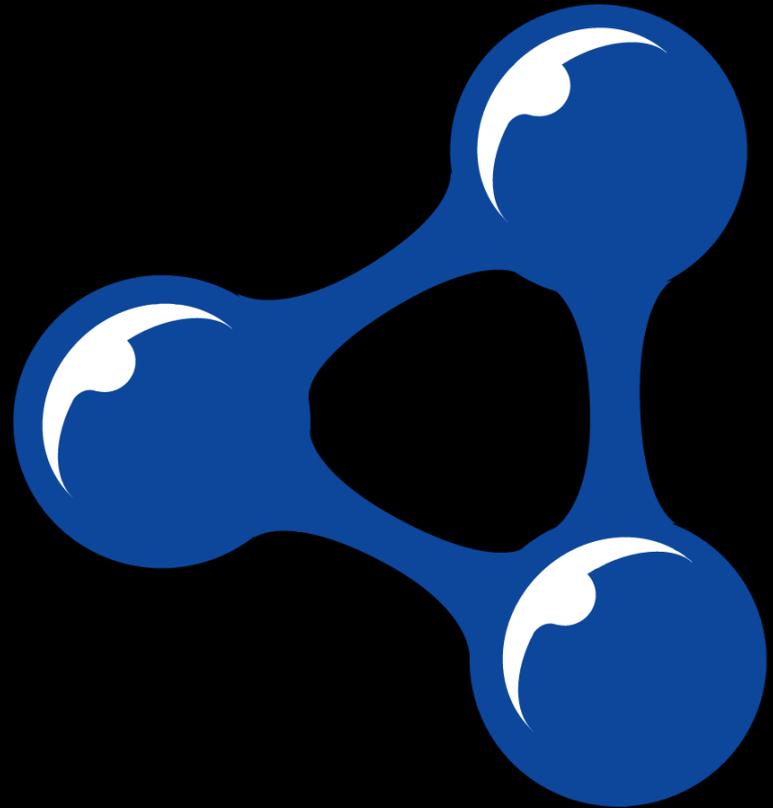
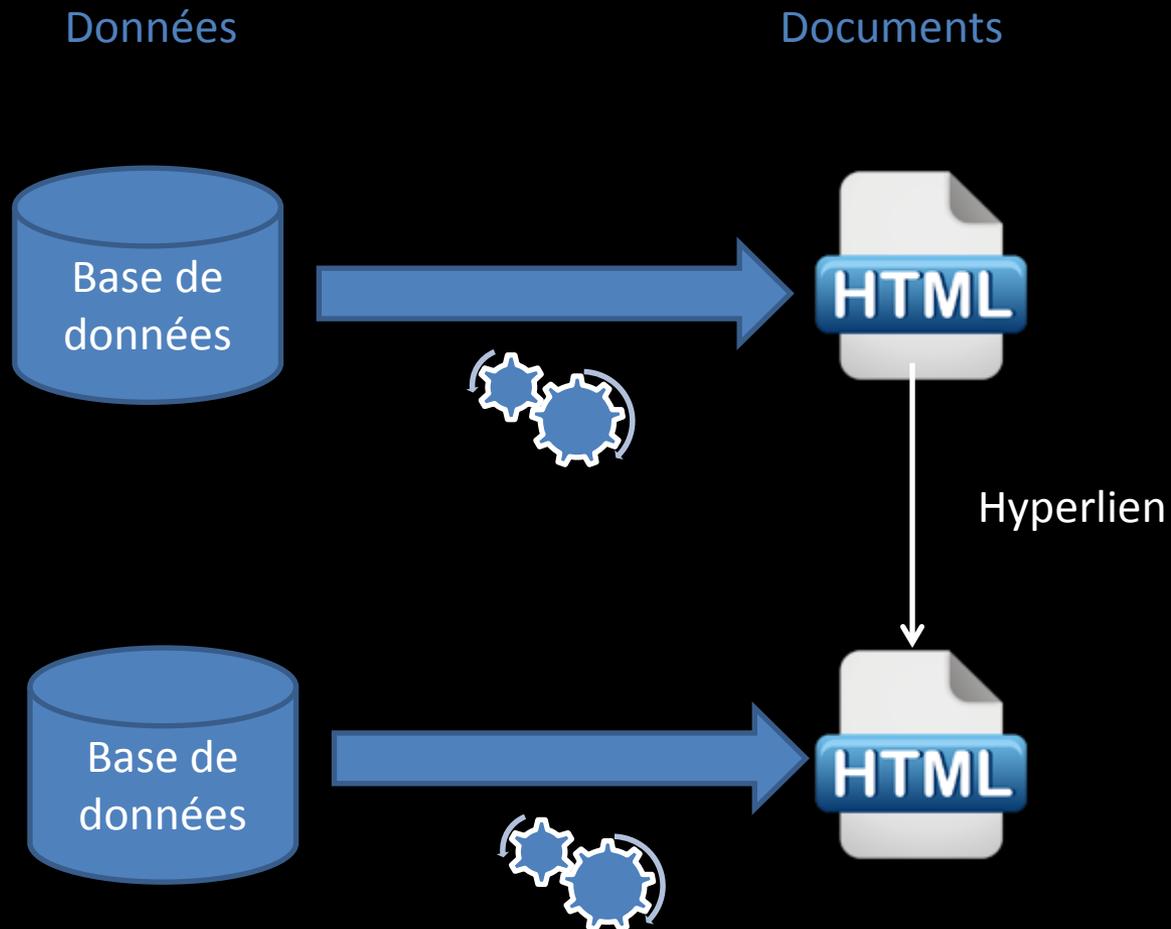


