

## 1. Objectifs

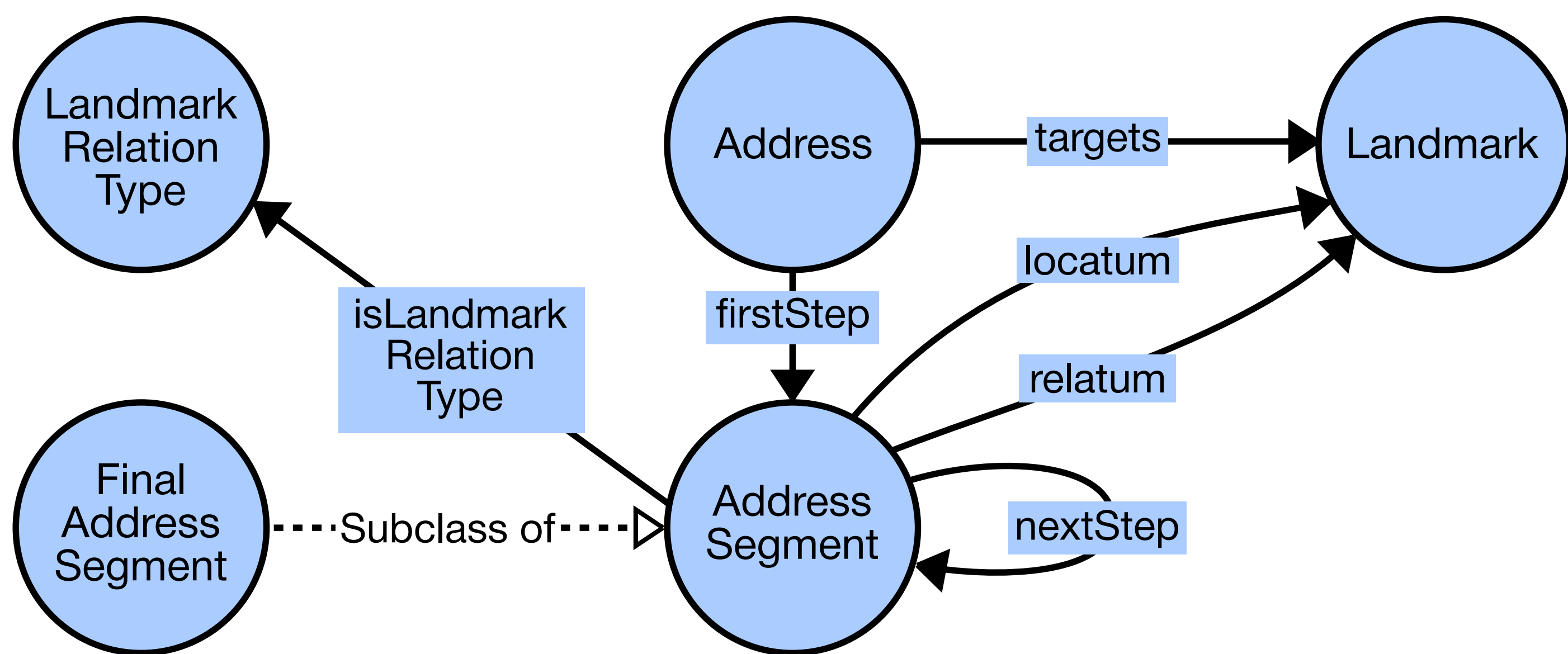
- Construction d'un référentiel géo-historique d'adresses à partir de différentes sources hétérogènes
- Adopter un modèle d'intégration de données historiques hétérogènes et fragmentaires : **graphe de connaissances**
- Représenter les états successifs des entités géographiques (rues, points d'intérêts, etc.) et les événements qui les font évoluer à partir des sources qui n'en décrivent pas.

