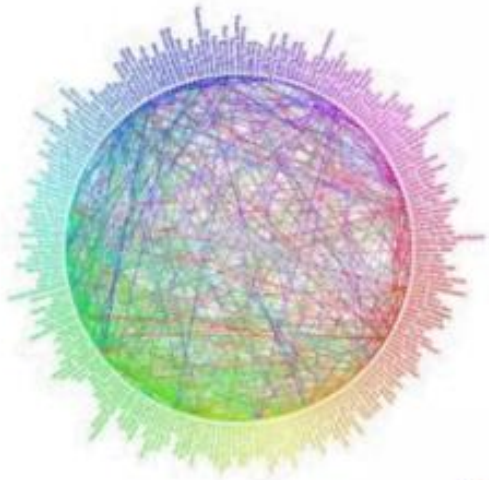
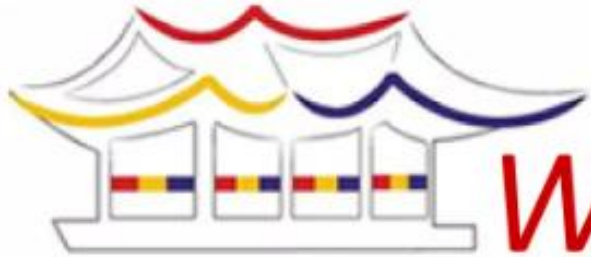
An Introduction to Semantic Web and Linked Data

Semantic Web and Linked Data

or how to link data and schemas on the web





WWW 2014

a little semantics in a world of links

semantic web: linked data and semantics of schemas

- The web we use today is largely a web of documents designed for human consumption. The **Semantic Web**, a vision proposed by **Tim Berners-Lee**, extends the current web by giving information **well-defined meaning**, enabling computers and people to work in cooperation.
- The **Linked Data** principles form the foundation of this Semantic Web by using standard web protocols and URIs to interconnect data across the web.

الويب اليوم عبارة عن شبكة من المستندات المصممة للاستخدام البشري. يُوسّع الويب الدلالي، وهو رؤية اقترحها تيم بيرنرز لي، نطاق الويب الحالي من خلال إعطاء المعلومات معنى واضحًا، مما يُمكّن أجهزة الكمبيوتر والبشر من العمل معًا. تُشكّل مبادئ **البيانات المترابطة** أساس هذا الويب الدلالي، باستخدام بروتوكولات الويب القياسية ومعرفات الموارد الموحدة (URIs)لربط البيانات عبر الويب.

What is the Semantic Web?

- A framework to make Internet data machine-readable.
- Represents data in a structured, meaningful way.
- Enables automated agents to access the Web more intelligently.

ما هو الويب الدلالي؟

1. إطار عمل لجعل بيانات الإنترنت قابلة للقراءة آليًا.
2. يُمثل البيانات بطريقة منظمة وذات معنى.
3. يُمكّن الوكلاء الآليين من الوصول إلى الويب بذكاء أكبر.

تدعم هذه التقنيات مجتمعةً شبكة بيانات تترابط فيها المعرفة، ويسهل الوصول إليها ومعالجتها بواسطة البشر والآلات.

Key Goals:

Interoperability between systems.

Data integration from multiple sources.

Reasoning over data using logic and ontologies.

الأهداف الرئيسية

- التوافق بين الأنظمة.
- تكامل البيانات من مصادر متعددة.
- تحليل البيانات باستخدام المنطق والأنطولوجيا.

Key Technologies of the Semantic Web

Technology	Description
RDF (Resource Description Framework)	A data model for describing resources and relationships using triples (subject-predicate-object).
OWL (Web Ontology Language)	Describes rich and complex knowledge about things, groups of things, and relations.
SPARQL	A query language for RDF data.
URI (Uniform Resource Identifier)	Identifies resources uniquely on the web.
RDFS (RDF Schema)	A vocabulary for describing properties and classes of RDF resources.

you loose!

we identify and interpret information,

machines don’t.