# Le web de données et les bibliothèques



# Le Web, web de documents

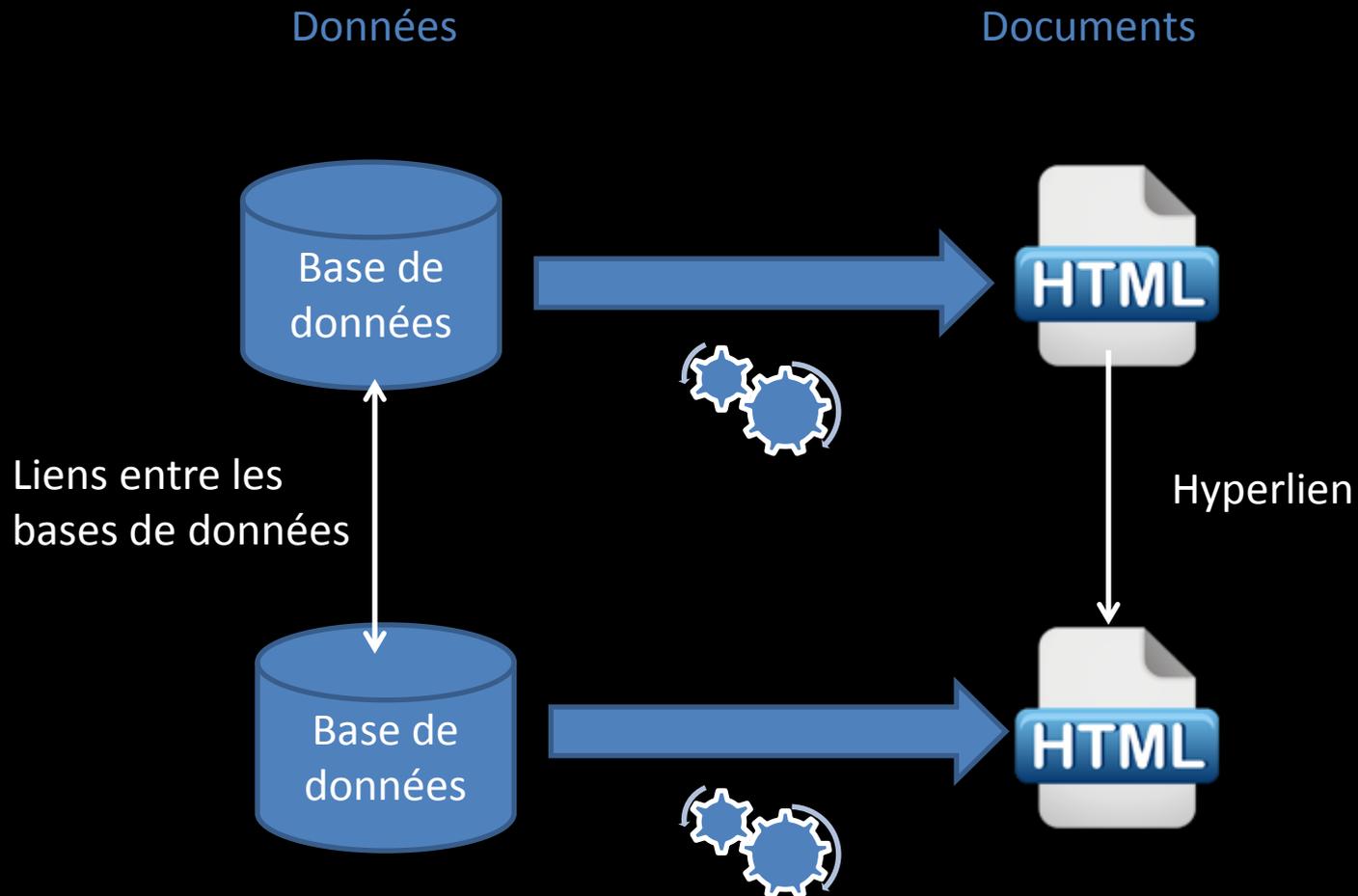




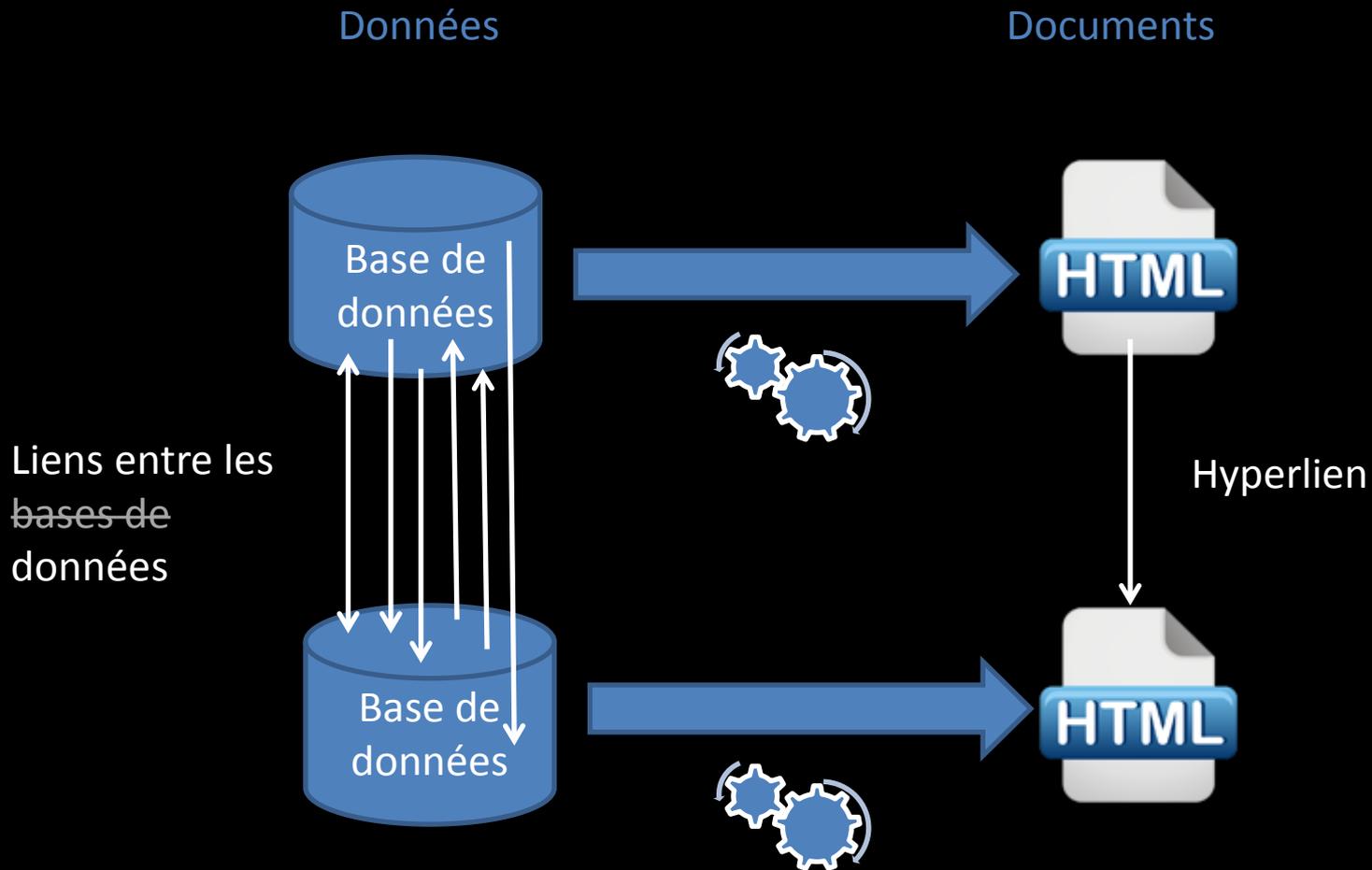
# Un archipel d'icebergs

- Les données sont cachées sous les pages HTML (*deep web*)
- Seules les pages HTML sont liées entre elles
- Les pages HTML sont faites pour les humains
- Ce que veulent les machines, ce sont des données brutes
  - et des données liées
    - car elles se complètent les unes les autres
    - car les liens permettent de naviguer et de découvrir
  - Et des liens qualifiés, signifiants
    - au-delà du « voir aussi » des hyperliens

# Le web de données liées



# Le web de données liées



Exemple

**WEB DE DOCUMENTS**

**ET**

**WEB DE DONNÉES**



## Eric Dolphy

[Tracks](#) [Albums](#) [Pictures](#) [Videos](#) [Events](#) [Biography](#) [More...](#)[+](#) Add to library[Buy](#)[Share](#)

1,363,770

scrobbles

169,689

listeners



Douglas\_Sabbath is listening to  
*Eric Dolphy – Straight Up and Down*

[jazz](#) [free jazz](#) [avant-garde](#) [avant-garde jazz](#) [saxophone](#) [more tags](#)

Los Angeles, United States (1928 – 1964)

Eric Allan Dolphy (June 20, 1928 – June 29, 1964) was a [jazz](#) musician who played [alto saxophone](#), [flute](#) and [bass clarinet](#) and was educated at Los Angeles City College. Dolphy was the first important bass clarinet soloist in jazz, and one of the first viable flute soloists in jazz. On early recordings, he occasionally played traditional B-flat clarinet. His un-

[Sources \(view history\)](#)[http://en.wikipedia.org/wiki/Eric\\_Dolphy](http://en.wikipedia.org/wiki/Eric_Dolphy)

MongoDB hosting, designed to scale, built by experts.

mongoHQ

Smooth scaling from shared to clustered





WIKIPEDIA  
The Free Encyclopedia

Navigation

Main page

Contents

Featured content

Current events

Random article

Donate to Wikipedia

Interaction

Help

About Wikipedia

Community portal

Recent changes

Contact page

Toolbox

What links here

Related changes

Upload file

Special pages

Permanent link

Article Talk

Read Edit View history

Search

## Eric Dolphy

From Wikipedia, the free encyclopedia

**Eric Allan Dolphy, Jr.** (June 20, 1928 – June 29, 1964) was an American jazz alto saxophonist, flautist, and bass clarinetist. On a few occasions, he also played the clarinet, piccolo, and baritone saxophone. Dolphy was one of several multi-instrumentalists to gain prominence in the 1960s. He was also the first important bass clarinet soloist in jazz, and among the earliest significant flute soloists.

His improvisational style was characterized by the use of wide intervals, in addition to using an array of extended techniques to reproduce human- and animal-like effects which almost literally made his instruments speak. Although Dolphy's work is sometimes classified as free jazz, his compositions and solos were often rooted in conventional (if highly abstracted) tonal bebop harmony and melodic lines that suggest the influences of modern classical composers Béla Bartók and Igor Stravinsky.

**Contents** [hide]

### 1 Biography

#### 1.1 Early life

#### 1.2 Early partnerships

##### 1.2.1 Charles Mingus

##### 1.2.2 John Coltrane

##### 1.2.3 Booker Little

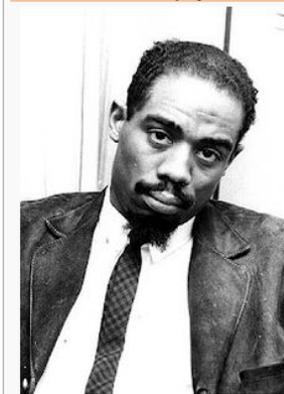
##### 1.2.4 Others

#### 1.3 As a leader

#### 1.4 Final months

### 2 Influence

Eric Dolphy



Information

Dolphy, Jr.  
928  
es, California, United

## Further reading [edit]

- Vladimir Simosko & Barry Tepperman: *Eric Dolphy: A Musical Biography and Discography*, Da Capo Press, New York, 1979, ISBN 0-306-80107-8
- Guillaume Belhomme: *Eric Dolphy*, Le mot et le reste, Marseille, 2008, ISBN 978-2-915378-53-5
- Raymond Horricks: "The Importance of Being Eric Dolphy", D J Costello Publishers Ltd., Great Britain, 1989. ISBN 0-7104-3048-5

## External links [edit]

# Session and discography

**Authority control** WorldCat  · VIAF:  66651545  · LCCN:  n81035956  · ISNI:  0000 0000 8146 531X  · GND:  118812866  · BNF:  cb13893337q

Categories: 1928 births | 1964 deaths | Avant-garde jazz musicians | African-American musicians | American people of Panamanian descent | American jazz clarinetists | American multi-instrumentalists | American jazz bass clarinetists | Deaths from diabetes | Musicians from Los Angeles | Bass clarinetists | Accidental deaths in Germany

## Eric Dolphy Discography - session index

### 1948 (age 20)

- Roy Porter's Seventeen Beboppers LA, 1948 or Jan. 19, 1949

### 1949

- Roy Porter Boppers LA, Feb. 23, 1949
- Charlie Mingus 22 Piece Be Bop Band Hollywood, CA, spr. 1949

### 1958 (age 30)

- Chico Hamilton 5 LA, Apr., 1958
- Chico Hamilton 5 Newport Jazz Fest., Newport, RI, July 7, 1958
- Chico Hamilton 5 w/ Strings LA, Oct. 26, 1958
- Chico Hamilton 5 w/ Strings LA, Oct. 27, 1958
- Chico Hamilton 5 LA, Dec. 29 & 30, 1958

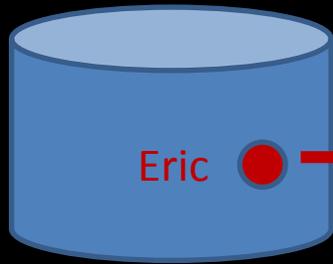
### 1959

- Chico Hamilton 8 LA, Feb. 25, 1959
- Chico Hamilton 5 LA, May 19, 1959
- Chico Hamilton 5 LA, May 20, 1959
- Chico Hamilton 5 LA, late Oct., 1959

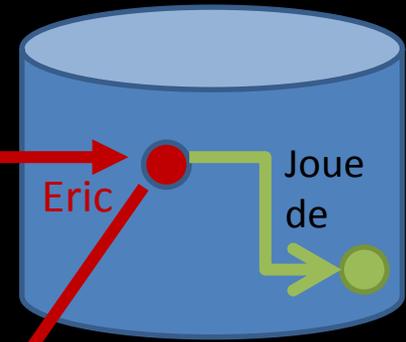
### 1960

- Sammy Davis Jr. w/ Sy Oliver Orch. NYC, early 1960
- Eric Dolphy 5 Rudy Van Gelder Studio, Englewood Cliffs, NJ, Apr. 1, 1960
- Ruth Brown NYC, Apr. 19, 1960

Last.fm



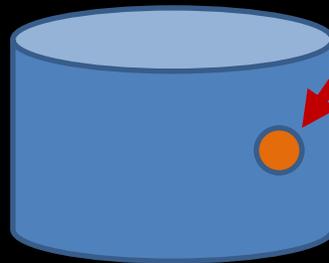
est le même que



Wikipedia

flute

a joué pendant



Session - LA, Feb. 25, 1959

Jazzdisco.org

Le web  
comme une grande  
base de données  
distribuée

# PARLER AUX MOTEURS DE RECHERCHE

schema.org

Without Markup   Microdata   RDFa   JSON-LD

GreatFood  
4 stars - based on 250 reviews  
1901 Lemur Ave  
Sunnyvale, CA 94086  
(408) 714-1489  
<a href="http://www.greatfood.com">www.gre  
Hours:  
Mon-Sat 11am - 2:30pm  
Mon-Thur 5pm - 9:30pm  
Fri-Sat 5pm - 10pm  
Categories: Middle Eastern, Mediterranean  
Price Range: \$\$  
Takes Reservations: Yes

Without Markup   Microdata   RDFa   JSON-LD

```
<script type="application/ld+json">
{
  "@context": "http://schema.org",
  "@type": "Restaurant",
  "address": {
    "@type": "PostalAddress",
    "addressLocality": "Sunnyvale",
    "addressRegion": "CA",
    "postalCode": "94086",
    "streetAddress": "1901 Lemur Ave"
  },
  "aggregateRating": {
    "@type": "AggregateRating",
    "ratingValue": "4",
    "reviewCount": "250"
  },
  "name": "GreatFood",
  "openingHours": [
    "Mo-Sa 11:00-14:30",
    "Mo-Th 17:00-21:30",
    "Fr-Sa 17:00-22:00"
  ],
  "priceRange": "$$",
  "servesCuisine": [
    "Middle Eastern",
    "Mediterranean"
  ],
  "telephone": "(408) 714-1489",
  "url": "http://www.dishdash.com"
}
```

@todo #bu

- . Site web, bib num, etc.
- . Utile et formateur

**UN EXEMPLE DE CHEZ NOUS**

# Une base de données

DOCUMENTS		
id	auteur	titre
doc1	pers1	Retractatio

Lien interne à la base

AUTEURS		
id	nom	
pers1	Paul Veyne	

# Une autre base de données

LIVRES			
id	z7XX	titrepropre	editeur
068391307	027182800	Le Pain et le cirque	Le Seuil

Lien interne à la base

PERSONNES		
id	nom	personnetype
027182800	Veyne, Paul (1930-....	0

DOCUMENTS		
id	auteur	titre
doc1	pers1	Retractatio

AUTEURS		
id	nom	
pers1	Paul Veyne	

#1 Comment lier la donnée d'une base à la donnée de l'autre base ?