## 2. Ontologie

Différentes manières de désigner un lieu avec une adresse :

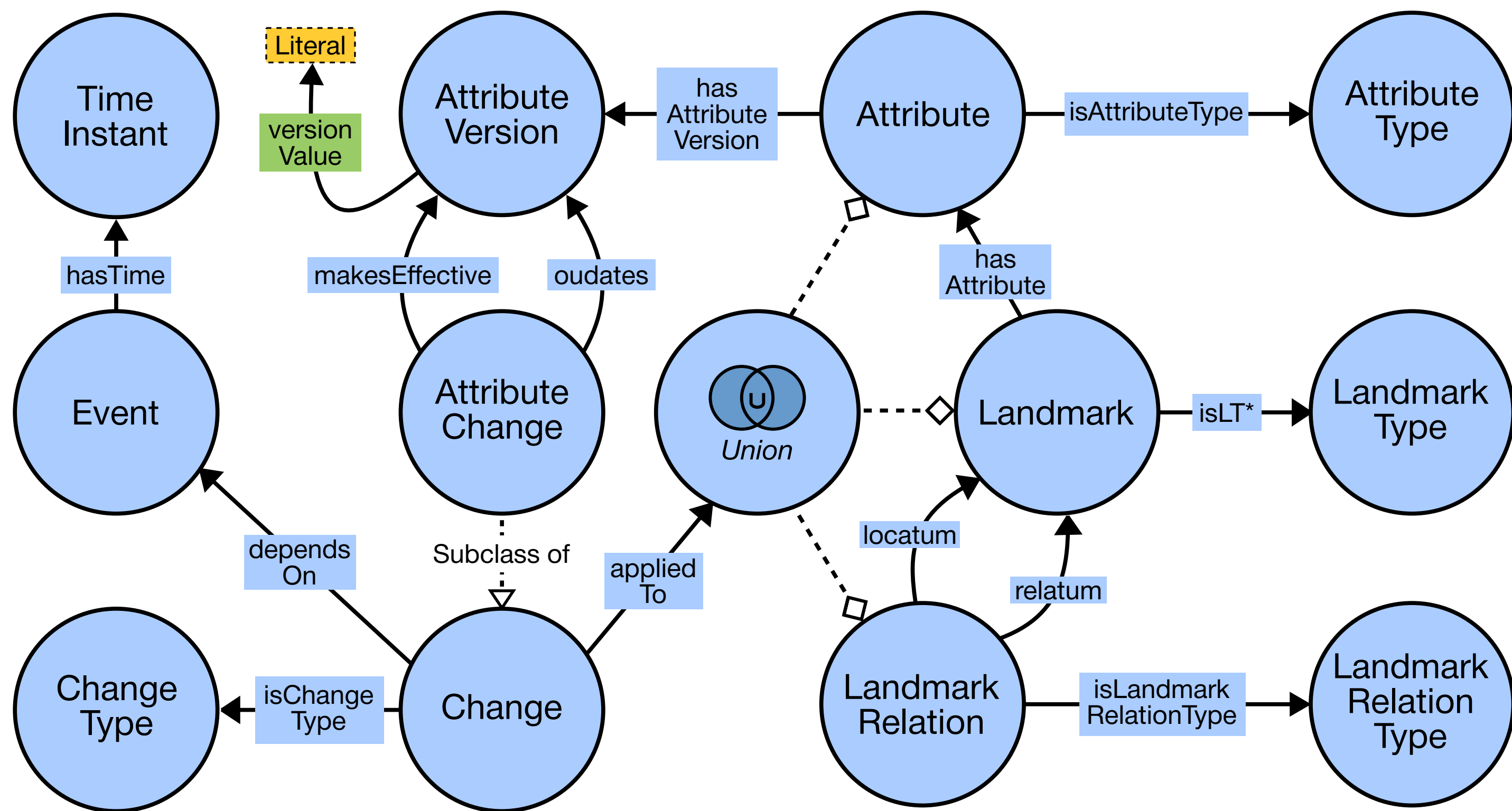
- 4 place de la Nation, 75004 Paris
- au coin du boulevard Diderot et de l'avenue Dorian

Adresse : énoncé structuré désignant un repère (ou entité géographique) via une suite ordonnée de relations spatiales (AddressSegment) entre repères (Landmark)[1]



