

An approach to extract structured data from the Napoleonic land registry maps and registers

Assia Ahmed-Abdi ⁽¹⁾, Solenn Tual ⁽¹⁾, Nathalie Abadie ⁽¹⁾, Joseph Chazalon ⁽³⁾, Bertrand Duméniou ⁽²⁾, Julien Perret ^(1,2)

GDR-IASIS : Intelligence artificielle pour le patrimoine image, 3 Octobre 2024, Saint-Mandé, France.

Licence Creative Commons BY SA 4.0. [[hal-04721201](https://hal.archives-ouvertes.fr/hal-04721201)]

IGN
INSTITUT NATIONAL
DE L'INFORMATION
GÉOGRAPHIQUE
ET FORESTIÈRE



(1)

L'ÉCOLE
DES HAUTES
ÉTUDES EN
SCIENCES
SOCIALES

(2)



(3)

cnrs **GDR** Groupement
de recherche
IASIS Information Apprentissage
Signal Image vISion

The Napoleonic land registry

- Why?
 - Making property tax more efficient and fair
- When?
 - Initiated by the law on 15th September 1807, at least one version was produced for each town between 1808 and 1845.
 - Replaced by the renovated land registry between 1930 and 1975 depending on the town.
 - First corpus of very high scale maps of France and first high precision referential of land use and cover



Cadastral index maps of Gentilly (94), Le Croisic (44) and Brest (29)

NATURE DES PROPRIÉTÉS.		COPIE DU TARIFF DÉFINITIF DU PROJET NET		
NATURE	POUR UN ARPENT DE CHAMP EN CARRÉ	1 ^{re}	2 ^e	3 ^e
Gravel				
Pierre				
Gravel & Pierre				
Bois				
Bois				
Gravel & Bois				
Pierre				
Pierre & Bois				
Bois				
Bois Impérial				
Bois d'alignement				
Bois de réserve				
Bois de réserve				

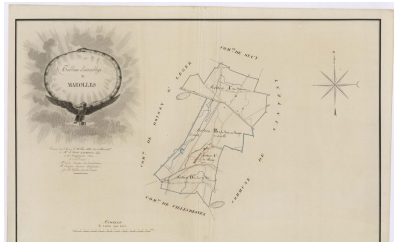
List of the plots natures and associated value. Change register of 1812, Marolles-en-Brie (94)



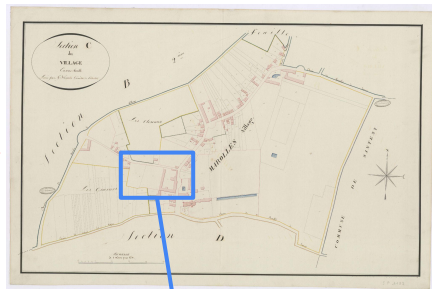
Land registry records (initial register and change register), Marolles-en-Brie, 1810-1932

The Napoleonic land registry: how does it work?

Assembly map of the cadastral index maps (1810)



Cadastral index map (1810)

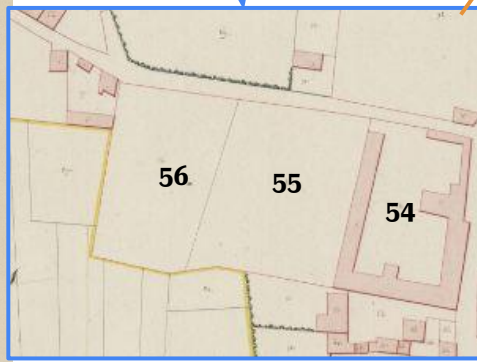


Initial register (1812)

CANTON, TRIAGES, ou Lieux dits.	NUMÉROS DU PLAN		NOMS, PROFESSIONS, DEMEURES des Propriétaires ou Usultiers.	NATURES des Propriétés.	CONTENANCE		
	provisoire	définitif			par Propriété.	par nature de culture.	
	+	54	Beaubrier, alias M ^r Prince de Neufchâteau	Maison et Cour	30 15		so. M. 2 ^e date
	+	55	Idem	Idem	37 60	37 00	2 ^e
	+	56	Idem	Vigne, arbray	29 00	29 00	1 ^e

Initial description of plots C-54, C-55 and C-56

Change registers (valid from 1812 to 1822)



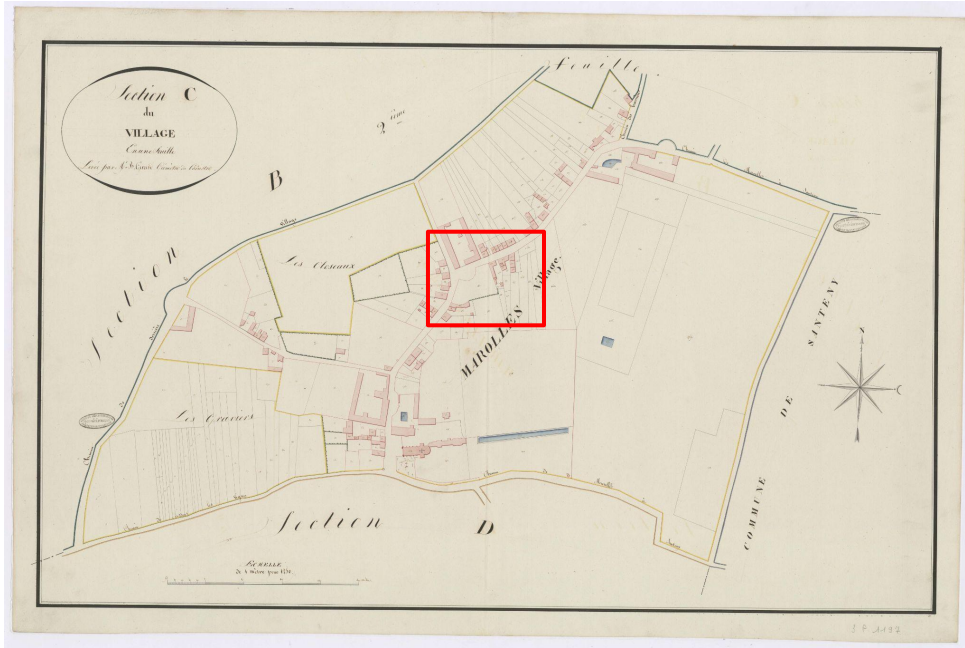
Town of Marolles-en-Brie, Section C

ARTICLES.	NOMS, PRÉNOMS, PROFESSIONS ET DEMEURES des Propriétaires et Usultiers.	INDICATION				NOMBRE D'ARPENS, PERCHES ET MÈTRES,		CLASSES.	REVENU		REVOIS pour les Mutations,	
		de la section.	du N ^o de la section.	DU CANTON, triage ou lieu dit.	DE LA NATURE de la propriété.	par article de section.	par article de MATRICE.		de chaque article de SECTION.	de chaque article de MATRICE.	tiré de l'article.	porté à l'article.
9	Beaubrier, Alexandre Prince de Neufchâteau à Gros Bois	B	54	hiefessard	Cour	arp. p. m.	arp. p. m.	2	fr. c.	fr. c.		
						2 06 ff		2	22 27			
						0 14 r.		2	12 02			
						0 30 al		1	19 60			
						0 37 60		2	29 70			
						29 00		1	26 67			

States of plots C-54, C-55 and C-56 at a given time in the first change register

A geohistorical knowledge graph of plots

Plot 191, Section C, Marolles-en-Brie (Val-de-Marne)



1810

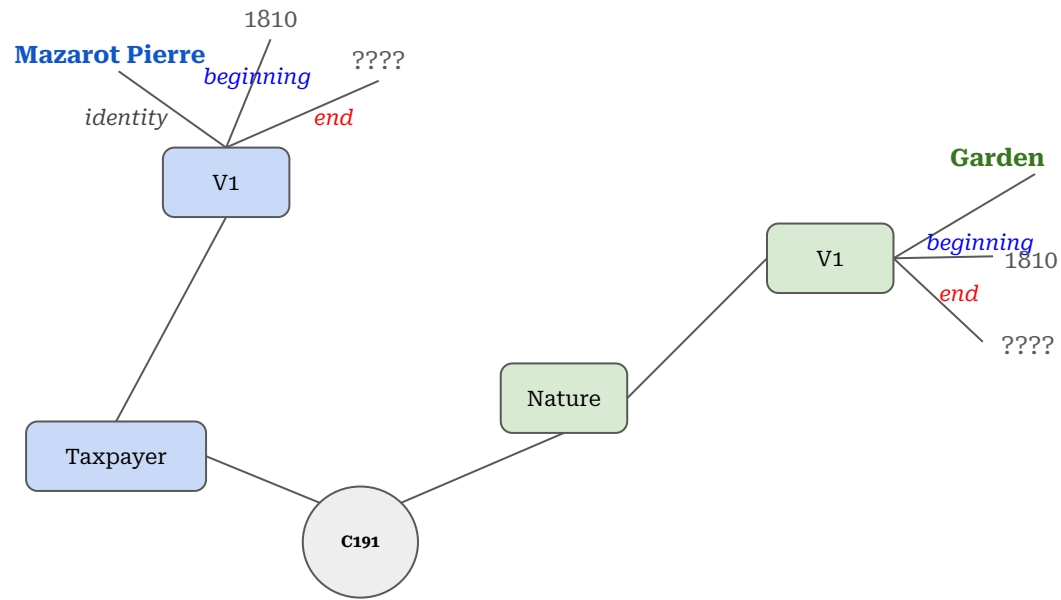
Knowledge graph construction

Initial register of Marolles en Brie (1812)

Page number	
Named place	Du Village
Taxpayer	Mazarot Pierre cabaretier à id
Nature	Jardin
Surface	2p. 35m.
Taken from page number X	
Beginning date	
Passed to page number Y	
End date	

