D17 - Metóda vyhľadávania middleware

Projekt financovaný Európskou úniou Next GenerationEU prostredníctvom Plánu obnovy a odolnosti SR v rámci projektu č. 09I05-03-V02-00049.

Financované Európskou úniou NextGenerationEU
PLÁN OBNOVY
ÚRAD PODPREDSEDU VLÁDY SLOVENSKEJ REPUBLIKY PRE PLÁN OBNOVY A ZNALOSTNÚ EKONOMIKU
VAIA VÝSKUMNÁ A INOVAČNÁ AUTORITA
VÝSKUMNÁ AGENTÚRA

ÚVOD

Projekt „**Automatizácia analýzy právnych textov na základe strojového učenia**“ (ďalej len „**ALTAML**“) predstavuje kľúčovú iniciatívu zameranú na integráciu inovatívnych prístupov v oblasti spracovania údajov a následnej analýzy, so zameraní na oblasť práva, ktorá zahŕňa aj právo informačných a komunikačných technológií. Cieľom tohto projektu je vyvinúť a overiť účinné metódy automatizovanej analýzy právnych textov pomocou techník strojového učenia. To zahŕňa najmä vývoj nástrojov, ktoré uľahčujú spracovanie a analýzu veľkých objemov údajov vo forme rôznych právnych dokumentov, ako aj extrakciu relevantných informácií (atribútov) z takýchto dokumentov vrátane identifikácie kľúčových pojmov, odkazov na právne predpisy a ďalších atribútov.

Cieľom projektu ALTAML je teda prispieť k efektívnejšiemu prístupu k právnym informáciám a k zrýchleniu právnych procesov, čím sa zabezpečí vyššia miera právnej istoty a zlepší sa dostupnosť právnych textov, výsledkov ich analýzy a relevantných právnych informácií pre odbornú aj širokú verejnosť.

V rámci pracovného balíka KPB3 bola vyvinutá **D17 – Metóda vyhľadávania middleware**, ktorej úlohou je podporiť používateľa pri formulovaní vyhľadávacích dopytov nad súdnymi rozhodnutiami a právnymi predpismi.

Pri vyhľadávaní právnych dokumentov používatelia často formulujú dopyty pomocou konkrétnych právnych alebo skutkových pojmov a odkazov na právne ustanovenia. Problém však nastáva v situáciách, keď používateľ pozná iba všeobecnú tému, ktorú chce vyhľadávať, ale pri zadávaní dopytu mu nemusia hneď napadnúť všetky súvisiace pojmy a právne ustanovenia, ktoré by mohli viesť k relevantným výsledkom. Vstupom metódy preto môže byť aj všeobecne formulovaná téma, napríklad „kyberkriminalita“.

Metóda pre takýto vstup najprv pomocou sémantického vyhľadávania vyhľadá významovo súvisiace časti súdnych rozhodnutí a právne predpisy. Zo sémanticky vyhľadaných častí súdnych rozhodnutí následne navrhne vyhľadávacie pojmy, ktoré sa v týchto rozhodnutiach priamo vyskytujú, napríklad „IP adresa“, „počítačová sieť“ alebo „internetová stránka“. Z vyhľadaných právnych predpisov zároveň vyberie relevantné právne ustanovenia, napríklad § 360b Trestného zákona týkajúci sa nebezpečného elektronického obťažovania alebo § 247a Trestného zákona upravujúci neoprávnený zásah do počítačového systému.

Výstupom metódy je štruktúrovaná odpoveď obsahujúca odporúčané vyhľadávacie pojmy a identifikátory relevantných právnych ustanovení. Používateľ si z navrhnutých hodnôt môže vybrať tie pojmy a právne ustanovenia, ktoré chce použiť pri ďalšom vyhľadávaní.

Metóda bola implementovaná ako spustiteľný prototyp nad množinou súdnych rozhodnutí a právnych predpisov.

