



HAL
open science

Outils informatiques et terminologie pour la valorisation du territoire : L'enrichissement d'une ontologie sur les énergies renouvelables en français et en italien

Virginia Carrella, Charlotte Darricades, Silvia Domenica Zollo, Christian
Sallaberry, Sébastien Laborie

► To cite this version:

Virginia Carrella, Charlotte Darricades, Silvia Domenica Zollo, Christian Sallaberry, Sébastien Laborie. Outils informatiques et terminologie pour la valorisation du territoire : L'enrichissement d'une ontologie sur les énergies renouvelables en français et en italien. XX Jornada Científica REALITER y XXIII Jornada AETER, Oct 2024, San Millán de la Cogolla, Espagne. 10.5281/zenodo.11210074 . hal-04752677

HAL Id: hal-04752677

<https://hal.science/hal-04752677v1>

Submitted on 24 Oct 2024

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.

Outils informatiques et terminologie pour la valorisation du territoire : L'enrichissement d'une ontologie sur les énergies renouvelables en français et en italien

Virginia Carrella ¹, Charlotte Darricades ^{2,3,4}, Silvia Domenica Zollo ¹, Christian Sallaberry ³,
Sébastien Laborie ³

¹ Università degli Studi di Napoli Parthenope
virginia.carrella001@studenti.uniparthenope.it, silvia.zollo@uniparthenope.it

² Université de Lille, GERiiCO
charlotte.darricades@univ-lille.fr

³ Université de Pau et des Pays de l'Adour, E2S UPPA, LIUPPA
christian.sallaberry@univ-pau.fr, sebastien.laborie@univ-pau.fr

⁴ Pôle d'Études et de Recherche de Lacq, TotalEnergies SE, BP 47, 64170 Lacq
charlotte.darricades@totalenergies.com

Mots-clés : *valorisation du territoire, ontologie, terminologie, médiation, communication, énergies renouvelables, énergies marines, technologies d'automatisation*

Dans cette proposition nous présentons les objectifs et les phases de mise en œuvre d'une étude pilote, qui s'inscrit dans le cadre d'une collaboration scientifique entre les équipes de recherche en informatique de l'Université de Pau et des Pays de l'Adour et en linguistique française de l'Université de Naples Parthenope.

Le premier projet, en Sciences de l'Information et de la Communication (SIC) initié à l'Université de Lille (GERiiCO), a développé une ontologie de domaine (Rawsthorne *et al.* 2022), celui des énergies renouvelables, selon la méthode SAMOD-PLUS (Darricades *et al.* 2023), une adaptation de la méthode SAMOD (Peroni 2016) : cette approche commence par la création d'un sociogramme détaillé de l'organisation de l'entreprise et la collecte d'informations via des entretiens semi-directifs. Le corpus est construit à partir des publications scientifiques des chercheurs, puis il est analysé (Sallaberry 2014) pour structurer une base de connaissances ciblant les thématiques énergétiques chez l'entreprise *TotalEnergies*. L'ontologie créée – sur les énergies renouvelables – vise à améliorer la compréhension et l'interprétation des informations, avec une expérimentation finale réalisée avec les chercheurs et les communicants pour affiner cette base grâce à leurs retours.

L'autre projet en linguistique française intitulé *Littératie océanique* auprès de l'Université de Naples Parthenope a pour but de collecter, encoder et analyser les lexiques et les discours technico-scientifiques du domaine de la biologie marine, à l'aide des approches et des outils de la linguistique de corpus et de la terminologie (L'Homme 2020) : notamment, il vise à promouvoir la protection et la valorisation du patrimoine naturel marin à travers la création

de ressources linguistiques (i.e. fiches terminographiques, glossaires, etc.) pour la vulgarisation des savoirs scientifiques et des terminologies en question (Zollo 2024a).

Malgré la divergence des approches et du cadre théorique-méthodologique, les projets visent les mêmes objectifs, y compris la médiation et la diffusion des terminologies (Grimaldi 2017) en vue de leur valorisation sur le territoire (Lino *et al.* 2020) : cela a engendré la collaboration, qui a abouti à la réalisation d'une étude pilote, ayant pour but de compléter l'ontologie initiale sur les énergies renouvelables au regard d'autres corpus, langues et domaines, à l'aide des nouveaux outils informatiques.