#2 Comment lier deux bases qui n'ont pas le même schéma

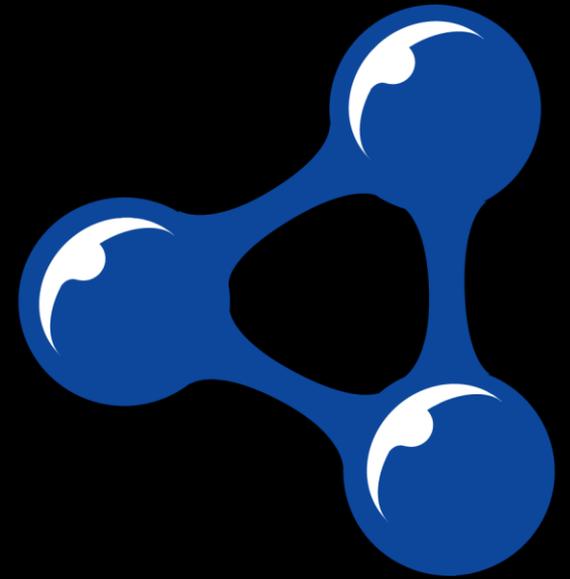
#3 Comment qualifier la nature de ce lien ?

#4 Comment faire tout ça simplement (simple et efficace comme le web [de documents])?

est la même entité que

LIVRES			
id	z7XX	titrepropre	editeur
068391307	027182800	Le Pain et le cirque	Le Seuil

PERSONNES		
id	nom	personnetype
027182800	Veyne, Paul (1930-....	0



RDF

**UN MODÈLE GÉNÉRAL POUR  
DÉCRIRE ... N'IMPORTE QUOI**

# En RDF, toute connaissance se résume à

## ça



Ceci est un **triplet**

NB : « prédictat » == « relation »

NB : les objets et les sujets sont des « ressources », des choses, entités, trucs...

et ça



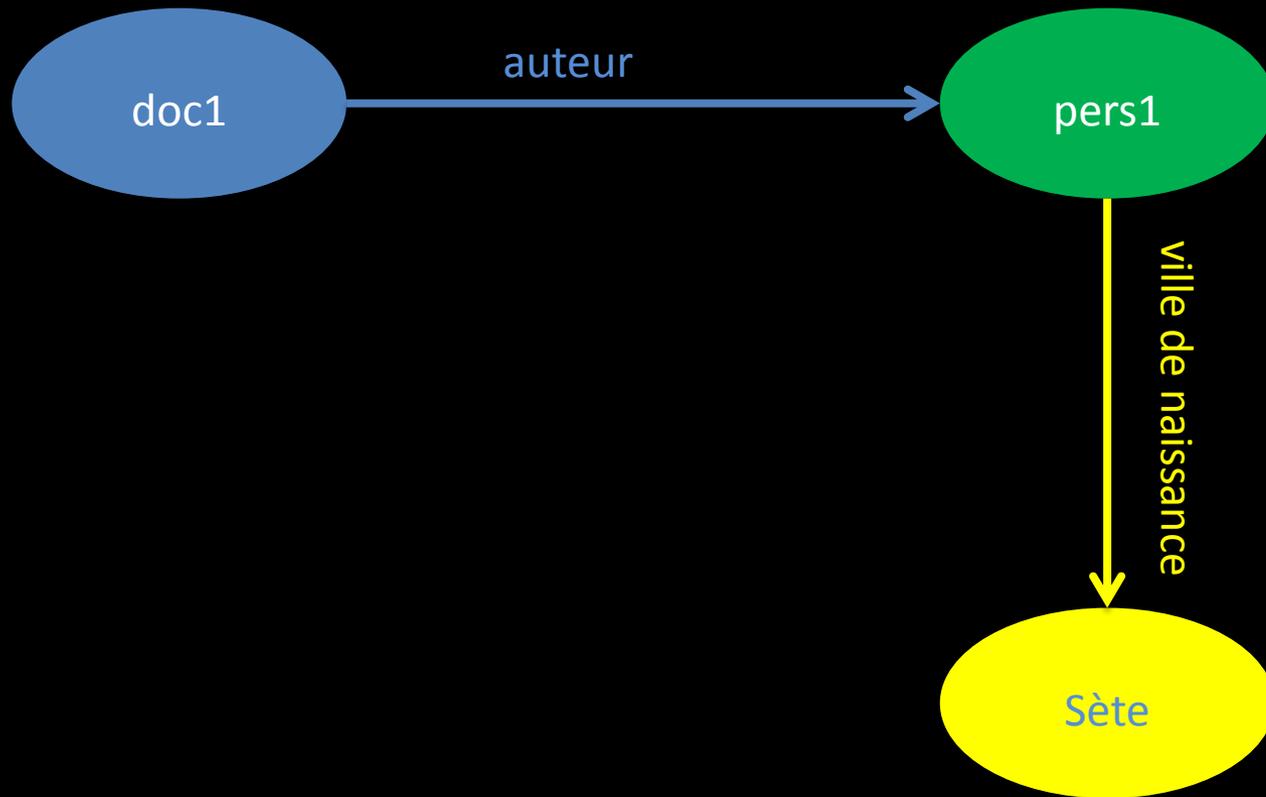
Ceci est un **triplet**

NB : « blabla » = du texte = un Littéral

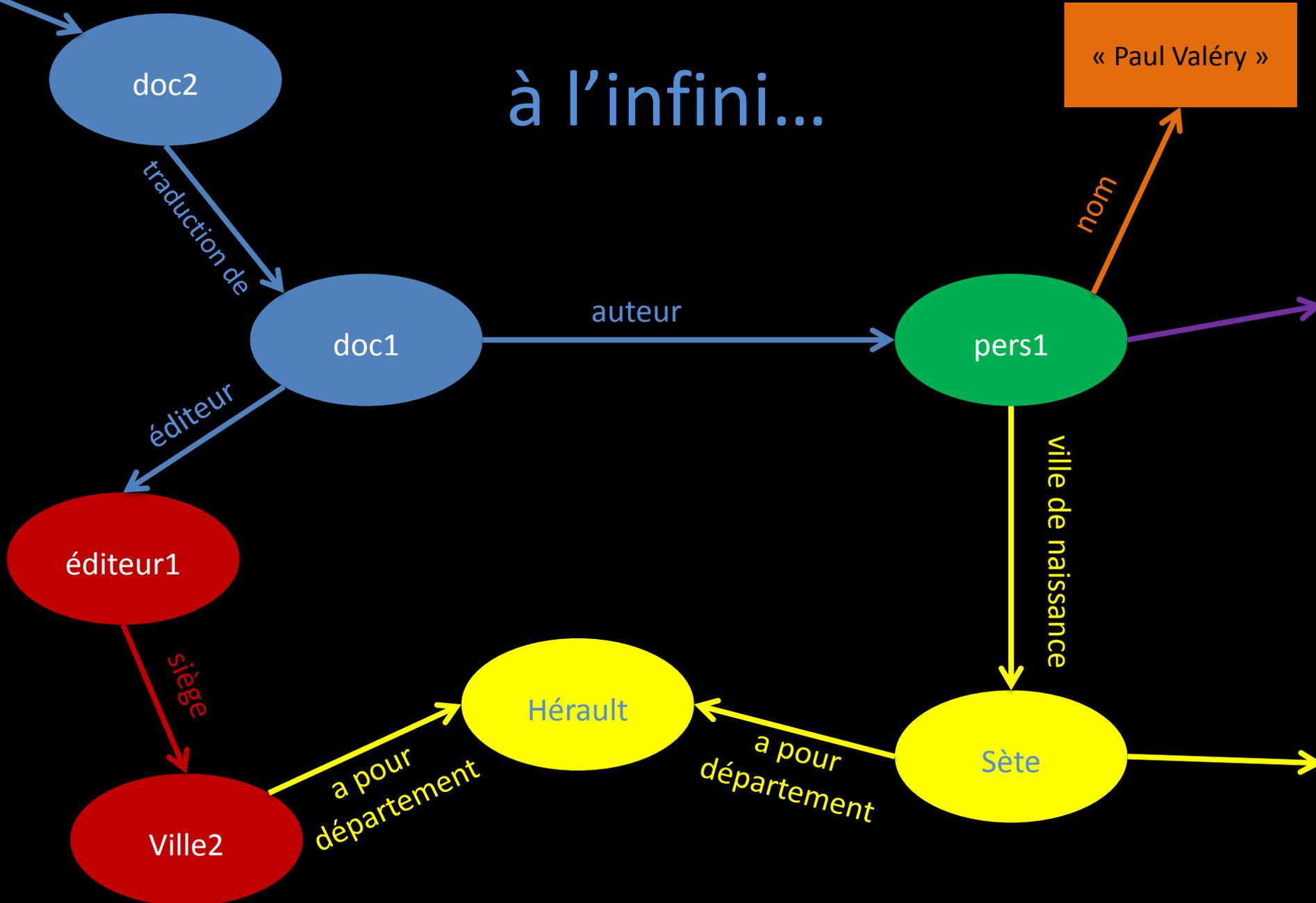
# Et les triplets, ça s'emboîte



# Et les triplets, ça s'emboîte

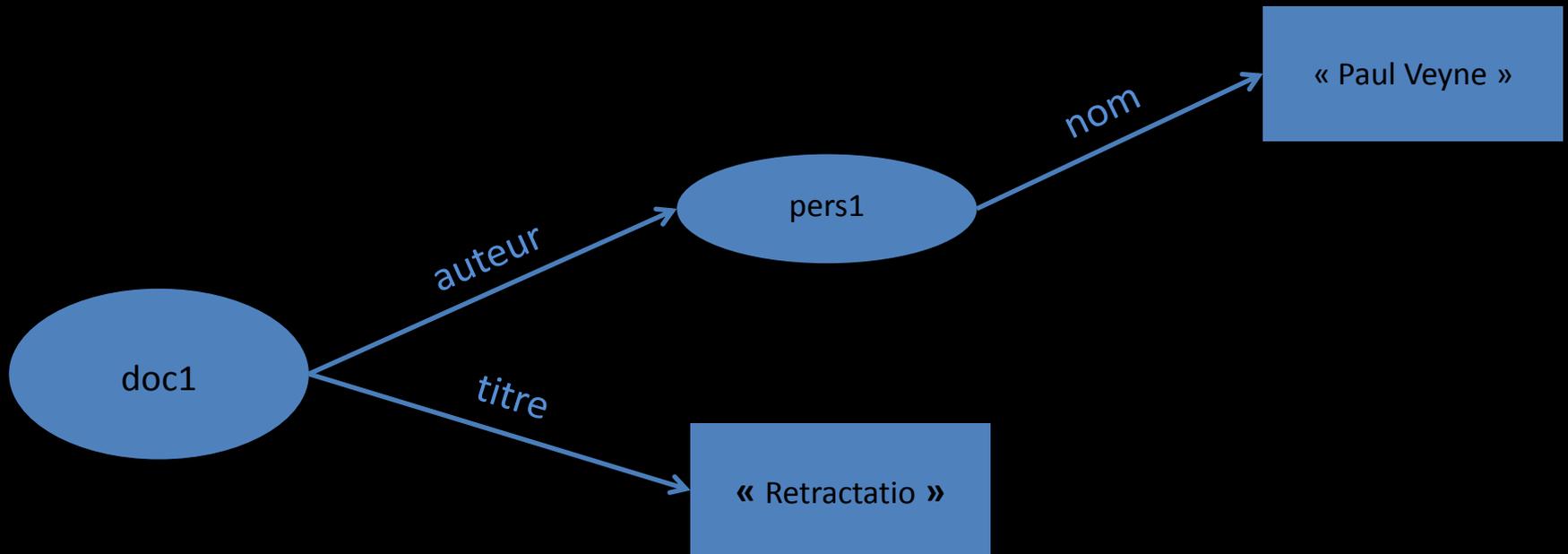


à l'infini...

