

<https://doi.org/10.5817/RPT2020-2-8>

SROVNÁVACÍ STUDIE PRÁVNÍCH INFORMAČNÍCH SYSTