

PeGazUs

A knowledge graph based approach to build urban perpetual gazetteers

Charly Bernard¹, Solenn Tual¹

Nathalie Abadie¹, Bertrand Duménieu², Julien Perret¹, Joseph Chazalon³

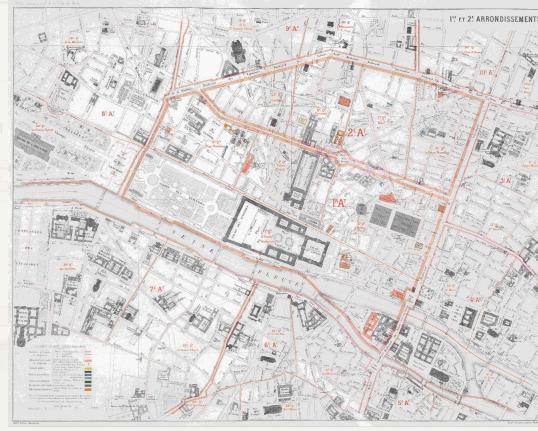
¹ LASTIG, Université Gustave Eiffel, IGN-ENSG, Saint-Mandé, France

² CRH, EHESS, Paris, France

³ LRE, EPITA, Le Kremlin-Bicêtre, France

Context

- Growing volume of digitized archival records available online
 - ⇒ significant amount of fragmentary geohistorical information
- Sources need to be indexed both spatially and temporally to enhance information retrieval.
 - ⇒ very useful for research in humanities and for GLAM



Administrative atlas of the districts of Paris, 1868

Non-Commerçans. (Paris):		269		
Chardin, R. Pavé, 16. — R. C.	Charviro, R. Chapon, 13.			
Chardin, R. Michel Lepelletier, 21.	Chimay, (Mme.) R. de Varennes, 31.			
Chardon, (Ve.) R. S. Marc, 15.	Choirat-Duplessis, R. de Turenne, 31.			
AMADOU ET ALLUMETTES. — Pour les ALLUMETTES OXYGÉNÉES. Voyez BRICKETS PHYSIQUES.				
DARRAS (Thomas), r. de la Vieille-Monnaie, 10.	GALLIENNE Je ^r , r. de la Héumerie, 3.			
Briques et veillennes, mèches à quinquettes, mèches souffrées, pierres, agatice de chêne, liège en planches, bouchons.	Brûle-tout, boîtes à briquet, mèches à vin et à quinquette, veillennes méches, souffrées ; pierres, agatice, bouchons, liège.			
	LEROY, r. Aubry-le-Boucher, 43.			
BAUDOUIN (dans). — Rue d'Orléans, pourtour St-Gervais, Saint-Antoine et Renaud-Lefèvre.				
1 Lissoty (Vve), vins, 2 ^e Privé, distillateur, Lamouroux et Leroy, nouveautés. Chantier, court-gourn.	2 ^e Longuyard, bijoutier en or et argent. Saint-Omer, émaillerie, Couture (A.), graveur-é- tamineur. Bouscassé (J.), bijoutier en or. Bouvet, orfèvre-faï. Lébèzy, dorure. Bouton, fab. de cuir ver- si pardon, vins.	3 ^e Bourguignon, fab., dr. Vauvain, pasticcerie. Perrin, gâteaux dans Rabé sind, fab. de bâ- tons. Gaudin, charpenter. Moisy, tablier. Cendrier ainsi prop. d'embalage, fab. botte- Ferrand, faridistre. 1 ^{er} Baduel (Vve), fab. de	Ecole communale de jeu- nes filles. Berthelot, vins. Verneuil, serviteur-ind- ianicien. Michel, brouster. Baudouin, servurier. Sacrez, vins. Baudoin, épicer. Leroy, chouettes et cr- pin. Baduel (Vve), fab. de	et tapissières. Laine jeune, vins. Jumelles omniales et en- cadrées, servante générale des Omnibus. Méhulay, vins en gros, et à la bouteille. Counaud, coiffure. Monmain (P.), vins en gros. Dufaillly, sculpt. fab. de carton-pierre.

Extracts of 19th century Paris trade directories 2

Motivations

Indexation requires to have **perpetual gazetteers** that describe landmarks located at **any coordinates** at **any date** including:

- different names
- properties
- sources that assert their existence
- evolution

Challenge

 **Build this perpetual gazetteer**

Representation and **integration of fragmentary data**
extracted from several sources (diverse and heterogeneous)

Challenge

→ Build this perpetual gazetteer

Representation and **integration of fragmentary data**
extracted from several sources (diverse and heterogeneous)

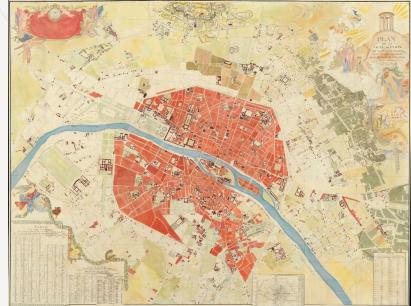
→ Knowledge graphs are a suitable solution!

