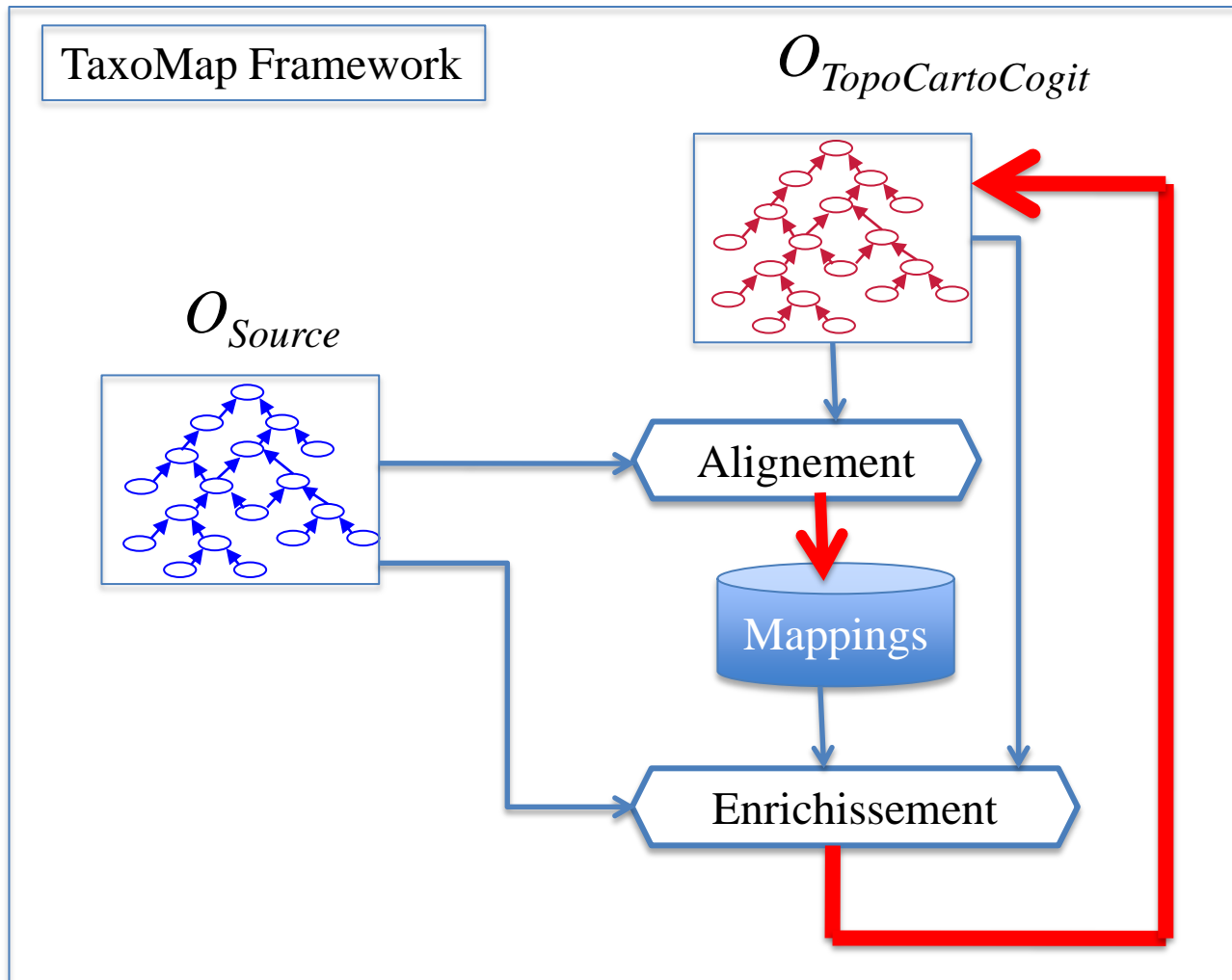


Géonto Lot 2.1 - Avancement semestre 5

- Affinement de mappings
 - Implémentation/validation des patterns (pour Carto → Topo)
6 patterns portant sur 4 techniques, 25 modifications, 23 OK
 - 3 publications RFIA, M&M, EKAW
- Intégration d'outils divers à l'environnement
- Implémentation des 1^{er} patterns d'enrichissement

