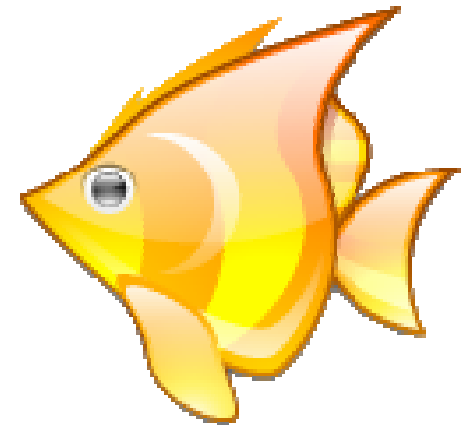


Publier les référentiels sur le Web sémantique

Tutoriel

BERMES, Emmanuelle
Centre G. Pompidou
Paris, France



Journée AFNOR / BnF
Mai 2011

Sommaire

1. Web sémantique : définitions, notions de base
2. Un vocabulaire pour les référentiels : SKOS
3. Référentiels, interopérabilité et Web de données

1998-2000 : Semantic Web road map



Tim Berners-Lee

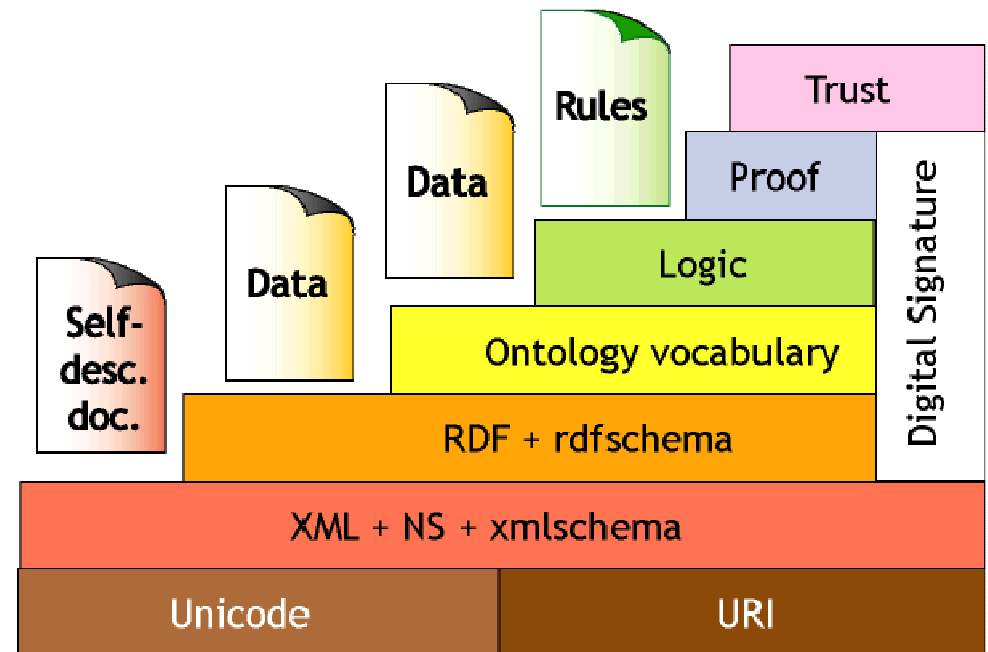
« The Semantic Web is a web of data, in some ways like a global database »

« The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation. »

Source : <http://www.w3.org/DesignIssues/Semantic.html>

Le layer cake ou Web semantic stack

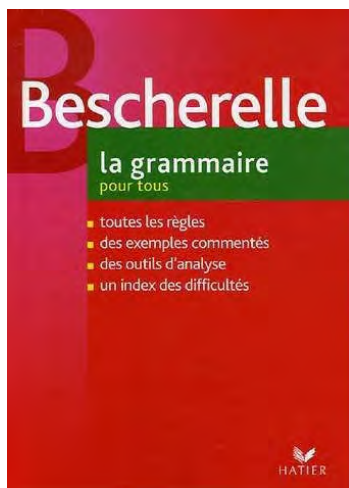
Formalisation graphique des différentes briques technologiques composant le Web sémantique d'après la vision de Tim Berners-Lee en 2000



Source : <http://www.w3.org/2000/Talks/1206-xml2k-tbl>

Le Web sémantique : un langage pour les machines

Une grammaire



RDF

Le vocabulaire



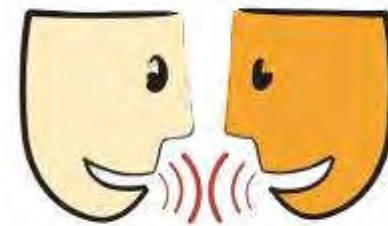
RDFS/OWL
Ontologie

Des règles
d'écriture



RDF/XML
N3, Turtle
RDFa

Des moyens
de communication



SPARQL

Les URI

UNIFORM RESOURCE IDENTIFIER

Syntaxe :

<scheme:chaîne/de.caractère>

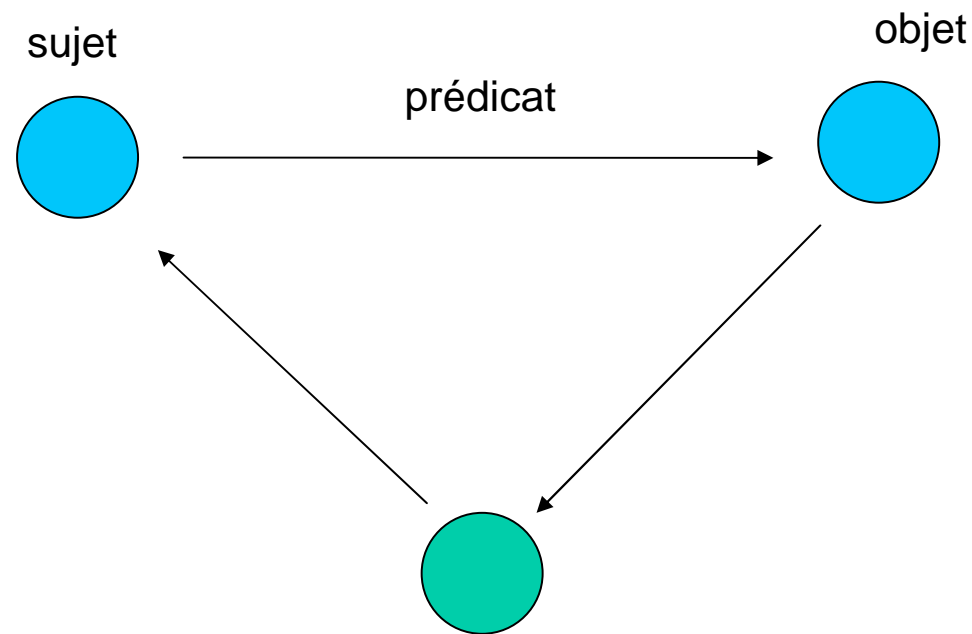
<http://monsite.com/dossier/fichier.html#ancree>