Research questions

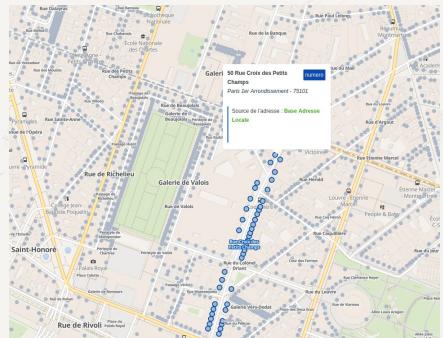
1. Which **ontologies** are suitable to build perpetual gazetteers?
 - structure named places that are useful to localize resources
 - retrieve all named places that refer to the location at any date
 - preserve the link between data and sources
2. How to **populate an ontology** of historical landmarks?
 - get the evolution of their attributes
 - get the evolution of their relations with other features

Data

Sources describe addresses and geographical entities over time



Verniquet, Plan de Paris (1791)



Base Adresse Nationale (2023)

rue Louis-Philippe (Q3448913)

élément Discussion

voie de Paris

[modifier](#)

+ Plus de langues

Configurer

Langue	Libellé	Description	Également connu comme
français	rue Louis-Philippe	voie de Paris	

Toutes les langues saisies

Déclarations

nature de l'élément

	rue	modifier
	+ 0 référence	
	+ ajouter une référence	

	anciennes voies de Paris	modifier
	+ 0 référence	
	+ ajouter une référence	

partie de

	anciennes voies de Paris	modifier
	+ 0 référence	
	+ ajouter une référence	

pays

	France	modifier
	+ 0 référence	
	+ ajouter une référence	

localisation administrative

	Paris	modifier
	+ 0 référence	
	+ ajouter une référence	

	Paris	modifier
	+ 0 référence	
	+ ajouter une référence	

	Paris	modifier
	+ 0 référence	
	+ ajouter une référence	

Wikidata page of Rue Louis-Philippe

Non-Commerçans. (Paris):	
Chardin, B. Pavé,	16. — B. C. Cheviron, R. Chapow, 13.
Chardin, R. Michel Lepelletier,	21. Chimay, (Mme.) R. de Varennes, 31.
Chardin, , (Ve.) R. S. Marc,	13. ChorDuplessis, R. de Turenne, 31.
AMADOU ET ALLUMETTES. — Pour les ALLUMETTES OXYGÉNÉES. Voyez Briquets physiques.	
DARRAS (Thomas), r. de la Vieille-Monnaie, 10.	GALLIENNE jr., r. de la Hausserie, 3.
Briquets et veillennes, mèches à quinquetes, à quinquet, veillenes mèches, souffrées ; mèches souffrées, pierres, agaric, bouchons, liège en planches, bouchon.	Brûle-tout, boîtes à briquet, mèches à vin et
	Lorrain, r. de la Hausserie, 43.
BAUDOYER (place).	
IX Arr., Hôtel-de-Ville.) + St-Gervais, Saint-Antoine et Renaud-Levire.	IX Longpré sind, bijouterie, perruques, verroterie, Vautain, passementier, Saint-Omer, émailleur, Geller (A.), graver-ciseleur, Baudoyer, fabr. de bouillons.
1 Lisonny (Vev), vins, 2 ^e étage, distillateur, Léonard Jusel et Leroy, nonveuvels. Chantrel, court-gourme.	1 Lisonny, orfèvre-fabr., Moisy, tablier, Lévy, dorur, 90 Drouot, fab. de cuir vernis. 5 ^e Gentefier sind, prop. Drouot, orfèvre-fabr., d'embellissement, Pardon, vins.
	10 Bourgogne sind, bijouterie, verroterie, Berthelot, vins. 6 Versain, serrurier-mécanicien, fabr. de bouillons.
	8 Michel, broisseur, 11 Melouay, vins et pros, et la Baudouin, 12 Baudouin, coiffer, 13 Monmain (P.), vins en gobelets, 14 Dufault, sculpt, fabr. de carton-pierre.
Non-Commerçans. (Paris):	
Chardin, B. Pavé,	16. — B. C. Cheviron, R. Chapow, 13.
Chardin, R. Michel Lepelletier,	21. Chimay, (Mme.) R. de Varennes, 31.
Chardin, , (Ve.) R. S. Marc,	13. ChorDuplessis, R. de Turenne, 31.
AMADOU ET ALLUMETTES. — Pour les ALLUMETTES OXYGÉNÉES. Voyez Briquets physiques.	
DARRAS (Thomas), r. de la Vieille-Monnaie, 10.	GALLIENNE jr., r. de la Hausserie, 3.
Briquets et veillennes, mèches à quinquetes, à quinquet, veillenes mèches, souffrées ; mèches souffrées, pierres, agaric, bouchons, liège en planches, bouchon.	Brûle-tout, boîtes à briquet, mèches à vin et
	Lorrain, r. de la Hausserie, 43.
BAUDOYER (place).	
IX Arr., Hôtel-de-Ville.) + St-Gervais, Saint-Antoine et Renaud-Levire.	IX Longpré sind, bijouterie, perruques, verroterie, Vautain, passementier, Saint-Omer, émailleur, Geller (A.), graver-ciseleur, Baudoyer, fabr. de bouillons.
1 Lisonny (Vev), vins, 2 ^e étage, distillateur, Léonard Jusel et Leroy, nonveuvels. Chantrel, court-gourme.	1 Lisonny, orfèvre-fabr., Moisy, tablier, Lévy, dorur, 90 Drouot, fab. de cuir vernis. 5 ^e Gentefier sind, prop. Drouot, orfèvre-fabr., d'embellissement, Pardon, vins.
	10 Bourgogne sind, bijouterie, verroterie, Berthelot, vins. 6 Versain, serrurier-mécanicien, fabr. de bouillons.
	8 Michel, broisseur, 11 Melouay, vins et pros, et la Baudouin, 12 Baudouin, coiffer, 13 Monmain (P.), vins en gobelets, 14 Dufault, sculpt, fabr. de carton-pierre.