Enrichissement d'Ontologie



Enrichissement

- L'enrichissement de O_{TCC} peut se faire avec les ontologies construites par l'IRIT à partir des spécifications
- ... ou à partir d'autres ontologies trouvées sur le web
- Problèmes à traiter
 - Evaluer l'intérêt des ontologies trouvées sur le web en les comparant à O_{TCC}
 - Extraire les parties intéressantes
 - Les utiliser pour l'enrichissement

Comparaison d'Ontologies

[Maedche & Staab 02]

→ Sémantique Cotopy d'un concept c dans une ontologie O
 $SC(c, O) =$ l'ensemble de ses descendants et généralisants

Taxonomic Overlap pour un même concept dans 2 ontologies

$$TO(C_i, H_1, H_2) = \frac{|Sc(C_i, H_1) \cap Sc(C_i, H_2)|}{|Sc(C_i, H_1) \cup Sc(C_i, H_2)|}$$

Quand C_i n'appartient pas à H_2 ?

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

$$\text{Sim}(O_1, O_2) = \frac{1}{|O_1|} \sum_{C_i \in O_1} TO(C_i, O_1, O_2)$$

Intérêts de cette mesure (1)

- Mesure structurelle

0 franchissement

1 **écluse**

0 entité_topographique_artificielle

1 infrastructure_de_transport

2 infrastructure_de_transport_maritime_ou_fluvial

3 **écluse**

4 bajoyer

4 cale_sèche

4 ascenseur_à_bateaux

4 radoub

$TO(\text{écluse}, \text{carto}, \text{topo}) = 1/9$

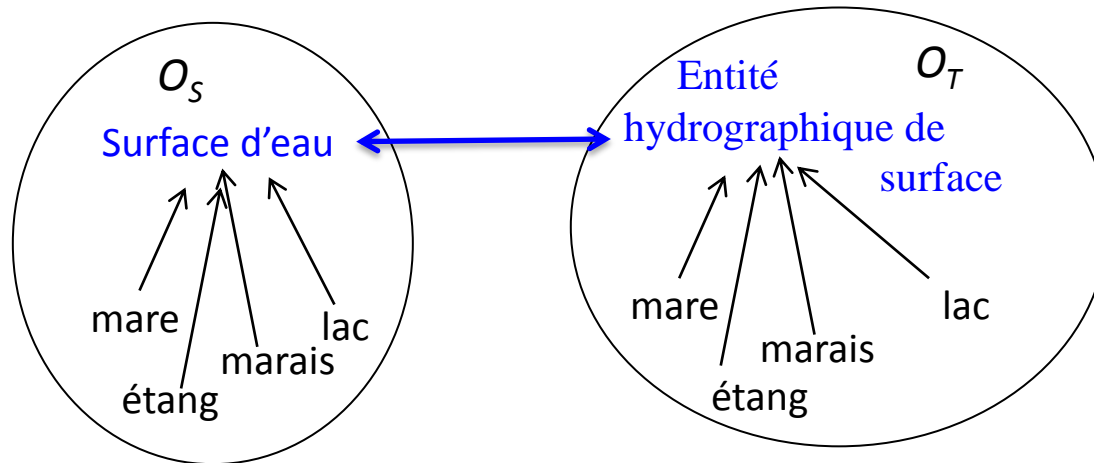
Faible similarité des 2 ontologies, car structure très dissemblable

Carto : très dispersée 65 Concepts à niv 0, dont 50 ont moins de 10 fils
peu profonde 429 à niv 1, 52 à niv 2 et 11 à niv 3

Intéressante pour comparer Topo à d'autres ontologies bien structurées

Intérêts de cette mesure (2)