1. Konfigurácia vyhľadávacích indexov

Sprostredkovacia vyhľadávacia metóda využíva samostatné Elasticsearch indexy pre súdne rozhodnutia a právne ustanovenia. Oba indexy sú pripravené tak, aby podporovali sémantické vyhľadávanie pomocou vektorových reprezentácií textu. Cieľom týchto indexov nie je pokryť všetky atribútové filtre definované v dopytovacom modeli D16, ale poskytnúť vstupné dokumenty pre sprostredkovaciu metódu, ktorá z nich následne generuje odporúčané vyhľadávacie pojmy a relevantné právne ustanovenia.

Na vytváranie vektorových reprezentácií je použitý model *sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2*, ktorý vytvára 384-rozmerné embeddingy. Rovnaký model sa používa pri indexovaní dokumentov aj pri spracovaní používateľskej témy, aby boli dokumenty a vstupný dopyt reprezentované v rovnakom vektorovom priestore. V nasledujúcich podkapitolách je preto opísaná konfigurácia indexu súdnych rozhodnutí a indexu právnych ustanovení.

1.1. Index súdnych rozhodnutí

Súdne rozhodnutia sú v indexe uložené na úrovni menších textových častí, ďalej označovaných ako chunky. Celé rozhodnutie môže byť rozsiahle, preto jedna vektorová reprezentácia celého dokumentu nemusí dostatočne presne zachytiť význam jeho jednotlivých častí.

Chunkovanie súdnych rozhodnutí prebieha nad hlavnými časťami dokumentu identifikovanými pri predspracovaní, konkrétne nad úvodom, výrokom, odôvodnením a poučením. Text každej časti sa najprv rozdelí na odseky textu. Ak je niektorý odsek dlhší ako maximálna veľkosť chunku, rozdelí sa ďalej podľa slov. Následne sa kratšie textové úseky skladajú do chunkov tak, aby neprekročili stanovený maximálny rozsah. V prototype je maximálna veľkosť chunku nastavená na 400 znakov. Ak má posledný chunk menej ako 100 znakov, spojí sa s predchádzajúcim chunkom, pokiaľ výsledný text neprekročí maximálnu veľkosť chunku.

Každý chunk uložený v indexe obsahuje identifikátor pôvodného rozhodnutia, označenie časti rozhodnutia (`section`), text chunku a jeho vektorovú reprezentáciu (embedding). Pole `section` umožňuje rozlíšiť, či chunk pochádza napríklad z úvodu, výroku, odôvodnenia alebo poučenia. Pri vyhľadávaní sa používateľská téma prevedie na embedding a porovnáva sa s embeddingmi jednotlivých chunkov.

Uvedené informácie sa premietajú aj do štruktúry dokumentov uložených v Elasticsearch indexe. Časť `mappings` určuje názvy polí, ich dátové typy a spôsob ich využitia pri vyhľadávaní. Detailnejší opis nastavení Elasticsearch indexov a použitých dátových typov je uvedený vo výstupe D16 – Dopytovací model pre vyhľadávanie právnych textov.

Konfiguráciu časti `mappings` pre index súdnych rozhodnutí obsahuje Ukážka 1.

```
{
  "mappings": {
    "properties": {
      "doc_id": {
        "type": "keyword"
      },
      "chunk_id": {
        "type": "integer"
```

```
        },
        "section": {
          "type": "keyword"
        },
        "text": {
          "type": "text"
        },
        "embedding": {
          "type": "dense_vector",
          "dims": 384,
          "index": true,
          "similarity": "cosine"
        }
      }
    }
  }
```

Ukážka 1 - Konfigurácia indexu súdnych rozhodnutí.

Tabuľka 1 uvádza základnú štruktúru dokumentu uloženého v indexe súdnych rozhodnutí.

Tabuľka 1 - Atribúty indexu súdnych rozhodnutí.

Atribút	**Typ**	**Význam**
`doc_id`	`keyword`	identifikátor pôvodného súdneho rozhodnutia
`chunk_id`	`integer`	poradové číslo textového chunku v rámci rozhodnutia
`section`	`keyword`	časť rozhodnutia, z ktorej chunk pochádza
`text`	`text`	text chunku použitý ako kontext pre jazykový model
`embedding`	`dense_vector`	vektorová reprezentácia chunku použitá pri sémantickom vyhľadávaní

1.2. Index právnych ustanovení

Právne ustanovenia sú v indexe uložené na rôznych úrovniach podrobnosti, konkrétne na úrovni paragrafu, odseku alebo písmena. Na rozdiel od súdnych rozhodnutí sa texty právnych ustanovení ďalej nedelia na menšie chunky, pretože jednotlivé ustanovenia sú spravidla dostatočne krátke na priame vytvorenie jednej vektorovej reprezentácie.