Représentation simplifiée de l'ontologie des adresses

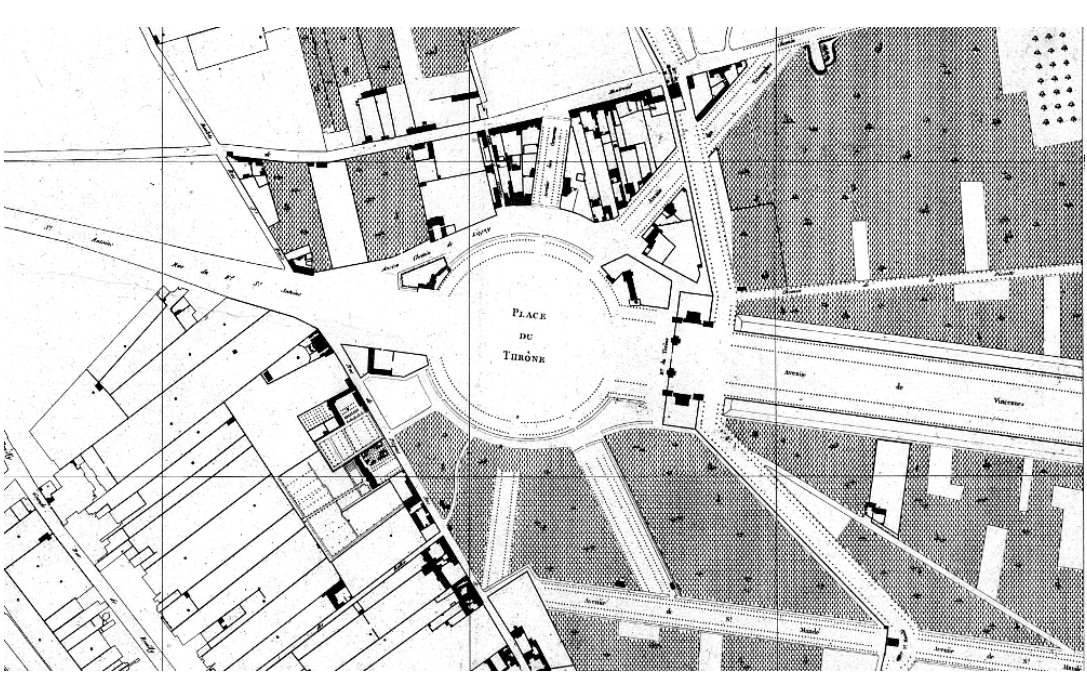
Évolution temporelle : description des événements (Event) représentant une évolution du territoire impliquant un ou plusieurs changements (Change) témoignant de l'évolution d'une ressource. [2]



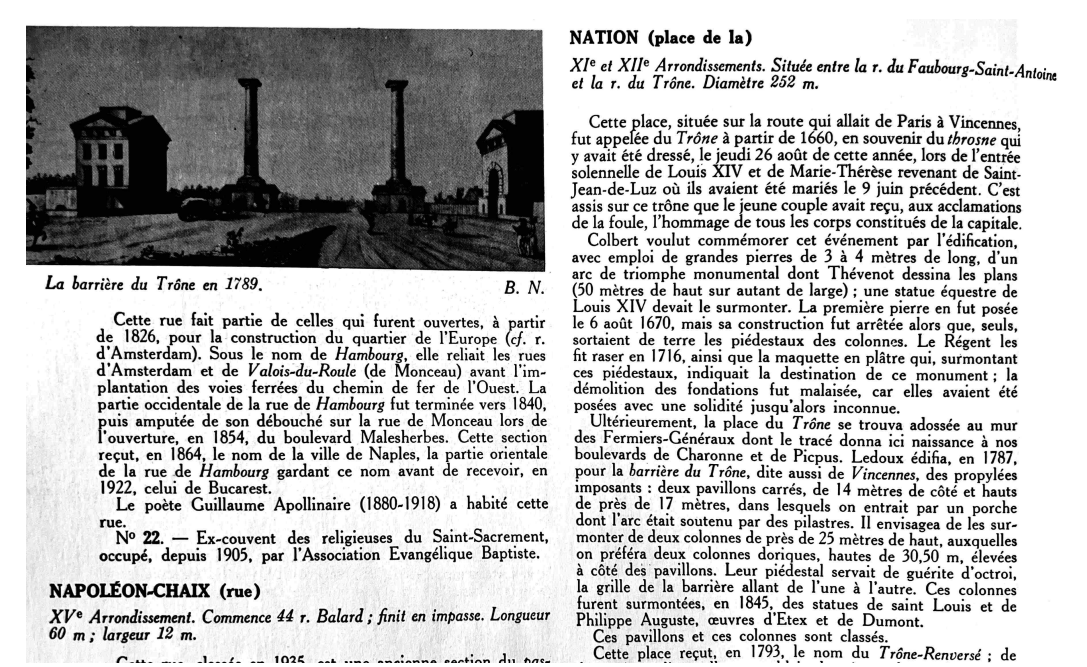
Modélisation des entités géographiques et de leur évolution temporelle

