

The Semantic Web Technology Stack (not a piece of cake...)

Most apps use only a subset of the stack

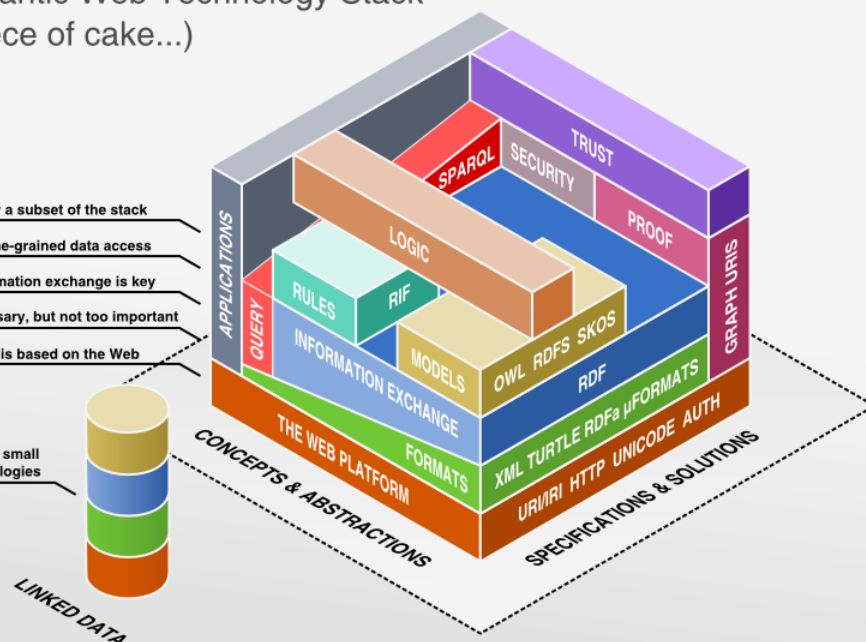
Querying allows fine-grained data access

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



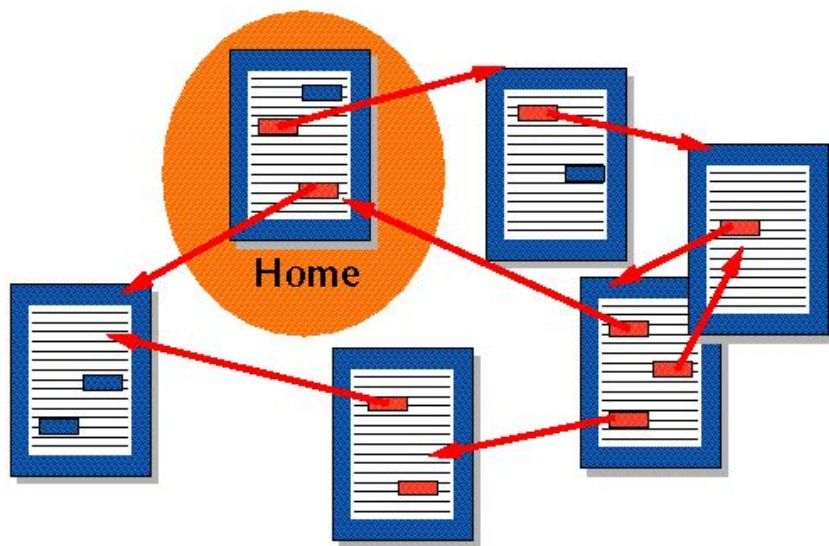
Partie 2 : modélisation des données sur le Web

Sylvie Calabretto, Mehdi Kaytoue et Aimene Belfodil

Objectif du cours

- Avant de commencer cette séance, qu'avez-vous retenu de la première séance ?
- A la fin du cours, vous êtes censé savoir répondre à :
 - Comment **représenter** les données dans le **Web 3.0**?
 - Qu'est ce que **RDF**?
 - Qu'est qu'une **ressource** dans le Web ? Comment est-elle **identifiée**? Qu'est ce qu'on entend par **URL, URI, IRI**?
 - Comment **publier** les **données sémantiques** sur le Web?
 - ...

Rappel (1) – La préhistoire du Web et la notion d'Hypertexte



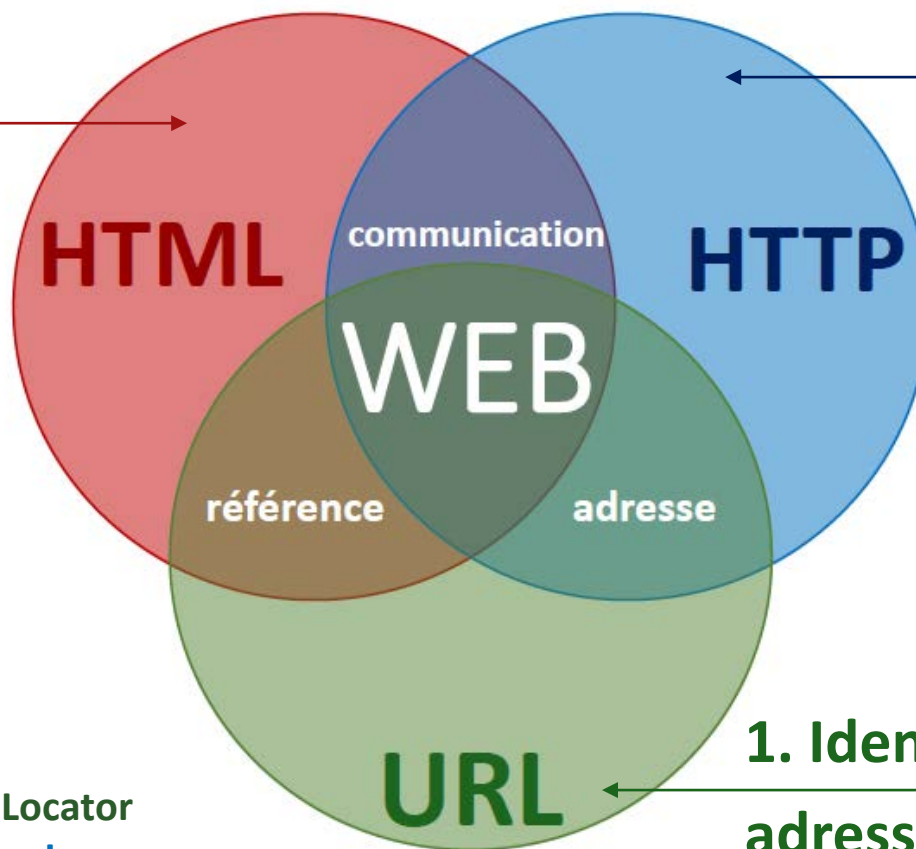
Hypertexte – ensemble de documents contenant des unités d'informations liées entre elles par des **Hyperliens (liens)**



Du **Memex (Vannevar Bush – 1945)** au **Web de Tim Berners-Lee (1989)** en passant par l'introduction du mot **hypertexte** par **Ted Nelson (1965)**.

Rappel (2) – Les 3 composants de l'architecture du Web des documents

2. Langage de représentation des ressources

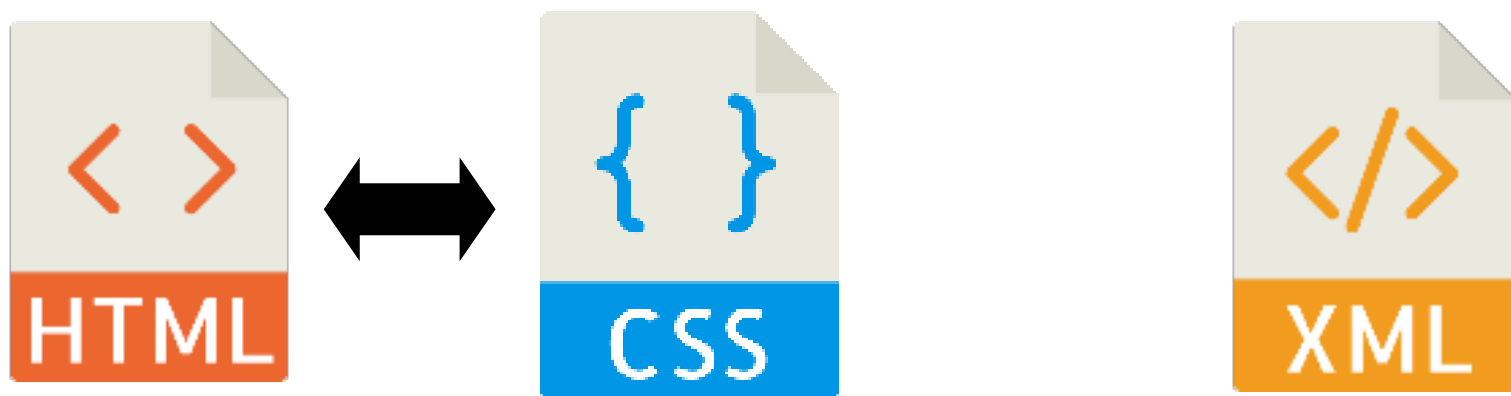


3. Protocole de communication de ces représentations

1. Identification et adressage des ressources

URL – Uniform Resource Locator
HTML – HyperText Markup Language
HTTP – HyperText Transfer Protocol

Rappel (3) – Initiatives avant le Web 3.0



Séparer le **fond**
de la **forme**

Echange de
données via **XML**

Rappel (4) – Histoire du Web

WEB 1.0

(Read Only Web)

Web documentaire

Connecter des documents
(documents statiques)



read



write



WEB 2.0

(Read/write Web)

Web social

Connecter des gens
(documents dynamiques)

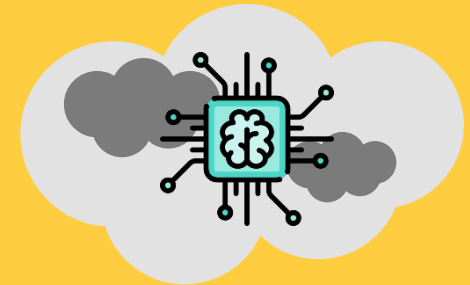


WEB 3.0

(Read/write Web)

Web sémantique

Connecter la connaissance



Aujourd'hui

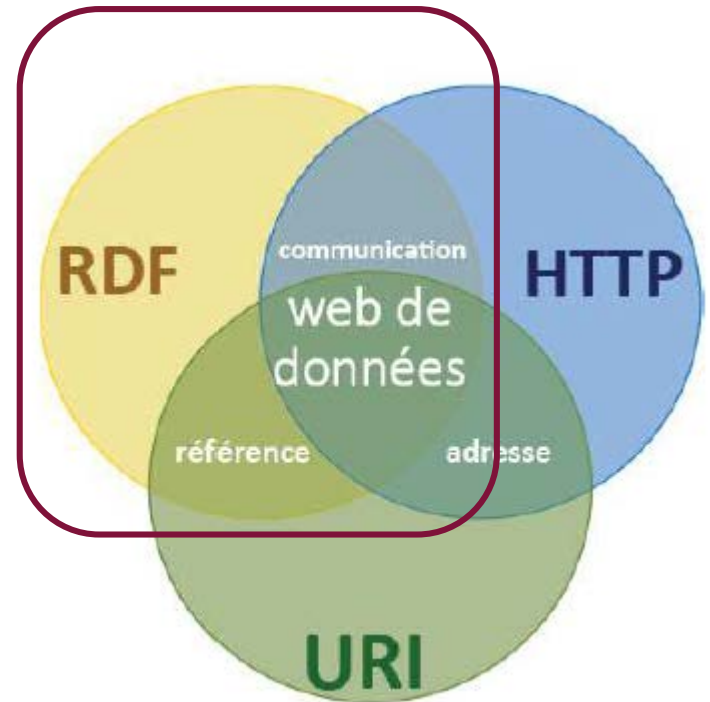


World Wide Web Consortium

Recommande l'utilisation de **RDF**
(Resource Description Framework) pour
représenter ces données sur le Web.

<https://www.w3.org/TR/rdf11-primer/>

Nouvelle représentation



Remarque: Une recommandation est le fruit du pipeline:
atelier → working draft → last call → candidate/proposed
recommendation → **recommandation**

Motivations d'un nouveau format de représentation des données

- **Métadonnées Web** : fournir des informations sur les ressources Web
- **Traitement automatique de l'information** : d'une information uniquement lisible par les humains (pages web) vers une information **interprétable par la machine** (e.g. un document correspond à un hôtel, un hôtel est un mode d'hébergement, dans un hôtel, ...)
- **Collaboration entre applications** : combiner les données de plusieurs applications, créer de nouvelles informations (**Raisonnement**).
- **Modèles d'information ouverts** : annotation de ressources par les différents utilisateurs du Web afin de créer un réseau mondial de connaissances. On parle de:

Slogan **AAA**: Anyone can say Anything about Anything.

RDF – Resource Description Framework:

Une recommandation fondamentale

Objectifs de RDF

- avoir un **modèle de données** simple
- avoir une **sémantique formelle** pour pouvoir raisonner sur les données, l'interpréter.
- Slogan **AAA**: **A**nyme can say **A**nthing about **A**nthing

Basé sur l'architecture classique du Web

- **URI** pour identifier les ressources décrites et les liens entre elles.
- **HTTP** pour la communication.
- **XML** (et **autres**) pour la syntaxe concrète derrière.

Gérer les données distribuées : représenter l'information de la manière la **moins contraignante** et la **plus flexible possible**

Exemple (1)

Tous Vidéos Images Actualités Maps Plus Paramètres Outils

Environ 502 000 résultats (0,71 secondes)

Vidéos

Les Triplettes de Belleville (Bande Annonce)

MagelisTV
YouTube - 13 déc. 2012

Les Triplettes de Belleville - Film complet en Français

Dylan PROD
YouTube - 25 janv. 2018

The Triplets of Belleville (2003) - Belleville Rendez-vous

Argus Argus
YouTube - 8 janv. 2010

Les Triplettes de Belleville — Wikipédia

https://fr.wikipedia.org/wiki/Les_Triplettes_de_Belleville

Les Triplettes de Belleville est un film d'animation franco-belgo-québécois écrit (d'après son œuvre) et réalisé par Sylvain Chomet, sorti en 2003, largement ...