Directories with addresses of Paris during XIXth century

346

TOPOGRAPHIE HISTORIQUE DU VIEUX PARIS.

situation, à la rue de ce nom et à celle de la Montagne-Sainte-Geneviève, qui seront décrites dans le volume suivant : RÉGION ORIENTALE DE L'UNIVERSITÉ.

CENSIVE DE L'ÉVÉCHÉ.

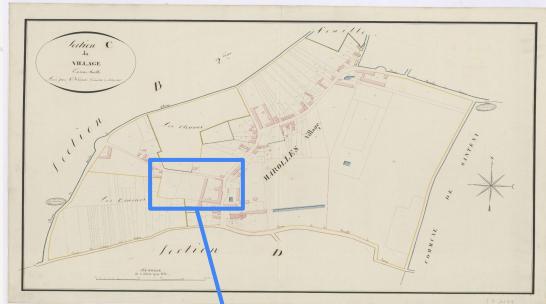
MAISON DE LA COURONNE (1575), faisant le coin occidental de la rue des Carmes. Elle est dite, en 1322, la MAISON DE LA GRANT COURONE et DE LA PETITE COURONE, « au coin de la ruë des Noërs et de la ruë Saint Hilaire »; et en 1553, « maison devant les Carmes, faisant angle de la rue Saint-Hilaire », laquelle maison fut depuis divisée en trois corps de logis : LA COURONNE, L'IMAGE SAINT MARTIN et LA BOUILLIE⁽¹⁾.

Adolphe Berty, Topographie historique du vieux Paris (1866-1897)

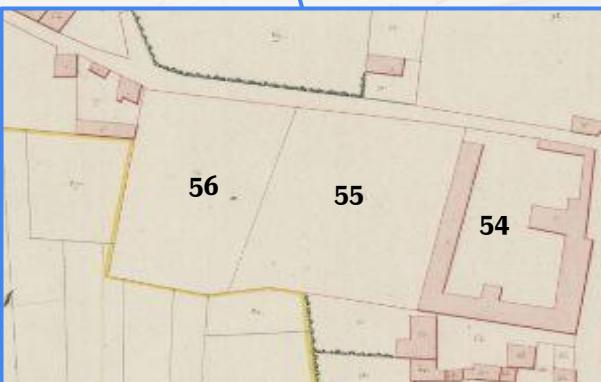
Data

Archival records: the Napoleonic Land Registry (1808-1950)

Cadastral index map (1810)



unique graphic representation of plots



legend of

handwritten
spatial temporal
referential of
plots

Initial register (1812)

CANTON, TRIAGES, ou Lieux dits.	NUMÉROS DU PLAN provisoire définitifs	NOMS, PROFESSIONS, DÉMEURES des Propriétaires ou Usufructuaires.	NATURES des Propriétés.	CONTENANCE par Propriété.	CONTENANCE par nature de culture.
+ - 54.	Berthier, alex. M. Prince de Montrichard	Maison et Cour	arp. p. m.	30 15	
+ - 55.	Jean	Terrain	arp. p. m.	37 00	37 00
+ - 56.	Jean	Vigne	arp. p. m.	29 00	29 00

Description of the plots C-54, C-55 and C-56 at the creation of the land registry

Change registers (here: 1812-1822)

ARTICLES,	NOMS, PRÉNOMS, PROFESSIONS ET DÉMEURES des Propriétaires et Usufructuaires.	INDICATION				CLASSES	REVENU	RENVOIS pour les Mutations, tiré de l'article, porté à l'article.	
		de la section.	du N° de la section.	du CANTON, triage ou lieu dit.	DE LA NATURE de la propriété.				
3	Berthier, alexandré Lima Le Raychard a Gros Bois	B	90	Beffard	5000	arp. p. m.	arp. p. m.	fr. c.	fr. e.
					2 6 ff			98 92	
					o 11 n			1 13 12	
					o 30 16			1 19 60	
					o 37 60			2 20 70	
					o 29 00			1 26 67	

