

# Enjeux et verrous des graphes de connaissances géohistoriques

**Nathalie ABADIE**  
LASTIG, Univ. Gustave Eiffel, IGN-ENSG

Cinquième symposium du GRD CNRS MaDICS  
Action RoCED  
24 mai 2023, Troyes

CC-BY 4.0

# L'information géographique au coeur du Web de données

*“Geography provides an intuitive way to align datasets — no surprise, as governments generate public sector information about the territory over which they have jurisdiction, so the data has an implicitly geographical dimension. The Linked- Data Web is well-stocked with geographical data; the Geonames service manages eight million URIs for geographical resources. **Therefore, where there is an authoritative geographical knowledge base available, [...], geography is an irresistible join point for datasets.**”*

# Référentiel de données géographiques pour l'IGN

Ensemble de données représentant l'état du territoire:

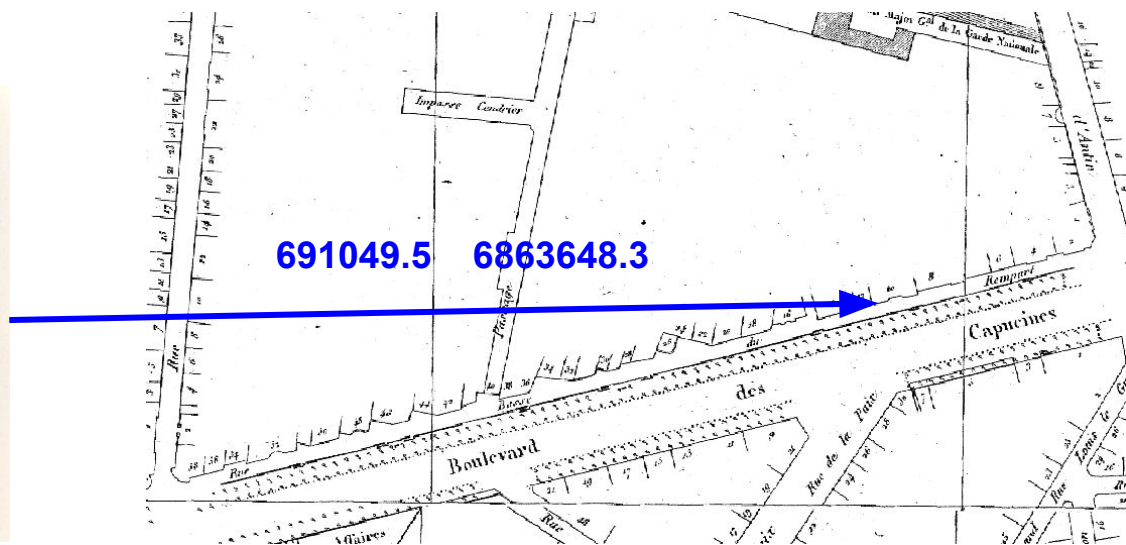
- homogènes,
- exhaustives,
- avec des références spatiales directes exactes, ou du moins, dont la qualité est connue, maîtrisée et documentée,
- avec des niveaux de détail variés (analyses du territoire à différentes échelles),
- à jour,
- disponibles à différentes dates, de façon aussi continue que possible,
- accessibles et interopérables.

# Localiser les ressources du Web, à la surface de la Terre

Hôtel Radix de Sainte-Foix,  
situé 10 Rue Basse-Du-Rempart



Source gallica.bnf.fr / Bibliothèque nationale de France



Par Jean-Baptiste Lallemand — Bibliothèque nationale de France, Domaine public,  
<https://commons.wikimedia.org/w/index.php?curid=13275379>