Scénario: Sylvain Chomet Sortie: 2003

Réalisation: Sylvain Chomet

Résumé · Fiche technique · Distinctions

Les Triplettes de Belleville - film 2002 - AlloCiné

www.allocine.fr > Cinéma > Films à l'affiche

★★★★☆ Note : 3,3 - 5 059 votes

Plus d'images

Les Triplettes de Belleville

2003 · Drame/Aventure · 1h 22m

[Regarder la bande-annonce sur YouTube](#)

3,3/5 · AlloCiné

91 % des utilisateurs ont aimé ce film

Utilisateurs de Google

Madame Souza élève seule son petit-fils, un garçonnet mélancolique. Quand elle découvre que l'enfant aime le vélo, elle l'encourage à s'entraîner, lui faisant même subir un programme rigoureux. Les années passent et le petit garçon devient un sportif accompli, à tel point qu'il participe au Tour de ... [PLUS](#)

Date de sortie : 11 juin 2003 (France)

Réalisateur : Sylvain Chomet

Personnages : Blanche Triplette, Mendiant à Belleville, [PLUS](#)

Exemple (2)



[Tous](#) [Actualités](#) [Shopping](#) [Images](#) [Vidéos](#) [Plus](#) [Paramètres](#) [Outils](#)

Environ 7 100 000 résultats (0,51 secondes)

Résultats pour **Christophe Castaner**
Essayez avec l'orthographe [Christophe Castane](#)

À la une



Sénat: première houleuse pour Christophe Castaner hué par les sénateurs...

BFMTV

Il y a 9 heures



Christophe Castaner, « communautariste » ? Même Marine Le Pen n'a pas osé

Le Point

Il y a 23 heures



Comment Castaner est devenu le premier marcheur de Macron

JDD

Il y a 35 minutes

[➔ Plus de résultats pour "Christophe Castaner"](#)

Christophe Castaner — Wikipédia
https://fr.wikipedia.org/wiki/Christophe_Castaner ▼
Christophe Castaner, né le 3 janvier 1966 à Ollioules (Var), est un homme politique français. Élu

Christophe Castaner

Homme politique



Christophe Castaner, né le 3 janvier 1966 à Ollioules, est un homme politique français. Élu député pour les Alpes-de-Haute-Provence en 2012, il quitte en 2016 le Parti socialiste — dont il était membre depuis 1986 — pour rejoindre le mouvement En marche, conduit par Emmanuel Macron. [Wikipédia](#)

Date et lieu de naissance : 3 janvier 1966 (Âge: 52 ans), Ollioules

Enseignement : Université d'Aix-Marseille

Parti : En Marche !

Recherches associées

[Voir d'autres éléments \(plus de 10\)](#)



Richard Ferrand



Benjamin Griveaux



Christian Estrosi



Olivier Faure



Stéphane Travert

[Revendiquer cette fiche info](#) [Commentaires](#)

Exemple (3)

[Tous](#) [Actualités](#) [Vidéos](#) [Shopping](#) [Images](#) [Plus](#) [Paramètres](#) [Outils](#)

Environ 129 000 000 résultats (0,55 secondes)

[Red Dead Redemption 2 | Prix de lancement 46,49€](#)
[\[Annonce\]](#) www.carrefour.fr/ ▼
Précommandez en ligne Read Dead 2 & Retirez-le en Magasin Carrefour le 26 Octobre !

Catalogue Carrefour
Accédez à l'intégralité Des Catalogues et Promotions en cours !

Votre Compte Carrefour
Profitez de tous les Avantages de l'Espace Perso: Chèque Fidélité

À la une



«Red Dead Redemption 2»: Le «crunch» ou bosser jusqu'à 100 heures par semaine s...

20 Minutes
Il y a 2 heures



«Red Dead Redemption 2»: un patron de Rockstar jette le doute sur les conditions de...

Libération
Il y a 23 heures



Red Dead Redemption 2 : La bande-annonce de lancement datée

Gameblog
Il y a 16 minutes

→ Plus de résultats pour "red dead redemption 2"



Plus d'images

Red Dead Redemption 2

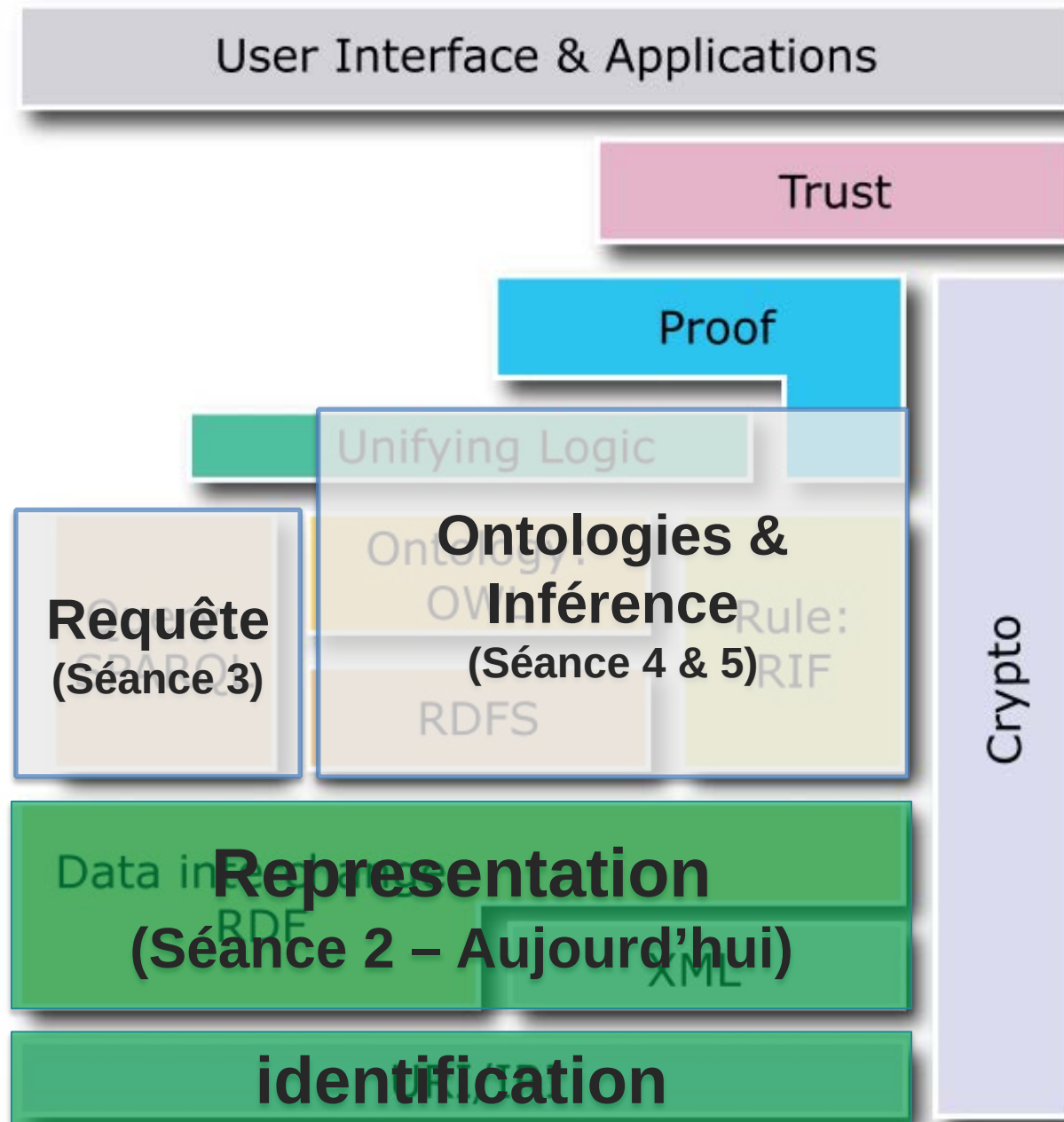
Jeu vidéo

Red Dead Redemption 2, typographié Red Dead Redemption II et également abrégé RDR 2, est un jeu vidéo d'action-aventure et de western multiplateforme, développé par Rockstar Studios et édité par Rockstar Games, dont la sortie est prévue pour le 26 octobre 2018 sur PlayStation 4 et Xbox One. [Wikipédia](#)

Date de sortie initiale : 26 octobre 2018
Mode de jeu : Solo, multijoueur
Début du projet : 2010
Développeur : [Rockstar Games](#)
Moteur : [Rockstar Advanced Game Engine](#)
Plates-formes : PlayStation 4, Xbox One

Sémantique associée à la recherche (film, homme politique, jeu, ...)

Semantic Web Layer Cake



Plan du cours

Vers un Web des données

1. Modélisation

- Décrire des données sur Web (triplets et graphes RDF)
- Syntaxes concrètes pour un modèle abstrait
- Identifier des ressources sur le Web (URI, espace de noms)
- Quelques détails sur les éléments d'un graphe RDF.

2. Publication des données

- RDF, SGBD2RDF, RDFa

3. Accès aux données

- Déréférencement et négociation de contenu.

4. Intégration des données

- Fusion, linked open data

1. Modélisation

1.a. Le modèle RDF

Resource Description Framework (RDF)

R

Resource : « entité », « chose », en fait tout ce à quoi on peut se référer est une ressource (page Web, image, personne réelle ou fictive, lieu, dispositif, évènement, organisation, service, ...)

D

Description : une **ressource** peut être décrite par ses **attributs** et les **valeurs** que ces derniers prennent mais aussi par les **relations (et leur type)** qu'elle peut avoir avec d'autres **ressources**.

F

Framework : ensemble de **modèles**, **langages** et **syntaxes** de ces **descriptions**

Recommandation W3C: <https://www.w3.org/TR/2004/REC-rdf-primer-20040210/>

La notion de ressource

Sur le Web, toute chose d'intérêt est désignée par une **ressource**

Une **ressource** est **identifiée par un URI (Uniform Resource Identifier)** de manière unique. Deux types de ressources:

- **De type 'information'** : on peut obtenir une **représentation** via **HTTP(S)** (une page Web, une image, etc.). On parle de **URL (Uniform Resource Locator)**
- **De type 'non-information'** : on ne peut pas obtenir directement de **représentation** mais on veut pouvoir en parler quand même (l'amphi Gaston Berger, le livre **urn:isbn:3540627715**). On parle de **URN (Uniform Resource Name)**. On parlera plus tard dans la séance de la **négociation de contenu**.

On parle parfois mais rarement de **IRI (Internationalized Resource Identifier)**

https://en.wikipedia.org/wiki/Order_theory

URI (de type **URL**)
Identifiant la ressource

Order Theory

Déréférencement

Représentation
obtenu via le
protocole **HTTPS**



Quel modèle utiliser pour décrire les ressources?

Une représentation classique : les tables

- **lignes** : sujet
- **colonnes** : attribut/propriété
- **cellules** : valeur de la propriété pour le sujet

Nom / Prénom		Adresse de courriel	Rôles
 el Omari Alaoui Abdelaziz		abdelaziz.el-omari-alaoui@insa-lyon.fr	Étudiant
 Sauvestre Agathe		agathe.sauvestre@insa-lyon.fr	Étudiant
 Belfodil Aimene		aimene.belfodil@insa-lyon.fr	Enseignant, Enseignant non éditeur, Inscripteur de cohorte
 Paugois Alan		alan.paugois@insa-lyon.fr	Étudiant
 Fournier Alexandre		alexandre.fournier@insa-lyon.fr	Étudiant

Il faut pouvoir aisément:

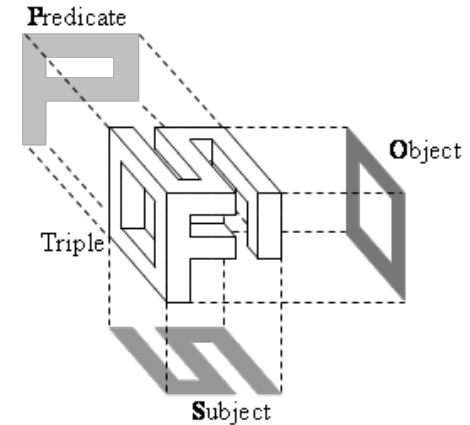
- **Distribuer** des données sur le Web
- **Fusionner/Intégrer** les données d'information multiple

*Il faut pouvoir **identifier les choses** et échanger des **fragments de connaissances** sur elles entre serveurs distants*



Les Triplets RDF – Les atomes de connaissances

Décrire et connecter des ressources avec des triplets RDF
(**sujet** , **prédictat** , **objet**)



- Un énoncé axiomatique, un fait affirmé, une **assertion**.
- On appelle le **prédictat** la **propriété** du triplet. C'est un **prédictat binaire** en logique de premier ordre (*cf. 4-IF-ALIA-Logique et 4-IF-ALIA-PR1-Prolog*).

