



Un cadre conceptuel et une approche de gestion du risque pour supporter l'interopérabilité sémantique des cubes de données spatiales

Étudiant : **Tarek Sboui**

Directeur : **Yvan Bédard**

Co-directeurs : **Jean Brodeur**

Thierry Badard

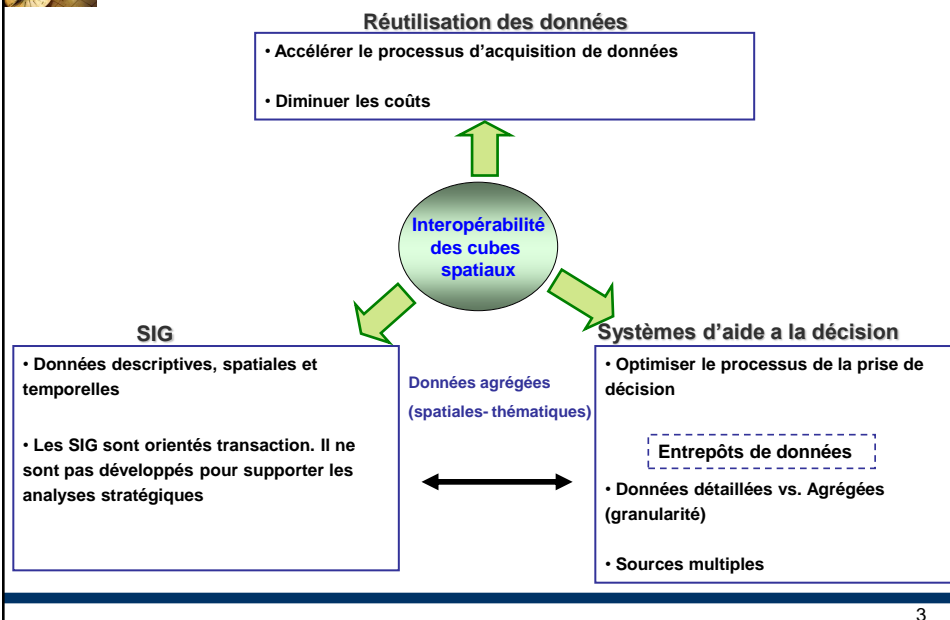


Plan

- Mise en contexte
- C'est quoi l'interopérabilité des cubes ?
- Conflits sémantiques liés à l'interopérabilité
- Approche proposée
 - Cadre conceptuel pour l'interopérabilité des cubes
 - Gestion du risque de mauvaise interprétation
- Résultats
- Conclusion



La géomatique décisionnelle et la réutilisation



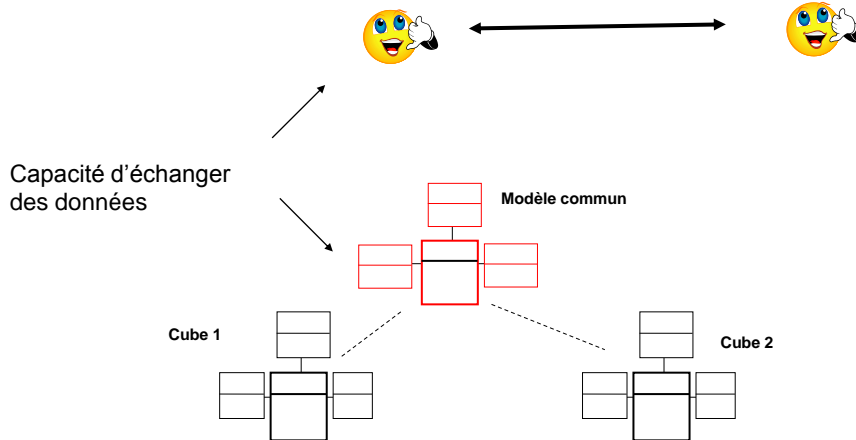
Besoin d'interopérer les cubes de données spatiales

- Navigation simultanée et rapide des cubes de données.
- Insertion rapide de données dans un cube de données.
- Comparaison rapide et interactif des données décisionnelles pour analyser les changements de phénomènes.



