



Manuel ATENCIA, Jérôme DAVID et Philippe GENOUD
Université Grenoble Alpes, CNRS, Grenoble INP, Inria, LIG

WEB DES DONNÉES: DONNÉES OUVERTES ET LIÉES

Les principes – Les technologies

Plan

- Introduction
- De l'Open Data au Linked Open Data (1^{ère} partie)
- Les principes des données ouvertes et liées
 - Nommer les données: URI
 - Modéliser les données: RDF
 - Interroger les données: SPARQL
- De l'Open Data au Linked Open Data (2^{ème} Partie)
- Conclusion
- Bibliographie

Du web des documents au web des données

- le Web peut être vu comme une (très) grande base de données (base de connaissances) décentralisée d'informations accessibles aux machines.

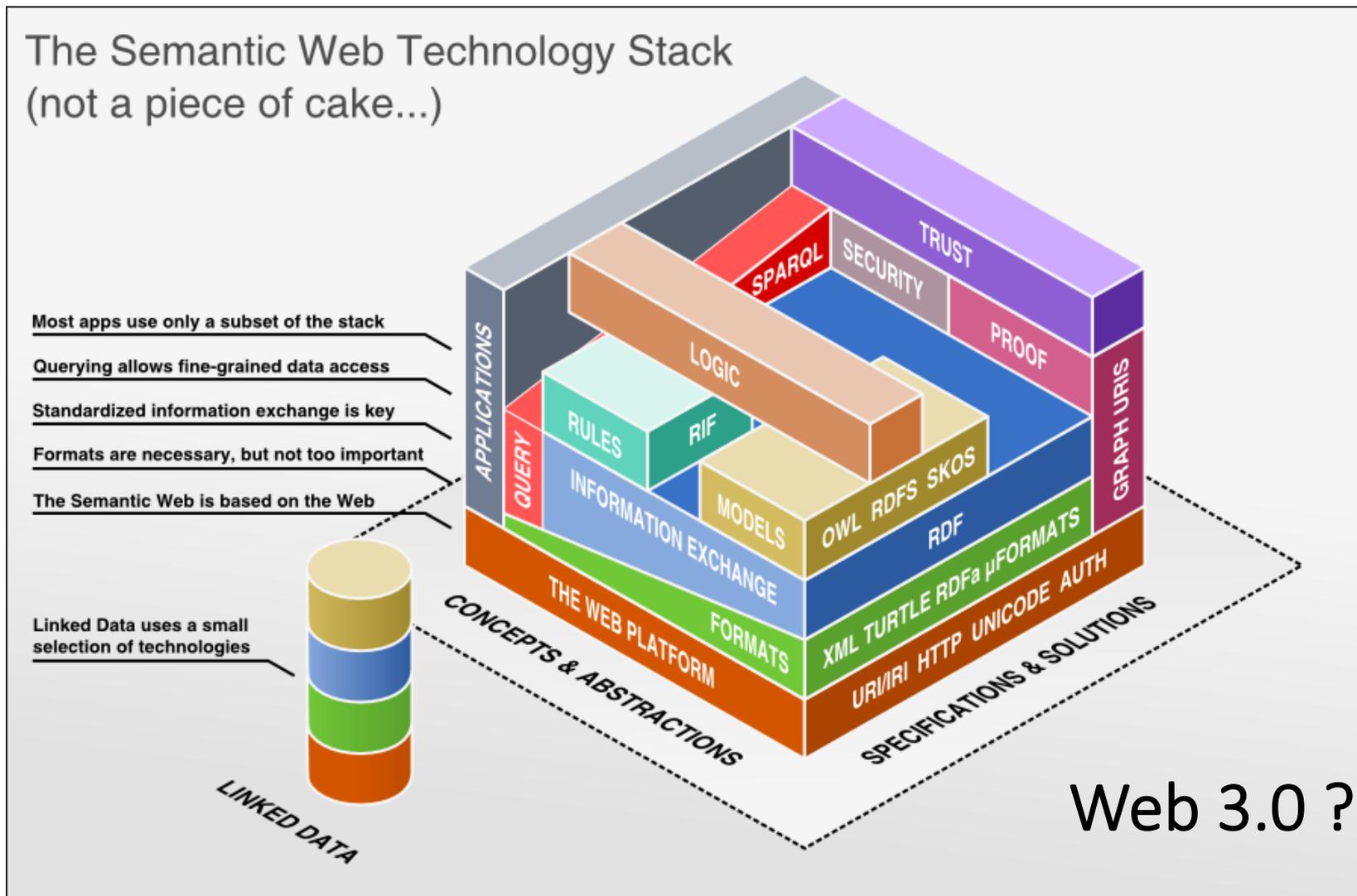
The Data Activity recognizes and works to overcome this diversity to facilitate potentially Web-scale data integration and processing. It does this by providing standard data exchange formats, models, tools, and guidance.

The screenshot shows the W3C Data Activity website. The main heading is "W3C DATA ACTIVITY Building the Web of Data". Below this, a paragraph states: "More and more Web applications provide a means of accessing data. From simple visualizations to sophisticated interactive tools, there is a growing reliance on the availability of data which can be 'big' or 'small', of diverse origin, and in different formats; it is usually published without prior coordination with other publishers — let alone with precise modeling or common vocabularies. The Data Activity recognizes and works to overcome this diversity to facilitate potentially Web-scale data integration and processing. It does this by providing standard data exchange formats, models, tools, and guidance." This paragraph is highlighted in a light green box. Below it, a red box contains the text: "The overall vision of the Data Activity is that people and organizations should be able to share data as far as possible using their existing tools and working practices but in a way that enables others to derive and add value, and to utilize it in ways that suit them. Achieving that requires a focus not just on the interoperability of data but of communities." To the left, a sidebar lists "ACTIVE GROUPS" including "Linked Data Platform Working Group", "Data on the Web Best Practices Working Group", "CSV on the Web Working Group", "Semantic Web Interest Group", "Semantic Web Health Care and Life Sciences Interest Group", and "Data Activity Coordination Group". A "NEARBY" section at the bottom left shows "Data Activity Blog" with the URL "www.w3.org/standards/".

people and organizations should be able to share data as far as possible using their existing tools and working practices but in a way that enables others to derive and add value, and to utilize it in ways that suit them

Web 3.0 ?

Du web des documents au web des données



<http://www.bnode.org/blog/tag/layer%20cake> (Benjamin Nowack)

Les Données Liées (Linked Data)



"The Semantic Web isn't just about putting data on the web. It is about making links, so that a person or machine can explore the web of data. With linked data, when you have some of it, you can find other, related, data."

Tim Berners-Lee - 2006

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

Linked Data Principles : un ensemble de bonnes pratiques pour publier et inter-relier des données structurées sur le web.

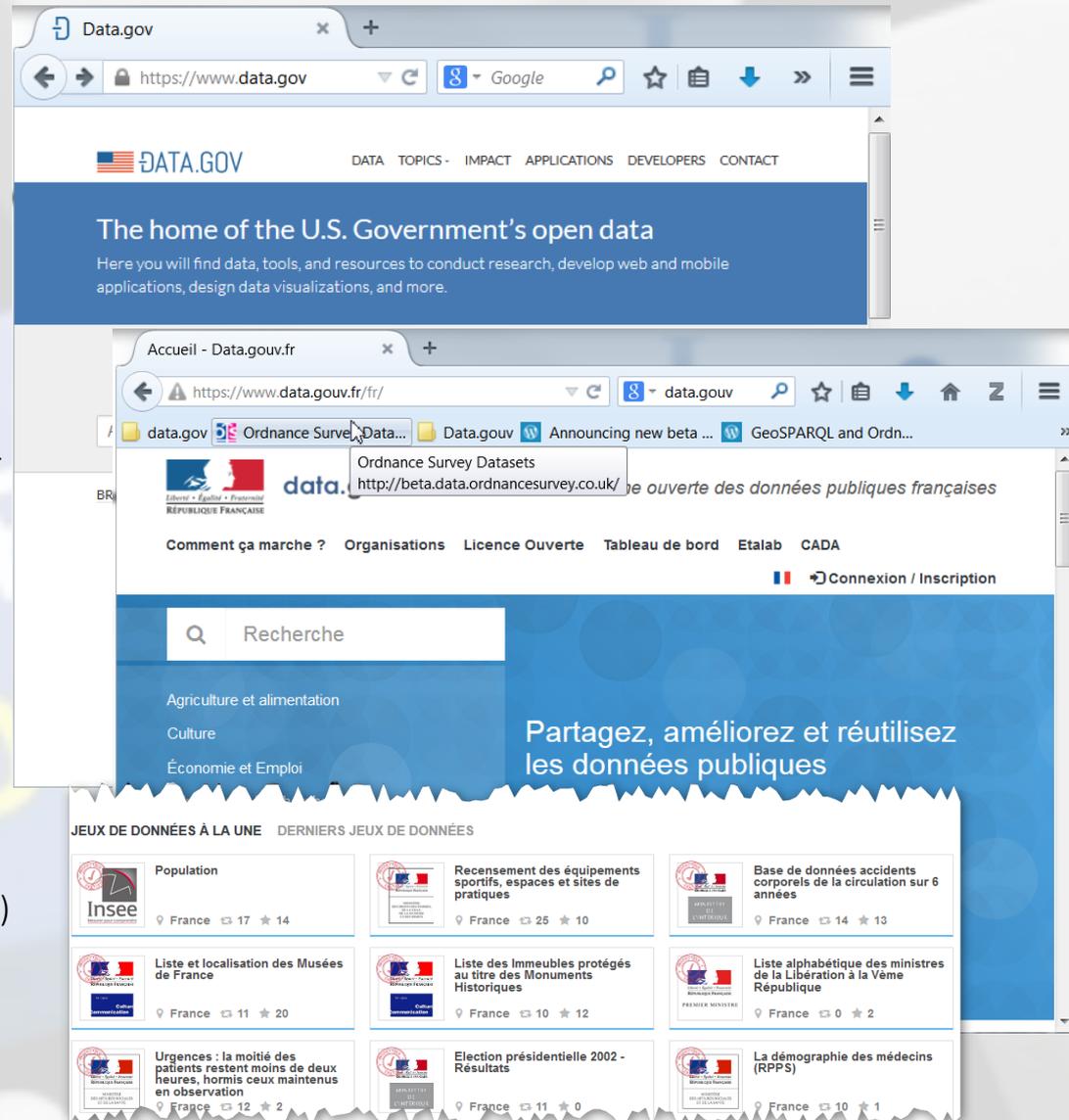
Les données ouvertes (Open Data)

- **Mouvement Open Data**

- “A piece of content or data is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike.”

<http://opendefinition.org/>

- une idée ancienne mais un terme récent gagnant en popularité
 - grâce au développement de l'Internet et du World Wide Web
 - avec le lancement d'initiatives gouvernementales pour l'ouverture des données (Data.gov, data.gouv.fr ...)

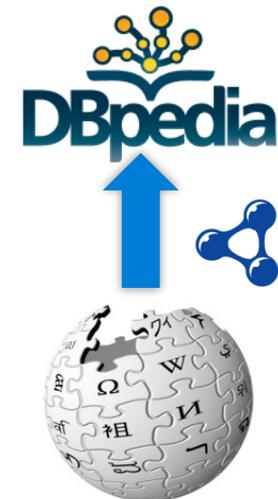
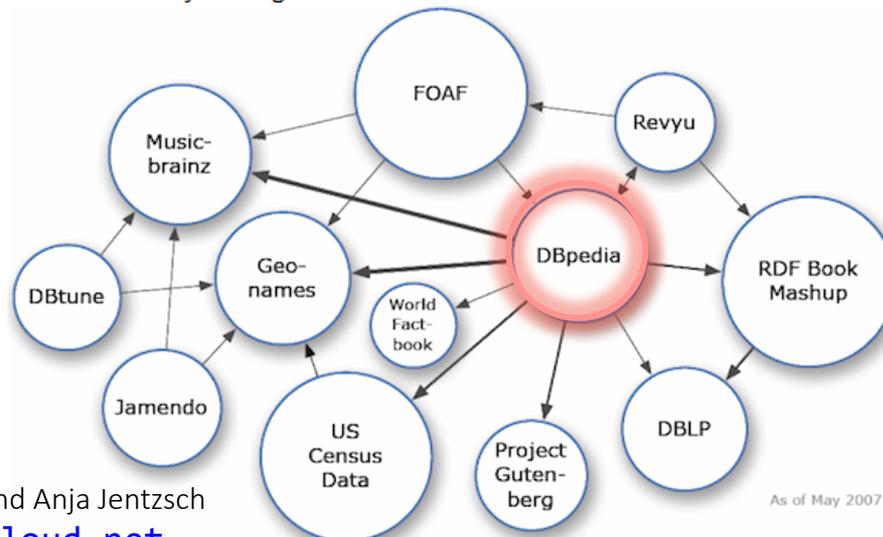


Ouvertes Les Données Liées (Linked Data) Open



The [Open Data Movement](#) aims at making data freely available to everyone. There are already various interesting open data sets available on the Web. Examples include [Wikipedia](#), [Wikibooks](#), [Geonames](#), [MusicBrainz](#), [WordNet](#), the [DBLP bibliography](#) and many more which are published under [Creative Commons](#) or [Talis](#) licenses.

The goal of the W3C SWEQ Linking Open Data community project is to extend the Web with a data commons by publishing various open data sets as RDF on the Web and by setting RDF links between data items from different data sources.

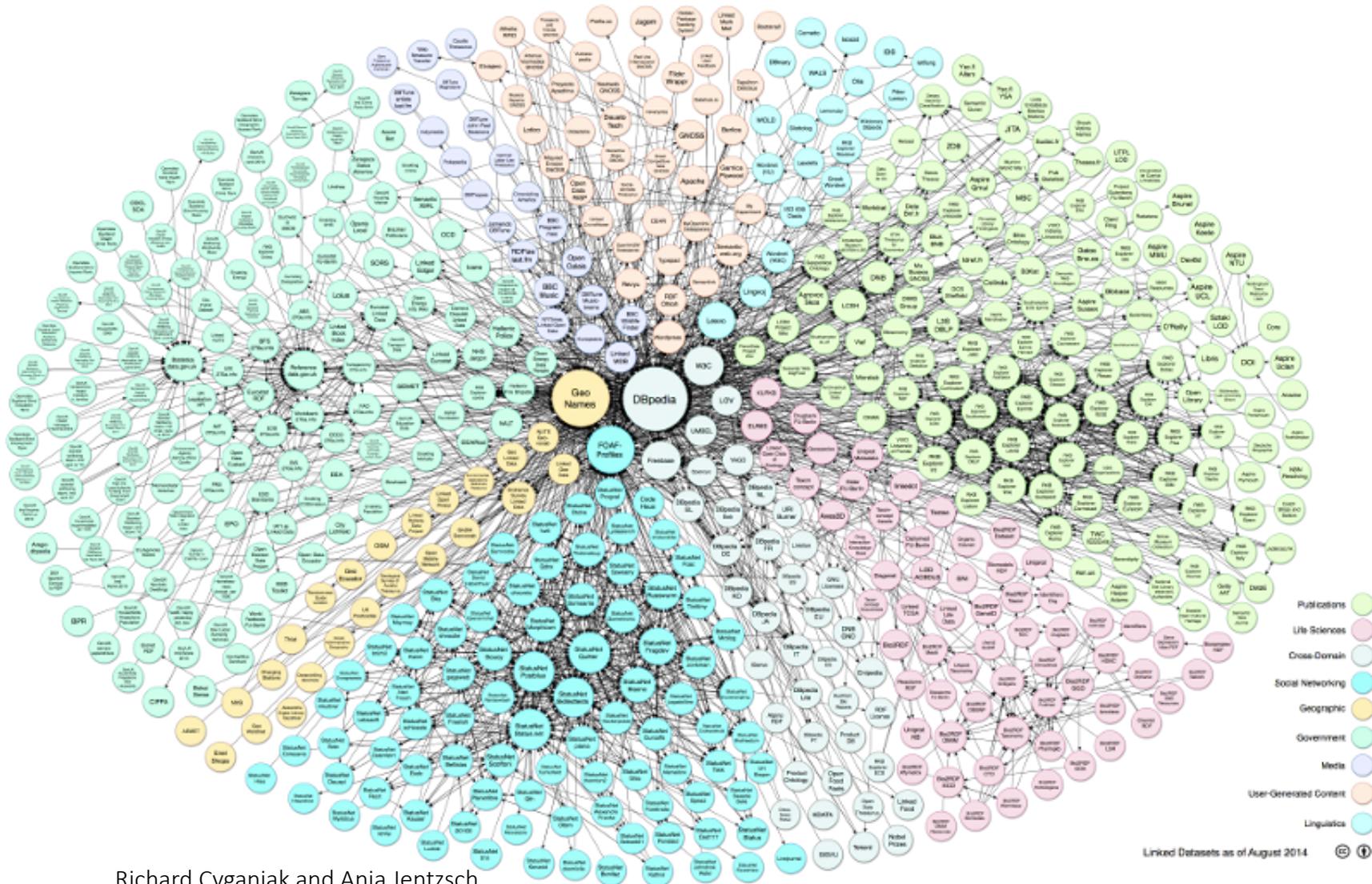


Richard Cyganiak and Anja Jentzsch
<http://lod-cloud.net>

As of May 2007

Ouvertes Les Données Liées (Linked Data)

Open Linked Data



Richard Cyganiak and Anja Jentzsch

<http://lod-cloud.net>

Plan

- Introduction
- De l'Open Data au Linked Open Data (1^{ère} partie)
- Les principes des données ouvertes et liées
 - Nommer les données: URI
 - Modéliser les données: RDF
 - Interroger les données: SPARQL
- De l'Open Data au Linked Open Data (2^{ème} Partie)
- Conclusion
- Bibliographie

Des données ouvertes aux données ouvertes et liées



- 2006: définit les principes de base pour les données liées
- 2010: ajout d'un système de notation à 5 ★ pour les données ouvertes et liées

"in order to encourage people -- especially government data owners -- along the road to good linked data..."