• **Triplet:** (<https://perso.liris.cnrs.fr/mehdi.kaytoue/>, foaf:name, 'Mehdi Kaytoue')

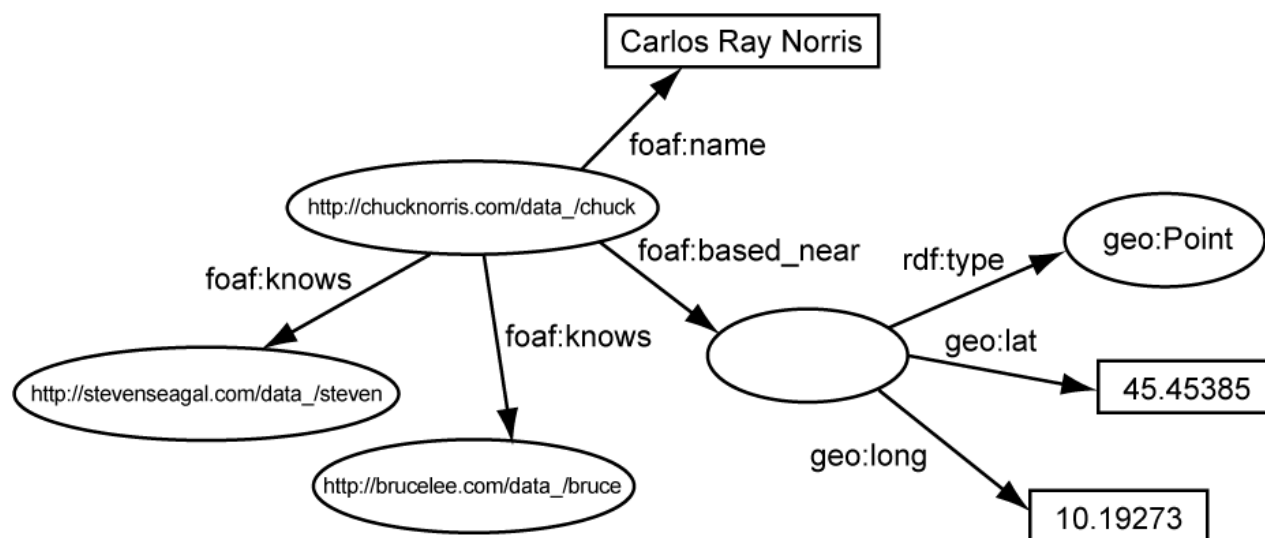
• **Diagramme:**

RDF - Modèle de données générique de type graphe (1)

Un ensemble de triplets RDF se voit comme un graphe RDF

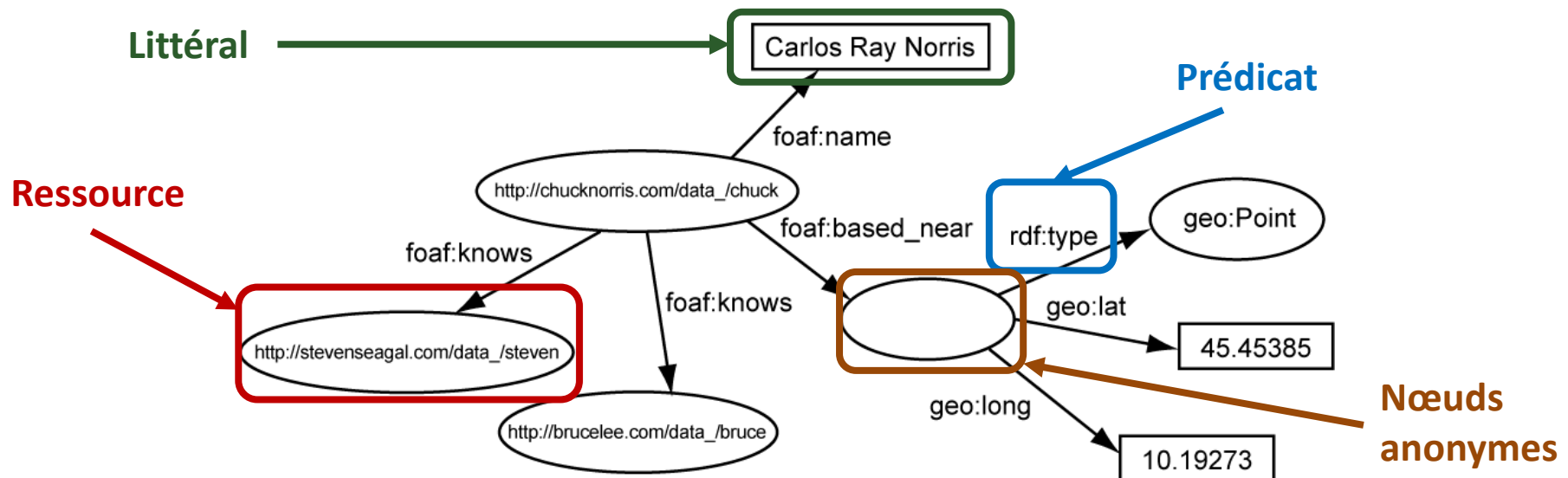
- **Noëuds étiquetés** : les **sujets** et **objets** des triplets
- **Arêtes étiquetées** : les **prédicats** des triplets, des relations entre les ressources
- **Multi-graphe** : différents **prédicats** peuvent lier un même **sujet** à un même **objet**
- **Orienté** : du **sujet** vers l'**objet**

Décrire les informations pour être utilisées par la machine, pas par l'humain



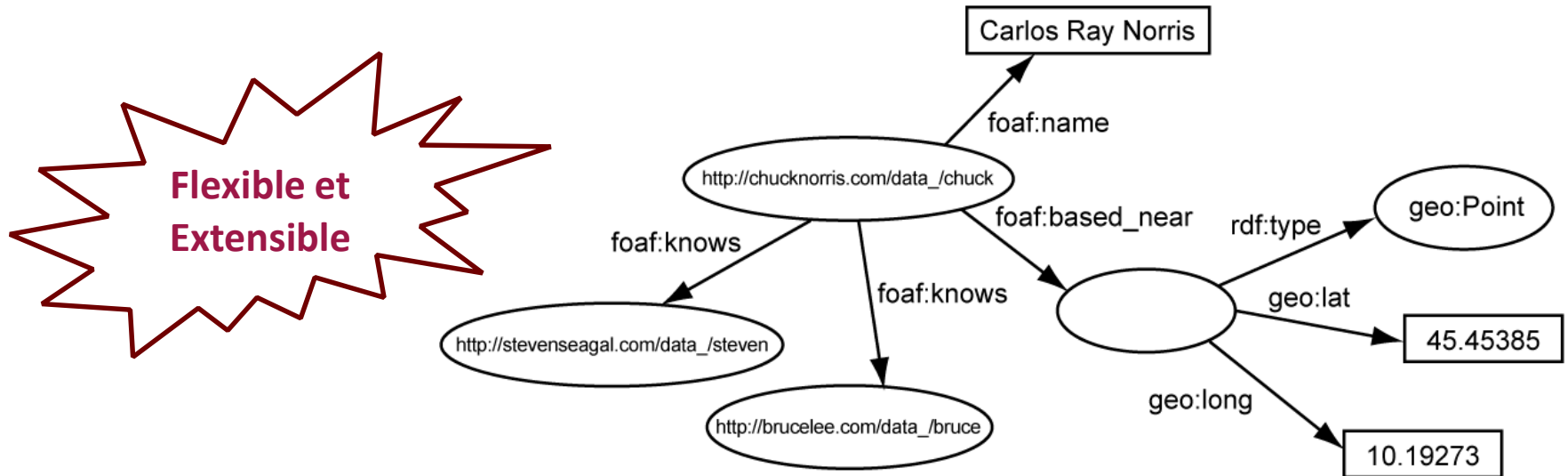
RDF - Modèle de données générique de type graphe (2)

- **Ressource** : ellipse étiquetée par son **URI**. Elle joue le rôle d'un **sujet** ou d'un **objet** dans les triplets RDF.
- **Littéral** : rectangle étiqueté par une chaîne **pouvant être typée** (exemple: date, nombre, ...). Il peut être uniquement un **objet** dans une assertion RDF.
- **Prédicat** : flèche orientée du sujet vers l'objet étiqueté par **URI**
- **Nœud anonyme** : ellipses non étiquetées. On peut connaître une ressource sans pouvoir la référencer (Exemple. lieu d'habitation de Chuck Norris).



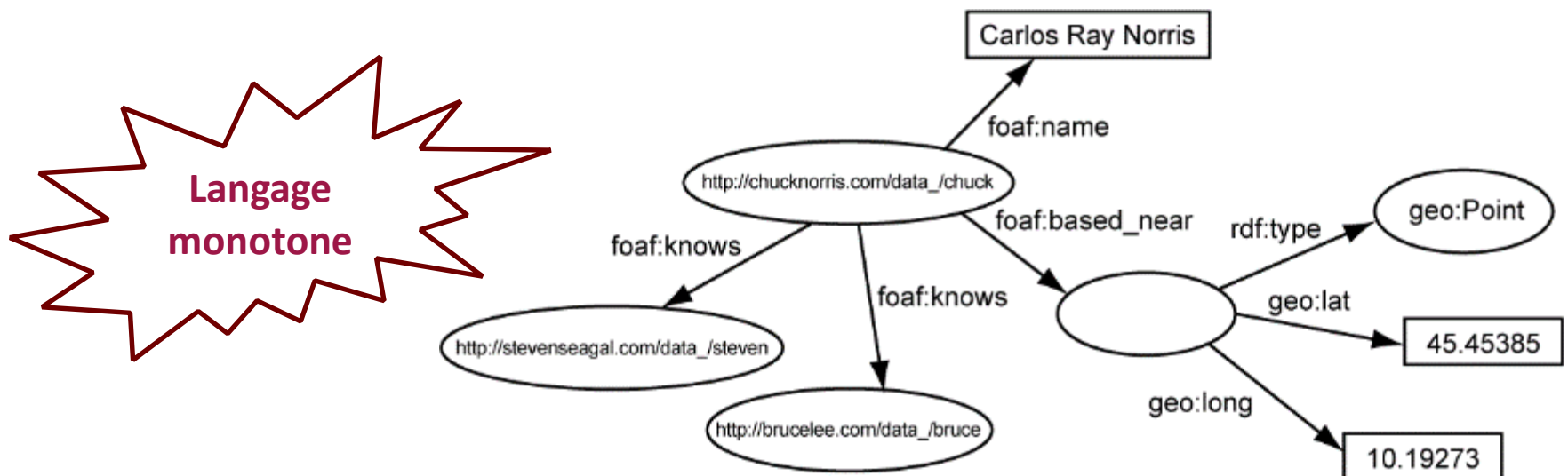
RDF - une syntaxe abstraite (1)

- **Syntaxe** sur laquelle la **sémantique formelle** est définie.
- La **syntaxe RDF (abstraite)** est à distinguer des **implémentations** (XML/RDF, JSON/RDF,...) qui constituent la **syntaxe concrète**. Ces implémentations sont libres de représenter les **graphes RDF** sous toute autre forme équivalente.
- **L'assertion d'un graphe RDF** est équivalente à l'**assertion de tous ses triplets**, et par conséquent la signification d'un graphe RDF est la **conjonction (ET logique)** des faits correspondant à tous ses triplets.



RDF - une syntaxe abstraite (2)

- **RDF** est un sous-ensemble de la logique du premier ordre
 - **Avec** : prédicats binaires, quantification existentielle ($\exists$), conjonction ($\wedge$)
 - **Sans** : disjonction, négation, quantification universelle ($\forall$), ...
- **RDF est monotone** : Tout énoncé RDF est considéré comme vrai. Une base de connaissances RDF ne peut être que étendue. Analogiquement à **Prolog**: RDF est une suite de **assert(...)** où il n'y a pas de **retract(...)**.



1.b. Syntaxes concrètes

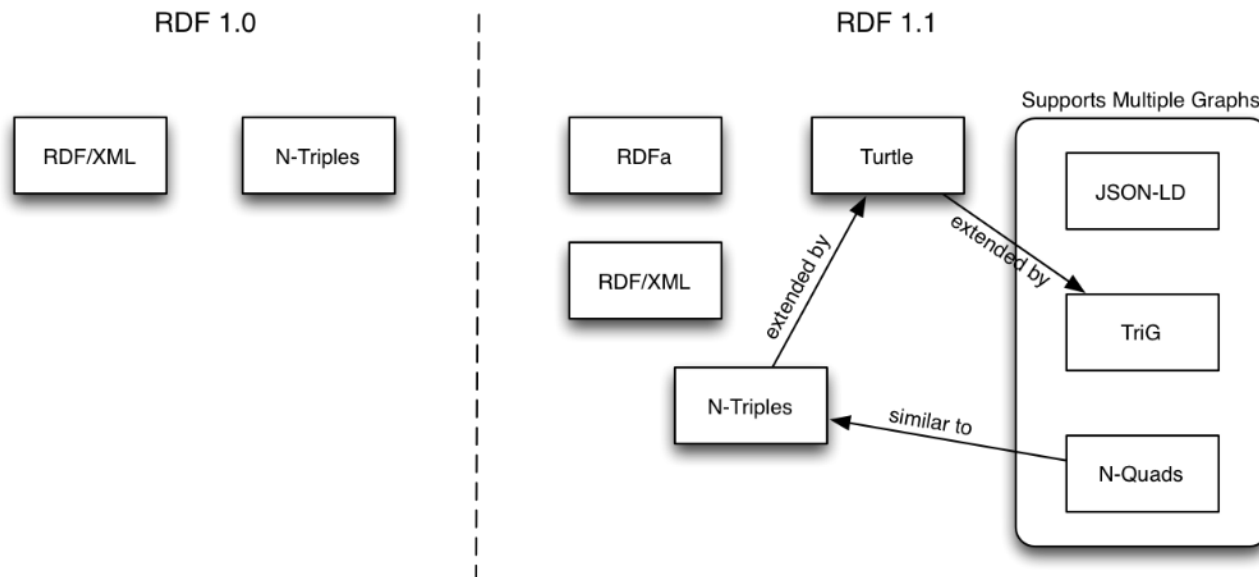
RDF - plusieurs syntaxes pour un même modèle

Le graphe RDF est un modèle abstrait

- Visuel et compréhensible facilement par l'homme
- Intuitivement extensif et flexible

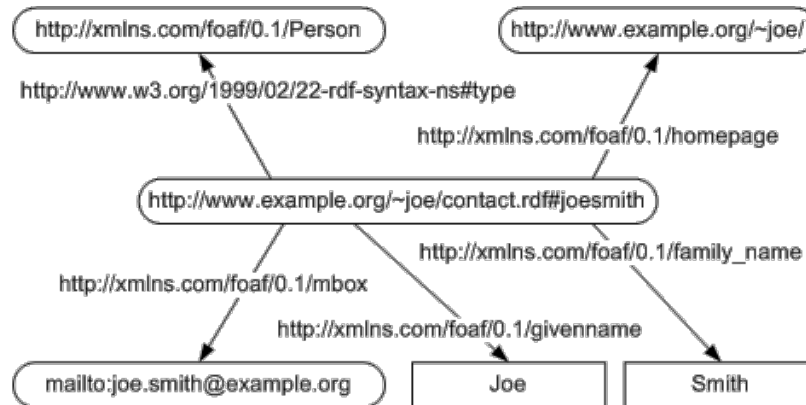
Une représentation machine nécessite une syntaxe concrète (sérialisation)

- **Différentes syntaxes** : RDF/XML, N-triples, Turtle, N3, JSON-LD
- Le code XML (arbre) d'un RDF (graphe) n'est pas destiné à être lu par l'utilisateur. Il constitue un format d'échange standardisé.



