

OPILAND : identification de la perception des territoires par la fouille de texte

Éric Kergosien^{*,**}, Bernard Laval^{**},
Mathieu Roche^{**}, Maguelonne Teisseire^{**}

* LIRMM, UMR 5506, 161 rue Ada, 34392 Montpellier - France
{prénom.nom}@lirmm.fr

** UMR TETIS, 500 rue Jean-François Breton, 34093 Montpellier - France
{prénom.nom}@teledetection.fr

Résumé. De nombreux travaux ont été réalisés en extraction d'informations et plus particulièrement en fouille de données d'opinions dans des contextes spécifiques tels que les critiques de films, les évaluations de produits commerciaux, les discours électoraux... Dans le cadre du projet SENTERRITOIRE, nous nous posons la question de l'adéquation de ces méthodes pour des documents associés à l'aménagement des territoires. Ces documents renferment différents types d'informations se rapportant à des acteurs, des opinions, des informations géographiques, et tout autre aspect lié plus généralement à la notion de territoire. Cependant, il est extrêmement difficile d'identifier puis de lier les opinions à ces informations. Après avoir souligné les limites des propositions actuelles et les verrous soulevés par les données textuelles associées, nous proposons la méthode semi-automatique nommée OPILAND (OPinion mIning from LAND-use planning documents) combinant une chaîne de Traitement Automatique du Langage Naturel et des techniques de Fouilles de Textes pour (1) détecter les entités nommées de type lieu et organisation, (2) construire un vocabulaire d'opinions relatif au domaine d'application, et (3) identifier les opinions relatives aux entités nommées traitées. Les expérimentations sont menées sur des données du bassin de Thau (France), puis appliquées sur trois corpus relatifs à d'autres domaines afin de mettre en avant la généralité de notre approche.

1 Introduction

Au-delà de sa stricte définition d'entité administrative et politique, le territoire, selon Guy Di Méo témoigne d'une « appropriation à la fois économique, idéologique et politique de l'espace par des groupes qui se donnent une représentation particulière d'eux-mêmes, de leur histoire, de leur singularité » (Di-Méo, 1998). Dans ce contexte éminemment subjectif, la caractérisation et la compréhension des perceptions d'un même territoire par les différents acteurs est difficile, mais néanmoins particulièrement intéressante dans une perspective d'aménagement du territoire (Derungs et Purves, 2013) et de politique publique territoriale. La notion de territoire, et plus spécifiquement d'aménagement du territoire, fait référence à différents concepts