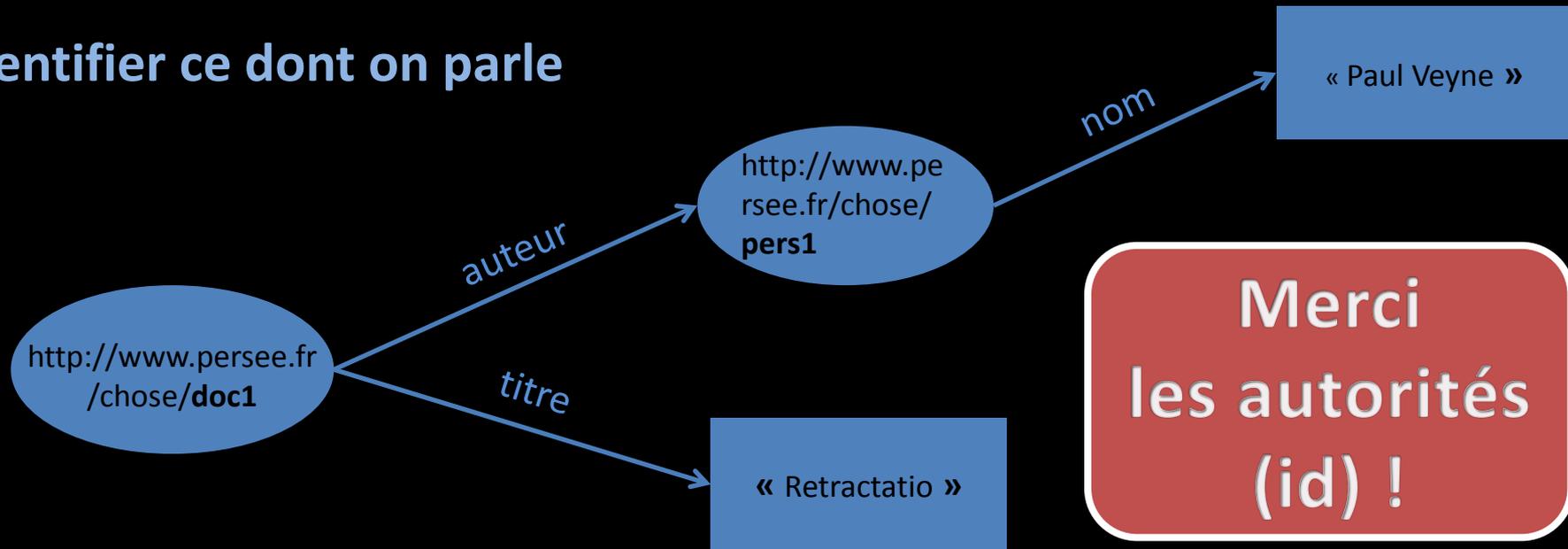


**REVENONS À PAUL VEYNE**

La base devient une collection de triplets connectés entre eux



## Identifier ce dont on parle



Merci  
les autorités  
(id) !

Donner aux choses  
un nom sans ambiguïté,  
pour faire des liens d'une base à  
l'autre, à l'échelle du web

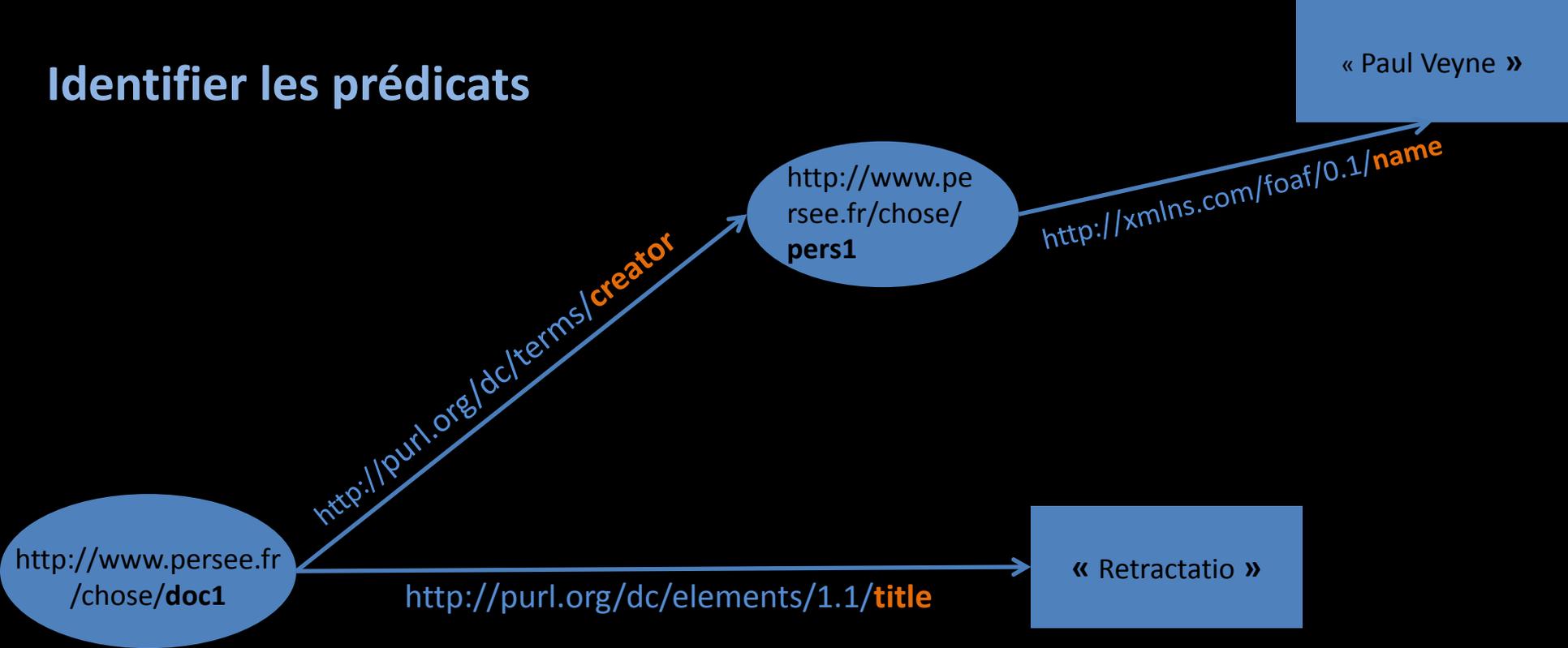
Les entités (*ressources*) ont un nom précis.

Ce nom est une **URL**.

ex : <http://www.persee.fr/chose/doc1>

ex : <http://www.sudoc.fr/092673007/id>

# Identifier les prédicats



Les relations (*prédicats*) ont un nom précis.  
Ce nom est une **URL**.

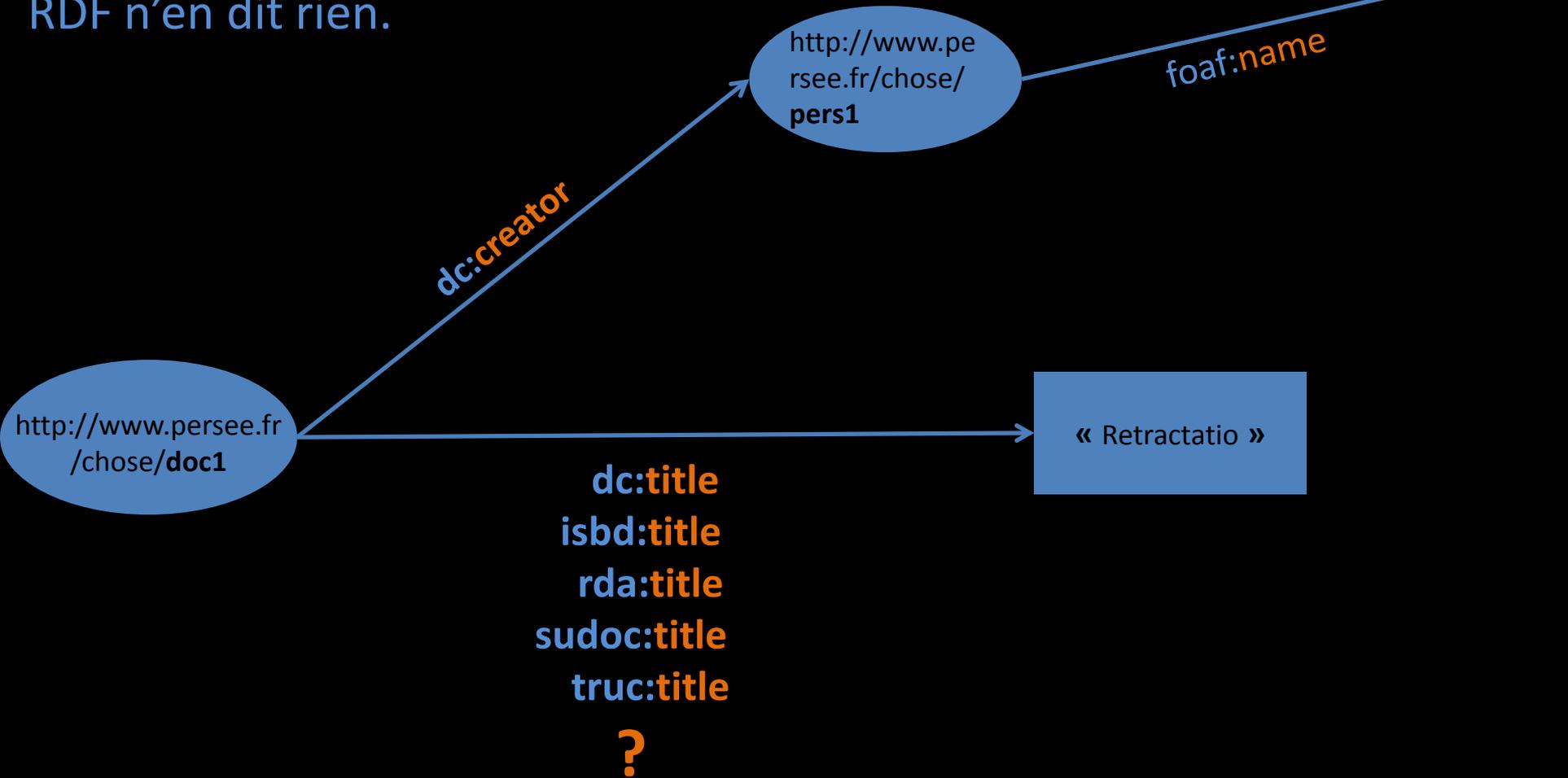
Le titre selon le vocabulaire Dublin Core = `http://purl.org/dc/elements/1.1/title`

@todo #bu

. Soigner les identifiants !

- IdRef
- Labos
- Lieux
- etc.