C'est quoi l'interopérabilité des cubes ?

L'interopérabilité est la capacité de deux ou plusieurs systèmes ou composants d'échanger des informations et de les utiliser

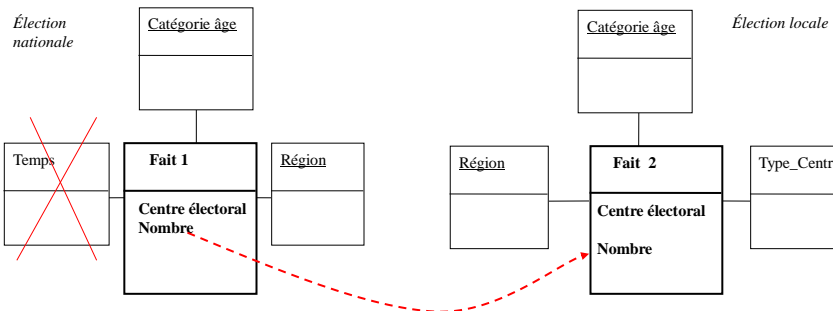


5



C'est quoi l'interopérabilité des cubes ?

Ajouter une mesure d'un cube à un autre

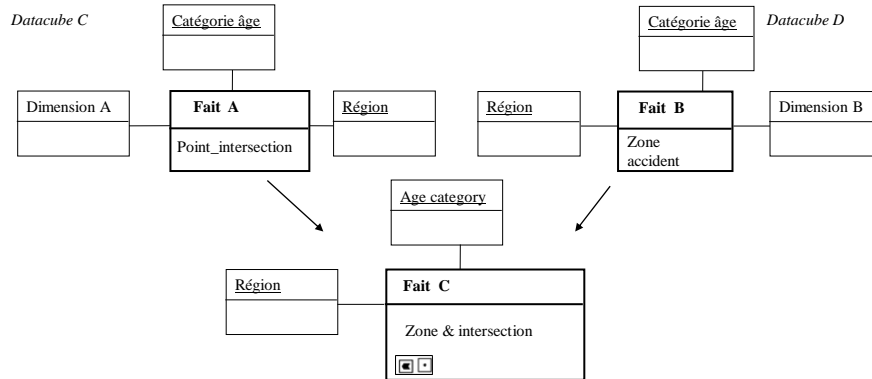


6



C'est quoi l'interopérabilité des cubes ?

Intégrer des mesures

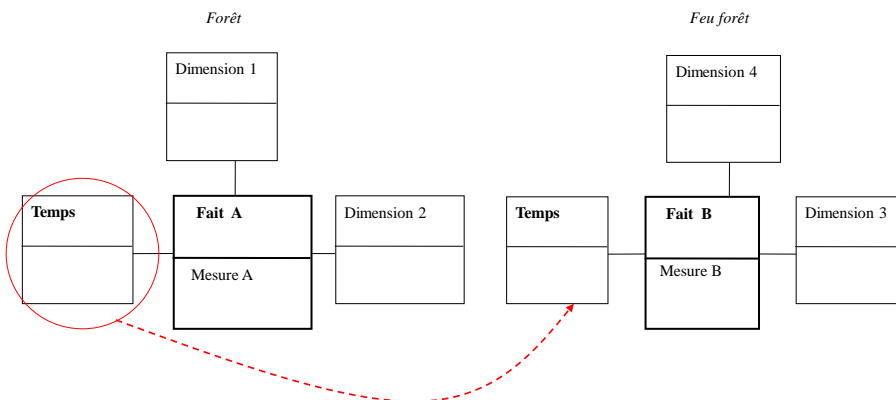


7



C'est quoi l'interopérabilité des cubes ?