35

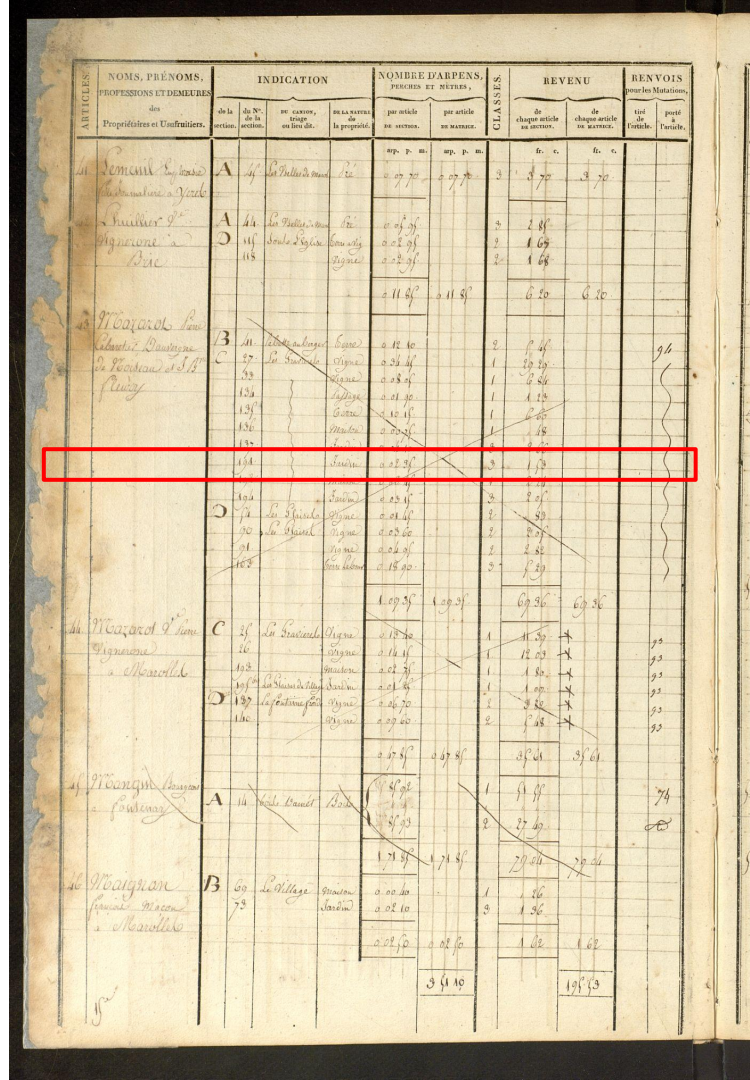
CANTON, TRIAGES, ou Lieux dits.	NUMÉROS DU PLAN provisoire, définitif.	N O M S, PROFESSIONS, DEMEURES des Propriétaires ou Usufruitiers.	NATURES des Propriétés.	CONTENANCE	
				par Propriété.	par aune de culture.
				a. p. m.	a. p. m.
Du Village	+ 186	Galland, Louis meuble et cabaretier à id	Bois Commune	4-95	0
	+ 187	Galland, Louis, meub. à id	Bâtiment	95	0
	+ 188	Id	Bois Commune	3-60	1-70 3-00
	+ 189	Idem	Bois Vigne	2-60	2-40
	+ 190	Idem	Bois Vigne	8-60	2-
	+ 191	Mazarot, Pierre Cabaretier à id	Jardin	2-35	10-75
	+ 192	Idem	Maison à id	3-45	6-10 à id
	+ 193	Mazarot, Pierre	Bois	2-75	6-10 à id
	+ 194	Mazarot, Pierre	Jardin	3-15	3-
	+ 195	Idem	Bois Commune	6-10	à id
+ 196	Mazarot, Pierre	Bois Commune	2-20	à id	
+ 197	Idem	Bois Commune	56-30	à id	
+ 198	Idem	Bois Commune	36-15	62-80 36-15	
+ 199	Idem	Bois Commune	1-01	1-01	
+ 200	Idem	Bois Commune	10-60	10-60	
+ 201	Idem	Bois Commune	63-00	63-00	
+ 202	Idem	Bois Commune	10	10	
+ 203	Idem	Bois Commune	28-36	28-36	
+ 204	Idem	Bois Commune	3-05	3-05	
+ 205	Idem	Bois Commune	4-00	4-00	
+ 206	Idem	Bois Commune	4-15-20	4-15-20	

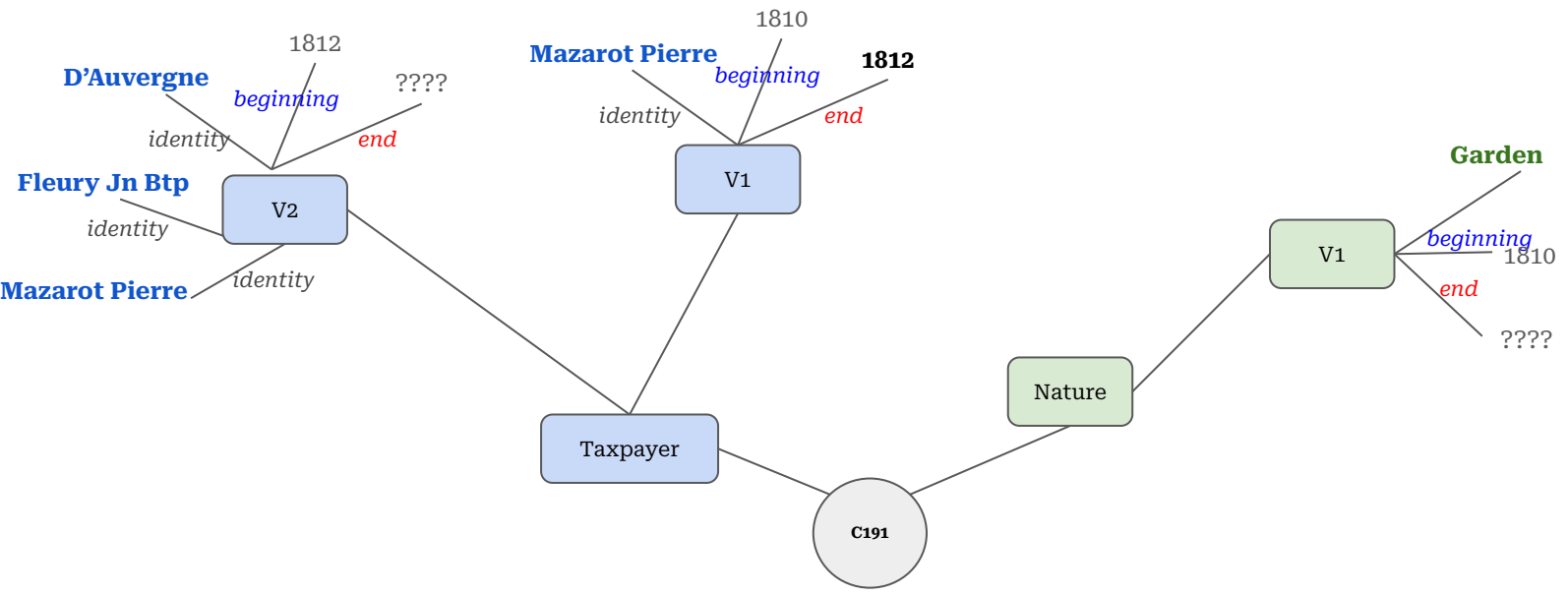


Knowledge graph construction

Change register (first one, used from 1812 to 1822)

Page number	43
Named place	Le Village
Taxpayer	Mazarot Pierre cabaretier, D'Auvergne de Noiseau et J ⁿ B ^t Fleury
Nature	Jardin
Surface	2p. 35m.
Taken from page number X	
Beginning date	
Passed to page number Y	94
End date	????



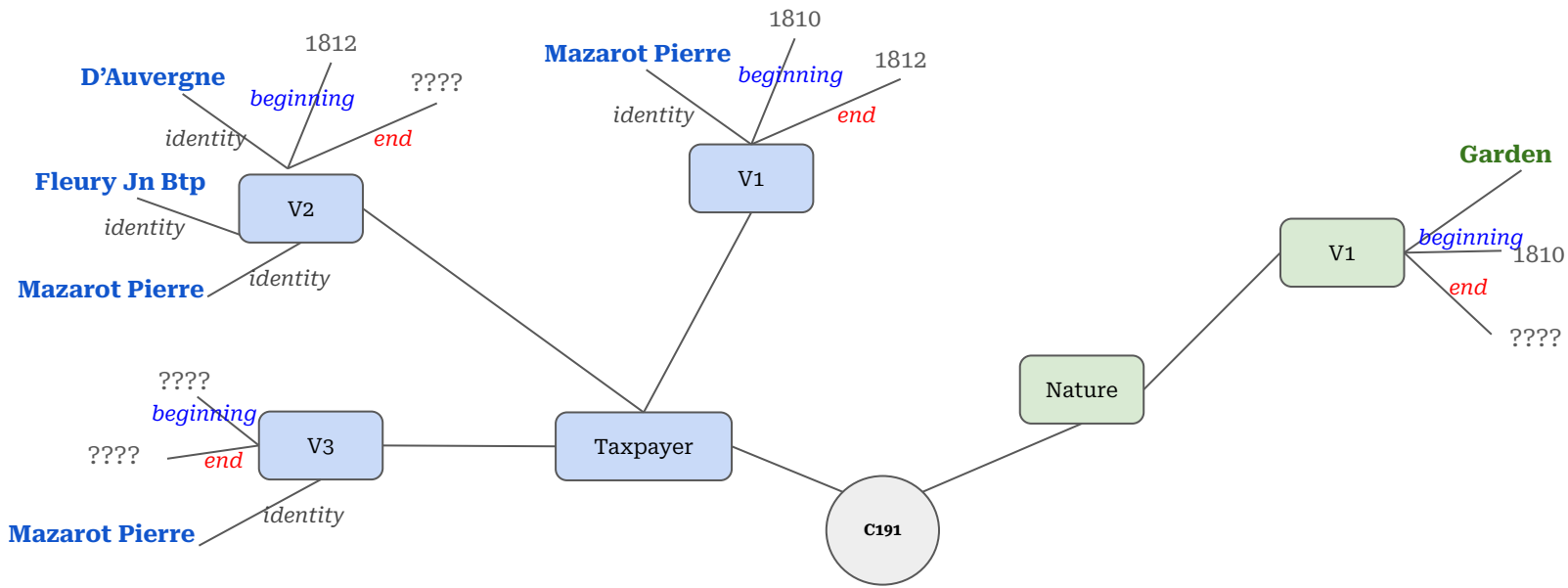


Knowledge graph construction

Change register (first one, used from 1812 to 1822)