RDF – exemple de syntaxes concrètes

Graphe RDF (Syntaxe abstraite)



Différentes
implémentations

```
<http://www.example.org/~joe/contact.rdf#joesmith> <http://xmlns.com/foaf/0.1/homepage> <http://www.example.org/~joe/> .
<http://www.example.org/~joe/contact.rdf#joesmith> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://xmlns.com/foaf/0.1/Person> .
<http://www.example.org/~joe/contact.rdf#joesmith> <http://xmlns.com/foaf/0.1/family_name> "Smith" .
<http://www.example.org/~joe/contact.rdf#joesmith> <http://xmlns.com/foaf/0.1/givenname> "Joe" .
<http://www.example.org/~joe/contact.rdf#joesmith> <http://xmlns.com/foaf/0.1/mbox> <mailto:joe.smith@example.org> .
```

N-Triples

```
@prefix : <http://www.example.org/~joe/contact.rdf#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
```

```
:joesmith a foaf:Person ;
  foaf:family_name "Smith" ;
  foaf:givenname "Joe" ;
  foaf:homepage <http://www.example.org/~joe/> ;
  foaf:mbox <mailto:joe.smith@example.org> .
```

Turtle

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
>
  <foaf:Person rdf:about="http://www.example.org/~joe/contact.rdf#joesmith">
    <foaf:givenname>Joe</foaf:givenname>
    <foaf:homepage rdf:resource="http://www.example.org/~joe/" />
    <foaf:mbox rdf:resource="mailto:joe.smith@example.org" />
    <foaf:family_name>Smith</foaf:family_name>
  </foaf:Person>
</rdf:RDF>
```

RDF/XML

Autres: JSON-LD, N3, Microdata ...

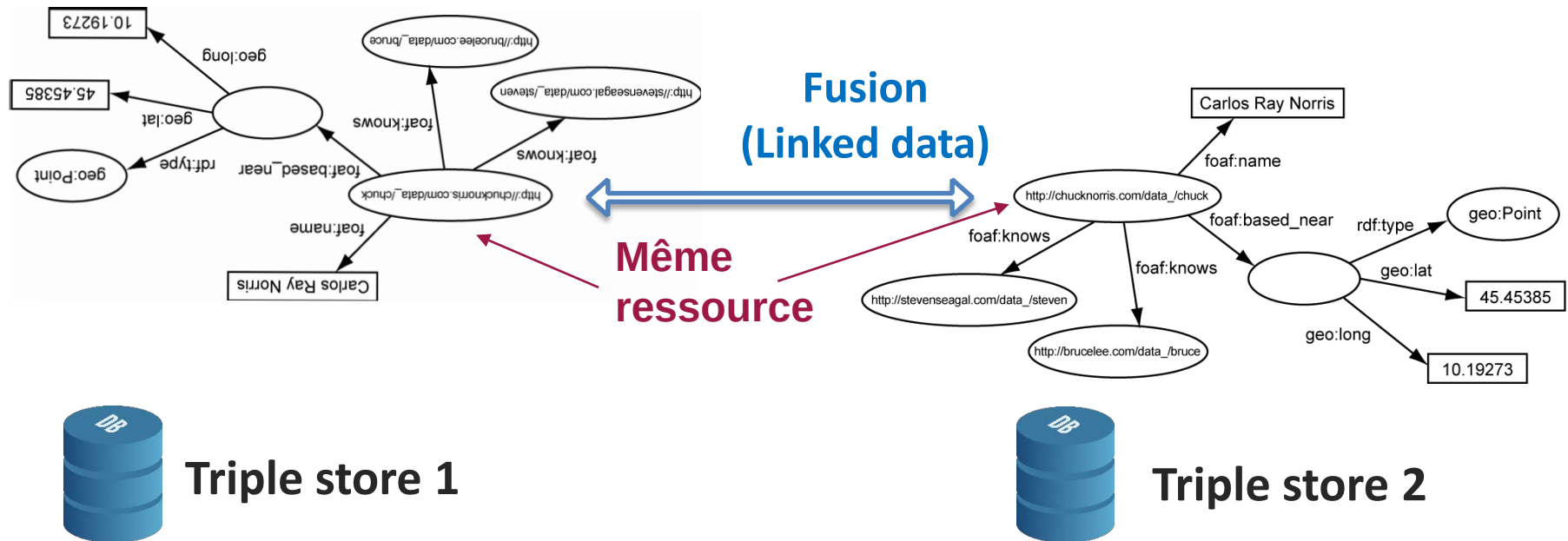
- <http://www.easyrdf.org/converter>
- <http://www.w3.org/wiki/ConverterToRdf>

1.c. Espace de nommages et vocabulaire

URI - un identifiant universel (1)

RDF fournit un cadre pour avoir une base de connaissances universelle

- Plusieurs **bases RDF (graphe RDF, Triple store)** peuvent avoir besoin de parler de la même chose.
- Il y a ainsi besoin d'identifier les ressources de manière universelle. On parle de **URI – Universal Resource Identifier**.

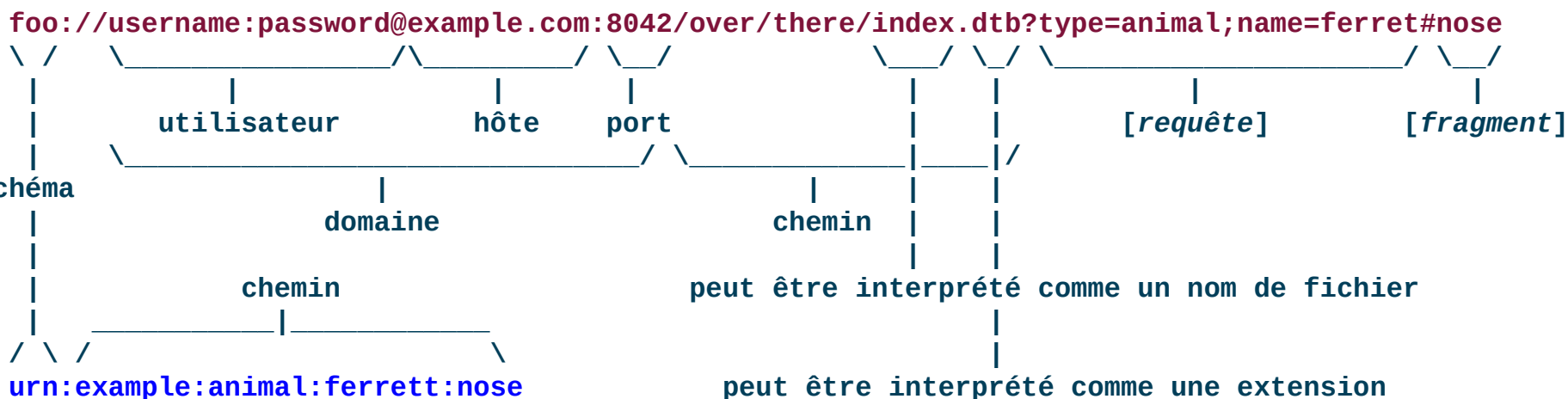


URI - un identifiant universel (2)

URI: Uniform Resource Identifier

- RFC 3986 (Berners-Lee et al., 2005) : décrit un schéma simple et extensif pour l'identification d'objets, abstraits ou physiques, à travers le monde.
- Séquence compacte de caractères suivant une syntaxe générique

<schéma>:<partie hiérarchique>[?<requête>][#<fragment>]



URI – Différents schémas

Un vaste panel enregistré par l'IANA



Internet Assigned Numbers Authority

- Différents schéma d'URI existent:
 - Des **schémas officiels**: file, ftp, dns, http, https, im, ldap, nfs, tel, ...
 - Des **schémas non officiels**: chrome, doi, ...
- Un schéma d'**URI n'est pas un protocole!** Mais en général, un schéma est associé à un **protocole** (http, ftp, irc, ...). Par exemple, le schéma **file** n'est pas associé à un **protocole**.
- **Exemple:**
 - **telnet**://<user>:<password>@<host>[:<port>/]
 - **file**://[host]/path
 - **irc**://<host>[:<port>]/[<channel>[?<password>]]
 - **lastfm**://<radio_stream>

URI & RDF – notion de vocabulaire (1)

‘Bordeaux’ : oui, mais dans quel vocabulaire ?



<http://www.insa-lyon.fr/if/ws/couleur#Bordeaux>

<http://www.insa-lyon.fr/if/ws/vin#Bordeaux>

<http://www.insa-lyon.fr/if/ws/geo#Bordeaux>

URI complet.

<http://www.insa-lyon.fr/if/ws/geo#bordeaux>

Espace de nommage (vocabulaire) – identifie le concept auquel on fait référence

Identifiant de la ressource

URI simplifié.

Utilisation des **Qualified names (qname)**:

qname:identifiant au lieu de **namespaceidentifiant**

Exemple: **geo** pour [http://www.insa-lyon.fr/if/ws/geo#](http://www.insa-lyon.fr/if/ws/geo#geo:Bordeaux)
geo:Bordeaux, **vin:Bordeaux**, **col:Bordeaux**

Remarque. Quand le vocabulaire est clair, on écrit: «**:Bordeaux**»

URI & RDF – notion de vocabulaire (2)

- De nombreux espaces standards
 - **xmnl:** **Définition d'espaces de nommages (i.e. qname)**
<http://www.w3.org/1999/xhtml>
 - **xsd:** **Définition de schémas XML**
<http://www.w3.org/2001/XMLSchema#>
 - **rdf:** **Pour définir des types et propriétés en RDF**
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
 - **rdfs:** **RDF Schema language (patience...)**
<http://www.w3.org/2000/01/rdf-schema#>
 - **owl:** **Web Ontology language (patience...)**
<http://www.w3.org/2002/07/owl#>
- **Remarque:** Déréférencer l'**URI de l'espace de nommage** pour comprendre le sens du vocabulaire ainsi que ses ressources standards.

URI & RDF – notion de vocabulaire (3)

Vocabulaire foaf (Friend Of A Friend):

Décrire les personnes et les relations entre amis.

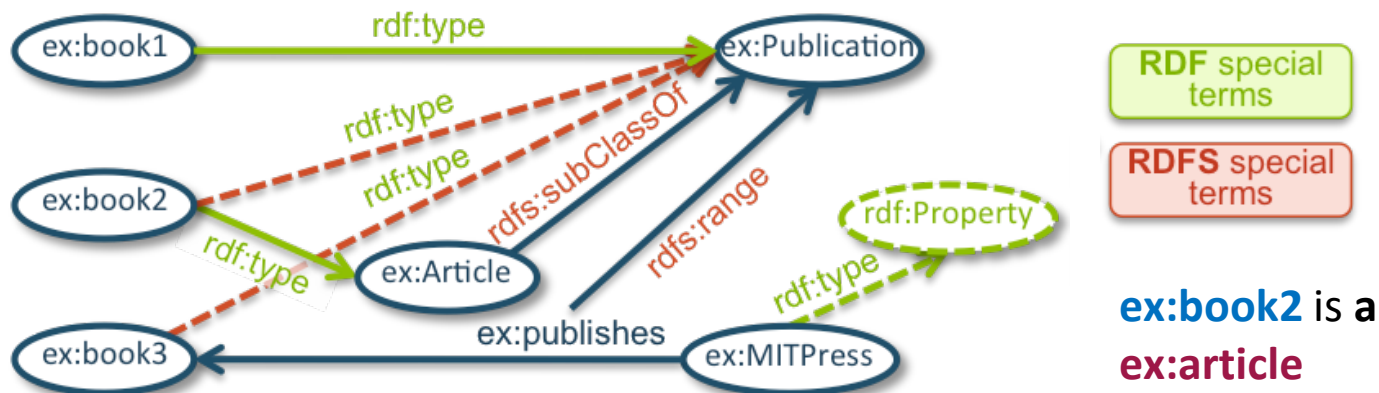
```
<?xml version="1.0" encoding="utf-8" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:dc="http://purl.org/dc/terms/">
  <foaf:Person rdf:about="https://liris.cnrs.fr/page-membre/adnene-belfodil">
    <foaf:name>Adnene Belfodil</foaf:name>
    <foaf:title>M.</foaf:title>
    <foaf:givenname>Adnene</foaf:givenname>
    <foaf:family_name>Belfodil</foaf:family_name>
    <foaf:nick>dino</foaf:nick>
    <foaf:dateOfBirth>1993-02-22</foaf:dateOfBirth>
    <foaf:homepage>
      <rdf:Description rdf:about="https://liris.cnrs.fr/page-membre/adnene-belfodil"
        <dc:title>Page personnelle</dc:title>
      </rdf:Description>
    </foaf:homepage>
    <foaf:img rdf:resource="https://.../private/photo_membre/belfodil.jpg"/>
    <foaf:schoolHomepage rdf:resource="http://www.insa-lyon.fr"/>
    <foaf:gender>male</foaf:gender>
  </foaf:Person>
</rdf:RDF>
```

Jouer avec <http://www.w3.org/RDF/Validator>

Remarque : Un autre vocabulaire souvent utilisé est **Dublin Core**

Vers les ontologies ...

- La primitive ***rdf:type*** permet d'associer un type à une ressource.
- **RDF est puissant** : La primitive ***rdf:type*** peut être impliquée dans plusieurs triplets avec le même sujet. On parle de multi-instanciation (e.g. $\neq$ Java).
- L'idée est de fournir les éléments de base pour définir des ontologies et des vocabulaires destinés à structurer les ressources RDF (**schéma RDF (RDFS)** et les **ontologies** en **Séance 4 du cours**). Pour le moment, nous nous limiterons à utiliser des vocabulaires simples et existants (**rdf, foaf, dublin core**).
- C'est aussi un point majeur afin de pouvoir définir un langage d'interrogation (requêtes SPARQL par exemple – **Séance 3 du cours**)



1.d. Les Littéraux

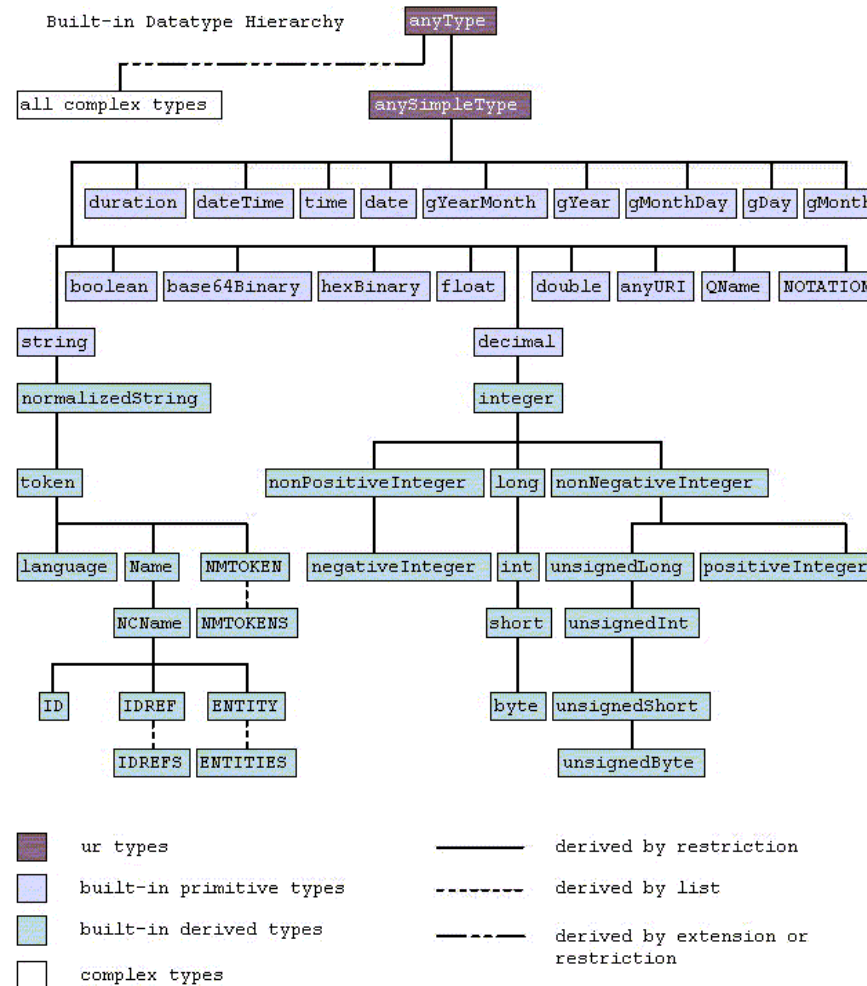
Age Group	Percentage
18-24	15%
25-34	20%
35-44	25%
45-54	20%
55-64	15%
65-74	10%
75-84	5%
85+	5%



Littéral (1)

- **Littéral** : Chaîne de caractère arbitraire **typée (voir XML Schema datatype)** ou **non typée** représentant des valeurs (texte, entiers, flottant, dates, ...). Exemple:
 - **Chaîne de caractères** : "Aimene Belfodil" ^^**xsd**:string.
 - **Entier** : "2017"^^**xsd**:integer.
 - **Spécification de la langue** : "Fiction pulpeuse"@fr, "Pulp fiction"@en.
- Tout ce qui est représenté par un **littéral** pourrait aussi être représenté par une **URI**, mais il est généralement plus adéquat ou intuitif d'employer des littéraux
- Un littéral peut être l'**objet** d'une **assertion RDF**, mais il ne peut en être ni le **sujet** ni le **prédicat**.
- **Rappel** : **xsd** est un espace standard pour la définition des schéma XML. Il réfère à: **<http://www.w3.org/2001/XMLSchema#>**

Littéral (2)



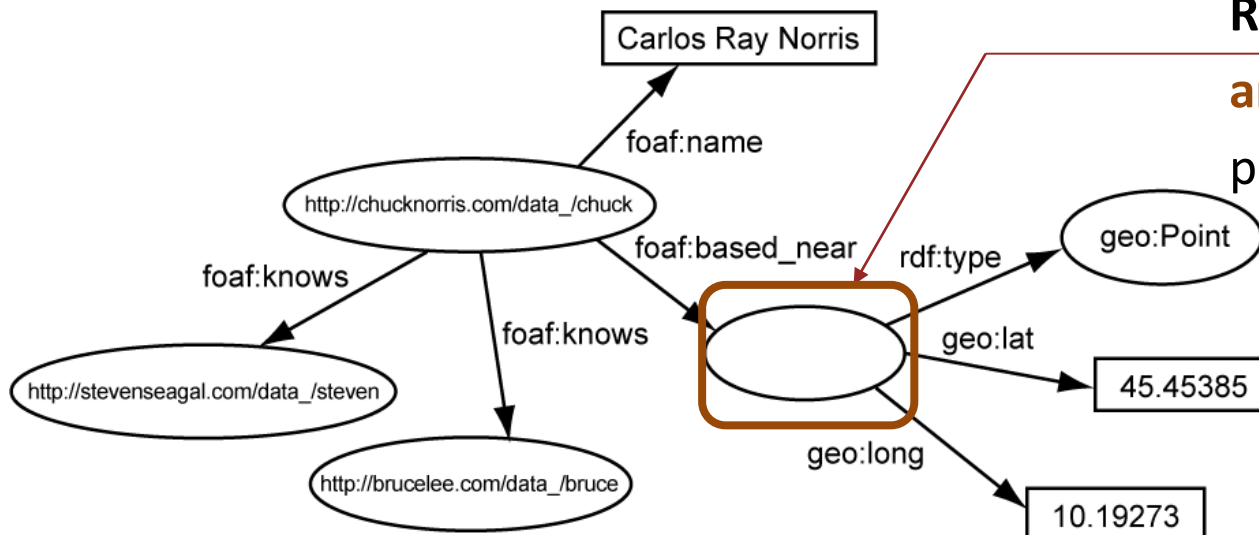
Types XML – XML Schema Datatype

<https://www.w3.org/TR/xmlschema-2/>

1.e Les nœuds anonymes et leurs utilités

Ressources anonymes

- L'existence d'une **ressource** peut être connue, sans que l'on puisse l'identifier par un **URI**. On parle de **ressource anonyme** (**nœud vide**, **blank-node**, **bnode**).
- Les **bnode** sont identifiés de manière **locale dans un document rdf** comme suit: **_:xyz** (préfixé par un trait de soulignement **_**)
- L'utilisation des **ressources anonymes est à éviter** sauf pour les **conteneurs** et les **collections**. Ils posent **beaucoup de problèmes** lorsque l'on désire fusionner deux graphes RDF (i.e. deux ressources anonymes dans deux documents RDF différents sont considérées comme différentes même si ils ont le même **identifiant**!).



Rappel. Une **ressource anonyme** est représentée par un **nœud non étiqueté**

RDF - les conteneurs [ouverts] (1)

Conteneurs de ressources/littéraux

- **Sac (rdf:Bag)** : non ordonné, doublons (e.g. contenu d'un colis)
- **Séquence (rdf:Seq)** : ordonné, doublons (e.g. vainqueurs de la FIFA WC depuis 1930)
- **Alternative (rdf:Alt)** : ensemble de valeurs dont une seule peut être tirée selon son contexte d'utilisation (e.g. les titres d'un seul et même livre en différentes langues)

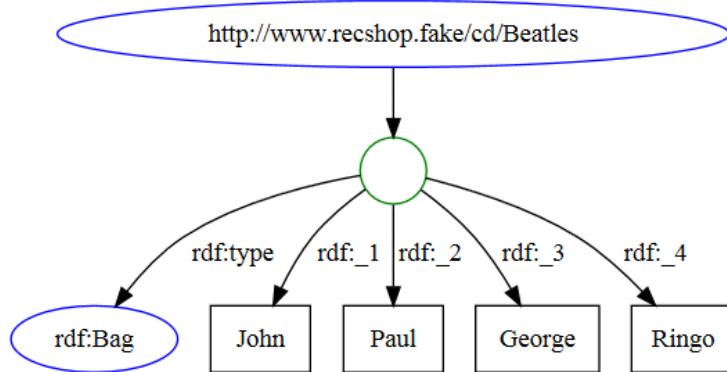
Utilisation des conteneurs

- Une ressource (souvent anonyme, le conteneur) est créée, et **typée** (en utilisant **rdf:type**) comme **sac**, **séquence** ou **alternative**
- Les éléments du conteneur lui sont attachés en utilisant **rdf:li**
- Chaque occurrence de **rdf:li** sera transformée en **rdf:_1**, **rdf:_2**, etc.

Ces conteneurs sont ouverts

- On ne peut pas affirmer que l'ensemble est complet et qu'il n'existe pas d'autres éléments ailleurs

RDF - les conteneurs [ouverts] (2)