Tim Berners-Lee

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



Publication des données sur le web



rendez vos données disponibles sur le web (quelque soit le format) sous une licence ouverte.



Temperature forecast for Galway, Ireland	
Day	Lowest Temperature (°C)
Saturday, 13 November 2010	2
Sunday, 14 November 2010	4
Monday, 15 November 2010	7

Publication des données sur le web

-  coûts et bénéfices
 - Pour l'utilisateur ...
 - ✓ Il peut les regarder.
 - ✓ Il peut les imprimer.
 - ✓ Il peut les stocker localement (sur un disque, une clé USB...).
 - ✓ Il peut réutiliser les données ... en les entrant dans un autre système.
 - ✓ S'il le souhaite il peut modifier les données réutilisées.
 - ✓ Il peut les partager avec qui il veut.
 - Pour celui qui publie ...
 - ✓ Facilité de publication.
 - ✓ Pas besoin d'expliquer aux autres qu'ils peuvent utiliser vos données.

Michael Hausenblas <http://5stardata.info/>

Accès aux données dans format structuré

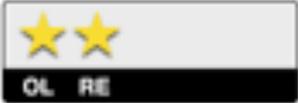
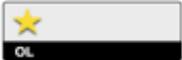


rendez les disponibles dans un format structuré (par exemple, à une image d'une table on préférera un format Excel)

The screenshot shows the Microsoft Excel interface with a table containing the following data:

	A	B	C
1	Temperature forecast for Galway, Ireland		
2			
3	Day	Lowest Temperature (°C)	
4	Saturday, 13 November 2010	2	
5	Sunday, 14 November 2010	4	
6	Monday, 15 November 2010	7	
7			

Accès aux données dans format structuré

-  coûts et bénéfices
 - Pour l'utilisateur... en plus des avantages de 
 - ✓ Il peut traiter directement les données avec le logiciel (propriétaire) pour les visualiser, effectuer des calculs, les agréger, etc.
 - ✓ Il peut les exporter dans un autre format structuré.
 - Pour celui qui publie ...
 - ✓ La publication est toujours facile.

Accès aux données dans format ouvert

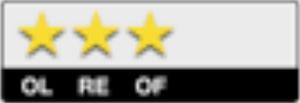


utilisez des formats non propriétaires (par exemple on préfère CSV à Excel)

```
gtd-3.csv
1 "Temperature forecast for Galway, Ireland",
2
3 "Day", "Lowest Temperature (C)"
4 "Saturday, 13 November 2010",2
5 "Sunday, 14 November 2010",4
6 "Monday, 15 November 2010",7

length:164 lines:6 Ln:1 Col:1 Sel:0 Macintosh ANSI as UTF-8 INS
```

Accès aux données dans format ouvert

-  coûts et bénéfices
 - Pour l'utilisateur... en plus des avantages de 
 - ✓ Possibilité de manipuler les données de la manière qu'il souhaite sans être confiné aux fonctionnalités d'un logiciel particulier.
 - Pour celui qui publie ...
 - ✗ L'exportation des données à partir du format propriétaire peut nécessiter des plug-ins ou convertisseurs.
 - ✓ Mais la publication demeure assez simple..

Accès aux données dans format ouvert

Excellent! The data is not only available via the Web but now everyone can use the data easily. On the other hand, it's still data on the Web and not [data in the Web](#).

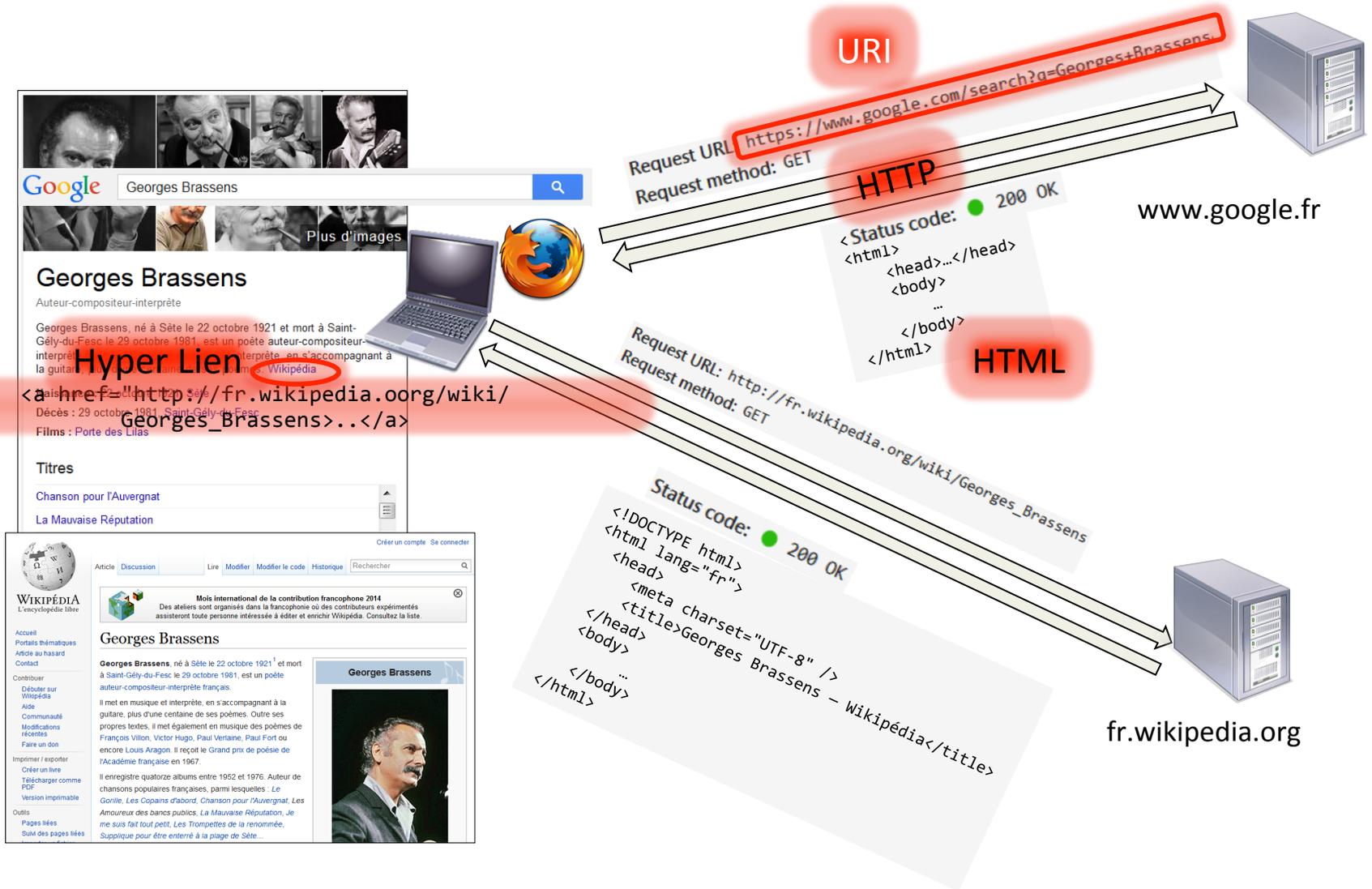




Le WEB: raisons d'un succès

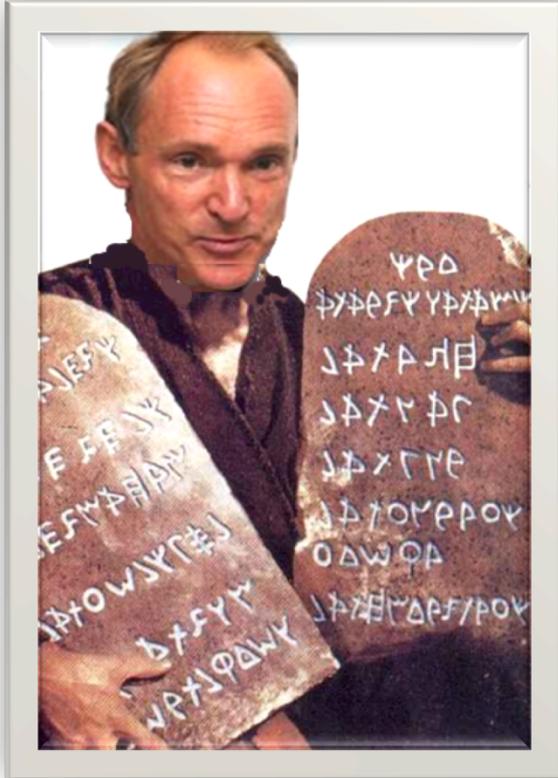
- WEB : un espace d'information global combinant simplicité avec décentralisation et ouverture
- succès du Web basé sur une architecture construite sur un ensemble de standards simples
 - un mécanisme d'identification global : URI (Uniform Resource Identifier)
 - un mécanisme d'accès universel : protocole HTTP
 - un format largement utilisé pour les contenus : HTML
 - possibilité de définir des **hyperliens** entre documents résidents sur différents serveurs

Architecture du Web





Données liées : les principes



Linked Data Principles

Tim Berners-Lee

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

idée de base: appliquer l'architecture générale du World Wide Web pour partager des données structurées dans un espace global.



1. Use **URIs** as names for things.
2. Use **HTTP URIs**, so that people can look up those names.
3. When someone looks up a URI, **provide useful information, using the standards** (RDF, SPARQL).



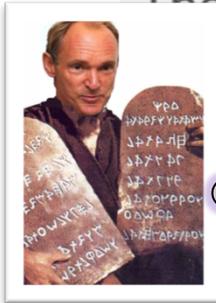
4. Include **links** to other URIs, so that they can discover more things.



Plan

- Introduction
- De l'Open Data au Linked Open Data (1^{ère} partie)
- Les principes des données ouvertes et liées
 - Nommer les données: URI
 - Modéliser les données: RDF
 - Interroger les données: SPARQL
- De l'Open Data au Linked Open Data (2^{ème} Partie)
- Quelques jeux de données du LOD cloud
- Conclusion
- Bibliographie

Uniform Resource Identifiers (URIs)



Données liées : 1^{er} Principe
Use URIs (Uniform Resource Identifiers) to name things

Most apps use only a subset of the stack

Querying allows fine-grained data access

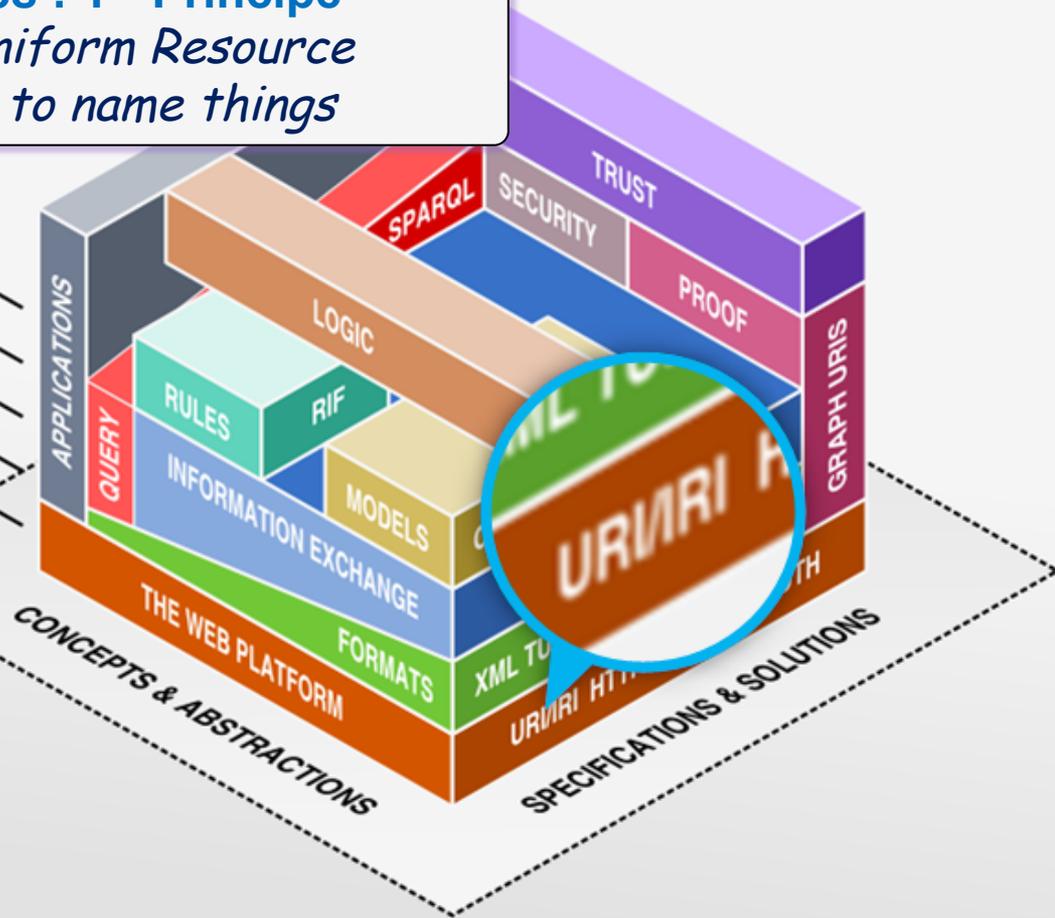
Standardized information exchange is key

Formats are necessary, but not too important

The Semantic Web is based on the Web

Linked Data uses a small selection of technologies

LINKED DATA



<http://www.bnode.org/blog/tag/layer%20cake> (Benjamin Nowack)

URI: définition

- "Un URI, de l'anglais Uniform Resource Identifier, soit littéralement identifiant uniforme de ressource, est une courte chaîne de caractères identifiant une ressource sur un réseau (par exemple une ressource Web) physique ou abstraite, et dont la syntaxe respecte une norme d'Internet mise en place pour le World Wide Web (voir RFC 3986)."

http://fr.wikipedia.org/wiki/Uniform_Resource_Identifier

Names in the Web...

10

**Uniform
Resource
Identifier**

- different types of resource identifiers all constructed according to a uniform schema
- whatever may be identified via URI
- to distinguish one resource from another

HPI Hasso Plattner Institut

Knowledge Engineering with Semantic Web Technologies, Dr. Harald Sack, Hasso-Plattner-Institut, Universität Potsdam

Knowledge Engineering with Semantic Web Technologies, Dr. Harald Sack, Hasso-Plattner-Institut, Universität Potsdam

URI: syntaxe

- URI : syntaxe générique

scheme ":" ["//" authority "/"] [path] ["?" query] ["#" fragment]

- **scheme:** http, ftp, mailto, ...
- **authority:** [userinfo@]host[:port]
 - **userinfo:** ex.: username:password
 - **host:** domaine-nom, adresse IP
 - **port:** numéro de port, ex: 80 le port HTTP standard
- **path:** séquence de segments (similaires à des répertoires), séparés par des /, ex. : chemin dans un le système de fichiers d'un serveur HTTP
- **query:** contient des informations complémentaires qui ne sont pas de nature hiérarchique (ex: suite de paires <cle>=<valeur> séparées par & pour requêtes HTTP)
- **fragment:** fournit des informations supplémentaires permettant d'accéder à une ressource secondaire (ex: id d'une ancre pour document HTML)



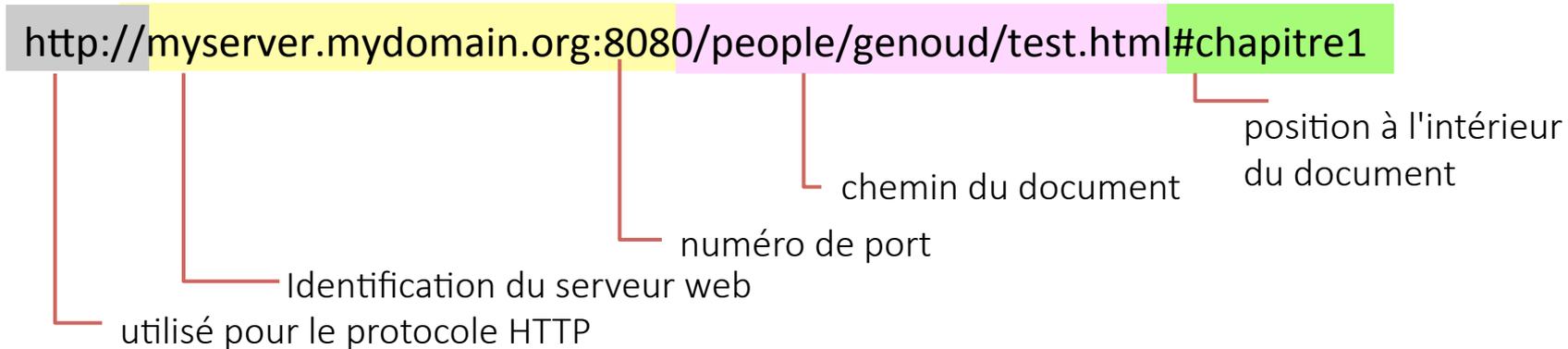
Internationalized Resource Identifier (RFC 3987): extension pour supporter encodage Unicode des caractères http://fa.dbpedia.org/resource/ب_رسنس

URI: exemples

- URI : syntaxe générique

scheme ":" ["//" authority "/"] [path] ["?" query] ["#" fragment]

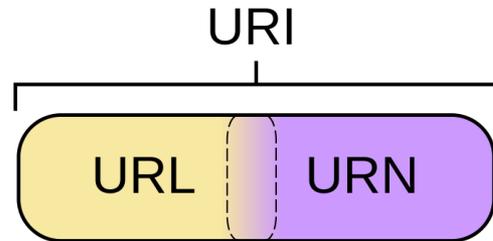
exemple :



autres exemples d'URI : `ftp://server.example.com/foo`
`mailto:person@example.fr`
`urn:isbn:978-0553283686`

URI - URL - URN

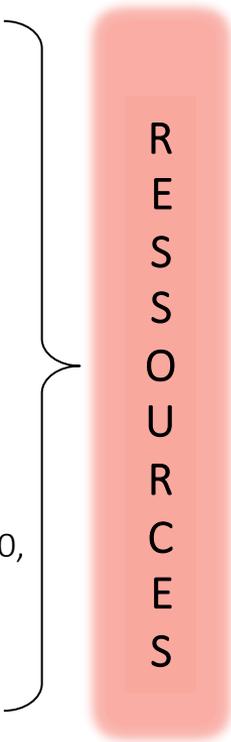
- une URI/IRI n'identifie pas nécessairement une ressource résolvable sur web



- Adresse (Locator)
 - Uniform Resource Locator (RFC 1738)
 - indique où une ressource peut être trouvée sur le web
 - *peut changer durant le cycle de vie de la ressource*
- Identité (Name)
 - Uniform Resource Name (RFC 2141)
 - identificateur persistant pour une ressource web
 - *demeure inchangé durant le cycle de vie de la ressource*

Ressources dans le web des données

- les données décrivent les éléments (things) d'un domaine d'intérêt au travers de leurs propriétés et de leurs relations.
- ces éléments peuvent être:
 - des documents web
 - pages, HTML, images, données accédées à travers un service web
 - mais aussi :
 - des entités du monde réel:
 - lieu : Sète
 - personne: Georges Brassens
 - ...
 - des concepts abstraits:
 - l'ensemble des auteurs/compositeurs français des années 60-80,
 - la relation du type *auteur de*
 - ...



