

Atelier DECOR
Namur, 22 janvier 2007

Appariement de schémas de BD géographiques à l'aide d'ontologies déduites des spécifications

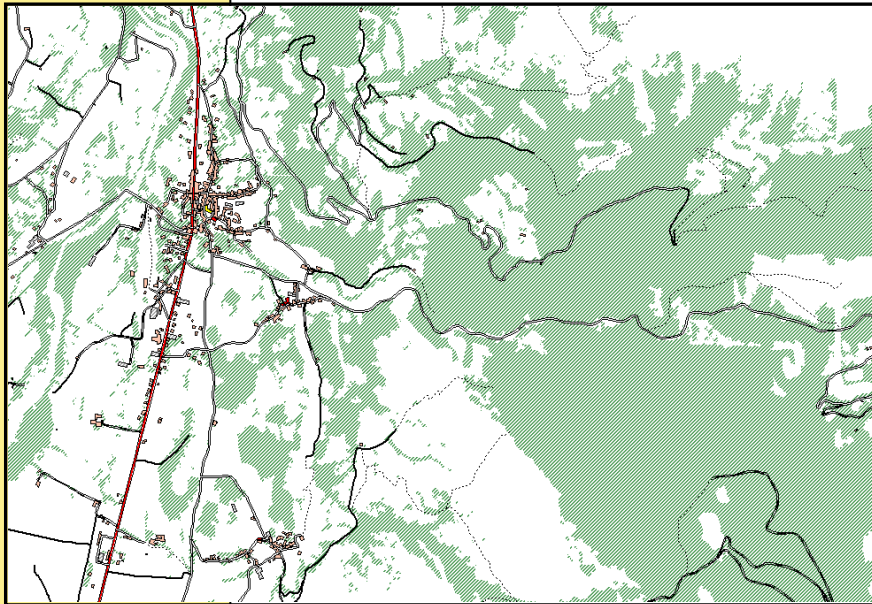
Sébastien Mustière
Nathalie Abadie
Frédéric Laurens

Laboratoire COGIT
Institut Géographique National
<http://recherche.ign.fr/labos/cogit/>

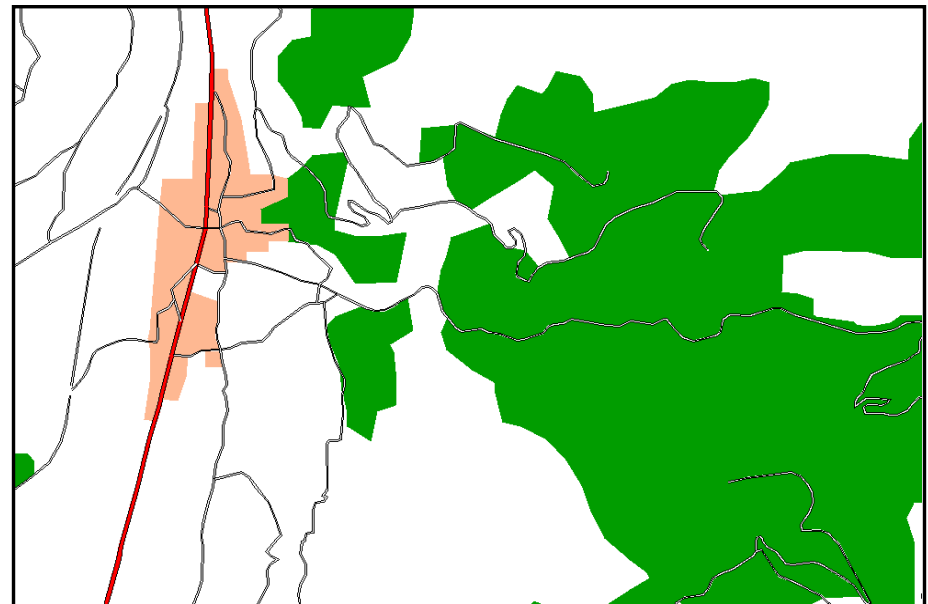


Objectif

- Intégrer BD géographiques
 - Appariement de schémas
 - Appariement de données



BD TOPO

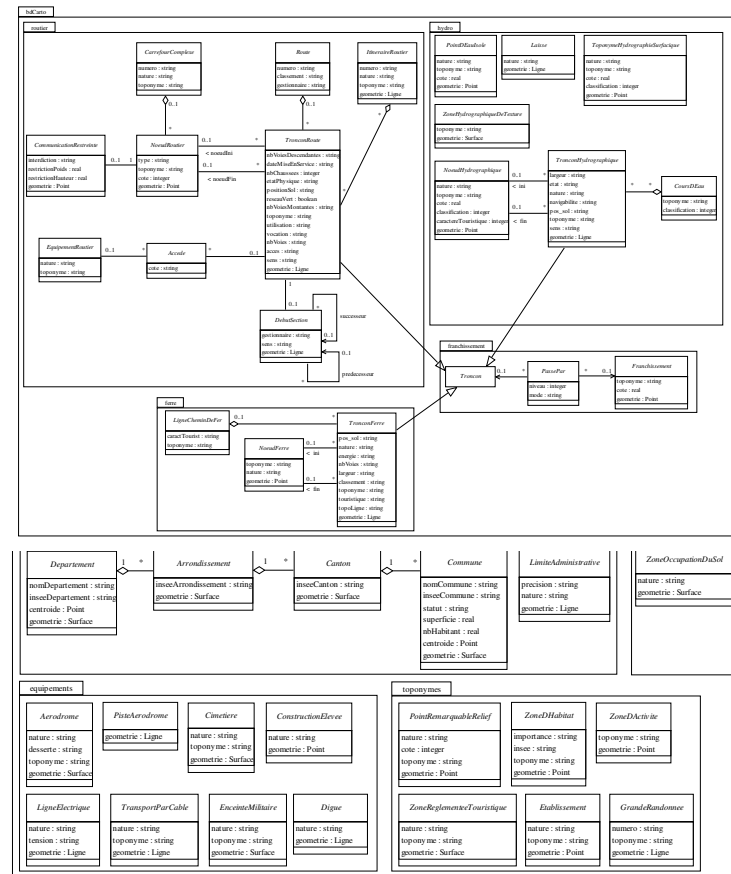
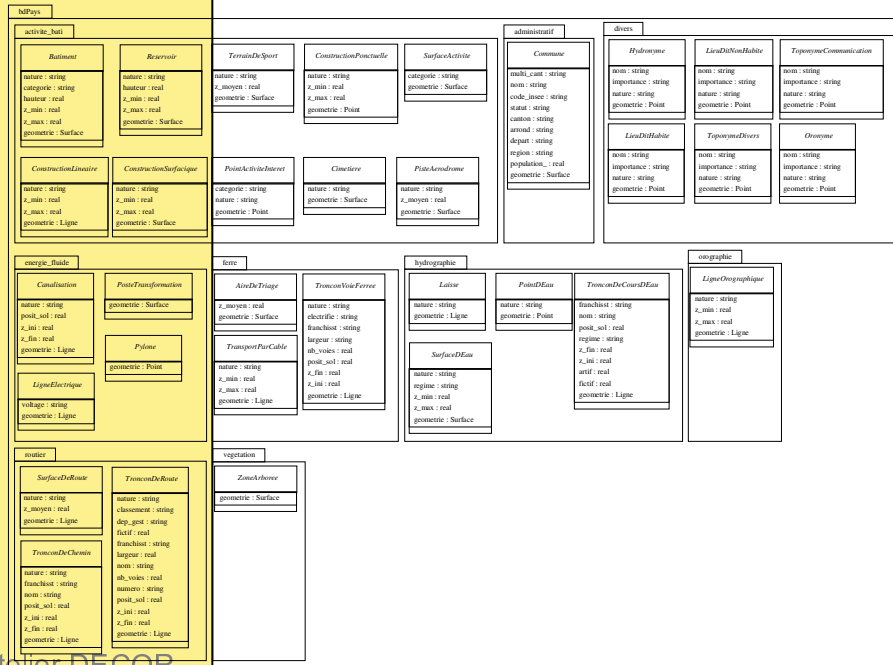


