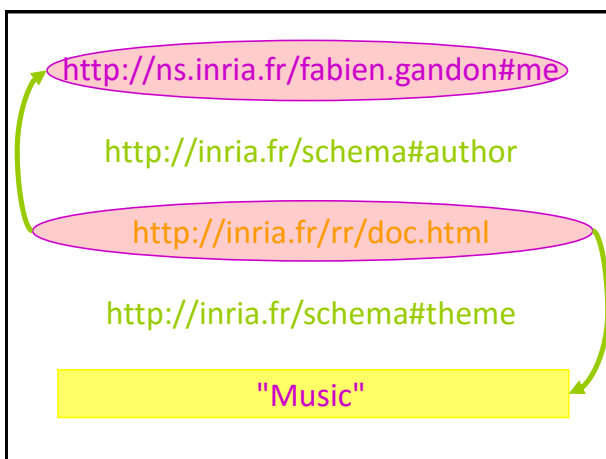
namespace

Definition: abstract space gathering names in a same set

http://inria.fr/sujets#compilation
http://mit.edu/org/Lab

e.g. a dictionary, a library index, a glossary, a standard, a thesaurus,...

73

Slide 74

NAMESPACES
- A collection of names identified by a URI
- Names belonging to a same namespace start with the same URI.

PREFIXES
- local shortcut to declare an namespace in a file, document, etc.
- localy use prefix instead of repeating namespace
- representation languages provide prefix declaration means

QUALIFIED NAMES
- prefix + ":" + local name
- e.g.   dc:title   instead of   <http://purl.org/dc/elements/1.1/title>

Namespaces, Prefixes, Qualified Names (in general)

74

Slide 75

http://ns.inria.fr/fabien.gandon#me

voc:author

http://inria.fr/rr/doc.html

http://inria.fr/rr/doc.html

voc:theme

Music

Slide 76
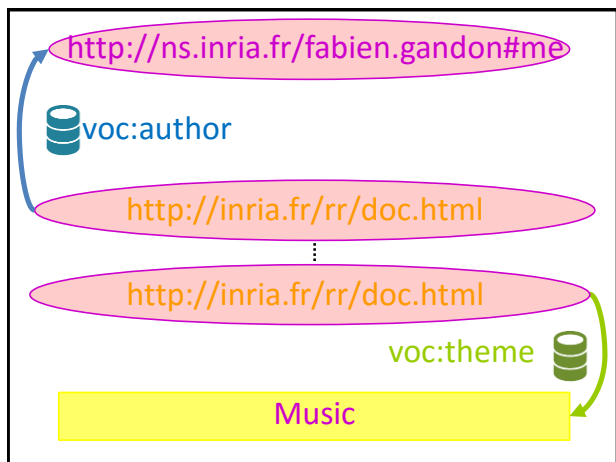
< RDF/> has an XML syntax

Slide 77

<RDF/> : graphs serialized in XML trees

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-
  syntax-ns#" xmlns:inria="http://inria.fr/schema#" >

<rdf:Description
 rdf:about="http://inria.fr/rr/doc.html">
  <inria:author rdf:resource=
   "http://ns.inria.fr/fabien.gandon#me"/>
  <inria:theme>Music</inria:theme>
</rdf:Description>

</rdf:RDF>
```

http://inria.fr/rr/doc.html
http://inria.fr/schema#author
http://ns.inria.fr/fabien.gandon#me
http://inria.fr/schema#theme
"Music"

77

Slide 78

RDF has other syntaxes
(Turtle, TriG, N-Triples, N-Quads, JSON, RDFa)

RDFa — Turtle — Supports Multiple Graphs — JSON-LD — extended by — TriG — RDF/XML — extended by — N-Triples — similar to — N-Quads

Slide 1 (N-Triples)
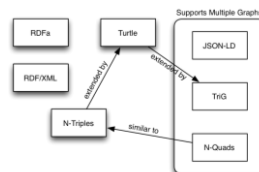
```
<http://inria.fr/rr/doc.html>
 <http://inria.fr/schema#author>
   <http://ns.inria.fr/fabien.gandon#me> .


<http://inria.fr/rr/doc.html>
 <http://inria.fr/schema#theme> "Music" .
```

just a list of triples: simple to load / parse
- URI between angle brackets   <...>
- literal values between double quotes   "..."
- triplets separated by a point   .
(but verbose)

N-Triples

Slide 2

RDF very concise syntax (Turtle/N3)

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
@prefix inria: <http://inria.fr/schema#> .

<http://inria.fr/rr/doc.html>
 inria:author <http://ns.inria.fr/fabien.gandon#me> ;
 inria:theme "Music" .
```

80

Slide 3

What is the historical syntax of RDF ?

?

Slide 4

validate & transform

Slide 5 (Practice)

Practice
1. Get the RDF data from:
   http://ns.inria.fr/fabien.gandon#me
2. What is the syntax used?
3. Validate it and see the graph:
   http://www.w3.org/RDF/Validator/
4. Translate into Turtle/N3:
   http://rdf-translator.appspot.com/
5. Visualize it also with:
   https://graves.cl/visualRDF/
6. Adapt to your data and do it again