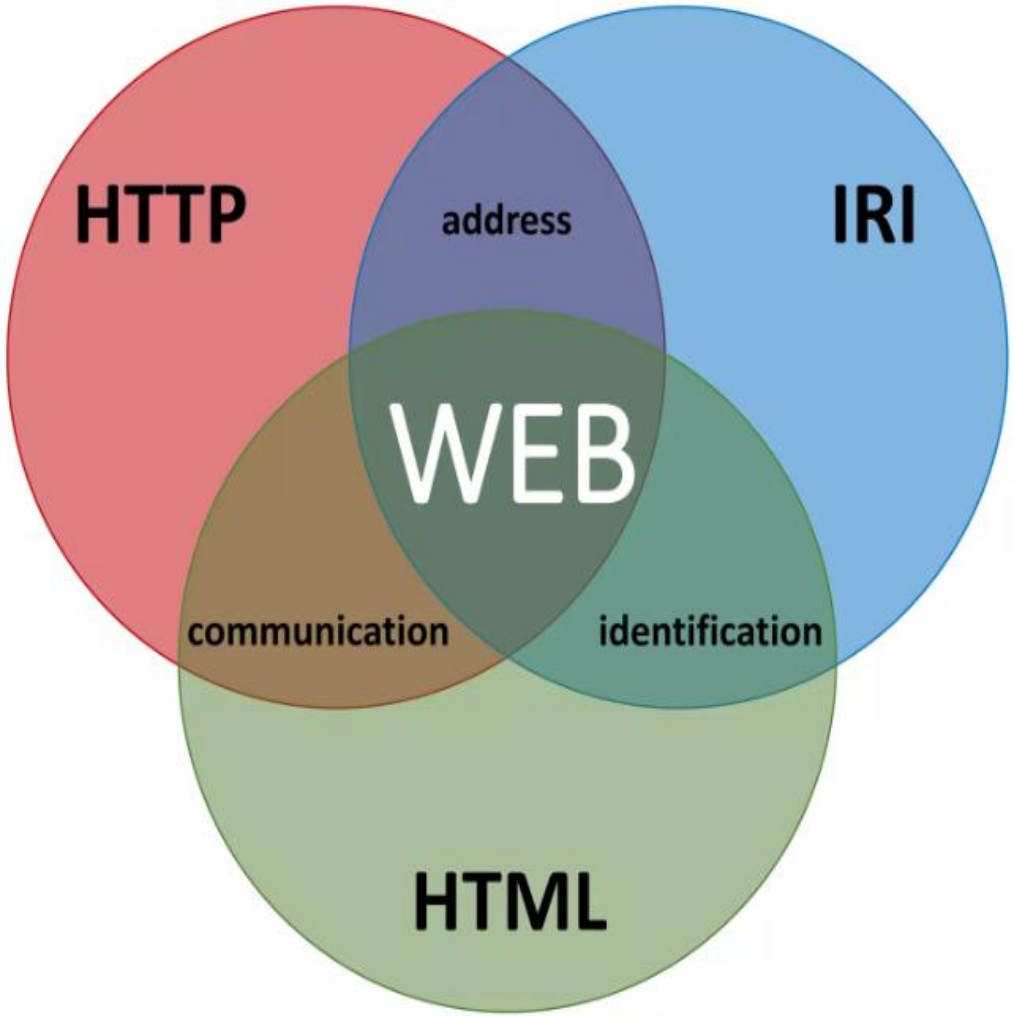
A WEB OF LINKED DATA

What is Linked Data?

A method of publishing structured data so that it can be **interlinked and become more useful**.

Principles of Linked Data (Tim Berners-Lee):

1. Use URIs to identify things.
2. Use HTTP URIs so people can look them up.
3. When someone looks up a URI, provide useful information using standards like RDF and SPARQL.
4. Include links to other URIs to discover more related data.



ما هي البيانات المرتبطة؟

طريقة لنشر البيانات المنظمة بحيث يمكن ربطها ببعضها البعض وزيادة فائدتها.

مبادئ البيانات المرتبطة

(تيم بيرنرز لي):

- استخدام عناوينURL لتحديد الأشياء.
- استخدام عناوين HTTP URIs

ليتمكن المستخدمون من البحث عنها.

- عند البحث عن عنوان URI، قدّم معلومات مفيدة باستخدام معايير مثل RDF وSPARQL
- أدرج روابط لعناوين URIs أخرى لاكتشاف المزيد من البيانات ذات الصلة.

Linked Data vs. Semantic Web

Linked Data is a **subset** and **practical implementation** of the Semantic Web.

Semantic Web includes **reasoning**, **inference**, and **ontologies** beyond just linking data.

البيانات المرتبطة مقابل الويب الدلالي

البيانات المرتبطة هي جزء من **الويب الدلالي** وتطبيق عملي له.

يشمل الويب الدلالي **الاستفهام والاستدلال والوجوديات**، بما يتجاوز مجرد ربط البيانات.

Example of RDF Triple

Statement: *"Dr. Luma Akram teaches Semantic Web."*

Subject	Predicate	Object
Dr. Luma Akram	teaches	Semantic Web

This is written in RDF as:

<http://example.org/people/LumaAkram> <http://xmlns.com/foaf/0.1/teaches> <http://example.org/topics/SemanticWeb> .

Real-world Applications

DBpedia – Structured data extracted from Wikipedia.

Wikidata – A free knowledge base that can be read and edited by humans and machines.

Healthcare and Life Sciences – Linking clinical trial data, genes, and literature.

Government Data – Open data initiatives, e.g., data.gov.uk.

تطبيقات واقعية

- DBpedia بيانات مُهيكلة مُستخرجة من ويكيبيديا.
- Wikidata قاعدة معارف مجانية يُمكن للبشر والآلات قراءتها وتحريرها.
- الرعاية الصحية وعلوم الحياة ربط بيانات التجارب السريرية والجينات والأدبيات.
- بيانات الحكومة - مبادرات البيانات المفتوحة، مثل data.gov.uk.