## 3. Cas d'application

Construire un graphe géo-historique d'adresses de la ville de Paris depuis le XVIII<sup>e</sup> siècle à partir de sources hétérogènes



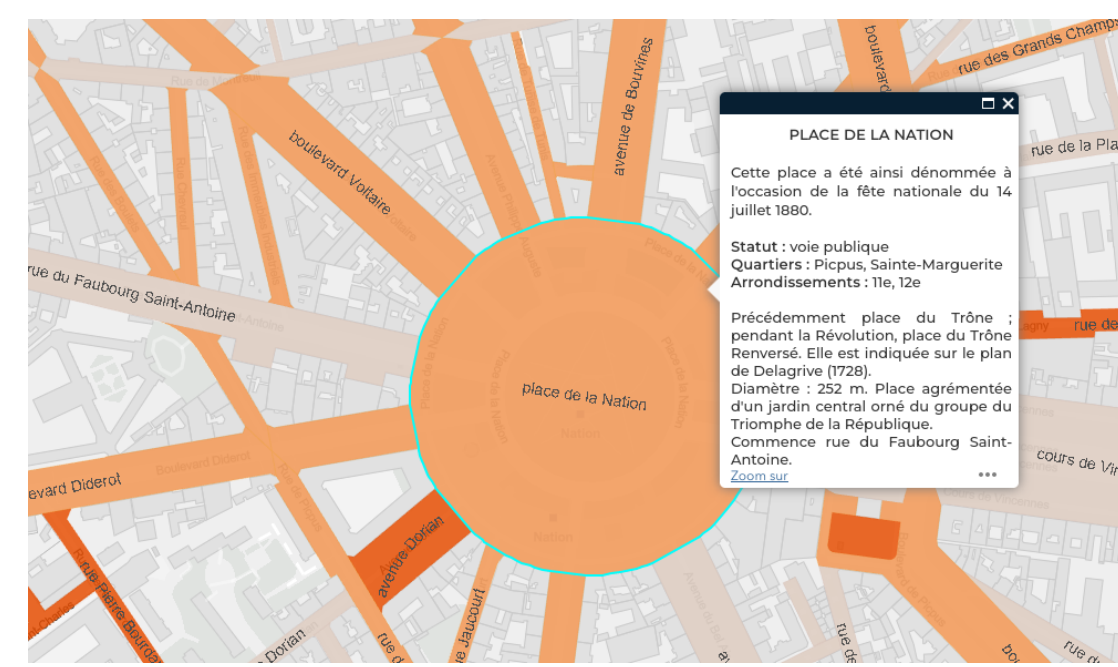
Edme Verniquet, Plan de Paris (1791)



Jacques Hillairet, Dictionnaire historique des rues de Paris (1960)

Bégault, *courtier gourmet*, q. Bourbon, 49.  
Begin, *Calandre*, 19.  
Begis, *f. colle-forte et huile de pieds de bœuf*, Ormeaux, 2, pl. du Trône.  
Béglé, *cordonnier*, Jacob, 56.  
Béglet et co., *commis. draps*, Thibautodé, 8.  
Valat, *vins en gros. pass. du Bureau*, 23.  
Valat fils, *vins et hôtel*, pl. de la Nation, 21, et av. du Trône, 1.  
Valax, *corsets*, r. Auber, 9.  
Valax, *tailleur*, r. Boudreau, 2.  
Valax, *vins*, r. Rochambeau, 10.

Extraits d'annuaires commerciaux (XIX<sup>e</sup> siècle)