<http://mondomaine.org/ressource/1234>

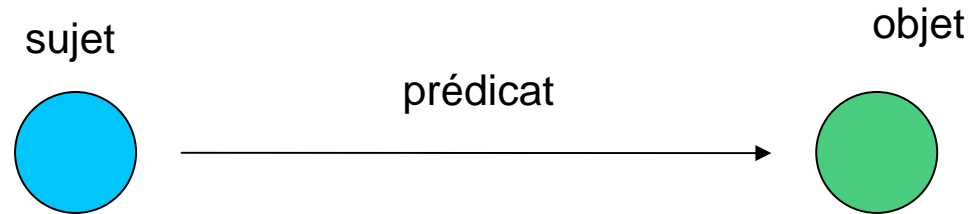
NB : Les URI sont des identifiants Web. Les URL sont des URI qui identifient la ressource par le moyen d'y accéder. Les URL peuvent être pérennes.

Le principe de RDF

RESOURCE DESCRIPTION FRAMEWORK



Le principe de RDF



Le SUJET est toujours une URI.

Toute “chose” sur laquelle on veut faire des assertions (sujet) doit avoir une URI.

Les “choses” ont toujours un type : une CLASSE.

Le PREDICAT est toujours une URI.

Il permet d'exprimer les PROPRIÉTÉS des “choses”, ou les RELATIONS des “choses” entre elles.

L'OBJET peut être un texte (LITTÉRAL) ou une URI.

Les CLASSES et les PROPRIÉTÉS sont déclarées dans des VOCABULAIRES pour être réutilisées

RDFS, OWL



- RDFS et OWL sont des standards RDF permettant de déclarer des classes, des propriétés, et leur comportement
=> on parle de vocabulaires RDF ou d'ontologies

- Ex. de classe : AGENT
- Ex de propriété de type de données : NAME
- Ex. de propriété d'objet : KNOWS
- Ex. de comportement : propriété symétrique



- Ex. de vocabulaire de classes et de propriétés : FOAF



La même chose avec des URI



- <http://xmlns.com/foaf/0.1/Agent>
- <http://xmlns.com/foaf/0.1/name>
- <http://xmlns.com/foaf/0.1/knows>
- <http://xmlns.com/foaf/0.1/>



ABC



Les référentiels et le Web sémantique : 3 définitions



Dataset (jeu de données)

“A dataset is a collection of structured metadata – descriptions of things, such as books in a library. Library records consist of statements about things (...)”

Metadata element set (jeu de métadonnées)

“A metadata element set defines classes of entities and attributes (elements) of entities. (...) such element sets are materialized through (RDF) schemas or (OWL) ontologies, the term "RDF vocabulary" being sometimes used as an umbrella for these. (...)”

Value vocabularies (référentiel de valeurs)

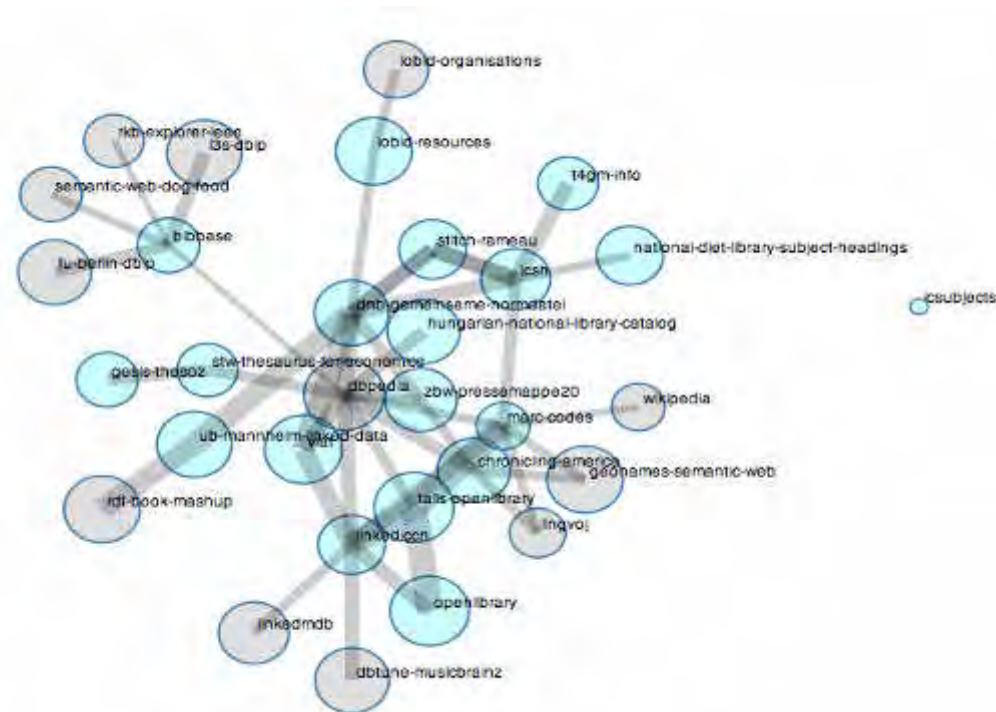
“A value vocabulary defines resources (topics, art styles, authors) that are used as values of elements in metadata records. (...) A value vocabulary thus represents a "controlled list" of allowed values for an element. Resources that can be considered as value vocabularies include: thesaurus, code list, term list, classification scheme, subject heading list, taxonomy, authority file, digital gazetteer, concept scheme, and other types of knowledge organisation system. (...) value vocabularies often have http URIs assigned to the value, which would appear in a metadata record instead of or in addition to the literal value.”

Exemples

- Jeux de métadonnées : Dublin Core, RDA, FOAF...
 - Référentiels de valeurs : LCSH, VIAF, Marc Code Lists...
 - Jeux de données : Dbpédia, Géonames...
- => la frontière peut être floue entre ces catégories !

