

# Extrakcia právnych odkazov pomocou modelov transformérov

Vedúci práce: RNDr. Peter Gurský, PhD.

Konzultant: RNDr. Dávid Varga

Študent: Bc. Nicol Fedurcová



# Prečo je extrakcia právnych odkazov dôležitá?

- Využitelnosť v súdnej vede :
  - Nájdenie rozhodnutí, týkajúce sa podobných prípadov
  - Skúmanie vzťahov medzi rozhodnutiami
- Aplikácia v systéme na vyhľadávanie súdnych rozhodnutí
- Zjednodušenie automatizovaného spracovania rozhodnutí

# Výzvy pri vyhľadávaní a anotovaní právnych odkazov

**HLAVNÝ PROBLÉM:** súdne rozhodnutia sú písané prirodzeným jazykom

- Odkazy nemajú jednotnú štruktúru
- Používanie skratiek, značiek, skratkových slov, ktoré nemusia byť vždy vysvetlené
- Každý sudca má vlastný štýl písania rozhodnutia
- Štruktúra rozhodnutí je naprieč rôznymi krajinami rôzna



# Motivácia

- Aktuálna tématika modelov transformérov
- Práca na datase Slovenských súdnych rozhodnutí
- Využitelnosť výsledku práce v súdnej vede
- Aplikácia metódy v systéme na vyhľadávanie súdnych rozhodnutí
- Vylepšenie presnosti existujúcej metódy

# Čo sme urobili

Prehľad použiteľných jazykových modelov

Hľadanie príčin zlyhania extrakcie odkazov pôvodnou metódou

Vytvorenie metódy na nájdenie rozhodnutí s viacerými možnosťami výberu pri aliase

Anotovanie rozhodnutí s viacerými možnosťami výberu pri aliase, Vytvorenie prvej verzie Choice checkera

Modifikovanie metódy na nájdenie rozhodnutí s viacerými možnosťami výberu pri aliase

# Čo sme urobili

Vytvorenie metódy na nájdenie a preformátovanie znenia konkrétneho odkazu

Vytvorenie test makera, ktorého výstup bude vstupom pre nový choice checker

Vytvorenie novej (zjednodušenej) verzie choice checkera (kontrolóra viacerých možností pri aliase – bez kontroly existencie odkazu)

Zostavenie a manuálne anotovanie testovacej sady na porovnanie choice checkerov s rôznymi parametrami

Porovnanie výsledkov rôznych parametrov choice checkera

Aktuálne dokončujeme

## Prehľad použiteľných jazykových modelov

Viac-jazyčné:

- paraphrase-multilingual-MiniLM-L12-v2
- distiluse-base-multilingual-cased-v2
- all-MiniLM-L6-v2

Slovenský:

- kinit/slovakbert-sts-stsb

# Hľadanie príčin zlyhania extrakcie odkazov pôvodnou metódou

## Príklad A)

Súd podľa § 101 ods. 2 **O.s.p.** konal a rozhodol v neprítomnosti navrhovateľa a jeho právneho zástupcu, ktorí svoju neúčast na pojednávaní ospravedlnili, súhlasili s pojednávaním v ich neprítomnosti.

```
_id: "O.s.p"  
▼ law_ids: Array  
  0: "99/1963"
```

## Príklad B)

§ 62 ods. 1 zákona č. 36/2005 Z. z. o rodine a o zmene a doplnení niektorých zákonov - ďalejlen \"**ZoR**\"

Podľa § 78 ods. 2 **ZoR** ak dôjde k zrušeniu alebo zníženiu výživného pre maloleté dieťa

```
_id: "ZoR"  
▼ law_ids: Array  
  0: "94/1963"  
  1: "530/2003"  
  2: "36/2005"
```

## Príklad C)

Uvedené ust. § 23 **ZoR** predpokladá objektívnuexistenciu rozvratu vzťahov bez ohľadu na subjektívne



# Vytvorenie metódy na nájdenie a preformátovanie znenia konkrétneho odkazu

[static.slov-lex.sk/static/SK/ZZ/2003/530/20160318.html#paragraf-14.odsek-4.pismo-a](http://static.slov-lex.sk/static/SK/ZZ/2003/530/20160318.html#paragraf-14.odsek-4.pismo-a)

