Semantic Web and knowledge engineering

Identify things and represent simple facts



Dr. Azanzi Jiomekong

University of Yaounde I, Department of Computer Science

10 avril 2023



Copyright (c) 2020 Azanzi Jiomekong.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. More information about the license is included in the section entitled "The GNU Free Documentation License".

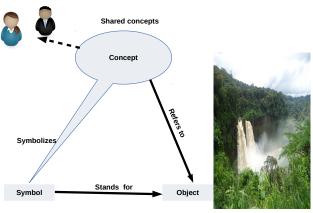
Semiotic triangle



个 個 असल 1



Semiotic triangle



Chutes jumelles Ekom Kam



Questions

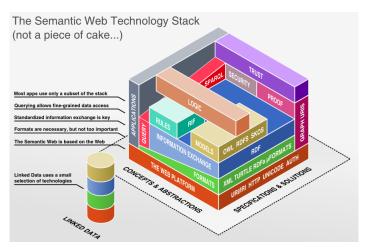
- How to identify things?
 Uniform Resource Identifier URI
- How to represent things?
 RDF Graph
- How to encode and represent facts?
 Graph serialization

The responses to these questions are given in this course



Semantic Web Stack

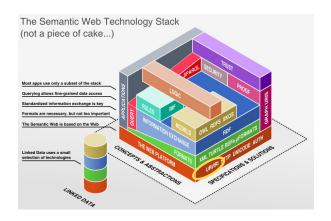
Represents Semantic Web Basic Architecture





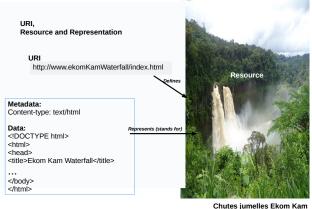
How to identify things?

Resp = Uniform Resource Identifier





How to identify things? URI





Representation (HTML) and presentation (Web browser)
Put the previous waterfall in a Web page to have his representation



HTTP Request / Response Use the previous image to show how HTTP works (with the representation and the presentation of information)



- What is the difference between the picture of a thing and the thing itself?
- ullet When we talk about a thing \longrightarrow we talk about things having the same property of this thing
- We share a common concepts

 we are able to communicate information and understand information
- ullet We must identify things that we are talking about \longrightarrow we need the representation of these things in the Semantic Web





- Simple and extensible schema for worldwide unique identification of abstract physical resources (RFC 3986)
- Uniform : different types of resources identifiers all constructed according to a uniform schema
- Resource:
 - Can be every object with a clear identity (according to the content of the application)
 - Whatever may be identified via URI
 - e.g., web pages, books, locations, persons, relations among objects, abstract concepts, etc.
- Identifier: to distinguish one resource from another

Semantic Web Link Open Data Orecioges

Uniform Resource Identifier

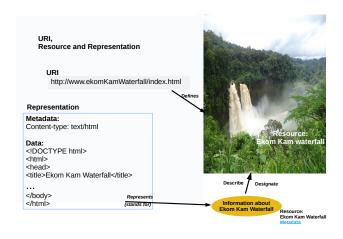
- URI = 2 things : Address and Identity
- Uniform : we can different types of resources
- Identifier: we can distinguish one resource from another in a unique way
- Representation of information :
 - We have a representation of a thing in the web page
 - We have the metadata of the representation of this thing
 - The browser transform this representation in a format that a human can understand
- When a user send a request, to a server, the representation of what he wants is delivered



Designator and Designatum

- The distinguishing characteristics of [information] resources is that all of their essential characteristics and can be conveyed in a messages - W3C: Architecture of the WWW, volume One
- A resource can be describeb (designated) via Metadata
- Even if the resource itself cannot be delivered from the web server, probably its representation might be available that describes the resource sufficiently







- Combines :
 - Address (Locator)
 - Uniform Resource Locator (URL, RFC 1738)
 - Denotes where a resource can be found in the Web by stating its primary access mechanism
 - Might change during life cycle
 - o Identity (Name):
 - Uniform Resource Name (URN, RFC 2141)
 - Persistent identifier for a Web resource
 - Remains unchanged during life cycle