Versions of the plots C-54, C-55 et C-56 at a given date (snapshot)

State of the art

Existing approaches

- Kauppinen et al. [1]: Finnish municipalities (since 1865)
 - Territorial Statistical Nomenclature (TSN) [2]: administrative and statistical units
 - Hierarchical Historical Territory (HHT) [3]: territorial subdivisions of France before 1789
- ⇒ **Change-Bridge** concept [1]: describes changes between landmark versions
- ⇒ Treat only landmarks versions **which attributes never evolve**
- ⇒ Use mostly data:
- **structured**
 - **normalized**
 - **aggregated and interpreted**

[1] Kauppinen et al. (2008). Creating and Using Geospatial Ontology Time Series in a Semantic Cultural Heritage Portal

[2] Bernard et al. (2018) Modeling changes in territorial partitions over time: Ontologies tsn and tsn-change

[3] Charles et al. (2023) : HHT: an approach for representing temporally-evolving historical territories

State of the art

Existing approaches

- Kauppinen et al. [1]: Finnish municipalities (since 1865)
- Territorial Statistical Nomenclature (TSN) [2]: administrative and statistical units
- Hierarchical Historical Territory (HHT) [3]: territorial subdivisions of France before 1789

⇒ **Change-Bridge** concept [1]: describes changes between landmark versions

⇒ Treat only landmarks versions **which attributes never evolve**

⇒ Use mostly data:

- structured
- normalized
- aggregated and interpreted

[1] Kauppinen et al. (2008). Creating and Using Geospatial Ontology Time Series in a Semantic Cultural Heritage Portal

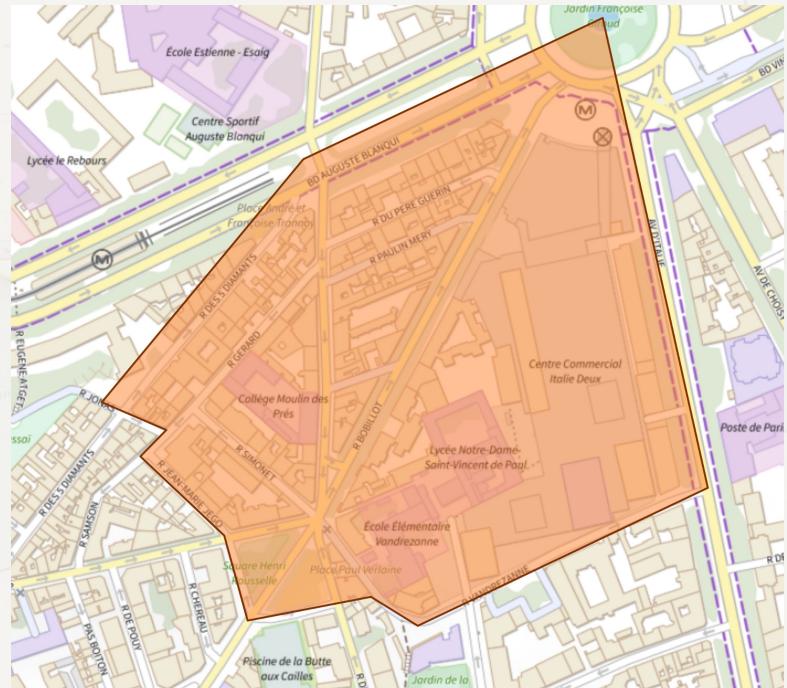
[2] Bernard et al. (2018) Modeling changes in territorial partitions over time: Ontologies tsn and tsn-change

[3] Charles et al. (2023) : HHT: an approach for representing temporally-evolving historical territories