```

<?xml version="1.0" encoding="utf-8" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:ns0="http://www.recshop.fake/cd#">

  <rdf:Description rdf:about="http://www.recshop.fake/cd/Beatles">
    <ns0:artist>
      <rdf:Bag>
        <rdf:li>John</rdf:li>
        <rdf:li>Paul</rdf:li>
        <rdf:li>George</rdf:li>
        <rdf:li>Ringo</rdf:li>
      </rdf:Bag>
    </ns0:artist>

  </rdf:Description>

</rdf:RDF>
  
```

RDF/XML

```

@prefix ns0: <http://www.recshop.fake/cd#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

<http://www.recshop.fake/cd/Beatles> ns0:artist [
  a rdf:Bag ;
  rdf:_1 "John" ;
  rdf:_2 "Paul" ;
  rdf:_3 "George" ;
  rdf:_4 "Ringo"
] .
  
```

Turtle

```

<http://www.recshop.fake/cd/Beatles> <http://www.recshop.fake/cd#artist> _:genid1 .
_:genid1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://www.w3.org/1999/02/22-rdf-syntax-ns#Bag> .
_:genid1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#_1> "John" .
_:genid1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#_2> "Paul" .
_:genid1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#_3> "George" .
_:genid1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#_4> "Ringo" .
  
```

N-Triples

RDF - les collections [fermées] (1)

Collections fermées

- Ces groupes ne contiennent que les membres spécifiés lors de leur déclaration.

Syntaxe abstraite:

- Les collections fermées sont des **listes chaînées** de membres : **rdf:list** lie une ressource à sa collection (ressource anonyme) qui est liée à une ressource tête (**rdf:first**) et aux restes du groupe (**rdf:rest**). **rdf:nil** est utilisé en fin de liste.

Représente la Même collection

```
<rdf:Description rdf:about="http://e.org/family/349">
  <s:familyMembers rdf:parseType="Collection">
    <rdf:Description rdf:about="http://e.org/person/Paolo"/>
    <rdf:Description rdf:about="http://e.org/person/Emanuele"/>
    <rdf:Description rdf:about="http://e.org/person/Maria"/>
    <rdf:Description rdf:about="http://e.org/person/Franco"/>
  </s:familyMembers>
</rdf:Description>
```

```
<rdf:Description rdf:about="http://e.org/family/349">
  <s:familyMembers rdf:nodeID="sch1"/>
</rdf:Description>

<rdf:Description rdf:nodeID="sch1">
  <rdf:first rdf:resource="http://e.org/person/Paolo"/>
  <rdf:rest rdf:nodeID="sch2"/>
</rdf:Description>

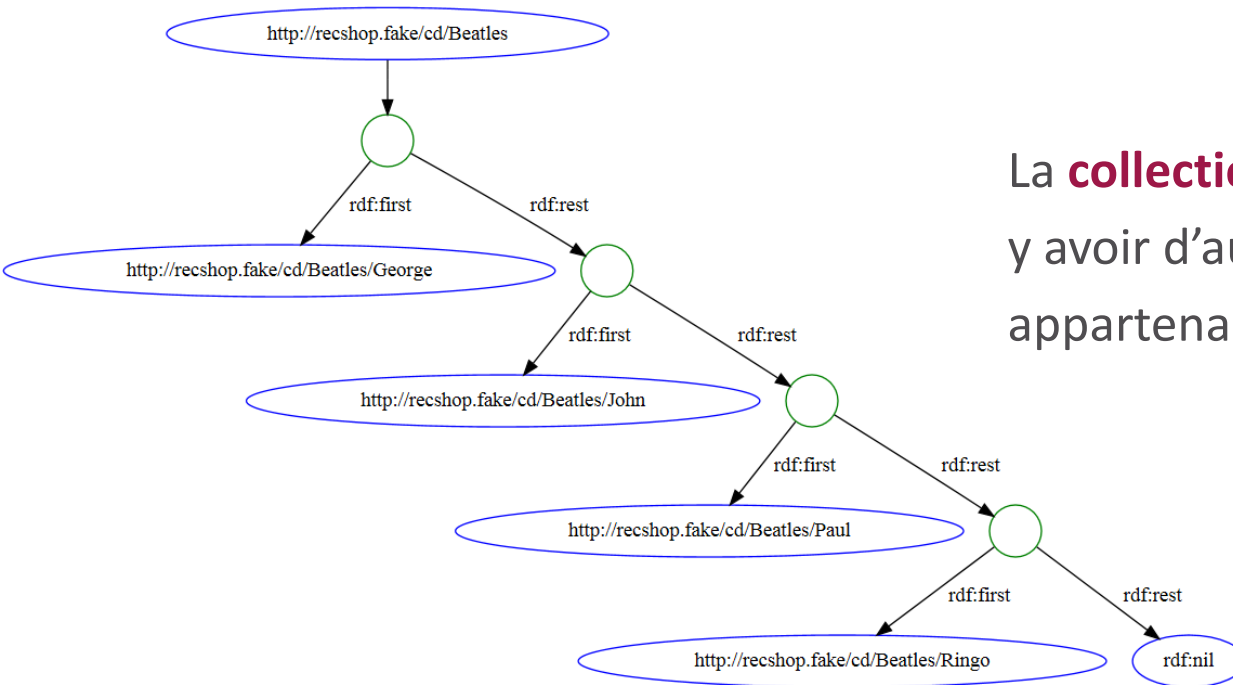
<rdf:Description rdf:nodeID="sch2">
  <rdf:first rdf:resource="http://e.org/person/Emanuele"/>
  <rdf:rest rdf:nodeID="sch3"/>
</rdf:Description>

<rdf:Description rdf:nodeID="sch3">
  <rdf:first rdf:resource="http://e.org/person/Maria"/>
  <rdf:rest rdf:nodeID="sch4"/>
</rdf:Description>

<rdf:Description rdf:nodeID="sch4">
  <rdf:first rdf:resource="http://e.org/person/Franco"/>
  <rdf:rest rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#nil"/>
</rdf:Description>
```

Remarque: Certaines **syntaxes concrètes** fournissent des **implémentations simplifiées**. RDF/XML fournit: **rdf:parseType="Collection"**.

RDF - les collections [fermées] (2)



La **collection étant fermée**, il ne peut y avoir d'autres membres ailleurs appartenant à cette collection.

N-Triples: implémentation simplifiée

```
@prefix ns0: <http://recshop.fake/cd#> .