- S'appuie sur les labels de concepts identiques (souvent les feuilles)
- mais permet de rapprocher (avec TO') les concepts intermédiaires de labels différents qui recouvrent les mêmes notions



point_remarquable_du_relief ↔ élément_du_relief

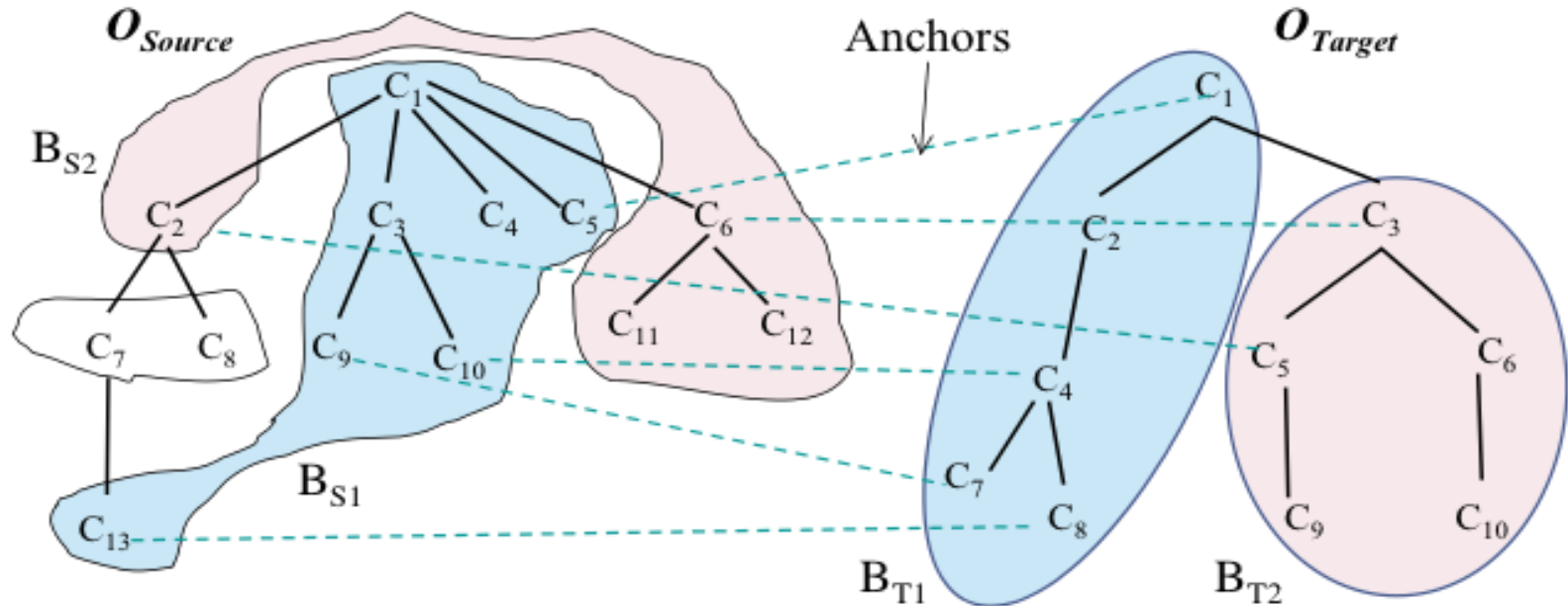
noeud_hydrographique ↔ extrémité_du_réseau_hydrographique_terrestre