Page number	94
Named place	Le Village
Taxpayer	Mazarot Pierre cabaretier à Marolles
Nature	Jardin
Surface	2p. 35m.
Taken from page number X	43
Beginning date	????
Passed to page number Y	
End date	

ARTICLES.	NOMS, PRÉNOMS, PROFESSIONS ET DEMEURÉS des Propriétaires et Concessionnaires.	INDICATION de la section, de la nature, de la charge ou lieu de la propriété.	NOMBRE D'ARTICLES, PAR ARTICLE ou SECTION.		CLASSE.	REVENU de chaque article ou section.		REVENUS pour les Métrés.		
			par article ou section.	par article ou section.		en fr.	en c.	par article ou section.	par article ou section.	
93	Jais Guillot père	C. 172	Le grenier	00 92	1	00 75		46		
			Le jardin	00 82	1	00 75		46		
			Le jardin	00 30	2	00 75		46		
			Le jardin	00 97	1	00 75		46		
			Le jardin	00 08	2	00 75		46		
			Le jardin	02 25	2	1 27 75		46		
			Le jardin	03 20	2	1 53 75		46		
			Le jardin	07 60	2	6 64		46		
			Le jardin	18 90	3	5 29		46		
			Le jardin	2 01 69	1	100 92		46		
94	Mazarot Pierre cabaretier à Marolles	C. 172	Le grenier	02 10	2	00 67		46		
			Le grenier	08 93	1	7 19		46		
			Le jardin	09 43	1	8 02		46		
			Le jardin	24 45	1	29 29		46		
			Le jardin	05 05	1	5 84		46		
			Le jardin	01 90	1	1 23		46		
			Le jardin	10 15	1	6 60		46		
			Le jardin	00 37	1	0 88		46		
			Le jardin	06 10	2	1 66		46		
			Le jardin	02 57	2	1 70		46		
			Le jardin	00 97	1	2 20		46		
			Le jardin	01 83	1	1 20 75		46		
			Le jardin	02 17	2	2 05		46		
			Le jardin	00 83	1	0 75 75		46		
			95	Bernard fuy	C. 172	Le grenier	01 67	2	0 83	
Le grenier	03 60	2				2 05		46		
Le grenier	06 97	2				2 82		46		
Le grenier	06 47	2				2 17		46		
Le grenier	06 80	2				3 05		46		
Le grenier	18 90	3				1 29		46		
Le grenier	07 70	1				0 03		46		
Le grenier	6 06 07	2				112 69		46		
Le grenier	08 50	1				0 02		46		
Le grenier	3 47 70	3				97 36		46		
96	Bernard fuy	C. 172	Le grenier	6 82 17	2	179 80		46		
			Le grenier	11 37	1	0 06		46		
			Le grenier	02 17	1	0 07		46		
			Le grenier	9 98 70	2	279 67		46		
			Le grenier	1 63 70	1	16 17		46		
			Le grenier	2 17 77	1	38 81		46		
			Le grenier	6 17 77	2	126 63		46		
			Le grenier	31 94 60	1	860 15		46		
			Le grenier							
			Le grenier							



Knowledge graph construction

Change register (second register of built properties, used from 1911 to 1932)

Page number	24
Named place	Village
Taxpayer	Guillot Pierre Louis Nicolas à Marolles Vandermasse Rémy commis de ferme à Marolles
Nature	Maison /// Fournil
Surface	
Taken from page number X	
Beginning date	
Passed to page number Y	
End date	

The image shows a page from a historical tax register (Case 23 and 24). The page is divided into two main sections, one for 'Case 23' and one for 'Case 24'. Each section contains a table for recording property details and a summary table for tax calculations.

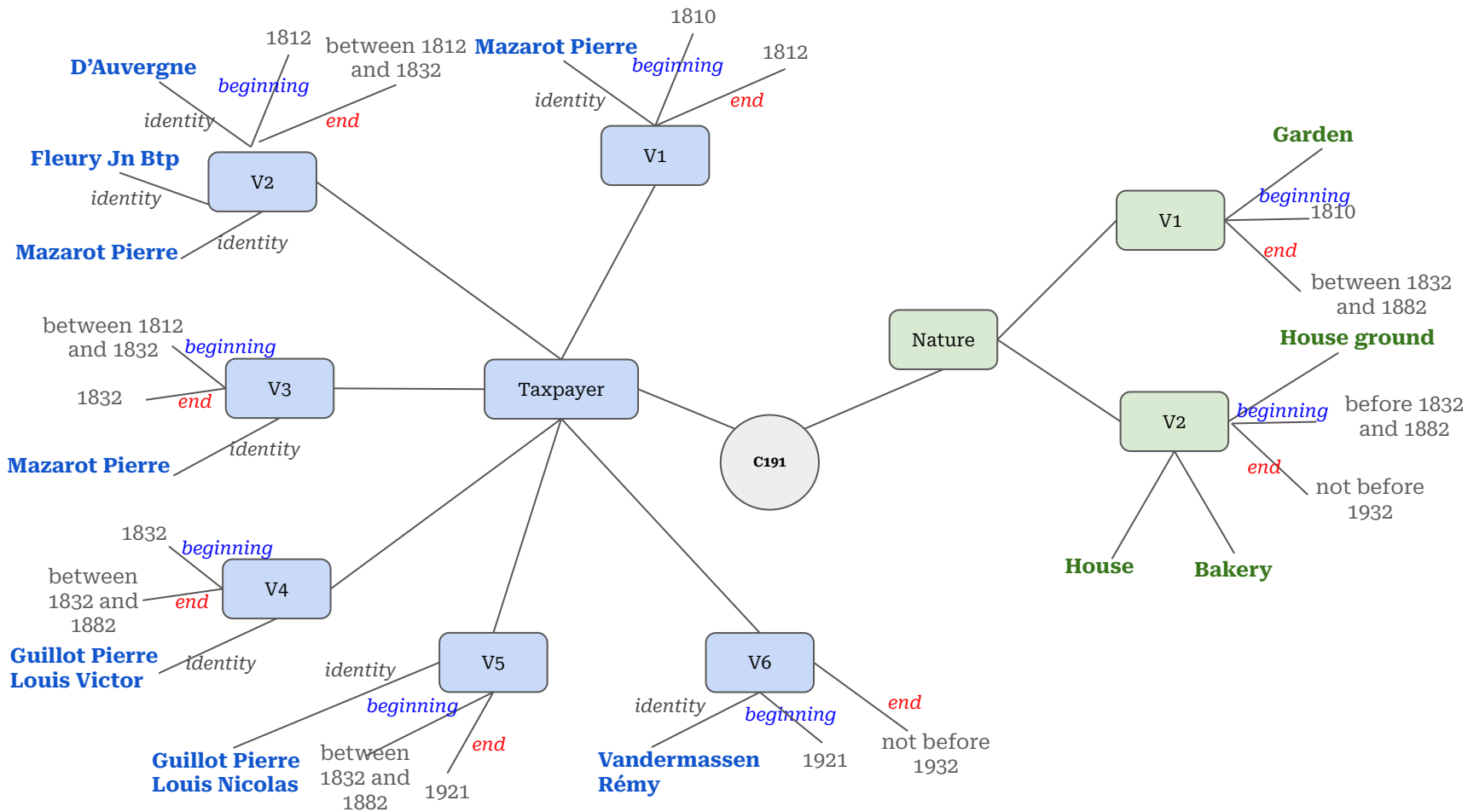
Case 23:

- MUTATIONS:** ANNEE de l'ASSURANCE (1911), VILLE (Village), NATURE (Maison /// Fournil).
- LIEU-DIT:** Village.
- NATURE:** Maison /// Fournil.
- REVENU NET IMPONABLE:** Table with columns for 'pour le période commençant le 1^{er} janvier' and 'pour le période terminant le 31 décembre'. Handwritten entries show values for 1911, 1912, and 1913.
- TOTAL:** Summary table at the bottom of the section.

Case 24:

- MUTATIONS:** ANNEE de l'ASSURANCE (1911), VILLE (Village), NATURE (Maison /// Fournil).
- LIEU-DIT:** Village.
- NATURE:** Maison /// Fournil.
- REVENU NET IMPONABLE:** Table with columns for 'pour le période commençant le 1^{er} janvier' and 'pour le période terminant le 31 décembre'. Handwritten entries show values for 1911, 1912, and 1913.
- TOTAL:** Summary table at the bottom of the section.

A red box highlights the entries for 'Village' and 'Fournil' in the table for Case 24.