<http://recshop.fake/cd/Beatles> ns0:artist (
  <http://recshop.fake/cd/Beatles/George>
  <http://recshop.fake/cd/Beatles/John>
  <http://recshop.fake/cd/Beatles/Paul>
  <http://recshop.fake/cd/Beatles/Ringo>
) .
```

N-Triples: pas d'implémentation simplifiée

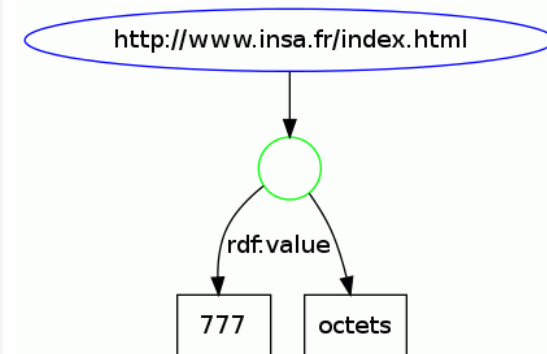
```
<http://recshop.fake/cd/Beatles> <http://recshop.fake/cd#artist> _:genid1 .
_:genid1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#first> <http://recshop.fake/cd/Beatles/George> .
_:genid1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#rest> _:genid2 .
_:genid2 <http://www.w3.org/1999/02/22-rdf-syntax-ns#first> <http://recshop.fake/cd/Beatles/John> .
_:genid2 <http://www.w3.org/1999/02/22-rdf-syntax-ns#rest> _:genid3 .
_:genid3 <http://www.w3.org/1999/02/22-rdf-syntax-ns#first> <http://recshop.fake/cd/Beatles/Paul> .
_:genid3 <http://www.w3.org/1999/02/22-rdf-syntax-ns#rest> _:genid4 .
_:genid4 <http://www.w3.org/1999/02/22-rdf-syntax-ns#first> <http://recshop.fake/cd/Beatles/Ringo> .
_:genid4 <http://www.w3.org/1999/02/22-rdf-syntax-ns#rest> <http://www.w3.org/1999/02/22-rdf-syntax-ns#nil> .
```

Exemple - Les valeurs complexes

Des bonnes pratiques (<http://www.w3.org/2001/sw/BestPractices/>)

- Toute relation n-aire doit être décomposée.
- **Exemple** : un document d'une taille de 777 octets

```
@prefix ns0: <http://www.insa.fr/voc#> .  
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .  
@prefix ns1: <http://www.insa.fr/voc#> .  
  
<http://www.insa.fr/index.html> ns0:taille [  
  rdf:value "777" ;  
  ns1:unit "octets"  
] .
```



2. Publication

RDFa – RDF in Attributes (1)

- Une **syntaxe pour injecter du RDF** dans une page Web et notamment indiquer et permettre l'extraction de données présentes dans une page.
- Une page Web peut être ainsi à la **fois la représentation d'un document visualisable** dans un navigateur et une **source de données sémantiques**.
- Utilise des attributs (X)HML et d'autres sont ajoutés

Existant

<i>rel</i>	<i>inv</i>	Liste de propriétés séparées par des espaces Utilisé pour exprimer des relations entre des ressources
<i>content</i>		Valeur littérale d'une propriété
<i>href</i>	<i>src</i>	URI d'une ressource utilisée comme objet d'une propriété

Nouveaux

<i>about</i>	Le sujet d'une propriété
<i>property</i>	Une liste de propriétés ayant une valeur littérale
<i>resource</i>	URI d'une ressource utilisée comme objet d'une propriété ($\neq$ href)
<i>datatype</i>	Permet de spécifier le type de données littérales
<i>typeof</i>	Une liste de types pour une ressource séparés par des espaces

RDFa – RDF in Attributes (2)

```
<?xml version="1.0" encoding="UTF-8"?>
<html
  xmlns="http://www.w3.org/1999/xhtml"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  version="XHTML+RDFa 1.0"
  xml:lang="en">
  <head>
    <meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
    <style type="text/css"> ... </style>
    <title>Homepage of Mehdi Kaytoue</title>
    <base href="https://perso.liris.cnrs.fr/mehdi.kaytoue/">
    <meta property="dc:creator" content="Mehdi Kaytoue"/>
  </head>
  <body about="https://perso.liris.cnrs.fr/mehdi.kaytoue/">
    <h1>Homepage of Mehdi Kaytoue</h1>
    <span property="foaf:title">Dr.</span>
    <span property="foaf:givenName">Mehdi</span>
    <span property="foaf:familyName">Kaytoue</span>
    <br>
    Institut National des Sciences Appliquées de Lyon<br>
    Département Informatique<br>
    LIRIS CNRS UMR 5205 - Team DM2L<br>
    Bâtiment Blaise Pascal<br>
    F-69621 Villeurbanne cedex, France<br>
    <span property="foaf:mbox">mehdi.kaytoue@insa-lyon.fr</span><br>
    <br>
    
    ...
  </body>
</html>
```

Ici, on utilise les
vocabulaires **dc** et **foaf**

https://perso.liris.cnrs.fr/me...

Applications Bookmarks INSA IF - Choix de l'emploi Autres favoris

Homepage of Mehdi Kaytoue

Dr. Mehdi Kaytoue
Institut National des Sciences Appliquées de Lyon
Département Informatique
LIRIS CNRS UMR 5205 - Team DM2L
Bâtiment Blaise Pascal
F-69621 Villeurbanne cedex, France
mehdi.kaytoue@insa-lyon.fr



This page is an extended CV under construction, you can have complementary information here:

- LinkedIn: <https://www.linkedin.com/in/mehdikaytoue>
- DBLP: <http://dblp.uni-trier.de/pers/hd/k/Kaytoue=Uberall:Mehdi>
- Google Scholar: <http://scholar.google.com/citations?user=TXvIWggAAAAJ>

Triplets RDF extraits avec
<https://rdfa.info/play/>

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
<https://perso.liris.cnrs.fr/mehdi.kaytoue/>
  <http://purl.org/dc/elements/1.1/creator> "Mehdi Kaytoue";
  foaf:title "Dr.";
  foaf:givenName "Mehdi";
  foaf:familyName "Kaytoue";
  foaf:mbox "mehdi.kaytoue@insa-lyon.fr";
  foaf:depiction <https://perso.liris.cnrs.fr/mehdi.kaytoue/small.jpg> .
```

3. Accès aux données

Déréférencement des ressources

Déréférencement d'une URI :

- Acte de retrouver la représentation de la ressource qu'il identifie.
- Rappel : on distingue deux types de ressources (et donc d'**URI**):
 - **Ressource de type 'information'** : la ressource est déréférencée directement et on récupère une représentation du document (e.g. un **code html** à afficher). On parle de **URL (Uniform Resource Locator)**.
 - **Ressource de type 'non-information'** : on récupère une description RDF de la ressource (RDF/XML), une **page HTML** (une autre **URL**), ou autres ... à la suite d'une étape de **négociation de contenu**. On parle de **URN (Uniform Resource Name)**.

Une **URN** est plutôt utilisé pour **identifier de manière pérenne** une ressource dans le web (ex: **isbn** d'un livre) alors qu'une **URL** fournit plutôt l'**adresse** pour obtenir cette ressource

Déréférencement des sources de type informa

Example:

- Page **Wikipédia** sur **Red Dead Redemption 2** (HTML) :

```
curl https://en.wikipedia.org/wiki/Red_Dead_Redemption_2
```

```
</div>
<!-- toclevel-1 tocsection-9 --><a href="https://en.wikipedia.org/wiki/Red_Dead_Redemption_2#Notes"><span class="tocnumber">6</span><span class="toctext">No
</a>
<!-- toclevel-1 tocsection-10 --><a href="https://en.wikipedia.org/wiki/Red_Dead_Redemption_2#References"><span class="tocnumber">7</span><span class="toctext">Refe
</a>
<!-- toclevel-1 tocsection-11 --><a href="https://en.wikipedia.org/wiki/Red_Dead_Redemption_2#External_links"><span class="tocnumber">8</span><span class="toctext">Ext
</a>
</div>

<h2><span class="my-headline" id="Gameplay">Gameplay</span></h2>
<div class="thumb tleft"><div class="thumbinner" style="width:252px;"><a href="https://en.wikipedi
<p><i>Red Dead Redemption 2</i> is a <a href="https://en.wikipedia.org/wiki/Western_genre" title
</p><p>The game will continue to use the Honor system from its predecessor, <i><a href="https://en
</p><p>The game's main story is expected to take 60 hours to complete.<sup id="cite_ref-11" class=
</p>
<h2><span class="my-headline" id="Premise">Premise</span></h2>
<p>In 1899, the era of the <a href="https://en.wikipedia.org/wiki/American_frontier" title="Americ
</p><p>Among the returning characters are John Marston and Dutch Van der Linde, the protagonist and
</p>
<h2><span class="my-headline" id="Development">Development</span></h2>
<p><i>Red Dead Redemption 2</i> is being developed by <a href="https://en.wikipedia.org/wiki/Rockst
</p>
<h3><span class="my-headline" id="Crunch_time_controversy">Crunch time controversy</span></h3>
<p>In an interview with Harold Goldberg of magazine <i><a href="https://en.wikipedia.org/wiki/New
</p>
<h2><span class="my-headline" id="Release">Release</span></h2>
<p><i>Red Dead Redemption 2</i> was announced by publisher <a href="https://en.wikipedia.org/wiki/E
</p><p>The game's debut trailer was released on October 20, 2016, depicting the open world environm
</p><p><An <a href="https://en.wikipedia.org/wiki/Online_multiplayer" class="my-redirect" title="On
</p>
<h2><span class="my-headline" id="Reception">Reception</span></h2>
<h3><span class="my-headline" id="Sales">Sales</span></h3>
<p>Since the previous installment in the series was among the bestselling games of the <a href="ht
```



WIKIPEDIA
The Free Encyclopedia

Not logged in | [Talk](#) | [Contributions](#) | [Create account](#) | [Log in](#)

[Article](#)
[Talk](#)

[Read](#)
[View source](#)
[View history](#)

Red Dead Redemption 2

From Wikipedia, the free encyclopedia

[Main page](#)
[Contents](#)
[Featured content](#)
[Current events](#)
[Random article](#)
[Donate to Wikipedia](#)
[Wikipedia store](#)

[Interaction](#)
[Help](#)
[About Wikipedia](#)
[Community portal](#)
[Recent changes](#)
[Contact page](#)

[Tools](#)

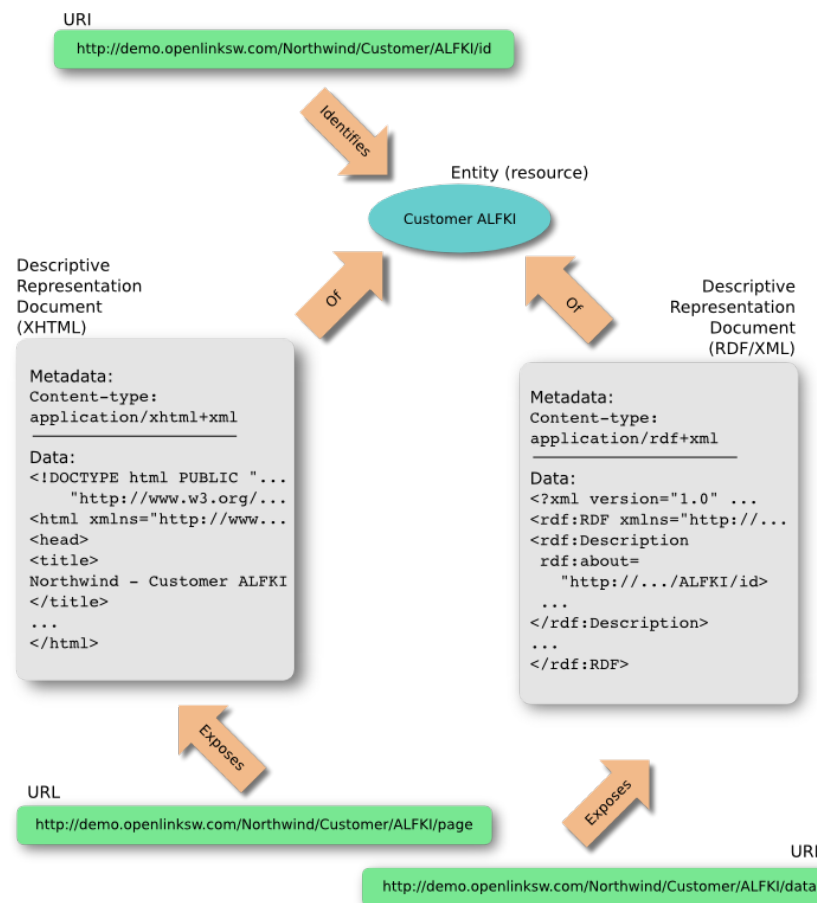
Red Dead Redemption 2^[b] is an upcoming [Western](#)-themed [action-adventure](#) video game developed and published by [Rockstar Games](#). It is scheduled to be released for [PlayStation 4](#) and [Xbox One](#) on October 26, 2018. The game is a [prequel](#) to the 2010 game [Red Dead Redemption](#), and will be the third entry in the [Red Dead](#) series. Set in 1899, it follows the story of outlaw Arthur Morgan, a member of the Dutch Van der Linde gang.