Contributions

PeGaZus : Perpetual Gazetteer of approach-address UtteranceS

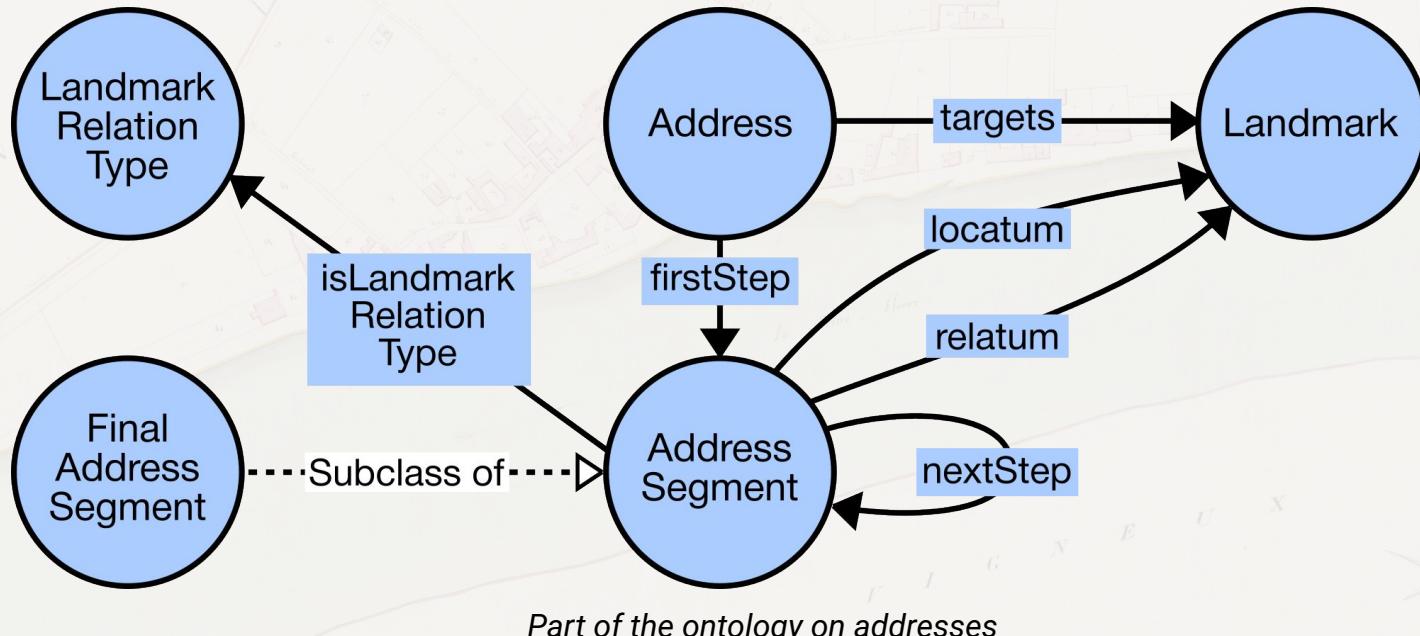
- **Ontology** which represents:
 - addresses and landmarks,
 - with **attributes that evolve** over time
- **Population method** retrieves evolutions of addresses, plots and of their attributes
- Applied on a whole neighborhood in Paris (France) since the 1800s
- Map-based evaluation



Delimiting the Butte aux Cailles neighborhood in Paris, France

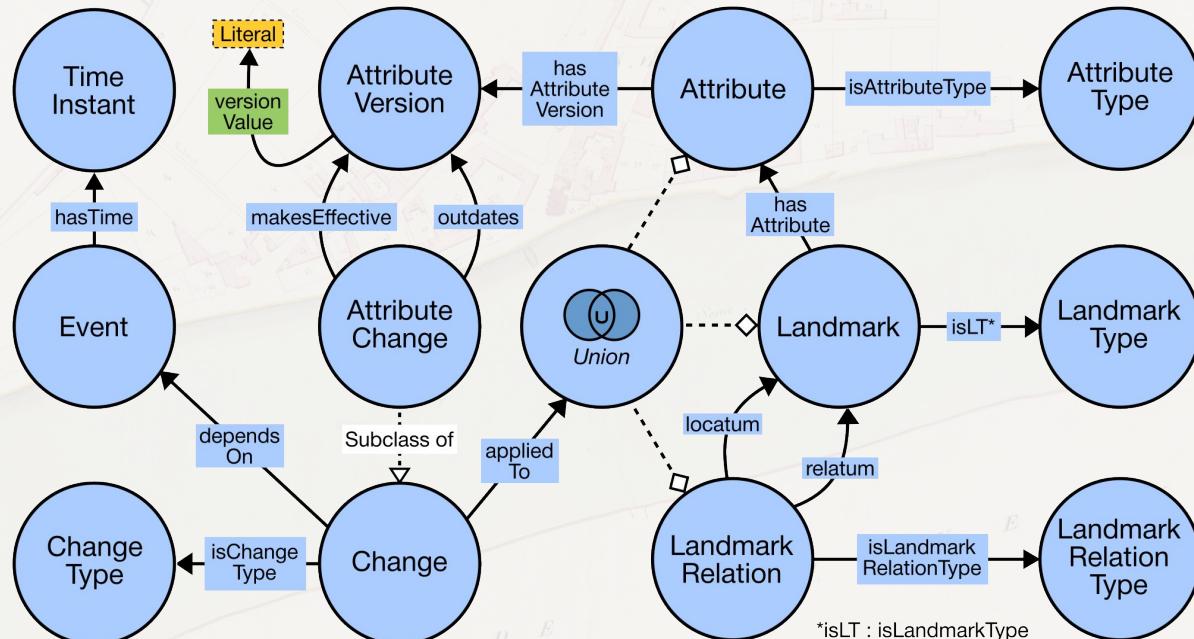
Ontology

Address: statement which targets a place described as an ordered sequence of landmark relations [4]



Ontology

Evolution: representation of events. An event is related to an evolution on the territory, may imply multiple changes on data.



