

TENET, un outil pour construire des ontologies OWL à partir de textes en langue naturelle

David Rouquet*, Aurélien Lamercherie**, Valérie Bellynck***
Christian Boitet***, Vincent Berment****, Guillaume de Malézieux****,****

*Tétras Libre
**UGA, LIG, GETALP, Grenoble INP
***UGA, LIG, GETALP
****CS Group

Résumé. Cet article présente la démonstration de TENET, un outil générique et Open Source pour construire des ontologies OWL à partir de textes écrits en langue naturelle. Nous illustrons les différentes étapes de la chaîne de traitement, notamment le passage par une représentation sémantique pivot des énoncés (sérialisée en RDF), ainsi que le paramétrage et la réalisation du processus d'extraction et de construction de l'ontologie cible. Enfin, nous montrons comment l'ontologie cible peut être exploitée, dans un contexte industriel, avec un exemple de vérification automatique d'exigences système.

1 Introduction

Cet article présente TENET¹, une plateforme générique pour construire des ontologies OWL à partir de textes écrits en langue naturelle (LN). Le code source est disponible sous licence CeCILL-B dans un dépôt Gitlab dédié². Cette application est basée sur les standards du Web Sémantique du W3C³ (RDF, OWL, SPARQL, SHACL). Elle implémente une *enconversion* des textes en UNL⁴, une représentation sémantique pivot et un procédé d'extraction fondé sur le concept de transduction sémantique compositionnelle (Lamercherie (2021)).

Le développement de ce prototype s'inscrit dans le cadre du projet RAPID UNSEL⁵, financé par la DGA⁶ de 2019 à 2021. Le but du projet est de fournir des outils pour accompagner la spécification de systèmes complexes (par exemple, un système de communications sol-air pour un aéroport ou un système de freinage d'urgence). Précisément, nous développons des processus permettant de raisonner sur un ensemble de spécifications, afin de détecter des incomplétudes ou des incohérences. Les énoncés sont traités en suivant une approche multilingue et interactive qui s'appuie sur une représentation pivot interlingue.

1. TENET : Tool for Extraction using Net Extension by (semantic) Transduction
2. <https://gitlab.tetras-libre.fr/unl/tenet>
3. <https://www.w3.org/standards/semanticweb/>
4. <http://www.unlweb.net/wiki/images/a/ab/Spec33.pdf>
5. UNsel : Universal Networking system engineering Language
6. Direction Générale de l'Armement, <https://www.defense.gouv.fr/dga>