[Contents](#) [hide]

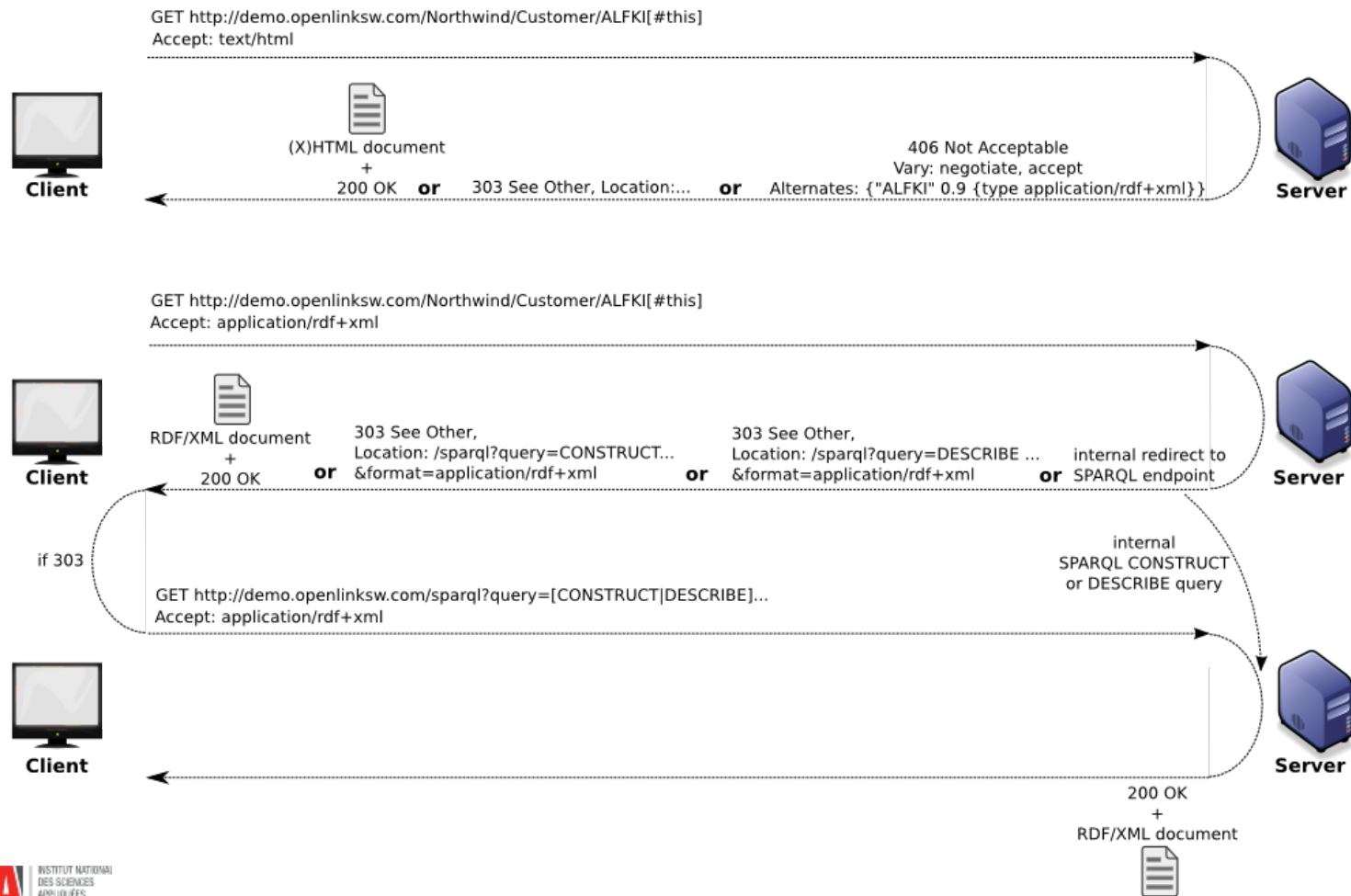
Déréférencement des ressources de type non information (1)

Différentes représentations : Une même **URI** peut avoir en effet plusieurs représentations (plusieurs **URL**) : Page **HTML**, Fichier **RDF/XML**, ...



Déréférencement des ressources de type non information (2)

Négociation de contenu: Pour avoir une des représentations possibles, il faut faire une négociation de contenu.



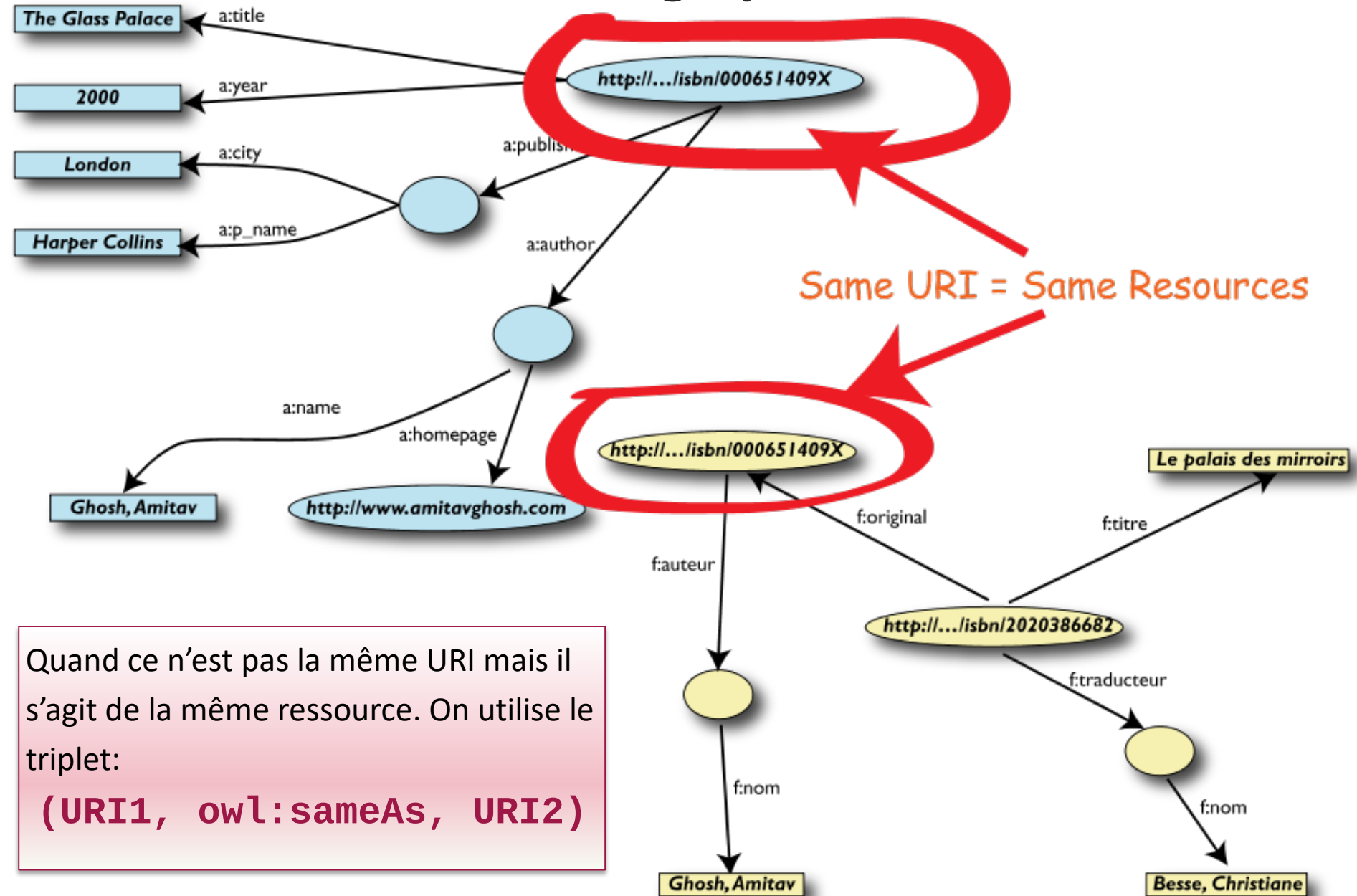
Déréférencement des ressources de type non information (3)

Exemple :

- On demande une représentation **HTML** :  Suivre la redirection
`curl -H "Accept: text/html" -L`
`http://dbpedia.org/resource/Red\_Dead\_Redemption\_2`
 - On est redirigé (**303 Redirect**) vers la page:
`http://dbpedia.org/page/Red\_Dead\_Redemption\_2`
-
- On demande une représentation **RDF/XML**:
`curl -H "Accept: application/rdf+xml" -L`
`http://dbpedia.org/resource/Red\_Dead\_Redemption\_2`
 - On est redirigé (**303 Redirect**) vers la page:
`http://dbpedia.org/data/Red\_Dead\_Redemption\_2`

4. Intégration de données

Fusion de graphes RDF



RDF 1.1 - Named graphs

Recommandation W3C: <https://www.w3.org/TR/rdf11-concepts/>

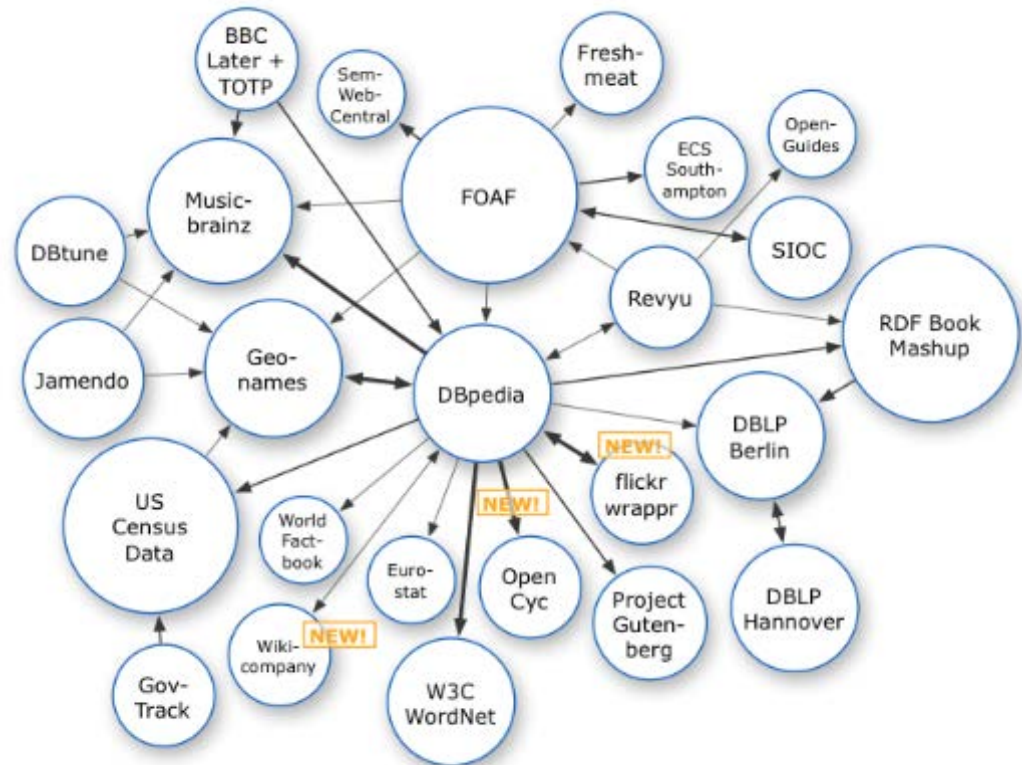
- Données séparées mais facilement intégrables
- Ajout de la notion de **Graphes nommés** dans RDF 1.1. Un **graphe nommé RDF** est tout simplement un ensemble de triplets RDF.
- **Syntaxes concrètes:**
 - **triG** extension de **Turtle** pour supporter la notion de graphe.
 - **N-quads** extension de **N-Triples** pour supporter la notion de graphe en utilisant une suite de **Quadruplets** (graphe, sujet, prédicat, objet).