Z pohľadu sprostredkovacej metódy sú pri právnom ustanovení dôležité tri údaje: identifikátor `doc_id`, text ustanovenia a jeho vektorová reprezentácia. Identifikátor sa poskytuje jazykovému modelu a zároveň tvorí hodnotu vracanú v poli `legal_provisions`. Text slúži ako kontext pre jazykový model a embedding sa používa pri sémantickom vyhľadávaní.

Ukážka 2 obsahuje konfiguráciu časti *`mappings`* pre index právnych ustanovení.

```
{
```

```
                "mappings": {
                  "properties": {
                    "doc_id": {
                      "type": "keyword"
                    },
                    "text": {
                      "type": "text"
                    },
                    "embedding": {
                      "type": "dense_vector",
                      "dims": 384,
                      "index": true,
                      "similarity": "cosine"
                    }
                  }
                }
              }
```

Ukážka 2 - Konfigurácia indexu právnych ustanovení.

Tabuľka 2 uvádza základnú štruktúru dokumentu uloženého v indexe právnych predpisov.

Tabuľka 2 - Atribúty indexu právnych ustanovení.

Atribút	Typ	Význam
`doc_id`	`keyword`	identifikátor právneho ustanovenia
`text`	`text`	text právneho ustanovenia použitý pri vytváraní embeddingu a ako kontext pre jazykový model
`embedding`	`dense_vector`	vektorová reprezentácia právneho ustanovenia použitá pri sémantickom vyhľadávaní

1.2.1. Hierarchické uloženie textu právnych ustanovení

Právne ustanovenia sú v indexe uložené na viacerých úrovniach podrobnosti. Rovnaká časť právneho textu teda môže byť reprezentovaná všeobecnejším aj konkrétnejším dokumentom. Všeobecnejší dokument obsahuje širší textový kontext, zatiaľ čo konkrétnejší dokument obsahuje iba príslušnú časť ustanovenia a jej nadradený kontext.

Na najvyššej úrovni môže byť uložený celý paragraf vrátane všetkých jeho odsekov a písmen:

```
{ "id": "/SK/ZZ/2005/300/20260701.html#paragraf-10",
  "text": "§ 10 Názov paragrafu alebo úvodná veta paragrafu.
            (1) Text prvého odseku.
            (2) Text druhého odseku.
                  a) Text písmena a).
                  b) Text písmena b)." }
```

Na nižšej úrovni môže byť uložený konkrétny odsek. V tomto prípade dokument obsahuje názov alebo úvodnú vetu paragrafu a text príslušného odseku vrátane jeho písmen:

{ "id": "/SK/ZZ/2005/300/20260701.html#paragraf-10.**odsek-2**",
"text": "**§ 10** Názov paragrafu alebo úvodná veta paragrafu.
(2) Text druhého odseku.
a) Text písmena a).
b) Text písmena b)." }

Ak sa odsek člení na písmená, môže byť uložený aj dokument pre konkrétne písmeno. V tomto prípade obsahuje nadradený kontext paragrafu a odseku, ale už neobsahuje ostatné písmená na rovnakej úrovni:

{ "id": "/SK/ZZ/2005/300/20260701.html#paragraf-10.odsek-2.**pismeno-a**",
"text": "**§ 10** Názov paragrafu alebo úvodná veta paragrafu.
(2) Text druhého odseku.
a) Text písmena a)."}

Tento spôsob ukladania je potrebný preto, aby aj konkrétnejšie časti právneho ustanovenia zostali významovo úplné. Právny text je tak v indexe reprezentovaný na viacerých úrovniach: všeobecnejšie dokumenty zachytávajú širší kontext ustanovenia, zatiaľ čo konkrétnejšie dokumenty reprezentujú užšiu a presnejšie vymedzenú časť právneho textu. Pri sémantickom vyhľadávaní to umožňuje nájsť buď širšie súvisiace ustanovenie, alebo detailnejšiu časť ustanovenia podľa významu používateľskej témy.