Benefits

- ❖ Better data integration.
- ❖ Enhanced data discovery and reuse.
- ❖ Improved decision-making with inference and analytics.
- ❖Transparent and explainable data relationships.

الفوائد

1. تكامل أفضل للبيانات.
2. اكتشاف البيانات وإعادة استخدامها بشكل أفضل.
3. تحسين عملية اتخاذ القرارات من خلال الاستدلال والتحليل.
4. علاقات بيانات شفافة وقابلة للتفسير.

Challenges

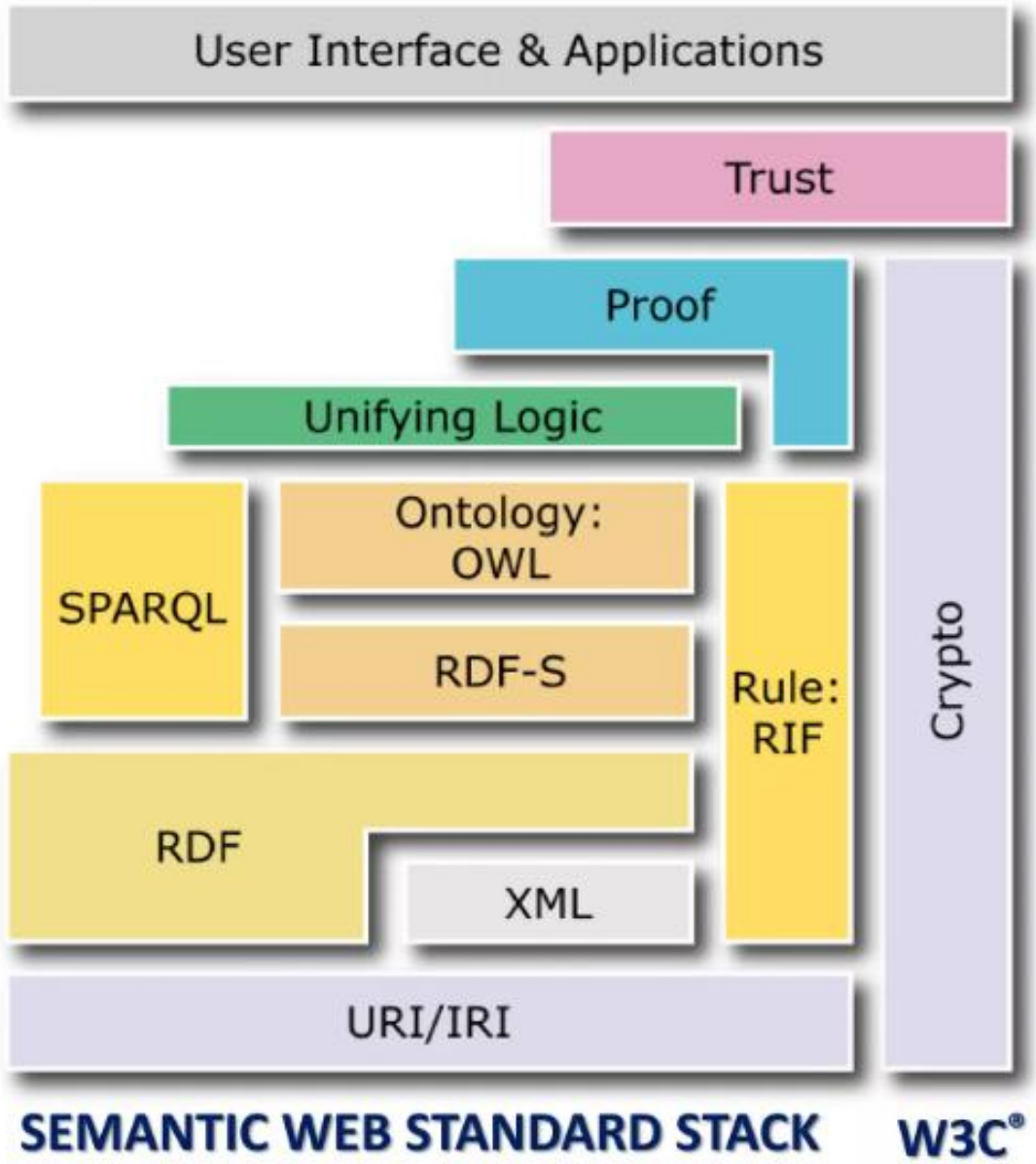
Data quality and consistency.

