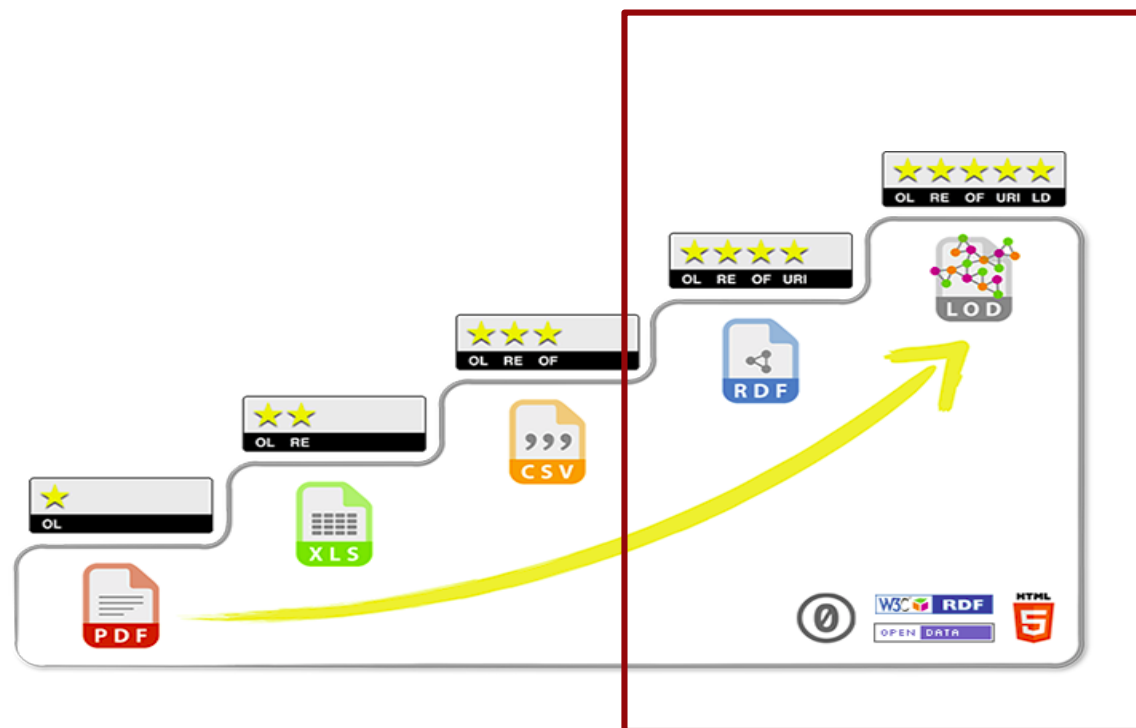




Données ouvertes et biodiversité



Vers des données ouvertes



Private
Data



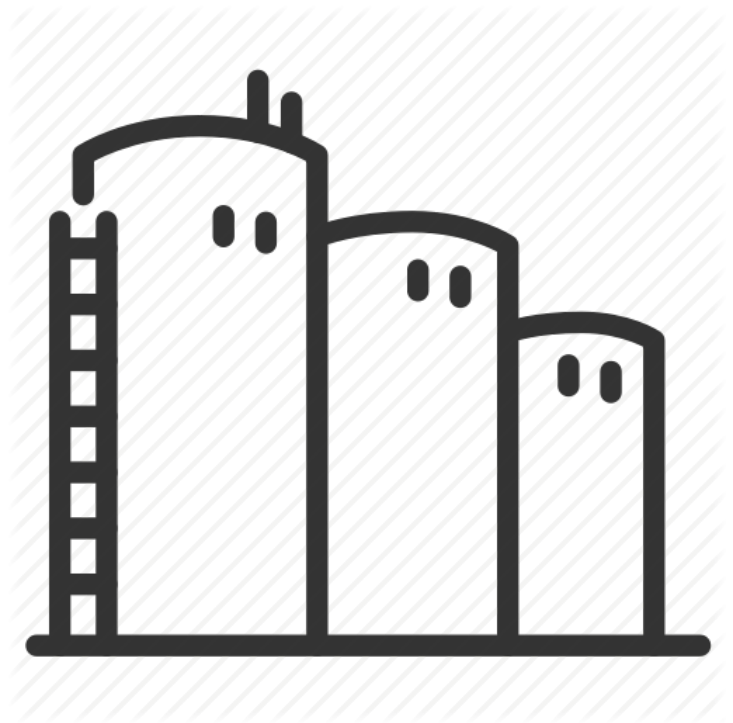
Open
Data



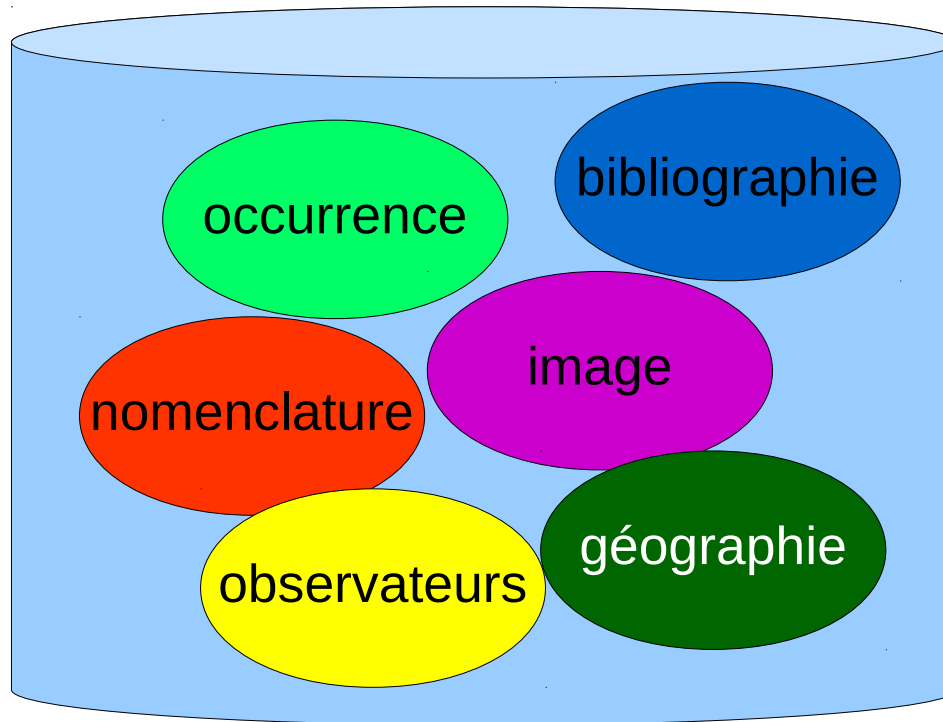
Linked
Open Data

Approche traditionnelle

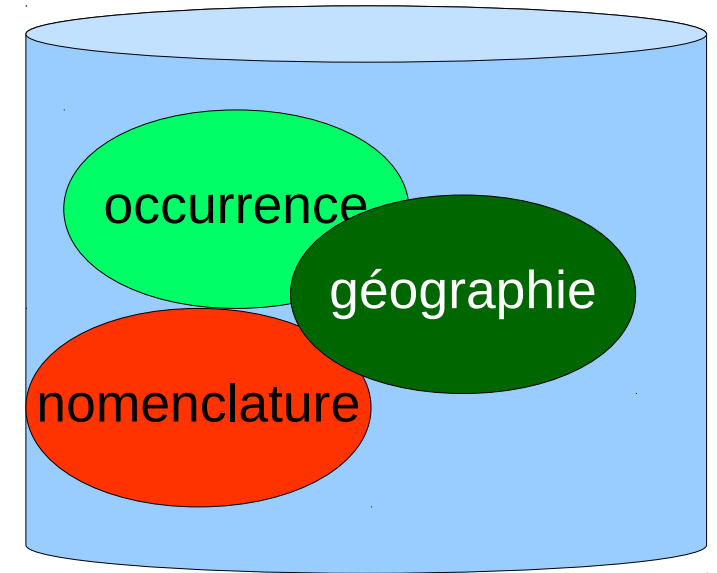
Silo de données



Silos de Données



Structure 1