## § 14

### Splnomocňovacie ustanovenia

- (1) Vzory tlačív na podávanie návrhov na zápis v listinnej podobe a zoznam listín, ktoré treba k návrhu na zápis priložiť, a to bez ohľadu na jeho podobu, ustanoví všeobecne záväzný právny predpis, ktorý vydá ministerstvo.
- (2) Elektronickú podobu tlačív podľa odseku 1 určí ministerstvo na svojej internetovej stránke.
- (3) Spôsob a podmienky zverejňovania údajov podľa tohto zákona sú upravené osobitným zákonom.<sup>23a)</sup>
- (4) Všeobecne záväzný právny predpis, ktorý vydá ministerstvo, ustanoví
  - a) postup pri overovaní osobných údajov podľa [§ 5a ods. 2](#),

× Zrušiť označenie
  - b) podrobnosti a spôsob zasielania, odovzdávania a prijímania údajov, listín a informácií elektronickými prostriedkami na účely obchodného registra a zbierky listín.

#### PROVISION:

SK/ZZ/2003/530/20160318.html#paragraf-14.odsek-4.pismo-a  
-----TEXT

Splnomocňovacie ustanovenia  
Všeobecne záväzný právny predpis,  
ktorý vydá ministerstvo,  
ustanoví postup pri overovaní osobných  
údajov podľa § 5a ods. 2,

# Vytvorenie prvej verzie Choice checker-a

```
..... PARAGRAPH PROCESSING ENDED  
OLD METHOD SUCESSFUL IN 7 out of 8 paragraphs  
CHOICE CHECKER FOUND 1 improvements:  
    old choice: /SK/ZZ/1963/94/20040101.html#paragraf-78.odsek-2  
    better choice: /SK/ZZ/2005/36/20160101.html#paragraf-78.odsek-2
```

# Testovací dataset

```
{
  "url": "/SK/ZZ/1977/61/19930822.html#paragraf-127.odsek-4",
  "law": "61/1977",
  "version": "19930822",
  "fragment": "paragraf-127.odsek-4",
  "section": 4,
  "letter": null,
  "start": 5937,
  "end": 5953,
  "paragraph": "127",
  "zakonnik_start": 5949,
  "zakonnik_end": 5953,
  "extracted_from": "global from db",
  "alias_text": "",
  "aliases": "ZoCP",
  "alias_options": [
    "61/1977",
    "526/1990",
    "566/2001",
    "326/2005",
    "8/2009"
  ],
  "annotated_url": "/SK/ZZ/2001/566/20130610.html#paragraf-127.odsek-4",
  "existing_options": [
    {
      "zakonnik": "566/2001",
      "provision_text": "566/2001, o cenných papieroach a investičných službách a o zmene a doplnení niektorých zákonov (zákon o cenných papieroach), och",
      "version": "20130610.html"
    },
    {
      "zakonnik": "8/2009",
      "provision_text": "8/2009, o cestnej premávke a o zmene a doplnení niektorých zákonov, evidencia vozidiel, evidovanie vozidiel a evidenčné čísla, evid",
      "version": "20160101.html"
    }
  ],
  "surrounding_text": "neskorších predpisov /ZoCP/ za porušenie ustanovenia § 129 ods. 3 v spojení sustanovením § 128 ods.1 písmeno b/ c/ h/ ZoCP"
}
```