BD CARTO

Difficultés (1)

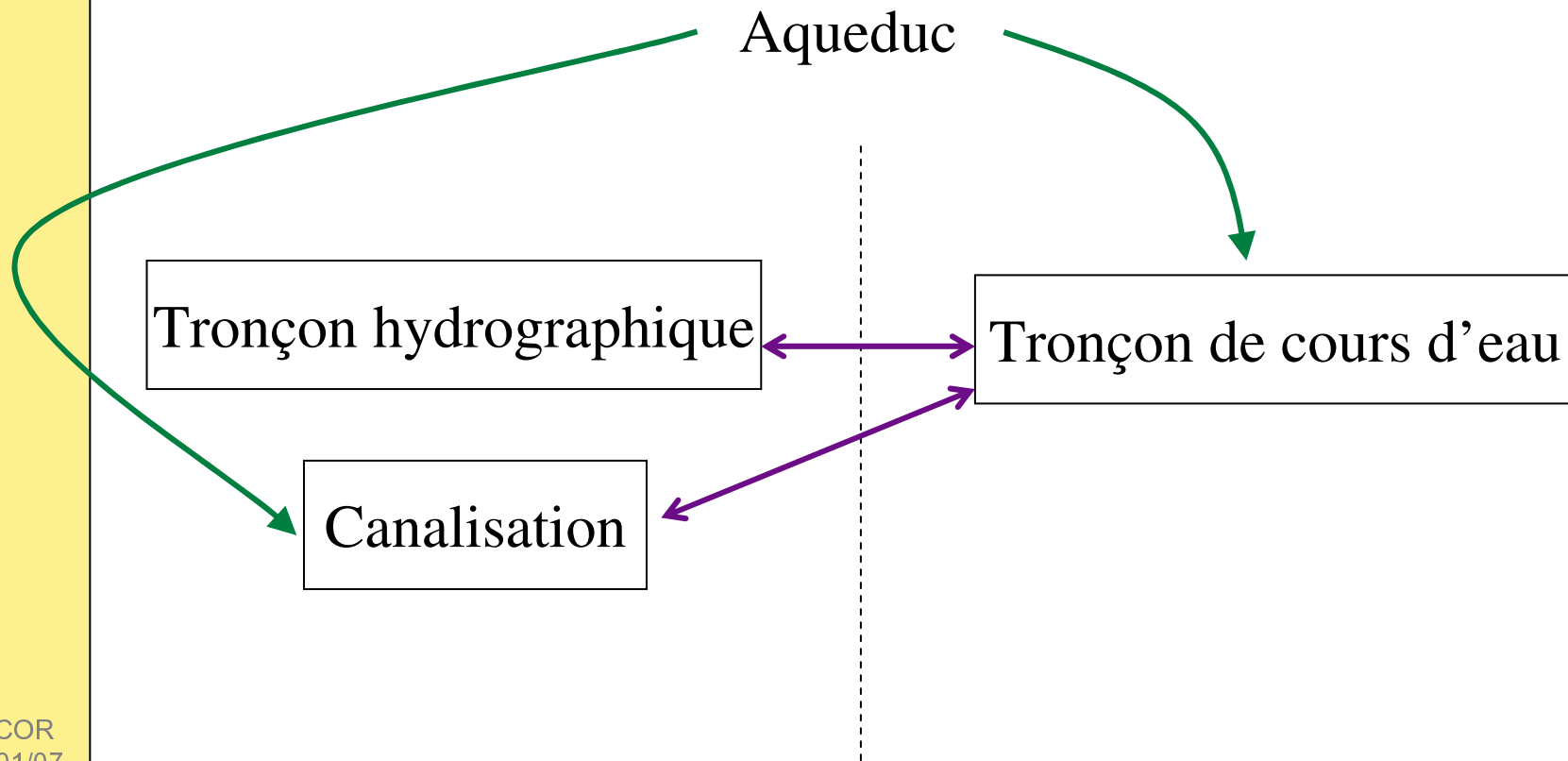
➤ Des schémas...

- complexes
- avec de nombreuses classes et valeurs d'attributs
- variés



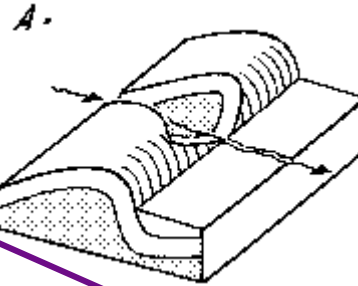
Difficultés (2)

- Des schémas à la sémantique insuffisante



Difficultés (2)

➤ Des schémas à la sémantique insuffisante



<i>Point Remarquable Relief</i>
nature : string
cote : integer
toponyme : string
geometrie : Point

BDCARTO

Cap, pointe
 Cirque
 Col, passage
 Volcan, cratère
 Sommet, crête, colline
 Coteau, Falaise
 Cuvette, dépression
 Vallée
 Ile
 Dune, plage
 Pic
 Plaine, plateau
 Récifs
 Rocher
 Espace marin
 Banc, haut fond

Cap
 Carrière
 Cirque
 Col
 Crête
 Dépression
 Dune
 Escarpement
 Gorges
 Grotte
 Ile
 Isthme
 Montagne
 Pic
 Plage
 Plaine ou plateau
 Récifs
 Rochers
 Sommet
 Vallée
 Versant
 Volcan

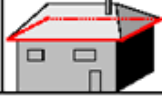

<i>Oronyme</i>
nom : string
importance : string
nature : string
geometrie : Point

BDTOPO

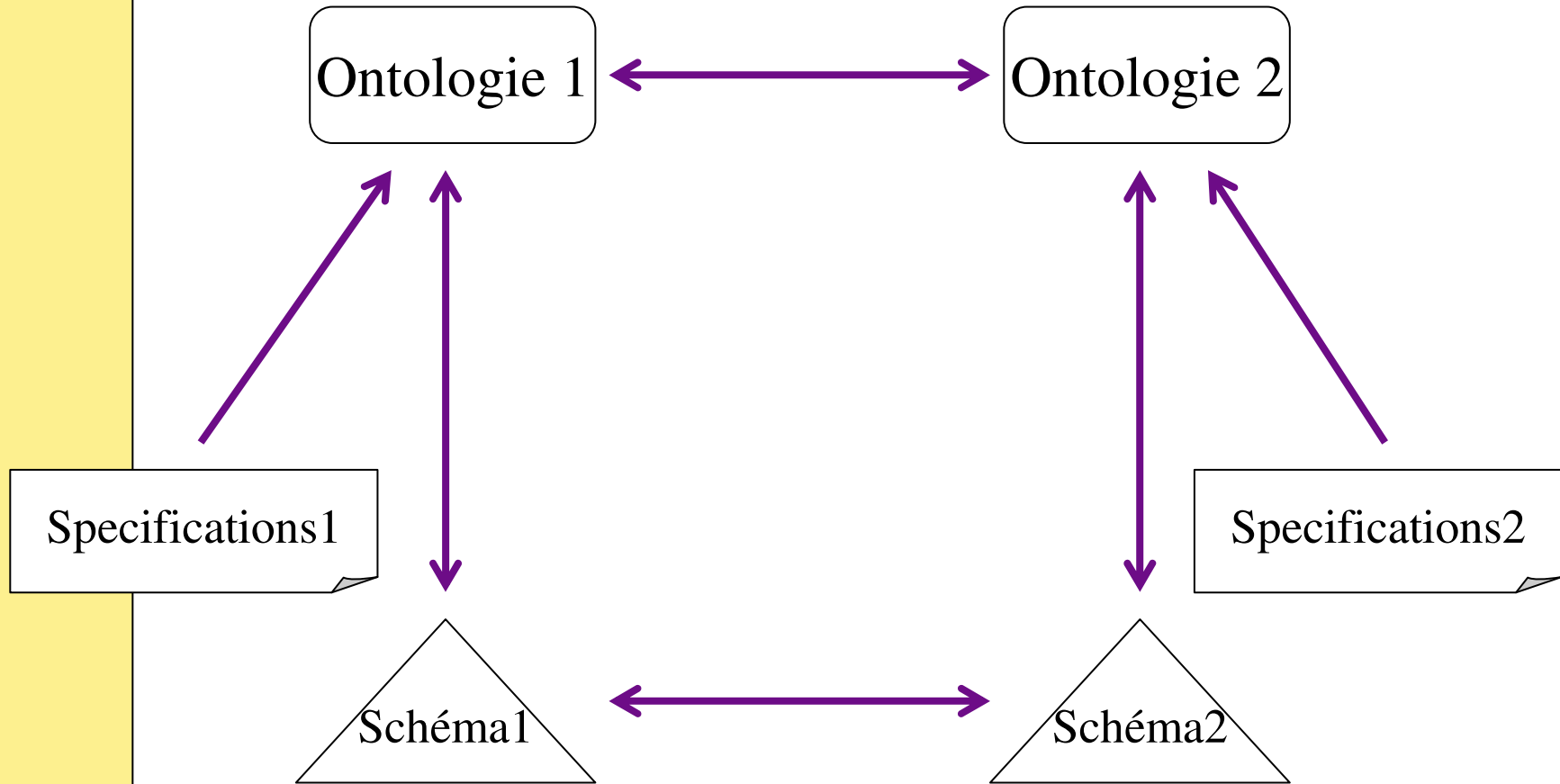
- Des schémas à la sémantique insuffisante
 - Regroupements opérés dans les classes
 - Critères de sélection
 - Critères de découpage
 - ...