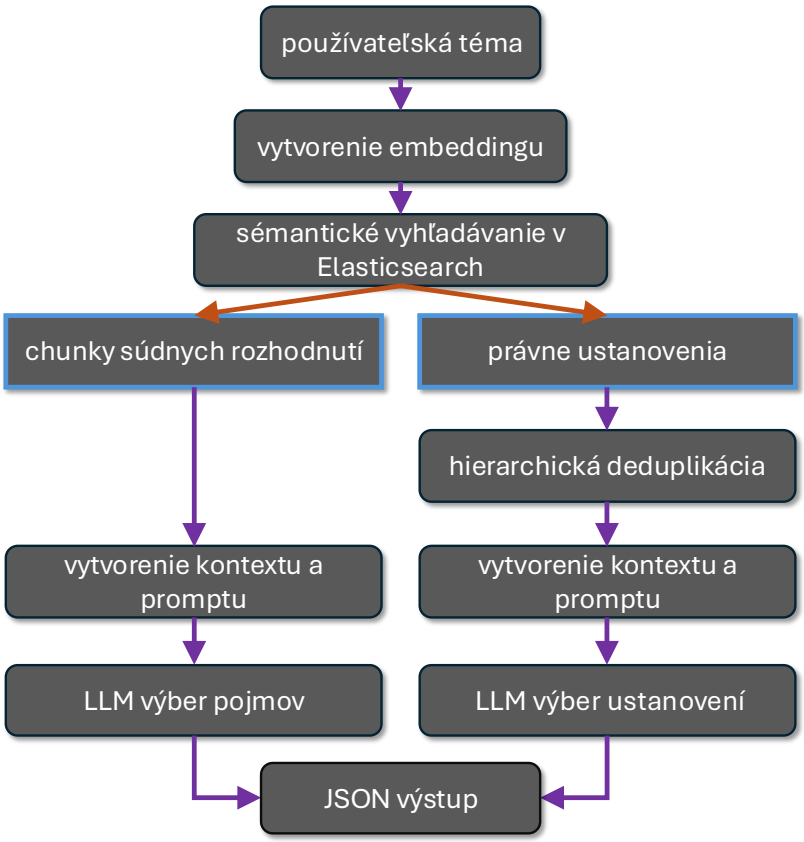
Ak by bolo písmeno uložené iba samostatne, bez názvu paragrafu, úvodnej vety alebo textu príslušného odseku, jeho význam by nemusel byť jednoznačný. Napríklad samotné písmeno:

{ "id": "/SK/ZZ/2005/300/20260701.html#paragraf-10.odsek-2.pismeno-a",
"text": "**a)** Text písmena a)." }

2. Postup sprostredkovacej metódy

Sprostredkovacia metóda spracováva vstup používateľa vo forme témy alebo stručne formulovaného dopytu. Výstupom metódy nie je finálny Elasticsearch DSL dopyt, ale štruktúrovaný návrh hodnôt, ktoré môže používateľ následne využiť pri ďalšom vyhľadávaní. Metóda vracia dve skupiny výsledkov: odporúčané vyhľadávacie pojmy a identifikátory relevantných právnych ustanovení.

Celkový postup sprostredkovacej metódy zobrazuje Obrázok 1. Metóda pozostáva z viacerých krokov. Najprv sa používateľský vstup prevedie na embedding. Ten sa použije na sémantické vyhľadanie súvisiacich chunkov súdnych rozhodnutí a právnych ustanovení. Z vyhľadaných výsledkov sa následne vytvoria textové kontexty, ktoré vstupujú do dvoch samostatných volaní jazykového modelu. Prvé volanie generuje odporúčané vyhľadávacie pojmy a druhé volanie vyberá relevantné právne ustanovenia. Výsledky oboch volaní sa nakoniec spoja do jedného JSON výstupu.



Obrázok 1 - Postup sprostredkovacej vyhľadávacej metódy.