Slide 6

RDF (named) graphs
      group triples in graphs named by IRIs

http://inria.fr/people

http://ns.inria.fr/fabien.gandon#me

http://inria.fr/schema#author

http://inria.fr/rr/doc.html

http://inria.fr/topics

http://inria.fr/rr/doc.html

http://inria.fr/schema#theme

Music

RDF 1.1 extends Turtle and N-Triples for named graphs

W3C ©

86
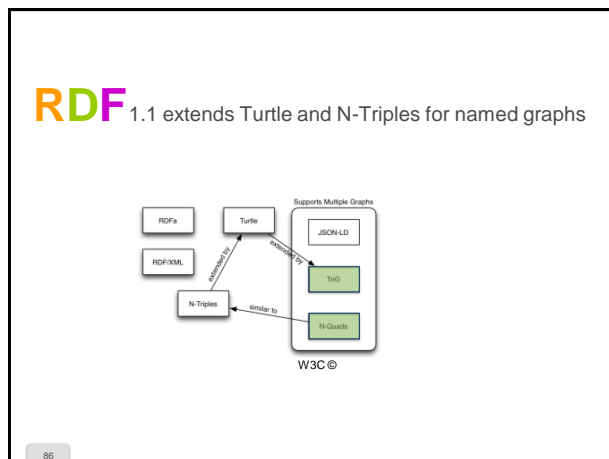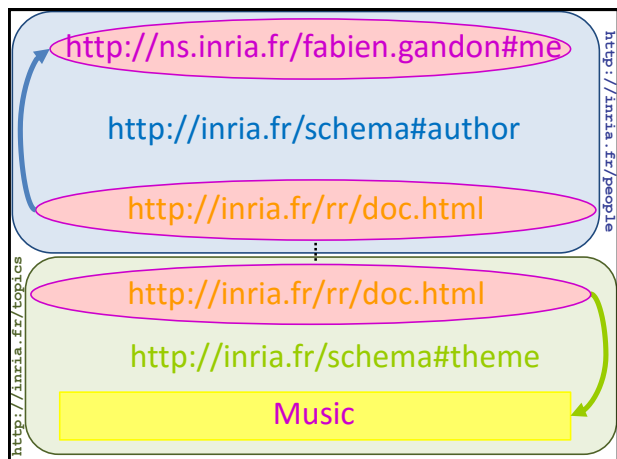
```
@prefix rdf:
  <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix inria: <http://inria.fr/schema#> .

GRAPH <http://inria.fr/people>
{ <http://inria.fr/rr/doc.html>
  inria:author
   <http://ns.inria.fr/fabien.gandon#me> .
}

GRAPH <http://inria.fr/topics>
{ <http://inria.fr/rr/doc.html>
 inria:theme "Music" .
}
```
TriG

```
<http://inria.fr/rr/doc.html>
 <http://inria.fr/schema#author>
  <http://ns.inria.fr/fabien.gandon#me>
   <http://inria.fr/people> .

<http://inria.fr/rr/doc.html>
 <http://inria.fr/schema#theme> "Music"
  <http://inria.fr/topics> .
```
N-Quads

Linked Data in JSON JSON-LD {⋅}

- JSON (JavaScript Object Notation)
  - hierarchy of name-value pairs
- JSON-LD (JSON for Linked Data)
  - designed around the notion of "context" to provide additional mappings from JSON to an RDF model.
  - a context can be embedded directly in a JSON-LD document or put into a separate file and referenced.
  - specific reserved names prefixed by @
    e.g. @context , @type

```
{



"firstName": "Fabien",
"headline": "Research Director at Inria",
"id": "Fg-fjekzI",
"lastName": "Gandon",
"siteStandardProfileRequest": {
  "url": "https://www.linkedin.com/profile/view?id=AAAAAA"
 }
}
```
e.g. LinkedIn JSON

15

```
{
  "@context": {
    "@vocab": "http://schema.org/",
    "@base" : "http://data.org/",
    "id" : "@id",
    "firstName": "givenName",
    "lastName": "familyName",
    "headline": { "@id": "jobTitle", "@language": "en" },
    "siteStandardProfileRequest" : null },

  "firstName": "Fabien",
  "headline": "Research Director at Inria",
  "id": "Fg-fjekzI",
  "lastName": "Gandon",
  "siteStandardProfileRequest": {
    "url": "https://www.linkedin.com/profile/view?id=AAAAAA"
  }
}
```

Mapping with @vocab e.g. LinkedIn JSON

```
{
  "@context": {
    "@vocab": "http://schema.org/",
    "@base" : "http://data.org/",
    "id" : "@id",
    "firstName": "givenName",
    "lastName": "familyName",
    "headline": { "@id": "jobTitle", "@language": "en" },
    "siteStandardProfileRequest" : null },

  "firstName": "Fabien",
  "headline": "Research Director at Inria",
  "id": "Fg-fjekzI",
  "lastName": "Gandon",
  "siteStandardProfileRequest": {
    "url": "https://www.linkedin.com/profile/view?id=AAAAAA"
  }
}
        @prefix : <http://schema.org/> .

        <http://data.org/Fg-fjekzI> :familyName "Gandon" ;
                :givenName "Fabien" ;
                :jobTitle "Research Director at Inria"@en .
```

Mapping with @vocab e.g. LinkedIn JSON

Test online

- Transform your FOAF profile in JSON-LD with the translator: http://rdf-translator.appspot.com/
- Use the following online tool to generate different variations of JSON-LD of your profile (expanded, collapsed, flattened, etc.) http://json-ld.org/playground/

Visit Leukocyte surface antigen CD53

1. See HTML data from: http://www.uniprot.org/uniprot/Q61451
2. Get RDF data from: http://www.uniprot.org/uniprot/Q61451.rdf
3. What is the syntax?
4. Translate into Turtle/N3: http://rdf-translator.appspot.com/
5. Any remark?
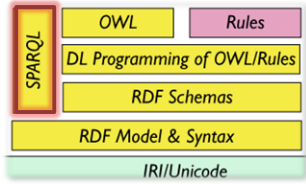
ACCESSING DATA ON THE WEB?

Query RDF data → SPARQL

96
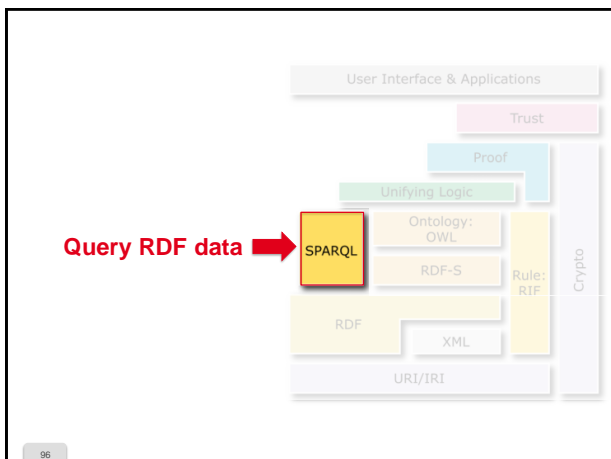
SPARQL
Protocol and
RDF
Query
Language

SPARQL in 3 parts
part 1: query language
part 2: result format
part 3: access protocol