Structure 2

- Redondance des données (taxonomie, géographie, ...)
- Nettoyage indépendant
- Mise à jour manuelle

Les efforts ne sont pas partagés

Structuration propre

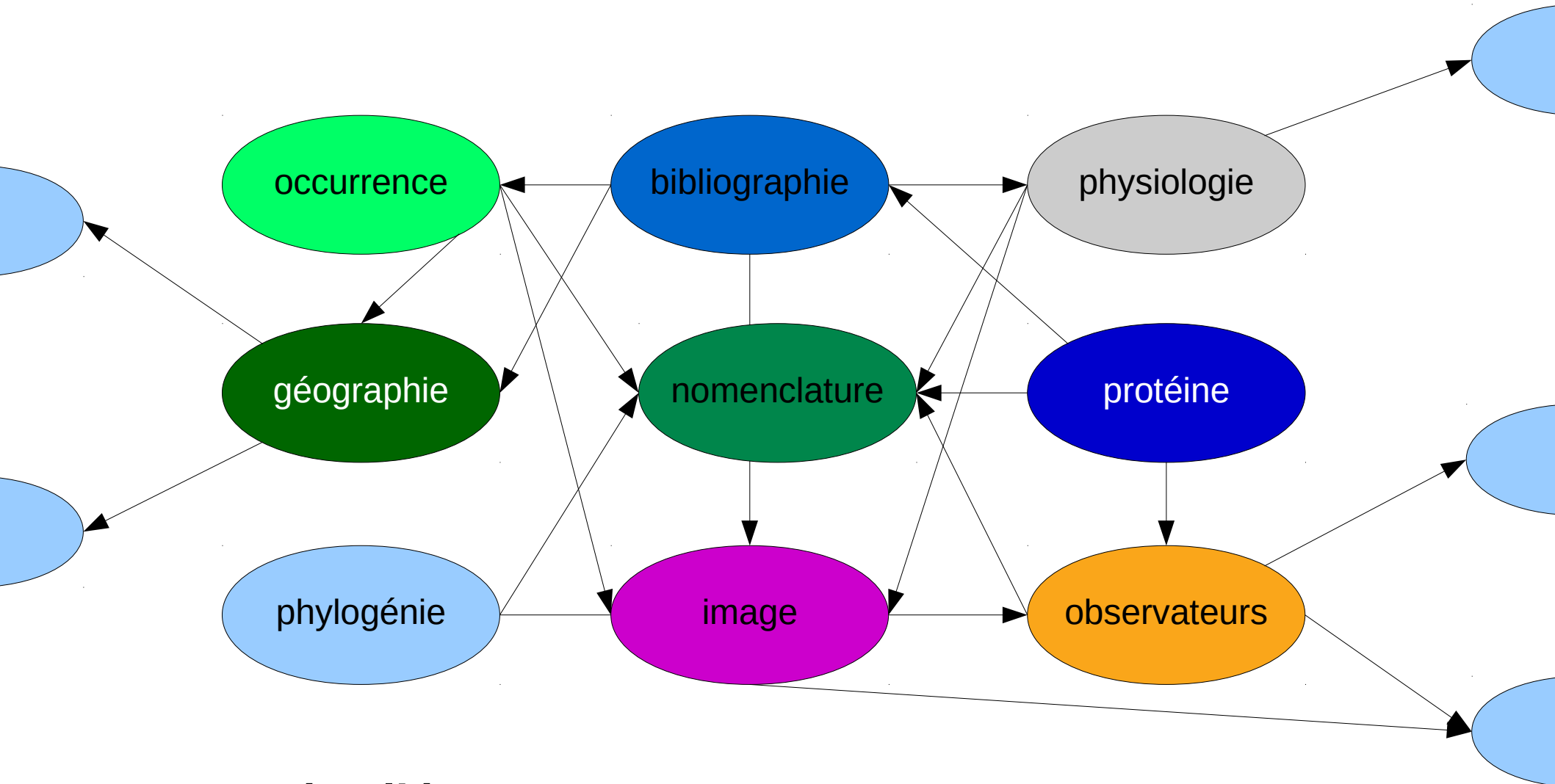
- Champs différents
- Différents nom pour le même champ / Même nom pour différent champs
- Différentes façon d'interpréter un champ
- Vocabulaires différents