R
E
S
S
O
U
R
C
E
S

Uniform Resource Identifiers (URIs)



représente

<http://dbpedia.org/resource/grenoble>

le nom (URI) qui dans DBpedia
représente Grenoble



GeoNames

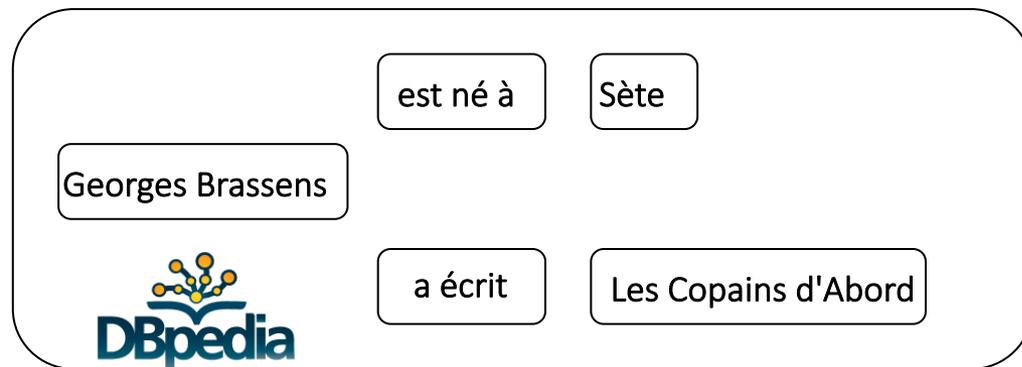
représente

<http://sws.geonames.org/3014728>

le nom (URI) qui dans GeoNames
représente Grenoble

Exemples de ressources : DBpedia

- données sur Georges Brassens disponible sur le Web

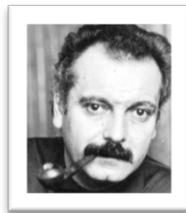


document contenant des données structurées sur Georges Brassens

une page web (HTML)



une photographie



une vidéo



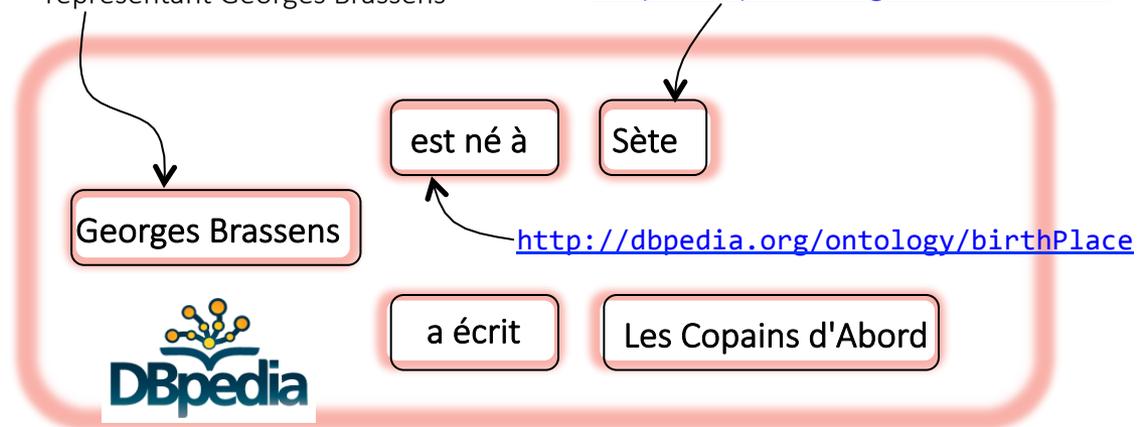
Exemples de ressources : DBpedia

- entités du monde réels et concepts abstraits décrits dans les données sont considérés comme des ressources → au même titre que les documents web ils sont identifiés par une URI

http://dbpedia.org/resource/Georges_Brassens

la ressource DBpedia représentant Georges Brassens

<http://dbpedia.org/resource/Sète>



document contenant des données structurées sur Georges Brassens
http://dbpedia.org/data/Georges_Brassens.xml

une page web (HTML)



http://fr.wikipedia.org/wiki/Georges_Brassens

une photographie



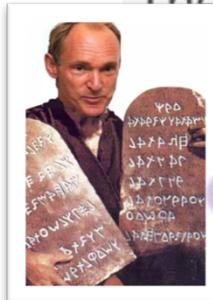
<http://culturetheque.org.uk/media/item/17545/800/brassens.jpg>

une vidéo



<https://www.youtube.com/watch?v=rs1ShTbqNbo>

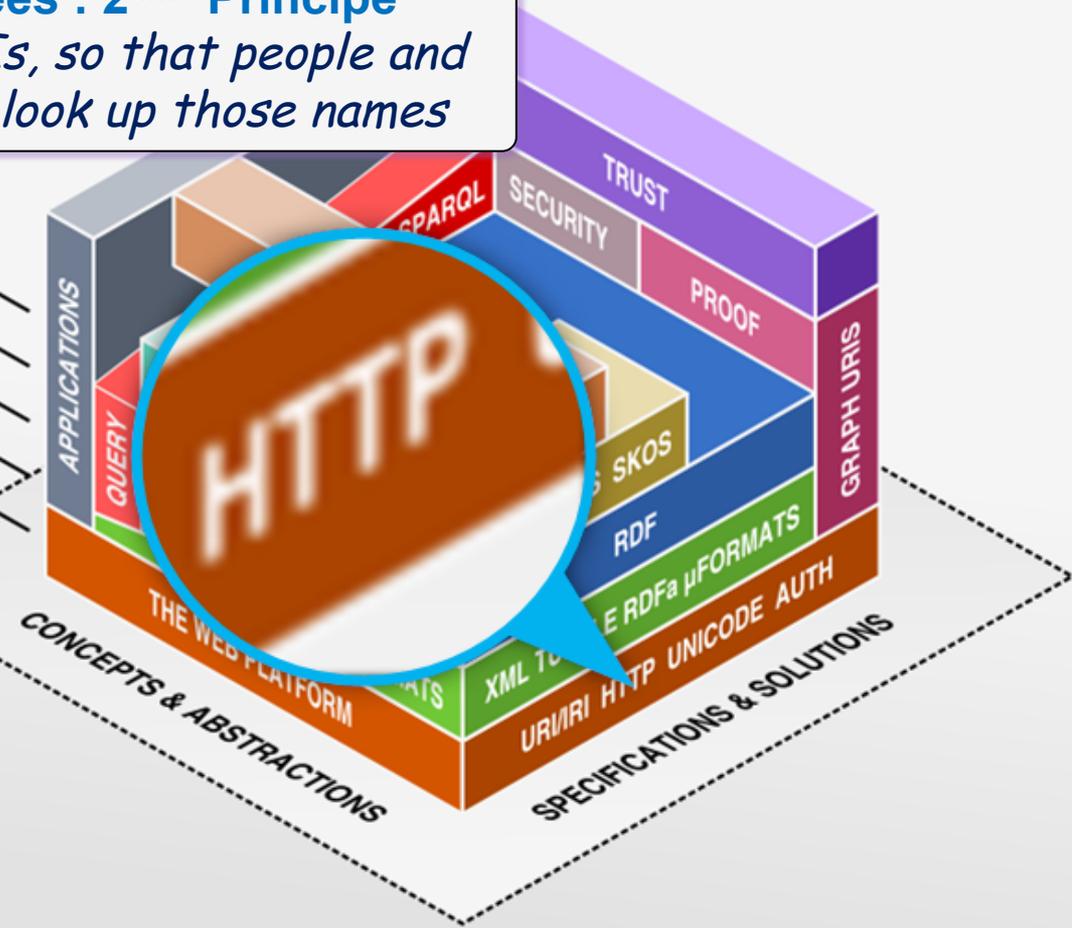
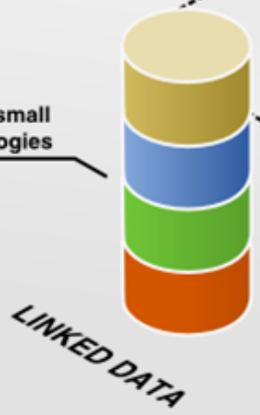
HTTP URIs



Données liées : 2^{ème} Principe
Use HTTP URIs, so that people and programs can look up those names

- Most apps use only a subset of the stack
- Querying allows fine-grained data access
- Standardized information exchange is key
- Formats are necessary, but not too important
- The Semantic Web is based on the Web

Linked Data uses a small selection of technologies



<http://www.bnode.org/blog/tag/layer%20cake> (Benjamin Nowack)

HTTP URI

- HTTP (Hyper Text Transfert Protocol) est le mécanisme universel d'accès du Web
 - les URI HTTP combinent un mécanisme d'identification global avec un mécanisme de récupération simple et bien compris : le protocole HTTP.
- Dans le contexte des données liées,
 - les URIs HTTP utilisées pour identifier des entités du monde réel ou des concepts abstraits peuvent être **déréférencées** via le protocole HTTP en une description de l'objet ou du concept identifié
 - *les descriptions destinées à être lues par des humains sont souvent représentées sous la forme de documents HTML.*
 - *les descriptions destinées à être consommées par des programmes sont représentées par des données RDF.*

Déréférencement d'URI

- Quand des URI identifient des entités du monde réel, il est essentiel de ne pas confondre ces entités elles-mêmes avec les documents web qui les décrivent.

c'est une URN mais pas une URL

URI de la ressource DBpedia représentant Georges Brassens

http://dbpedia.org/resource/Georges_Brassens

HTTP



http://dbpedia.org/page/Georges_Brassens

http://dbpedia.org/data/Georges_Brassens.xml

Property	Value
dbpedia-owl:abstract	<ul style="list-style-type: none">Georges Brassens (French: [ʒoʁʒ bʁasɛ̃s]; 22 October 1921 – 29 Oct. poet.Georges Brassens, né à Sète le 22 octobre 1921 et mort à Saint-Gély-du-Fesc le 29 octobre 1981, est un poète auteur-compositeur-interprète français. Il met en musique et interprète, en s'accompagnant à la guitare, plus d'une centaine de ses poèmes. Outre ses propres textes, il met également en musique des poèmes de François Villon, Paul Verlaine, Paul Fort ou encore Louis Aragon. Il reçoit le Grand prix de poésie de l'Académie française en 1967. Il enregistre quatorze albums entre 1952 et 1976.
dbpedia-owl:activeYearsEndYear	1981-01-01 (xsd:date)
dbpedia-owl:activeYearsStartYear	1951-01-01 (xsd:date)
dbpedia-owl:associatedBand	dbpedia:Pierre_Nicolas
dbpedia-owl:associatedMusicalArtist	dbpedia:Pierre_Nicolas
dbpedia-owl:background	solo_singer
dbpedia-owl:birthDate	1921-10-22 (xsd:date)
dbpedia-owl:birthPlace	dbpedia:Sète

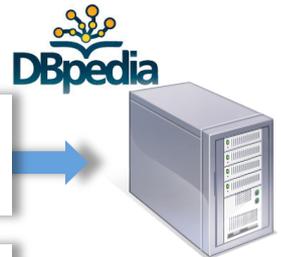
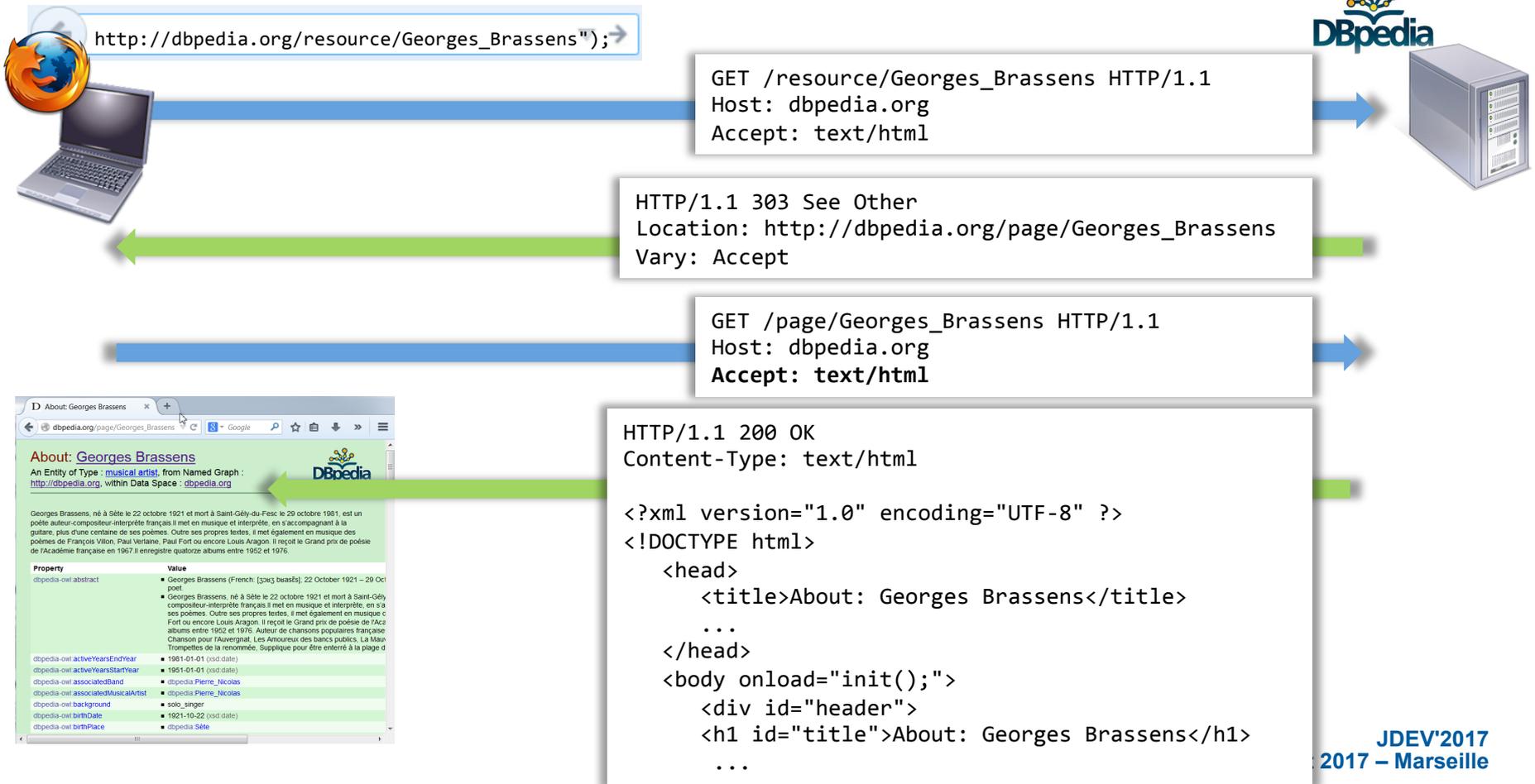
```
<rdf:RDF>
- <rdf:Description rdf:about="http://dbpedia.org/resource/6587_Brassens">
  <dbpprop:namedAfter rdf:resource="http://dbpedia.org/resource/Georges_Brassens"/>
</rdf:Description>
- <rdf:Description rdf:about="http://dbpedia.org/resource/Volume_3_(Fabrizio_De_Andr%C3%A9_album)">
  <dbpprop:music rdf:resource="http://dbpedia.org/resource/Georges_Brassens"/>
</rdf:Description>
- <rdf:Description rdf:about="http://dbpedia.org/resource/Le_D%C3%A9ner_de_Cons_(film)">
  <dbpprop:extra rdf:resource="http://dbpedia.org/resource/Georges_Brassens"/>
</rdf:Description>
- <rdf:Description rdf:about="http://dbpedia.org/resource/Georges_Brassens">
  <rdf:type rdf:resource="http://dbpedia.org/class/yago/PeopleFromS%C3%A8te"/>
  <rdf:type rdf:resource="http://dbpedia.org/class/yago/Creator109614315"/>
  <rdf:type rdf:resource="http://dbpedia.org/class/yago/Artist109812338"/>
  <rdf:type rdf:resource="http://dbpedia.org/class/yago/Musician110339966"/>
  <rdf:type rdf:resource="http://dbpedia.org/class/yago/FrenchPeopleOfItalianDescent"/>
  <rdf:type rdf:resource="http://schema.org/MusicGroup"/>
  <rdf:type rdf:resource="http://dbpedia.org/class/yago/YagoLegalActor"/>
  <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person"/>
  <rdf:type rdf:resource="http://dbpedia.org/ontology/Person"/>
  <rdf:type rdf:resource="http://dbpedia.org/class/yago/CausalAgent100007347"/>

```

document
RDF/XML

Négociation de contenu (303 URI)

- Utilisation du mécanisme de négociation de contenu du protocole HTTP pour fournir à un client HTTP une représentation associée à une URI (URN) identifiant un objet du monde réel ou un concept.