Objectives

Applications

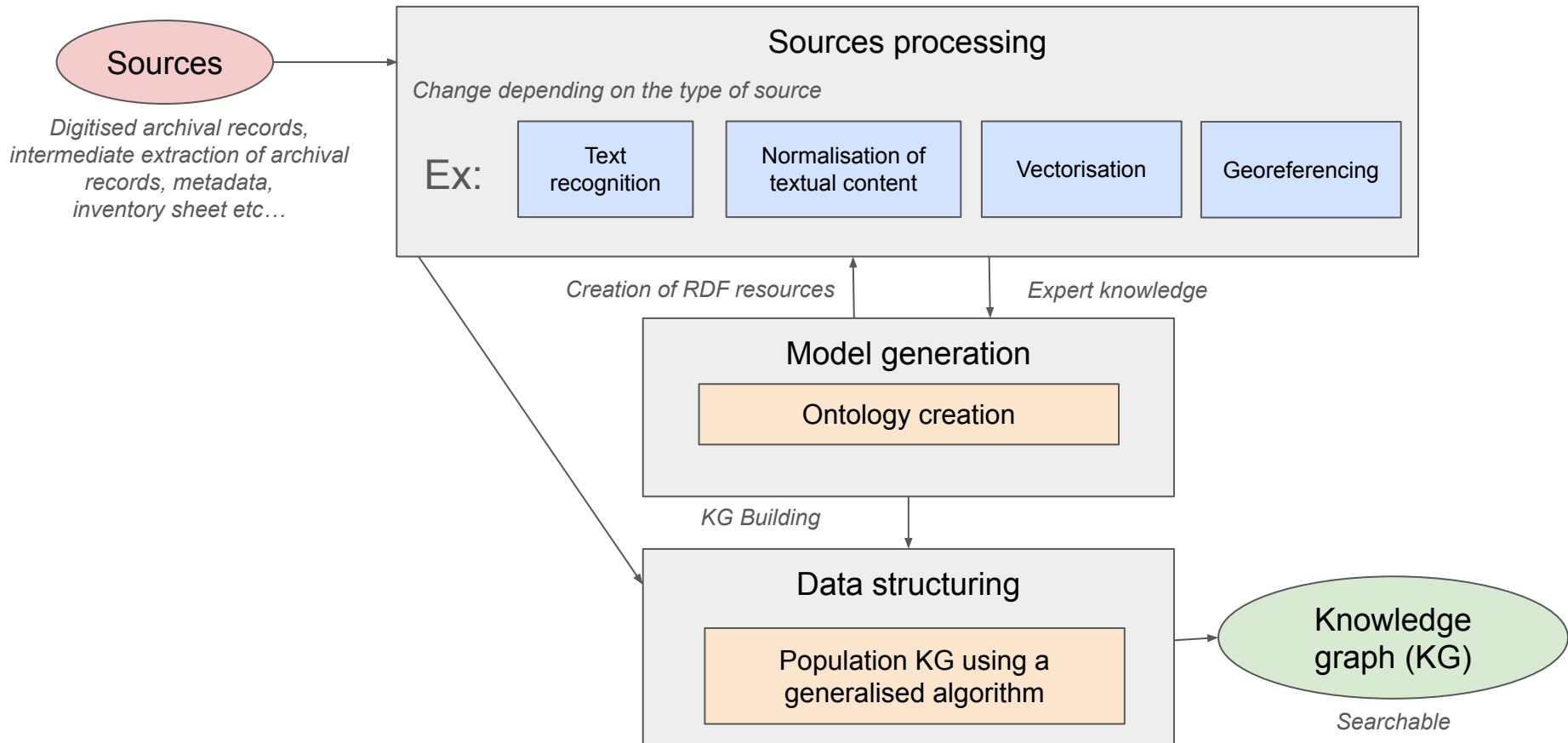
- Create a geo-historical knowledge graph useful to track changes in land use and land cover from the early 19th century onwards.
- Complete and refine the building construction dates in the IGN's databases.
- To enhance the value of a significant cartographic and archival heritage.

Goals

- Setting up an automatic processing chain to extract information from old land registry documents (maps and registers).
- Creating a geohistorical knowledge graph of plots using these information

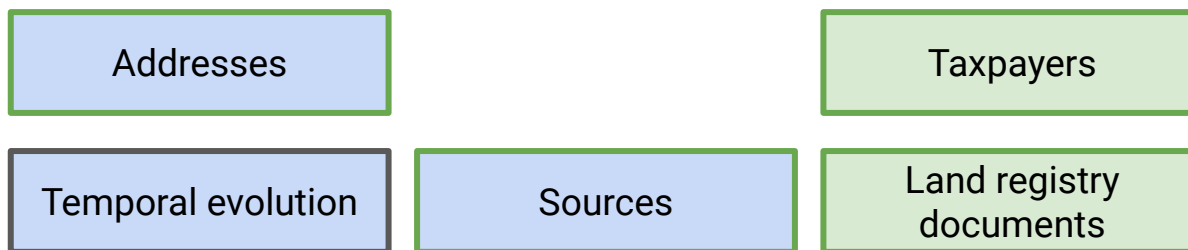
Open Science

Pipeline



Ontology

PERpetual GAZeteer of approach-address UtteranceS [4] [5] : ontology



In **blue**: modelets in common with the ontology of [4] + specific additions to describe the cadastre.

In **green**: modelets specific to the cadastre

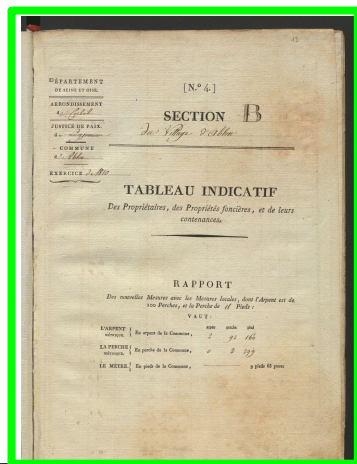
[4] Charly Bernard, Solenn Tual, Nathalie Abadie, Bertrand Duménieu, Joseph Chazalon, Julien Perret. **PeGazUs: A knowledge graph based approach to build urban perpetual gazetteers**, European Knowledge Acquisition Workshop (EKAW-2024), November 26-28 2024, Amsterdam, Netherlands. <https://github.com/umrlastig/pegazus-ontology>

[5] 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>

Information extraction from digitised registers

Multiple page types (e.g. section cover, main table, intermediate resume, final resumes...).

=> Classification step



NUMERO	NOM, DESCRIPTION, SITUATION	NATURE DE LA PROPRIÉTÉ	CONTENANCE	PROPRIÉTAIRE	CLASSE
1	Château de Chateaubriant	Terre	2 25 1/2		5
2	Château de Chateaubriant	Terre	2 25 1/2		5
3	Château de Chateaubriant	Terre	2 25 1/2		5
4	Château de Chateaubriant	Terre	2 25 1/2		5
5	Château de Chateaubriant	Terre	2 25 1/2		5
6	Château de Chateaubriant	Terre	2 25 1/2		5
7	Château de Chateaubriant	Terre	2 25 1/2		5
8	Château de Chateaubriant	Terre	2 25 1/2		5
9	Château de Chateaubriant	Terre	2 25 1/2		5
10	Château de Chateaubriant	Terre	2 25 1/2		5

CLASSEMENT		APPLICATION DU TAUX		TOTAL	
CLASSE	CONTENANCE	Taux	Produit	CLASSE	Produit
1	2 25 1/2	100	2 25 1/2	1	2 25 1/2
2	2 25 1/2	100	2 25 1/2	2	4 50
3	2 25 1/2	100	2 25 1/2	3	7 25 1/2
4	2 25 1/2	100	2 25 1/2	4	10 00
5	2 25 1/2	100	2 25 1/2	5	12 50
6	2 25 1/2	100	2 25 1/2	6	15 00
7	2 25 1/2	100	2 25 1/2	7	17 50
8	2 25 1/2	100	2 25 1/2	8	20 00
9	2 25 1/2	100	2 25 1/2	9	22 50
10	2 25 1/2	100	2 25 1/2	10	25 00

RÉCAPITULATION DES CONTENANCES										
CLASSE	CONTENANCE	PRODUIT	CLASSE	PRODUIT	CLASSE	PRODUIT	CLASSE	PRODUIT	CLASSE	PRODUIT
1	2 25 1/2	2 25 1/2	2	4 50	3	7 25 1/2	4	10 00	5	12 50
6	2 25 1/2	15 00	7	17 50	8	20 00	9	22 50	10	25 00
Total										

Conte de Chateaubriant sur papier de la Commune de Chateaubriant le 25 Mars 1808

RELEVÉ des Nombres de déca mètres									
SECTION	CLASSEMENT	APPLICATION DU TAUX		TOTAL		PRODUIT			
1	2 25 1/2	100	2 25 1/2	1	2 25 1/2	100	2 25 1/2	100	2 25 1/2
2	2 25 1/2	100	2 25 1/2	2	4 50	100	4 50	100	4 50
3	2 25 1/2	100	2 25 1/2	3	7 25 1/2	100	7 25 1/2	100	7 25 1/2
4	2 25 1/2	100	2 25 1/2	4	10 00	100	10 00	100	10 00
5	2 25 1/2	100	2 25 1/2	5	12 50	100	12 50	100	12 50
6	2 25 1/2	100	2 25 1/2	6	15 00	100	15 00	100	15 00
7	2 25 1/2	100	2 25 1/2	7	17 50	100	17 50	100	17 50
8	2 25 1/2	100	2 25 1/2	8	20 00	100	20 00	100	20 00
9	2 25 1/2	100	2 25 1/2	9	22 50	100	22 50	100	22 50
10	2 25 1/2	100	2 25 1/2	10	25 00	100	25 00	100	25 00