➤ S'appuyer sur les spécifications textuelles

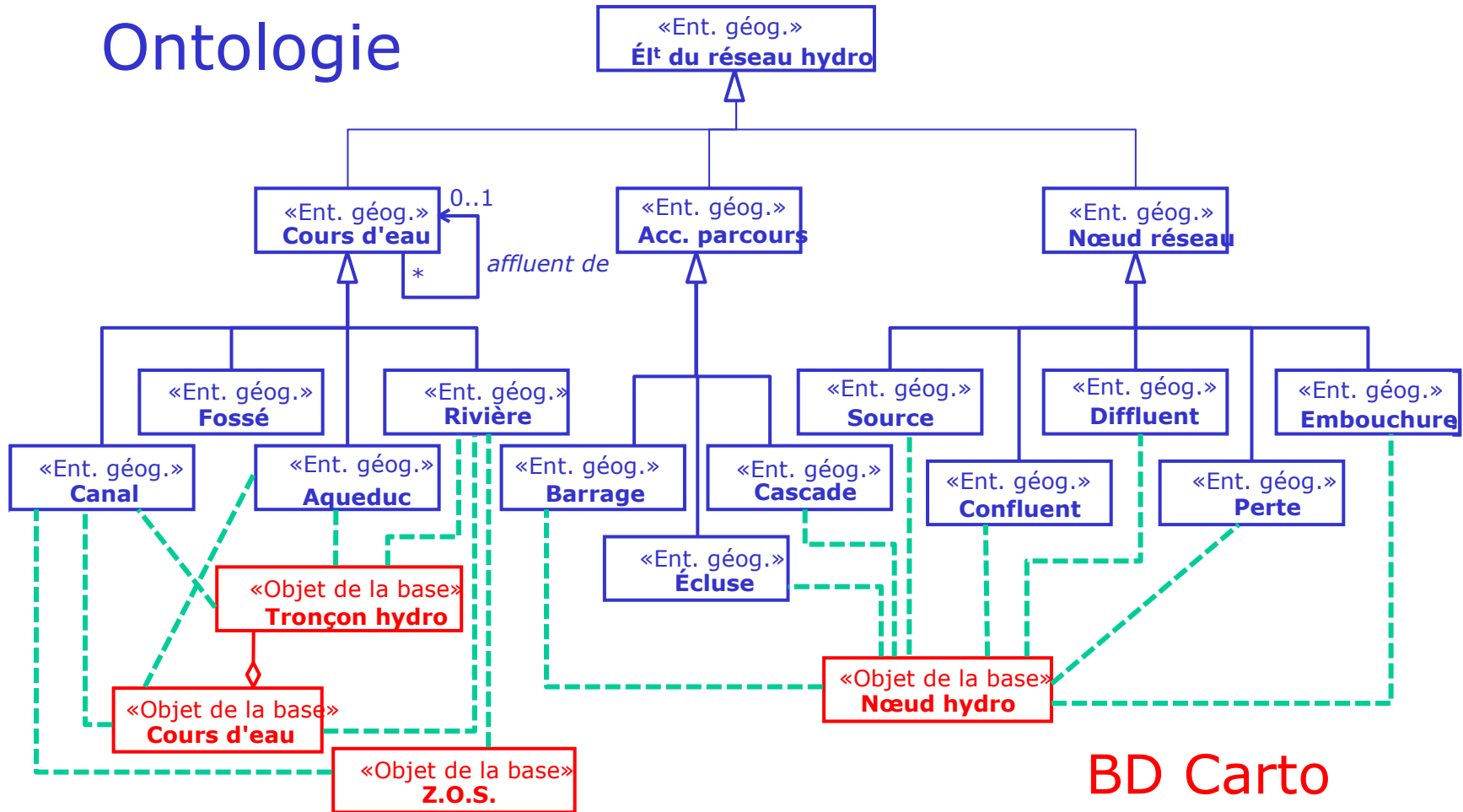
- Très riches en information
- Très organisées / texte libre
- Très volumineuses
- Avec beaucoup de langage naturel
- Avec des organisations sensiblement différentes

BDTopo Pays/Agglo SPÉCIFICATIONS DE CONTENU Domaine E		Institut Géographique National Service des Bases de Données Vecteurs	Page : 76/144 Version : 1.2 Date : 10 juillet 2002
E1 Bâtiment			
Type :	Simple	Attributs (* voir les spécifications générales)	
Localisation :	Surfacique tridimensionnelle	<ul style="list-style-type: none"> • Signature électronique* • Nature • Fonction • Altitude sol bâtiment • Source géométrique des données* 	
Liens :			
Définition			
Bâtiment de plus de 20 m ² .			
Regroupement : Voir les différentes valeurs des attributs <nature> et <fonction>.			
Sélection			
Tous les bâtiments de plus de 50 m ² sont inclus.			
Les bâtiments faisant entre 20 et 50 m ² sont sélectionnés en fonction de leur environnement* et de leur aspect**.			
Les bâtiments de moins de 20 m ² sont représentés par un objet de classe <construction ponctuelle> s'ils sont très hauts, ou s'ils sont spécifiquement désignés sur la carte au 1 : 25 000 en cours (ex. monument, antenne, ...).			
* Les petits bâtiments isolés (plus de 100 m d'une habitation) de plus de 20 m ² sont inclus, alors que les petits bâtiments situés en ville ne le sont pas (ex. petit garage individuel, petit atelier, annexes diverses).			
** Les petits bâtiments d'aspect précaire (cabanes de chantier, petits abris pour animaux, ...) sont exclus.			
Modélisation géométrique			
Contour extérieur du bâtiment tel qu'il apparaît vu d'avion (le plus souvent, ce contour correspond à celui du toit); altitude* correspondant à ce contour (généralement l'altitude des gouttières).			
* altitude de l'arête supérieure en cas de face verticale.			
Seules les cours intérieures de plus de 10 m de large sont représentées par un trou dans la surface bâtie.			
Description	Monde réel et modélisation	Modélisation géométrique	
Modélisation d'une maison			
Plusieurs bâtiments contigus ou superposés de même « nature » et de même « fonction » sont généralement considérés comme un seul et même objet (seul le contour extérieur est saisi). Deux objets contigus ou superposés sont cependant représentés s'ils présentent les caractéristiques suivantes :			
<ul style="list-style-type: none"> - différence de hauteur entre les deux bâtiments > 10 m environ (ou 3 étages) ; - surface de chaque objet résultant de 400 m² environ ou plus 			

Approche [Gesbert 2005]