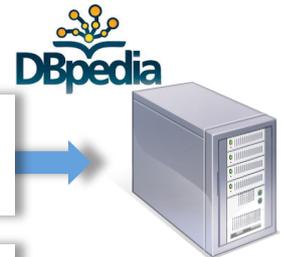
Négociation de contenu (303 URI)

- Utilisation du mécanisme de négociation de contenu du protocole HTTP pour fournir à un client HTTP une représentation associée à une URI (URN) identifiant un objet du monde réel ou un concept.

```
final String url = "http://dbpedia.org/resource/Georges_Brassens";  
final Model model = ModelFactory.createDefaultModel();  
model.read(url);
```



```
GET /resource/Georges_Brassens HTTP/1.1  
Host: dbpedia.org  
Accept: application/rdf+xml
```



```
HTTP/1.1 303 See Other  
Location: http://dbpedia.org/data/Georges_Brassens.xml  
Vary: Accept
```

```
GET /data/Georges_Brassens.xml HTTP/1.1  
Host: dbpedia.org  
Accept: text/html
```

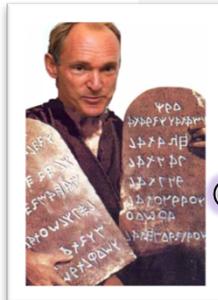
```
HTTP/1.1 200 OK  
Content-Type: application/rdf+xml
```

```
<?xml version="1.0" encoding="utf-8" ?>  
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"...>  
  <rdf:Description rdf:about="http://dbpedia.org/resource/Georges_Brassens">  
    <rdf:type rdf:resource="http://dbpedia.org/class/yago/PeopleFromS%C3%A8te" />  
    <rdf:type  
      rdf:resource="http://dbpedia.org/class/yago/FrenchPeopleOfItalianDescent" />  
    ...
```

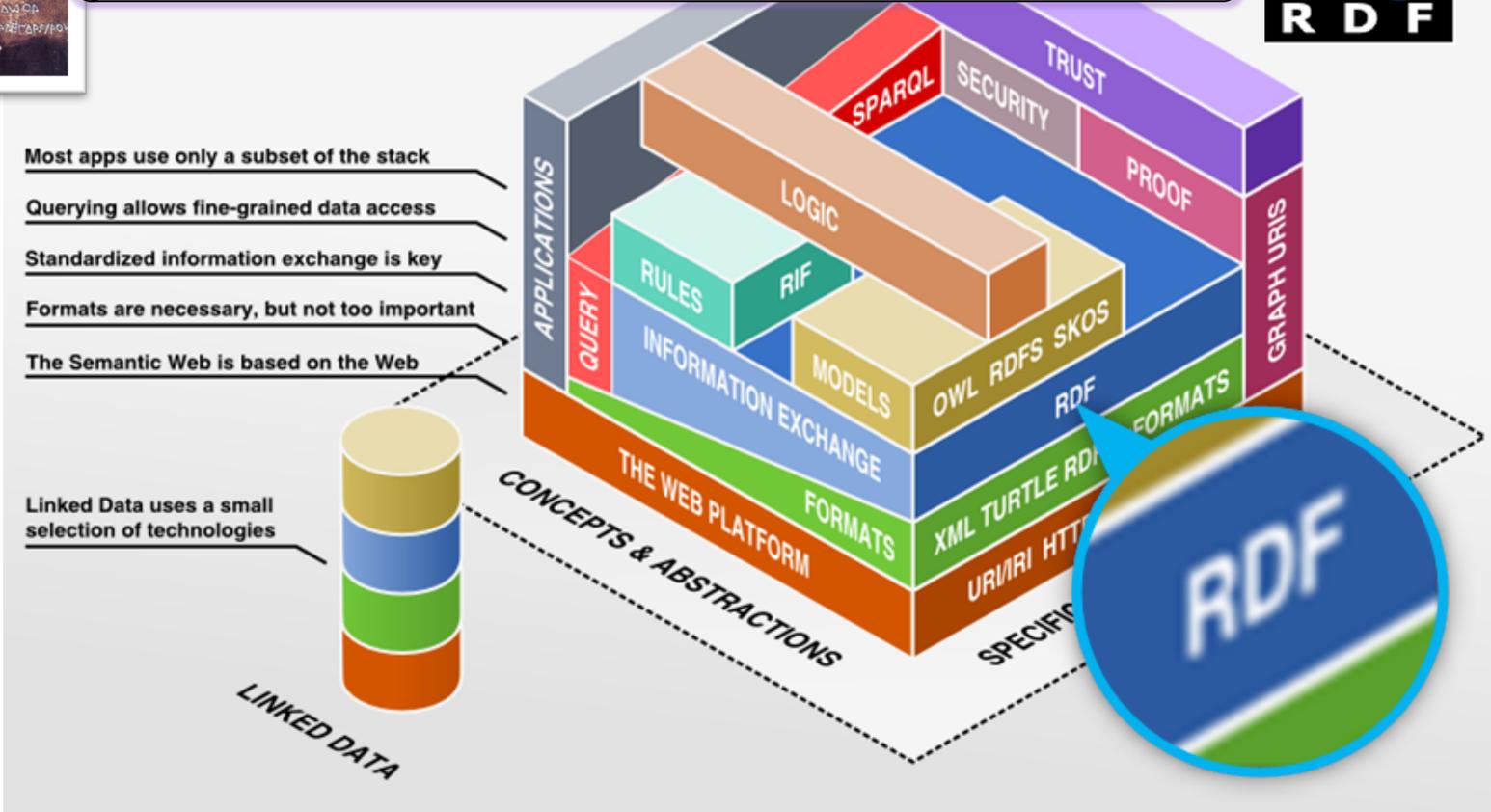
Plan

- Introduction
- De l'Open Data au Linked Open Data (1^{ère} partie)
- Les principes des données ouvertes et liées
 - Nommer les données: URI
 - **Modéliser les données: RDF**
 - Interroger les données: SPARQL
- De l'Open Data au Linked Open Data (2^{ème} Partie)
- Conclusion
- Bibliographie

Resource Description Framework (RDF)



Données liées : 3^{ème} Principe
When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL).



<http://www.bnode.org/blog/tag/layer%20cake> (Benjamin Nowack)

Resource Description Framework

- Pour permettre à une grande variété d'applications différentes de manipuler et traiter les données issues du web il est important de s'accorder sur des formats **standards** de contenu .
 - l'accord sur HTML comme format de document dominant a été un facteur important qui a permis le passage à l'échelle du Web.
- Dans le contexte des données liées le modèle pour publier des données structurées sur le web est défini par **RDF (Resource Description Framework)**
- RDF fait partie de l'activité Web Sémantique du W3C
 - recommandation du W3C en février 2004 (RDF 1.0)
 - mise à jour en février 2014 (RDF 1.1)

<http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/> [1]

<http://www.w3.org/TR/2014/NOTE-rdf11-primer-20140624/> [2]

"The **Resource Description Framework (RDF)** is a framework for representing information in the **Web**." [1]

Modèle de données RDF

- En RDF, la connaissance est représentée par une liste d'assertions (statements)
- Toutes les assertions RDF suivent un même et simple schéma composé de 3 parties:
 - la chose (l'entité) que l'assertion décrit
 - la propriété de la chose sur laquelle porte l'assertion
 - la valeur de cette propriété

l'entité décrite	propriété	valeur
Georges Brassens	est né à	Sète

Modèle de données RDF

- Une assertion RDF est définie comme un triplet (*triple*)

Sujet

URI

Prédicat

URI

Objet

URI/Littéral

- sujet et prédicat sont des **ressources** : ils sont identifiés de manière unique par URI
- l'objet peut être une **ressource** (URI) ou une **valeur littérale** (constante n'ayant pas d'autres attributs la décrivant)

Georges Brassens **est né à** **Sète**

sujet

prédicat

valeur

http://dbpedia.org/resource/Georges_Brassens

<http://dbpedia.org/ontology/birthPlace>

<http://dbpedia.org/resource/Sète>

Georges Brassens **est né le** **22 octobre 1921**

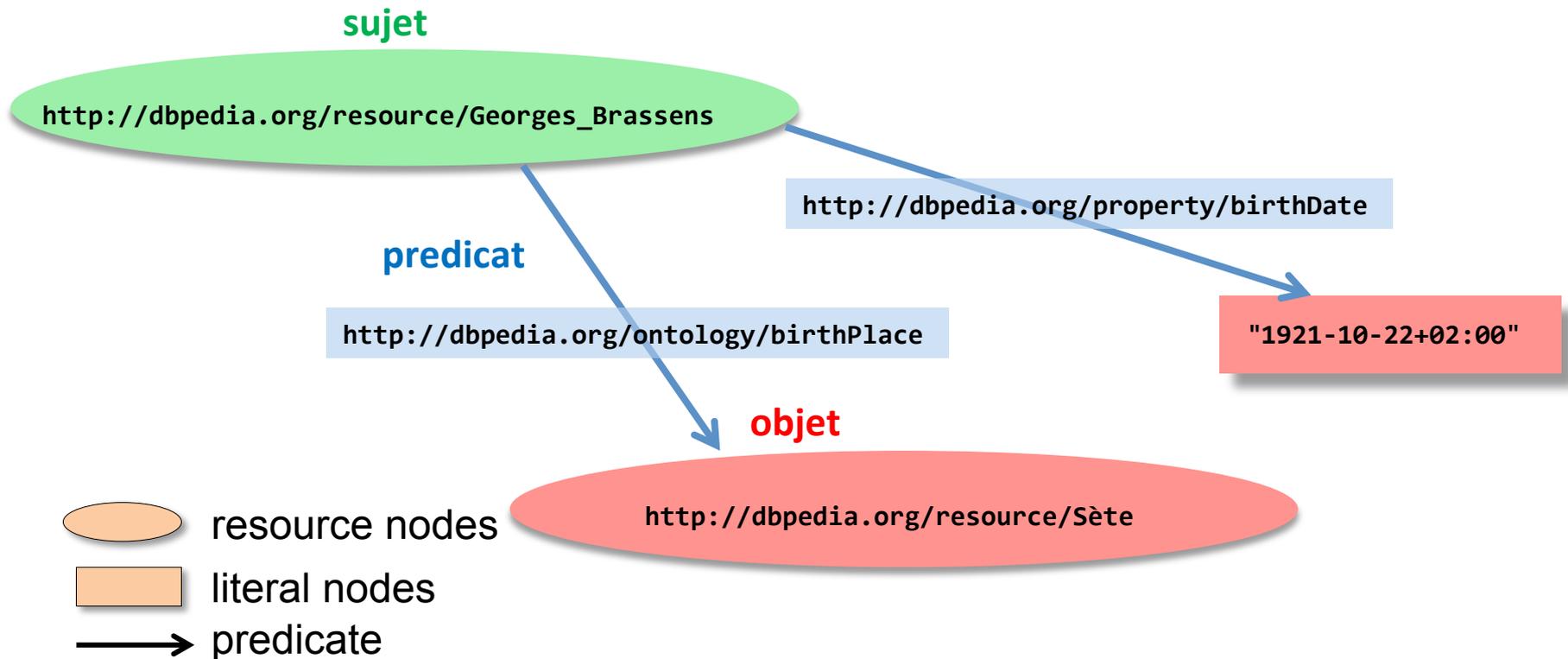
http://dbpedia.org/resource/Georges_Brassens

<http://dbpedia.org/property/birthDate>

"1921-10-22+02:00"

Modèle de données RDF

- les données RDF peuvent être vues comme un graphe étiqueté orienté.
 - sujets et objets sont les noeuds (sommets)
 - prédicats sont les arrêtes orientées (arcs)



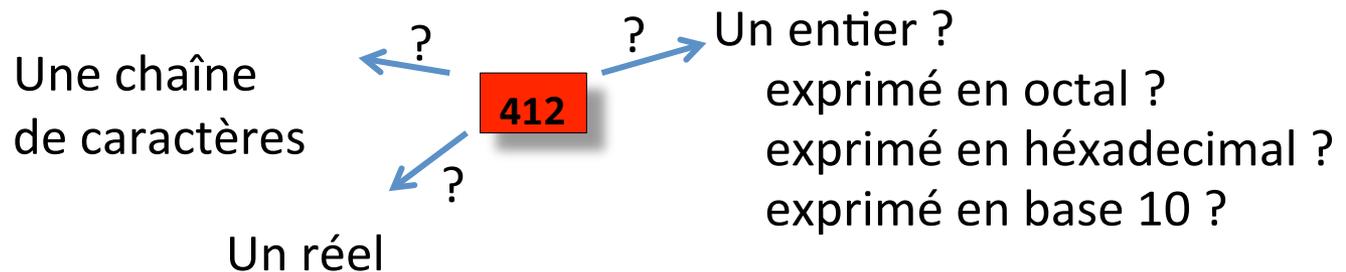
Littéraux Typés

- Les littéraux ne sont pas des ressources: ce sont des valeurs



En regardant cette description un humain peut facilement deviner qu'il s'agit d'un entier

Mais qu'en est-il d'un programme ?



Nécessité de fournir un certain **contexte** si vous avez l'intention d'utiliser la valeur d'une autre manière qu'une simple visualisation sur une page web → **littéraux typés**

Littéraux Typés : XSD

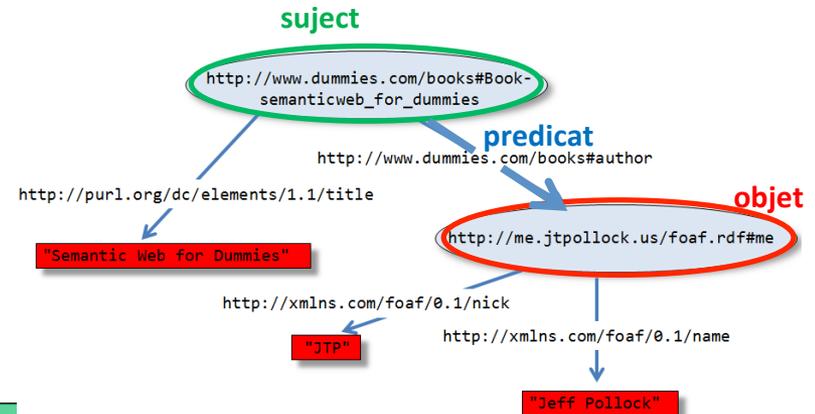
- utilisation de XML Schéma (XSD) pour associer un type aux littéraux.
 - W3C recommandations :
 - <http://www.w3.org/TR/xmlschema-2/>
 - <http://www.w3.org/TR/rdf-mt/>
 - XML schéma définit une hiérarchie prédéfinie de type de données
 - types primitif (string, float, decimal, etc.)
 - types dérivés (integer, long, etc.)
 - de nouveaux types peuvent être définis par dérivation
 - restriction
 - listes
 - union
 - extension

Sérialisations de RDF

- Graphe RDF : modèle abstrait bien adapté à une interprétation humaine
- sérialisation RDF
 - fournir un moyen de conversion entre le modèle abstrait et un format concret (format de fichier ou flux d'octets (bytestream))
 - plusieurs formats de sérialisation d'expressivité égale
 - XML/RDF
 - N-Triples
 - Turtle (Terse RDF Triple Language)
 - N3 (Notation3)
 - RDF/JSON
 - RDFa

Sérialisation de RDF - N-Triples

- notation la plus simple
 - chaque ligne représente un triplet et est terminé par '.'
 - ressources (sujet, prédicat, ressource objet) sont représentées par leur URI complète délimitée par '<' et '>'
 - valeurs littérales entre " "



example.nt

subject

```
<http://www.dummies.com/books#Book-semanticweb_for_dummies>
<http://www.dummies.com/books#author> <http://me.jtpollock.us/foaf.rdf#me>.
<http://www.dummies.com/books#Book-semanticweb_for_dummies>
  <http://purl.org/dc/elements/1.1/title> "Semantic Web for Dummies".
<http://me.jtpollock.us/foaf.rdf#me> <http://xmlns.com/foaf/0.1/name> "Jeff Pollock".
<http://me.jtpollock.us/foaf.rdf#me> <http://xmlns.com/foaf/0.1/nick> "JTP".
```