Ajouter une dimension d'un cube à un autre

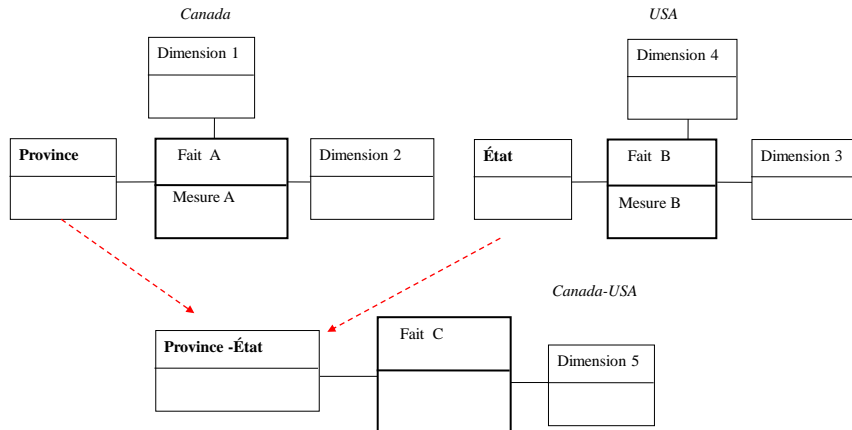


8



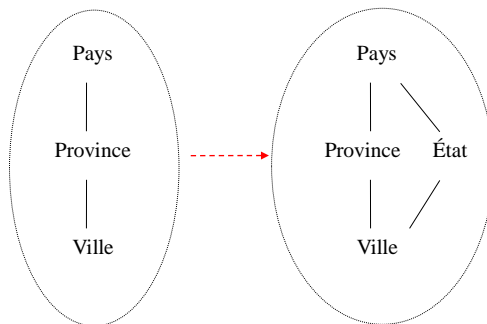
C'est quoi l'interopérabilité des cubes ?

Intégrer des dimensions



C'est quoi l'interopérabilité des cubes ?

Intégrer des dimensions



Une ou plusieurs décisions doivent être prises. Ex: décider de considérer ou non un élément à ajouter dans un autre cube.



Interopérabilité entre les cubes VS Interopérabilité entre les BD transactionnelles

interopérabilité sémantique entre les BD transactionnelles	interopérabilité sémantique entre les cubes
Similarités	
Réutilisation de données	
Différences	
Traiter les hétérogénéités des concepts transactionnels (i.e. tables, attribues, relations, etc.).	Traiter les hétérogénéités des concepts décisionnels : <ul style="list-style-type: none">- faits, mesures, dimensions, niveaux- méthodes et condition d'agrégation, méthodes pour calculer les mesures.

11



Problème

Absence d'une solution spécifique à l'interopérabilité sémantique entre les cubes de données spatiales.

