



HAL
open science

Reconstitution automatique de la généalogie des parcelles et des bâtiments à partir de sources historiques

Solenn Tual

► To cite this version:

Solenn Tual. Reconstitution automatique de la généalogie des parcelles et des bâtiments à partir de sources historiques. 32e Journées de la Recherche UGE-IGN-ENSG & Journée des doctorants de l'école doctorale MSTIC, Mar 2023, Champs-sur-Marne, France. 2023. <hal-04119250>

HAL Id: hal-04119250

<https://hal.science/hal-04119250v1>

Submitted on 6 Jul 2023

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 CC BY 4.0 - Attribution - International License

RECONSTITUTION AUTOMATIQUE DE LA GÉNÉALOGIE DES PARCELLES ET DES BÂTIMENTS À PARTIR DE SOURCES HISTORIQUES

DES CADASTRES ANCIENS AUX RÉFÉRENTIELS DE DONNÉES GÉOGRAPHIQUES MODERNES



Solenn Tual
LaSTIG, Univ. Gustave Eiffel,
IGN-ENSG



Contexte

Les progrès rapides apportés par les réseaux de neurones profonds dans le domaine de l'extraction d'informations à partir de sources non structurées (textes et images) ouvrent de multiples possibilités quant à la production automatique de données à partir de ces documents encore inexploités.

Objectifs

- Mise en place d'une chaîne de traitement automatique pour procéder à l'extraction des informations contenues dans les documents du cadastre ancien ;
- Structuration des données extraites sous la forme d'un graphe de connaissance géohistorique.

Applications

- Constitution de bases de données spatio-temporelles permettant de suivre l'évolution de l'occupation du sol pour des périodes antérieures à la couverture globale du territoire par photographie aérienne et ainsi de contribuer à la cartographie de l'Anthropocène dès le XIX^{ème} siècle ;
- Enrichissement des attributs relatifs aux dates des bâtiments dans les bases de données de l'IGN, utiles à la mise en oeuvre des politiques de rénovation thermique et au suivi de l'artificialisation des sols ;
- Valorisation d'un important patrimoine cartographique et archivistique.

Sources historiques

Le cadastre napoléonien est le premier cadastre parcellaire produit en France entre 1808 et 1850. Il est composé, pour chaque commune, de quatre types de documents.

- Le **tableau d'assemblage** : plan général de la commune, divisée en sections, généralement à l'échelle 1/10 000, qui représente l'assemblage des feuilles du plan parcellaire.
- Les **plans parcellaires** : plans représentant en détail les parcelles (numérotées) et bâtiments d'une portion de la commune. Il peut être réalisé aux échelles 1/2 500 (cas général), 1/1 250 (zones urbaines ou rurales denses) ou 1/5 000 (zones très peu denses).
- Le **tableau indicatif des propriétaires et propriétés foncières (état de section)** : tableau détaillant les informations associées à chaque parcelle à la date de production du plan (propriétaire, nature de l'occupation du sol, superficie, montant de l'impôt).
- La **matrice des propriétaires** : document divisé en deux parties. La liste alphabétique des propriétaires associée à chaque nom le ou les numéros de folios (pages) qui détaillent les transactions foncières de chaque individu.