Generic Syntax

- US-ASCII Encoding
- Percent Encoding for reserved characters, or characters that do not exist in US-ASCII encoding pct-encoded="%" HEXDIG HEXDIG
- Reserved characters with special function reserved = gen-delims/sub-delims gen-delims=" : "/"/"/"?"/"#"/"["/"]"/"@" sub-delims="!"/"\$"/"&"/""/"("/")"/"*"/"+"/","/";"/"="
- Permitted characters
 unreserved = ALPHA/DIGIT/" "/"."/"."/" "
- Extension to Universal Character Code (Unicode/ISO 10646)
 International Resource Identifier (IRI, RFC 3987), e.g., http address in Japanese language



Generic syntax

schema"://[userinfo"@"]host[:port][path]["?"query]["#"fragment]

- schema : e.g. http, ftp, mailto, etc.
- userinfo : e.g. username :password
- host : e.g. Domain-Name, IPV4/IPV6 Address
- port : e.g. 80 for standard HTTP port
- path: e.g. path in file system of www server
- query : e.g. parameters to be passed over to applications
- fragment : e.g. determines a specific fragment of a document



In the Semantic Web

- What if a URI for a resource does not exist?
 - Define a URI by yourself :
 - ullet avoid overlaps \longrightarrow use your own website
 - enables documentation at the same place (→Content negociation)
 - Separate URI for resource (Designatum) and its documentation (Designator) via URI references (i.e. via "#" fragments) or content negociation

http://www.facsciences-uy1/azanzi.foaf.rdf#me



- Already establish in various domains :
 - o The Web: URL, PRN
 - $\circ~$ Books and publications : ISBN, ISSN
 - Digital Object Identifier : DOI



Knowledge

Facts, information or experience



How to represent simple facts?

Example: How to represent the fact: "Dr. Jiomekong has the email fidel.jiomekong@facsciences-uy1@uninet.cm"?:

Response:

- Natural Language representation
- Graph representation
- XML Serialization



A simple example with XML

How do I represent: "Dr. Jiomekong has the email fidel.jiomekong@facsciences-uy1.uninet.cm?

<emailaddr>

<owner> Dr. Jiomekong </owner>
<email>fidel.jiomekong@facsciences-uy1.uninet.cm</email>
</emailaddr>

<person name= "Dr. Jiomekong" email= fidel.jiomekong@facsciences-uy1.uninet.cm/>



Knowledge representation A simple example with Natural Language

How do I represent: "Dr. Jiomekong has the email fidel.jiomekong@facsciences-uy1.uninet.cm?

Dr. Jiomekong
Has the email
fidel.jiomekong@...



A simple example with a graph

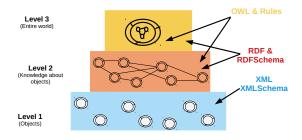
How do I represent: "Dr. Jiomekong has the email fidel.jiomekong@facsciences-uy1.uninet.cm?

Simple and intuitive knowledge representation with directed graphs





In the Semantic Web





In the Semantic Web

Three steps of Semantic knowledge representation: XML & XMLSchema, RDF & RDFSchema, OWL & Rules:

- XML & XMLSchema:
 - Available in the Web
 - Used to describe simple objects
 - o Are not connected to each other
 - o It is difficult to decide if they represent the same or difference
- RDF & RDFSchema :
 - Give the possibility to define relations among objects
 - o First knowledge, first facts that we can express
 - Used to take objects into relations
 - \circ Used to construct little world that is open and based on RDF and RDFS

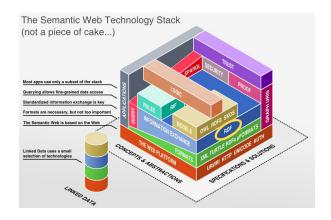


In the Semantic Web

Three steps of Semantic knowledge representation : XML & XMLSchema, RDF & RDFSchema, OWL & Rules :

