

Master thesis : Knowledge Graph Construction to Facilitate Chemical Compound Hazard Assessment in the TOXIN Project

Auteur : Vrijens, Guillaume

Promoteur(s) : Debruyne, Christophe

Faculté : Faculté des Sciences appliquées

Diplôme : Master : ingénieur civil en science des données, à finalité spécialisée

Année académique : 2022-2023

URI/URL : <http://hdl.handle.net/2268.2/16763>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

Knowledge Graph Construction to Facilitate Chemical Compound Hazard Assessment in the TOXIN Project

Master's thesis completed in order to obtain the degree of
Master of Science in Data Science and Engineering

Abstract

This master thesis presents a method for integrating multiple data sources from the field of toxicology into a knowledge graph and linking it with the TOXIN knowledge graph to facilitate the hazard assessment of new compounds. The proposed method uses a hybrid approach, combining an ontology and Linked Data to capture the granularity of the toxicological domain and provide a consistent representation while maintaining the flexibility of Linked Data. The ontology used in the method is the ToXic Process Ontology (TXPO), which offers a structured and reliable representation of the relationships between toxicological processes. The method also incorporates the use of named graphs and provenance information to store different opinions on data and track the integration of different sources. The feasibility and utility of the proposed method for building the knowledge graph are demonstrated through the development of a prototype, the TOXIN enriched knowledge graph (TEKG). Finally, this project illustrates the potential value and usefulness of a knowledge graph such as TEKG for improving access to relevant information, offering a satisfactory representation of the toxicological domain and supporting domain-specific tagging mechanisms.