D23 – Hodnotiaca správa

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

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 zameraním 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 KPB5 bol vytvorený výstup **D23 - Hodnotiaca správa**. Tento výstup obsahuje analýzu metód a modelov založenú na štandardnej metodológii empirického právneho výskumu. Účelom tohto dokumentu je získať spätnú väzbu od právnych výskumných tímov. Vybrané sú konkrétne prípadové štúdie (právne výskumné témy). Právne výskumné tímy v rámci tejto správy uvádzajú hodnotenia použitej metodológie, jej výsledkov ako aj odporúčania a návrhy na ďalší rozvoj použitých metód a modelov.

Obsah

Úvod **1**

1. Úvod do právneho výskumu **3**

2. Metóda systematickej kontextovej analýzy práva a jej jednotlivé štádiá **5**

3. Metodológia tvorby hodnotiacej správy **8**

4. Právne tímy a stanovenie vhodných výskumných otázok / hypotéz **12**

5. Identifikácia a zber vhodnej vzorky údajov **18**

6. Kódovanie právneho textu **24**

7. Formulácia záverov, resp. pozorovaní na základe dátovej analýzy **30**

7.1 Deskriptívna štatistická analýza *31*

7.2 Analytická metóda - Komparatívna analýza skupín *32*

7.3 Analytická metóda – Korelačná analýza medzi atribútmi *34*

7.4 Analytická metóda – Zhluková analýza *35*

7.5 Analytická metóda – Detekcia anomálií *37*

7.6 Analytická metóda – Asociačné pravidlá *38*

7.7 Analytická metóda – Prediktívna analýza *39*

8. Informovanie o dosiahnutých výsledkoch spôsobom pochopiteľným pre právnu komunitu **42**

Záver **53**

Príloha č. 1 ..... **56**

1. Úvod do právneho výskumu

Vymedzenie metodologického prístupu a konkrétnych metodologických postupov, ktoré boli použité v rámci výskumu, predstavuje nevyhnutnú súčasť akéhokoľvek vedeckého skúmania. Napriek uvedenému sa možno v odbornej spisbe stretnúť s nedostatočnou špecifikáciou výskumných metód, na základe ktorých sú formulované závery vo vzťahu k analyzovaným výskumným otázkam, oslabujúc vedeckosť konkrétneho skúmania. Právny výskum pritom môže byť založený na rôznych vedeckých metódach. Štandardný prístup k právnemu skúmaniu predstavuje tzv. **doktrinálny právny výskum** založený na normatívnej analýze práva, ktorého cieľom je identifikovať, analyzovať a syntetizovať obsah práva[1], a to spravidla na základe analýzy relevantnej legislatívy a judikatúry umožňujúcej interpretáciu práva,[2] ako aj hľadanie vzájomných vzťahov, problematických oblastí a možných hypotéz o budúcom právnom vývoji.[3] Rozlíšiť však možno aj kategóriu **empirického právneho výskumu** implementujúceho rôzne empirické vedecké metódy, ktorých využitie v rámci právneho výskumu naberá na dôležitosti.[4] Požiadavka na výraznejšie zahrnutie metód empirického výskumu do rámca právneho výskumu bola vyjadrená vo viacerých štúdiách.[5] Empirický právny výskum sa považuje za interdisciplinárnu verziu právneho výskumu, v rámci ktorého sa na právne štúdie aplikujú metódy zvyčajne prevzaté zo spoločenských vied.[6] Empirický právny výskum nenahrádza doktrinálny právny výskum, ale dopĺňa ho, pričom oba prístupy možno využiť súčasne za účelom preskúmania stanovených výskumných otázok.[7] Významnú súčasť empirického

[1] HUTCHINSON, Terry. Doctrinal Research: researching the jury. In WATKINS, Dawn, ed. Research Methods in Law. London: Routledge, 2013, s. 7–33. https://doi.org/10.4324/9780203489352
[2] MCCONVILLE, Michael. CHUI, Wing Hong, eds. Research methods for law. Edinburgh: Edinburgh University Press, 2017. https://doi.org/10.1515/9781399515092
[3] HUTCHINSON, Terry. Doctrinal Research: researching the jury. In WATKINS, Dawn, ed. Research Methods in Law. London: Routledge, 2013, s. 7–33. https://doi.org/10.4324/9780203489352
[4] MONAHAN, John. WALKER, Laurens. Twenty-Five Years of Social Science in Law. In *Law and Human Behavior*, 2011, roč. 35, č. 1, s. 72–82. https://doi.org/10.1007/s10979-009-9214-8
[5] Pozri napríklad 1) BALDWIN, John. DAVIS, Gwynn. Empirical Research in Law. In CANE, Peter. TUSHNET, Mark, eds. The Oxford Handbook of Legal Studies. Oxford: Oxford University Press, 2003. https://doi.org/10.1093/oxfordhb/9780199248179.013.0039, 2) LEEUW, Frans L. Empirical Legal Research: The Gap between Facts and Values and Legal Academic Training. In *Utrecht Law Review*, 2015, roč. 11, č. 2, s. 19-33. https://doi.org/0.18352/ulr.315
[6] EPSTEIN, Lee. KING, Gary. The Rules of Inference. In *The University of Chicago Law Review*, 2002, roč. 69, č. 1, s. 1–3. https://doi.org/10.2307/1600349
[7] BALDWIN, John. DAVIS, Gwynn. Empirical Research in Law. In CANE, Peter. TUSHNET, Mark, eds. The Oxford Handbook of Legal Studies. Oxford: Oxford University Press, 2003. https://doi.org/10.1093/oxfordhb/9780199248179.013.0039

právneho výskumu pritom tvorí využitie **kvantitatívnych metód skúmania**, ktoré sledujú systematické vedecké preskúmanie kvantitatívnych vlastností javu a ich vzájomných vzťahov. Jednou z relevantných kvantitatívnych metód empirického výskumu, ktorá je aplikáciou kontextovej analýzy v práve, je tzv. **metóda systematickej kontextovej analýzy práva** (*systematic content analysis*), ktorá umožňuje systematickú analýzu textov rôzneho druhu štandardizovaným spôsobom. Predmetná metóda prináša kvantitatívny pohľad na analyzované výskumné otázky reflektujúc obsah právnych textov, ktoré sú predmetom skúmania v konkrétnom prípade, a umožňuje dosiahnutie exaktnejších a predovšetkým replikovateľných a opätovne verifikovateľných výsledkov bez ohľadu na osobu, ktorá výskum vykonáva. Využitie tejto metódy sa javí ako zvlášť vhodné v prípade analýzy väčšieho množstva právnych textov, ktorých komplexné preskúmanie nevyhnutne predpokladá osobitný prístup umožňujúci štandardizovanú analýzu a jednotné vyhodnotenie skutočností skúmaných v konkrétnom prípade.

2. Metóda systematickej kontextovej analýzy práva a jej jednotlivé štádiá

Metóda systematickej kontextovej analýzy práva predstavuje metodickú, replikovateľnú techniku na kompresiu mnohých slov právneho textu do menšieho počtu obsahových kategórií na základe explicitných pravidiel kódovania.[8] Predmetná metóda umožňuje produkovať kvalitatívne aj kvantitatívne výsledky tým, že umožňuje výskumníkovi komprimovať (právny) text do menšieho počtu obsahových kategórií na základe explicitných pravidiel kódovania.[9] Poskytuje tak objektívne a vedecky verifikovateľné výsledky vyplývajúce z analýzy rôznych druhov textov, počínajúc prepismi rozhovorov až po právne texty vrátane právnych predpisov a súdnych rozhodnutí.[10]

Vhodnosť výberu tejto výskumnej metódy pre účely právneho výskumu je preukázaná jej podobnosťou k tradičnej právnej analýze spočívajúcej v preskúmaní zbierky právnych textov, napríklad súdnych rozhodnutí, s cieľom identifikovať spoločné prvky a zhodnotiť ich relevantnosť, so súčasnou inkorporáciou prísnosti typickej pre spoločenské vedy umožňujúcej pochopenie analyzovanej judikatúry.[11] V porovnaní s tradičným právnym výskumom, ktorý sa spravidla zameriava na analýzu konkrétneho súdneho rozhodnutia alebo menšieho počtu osobitne dôležitých rozhodnutí, táto výskumná metóda umožňuje preskúmanie väčšieho množstva podobne dôležitých rozhodnutí s cieľom identifikácie spoločných znakov, resp. všadeprítomných vzorcov.[12] Ďalšou výhodou tejto výskumnej metódy je možnosť jej použitia vo vzťahu k akémukoľvek druhu právneho textu, vrátane právnych predpisov, súdnych rozhodnutí, rozhodnutí iných príslušných štátnych orgánov, zmlúv, obchodných podmienok a pod., za predpokladu dispozície výskumníka s nimi.

---

[8] WEBER, Robert. Basic content analysis. Roč. 49. Newbury Park: Sage, 1990. https://doi.org/10.4135/9781412983488

[9] BURNS, Kylie. HUTCHINSON, Terry. The impact of "empirical facts" on legal scholarship and legal research training. In *The Law Teacher*, 2009, roč. 43, č. 2, s. 153–178. https://doi.org/10.1080/03069400903004236

[10] SALEHIJAM, Maryam. The Value of Systematic Content Analysis in Legal Research. In *Tilburg Law Review*, 2018, roč. 23, č. 1-2, s. 36-38. https://doi.org/10.5334/tilr.5

[11] HALL, Mark A. WRIGHT, Ronald F. Systematic content analysis of judicial opinions. In *California Law Review*, 2008, roč. 96, č. 1, s. 63–79.

[12] HALL, Mark A. WRIGHT, Ronald F. Systematic content analysis of judicial opinions. In *California Law Review*, 2008, roč. 96, č. 1, s. 63–79.

Metóda systematickej kontextovej analýzy práva má viaceré, na seba nadväzujúce štádiá. Hall a Wright rozlišujú tri osobitné zložky tejto výskumnej metódy, konkrétne

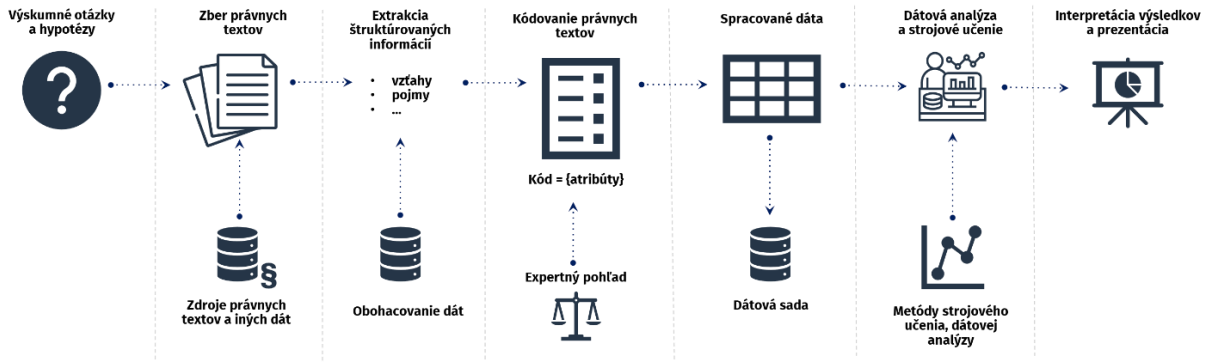
1) výber súdnych rozhodnutí,
2) kódovanie súdnych rozhodnutí a
3) analýza kódovaných súdnych rozhodnutí, často prostredníctvom štatistických metód.[13]

Salehijam bližšie rozlišuje nasledujúce štádiá predmetnej výskumnej metódy:

1) stanovenie vhodných výskumných otázok alebo hypotéz,
2) identifikácia a zber vhodnej vzorky údajov za účelom analýzy,
3) kódovanie údajov, ktoré má svoje vlastné štádiá,
4) formulácia záverov, resp. pozorovaní, a
5) informovanie o výsledkoch spôsobom pochopiteľným pre právnu komunitu.[14]

Obrázok č. 1 vizuálne zobrazuje postup pri použití skúmanej výskumnej metódy. Z klasifikácie jednotlivých štádií metódy systematickej kontextovej analýzy práva podľa Salehijam vychádzame s menšími úpravami a doplneniami. Postup metódy začína stanovením vhodnej výskumnej otázky, výberom vhodných zdrojov právnych textov a ich následným výberom. Metóda pokračuje v extrakcii štruktúrovaných údajov, prípadným doplnením externých, kontextových informácií. Jadro metódy predstavuje kódovanie právnych textov, pri ktorom je nevyhnutný vstup experta pre konkrétnu doménu, resp. oblasť práva. Takto vytvorená tabuľka kódov je následne spracovaná do podoby, ktorá je vhodná pre automatizované spracovanie. Spôsob spracovania závisí od výberu metódy. Oproti pôvodnému návrhu od Salehijam, dopĺňame fázu dátovej analýzy a strojového učenia. Výsledky z tejto fázy sú následne interpretované, formulované v záveroch a prezentované pre odbornú, či širokú verejnosť.

[13] Ibid. s. 79.
[14] SALEHIJAM, Maryam. The Value of Systematic Content Analysis in Legal Research. In *Tilburg Law Review*, 2018, roč. 23, č. 1-2, s. 36-38. https://doi.org/10.5334/tilr.5



Obrázok 1 Vizualizácia systematickej kontextovej analýzy práva

3. Metodológia tvorby hodnotiacej správy

V rámci tejto kapitoly uvádzame metodologický postup pre realizáciu právneho výskumu a samotné vytvorenie tejto hodnotiacej správy. Tento postup pozostával z nasledujúcich krokov, ktoré si v nasledujúcich kapitolách bližšie priblížime:

1) tvorba výskumných tímov a definovanie výskumných otázok,
2) hrubé a jemné stanovenie množiny rozhodnutí,
3) stanovenie atribútov pre kódovanie rozhodnutí,
4) tvorba parciálnych analýz súdnych rozhodnutí a
5) spätná väzba od právnych výskumných tímov.

1. Tvorba výskumných tímov a definovanie výskumných otázok

Prvým krokom metodológie bolo zostavenie špecializovaných právnych výskumných tímov pokrývajúcich štyri odborné domény, a to nekalú súťaž, zmluvnú pokutu, trestnoprávne zaistenie počítačových údajov a autorské právo. Každý tím bol zložený z akademických právnych expertov z Právnickej fakulty Univerzity Pavla Jozefa Šafárika (PravF UPJŠ) a Právnickej fakulty Univerzity Komenského (PravF UK), pričom každý tím formuloval konkrétnu výskumnú otázku, ktorá tvorila základ ďalšej práce. Výskumné otázky smerovali k empiricky overiteľným javom v rozhodovacej praxi slovenských súdov. Napríklad k tomu, ako súdy pristupujú k moderácii zmluvnej pokuty alebo pri akých trestných činoch sa najčastejšie vydávajú príkazy na zaistenie elektronických údajov. Bližší popis právnych výskumných tímov ako aj ich výskumných otázok je uvedený v kapitole č. 4 tejto hodnotiacej správy.

2. Hrubé a jemné stanovenie množiny rozhodnutí

V druhom kroku sa pre každú doménu vymedzil korpus relevantných súdnych rozhodnutí prostredníctvom dvojúrovňového filtrovania. Na úrovni hrubého vymedzenia sa stanovili základné vyhľadávacie kritériá. Išlo predovšetkým odkazy na konkrétne zákonné ustanovenia (napr. § 301 Obchodného zákonníka, § 544 ods. 2 Občianskeho zákonníka, § 91 a nasl. Trestného poriadku) a relevantné právne predpisy vrátane európskych smerníc a medzinárodných dohovorov. Na úrovni jemného vymedzenia sa aplikovali kumulatívne kritériá kombinujúce viacero podmienok súčasne. Napríklad sa použil odkaz na konkrétny paragraf v spojení s prítomnosťou kľúčového právneho pojmu v texte rozhodnutia (napr. „zmluvná pokuta“ a „primeranosť“ alebo „počítačový údaj“ a „mobilný telefón“). Táto dvojúrovňová selekcia umožnila dosiahnuť tematickú homogenitu datasetu a zároveň pokryť dostatočne široký rozsah rozhodnutí pre následnú analýzu.

3. Stanovenie atribútov pre kódovanie rozhodnutí

Tretím krokom bolo spoločné zadefinovanie sady atribútov pre každý dataset, teda vlastností, ktoré sa systematicky zachytávali z každého súdneho rozhodnutia. Atribúty boli navrhované v súčinnosti právnych expertov s dátovými analytikmi tak, aby čo najvernejšie zachytávali právne relevantné informácie. Napr. výška zmluvnej pokuty, výsledok konania, typ súdu, aplikovaná skutková podstata alebo druh trestnej činnosti. Pre každý atribút bol vypracovaný popis jeho obsahu, rozsahu hodnôt (napr. binárny, kategorický, numerický) a spôsobu jeho identifikácie a extrakcie z textu rozhodnutia. Súčasťou tohto kroku bola aj príprava kódovacieho manuálu, ktorý mal zabezpečiť jednotné a konzistentné kódovanie naprieč členmi tímu.

4. Tvorba parciálnych analýz súdnych rozhodnutí

Na základe takto vytvorených datasetov sa aplikoval súbor analytických metód pokrývajúcich rôzne aspekty skúmaných dátových sád súdnych rozhodnutí. Konkrétne boli využité nasledujúce analytické metódy (bližší popis je uvedený v kapitole č. 7):

1) deskriptívna štatistická analýza (frekvencie, podiely, distribúcie atribútov),
2) komparatívna analýza (porovnanie skupín rozhodnutí podľa súdu, regiónu či obdobia),
3) korelačná analýza medzi atribútmi (vzájomné závislosti medzi atribútmi),
4) zhluková/klastrovacia analýza (objavovanie prirodzených skupín rozhodnutí bez vopred určených kategórií),
5) detekcia anomálií (identifikácia rozhodnutí výrazne odchyľujúcich sa od typickej praxe),
6) asociačné pravidlá (vzory spoluvýskytu atribútov v tvare „ak X, potom Y") a
7) prediktívna analýza (modelovanie výsledku konania na základe ostatných atribútov s identifikáciou najvýznamnejších prediktorov).

Výstupy týchto analýz tvorili základ pre spätnú väzbu právnych tímov.

