

Posudek vedoucího diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce: Bc. Martin Gora
Název práce: Searching classes in the Wikidata ontology
Rok odevzdání: 2024
Studijní program: Informatika
Studijní obor: Softwarové a datové inženýrství

Autor posudku: doc. Mgr. Martin Nečaský, Ph.D.
Pracoviště: KSI

Diplomová práce se věnuje využití ontologie vytvářené v komunitním projektu Wikidata v rámci softwarového nástroje Dataspecer. Konkrétně navrhuje metody pro vyhledávání tříd v ontologii Wikidata a implementuje je do prostředí nástroje. **Těžké** na tomto problému **je, jakým způsobem bychom měli v ontologii Wikidata vyhledávat vzhledem k její velikosti a také nekonzistentnosti**. V předložené diplomové práci se autorovi podařilo:

- Analyzovat Wikidata a ontologii obsaženou v tomto znalostním grafu a navrhnout metodu extrakce použitelné ontologie.
- Navrhnout, implementovat a otestovat nástroj, který metodu extrakce ontologie implementuje.
- Nad extrahovanou ontologií navrhnul metodu pro vyhledávání tříd. Metoda vychází z existujících metod „text retrieval“ a „retrieval in knowledge graphs“, včetně metod založených na „pre-trained language models“.
- Metodu implementoval do prostředí Dataspecer a otestoval, včetně uživatelského testování.

Textová i softwarová část práce je zpracována kvalitně. Práce splňuje zadání. Autor v práci provedl všechny nutné kroky z pohledu softwarového inženýrství. Z práce společně s vedoucím práce zpracovává vědecký článek, který plánujeme během roku 2025 zaslat k recenznímu řízení ve vědeckém časopisu.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

Datum 3.2.2025