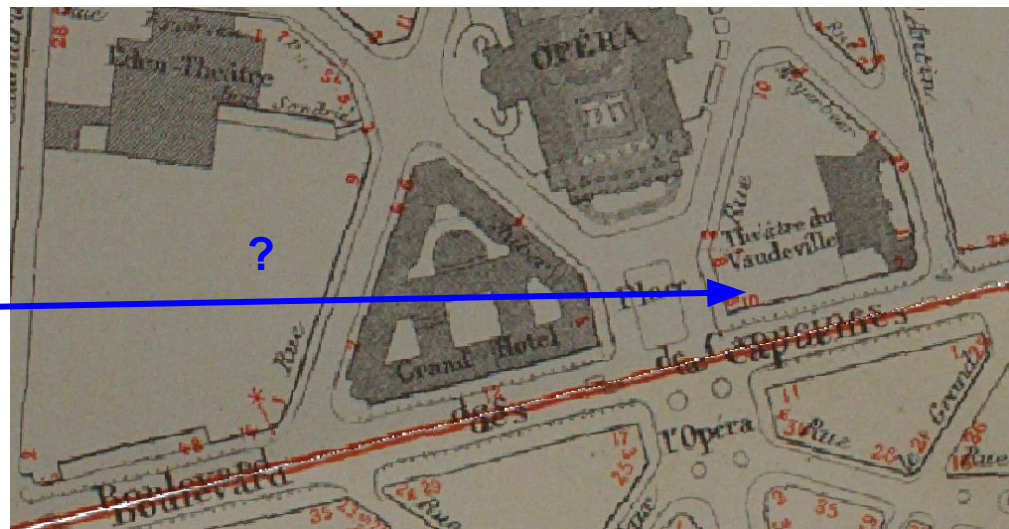
Atlas de Jacoubet géoréférencé,  
<http://geohistoricaldata.org/geoserver/wms>  
Temps valide: 1827 - 1836

# Localiser les ressources du Web, dans le temps

Hôtel Radix de Sainte-Foix,  
situé **10 Rue Basse-Du-Rempart**



Source gallica.bnf.fr / Bibliothèque nationale de France



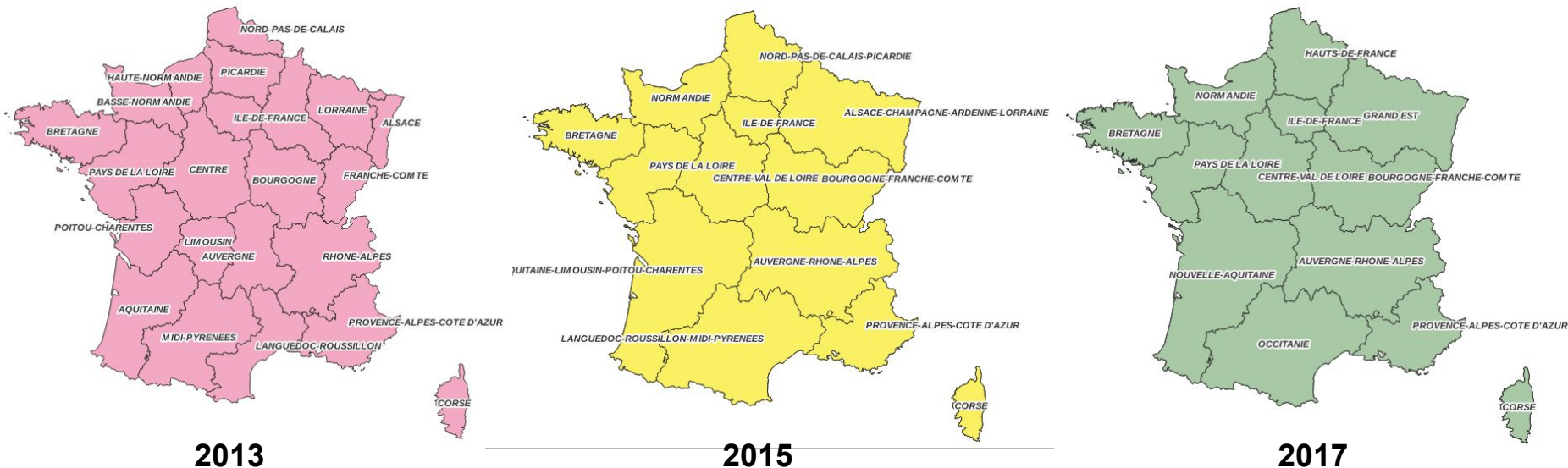
Par Jean-Baptiste Lallemand — Bibliothèque nationale de France, Domaine public,  
<https://commons.wikimedia.org/w/index.php?curid=13275379>