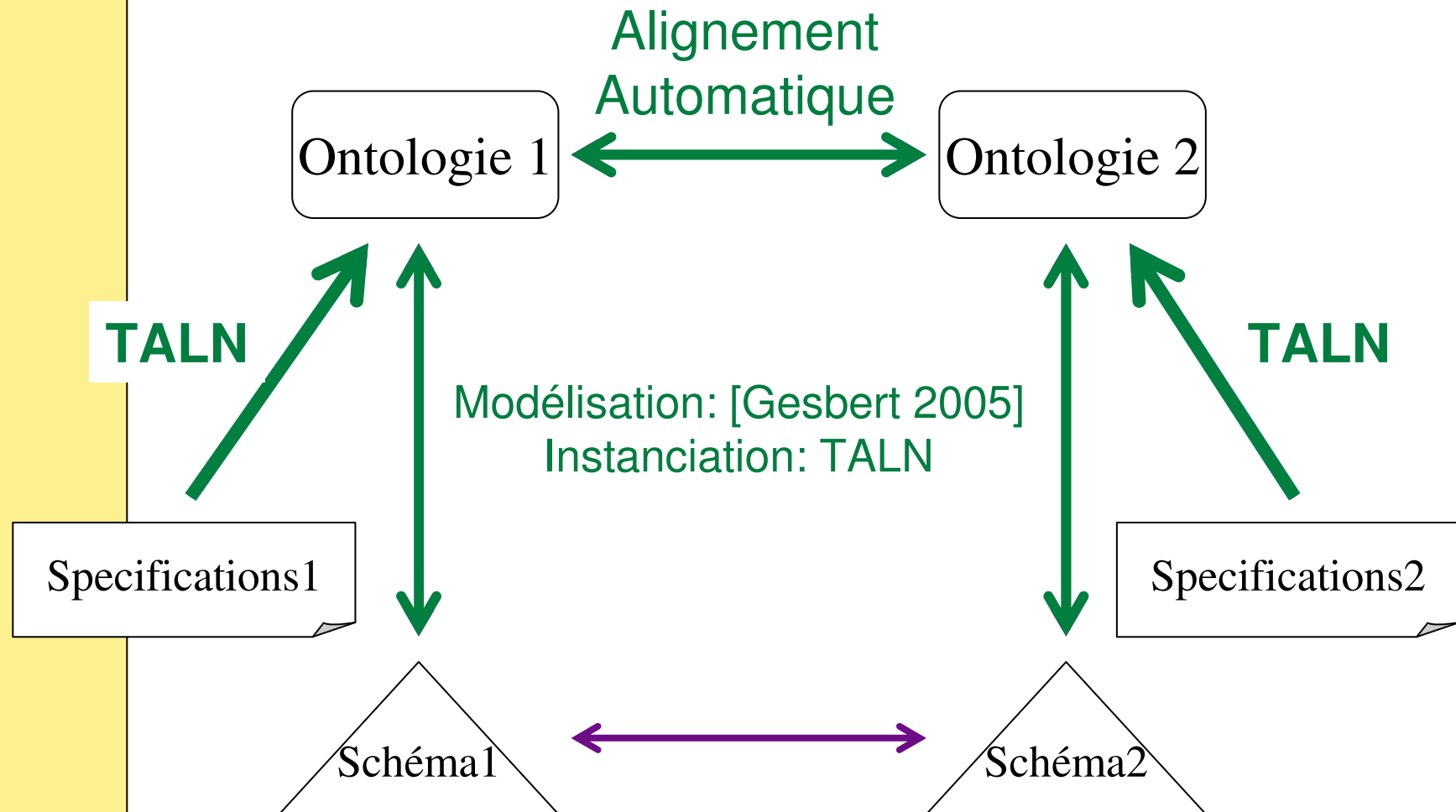


Ontologie

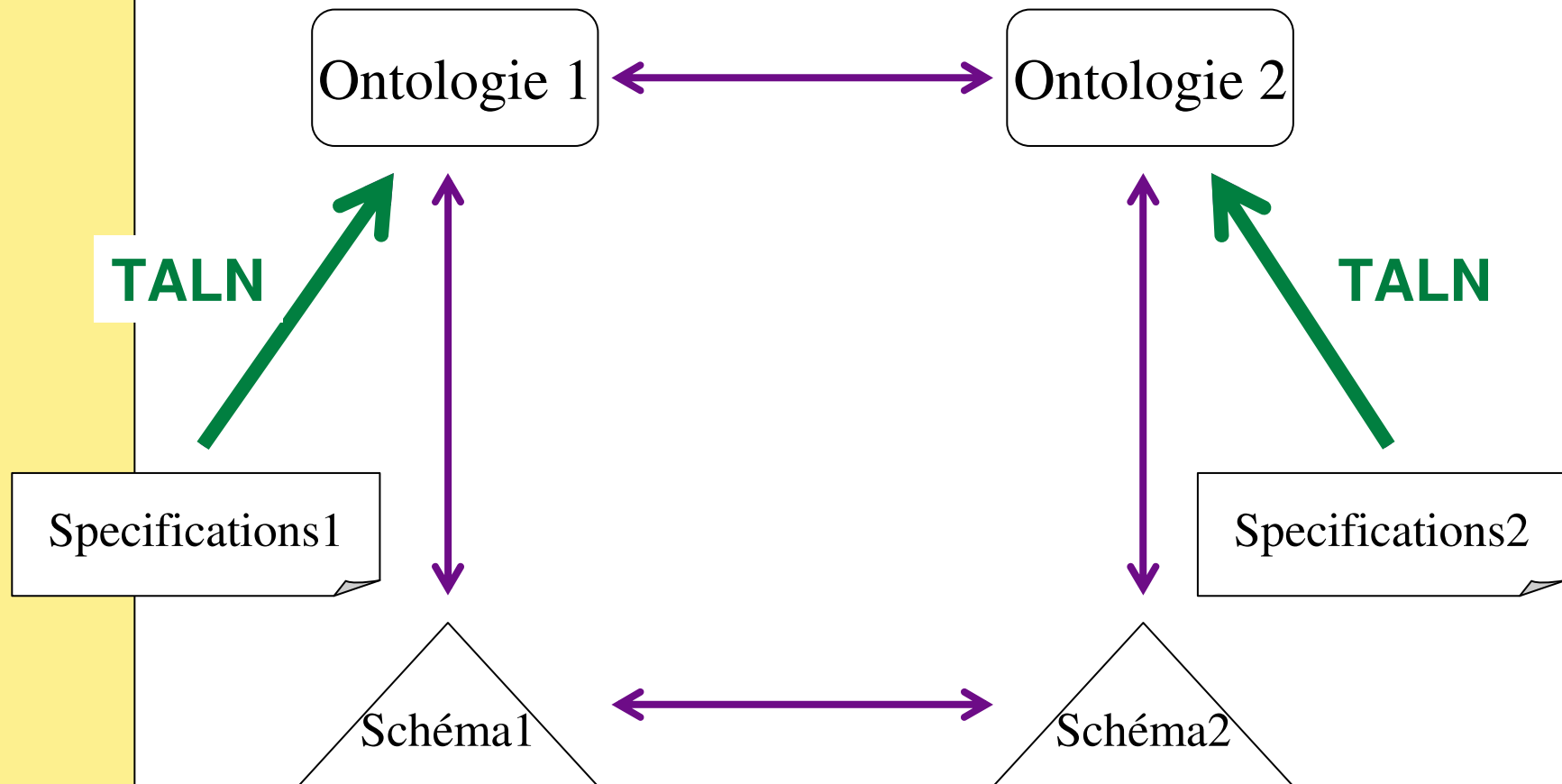


BD Carto

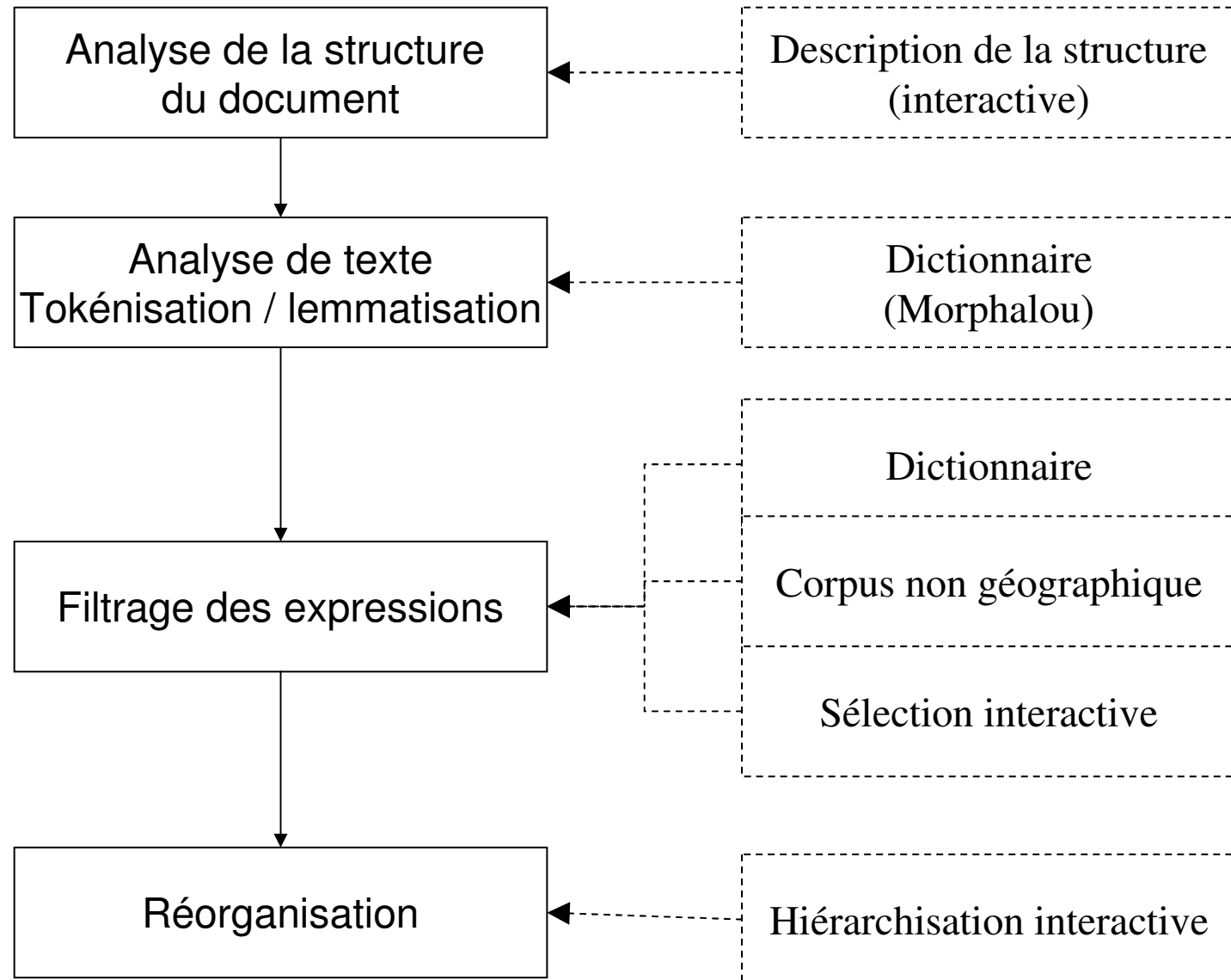
Approche: points clefs