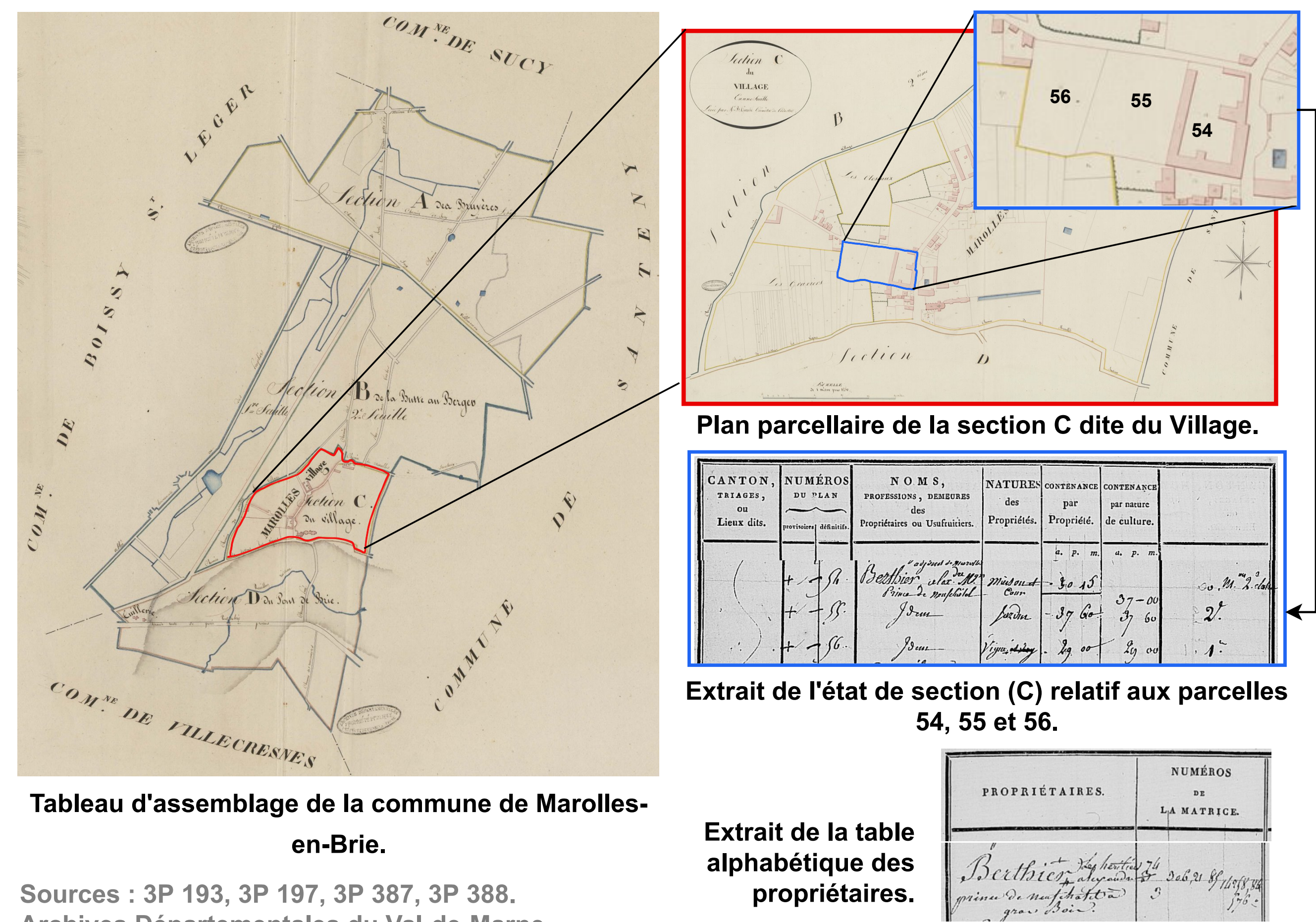
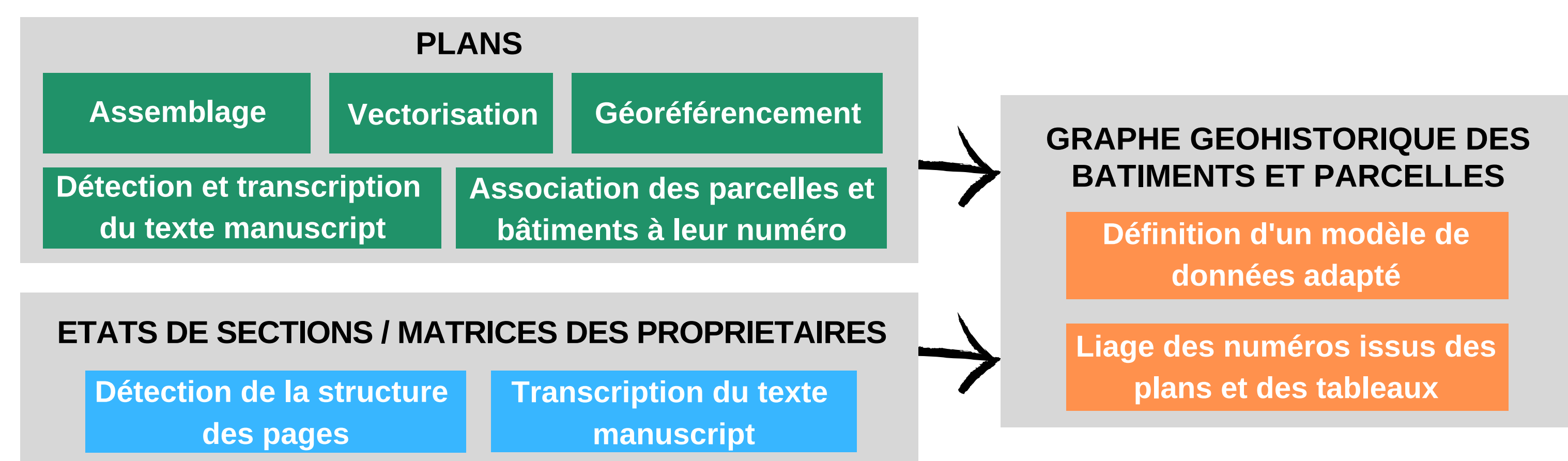