Atlas Municipal géoréférencé,  
<http://geohistoricaldata.org/geoserver/wms>  
**Temps valide : 1888**



# Limites des référentiels de données géographiques actuels

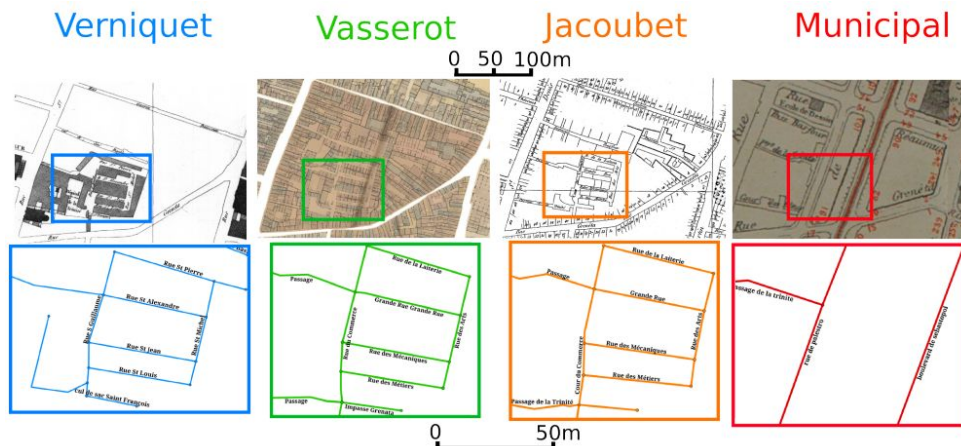
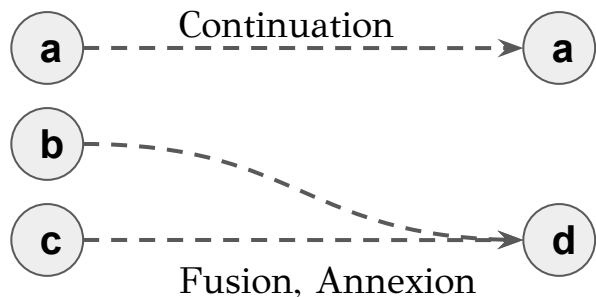
- Profondeur temporelle des référentiels de données géographiques structurées très faible (ex. Pour la BD TOPO IGN ~ début des années 2000).
- La logique de *snapshots* prévaut: pas ou peu de données spatio-temporelles.



# Approches pour la création de graphes spatio-temporels

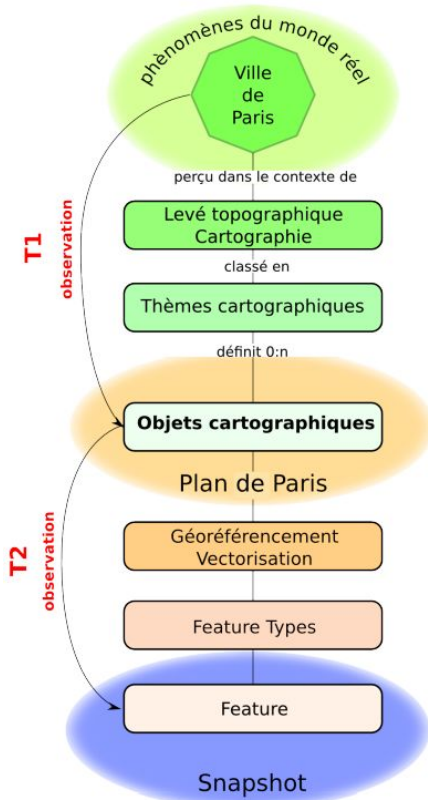
A partir de *snapshots* (1, 2, 3, 4) :

- Détection de relations de filiation entre entités géographiques par liage.
- Interprétation des graphes de filiation en processus spatio-temporels.