occupation_du_sol ↔ zone_de_culture

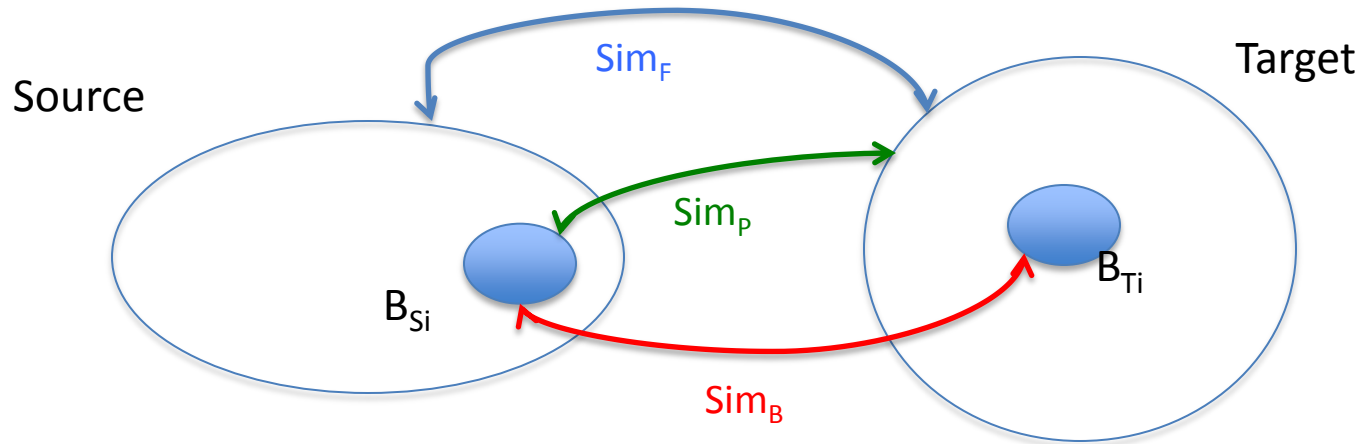
entité_hydrographique_de_surface ↔ Surface d'eau

Intérêts de cette mesure (3)

- S'associe bien à la technique de partitionnement qui s'appuie aussi sur les « ancres » et sur une mesure de similarité structurelle [Wu & Palmer]



Comparaison d'ontologies et partitionnement



3 calculs de similarité (mesure Maedche-Staab)

Sim_{Full} entre O complètes

$Sim_{Partial}$ entre un bloc source et la cible complète

Sim_{Bloc} entre un bloc source et le bloc cible correspondant

Globalement pour (presque) tous les blocs issus de la source, par partitionnement

$$Sim_{Bloc} > Sim_{Partial} \geq Sim_{Full}$$

Méthodologie pour l'enrichissement

1. Partitionner manuellement notre ontologie (O_{TCC}) en blocs thématiques précis
2. Partitionner une ontologie extérieure en fonction de nos blocs
3. Comparer les paires de blocs trouvées
4. N'aligner blocs à blocs que ceux avec une mesure de comparaison élevée
5. Utiliser les mappings trouvés pour l'enrichissement

Des (petits) outils pour l'enrichissement intégrés au Framework

- 1) Mesure de comparaison (Maedche-Staab)
- 2) Visualisation d'ontologies
- 3) Analyseur d'ontologie calculant sur les 3 premiers niveaux
 - nb de fils d'un niveau
 - Taille du bloc correspondant à chaque fils
- 4) Partitionnement à partir de blocs cibles manuels

TaxoMap Framework : Module Enrichissement

Fayçal HAMDI

Module Enrichissement

1. Des patterns prédéfinis qui peuvent être appliqués sur des ontologies géographiques (ou sur d'autres ontologies).
2. Des outils (visualisation, partitionnement, ...etc.) pour aider l'expert/ingénieur à créer de nouveaux patterns.

Importer ou pas un nouveau concept ?

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

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

\langle Arène *ou* théâtre antique *isA* théâtre antique \rangle

\langle Plaine *ou* plateau *isA* plateau \rangle

 Éliminer les concepts portant sur 2 concepts existant dans la cible

Exemple contenant des « *mots vides* »