Résultats (1): création des ontologies [Laurens 2006]



Résultats (1): création des ontologies [Laurens 2006]

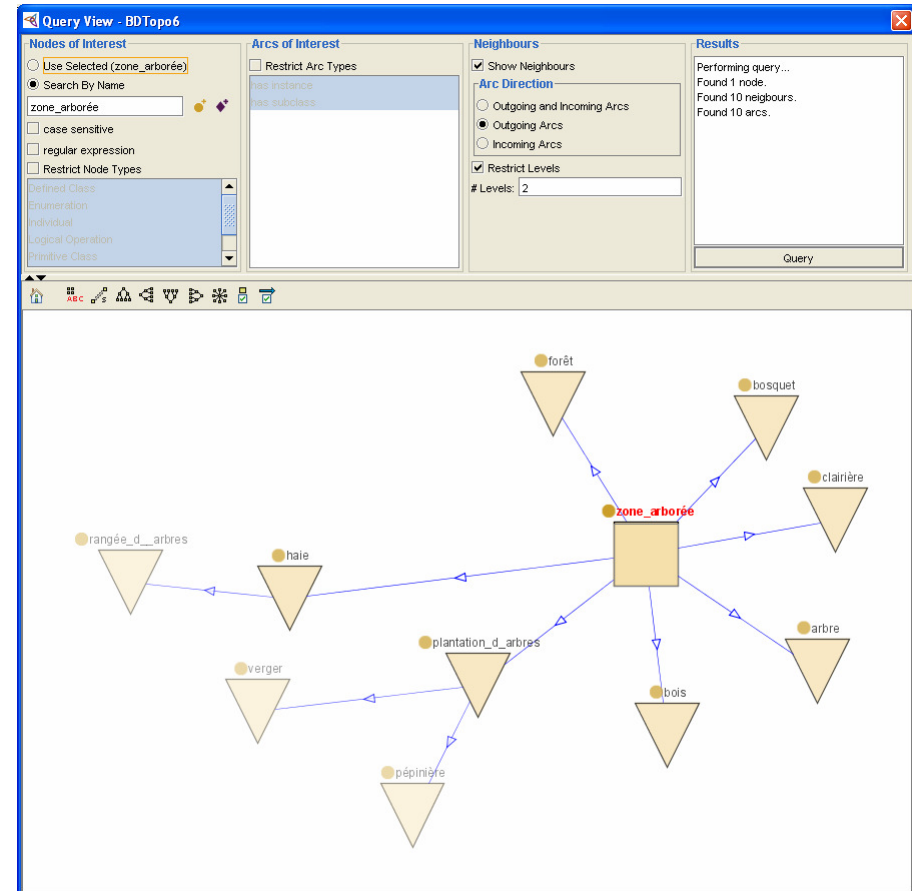
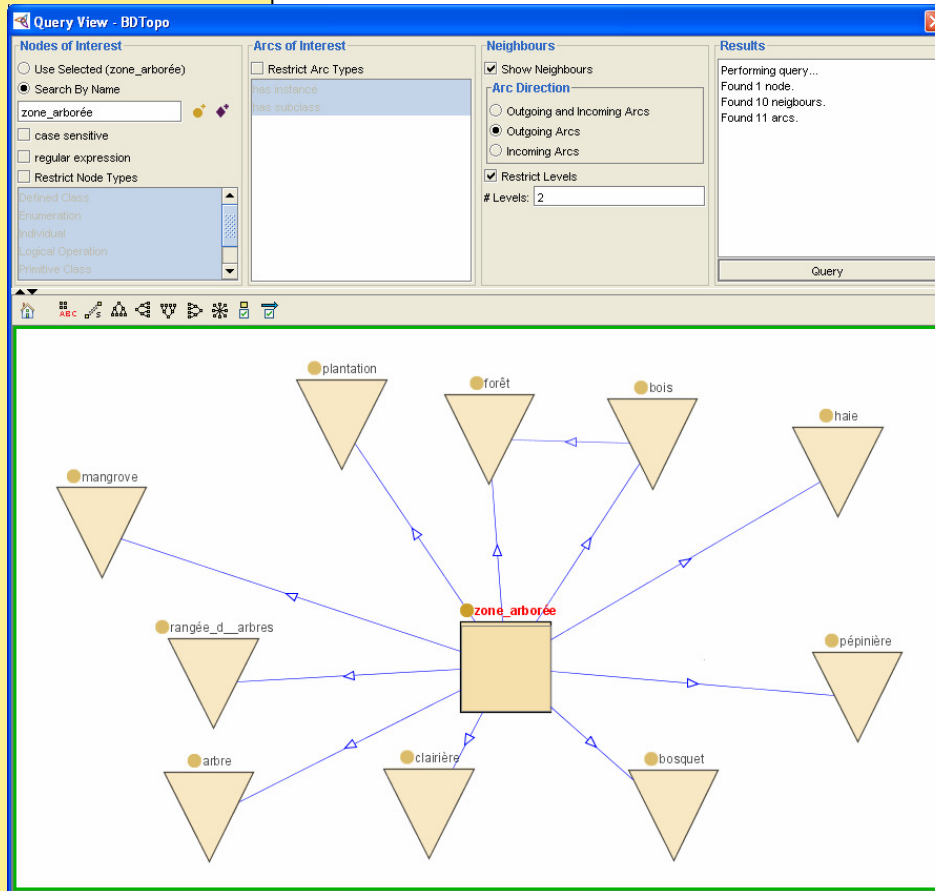


