



HAL
open science

Ontologie et méthodologie de peuplement d'un graphe de connaissances géo-historique d'adresses

Charly Bernard, Nathalie Abadie, Julien Perret, Bertrand Duménieu

► To cite this version:

Charly Bernard, Nathalie Abadie, Julien Perret, Bertrand Duménieu. Ontologie et méthodologie de peuplement d'un graphe de connaissances géo-historique d'adresses. Symposium MaDICS, May 2024, Blois, France. 2024. hal-04601077

HAL Id: hal-04601077

<https://hal.science/hal-04601077v1>

Submitted on 4 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

1. Objectifs

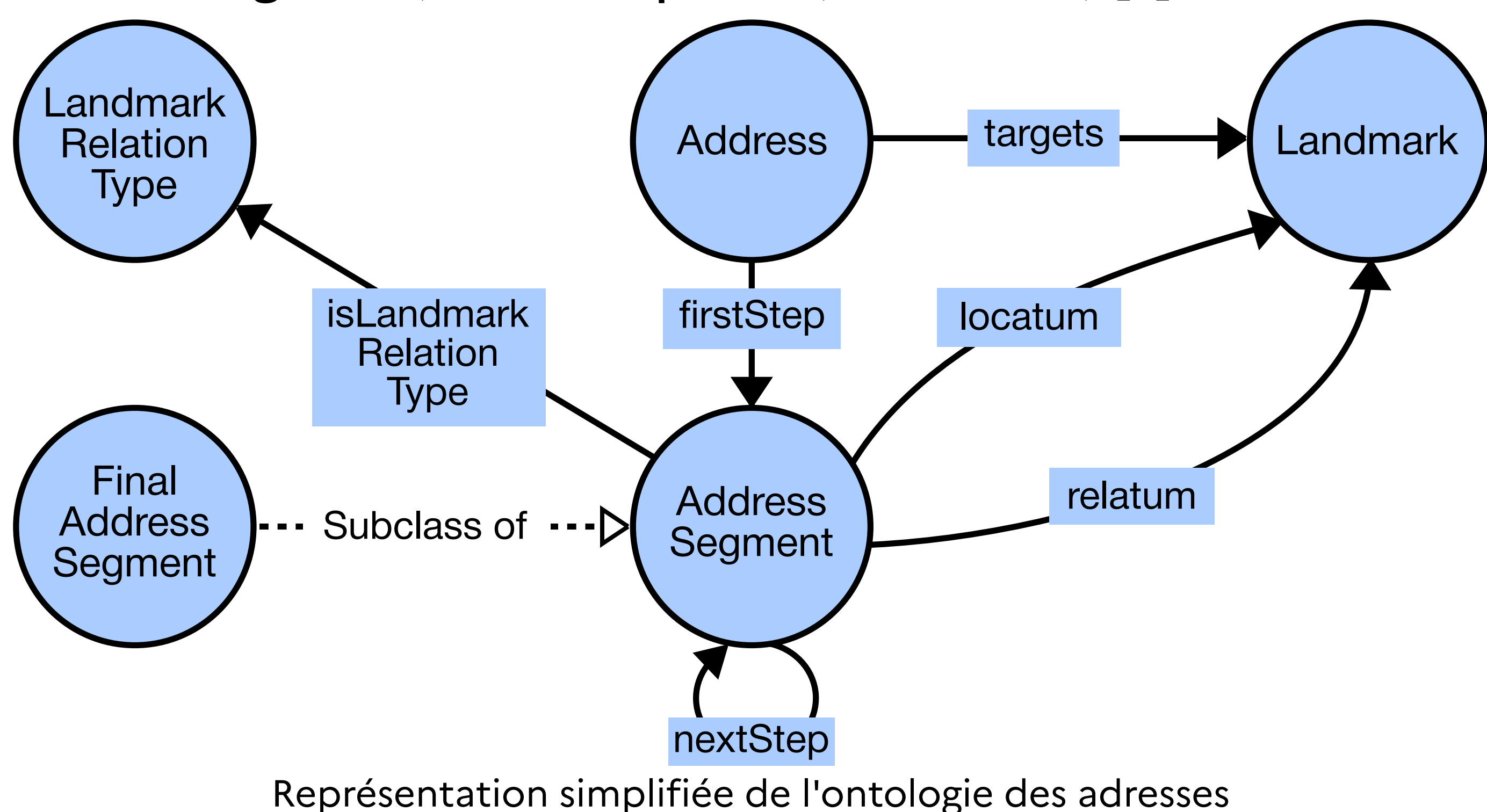
- Proposer une méthodologie de construction d'un référentiel géo-historique d'adresses à partir de différentes sources hétérogènes
- Adopter un modèle souple capable d'intégrer des données historiques qui sont hétérogènes et fragmentaires : **graphe de connaissances**
- Représenter les états successifs des entités géographiques (rues, points d'intérêts, etc.) et les événements qui les font évoluer (changement de nom, extension, fusion, destruction...) à partir des sources qui n'en décrivent pas.