Recensement des jeux de données en relation avec les bibliothèques dans CKAN :

<http://ckan.net/group/ild>



Un vocabulaire pour les référentiels : SKOS

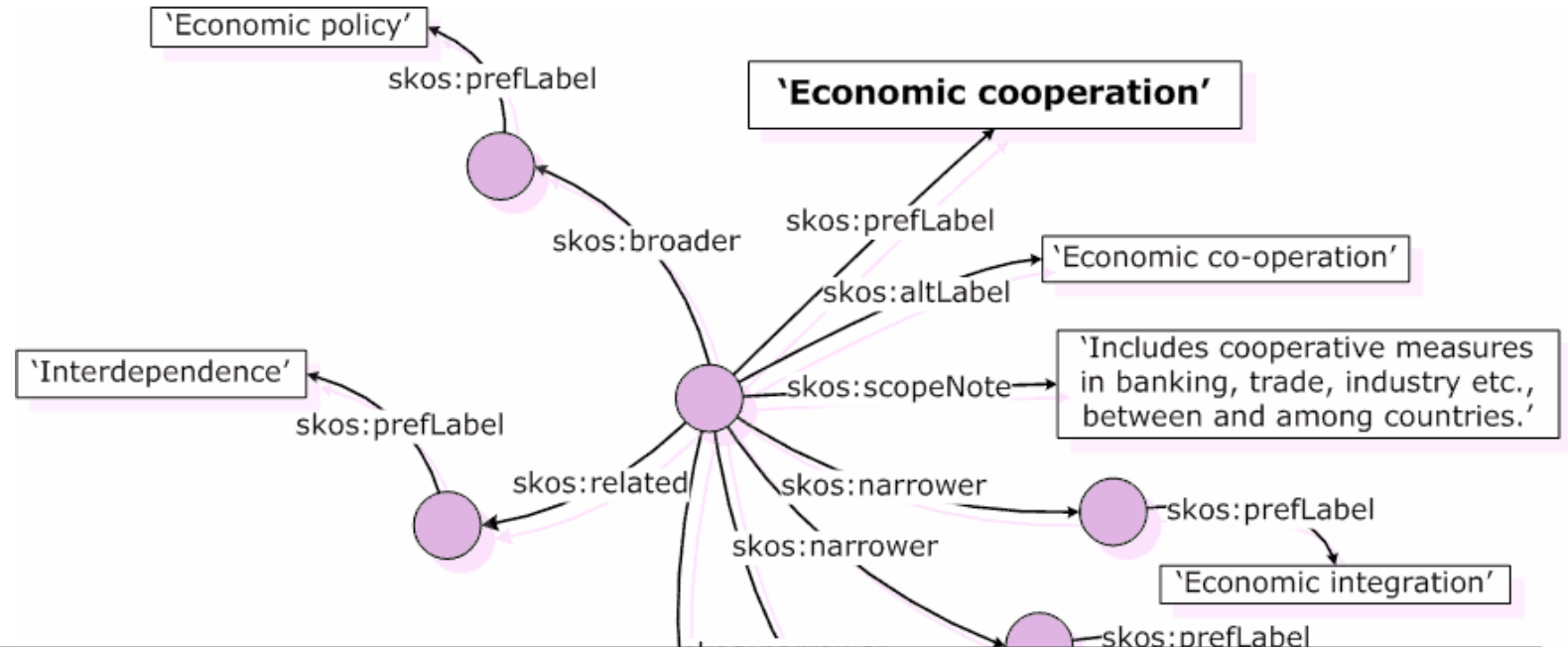
SKOS est un vocabulaire RDF permettant de décrire des systèmes d'organisation des connaissances simples : par ex. vocabulaires contrôlés ou des thésaurus. SKOS est normalisé au W3C.

S
Simple

K
Knowledge

O
Organization

S
System



Un thésaurus est un type particulier d'ontologie.
SKOS est un vocabulaire RDF permettant de décrire des thésaurus.
SKOS n'est pas un concurrent de RDFS/OWL.
La structure de SKOS est décrite en OWL

prefix skos: <<http://www.w3.org/2004/02/skos/core#>>

skos:prefLabel

'Industrial cooperation'

Les principaux éléments de SKOS

ConceptScheme : Rameau

Les LABELS :



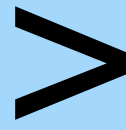
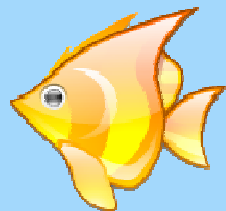
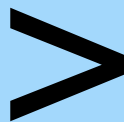
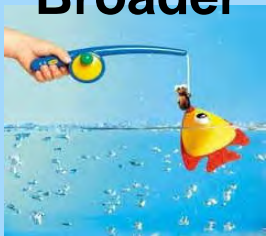
PrefLabel :
“Poisson”@fr
“Fish”@en

AltLabel :
“Poisson (aliment)”

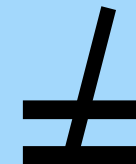
HiddenLabel :
“Poissno”

Les RELATIONS :

Broader



Narrower

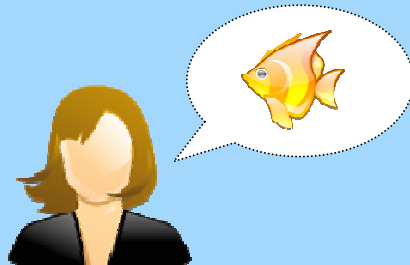


Related



Les NOTES :

Note
ScopeNote
Definition...



Les éléments pour
l'ALIGNEMENT :

CloseMatch
ExactMatch



La même chose avec des URI

 <<http://catalogue.bnf.fr/ark:/12148/cb11975896k>>

<<http://www.w3.org/2004/02/skos/core#inScheme>> <<http://rameau.bnf.fr>>

<<http://www.w3.org/2004/02/skos/core#prefLabel>> “Poisson” @fr

<<http://www.w3.org/2004/02/skos/core#prefLabel>> “Fish” @en

<<http://www.w3.org/2004/02/skos/core#altLabel>> “Poisson (aliment)” @fr

<<http://www.w3.org/2004/02/skos/core#hiddenLabel>> “Poissno”

<<http://www.w3.org/2004/02/skos/core#broader>> <<http://catalogue.bnf.fr/ark:/12148/cb11946987h>>



<<http://www.w3.org/2004/02/skos/core#narrower>> <<http://catalogue.bnf.fr/ark:/12148/cb12004710n>>



<<http://www.w3.org/2004/02/skos/core#related>> <<http://catalogue.bnf.fr/ark:/12148/cb11933179r>>