- + très simple à utiliser
- verbeux (information redondante pénalisant aux niveau des temps de transmission et de parsing)

Sérialisation de RDF - Turtle [1]

- Turtle (Terse RDF Triple Language)
 - condense la plupart des répétitions de N-Triples
 - préfixes déclarés au début du document permettent d'abrégéer les URIs

```
<http://www.dummies.com/books#Book-semanticweb_for_dummies>  
<http://www.dummies.com/books#author> <http://me.jtpollock.us/foaf.rdf#me>.
```



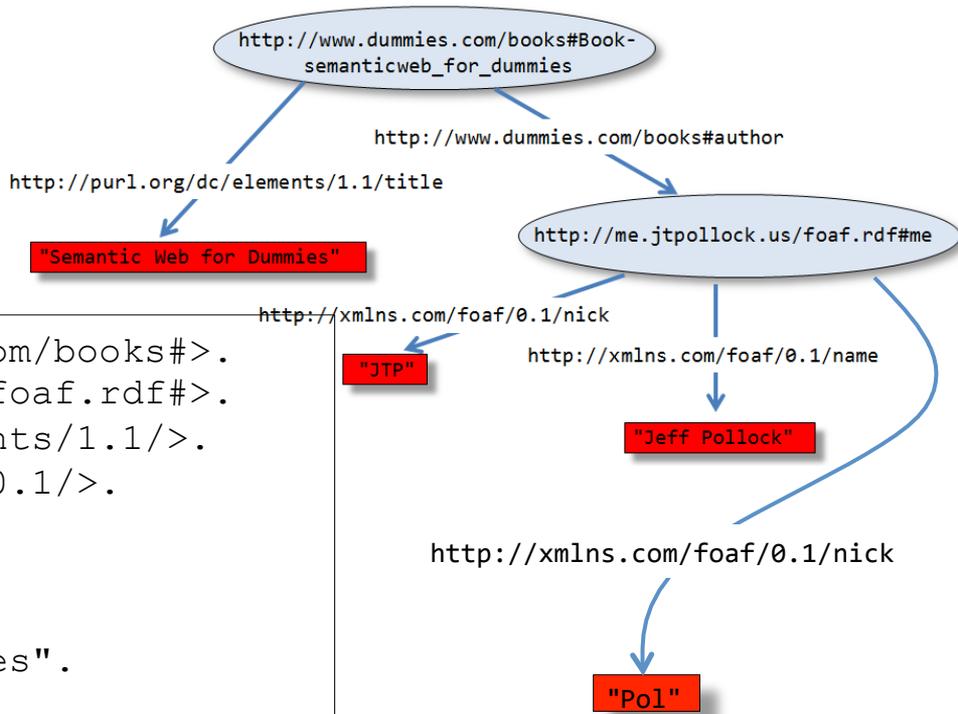
```
@prefix swbook: <http://www.dummies.com/books#>.
```

```
@prefix jtp: <http://me.jtpollock.us/foaf.rdf#>.
```

```
swbook:Book-semanticweb_for_dummies swbook:author jtp:me.
```

Sérialisation de RDF - Turtle

- ';' permet de combiner plusieurs assertions concernant le même sujet
- ',' permet de combiner plusieurs assertions concernant le même sujet et le même prédicat



example.ttl

```
@prefix swbook: <http://www.dummies.com/books#>.
@prefix jtp: <http://me.jtpollock.us/foaf.rdf#>.
@prefix dc: <http://purl.org/dc/elements/1.1/>.
@prefix foaf: <http://xmlns.com/foaf/0.1/>.

swbook:Book-semanticweb_for_dummies
  swbook:author jtp:me;
  dc:title "Semantic Web for Dummies".

jtp:me
  foaf:name "Jeff Pollock";
  foaf:nick "JTP", "Po1" .
```

Sérialisation de RDF : RDF/XML

- RDF/XML , 1^{ère} syntaxe standardisée par W3C (204)
- Assertions sur une ressource sont regroupée dans un élément `rdf:Description`
- forme générale

gives the subject of all statement within the description

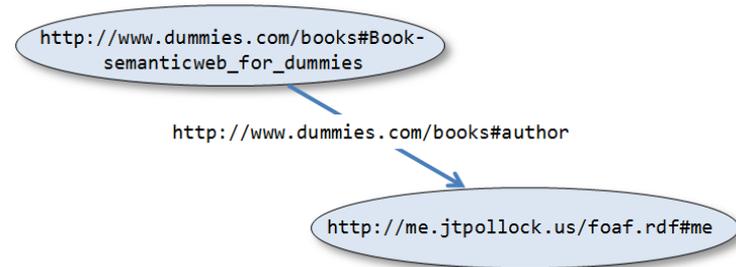
```
<rdf:Description rdf:about="subject">  
  <predicate rdf:resource="object" />  
  <predicate>literal value</predicate>  
</rdf:Description>
```

name of the internal tags represent a predicate

the object is represented differently depending on whether it is a resource or a literal

Sérialisation de RDF : RDF/XML

- exemple



déclaration XML (stipule qu'il s'agit d'un document XML)

```
<?xml version="1.0"?>
```

```
<rdf:RDF
```

```
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:books="http://www.dummies.com/books#">
```

```
  <rdf:Description
```

```
    rdf:about="http://www.dummies.com/books#Book-semanticweb_for_dummies">
```

```
    <books:author
```

```
      rdf:resource="http://me.jtpollock.us/foaf.rdf#me" />
```

```
  </rdf:Description>
```

```
</rdf:RDF>
```

espaces de
nommage XML

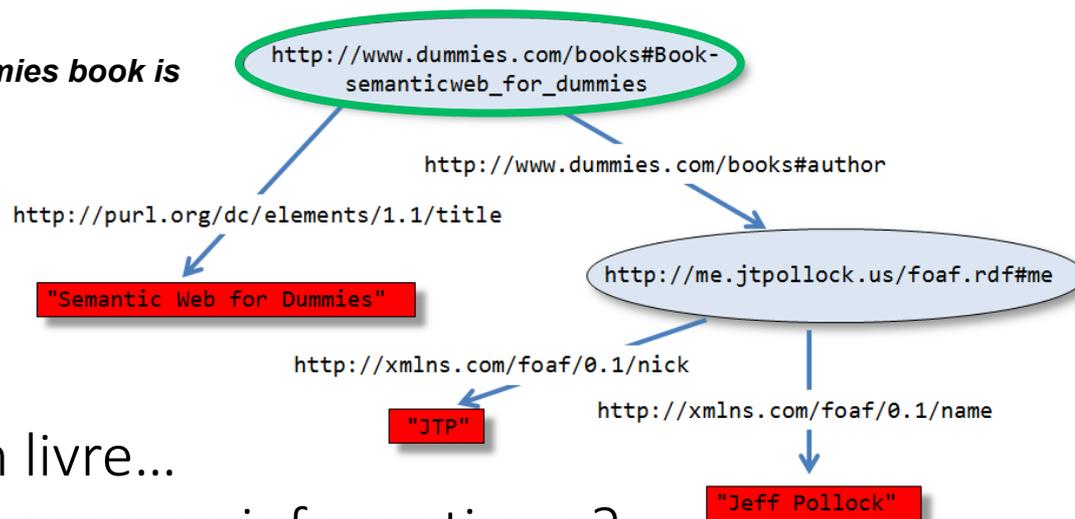
balise
racine

rdf:DescriptionElement
pour les assertions à propos d'une
ressource

Typing resources: `rdf:type`

- possibility of associating a type to a resource

The Semantic Web For Dummies book is authored by Jeff Pollock

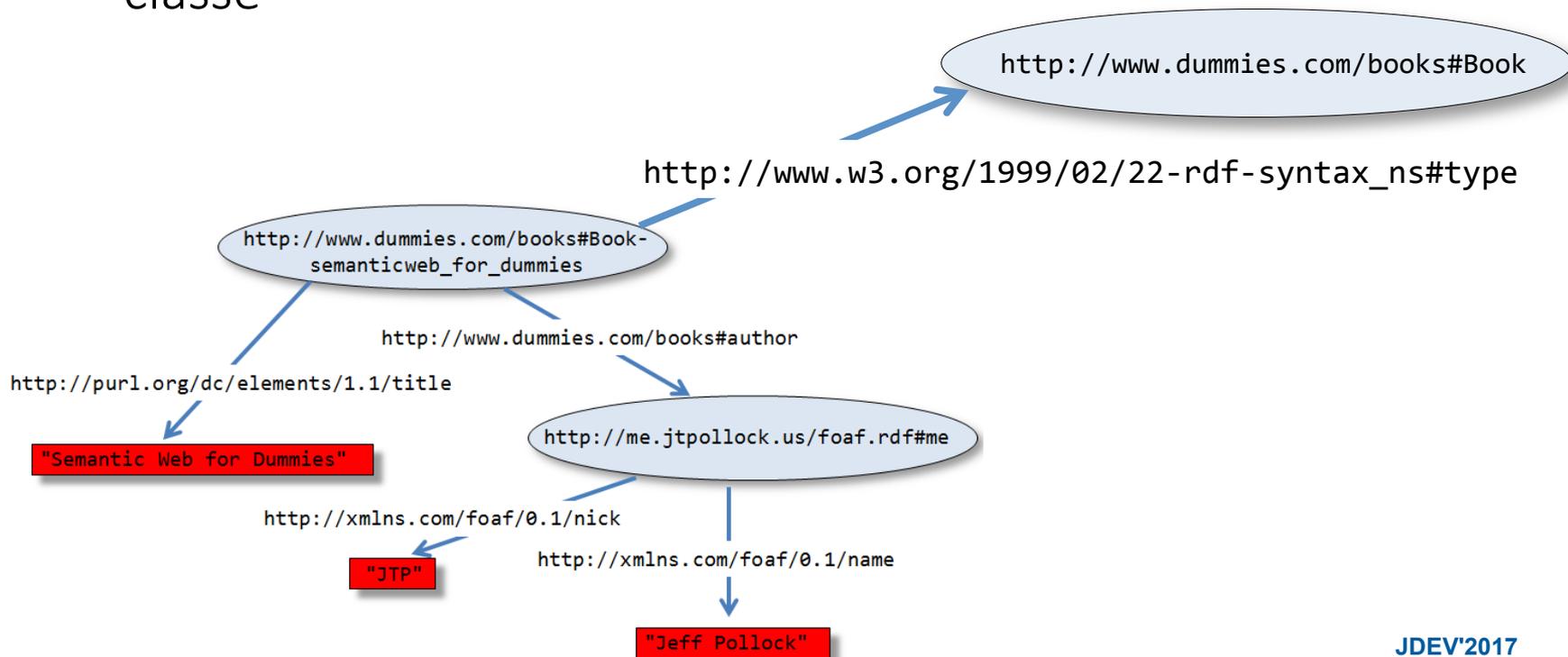


Nous savons que c'est un livre...
mais qu'en est-il d'un programme informatique ?

Pour résoudre le problème de classer les ressources d'une manière compréhensible par un logiciel, le vocabulaire RDF dispose d'un prédicat prédéfini : **`rdf:type`**

Sémantique de `rdf:type`

- sémantique **`rdf:type`**
 - la valeur de ce prédicat est une ressource et représente une classe d'objets
 - le sujet de ce prédicat est aussi une instance de cette classe



Identifying the type of a resource

- N3 - Turtle

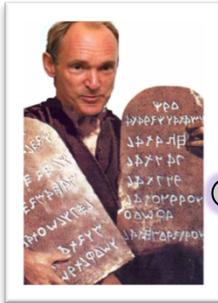
```
@prefix swbook: <http://www.dummies.com/books#>.
```

```
swbook:Book-semanticweb_for_dummies  
  swbook:author <http://me.jtpollock.us/foaf.rdf#me>;  
  a swbook:Book.
```

- XML/RDF

```
<?xml version="1.0"?>  
<rdf:RDF  
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"  
  xmlns:books="http://www.dummies.com/books#">  
  <rdf:Description  
    rdf:about="http://www.dummies.com/books#Book-semanticweb_for_dummies">  
    <rdf:type rdf:resource="http://www.dummies.com/books#Book"/>  
    <books:author  
      rdf:resource="http://me.jtpollock.us/foaf.rdf#me" />  
  </rdf:Description>  
</rdf:RDF>
```

RDF et intégration des données



Données liées : 4^{ème} Principe
Include links to other URIs, so that they can discover more things.

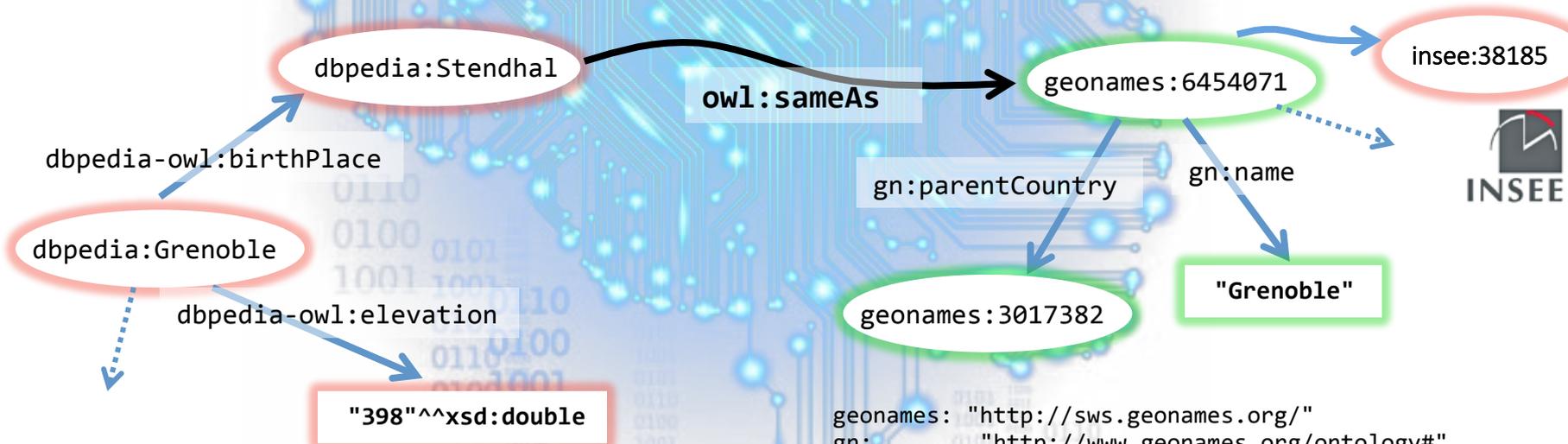
L'utilisateur peut maintenant faire des requêtes sur des informations qui n'étaient pas présentes dans son jeu de données initial.



GeoNames

<http://dbpedia.org/resource/grenoble>

<http://sws.geonames.org/6454071>



Préfixes

dbpedia: "http://dbpedia.org/resource/"
 dbpedia-owl: "http://dbpedia.org/ontology/"

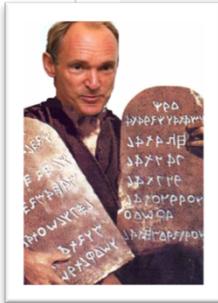
geonames: "http://sws.geonames.org/"
 gn: "http://www.geonames.org/ontology#"
 insee: "http://id.insee.fr/geo/commune/"



Plan

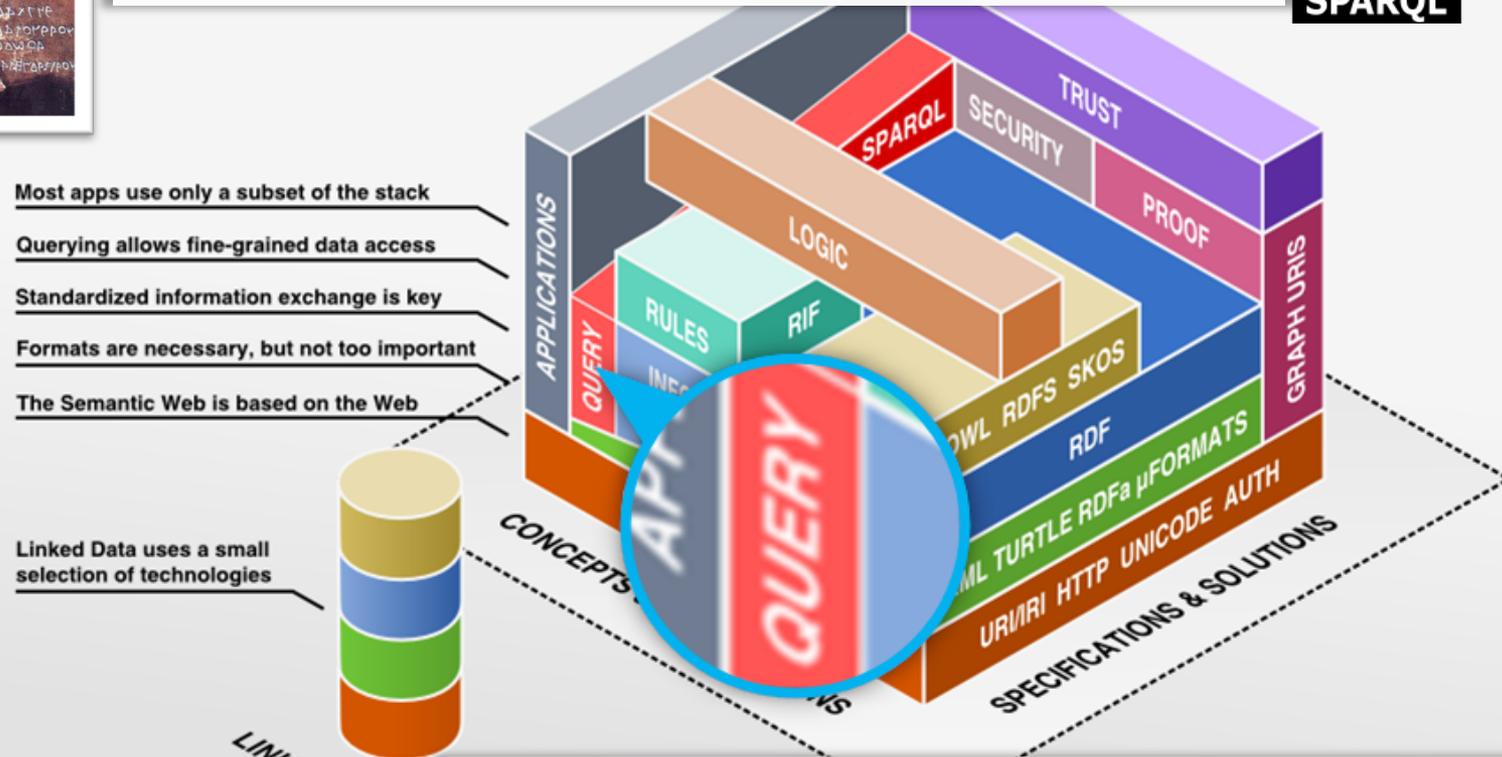
- Introduction
- De l'Open Data au Linked Open Data (1^{ère} partie)
- Les principes des données ouvertes et liées
 - Nommer les données: URI
 - Modéliser les données: RDF
 - Interroger les données: SPARQL
- De l'Open Data au Linked Open Data (2^{ème} Partie)
- Conclusion
- Bibliographie