5. Spätná väzba od právnych výskumných tímov

Záverečným krokom metodológie bolo systematické získanie spätnej väzby od členov právnych výskumných tímov prostredníctvom štruktúrovaného dotazníka. Jeho cieľom bolo zistiť, nakoľko sú výstupy automatizovanej analýzy skutočne použiteľné v

právnom výskume, kde sú ich limity a čo by bolo potrebné zmeniť pre budúce projekty. Dotazník bol členený do siedmich tematických okruhov:

- **Kontext a východiská výskumu** - otázky zisťujú, aké výskumné otázky právne výskumné tímy sledovali ešte pred zapojením automatizovaných metód, ako vyzerala ich doterajšia metodológia a aké obmedzenia plynuli z obsahu dostupných právnych textov.
- **Kódovanie atribútov a tvorba datasetu** - táto časť sa zameriava na to, nakoľko zvolené atribúty zodpovedali právnej realite a ustálenej terminológii, či niektoré kategórie boli nesprávne, nejasné alebo zavádzajúce, a aké zmeny by členovia výskumných tímov navrhli do budúcnosti.
- **Hodnotenie jednotlivých analytických výstupov** - pre každú z použitých analytických metód respondenti hodnotili obsahový prínos a limity výstupu (formou otvorenej odpovede), ako aj jeho zrozumiteľnosť a užitočnosť na päťbodovej škále. Myšlienkou tejto časti je zachytiť, ktoré metódy sú pre členov právnych výskumných tímov intuitívne uchopiteľné a ktoré si vyžadujú dodatočnú interpretačnú podporu.
- **Interpretácia výsledkov a prepojenie s právnou argumentáciou -** otázky skúmajú, či a ako možno výstupy priamo integrovať do právneho textu alebo argumentácie, kde je hranica medzi štatisticky zaujímavým a právne relevantným, a aká je rola kvantitatívnych výstupov vo vzťahu k tradičnej normatívnej analýze práva.
- **Limitácia metódy a manuálne kódovanie** - táto časť sa zameriava na identifikáciu aspektov právneho textu, ktoré sú len ťažko alebo vôbec automatizovateľné, na konkrétne prípady zavádzajúcich alebo nesprávnych záverov automatizovanej analýzy a na hodnotenie konzistentnosti kódovania naprieč členmi tímu (inter-rater reliability).
- **Perspektívy a odporúčania** – respondenti navrhli konkrétne zmeny v procese výskumu, identifikovali právne oblasti s najväčším potenciálom pre túto metodológiu a vyjadrili sa k zapojeniu veľkých jazykových modelov (LLM). Súčasne hodnotili, či by boli schopní metódu využívať samostatne bez podpory dátového analytika.
- **Záverečné poznámky** - otvorený priestor pre akékoľvek doplňujúce postrehy, ktoré neboli pokryté predchádzajúcimi otázkami, vrátane etických aspektov, inštitucionálnych bariér alebo vízie budúceho vývoja metódy.

Celé znenie dotazníka je súčasťou Prílohy č. 1 tejto hodnotiacej správy.

4. Právne tímy a stanovenie vhodných výskumných otázok / hypotéz

Metóda systematickej kontextovej analýzy práva môže byť použitá s cieľom zodpovedať výskumné otázky, resp. potvrdiť alebo vyvrátiť stanovené vedecké hypotézy z rôznych právnych oblastí. Výskumné otázky alebo vedecké hypotézy sa pritom môžu zameriavať na úzko formulované právne problémy alebo konkrétne právne inštitúty a spôsob ich využitia v praxi, alebo môžu byť formulované širšie, napríklad s cieľom analyzovať rozhodovaciu prax konkrétneho štátneho orgánu reflektujúc vymedzené ustanovenia určitého právneho predpisu alebo aplikovateľnej právnej úpravy ako takej. Ide o známy problém medzi **law in book vs law in action**.[15]

Formulácia výskumných otázok alebo hypotéz je kľúčová pre všetky nasledujúce štádiá systematickej kontextovej analýzy práva. Určujúca je predovšetkým vo vzťahu k identifikácii nevyhnutnej vzorky právnych textov, ktoré majú byť podrobené analýze. Stanovené výskumné otázky, resp. hypotézy pritom smerujú najmä k **preskúmaniu skutočného stavu aplikácie práva** na základe podrobnej analýzy vytvorenej vzorky právnych textov. S uvedeným súčasne súvisia hlavné nevýhody skúmanej výskumnej metódy spočívajúce v **(ne)dostupnosti požadovaných právnych textov**[16], ako aj v **kvalite získaných právnych textov**, ktoré majú byť podrobené bližšej analýze. Hlavným problémom, s ktorým sa možno stretnúť pri praktickom využívaní tejto metódy, je potreba úpravy pôvodne stanovených výskumných otázok spôsobom, ktorý umožní reflektovať skutočný obsah skúmanej právnej dokumentácie. V mnohých prípadoch totiž nie je možné vo vzorke právnych textov nájsť odpovede na stanovené výskumné otázky vôbec, prípadne v požadovanom rozsahu. Vo všeobecnosti platí pravidlo, že nie je možné hľadať odpoveď v dátach, v ktorých dané poznanie, resp. odpoveď obsiahnuté nie sú (napríklad skúmanie spôsobu preukazovania viny obvineného v konaní pred súdom v prípade rozhodnutí, ktoré majú povahu dohody o vine a treste alebo trestného rozkazu, ktoré neobsahujú odôvodnenie v plnom rozsahu v súlade s § 168 ods. 1 zákona č. 301/2005 Z. z. Trestný

[15] POUND, Roscoe. Law in books and law in action. In *American Law Review*, 1910, roč. 44, s. 12., HALPERIN, Jean. Law in books and law in action: The problem of legal change. In *Maine Law Review*, 2011, roč. 64, s. 45.
[16] K tomu bližšie v nasledujúcej kapitole.

poriadok). Zjednodušene povedané, **limitom metódy systematickej kontextovej analýzy práva je skutočný obsah analyzovaných právnych textov**, ktorý dokáže významne obmedziť výskumné ciele stanovené v konkrétnom prípade. Súčasne však platí, že v niektorých prípadoch bude aj absencia určitých skutočností v skúmaných právnych textoch signifikantná pre vyvodenie určitých právnych záverov.

V rámci projektu sa identifikovali 4 právne výskumné tímy, ktorých expertíza a zapojenie do projektu malo vplyv na tvorbu konkrétnych výstupov projektu. Tieto výskumné tímy sa venujú nasledujúcim doménam:

Doména č. 1: Nekalá súťaž

Výskumný tím z Právnickej fakulty Univerzity Pavla Jozefa Šafárika sa venuje problematike nekalej súťaže v kontexte obchodného práva. Výskum sa zameriava na ustanovenia § 41 až § 55d zákona č. 513/1991 Zb. Obchodného zákonníka (ďalej len „Obchodný zákonník"), ktoré upravujú rôzne formy nekalosúťažného konania. Osobitná pozornosť je venovaná generálnej klauzule, klamlivej reklame, parazitovaniu na povesti, zľahčovaniu či porušeniu obchodného tajomstva. Tím sa taktiež zaoberá tzv. sudcovskými skutkovými podstatami, teda prípadmi nekalej súťaže, ktoré nie sú explicitne zakotvené v zákone, ale boli vyvinuté súdnou praxou. Členovia tímu skúmajú sa aj procesnoprávne nástroje ochrany, akými sú neodkladné opatrenia, nápravné opatrenia či zverejňovanie súdnych rozhodnutí.

Doména č. 2: Zmluvná pokuta

Výskumný tím z Právnickej fakulty Univerzity Pavla Jozefa Šafárika sa zaoberá inštitútom zmluvnej pokuty ako zabezpečovacieho prostriedku v kontexte zákona č. 40/1964 Zb. Občianskeho zákonníka (ďalej len „Občianskeho zákonníka" aj Obchodného zákonníka. Výskum sa sústreďuje na tri kľúčové okruhy, a to požiadavku písomnej formy ako podmienku platnosti dohody o zmluvnej pokute, súlad výšky zmluvnej pokuty s dobrými mravmi a poctivým obchodným stykom a moderačné právo súdu znížiť neprimerane vysokú pokutu podľa § 301 Obchodného zákonníka. Analyzujú sa rozhodnutia, v ktorých súdy posudzujú, či zmluvná pokuta plní svoju preventívnu a sankčnú funkciu, alebo či naopak

predstavuje neplatný právny úkon pre rozpor s dobrými mravmi. Výskumný tím sleduje aj vzťah zmluvnej pokuty k paušalizovanej náhrade škody a jej primeranosti vo vzťahu k hodnote a významu zabezpečovaného záväzku.

Doména č. 3: Trestné právo – zaistenie počítačových údajov

Výskumný tím z Právnickej fakulty Univerzity Komenského skúma trestnoprávne rozhodnutia, v ktorých sa orgány činné v trestnom konaní zaoberajú zaisťovaním elektronických a počítačových údajov. Výskumná otázka smeruje k identifikácii toho, pri akých druhoch trestnej činnosti sa v najväčšej miere vydávajú príkazy na zabezpečenie počítačových údajov. Relevantný právny rámec zahŕňa ustanovenia zákona č. 301/2005 Z. z. Trestného poriadku, zákon č. 236/2017 Z. z. o európskom vyšetrovacom príkaze a medzinárodné dokumenty, ako je Dohovor o počítačovej kriminalite z roku 2001. Tím sleduje aj aplikáciu európskych nástrojov cezhraničného získavania elektronických dôkazov vrátane pripravovaného nariadenia EÚ 2023/1543, ktoré nadobudne účinnosť v auguste 2026. Kľúčovými pojmami pri vyhľadávaní sú napríklad obsah komunikácie, prevádzkový a lokalizačný údaj, dátový nosič či cloudové prostredie.

Doména č. 4: Autorské právo – neoprávnené užívanie diel

Výskumný tím z Právnickej fakulty Univerzity Komenského sa zaoberá prípadmi, v ktorých dochádza k verejnému prenosu autorských diel bez uzavretia licenčnej zmluvy s organizáciou kolektívnej správy (OKS), čím vzniká bezdôvodné obohatenie na strane neoprávneného užívateľa. Výskumná otázka sa pýta, pri akých kategóriách diel dochádza k takémuto porušovaniu práv autorov v najväčšej miere, a to z perspektívy OKS a koncových vysielateľov. Právny základ tvorí najmä zákon č. 185/2015 Z. z. Autorský zákon, Občiansky zákonník a relevantné európske smernice o autorských právach v informačnej spoločnosti a nájomnom práve. Tím sa osobitne zameriava na inštitút rozšírenej hromadnej licenčnej zmluvy a na judikatúru Súdneho dvora EÚ dotýkajúcu sa pojmu verejný prenos. Analyzovaná je aj otázka primeranej odmeny a hranice medzi oprávneným a neoprávneným užívaním chránených diel.

Otázka 1 Aké výskumné otázky alebo hypotézy ste sledovali v rámci vášho právneho výskumu pred zapojením automatizovanej analýzy?

Uveďte konkrétne výskumné otázky, ktoré ste sa snažili zodpovedať. Je dôležité pochopiť, do akej miery boli tieto otázky formulované dostatočne pred spustením analýzy alebo sa vyvíjali v priebehu projektu.

Príklad: *Napr.: Zaujímalo ma, ako súdy v SR pristupujú k moderácii zmluvnej pokuty – konkrétne: aké faktory berú do úvahy.*

Doména č. 1 – Nekalá súťaž	V rámci formulácie výskumných otázok som sa zaoberala predovšetkým tým, ktoré dôkazné prostriedky sú v konaniach najčastejšie využívané, akých nárokov sa dotknuté subjekty najčastejšie domáhajú a s akým výsledkom, ktoré skutkové podstaty nekalej súťaže sú najčastejšie vzájomne kombinované a taktiež, ktoré hospodárske odvetvie je najviac nekalosúťažným konaním dotknuté a iné. Vzhľadom k tomu, že zber atribútov bol rozsiahly, dataset poskytuje priestor aj pre zodpovedanie iných výskumných otázok, ktoré sa v rámci vyhodnocovania výsledkov môžu ukázať ako relevantné.
Doména č. 2 – Zmluvná pokuta	Výskum sa sústredil na rozhodovaciu prax slovenských súdov pri uplatňovaní zmluvnej pokuty v obchodnoprávnych vzťahoch. Sledovali sme najmä to, ako súdy posudzujú primeranosť zmluvnej pokuty, v ktorých prípadoch pristupujú k jej zníženiu podľa § 301 Obchodného zákonníka a aké právne alebo skutkové okolnosti považujú za rozhodujúce. Výskumné otázky boli formulované už pred analýzou, ale v priebehu práce sa spresňovali. Ukázalo sa, že nestačí sledovať iba výsledok sporu, ale je potrebné rozlišovať aj spôsob argumentácie súdu, povahu zabezpečenej povinnosti, výšku zmluvnej pokuty, postavenie strán a to, či súd skutočne aplikoval moderačné oprávnenie alebo iba všeobecne hodnotil primeranosť.
Doména č. 3 – Trestné právo	Zaujímalo nás aké digitálne údaje používajú súdy pri dokazovaní kybernetickej kriminality a čo je právnym titulom na ich získanie a aké druhy trestov a ich výmery sa ukladali pri počítačových trestných činoch.
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Výskum bol zameraný najmä na identifikáciu rozhodovacích trendov slovenských súdov v sporoch týkajúcich sa ochrany autorských práv a neoprávneného verejného prenosu diel. Zaujímalo nás, aké faktory vplývajú na úspešnosť žalôb kolektívnych správcov práv, ako súdy pristupujú k otázke bezdôvodného obohatenia a aké typy právnej argumentácie sa v rozhodnutiach opakujú.

	Výskumné otázky sa počas projektu postupne vyvíjali. Pôvodne boli formulované pomerne všeobecne, avšak po vytvorení datasetu a oboznámení sa so štruktúrou rozhodnutí bolo potrebné ich spresniť podľa reálne dostupných údajov a kvality odôvodnení jednotlivých rozhodnutí.

Právne výskumné tímy si pred analýzou stanovili výskumné otázky rôznej šírky a miery konkrétnosti, pričom všetky smerovali k empiricky overiteľným javom v rozhodovacej praxi slovenských súdov. Zhodným znakom naprieč výskumnými tímami bola určitá miera iteratívneho vývoja výskumných otázok. Žiaden z týchto tímov nezostal striktne pri pôvodnej formulácii a priebežne ich prispôsoboval obsahu a štruktúre reálne dostupných rozhodnutí.

Otázka 2 Akým spôsobom ste typicky vykonávali právny výskum v tejto oblasti pred využitím metód systematickej kontextovej analýzy?

Opíšte váš predchádzajúci metodologický prístup – napr. manuálna analýza kľúčových rozhodnutí, doktrinálna analýza, konzultácia s kolektívom, komparatívna analýza. Pomôže nám to pochopiť, v akom kontexte sú nové metódy hodnotené.

Príklad: *Napr.: Zvyčajne som čítal 10–20 rozhodnutí, ktoré som vybral podľa vlastného uváženia a z nich vyvodzoval závery. Teraz som po prvýkrát pracoval s datasetom 100 rozhodnutí.*

Doména č. 1 – Nekalá súťaž	V rámci práce s judikatúrou som zvyčajne som realizovala doktrinálny výskum, skrze ktorého boli analyzované konkrétne rozhodnutia, ktoré s problematikou súviseli, prípadne bol analyzovaný vývoj judikatúry v určitej oblasti.
Doména č. 2 – Zmluvná pokuta	Pred využitím systematickej kontextovej analýzy sme právny výskum vykonávali najmä tradičným spôsobom. Išlo o manuálne vyhľadávanie relevantných rozhodnutí, čítanie vybraných rozsudkov, doktrinálnu analýzu odbornej literatúry a porovnávanie argumentácie súdov. Výber rozhodnutí bol do určitej miery ovplyvnený dostupnosťou judikatúry a tým, ktoré rozhodnutia sa javili ako právne významné. Tento prístup umožňoval hlbokú právnu analýzu, ale bol časovo náročný a menej vhodný na overenie širších trendov v rozhodovacej praxi. Práca s datasetom umožnila pozerať sa na väčší súbor rozhodnutí systematickejšie.
Doména č. 3 – Trestné právo	Manuálna analýza rozhodnutí na základe kľúčových slov.

Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Pred využitím systematickej kontextovej analýzy bol právny výskum vykonávaný prevažne klasickými doktrinálnymi metódami. Typicky išlo o manuálnu analýzu menšieho počtu rozhodnutí vybraných na základe právnej relevancie alebo citovanosti v odbornej literatúre. Pri tejto metodológii však existovalo riziko selektívneho výberu rozhodnutí a obmedzenej reprezentatívnosti záverov. Práca s väčším datasetom umožnila identifikovať širšie rozhodovacie trendy a opakujúce sa vzory, ktoré by pri manuálnom výskume nemuseli byť zjavné.

Všetky právne výskumné tímy sa pred zapojením automatizovaných metód spoliehali prevažne na tradičný doktrinálny výskum, teda na manuálne čítanie a analýzu menšieho počtu rozhodnutí vybraných na základe právnej relevancie alebo citovanosti v odbornej literatúre. Spoločným menovateľom všetkých prístupov bola síce možnosť hlbokej právnej analýzy, ale aj časová náročnosť a obmedzená reprezentatívnosť záverov. Výskumná práca s väčším datasetom umožnila po prvý raz systematicky sledovať širšie rozhodovacie trendy, ktoré by pri manuálnom výskume nemuseli byť zjavné.

5. Identifikácia a zber vhodnej vzorky údajov

So stanovením výskumných otázok, resp. vedeckých hypotéz je úzko spojený výber vhodnej vzorky právnych textov, ktoré budú podrobené bližšiemu preskúmaniu. Vo všeobecnosti možno metódu systematickej kontextovej analýzy aplikovať **vo vzťahu k akémukoľvek druhu právnej dokumentácie**. Analyzovať možno ustanovenia právnej úpravy, text súdnych rozhodnutí alebo rozhodnutí iných príslušných štátnych orgánov, uzatvorených zmlúv, obchodných podmienok, a pod. Základným predpokladom v tejto súvislosti je predovšetkým **dostupnosť právneho textu.**

Právne predpisy sú povinne zverejňované a jednoducho prístupné prostredníctvom oficiálnych webových sídiel, konkrétne vo vnútroštátnych podmienkach prostredníctvom informačného systému verejnej správy Slov-Lex, ktorého správcom a prevádzkovateľom je podľa § 8 ods. 1 zákona č. 400/2015 Z. z. o tvorbe právnych predpisov a o Zbierke zákonov Slovenskej republiky Ministerstvo spravodlivosti SR, a vo vzťahu k právu Európskej únie na portáli EUR-Lex prevádzkovanom Úradom pre vydávanie publikácií Európskej únie, ktorý je k dispozícii vo všetkých 24 úradných jazykoch Európskej únie.

Pokiaľ ide o dostupnosť **súdnych rozhodnutí**, tieto sú spravidla zverejňované buď samotnou súdnou autoritou, ktorá rozhodnutia vydáva, alebo iným príslušným orgánom. V slovenských podmienkach ide napríklad o Ministerstvo spravodlivosti SR, ktoré podľa § 82a ods. 1 zákona č. 757/2004 Z. z. o súdoch na svojom webovom sídle zabezpečuje zverejnenie právoplatných rozhodnutí vo veci samej, rozhodnutí, ktorými sa končí konanie, rozhodnutí o neodkladnom opatrení a rozhodnutí o odklade vykonateľnosti rozhodnutia správneho orgánu v anonymizovanej podobe, a to do 15 pracovných dní odo dňa nadobudnutia právoplatnosti rozhodnutia. Podľa § 82a ods. 5 zákona č. 757/2004 Z. z. o súdoch sú súdy súčasne povinné sprístupniť verejnosti na základe žiadosti o sprístupnenie informácií podľa § 14 zákona č. 211/2000 Z. z. o slobodnom prístupe k informáciám (ďalej len „zákon o slobode informácií“) **všetky súdne rozhodnutia**, **vrátane neprávoplatných rozhodnutí a rozhodnutí, ktoré nie sú rozhodnutiami vo veci samej** s výnimkou:

- súhlasu s poskytnutím údajov, ktoré sú predmetom telekomunikačného tajomstva podľa osobitného predpisu,

- súhlasu na použitie informačno-technických prostriedkov a príkazu podľa Trestného poriadku;

súhlas však sprístupnia osobe, ktorej sa súhlas týka, ak tým nebude ohrozené alebo zmarené plnenie úloh štátu. Súdy pritom robia opatrenia na ochranu práv a právom chránených záujmov.