- OWL & Rules :
 - Use if we want to close our world
 - Used to define entire world with things inside that world
 - Enable the Semantic expressivity with OWL and Rules
 - Used to formulate axioms that contains knowledge, restrictions, constrainst about things we are modeling







- An RDF Model is a set of RDF statements
- Statement :=(subject, property, object)
- Subject is a resource
- Property is a resource
- Object is either a resource or a literal
- Knowledge is expressed as a list of statements
- All RDF statements follow the same simple schema



Resource

- Can be everything
- Must be uniquely identified and be referencable
- ...simply via URI

All objects that can be addressed via URI



Description

- =Description of resource
- ...via representing properties and relationships among resources
- ...relationships can be represent as graphs

Properties / attributes for the description of ressources



Framework

- =Combination of Web based protocols (URI, HTTP, XML, ...)
- Based on formal model (semantics)
- Defines all allowed relationships among resources



How do I represent: "Dr. Jiomekong has the email fidel.jiomekong@facsciences-uv1.uninet.cm?

Dr. Jiomekong

Has the email

Froperty

Fidel.jiomekong

Jobject

Semantic Web Link Open Data Creiciogas C

Resource Description Framework

- RDF
- Subject Property Value (the value of that property)
- The value of the property can be another subject or a value
- Constituents of RDF language : Resource, literal and blank nodes
- Several representation of RDF language :
 - Node-Edge-Node Triple : graph representation
 - o N3 Notation
 - o Turtle (Terse RDF Triple Language) simplification of N3
 - o RDF XML-Serialization is the standard



Example of a statement

Suppose I want to make the following statement: This is a lecture and this lecture has a name which is "Semantic Web and Applications". The property of this lecture is the lecture name and the lecture name contains a certain amount of time comprises four hours per week (Property for the duration).

Represent this statement with the different formalism



Constituents of RDF graph

• Resources:

Objects that can be addressed via **URI**

• Properties :

Attributes for the description of resources

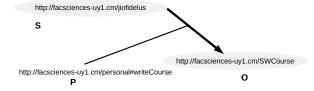
Statements(RDF-Triple) :



Constituents of RDF graph: URI

How do I represent: "Dr. Jiomekong write the course Semantic Web and applications ?

Used to reference resources uniquely

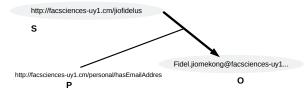




Constituents of RDF graph: Literals

- Describe data values that do not have a seperate existence
- Strings, interpretation via datatypes

How do I represent: "Dr. Jiomekong has the email fidel.jiomekong@facsciences-uy1.uninet.cm?





Constituents of RDF graph: Literals

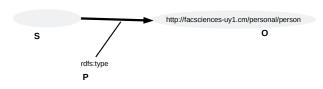
- Typed literals can be expressed via XML Schema datatypes
- Namespace for typed literals: http://www.w3c.org/2001/XMLSchema#
 e.g.: "Semantics" 'http://www.w3c.org/2001/XMLSchema#string
- Language Tags denote the (natural) language of the text : e.g. : "Language" @en, "Language" @fr



Constituents of RDF graph: Blank Nodes

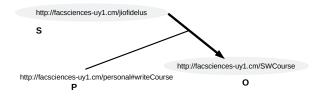
Denote existence of an individual with specific attributes, but without providing an identification or reference

How do I represent: "Existence of things of type person"





 ${\sf RDF}\ representation: Node-Edge-Node\ Triple$





RDF representation: N3 Notation

Simple listing of triples

```
{ http://facsciences-uy1.cm/jiofidelus,
http://facsciences-uy1.cm/personal#writeCourse,
http://facsciences-uy1.cm/SWCourse
}
```



RDF representation : Turtle

Simple listing of triples

- Terse RDF triple Language
- Extension of N3
- URIs in angle brackets
- Literals in quotation marks
- Triple ends with a period
- White space will be ignored

```
<Subject> <Property> <Object>. <Subject> <Property> "Object".
```