2. Ontologie

Différentes manières de désigner un lieu avec une adresse :

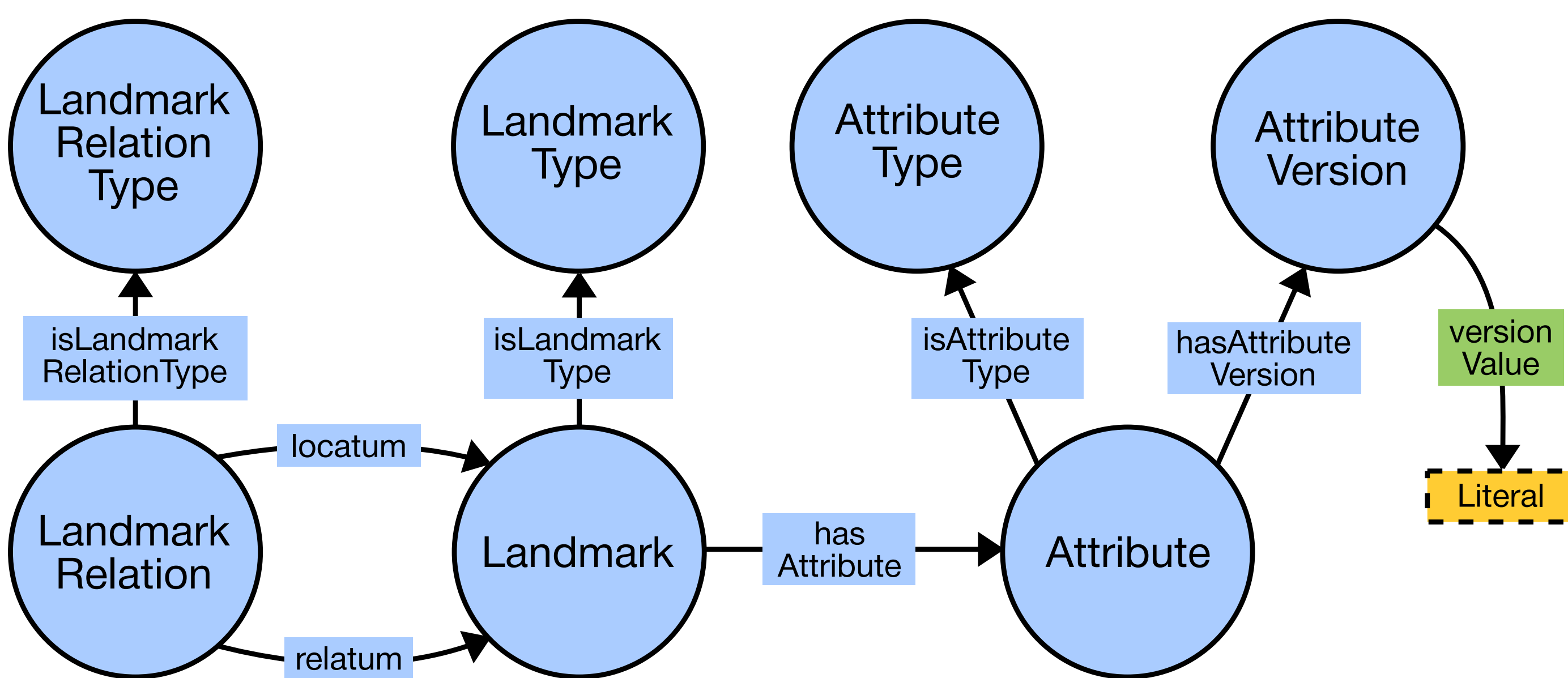
- 4 place de la Nation, 75012 Paris
- au coin du boulevard Diderot et de l'avenue Dorian, en face du kiosque à musique