Okrem súdov disponujú oprávnením prijímať **rozhodnutia** aj ďalšie štátne orgány, získanie prístupu ku ktorým však môže byť problematické. Kým niektoré štátne orgány nimi prijímané rozhodnutia zverejňujú na svojich webových sídlach, iné si takýto transparentný postup neosvojili. Prístup k takýmto rozhodnutiam však nie je *a priori* vylúčený. Do úvahy totiž prichádza možnosť využitia inštitútu žiadosti o sprístupnenie informácií podľa § 14 zákona č. 211/2000 Z. z. o slobodnom prístupe k informáciám (zákon o slobode informácií) s cieľom sprístupnenia rozhodnutí dotknutej autority výskumníkovi. Súčasne je však potrebné konštatovať, že samotné podanie žiadosti o sprístupnenie informácií automaticky neznamená, že požadované rozhodnutia budú aj skutočne sprístupnené. Zverejňovanie súdnych alebo iných rozhodnutí pritom nie je štandardom, ale skôr špecifikom slovenského právneho prostredia.[17] V iných štátoch sú pravidelne zverejňované predovšetkým rozhodnutia vyšších súdnych inštancií, pričom rozhodnutia ostatných súdov spravidla nie sú voľne sprístupnené.

Relevantný druh právneho textu, ktorý môže slúžiť na účely systematickej kontextovej analýzy práva, predstavujú aj **zmluvy,** získanie prístupu ku ktorým je však veľmi náročné. Vo všeobecnosti neexistuje zákonná povinnosť zverejňovať zmluvy uzatvárané medzi súkromnoprávnymi subjektmi v centralizovanej databáze. Výnimku tvoria v tejto súvislosti iba tzv. **povinne zverejňované zmluvy**, ktoré sú podľa § 5a ods. 1 zákona o slobode informácií uzatvárané povinnými osobami a ktoré obsahujú informáciu, ktorá sa získala za finančné prostriedky, s ktorými hospodária právnické osoby verejnej správy vrátane neštátnych účelových fondov, alebo sa týka používania týchto finančných prostriedkov, nakladania s majetkom štátu, majetkom obce, majetkom vyššieho územného celku alebo majetkom právnických osôb zriadených zákonom alebo na základe zákona

[17] Súčasne však platí, že v poslednom období možno identifikovať iniciatívy publikácie rozhodnutí všetkých súdnych autorít aj v iných štátoch, vrátane Českej republiky. Dostupné: https://msp.gov.cz/web/msp/rozhodnuti-soudu-judikatura-

alebo nakladania s finančnými prostriedkami Európskej únie. Aj v prípade týchto zmlúv však aplikovateľná úprava ustanovuje viaceré výnimky, kedy sa týmto spôsobom nezverejňujú napríklad služobné zmluvy, pracovné zmluvy, zmluvy týkajúce sa burzových obchodov a ich sprostredkovania, zmluvy týkajúce sa cenných papierov alebo iných finančných nástrojov, atď. Vo všeobecnosti majú zmluvné strany záujem predovšetkým na ochrane dôvernosti svojich zmluvných práv a povinností vyplývajúcich z uzavretých zmluvných dojednaní, za účelom čoho sú v zmluvách často inkorporované ustanovenia týkajúce sa ochrany dôvernosti informácií (§ 271 Obchodného zákonníka), obchodného tajomstva (§ 17 a nasl. Obchodného zákonníka), know-how, osobných údajov a pod. Uvedené automaticky nevylučuje možnosť použitia metódy systematickej kontextovej analýzy aj vo vzťahu k zmluvám, predpokladom však bude získanie prístupu k dostatočnej vzorke zmlúv, ktoré majú byť podrobené analýze. Na tomto mieste je nutné upozorniť aj na kvalitu zdrojových dát. Zmluvy sa ukladajú síce vo formáte PDF (Portable Document Format), ale viackrát ako uložené obrázky alebo v kvalite, ktorá je pre automatizované spracovanie problematická.[18] Nástroje pre OCR (Optical Character Recognition) môžu mať výrazný problém s takouto kvalitou vstupných dokumentov.

Okrem Centrálneho registra zmlúv môžu ako zdroj zmluvných dojednaní slúžiť napríklad databázy zmlúv advokátskych kancelárií alebo obchodných spoločností za predpokladu spôsobilosti preukázania oprávneného prístupu k takýmto zmluvám a ich dôkladnej anonymizácie. Okrem zmlúv však ako podklad k analýze môžu byť použité aj **všeobecné obchodné podmienky** (§ 273 Obchodného zákonníka), ktoré sú spravidla voľne dostupné na webových sídlach rôznych spoločností, prevádzkovateľov e-commerce platforiem alebo poskytovateľov sprostredkovateľských služieb. Tabuľka č. 1 obsahuje prehľad dostupných zdrojov právnych textov, ktoré môžu slúžiť ako zdroj pre účely systematickej kontextovej analýzy práva.

	Názov portálu	Dostupné
Právne predpisy		

[18] Podľa Metodického usmernenia Ministerstva investícií, regionálneho rozvoja a informatizácie Slovenskej republiky č. 006408/2021/oMK-3 zo dňa 25.03.2021 k vykonávaniu zaručenej konverzie je povolené rozmedzie medzi 200 až 300 DPI pre čiernobiely text.

Právna úprava Európskej únie	EUR-Lex	https://eur-lex.europa.eu/
Vnútroštátna legislatíva	Slov-Lex	https://www.slov-lex.sk/domov
Súdne rozhodnutia		
Európsky súd pre ľudské práva	HUDOC	https://hudoc.echr.coe.int/
Súdny dvor, Všeobecný súd a Súd pre verejnú službu	Curia	https://curia.europa.eu/
Najvyšší súd SR (103 238 rozhodnutí)	Rozhodnutia	https://www.nsud.sk/rozhodnutia/
Ústavný súd SR (50 103 rozhodnutí)	Rozhodnutia	https://www.ustavnysud.sk/rozhodovacia_cinnost/rozhodnutia
Najvyšší správny súd SR (15 430 rozhodnutí)	Rozhodnutia	https://www.nssud.sk/#
Rozhodnutia okresných súdov, krajských súdov, Mestského súdu, Najvyššieho súdu SR, Správneho súdu a Špecializovaného trestného súdu (4 509 025 rozhodnutí)	Rozhodnutia	https://www.justice.gov.sk/sudy-a-rozhodnutia/sudy/rozhodnutia/

Tabuľka 1 – Prehľad dostupných zdrojov právnych textov (stav k 14.9.2025)

Okrem dostupnosti právneho textu je potrebné zohľadniť aj ďalšie skutočnosti, ktoré môžu ovplyvniť praktické využitie metódy systematickej kontextovej analýzy práva. Ide predovšetkým o možnosti vyhľadávania relevantného právneho textu v dostupných databázach, štruktúrovanosť dostupných údajov, technické spracovanie dokumentácie, spôsob stiahnutia identifikovaného právneho textu z databázy a pod. V tomto smere je vhodné poukázať na niekoľko špecifík slovenskej právnej úpravy. Podľa ustanovenia § 21g (1) zákona o slobode informácií povinná osoba sprístupňuje informácie na účely ich opakovaného použitia v podobe a spôsobom, ktoré umožňujú jej technické podmienky, prednostne v elektronickej podobe ako otvorené údaje umožňujúce automatizované spracovanie spolu s ich metaúdajmi. Formáty a metaúdaje musia v čo najväčšom rozsahu spĺňať formálne otvorené štandardy zákona o slobode informácií. Inými slovami, ak je to možné, povinná osoba sprístupní tzv. API (application programming interface), ktoré umožňuje výskumníkom pristupovať k údajom aj do budúcna, bez nutnosti si neustále vyžiadať sprístupnenie nových právnych textov. Druhým aspektom je štruktúra, formát

údajov a ich správa verejnými subjektami. To upravuje vyhláška Úradu podpredsedu vlády Slovenskej republiky pre investície a informatizáciu č. 78/2020 Z. z. o štandardoch pre informačné technológie verejnej správy. Predmetná vyhláška rieši aj otázku otvorených údajov (§ 39), čo možno využiť v rámci výskumu ako doplňujúce, kontextové informácie. Zoznam otvorených datasetov je možné nájsť v Národnom katalógu otvorených dát.[19]

Ďalšou dôležitou požiadavkou v tejto súvislosti je schopnosť **odôvodniť výber konkrétnej vzorky právneho textu** podrobenej analýze. V tradičnom právnom výskume je totiž výber skúmaných právnych zdrojov daný výlučne výskumníkom, keďže *„právni experti, ktorí interpretujú rozhodnutia, skúmajú predovšetkým prípady, ktoré ich zaujímajú, často bez akejkoľvek diskusie o tom, kde tieto prípady našli alebo prečo si vybrali práve ich a nie iné. Čitateľ sa musí spoliehať na úsudok autora o tom, ktoré rozhodnutia stoja za prečítanie, tzn. ktoré rozhodnutia sú „určujúce", tzn. najlepšie ilustrujúce právny problém v dotknutom prípade."*[20] Problematickým môže byť identifikácia všetkých právnych textov relevantných pre zodpovedanie stanovených výskumných otázok, alebo vymedzenie reprezentatívnej vzorky právnych textov. Odôvodnenie výberu konkrétnej vzorky právnych zdrojov môže byť dané rôznymi skutočnosťami, najčastejšie vymedzením určitého časového obdobia (napríklad rozhodovacia činnosť príslušného orgánu za celé obdobie pôsobnosti, prípadne výlučne od určitého určujúceho okamihu, napr. po prijatí novej právnej úpravy), príslušných orgánov (v prípade postihovania porušení autorských práv možno analyzovať rozhodnutia podľa využitých prostriedkov ochrany – civilné, administratívne alebo trestnoprávne), stupňa rozhodovacej činnosti (rozhodnutia prvostupňových súdov, prípadne rozhodnutia prijaté súdmi vyššej inštancie), a pod. Samotná identifikácia relevantných právnych textov z celej množiny (populácie) právnych textov je zaujímavá výskumná otázka. Napríklad ak ide o výber rozhodnutí dotýkajúcich sa oblasti počítačovej kriminality z celkovej množiny všetkých rozhodnutí v trestných veciach, pôjde rádovo o státisíce rozhodnutí. Z tohto dôvodu je nevyhnutný automatizovaný, resp. poloautomatizovaný výber týchto rozhodnutí.

[19] Dostupné: https://data.slovensko.sk/

[20] HALL, Mark A. WRIGHT, Ronald F. Systematic content analysis of judicial opinions. In *California Law Review*, 2008, roč. 96, č. 1, s. 63–79.

Pre identifikáciu a zber vhodnej vzorky údajov bola relevantná nasledujúca otázka a odpovede právnych výskumných tímov.

Otázka 3 Narazili ste na obmedzenia plynúce z obsahu alebo dostupnosti právnych textov vo vašom datasete? Ak áno, ako tieto obmedzenia ovplyvnili vaše výskumné otázky?

Metóda systematickej kontextovej analýzy je limitovaná skutočným obsahom analyzovaných textov. Chceme vedieť, či ste museli vaše výskumné otázky upraviť alebo zúžiť z dôvodu absencie určitých informácií v rozhodnutiach, zmluvách alebo iných právnych textoch.

Príklad: *Napr.: Chcel som skúmať, aké dôkazy súdy brali do úvahy pri hodnotení digitálnej stopy, ale väčšina rozhodnutí neobsahovala podrobné odôvodnenie. Išlo o skrátené formy (trestné rozkazy, dohody o vine a treste).*

Doména č. 1 – Nekalá súťaž	Pri rozhodnutiach v obchodnoprávnych veciach, resp. civilnej agende, ako najproblematickejšie vnímam asi štruktúru odôvodnenia, keďže niektoré rozhodnutia sú „systematizovanejšie“ ako iné – napríklad v otázkach sumarizácie vykonaných dôkazov a popisu skutočností, ktoré z nich vyplývajú.
Doména č. 2 – Zmluvná pokuta	Áno. Najväčším obmedzením bola rôzna kvalita a podrobnosť odôvodnení súdnych rozhodnutí. Nie všetky rozhodnutia obsahovali dostatok informácií o tom, prečo súd považoval zmluvnú pokutu za primeranú alebo neprimeranú. V niektorých rozhodnutiach bola zmluvná pokuta riešená iba okrajovo, hoci formálne bola predmetom sporu. To ovplyvnilo možnosť presne kódovať dôvody moderácie alebo odmietnutia moderácie. Výskumné otázky sa preto museli zúžiť z ambície hodnotiť kvalitu právnej argumentácie na identifikáciu viditeľných znakov, ktoré sa v texte rozhodnutia dali spoľahlivo zachytiť.
Doména č. 3 – Trestné právo	Áno, spravidla relevantná vzorka nebola použiteľná pre kvantitatívne metódy skúmania, z dôvodu, že počet rozhodnutí bol štatisticky nevýznamný. Alebo počet rozhodnutí v relevantnej vzorke bol príliš vysoký.
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Áno. Významným obmedzením bola kvalita a rozsah odôvodnení jednotlivých rozhodnutí. Viaceré rozhodnutia obsahovali len stručné alebo formalizované odôvodnenie, pričom absentovali podrobnejšie úvahy súdu o právnej argumentácii strán alebo o dôvodoch priznania či nepriznania nároku. Obmedzenia sa prejavili aj pri niektorých atribútoch, ktoré nebolo možné spoľahlivo identifikovať z každého rozhodnutia. V dôsledku toho bolo potrebné niektoré výskumné otázky

	zúžiť najmä na tie aspekty, ktoré boli v datasete dostatočne zachytené.

Všetky výskumné tímy, ktoré poskytli odpoveď, sa s obmedzeniami stretli a považujú ich za jeden z hlavných praktických problémov aplikácie danej metódy. Najčastejšie identifikovaným obmedzením bola premenlivá kvalita a podrobnosť odôvodnení súdnych rozhodnutí. Mnohé rozhodnutia obsahovali len stručné alebo formalizované odôvodnenie bez hlbšej právnej argumentácie. Naprieč tímami je teda zrejmé, že skutočný obsah analyzovaných právnych textov je kľúčovým limitom celej metódy a výskumné otázky sa musia prispôsobovať tomu, čo je v rozhodnutiach reálne obsiahnuté. Inými slovami, metódy nedokážu odpovedať na otázky, ktoré nie sú obsiahnuté v právnych textoch.

6. Kódovanie právneho textu

Najdôležitejšou fázou metódy systematickej kontextovej analýzy právneho textu je jeho samotné kódovanie. Cieľom tohto kódovania je vytvorenie vhodnej reprezentácie jednotlivých právnych textov v kóde, ktorý reprezentuje vymedzené **atribúty (vlastnosti)**. Kód predstavuje množinu jednotlivých prvkov, ktoré označujeme ako atribúty. Ide o konkrétne vlastnosti právneho textu, ktoré môžu reprezentovať právny text ako taký (tzv. metadáta) alebo vychádzať z obsahu právneho textu.

Metadáta zahŕňajú informácie, ktoré slúžia na identifikáciu konkrétneho právneho textu, pri súdnych rozhodnutiach pôjde napríklad o informáciu identifikujúcu súd, ktorý vydal rozhodnutie, dátum prijatia rozhodnutia, druh rozhodnutia, spisová značka a pod. **Atribúty vychádzajúce zo samotného obsahu** právnych textov môžu zahŕňať napríklad informáciu o prítomnosti odkazu na konkrétny právny predpis, resp. jeho jednotlivé ustanovenia, pojednávanie o konkrétnom právnom inštitúte, druh uložených sankcií atď., reflektujúc skutkové ako aj právne aspekty vyjadrené v skúmanom právnom texte. Alternatívou k týmto atribútom sú externé (kontextové) atribúty, ktoré dopĺňajú kontext daného rozhodnutia, resp. právneho vzťahu. Príkladom môže byť výnos subjektu za daný rok, počet zamestnancov alebo informácia o sídle podnikania, resp. trvalého bydliska.

Jednotlivé atribúty tiež možno kategorizovať na základe formátu, v akom reprezentujú údaje. V tejto súvislosti možno rozlišovať napríklad **numerické atribúty** (výška uloženej peňažnej pokuty, počet poľahčujúcich alebo priťažujúcich okolností), **kategorické**

atribúty (povaha páchateľa trestného činu, spôsob úhrady ceny predmetu zmluvy – jednorazovo, mesačne a pod.), alebo **binárne atribúty**, ktoré konštatujú prítomnosť alebo neprítomnosť určitej informácie v skúmanom právnom texte (druh trestu uloženého v trestnom konaní, napr. uloženie trestu odňatia slobody, výskyt konkrétneho právneho inštitútu v zmluve – atribúty ÁNO/NIE).

Okrem vyššie uvedeného rozoznávame aj **nezávislé a závislé atribúty**. Nezávislé (vysvetľujúce) atribúty predstavujú vstupy, ktoré používame na analýzu alebo vysvetlenie danej otázky (napr. konkrétna poľahčujúca okolnosť v rámci trestného konania). Na druhej strane, závislé (cieľové) atribúty predstavujú výstup, ktorý analyzujeme, resp. vysvetľujeme (napr. výrok súdu - vinný/nevinný). Závislé atribúty sú nevyhnutné pre analýzu, v rámci ktorej potrebujeme ohodnotené údaje (označenia). Ide o tzv. učenie s učiteľom a tieto označenia umožňujú metóde sa naučiť. Napr. na základe označení vinný/nevinný sa model môže naučiť, ktorá kombinácia nezávislých (vysvetľujúcich) atribútov vplýva na určenie výroku súdu v merite veci. Opakom učenia s učiteľom je učenie bez učiteľa. To si nevyžaduje pre samotnú analýzu závislé atribúty. Tie sú ale nevyhnutné pre overenie a vyhodnotenie správnosti výsledkov.

Hall a Wright rozlišujú nasledujúce kategórie informácií, ktoré možno kódovať:

a) identita jednotlivých strán a ich vlastnosti,
b) typ právnych otázok, ktoré boli vznesené v konkrétnom prípade a za akých okolností,
c) spôsob vyriešenia prípadu alebo problému,
d) základy pre vydanie daného rozhodnutia.[21]

Vo všeobecnosti platí, že kategórie určené pre účely kódovania môžu byť rozmanité, prihliadajúc na rôznorodosť otázok, ktoré si kladú výskumníci.[22] V určitých prípadoch môžu zahŕňať aj **informácie, ktoré nevyplývajú priamo z právneho textu**, ale ich prítomnosť možno odvodiť implicitne (napríklad aplikácia určitej zásady).

Identifikácia atribútov, ktoré majú byť extrahované z právneho textu, predstavuje pomerne náročnú úlohu pre výskumníka. Zvolené atribúty musia byť dostačujúce na účely

[21] HALL, Mark A. WRIGHT, Ronald F. Systematic content analysis of judicial opinions. In *California Law Review*, 2008, roč. 96, č. 1, s. 63–79.
[22] Ibidem.

zodpovedania stanovených výskumných otázok, resp. na potvrdenie/vyvrátenie vymedzených vedecký hypotéz. V každom jednotlivom prípade je pritom potrebné posudzovať **úplnosť atribútov**. Nezahrnutie relevantných atribútov do analýzy môže viesť k skresleným výsledkom, ktoré nezohľadňujú právny problém v jeho celosti. Výskumník musí zvážiť, či jednotlivý právny zdroj poskytuje dostatočný základ (obsahuje potrebné informácie) na preskúmanie stanovených výskumných otázok alebo overenie výskumných hypotéz.