\langle Bâtiment *de type* industriel *isA* bâtiment industriel \rangle

 Éliminer les concepts

Exemple contenant des adjectifs parenthésés

\langle canal (*large*) *isA* canal \rangle

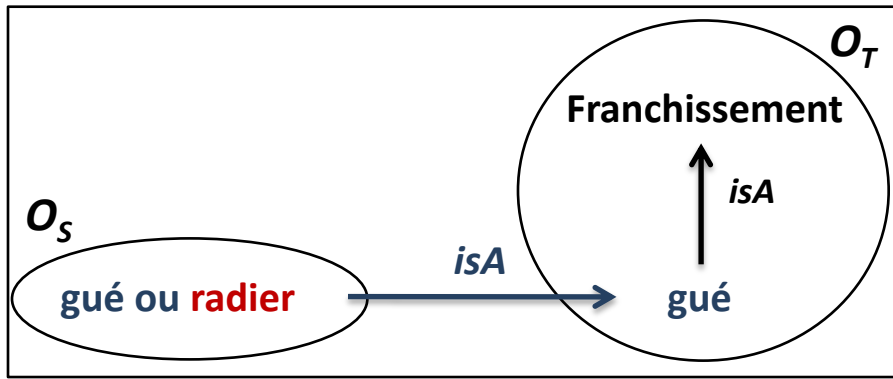
\langle atelier (*grand*) *isA* atelier \rangle



Ne pas introduire le concept
mais garder une propriété

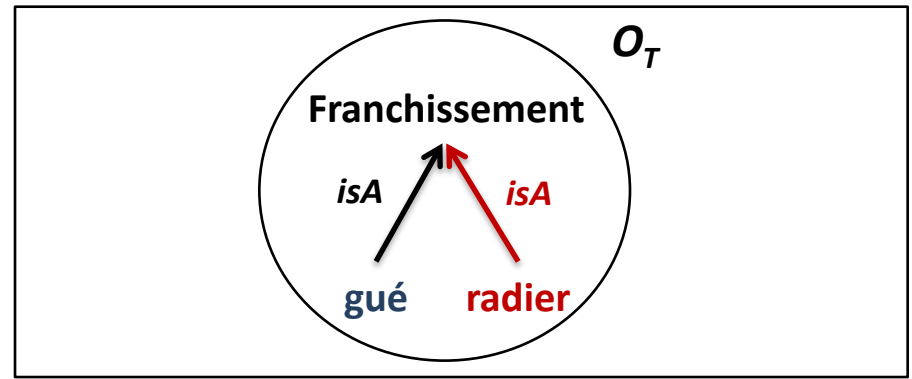
Patterns d'enrichissement (exemple)

Onto_IRIT



Mapping proposé par *TaxoMap*

BDTopo



Enrichissement suggéré par l'expert

Pattern :

Partie contexte :

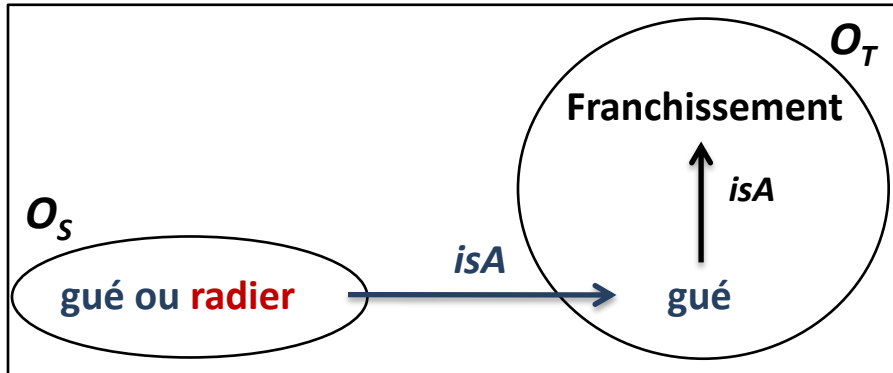
Pour un mapping identifié par la technique T_2 . Si :

- Le concept de O_S (**gué ou radier**) contient :
 - Une conjonction (et, ou, /), un concept de O_T (**gué**) et un concept de O_S (**radier**) qui n'est pas aligné.