Résultats (1): création des ontologies [Laurens 2006]

- 580 concepts dans la taxonomie créée
- Comparaison à un index manuel
 - 380 identiques
 - 200 en plus dans la taxonomie créée
 - 80 en plus dans l'index
- Hiérarchisation insuffisante
 - Ex: pic, sommet, gorges, plage...
- Trop proche de la BD
 - Ex: Bâtiment ponctuel, surfacique, linéaire
- Un travail interactif est nécessaire
 - Pour filtrer (rapide)
 - Pour réorganiser (plus long)

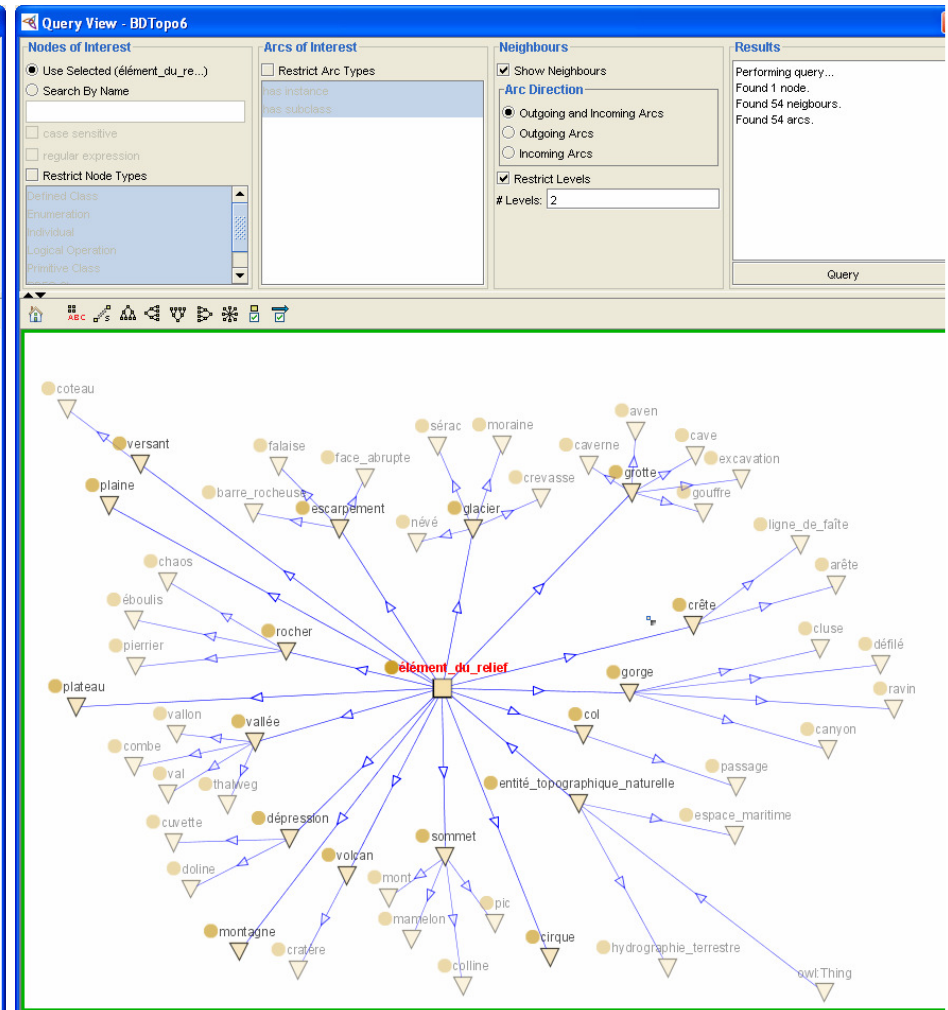
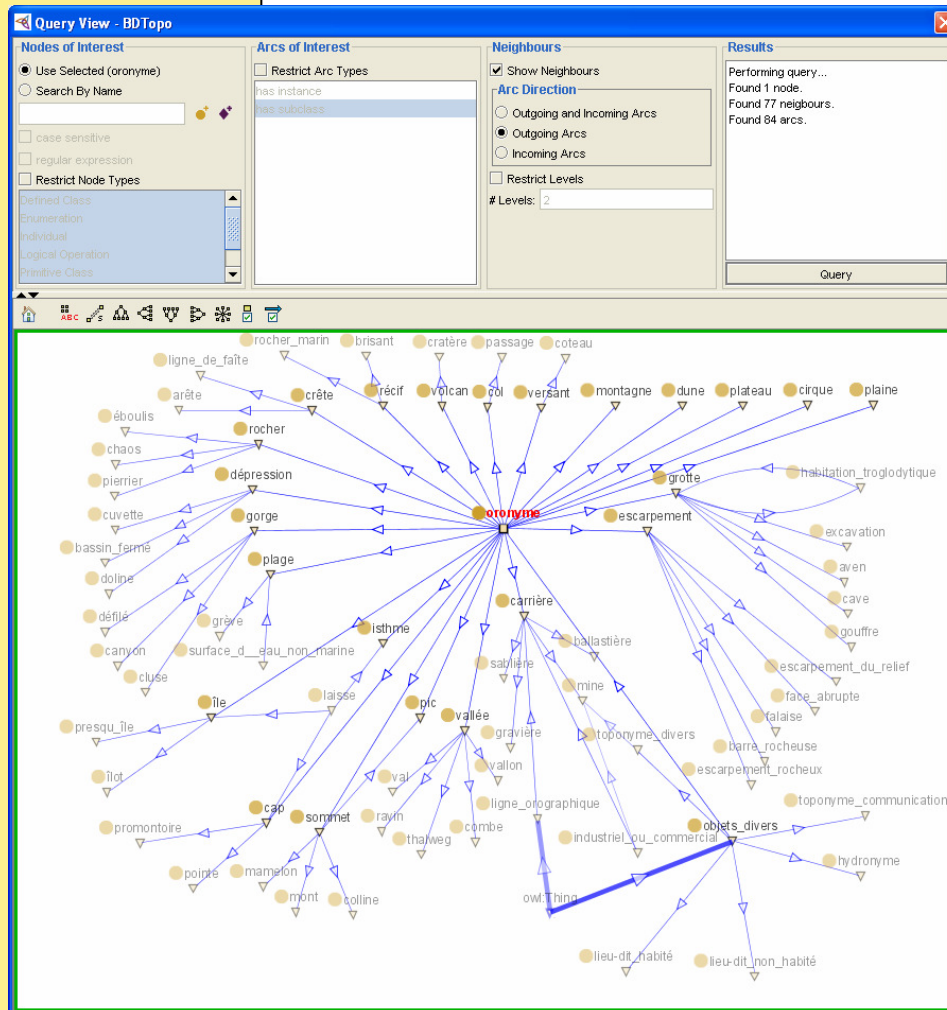
Résultats (1): création des ontologies [Laurens 2006]