RDF representation : Turtle

```
<http://facsciences-uy1.cm/Jiomekong>
<http://facsciences-uy1.cm/personal#writeCourse>
<http://facsciences-uy1.cm/SWCourse>.

<http://facsciences-uy1.cm/Jiomekong>
<http://facsciences-uy1.cm/personal#hasEmailAddress>
  "fidel.jiomekong@facsciences-uy1.uninet.cm".
```



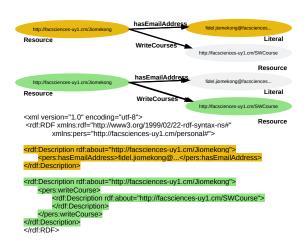
RDF representation : Turtle



RDF representation: RDF XML-Serialization



RDF representation: RDF XML-Serialization





RDF representation: RDF XML-Serialization





RDF representation : RDF XML-Serialization



<xml version="1.0" encoding="utf-8">
<rdf:RDF xmlns:rdf="http://www3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:pers="http://facsciences-uy1.cm/personal#"
 xml:base="http://facsciences-uy1.cm/Lecturer" >



RDF representation : Turtle representation



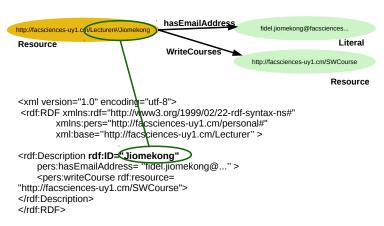
Resource

- @prefix rdf: http://www3.org/1999/02/22-rdf-syntax-ns#
- @prefix pers: http://facsciences-uy1.cm/personal#>
- @base base http://facsciences-uy1.cm/Lecturer>.

:Jiomekong pers:hasEmailAddress "fidel.jiomekong@..."; pers:writeCourse http://facsciences-uy1.cm/SWCourse.

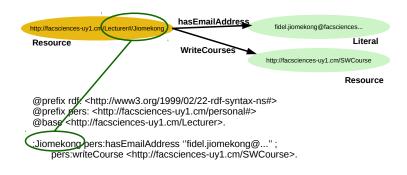


RDF representation: RDF XML-Serialization



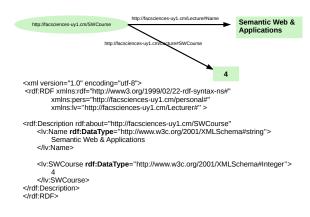


RDF representation : Turtle Serialization





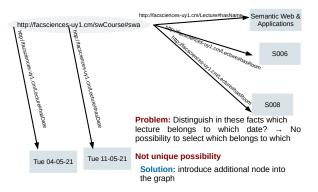
RDF representation : RDF XML-Serialization





Blank nodes

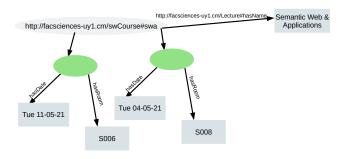
- A lecture takes place twice a week in two different rooms
- How to model this in RDF?





Blank nodes

- A lecture takes place twice a week in two different rooms
- How to model this in RDF?



The blank node represent a place and a time

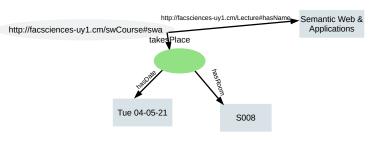


Blank nodes

- Bnodes
- Also called Empty node
- Object which do not have URI associated with
- Used to make a statement about things that exist without telling which individual directly are addressed with this information
- Existential assertion without denoting specific individual (There exist at least one)
- Used to represent multivalue relations



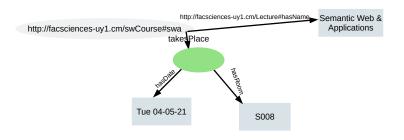
Anonymous Blank Nodes



<xml version="1.0" encoding="utf-8">
<rdf:RDF xmlns:rdf="http://www3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:fs-uy1="http://facsciences-uy1.cm/Lecturer#" >



Anonymous Blank Nodes: Turtle representation



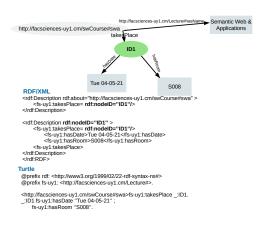
@prefix rdf: http://www3.org/1999/02/22-rdf-syntax-ns# @prefix fs-uy1: http://facsciences-uy1.cm/Lecture#>.

