



## GéOnto 4 décembre : autres points

---

- Sur l'enrichissement de TopoCogit à partir de  $O_{Irit}$ 
  - Alignement de  $O_{Irit} \rightarrow$  TopoCogit
  - Comment utiliser les mappings pour l'enrichissement
  - Eventuelle amélioration de  $O_{irit}$
- Identification de mots vides
- Enrichissement par rapport à  $O_{pau}$  ?



# Mots vides ?

---

- Mots de modélisation
  - Nature, valeur, qualité, type, à caractère
- Mots constructeurs en cartographie
  - Surface, aire, zone (à considérer comme équivalents ??)
  - Tronçon, équipement, établissement, station, limite, liaison
  - Primitives géométriques (ligne, point, surface)
  - **bâtiment, bassin** apparaîtront comme des concepts
- A identifier en tant que tel



# Enrichissement de l'ontologie

---

A partir des concepts élaborés par l'Irit dans  $O_{Irit}$   
et des mappings issus de l'alignement de  $O_{irit}$  vers Topo-Cogit

## 3 phases

- Evaluer la qualité des concepts extraits par la phase d'acquisition (et retour vers cette phase)
- Juger de l'opportunité d'importer ou pas un concept
- Interface d'aide à l'enrichissement interactif

Étudier ce qui pourra être automatisé



# Nettoyage des concepts extraits par l'Irit

---

Aire de péage *isEq* aire de péage Sim: 1.0

Aire de péage *isEq* aire de péage Sim: 1.0

Etablissement thermal / *isEq* établissement thermal Sim: 0.999

Etablissement thermal *isEq* établissement thermal Sim: 0.999

Cathédrale *isEq* cathédrale Sim: 1.0

cathédrale *isEq* cathédrale Sim: 1.0

Terrain de sport *isEq* terrain de sport Sim: 1.0

Terrain de sports *isEq* terrain de sport Sim: 1.0

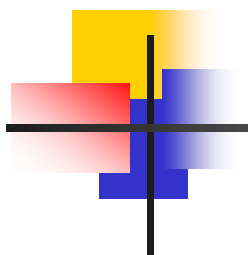
Camping / Dolmen *isA* camping Sim: 0.716

Télécabine téléphérique télésiège *isA* téléphérique Sim: 0.644

Parc municipal intercommunal départemental interdépartemental  
*isA* parc municipal Sim: 0.457

Industriel ou commercial *isA (struct)* bâtiment industriel

Label de concepts bizarres : 90 k V, Importance, Inconnu, nature, oui, artificialisé



# Importer ou pas un nouveau concept ?

Mapping de type  $\langle S \text{ isA } C_{\max} \rangle$  avec  $\text{Label}(C_{\max}) \subseteq \text{Label}(S)$

Exemple des énumérations contenant « ou » « et » « / »

$\langle \text{Arène ou théâtre antique } \textit{IsA} \text{ théâtre antique} \rangle$

$\langle \text{Plaine ou plateau } \textit{IsA} \text{ plateau} \rangle$

éliminer les concepts portant sur 2 concepts existant dans la cible

$\langle \text{gué ou radier } \textit{IsA} \text{ gué} \rangle$

$\langle \text{Funiculaire ou crémaillère } \textit{IsA} \text{ Funiculaire} \rangle$

Introduire le concept  
non présent dans  $O_C$   
comme un frère de  $c_c$

Exemple contenant des « mots vides »

$\langle \text{Batiment de type industriel } \textit{isA} \text{ bâtiment industriel} \rangle$

éliminer  
les concepts

Exemple contenant des adjectifs parenthésés

$\langle \text{canal (large) } \textit{IsA} \text{ canal} \rangle$

$\langle \text{atelier (grand) } \textit{IsA} \text{ atelier} \rangle$

Ne pas introduire le concept  
mais garder une propriété



# Enrichissement basé sur les alignements

Regrouper ensemble les mappings associés  
à un même concept de la cible

- 1) Bâtiment industriel    *isEq*    bâtiment industriel    Sim: 1.0
- 2) Bâtiment d élevage industriel    *isA*    bâtiment industriel Sim: 0.817
- 3) Batiment de type industriel    *isA*    bâtiment industriel    Sim: 0.882
- 4) Bâtiment industriel (grand )    *isA*    bâtiment industriel    Sim: 0.886
- 5) Coopérative (vinicole céréalière ?)    *isA (AfterEq)*    bâtiment industriel
- 6) Construction technique    *isA (AfterEq)*    bâtiment industriel
- 7) Divers industriel    *isA (struct)*    bâtiment industriel
- 8) Elevage industriel    *isA (struct)*    bâtiment industriel
  
