

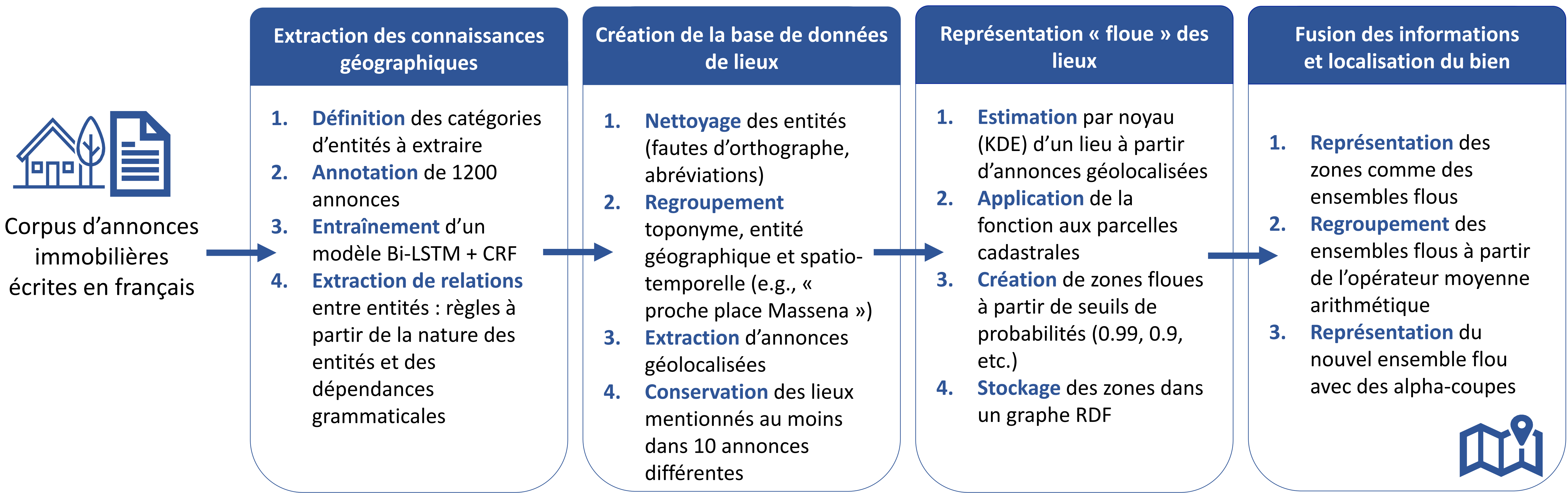
Extraction de connaissances géographiques et localisation incertaine à partir d'un corpus d'annonces immobilières



Introduction

Les annonces immobilières regorgent de **connaissances géographiques et spatiales** qui mettent en avant la localisation d'un bien et qui peuvent enrichir les dictionnaires géographiques (e.g., IGN, OpenStreetMap, GeoNames). Ces connaissances peuvent être des noms de lieux connus, mais aussi des lieux non-officiels (e.g., la « banane » à Cannes) ou non-nommés (e.g., la plage, la gare, etc.). De plus, des **relations spatiales qualitatives** (e.g., proche, non loin, à deux pas) sont souvent employées pour localiser le bien mais sont difficile à modéliser. Enfin, les limites des lieux mentionnés sont souvent incertaines et ne correspondent pas aux limites officielles mais plutôt à la perception que les agents immobilier en ont. Dans cette étude, nous proposons une **méthode pour extraire les connaissances géographiques et spatiales** dans les annonces immobilières écrites en français. Ensuite, nous développons une méthode pour **représenter de manière floue** les lieux et relations spatiales. Enfin, nous utilisons les informations extraites pour retrouver une approximation de la localisation du bien.

Méthode



Application de la méthode

Annonce immobilière → Lieux extraits → Représentation floue → Fusion des lieux

NICE – Appartement situé sur la célèbre Promenade des Anglais à deux pas de la place Massena. La plage est à 5 minutes à pied.