Interroger les données : SPARQL



Données liées : 3^{ème} Principe

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

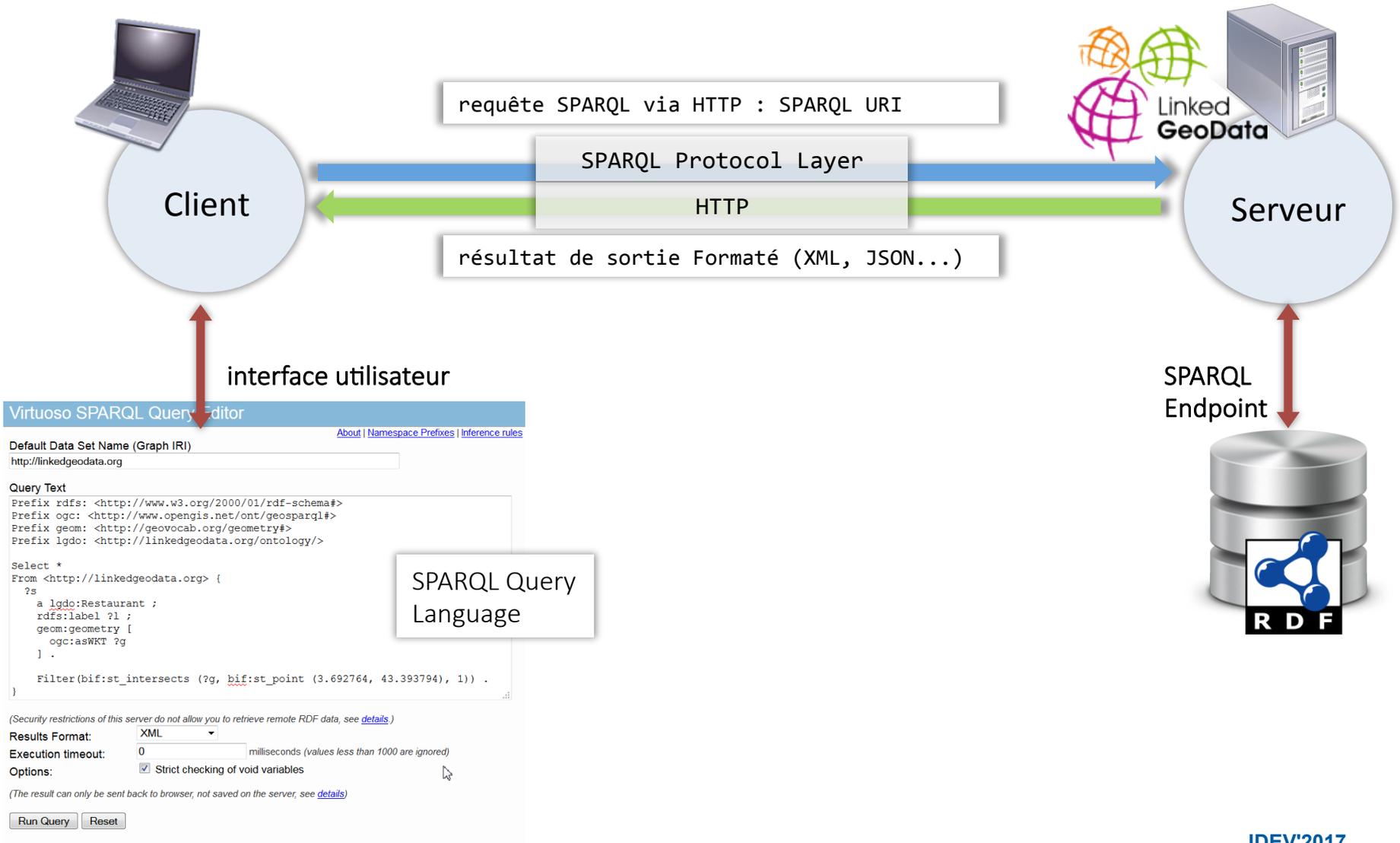


SPARQL Protocol And RDF Query Language

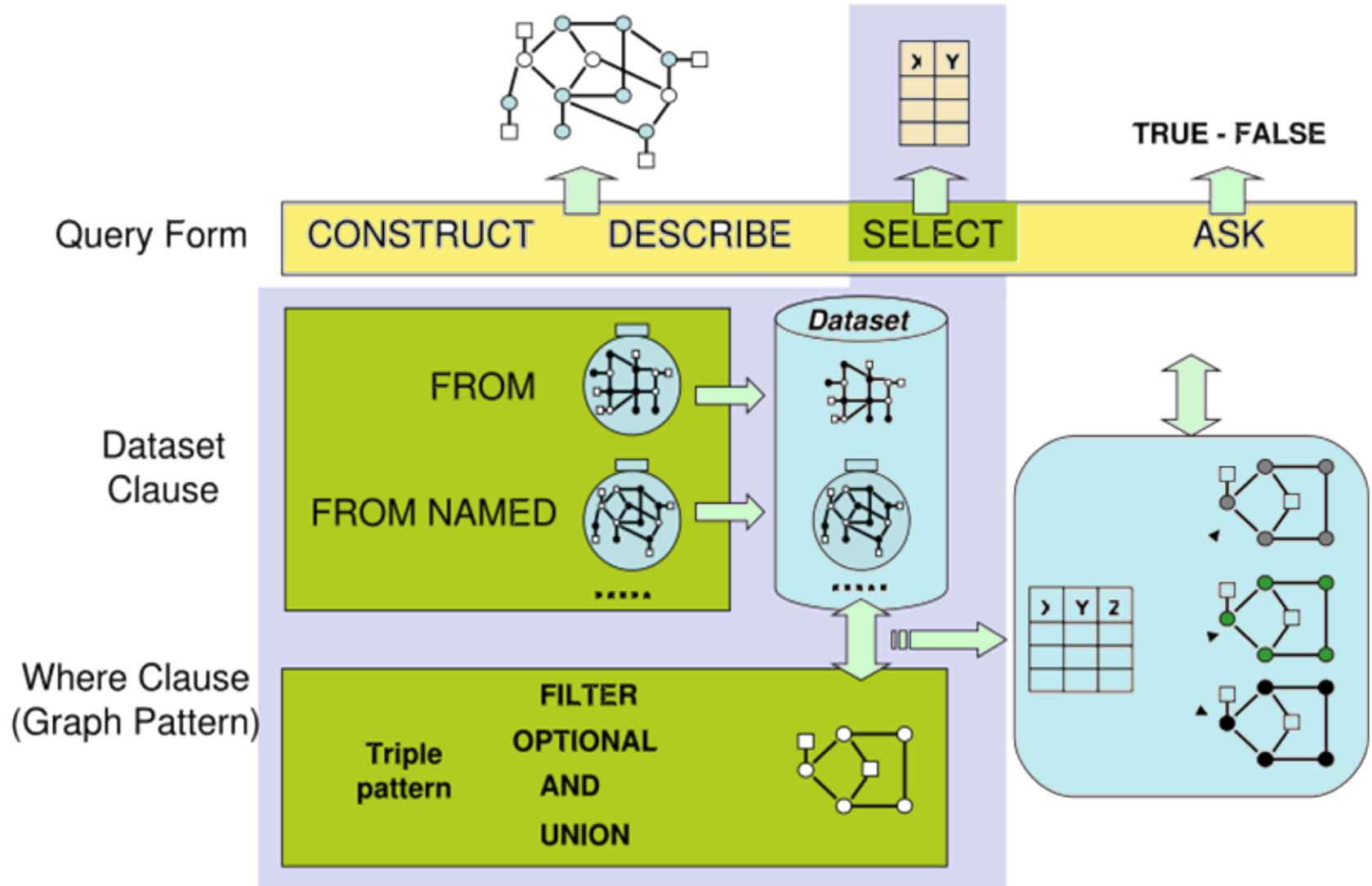
SPARQL introduction

- SPARQL : SPARQL Protocol And RDF Query Language
 - un langage de requêtes pour interroger un graphe RDF (SPARQL Query Language Specification)
 - un protocole pour soumettre des requêtes SPARQL via HTTP (SPARQL protocol for RDF Specification)
 - un format de sortie XML (et maintenant aussi JSON) pour les résultats de requêtes SPARQL (SPARQL Query XML Results Format)
 - un standard du W3C
 - Recommandations du W3C ,
 - SPARQL 1.0 Janvier 2008,
 - SPARQL 1.1 Mars 2013
 - inspiré par SQL

SPARQL vue générale



SPARQL Query Language (1.0)



Source: Pérez, Arenas and Gutierrez, Chapter 1: On the Semantics of SPARQL, Semantic Web Information Management: A Model Based Perspective, Springer 2010

SPARQL Query Language

- retrouver tous les restaurants à moins d'un kilomètre de Fort Saint-Pierre (Théâtre de la Mer – Sète)

```
Prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
Prefix ogc:  <http://www.opengis.net/ont/geosparql#>
Prefix geom: <http://geovocab.org/geometry#>
Prefix lgdo: <http://linkedgedata.org/ontology/>
```

```
Select ?s, ?l From <http://linkedgedata.org>
```

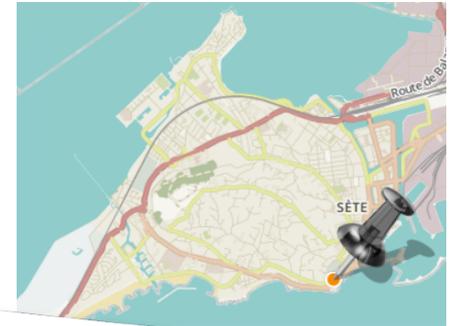
```
Where
```

```
{
```

```
  ?s    a lgdo:Restaurant ;
        rdfs:label ?l ;
        geom:geometry [
          ogc:asWKT ?g
        ] .
```

```
  Filter(bif:st_intersects (?g, bif:st_point (3.692764, 43.393794), 1)) .
```

```
}
```



SPARQL Query Language

- SPARQL basé sur:
 - la sérialisation de RDF en Turtle
 - la mise en correspondance de motifs de graphe (Graph Pattern Matching)

Prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>

Prefix ogc: <http://www.opengis.net/ont/geosparql#>

Prefix geom: <http://geovocab.org/geometry#>

Prefix lgdo: <http://linkedgeodata.org/ontology/>

Select ?s, ?g From <http://linkedgeodata.org>

Where

{

?s

variable

a lgdo:Restaurant ;

rdfs:label ?l ;

geom:geometry [

ogc:asWKT ?g

] .

Graph Pattern: triplet RDF contenant une ou plusieurs variables à n'importe quelle position (sujet, prédicat, objet)

Filter(bif:st_intersects (?g, bif:st_point (3.692764, 43.393794), 1)) .

}

SPARQL Query Language

- Les motifs de graphe (graph patterns) peuvent être combinés pour construire des requêtes (conjonctives) complexes.

Prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>

Prefix ogc: <http://www.opengis.net/ont/geosparql#>

Prefix geom: <http://geovocab.org/geometry#>

Prefix lgdo: <http://linkedgedata.org/ontology/>

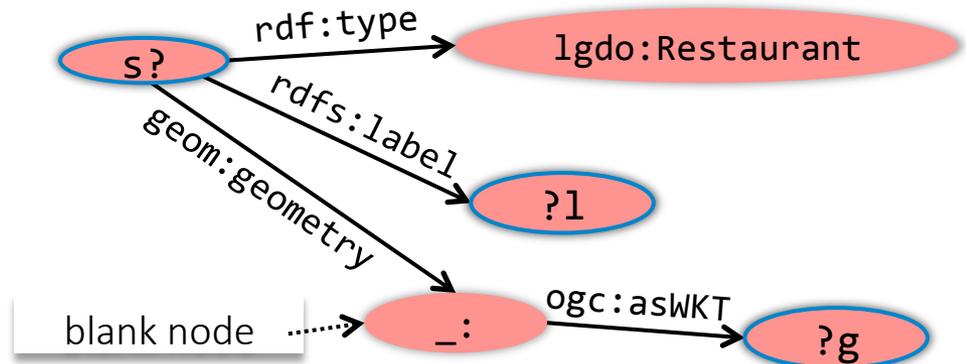
Select ?s, ?l From <http://linkedgedata.org>

Where

{

?s a lgdo:Restaurant ;
rdfs:label **?l** ;
geom:geometry [ogc:asWKT **?g**] .

variable



Filter(bif:st_intersects (?g, bif:st_point (3.692764, 43.393794), 1)) .

}

SPARQL Query Language

- Possibilité de filtrer les résultats
 - Filtre permet de contraindre les valeurs dans la solution

Prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>

Prefix ogc: <http://www.opengis.net/ont/geosparql#>

Prefix geom: <http://geovocab.org/geometry#>

Prefix lgdo: <http://linkedgedata.org/ontology/>

Select ?s, ?l From <http://linkedgedata.org>

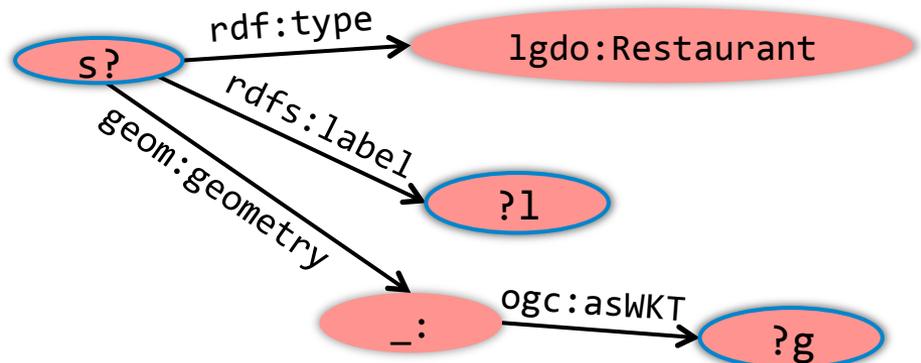
Where

{

```
?s a lgdo:Restaurant ;  
  rdfs:label ?l ;  
  geom:geometry [  
    ogc:asWKT ?g  
  ] .
```

```
Filter(bif:st_intersects (?g, bif:st_point (3.692764, 43.393794), 1)) .  
}
```

Graph Pattern



filtre sur la variable g

SPARQL Query Language : Filtres

- Filtres : permettre de restreindre les valeurs dans une solution
 - expressions booléennes que doivent satisfaire les solutions de la requête.
 - langage d'expressions riche basé sur Xpath, XQuery et des opérateurs spéciaux définis par SPARQL.

(voir section 11 du document de spécifications SPARQL <http://www.w3.org/TR/rdf-sparql-query/#tests>)

- Opérateurs relationnels : **<**, **>**, **=**, **<=**, **>=**, **!=**
- Opérateurs booléens : **&&**, **||**, **!**
- Opérateurs arithmétiques : **+** ***** **-** **/**
- Testes de liaison des variables: **isURI(?x)**, **isBlank(?x)**, **isLiteral(?x)**, **bound(?x)**
- Expressions régulières: **regex(?x, "A.*")**
- Accès aux attributs/valeur **lang()**, **datatype()**, **str()**
- Casting (fonctions de transtypage) **xsd:integer(?x)**
- Fonctions externes / extensions

exemple de requête Select (Linked Geodata)

- la clause SELECT indique les variables à prendre en compte dans le résultat
- * toutes les variables (comme en SQL)

Prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>

Prefix ogc: <http://www.opengis.net/ont/geosparql#>

Prefix geom: <http://geovocab.org/geometry#>

Prefix lgdo: <http://linkedgeodata.org/ontology/>

```
Select ?s, ?l From <http://linkedgeodata.org>
```

```
Where
```

```
{
```

```
  ?s    a lgdo:Restaurant ;
```

```
  rdfs:label ?l ;
```

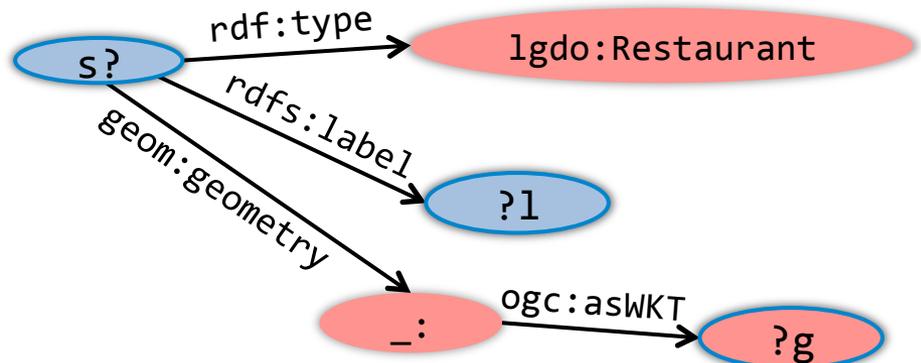
```
  geom:geometry [
```

```
    ogc:asWKT ?g
```

```
  ] .
```

```
  Filter(bif:st_intersects (?g, bif:st_point (3.692764, 43.393794), 1)) .
```

```
}
```



SPARQL Query Language: DataSet

- le service (RDF datastore) pouvant gérer un ou plusieurs graphes RDF, la requête SPARQL est exécutée sur un jeu de données (RDF Dataset) qui représente une collection de un ou plusieurs graphes.

Prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>

Prefix ogc: <http://www.opengis.net/ont/geosparql#>

Prefix geom: <http://geovocab.org/geometry#>

Prefix lgdo: <http://linkedgeodata.org/ontology/>

Select ?s, ?l From <http://linkedgeodata.org>

Where

{

 ?s a lgdo:Restaurant ;

 rdfs:label ?l ;

 geom:geometry [

 ogc:asWKT ?g

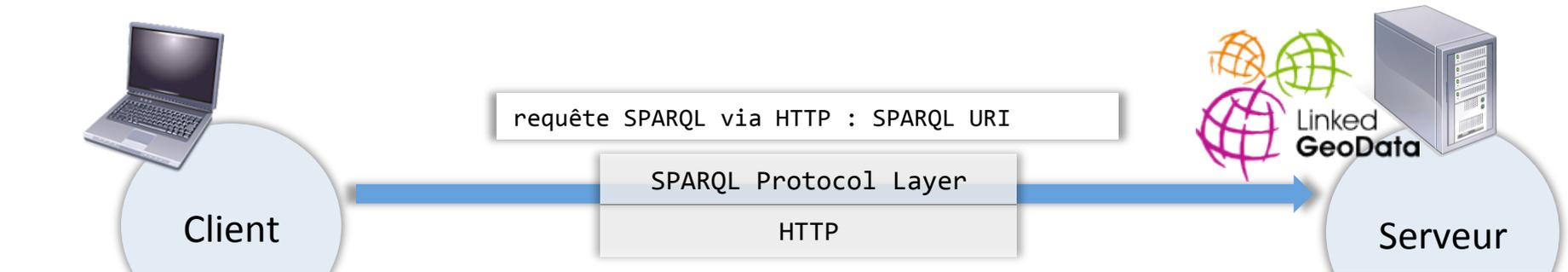
] .

 Filter(bif:st_intersects (?g, bif:st_point (3.692764, 43.393794), 1)) .

}

graphe sur lequel
porte la requête

Le protocole SPARQL



SPARQL Query Language

```
Prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
Prefix ogc: <http://www.opengis.net/ont/geosparql#>
Prefix geom: <http://geovocab.org/geometry#>
Prefix lgdo: <http://linkedgeo.org/ontology/>
```

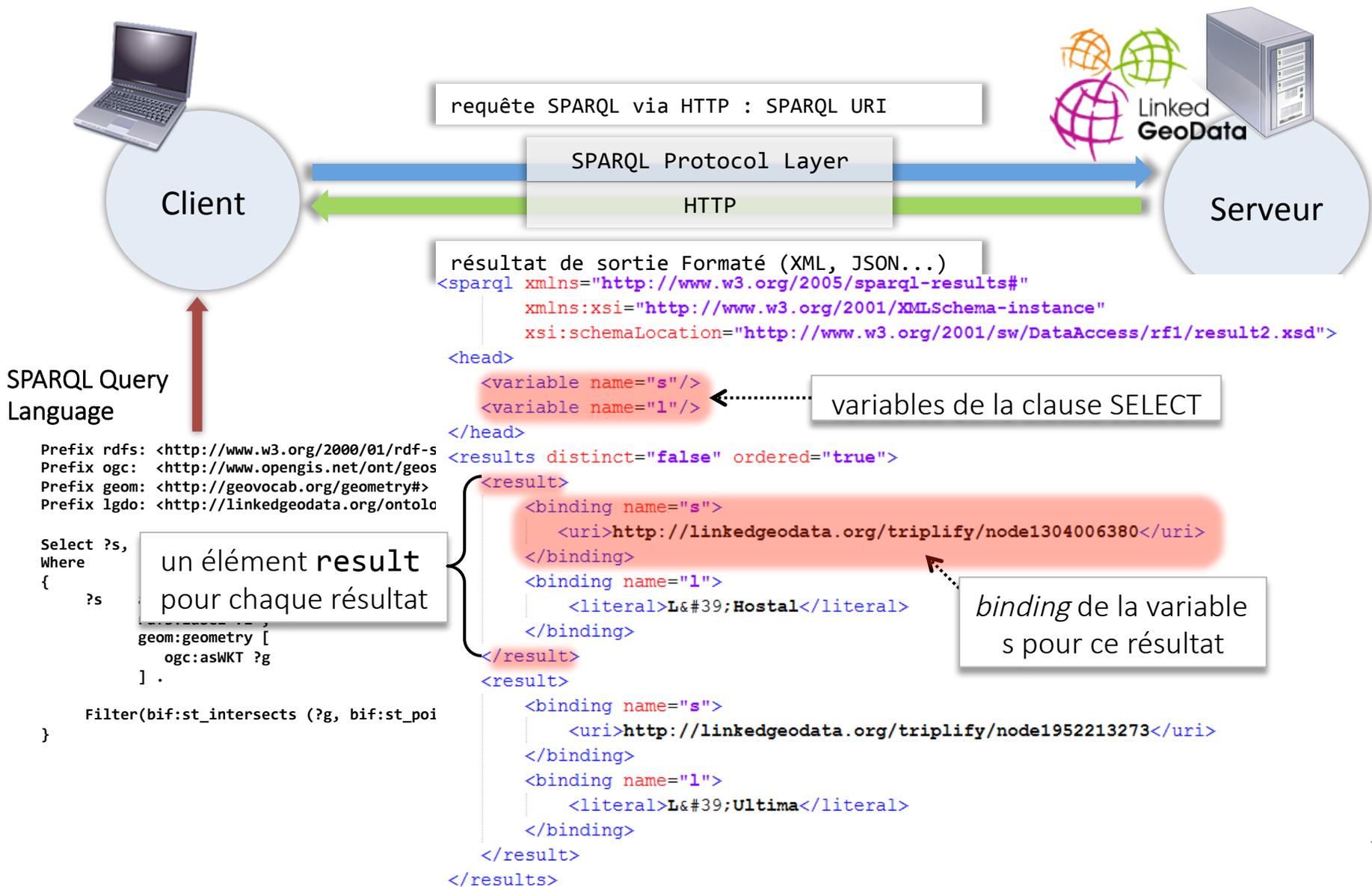
```
Select ?s, ?l From <http://linkedgeo.org>
Where
{
  ?s a lgdo:Restaurant ;
  rdfs:label ?l ;
  geom:geometry [
    ogc:asWKT
  ] .
  Filter(bif:st_intersects(

```

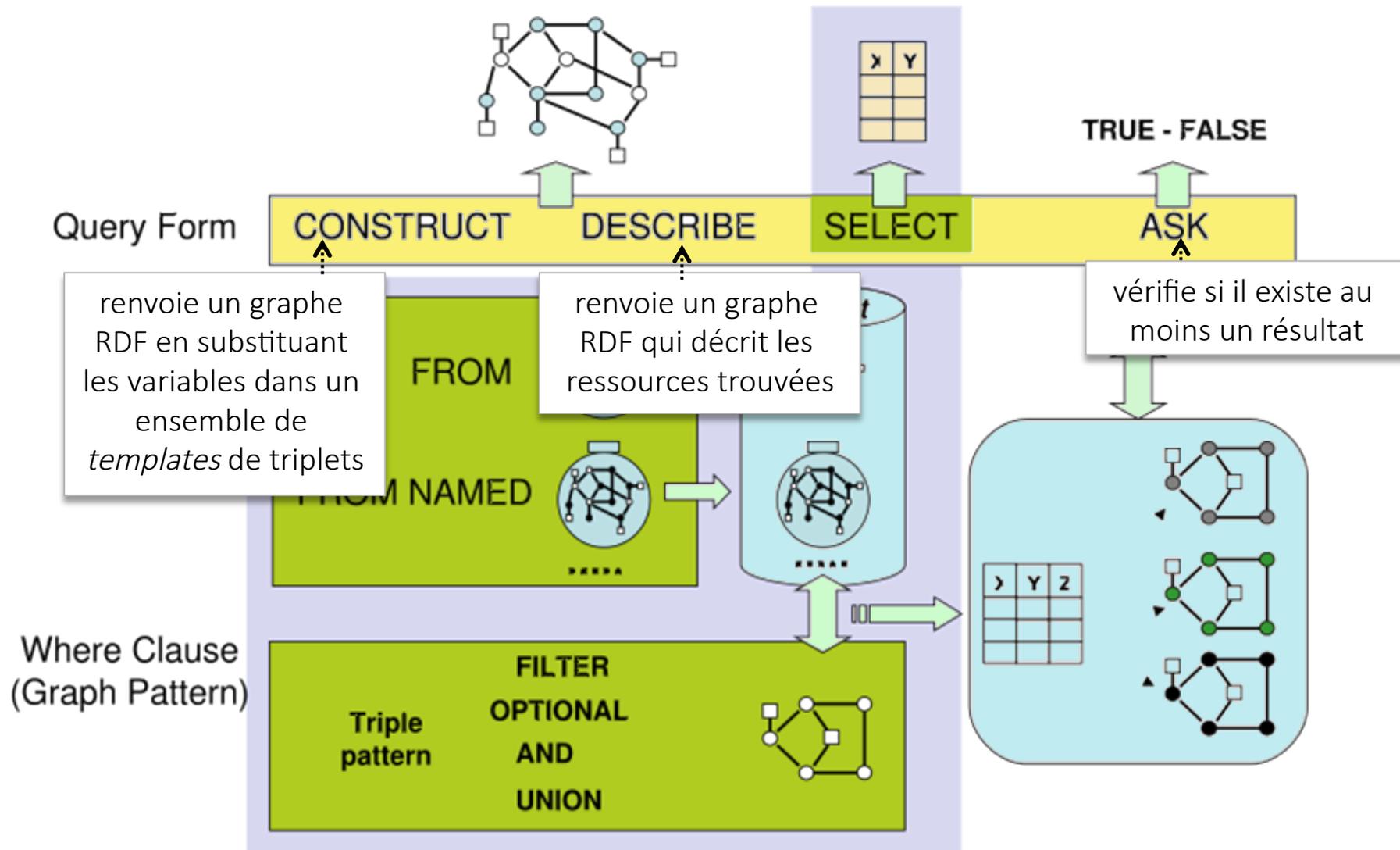
- requête SPARQL (SPARQL URI) composée de 3 parties:
 - URL du SPARQL endpoint <http://linkedgeo.org/sparql>
 - le(s) graphe(s) RDF à requêter (optionnel)
 - **default-graph-uri**=<http://linkedgeo.org>
 - **named-graph-uri**=...
 - la chaîne correspondant à la requête
 - **query**= ...
 - le format pour le résultat
 - **format**=application/sparql-results+xml (text/html, json...)

```
http://linkedgeo.org/sparql?default-graph-uri=http%3A%2F%2Flinkedgeo.org&query=Prefix+rdfs%3A+%3Chttp%3A%2F%2Fwww.w3.org%2F2000%2F01%2Frdf-schema%23%3E%0D%0APrefix+ogc%3A+%3Chttp%3A%2F%2Fwww.opengis.net%2Font%2Fgeosparql%23%3E%0D%0APrefix+geom%3A+%3Chttp%3A%2F%2Fgeovocab.org%2Fgeometry%23%3E%0D%0APrefix+lgdo%3A+%3Chttp%3A%2F%2Flinkedgeo.org%2Fontology%2F%3E%0D%0A%0D%0ASelect+*%0D%0AFrom+%3Chttp%3A%2F%2Flinkedgeo.org%3E+%7B%0D%0A+++%3Fs%0D%0A++++a+lgdo%3ARestaurant+%3B%0D%0A++++rdfs%3Alabel+%3FI+%3B+
+++%0D%0A++++geom%3Ageometry+%5B%0D%0A++++ogc%3AasWKT+%3Fg%0D%0A++++%5D+.%0D%0A%0D%0A++++Filter%28bif%3Ast_intersects+%28%3Fg%2C+bif%3Ast_point+%283.692764%2C+43.393794%29%2C+1%29%29+.%0D%0A%27D&format=application%2Fsparql-results%2Bxml
```

SPARQL : format de sortie (XML)



SPARQL Query Language (1.0)



Source: Pérez, Arenas and Gutierrez, Chapter 1: On the Semantics of SPARQL, Semantic Web Information Management: A Model Based Perspective, Springer 2010

SPARQL 1.1: nouvelles fonctionnalités

- Recommandation du W3C, 21 Mars 2013
 - <http://www.w3.org/TR/2013/REC-sparql11-overview-20130321/>
- Query
 - fonctions d'agrégation des résultats (count, min, max, group by, etc.)
 - affectation de variables
 - `SELECT (COUNT(DISTINCT ?s)) AS ?num` *nombre de restaurants distincts*
 - path expressions (chemin possible entre deux nœud du graphe RDF)
 - subqueries
 - négation (NOT EXIST, EXIST, MINUS)
 - `FILTER NOT EXISTS {?author dbpedia-owl:notableWork ?work .}`
 - requêtes fédérées basiques (SERVICE,BINDING)
 - exécuter des requêtes distribuées sur différents endpoints SPARQL
- nouveaux formats de sérialisation pour résultats de requêtes (JSON...)
- CRUD
 - Mise à jour du graphe (INSERT, INSERT DATA, DELETE DATA, DELETE, DELETE WHERE, LOAD, CLEAR)
 - Gestion des graphes (CREATE, DROP, COPY, MOVE, ADD)
- Entailments for RDF, RDFS, OWL, RIF

Plan

- Introduction
- De l'Open Data au Linked Open Data (1ère partie)
- Les principes des données ouvertes et le Linked Open Data
 - Nommer les données: URI
 - Modéliser les données: RDF
 - Interroger les données: SPARQL
- De l'Open Data au Linked Open Data (2ème partie)
- Conclusion
- Bibliographie



LE RETOUR

Données liées ★★★★★



http://data...

1. Use URIs as names for things.
2. Use HTTP URIs, so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards(RDF, SPARQL).