**Demande un effort de compréhension
pour chaque source de données**

Vers des données liées



Données liées



Données liées :

- éviter la redondance
- facilité l'exploration et l'émergence de nouvelles données

Vers des données liées

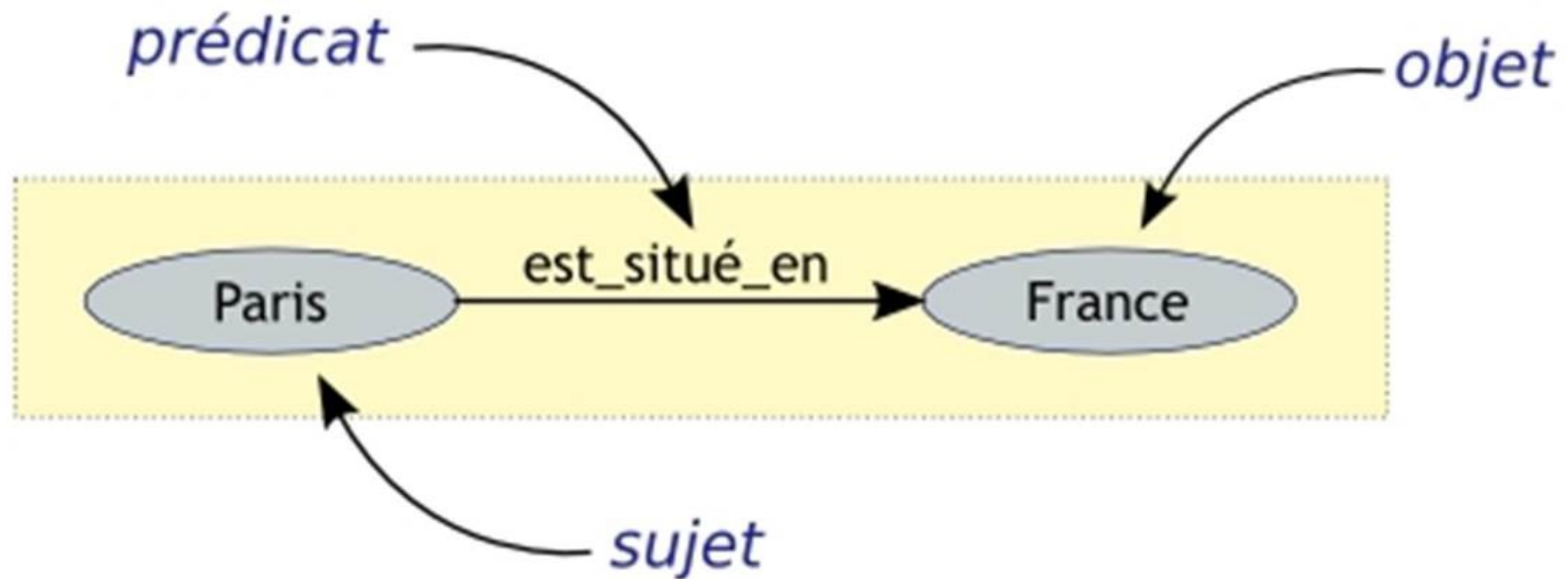
- **Utilisation de standards :**
 - > parler le même langage
 - > profiter de la réflexion de la communauté
 - > favoriser l'échange des données



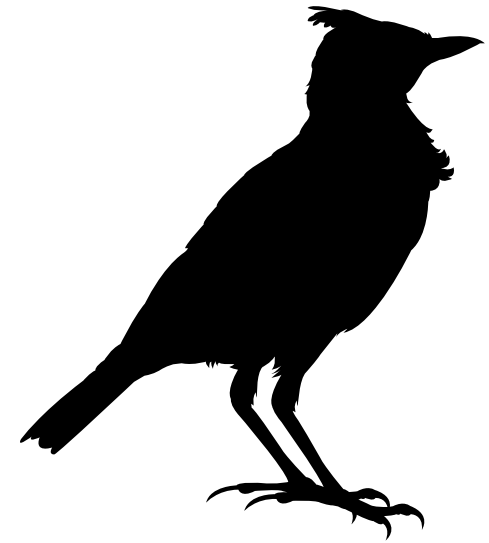
« le Web sémantique fournit un modèle qui permet aux données d'être partagées et réutilisées entre plusieurs applications, entreprises et groupes d'utilisateurs » W3C