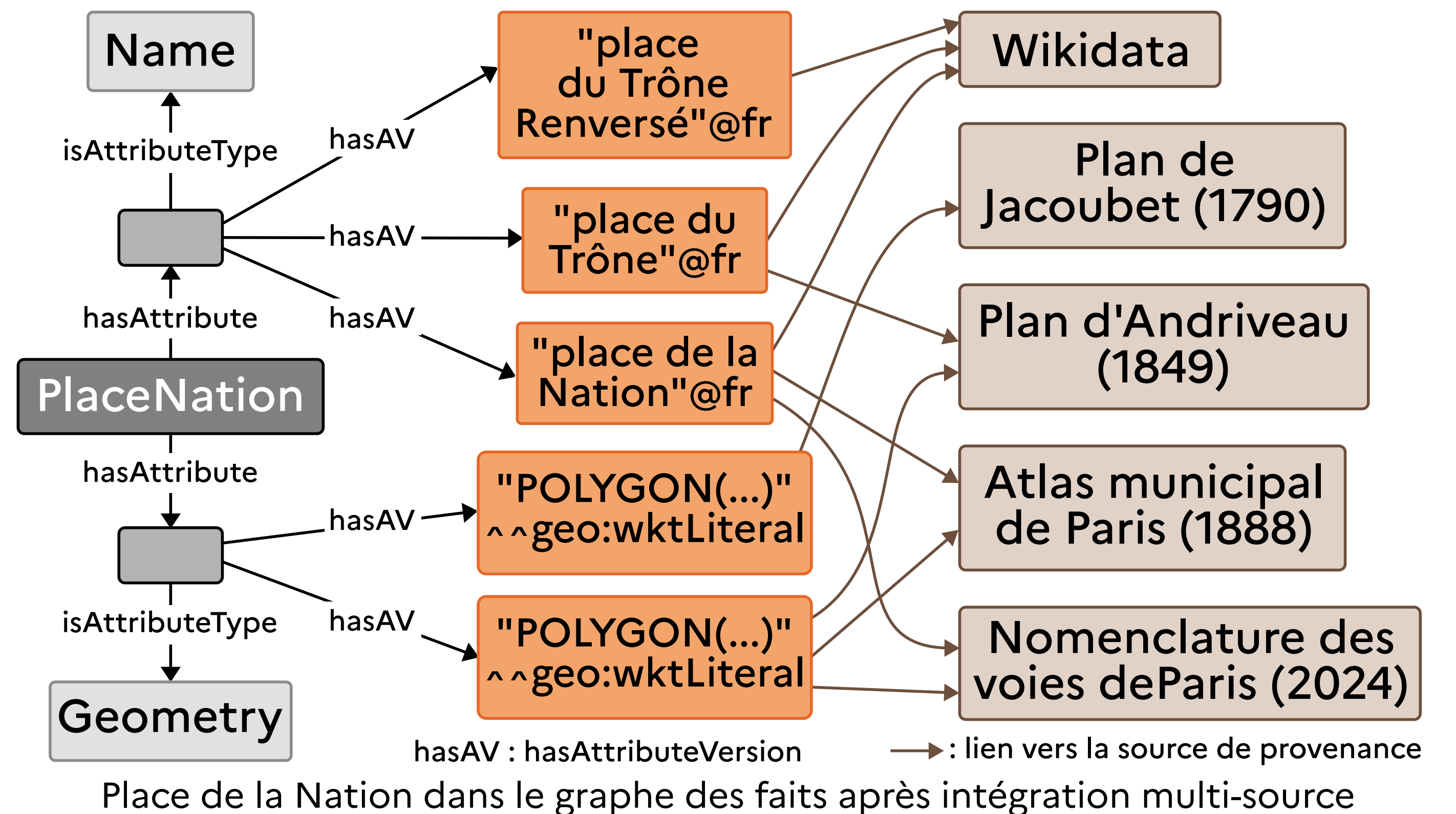
Ville de Paris, Dénominations des emprises des voies actuelles (2024)

## 4. Peuplement itératif

Comment intégrer les données provenant des diverses sources dans le graphe géo-historique ?

Approche itérative :