États successifs de l'îlot de la Trinité et réseaux de rues vectorisés correspondants - Source : Thèse de Bertrand Duménieu (2)

# Limites des sources de données géographiques anciennes



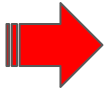
- Sources indirectes sur les états passés du territoire (à l'exception des photographies),
- Sources non structurées (textes ou images),
- Sources non géoréférencées,
- Sources le plus souvent fragmentaires (spatialement ou temporellement).



# Construction de graphes de connaissances géohistoriques: un enjeu d'intégration légère de données anciennes

Ensemble de données représentant l'état du territoire:

- ~~homogènes,~~
- ~~exhaustives,~~
- avec des références spatiales directes ~~exactes, ou du moins,~~ dont la qualité est connue, maîtrisée et documentée,
- avec des niveaux de détail variés (analyses du territoire à différentes échelles),
- ~~à jour~~ **de temps valide connu,**
- disponibles à différentes dates, de façon aussi continue que possible, **avec des liens spatio-temporels entre les représentations des entités géographiques à différentes dates,**
- accessibles et interopérables.



**Mettre à profit l'infrastructure du Web de données pour sortir des logiques de production de référentiels géohistoriques en silos.**

# Des sources géographiques anciennes à l'analyse géohistorique

Reconnaissance

Image → Géométries



Scan → Texte

Ravrio et comp., fabr. de bronzes et curiosités, r. Richelieu, 93; la fabrique rue Montmartre, 161.

Classification

Géométries → Annotations



Texte → Entités nommées spatiales

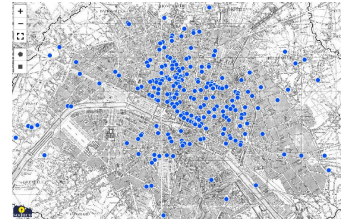
Ravrio et comp. PESE , fabr. de bronzes et curiosités ACT  
r. Richelieu LOC , 93 CARDINAL ; la fabrique FT  
rue Montmartre LOC , 161 CARDINAL

Géo-référencement

Données géométriques → Données géographiques

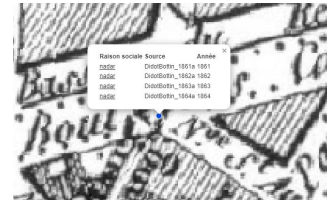
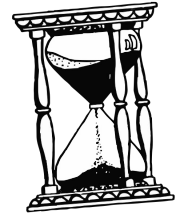