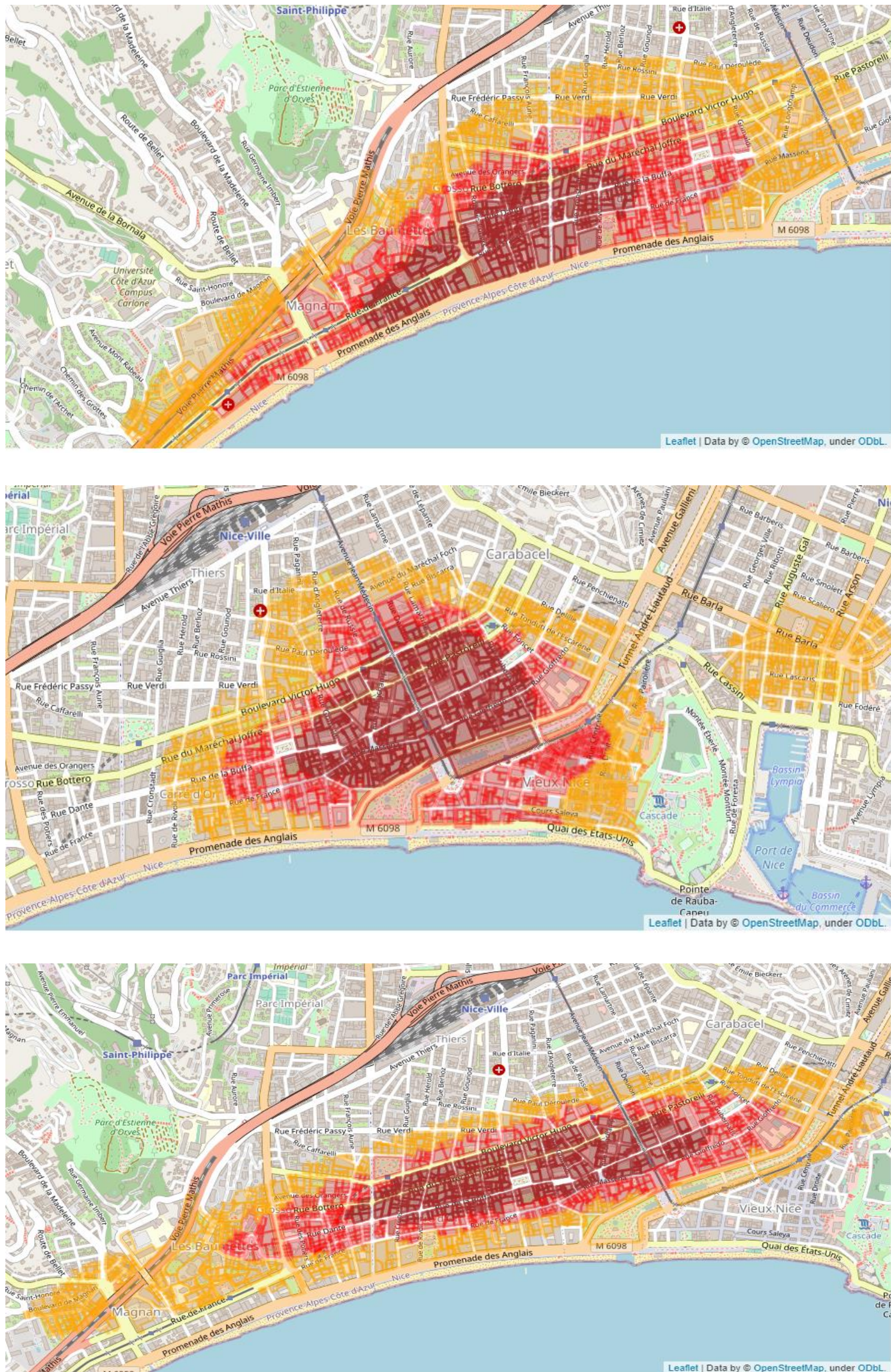
Lieu 1 : « Promenade des Anglais »

Lieu 2 : « A deux pas de la place Massena »