Tableau d'assemblage de la commune de Marolles-en-Brie.
Sources : 3P 193, 3P 197, 3P 387, 3P 388.
Archives Départementales du Val-de-Marne.

Extrait de la table alphabétique des propriétaires.

Chaîne de traitement



Verrous et méthodes envisagées

- **Vectorisation** : Approche de segmentation automatique des plans à l'aide de réseaux de neurones profonds [1] adaptée aux plans du cadastre ancien. Validation participative ;
- **Géoréférencement** : Détection automatique des points de contrôle à partir des angles des objets vectorisés [3] sélectionnés selon leur pertinence. Adaptation à des cartes à très grande échelle ;
- **Détection du texte sur les cartes** : toponymes [2] et des numéros des parcelles manuscrits de typographie et lisibilité variable ;
- **Détection de la structure des tableaux** [4] ;
- **Transcription du texte manuscrit** : Modèles d'HTR (Handwritten Text Recognition) utilisés avec des documents du XIX^{ème} siècle.

Variabilité des sources

La mise en place d'une chaîne de traitement des cadastres anciens nécessite d'établir des méthodes résistantes aux multiples variabilités de ces sources :

Production	Conservation	Numérisation
Qualité géométrique Méthodes, outils Rigueur des levés	Nombreuses modifications des tableaux Documents très manipulés	Résolution numérique variable Numérisation d'une sélection de documents
Étalement temporel des levés Nombreux opérateurs Évolution des règles de production	Conditions de conservation variables Disparition de certains documents	

Informations

- Thèse financée par l'IGN et l'Agence d'innovation de défense (AID)
- Laboratoire : LaSTIG, Univ. Gustave Eiffel, IGN-ENSG
- Encadrement : Nathalie Abadie (IGN), Joseph Chazalon (EPITA), Bertrand Duméniou (EHESS), Julien Perret (IGN)
- Avec les Archives Départementales du Val-de-Marne (94)

Références

- [1] Y. Chen, E. Carlinet, J. Chazalon, C. Mallet, B. Duméniou, and J. Perret. Vectorization of historical maps using deep edge filtering and closed shape extraction. In *ICDAR 2021*, volume 12824 of *Lecture Notes in Computer Science*, pages 510–525, 2021.
- [2] L. Lenc, J. Martínek, J. Baloun, M. Prantl, and P. Král. Historical map toponym extraction for efficient information retrieval. In *Document Analysis Systems*, volume 13237, pages 171–183, 2022.
- [3] J. Luft and J. Schiewe. Automatic content-based georeferencing of historical topographic maps. *Transactions in GIS*, 25(6):2888–2906, 2021.
- [4] J. Ziomek and S. E. Middleton. GloSAT historical measurement table dataset: Enhanced table structure recognition annotation for downstream historical data rescue. In *Historical Document Imaging and Processing 2021*, pages 49–54, 2021.