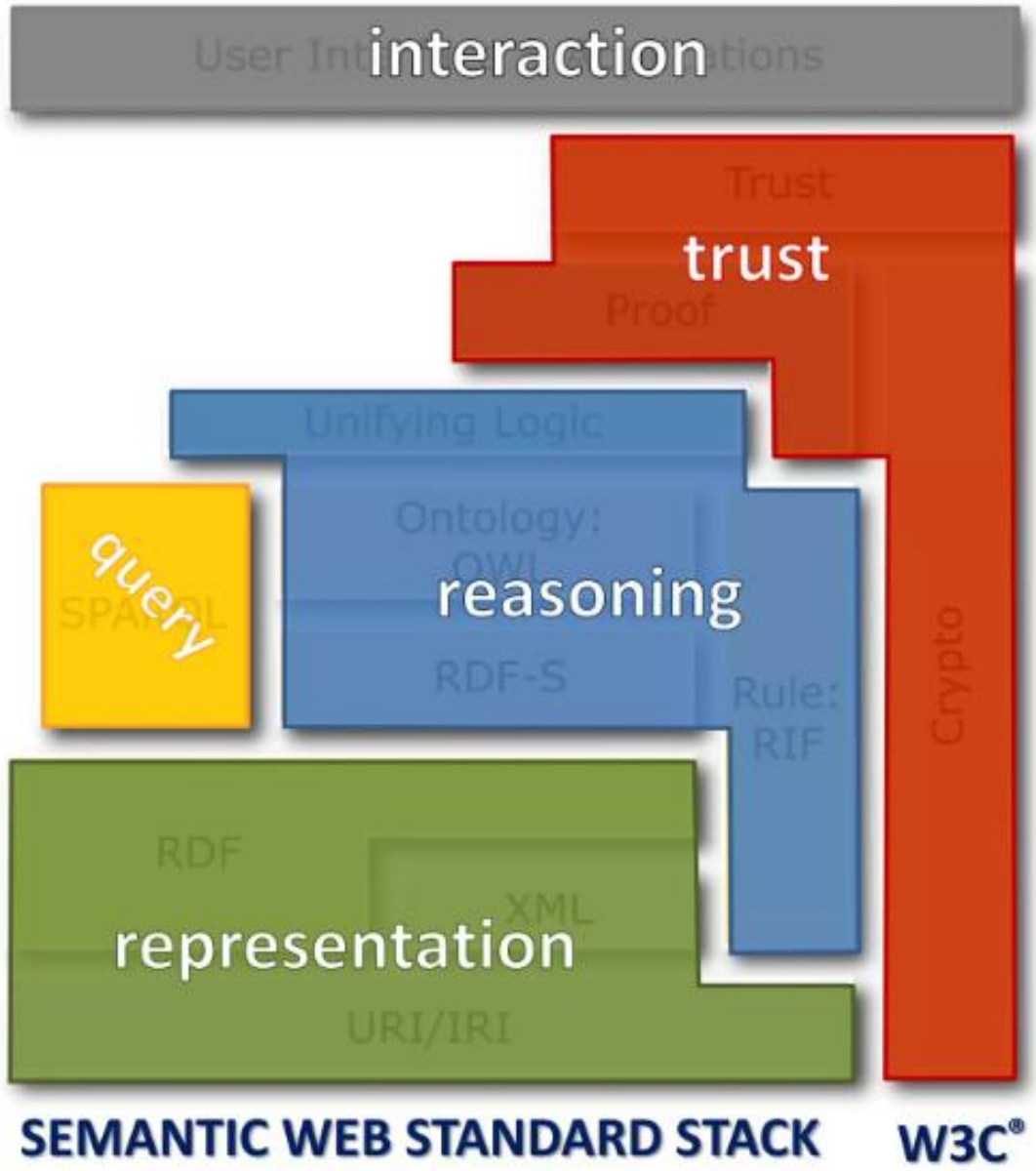
Scalability of reasoning engines.

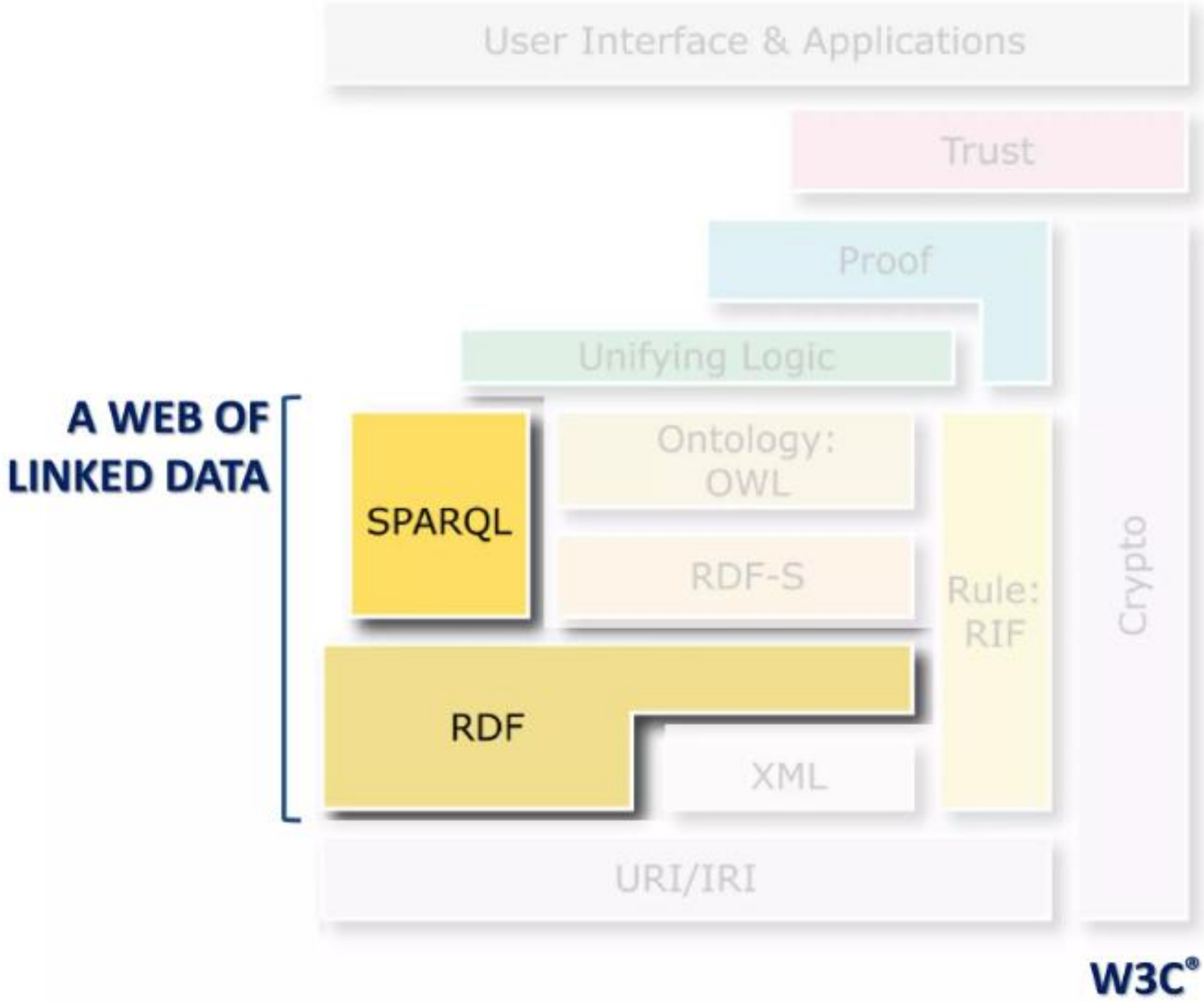
Security and trust issues.

Lack of widespread adoption.

- جودة البيانات واتساقها.
- قابلية توسع محركات الاستدلال.
- مشاكل تتعلق بالأمن والثقة.
- قلة التبني على نطاق واسع.


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


interaction
trust
query
reasoning
representation
Crypto
SEMANTIC WEB STANDARD STACK
W3C®


User Interface & Applications
Trust
Proof
Unifying Logic
A WEB OF LINKED DATA
SPARQL
Ontology: OWL
RDF-S
Rule: RIF
Crypto
RDF
XML
URI/IRI
W3C®

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

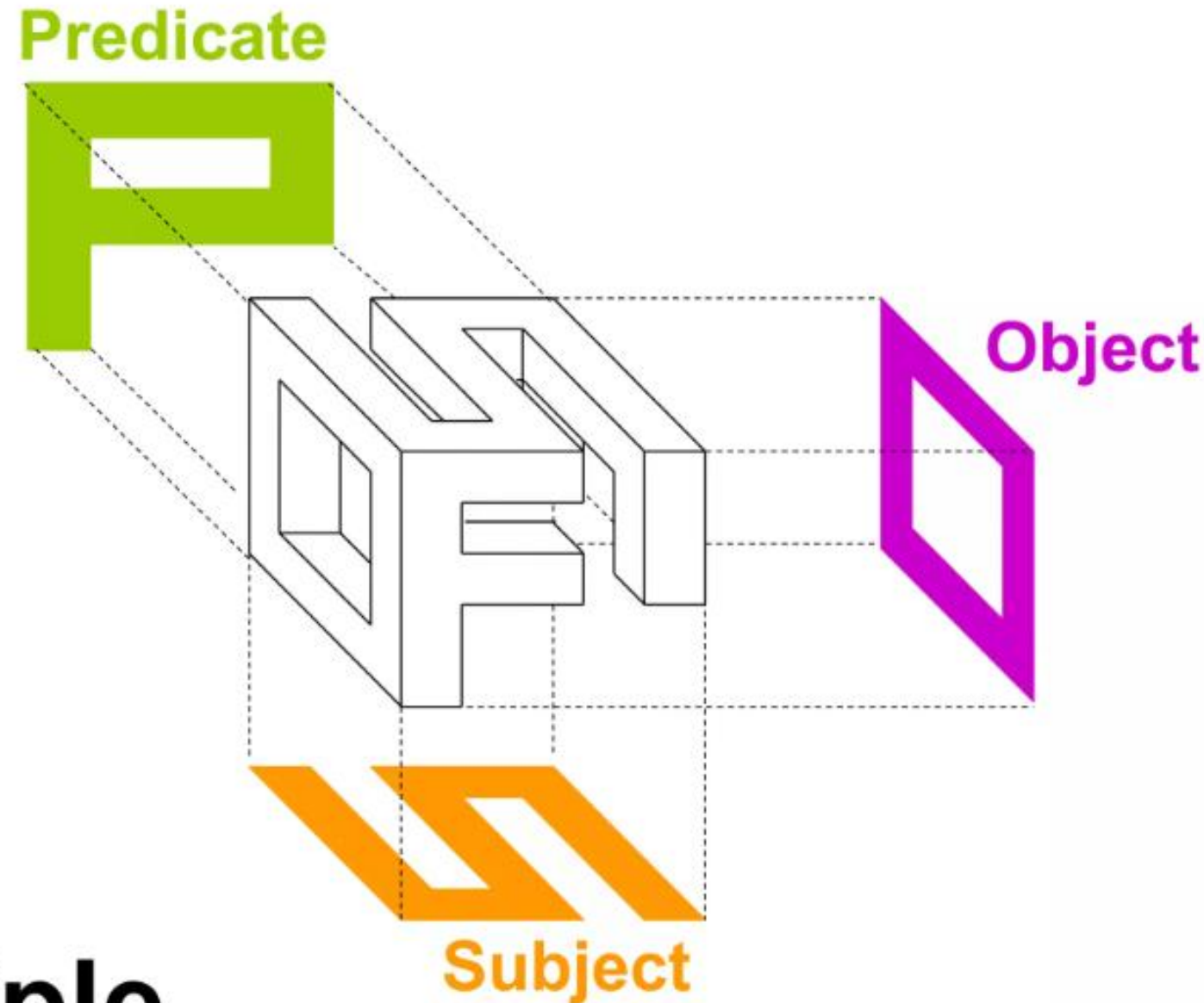
(subject , predicate , object)