RELEVÉ des Nombres de déca mètres									
SECTION	CLASSEMENT	APPLICATION DU TAUX		TOTAL		PRODUIT			
1	2 25 1/2	100	2 25 1/2	1	2 25 1/2	100	2 25 1/2	100	2 25 1/2
2	2 25 1/2	100	2 25 1/2	2	4 50	100	4 50	100	4 50
3	2 25 1/2	100	2 25 1/2	3	7 25 1/2	100	7 25 1/2	100	7 25 1/2
4	2 25 1/2	100	2 25 1/2	4	10 00	100	10 00	100	10 00
5	2 25 1/2	100	2 25 1/2	5	12 50	100	12 50	100	12 50
6	2 25 1/2	100	2 25 1/2	6	15 00	100	15 00	100	15 00
7	2 25 1/2	100	2 25 1/2	7	17 50	100	17 50	100	17 50
8	2 25 1/2	100	2 25 1/2	8	20 00	100	20 00	100	20 00
9	2 25 1/2	100	2 25 1/2	9	22 50	100	22 50	100	22 50
10	2 25 1/2	100	2 25 1/2	10	25 00	100	25 00	100	25 00

RELEVÉ des Nombres de déca mètres									
SECTION	CLASSEMENT	APPLICATION DU TAUX		TOTAL		PRODUIT			
1	2 25 1/2	100	2 25 1/2	1	2 25 1/2	100	2 25 1/2	100	2 25 1/2
2	2 25 1/2	100	2 25 1/2	2	4 50	100	4 50	100	4 50
3	2 25 1/2	100	2 25 1/2	3	7 25 1/2	100	7 25 1/2	100	7 25 1/2
4	2 25 1/2	100	2 25 1/2	4	10 00	100	10 00	100	10 00
5	2 25 1/2	100	2 25 1/2	5	12 50	100	12 50	100	12 50
6	2 25 1/2	100	2 25 1/2	6	15 00	100	15 00	100	15 00
7	2 25 1/2	100	2 25 1/2	7	17 50	100	17 50	100	17 50
8	2 25 1/2	100	2 25 1/2	8	20 00	100	20 00	100	20 00
9	2 25 1/2	100	2 25 1/2	9	22 50	100	22 50	100	22 50
10	2 25 1/2	100	2 25 1/2	10	25 00	100	25 00	100	25 00

RÉCAPITULATION de la somme de la somme liquidée de la section										
PROPRIÉTÉS NON BATIES										
CLASSE	CONTENANCE	PRODUIT	CLASSE	PRODUIT	CLASSE	PRODUIT	CLASSE	PRODUIT	CLASSE	PRODUIT
1	2 25 1/2	2 25 1/2	2	4 50	3	7 25 1/2	4	10 00	5	12 50
6	2 25 1/2	15 00	7	17 50	8	20 00	9	22 50	10	25 00
Total										

PROPRIÉTÉS BATIES										
RÉSUMÉ										
NATURE DES PROPRIÉTÉS		CLASSE	CONTENANCE	PRODUIT	CLASSE	CONTENANCE	PRODUIT	CLASSE	CONTENANCE	PRODUIT
Bâties		1	2 25 1/2	2 25 1/2	2	4 50	3	7 25 1/2	4	10 00
Non bâties		5	22 50	22 50	6	15 00	15 00	7	17 50	17 50
Total		6	24 75 1/2	24 75 1/2	8	20 00	20 00	9	22 50	22 50
ADRESSE DE MOYEN										

Information extraction from digitised registers

Multiple table structures of different initial registers from Seine and Seine-et-Oise departments (near Paris).

Second cadaster of Seine dept.(1835-1845)

Before 1822

CANTONS, TERRAGE ou lieux-dits.	NUMÉROS de la table hypothécaire de section.	NOMS, PROFESSIONS, demeures des Propriétaires et Usufruitiers.	NATURE des parcelles de propriétés, par nature de culture.	CONTENANCE par parcelle de propriétés, en ares, centiares, millièmes.	provisionnel DU CLASSEMENT.
Chemin de Brée	96. 166	Lefèvre	terre	27510	1
Chemin de Brée	29. 167	Ducloson pignon	terre	1135	1
	29. 168	Caron.	terre	9610	1
	96. 169	Lefèvre	terre	2590	1
	29. 170	Ducloson, pignon	terre	0720	1
	29. 171	Serau	terre	0560	1
	81. 172	Travis.	terre	0615	1
	81. 173		terre	0610	1

After 1822

NOMS, PRÉNOMS, PROFESSIONS ET DEMIERS DES PROPRIÉTAIRES.	N° de parcelle.	CANTONS ou lieux-dits.	NATURE des propriétés.	CONTENANCE en ares, centiares, millièmes.	CLASSE.	REVENU en francs, centimes.
Paris Denis vicent	146	St Jean	terre	2 3/4	1	10
Millevoy guaren gultann	147		terre	1 1/2	1	10
Voisier Jopeur.	148		terre	1 1/4	1	10
Millevoy Jean Jean	149		terre	2 1/2	1	10
Chabot, Jean Jopeur	150		terre	2 1/2	1	10
Millevoy Jean millevoy	151		terre	2 1/2	1	10
Millevoy Jean millevoy	152		terre	2 1/2	1	10
Millevoy Jean millevoy	153		terre	2 1/2	1	10
Millevoy Jean millevoy	154		terre	2 1/2	1	10
Millevoy Jean millevoy	155		terre	2 1/2	1	10
Millevoy Jean millevoy	156		terre	2 1/2	1	10
Millevoy Jean millevoy	157		terre	2 1/2	1	10
Millevoy Jean millevoy	158		terre	2 1/2	1	10
Millevoy Jean millevoy	159		terre	2 1/2	1	10

NOMS, PRÉNOMS, PROFESSIONS ET DEMIERS DES PROPRIÉTAIRES.	N° de parcelle.	CANTONS, TERRAGE ou lieux-dits.	NATURE des propriétés.	CONTENANCE en ares, centiares, millièmes.	INDICATION DE LA NATURE DE LA CULTURE.		ES EN CULTURE.	ES EN CULTURE.	ES EN CULTURE.	ES EN CULTURE.
					TERRE	TERRE				
Paris Denis vicent	146	St Jean	terre	2 3/4	1	10	1	1	1	1
Millevoy guaren gultann	147		terre	1 1/2	1	10	1	1	1	1
Voisier Jopeur.	148		terre	1 1/4	1	10	1	1	1	1
Millevoy Jean Jean	149		terre	2 1/2	1	10	1	1	1	1
Chabot, Jean Jopeur	150		terre	2 1/2	1	10	1	1	1	1
Millevoy Jean millevoy	151		terre	2 1/2	1	10	1	1	1	1
Millevoy Jean millevoy	152		terre	2 1/2	1	10	1	1	1	1
Millevoy Jean millevoy	153		terre	2 1/2	1	10	1	1	1	1
Millevoy Jean millevoy	154		terre	2 1/2	1	10	1	1	1	1
Millevoy Jean millevoy	155		terre	2 1/2	1	10	1	1	1	1
Millevoy Jean millevoy	156		terre	2 1/2	1	10	1	1	1	1
Millevoy Jean millevoy	157		terre	2 1/2	1	10	1	1	1	1
Millevoy Jean millevoy	158		terre	2 1/2	1	10	1	1	1	1
Millevoy Jean millevoy	159		terre	2 1/2	1	10	1	1	1	1

Named place where the plot is located in a section

Plot number

Taxpayer index number in the change register

Taxpayer identity

Plot nature

Former plot number (previous cadaster)

Former plot nature (previous cadaster)

Information extraction from digitised registers

Document Attention Network (DAN [1]), a document analysis and text recognition model (TEKLIA implementation [2])
Dataset “Initial registers”: around 150 annotated pages, *key:value* annotation with Callico [3]

The image shows a digitized page from a cadastral register with a digital annotation interface. The register table has the following columns: CANTONS, NUMÉROS, NOMS, PROFESSIONS, NATURE, CONTENANCE, CONTENANCE, and COLONNE. The interface on the right shows the following extracted data:

CANTONS	NUMÉROS	NOMS, PROFESSIONS	NATURE	CONTENANCE	CONTENANCE	COLONNE
TRIAGES ou lieux-dits.	de la table alphabétique de section.	demeures des Propriétaires et Usufruitiers.	des propriétés.	par parcelles de propriété impossible. sans impossible.	par nature de culture.	provisoire DU CLASSEMENT.
chemin de Brie	95. 166	Lefevre	terre	2 75 10	1	1
chemin de Brie	28. 167	Duchon propriétaire	terre	11 35	1	1
	29. 168	Carton	terre	9 10	1	1
	95. 169	Lefevre	terre	25 90	1	1
	28. 170	Duchon, propriétaire	terre	07 20	1	1
	28. 171	Serun	terre	05 10	1	1
	81. 172	Sarrieu	terre	06 15	1	1
	81. 173		terre	06 10	1	1

The digital interface on the right shows the following fields and their values:

- Canton, triage, lieu-dit, rue ou tout autre entité géographique: chemin—de Brie—
- Numéro du propriétaire dans la table alphabétique: 95
- Identité du propriétaire ou de l'usufruitier: Lefevre
- Numéro de la parcelle: 166
- Nature de la parcelle: terre
- Ancien(s) numéro(s) de la parcelle: indiqué(s) uniquement dans le cadastre rectifié de la Seine
- Ancienne(s) nature(s) de la parcelle: indiquée(s) uniquement dans le cadastre rectifié de la Seine

[1] Denis Coquenat, Clément Chatelain, Thierry Paquet, **DAN: a Segmentation-free Document Attention Network for Handwritten Document Recognition**, *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2023.