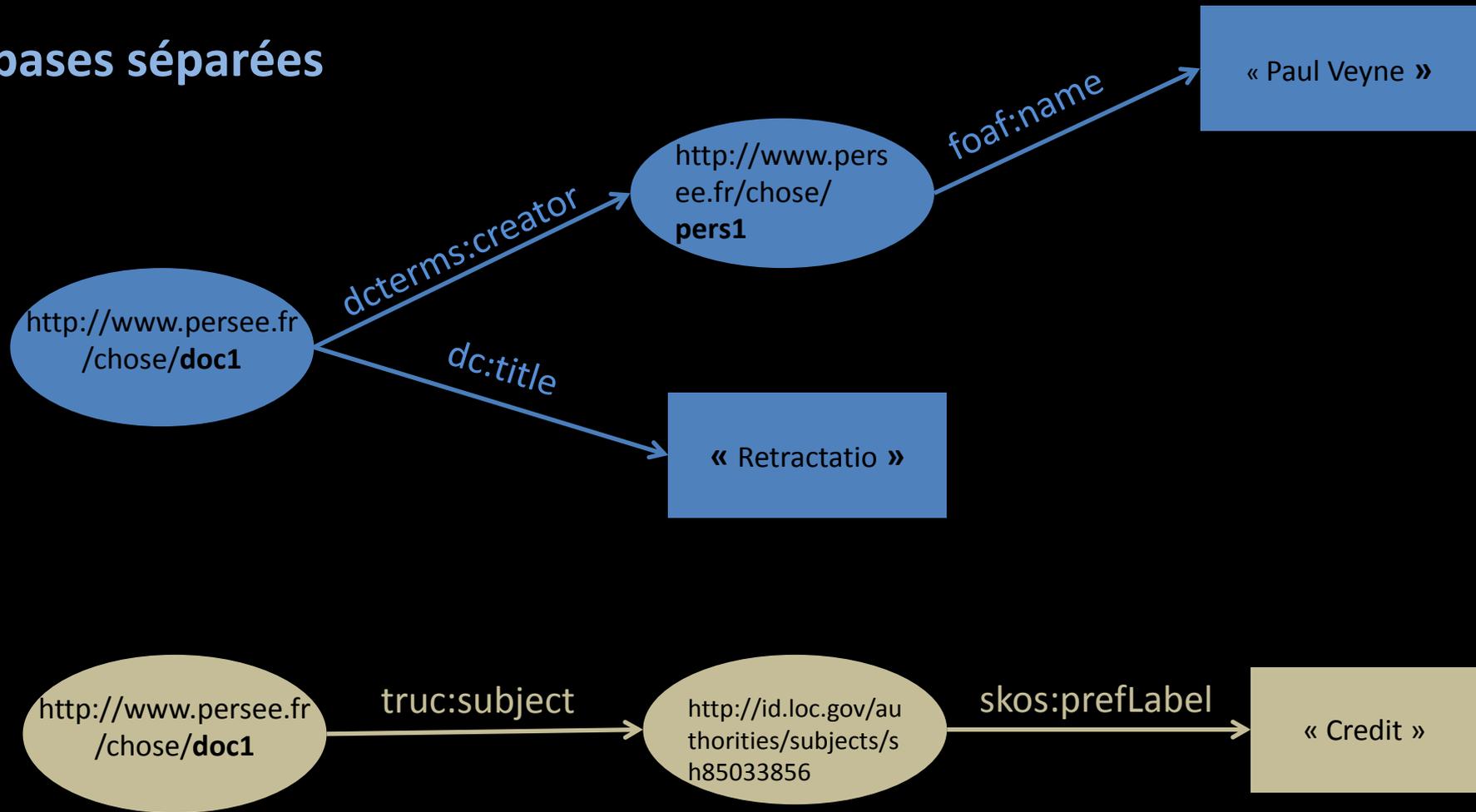
Quel vocabulaire utiliser ?  
RDF n'en dit rien.





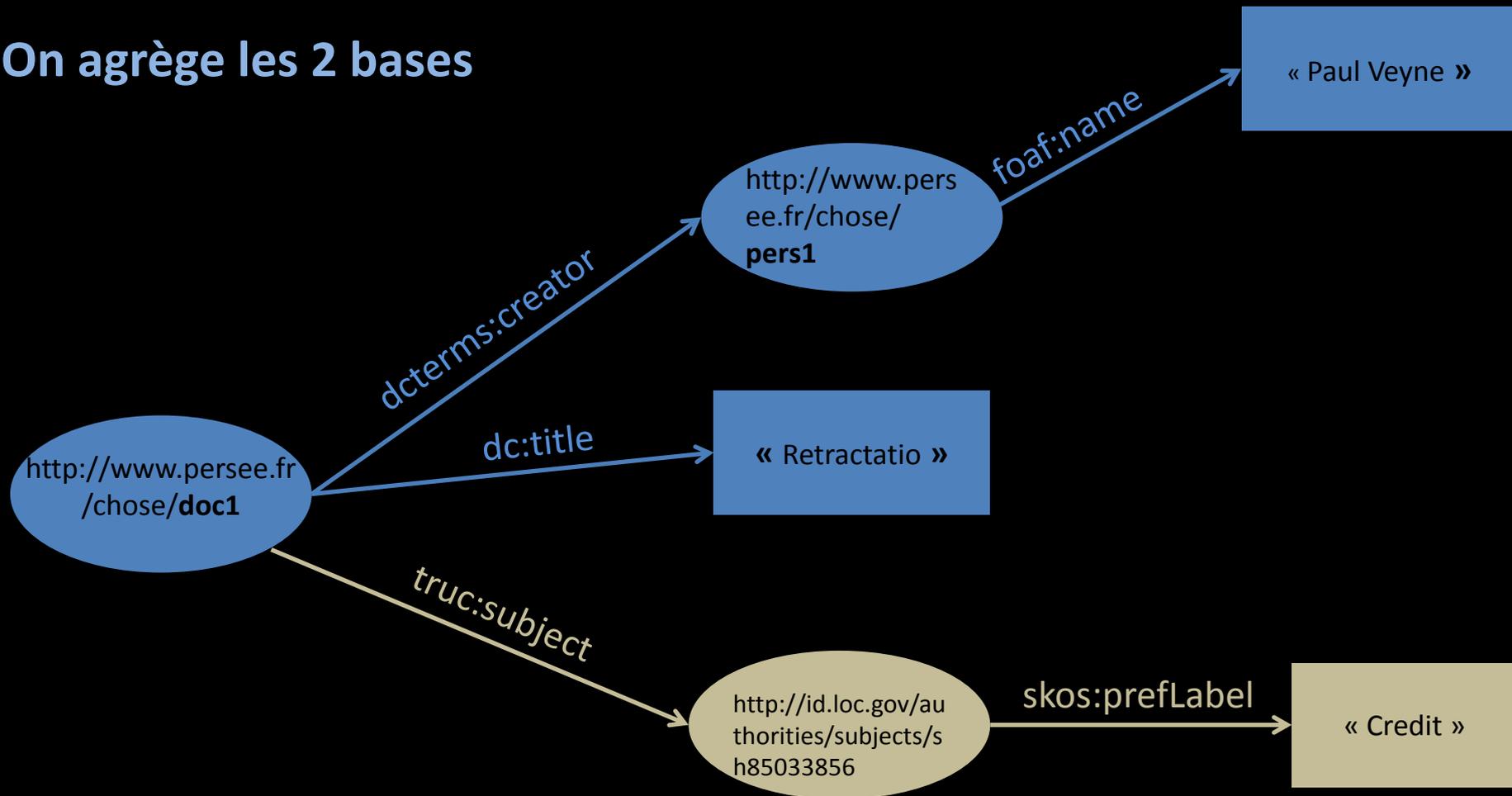
**REVENONS À NOTRE PROBLÈME :**  
**LIER DEUX BASES**

## 2 bases séparées



Deux sites en RDF peuvent parler de la même chose – explicitement ici (même nom (URL)).

# On agrège les 2 bases

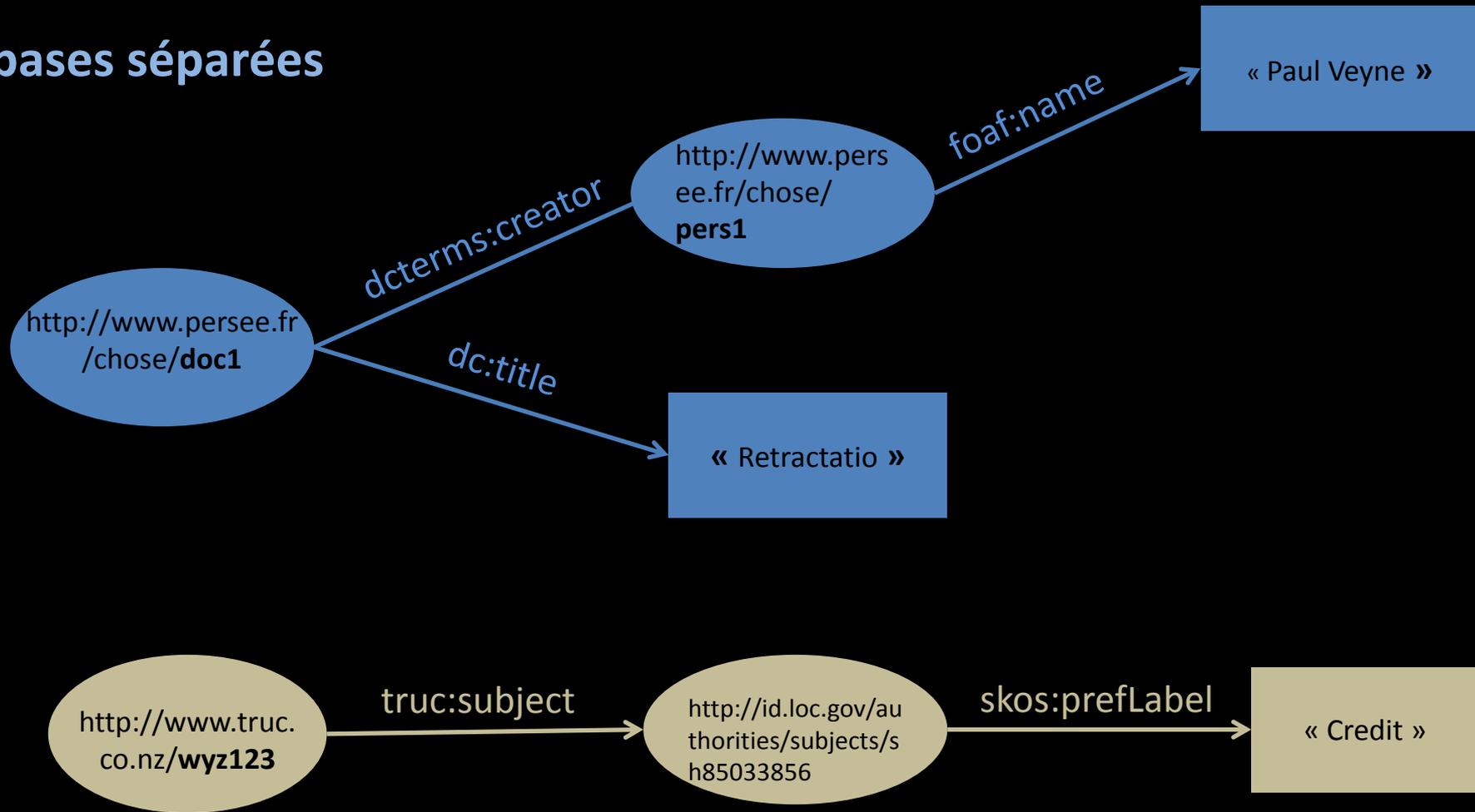


L'agrégation des données est immédiate

Deux sites en RDF peuvent parler de la même chose.