SPARQL query

```
SELECT ...
FROM ...
WHERE { ... }
```

---

query syntax based on Turtle
e.g. persons at least 18-year old
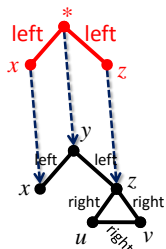
```
PREFIX ex: <http://inria.fr/schema#>
SELECT ?person ?name
WHERE {
 ?person rdf:type ex:Person .
 ?person ex:name ?name .
 ?person ex:age ?age .
 FILTER (?age > 17)
}
```

---

left(x,y)   right(z,v)
left(y,z)   right(z,u)
            right(u,v)

$\models$ $\Longleftrightarrow$

left(x,?p) left(?p,z)

---

question:
- Query:
```
SELECT ?name WHERE {
 ?x name ?name .
 ?x email ?email .
 }
```
- Base:
```
_:a name "Fabien"      x2
_:b name "Thomas"
_:c name "Lincoln"
_:d name "Aline"
_:b email <mailto:thom@chaka.sn>
_:a email <mailto:Fabien.Gandon@inria.fr>
_:d email <mailto:avalandre@pachinko.jp>
_:a email <mailto:bafien@fabien.info>
```
- Results ?

same shortcuts as Turtle

triples with a common subject:

```
SELECT ?name ?fname        SELECT ?name ?fname
WHERE {                    WHERE {
  ?x a Person;               ?x rdf:type Person .
     name ?name ;            ?x name ?name .
     firstname ?fname ;      ?x firstname ?fname .
     author ?y . }           ?x author ?y .
                           }
```

list of values

```
?x firstname "Fabien", "Lucien" .
```

blank node

```
[firstname "Fabien"] or [] firstname "Fabien"
```

Query DBpedia
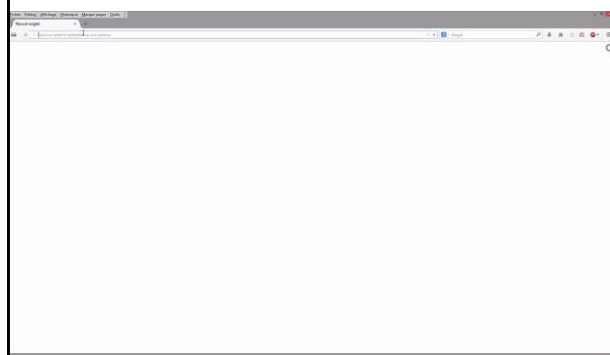
Test on DBpedia

- Connect to:
  http://dbpedia.org/snorql/   or
  http://fr.dbpedia.org/sparql   or …
  http://wiki.dbpedia.org/Internationalization/Chapters

- Query:
```
SELECT * WHERE {
  ?x rdfs:label "Paris"@fr .
  ?x ?p ?v .
}
LIMIT 10
```

American presidents on Wikidata
https://query.wikidata.org/

WIKIDATA

```
SELECT ?president  ?name WHERE
{     wd:Q30 p:P6/ps:P6 ?president .
      ?president p:P734 ?pLabel .
      ?pLabel ?x ?v .
      ?v rdfs:label ?name .
      FILTER (lang(?name)="en")
}
```

SPARQL Update
Update language for RDF graphs
CRUD: Create Read Update Delete

108

18

SPARQL Update

Search, remove and add triples

```
PREFIX ex: <http://example.org/>
DELETE {
    ?x a ex:Musician
}
INSERT {
    ?x a ex:Artist
}
WHERE {
    ?x a ex:Musician
}
```

109

HTTP   **SPARQL**
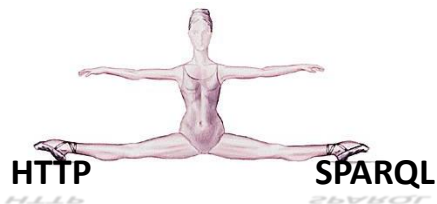
Linked Data Platform

HTTP access to LD resources & containers
get, post, put, delete resources from LD servers.

```
GET /people/fab HTTP/1.1
Host: data.inria.fr
```
?

```
PUT http://data.inria.fr/people/fab  HTTP/1.1
Host: data.inria.fr
Content-Type: text/turtle

<fab> a foaf:Person ;
      rdfs:label "Fabien" ;
      foaf:mbox  <fabien.gandon@inria.fr> .
```
!

Corese KGram

File Edit Engine Debug Query Template Explain ?

System

Loaded files:

/Users/ereal_5/Desktop/demoquery/history.ttl

logs:

reset...

done.
Loading ttl file from path : /Users/ereal_5/Desktop/demoquery/history.ttl

---
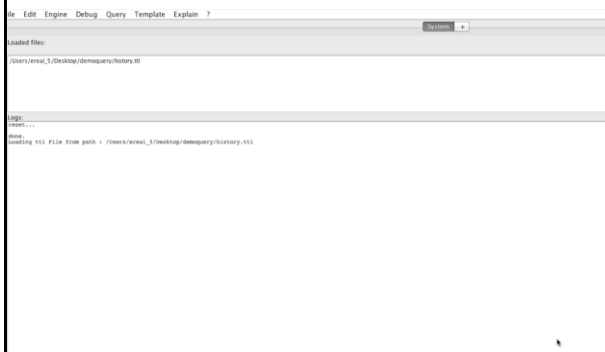
SEMANTIC WEB

---

do not read
the following sign

19

you loose

we identify and interpret information, **machines don't.**

know the meaning of data
to find out what can be done with it.