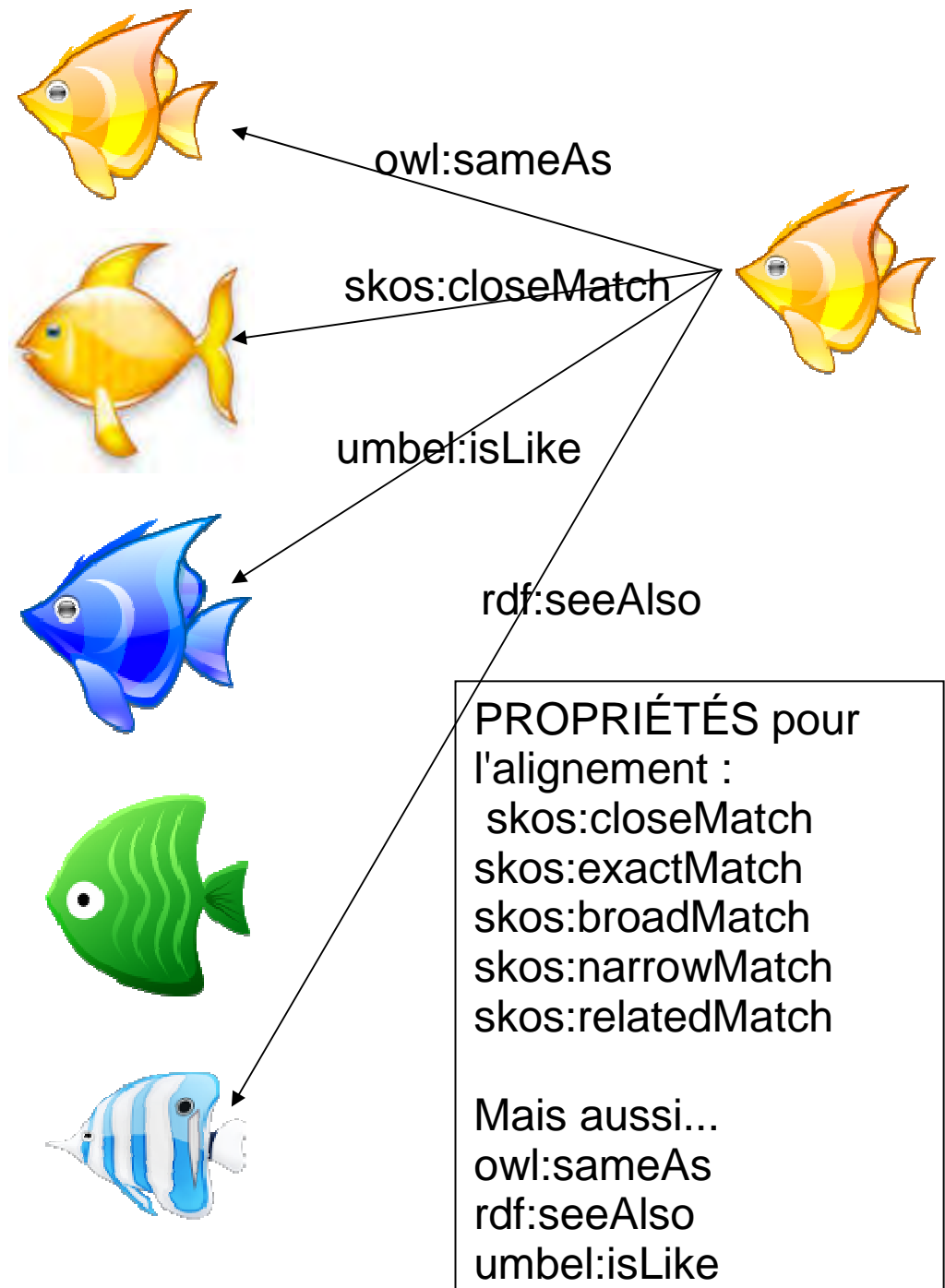


<<http://www.w3.org/2004/02/skos/core#exactMatch>> <<http://id.loc.gov/authorities/sh85048617#concept>>



Les alignements...

- Méthodes
 - Manuelle
 - Automatique
 - Semi-automatique
- Les aides pour l'alignement automatique :
 - Identifiants
 - Contexte (ex. de VIAF)

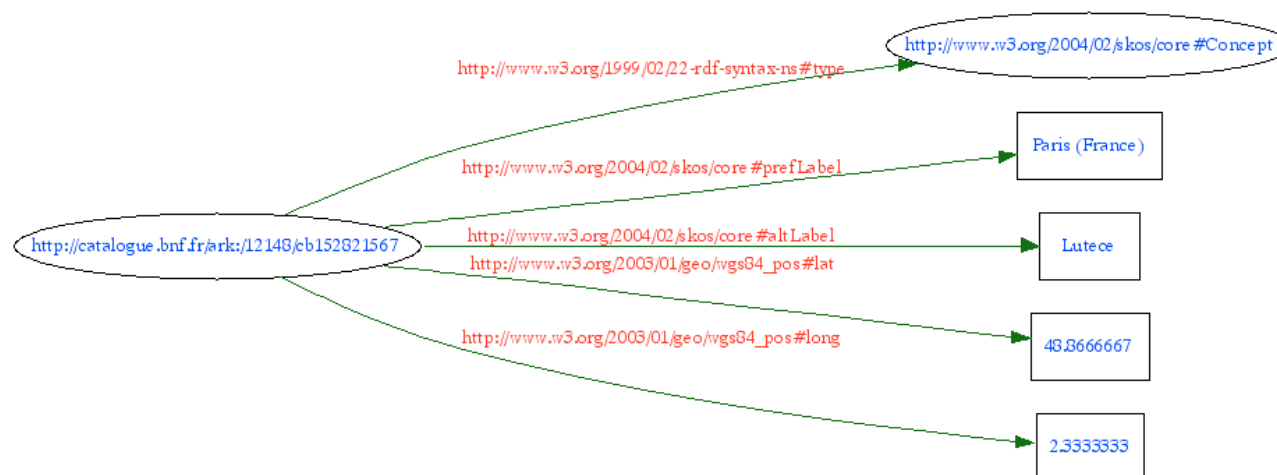


Étendre SKOS

- Avantage du SemWeb : on n'est pas tenu à l'usage rigide d'une seule grammaire comme en XML...
- Mais on peut combiner différents vocabulaires

Exemple avec un référentiel géographique

Number	Subject	Predicate	Object
1	http://catalogue.bnf.fr/ark:/12148/cb152821567	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/2004/02/skos/core#Concept
2	http://catalogue.bnf.fr/ark:/12148/cb152821567	http://www.w3.org/2004/02/skos/core#prefLabel	"Paris (France) "
3	http://catalogue.bnf.fr/ark:/12148/cb152821567	http://www.w3.org/2004/02/skos/core#altLabel	"Lutece"@fr
4	http://catalogue.bnf.fr/ark:/12148/cb152821567	http://www.w3.org/2003/01/geo/wgs84_pos#lat	"48.8666667"
5	http://catalogue.bnf.fr/ark:/12148/cb152821567	http://www.w3.org/2003/01/geo/wgs84_pos#long	"2.3333333"