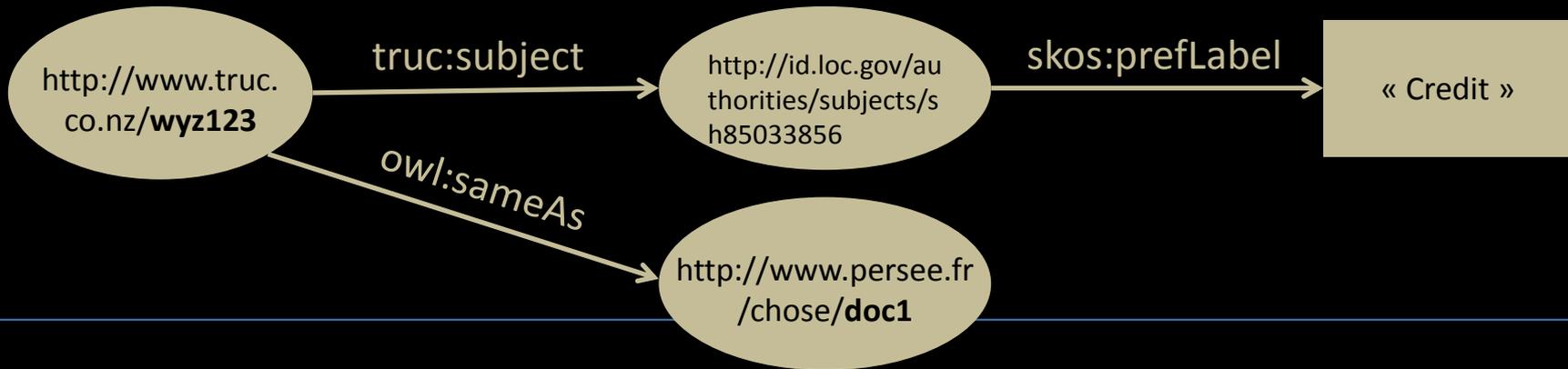
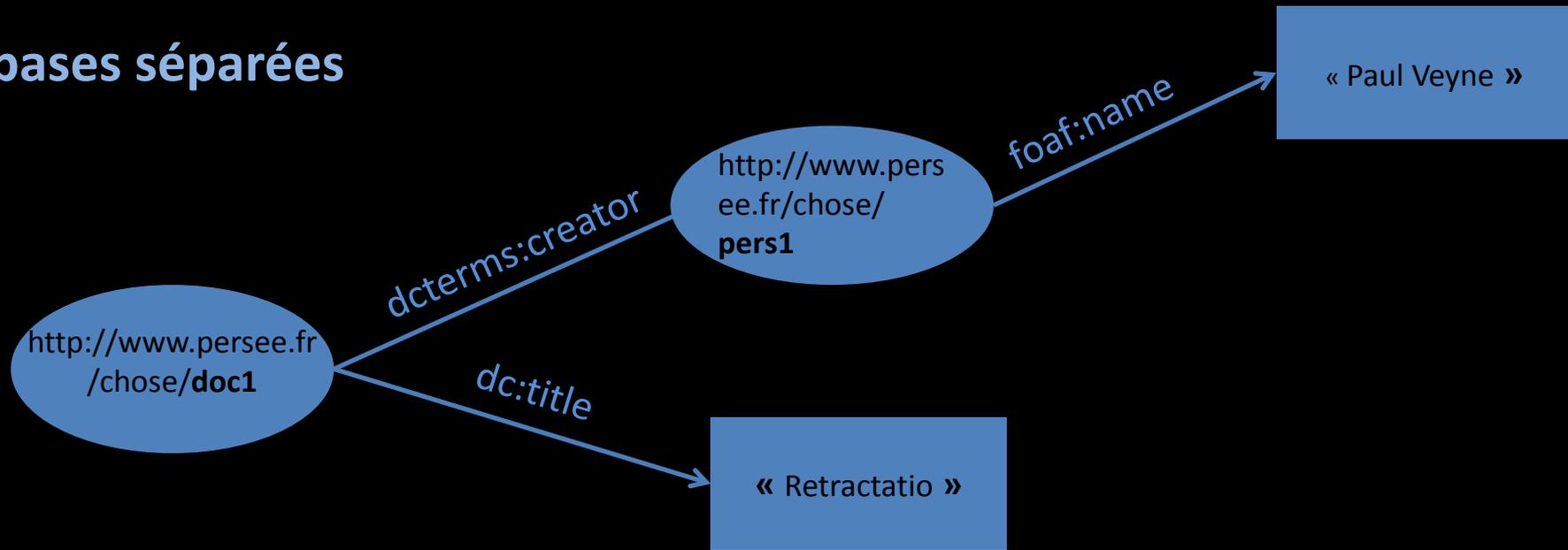
Mais pas forcément de la même manière (pas de schéma ou de vocabulaire unique)

## 2 bases séparées



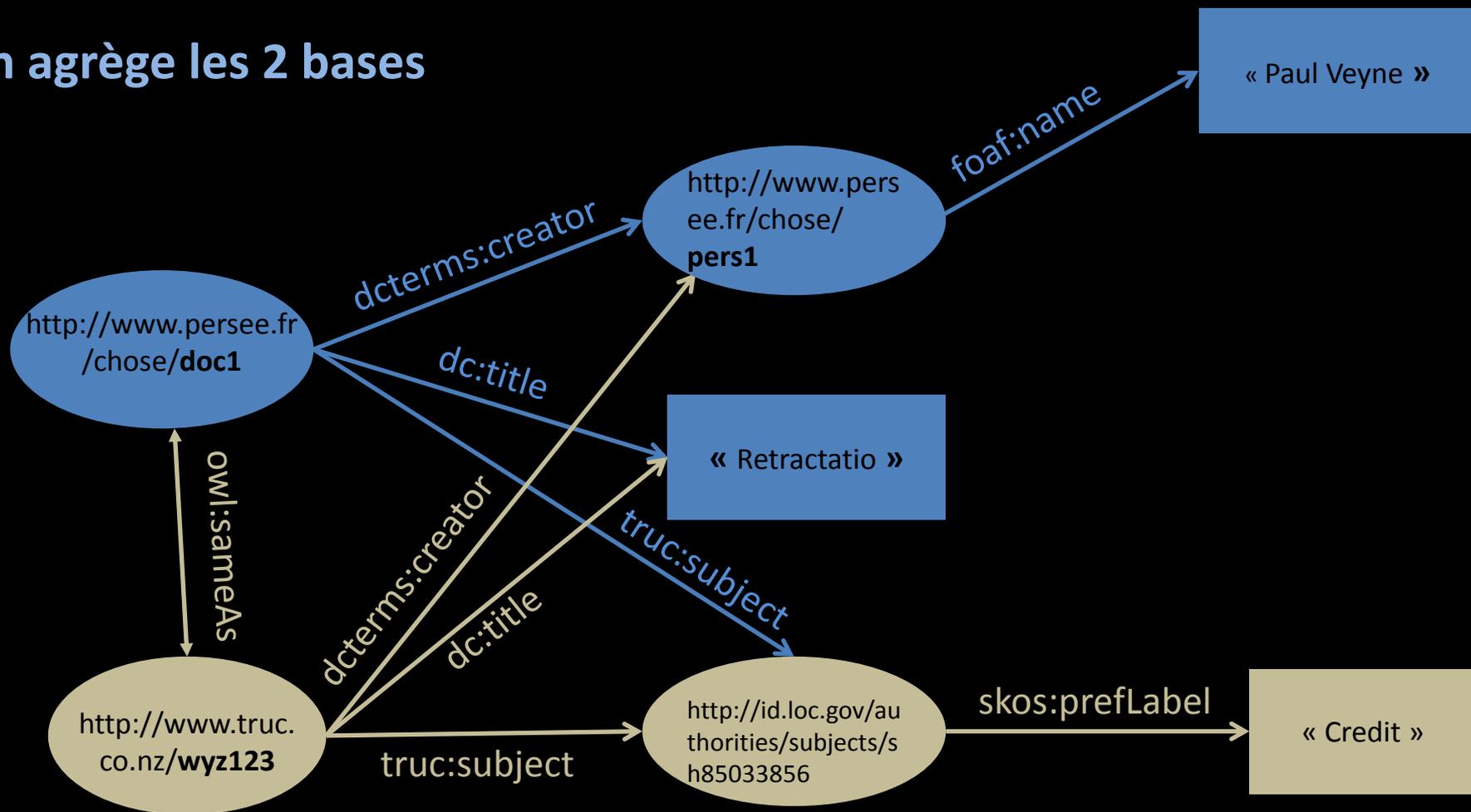
Deux sites en RDF peuvent parler de la même chose – **implicitement** ici (deux noms différents !)  
→ Problème pour les agréger

## 2 bases séparées



→ On dit **explicitement** qu'il s'agit de la même entité avec `owl:sameAs`

# On agrège les 2 bases



Grâce à `owl:sameAs`, la fusion est totale : tout ce qu'on dit de l'un est aussi vrai de l'autre

# Paul Veyne dans le RDF de VIAF

<http://viaf.org/viaf/108250528/rdf.xml>

```
- <rdf:RDF xml:base="http://viaf.org/">
- <rdf:Description rdf:about="http://viaf.org/viaf/108250528">
- <!--
  The primary entity (formerly a viaf:NameAuthorityCluster)
-->
  <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person"/>
  <rdf:type rdf:resource="http://rdvocab.info/uri/schema/FRBRentitiesRDA/Person"/>
  <foaf:name>Veyne, Paul, 1930-....</foaf:name>
  <foaf:name>Veyne, Paul</foaf:name>
  <foaf:name>1930، بول، فيين</foaf:name>
  <foaf:name>Ven, Pöl, 1930-</foaf:name>
  <foaf:name>Vein, Pol</foaf:name>
  <foaf:name>Bein, Pol</foaf:name>
  <foaf:name>1930، بول فيين</foaf:name>
  <rdaGr2:dateOfBirth>1930-06-13</rdaGr2:dateOfBirth>
  <owl:sameAs rdf:resource="http://dbpedia.org/resource/Paul_Veyne"/>
  <owl:sameAs rdf:resource="http://libris.kb.se/resource/auth/309710"/>
  <owl:sameAs rdf:resource="http://www.idref.fr/027182800/id"/>
  <owl:sameAs rdf:resource="http://d-nb.info/gnd/http://d-nb.info/gnd/119421194"/>
</rdf:Description>
- <!--
  Controlled and alternate headings from various authority files
-->
- <skos:Concept rdf:about="http://viaf.org/viaf/sourceID/BNF%7C11928105#skos:Concept">
  <skos:inScheme rdf:resource="http://viaf.org/authorityScheme/BNF"/>
  <skos:prefLabel>Veyne, Paul, 1930-....</skos:prefLabel>
  <foaf:focus rdf:resource="http://viaf.org/viaf/108250528"/>
</skos:Concept>
```

<http://viaf.org/viaf/108250528>

owl:sameAs

<http://www.idref.fr/027182800/id>

Le Veyne de VIAF, le Veyne d'IdRef et le Veyne de la DNB ne font qu'un

# Paul Veyne dans

F

28/rdf.xml

On peut se promener d'une base à l'autre :

« Je veux les livres du catalogue de la BN allemande dont l'auteur est [www.idref.fr/027182800/id](http://www.idref.fr/027182800/id) »