Po finalizácii zoznamu atribútov je potrebné preskúmať individuálne právne texty s cieľom určenia informácií zaznamenaných vo zvolenom spôsobe reprezentácie údajov. Metódy umožňujúce extrakciu atribútov zo skúmaných právnych zdrojov sú rôzne. Vo väčšine prípadov výskumník manuálne prečíta vybrané právne zdroje a zaznamená svoj popis atribútov objektívnym a reprodukovateľným spôsobom. Na urýchlenie tohto procesu a uľahčenie preskúmania väčšieho súboru právnych zdrojov je možné použiť automatizované metódy extrakcie atribútov. Podobne rôznorodé sú aj modely reprezentácie údajov, ktoré si môže výskumník zvoliť na zaznamenanie informácií extrahovaných z vybraných právnych zdrojov. Vo väčšine prípadov sú extrahované informácie uložené vo forme dátovej tabuľky. Riadky tejto tabuľky predstavujú jednotlivé objekty (napr. súdne rozhodnutia, zmluvy) a stĺpce príslušné atribúty.

Pre identifikáciu a zber vhodnej vzorky údajov boli relevantné nasledujúce otázky a odpovede právnych výskumných tímov.

Otázka 4 **Nakoľko zachytávali zvolené atribúty právne relevantné informácie pre vašu výskumnú otázku?**

Kódovanie právnych textov stojí a padá s výberom atribútov. Kódovanie ste vykonávali po predchádzajúcej diskusii. Uveďte, ktoré atribúty si myslíte, že by ste inak zadefinovali, resp. napokon Vám chýbali.

Príklad: *Napr.: Atribút „výška zmluvnej pokuty" bol veľmi dôležitý, ale chýbal mi atribút rozlišujúci, či bola pokuta moderovaná z úradnej povinnosti alebo na základe návrhu strany – toto rozlíšenie má právny dosah, ale nebolo zachytené.*

Doména č. 1 – Nekalá súťaž	Ako čiastočne problematické som vnímala „kódovanie" hospodárskeho odvetvia, ktoré s nekalou súťažou súviselo, keďže som očakávala vyššiu mieru typizácie odvetví, avšak tieto boli pomerne rôznorodé a doposiaľ uvažujem nad ich kategorizáciou.

Doména č. 2 – Zmluvná pokuta	Zvolené atribúty zachytávali významnú časť právne relevantných informácií, najmä výšku zmluvnej pokuty, výsledok konania, typ súdu, právny základ nároku a prípadnú aplikáciu § 301 Obchodného zákonníka. Ako právne dôležité sa ukázalo rozlišovať, či súd zmluvnú pokutu priznal v plnej výške, znížil ju, zamietol nárok alebo sa otázkou primeranosti zaoberal iba nepriamo. Chýbalo nám však jemnejšie rozlíšenie dôvodov moderácie, napríklad nepomer medzi výškou pokuty a zabezpečenou povinnosťou, dĺžka omeškania, zavinenie, ekonomické postavenie strán alebo opakované porušenie povinnosti. Tieto faktory majú právny význam, ale často sa nedajú jednoducho zachytiť jedným atribútom.
Doména č. 3 – Trestné právo	Áno zachytávali, väčšia vnútorná previazanosť jednotlivých atribútov a ich analýza by priniesla vyššiu pridanú hodnotu (napr. spojenie lokalizačného údaja, alebo prevádzkového údaja s miestom činu, resp. prepojenie konkrétneho lokalizačného údaju s konkrétnym titulom získania, ktorý v mnohých analyzovaných rozhodnutiach chýbal.
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Zvolené atribúty vo všeobecnosti zachytávali základné právne relevantné informácie potrebné pre analýzu sporov z oblasti autorského práva. Za najvýznamnejšie považujem atribúty týkajúce sa typu nároku, existencie licenčného vzťahu, výsledku konania a typu porušenia autorských práv. Na druhej strane v datasete absentovali niektoré jemnejšie kvalitatívne aspekty právnej argumentácie, napríklad spôsob interpretácie pojmu verejný prenos alebo intenzita dokazovania zo strany žalobcu. Tieto aspekty sa pri binárnom alebo kategorickom kódovaní zachytávajú len veľmi obmedzene.

Právne výskumné tímy sa zhodli, že zvolené atribúty zachytávali základnú vrstvu právne relevantných informácií, avšak pri bližšom skúmaní sa ukázali viaceré medzery. Naprieč výskumnými tímami prevláda záver, že atribúty pokrývali formálnu úroveň rozhodnutí dostatočne, ale substantívna, hodnotová vrstva právneho textu ostávala mimo dosahu.

Otázka 5 **Stretli ste sa s prípadmi, kde bol atribút definovaný spôsobom, ktorý nezodpovedal právnej realite alebo ustálenej právnej terminológii?**

Pri automatizovanom spracovaní právneho textu môže dôjsť k zjednodušeniu alebo skomoleniu právnych pojmov. Zaujíma nás, či kategórie atribútov zodpovedali skutočnému právnemu obsahu, alebo či boli niektoré zaradenia nesprávne, nejasné alebo zavádzajúce.

Príklad: *Napr.: Atribút „typ konania“ mal kategóriu „civilné“, ale niektoré konania boli svojou povahou skôr správne. To záviselo to od interpretácie. Alebo: kategória „vyhovenie žalobe“ bola binárna, hoci reálne súdy rozhodujú v spektre (čiastočné vyhovenie, zmena formulácie a pod.).*

Doména č. 1 – Nekalá súťaž	Áno. V jednom rozhodnutí nebolo definované či ide o rozsudok, alebo uznesenie, ale bolo uvedené, že ide o rozhodnutie. S inými anomáliami som sa asi nestrela.
Doména č. 2 – Zmluvná pokuta	Áno. V niektorých prípadoch sa ukázalo, že kategórie boli pre právnu realitu príliš zjednodušené. Napríklad samotné označenie výsledku ako vyhovenie alebo zamietnutie nevystihuje situácie, keď súd priznal časť zmluvnej pokuty, znížil ju podľa § 301 Obchodného zákonníka alebo nárok posúdil cez iný právny základ. Podobne atribút aplikácie § 301 vyžaduje právne citlivé čítanie, pretože súd môže hovoriť o primeranosti bez výslovnej aplikácie moderačného oprávnenia. Preto je potrebné pri takýchto atribútoch doplniť metodické vysvetlenie a príklady hraničných situácií.
Doména č. 3 – Trestné právo	Nie.
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Áno. V niektorých prípadoch boli atribúty definované príliš všeobecne alebo binárne, hoci právna realita bola podstatne komplexnejšia. Napríklad výsledok konania bol často redukovaný na úspech alebo neúspech žaloby, hoci v praxi existovali aj čiastočné úspechy alebo procesné rozhodnutia s odlišným významom. Podobne pri kategorizácii právnych základov nárokov bolo niekedy náročné jednoznačne rozlíšiť medzi nárokmi z bezdôvodného obohatenia a nárokmi s licenčným základom, keďže súdy tieto otázky v odôvodneniach prelínali.

Väčšina právnych výskumných tímov sa s prípadmi, kedy množina atribútov nezodpovedala právnej realite, stretla, hoci v rôznej miere. Spoločným záverom je, že právna realita je spravidla komplexnejšia, ako umožňuje binárna alebo kategorická klasifikácia, a že kódovacie kategórie musia byť doplnené metodickým vysvetlením a príkladmi hraničných situácií.

Otázka 6 Ak by ste mohli navrhnúť zmeny v zozname atribútov pre váš dataset, aké by to boli? Čo by ste pridali, odobrali alebo inak formulovali?

Táto otázka je kľúčová pre budúci výskum. Vaše odborné poznanie oblasti práva je nenahraditeľné pri identifikácii toho, čo skutočne rozlišuje právne relevantné prípady a čo je z právneho hľadiska bezvýznamné.

Príklad: *Napr.: Pridal by som atribút „aplikácia generálnej klauzuly nekalej súťaže verzus konkrétnej skutkovej podstaty“, pretože tieto dve cesty majú diametrálne odlišné procesné dôsledky, ale v datasete sa zlievali do jednej kategórie.*

Doména č. 1 – Nekalá súťaž	Pridala by som ešte jeden atribút a to či pri neodkladných opatreniach, prípadne priebežných opatreniach (za predchádzajúcej právnej úpravy OSP) bol žalobca vyzvaný na podanie žaloby vo veci samej, resp. možno ďalšie atribúty, ktoré by uchopili, akých nárokov sa najčastejšie dotknuté subjekty domáhajú skrze neodkladné opatrenia a čo im súd napokon prizná a čoho sa domáhajú žalobou vo veci samej.
Doména č. 2 – Zmluvná pokuta	Navrhli by sme doplniť najmä tieto atribúty: výslovná aplikácia § 301 Obchodného zákonníka, implicitné hodnotenie primeranosti bez výslovného odkazu na § 301, pomer medzi výškou zmluvnej pokuty a zabezpečenou povinnosťou, dĺžka omeškania alebo trvania porušenia, typ zabezpečenej povinnosti, dôvod moderácie uvedený súdom, procesná iniciatíva strany pri námietke neprimeranosti a charakter strán. Užitočné by bolo aj rozlíšenie medzi obchodnoprávnym a prípadne spotrebiteľským alebo zmiešaným kontextom, pretože právne hodnotenie zmluvnej pokuty sa môže líšiť podľa povahy vzťahu.
Doména č. 3 – Trestné právo	Nie.
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Pridali by sme atribúty zachytávajúce: • kvalitu a rozsah odôvodnenia rozhodnutia, • typ právnej argumentácie súdu, • existenciu ustálenej judikatúry, • mieru aktivity kolektívneho správcu pri dokazovaní, • a rozlíšenie medzi individuálnou a kolektívnou správou práv. Zároveň by bolo vhodné precizovať kategorizáciu výsledku konania a vytvoriť jemnejšie rozlíšenie jednotlivých typov porušení autorských práv.

Právne výskumné tímy navrhujú predovšetkým doplnenie jemnejších, právne zmysluplnejších atribútov, ktoré by lepšie zachytili substantívnu úroveň rozhodnutí. Naprieč výskumnými tímami sa opakuje požiadavka na precíznejšiu kategorizáciu výsledku konania a na pridanie atribútov, ktoré by umožnili rozlíšiť formálne a reálne uplatnenie konkrétnych právnych inštitútov.

7. Formulácia záverov, resp. pozorovaní na základe dátovej analýzy

Výstupom fázy kódovania právnych textov je reprezentácia extrahovaných informácií vo vytvorenom datasete, ktorý je podkladom pre ďalšiu analýzu, v rámci ktorej možno využiť rôzne metódy. Tieto môžu zahŕňať jednoduchú štatistickú analýzu údajov až po zložitejšie metódy strojového učenia. Využitie komplexných štatistických alebo iných metód však nie je podmienkou,[23] keďže aj jednoduché počty reprezentujúce výskyt skúmanej skutočnosti môžu mať relevantnú výpovednú hodnotu.[24] Výber metód použitých v konkrétnom prípade je na výskumníkovi, ktorý musí zohľadniť viaceré skutočnosti, vrátane množstva kódovaného právneho textu (metódy strojového učenia často vyžadujú datasety obsahujúce väčšie množstvo právneho textu), povahy stanovených výskumných otázok (napríklad ak je cieľom autora identifikácia hlavných prúdov rozhodovacej činnosti, bude nevyhnutné vytvorenie určitých skupín rozhodnutí na základe identifikovaných spoločných vlastností), a pod.

Podľa Sarkera[25] rozdeľujeme metódy strojového učenia podľa toho, akú úlohu v rámci strojového učenia riešia. Medzi jednotlivé úlohy, ktoré sa vzťahujú aj na analýzu právnych textov zaraďujeme:

a) **klasifikáciu objektov** – príslušné metódy strojového učenia umožňujú zatriediť objekty (napr. súdne rozhodnutia) podľa zvoleného kritéria (napr. podľa rozhodnutia súdu – vinný/nevinný).

b) **zhlukovanie (klastrovanie) objektov** – príslušné metódy strojového učenia umožňujú oddelenie jednotlivých objektov záujmu (napr. rozhodnutí, zmlúv) do jednotlivých skupín bez toho, aby sme poznali výslednú skupinu (učenie bez učiteľa). To nám umožňuje napríklad určiť smerovania rozhodovacej činnosti súdov.

c) **detekciu anomálií** (*outliers*) – príslušné metódy strojového učenia identifikujú tie objekty, ktoré predstavujú najväčší extrém od pomyselného priemeru naprieč všetkými skúmanými objektami a atribútmi. Identifikácia anomálnych hodnôt

[23] HALL, Mark A. WRIGHT, Ronald F. Systematic content analysis of judicial opinions. In *California Law Review*, 2008, roč. 96, č. 1, s. 63–79.
[24] SALEHIJAM, Maryam. The Value of Systematic Content Analysis in Legal Research. In *Tilburg Law Review*, 2018, roč. 23, č. 1-2, s. 36-38. https://doi.org/10.5334/tilr.5
[25] SARKER, Ihsan H. Machine learning: Algorithms, real-world applications and research directions. In *SN Computer Science*, 2021, roč. 2, č. 3, art. 160. https://doi.org/10.1007/s42979-021-00592-x

pomáha presnejšiemu zhlukovaniu, ale súčasne poskytuje spôsob, ako identifikovať tie prvky (napr. rozhodnutia, zmluvy), ktoré sú najviac vzdialené od ustálenej právnej praxe (napr. od ustálenej praxe rozhodovacieho orgánu).

d) **tvorbu asociačných pravidiel** – príslušné metódy umožňujú na základe objektov (súdnych rozhodnutí, zmlúv) a ich atribútov identifikovať určité vlastnosti v správaní atribútov. Ak nastane atribút 1 a súčasne nastane atribút 2, potom s určitou pravdepodobnosťou nastane atribút 3. Napríklad ak zmluva obsahuje určitý typ vylúčenia zodpovednosti a jej predmet sa dotýka určitého plnenia, potom s určitou pravdepodobnosťou je zmluva uzavretá na dobu neurčitú.

Pre výber vhodných metód, výstupov a záverov sme zvolili metódy, ktoré bližšie popisujeme v nasledujúcich podkapitolách. Ku každej metóde uvádzame aj relevantné odpovede od právnych výskumných tímov.

7.1 Deskriptívna štatistická analýza

Krátky popis: Základná analýza frekvencií, podielov, priemerov, mediánov, rozsahov a časových trendov jednotlivých atribútov v datasete. Vhodná je pre binárne, kategorické aj numerické atribúty. Predstavuje východiskový krok pri každom datasete, pretože umožňuje pochopiť jeho štruktúru, kvalitu a základné rozdelenie znakov.

Cieľ analýzy: Získať základný prehľad o skúmanom datasete, identifikovať dominantné javy, frekvenciu výskytu vybraných právnych inštitútov, rozhodovacích vzorcov, sankcií, typov zmluvných klauzúl alebo iných relevantných atribútov.

Otázka 7 Deskriptívna štatistická analýza (frekvencie, podiely, priemery, distribúcie atribútov)

Tento výstup ukazuje, aké hodnoty atribútov sa vyskytujú, ako často a v akom pomere. Napríklad: koľko percent rozhodnutí skončilo vyhovením, aká bola priemerná výška pokuty, ako sa rozkladajú typy konaní.

> **Príklad:** *Napr.: Zistil som, že 73 % rozhodnutí o zmluvnej pokute skončilo aspoň čiastočným vyhovením žalobe – to bol zaujímavý základ pre ďalšie otázky. Ale samotné číslo mi nepovedalo, prečo to tak je.*

Doména č. 1 – Nekalá súťaž	Zistila som, že úspešnosť v spore v pomere žalobca a žalovaný je približne 50% na 50% - keďže niekedy je návrhu vyhovené čiastočne. Užitočnosť výsledkov: 5
Doména č. 2 – Zmluvná pokuta	Deskriptívna štatistika bola veľmi užitočná ako prvý obraz o datasete. Umožnila zistiť, ako často sa v rozhodnutiach objavuje zmluvná pokuta, aký je výsledok konania, koľko rozhodnutí pracuje s § 301 Obchodného zákonníka a v akých typoch sporov sa zmluvná pokuta najčastejšie vyskytuje. Pre právny výskum je to dobrý základ, pretože ukazuje, či ide o ojedinelý jav alebo stabilnejšiu rozhodovaciu prax. Limitom je, že samotné čísla nevysvetľujú právne dôvody rozhodnutia. Percentuálny podiel znížených zmluvných pokút je preto potrebné vždy doplniť čítaním odôvodnení. Zrozumiteľnosť výsledkov: 5 Užitočnosť výsledkov: 4
Doména č. 3 – Trestné právo	Zrozumiteľnosť výsledkov: 4 Užitočnosť výsledkov: 4
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Deskriptívna analýza bola z praktického hľadiska veľmi užitočná. Poskytla základný prehľad o štruktúre datasetu, frekvencii jednotlivých typov sporov a úspešnosti jednotlivých typov nárokov. Pomohla identifikovať dominantné typy konaní a najčastejšie právne problémy. Limitom bolo, že samotné frekvencie nevysvetľovali dôvody rozhodovacej praxe. Zrozumiteľnosť výsledkov: 5 Užitočnosť výsledkov: 5

Deskriptívna štatistická analýza bola právnymi výskumnými tímami hodnotená mimoriadne pozitívne a získala najvyššie hodnotenia užitočnosti spomedzi všetkých analytických metód. Spoločným záverom naprieč tímami je, že deskriptívna štatistika predstavuje nevyhnutný a prirodzený prvý krok každej analýzy, ktorý je zároveň právne najintuitívnejšie uchopiteľný, avšak musí byť vždy doplnený hlbšou normatívnou analýzou, keďže frekvencie samy osebe nevysvetľujú príčiny sledovaných javov.

7.2 Analytická metóda - Komparatívna analýza skupín

Krátky popis: Porovnanie rozdielov medzi vopred definovanými skupinami dokumentov, napríklad medzi rozhodnutiami rôznych súdov, rôznych orgánov, rôznych období, rôznych typov konaní alebo rôznych kategórií zmlúv.

Cieľ analýzy: Určiť, či sa jednotlivé skupiny právnych textov od seba systematicky odlišujú, a ak áno, v ktorých atribútoch a akou mierou.

Otázka 8 Komparatívna analýza (rozdiely medzi skupinami – súdy, kraje, obdobia, typy konaní)

Táto analýza porovnáva skupiny dokumentov podľa metadát a hľadá systematické rozdiely. Napríklad: líšia sa rozhodnutia krajských súdov od okresných súdov? Existujú regionálne rozdiely v rozhodovaní?