Visu: Jambalaya

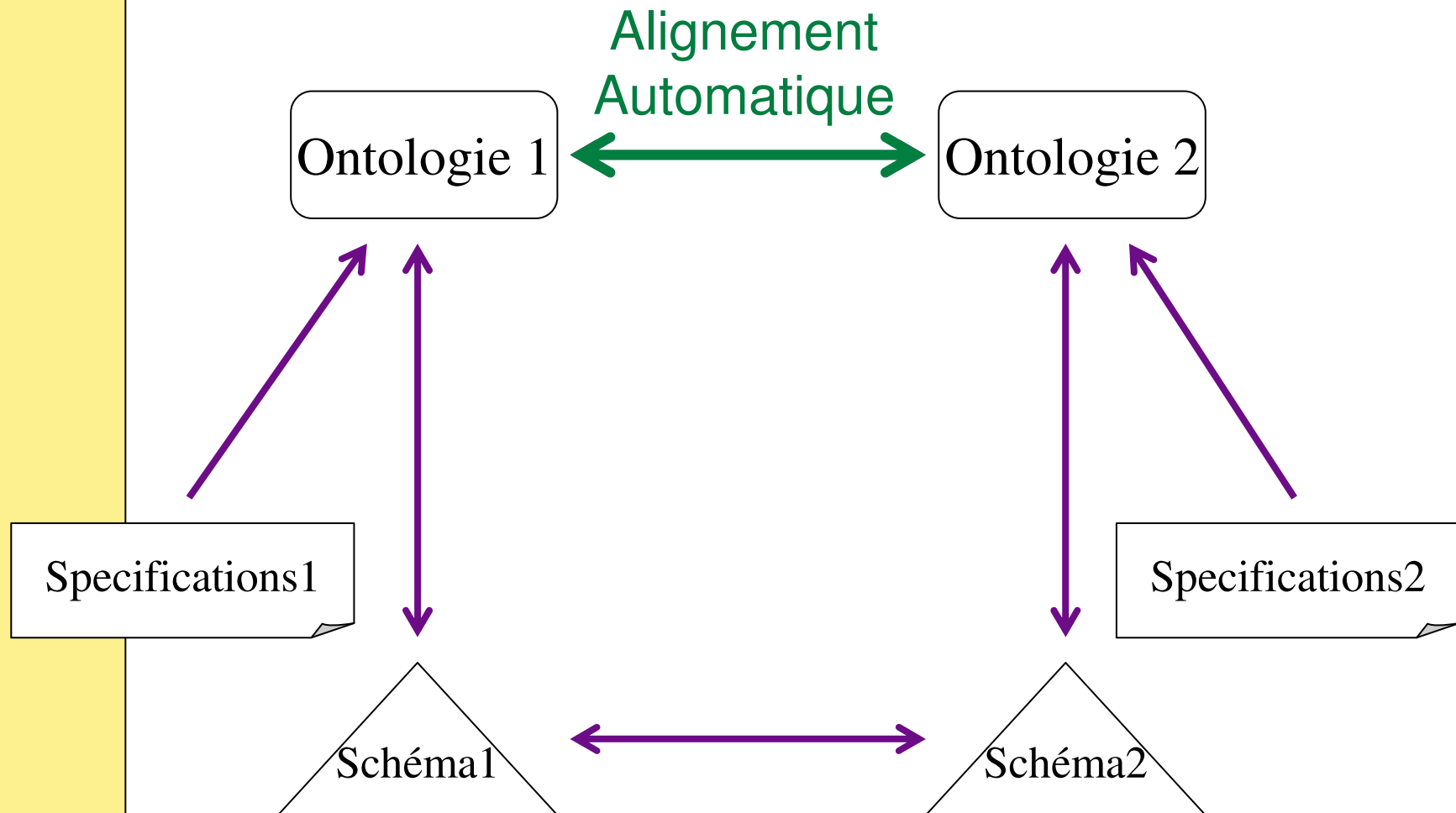


Résultats (1): création des ontologies [Laurens 2006]

Visu: Jambalaya



Résultats (2): Alignement d'ontologie [Abadie 2007]

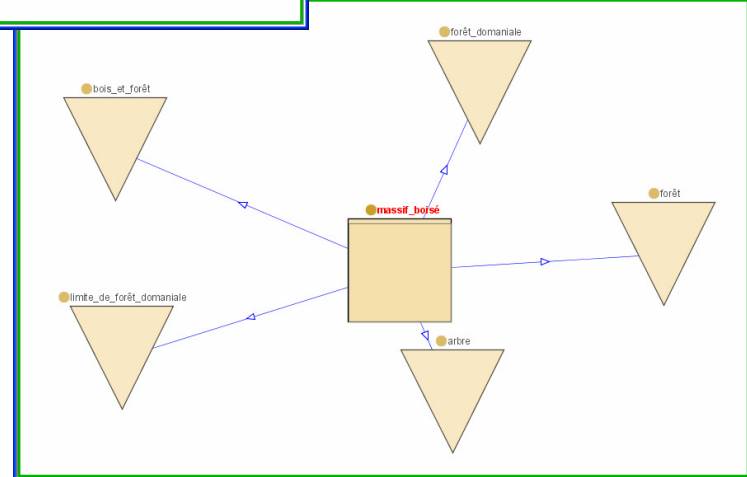
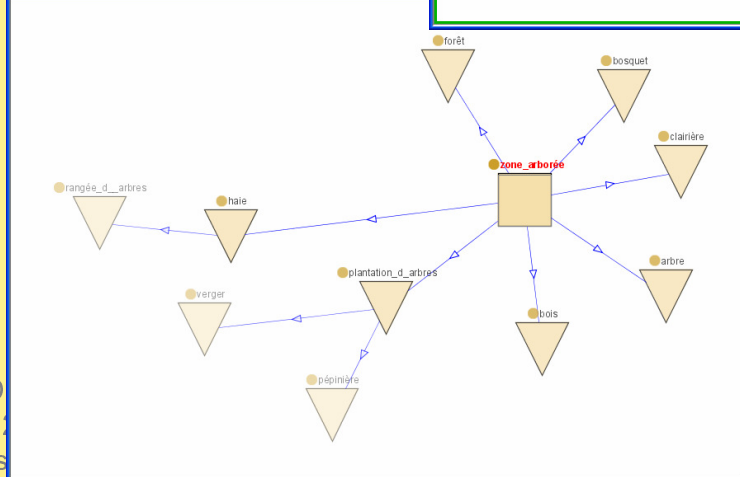
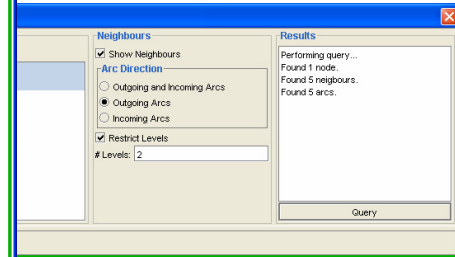
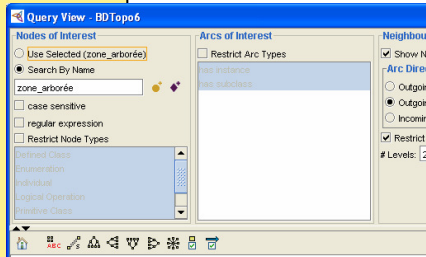
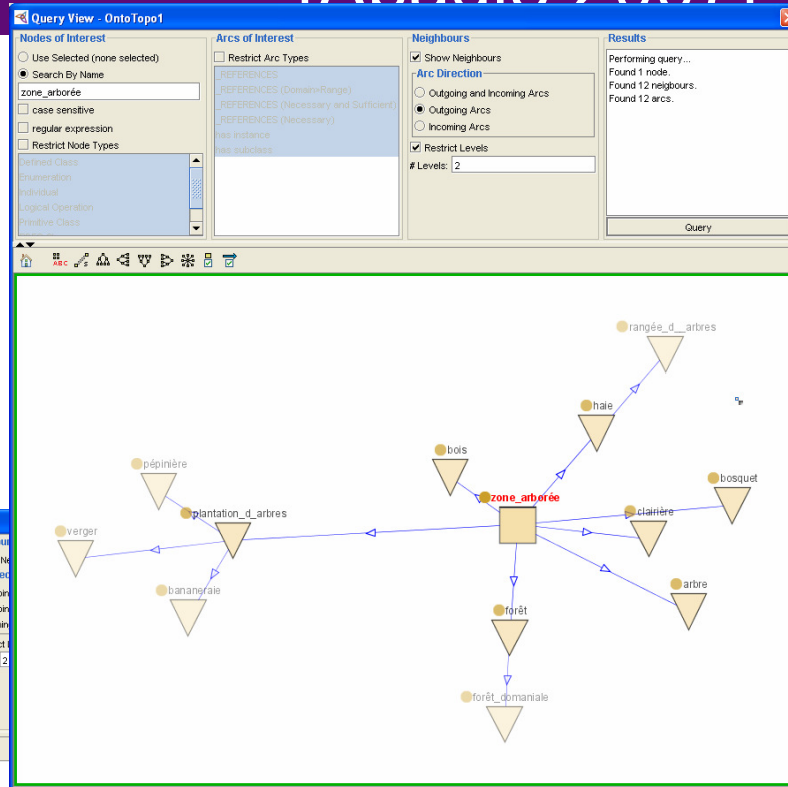