+ =

117
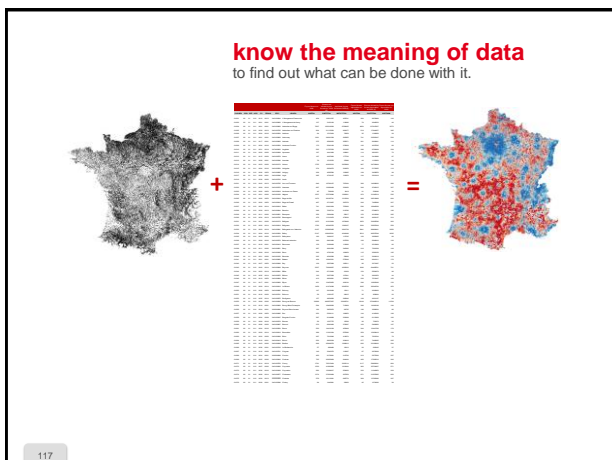
what is the last **document** you read?

documents

{ }
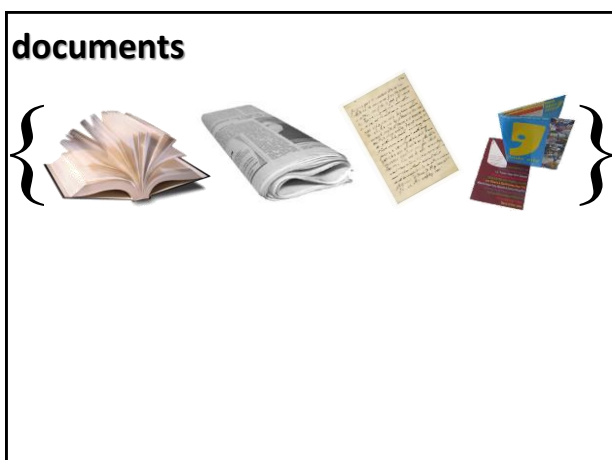
your answer relies on a **shared** ontology

you infer from it

we all understood

Slide 1:

kind
of

Document

Book

Novel   Short story

Slide 2:

kind
of

"document"   #12

"book"
"livre"   #21

#47   #48

"novel"   "short story"
"roman"   "nouvelle"

Slide 3:

#12

#21 ⟹ #12

#21

#47 ⟹ #21   #48 ⟹ #21

#47   #48

formalized ontological knowledge

Slide 4:

languages
to formalize
ontologies

Slide 5:

User Interface & Applications

Trust

Proof

Unifying Logic

Ontology: OWL

RDF-S

SPARQL

RIF

RDF

XML

URI/IRI

PUBLISH SEMANTICS OF SCHEMAS

W3C®

Slide 6:

stack of standards

RDFS   communication   HTTP
web de données
référence   adresse

URI

User Interface & Applications

Trust

Proof

Unifying Logic

Ontology: OWL

SPARQL   RDF-S   Rule: RIF   Crypto

RDF

XML

URI/IRI

W3C®

126

21

RDFS means RDF Schema

RDFS provides primitives to 𝒲rite lightweight ontologies

RDFS to define classes of resources and organize their hierarchy

RDFS to define relations between resources, their signature and organize their hierarchy

---
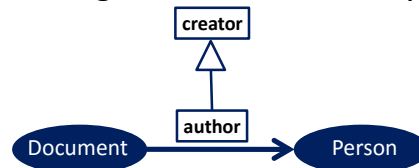
instances of rdfs:Class

the class of classes is in RDFS namespace.

```
<rdf:RDF xml:base="http://inria.fr/2005/humans.rdfs"
 xmlns:rdf ="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
 <rdfs:Class rdf:ID="Man">
  <rdfs:subClassOf rdf:resource="#Person"/>
  <rdfs:subClassOf rdf:resource="#Male"/>
 </rdfs:Class>
</rdf:RDF>
```

```
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@base <http://inria.fr/2005/humans.rdfs> .
<Man> a rdfs:Class ;
     rdfs:subClassOf <Person>, <Male> .
```

131

---

instances of rdf:Property

the class of properties was placed in the RDF namespace because triples are a construction of RDF.

```
<rdf:RDF xml:base="http://inria.fr/2005/humans.rdfs"
 xmlns:rdf ="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
 <rdf:Property rdf:ID="hasMother">
  <rdfs:subPropertyOf rdf:resource="#hasParent"/>
 </rdf:Property>
</rdf:RDF>
```

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@base <http://inria.fr/2005/humans.rdfs> .
<hasMother> a rdf:Property ;
  rdfs:subPropertyOf <hasParent> .
```

132

domain and range

Class of departure or domain: rdfs:domain
Class of arrival, co-domain or range: rdfs:range

```
<rdf:RDF xml:base="http://inria.fr/2005/humans.rdfs"
 xmlns:rdf ="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
 <rdf:Property rdf:ID="hasMother">
  <rdfs:subPropertyOf rdf:resource="#hasParent"/>
  <rdfs:domain rdf:resource="#Human"/>
  <rdfs:range  rdf:resource="#Woman"/>
 </rdf:Property>
</rdf:RDF>
```

133

domain and range

Class of departure or domain: rdfs:domain
Class of arrival, co-domain or range: rdfs:range

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@base <http://inria.fr/2005/humans.rdfs> .
<hasMother> a rdf:Property ;
 rdfs:subPropertyOf <hasParent> ;
 rdfs:domain <Human> ;
 rdfs:range  <Woman> .
```

134

multiple domains and ranges

- conjunction of domains and ranges.
- the effective domain is the intersection of declared and inherited domains.
- the effective range is the intersection of declared and inherited ranges.
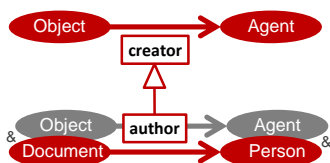
135

semantics

1. Type inference (domain)
   IF p rdfs:domain d AND x p y
   THEN x rdf:type d

2. Type inference (range)
   IF p rdfs:range r AND x p y
   THEN y rdf:type r

136

question

If I use the property author on a car what happens?
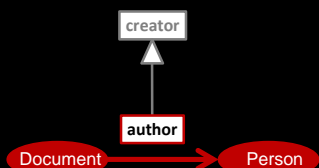
OWL provides **additional primitives** for **heavyweight ontologies**

Web Ontology Language (OWL)

- a W3C recommendation
- additional primitives for more complex ontologies.
- richer definitions of classes and properties.
- perform more inferences, draw more conclusions.
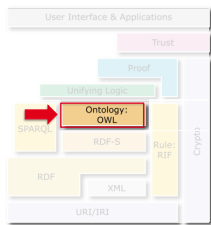
namespace and prefix for OWL

http://www.w3.org/2002/07/owl#

- namespace of the OWL primitives
- same principle as RDFS
- `owl:` prefix in the rest of the slides

OWL in one…

- algebraic properties
- disjoint properties
- qualified cardinality
- individual prop. neg
- chained prop.

union
disjunction
intersection
complement
restriction
cardinality
equivalence
enumeration
value restrict.
disjoint union
keys
…

OWL profiles

- Each profile is a sub-set of the OWL primitives.
- Choosing a profile is choosing a level of expressivity.
- The higher the expressivity the more complex the inferences.
- The more complex the expressivity, the longer it takes to compute the results.