Structuration des données : Principes

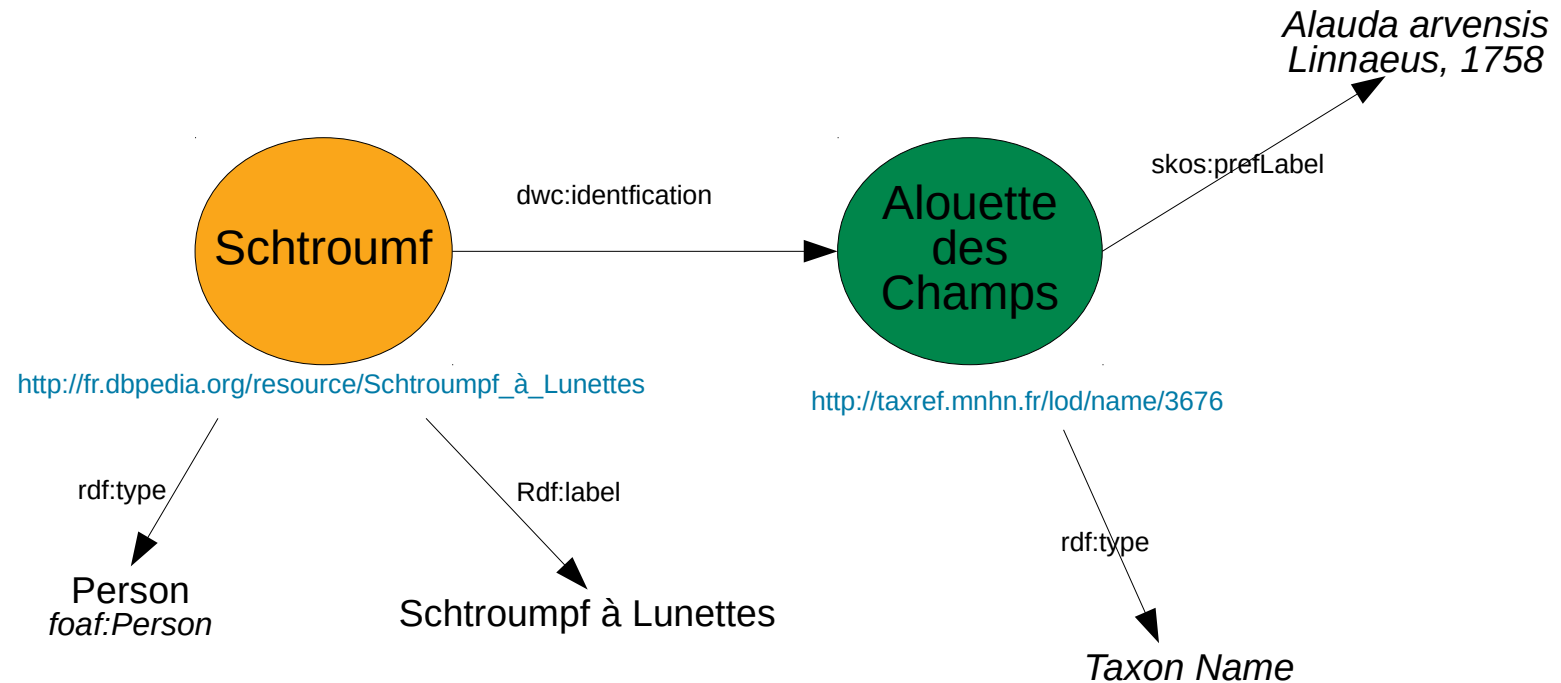
- Triplet de données
- Données sous forme de graph
- Utilise des vocabulaires existants (ontologies)
- Utiliser des adresses URI HTTP qui existent sur le Web



Schtroumf à lunette a identifié une alouette



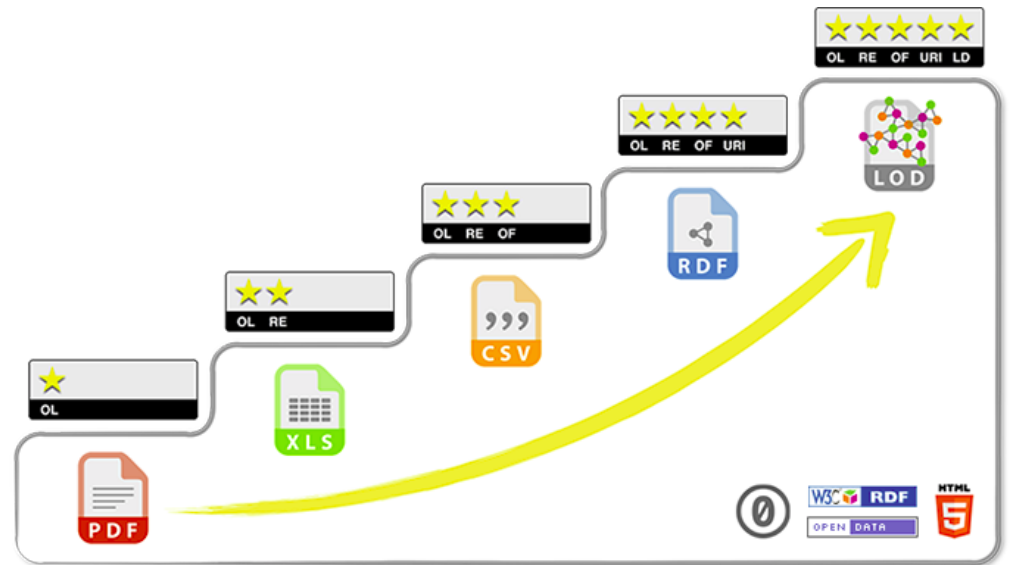
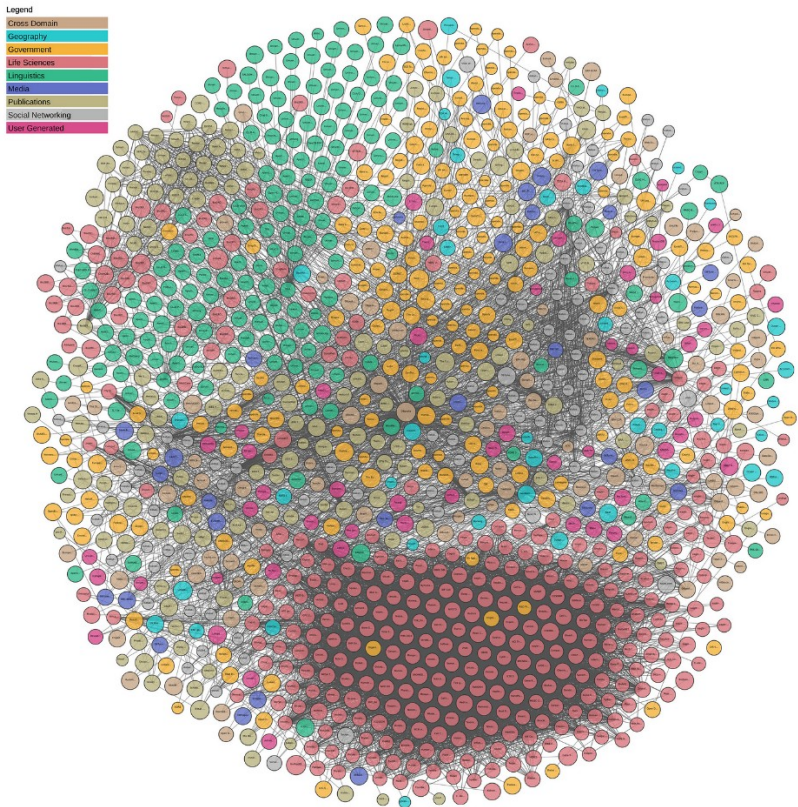
Structuration des données