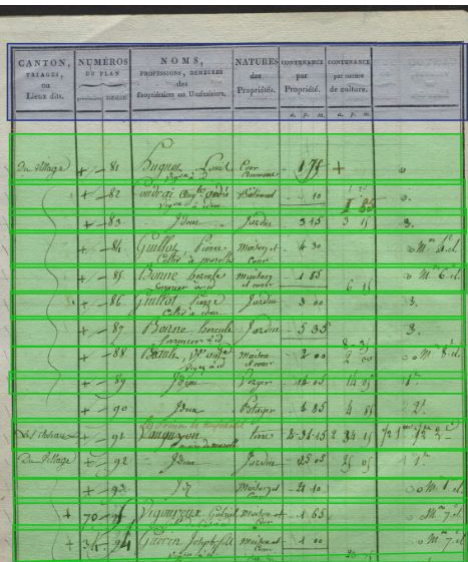
[2] <https://gitlab.teklia.com/atr/dan>

[3] Christopher Kermorvant, Eva Bardou, Manon Blanco, Bastien Abadie, **Callico: a Versatile Open-Source Document Image Annotation Platform**, *International Conference on Document Analysis and Recognition (ICDAR)*, Athens, Greece, 2024

Information extraction from digitised registers

Initial register transcribed at page level (ground-truth)

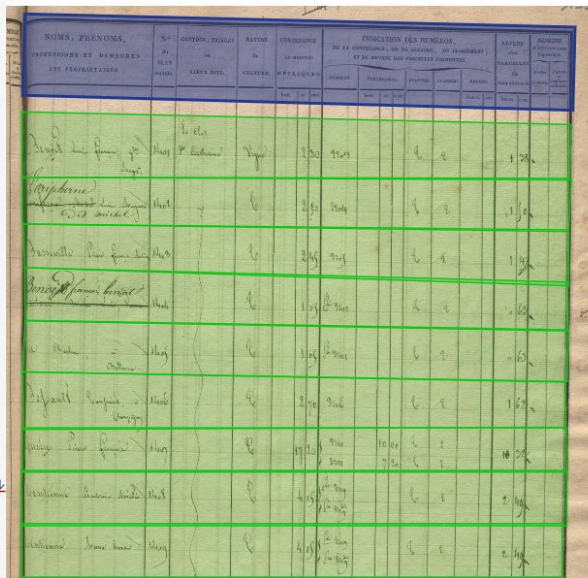
- Named place where the plot is located in a section
- Taxpayer identity
- Plot number
- Former plot number
- Plot nature
- Former plot nature



A scanned page from a cadastre register before 1822. The table has columns for 'CANTON', 'NUMEROS', 'NOMS', 'NATURES', 'CONTENANCE', 'COTISATION', and 'REMARQUES'. The rows contain handwritten entries with plot numbers and names.

[lieu-dit](#)Du Village [identite](#)Huguet
[Louis](#)→[vig](#)↑[on](#)↓ à [id](#) [numero_parcelle](#)81 [nature](#)
[Cour](#)→[Commune](#)
[lieu-dit](#)\$ [identite](#)Coudrau [aug](#)↑[te](#)↓
[andré](#)→[vig](#)↑[on](#)↓ à [idem](#) [numero_parcelle](#)82
[nature](#)Bâtiment
[lieu-dit](#)\$ [identite](#)Idem [numero_parcelle](#)83 [nature](#)
[Jardin](#)
[lieu-dit](#)\$ [identite](#)Guillot, Pierre→[Cultiv](#) à
[marolles](#) [numero_parcelle](#)84 [nature](#)maison
[et](#)→[cour](#)
[lieu-dit](#)\$ [identite](#)Bonne hercule→[serrurier](#) à [id](#)
[numero_parcelle](#)85 [nature](#)maison→[et](#) [cour](#)
[lieu-dit](#)\$ [identite](#)Guillot Pierre→[cultiv](#) à [idem](#)
[numero_parcelle](#)86 [nature](#)Jardin
[lieu-dit](#)\$ [identite](#)Bonne hercule→[serrurier](#) à [id](#)
[numero_parcelle](#)87 [nature](#)Jardin
[lieu-dit](#)\$ [identite](#)Beaulh, Vre↓[ant](#)↑[e](#)↓→[vig](#)↑[on](#)↓
à [id](#) [numero_parcelle](#)88 [nature](#)maison→[et](#) [cour](#)
[lieu-dit](#)\$ [identite](#)Idem [numero_parcelle](#)89 [nature](#)
[Verger](#)

Initial register table before 1822



A scanned page from a cadastre register for the second cadaster of Seine (1835-1845). The table has columns for 'NOMS, PRENOMS, SURNOMS, QUALITES ET HERAULTS DES PROPRIETAIRES', 'SECTION', 'NATURE', 'CONTENANCE', 'INDICATION DES SURCROISS, DE LA DIMINUTION DE LA CONTENANCE, DE LA CHANGEMENT, ET DE MOUVEN DES PARCELLES SUIVANTES', and 'NATURE DES CULTURES'. The rows contain handwritten entries with plot numbers and names.

[lieu-dit](#)Le Clos→[S](#)↑[te](#)↓ Catherine [identite](#)Benoit
[Louis François, g](#)↑[dre](#)↓→[Rouyer](#) [numero_parcelle](#)
1401 [nature](#)Vigne [ancien_numero_parcelle](#)2203
[ancienne_nature](#)I
[lieu-dit](#)Id [identite](#)Loriferne→×[Loriferne](#)
[Michel](#)±[Louis Auguste](#)→[id](#) et [Michel](#)
[numero_parcelle](#)1402 [nature](#)I
[ancien_numero_parcelle](#)2204 [ancienne_nature](#)I
[lieu-dit](#)\$ [identite](#)Basseville Pierre François
[numero_parcelle](#)1403 [nature](#)I
[ancien_numero_parcelle](#)2205 [ancienne_nature](#)I
[lieu-dit](#)\$ [identite](#)Benoit François benoit→×[actar](#)
[nicolas louis, Vanve](#)± [numero_parcelle](#)1404 [nature](#)
I [ancien_numero_parcelle](#)P↑[tie](#)↓ 2205
[ancienne_nature](#)I
[lieu-dit](#)\$ [identite](#)Charles à→[Bellevue](#)
[numero_parcelle](#)1405 [nature](#)I
[ancien_numero_parcelle](#)P↑[tie](#)↓ 2201
[ancienne_nature](#)I

Initial register table, second cadaster of Seine
(1835-1845)

Information extraction from digitised registers

Initial register transcribed at page level (WIP results)

CANTONS	NUMEROS	NOMS, PROFESSIONS,	NATURE	CONTENANCE	CONTENANCE	COLONNE
ou lieux-dits.	de la table	des Propriétaires et Descriptions.	des	par parcelle de propriété	par mètres	provisoire
	de			superficie.	de surface.	ou CLASSÉS
	146 206	Tremelet Em.	Terre	30 50	30 50	3 3 3
	147 207	Berthier	Terre	12 30	12 30	2 2 2
	148 208	Debeste	Terre	20 60	20 60	3 3 3
	149 209		Terre	19 45	19 45	3 3 3
	150 210		Terre	19 45	19 45	3 3 3
	151 211		Terre	19 45	19 45	3 3 3
	152 212		Terre	19 45	19 45	3 3 3
	153 213		Terre	19 45	19 45	3 3 3
	154 214		Terre	19 45	19 45	3 3 3
	155 215		Terre	19 45	19 45	3 3 3
	156 216		Terre	19 45	19 45	3 3 3
	157 217		Terre	19 45	19 45	3 3 3
	158 218		Terre	19 45	19 45	3 3 3
	159 219		Terre	19 45	19 45	3 3 3
	160 220		Terre	19 45	19 45	3 3 3
	161 221		Terre	19 45	19 45	3 3 3
	162 222		Terre	19 45	19 45	3 3 3
	163 223		Terre	19 45	19 45	3 3 3
	164 224		Terre	19 45	19 45	3 3 3
	165 225		Terre	19 45	19 45	3 3 3
	166 226		Terre	19 45	19 45	3 3 3
	167 227		Terre	19 45	19 45	3 3 3
	168 228		Terre	19 45	19 45	3 3 3
	169 229		Terre	19 45	19 45	3 3 3
	170 230		Terre	19 45	19 45	3 3 3

Taxpayer
first folio

Taxpayer identity

Plot
number

Nature

M146N	E Tremelet laimé F	K256L	I Terre J
M147N	E Derthier - J7 N guyon F	K257L	I Pré J
M32N	E Debeste F	K258L	I Terre J
M32N	E id F	K259L	I Pd J
M32N	E id F	K260L	I Terre J
M32N	E id F	K261L	I id J
M32N	E id F	K262L	I id J
M143N	E Tourneur barthelemy V ↑ es ↓ F	K263L	I id J
M147N	E Vaugergon F	K264L	I id J
M99N	E Daspine alexandre F	K265L	I id J
M147N	E Berthier F	K266L	I id J
M147N	E id F	K267L	I id J
M2N	E V ↑ e ↓ Bault F	K268L	I id J
M2N	E id F	K269L	I id J
M111N	E Masareau P ↑ re ↓ F	K270L	I id J
M111N	E Vasareau P ↑ re ↓ F	K270L	I id J
M147N	E Vaugerger F	K270L	I id J
M94N	E Dasareau P ↑ re ↓ F	K270L	I id J

Wrong additional NER token

Missing crossed out text tokens and
missing not crossed out text part

Missing crossed out text

Hallucinations

Knowledge graph building

Input

DAN-like output + text normalisation (produced by hand for this work)