OWL 2 profiles

EL: large numbers of properties and/or classes and polynomial time.

QL: large volumes of instance data, and conjunctive query answering using conventional relational database in LOGSPACE

RL: scalable reasoning without sacrificing too much expressive power using rule-based reasoning in polynomial time

DL: the most expressive with complete reasoning

FAMOUS SCHEMAS

Slide: reuse
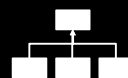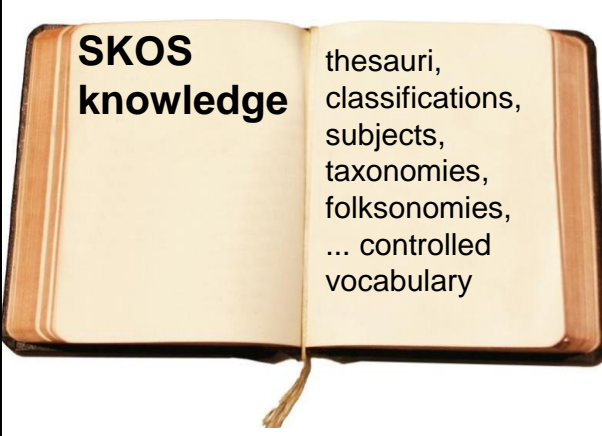
- **Dublin core**
- **Creative Commons**
- **FOAF**

...

Slide: SKOS

SKOS knowledge

thesauri, classifications, subjects, taxonomies, folksonomies, ... controlled vocabulary

146

Slide: SKOS schema

SKOS schema : 4 classes and 28 properties in OWL

skos:Concept	
URI:	http://www.w3.org/2004/02/skos/core#Concept
Definition:	Section 3. The skos:Concept Class
Label:	Concept
Disjoint classes:	skos:Collection skos:ConceptScheme

skos:ConceptScheme	
URI:	http://www.w3.org/2004/02/skos/core#ConceptScheme
Definition:	Section 4. Concept Schemes
Label:	Concept Scheme
Disjoint classes:	skos:Collection skos:Concept

http://www.w3.org/2004/02/skos/core

147

Slide: labels

natural language expressions to refer to concepts

```
inria:CorporateSemanticWeb
  skos:prefLabel "corporate semantic web"@en;
  skos:prefLabel "web sémantique d'entreprise"@fr;
  skos:altLabel "corporate SW"@en;
  skos:altLabel "CSW"@en;
  skos:hiddenLabel "web semantique d'entreprise"@fr.
```

148

Slide: relations

between concepts

```
inria:CorporateSemanticWeb
  skos:broader w3c:SemanticWeb;
  skos:narrower inria:CorporateSemanticWiki;
  skos:related inria:KnowledgeManagement.
```

Slide: notes

```
inria:CorporateSemanticWeb
  skos:scopeNote "only within KM community";
  skos:definition "a semantic web on an intranet";
  skos:example "Nokia's internal use of RDF gateway";
  skos:historyNote "semantic intranet until 2006";
  skos:editorialNote "keep wikipedia def. uptodate";
  skos:changeNote "acronym added by fabien".
```

CC (*Creative Commons*)

a very popular schema to describe rights associated to a resource

6+12 classes and 11 properties to :
- describe the rights associated with a resources
- describe a license and associated to a resource
- state the rights, conditions and prohibitions to use a resource

```
@prefic cc: <http://creativecommons.org/ns#>
```

CC REL : Creative Commons Rights Expression Language

http://creativecommons.org/ns

151

FOAF : *Friend of a Friend*

a very popular schema to describe persons and social networks

13 classes and 62 properties to describe:
- user profiles (your RDF homepage)
- social networks (persons you know)
- social activities (accounts, actions)

```
@prefix foaf: <http://xmlns.com/foaf/0.1/>
```

FOAF = Core + Social Web + Linked Data Utilities

http://xmlns.com/foaf/spec/

152

VoID: describing RDF datasets/linksets

DCAT: describe any dataset (not just RDF)

Data Cube: publish multi-dimensional data (statistics)

Provenance: PROV-DM & PROV-O
describe entities and activities
involved in providing a resource
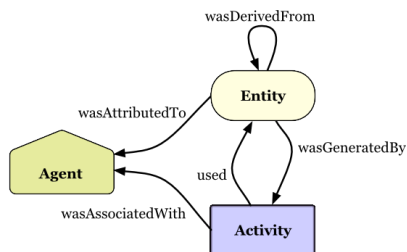
e.g. a chart produced from two sources of data

```
ex:illustrate prov:used ex:composition .
ex:chart1 prov:wasGeneratedBy ex:illustrate
```
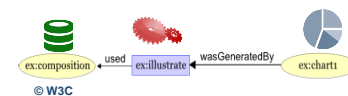

158

e.g. a chart produced from two sources of data

```
ex:compose prov:used ex:dataSet1 ;
          prov:used ex:regionList .
ex:composition prov:wasGeneratedBy ex:compose .
ex:illustrate prov:used ex:composition .
ex:chart1 prov:wasGeneratedBy ex:illustrate
```


159

LOV LOV LOV…
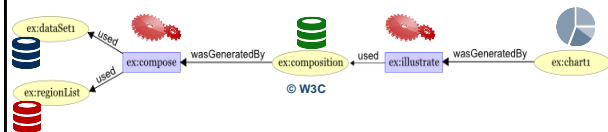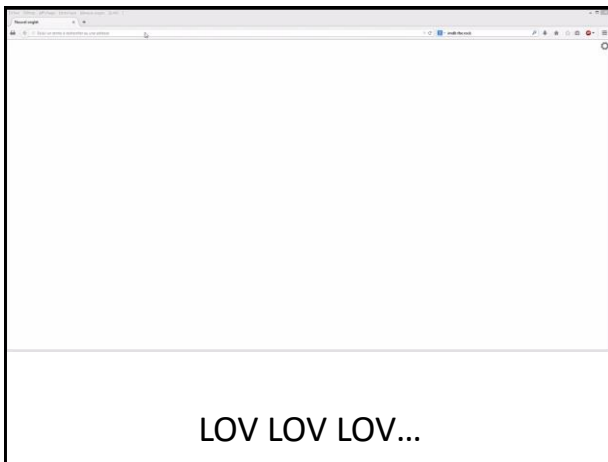
semantic waste separation
the web is a garbage can,
the semantic web will be a semantic garbage can.