LOD

● 2 niveaux

- > Structuration sémantique des données ★★☆☆
=> Utilisation de standards
- > Lier les données ★★★★★
=> Utilisation de ressources existantes



LOD et biodiversité

Steller's sea eagle

Steller's sea eagles are native to eastern Russia, inhabiting coastal cliffs and estuaries where they can easily access good fishing territories. They feed primarily on salmon, which they catch by swooping from perches located by the water's edge. Pairs are monogamous and hatch an average of two chicks each season, although crows and martens commonly take both eggs and young birds from the nest. During winter a small number of birds remain in Russia to tough it out, but the majority fly south to Japan.



Scientific name: *Haliaeetus pelagicus*

- http://dbpedia.org/resource/Steller's_Sea_Eagle

programme

Steller's sea eagle

www.bbc.co.uk

Lunch on the wing

www.bbc.co.uk

elsewhere on the web

animaldiversity.ummz.umic.edu

en.wikipedia.org

www.arkive.org

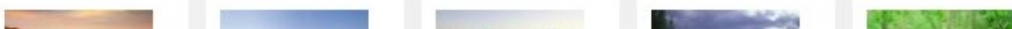
www.birdlife.org

www.flickr.com

www.iucnredlist.org

www.natural-research.org

habitat



Standards en biodiversité - Occurences

- Darwin-SW : DEE
- OBOE ; PO, ... : Détail des informations collectés

Graph model basis of 2014 Darwin-SW (DSW) ontology (version 0.4)

Steven J. Baskauf – Vanderbilt University
Campbell O. Webb – Arnold Arboretum of Harvard University

Basic class relationships laid out by Richard L. Pyle
<http://lists.tdvw.org/pipermail/tdvwg-content/2010-October/001703.html>

Key:

namespace:property property (italicized)
namespace:Class instance of named class
 → property arc from subject to object
 ↔ inverse property pair; arrow with property name shows direction

Colors:

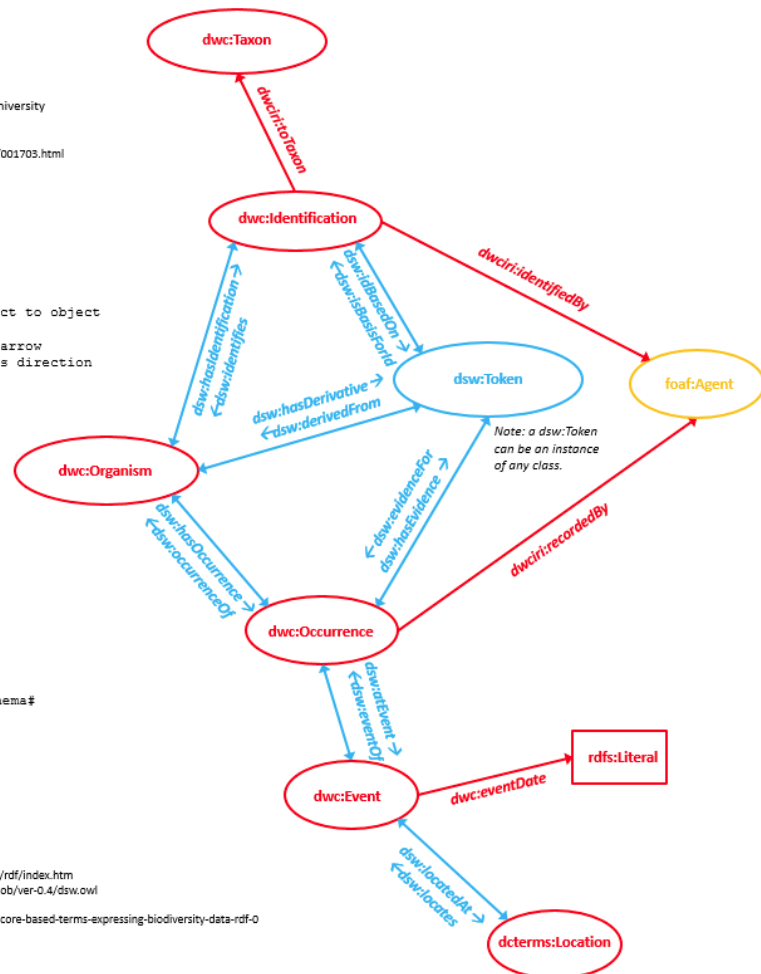
red=Darwin Core
 blue=Darwin-SW
 yellow=FOAF vocabulary