Entités nommées spatiales → Données géographiques → Données géohistoriques

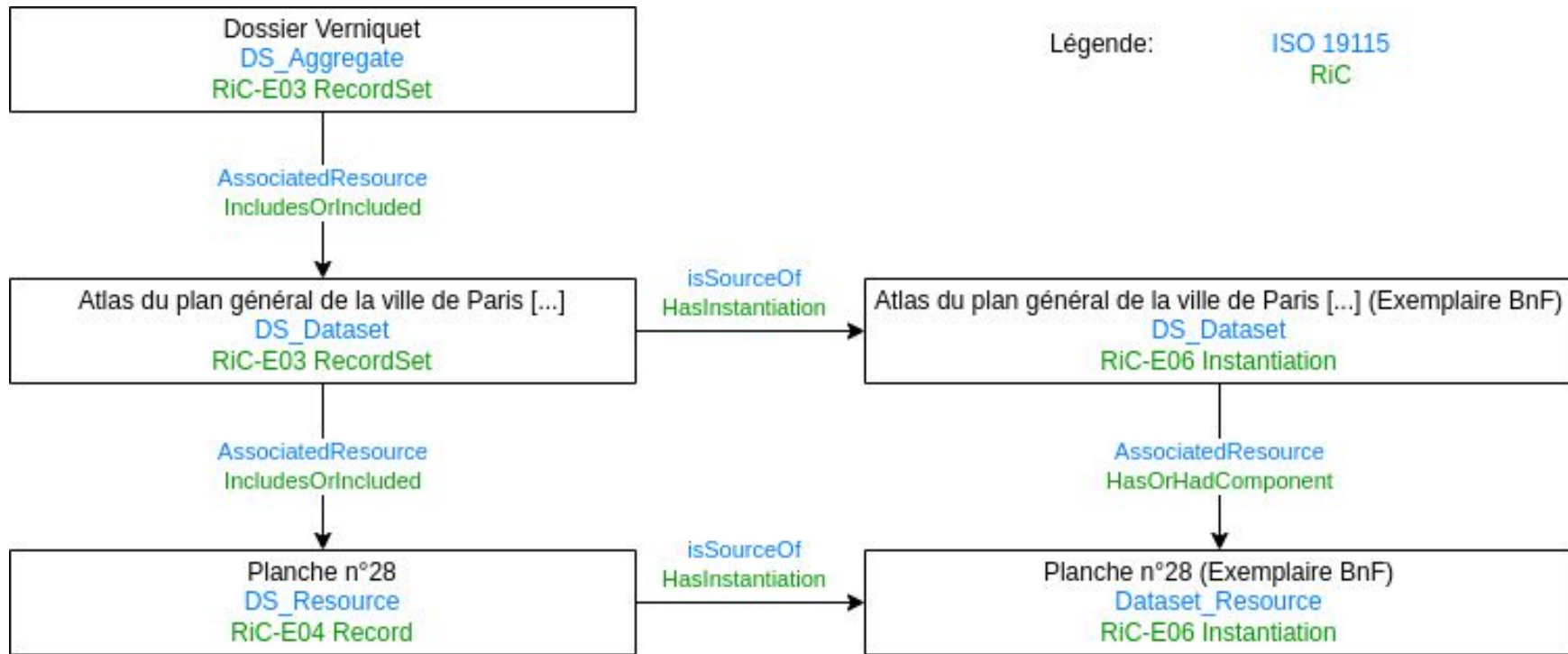


Structuration

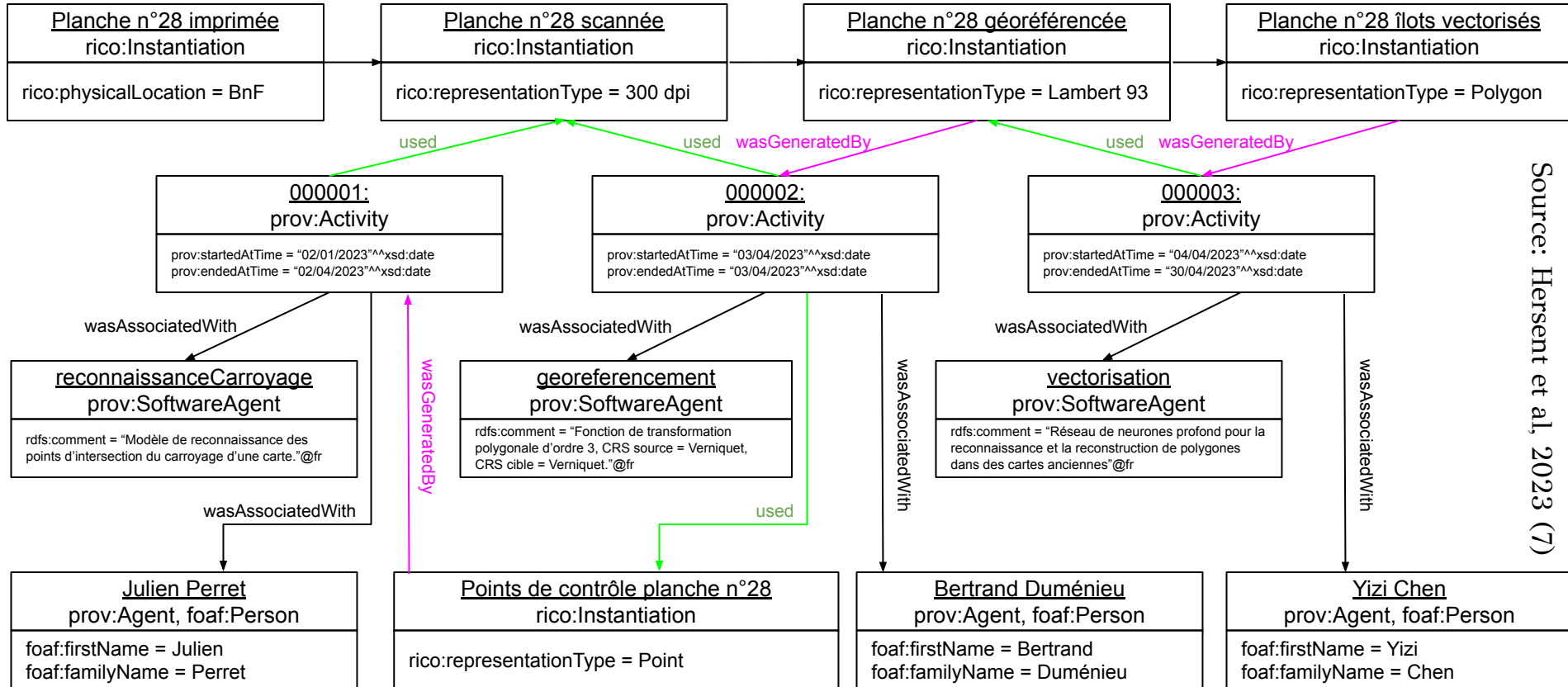
Données géographiques → Données géohistoriques



# Documenter les sources de données géographiques anciennes



# Documenter les références spatiales extraites, les ressources intermédiaires, les processus d'extraction...



# Extraire des références spatiales à partir de sources anciennes

## Sources cartographiques :

- Les représentations cartographiques varient énormément et les modèles de vectorisation automatique peinent à généraliser.



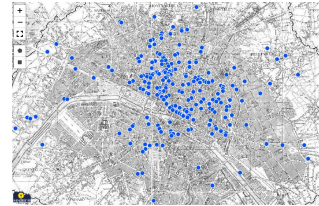
Thèse de Yizi Chen (2023, soutenue) Vectorization of historical maps using deep edge filtering.

<https://www.theses.fr/s234570>



## Sources textuelles :

- Détection de mise en page et HTR/OCR,
- NER pour reconnaître les lieux nommés,
- NEL pour les désambiguïser/géocoder,
- Reconnaissance de relations spatiales,
- Couplage de modèles.



GEOCODING

Ravrio et comp., *fabr. de bronzes et curiosités*, r. Richelieu, 93; la fabrique rue Montmartre, 161.

OCR

Ravrio et comp. INH . fabr. de bronzes et cu- riosités ACT .  
r. Richelieu LOC . 93 CARDINAL : la fabrique FT  
rue Montmartre LOC . 161 CARDINAL .