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



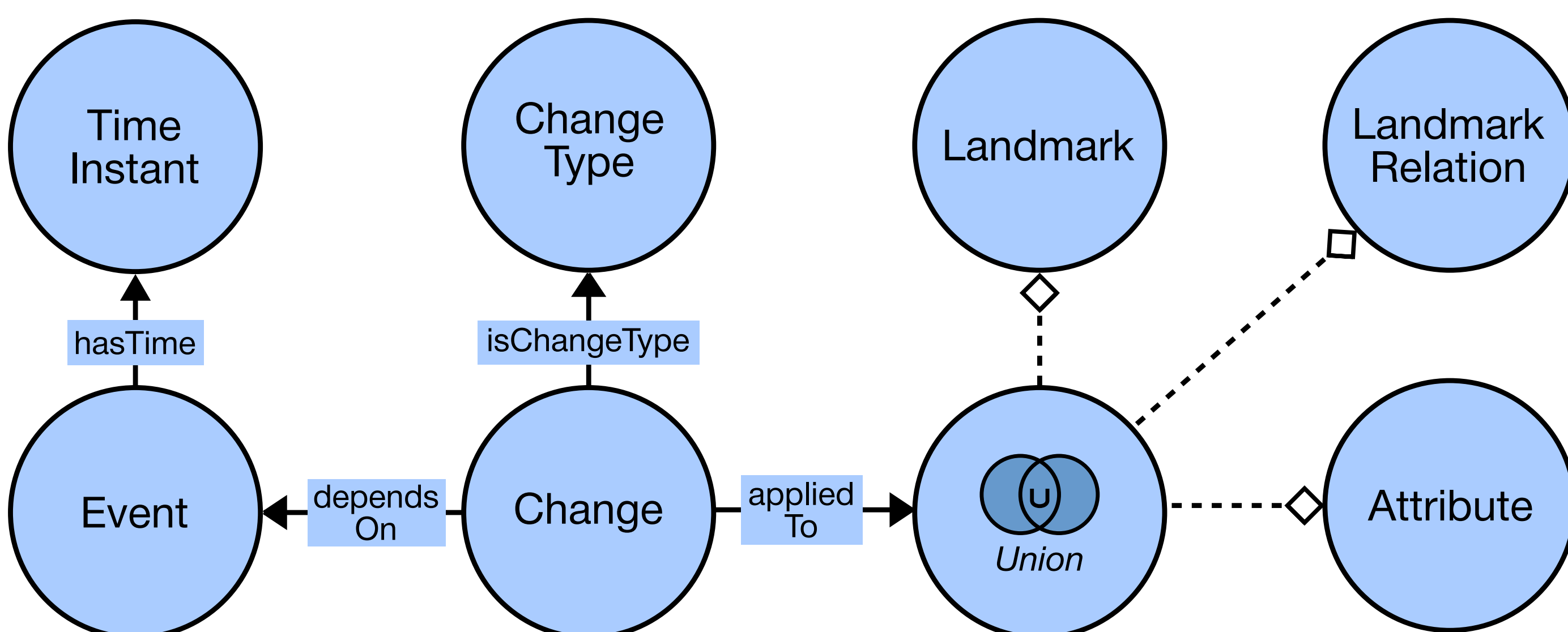
Représentation simplifiée de l'ontologie des adresses

Les repères qui composent les adresses peuvent évoluer au cours du temps. Par exemple, une rue peut être prolongée, ou un point d'intérêt peut changer de dénomination. Il faut pouvoir représenter ces évolutions.



Modélisation des entités géographiques

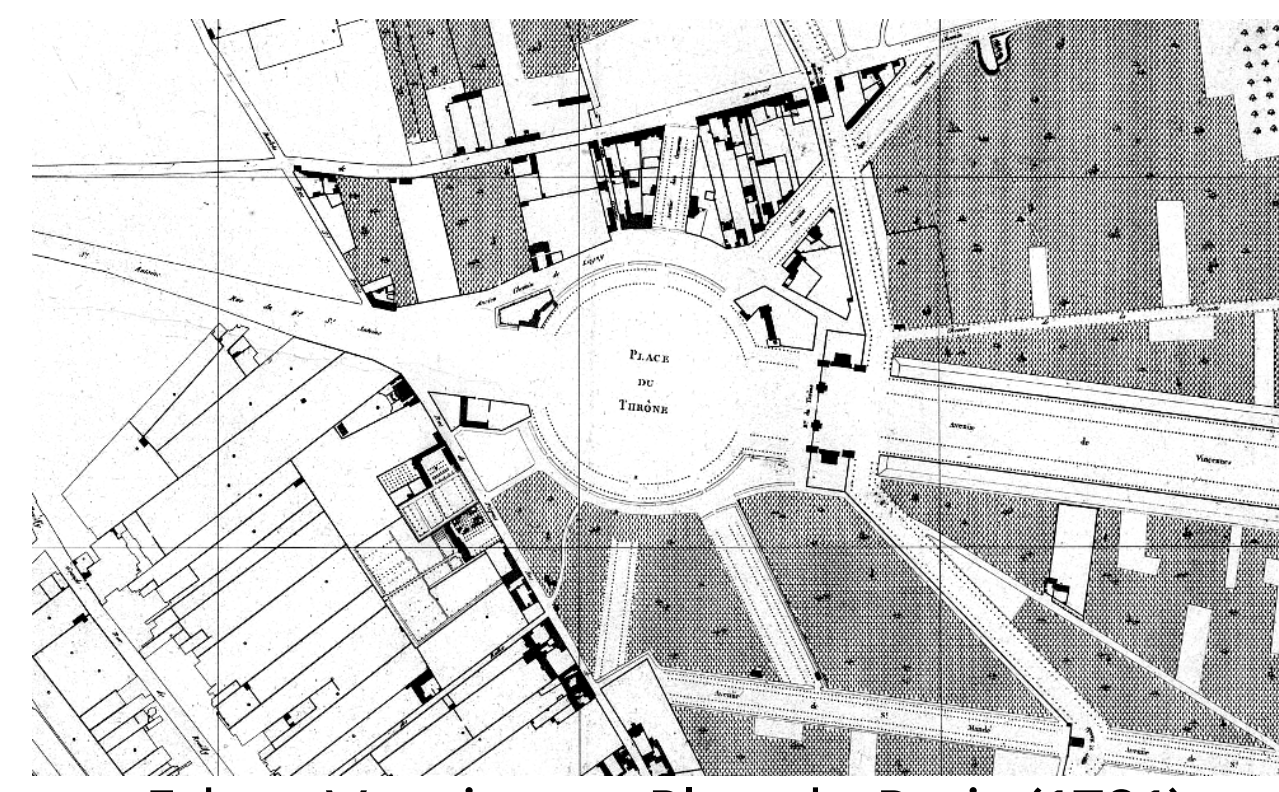
Évolution temporelle : **description des événements (Event)** représentant une évolution du territoire impliquant un ou plusieurs changements (**Change**) témoignant de l'évolution d'une ressource.