2.1. Vytvorenie embeddingu témy zo vstupu

Vstupná téma používateľa sa najprv prevedie na vektorovú reprezentáciu pomocou rovnakého embedding modelu, aký bol použitý pri indexovaní súdnych rozhodnutí a právnych ustanovení. V prototype je použitý model *sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2,* ktorý vytvára 384-rozmerné vektorové reprezentácie.

Ukážka 3 demonštruje vytvorenie embeddingu pre vstupnú tému „kyberkriminalita".

```
from sentence_transformers import SentenceTransformer

model = SentenceTransformer("sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2")
query_vector = model.encode("kyberkriminalita").tolist()
```

Ukážka 3 - Vytvorenie embeddingu témy zo vstupu.

2.2. Sémantické vyhľadanie súvisiacich dokumentov

Následne sa rovnaký embedding (**`query_vector`**) použije na vyhľadávanie v dvoch indexoch: v indexe súdnych rozhodnutí a v indexe právnych ustanovení. V oboch prípadoch ide o knn vyhľadávanie, teda vyhľadávanie najbližších vektorov k vstupnému vektoru používateľskej témy.

Pri súdnych rozhodnutiach sa embedding (**`query_vector`**) porovnáva s embeddingami jednotlivých chunkov rozhodnutí. Ukážka 4 zobrazuje Elasticsearch dopyt pre vyhľadanie najpodobnejších chunkov súdnych rozhodnutí.

```
{
  "knn": {
    "field": "embedding",
    "query_vector": [0.12, ..., 0.44], /*384-rozmerný vektor */
    "k": 10,
    "num_candidates": 100
  },
  "_source": [
    "doc_id",
    "chunk_id",
    "section",
    "text"
  ]
}
```

Ukážka 4 - Elasticsearch knn dopyt pre vyhľadanie chunkov súdnych rozhodnutí.

Pole `field` určuje, voči ktorému vektorovému poľu sa má vyhľadávať. V tomto prípade ide o pole `embedding`. Hodnota `query_vector` predstavuje embedding vstupnej témy používateľa. Parameter k určuje počet výsledkov, ktoré majú byť vrátené. V prototype je predvolená hodnota 10, teda z každého indexu sa vyhľadá desať najpodobnejších výsledkov. Parameter `num_candidates` určuje počet kandidátnych vektorov, z ktorých Elasticsearch vyberá najbližšie výsledky.

Časť `_source` určuje, ktoré polia dokumentov majú byť vrátené z indexu. Pri súdnych rozhodnutiach sa vracia identifikátor pôvodného rozhodnutia, poradové číslo chunku, časť rozhodnutia a text chunku. Embedding samotný sa vo výsledku nevracia, pretože je potrebný iba na výpočet podobnosti.

Rovnaký typ knn dopytu sa používa aj pri indexe právnych ustanovení. Rozdiel je iba v tom, že právne ustanovenia nie sú členené na chunky, preto sa z indexu vracia iba identifikátor ustanovenia a jeho text. Ukážka 5 zobrazuje zjednodušený dopyt pre index právnych ustanovení.

```
{
  "knn": {
  "field": "embedding",
  "query_vector": [0.12, ..., 0.44], /*384-rozmerný vektor */
      "k": 10,
      "num_candidates": 100
    },
    "_source": [
      "doc_id",
      "text"
    ]
 }
```

Ukážka 5 - Elasticsearch knn dopyt pre vyhľadanie právnych ustanovení.

Výsledkom tohto kroku sú dve množiny kandidátnych dokumentov: významovo podobné chunky súdnych rozhodnutí a významovo podobné právne ustanovenia. Tieto výsledky ešte nepredstavujú finálny výstup metódy. Slúžia ako vstup pre nasledujúce kroky, v ktorých jazykový model navrhne vyhľadávacie pojmy a vyberie relevantné identifikátory právnych ustanovení.

2.3. Generovanie odporúčaných vyhľadávacích pojmov

Výsledky sémantického vyhľadávania nad súdnymi rozhodnutiami sa používajú ako vstup pre jazykový model, ktorého úlohou je navrhnúť súvisiace vyhľadávacie pojmy rozširujúce pôvodnú tému používateľa. Pri téme „kyberkriminalita“ môže model navrhnúť napríklad pojmy ako „IP adresa“, „počítačová sieť“, „internetová stránka“ alebo „počítačové dáta“.