EXTENDING TO OTHER SOURCES

Slide 1

toward all forms of data on the web

W3C DATA ACTIVITY *Building the Web of Data*

Slide 2

schema

mapping

R2RML
a standard transformation of a relationnal database in RDF

Slide 3

two types of transformations

- Default transformation
  [A Direct Mapping of Relational Data to RDF]

- Customized transformation
  [R2RML: RDB to RDF Mapping Language]

  schema

  mapping

165

Slide 4

many data
buried and dormant in web pages

Slide 5

RDFa means RDF in HTML attributes

```
<body vocab="http://purl.org/dc/terms/>

<div resource="http://lib.com/books/0684853949">

<h2 property="title">The Man Who Mistook His
                Wife For a Hat</h2>
<h3 property="creator">Oliver Sacks</h3>
...
```

167

Slide 6

Do it...

- Look at the Web Page
  http://schema.openspring.net/person/dries-buytaert
- Call the translator on this Web page to get Turtle:
  http://rdf-translator.appspot.com/
- What are the types of the main resource extracted?
- Do the same with:
  http://schema.openspring.net/event/2014-winter-olympics
  http://schema.openspring.net/recipe/apple-pie
  http://schema.openspring.net/events/drupalcamps

Do it…

Use the online tool to play with RDFa adding for instance a "creator" property
https://rdfa.info/play/

```
<div about="" typeof="cc:Work"
     xmlns:cc="http://creativecommons.org/ns#"
     xmlns:dc="http://purl.org/dc/elements/1.1/" align="center">
<img alt="Creative Commons License"
 src="http://i.creativecommons.org/l/by/3.0/us/88x31.png" /> <br />
<span property="dc:title">The Lessig Blog</span>, a
<span rel="dc:type" href="http://purl.org/dc/dcmitype/Text">
 collection of texts </span> by
<a property="cc:attributionName" rel="cc:attributionURL"
   href="http://lessig.org/"> Lawrence Lessig </a>,<br />
 is licensed under a <a rel="license"
  href="http://creativecommons.org/licenses/by/3.0/"> Creative
       Commons Attribution License </a>.<br />
 There are <a rel="cc:morePermissions"
 href="http://lessig.org/blog/other-license"> alternative
 licensing options </a>. </div>
```

CC REL in RDFa

---

schema.org = bing + Google + YAHOO! + Яндекс Yandex

schemas to improve index, search and display e.g:

- **Creative works**, Book, Movie, MusicRecording, Recipe, TVSeries …
- **Embedded non-text objects**, AudioObject, ImageObject, VideoObject
- **Event**
- **Organization**
- **Person**
- **Place**, LocalBusiness, Restaurant ...
- **Product**, Offer, AggregateOffer
- **Review**, AggregateRating

OGP code

```
<html xmlns="http://www.w3.org/1999/xhtml" dir="ltr" lang="en-US"
   xmlns:fb="https://www.facebook.com/2008/fbml">
<head prefix="og: http://ogp.me/ns# fb: http://ogp.me/ns# YOUR_NAMESPACE:
        http://ogp.me/ns/apps/YOUR_NAMESPACE#">
   <meta property="fb:app_id" content="YOUR_APP_ID" />
   <meta property="og:type" content="YOUR_NAMESPACE:recipe" />
   <meta property="og:title" content="Stuffed Cookies" />
   <meta property="og:image" content="http://example.com/cookie.jpg" />
   <meta property="og:description" content="The Turducken of Cookies" />
   <meta property="og:url" content="http://example.com/cookie.html">
   <script type="text/javascript">
   function postCook()
   {  FB.api('/me/YOUR_NAMESPACE:cook' + '?recipe=http://example.com/cookie.html','post', (…)    );  }
   </script>
</head>
<body>
(…)
   <form>
     <input type="button" value="Cook" onclick="postCook()" />
   </form>
</body>
</html>
```
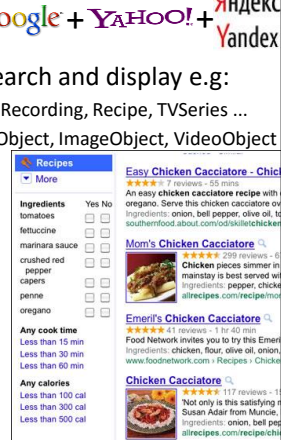
172

These data are accessible to everyone !

If you apply a parser to these pages, you will get their data...

173

Test online

- IMDB uses RDFa – OGP for the I like button
- Choose a movie on IMDB http://www.imdb.com
- Copy the URL of the page of the movie
- Go to the RDFa 1.0 RDFa Distiller and Parser:
  https://www.w3.org/2007/08/pyRdfa/
- Open the URI option, past the URL of the movie page and configure and perform the extraction to get Turtle
- Try also the transformation on the translator:
  http://rdf-translator.appspot.com/

Call the translator on…

- A product on eBay
- A movie in Dailymotion
- An article on LeMonde.fr
- A recipe on Marmiton.org
- A hotel on Booking.com
…