Modélisation de l'évolution temporelle

3. Cas d'application

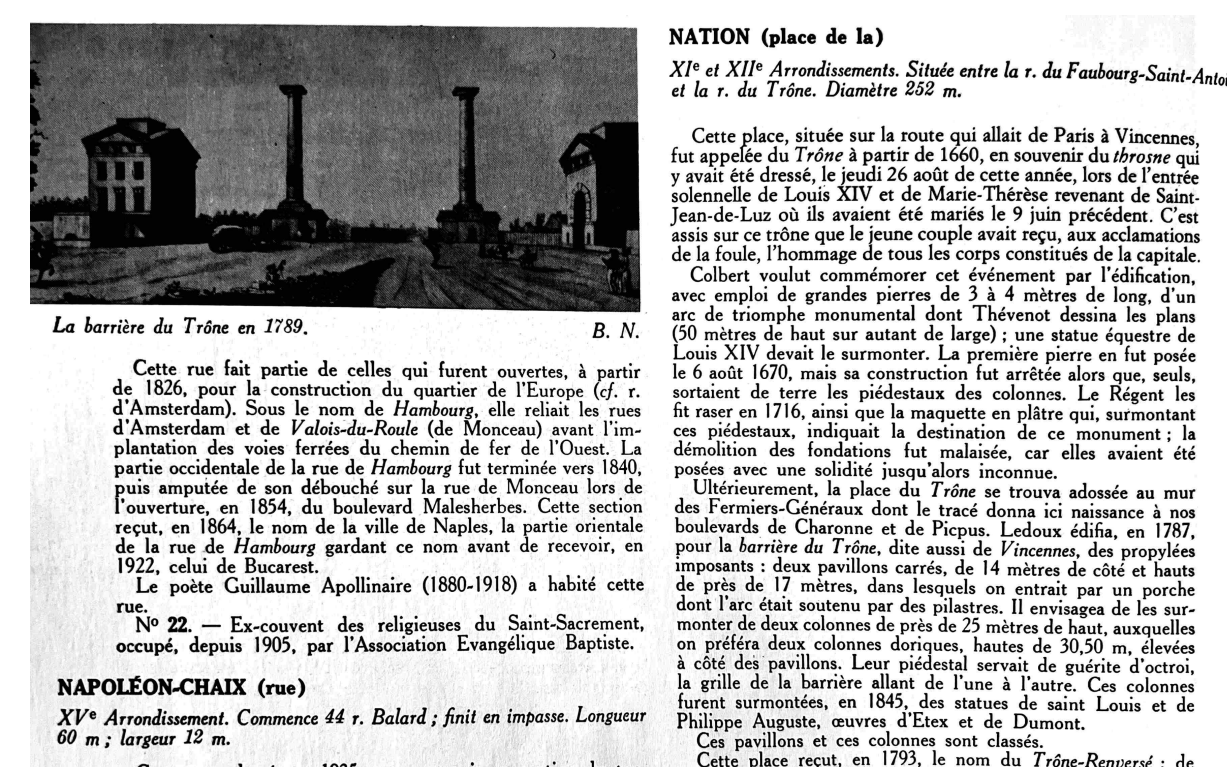
Construction d'un graphe géo-historique d'adresses de Paris de la fin du XVIII^e siècle à nos jours à partir de sources hétérogènes