Z vyhľadaných chunkov súdnych rozhodnutí sa najprv vytvorí textový kontext. Každý chunk je označený identifikátorom vo forme [Rozhodnutie #] a obsahuje text príslušného chunku. Jednotlivé chunky sú spojené do jedného textového kontextu a oddelené značkou ---. V prototype sa používa hodnota 10, preto sa do kontextu vkladá najviac desať najrelevantnejších chunkov súdnych rozhodnutí. Skrátený príklad takéhoto kontextu zobrazuje Ukážka 6.

```
[Rozhodnutie 1]
Obvinený neoprávnene získal prístup k používateľským údajom prostredníctvom internetovej stránky...
---
[Rozhodnutie 2]
Súd posudzoval neoprávnený zásah do počítačového systému a manipuláciu s počítačovými dátami...
---
• • •
---
[Rozhodnutie 10]
IP adresa bola použitá na identifikáciu zariadenia použitého pri útoku...
```

Ukážka 6 - Kontext vytvorený z vyhľadaných častí súdnych rozhodnutí.

Volanie jazykového modelu je rozdelené na systémový prompt a používateľský prompt. Systémový prompt definuje všeobecnú rolu modelu, čo zobrazuje Ukážka 7.

```
Si komponent právneho RAG systému určený na extrakciu vyhľadávacích pojmov.
```

Ukážka 7 - Systémový prompt pre generovanie odporúčaných vyhľadávacích pojmov.

Používateľský prompt obsahuje pôvodnú tému používateľa, konkrétnu úlohu, pravidlá generovania vyhľadávacích pojmov a kontext vyhľadaných súdnych rozhodnutí. Kontext je vložený v závere promptu medzi značky označujúce začiatok a koniec vyhľadaných dokumentov, čo zobrazuje Ukážka 8.

```
Pôvodná otázka používateľa:
kyberkriminalita

Úloha:
Navrhni 5 až 10 relevantných vyhľadávacích pojmov k používateľskej otázke na základe vyhľadaných súdnych rozhodnutí.

Pokyny:
- Použi iba vyhľadané súdne rozhodnutia.
- Všetky search_phrases musia byť po slovensky.
- Neprekladaj slovenské výrazy do angličtiny.
```

```
- Ak je pojem vo vyhľadaných rozhodnutiach uvedený po slovensky, zachovaj slovenskú
formuláciu.
- Anglické výrazy vráť iba vtedy, ak sa presne v tejto podobe nachádzajú vo vyhľadaných
rozhodnutiach a nemajú vhodný slovenský ekvivalent.
- Vyhľadávacie pojmy odvoď z významu a terminológie vyhľadaných rozhodnutí.
- Každý vyhľadávací pojem má byť jednoslovný pojem alebo krátke terminologické spojenie s
dĺžkou najviac 4 slová.
- Nevracaj duplicitné alebo významovo totožné výrazy.
- Nevkladaj názvy právnych predpisov, paragrafy, čísla ustanovení, spisové značky ani citácie.
- Ak vyhľadané dokumenty neobsahujú relevantné pojmy k otázke, vráť prázdny zoznam
search_phrases.
- Nevymýšľaj vyhľadávacie pojmy bez opory vo vyhľadaných dokumentoch.

--- ZAČIATOK VYHĽADANÝCH DOKUMENTOV ---
{documents_context}
--- KONIEC VYHĽADANÝCH DOKUMENTOV ---
```

Ukážka 8 - Používateľský prompt pre generovanie odporúčaných vyhľadávacích pojmov.

Na podporu štruktúrovaného výstupu sa pri volaní modelu používa JSON schéma. JSON schéma definuje očakávanú štruktúru odpovede, teda aké polia má výstup obsahovať a aký dátový typ majú mať ich hodnoty. Tá povoľuje iba pole `search_phrases`, ktorého hodnotou je zoznam reťazcov. Schému použitú pre tento krok zobrazuje Ukážka 9.

```
{
  "type": "object",
  "properties": {
     "search_phrases": {
        "type": "array",
        "items": {
           "type": "string"
         }
      }
   },
   "required": ["search_phrases"],
   "additionalProperties": false
  }
```

Ukážka 9 - JSON schéma výstupu pre odporúčané vyhľadávacie pojmy.