Part of the ontology on temporal evolution

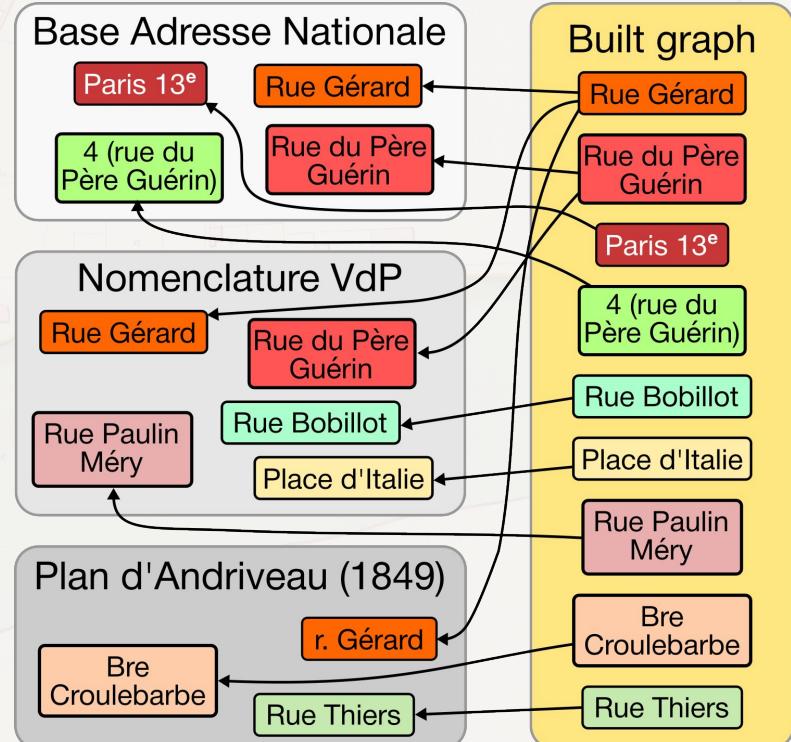
Ontology

Additional modelets:

- **sources**
- **taxpayers**
- **land registry documentation**

Population

1. Link similar landmark assertions from sources with a **root landmark**: use (pseudo-)identity criterion to link them
2. Order landmark assertions and their attributes versions
3. Infer changes and events related to landmarks (Appearance, Disappearance)

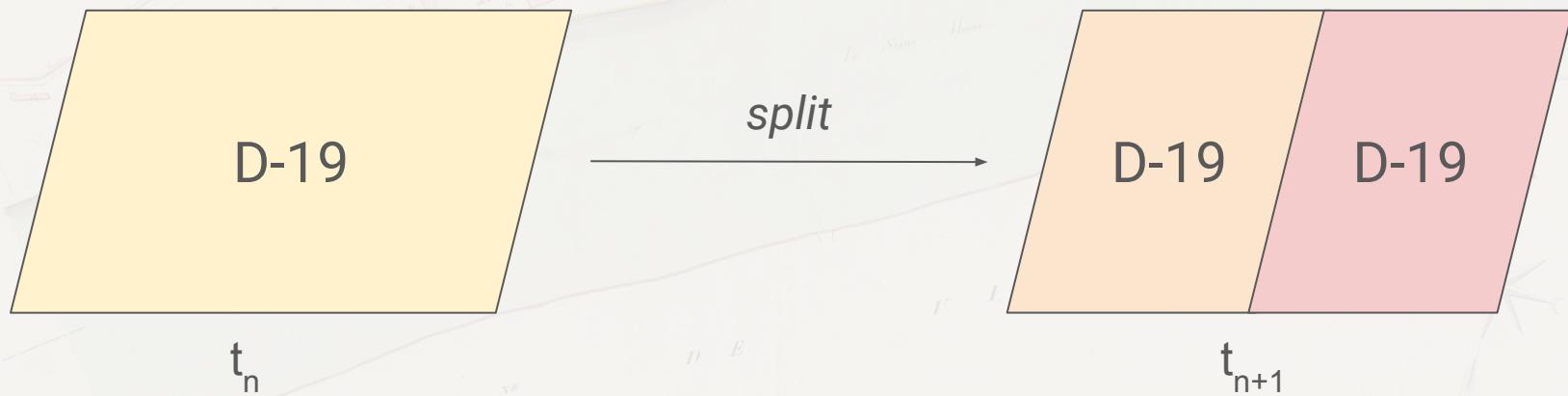


Detection of similar landmark within sources

Population

4. [OPTIONAL] Inferring landmarks identity (*if not known at step 1*)