ID	UUID	Type_CF	Num_Folio	Parcelle_treated	Lieu-dit_transcript	Lieu-dit_treated	Lieu-dit_type	Propriétaires_transcript	Nature_clean	Nature_treated	Date entrée_Date sortie	Tiré de	Tiré de_treated
17	5bb8bf	Non Bâti	138	224	Rue Thiers	Rue Thiers	Thoroughfare	Delaunay Elisabeth→Rue de la monnaie 17 à Paris	Maison	Maison	1836	293	293
18	c1c226	Non Bâti	138	224	id	Rue Thiers	Thoroughfare	Delaunay Elisabeth→Rue de la monnaie 17 à Paris	terre	Terre	1836	440	440
49	5a399c	Non Bâti	268	224	id	Les Girantiers	District	Leroy→automne fets ne Charles→à Villejuif	terre	Terre	1836		
55	99b681	Non Bâti	293	224	∅	Les Girantiers	District	Mériot→Charles froids→Claudet (1834)→rue tiers n°03	Maison	Maison	1831	1836	C'ons, N'elle, ConstructionNo
58	d62ba8	Non Bâti	302	224	rue Thiers	Rue Thiers	Thoroughfare	Meunier→Jean Charles	Maison	Maison	1831	1836	C'ons, N'elle, ConstructionNo
67	9127de	Non Bâti	440	224	id	Les Girantiers	District	Thiersse→à paris	terre	Terre	1836		
13	1c302ft	Non Bâti	107bis	224	id	Les Girantiers	District	~~Cornillon→vin primeur en caractère~~→1832→Cosson notaire→à Paris	Terre	Terre	1836	268	268
14	60a6b0	Non Bâti	107bis	224	§	Les Girantiers	District	~~Cornillon→vin primeur en caractère~~→1832→Cosson notaire→à Paris	Maison	Maison	1836	274	274
1	c2da61	Bâti	108	226				Legendre H↑re↓ de →Fontainebleau	maison	Maison			
46	969143	Non Bâti	260	226	∅	Les Girantiers	District	Legendre→B↑re↓ fontainebleau	M↑on↓	Maison			
47	ff016dd	Non Bâti	260	226	Les girantiers	Les Girantiers	District	Legendre→B↑re↓ fontainebleau	maison	Maison			
3	552180	Bâti	82	227				Hardon, Bourgeois ~nourisseur~~→à Paris	maison	Maison	1814	114	114
2	1843ae	Bâti	114	227				Louves	M↑on↓	Maison	1814		
48	4c5395	Non Bâti	260	227	id	Les Girantiers	District	Legendre→B↑re↓ fontainebleau	maison	Maison	1814		

Input data: Dataset containing the parcel status of the study area from the 3 Gently matrices before 1860 (produced manually)

Knowledge graph building

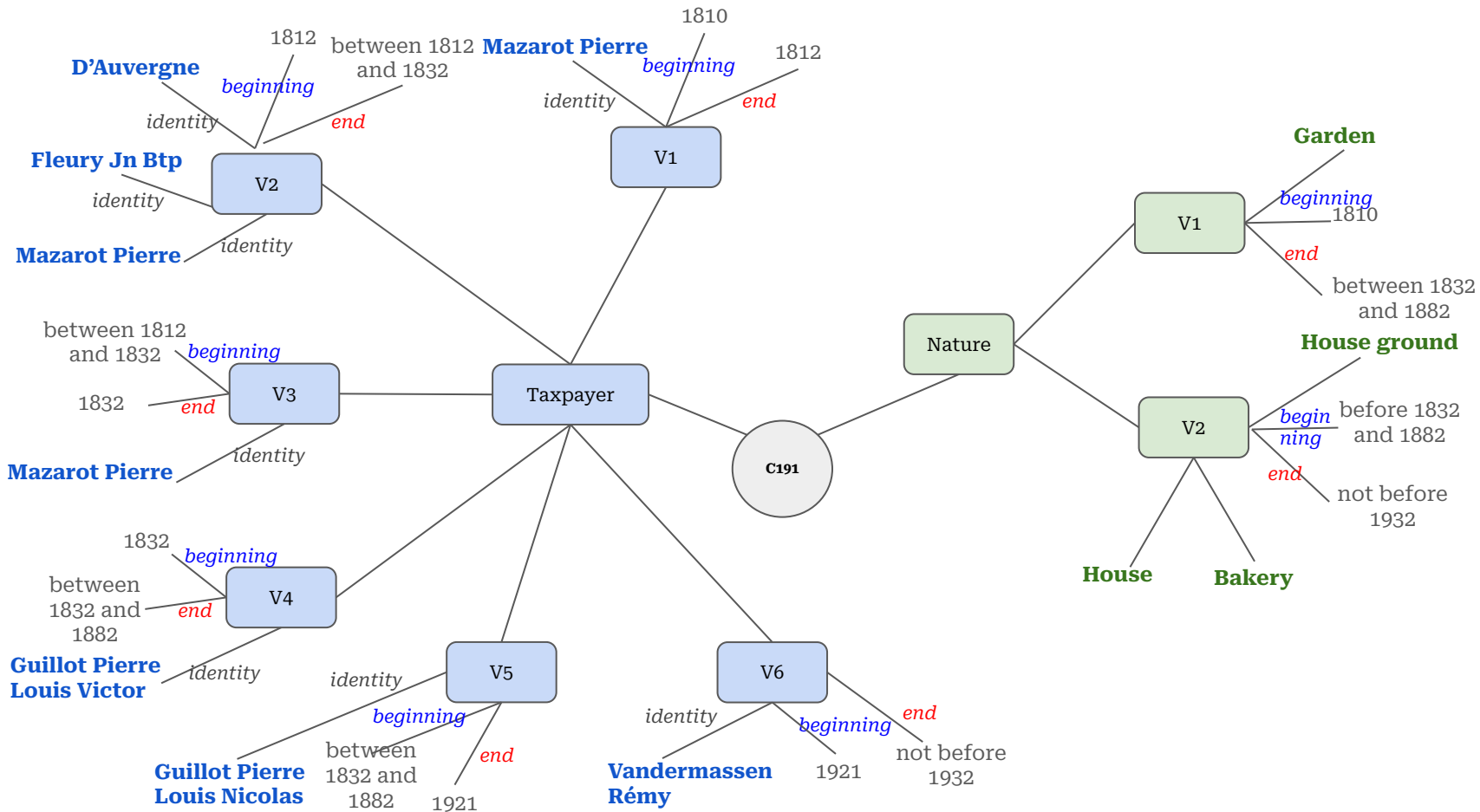
PERpetual GAZetteer of approach-address UtteranceS [4] : method

1. Gather all factoids relative to the same plot
2. Reorder factoids in the chronological order
3. Generate changes (in data) and (real-world) events related to the plots
 - a. Changes : LandmarkAppearance, LandmarkDisappearance
 - b. Events : Split of a plot, Merge of plots etc.
4. Plot identity disambiguation
5. Merge successive attribute versions if they are equal
6. Generate changes (in data) and events related to the plots attributes'
 - a. AttributeVersionAppearance, AttributeVersionTransition, AttributeVersionDisappearance
 - b. Event : Plot sale, Nature change, New building etc.

[4] Charly Bernard, Solenn Tual, Nathalie Abadie, Bertrand Duméniéu, Joseph Chazalon, Julien Perret. **PeGazUs: A knowledge graph based approach to build urban perpetual gazetteers**, International Conference on Knowledge Engineering and Knowledge Management (EKAW-2024), November 26-28 2024, Amsterdam, Netherlands.