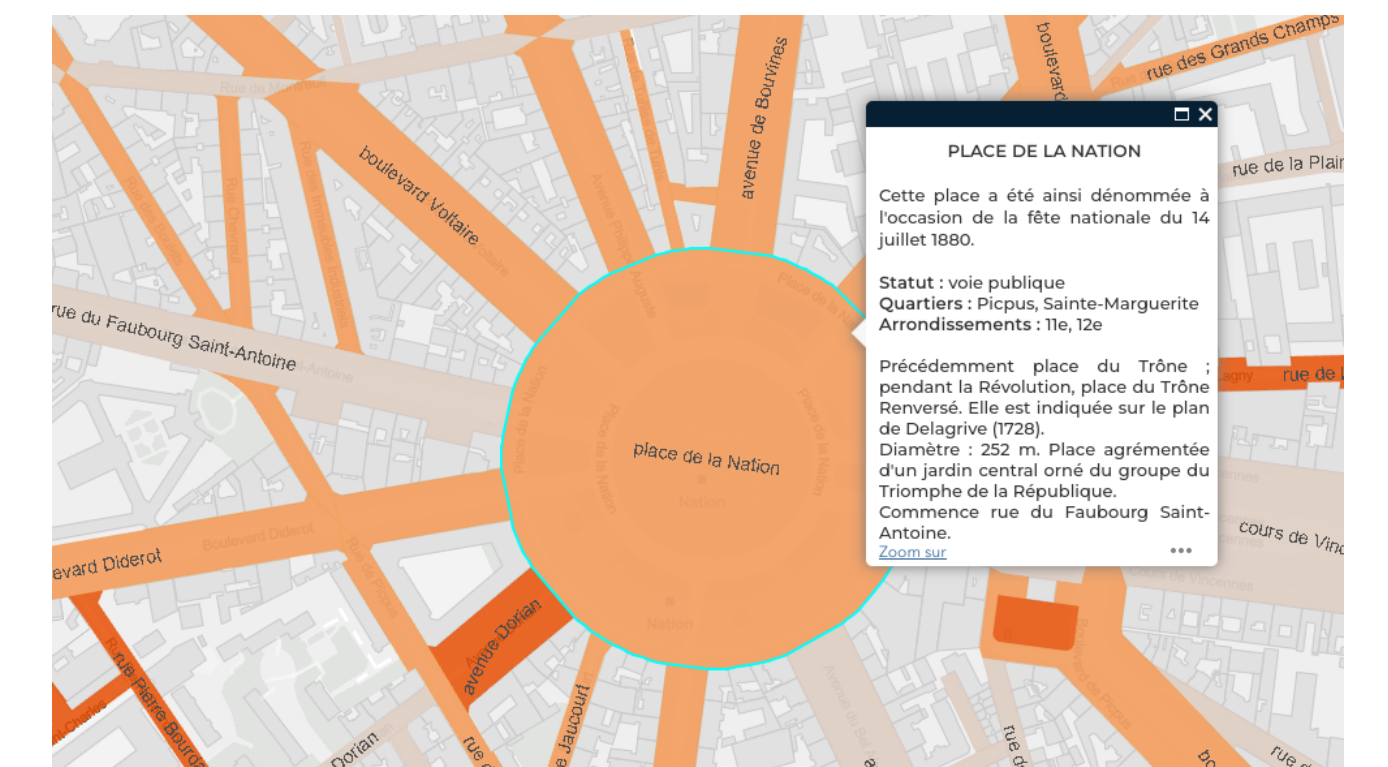
Edme Verniquet, Plan de Paris (1791)

- Bégauli, *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.
- Béglé 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^e siècle)