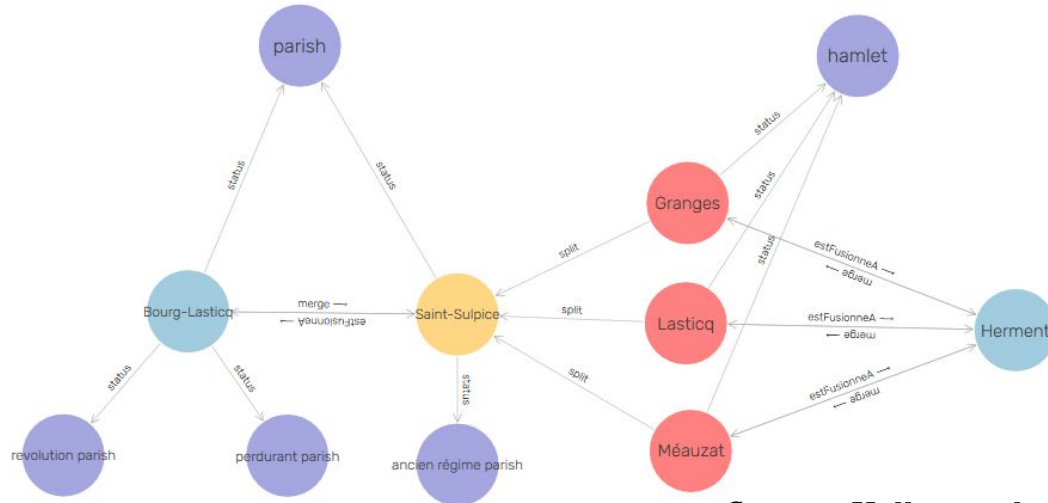
Ravrio et comp., fabr. de bronzes et curiosités, r. Richelieu, 93; la fabrique rue Montmartre, 161.