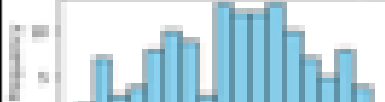
## Výsledok choice checkera V2

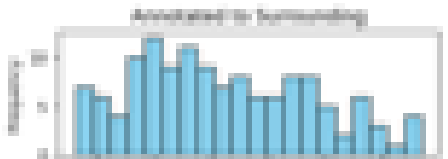
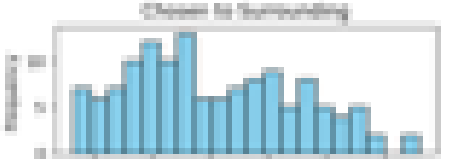
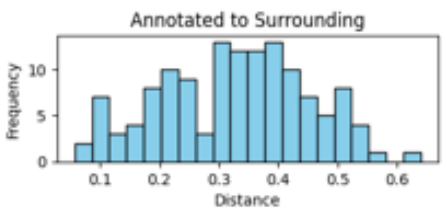
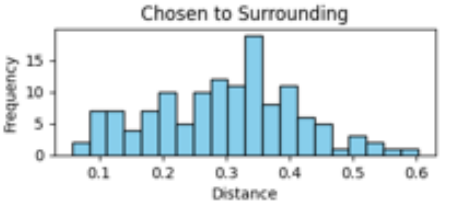
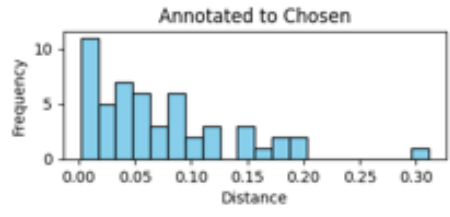
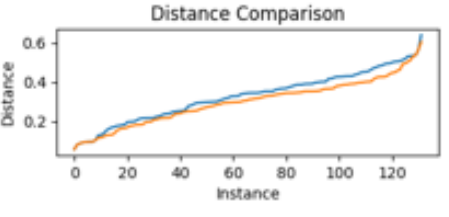
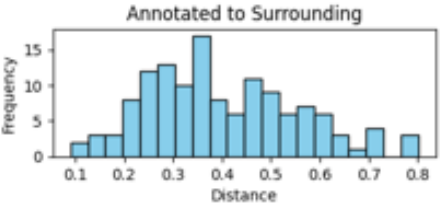
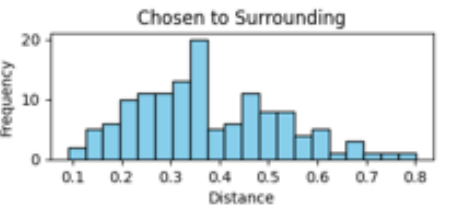
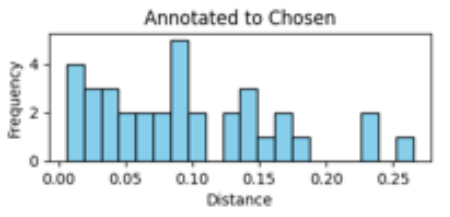
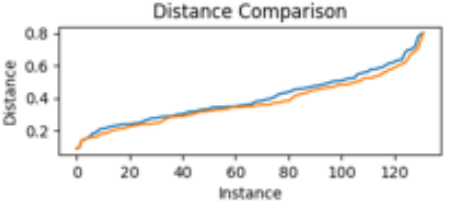
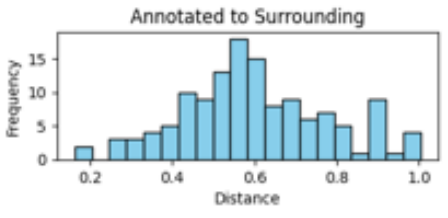
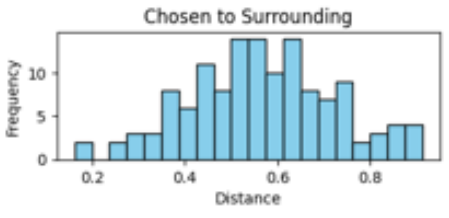
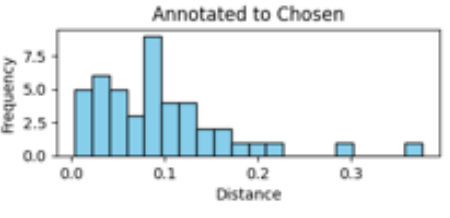
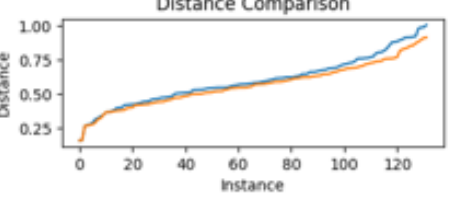
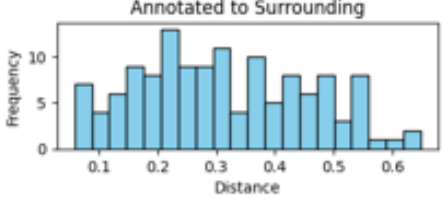
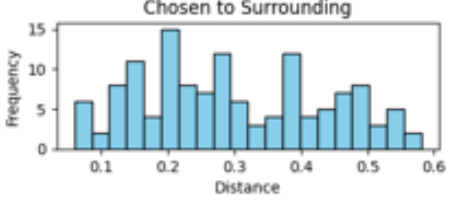
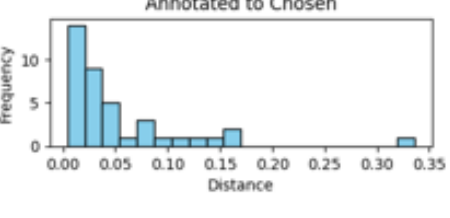
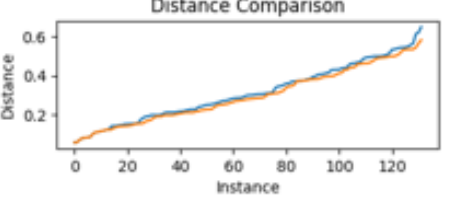
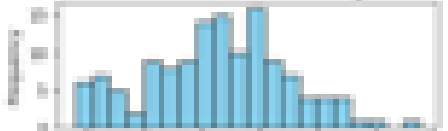
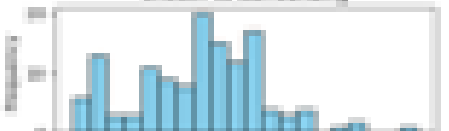
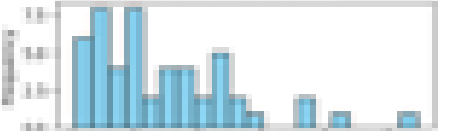
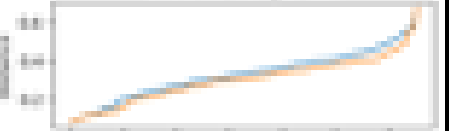
```
"5C/265/2015__2016-06-02__ECLI:SK:OSNO:2016:5815205480.3": [  
  {  
    "start": 5085,  
    "end": 5100,  
    "checker_url": "/SK/ZZ/2005/36/20160101.html#paragraf-78.odsek-2",  
    "surrounding_text": "nosti rodičov k deťom je ich zákonná povinnosť, ktorá trvá dočasu, kým deti nie sú sch",  
    "existing_options": {  
      "94/1963": {  
        "zakonnik": "94/1963",  
        "distance_from_surrounding_text": "0.64754695",  
        "version": "20040101.html",  
        "provision_text": "94/1963, Funkcia opatrovníka je čestnou povinnosťou."  
      },  
      "36/2005": {  
        "zakonnik": "36/2005",  
        "distance_from_surrounding_text": "0.15579677",  
        "version": "20160101.html",  
        "provision_text": "36/2005, o rodine a o zmene a doplnení niektorých zákonov, výživné, spoločné ustanoven",  
      }  
    }  
  }  
],  
"44C/54452/2015__2016-02-05__ECLI:SK:KSCF:2016:7014204000.4": [
```