Namespace abbreviations:

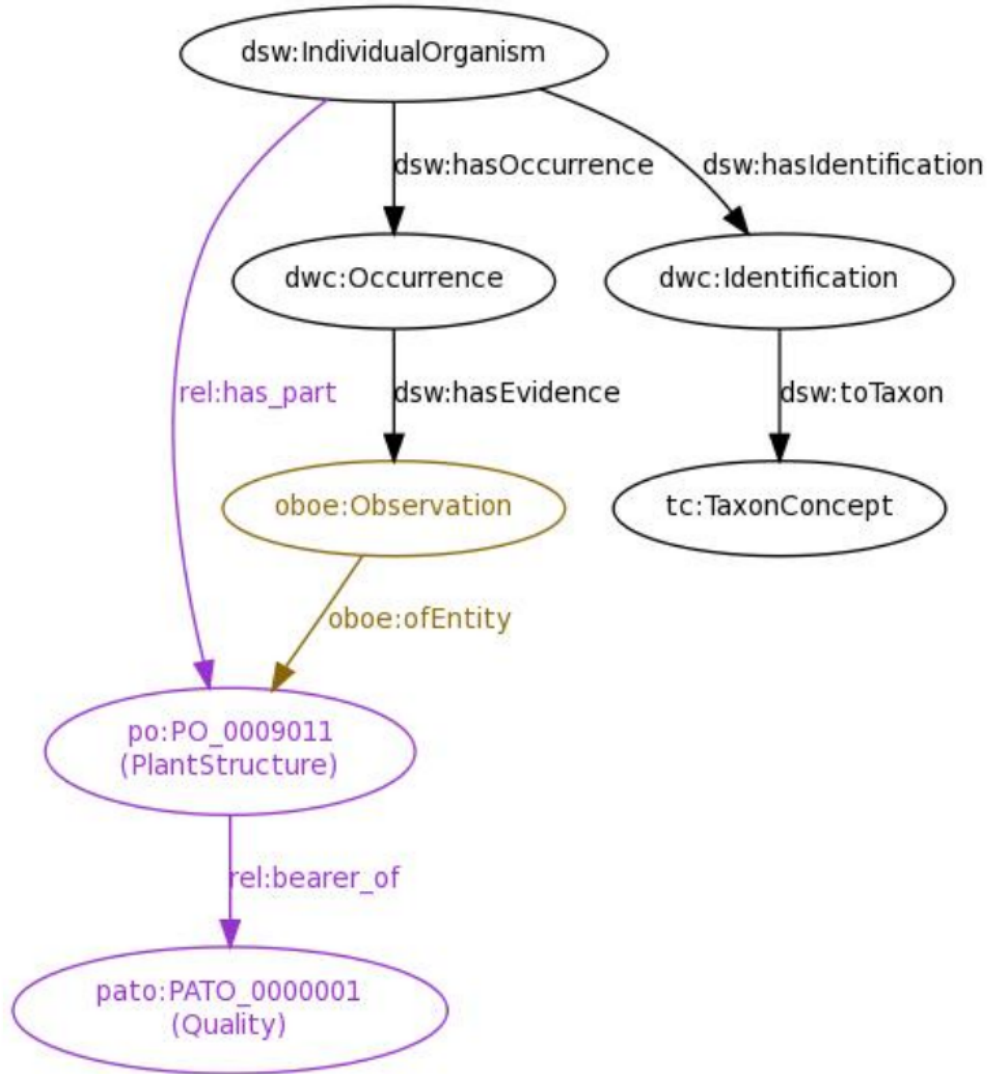
rdfs: <http://www.w3.org/2000/01/rdf-schema#>
dwc: <http://rs.tdvw.org/dwc/terms/>
dwciri: <http://rs.tdvw.org/dwc/iri/>
dsw: <http://purl.org/dsw/>
dcterms: <http://purl.org/dc/terms/>
foaf: <http://xmlns.com/foaf/0.1/>

Sources:

Darwin Core RDF Guide from <http://rs.tdvw.org/dwc/terms/guides/rdf/index.htm>
 Darwin-SW version 0.4 from <https://github.com/darwin-sw/dsw/blob/ver-0.4/dsw.owl>
 Darwin-SW is described in <http://www.semantic-web-journal.net/content/darwin-sw-darwin-core-based-terms-expressing-biodiversity-data-rdf-0>