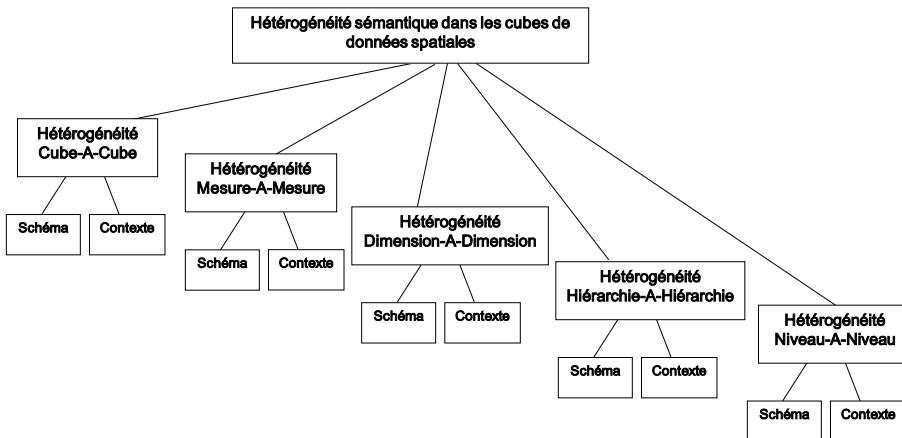
12



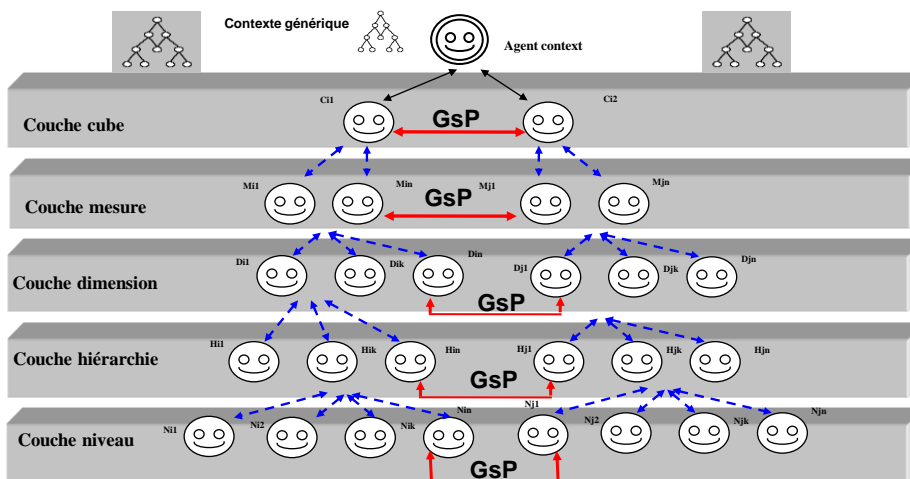
Conflits sémantiques

Hétérogénéité sémantique

Différence de signification, d'interprétation et d'utilisation de données



Approche - Cadre conceptuel





Ontologies: résolution imparfaite des conflits sémantiques

- Souvent, la mesure de similarité sémantique entre les concepts (appariement des ontologies) est une estimation. i.e. résolution imparfaite des problèmes d'hétérogénéité.

→ Une interprétation est **plus ou moins fidèle** à la signification d'origine.

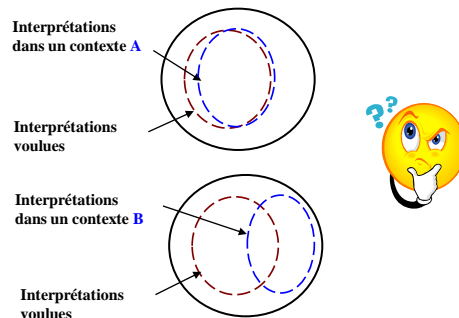
15



Ontologies: interprétation plus ou moins pertinente

- Dans des contextes différents, une ontologie peut référer à des interprétations différentes.

→ Une interprétation est **plus ou moins pertinente** au contexte courant. Les agents ignorent souvent le degré de pertinence.



16



Risques de mauvaise interprétation

→ **Risque de mauvaise interprétation: fausse interprétation ou incertitude à propos de l'interprétation appropriée.**

17



Risques de mauvaise interprétation : impact

- Une mauvaise interprétation affecte le processus de l'interopérabilité sémantique entre les datacubes géospatiales.
- À chaque niveau, une mauvaise interprétation peut avoir un impact sur les décisions des agents.
 - Ex. ajouter ou non un élément d'un cube à un autre.

18



Approche - Gestion des risques de mauvaise interprétation

- On se base sur le paradigme de gestion du risque :
 - Définir les risques.
 - Évaluer les risques.
 - Répondre aux risques.

19



Evaluation de la qualité externe du modèle conceptuel

- Une bonne qualité externe du modèle de données (schéma et contexte de production de données) indique un faible risque.
- Une mauvaise qualité externe du modèle de données indique un risque élevé.
- On évalue la qualité externe du modèle de données pour définir et évaluer les risques.

