

Données géohistoriques ouvertes et liées : les transformations territoriales de Montréal sous la loupe des graphes de connaissances

David Valentine, doctorat en sciences de l'information — david.valentine@umontreal.ca



Résumé

Ce projet de recherche porte sur la représentation sémantique des informations géohistoriques extraites de cartes historiques. Il vise à décrire les transformations territoriales de Montréal à l'aide des outils de modélisation du Web sémantique pour ouvrir l'écosystème montréalais des informations historiques au monde des données liées.

Problématique

- L'enrichissement sémantique des données géohistoriques sur Montréal demeure à ce jour passablement inexploité.
- La représentation des informations contenues dans les cartes historiques de la ville est toujours peu formalisée.
- Les informations géographiques à elles seules ne suffisent pas à générer du sens historique : d'autres informations doivent accompagner les formes et les toponymes [1].
- La sémantisation des données géohistoriques contribue à la construction des connaissances historiques [2].

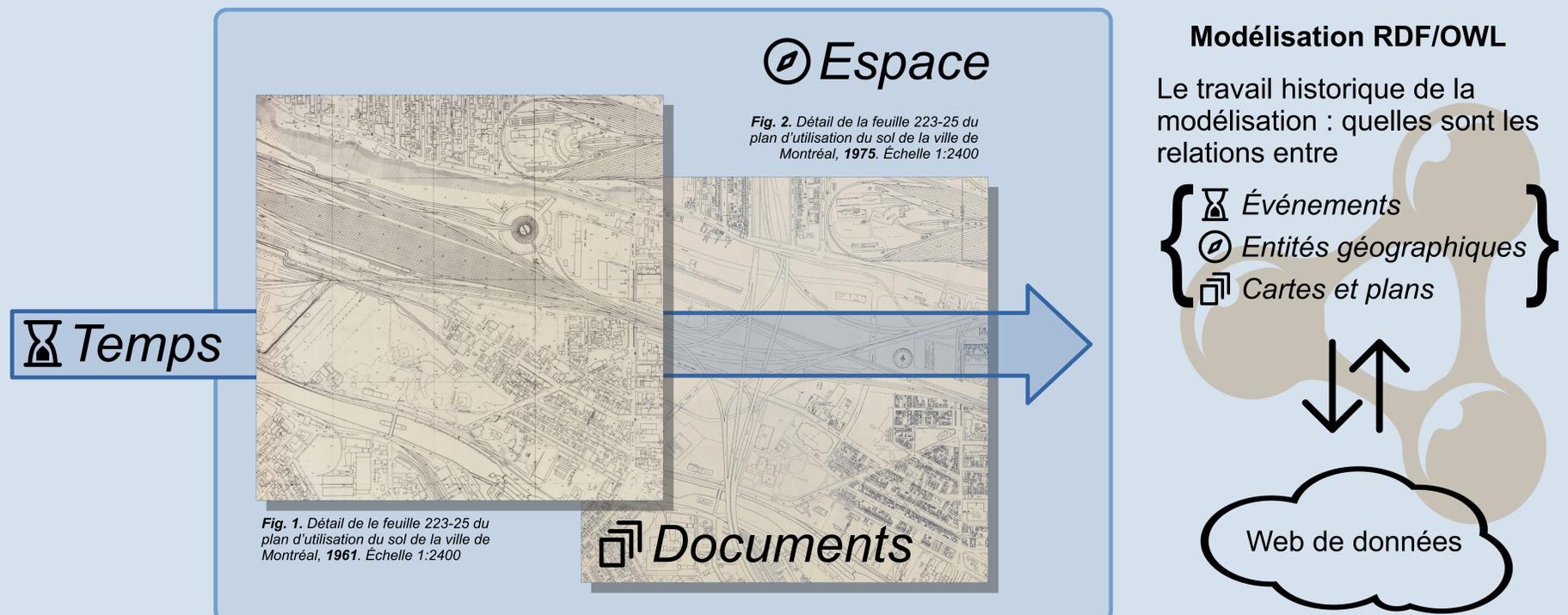
But de la recherche

- **Évaluer** l'apport potentiel des modèles du Web sémantique pour la représentation des informations extraites d'archives cartographiques, dans le contexte de l'écosystème des données géohistoriques à Montréal.
- **Expliquer** les possibilités et les limites de ces technologies sous l'angle des dynamiques liées aux transformations historiques du territoire de la ville, tel que celles-ci peuvent être représentées par une approche diachronique du corpus.

Objectifs spécifiques

1. **Formaliser** les transformations du territoire montréalais en constituant un jeu de données ouvertes et liées basé sur des graphes de connaissances spatiotemporelles.
2. **Comparer** le jeu de données constitué avec les structures en usage sur le terrain afin d'évaluer les possibilités et les pistes de réconciliation et d'intégration des données.
3. **Formuler** des recommandations sur la direction que pourrait prendre le développement des jeux de données ouvertes et liées pour l'intégration des informations géohistoriques.

Une approche diachronique structurée par le modèle RDF



Méthodologie

- **Devis explicatif** : comprendre, décrire et expliquer
- **Échantillonnage** :
 - ▶ **Constitution d'un corpus** de cartes et de plans historiques basée sur un échantillonnage théorique (période historique : 1840-1980).
- **Collecte de données** :
 - ▶ **Extraction des informations** : approche diachronique des documents soutenue par une grille d'observation structurée et un protocole de documentation.
 - ▶ **Observation** de l'ensemble des systèmes Web de l'écosystème des données géohistoriques à Montréal.
- **Modélisation sémantique** : basée sur les outils du Web sémantique et structurée par le protocole SAMOD.
- **Analyse** : visualisations d'information basées sur des graphes de connaissances et sur des systèmes d'information géographique.

Qualité de la recherche

- **Modélisation** : systématisme garanti par un protocole.
- **Données** : compatibles avec les principes FAIR.
- **Audit** : sollicitation des communautés d'intérêts.

Importance et retombées

- Perspectives nouvelles sur les transformations historiques du territoire montréalais.
- Proposition pour la sémantisation des données géohistoriques.
- Amélioration de l'accès aux contenus des documents.

Références

- [1] Grossner, K., Grunewald, S. et Mostern, R. (2022). Bringing places from the distant past to the present: A report on the World Historical Gazetteer. *International Journal on Digital Libraries*.
- [2] Beretta, F. (2022). Interopérabilité des données de la recherche et ontologies fondationnelles : un écosystème d'extensions du CIDOC CRM pour les sciences humaines et sociales. Dans N. Lasolle, O. Bruneau et J. Lieber (dir.), *Actes des journées humanités numériques et Web sémantique* (p. 2-22).