Résultats (2): Alignement d'ontologie [Abadie 2007]

- Utilisation de Prompt (plug-in Protégé)
 - Prometteur
 - Alignement automatique possible
 - Inférences intéressantes
 - A affiner
 - Erreurs: termes proches
 - Modification de la hiérarchisation pas toujours pertinente

Résultats (2): Alignement d'ontologie [Abadie 2007]

Visu: Jambalaya



Résultats (2): Alignement d'ontologie [Abadie 2007]

OntoTopo1 Protégé 3.2.1 (file:\E:\Users\Abadie\Ontologies\owl\OntologieGlobale\VersionComplete\OntoTopo1.pprj, OWL / RDF Files)

File Edit Project QWL Code Tools Window Prompt Help

Metadata (Ontology1169118417.owl) ● OWLClasses ■ Properties ◆ Individuals ■ Forms Prompt

Table view Tree view

Image table

f1	f2	renamed	operation	map level	rename explanation
	● aquarium	No	Add		<null>
	● aérotrain	No	Add		<null>
	● bac	No	Add		<null>
	● bananeraie	No	Add		<null>
	● bassin_de_compensation	No	Add		<null>
	● bassin_d_écluse	No	Add		<null>
	● bassin_portuaire_fluvial	No	Add		<null>
	● bras	No	Add		<null>
	● bras_de_décharge	No	Add		<null>
	● bras_naturel	No	Add		<null>
	● bras_secondaire	No	Add		<null>
	● bretelle_d_accès	No	Add		<null>
	● bretelle_d_accès_aux_...	No	Add		<null>
	● bretelle_d_échangeur	No	Add		<null>
	● broussaille	No	Add		<null>
	● bâtiment_de_caractère	No	Add		<null>
	● bâtiment_remarquable_ε...	No	Add		<null>
	● campanile	No	Add		<null>
	● canal_d_alimentation	No	Add		<null>
	● canal_de_décharge	No	Add		<null>
	● canal de dérivation	No	Add		<null>

Changed by:
Differences

Operation	Property	Restriction	Old Value	New Value

Résultats (2): Alignement d'ontologie [Abadie 2007]

BDTopo6 Protégé 3.2.1 (file:\E:\Users\Abadie\Ontologies\low\BDTopo\Alignement\BDTopo6.pprj, OWL / RDF Files)

File Edit Project OWL Code Tools Window Prompt Help

Table view Tree view

Image table

f1	f2	renamed	operation	map level	rename explanation
cours_d_eau	cours_d_eau	Yes	Map	Directly-changed	different delimiters
noeud_du_reseau_hydrographique_terrestre	extrémité_du_reseau_hydrographique_terrestre	Yes	Map	Directly-changed	Same superclass and subclasses
ancien_chef_lieu_de_commune	ancien_chef_lieu_de_commune	No	Map	Directly-changed	frame name and type are the same
arc_de_triomphe	arc_de_triomphe	No	Map	Directly-changed	frame name and type are the same
arrondissement_municipal	arrondissement_municipal	No	Map	Directly-changed	frame name and type are the same
autoroute	autoroute	No	Map	Directly-changed	frame name and type are the same
barrage_voûte	barrage_voûte	No	Map	Directly-changed	frame name and type are the same
berge	berge	No	Map	Directly-changed	frame name and type are the same
bretelle	bretelle	No	Map	Directly-changed	frame name and type are the same
cascade	cascade	No	Map	Directly-changed	frame name and type are the same
chef_lieu_de_commune	chef_lieu_de_commune	No	Map	Directly-changed	frame name and type are the same
château	château	No	Map	Directly-changed	frame name and type are the same
clocher	clocher	No	Map	Directly-changed	frame name and type are the same
collectivité_territoriale	collectivité_territoriale	No	Map	Directly-changed	frame name and type are the same
digue	digue	No	Map	Directly-changed	frame name and type are the same
dolmen	dolmen	No	Map	Directly-changed	frame name and type are the same
déversoir	déversoir	No	Map	Directly-changed	frame name and type are the same
habitation	habitation	No	Map	Directly-changed	frame name and type are the same
habitation_isolée	habitation_isolée	No	Map	Directly-changed	frame name and type are the same
hameau	hameau	No	Map	Directly-changed	frame name and type are the same
hippodrome	hippodrome	No	Map	Directly-changed	frame name and type are the same
hôpital_psychiatrique	hôpital_psychiatrique	No	Map	Directly-changed	frame name and type are the same
jardinierie	jardinierie	No	Map	Directly-changed	frame name and type are the same
laisse_de_hautes_mers	laisse_de_hautes_mers	No	Map	Directly-changed	frame name and type are the same
levée_de_terre	levée_de_terre	No	Map	Directly-changed	frame name and type are the same
maison_du_parc	maison_du_parc	No	Map	Directly-changed	frame name and type are the same
manufacture	manufacture	No	Map	Directly-changed	frame name and type are the same
marais_salant	marais_salant	No	Map	Directly-changed	frame name and type are the same
menhir	menhir	No	Map	Directly-changed	frame name and type are the same
moulin	moulin	No	Map	Directly-changed	frame name and type are the same
musée	musée	No	Map	Directly-changed	frame name and type are the same

Changed by:

Differences

Operation	Property	Restriction	Old Value	New Value
-----------	----------	-------------	-----------	-----------

Résultats (2): Alignement d'ontologie [Abadie 2007]

Visu: Jambalaya

The image displays three instances of the Jambalaya ontology alignment tool, each showing a different query configuration and its resulting network visualization. The tool's interface includes several control panels:

- Nodes of Interest:** Allows searching for specific nodes (e.g., 'élément_hydrographique_terrestre') and filtering by type (e.g., 'Defined Class', 'Enumeration').
- Arcs of Interest:** Filters relationships based on types like 'REFERENCES (Domain=Range)' or 'has instance'.
- Neighbours:** Controls the search for related nodes, including 'Arc Direction' (Outgoing, Incoming, or Both) and the number of levels to explore.
- Results:** Displays the outcome of the query, such as the number of nodes and arcs found.

The network visualizations show a central node, 'élément_hydrographique_terrestre', which is connected to a wide variety of other hydrological features. The connections are represented by blue arrows, indicating the direction of the relationships. The nodes are color-coded: red for the central node and yellow for other nodes. The three visualizations represent different query results, showing how different filters and parameters affect the network structure.

Bilan sur le passage à l'échelle: quelle automatisation ?

- Création de taxonomie/ontologie
 - Contrôle interactif nécessaire
 - Mais de moins en moins long / richesse d'ontologies existantes

- Instanciation des liens schéma – ontologie
 - Liens simples:
 - globalement OK,
 - à gérer : « à l'exception de... », « en zone de... »
 - Liens riches ?

- Alignement d'ontologie
 - OK ?
 - Contrôle interactif nécessaire

- Inférence pour apparier les schémas
 - non testé

Suites prévues

- D'une taxonomie vers une ontologie
 - Enrichissement des liens entre concepts
 - Quelle source de connaissances pour cela ?

- Liens entre schémas et ontologie
 - Modélisation OK
 - Instanciation par analyse des spécifications (TALN)

- Alignement
 - Peu de problèmes soulevés pour l'instant
 - Mais cela est sûrement dû à la relative pauvreté des taxonomies qui doivent être enrichies

- Collaborations bienvenues !!!

Autres voies à explorer

- Appariier les données pour appariier les schémas
 - Appariement géométrique [Mustière 2002]
 - Boucle récursive appariement de données – appariement de schémas [Olteanu 2007]

- Autres utilisations de l'ontologie
 - Accès aux données
 - Expression de besoin