Po získaní odpovede sa z výsledku odstránia prípadné prázdne hodnoty a počet vrátených pojmov sa obmedzí na hodnotu 10. Príklad výsledného výstupu zobrazuje Ukážka 10.

```
{
  "search_phrases": [
     "IP adresa",
     "počítačová sieť",
     "internetová stránka",
     "počítačové dáta"
   ]
 }
```

Ukážka 10 - Príklad výstupu generovania odporúčaných vyhľadávacích pojmov.

2.4. Výber relevantných právnych ustanovení

Druhým výstupom metódy je zoznam relevantných právnych ustanovení. Tento krok nadväzuje na sémantické vyhľadávanie právnych ustanovení opísané v podkapitole 2.2 a pracuje s výsledkami dopytu, ktorý zobrazuje Ukážka 5.

Pri právnych ustanoveniach sa po sémantickom vyhľadaní kandidátov vykoná hierarchická deduplikácia. Tá rieši situáciu, keď sa vo výsledkoch nachádza všeobecnejšie aj konkrétnejšie ustanovenie. Ak výsledky obsahujú napríklad identifikátor paragrafu aj identifikátor jeho odseku, všeobecnejší identifikátor paragrafu sa odstráni a ponechá sa konkrétnejšie ustanovenie.

Z takto upravených právnych ustanovení sa vytvorí textový kontext pre jazykový model. Každé ustanovenie obsahuje poradové označenie, identifikátor ID a text ustanovenia. Príklad skráteného kontextu zobrazuje Ukážka 11.

```
[Právne ustanovenie 1]
ID: /SK/ZZ/2005/300/20260701.html#paragraf-247a.odsek-2.pismeno-c
Text:
Neoprávnený zásah do počítačového systému
(1) Odňatím slobody na jeden rok až päť rokov sa páchateľ potrestá, ak spácha čin uvedený v odseku 1
  c) tak, že zneužije osobné údaje iného s cieľom získať dôveru tretej strany,
---
[Právne ustanovenie 2]
ID: /SK/ZZ/2005/300/20260701.html#paragraf-360b.odsek-1.pismeno-a
Text:
Nebezpečné elektronické obťažovanie
(1) Kto úmyselne prostredníctvom elektronickej komunikačnej služby...
 a) ho dlhodobo ponižuje, zastrašuje, neoprávnene koná v jeho mene...
---
...
```

Ukážka 11 - Kontext vytvorený z vyhľadaných právnych ustanovení.

Rovnako ako pri generovaní vyhľadávacích pojmov, aj tento krok používa samostatný systémový a používateľský prompt. Systémový prompt definuje rolu modelu ako komponentu určeného na výber relevantných právnych ustanovení z vyhľadaného kontextu, čo zobrazuje Ukážka 12.

```
Si komponent právneho RAG systému určený na výber relevantných ustanovení z vyhľadaného zoznamu.
```

Ukážka 12 - Systémový prompt pre výber relevantných právnych ustanovení.

Používateľský prompt pre výber právnych ustanovení zobrazuje Ukážka 13. Model má vyberať iba z predložených ustanovení v kontexte zobrazenom v Ukážka 11 a do výstupu vrátiť iba hodnoty uvedené za **ID:**.

```
Pôvodná otázka používateľa:
kyberkriminalita

Úloha:
Vyber relevantné právne ustanovenia k používateľskej otázke podľa priloženého zoznamu
```

```
právnych ustanovení.

Pokyny:
- Do poľa legal_provisions vráť iba hodnoty uvedené za "ID:".
- Nevracaj text právneho ustanovenia.
- Nevymýšľaj ID právnych ustanovení.
- Ak vyhľadané právne ustanovenia neobsahujú relevantné ustanovenia k otázke, vráť prázdny zoznam legal_provisions.

--- ZAČIATOK VYHĽADANÝCH PRÁVNYCH USTANOVENÍ ---
{laws_context}
--- KONIEC VYHĽADANÝCH PRÁVNYCH USTANOVENÍ ---
```

Ukážka 13 - Používateľský prompt pre výber relevantných právnych ustanovení.