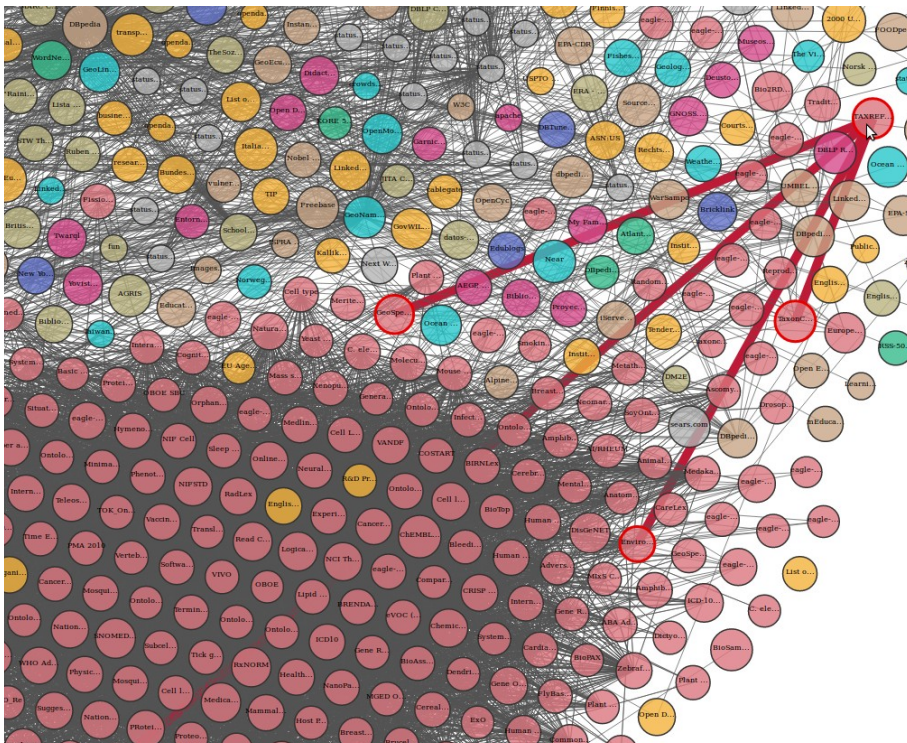
Mixer les standards



Standards en biodiversité - Taxonomie

- TaxonConcept, NCBI,
- Dbpedia...
- Taxref-LD

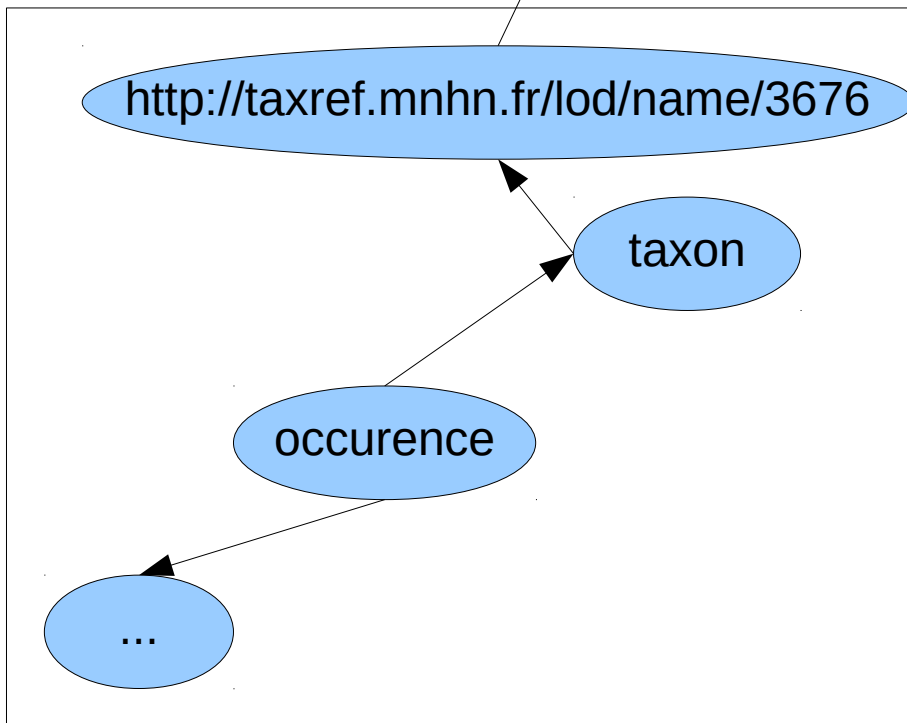
⇒ Jeux de données liés ouvre toutes les passerelles



Alouette des champs



<https://www.wikidata.org/wiki/Q25961>



https://inpn.mnhn.fr/espece/cd_nom/3676



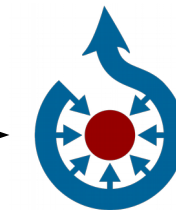
Encyclopedia of Life

<http://eol.org/pages/916864>



GBIF

<https://www.gbif.org/species/8077224>



WIKIMEDIA
COMMONS

https://commons.wikimedia.org/wiki/File:Alauda_arvensis_2.jpg

Enrichissement : exemple taxonomique



- Utilisation de wikidata pour récupérer des médias sur les taxons

```
SELECT
  ?item ?image
WHERE {
  ?item wdt:P18 ?image.
  ?item wdt:P3186 '3676'
}
```