Príklad: *Napr.: Výstup ukázal, že Krajský súd v Bratislave mal výrazne vyšší podiel zamietnutých žalôb v oblasti nekalej súťaže ako ostatné krajské súdy – toto by stálo za ďalší výskum, ale neviem, či je to dôvod metodiky alebo iný faktor.*

Doména č. 1 – Nekalá súťaž	Zrozumiteľnosť výsledkov: 3 Užitočnosť výsledkov: 3
Doména č. 2 – Zmluvná pokuta	Komparatívna analýza bola prínosná najmä pri porovnaní rozhodovania rôznych súdov a období. Umožnila identifikovať, či sa aplikácia § 301 alebo úspešnosť nárokov zo zmluvnej pokuty líši podľa typu súdu alebo podľa časového obdobia. Výstup pomohol formulovať ďalšie otázky, napríklad či niektoré rozdiely súvisia s typom sporov, ktoré sa dostávajú na konkrétne súdy, alebo s odlišnou argumentačnou praxou. Zrozumiteľnosť výsledkov: 4 Užitočnosť výsledkov: 4
Doména č. 3 – Trestné právo	Analýza poskytla iba čiastočný vhľad do komparatívnej analýzy, vzhľadom na zvolené atribúty. Závery boli očakávateľné. Zrozumiteľnosť výsledkov: 4 Užitočnosť výsledkov: 4
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Komparatívna analýza umožnila porovnávať rozhodovaciu prax medzi jednotlivými súdmi a časovými obdobiami. Zaujímavé boli rozdiely v prístupe jednotlivých senátov k otázke výšky náhrady alebo posudzovania verejného prenosu diel. Výsledky však bolo potrebné interpretovať opatrne vzhľadom na rozdielny počet rozhodnutí a kvalitu datasetu.

	Zrozumiteľnosť výsledkov: 4 Užitočnosť výsledkov: 4

Komparatívna analýza bola hodnotená prevažne pozitívne. Niektoré právne výskumné tímy narážali na problém s úplnosťou výstupov. Celkovo je komparatívna analýza vnímaná ako nástroj vhodný na formulovanie hypotéz a identifikáciu oblastí hodných ďalšieho skúmania, nie ako metóda schopná poskytnúť záverečné právne závery.

7.3 Analytická metóda – Korelačná analýza medzi atribútmi

Krátky popis: Skúmanie vzťahov medzi dvoma alebo viacerými atribútmi, napríklad medzi typom konania a výsledkom rozhodnutia, medzi druhom porušenia a výškou sankcie, alebo medzi predmetom zmluvy a trvaním zmluvy. Pri binárnych a kategorických dátach možno využiť kontingenčné tabuľky, Pearsonov koeficient pre binárne dáta, chí-kvadrát test nezávislosti, prípadne Cramerovo V.

Cieľ analýzy: Identifikovať právne alebo faktické súvislosti medzi atribútmi a zistiť, ktoré znaky sa typicky vyskytujú spolu.

Otázka 9 **Korelačná analýza (vzájomné závislosti medzi atribútmi)**

Korelačná analýza meria, nakoľko spolu súvisia rôzne atribúty. Napríklad: existuje vzťah medzi výškou zmluvnej pokuty a výsledkom konania? Je spojitosť medzi typom porušenia a druhom sankcie?

Príklad: *Napr.: Analýza ukázala silnú pozitívnu koreláciu medzi atribútom „aplikácia generálnej klauzuly“ a „zamietnutím žaloby“. To bolo pre mňa právne prekvapujúce a viedlo ma k ďalšej hypotéze.*

Doména č. 1 – Nekalá súťaž	Potvrdili sa viaceré aspekty, ktoré spolu súvisiaca – napríklad, že pri rozhodnutiach na základe návrhu na neodkladné opatrenie nie sú prítomné výsluchy svedkov a strán. Zrozumiteľnosť výsledkov: 5 Užitočnosť výsledkov: 5
Doména č. 2 – Zmluvná pokuta	Korelačná analýza bola užitočná pri hľadaní vzťahov medzi atribútmi, napríklad medzi výškou zmluvnej pokuty, výsledkom konania, aplikáciou § 301 a typom zabezpečenej povinnosti. Pre právnika je prínosné najmä to, že korelácie dokážu upozorniť na súvislosti, ktoré by pri manuálnom čítaní jednotlivých rozhodnutí

	nemuseli byť zjavné. Zároveň však ide o výstup, ktorý vyžaduje veľkú opatrnosť. Korelácia neznamená príčinnú súvislosť a právny záver nemožno postaviť iba na číselnom vzťahu. Výsledok je vhodný skôr ako podklad na formulovanie hypotéz. Zrozumiteľnosť výsledkov: 3 Užitočnosť výsledkov: 4
Doména č. 3 – Trestné právo	Vzhľadom na zvolené atribúty korelačná analýza vykonávaná nebola. Zrozumiteľnosť výsledkov: 4 Užitočnosť výsledkov: 4
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Korelačná analýza pomohla identifikovať určité vzťahy medzi atribútmi, napríklad medzi existenciou licenčného vzťahu a úspešnosťou žaloby. Bola užitočná pri formulovaní ďalších hypotéz. Zrozumiteľnosť výsledkov: 3 Užitočnosť výsledkov: 4

Korelačná analýza bola hodnotená ako užitočná pri odhaľovaní neočakávaných súvislostí, ale zároveň ako metodologicky najnáročnejšia pre právnu interpretáciu. Naprieč právnymi výskumnými tímami prevláda záver, že korelačná analýza je cenná pre generovanie nových výskumných hypotéz, ale vyžaduje silný filter právnej relevancie a nemôže byť použitá ako samostatný zdroj právnych záverov.

7.4 Analytická metóda – Zhluková analýza

Krátky popis: Metóda učenia bez učiteľa, ktorá rozdeľuje právne texty do skupín na základe podobnosti atribútov. Pre kategorické a binárne dáta je vhodný napríklad K-modes, prípadne hierarchické zhlukovanie s vhodnou metrikou vzdialenosti.

Cieľ analýzy: Objaviť prirodzené skupiny právnych textov bez vopred stanovených tried, napríklad typické rozhodovacie línie, typové vzory zmlúv alebo opakujúce sa kombinácie právnych znakov.

Otázka 10 Zhluková/klastrovacia analýza (objavovanie prirodzených skupín právnych textov/rozhodnutí)

Zhlukovanie bez vopred definovaných kategórií hľadá podobné právne texty/rozhodnutia a zoskupuje ich. Cieľom je objaviť opakujúce sa vzory a typické prípady bez toho, aby výskumník vopred vedel, aké skupiny existujú.

Príklad: *Napr.: Zhlukovanie odhalilo tri opakujúce sa „typy" rozhodnutí o nekalej súťaži, ktoré som predtým vnímal skôr intuitívne. Teraz vidím ich konkrétne atribútové charakteristiky.*

Doména č. 1 – Nekalá súťaž	V rámci klastrovej analýzy nevznikli také klastre, ktoré by tvorili homogénne skupiny, skôr veľké množstvo malých klastrov po 2-3 rozhodnutia. Čítam to tak, že jednotlivé spory vykazujú vysokú mieru špecifickosti a nie je možné ich jednoducho kategorizovať. Zrozumiteľnosť výsledkov: 5 Užitočnosť výsledkov: 5
Doména č. 2 – Zmluvná pokuta	Zhlukovanie pomohlo uvažovať o typických skupinách rozhodnutí. Pri zmluvnej pokute sa ukázalo, že rozhodnutia možno rozdeliť napríklad na prípady štandardného priznania nároku, prípady moderácie podľa § 301, prípady zamietnutia pre procesné alebo dôkazné dôvody a prípady, v ktorých bola zmluvná pokuta posudzovaná spolu s inými nárokmi. Výstup bol zaujímavý najmä tým, že ukázal podobnosti medzi rozhodnutiami, ktoré by právnik intuitívne vnímal oddelene. Limitom je, že pomenovanie klastrov je už právna interpretácia a nemôže ho urobiť algoritmus sám. Zrozumiteľnosť výsledkov: 3 Užitočnosť výsledkov: 3
Doména č. 3 – Trestné právo	Zhlukovacia analýza predviedla možnosť klastrovať rozhodnutia podľa jednotlivých atribútov. Zrozumiteľnosť výsledkov: 4 Užitočnosť výsledkov: 4
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Zhlukovanie bolo zaujímavé najmä pri identifikácii opakujúcich sa typov sporov a rozhodnutí. Pomohlo odhaliť podobné prípady, ktoré by pri manuálnom čítaní nemuseli byť okamžite zjavné. Na druhej strane interpretácia jednotlivých klastrov bola právne náročnejšia a vyžadovala dodatočné manuálne čítanie rozhodnutí. Zrozumiteľnosť výsledkov: 3 Užitočnosť výsledkov: 3

Zhluková analýza sa ukázala ako metodologicky zaujímavá, no interpretačne náročná metóda, ktorá tímom priniesla pohľad na podobnosti medzi rozhodnutiami, ktoré by inak zostali skryté. Zrozumiteľnosť a užitočnosť tejto metódy boli hodnotené nižšie ako pri deskriptívnej štatistike, čo naznačuje, že ide o metódu vhodnejšiu pre pokročilejšiu fázu výskumu.

7.5 Analytická metóda – Detekcia anomálií

Krátky popis: Vyhľadávanie rozhodnutí, ktoré sa významne odlišujú od ostatných. Možno využiť metódy ako LOF, Isolation Forest, ECOD, COPOD alebo PCA-based detection. Táto analýza je vhodná najmä tam, kde chceme nájsť neštandardné rozhodnutia, netypické zmluvy alebo neobvyklé prípady.

Cieľ analýzy: Identifikovať odľahlé prípady, ktoré môžu predstavovať buď chybu, alebo právne významnú výnimku z ustálenej praxe.

Otázka 11 **Detekcia anomálií (identifikácia odchýlok od typickej praxe)**

Táto metóda identifikuje právne texty/rozhodnutia, ktoré sa výrazne líšia od ostatných. Môže ísť o netypické rozhodnutia, extrémne hodnoty sankcií, neštandardné zmluvné dojednania alebo rozhodnutia odchyľujúce sa od ustálenej praxe.

Príklad: *Napr.: Analýza označila jedno rozhodnutie o zmluvnej pokute ako anomálne. Po prečítaní som zistil, že súd použil neobvyklú argumentáciu priamo odkazujúcu na doktrínu iného štátu. To by som inak prehliadol.*

Doména č. 1 – Nekalá súťaž	V rámci detekcie anomálii sa ako najanomálnejší ukázal atribút smerujúci ku hospodárskemu odvetviu a rovnako tak prítomnosť iných dôkazných prostriedkov ako tie, ktoré pomenúva (rozoznáva) CSP. Zrozumiteľnosť výsledkov: 4 Užitočnosť výsledkov: 5
Doména č. 2 – Zmluvná pokuta	Detekcia anomálií bola užitočná najmä ako nástroj na vyhľadanie rozhodnutí, ktoré si zaslúžia bližšie manuálne čítanie. Pri zmluvnej pokute môže ísť o extrémne vysokú pokutu, neobvyklý pomer k zabezpečenej povinnosti, netypickú argumentáciu alebo neštandardné použitie § 301 Obchodného zákonníka. Takéto prípady sú výskumne hodnotné, pretože často ukazujú hranice ustálenej rozhodovacej praxe. Limitom je, že anomália v dátach nemusí byť právne anomálna. Niekedy ide iba o technický

	dôsledok kódovania alebo neúplnosti textu rozhodnutia. Zrozumiteľnosť výsledkov: 4 Užitočnosť výsledkov: 4
Doména č. 3 – Trestné právo	Analýza označila pomerne veľa anomálií vzhľadom na chýbajúce dáta, ktoré budú doplnené. V zaslanom exceli neboli pri určitých atribútoch žiadne dáta. Zrozumiteľnosť výsledkov: 5 Užitočnosť výsledkov: 5
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Detekcia anomálií bola prakticky veľmi užitočná pri identifikovaní rozhodnutí, ktoré sa výrazne odlišovali od bežnej rozhodovacej praxe. Takéto rozhodnutia často obsahovali špecifickú argumentáciu alebo netypický procesný postup. Zrozumiteľnosť výsledkov: 4 Užitočnosť výsledkov: 4

Detekcia anomálií patrila spolu s deskriptívnou štatistikou medzi najlepšie hodnotené metódy a tímy ju vnímajú predovšetkým ako navigačný nástroj smerujúci pozornosť výskumníka na rozhodnutia hodné bližšieho skúmania. Spoločným záverom právnych výskumných tímov je, že detekcia anomálií má najväčšiu hodnotu ako vstupný filter pre hlbšiu kvalitatívnu analýzu, nie ako metóda priamo produkujúca právne závery.

7.6 Analytická metóda – Asociačné pravidlá

Krátky popis: Hľadanie pravidelností typu „ak nastane atribút A a súčasne atribút B, potom často nastáva aj atribút C“. Vhodné najmä pri rozsiahlejších binárnych alebo kategorických datasetoch. Použiť možno napríklad Apriori alebo FP-Growth.

Cieľ analýzy: Odhaľovať opakujúce sa kombinácie atribútov a identifikovať pravidlá, ktoré charakterizujú rozhodovaciu alebo zmluvnú prax.

Otázka 12 Asociačné pravidlá (vzory tvaru: AK x → POTOM y, napr. ak dlhý trest → nepodmienečný)

Asociačné pravidlá hľadajú pravidelnosti v tom, čo sa v textoch spoluvyskytuje. Výstupom sú pravidlá s mierou podpory (support), spoľahlivosti (confidence) a relevancie (lift).

Príklad: *Napr.: Pravidlo „ak IP adresa + dlhé trvanie skutku → nepodmienečný trest odňatia slobody" bolo štatisticky silné, ale právne som ho musel interpretovať opatrne. Korelácia nie je kauzalita.*

Doména č. 1 – Nekalá súťaž	Tieto výsledky boli formulované veľmi zrozumiteľne. Ich interpretácia si ale bude vyžadovať pre účely aplikovateľnosti (publikovateľnosti) precizáciu. Zrozumiteľnosť výsledkov: 5 Užitočnosť výsledkov: 5
Doména č. 2 – Zmluvná pokuta	Asociačné pravidlá boli zaujímavé pri hľadaní opakujúcich sa kombinácií znakov, napríklad ak sa určitý typ porušenia alebo určitá výška pokuty často spája s úspechom žalobcu alebo s moderáciou. Právne však ide o metódu, ktorú treba interpretovať veľmi opatrne. Pravidlo typu 'ak x, potom y' môže navodzovať dojem automatického právneho dôsledku, čo v rozhodovaní súdov neplatí. Pre právny výskum sú tieto pravidlá použiteľné skôr ako navigačný nástroj, ktorý ukáže, ktoré kombinácie prípadov majú byť ďalej manuálne analyzované. Zrozumiteľnosť výsledkov: 3 Užitočnosť výsledkov: 3
Doména č. 3 – Trestné právo	Asociačné pravidla potvrdili vlastnú empíriu, jednak čo sa týka atribútov trestov, ale i procesných atribútov jednotlivých digitálnych dôkazov. Zrozumiteľnosť výsledkov: 4 Užitočnosť výsledkov: 4
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Asociačné pravidlá boli zaujímavé najmä pri identifikovaní opakujúcich sa kombinácií atribútov. Nie všetky pravidlá však mali reálnu právnu relevanciu a viaceré výsledky bolo potrebné interpretovať veľmi opatrne. Zrozumiteľnosť výsledkov: 3 Užitočnosť výsledkov: 3

Asociačné pravidlá tímy hodnotia rozdielne. Pre niektoré domény boli veľmi zrozumiteľné a potvrdzujúce, pre iné vyžadujú opatrnú interpretáciu a majú skôr navigačnú hodnotu. Hodnotenia zrozumiteľnosti a užitočnosti sa líšia od 3 do 5, čo odráža heterogenitu skúseností naprieč doménami.

7.7 Analytická metóda – Prediktívna analýza

Krátky popis: Metóda učenia s učiteľom, pri ktorej sa model učí predikovať cieľový atribút na základe ostatných atribútov. Napríklad výsledok konania, typ sankcie alebo kategóriu rozhodnutia. Možno použiť rozhodovacie stromy, logistickú regresiu, random forest alebo iné interpretovateľné modely. Cieľové atribúty podľa datasetov napríklad:

- výška uloženého trestu
- výška zmluvnej pokuty po uplatnení moderačného práva súdu
- uplatnenie generálnej klauzuly alebo niektorej zo skutkových podstát nekalej súťaže

Cieľ analýzy: Zistiť, ktoré atribúty najviac vplývajú na určitý výsledok a overiť, či je možné rozhodovaciu alebo inú právnu prax modelovať na základe identifikovaných znakov.

Otázka 13 Prediktívna analýza (predikovanie výsledku konania alebo iného cieľového atribútu na základe ostatných atribútov)

Model strojového učenia sa snaží predpovedať, ako dopadne konanie alebo aká bude sankcia. Zároveň ukazuje, ktoré atribúty majú na výsledok najväčší vplyv (feature importance).

Príklad: *Napr.: Model predikoval výsledok konania s presnosťou F1=0.72. Pre mňa ako právnika bola zaujímavá nie samotná predikcia, ale zoznam najdôležitejších atribútov: „výška škody“ a „opakované porušenie“ boli kľúčové, čo zodpovedá právnemu cíteniu.*

Doména č. 1 – Nekalá súťaž	NA
Doména č. 2 – Zmluvná pokuta	Prediktívna analýza bola metodologicky zaujímavá, ale z právneho hľadiska ju vnímame skôr ako doplnkový nástroj. Užitočná nebola ani tak samotná predikcia výsledku konania, ale skôr identifikácia atribútov, ktoré model považoval za dôležité. Pri zmluvnej pokute to môže pomôcť overiť, či model pracuje s právne relevantnými premennými, napríklad výškou pokuty, typom povinnosti alebo aplikáciou § 301. Limitom je, že rozhodnutie súdu nie je mechanický výsledok atribútov. Predikcia preto nemôže nahradiť právnu analýzu a musí byť vysvetlená právnym jazykom. Zrozumiteľnosť výsledkov: 3 Užitočnosť výsledkov: 3
Doména č. 3 – Trestné právo	Vzhľadom na zvolené atribúty predikatívna analýza mala povahu skôr ad informandum. Užitočnosť výsledkov: 5
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Prediktívna analýza bola zaujímavá skôr z metodologického než praktického hľadiska.

	Najväčší prínos nebol samotný predikčný model, ale identifikácia atribútov, ktoré mali najväčší vplyv na výsledok konania. Zrozumiteľnosť výsledkov: 3 Užitočnosť výsledkov: 4

Prediktívna analýza bola hodnotená ako metodologicky zaujímavá, no z právneho hľadiska skôr doplnková metóda, pričom jej hlavná hodnota spočívala nie v samotnej predikcii, ale v identifikácii kľúčových atribútov. Zrozumiteľnosť prediktívnej analýzy je hodnotená nižšie ako pri jednoduchších metódach, čo naznačuje, že pre právne výskumné tímy bez štatistického vzdelania predstavuje interpretačnú bariéru, ktorú je potrebné prekonávať dôsledným právnym jazykovým prekladom výstupov.

8. Informovanie o dosiahnutých výsledkoch spôsobom pochopiteľným pre právnu komunitu

Záverečná fáza systematickej kontextovej analýzy sa zameriava na interpretáciu a diskusiu dosiahnutých výsledkov právne zrozumiteľným spôsobom. Výsledky musia byť interpretované vo vzťahu k stanoveným výskumným otázkam a/alebo vedeckým hypotézam v kontexte príslušnej oblasti právneho výskumu, nakoľko výsledky získané prostredníctvom štatistickej analýzy alebo rôznych metód strojového učenia predstavujú špecifické hodnoty, ktoré nemožno skúmať bez poskytnutia relevantného kontextu. Tieto metódy nešpecifikujú význam získaných čísel ani spôsob ich interpretácie, ale nevyhnutnou je v tejto súvislosti právna expertíza.