En effet, à cette fin, la collaboration a conduit à la réalisation des phases suivantes :

1. L'identification d'un nouveau sous domaine, motivé par une lacune dans l'ontologie initiale : l'étude des termes – en français et en italien – sur la sauvegarde et la valorisation du patrimoine naturel marin en cours à Naples depuis 2022 (Zollo 2024a) a donné l'idée de s'attarder sur le sous-domaine des énergies marines renouvelables, un secteur en croissance et de grande importance pour la durabilité environnementale ;
2. L'expérimentation de nouveaux outils informatiques pour la collecte de textes et l'extraction automatique des connaissances : plus précisément, au laboratoire LIUPPA – dans le cadre de la collaboration entre les deux universités et sur la base des critères proposés pour la constitution du corpus *ZooCor* (Zollo 2024b) – a été développé *Corpus Builder*, une application permettant d'automatiser les processus de collecte des textes (français et italien), d'extraction de mots-clés et de résumés appartenant au sous-domaine des énergies marines renouvelables ;
3. La création d'un lexique spécialisé : la classification et l'introduction des nouveaux termes extraits – et des équivalents dans d'autres langues – dans l'ontologie initiale sur les énergies renouvelables.

Ainsi, nous voulons montrer comment, grâce à la comparaison interdisciplinaire entre les deux projets de recherche, il est possible d'utiliser des outils informatiques pour l'enrichissement d'une ontologie de domaine liée à un domaine d'activités spécialisées (Slodzian 1995 ; Zanolà 2018) et l'implémentation de réseaux conceptuels existants créés avec *Protégé*.

Référence bibliographique :

Darricades Charlotte, Laborie Sébastien, Sallaberry Christian, Kergosien Eric, De La Broise Patrice (2023). "Identifier les traces pertinentes dans la documentation scientifique d'une entreprise à des fins de communication institutionnelle. Le cas de la Compagnie TotalEnergies.", Inforsid, La Rochelle ;

Grimaldi Claudio (2017), *Discours et terminologie dans la presse scientifique française (1699–1740). La construction des lexiques de la botanique et de la chimie*, Bruxelles, Peter Lang ;

L'Homme Marie-Claude (2020), *La terminologie : principes et techniques*, Montréal, Presses de l'Université de Montréal ;

Lino Teresa, Fangour Robalo Karima, Barbosa Sílvia, Heitor Olga, Ferreira Fátima, Lino Catarina (2020), « Terminologie vulgarisée des énergies renouvelables », en Xavier Blanco Escoda (ed.), *La lexicographie en tant que méthodologie de recherche en linguistique, Langue(s) & Parole : revista de filología francesa y románica*, CIPA, vol. 5, Mons, pp. 231-253;

Peroni Silvio (2016), "A Simplified Agile Methodology for Ontology Development", Owledore;

Rawsthorne Helen Mair ; Abadie, Nathalie ; Kergosien, Eric ; Duchêne, Cécile ; Saux, Eric (2022), « ATLANTIS : Une ontologie pour représenter les Instructions nautiques. Journées Francophones d'Ingénierie des Connaissances (IC) », Plate-Forme Intelligence Artificielle, Saint-Etienne, pp. 154-163 ;

Sallaberry Christian (2014), *Recherche d'information géographique dans des corpus textuels*, ISTE Group ;

Slodzian Monique (1995), « Comment revisiter la doctrine terminologique aujourd'hui ? » La Banque des mots : Terminologie et Intelligence Artificielle, vol. 7, pp. 11-18 ;

Zanola Maria Teresa (2018), *Che cos'è la terminologia*, Roma, Carocci;

Zollo Silvia Domenica (2024a), « Lexiques et corpus au service de la littérature océanique : propriétés et relations lexicales dans le domaine de la faune marine », *Studia Universitatis Babeş-Bolyai-Philologia*, pp. 10-29 ;

Zollo Silvia Domenica (2024b), *ZooCor Corpus (fr-it) (First version)* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.11210074>.