# Výsledky

[illegible]

| cs/cslv2                              |       |         |           |          |             |             |                 |                 |                  |                  |                 |                 |  |     |
|---------------------------------------|-------|---------|-----------|----------|-------------|-------------|-----------------|-----------------|------------------|------------------|-----------------|-----------------|--|-------------------------------------------------------------------------------------|
| cs/cslv2                              | 132   | 93      | 40        | cosine   | 800         | 69,66666667 | 0,32067721      | 0,34772727      | 0,29444444       | 0,30333333       | 0,29166667      | 0,32222222      |  |   |
| Model                                 | Total | Matched | Unmatched | Distance | Window size | Accuracy    | Avg_Ann_to_Surr | Med_Ann_to_Surr | Avg_Chos_to_Surr | Med_Chos_to_Surr | Avg_Ann_to_Chos | Med_Ann_to_Chos |  |  |
| all-MiniLM-L6-v2                      | 132   | 80      | 52        | cosine   | 800         | 60,60606061 | 0,327162382     | 0,34336418      | 0,298077145      | 0,308030395      | 0,073831754     | 0,06124678      |  |  |
| paraphrase-multilingual-MiniLM-L12-v2 | 132   | 97      | 35        | cosine   | 800         | 73,48484848 | 0,399497231     | 0,361825885     | 0,373926797      | 0,3483851        | 0,096437067     | 0,0884207       |  |  |
| distiluse-base-multilingual-cased-v2  | 132   | 87      | 45        | cosine   | 800         | 65,90909091 | 0,597916833     | 0,580486025     | 0,565329815      | 0,56241157       | 0,095588588     | 0,08853113      |  |  |
| kinit/slovakbert-sts-stsb             | 132   | 93      | 39        | cosine   | 800         | 70,45454545 | 0,312877411     | 0,299257395     | 0,297329678      | 0,27970916       | 0,052623097     | 0,02573776      |  |  |
|                                       |       |         |           |          |             |             |                 |                 |                  |                  |                 |                 |  |  |

|      |             |             |            |                                                                                     |                                                                                       |                                                                                       |                                                                                       |
|------|-------------|-------------|------------|-------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
|      |             |             |            |     |     |     |     |
|      |             |             |            | plot: Annotated to Surrounding                                                      | plot: Chosen to Surrounding                                                           | plot: Annotated to Chosen                                                             | plot: Distance Comparison<br>(orange=chosen:surr, blue=annot:surr)                    |
| 7145 | 0,308030395 | 0,073831754 | 0,06124678 |    |    |    |    |
| 5797 | 0,3483851   | 0,096437067 | 0,0884207  |    |    |    |    |
| 9815 | 0,56241157  | 0,095588588 | 0,08853113 |   |   |   |   |
| 9678 | 0,27970916  | 0,052623097 | 0,02573776 |  |  |  |  |
|      |             |             |            |  |  |  |  |