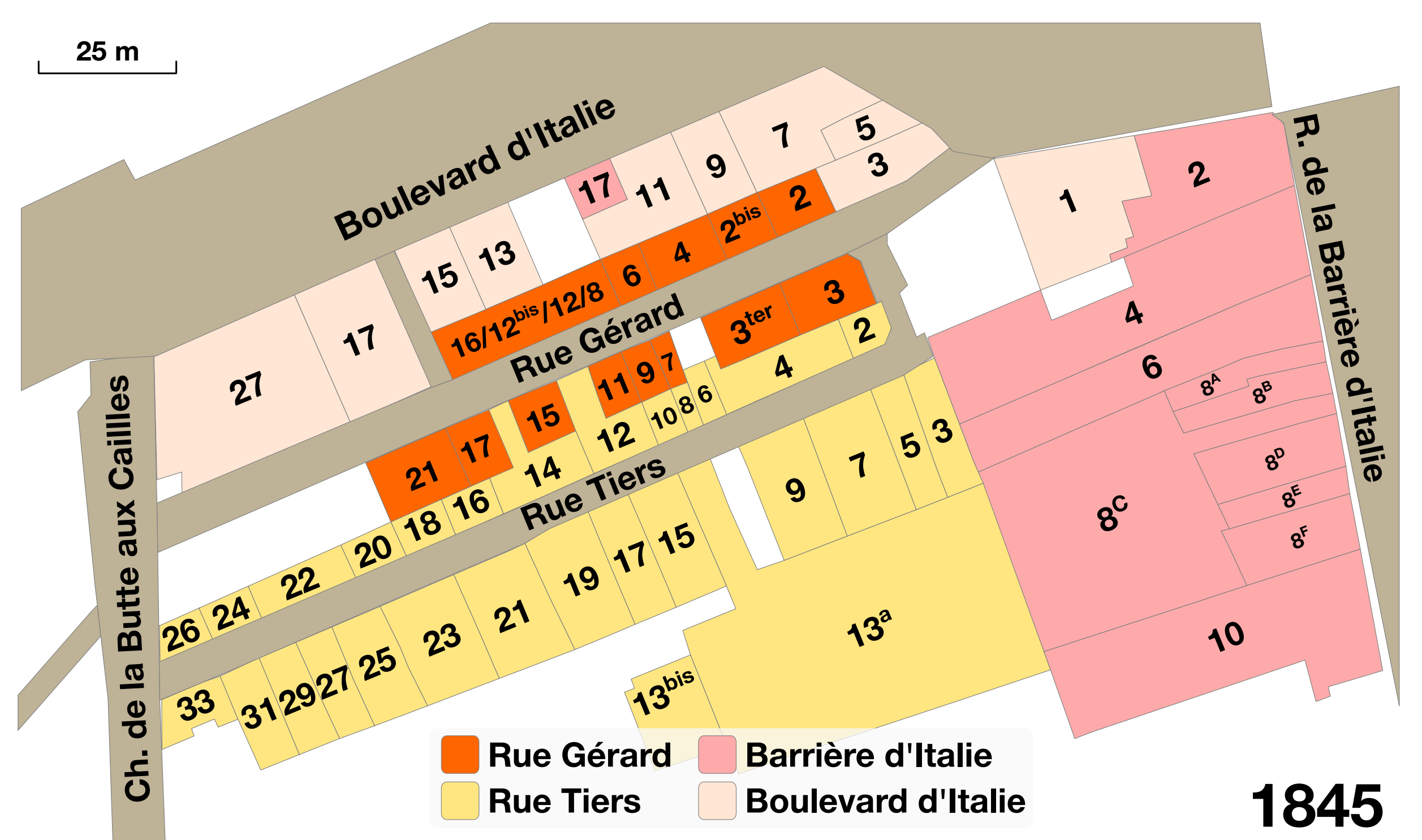
- pour chaque source, ses données sont structurées selon l'ontologie puis intégrées dans un graphe lié à cette source (graphe de factoides) [3] ;
- le graphe géo-historique (ou graphe de faits) centralise les données provenant des différentes sources ;
- construction itérative : chaque graphe de factoides est connecté à celui des faits, les ressources équivalentes sont liées via un lien owl:sameAs qui propage les informations des factoides vers les faits. Le graphe des factoides est ensuite déconnecté.



Place de la Nation dans le graphe des faits après intégration multi-source

## 5. Évaluation

- Tester la cohérence interne des données : règles SHACL et requêtes SPARQL vérifient si la construction itérative n'a généré aucune incohérence.
- Pas de vérité terrain sur laquelle se baser. Il n'existe pas de base complète qui permette
- Générer des snapshots du territoire et les comparer des données existantes et contemporaines.



## Références

- [1] Charly Bernard, Nathalie Abadie, Julien Perret, Bertrand Duménieu. Création d'un référentiel géo-historique d'adresses à partir de sources multiples. GAST, Jan 2024, Dijon, France.
- [2] Charly Bernard, Solenn Tual, Nathalie Abadie, Bertrand Duménieu, Joseph Chazalon, et al.. PeGazUs: A knowledge graph based approach to build urban perpetual gazetteers. International Conference on Knowledge Engineering and Knowledge Management, Nov 2024, Amsterdam, Netherlands.
- [3] Michele Pasin, John Bradley, Factoid-based prosopography and computer ontologies: towards an integrated approach, Digital Scholarship in the Humanities, Volume 30, Issue 1, April 2015, Pages 86–97, <https://doi.org/10.1093/llc/fqt037>