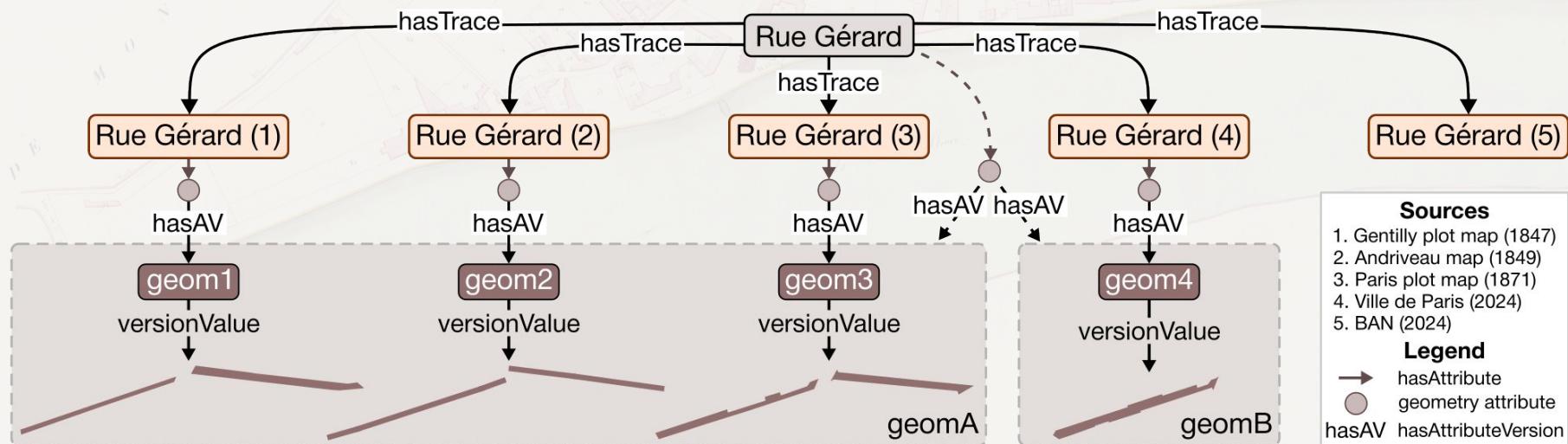
Example : in the land registry, if a plot is split, its successors keeps the same number in the registers



Population

5. Merge landmark attribute versions
6. Create changes and events that affect attributes of landmarks

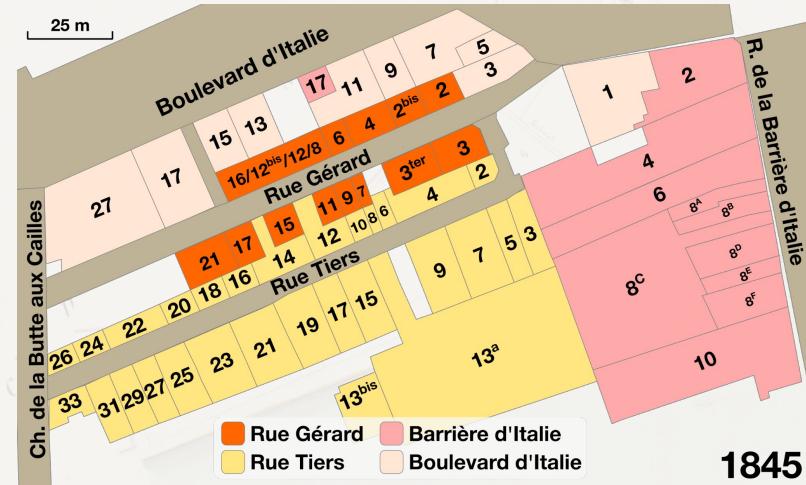
The process can be repeated to infer new knowledge.



Example of merging attribute versions for Rue Gérard

Evaluation

Which streets already exist in 1845?
What were the plot addresses in 1845?



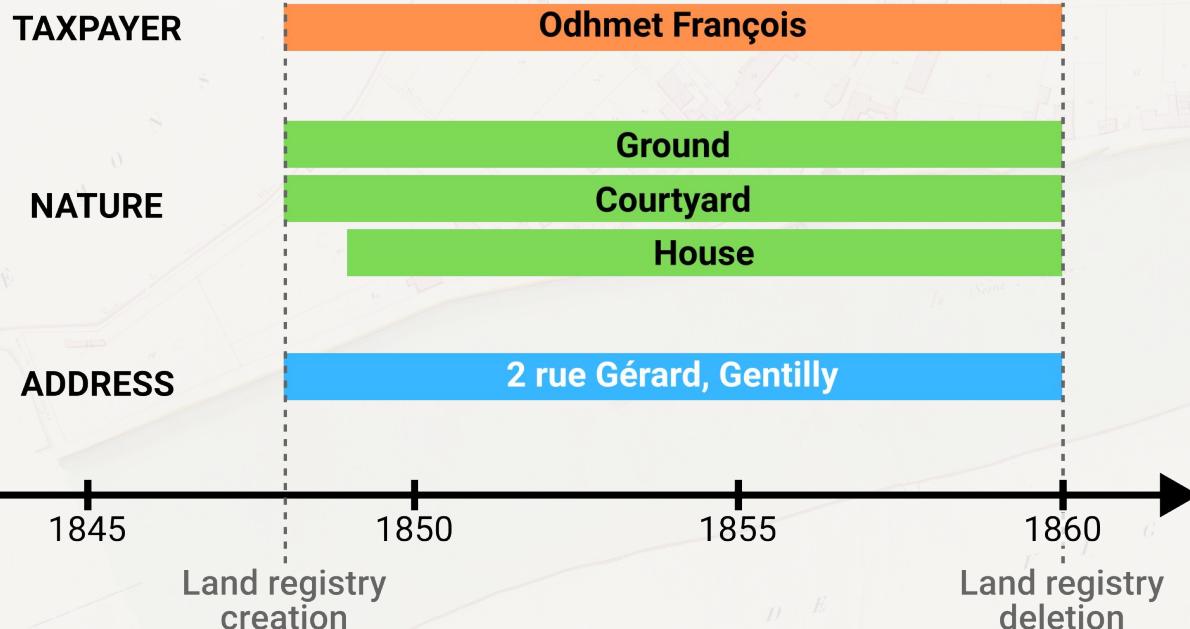
Snapshot of the Butte aux Cailles neighborhood in 1845