NER



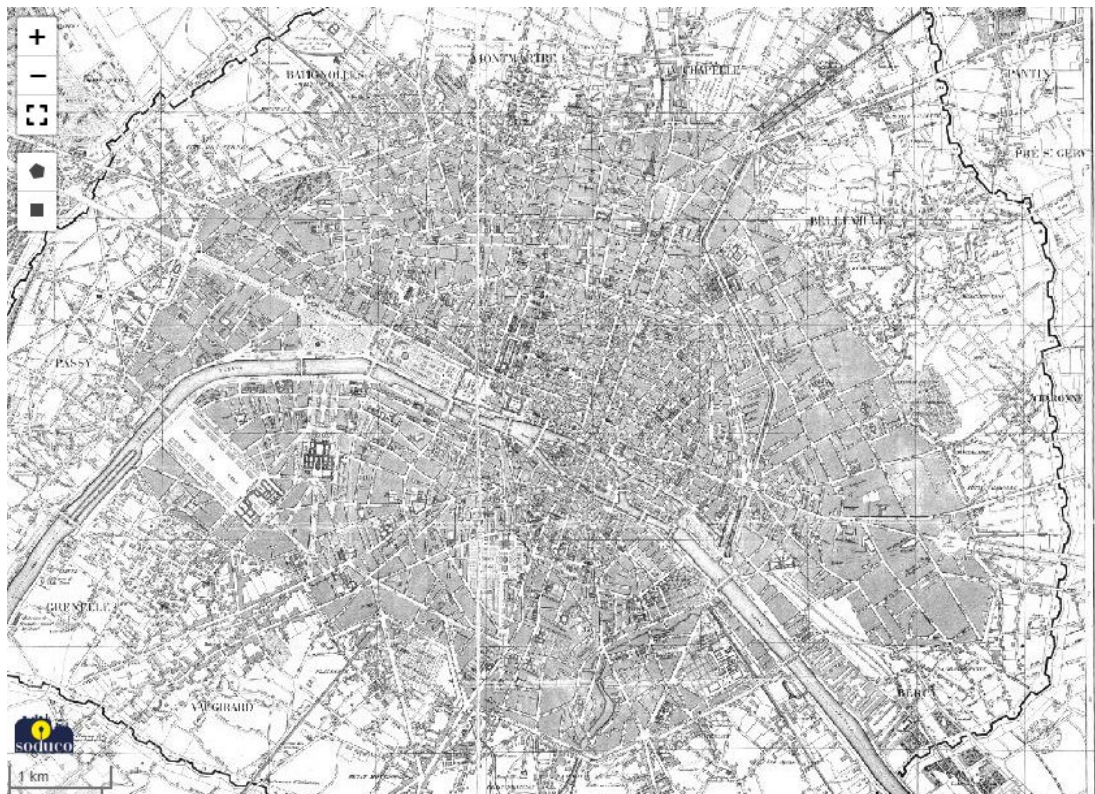
# Inférences complexes

- Inférer une localisation (même approximative) à partir de relations spatiales par rapport à un lieu de localisation connue
- Inférer des éléments de qualité des données à partir des métadonnées
- Inférer des événements/processus spatio-temporels à partir d'états (et réciproquement)



Source: Keller et al, 2018 (6)

# Un graphe de connaissances géohistoriques sur les photographes parisiens du 19ème siècle - Tual et al., 2023 (5)



Entrées extraites des annuaires du commerce de Paris (XIXème siècle) relatives aux photographes et aux professions associées

Filtres

Propriétés

Raison sociale

Ex : nadar

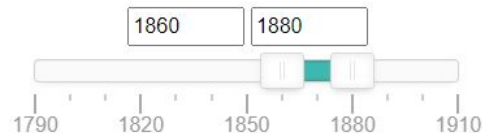
Description

Ex : photo

Adresse

Ex : rivoli

Période



? Le filtre temporel permet de faire varier l'affichage des points préalablement chargés sur la carte sans lancer une nouvelle recherche.

# Rejoignez l'action de recherche “Graphes de connaissances géohistoriques” du GdR CNRS MAGIS !

Site Web: <https://gcgh-magis.github.io/>

Liste de diffusion: [ap-gcgh-magis@ehess.fr](mailto:ap-gcgh-magis@ehess.fr)

Pour s'inscrire, écrire à [sympa@ehess.fr](mailto:sympa@ehess.fr)

Objet du mail : *subscribe ap-gcgh-magis Prénom Nom*



# Références

- 1) Del Mondo Géraldine. Un modèle de graphe spatio-temporel pour représenter l'évolution d'entités géographiques. Thèse de doctorat, Université de Brest, 2011.
- 2) Duménieu Bertrand. Un système d'information géographique pour le suivi d'objets historiques urbains à travers l'espace et le temps. Thèse de doctorat, Ecole des Hautes Etudes en Sciences Sociales, 2015.
- 3) Aurélie Leborgne, Adrien Meyer, Henri Giraud, Florence Le Ber, and Stella Marc-Zwecker. Un graphe spatio-temporel pour modéliser l'évolution de parcelles agricoles. In *Conférence internationale francophone en analyse spatiale et géomatique SAGEO*, 2019.
- 4) Camille Bernard, Marlène Villanova-Oliver, and Jérôme Gensel. Theseus : A framework for managing knowledge graphs about geographical divisions and their evolution. *Transactions in GIS*, 2022.
- 5) Solenn Tual, Nathalie Abadie, Bertrand Duménieu, Joseph Chazalon and Edwin Carlinet. Création d'un graphe de connaissances géohistorique à partir d'annuaires du commerce parisien du 19ème siècle: application aux métiers de la photographie. *IC 2023, 34èmes journées francophones d'Ingénierie des connaissances*, Strasbourg, France, 3-7 Juillet 2023.
- 6) Antoine Keller, Nathalie Abadie, Bertrand Dumenieu, Stéphane Baciocchi, Eric Kergosien. Vers la construction d'une base de connaissances sur la réorganisation territoriale française à la Révolution. Conférence Sagéo 2018 Atelier Exces, 2018, Montpellier, France
- 7) Melvin Hersent, Nathalie Abadie, Bertrand Duménieu, Julien Perret. Modèles et outils pour la publication de métadonnées d'archives géographiques et de leurs données dérivées. *Humanistica 2023*, Association francophone des humanités numériques, Juin 2023, Genève, Suisse.