<http://viaf.org/viaf/108250528/rdf.xml>

owl:sameAs

<http://www.idref.fr/027182800/id>

Veyne  
VIAF

eyne  
Ref  
ne font  
qu'un

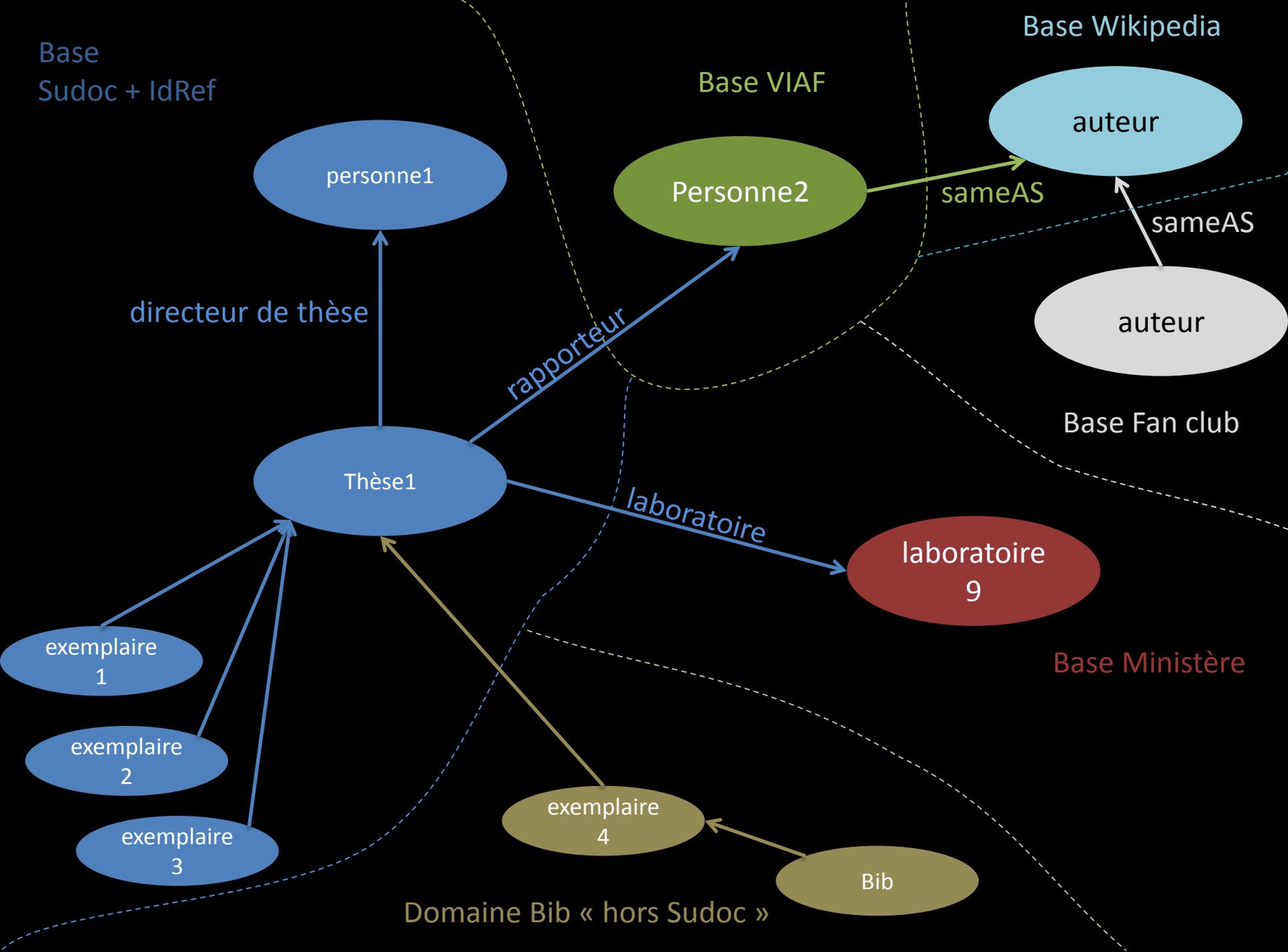
```

...<skos:Concept rdf:about="http://www.idref.fr/027182800/id"/>
...<skos:Concept rdf:about="http://d-nb.info/gnd/http://d-nb.info/gnd/119421194"/>
...</skos:Concept>
<!--
Controlled and alternate headings from various authority files
-->
- <skos:Concept rdf:about="http://viaf.org/viaf/sourceID/BNF%7C11928105#skos:Concept">
  <skos:inScheme rdf:resource="http://viaf.org/authorityScheme/BNF"/>
  <skos:prefLabel>Veyne, Paul, 1930-....</skos:prefLabel>
  <foaf:focus rdf:resource="http://viaf.org/viaf/108250528"/>
</skos:Concept>

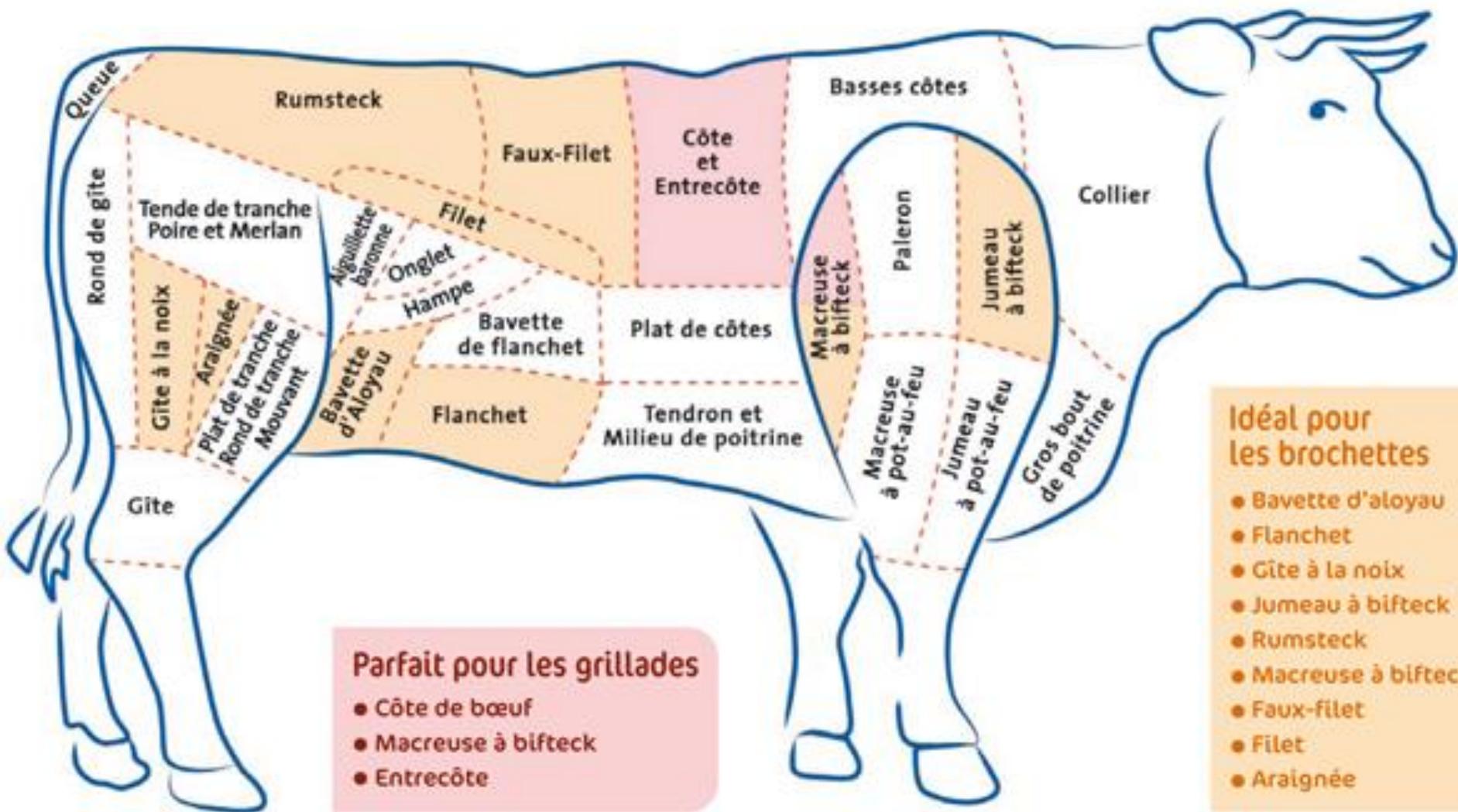
```

illustration

# APPLICATION AU MONDE DES BIBLIOTHÈQUES



# « Chacun son métier et les vaches seront bien gardées »



- Parfait pour les grillades**
- Côte de bœuf
  - Macreuse à bifteck
  - Entrecôte

- Idéal pour les brochettes**
- Bavette d'aloïyou
  - Flanchet
  - Gîte à la noix
  - Jumeau à bifteck
  - Rumsteck
  - Macreuse à bifteck
  - Faux-filet
  - Filet
  - Araignée

# Application au monde des bib.

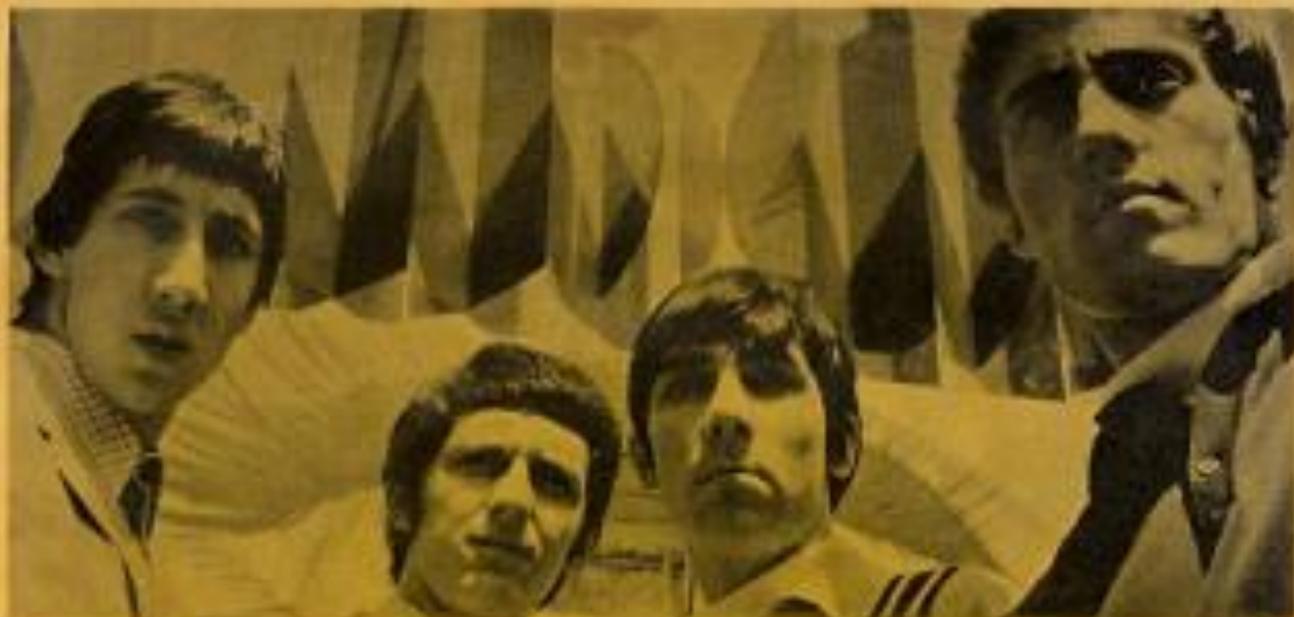
- Chacun peut/devrait se spécialiser
  - sur son cœur de métier
  - sur son cœur de métadonnées
- On coopère en faisant des liens
  - ce qui exige une coordination faible
  - ce qui suppose que l'autre ouvre ses données
    - manquent : référentiels de labos, entreprises, œuvres (!), éditeurs, etc.
- On n'est plus confiné au monde des bib.
  - technologies universelles (innovation, marché)
  - Liens
    - depuis nos données
    - vers nos données (données bib = données de référence ?)