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .  
@prefix inria: <http://inria.fr/schema#> .  
GRAPH <http://inria.fr/people>  
{ <http://inria.fr/rr/doc.html>  
  inria:author <http://ns.inria.fr/catherine.faron#me> .  
}  
GRAPH <http://inria.fr/topics>  
{ <http://inria.fr/rr/doc.html> inria:topic "Web of Data" . }
```

Syntaxe triG

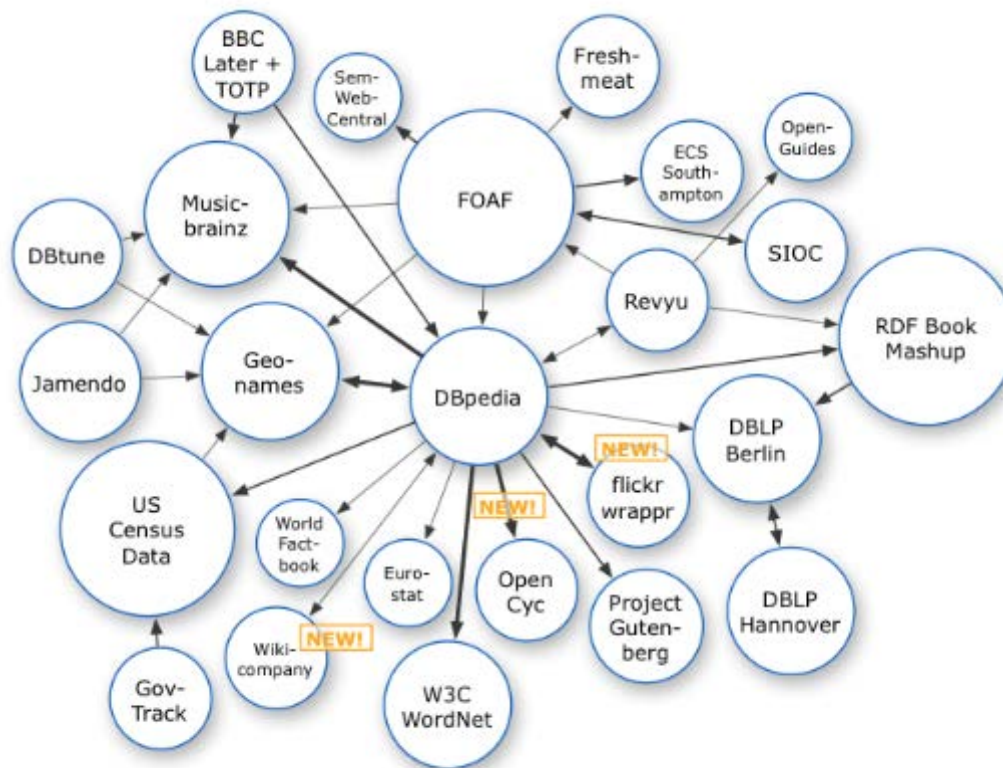
LOD - Linking Open Data (1)

- Constituer une base de connaissance universelle
- Divers domaines concernés
 - publique (compagnies)
 - scientifiques
 - gouvernements



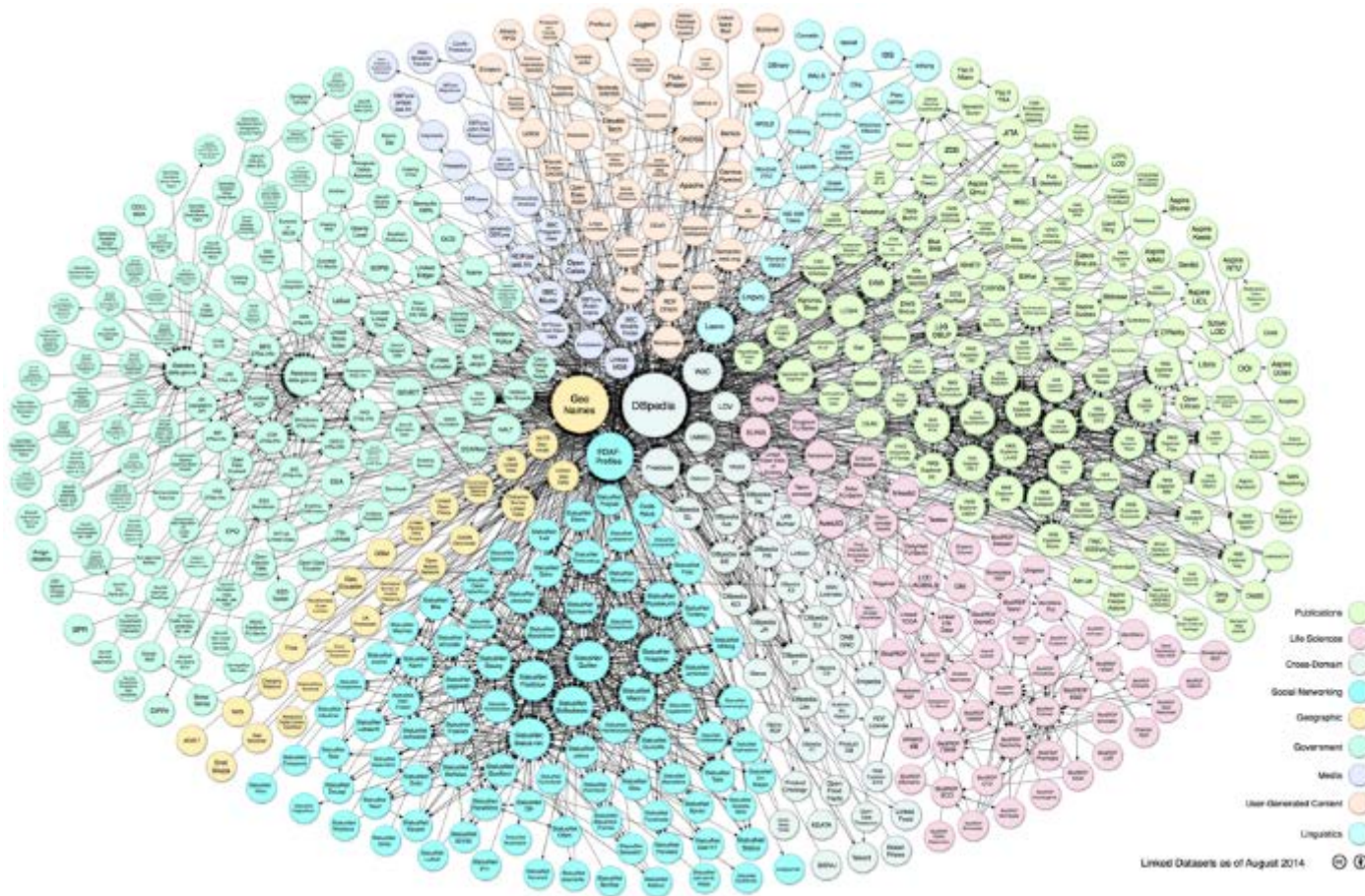
LOD - Linking Open Data (2)

- **En Octobre 2007 : 2 milliards** de triplets RDF et **12** bases liées.



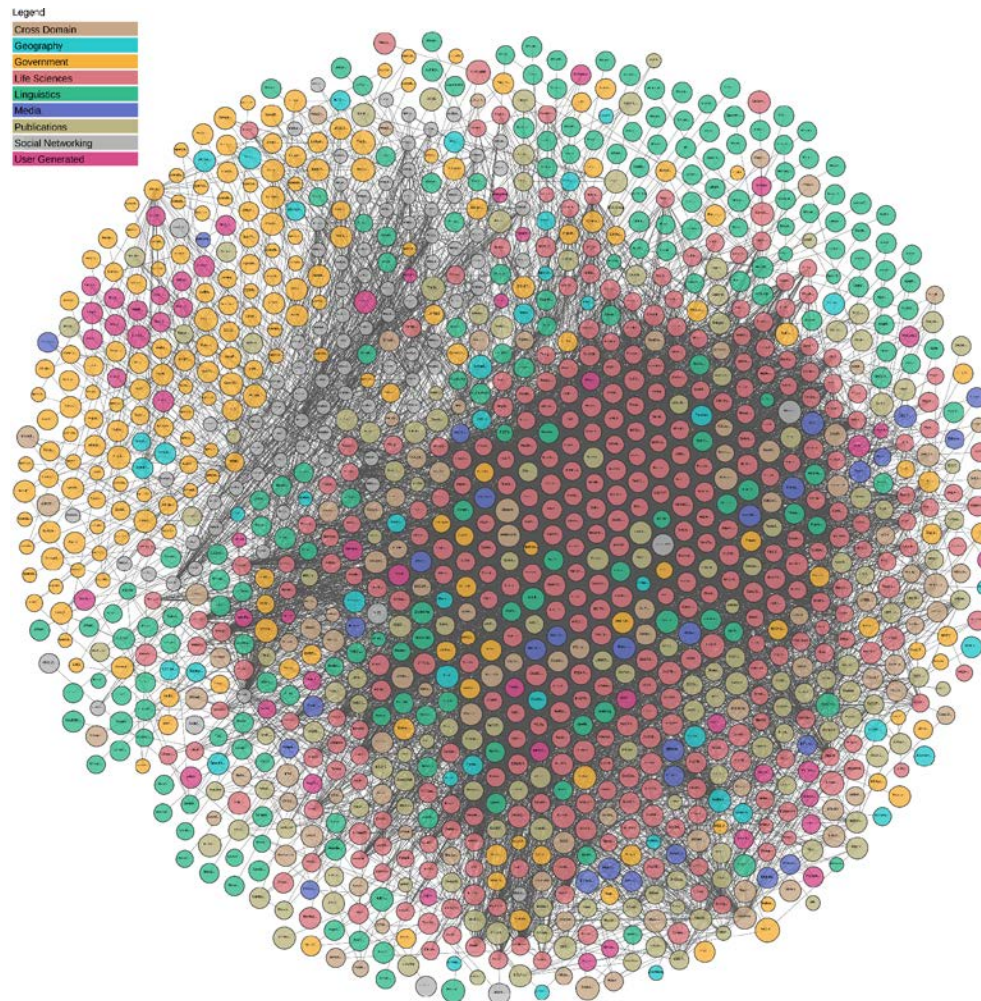
LOD - Linking Open Data (3)

- En septembre 2011: **31** milliards de triplets RDF et **295** bases liées.



LOD - Linking Open Data (4)

- **En Juillet 2018: 1220** bases liées.



Epilogue

Faisons le point

- **RDF** : un cadre de modélisation de données dans le **Web 3.0**
- **Triplet RDF** : atome d'information : (**sujet**, **prédicat**, **objet**) avec **sujet** et **prédicat** sont des **ressources identifiées** par un **URI** et **objet** pouvant être une **ressource** ou un **littéral**.
- **Modèle simple** : Triplets/Graphes. On parle de **Triple store**.
- **Infrastructure du Web** : Utilisation des URI et espace de nommage
- **Syntaxes concrètes de RDF** indépendantes du **modèle abstrait**
- **Publication facile de données RDF**: via RDFa par exemple.
- Notion de **déréférencement** de ressources et **négociation de contenu** (Distinction entre **URI (IRI)**, **URL** et **URN**)
- **Fusion des données aisée** : agrégation de triplets.
- **Slogan AAA** : **A**nycne can say **A**nycnng about **A**nycnng.
- **Monde ouvert** et **logique monotone**.

Séance 3 - Interroger avec SPARQL (1)

The Semantic Web Technology Stack
(not a piece of cake...)

Most apps use only a subset of the stack

Querying allows fine-grained data access

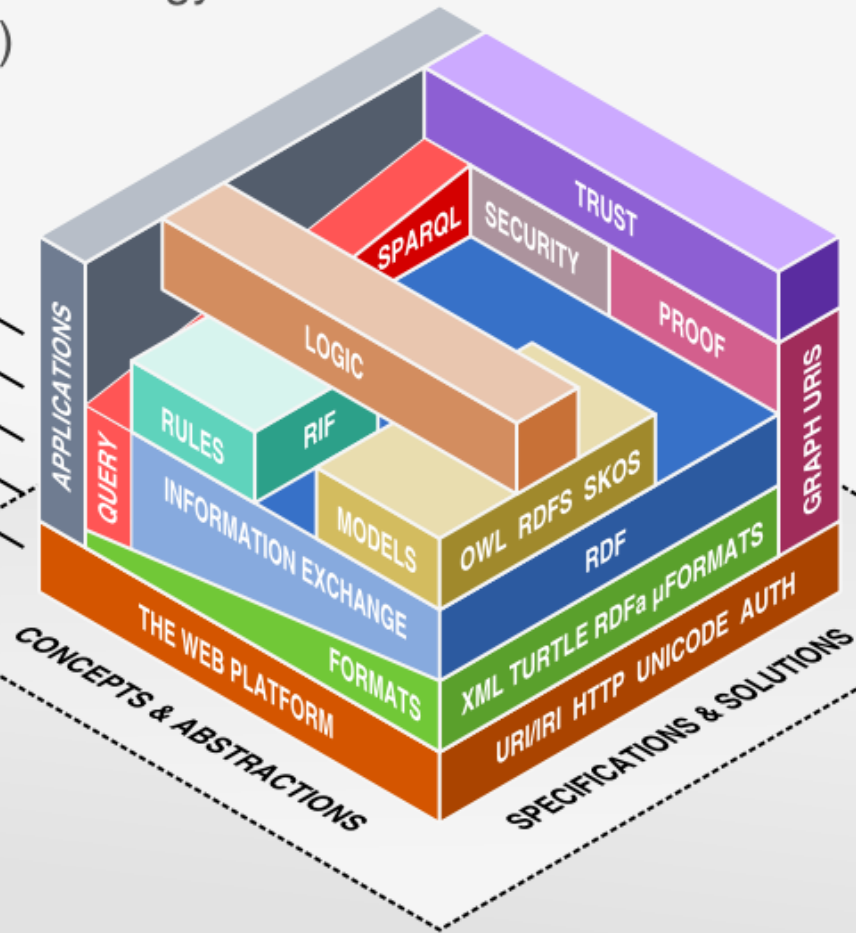
Standardized information exchange is key

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



Séance 3 - Interroger avec SPARQL (2)

- Nous avons vu comment constituer des **bases de données RDF**.
- **SPARQL (SPARQL Protocol And RDF Query Language)** est un langage de requêtes qui permet d'interagir avec les bases de données RDF.
- **Exemple: Quelles sont les communes d'île de France >100 000 habitants?**

Requête SPARQL:

```
prefix db-owl: <http://dbpedia.org/ontology/>
select * where {
  ?ville db-owl:region <http://fr.dbpedia.org/resource/Île-de-France> .
  ?ville rdf:type db-owl:Settlement .
  ?ville db-owl:populationTotal ?population .
  FILTER (?population > "100000"^^xsd:long)
}
```

ville	population
http://fr.dbpedia.org/resource/Saint-Denis_(Seine-Saint-Denis)	106 785
http://fr.dbpedia.org/resource/Boulogne-Billancourt	114 205
http://fr.dbpedia.org/resource/Montreuil_(Seine-Saint-Denis)	102 770
http://fr.dbpedia.org/resource/Argenteuil_(Val-d'Oise)	103 125
http://fr.dbpedia.org/resource/Paris	2 243 833

Courte biblio-webographie

- Harald Sack. **Semantic Web technologies**. Open course Hasso Plattner Institut (2013) : <https://open.hpi.de/course/semanticweb>
- Gandon F., Faron-Zucker C., Corby O. **Le Web sémantique. Comment lier les données et les schémas sur le Web ?** Dunod eds. 2012. (et MOOC INRIA correspondant).
- Allemang D., Hendler J. **Semantic Web for the working ontologist : Effective modelling in RDFS and OWL**. Morgan Kaufmann eds. 1st edition (2008)
- Cours de F. Badra (fadi.lautre.net) et références y figurant
- Actes annuels des conférences **WWW, ISWC, ESWC**:
[voir www.informatik.uni-trier.de/~ley/db](http://www.informatik.uni-trier.de/~ley/db)

*Les **recommandations** et **RFC** sont d'**excellentes sources** !*

- <https://www.w3.org/TR/rdf11-primer/>
- <https://www.w3.org/TR/rdf11-concepts/>