pages || data

page && data
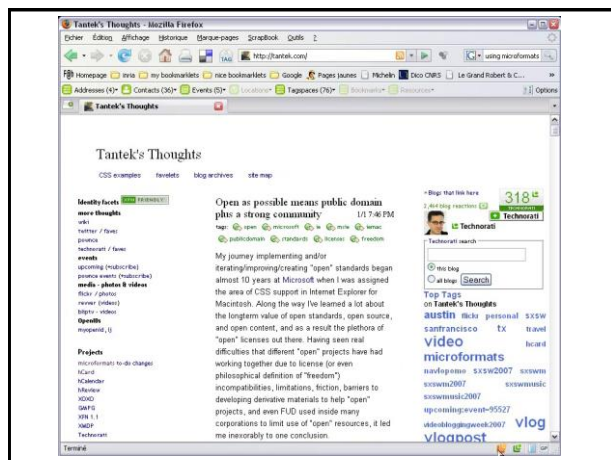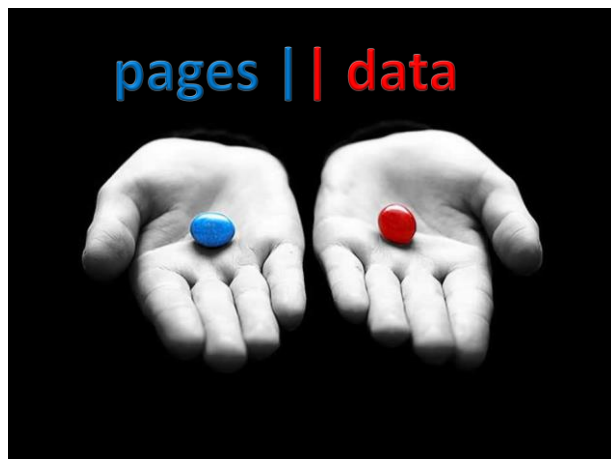
CSV-LD & Linked CSV

- contexts to interpret and generate CSV
- conventions for CSV to be linked in RDF
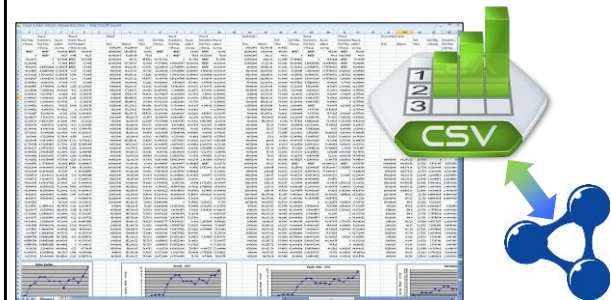
country	country group	name (en)	name (fr)	name (de)	latitude	longitude
at	eu	Austria	Autriche	Österreich	47.6965545	13.34598005
be	eu	Belgium	Belgique	Belgien	50.501045	4.47667405
bg	eu	Bulgaria	Bulgarie	Bulgarien	42.72567375	25.4823218

```
"country","country group","name (en)","name (fr)","name (de)","latitude","longitude"
"at","eu","Austria","Autriche","Österreich","47.6965545","13.34598005"
"be","eu","Belgium","Belgique","Belgien","50.501045","4.47667405"
"bg","eu","Bulgaria","Bulgarie","Bulgarien","42.72567375","25.4823218"
```

Spreadsheet to CSV

Slide 1 (top-left):

```
"country","country group","name (en)","name (fr)","name (de)","latitude","longitude"
"at","eu","Austria","Autriche","Österreich","47.6965545","13.34598005"
"be","eu","Belgium","Belgique","Belgien","50.501045","4.47667405"
"bg","eu","Bulgaria","Bulgarie","Bulgarien","42.72567375","25.4823218"

{
  "@context": "http://www.w3.org/ns/csvw",       "aboutUrl":
  "url": "countries.csv",                    "http://example.org/country/{code}#geo",
  "tableSchema": {                                "propertyUrl": "schema:latitude"
    "aboutUrl":                                 },{
"http://example.org/country/{code}",            "titles": "longitude",
    "columns": [{                                 "datatype": "number",
      "titles": "country",                        "aboutUrl":
      "name": "code",                   "http://example.org/country/{code}#geo",
      "suppressOutput": true                      "propertyUrl": "schema:longitude"
    },{                                         },{
      "titles": "country group",                  "virtual": true,
      "suppressOutput": true                      "propertyUrl": "rdf:type",
    },{                                           "valueUrl": "schema:Country"
      "titles": "name (en)",                    },{
      "lang": "en",                               "virtual": true,
      "propertyUrl": "schema:name"                "propertyUrl": "schema:geo",
    },{                                           "valueUrl":
      "titles": "name (fr)",            "http://example.org/country/{code}#geo"
      "lang": "fr",                             },{
      "propertyUrl": "schema:name"                "virtual": true,
    },{                                           "aboutUrl":
      "titles": "name (de)",           "http://example.org/country/{code}#geo",
      "lang": "de",                               "propertyUrl": "rdf:type",
      "propertyUrl": "schema:name"                "valueUrl": "schema:GeoCoordinates"
    },{                                         }]
      "titles": "latitude",                   }
      "datatype": "number",
```

Based on JSON contexts

Slide 2 (top-right):

```
"country","country group","name (en)","name (fr)","name (de)","latitude","longitude"
"at","eu","Austria","Autriche","Österreich","47.6965545","13.34598005"
"be","eu","Belgium","Belgique","Belgien","50.501045","4.47667405"
"bg","eu","Bulgaria","Bulgarie","Bulgarien","42.72567375","25.4823218"

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix schema: <http://schema.org/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

<http://example.org/country/at> a schema:Country;
    schema:geo <http://example.org/country/at#geo>;
    schema:name "Austria"@en, "Autriche"@fr, "Österreich"@de .

<http://example.org/country/be> a schema:Country;
    schema:geo <http://example.org/country/be#geo>;
    schema:name "Belgium"@en, "Belgique"@fr, "Belgien"@de .

<http://example.org/country/bg> a schema:Country;
    schema:geo <http://example.org/country/bg#geo>;
    schema:name "Bulgaria"@en, "Bulgarie"@fr, "Bulgarien"@de .

<http://example.org/country/at#geo> a schema:GeoCoordinates;
    schema:latitude 4.76965545e1;
    schema:longitude 1.334598005e1 .

<http://example.org/country/be#geo> a schema:GeoCoordinates;
    schema:latitude 5.0501045e1;
    schema:longitude 4.47667405e0 .

<http://example.org/country/bg#geo> a schema:GeoCoordinates;
    schema:latitude 4.272567375e1;
    schema:longitude 2.54823218e1 .
```
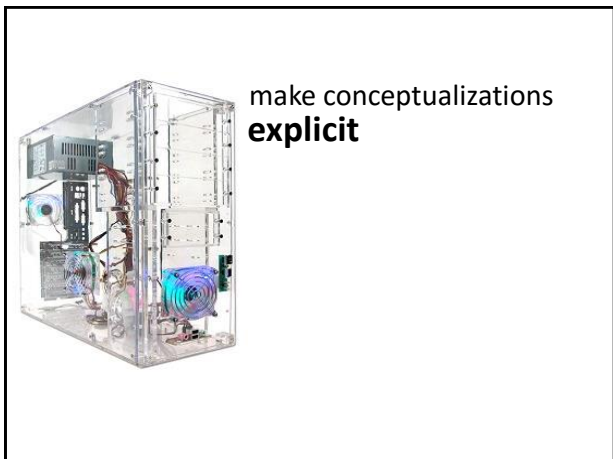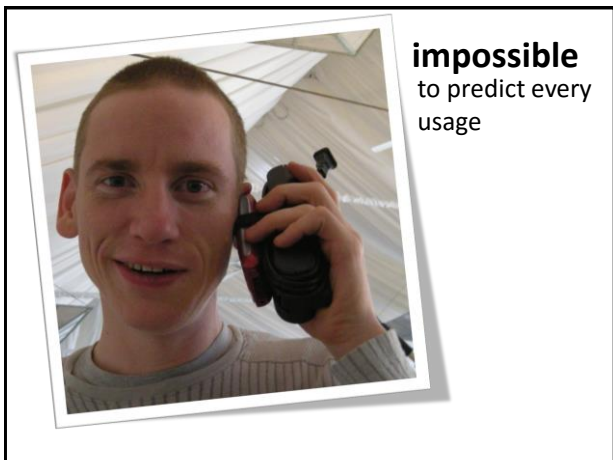
RDF Result