Otázka 14 Dokázali by ste výstupy dátovej analýzy priamo použiť vo vašom právnom výskume? Ak áno, akým spôsobom?

Zaujíma nás, či je možné výstupy analýzy skutočne integrovať do právneho textu, argumentácie alebo záverov, alebo zostávajú oddelené. Opíšte konkrétny prípad, ak je to možné.

Príklad: *Napr.: Deskriptívnu štatistiku by som vedel použiť v článku ako empirický základ pre tvrdenie o dominantnom type rozhodnutí. Zhlukovanie mi pomohlo štruktúrovať kapitoly podľa identifikovaných typov prípadov.*

Doména č. 1 – Nekalá súťaž	Určite, jednotlivé výstupy plánujem využiť v rámci svojej publikačnej činnosti vo viacerých výstupoch, reflektujúc na nastolenie a zodpovedanie rôznych výskumných otázok.
Doména č. 2 – Zmluvná pokuta	Áno. Výstupy dátovej analýzy sa dajú použiť najmä ako empirický základ právneho výskumu. Deskriptívnu štatistiku možno využiť pri úvodnom zmapovaní rozhodovacej praxe o zmluvnej pokute. Komparatívne výstupy môžu podporiť tvrdenia o rozdieloch medzi skupinami rozhodnutí. Zhlukovanie a detekcia anomálií môžu pomôcť pri štruktúrovaní kapitoly o typových prípadoch alebo pri výbere rozhodnutí na hlbšiu kvalitatívnu analýzu. Výstupy by som však nepoužila izolovane. Ich hodnota vzniká až vtedy, keď sú prepojené s normatívnou analýzou § 301 Obchodného zákonníka a s konkrétnou argumentáciou súdov.
Doména č. 3 – Trestné právo	Pri dodržaní väčšieho počtu rozhodnutí a vzorky určite áno.
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Áno. Výstupy bolo možné využiť ako empirický základ pre právnu argumentáciu a formulovanie všeobecných záverov o rozhodovacej praxi slovenských súdov.

	Prakticky využiteľné boli najmä deskriptívne a komparatívne analýzy. Analytické výstupy zároveň pomohli identifikovať oblasti, ktoré si vyžadovali podrobnejšie manuálne preskúmanie.

Všetky právne výskumné tímy odpovedajú kladne, avšak s dôležitými podmienkami a výhradami. Spoločným záverom je, že výstupy sú v právnom výskume použiteľné, ale nie ako autonómne zdroje záverov. Hodnota výstupov spočíva v doplňovaní, nie v nahrádzaní klasickej normatívnej analýzy.

Otázka 15 Existujú výstupy analýzy, ktoré by ste nedokázali alebo nevedeli zmysluplne interpretovať v právnom kontexte? Opíšte, prečo.

Nie každý štatistický výsledok má automaticky právny zmysel. Zaujíma nás, kde je hranica medzi štatisticky zaujímavým a právne relevantným, a kde narážate na interpretačné ťažkosti.

Príklad: *Napr.: Korelačná matica ako celok bola pre mňa ťažko čitateľná. Nevedel som, ktorú koreláciu mám brať vážne a ktorú ignorovať. Chýbal mi filter „právnej relevancie“.*

Doména č. 1 – Nekalá súťaž	Ako som zmieňovala aj vyššie v tomto dotazníku, napríklad výsledok o korelácii vo vzťahu výsluchu svedka ako dôkazu a návrhu na nariadenie neodkladného opatrenia sa síce ukázal štatisticky „pekne“, avšak z právneho hľadiska je to výsledok prirodzený a zrejmý.
Doména č. 2 – Zmluvná pokuta	Áno. Najťažšie sa interpretujú výstupy, ktoré sú štatisticky zaujímavé, ale právne nejasné. Typickým príkladom je korelačná matica alebo asociačné pravidlá bez vysvetlenia, prečo by daná súvislosť mala mať právny význam. Problémom je aj situácia, keď atribút zachytáva formálny znak, ale nie právny obsah. Napríklad samotný výsledok konania nemusí ukazovať, či súd skutočne riešil primeranosť zmluvnej pokuty. Pre právnu interpretáciu je preto potrebný filter právnej relevancie a možnosť rýchlo sa vrátiť ku konkrétnym rozhodnutiam, z ktorých výsledok vychádza.
Doména č. 3 – Trestné právo	Z tých, ktoré sú v dotazníku nie, Nevylučujem, že existujú kvantitatívne metódy analýzy, ktoré sú právne horšie interpretovateľné, alebo vôbec neinterpretovateľné.
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Niektoré komplexnejšie analytické výstupy, najmä rozsiahle korelačné matice alebo niektoré výstupy zhlukovej analýzy, boli právne interpretovateľné len obmedzene. Chýbal pri nich jasný filter právnej relevancie. V praxi bolo potrebné kombinovať analytické výstupy s manuálnym čítaním rozhodnutí a právnou expertízou.

Právne výskumné tímy sa zhodujú, že interpretačné ťažkosti sa objavujú najmä pri komplexnejších výstupoch, ktorým chýba „filter právnej relevancie“. Spoločnou požiadavkou všetkých právnych výskumných tímov je potreba prepojenia analytických výstupov s možnosťou návratu ku konkrétnym rozhodnutiam a doplnenie právnicky formulovanej interpretácie výsledkov.

Otázka 16 Môžu výstupy analýzy nahradiť právnu argumentáciu, alebo sú výlučne jej doplnkom? Vysvetlite vaše stanovisko.

Ide o zásadnú metodologickú otázku. Chceme vedieť, ako právni výskumníci vnímajú rolu kvantitatívnych výstupov vo vzťahu k tradičnej normatívnej analýze práva.

Príklad: *Napr.: Podľa môjho názoru výstupy nikdy nenahradia právnu argumentáciu. Môžu identifikovať problém, ale nie ho vysvetliť. Súd nepochopí „F1 skóre 0,7“. Musím to preložiť do právneho jazyka.*

Doména č. 1 – Nekalá súťaž	Myslím si, že tento typ výskumu má zásadný komplementárny význam, možno aj s ohľadom na to, že nie všetci právnici sú zvyknutí na túto formu interpretácie výsledkov a bude si to zrejme vyžadovať aj v rámci prezentovania výsledkov prehĺbenie diskusie.
Doména č. 2 – Zmluvná pokuta	Výstupy analýzy nemôžu nahradiť právnu argumentáciu. Môžu ju významne podporiť, spresniť a empiricky ukotviť, ale právny záver musí vychádzať z výkladu právnej normy, judikatúry, doktríny a konkrétneho skutkového kontextu. Pri zmluvnej pokute algoritmus môže ukázať, v ktorých prípadoch súdy častejšie znižujú pokutu, ale nevysvetlí, či bol postup súdu právne správny. Výstupy preto vnímame ako doplnok právnej argumentácie, nie ako jej náhradu. Ich hlavnou funkciou je identifikovať vzory, odchýlky a otázky, ktoré má právnik ďalej interpretovať.
Doména č. 3 – Trestné právo	Nemyslím si, že výstupy analýzy nahradia právnu argumentáciu, sú jej doplnkom, alebo jej základom (argumentácia z nich vychádza).
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Výstupy dátovej analýzy podľa môjho názoru nemôžu nahradiť právnu argumentáciu. Môžu však významne dopĺňať právny výskum a pomôcť identifikovať trendy, vzory alebo problematické oblasti, ktoré by pri klasickom výskume zostali nepovšimnuté. Samotné analytické výsledky však vždy vyžadujú právnu interpretáciu a zasadenie do normatívneho kontextu

Všetky právne výskumné tímy jednoznačne odmietajú, že by výstupy automatizovanej analýzy mohli nahradiť právnu argumentáciu. Zhodne tieto výstupy definujú ako doplnok a nie ako náhradu klasického právneho výskumu. Spoločným záverom je, že hlavnou funkciou automatizovanej analýzy je identifikovanie vzorov, odchýliek a výskumných otázok, ktoré má právnik ďalej interpretovať.

Otázka 17 Aké aspekty právneho textu sú podľa vás len ťažko alebo vôbec automatizovateľné a nevyhnutne vyžadujú manuálny pohľad experta?

Niektoré právne informácie sú implicitné, kontextuálne alebo vyžadujú interpretáciu. Nedajú sa zachytiť binárnym alebo kategorickým atribútom. Zaujíma nás, kde je podľa vás hranica automatizácie v právnom texte.

Príklad: *Napr.: Kvalitu právnej argumentácie súdu, aplikáciu diskrečného práva alebo hodnotenie proporcionality sankcie nedokáže zachytiť žiadny atribút. Vždy si to vyžaduje prečítanie celého rozhodnutia.*

Doména č. 1 – Nekalá súťaž	S ohľadom na môj konkrétny dataset a otázku, ktorá v praxi v podstate v otázkach nekalej súťaže rezonuje najviac je to, ktoré konkrétne konanie je spôsobilé naplniť konkrétnu skutkovú podstatu nekalej súťaže. Taktiež by bolo možno problematické „zakódovať", ktoré konkrétne skutočnosti súd viedli ku konštatovaniu, že došlo k nekalosúťažnému konaniu – jednoducho povedané, toto je ešte povolené v rámci dobrých mravov súťaže, ale toto konanie je už nepovolené a bude sankcionované.
Doména č. 2 – Zmluvná pokuta	Ťažko automatizovateľné sú najmä hodnotiace a kontextové aspekty právneho textu. Ide napríklad o kvalitu právnej argumentácie, posúdenie proporcionality, diskrečnú úvahu súdu pri moderácii zmluvnej pokuty, implicitné hodnotenie dobrých mravov alebo poctivého obchodného styku a význam skutkových okolností, ktoré súd neopíše explicitne. Automatizácia vie zachytiť výskyt určitého pojmu alebo kategórie, ale nevie spoľahlivo určiť, či súd použil presvedčivú právnu argumentáciu. Práve pri § 301 Obchodného zákonníka je ľudský právny pohľad nevyhnutný, pretože nejde len o výsledok, ale o dôvody a mieru úvahy súdu.
Doména č. 3 – Trestné právo	Voľné hodnotenie dôkazov, odôvodnenosť rozhodnutia, atribút spravodlivosti rozhodnutia, právo na spravodlivý proces apod.
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Za ťažko automatizovateľné považujeme najmä: • kvalitu právnej argumentácie, • interpretáciu neurčitých právnych pojmov, • hodnotenie proporcionality, • implicitné úvahy súdu, • a význam kontextu konkrétneho prípadu. Tieto aspekty si vo väčšine prípadov vyžadujú manuálne čítanie a odborný právny úsudok.

Právne výskumné tímy sa zhodujú, že automatizácia má jasné limity pri zachytávaní hodnotových, kontextuálnych a diskrečných aspektov právneho textu. Naprieč výskumnými tímami je konsenzuálnym záverom, že automatizácia dokáže spoľahlivo zachytiť výskyt

pojmov a formálne kategórie, ale hodnotová a argumentačná vrstva právneho textu si nevyhnutne vyžaduje manuálne prečítanie a odborný právny úsudok.

Otázka 18 **Narazili ste pri kódovaní alebo interpretácii výsledkov na príklady, kde automatizácia produkovala zavádzajúci alebo nesprávny záver? Opíšte prípad.**

Chceme poznať konkrétne zlyhania alebo skreslenia, aby sme mohli zlepšiť budúce metódy. Nie každá chyba je zásadná. Zaujíma nás aj malá odchýlka, ktorá mohla mať právny dosah.

Príklad: *Napr.: Analýza označila konanie ako „civilné“, hoci išlo o správne súdnictvo. Kódovanie vychádzalo z formálneho označenia bez zohľadnenia vecnej povahy konania. Výsledky komparatívnej analýzy boli preto skreslené.*

Doména č. 1 – Nekalá súťaž	Pri prvotnom skúmaní výsledkom som sa som skreslením výsledkov nestretla. Určite však pre účely publikovania budem selektovať, ktoré výsledky majú pridanú hodnotu a ktoré možno považovať za istú formu „notoriety“.
Doména č. 2 – Zmluvná pokuta	Pri interpretácii sa objavili situácie, v ktorých automatizovaný alebo príliš formálny prístup mohol viesť k zavádzajúcemu záveru. Napríklad rozhodnutie mohlo byť označené ako prípad aplikácie § 301 iba preto, že sa paragraf v texte spomínal, hoci súd moderačné oprávnenie reálne nepoužil. Naopak, niektoré rozhodnutia hodnotili primeranosť zmluvnej pokuty bez výslovného odkazu na § 301. Takéto prípady ukazujú, že samotný výskyt slov alebo ustanovení nestačí. Pri právnych dátach je potrebná kontrola právnikom a jasné pravidlá kódovania.
Doména č. 3 – Trestné právo	Nie.
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Áno. V niektorých prípadoch automatizácia vytvárala príliš zjednodušený obraz rozhodnutí. Napríklad formálne podobné prípady mohli mať z právneho hľadiska odlišný význam, ktorý nebolo možné zachytiť len prostredníctvom atribútov. Problém vznikal aj pri rozhodnutiach s veľmi stručným odôvodnením, kde bolo náročné správne identifikovať právne relevantné faktory.

Právne výskumné tímy reportujú zavádzajúce závery predovšetkým tam, kde sa zamenila formálna prítomnosť právneho pojmu v texte za jeho skutočnú aplikáciu. Celkový záver je, že výskyt slov alebo právnych ustanovení nestačí ako indikátor ich skutočnej aplikácie. Pre právne dáta je nevyhnutná kontrola právnikom a jasné pravidlá kódovania.

Otázka 19 Ako hodnotíte spoľahlivosť a opakovateľnosť kódovania atribútov v rámci vášho tímu (tzv. inter-rater reliability)? Existoval postup, ako riešiť nezrovnalosti?

Replikovateľnosť je jednou z kľúčových výhod metódy systematickej kontextovej analýzy. Zaujíma nás, ako ste v praxi zabezpečili konzistentné kódovanie a ako ste riešili prípady, kde si viacero kóderov nevedelo zhodnúť.

Príklad: *Napr.: Mali sme coding guide, ale niektoré atribúty (napr. „závažnosť porušenia") boli napriek tomu nekonzistentne kódované. Každý z nás mal iné právne intuitívne chápanie toho, čo je „závažné".*

Doména č. 1 – Nekalá súťaž	Keďže som dataset spracovávala sama, popisované problémy nevznikli.
Doména č. 2 – Zmluvná pokuta	NA
Doména č. 3 – Trestné právo	Nenastal taký prípad.
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Spoľahlivosť kódovania bola vo všeobecnosti dobrá, keďže tím pracoval s vopred dohodnutými pravidlami a coding guide. Napriek tomu sa pri niektorých atribútoch objavovali interpretačné rozdiely, najmä pri kvalitatívnych právnych kategóriách. Nezrovnalosti sme riešili spoločnou diskusiou v tíme a následným zjednotením metodiky kódovania.

Právne výskumné tímy reportujú pomerne odlišné skúsenosti, čo súvisí aj s tým, že niektoré domény pracovali s jedným kóderom, iné s viacerými. Spoločným záverom je, že kódovanie individuálnymi expertmi síce eliminuje nesúlad, ale zároveň obmedzuje kontrolu konzistencie. Pri viacčlenných tímoch je coding guide nevyhnutný, ale nepostačujúci pre atribúty s hodnotovým alebo interpretačným obsahom, kde je potrebná aktívna diskusia a priebežná harmonizácia pravidiel.

Otázka 20 Aké zmeny v procese (výber rozhodnutí, definícia atribútov, analytické metódy, prezentácia výsledkov) by zvýšili použiteľnosť výstupov pre váš právny výskum?

Ide o kľúčovú otázku pre zlepšenie metodológie. Vaše konkrétne odporúčania budú priamo zapracované do záverov výstupu projektu ALTAML. Prosíme o čo najkonkrétnejšie návrhy.

Príklad: *Napr.: Privítal by som iteratívny proces. Najprv ukázať výsledky deskriptívnej analýzy, diskutovať ich s právnymi expertmi a až potom pokračovať s predikčnou analýzou. Právnik by tak mohol korigovať kódovanie pred tým, než sa vygenerujú konečné výsledky.*

Doména č. 1 – Nekalá súťaž	V rámci zlepšenia výsledkov by som zrejme išla do anotovania viacerých rozhodnutí, na väčšej dátovej sade.
Doména č. 2 – Zmluvná pokuta	Použiteľnosť výstupov by zvýšil najmä iteratívny proces. Najprv by sa mala vytvoriť menšia testovacia vzorka rozhodnutí, na ktorej

	právny tím overí atribúty a pravidlá kódovania. Až potom by sa mal kódovať celý dataset. Užitočné by bolo aj prepojenie každého štatistického výstupu s konkrétnymi rozhodnutiami, aby si právnik vedel overiť, z čoho záver vychádza. Pri prezentácii výsledkov by pomohli stručné právne interpretácie grafov, upozornenia na malé vzorky a jasné odlíšenie medzi štatistickým vzťahom a právnym záverom. Dôležité je tiež zachovať možnosť manuálnej korekcie atribútov po prvotnej analýze.
Doména č. 3 – Trestné právo	Zvýšiť počet rozhodnutí na skúmanú analýzu, tak aby išlo štatisticky relevantnú vzorku a závery bolo možné formulovať pre celý systém rozhodnutí. Využitie Bayesovskej pravdepodobnosti napr. na hodnotenie dôkazov a predikcie. Entropická teória v spoločenských vedách použiteľná najmä v/na legislatívnom procese (analýza faktorov, ktoré zvyšujú/znižujú neusporiadanosť – počet výkladov danej normy, v systéme).
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Za najdôležitejšie považujeme: • iteratívny proces medzi právnikmi a dátovými analytikmi, • precíznejšiu definíciu atribútov, • priebežnú validáciu datasetu, • a lepšiu vizualizáciu výsledkov pre právnikov. Právny expert by mal byť intenzívnejšie zapojený už pri návrhu atribútov a interpretácii analytických výstupov.

Právne výskumné tímy konvergujú ku konkrétnym, navzájom sa dopĺňajúcim odporúčaniam. Spoločným menovateľom všetkých odporúčaní je intenzívnejšie zapojenie právneho experta už od fázy návrhu atribútov a interpretácie výsledkov.

Otázka 21 V akých konkrétnych právnych oblastiach alebo typoch výskumu vidíte najväčší potenciál pre využitie tejto metodológie na Slovensku?

Metóda sa javí vhodná predovšetkým pri väčšom množstve homogénnych právnych textov. Zaujíma nás, v ktorých oblastiach práva vidíte najväčšiu pridanú hodnotu z pohľadu vášho výskumného záujmu.

Príklad: *Napr.: Podľa mňa má najväčší potenciál v oblasti ochrany spotrebiteľa (veľký počet typizovaných prípadov), verejného obstarávania (povinne zverejňované rozhodnutia) a trestného práva (dostupné rozhodnutia).*

Doména č. 1 – Nekalá súťaž	Toto je náročná otázka, keďže podľa mňa nejestvuje konkrétna jediná oblasť, alebo oblasti práva, pre ktoré by táto metodológia bola vhodná. Som toho názoru, že metodológia je využiteľná (takmer) na všetky oblasti práva a typy rozhodnutí, avšak kľúčové na precizovať výskumné otázky a vhodne definovať atribúty, ktoré majú byť zbierané.

