

Využitie LLM pre analýzu právnych dokumentov

Prezentácia BP1

Autor: Samuel Bagín

Vedúci bakalárskej práce: Ing. Marek Vančo, PhD.

Zadanie bakalárskej práce

Cieľom projektu je vytvoriť AI systém na automatickú extrakciu a prepojenie informácií z právnych textov do znalostného grafu s pokročilým sémantickým vyhľadávaním. Systém kombinuje grafovú databázu s vektorovým úložiskom pre hybridné vyhľadávanie, ktoré umožňuje používateľom klášť otázky v prirodzenom jazyku a získavať presné odpovede na základe štruktúrovaných vzťahov aj sémantickej podobnosti.

Overené podobné riešenia

Blackstone – britské riešenie, využívajúce strojové učenie

Problém

Problematika halucinácie LLM pri pýtaní sa otázok na legálne pojmy.

Príklad: daň z pridanej hodnoty – daň z príjmu

LLM s možnosťou vyhľadávania na internete, nájdu nerelevantné a neaktuálne dokumenty a informácie.

O projekte

1

RecursiveCharacterTextSplitter

LLMGraphTransformer

extrahovanie entít a vzťahov

chunk nodes

vloženie do databáz

2

3-5 variácií otázky

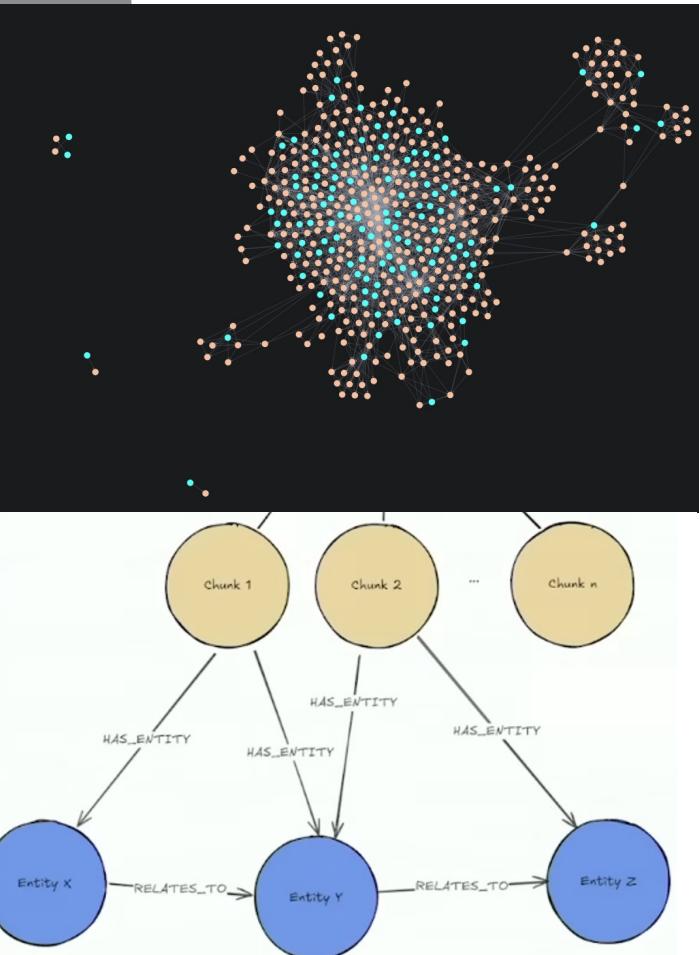
podmet, predmet, prísudok

Cypher query → KG

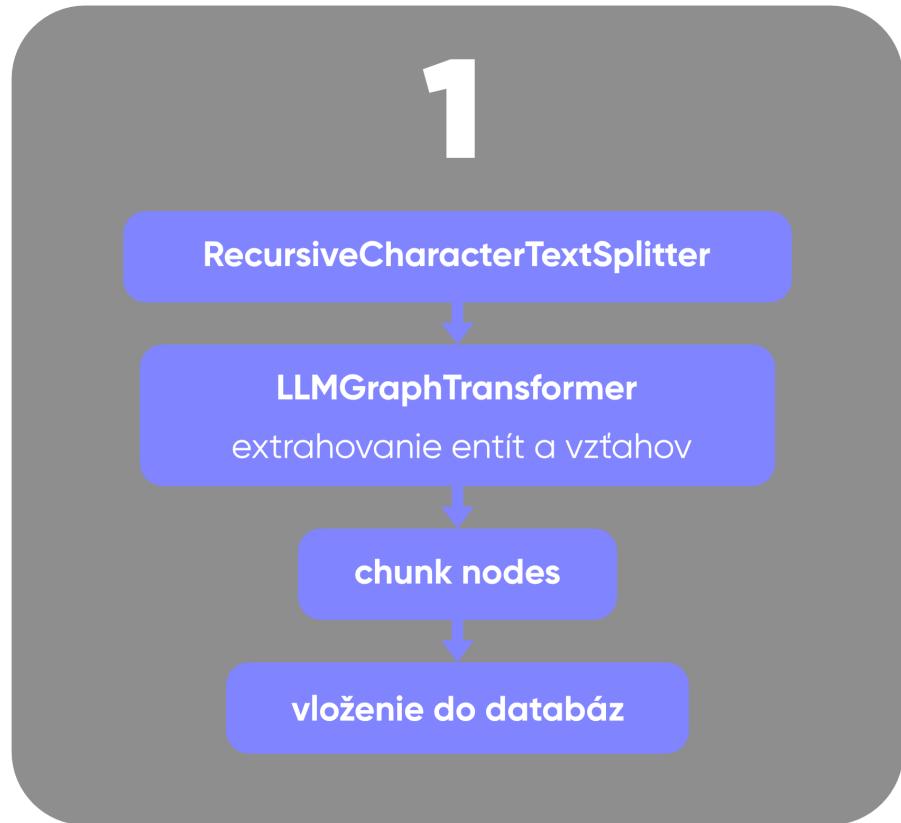
nájdenie chunkov, kde sa nachádzajú vrcholy z KG

UQ + VQs + KGs + chunky + SP

Neo4j, LangChain, OpenAI



O projekte



sémantické vyhľadávanie:



Plán na letný semester

- Implementovanie multi-hop vyhľadávania
 - Jožkov bratranec má rád hot-dog.
- Zobrazovanie nájdených znalostných grafov
 - Vizualizácia pomocou PyVis alebo D3.js
- Vytvorenie používateľského rozhrania
 - Webové rozhranie alebo Jupyter Notebook
- Optimalizácia promptov
 - Zjednodušenie, skrátenie a zrozumiteľnejšie prompty

Ďakujem za pozornosť



SLOVENSKÁ TECHNICKÁ
UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY
A INFORMATIKY