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

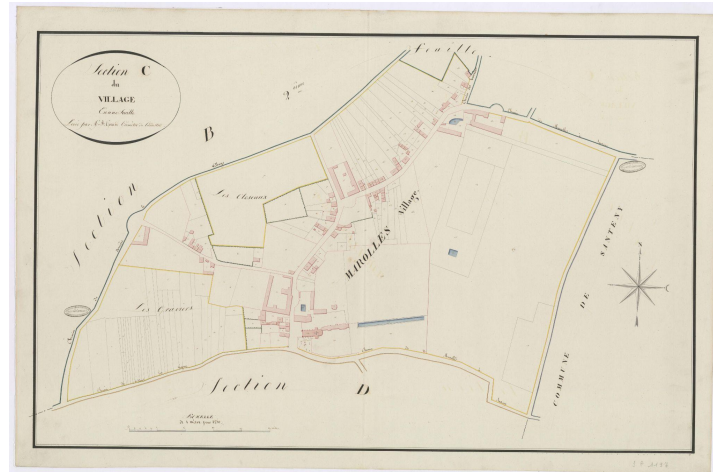
Knowledge graph building



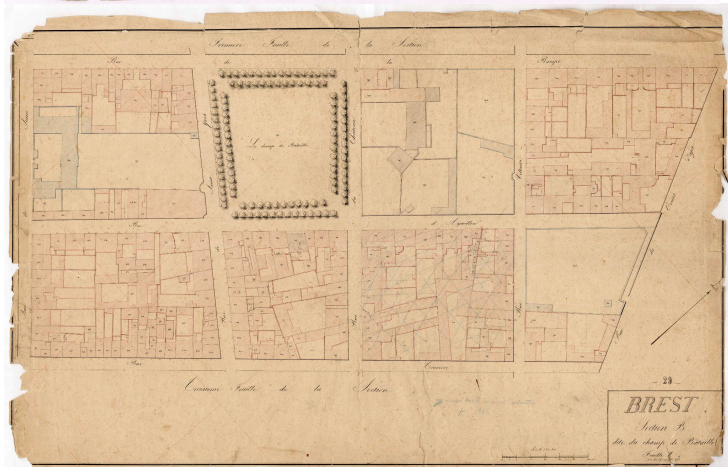
Cadastral index maps



*Le Croisic (A1),
Departement of
Loire-Atlantique,
1809*



*Marolles-en-Brie
(C), Departement of
Val-de-Marne, 1810*



*Brest (B),
Departement of
Finistère, 1830*



*Savine (B10),
Departement of
Hautes-Alpes, 1843*

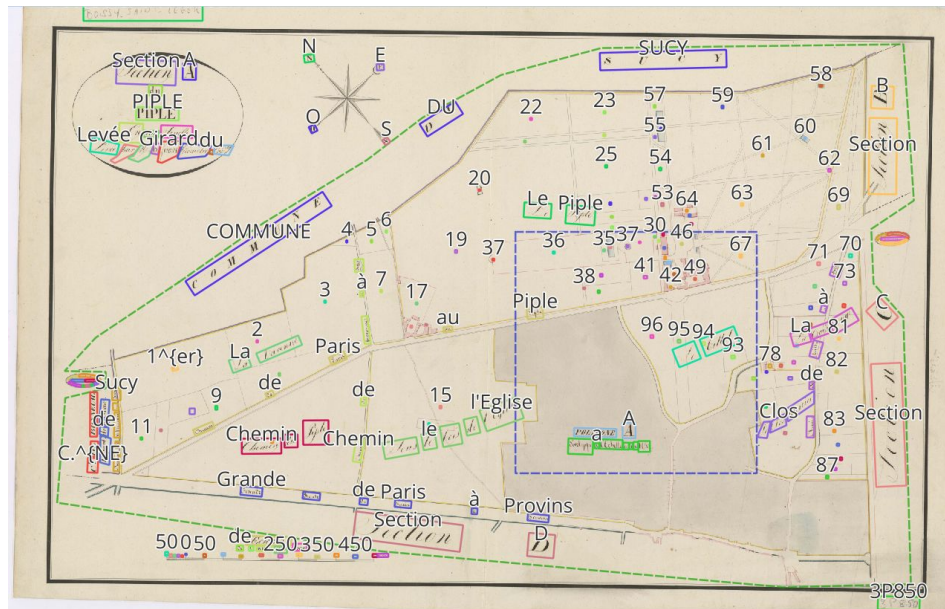
Detection, recognition and linking of text on the cadastral index maps

Dataset “French cadaster”

- Dataset created for the ICDAR 2024 MapText Competition on Historical Map Text Detection, Recognition, and Linking [5].
 - 145 crops (2000x2000 pixels)
 - 8 methods (1 to 4 tasks) submitted
- Dataset Part 2 is coming!
 - 110 maps from several administrative units, entirely annotated

ICDAR MapText Competition on Historical Map Text Detection, Recognition, and Linking should return in 2025 !

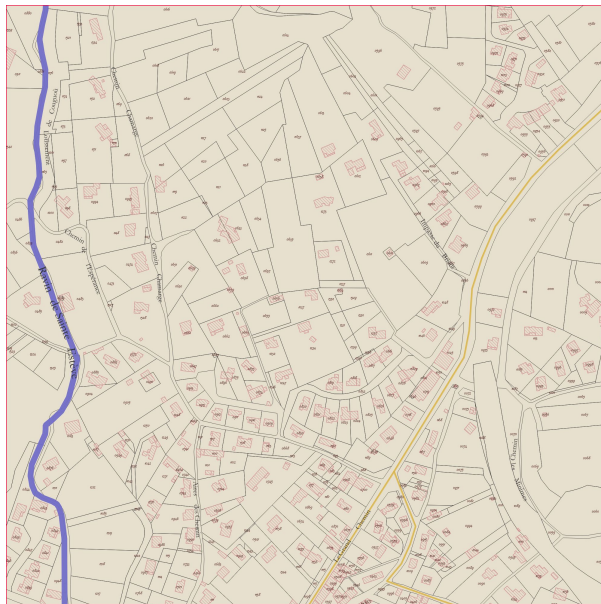
<https://rrc.cvc.uab.es/?ch=28>



Sheet number 1 of section A in Boissy-Saint-Léger. AD 94

Detection, recognition and linking of text on the cadastral index maps

Dataset “French cadaster synthetic maps”



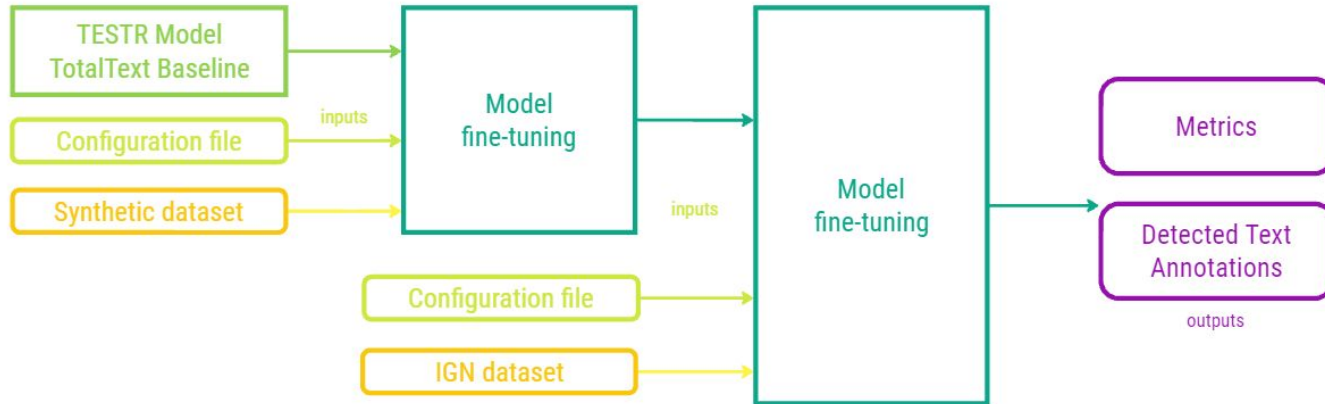
- Groups of 2000x2000 crops
- Produced with IGN data (BDTOPO, PCI Express) display with old land registry styles inspired by true old maps
- 5 french administrative units : variable topography, plots shapes and density

Examples of synthetic maps

Detection, recognition and linking of text on the cadastral index maps

Focus on text detection

- Internship of Assia Ahmed-Abdi (may to august 2024)
- Method based on TESTR model (pre-trained on natural scene images)



Detection, recognition and linking of text on the cadastral index maps

Evaluation

Rank	Method name	Quality	Tightness	fScore	Precision	Recall
1	dino_map	64.7	72.2	89.7	88.7	90.8
2	MapTest	64.2	69.9	91.9	90.8	93.0
3	TESTR - Synth Val de Marne - IGN transformed	57.5	67.2	85.5	85.6	85.4
4	ensem	51.8	71.4	72.5	90.3	60.6
5	DS-LP	44.1	65.0	67.8	64.8	71.0
6	MapText Dete...Strong Pipeline	42.3	69.3	61.1	82.5	48.5
7	MapDet	35.7	65.3	54.7	70.1	44.8
8	Baseline TESTR Checkpoint	20.6	70.5	29.2	86.4	17.6

Evaluation of the detection task



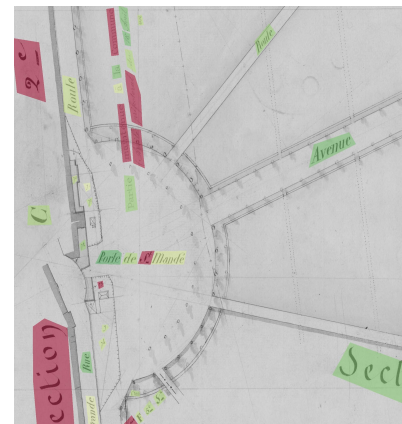
Precision map



Recall map



Precision map



Recall map

Future work

For the registers

1. improve DAN training (increase dataset size + optimise training parameters)
2. develop text normalisation process
3. semantic annotation of table to build RDF resources
 - + replicate the process for the change registers
 - + use KG building algorithm with data produced automatically

For the maps

- automated text detection, recognition and linking
 - links between the plots described in the registers and their location on the cadastral index map
- automated vectorization
- (automated) georeferencing

Thank you !

Solenn Tual, Nathalie Abadie, Joseph Chazalon, Bertrand Duménieu, Julien Perret. **An approach to extract structured information from maps and registers guided by the history of the Napoleonic cadaster**. 30e Conférence internationale sur l'histoire de la cartographie (ICHC 2024), Lyon, Jul 2024, Lyon, France. (hal-04637689)

Zekun Li, Yijun Lin, Yao-Yi Chiang, Jerod Weinman, Solenn Tual, Joseph Chazalon, Julien Perret, Bertrand Duménieu; Nathalie Abadie. **ICDAR 2024 Competition on Historical Map Text Detection, Recognition, and Linking**. In: Barney Smith, E.H., Liwicki, M., Peng, L. (eds) Document Analysis and Recognition - ICDAR 2024. ICDAR 2024. Lecture Notes in Computer Science, vol 14809. Springer, Cham. https://doi.org/10.1007/978-3-031-70552-6_22

Accepted :

Charly Bernard, Solenn Tual, Nathalie Abadie, Joseph Chazalon, Bertrand Duménieu, Julien Perret. **PeGazUs: A knowledge graph based approach to build urban historical gazetteers**. 24th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2024), Amsterdam, Netherlands, Nov 2024. <https://github.com/umrlastig/pegazus-ontology>

Assia Ahmed Abdi ⁽¹⁾, Solenn Tual ⁽¹⁾, Nathalie Abadie ⁽¹⁾, Joseph Chazalon ⁽³⁾, Bertrand Duménieu ⁽²⁾, Julien Perret ^(1,2)
solenn.tual@ign.fr nathalie-f.abadie@ign.fr joseph.chazalon@epita.fr
bertrand.dumenieu@ehess.fr julien.perret@ign.fr



Please refer to Hal to cite this presentation : <https://hal.science/hal-04721201>



(1)



(2)



(3)