Doména č. 2 – Zmluvná pokuta	Najväčší potenciál vidíme v oblastiach, kde existuje väčší počet relatívne homogénnych rozhodnutí alebo dokumentov. V obchodnom práve ide najmä o zmluvnú pokutu, nekalú súťaž, obchodné záväzky, zodpovednosť za škodu a spory z obchodných zmlúv. Veľký potenciál má aj spotrebiteľské právo, pracovné právo, verejné obstarávanie, správne trestanie a ochrana osobných údajov. Metodológia je vhodná najmä tam, kde výskumník nechce analyzovať iba niekoľko známych rozhodnutí, ale potrebuje zistiť širšie trendy v aplikačnej praxi.
Doména č. 3 – Trestné právo	Rozhodnutia a judikatúra všeobecne, dokazovanie a hodnotenie dôkazov
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Najväčší potenciál vidíme najmä v oblastiach s veľkým počtom relatívne homogénnych rozhodnutí, napríklad: • ochrana spotrebiteľa, • autorské právo, • správne trestanie, • verejné obstarávanie

Právne výskumné tímy sa výrazne líšia v miere konkrétnosti odporúčaní, zhodujú sa však na tom, že metodológia je pri dostatočnej dostupnosti textov uplatniteľná naprieč právnymi odvetviami. Naprieč výskumnými tímami prevláda záver, že metodológia má najväčší potenciál tam, kde výskumník potrebuje sledovať širšie trendy v aplikačnej praxi, nie analyzovať iba niekoľko kľúčových rozhodnutí.

Otázka 22 Ako by ste uvítali zapojenie veľkých jazykových modelov (LLM, napr. GPT, Gemini) do procesu analýzy právnych textov? Aké vidíte výhody a riziká v porovnaní s klasickými metódami strojového učenia?

Autori projektu sa vedome sústredili na klasické metódy strojového učenia (ML). Výsledky sú vysvetliteľné cez atribúty. Zaujíma nás však, aký je pohľad právnikov na generatívnu AI a veľké jazykové modely v tomto kontexte.

Príklad: *Napr.: LLM by mohol automatizovať kódovanie, ale bojím sa, že nebude konzistentný a jeho odpovede nebudú verifikovateľné. Na druhej strane by mohol zachytiť nuansy, ktoré binárne atribúty jednoducho nezvládnu.*

Doména č. 1 – Nekalá súťaž	Nie som si istá do akej miery je to realizovateľné, ale zrejme pokiaľ by došlo k anotovaniu v systéme Docano, bolo by vhodné, keby podľa určitej vzorky vedela AI sama odanotovať väčšiu vzorku rozhodnutí, podľa predloženého vzoru. Veľké jazykové modely považujem za vhodnejšie na vyhľadávanie informácií, než na takúto analytickú činnosť, keďže ja osobne by som to nevedela odkontrolovať a plne sa na to spoľahnúť.

Doména č. 2 – Zmluvná pokuta	LLM by mohli byť veľmi užitočné najmä pri predbežnom čítaní rozhodnutí, návrhu atribútov, sumarizácii skutkového stavu a identifikácii pasáží, ktoré sa týkajú zmluvnej pokuty alebo § 301 Obchodného zákonníka. Výhodou oproti klasickému strojovému učeniu je schopnosť pracovať s jazykovým kontextom a právnymi nuansami. Rizikom je však nekonzistentnosť, možnosť halucinácií a nižšia verifikovateľnosť odpovedí. Preto by LLM nemali samostatne rozhodovať o právnych kategóriách. Najvhodnejší model je kombinácia: LLM ako asistent pri predspracovaní a extrakcii informácií, klasické ML a štatistika pre vysvetliteľné výstupy a právnik ako konečný kontrolór právnej interpretácie.
Doména č. 3 – Trestné právo	Nevýhodou je jednoznačne fenomenén blackbox fungovania modelu. Ako čiastočné riešenie vidím obmedzenie funkcionality iba na rešerše a analýzy cez atribúty, povinne obsiahnuté v prompte.
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	LLM modely by mohli výrazne pomôcť pri predspracovaní textov, automatickom sumarizovaní rozhodnutí alebo predbežnom kódovaní atribútov. Ich výhodou je schopnosť zachytiť širší jazykový a argumentačný kontext. Rizikom je však nižšia vysvetliteľnosť, nekonzistentnosť odpovedí a problematická verifikovateľnosť záverov. V právnom výskume je preto podľa môjho názoru potrebné kombinovať LLM s klasickými transparentnými analytickými metódami

Právne výskumné tímy oceňujú potenciál LLM pri predspracovaní a sumarizácii, ale zároveň vyjadrujú obavy o konzistentnosť, overiteľnosť a transparentnosť výstupov. Spoločným záverom je, že LLM predstavujú sľubný doplnkový nástroj, nie autonómny rozhodovací systém. Ich zapojenie si vyžaduje jasné vymedzenie rolí a zachovanie právnej kontroly výstupov.

Otázka 23 **Využili by ste metódu systematickej kontextovej analýzy vo svojom ďalšom výskume aj bez podpory dátového analytika? Čo by na to potreboval právny výskumník z hľadiska kompetencií alebo nástrojov?**

Otázka sa týka udržateľnosti metodológie. Zaujíma nás, či by právni výskumníci boli schopní a ochotní využívať tieto metódy samostatne, a čo by im k tomu chýbalo.

Príklad: *Napr.: Áno, ale potreboval by som jednoduchý nástroj na kódovanie (aplikácia Excel nestačí), základný statistický softvér s právnicky zrozumiteľným výstupom a metodický sprievodca krok za krokom (nie kód v programovacom jazyku Pythone).*

Doména č. 1 – Nekalá súťaž	Popravde si to sama úplne neviem predstaviť, keďže to považujem za pomerne náročné. Možno pokiaľ by som realizovala najprv tie kroky pod dohľadom, z čoho by som si urobila vlastné poznámky

	(postup) a následne by som teoreticky možno vedela ako postupovať, ale zrejme úplne sama by som to nezvládla berúc do úvahy aj absenciu znalostí z oblasti štatistiky, teórii grafov a podobne.
Doména č. 2 – Zmluvná pokuta	Metódu by sme využili aj v ďalšom výskume, ale nie úplne bez metodickej a technickej podpory. Právny výskumník by potreboval jednoduchý nástroj na kódovanie rozhodnutí, jasný metodický manuál, šablónu coding guide, základné školenie o interpretácii štatistických výstupov a vizualizácie pripravené v právnicky zrozumiteľnej forme. Excel môže postačovať na začiatok, ale pri väčších datasetoch je limitujúci. Kľúčové je, aby nástroj umožnil návrat ku konkrétnemu textu rozhodnutia a aby výstupy neboli prezentované iba technickým jazykom.
Doména č. 3 – Trestné právo	Určite áno. Jednoduché spracovanie atribútov na veľký počet rozhodnutí bez potreby anotovať každé rozhodnutie samostatne.
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Áno. Metódu by bolo možné využívať aj bez priamej podpory dátového analytika, avšak právny výskumník by potreboval: • jednoduchšie používateľské nástroje, • základné metodologické školenie, • vizualizácie zrozumiteľné pre právnikov, • a metodický manuál vysvetľujúci jednotlivé analytické výstupy

Právne výskumné tímy sa líšia v miere sebadôvery pri samostatnom využití metódy, zhodujú sa však na tom, že určitá forma metodickej a technickej podpory zostáva nevyhnutná. Spoločným záverom je, že právni výskumníci sú ochotní metódu prijať, ale jej autonómne využitie si vyžaduje zásadné zníženie technickej bariéry a prípravu právnicky orientovaných nástrojov a materiálov.

Otázka 24 Existujú ďalšie otázky alebo aspekty, ktoré by ste v súvislosti s využitím automatizovanej analýzy právnych textov chceli zdôrazniť alebo pridať?

Tento otvorený priestor je pre akékoľvek myšlienky, ktoré neboli pokryté predchádzajúcimi otázkami. Od etických otázok, cez inštitucionálne bariéry, až po vízie budúceho vývoja metódy.

> **Príklad:** *Napr.: Chcel by som upozorniť na otázku anonymizácie. Pracujeme s rozhodnutiami, kde sú osobné údaje anonymizované, ale atribúty môžu de-anonymizáciu umožniť (napr. kombinácia „výška pokuty + región + rok"). To je etická otázka, ktorú projekt nerieši.*

Doména č. 1 – Nekalá súťaž	Chcela by som sa poďakovať všetkým členom teamu, ktorí v tomto procese participovali za ich prácu na mojom datasete.
Doména č. 2 – Zmluvná pokuta	Za dôležité považujeme zdôrazniť, že automatizovaná analýza právnych textov má veľký potenciál, ale jej kvalita závisí od kvality

	vstupných dát, atribútov a právnej kontroly. Pri rozhodnutiach o zmluvnej pokute je osobitne dôležité rozlišovať medzi formálnym výskytom právneho ustanovenia a jeho skutočnou aplikáciou. Metóda by sa nemala prezentovať ako náhrada právnika, ale ako nástroj, ktorý umožňuje právnikovi vidieť širšie vzory a klásť presnejšie výskumné otázky. Do budúcnosti bude dôležité riešiť aj etické otázky, ochranu údajov, transparentnosť modelov a udržateľnosť metodiky po skončení projektu.
Doména č. 3 – Trestné právo	NA
Doména č. 4 – Autorské právo – neoprávnené užívanie diel	Projekt ukázal, že systematická kontextová analýza má významný potenciál pre empirický právny výskum aj na Slovensku. Zároveň sa však ukázalo, že kvalita výsledkov závisí predovšetkým od kvality datasetu, správne nastavenej metodiky a úzkej spolupráce medzi právnikmi a dátovými analytikmi. Do budúcna by bolo vhodné viac sa zamerať aj na otázky štandardizácie právnych datasetov, vysvetliteľnosti analytických modelov a etických aspektov práce s rozhodnutiami a právnymi dátami.

Právne výskumné tímy využili záverečný priestor na vyjadrenie širšieho kontextu a systémových odporúčaní. Napríklad vyjadrenie myšlienky, že automatizovaná analýza má veľký potenciál, ale jej kvalita závisí od kvality vstupných dát, atribútov a právnej kontroly. Metóda by sa nemala prezentovať ako náhrada právnika, ale ako nástroj umožňujúci právnikovi vidieť širšie vzory.

Záver

Metóda systematickej kontextovej analýzy práva umožňuje odlišný prístup k právnemu výskumu spočívajúcemu v analýze právneho textu. Špecifikum tejto metódy spočíva predovšetkým v jej založení na identifikácii prítomnosti konkrétnych atribútov prostredníctvom kódovania právnych textov (*data-based approach*) právnym expertom, prípadne na základe využitia automatizovaných metód extrakcie atribútov, za účelom následnej analýzy identifikovaných atribútov s možnosťou využitia rôznych empirických metód, počínajúc jednoduchými štatistickými metódami a končiac náročnejšími metódami strojového učenia, reflektujúc povahu, ako aj kvantitu jednotlivých atribútov. Osobitnou výhodou tejto metódy je predovšetkým jej exaktnosť a možnosť verifikovať dosiahnuté výsledky treťou, nezávislou osobou na základe vytvoreného datasetu atribútov. Medzi identifikované nevýhody tejto metódy možno zaradiť najmä dostupnosť právnych textov, ktoré majú byť podrobené analýze, ako aj ich obsah a kvalitu, ktoré dokážu významne determinovať a limitovať výskumné otázky, resp. vedecké hypotézy, na riešenie ktorých možno použiť predmetnú metódu. Aspektom, s ktorým sa je potrebné pri využití tejto metódy osobitne vysporiadať, je tiež úplnosť vymedzených atribútov, nakoľko nezahrnutie relevantných atribútov do analýzy môže viesť k skresleným výsledkom, ktoré nezohľadňujú právny problém v jeho celosti. Výskumník si tak musí zvážiť, či jednotlivý právny zdroj poskytuje dostatočný základ (obsahuje potrebné informácie) na preskúmanie stanovených výskumných otázok. Využitie tohto metodologického rámca považujeme za osobitne vhodné vo vzťahu k analýze väčšieho množstva právnych textov rôzneho druhu s cieľom priblížiť skutočný stav aplikácie práva v posudzovanom prípade. Skúmaná metodológia môže tiež slúžiť k doplneniu klasického doktrinálneho právneho výskumu, napríklad identifikáciou problémov vyskytujúcich sa v praxi, ktoré vyžadujú hlbšie preskúmanie. V tomto kontexte sme sa nevenovali stále viac populárnejšej generatívnej umelej inteligencii a veľkým jazykovým modelom. Automatizácia v práve vrátane automatizácie právnych textov má svoje pravidlá a požiadavky. Jednou z nevyhnutných požiadaviek je možnosť vysvetliť výsledok, ktorý poskytne daná metóda. V tomto smere majú stále navrch „klasické“ metódy strojového učenia, pri väčšine ktorých vieme vysvetliť výsledok metódy prostredníctvom kódu, resp. atribútov.

Záverečné zistenia hodnotiacej správy vychádzajúce zo spätnej väzby štyroch právnych výskumných tímov pôsobiacich v doménach nekalej súťaže, zmluvnej pokuty, trestného práva a autorského práva tieto teoretické východiská potvrdzujú a zároveň ich obohacujú o praktickú dimenziu. Právne výskumné tímy sa pred zapojením automatizovanej analýzy spoliehali výlučne na tradičný doktrinálny výskum spočívajúci v manuálnom čítaní obmedzeného počtu rozhodnutí, pričom systematická kontextová analýza im po prvý raz umožnila sledovať rozhodovacie trendy vo väčšom rozsahu a s vyššou mierou replikovateľnosti. Výskumné otázky pritom žiadny z tímov neudržal v pôvodnej formulácii. Všetky boli priebežne prispôsobované skutočnému obsahu dostupných rozhodnutí, čo samo osebe potvrdzuje jednu z kľúčových téz metódy. Limity metódy sú limity samotných právnych textov.

Z hľadiska hodnotenia jednotlivých analytických metód sa ako najprístupnejšie a najužitočnejšie ukázali deskriptívna štatistická analýza a detekcia anomálií, ktoré právne výskumné tímy označujú za prirodzený vstupný bod každej analýzy a za nástroj navigujúci pozornosť výskumníka na rozhodnutia hodné hlbšieho manuálneho skúmania. Komparatívna a korelačná analýza boli hodnotené pozitívne ako nástroje generovania nových hypotéz, avšak s dôrazom na nevyhnutnosť filtra právnej relevancie pri interpretácii výstupov. Zhlukovacia analýza a asociačné pravidlá priniesli zaujímavé postrehy, no vyžadujú vyššiu mieru interpretačnej zdatnosti a priebežnej spolupráce s právnym expertom. Prediktívna analýza bola vnímaná skôr ako metodologický impulz než záverečný nástroj, pričom jej prínos spočíval najmä v identifikácii atribútov s najvyššou vysvetľovacou silou.

Všetky právne výskumné tímy sa zhodli na niekoľkých zásadných záveroch pre budúce využitie metódy. Po prvé, výstupy automatizovanej analýzy nemôžu nahradiť právnu argumentáciu. Sú jej cenným, ale výhradne doplnkovým nástrojom, keďže hodnotová, diskrečná a argumentačná vrstva právneho textu si nevyhnutne vyžaduje manuálny pohľad experta. Po druhé, kvalita a úplnosť atribútov sú rozhodujúce pre výpovednú hodnotu výsledkov. Právne výskumné tímy identifikovali potrebu jemnejších atribútov zachytávajúcich nielen formálnu prítomnosť právnych inštitútov, ale aj ich skutočnú aplikáciu. Po tretie, právne výskumné tímy zhodne odporúčajú zavedenie iteratívneho procesu, v rámci ktorého je právny expert intenzívne zapojený už pri návrhu atribútov a priebežne koriguje kódovanie pred generovaním záverečných výsledkov. Po štvrté, veľké jazykové modely sú vnímané ako sľubný doplnkový nástroj predovšetkým pre predspracovanie textov a návrh atribútov, nie však ako autonómny analytický systém. Ich zapojenie si vyžaduje zachovanie právnej kontroly a transparentnosti výstupov. Metodológia bola tímami identifikovaná ako osobitne hodnotná v oblastiach s väčším počtom homogénnych rozhodnutí, medzi ktoré zaraďujú spotrebiteľské právo, verejné obstarávanie, správne trestanie, obchodné záväzky aj trestné právo. Predpokladom jej širšieho a udržateľného využitia v slovenskom právnom výskume je zníženie technickej

bariéry prostredníctvom právnicky orientovaných nástrojov, metodických manuálov a základného vzdelávania právnych výskumníkov v oblasti interpretácie štatistických výstupov.

Príloha č. 1

D O T A Z N Í K

Použiteľnosť výstupov automatizovanej analýzy právnych textov v právnom výskume a praxi

Projekt: Automatizácia analýzy právnych textov (ALTAML) financovaná EÚ – NextGenerationEU / Plán obnovy SR (č. 09I05-03-V02-00049)

Výskum: Systematická kontextová analýza práva (systematic content analysis in law)

Určený pre: Členov právnych tímov projektu, ktorí dostali výstupy dátovej analýzy a pracujú s právnymi datasetmi

Tento dotazník je určený pre členov právnych tímov projektu ALTAML, ktorí pracujú s výstupmi automatizovanej analýzy právnych datasetov (deskriptívna štatistika, korelačná analýza, zhlukovanie, detekcia anomálií, prediktívna analýza, asociačné pravidlá a ďalšie metódy aplikované na datasety z oblastí zmluvnej pokuty, nekalej súťaže, ochrany osobných údajov, digitálnej stopy a ďalších).

Cieľom dotazníka je zistiť, nakoľko sú tieto výstupy použiteľné pre váš právny výskum, čo vám pomáha, čo chýba a aké sú limity tejto metodológie z pohľadu právneho experta. Vaše odpovede budú súčasťou výstupu projektu ALTAML a pomôžu zlepšiť aplikovateľnosť metód v budúcich výskumných projektoch.

Predpokladaný čas vyplnenia: 90 minút. Väčšina otázok je otvorená. Prosíme o konkrétne odpovede.

Meno a priezvisko (voliteľné):	
Právna oblasť, na ktorej pracujete:	
Dataset(y), s ktorými ste pracovali:	
Dátum vyplnenia:	

A. Kontext a východiská právneho výskumu

Čo skúmate, aké sú vaše výskumné otázky a prístupy pred zapojením automatizovaných metód

Otázka 1 Aké výskumné otázky alebo hypotézy ste sledovali v rámci vášho právneho výskumu pred zapojením automatizovanej analýzy?

Uveďte konkrétne výskumné otázky, ktoré ste sa snažili zodpovedať. Je dôležité pochopiť, do akej miery boli tieto otázky formulované dostatočne pred spustením analýzy alebo sa vyvíjali v priebehu projektu.

> **Príklad:** *Napr.: Zaujímalo ma, ako súdy v SR pristupujú k moderácii zmluvnej pokuty – konkrétne: aké faktory berú do úvahy.*

Otázka 2 Akým spôsobom ste typicky vykonávali právny výskum v tejto oblasti pred využitím metód systematickej kontextovej analýzy?