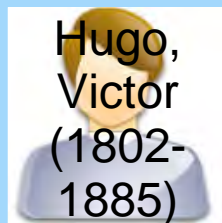
Extensions de SKOS

SKOS XL



FOAF:FOCUS

Créé par la BnF
Formes préférées
Formes rejetées



`foaf:focus`



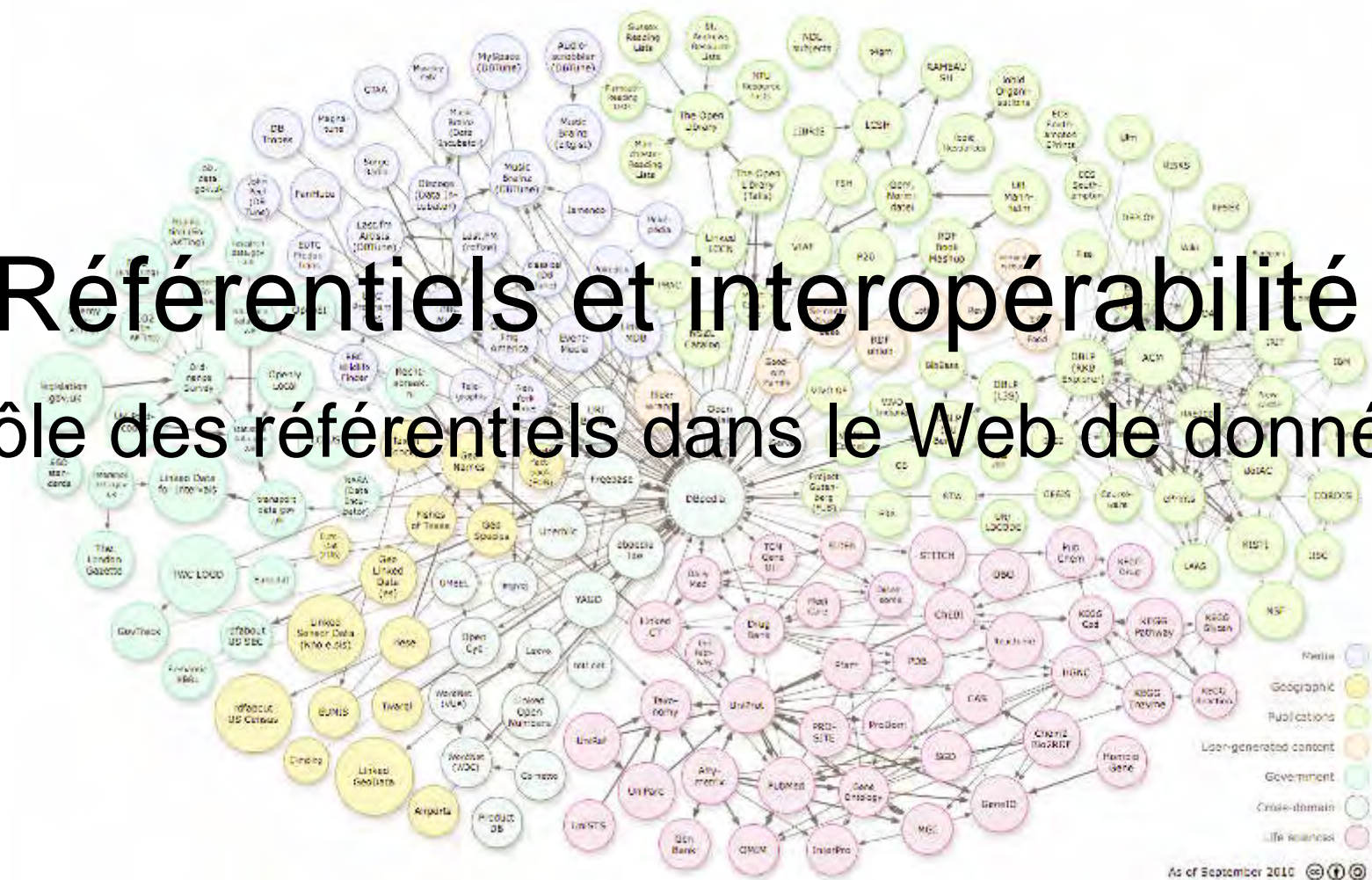
Né en 1802
Mort en 1855
Écrivain

Une entité "concept" à laquelle se rattachent les assertions spécifiques au contexte du thésaurus

Une entité "personne" à laquelle se rattachent les informations qui sont vraies dans tout contexte

Référentiels et interopérabilité

Rôle des référentiels dans le Web de données



Principes du « Linked Data »