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

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

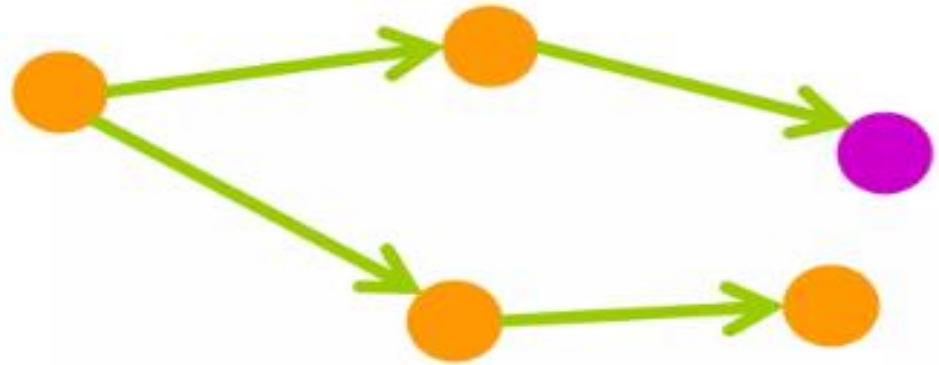
(doc.html , author , Fabien)
(doc.html , theme , Music)

(subject , predicate , object)



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)