- 10) Industriel ou commercial    *isA (struct)*    bâtiment industriel
- 11) Industriel agricole ou commerciale    *isA (struct)*    bâtiment industriel



# Aide à l'enrichissement

## Position du concept cible dans $O_{\text{Cible}}$

Surfaces d'activités et bâti

-Bâtiment

-Nature

-Bâtiment industriel

-Construction technique

## Position de concepts sources dans $O_{\text{Source}}$

Surfaces d'activités et bâti

-Point d'activité ou d'intérêt

-Nature

-Divers industriel

-Coopérative (vinicole, céréalière)

1 entité\_à\_vocation\_industrielle  
2 centrale\_électrique  
2 infrastructure\_de\_gestion\_des\_eaux  
2 carrière  
2 zone\_d\_activité  
**2 bâtiment\_industriel**  
3 coopérative  
3 cheminée  
3 abattoir  
3 minoterie  
3 hangar\_industriel  
3 réservoir\_industriel  
4 réservoir\_de\_matériaux\_de\_construction  
4 gazomètre  
4 réservoir\_d\_hydrocarbure  
3 torchère  
3 tapis\_roulant\_industriel  
3 haut\_fourneau  
3 usine  
4 manufacture  
4 scierie  
4 atelier  
4 fabrique  
3 entrepôt



# Pbs des concepts de même label

ci-dessous les 6 concepts de label « Péage »  
alignement,

< Péage *isClose* aire de Péage >

#Surfaces d'activités et bâti - Point d'activité ou d'intérêt - Nature-Péage  
(Péage d'autoroute)

#Surfaces d'activités et bâti - Point d'activité ou d'intérêt – Catégorie – Transport - Péage

#Surfaces d'activités et bâti - Surface d'activité-Catégorie – Transport - Péage

#Surfaces d'activités et bâti-Bâtiment – Catégorie – Transport - Péage

#Surfaces d'activités et bâti – Bâtiment – Nature - Péage

#Voies de communication routière-Surface de route – Nature - Péage  
(aire de Péage)

??? à comparer au chemin menant à « aire de Péage » dans la  
cible ???

0- entité\_topographique\_artificielle

1- infrastructure\_de\_transport

2- infrastructure\_de\_transport\_routier

3- aire\_de\_péage

## Nouvelle technique : prise en compte des descendants

Sémantique Cotopy d'un concept  $c$  dans une ontologie  $O$

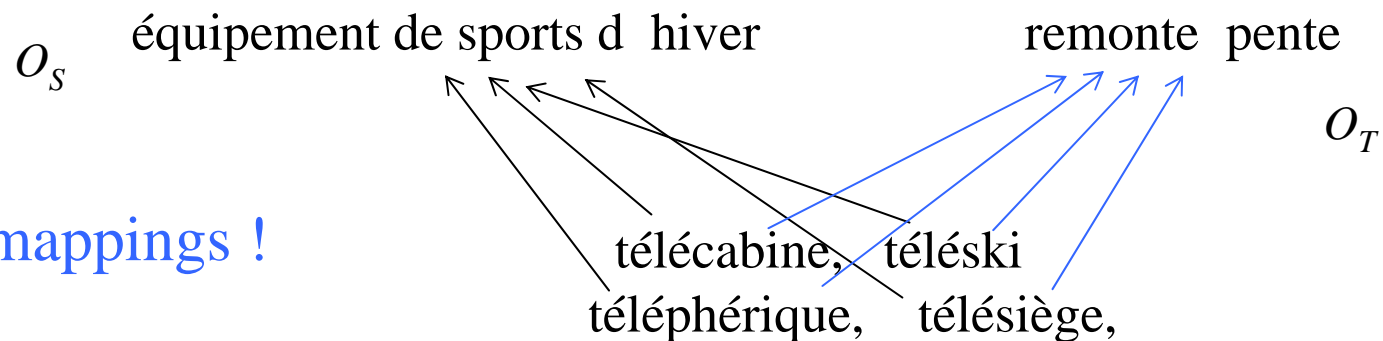
→  $SC(c, O) =$  l'ensemble de ses descendants et généralisants

### Taxonomic Overlap

$$TO(C_i, H_1, H_2) = \frac{|Sc(C_i, H_1) \cap Sc(C_i, H_2)|}{|Sc(C_i, H_1) \cup Sc(C_i, H_2)|} \quad [Maedche \& Staab 02]$$

Quand  $C_i$  n'appartient pas à  $H_2$  ?

$$TO'(C_i, H_1, H_2) = \text{Max}_{C \in H_2} \left\{ \frac{|Sc(C_i, H_1) \cap Sc(C, H_2)|}{|Sc(C_i, H_1) \cup Sc(C, H_2)|} \right\}$$



14 nouveaux mappings !