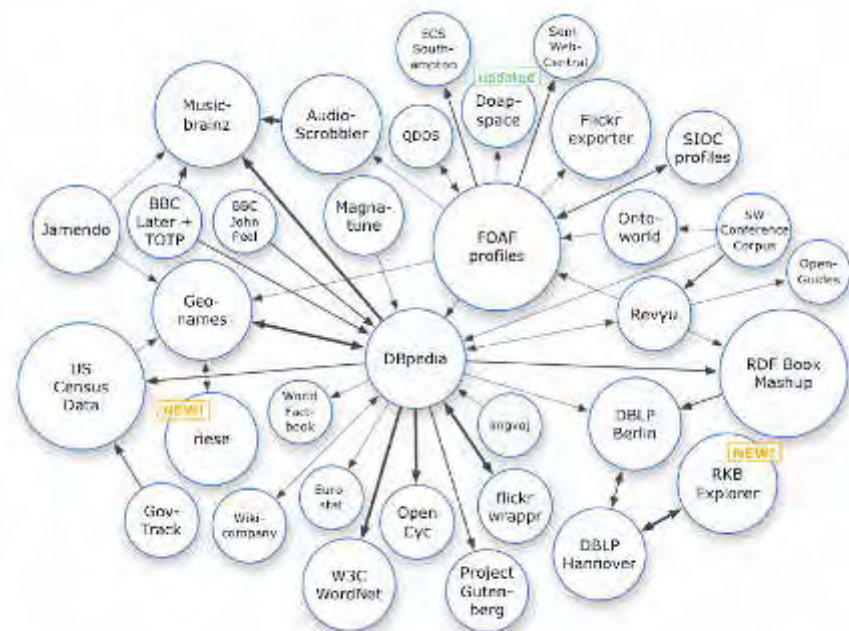


Utiliser des URIs

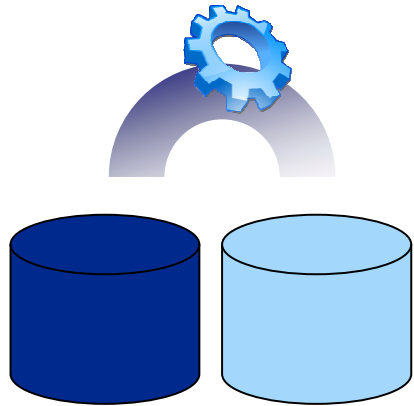
Utiliser des URIs accessibles via HTTP

Donner l'accès aux données utiles en utilisant les standards SPARQL et RDF

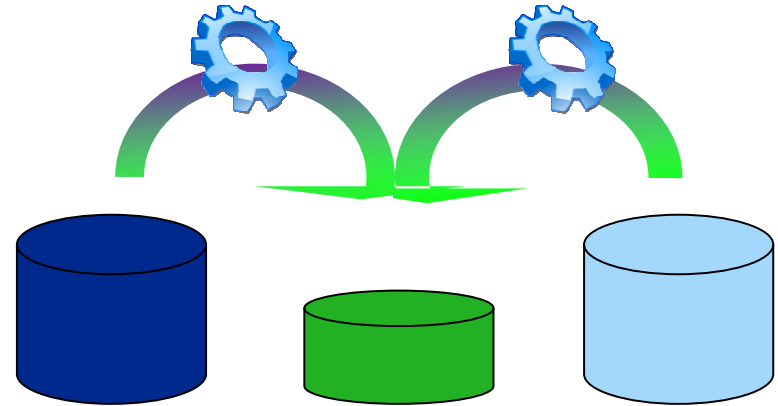
Exprimer l'URI des objets liés



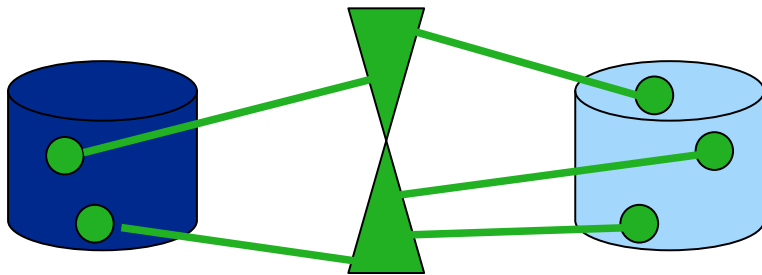
Des “styles” d'interopérabilité



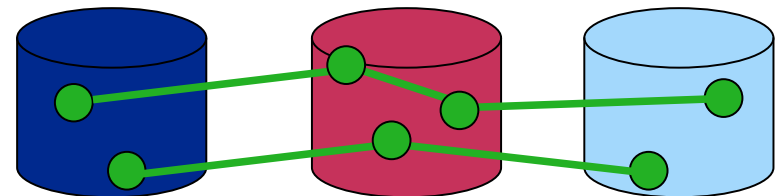
Conversion complexe
Fusion des bases sources