Opíšte váš predchádzajúci metodologický prístup – napr. manuálna analýza kľúčových rozhodnutí, doktrinálna analýza, konzultácia s kolektívom, komparatívna analýza. Pomôže nám to pochopiť, v akom kontexte sú nové metódy hodnotené.

> **Príklad:** *Napr.: Zvyčajne som čítal 10–20 rozhodnutí, ktoré som vybral podľa vlastného uváženia a z nich vyvodzoval závery. Teraz som po prvýkrát pracoval s datasetom 100 rozhodnutí.*

Otázka 3 Narazili ste na obmedzenia plynúce z obsahu alebo dostupnosti právnych textov vo vašom datasete? Ak áno, ako tieto obmedzenia ovplyvnili vaše výskumné otázky?

Metóda systematickej kontextovej analýzy je limitovaná skutočným obsahom analyzovaných textov. Chceme vedieť, či ste museli vaše výskumné otázky upraviť alebo zúžiť z dôvodu absencie určitých informácií v rozhodnutiach, zmluvách alebo iných právnych textoch.

> **Príklad:** *Napr.: Chcel som skúmať, aké dôkazy súdy brali do úvahy pri hodnotení digitálnej stopy, ale väčšina rozhodnutí neobsahovala podrobné odôvodnenie. Išlo o skrátené formy (trestné rozkazy, dohody o vine a treste).*

B. Kódovanie atribútov a tvorba datasetu

Kvalita, úplnosť a relevantnosť atribútov z pohľadu právneho experta

Otázka 4 Nakoľko zachytávali zvolené atribúty právne relevantné informácie pre vašu výskumnú otázku?

Kódovanie právnych textov stojí a padá s výberom atribútov. Kódovanie ste vykonávali po predchádzajúcej diskusii. Uveďte, ktoré atribúty si myslíte, že by ste inak zadefinovali, resp. napokon Vám chýbali.

> **Príklad:** *Napr.: Atribút „výška zmluvnej pokuty" bol veľmi dôležitý, ale chýbal mi atribút rozlišujúci, či bola pokuta moderovaná z úradnej povinnosti alebo na základe návrhu strany – toto rozlíšenie má právny dosah, ale nebolo zachytené.*

Otázka 5 Stretli ste sa s prípadmi, kde bol atribút definovaný spôsobom, ktorý nezodpovedal právnej realite alebo ustálenej právnej terminológii?

Pri automatizovanom spracovaní právneho textu môže dôjsť k zjednodušeniu alebo skomoleniu právnych pojmov. Zaujíma nás, či kategórie atribútov zodpovedali skutočnému právnemu obsahu, alebo či boli niektoré zaradenia nesprávne, nejasné alebo zavádzajúce.

> **Príklad:** *Napr.: Atribút „typ konania" mal kategóriu „civilné", ale niektoré konania boli svojou povahou skôr správne. To záviselo to od interpretácie. Alebo: kategória „vyhovenie žalobe" bola binárna, hoci reálne súdy rozhodujú v spektre (čiastočné vyhovenie, zmena formulácie a pod.).*

☐ ÁNO	☐ NIE	*Komentár:* _____

Otázka 6 **Ak by ste mohli navrhnúť zmeny v zozname atribútov pre váš dataset, aké by to boli? Čo by ste pridali, odobrali alebo inak formulovali?**

Táto otázka je kľúčová pre budúci výskum. Vaše odborné poznanie oblasti práva je nenahraditeľné pri identifikácii toho, čo skutočne rozlišuje právne relevantné prípady a čo je z právneho hľadiska bezvýznamné.

> **Príklad:** *Napr.: Pridal by som atribút „aplikácia generálnej klauzuly nekalej súťaže verzus konkrétnej skutkovej podstaty", pretože tieto dve cesty majú diametrálne odlišné procesné dôsledky, ale v datasete sa zlievali do jednej kategórie.*

C. Hodnotenie jednotlivých analytických výstupov

Čo vám pomohlo a čo nie – konkrétne pre každý typ analýzy

Pre každú analytickú metódu, s ktorou ste pracovali, uveďte:

(1) čo konkrétne vám výstup z analýzy priniesol a aké boli jeho limity (rozpíšte sa do textového poľa),

(2) nakoľko sú výsledky analýzy zrozumiteľné pre právneho vedca bez hlbšej znalosti štatistiky alebo strojového učenia (výber z možností 1-5),

(3) Nakoľko vám je/bol výstup užitočný (výber z možností 1-5)

Otázka 7 **Deskriptívna štatistická analýza (frekvencie, podiely, priemery, distribúcie atribútov)**

Tento výstup ukazuje, aké hodnoty atribútov sa vyskytujú, ako často a v akom pomere. Napríklad: koľko percent rozhodnutí skončilo vyhovením, aká bola priemerná výška pokuty, ako sa rozkladajú typy konaní.

> **Príklad:** *Napr.: Zistil som, že 73 % rozhodnutí o zmluvnej pokute skončilo aspoň čiastočným vyhovením žalobe – to bol zaujímavý základ pre ďalšie otázky. Ale samotné číslo mi nepovedalo, prečo to tak je.*

Zrozumiteľnosť výsledkov:

1	2	3	4	5
Vôbec nie	Málo	Čiastočne	Prevažne	Úplne áno

Užitočnosť výsledkov:

1	2	3	4	5
Vôbec nie	Málo	Čiastočne	Prevažne	Úplne áno

Otázka 8 Komparatívna analýza (rozdiely medzi skupinami – súdy, kraje, obdobia, typy konaní)

Táto analýza porovnáva skupiny dokumentov podľa metadát a hľadá systematické rozdiely. Napríklad: líšia sa rozhodnutia krajských súdov od okresných súdov? Existujú regionálne rozdiely v rozhodovaní?

> **Príklad:** *Napr.: Výstup ukázal, že Krajský súd v Bratislave mal výrazne vyšší podiel zamietnutých žalôb v oblasti nekalej súťaže ako ostatné krajské súdy – toto by stálo za ďalší výskum, ale neviem, či je to dôvod metodiky alebo iný faktor.*

Zrozumiteľnosť výsledkov:

1	2	3	4	5
Vôbec nie	Málo	Čiastočne	Prevažne	Úplne áno

Užitočnosť výsledkov:

1	2	3	4	5
Vôbec nie	Málo	Čiastočne	Prevažne	Úplne áno

Otázka 9 Korelačná analýza (vzájomné závislosti medzi atribútmi)

Korelačná analýza meria, nakoľko spolu súvisia rôzne atribúty. Napríklad: existuje vzťah medzi výškou zmluvnej pokuty a výsledkom konania? Je spojitosť medzi typom porušenia a druhom sankcie?

> **Príklad:** *Napr.: Analýza ukázala silnú pozitívnu koreláciu medzi atribútom „aplikácia generálnej klauzuly" a „zamietnutím žaloby". To bolo pre mňa právne prekvapujúce a viedlo ma k ďalšej hypotéze.*

Zrozumiteľnosť výsledkov:

1	2	3	4	5
Vôbec nie	Málo	Čiastočne	Prevažne	Úplne áno

Užitočnosť výsledkov:

1	2	3	4	5
Vôbec nie	Málo	Čiastočne	Prevažne	Úplne áno

Otázka 10 Zhluková/klastrovacia analýza (objavovanie prirodzených skupín právnych textov/rozhodnutí)

Zhlukovanie bez vopred definovaných kategórií hľadá podobné právne texty/rozhodnutia a zoskupuje ich. Cieľom je objaviť opakujúce sa vzory a typické prípady bez toho, aby výskumník vopred vedel, aké skupiny existujú.

> **Príklad:** *Napr.: Zhlukovanie odhalilo tri opakujúce sa „typy" rozhodnutí o nekalej súťaži, ktoré som predtým vnímal skôr intuitívne. Teraz vidím ich konkrétne atribútové charakteristiky.*

Zrozumiteľnosť výsledkov:

1	2	3	4	5

Vôbec nie	Málo	Čiastočne	Prevažne	Úplne áno

Užitočnosť výsledkov:

1 Vôbec nie	2 Málo	3 Čiastočne	4 Prevažne	5 Úplne áno

Otázka 11 Detekcia anomálií (identifikácia odchýlok od typickej praxe)

Táto metóda identifikuje právne texty/rozhodnutia, ktoré sa výrazne líšia od ostatných. Môže ísť o netypické rozhodnutia, extrémne hodnoty sankcií, neštandardné zmluvné dojednania alebo rozhodnutia odchyľujúce sa od ustálenej praxe.

> **Príklad:** *Napr.: Analýza označila jedno rozhodnutie o zmluvnej pokute ako anomálne. Po prečítaní som zistil, že súd použil neobvyklú argumentáciu priamo odkazujúcu na doktrínu iného štátu. To by som inak prehliadol.*

Zrozumiteľnosť výsledkov:

1 Vôbec nie	2 Málo	3 Čiastočne	4 Prevažne	5 Úplne áno

Užitočnosť výsledkov:

1 Vôbec nie	2 Málo	3 Čiastočne	4 Prevažne	5 Úplne áno

Otázka 12 Asociačné pravidlá (vzory tvaru: AK x → POTOM y, napr. ak dlhý trest → nepodmienečný)

Asociačné pravidlá hľadajú pravidelnosti v tom, čo sa v textoch spoluvyskytuje. Výstupom sú pravidlá s mierou podpory (support), spoľahlivosti (confidence) a relevancie (lift).

> **Príklad:** *Napr.: Pravidlo „ak IP adresa + dlhé trvanie skutku → nepodmienečný trest odňatia slobody" bolo štatisticky silné, ale právne som ho musel interpretovať opatrne. Korelácia nie je kauzalita.*

Zrozumiteľnosť výsledkov:

1	2	3	4	5
Vôbec nie	Málo	Čiastočne	Prevažne	Úplne áno

Užitočnosť výsledkov:

1	2	3	4	5
Vôbec nie	Málo	Čiastočne	Prevažne	Úplne áno

Otázka 13 Prediktívna analýza (predikovanie výsledku konania alebo iného cieľového atribútu na základe ostatných atribútov)

Model strojového učenia sa snaží predpovedať, ako dopadne konanie alebo aká bude sankcia. Zároveň ukazuje, ktoré atribúty majú na výsledok najväčší vplyv (feature importance).

> **Príklad:** *Napr.: Model predikoval výsledok konania s presnosťou F1=0.72. Pre mňa ako právnika bola zaujímavá nie samotná predikcia, ale zoznam najdôležitejších atribútov: „výška škody“ a „opakované porušenie“ boli kľúčové, čo zodpovedá právnemu cíteniu.*

Zrozumiteľnosť výsledkov:

1	2	3	4	5
Vôbec nie	Málo	Čiastočne	Prevažne	Úplne áno

Užitočnosť výsledkov:

1	2	3	4	5

Vôbec nie	Málo	Čiastočne	Prevažne	Úplne áno

D. Interpretácia výsledkov a prepojenie s právnou argumentáciou

Ako pracujete s výstupmi vo vzťahu k právnemu výskumu

Otázka 14 Dokázali by ste výstupy dátovej analýzy priamo použiť vo vašom právnom výskume? Ak áno, akým spôsobom?

Zaujíma nás, či je možné výstupy analýzy skutočne integrovať do právneho textu, argumentácie alebo záverov, alebo zostávajú oddelené. Opíšte konkrétny prípad, ak je to možné.

> **Príklad:** *Napr.: Deskriptívnu štatistiku by som vedel použiť v článku ako empirický základ pre tvrdenie o dominantnom type rozhodnutí. Zhlukovanie mi pomohlo štruktúrovať kapitoly podľa identifikovaných typov prípadov.*

Otázka 15 Existujú výstupy analýzy, ktoré by ste nedokázali alebo nevedeli zmysluplne interpretovať v právnom kontexte? Opíšte, prečo.

Nie každý štatistický výsledok má automaticky právny zmysel. Zaujíma nás, kde je hranica medzi štatisticky zaujímavým a právne relevantným, a kde narážate na interpretačné ťažkosti.

> **Príklad:** *Napr.: Korelačná matica ako celok bola pre mňa ťažko čitateľná. Nevedel som, ktorú koreláciu mám brať vážne a ktorú ignorovať. Chýbal mi filter „právnej relevancie“.*

Otázka 16 Môžu výstupy analýzy nahradiť právnu argumentáciu, alebo sú výlučne jej doplnkom? Vysvetlite vaše stanovisko.

Ide o zásadnú metodologickú otázku. Chceme vedieť, ako právni výskumníci vnímajú rolu kvantitatívnych výstupov vo vzťahu k tradičnej normatívnej analýze práva.

> **Príklad:** *Napr.: Podľa môjho názoru výstupy nikdy nenahradia právnu argumentáciu. Môžu identifikovať problém, ale nie ho vysvetliť. Súd nepochopí „F1 skóre 0,7“. Musím to preložiť do právneho jazyka.*

E. Limitácia metódy a manuálne kódovanie

Kde má automatizácia hranice a čo vyžaduje nevyhnutne ľudský pohľad

Otázka 17 Aké aspekty právneho textu sú podľa vás len ťažko alebo vôbec automatizovateľné a nevyhnutne vyžadujú manuálny pohľad experta?

Niektoré právne informácie sú implicitné, kontextuálne alebo vyžadujú interpretáciu. Nedajú sa zachytiť binárnym alebo kategorickým atribútom. Zaujíma nás, kde je podľa vás hranica automatizácie v právnom texte.

> **Príklad:** *Napr.: Kvalitu právnej argumentácie súdu, aplikáciu diskrečného práva alebo hodnotenie proporcionality sankcie nedokáže zachytiť žiadny atribút. Vždy si to vyžaduje prečítanie celého rozhodnutia.*

Otázka 18 Narazili ste pri kódovaní alebo interpretácii výsledkov na príklady, kde automatizácia produkovala zavádzajúci alebo nesprávny záver? Opíšte prípad.

Chceme poznať konkrétne zlyhania alebo skreslenia, aby sme mohli zlepšiť budúce metódy. Nie každá chyba je zásadná. Zaujíma nás aj malá odchýlka, ktorá mohla mať právny dosah.

> **Príklad:** *Napr.: Analýza označila konanie ako „civilné“, hoci išlo o správne súdnictvo. Kódovanie vychádzalo z formálneho označenia bez zohľadnenia vecnej povahy konania. Výsledky komparatívnej analýzy boli preto skreslené.*

☐ ÁNO	☐ NIE	*Komentár: _____*

Otázka 19 **Ako hodnotíte spoľahlivosť a opakovateľnosť kódovania atribútov v rámci vášho tímu (tzv. inter-rater reliability)? Existoval postup, ako riešiť nezrovnalosti?**

Replikovateľnosť je jednou z kľúčových výhod metódy systematickej kontextovej analýzy. Zaujíma nás, ako ste v praxi zabezpečili konzistentné kódovanie a ako ste riešili prípady, kde si viacero kóderov nevedelo zhodnúť.

> **Príklad:** *Napr.: Mali sme coding guide, ale niektoré atribúty (napr. „závažnosť porušenia") boli napriek tomu nekonzistentne kódované. Každý z nás mal iné právne intuitívne chápanie toho, čo je „závažné".*

F. Perspektívy a odporúčania

Ako ďalej – čo by metódu urobilo hodnotnejšou

Otázka 20 **Aké zmeny v procese (výber rozhodnutí, definícia atribútov, analytické metódy, prezentácia výsledkov) by zvýšili použiteľnosť výstupov pre váš právny výskum?**

Ide o kľúčovú otázku pre zlepšenie metodológie. Vaše konkrétne odporúčania budú priamo zapracované do záverov výstupu projektu ALTAML. Prosíme o čo najkonkrétnejšie návrhy.

> **Príklad:** *Napr.: Privítal by som iteratívny proces. Najprv ukázať výsledky deskriptívnej analýzy, diskutovať ich s právnymi expertmi a až potom pokračovať s predikčnou analýzou. Právnik by tak mohol korigovať kódovanie pred tým, než sa vygenerujú konečné výsledky.*

Otázka 21 **V akých konkrétnych právnych oblastiach alebo typoch výskumu vidíte najväčší potenciál pre využitie tejto metodológie na Slovensku?**

Metóda sa javí vhodná predovšetkým pri väčšom množstve homogénnych právnych textov. Zaujíma nás, v ktorých oblastiach práva vidíte najväčšiu pridanú hodnotu z pohľadu vášho výskumného záujmu.

> **Príklad:** *Napr.: Podľa mňa má najväčší potenciál v oblasti ochrany spotrebiteľa (veľký počet typizovaných prípadov), verejného obstarávania (povinne zverejňované rozhodnutia) a trestného práva (dostupné rozhodnutia).*

Otázka 22 **Ako by ste uvítali zapojenie veľkých jazykových modelov (LLM, napr. GPT, Gemini) do procesu analýzy právnych textov? Aké vidíte výhody a riziká v porovnaní s klasickými metódami strojového učenia?**

Autori projektu sa vedome sústredili na klasické metódy strojového učenia (ML). Výsledky sú vysvetliteľné cez atribúty. Zaujíma nás však, aký je pohľad právnikov na generatívnu AI a veľké jazykové modely v tomto kontexte.

> **Príklad:** *Napr.: LLM by mohol automatizovať kódovanie, ale bojím sa, že nebude konzistentný a jeho odpovede nebudú verifikovateľné. Na druhej strane by mohol zachytiť nuansy, ktoré binárne atribúty jednoducho nezvládnu.*

Otázka 23 **Využili by ste metódu systematickej kontextovej analýzy vo svojom ďalšom výskume aj bez podpory dátového analytika? Čo by na to potreboval právny výskumník z hľadiska kompetencií alebo nástrojov?**

Otázka sa týka udržateľnosti metodológie. Zaujíma nás, či by právni výskumníci boli schopní a ochotní využívať tieto metódy samostatne, a čo by im k tomu chýbalo.

> **Príklad:** *Napr.: Áno, ale potreboval by som jednoduchý nástroj na kódovanie (aplikácia Excel nestačí), základný statistický softvér s právnicky zrozumiteľným výstupom a metodický sprievodca krok za krokom (nie kód v programovacom jazyku Pythone).*

☐ ÁNO	☐ NIE	*Komentár:* _____

G. Záverečné poznámky

Priestor pre vaše doplňujúce myšlienky

Otázka 24 Existujú ďalšie otázky alebo aspekty, ktoré by ste v súvislosti s využitím automatizovanej analýzy právnych textov chceli zdôrazniť alebo pridať?

Tento otvorený priestor je pre akékoľvek myšlienky, ktoré neboli pokryté predchádzajúcimi otázkami. Od etických otázok, cez inštitucionálne bariéry, až po vízie budúceho vývoja metódy.

> **Príklad:** *Napr.: Chcel by som upozorniť na otázku anonymizácie. Pracujeme s rozhodnutiami, kde sú osobné údaje anonymizované, ale atribúty môžu de-anonymizáciu umožniť (napr. kombinácia „výška pokuty + región + rok"). To je etická otázka, ktorú projekt nerieši.*

Ďakujeme za vaše odpovede. Vaše skúsenosti a odporúčania budú priamo zapracované do jedného z výstupov projektu ALTAML.

Kontakt: doc. RNDr. JUDr. Pavol Sokol, PhD. et PhD. | JUDr. Laura Bachňáková Rózenfeldová, PhD.
Projekt ALTAML