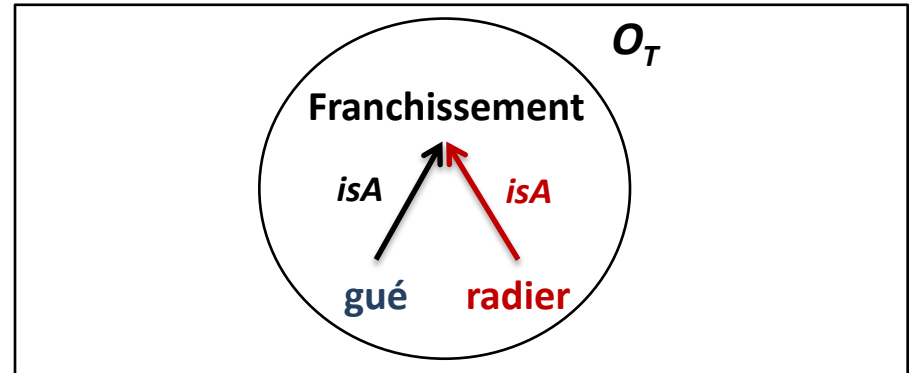
Partie solution:

Introduire une relation de subsomption *isA* entre **radier** et le père de **gué** (**franchissement**).

Patterns d'enrichissement (exemple)



Mapping proposé par *TaxoMap*



Enrichissement suggéré par l'expert

Pattern :

Partie contexte :

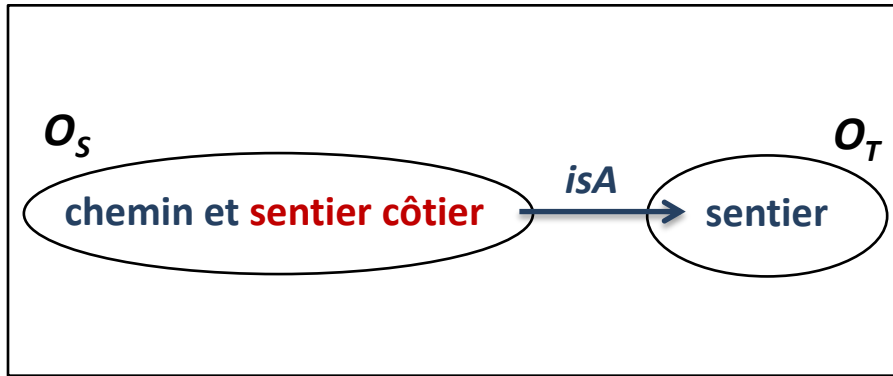
$$\exists x \exists y \exists z (isStrictInclusion(x, y) \wedge isParentof(y, z, O_T) \wedge extractFromLabel(x, _, y, r))$$
$$\wedge \exists c (labelOfConcept(r, O_S, c) \wedge notMapped(c))$$

Partie solution:

$$Add_Relation(c, z, \ll isA \gg)$$

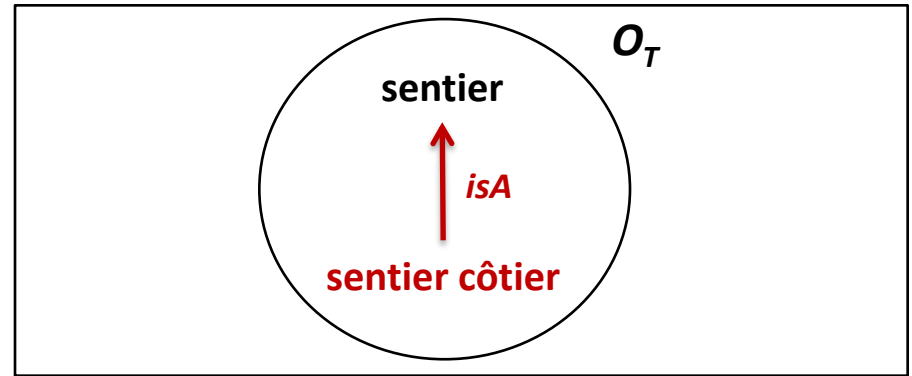
Patterns d'enrichissement (exemple)

Onto_IRIT



Mapping proposé par *TaxoMap*

BDTopo



Enrichissement suggéré par l'expert

Pattern :

Partie contexte :

Pour un mapping identifié par la technique T_2 . Si :

- Le concept de O_S (**chemin et sentier côtier**) contient :
 - Une conjonction (et, ou, /), Deux concepts de O_T (**chemin**) et (**sentier**) et une partie du concept (**sentier côtier**) $\notin \{O_S, O_T\}$

Partie solution:

Proposer **sentier côtier** comme spécialisant du concept O_T contenu (**sentier**).