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



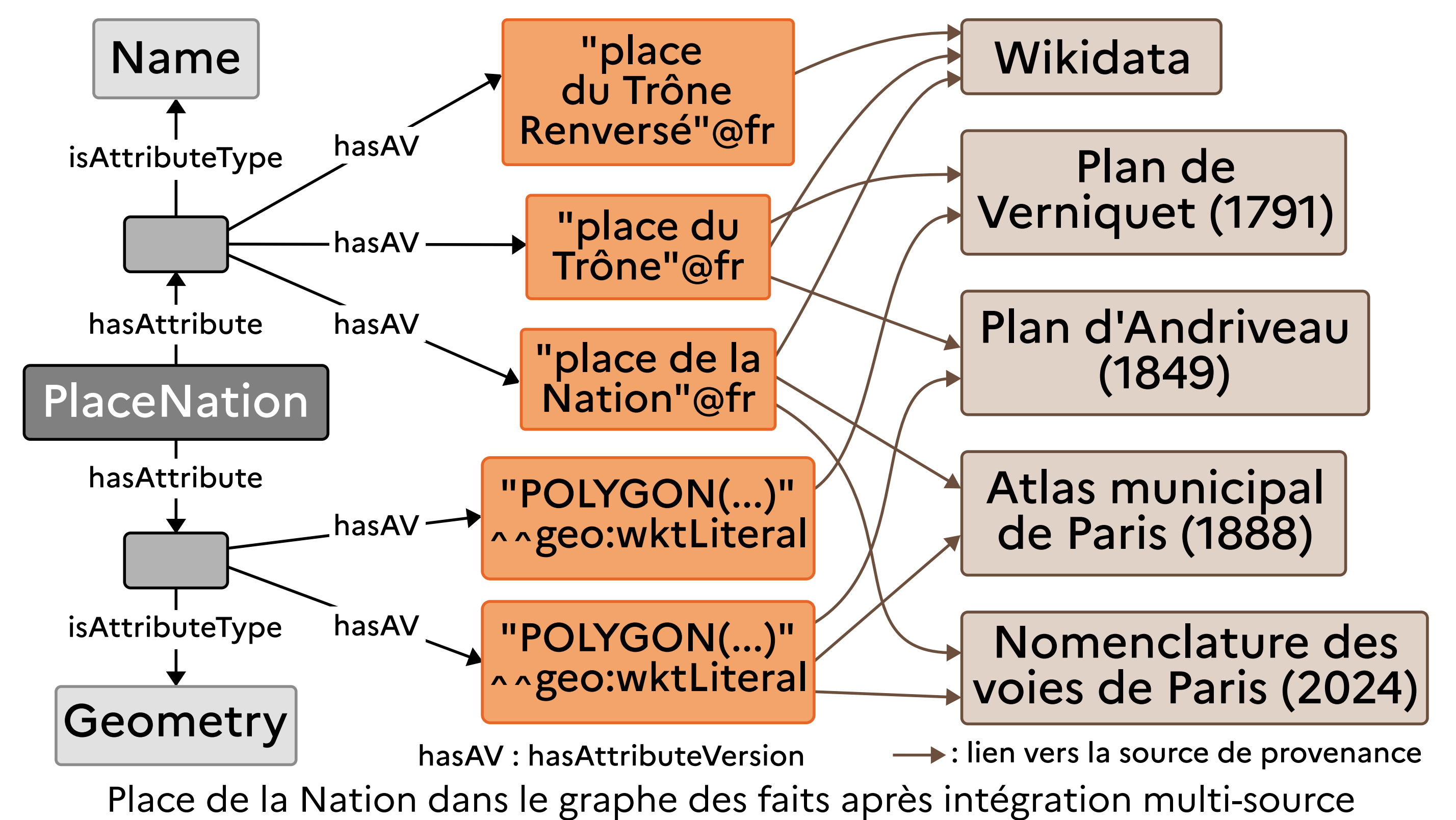
Ville de Paris, Dénominations des emprises des voies actuelles (2024)

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 factoïdes**) [2] ;
- le graphe géo-historique (ou **graphe de faits**) centralise les données provenant des différentes sources ;
- construction itérative : chaque graphe de factoïdes est connecté à celui des faits, les ressources équivalentes sont liées via un lien **owl:sameAs** qui propage les informations des factoïdes vers les faits. Le graphe des factoïdes est ensuite déconnecté.



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

5. Perspectives

- Amélioration de la stratégie d'intégration multi-source
- Évaluation de la qualité du graphe
- Inférence de données grâce aux informations temporelles
- Création de liens de filiation entre repères

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] 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>