20



Evaluation de la qualité externe du modèle conceptuel

- Ensemble d'indicateurs et une approche qualitative pour évaluer la qualité externe des éléments du modèle par rapport aux besoins de l'utilisation.

21



Evaluation de la qualité externe du modèle conceptuel

- Ces indicateurs aident les agents à prendre une décision appropriée. Ex:
 - ne pas considérer un élément qui a une qualité inférieure à un seuil.
 - choisir l'élément qui a une meilleure qualité de schéma et de contexte (i.e. l'élément qui convient le mieux au contexte courant).

22



Evaluation de la qualité externe du modèle conceptuel

- **Pertinence de la primitive géométrique** : cet indicateur indique la pertinence de la primitive géométrique utilisée pour représenter un élément du modèle de cube par rapport à la primitive géométrique requise par l'utilisateur final.

Élément des modèles de cubes	Besoins			
	0D	1D	2D	3D
0D	1	0	0.5	0.75
1D	0	1	0	0
2D	0.75	0	1	0.5
3D	0.5	0	0.75	1

23



Evaluation de la qualité externe du modèle conceptuel

- **Pertinence de structure** : Il indique le ratio des composants de l'élément du modèle qui sont sémantiquement liés avec ceux de l'élément requis.

$$P_s(E, R) = \begin{cases} \frac{N_{CE}}{N_{CR}} ; & \text{si } N_{CE} < N_{CR} \\ 1 ; & \text{si } N_{CE} = N_{CR} \end{cases}$$

- N_{CR} est le nombre total de composants requis par l'utilisateur,
- N_{CE} est le nombre de composants de l'élément qui sont sémantiquement liés avec les composants requis.

24



Evaluation de la qualité externe du modèle conceptuel

- **Pertinence de l'ordre de la hiérarchie** : indique la pertinence de l'ordre de chacune des hiérarchies hétérogènes par rapport à la structure de la hiérarchie requise par l'utilisateur final.

$$o_s : \forall n_1, n_2 \in H, \forall n_1', n_2' \in B : [n_1 < n_2] \Rightarrow n_1' < n_2'$$

- Où n_1, n_1' sont sémantiquement liés, and n_2, n_2' sont sémantiquement liés. La valeur 1 (ou 0) est assignée à la pertinence élémentaire d'ordre (o_s) entre chaque paire de niveaux lorsque l'expression est vraie (ou fausse).

25



Evaluation de la qualité externe du modèle conceptuel

- **Pertinence du contexte de production** : indique le degré de pertinence du contexte de production par rapport aux besoins de l'utilisateur final.

$$P_m = \begin{cases} w \times \frac{N_{Elt}}{N_{ReqElm}} & ; \text{si } N_{Elt} \leq N_{ReqElm} \\ 1 & ; \text{sinon} \end{cases}$$

- N_{Elt} est le nombre d'éléments thématiques, spatiaux, ou temporels de contexte de production sémantiquement liés aux éléments requis par l'utilisateur, N_{ReqElm} est le nombre total d'éléments de contexte de production requis, et w une valeur prédéfinie entre 0 et 1 qui indique le niveau d'importance de chaque type (i.e., thématique, spatial, et temporel) pour l'utilisateur final.

26



Evaluation de la qualité externe du modèle conceptuel

- **Actualité du contexte de production** : indique le degré d'actualité des contextes de production des cubes de données spatiales. Il est évalué en fonction de l'âge des contextes de production par rapport à leur durée de vie.