Mappings restants ?

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)
- ~~11)~~ Industriel agricole ou commerciale *isA (struct)* bâtiment industriel


Module Enrichissement


TaxoMap Framework


Alignment Refinement Enriching


Partitioning Visualizing Other tools Rules


Partitioning


Source Ontology 

Target Ontology 

Target Folder 

Reference Alignments 

Output Folder 

Language en 

Max size of Generated Blocks

Target Blocks Source Blocks

Partitioning Methods

PAP

APP

Start Partitioning

Module Enrichissement

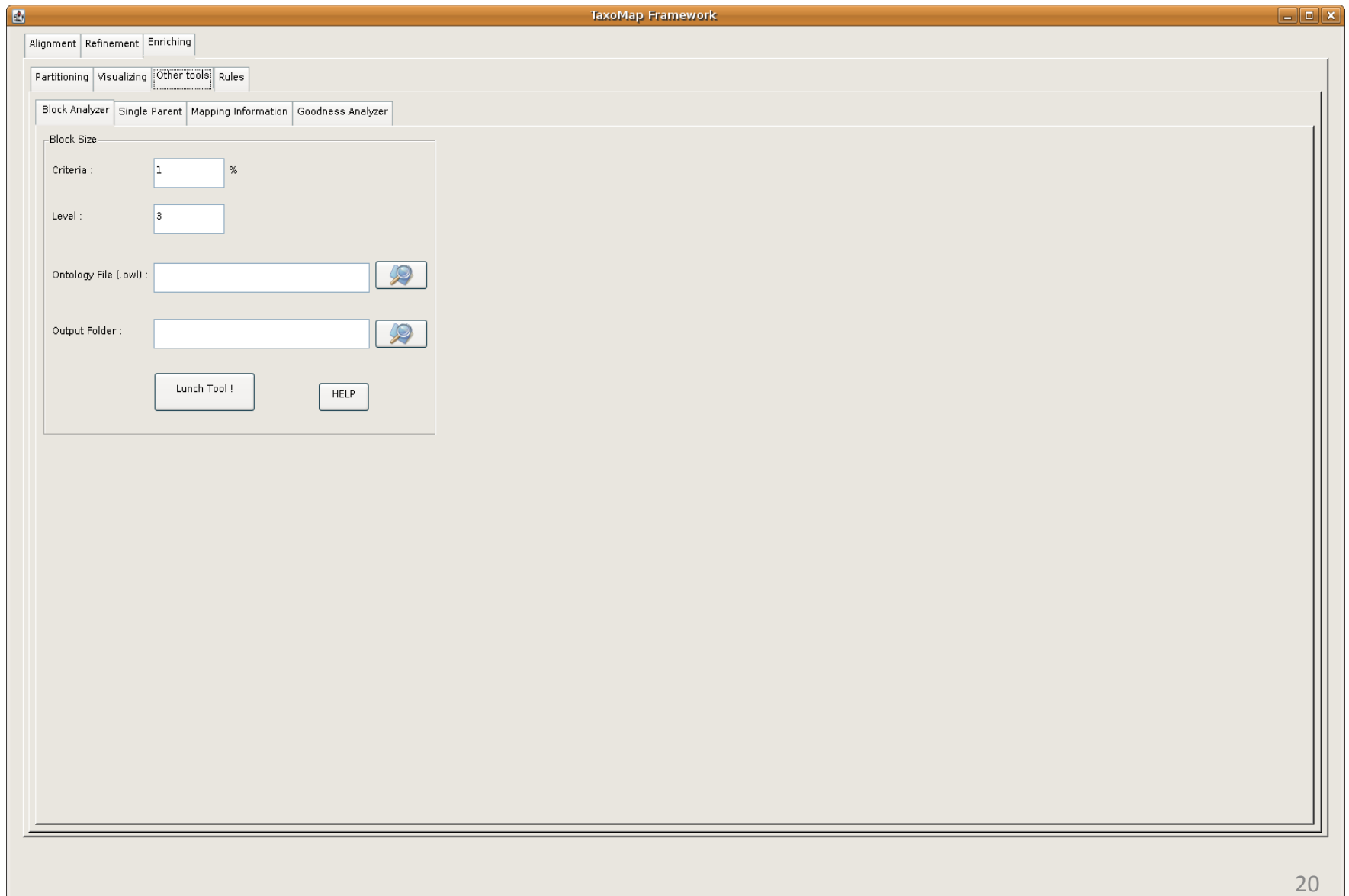
The screenshot displays the TaxoMap Framework interface, which is used for ontology alignment. The window title is "TaxoMap Framework". At the top, there are tabs for "Alignment", "Refinement", and "Enriching". Below these are sub-tabs for "Partitioning", "Visualizing", "Other tools", and "Rules".