**anyone can say anything anywhere**

**AAA**

anyone can say anything anywhere  
anyhow

AAAA



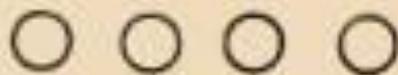
anyway,  
anyhow,  
anywhere

DADDY  
ROLLING  
STONE

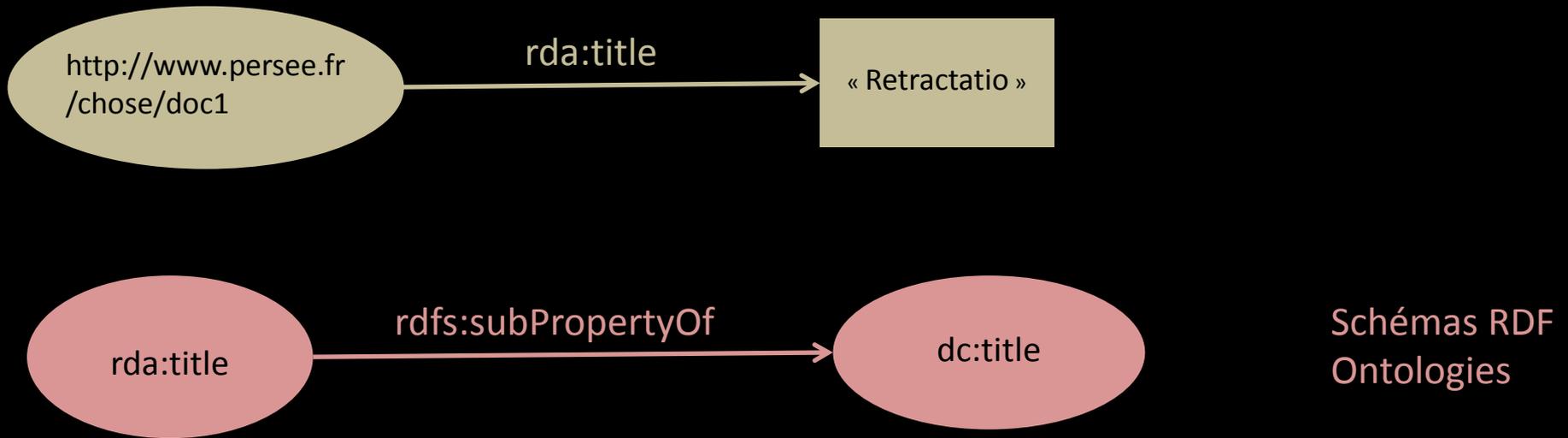
Warner

© 5935

the  
who



# Données + sémantique



Le schéma RDF ajoute des informations sur le vocabulaire utilisé (propriétés des propriétés et des classes)

Ce qui permet de faire du **raisonnement** →

## On en déduit que...



Schémas RDF  
Ontologies



Triplets déduits

Le raisonnement permet d'ajouter des informations (triplets).  
En l'occurrence, il permet d'exprimer automatiquement en Dublin Core un triplet qui utilisait un vocabulaire RDA

Exemple  
Depardon  
Stardog

# Quel(s) vocabulaire(s) métier ?

Devons-nous utiliser Bibframe, RDA, Dublin Core ou un vocabulaire bibliographique maison?

- Ils peuvent cohabiter
- On peut expliciter les relations entre eux, ce qui permet de passer automatiquement de l'un à l'autre, par raisonnement (en théorie...)

➔ soulage le dilemme de l'interopérabilité :  
= (précis + ad hoc) vs (imprécis + standard)

**VICES**

**PÉCHÉS DE  
JEUNESSE  
DE RDF**



# Limites actuelles

- Performances
- Compétences
- Qualité des données / Confiance
- Synchronisation des données
  - Pérennité des données liées/réutilisées
- Equilibre entre diversité des besoins et des vocabulaires et normalisation/convergence
- Fétichisme RDF
  - Utiliser RDF à bon escient

**QUE FAIT L'ABES ?**

- Ouvrir ses données en RDF (2008- )
  - » pour partager (≠ échanger)
- Agréger et traiter des données hétérogènes en RDF
  - » Hub de métadonnées 2013-
- Cataloguer dans le contexte du web de données (201?- )
  - » Catalinking, catalogic, cataLOD\*...

# Ouvrir ses données en RDF (2008- )

Base	Format d'origine	Quand ?	Format sémantiq.	Dump ?	SPARQL ?	URI déréférençables	Liens externes
Calames	XML (EAD)	2008	RDFa	Pas de dump	Non	Non	*
IdRef	MARC	2010	RDF/XML	Dump (non public)	Non	Oui	**
Sudoc	MARC	2011	RDF/XML Microdata	Dump (non public)	Non	Oui	**
theses.fr	XML (TEF)	2011	RDF/XML RDfa	Pas de dump	Non	Oui	**

# Sudoc, IdRef et theses.fr dans le web de données (2010-2011) : principes pour un premier pas

- Prendre les données telles qu'elles sont
  - Pas de transformations complexes
- Prendre des vocabulaires du marché
  - Pas forger des classes ou propriétés à nous
- On ne cherche pas à tout exprimer en RDF
- Faire des liens réciproques entre IdRef et Sudoc
- Génération du RDF à la volée
  - MARC/XML → RDF/XML par XSLT dans DB Oracle
  - Pas de triple store
- Respect des pratiques *Linked data* (5 étoiles)

# Sudoc en RDF

<http://www.sudoc.fr/13414442> page HTML

<http://www.sudoc.fr/13414442.rdf> données RDF

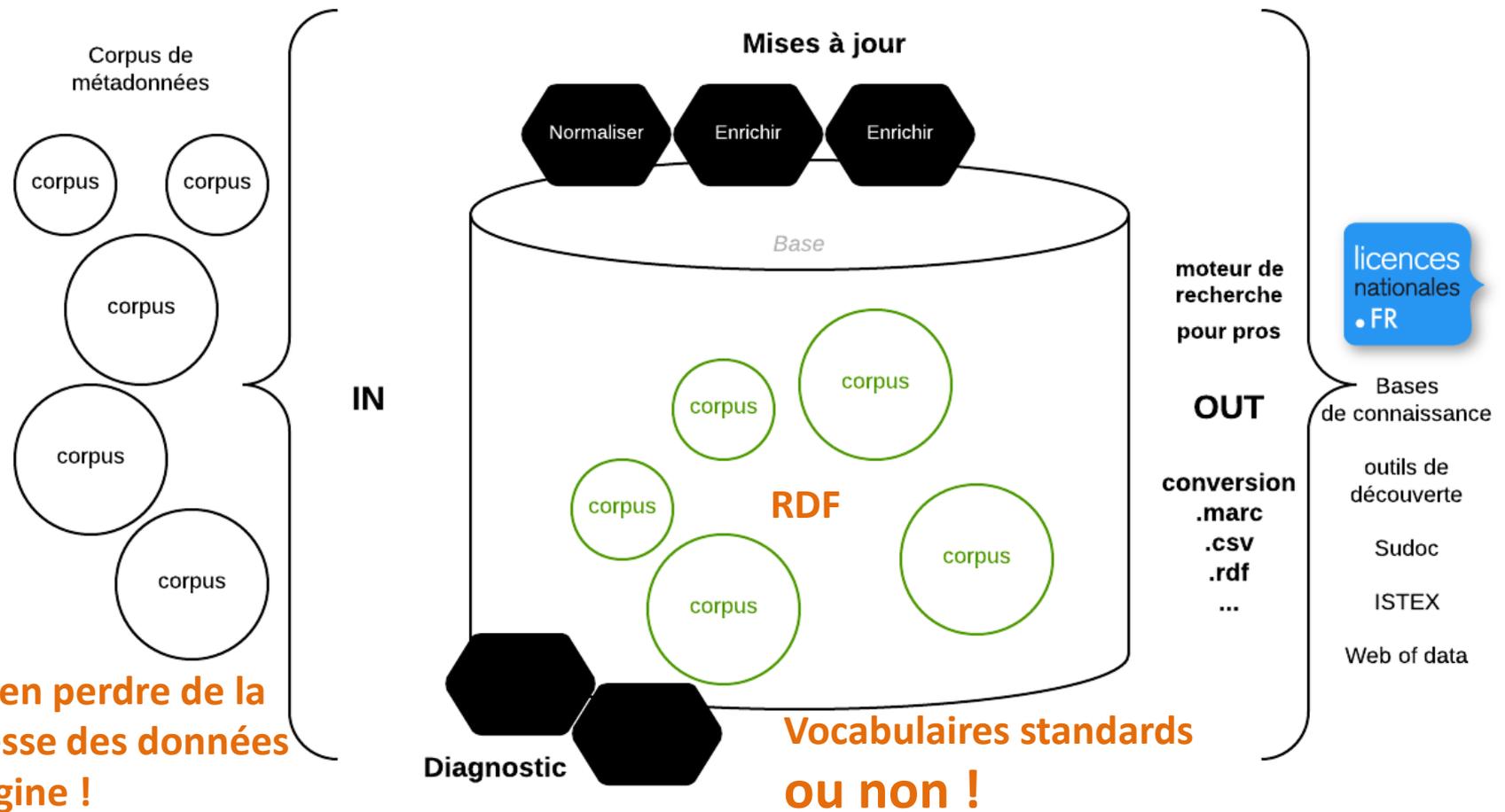
~~Données RDF exportées et interrogeables en  
SPARQL chez Sindice :~~

~~<http://punktokomo.abes.fr/2012/11/23/un-serveur-sparql-pour-le-sudoc/>~~

# Prochaines étapes

1. Enrichir la conversion en RDF en utilisant des vocabulaires RDF métier : ISBD, FRBR, RDA
    - Travail en cours
    - Mais ces vocabulaires ne sont pas mûrs
  2. Vers une conversion MARC/RDF sans perte
    - Pouvoir faire l'aller-retour
    - Pour se passer de MARC
    - Il faudra sacrifier une partie des données, mal structurée *sudoc:contenuHistorique*
- Interroger tout le Sudoc en Sparql

# Agréger et traiter des données hétérogènes en RDF (2013- )



« hub de métadonnées » en construction » à l'ABES

# Cataloguer dans le contexte du web de données (201?- )

- Tout est à imaginer et à faire
- Beaucoup d'obstacles et d'inconnues
- S'inspirer des autres métiers : industrie automobile, sciences et industries du vivant...

# Cataloguer dans le contexte du web de données (201?- )

<http://editor.bibframe.zepheira.com/static/>

- Partir du prototype Bibframe editor
  - Outil RDF générique (pas slt Bf)
  - Gestion des profils (vocabulaires)
  - Gestion des lookups (autorités)
- Aller plus loin :
  - Dériver (sans forcément recopier en local)
    - Sans modifier
    - En modifiant
      - Ajouter une info
      - Lier à IdRef > pas modifier la notice biblio, mais enrichir les référentiels
      - Ajouter un nouveau champ (= on modifie le formulaire ! pas slt le remplir)
  - Exemplariser
    - Exemple
    - Champs locaux

## BIBFRAME Editor

Get the code [here](#). In the meantime, select a Profile below to get started.

### BF Generic Profiles

[New Helditem](#)

[New Instance](#)

[New Work](#)

[New Work, Instance,  
& Helditem](#)

### Instance

Instance of

Work

Authorized access point

Authorized access point

Property value

Property value

Any title

Any title

Title variation

Title Entity

Transcribed title

Transcribed title

Abbreviated title

Title Entity

Key title

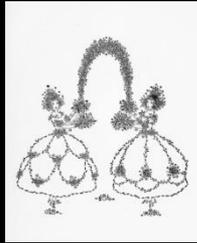
Title Entity

**QUE FAIRE EN ATTENDANT ?**

# En savoir plus

- **Infos** : <http://punktokomo.abes.fr/tag/semantique-web/>
- **Outil** : [inspector.sindice.com](http://inspector.sindice.com)
- **Tutos** : <http://web-semantique.developpez.com/tutoriels/>





par J. Paget-Frederick