RDFa = RDF dans attributs HTML

- permet annotations RDF dans documents (X)HTML
- réutilise des attributs existants (ex. href, src) et introduit de nouveaux attributs (vocab, typeof, property, resource, prefix)

Day	Lowest Temperature (°C)
Saturday, 13 November 2010	2
Sunday, 14 November 2010	4
Monday, 15 November 2010	7

```
<h1 property="dcterms:title">Temperature forecast for Galway, Ireland</h1>
<div id="data" about="#Galway" typeof="meteo:Place">
<table border="1px">
  <tr>
    <th>Day</th>
    <th>Lowest Temperature (<deg>C)</th>
  </tr>
  <tr rel="meteo:forecast" resource="#forecast20101113">
    <td>
      <div about="#forecast20101113">
        <span property="meteo:predicted" content="2010-11-13T00:00:00Z" datatype="xsd:dateTime">Saturda
      </div>
    </td>
    <td rel="meteo:temperature">
      <div about="#temp20101113">
        <span property="meteo:celsius" datatype="xsd:decimal">2</span>
      </div>
    </td>
  </tr>
</table>
```



RDFa

```
<h1 property="dcterms:title">Temperature forecast for Galway, Ireland</h1>
```

```
<div id="data" about="#Galway" typeof="meteo:Place">
```

```
<table border="1px">
```

```
<tr>
```

```
<th>Day</th>
```

```
<th>Lowest Temperature (&deg;C)</th>
```

```
</tr>
```

```
<tr rel="meteo:forecast" resource="#forecast20101113">
```

```
<td>
```

```
<div about="#forecast20101113">
```

```
<span property="meteo:predicted" content="2010-11-13T00:00:00Z" datatype="xsd:dateTime">Saturday, 13 November 2010</span>
```

```
</div>
```

```
</td>
```

```
<td rel="meteo:temperature">
```

```
<div about="#temp20101113">
```

```
<span property="meteo:celsius" data-
```

```
</div>
```

```
</td>
```

```
</tr>
```



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

```
<http://rdfa.info/play/>  
  dc:title "Temperature forecast for Galway, Ireland";  
  dc:date "2012-01-22"^^xsd:date;  
  dc:creator <http://sw-app.org/mic.xhtml#i> .  
<http://rdfa.info/play/#Galway>  
  rdf:type <http://purl.org/ns/meteo#Place>;  
  <http://purl.org/ns/meteo#forecast> <http://rdfa.info/play/#forecast20101113>;  
  <http://purl.org/ns/meteo#forecast> <http://rdfa.info/play/#forecast20101114>;  
  <http://purl.org/ns/meteo#forecast> <http://rdfa.info/play/#forecast20101115> .  
<http://rdfa.info/play/#forecast20101113>  
  <http://purl.org/ns/meteo#predicted> "2010-11-13T00:00:00Z"^^xsd:dateTime;  
  <http://purl.org/ns/meteo#temperature> <http://rdfa.info/play/#temp20101113> .  
<http://rdfa.info/play/#temp20101113>  
  <http://purl.org/ns/meteo#celsius> "2"^^xsd:decimal .  
<http://rdfa.info/play/#forecast20101114>  
  <http://purl.org/ns/meteo#predicted> "2010-11-14T00:00:00Z"^^xsd:dateTime;  
  <http://purl.org/ns/meteo#temperature> <http://rdfa.info/play/#temp20101114> .  
<http://rdfa.info/play/#temp20101114>  
  <http://purl.org/ns/meteo#celsius> "4"^^xsd:decimal .  
<http://rdfa.info/play/#forecast20101115>  
  <http://purl.org/ns/meteo#predicted> "2010-11-15T00:00:00Z"^^xsd:dateTime;  
  <http://purl.org/ns/meteo#temperature> <http://rdfa.info/play/#temp20101115> .  
<http://rdfa.info/play/#temp20101115>  
  <http://purl.org/ns/meteo#celsius> "7"^^xsd:decimal .
```

RDF Extrait
de la page HTML

RDFa

```
<h1 property="dcterms:title">Temperature forecast for Galway, Ireland</h1>
```

```
<div id="data" about="#Galway" typeof="meteo:Place">
```

```
<table border="1px">
```

```
<tr>
```

```
<th>Day</th>
```

```
<th>Lowest Temperature (&deg;C)</th>
```

```
</tr>
```

```
<tr rel="meteo:forecast" resource="#forecast20101113">
```

```
<td>
```

```
<div about="#forecast20101113">
```

```
<span property="rdfs:label" value="Visualization" />
```

```
<span property="rdfs:label" value="Raw Data" />
```

```
</div>
```

```
</td>
```

```
<td rel="meteo:temp" resource="#temp20101113">
```

```
<div about="#temp20101113">
```

```
<span property="rdfs:label" value="2" />
```

```
</div>
```

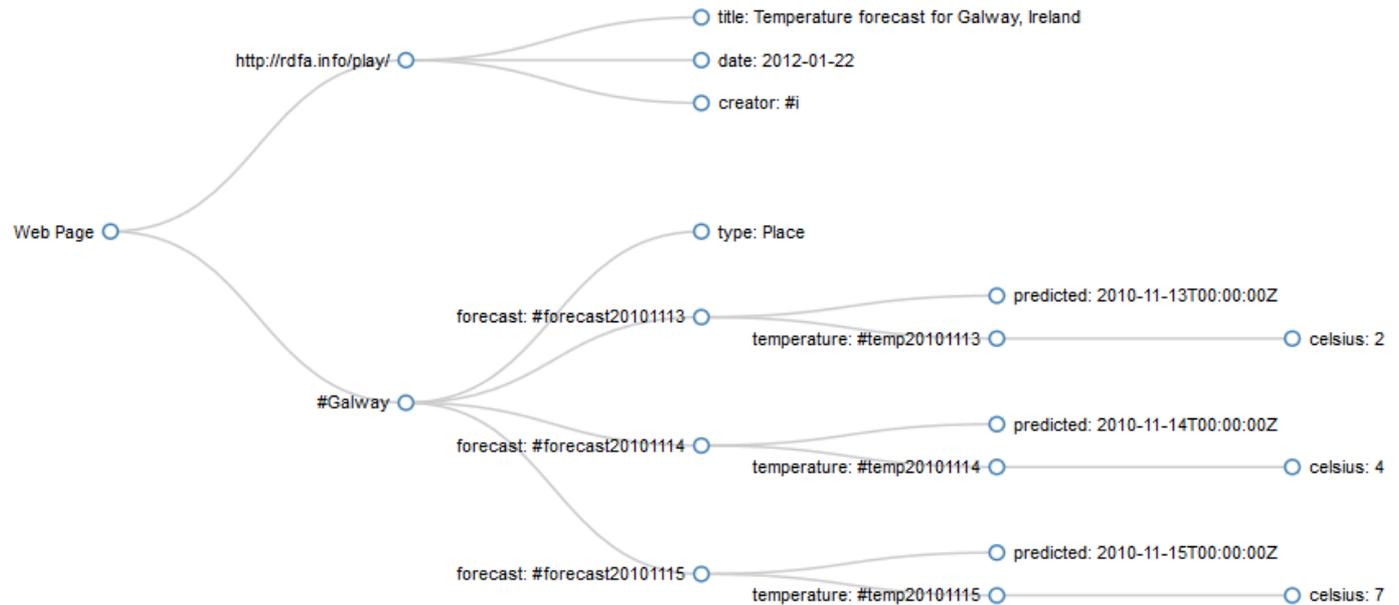
```
</td>
```

```
</tr>
```



extracteur RDFa

<http://rdfa.info/play/>



RDFa

English | Bienvenue. Veuillez créer un compte ou ouvrir une session. | Mon compte | État de la commande | Service à la clientèle

Produits, publications et Trouver un magasin Circulaire Cartes-cadeaux Zone récompenses

CELLULAIRES CARRIERS ACCESSOIRES SERVICES COMMUNAUTÉ BESTBUY.CA

Accueil : Best Buy Mobile : iPhone : iPhone 6 : Information du produit

iPhone 6 64GB - Space Grey - 2 Year Agreement

Modèle n° : MG3H2CLA | Code Web : 10320768

★★★★☆ 3,6 (7 classements) | Notez et évaluez ce produit

Authorized Reseller

374,99 \$

- Livraison gratuite* sur les commandes de plus de 25 \$ (sauf les produits en [livraison spéciale](#))
- Gagnez des points de la Zone récompenses sur tous vos achats
- Retours faciles

En ligne Rupture de stock

Quantité restante : 0

En magasin Réservation impossible

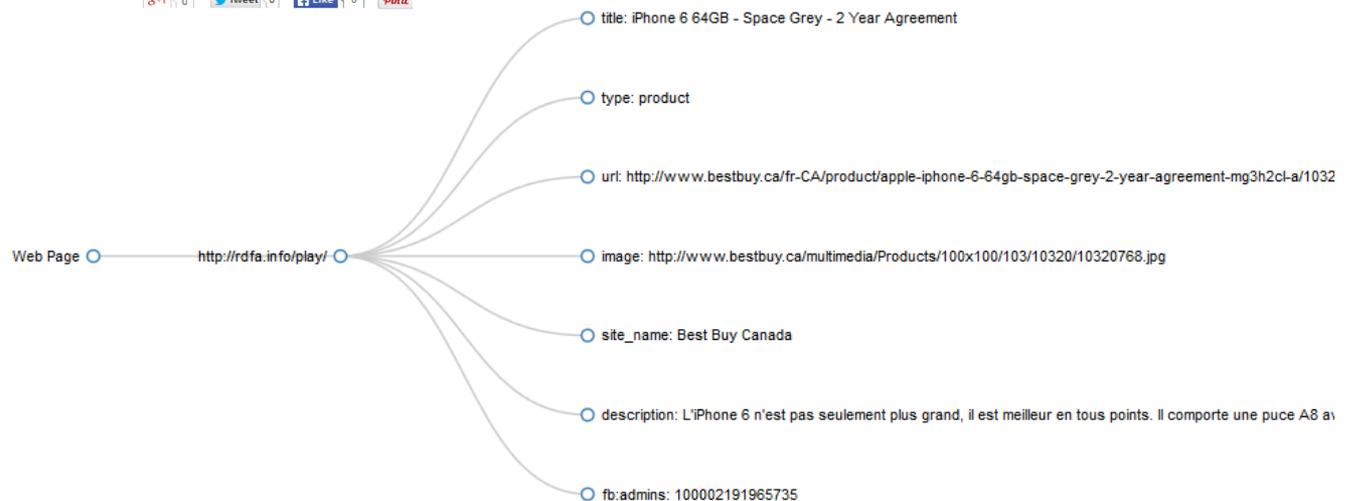
Non disponible dans les magasins les plus près
Vérifier dans d'autres magasins

Ajouter aux souhaits | Suivi du prix

3+1 | 0 | Tweet | 0 | Like | 0 | Pin it

RDFa utilisé par :

- Google, Yahoo
- Facebook, MySpace, LinkedIn
- Best Buy, O'Reilly ...
- Newsweek, BBC
- WhiteHouse.gov, UK government, Library of Congress



Données liées et ouvertes ★★★★★



coûts et bénéfices

– Pour l'utilisateur... en plus des avantages de



- ✓ You can link to it from any other place (on the Web or locally).
- ✓ You can bookmark it.
- ✓ You can reuse parts of the data.
- ✓ You may be able to reuse existing tools and libraries, even if they only understand parts of the pattern the publisher used.
- ✗ Understanding the structure of an RDF "Graph" of data can be more effort than tabular (Excel/CSV) or tree (XML/JSON) data.
- ✓ You can combine the data safely with other data. URIs are a global scheme so if two things have the same URI then it's intentional, and if so that's well on it's way to being 5 star data!

Michael Hausenblas <http://5stardata.info/>

Données liées et ouvertes



-  coûts et bénéfices

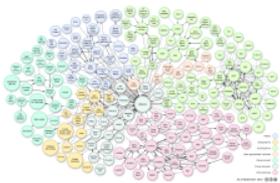
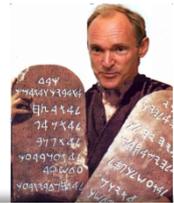
- Pour celui qui publie... en plus des avantages de 
 - ✓ You have fine-granular control over the data items and can optimise their access (load balancing, caching, etc.)
 - ✓ Other data publishers can now link into your data, promoting it to 5 star!
 - ✗ You typically invest some time slicing and dicing your data.
 - ✗ You'll need to assign URIs to data items and think about how to represent the data.
 - ✗ You need to either find existing patterns to reuse or create your own.

Michael Hausenblas <http://5stardata.info/>

Données liées et ouvertes



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



```
<h1 property="dcterms:title">Temperature forecast for Galway, Ireland</h1>
<div id="data" about="#Galway" typeof="meteo:Place">
  <span rel="owl:sameAs" resource="http://dbpedia.org/resource/Galway"></span>
  <table border="lpx">
    <tr>
      <th>Day</th>
      <th>
        <div about="#temp">
          Lowest
          <a rel="rdfs:seeAlso" href="#">
            <span rel="owl:sameAs" resource="http://dbpedia.org/resource/Galway"></span>
          </a>
        </div>
      </th>
    </tr>
    <tr rel="meteo:forecast" resource="#forecast">
      <td>
        <div about="#forecast20101113">
          <span property="meteo:prediction">
            </span>
        </div>
      </td>
      <td rel="meteo:temperature">
        <div about="#temp20101113">
          <span property="meteo:celsius">
            </span>
        </div>
      </td>
    </tr>
  </table>

```

dbpedia.org/page/Galway

Ontologies and Vocabu... What is URI (Uniform R...

About: Galway

An Entity of Type : [populated place](#), from Named Graph : [http://dbpedia.org](#), within Data Space : [dbpedia.org](#)

Galway est une ville de la province de Connacht, dans le Comté de Galway, sur la côte ouest de l'Irlande. Son nom vient de la rivière Corrib (Gaillimh) qui traverse la ville. Elle est également surnommée la « ville des tribus » en référence aux quatorze tribus qui se partageaient la ville à l'époque anglo-normande. La population de la ville est de habitants. L'agglomération de Galway est la quatrième du pays par le nombre d'habitants, après celles de Dublin, de Cork et de Limerick.

Property	Value
dbpedia-owl:abstract	<ul style="list-style-type: none">Galway és una ciutat d'Irlanda, capital del comtat de Galway, a la p...és una institució local de Galway. Antigament es digué Galway Cor...el desembre del 1484 pel rei Ricard III d'Anglaterra. El primer alcald...Galway, però revifat el 1937 com a municipi i el 1985 com a municip...Galway (írsky Gaillimh, nebo také An Gaillimh) je město v západoirs...rozkládá na západním pobřeží Írska. V irštině je označováno také ja...rostoucí írské město. Počet obyvatel v roce 2006 činil rovných 70 00...Galway ist die Hauptstadt der Grafschaft Galway in der Provinz Con...Galway or City of Galway (Cathair na Gaillimhe) is a city on the wes...Galway Bay and is surrounded by County Galway. It is the third larg...into fourth place behind Limerick. The population of Galway city at t...Galway es la capital del condado de Galway, en Irlanda. La ciudad :

Firefox

Temperature forecast for Galwa... +

5stardata.info/gtd-5.html

LarKC weblog » Blog A... Linking Open D...

Temper for Galway, Ireland

Day	Lowest Temperature (°C)
Saturday, 13 November 2010	2
Sunday, 14 November 2010	4
Monday, 15 November 2010	7

Last update: 2012-01-22 by Michael | Code available via GitHub

zotero

Données liées et ouvertes



-  coûts et bénéfices

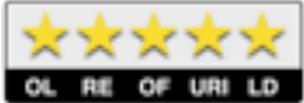
– Pour l'utilisateur... en plus des avantages de



- ✓ You can discover more (related) data while consuming the data.
- ✓ You can directly learn about the data schema.
- ✗ You now have to deal with broken data links, just like 404 errors in web pages.
- ✗ Presenting data from an arbitrary link as fact is as risky as letting people include content from any website in your pages. Caution, trust and common sense are all still necessary.

Michael Hausenblas <http://5stardata.info/>

Données liées et ouvertes ★★★★★

-  coûts et bénéfices

– Pour celui qui publie

- ✓ You make your data discoverable.
- ✓ You increase the value of your data.
- ✓ Your own organisation will gain the same benefits from the links as the consumers.
- ✗ You'll need to invest resources to link your data to other data on the Web.
- ✗ You may need to repair broken or incorrect links.

Michael Hausenblas <http://5stardata.info/>

Pour
conclure

The importance of Linked Data

Ivan Herman <http://www.w3.org/2009/Talks/0615-SanJose-tutorial-IH/>

- It provides a core set of data that Semantic Web applications can build on
 - stable references for “things”,
 - e.g., <http://dbpedia.org/resource/Grenoble>
 - many relationships that applications may reuse
 - e.g., the BBC application!
 - a “nucleus” for a larger, semantically enabled Web!
- For many, publishing data may be the first step into the world of Semantic Web

Some things to remember if you publish data

Ivan Herman <http://www.w3.org/2009/Talks/0615-SanJose-tutorial-IH/>

- Publish your data first, care about sexy user interfaces later!
 - the “raw data” can become useful on its own right and others may use it
 - you can add your added value later by providing nice user access
- If possible, publish your data in RDF but if you cannot, others may help you in conversions
 - trust the community...
- Add links to other data. “Just” publishing isn’t enough...

Open challenges for LOD

- Contradictions:
Population of Grenoble
 - DBpedia: 156,659
 - Freebase: **155,632**
- **Identity crisis:** is the capital of the Roman Empire the *same* As the capital of modern Italy?
- Data ownership, copyright, access control



Nombreux outils (pas une liste exhaustive !)

■ Quelques noms:

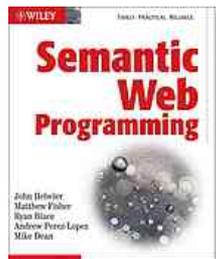
- Jena, AllegroGraph, Mulgara, Sesame, flickurl, 4Store, ...
- TopBraid Suite, Virtuoso environment, Falcon, Drupal 7, Redland, Pellet, Hermit ...
- Disco, Oracle 11g, RacerPro, IODT, Ontobroker, OWLIM, Talis Platform, ...
- RDF Gateway, RDFLib, Open Anzo, DartGrid, Zitgist, Ontotext, Protégé, ...
- Thetus publisher, SemanticWorks, SWI-Prolog, RDFStore...
- ...

plus sur <http://www.w3.org/2001/sw/wiki/Tools>

■ Catégories:

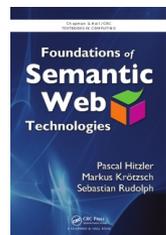
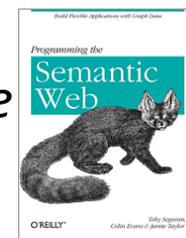
- Triple Stores
- Inference engines
- Converters
- Search engines
- Middleware
- CMS
- Semantic Web browsers
- Development environments
- Semantic Wikis
- ...

Bibliographie



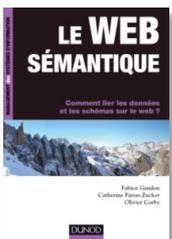
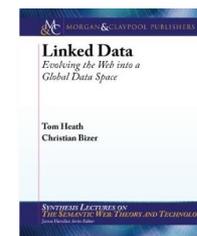
- John Hebel, Matthew Fisher, Ryan Blace, Andrew Perez-Lopez, Mike Dean, *Semantic Web Programming*, Wiley, (2009)

- Toby Segaran, Jamie Taylor, and Colin Evans, *Programming the Semantic Web*, O'Reilly, (2009)



- P. Hitzler, R. Sebastian, and M. Krötzsch: *Foundation of Semantic Web Technologies*, (2009)

- T. Heath and C. Bizer: *Linked Data: Evolving the Web into a Global Data Space*, (2011)



- F. Gandon, C. Faron-Zucker, O. corby : *Le Web Sémantique* (Dunod 2012)

- ... page wiki du W3C wiki plus de references : <http://www.w3.org/2001/sw/wiki/Books>

Merci de
votre attention