Slide 3 (middle-left):

doggy-bag

Slide 4 (middle-right):

impossible
to predict every usage

Slide 5 (bottom-left):

avoid building
black boxes

Slide 6 (bottom-right):

make conceptualizations
explicit

open your data
to those who could use them
W3C ©

#WatchDogs #WeAreData @ubisoft

66 FOAF primitives   3 475 908 348 references (2)
x 52 millions
"a small tree ruling a big graph"(1)
(1) Franck Van Harmelen, ISWC 2011
(2) Libby Miller, 2009
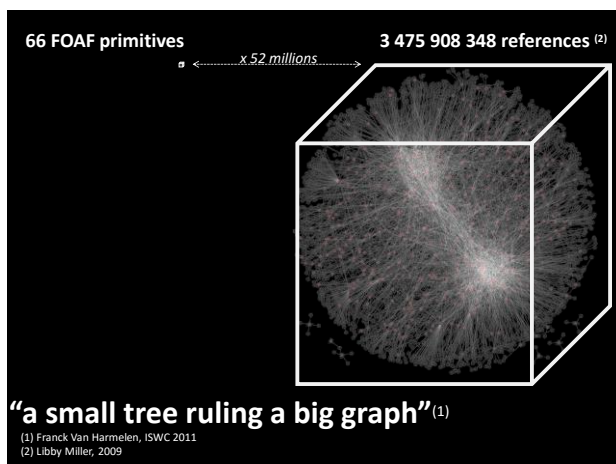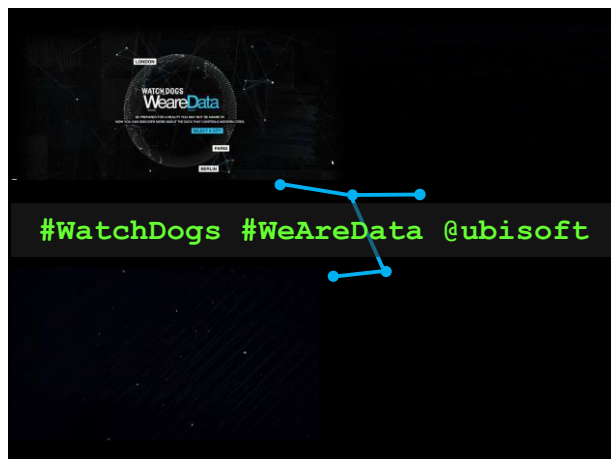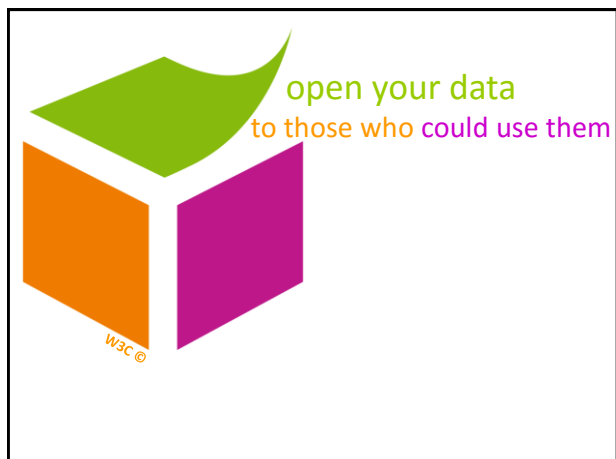
"semantic **web**"
and not
"**semantic** web"
[C. Welty, ISWC 2007]

"a lightweight ontology
allows us to do
lightweight reasoning"
[J. Hendler, ISWC 2007]

WWW 2017
Perth, Australia
3-7 April, 2017
26th & 27th International
World Wide Web Conferences
Lyon, France
23-27 April, 2018
THE WEB CONFERENCE

linked **data** streams → web of sensors, things, …
linked enterprise **data**
VISIBILITY
enterprise **data**
big **data** streams
VELOCITY
personal **data**
data streams
big linked **data**
linked healthcare **data**
data mining
closed **data**
big **data**
VOLUME
open **data**
data bases
data structures
linked open **data**
data
data type
linked **data**
VICINITY
data models
VARIETY
linked **data** schemas
web of **data**
data schemas
semantic web of **data**
VVeb data

identify	URI
describe & link	RDF
query	HTTP, SPARQL, LDP
reasoning	RDFS & OWL
trace	PROV-O

GOALS AND MEANS

identify	http://fabien.fr#me
describe & link	#me type man
query	select * {?r type ?t}
reasoning	man subClassOf male
trace	wasAttributedTo #me

GOALS AND MEANS

web 1, 2

web 1, 2, 3

he who controls metadata, controls the web
and through the *world-wide* web many things in our world.

fabien, gandon, @fabien_gandon, http://fabien.info