Fabien
author
doc.html
theme
Music

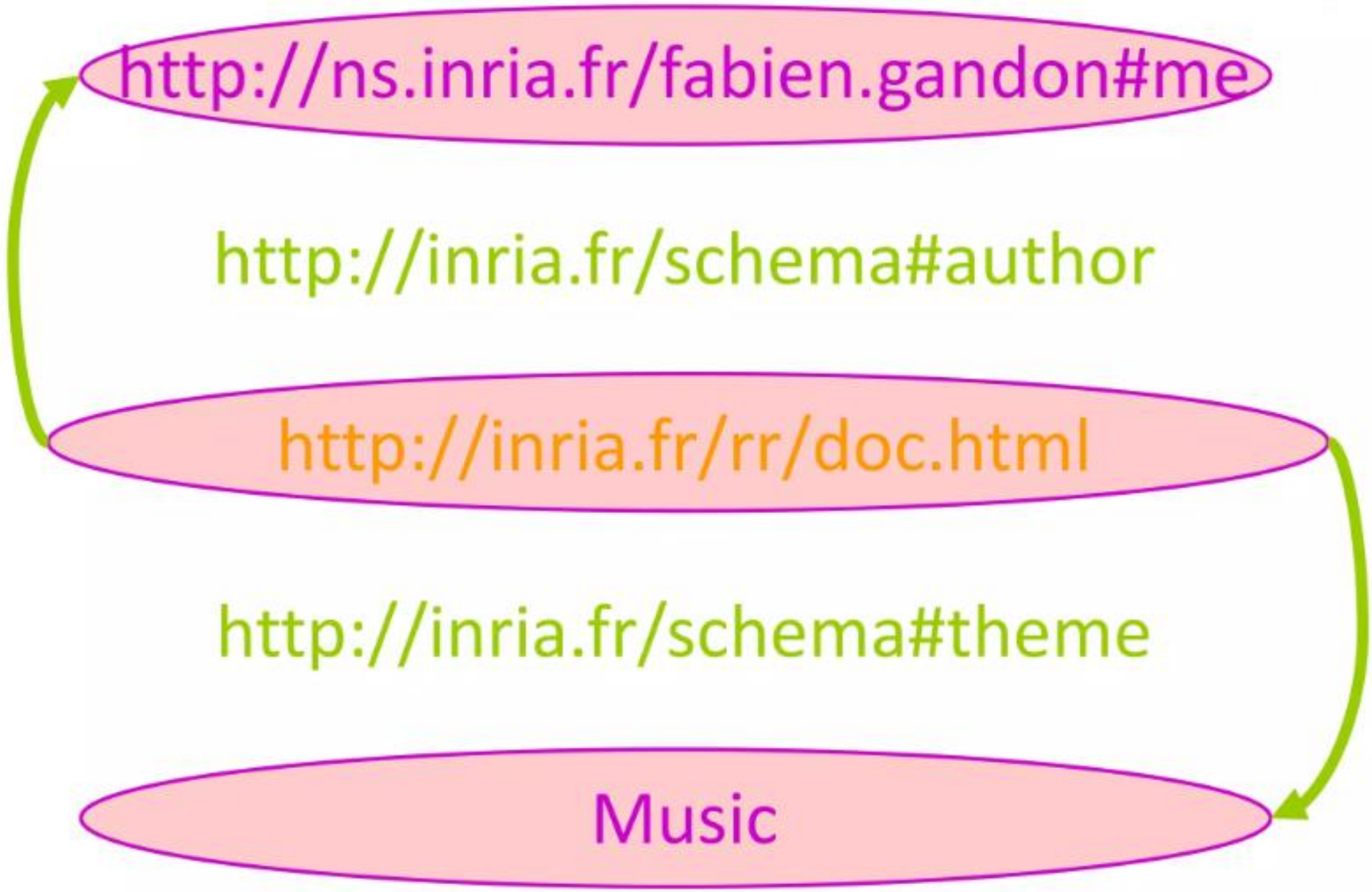
URL → IRI

identify what exists on the web

http://my-site.fr

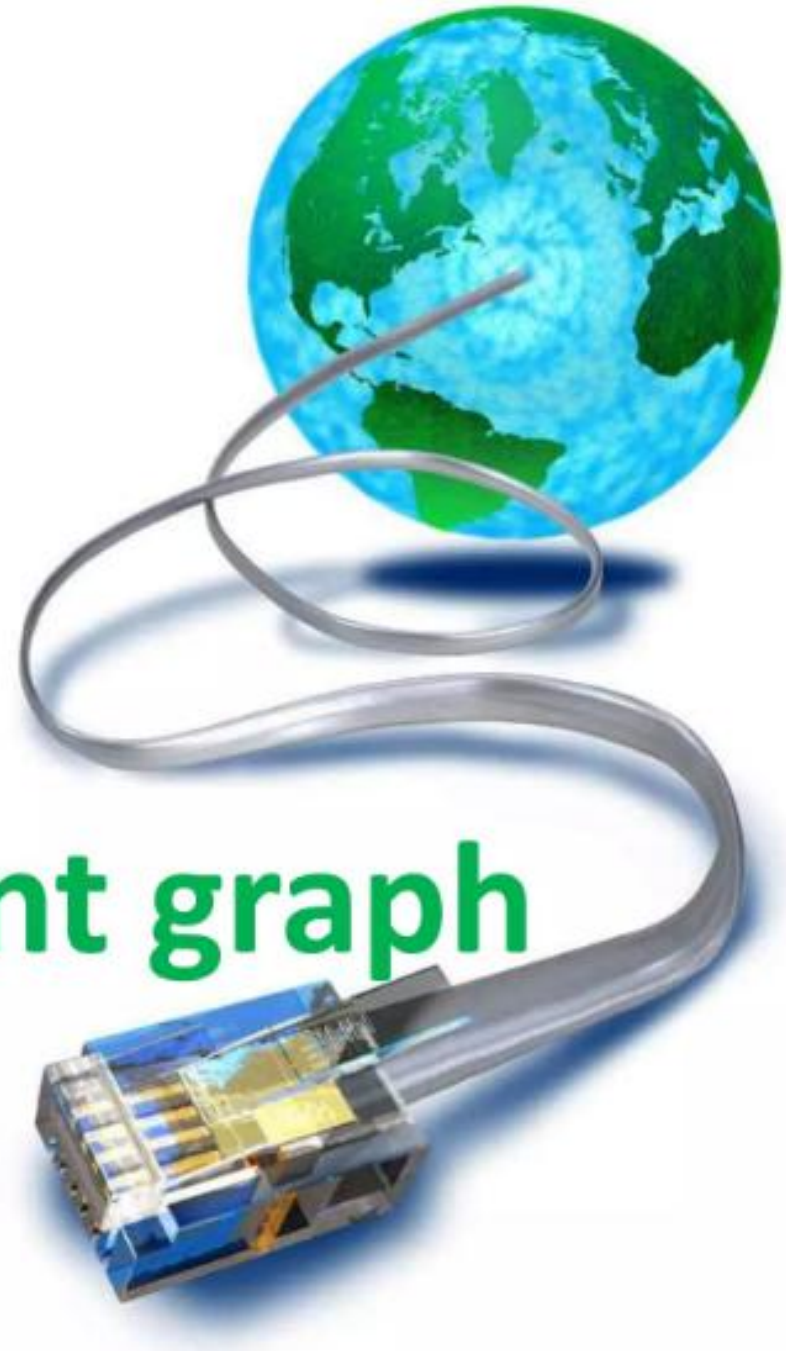
identify, on the web, what exists

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

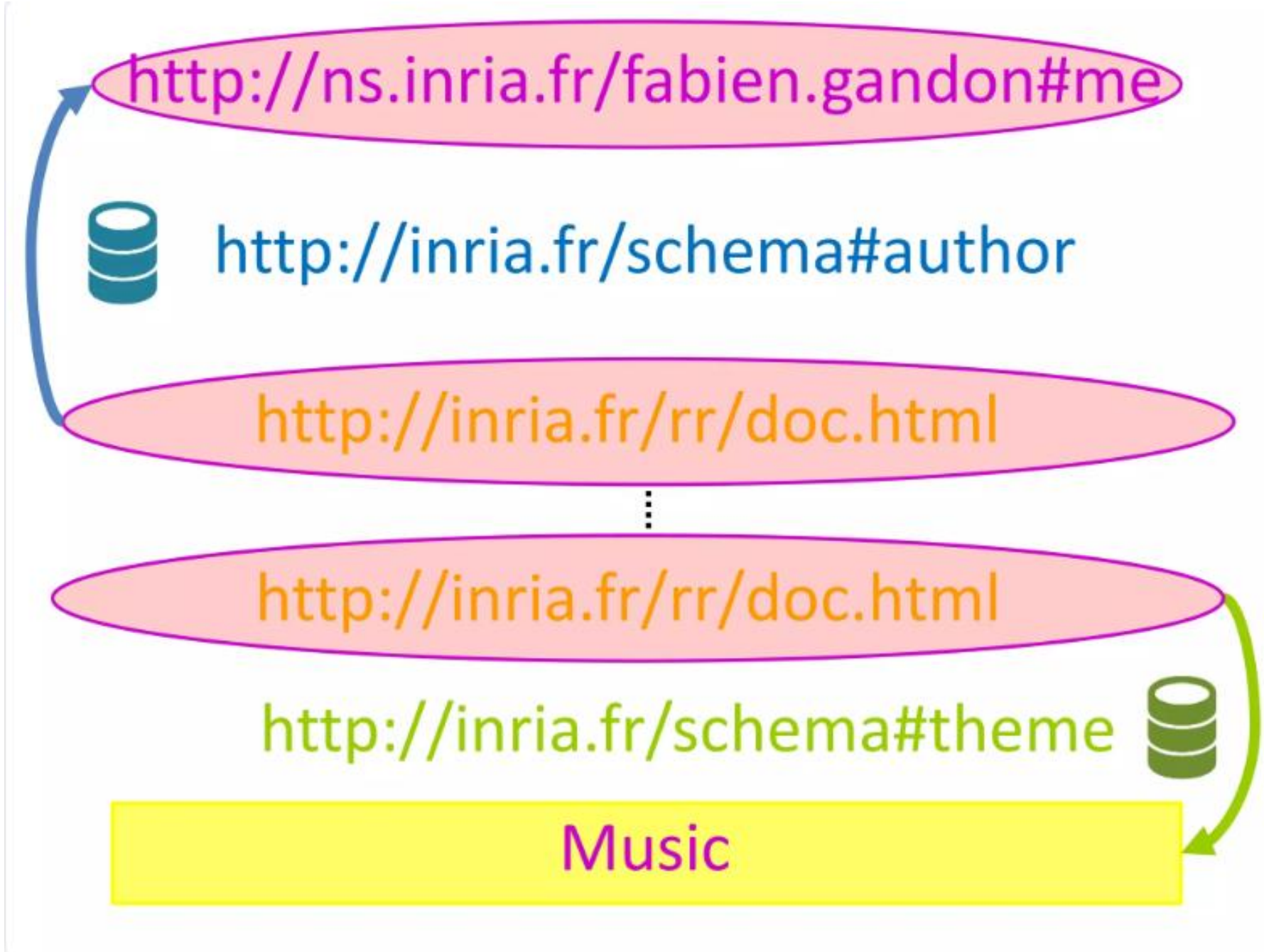
global giant graph



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/rr/doc.html
http://inria.fr/schema#theme
Music

< RDF /> has an XML syntax

```
<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>
```

RDF has other syntaxes

(Turtle, TriG, N-Triples, N-Quads, JSON, RDFa)