“Plus petit dénominateur commun”
Création d'une nouvelle base avec format pivot

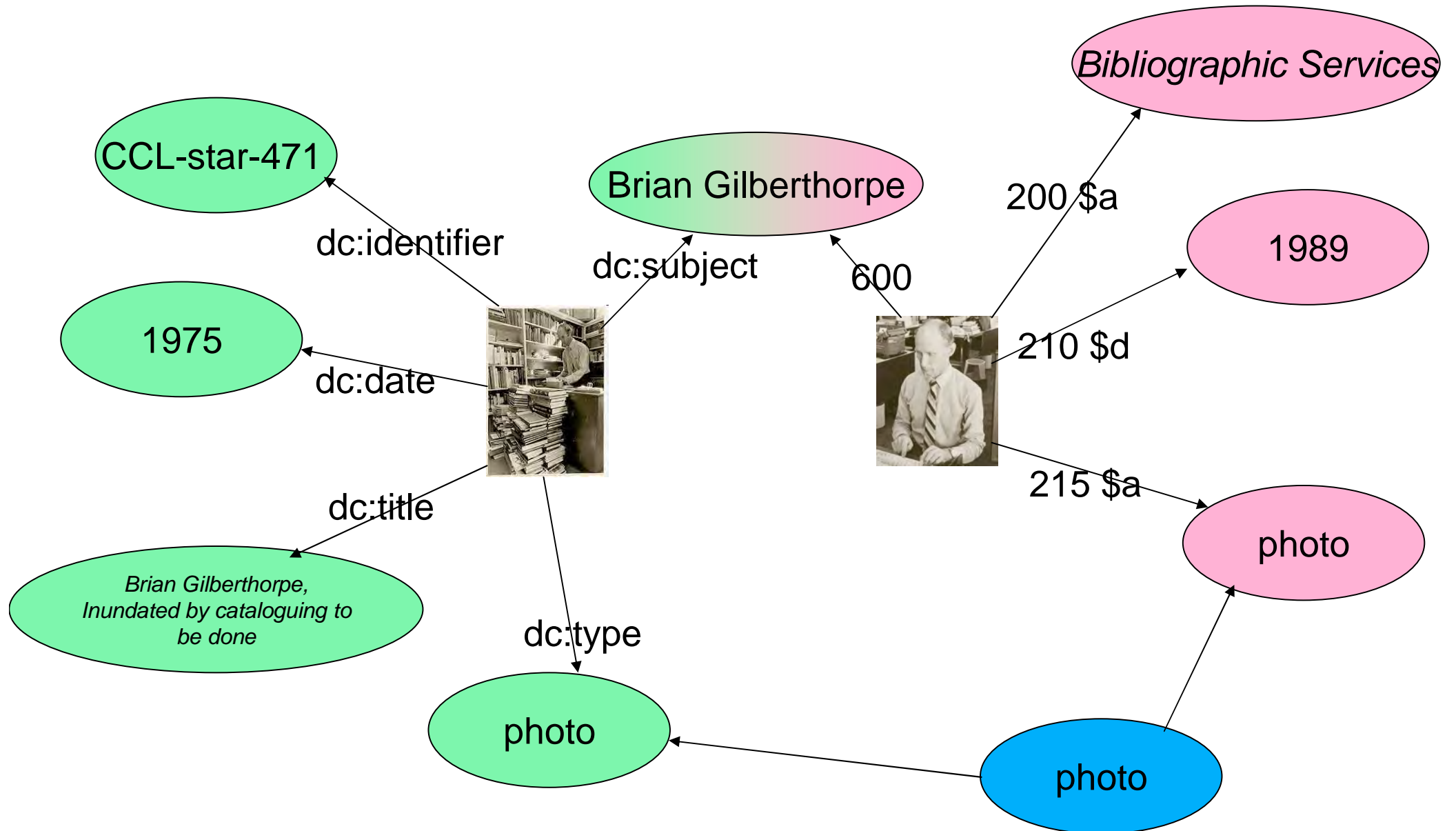


“hub and spoke”
Bases différentes, référentiels communs

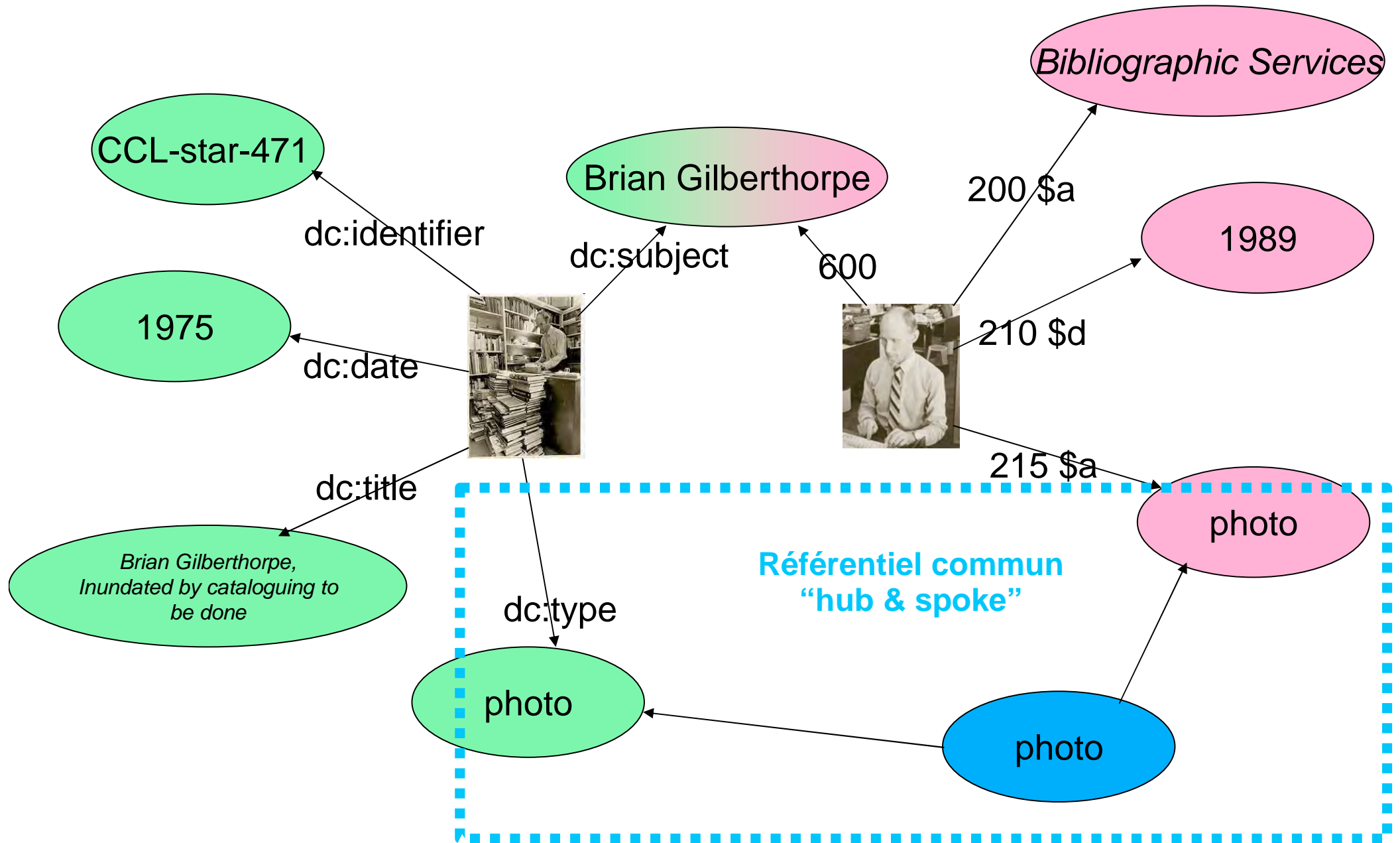


“follow your nose”
Interopérabilité par les liens

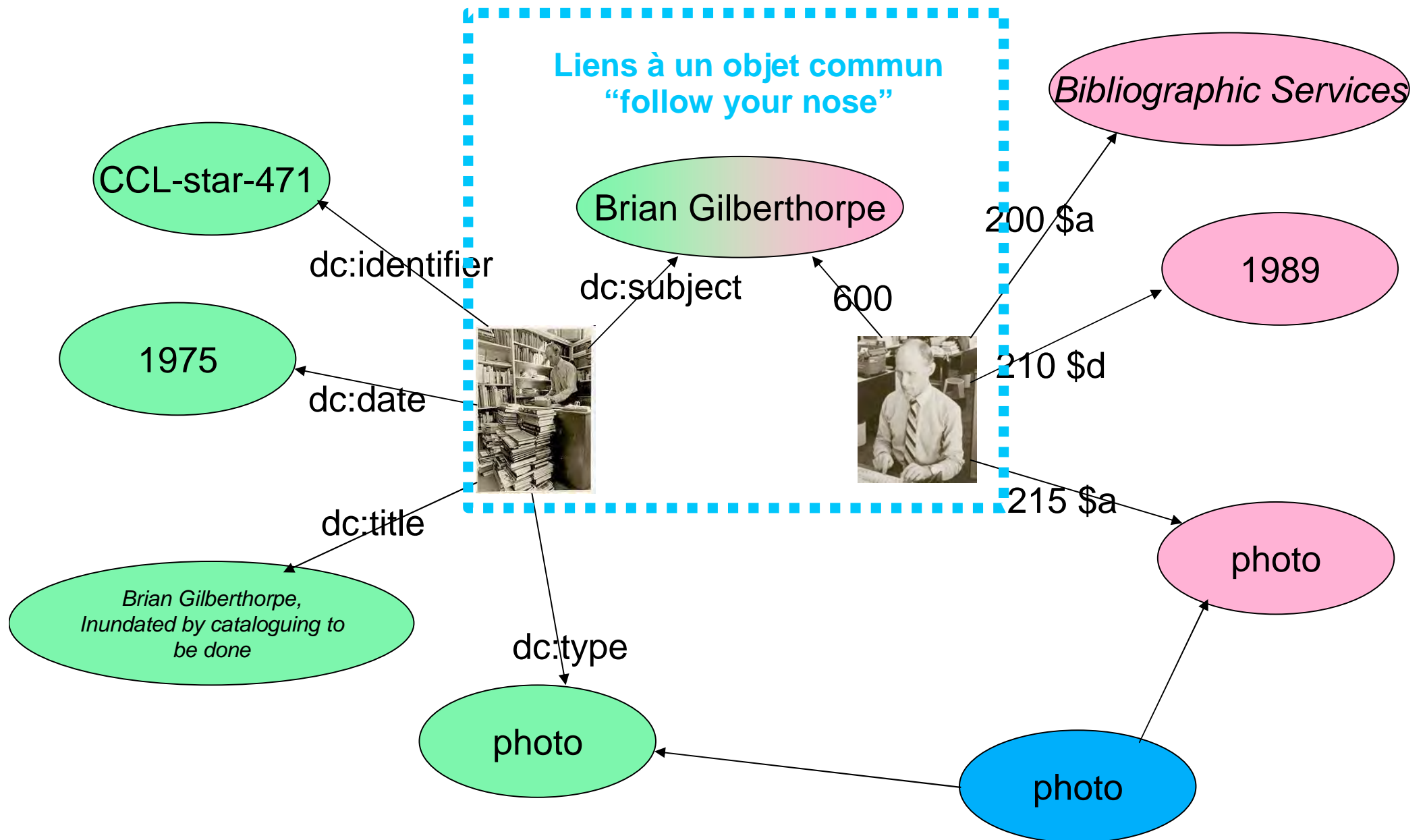
Interopérabilité par les liens



Interopérabilité par les liens

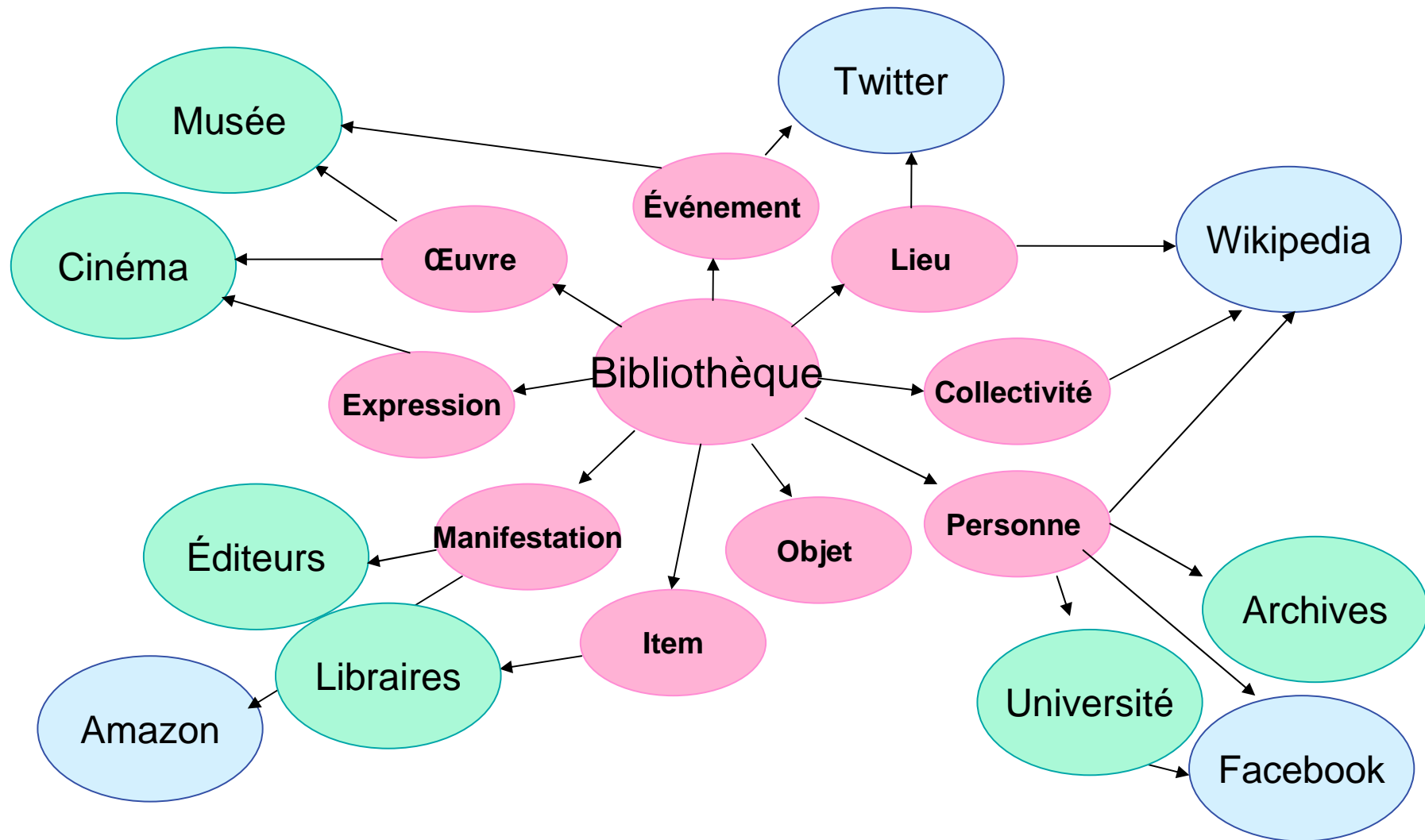


Interopérabilité par les liens



A quoi ça sert ?

Un espace global d'information



Références

- Skos primer (2009) : <http://www.w3.org/TR/skos-primer/>
- Vocabulary mapping framework (2009) :
<http://cdlr.strath.ac.uk/VMF/documents.htm>
- Bob Ducharme, *Improve your taxonomy management using the W3C SKOS standard* (Mai 2011) :
<http://www.ibm.com/developerworks/xml/library/x-skostaxonomy/index.html>
- Tom Heath, Chris Bizer, *Linked Data: Evolving the Web into a Global Data Space* (2011) :
<http://linkeddatabook.com/editions/1.0/>
- Library Linked Data incubator group (2011) :
http://www.w3.org/2005/Incubator/ld/wiki/Main_Page