# Výzvy

- Porovnanie rôznych parametrov a rôznych modelov
- Pridanie vektora „zvyčajného použitia“ k rozhodovaniu
- Pripravenie BERT modelu pretrénovanom na slovenských súdnych textoch na použitie
- Skonštruovanie modelu na extrakciu odkazov na rozhodnutia Európskeho súdneho dvora
  - NER úloha doplnená o pravidlový systém

# Výzvy

mimoriadneho opravného prostriedku, ktorý nebol dostupný účastníkom konania, ale len generálnemu

prokurátorovi, ktorý mal, čo sa týka aplikácie tohto inštitútu voľnú úvahu (rozsudok ESLP vo veci Brumarescu

•reference

•office

•sides

proti Rumunsku, sťažnosť č. 28342/95, rozsudok zo dňa 28. októbra 1999). Z judikatúry ESLP ďalej vyplýva, že

•id

•date

prieskum právoplatných súdnych rozhodnutí musí rešpektovať princíp právnej istoty, keď konečné rozsudky by

mali vo všeobecnosti zostať nedotknuté, zrušené môžu byť iba za účelom nápravy fundamentálnych chýb

(rozsudok ESLP vo veci Tishkevich proti Rusku, sťažnosť č. 2202/05, rozsudok zo dňa 4. decembra 2008).

•reference

•office

•sides

•id

•date

Fundamentálne pochybenie, ktoré opodstatňuje zrušenie rozhodnutia, môže podľa názoru ESLP predstavovať

napríklad omyl v právomoci, vážne porušenie súdneho procesu alebo zneužitie právomoci (rozsudok vo veci

•reference

Luchkina proti Rusku, sťažnosť č. 3548/04, rozsudok zo dňa 10. apríla 2008). Mimoriadny prieskum by ale

•sides

•id

•date



# Výzvy

```
{'7510/04', '3548/04', '11867/09', '18968/07', '30189/07', '28342/95', '42872/02', '9.6.2015', '25132/13',  
'7711/06', '33809/08', '36997/05', '72493/10', '46092/06', '2202/05'}
```

```
{'Európskym súdom pre ľudské práva', 'Európskom súde pre ľudské práva (ďalej len ESLP)', 'ESLP', 'ESLP '}
```

```
{'Tishkevich proti Rusku',  
'Tripon proti Rumunsku',  
'Kontrová v. SR',  
'V.C. v. SR',  
'PSMA, s.r.o. proti Slovenskej republike',  
'Soltész v. SR',  
'Sutyazhnik proti Rusku',  
'Abdullayev proti Rusku',  
'Luchkina proti Rusku',  
'Brumarescu proti Rumunsku',  
'Ištván a Ištvánová v. SR',  
'COMPCAR, s.r.o. proti Slovenskej republike',  
'Labsí v. SR',  
'DRAFT-OVA, a.s. proti Slovenskej republike',  
'Cornif proti Rumunsku',  
'Kutepov a Anikeyenko proti Rusku',  
'Černák v. SR',  
'Roseltrans proti Rusku',  
'Žúbor v. SR',  
'DRAFT-OVA, a.s., v. Slovenská republika',  
'Kormoš v. SR'}
```

# Literatúra

1. Dávid Varga, et al. „Extraction of Legal References from Court Decisions“. Vol. 3498, 2023, pp. 89–95,
2. Darji Harshil, Jelena Mitrović, and Michael Granitzer. "A Dataset of German Legal Reference Annotations." Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law. 2023.
3. Elena Leitner, Georg Rehm and Julian Moreno-Schneider. “A Dataset of German Legal Documents for Named Entity Recognition“. Proceedings of the 12th Language Resources and Evaluation Conference (LREC 2020)
4. Malte Ostendorff, Till Blume, and Saskia Ostendorff. 2020. „Towards an open platform for legal information“. In Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020. 385–388.