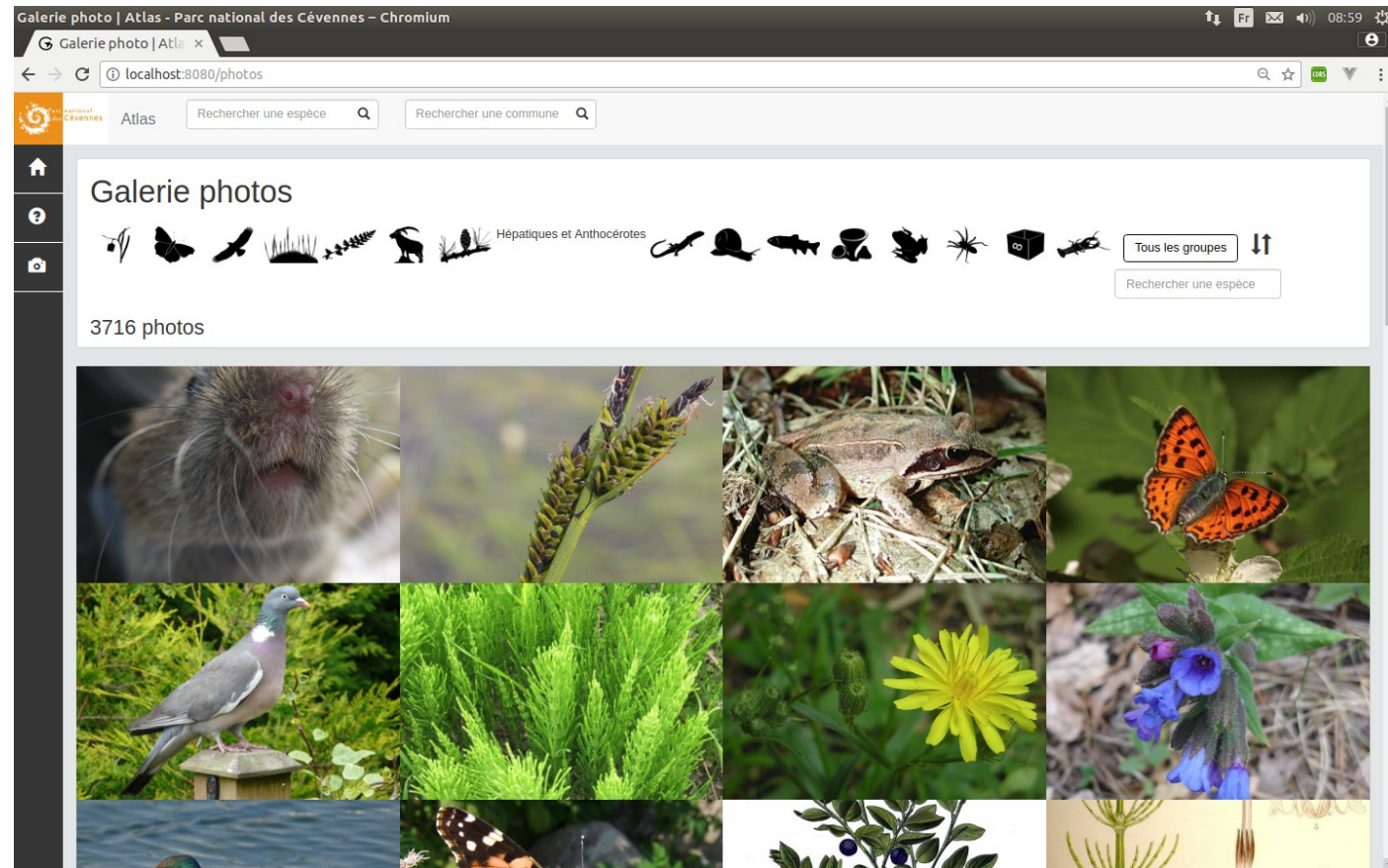


<https://query.wikidata.org>

Enrichissement : exemple taxonomique



- Exemple au PNC : Récupération de 6670 médias pour 4164 espèce sur 6680



<https://github.com/PnX-SI/TaxHub/issues/150>
import_wikimedia_commons

Publication

- Publier son jeux de données en rdf => google refine

- Sparql

 - > Requête de graph

 - > Attaque plusieurs jeux de données

Mon jeu de données → Taxref-LD → Dbpedia → UICN

```
# Récupération de données sur dbpedia de la liste des taxons des occurrences concernés
prefix skos: <http://www.w3.org/2004/02/skos/core#>
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX taxrefrk: <http://taxref.mnhn.fr/ld/taxrank/>




PREFIX dwc: <http://rs.tdwg.org/dwc/terms/>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX dsw: <http://purl.org/dsw/>
PREFIX dbo: <http://dbpedia.org/ontology/>

SELECT DISTINCT * where {
  ?taxon a dwc:Taxon; owl:sameAs ?taxrefName.
  SERVICE <http://taxref.mnhn.fr/sparql/> {
    ?taxrefName ?p ?dbpediaLink
    FILTER(regex(str(?dbpediaLink), "dbpedia", "i"))
  }.
  OPTIONAL { SERVICE <http://dbpedia.org/sparql/> {
    ?dbpediaLink dbo:conservationStatus ?uicn.
  } }.
  OPTIONAL { SERVICE <http://dbpedia.org/sparql/> {
    ?dbpediaLink dbo:thumbnail ?img.
  } }.
  OPTIONAL { SERVICE <http://dbpedia.org/sparql/> {
    ?dbpediaLink dbo:abstract ?abstract.
    FILTER (lang(?abstract) = 'fr')
  } }.
}
LIMIT 100
```

Liste des passériformes

Requete

Show 10 entries

| ID | Img | UICN | Nom scientifique | Description |
|----|--|------|-------------------------------------|---|
| 95 |  | NT | Turdus iliacus Linnaeus, 1766 | Turdus iliacus Turdus iliacusFichier:Redwing Turdus iliacus.jpg Grive maui 1758 Statut de conservation UICN LC : Préoccupation mineure La Grive m... passereaux de taille moyenne (la plus petite des grives) appartenant à la fa... n'est visible qu'en période de migration et d'hivernage ; Elle se reproduit pl |
| 3 |  | LC | Lullula arborea (Linnaeus, 1758) | Lullula arborea Pour les articles homonymes, voir Lulu. Lullula arboreaFich... arborea) Genre LullulaKaup, 1829 Nom binominal Lullula arborea(Linnaeu... : Préoccupation mineure L'Alouette lulu (Lullula arborea) est une espèce d'... elle appartient à la famille des Alaudidae. |
| 26 |  | LC | Emberiza | Emberiza citrinella Emberiza citrinellaFichier:Yellowhammer.jpg Bruant jaur |

Comment contribuer ?

- **Microdata :**

- > Insérer du sémantique dans les pages web



- **Publier ses jeux de données**

- > En RDF Darwin-SW avec Taref-LD



Conclusion

- **Open data :**

- > Il n'y a pas de petite contribution
- > Nouveau usage

- **LOD :**

- > Pratique pour l'enrichissement de données
- > Publication des données. Qui ?