<http://facsciences-uy1.cm/swCourse#swa>fs-uy1:takesPlace[fs-uy1:hasDate "Tue 04-05-21"; fs-uy1:hasRoom "S008"].



Deferencable Blank Nodes

It might be important to dereference a node \longrightarrow Give this blank node a name



- Aggregation of facts
- Use to talk about single individual and aggregate into sets
- General Data structure to enumerate any resources or literals
- Only shortcuts, no additional semantic expressivity

Lists

Two types of lists :

• Container:

Lists

- o pen list, can be extends afterwards
- o extension (new entries) possible
- $\circ\,$ rdf :sequential : the sequence of the list is important

• Collection:

- simple linear list
- closed list, no further extension is possible
- o rdf:nil is for close lit



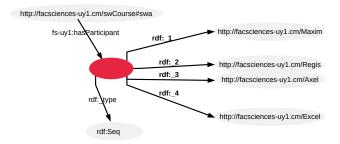
Lists: RDF Container

- The root node is assigned a container-type via rdf:type
- rdf :Bag
 - unordered set of elements
 - o there is no given order of elements
- rdf :Seq
 - o ordered set of elements
- rdf :Alt
 - o defines alternatives of elements
 - only one element of the given alternatives is relevant for the application



Lists: RDF Container

RDF Container



@prefix rdf: http://www3.org/1999/02/22-rdf-syntax-ns# @prefix fs-uy1: http://facsciences-uy1.cm/Lecture#>.



Lists: RDF Container - Turtle representation

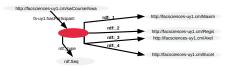


 $Lists: RDF\ Container\ -\ RDF/XML\ representation$



Lists: RDF Container

RDF Container



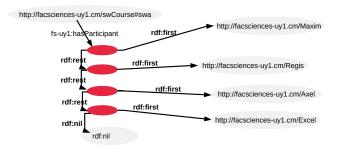
@prefix rdf: http://facsciences-uy1.cm/Lecture# @base http://facsciences-uy1.cm/.

```
:swa fs-uy1:hasParticipants[
a rdf:Seq;
rdf:_1 <Maxim>;
rdf:_2 <Regis>;
rdf:_3 <Axel>;
rdf:_4 <Excel>.
].
```



Lists: RDF-Collection

RDF Container: rdf:collection

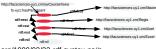


@prefix rdf: http://www3.org/1999/02/22-rdf-syntax-ns#>
@prefix fs-uy1: http://facsciences-uy1.cm/Lecture#>.



Lists: RDF-Collection

RDF Container: rdf:collection



@prefix rdf: ">http://facsciences-uy1.cm/Lecture#>"http://facsciences-uy1.cm/">http://facsciences-uy1.cm/.

```
:swa fs-uy1:hasParticipants[
rdf:first <Rayins; rdf:rest[
rdf:first <Axel>; rdf:rest[
rdf:first <Excel>; rdf:rest[
rdf:rest rdf:nil.
]]]]]
```

@prefix rdf: ">http://facsciences-uy1.cm/Lecture#>"http://facsciences-uy1.cm/">http://facsciences-uy1.cm/.

```
:swa fs-uy1:hasParticipants
(<Maxim> <Regis> <Axel> <Excel>).
```



Lists: RDF Collection - Turtle representation



Lists: RDF Collection - Turtle representation

Reification

- Consider the following statement of Sherlock Holmes: Dear Watson, ... I suppose that the Gardner has killed the Butler":
 - o There are several facts in there
 - How to represent these fact?
 - How to address the fact that Sherlock Holmes is referring to?
 - Connected facts together to a new fact: "the Gardner has killed the butler" and "Holmes suppose that the Gardner has killed the butler"
- Solution: Use RDF statement/RDF Reification in which a fact may become a subject



RDF-Reification

- Used for meta-modeling
- Used to model facts about facts
- Permits interleaving statements :
 i.e. to make statements about statements
- Example :
 - "Dear Watson, ... I suppose that the Gardner has killed the Butler"
 - Part 1: The Gardener has killed the Butler exv: Gardener exv: has Killed exv: Butler.
 - Part 2 : Sherlock Holmes supposes
 exv :SherlockHolmes exv :supposes????.