The main interface is divided into several sections:

- Charger Ontologie 1...:** Contains the file path "ers/iasi/hamdi/workspace/testnewtaxo/BDCarto.owl".
- Charger Ontologie 2...:** Contains the file path "ers/iasi/hamdi/workspace/testnewtaxo/BDTopo.owl".
- Afficher les mappings:** A central area showing two side-by-side class hierarchies with arrows indicating mappings between them.
- EXPLORER LES CLASSES (Left Panel):** Shows a class hierarchy for "owl:Thing" with classes like "agglomération", "arcade_ou_mur", "axe_du_lit", "aérodrome", and "aéroport".
- EXPLORER LES CLASSES (Right Panel):** Shows a class hierarchy for "owl:Thing" with classes like "entité_topographique_artificielle", "entité_topographique_naturelle", "owl:AllDifferent", "owl:DataRange", and "owl:Nothing".

The central mapping area shows two class hierarchies. The left hierarchy has a root node "zone_d_habitat" with children: "équipement_de_sports_d_hiver", "pays_d_art", "zone_bâtie", "chef-lieu_de_commune", "ville", "station_de_sports_d_hiver", "station_balnéaire", "quartier", "station_thermale", "village", "station_verte", and "village_pittoresque_au_niveau_de_l_architecture_o". The right hierarchy has a root node "élément_du_relief" with children: "escarpement", "plateau", "glacier", "versant", "vallée", "montagne", and "cirque". Arrows indicate mappings between these classes, such as "zone_d_habitat" mapping to "ville" and "station_de_sports_d_hiver", and "élément_du_relief" mapping to "montagne" and "cirque".

Module Enrichissement



Module Enrichissement

The screenshot displays the 'TaxoMap Framework' application window, specifically the 'Rules' tab. The interface is organized into several sections:

- Alignment | Refinement | Enriching**: Navigation tabs at the top.
- Partitioning | Visualizing | Other tools | Rules**: Sub-navigation tabs, with 'Rules' selected.
- Create New Rules**: The main area for defining new rules, divided into:
 - Predicates**:
 - Inter-Ontologies Relations**: isA(X, Y), isClose(X, Y), isEquivalent(X, Y).
 - Intra-Ontologies Relations**: isSubClassOf(X, Y), isParentOf(X, Y).
 - Properties**: strictInclusionLabel(X, Y), inclusionInLabel(X, c, Z).
 - Others**: inSameBlock(X, Y), isConcept(X).
 - Operations**:
 - Ontology Operations**: Add_Relation(X, Y, R).
 - Variables**:
 - Domain**: 1.
 - Variables**: (empty).
 - Buttons: Add variables, Select.
- Rule**: A text input field with **Validate** and **Reset** buttons.
- Restructuring Rules**: A large text area with **Execute** and **Remove** buttons.

Predefined Rules: A list of existing rules, including:
isAStrictInclusion(X,Y) ^ extractFromLabel(X,et,Y,R) ^ labelOfConcept(R,C) ^
isAStrictInclusion(X,Y) ^ extractFromLabel(X,ou,Y,R) ^ labelOfConcept(R,C)

At the bottom right, there is an **Execute** button for the predefined rules section.

Merci