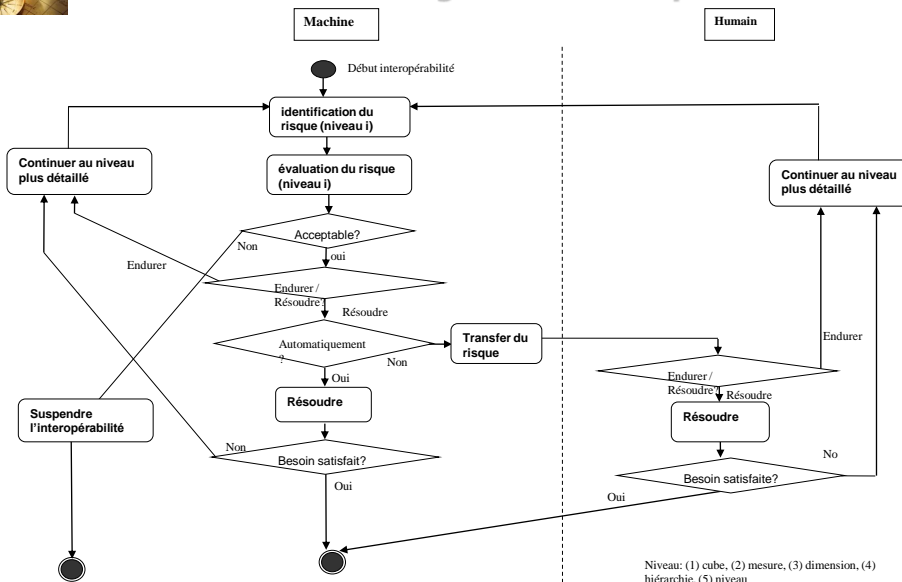
$$A_m = \begin{cases} 1 - \frac{|T_{requis} - T_{def}|}{DV} ; & \text{si } |T_{requis} - T_{def}| < DV \\ 0 ; & \text{sinon} \end{cases}$$

- Où DV est la durée de vie des contextes de production. La durée de vie et la date de définition des contextes de production peuvent être fournies par le producteur de contextes de production.


27



Processus de gestion du risque




28



Résultats


Élément 1	Élément 2	GSP	Résultat de l'évaluation	Proposition
Région administrative	Région	Égal	<ul style="list-style-type: none"> - Qs (Région administrative) = 1 - Qm (Région administrative) = 0.95 - Qs (Région) = 1 - Qm (Région) = 0.365 	<ul style="list-style-type: none"> - Pour la dimension Région administrative : continuer au niveau le plus détaillé. - Pour la dimension Région : continuer au niveau le plus détaillé. Mais faire attention parce que la qualité est faible.
H1: Ville, province, territoire, pays	H2: Ville, province, territoire, pays	Égal	<ul style="list-style-type: none"> - Qs (H1) = 1 - Qm (H1) = 1 - Qs (H2) = 1 - Qm (H2) = 0 	<ul style="list-style-type: none"> - Pour l'hierarchie H1 : continuer au niveau le plus détaillé. - Pour l'hierarchie H1: a ne pas considérer H2 dans le processus de l'interopérabilité.
Ville	Ville	Égal	<ul style="list-style-type: none"> - Qs (Ville) = 0.5 - Qm (Ville) = 1 	<ul style="list-style-type: none"> - L'utilisateur peut considérer cet élément.

29



Résultats

Warning



Message content

The fitness-for-use of schema is not very good, be careful if you continue the interoperability with the corresponding level

30



Résultats



31



Conclusion

- Définition de l'interopérabilité des cubes
- Catégorisation des conflits sémantiques dans les cubes spatiaux
- Définition d'un cadre conceptuel d'interopérabilité des cubes spatiaux
- Définition d'une approche pour gérer le risques de mauvaise interprétation
 - Identification et évaluation du risque de mauvaise interprétation de données.
 - Aider à protéger les agents contre le risque de mauvaise interprétation.

32



Conclusion

- Proposition d'un ensemble d'indicateurs et une méthode permettant d'évaluer la qualité externe des schémas et du contexte de production.
- Permet d'identifier un certain nombre de risques.
- Aide les agents (humain ou machine) à mieux gérer les risques, pas nécessairement les résoudre entièrement.

33



Remerciements

- Nous remercions la Chaire CRSNG de recherche industrielle en bases de données géospatiales décisionnelles ainsi que ses partenaires financiers :



34



Merci pour votre attention