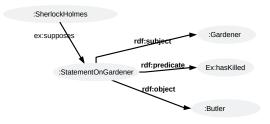
rdf:Statement Defines an RDF Statement, consisting of Subject, Predicate and Object

- rdf :subject the described resource
- rdf :predicate the original property
- rdf :object the value of the property

RDF-Reification



RDF-Reification



Sherlock Holmes supposes that the Gardener has killed the Butler



RDF-Reification

RDF-Reification

Sherlock Holmes supposes that the Gardener has killed the Butler



@prefix rdf: http://www3.org/1999/02/22-rdf-syntax-ns# @ex http://facsciences-uy1.cm/Crime# .

:SherlockHolmes ex:supposes:StatementOnGardener . :StatementOnGardener a rdf:Statement ;

rdf:subject :Gardener ; rdf:predicate ex:hasKilled ; rdf:object:Butler .



RDF-Reification: Turtle representation



 ${\sf RDF\text{-}Reification}: {\sf RDF/XML} \ representation$



RDF-Reification

- Used to model data provenance
- Formalizing statements about Reliability (Trust)
- Definition of Metadata about statements (Assertions, Statements)
- For (general) knowledge representation (ontologies construction)
 Used to transform Relations into Classes



Reification

- Used to model data provenance : Who has made a statement
- Used to formalize statements about reliability and trust

Warning : can cause the problems of decidability and type conflicts (classes, individuals, statements)



RDF and Data Integration

Example of a Bibliography Database

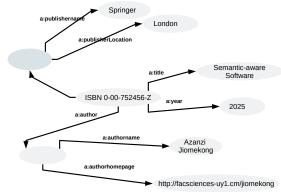
	ID			Author		Title		Publisher	Year
Books	ISBN 0-00-752456-Z			AJ-1254		Semantic-Aw Software	are	S-001	2025
Authors	ID	Name			Homepage				
	AJ-123	Azan	zi Jiome	ekong	htt	http://facsciences-uy1.cm/jiomekong			ng
					_				
	ID		Publisher		Lo	ocation			
Publisher	Publisher S-001 Sp		Spring	er		ondon			

Suppose we want to transform this database in a RDF knowledge base



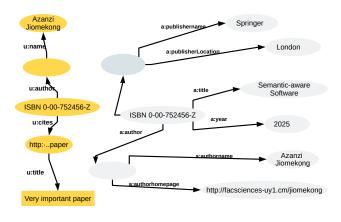
RDF and Data Integration

Database export into a set of relations





RDF and Data Integration





RDF and Data Integration



- Two tables with different schema
- Schema mapping is a difficult problem
- · Easy to do in RDF





Element	Description			
URI/IRI	Unique string used to identify a resource			
HTTP	Protocol used to fetch resources on the Web			
XML	Defines a set of rules for encoding document for data exchange			
RDF	W3C standard data model for data transfert on the Web			



Summary

RDF XML serialization

Element	Description
rdf :RDF	The root of an RDF document
rdf :Description	Container for the description of a resource
rdf :about	Defines the resource being described
rdf :resource	Defines a resource to identify a property
rdf :ID	Defines the ID of an element
rdf :DataType	Defines
rdf :parseType	Defines how an element should be parsed
rdf :nodeID	Defines the ID of an element node
rdf :li	Defines a list
rdf_n	Defines a node
rdf :Statement	Represent the class of RDF statement



Summary and further questions

In this course, we replied to the following questions:

- How to represent things?
- How to encode and represent facts?

But questions remain:

- How to represent knowledge on the Web?
- How to store RDF and RDFS data?

The responses to these questions are given in the next course