- Answer competency questions
- Check internal consistency of data with SHACL rules and SPARQL query
- Evaluate information consistency:
 - difficult to assess quality as no ground truth exists
 - ⇒ snapshots
 - ⇒ multi-source mapping (ex: addresses in several sources)

Evaluation

Evolution of the attributes of a landmark

*What are the successive natures/addresses/taxpayers
of plot D-14 in Gentilly?*



Plot D-14, Gentilly (second land registry)

Contributions

- **Ontology**

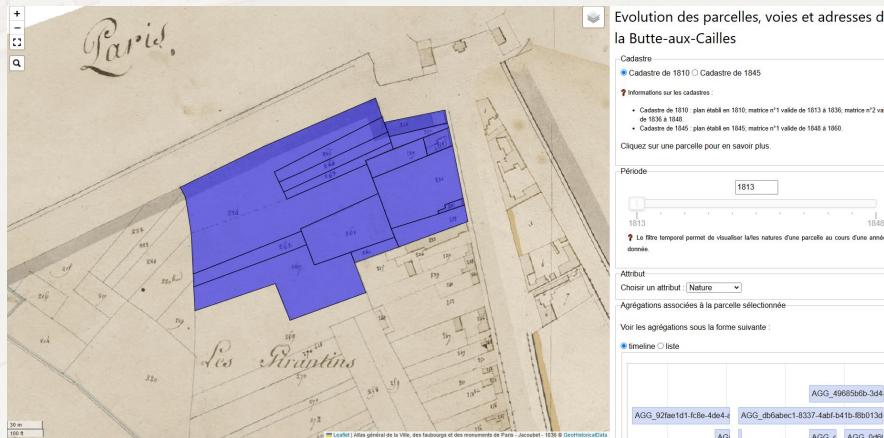
- describes addresses and landmarks
- **represents the evolution of their attributes**

- **Approach** to build a perpetual gazetteer as a KG

- treats several heterogeneous sources
- which handles heterogeneous fragmentary data

Future work

- Use case on Butte aux Cailles neighborhood:
 - intend to **enlarge spatial cover** (whole Paris?)
 - add **new sources** (notary minutes)
- **Critical analysis** of the knowledge graph (detect inconsistencies)
- Tools to **easily ask questions** to the graph (web map, natural language)



A web map to query the KG
(under developpement)

PeGaZus

- **Ontology**
 - describes addresses and landmarks
 - **represents the evolution of their attributes**
- **Approach** to build a perpetual gazetteer as a KG
 - several heterogeneous sources
 - fragmentary data

Thank you for
your attention !

Ontology, data and scripts are available

<https://github.com/umrlastig/pegazus-ontology>

charly.bernard@ign.fr ; solenn.tual@ign.fr



References

- [1] Kauppinen, T., Väätäinen, J., Hyvönen, E. (2008). Creating and Using Geospatial Ontology Time Series in a Semantic Cultural Heritage Portal. In: Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M. (eds) The Semantic Web: Research and Applications. ESWC 2008. Lecture Notes in Computer Science, vol 5021. Springer, Berlin, Heidelberg.
https://doi.org/10.1007/978-3-540-68234-9_11
- [2] Bernard, Camille, Villanova-Oliver, Marlène, Gensel, Jérôme, Dao, Hy. : Modeling changes in territorial partitions over time: Ontologies tsn and tsn-change. In Proceedings of the 33rd Annual ACM Symposium on Applied Computing, pp. 866–875. Pau, France (2018).
- [3] Charles, William, Aussenac-Gilles, Nathalie, Hernandez, Nathalie. : HHT: an approach for representing temporally-evolving historical territories. In European Semantic Web Conference, pp. 419–435. Cham: Springer Nature Switzerland, (2023).
- [4] Charly Bernard, Nathalie Abadie, Julien Perret, Bertrand Duménieu. Création d'un référentiel géo-historique d'adresses à partir de sources multiples. GAST - Gestion et l'Analyse de données Spatiales et Temporelles, Jan 2024, Dijon, France. <hal-04490732>

Ontology, data and scripts are available

<https://github.com/umrlastig/pegazus-ontology>

charly.bernard@ign.fr ; solenn.tual@ign.fr