Na podporu štruktúrovaného výstupu sa používa JSON schéma s poľom `legal_provisions`, ktorého hodnotou je zoznam reťazcov. Schému zobrazuje Ukážka 14.

```
{
  "type": "object",
  "properties": {
    "legal_provisions": {
        "type": "array",
        "items": {
           "type": "string"
         }
     }
  },
  "required": ["legal_provisions"],
  "additionalProperties": false
}
```

Ukážka 14 - JSON schéma výstupu pre relevantné právne ustanovenia.

Po získaní odpovede sa z výsledku odstránia prípadné prázdne hodnoty. Príklad výstupu výberu relevantných právnych ustanovení zobrazuje Ukážka 15.

```
{
  "legal_provisions": [
    "/SK/ZZ/2005/300/20260701.html#paragraf-247a.odsek-2.pismeno-c",
    "/SK/ZZ/2005/300/20260701.html#paragraf-247b.odsek-3.pismeno-b",
    "/SK/ZZ/2005/300/20260701.html#paragraf-247b.odsek-2.pismeno-c",
    "/SK/ZZ/2005/300/20260701.html#paragraf-360b.odsek-1.pismeno-b",
    "/SK/ZZ/2005/300/20260701.html#paragraf-360b.odsek-1.pismeno-a"
  ]
}
```

Ukážka 15 - Príklad výstupu výberu relevantných právnych ustanovení

Identifikátor právneho ustanovenia obsahuje informácie o právnom predpise, jeho verzii a konkrétnej časti ustanovenia. Časť /SK/ZZ/2005/300/20260701.html označuje právny predpis a jeho verziu – v uvedenom príklade 2005/300 predstavuje identifikátor Trestného zákona a 20260701 označuje verziu právneho predpisu. Časť za znakom # následne určuje konkrétne ustanovenie, napríklad § 247a, odsek 2 a pismeno c).

2.5. Výsledný výstup sprostredkovacej metódy

Výstupy z predchádzajúcich krokov sa na konci spoja do jedného JSON objektu. Pole `search_phrases` obsahuje odporúčané vyhľadávacie pojmy získané zo súdnych rozhodnutí a pole `legal_provisions` obsahuje identifikátory relevantných právnych ustanovení vybrané z vyhľadaného zoznamu právnych ustanovení.

Príklad výsledného výstupu sprostredkovacej metódy zobrazuje Ukážka 16.

```
{
  "search_phrases": [
    "IP adresa",
    "počítačová sieť",
    "internetová stránka",
    "počítačové dáta"
  ],
  "legal_provisions": [
    "/SK/ZZ/2005/300/20260701.html#paragraf-247a.odsek-2.pismeno-c",
    "/SK/ZZ/2005/300/20260701.html#paragraf-247b.odsek-3.pismeno-b",
    "/SK/ZZ/2005/300/20260701.html#paragraf-247b.odsek-2.pismeno-c",
    "/SK/ZZ/2005/300/20260701.html#paragraf-360b.odsek-1.pismeno-b",
    "/SK/ZZ/2005/300/20260701.html#paragraf-360b.odsek-1.pismeno-a"
  ]
}
```

Ukážka 16 - Výsledný JSON výstup sprostredkovacej metódy

3. Spustenie demo projektu

Spustiteľný prototyp sprostredkovacej metódy je dostupný v GitLab repozitári projektu: https://gitlab.science.upjs.sk/altaml/middleware-search-method.

Súčasťou repozitára je aj demo skript **demo_cybercrime.py**, ktorý slúži na overenie funkcionality metódy na téme „kyberkriminalita“.

Pred spustením demo skriptu je potrebné postupovať podľa súboru **README.md**, ktorý obsahuje informácie o príprave prostredia, konfigurácii pripojenia k Elasticsearch, pripojení k jazykovému modelu a vytvorení vyhľadávacích indexov. Demo dáta predstavujú podmnožinu súdnych rozhodnutí a právnych predpisov určenú na rýchle overenie funkcionality prototypu.