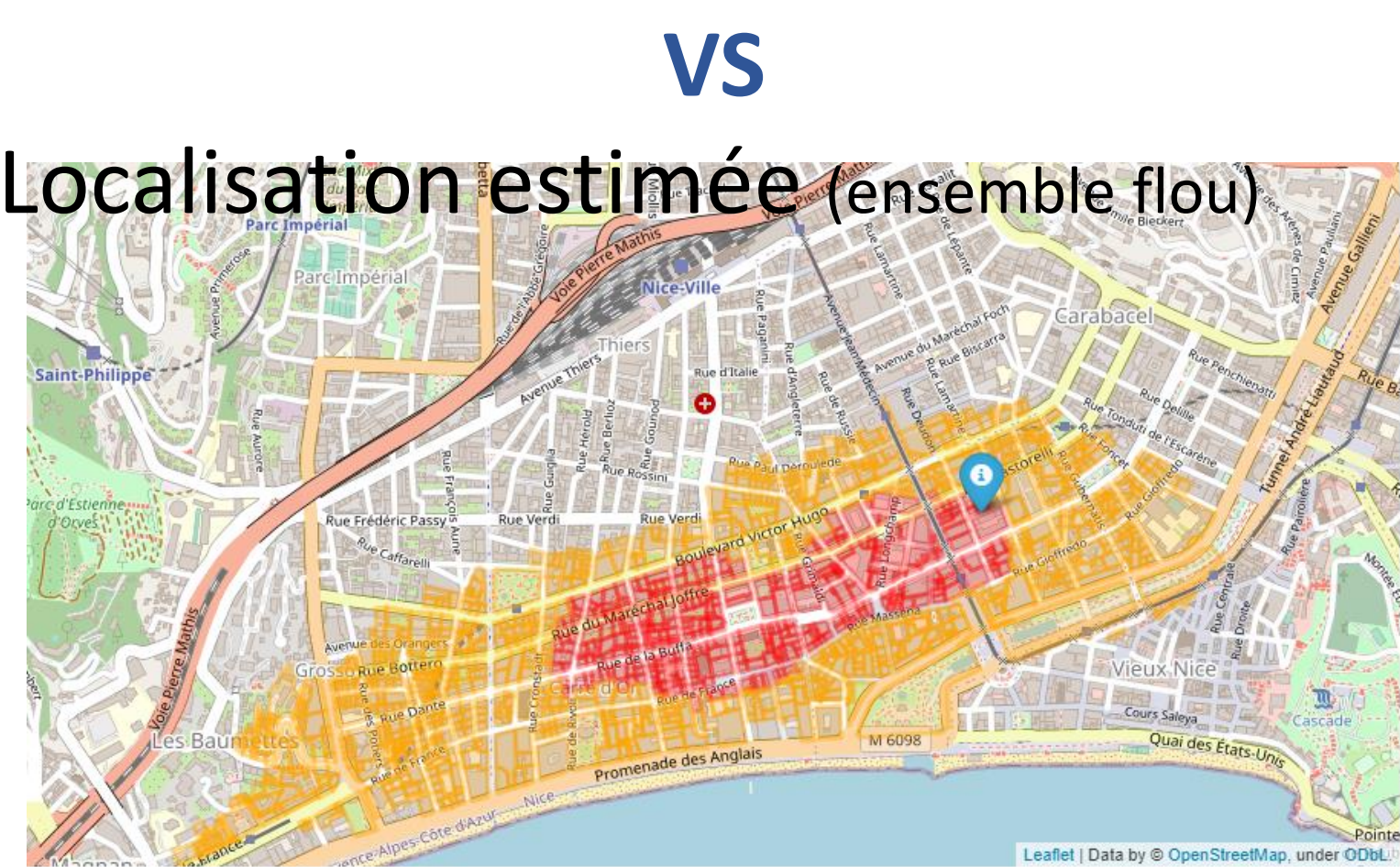
Lieu 3 : « A 5 minutes de la plage »

Catégories d'entités :

- Toponyme
- Entité géographique
- Entité Spatio-temporelle
- Mode de transport



Localisation précise du bien (point bleu)



Conclusion

Dans ce travail, nous avons développé une méthode qui permet d'extraire, représenter et fusionner des informations géographiques et spatiales floues dans un corpus d'annonces immobilières écrites en français. Cette méthode peut s'appliquer à d'autres types de textes. Aussi, l'extraction et la résolution des lieux peuvent contribuer à l'enrichissement des bases de données géographiques.

Contact et Partenaires

Andrea G. B. Tettamanzi, andrea.tettamanzi@univ-cotedazur.fr
Lucie Cadorel, lucie.cadorel@inria.fr

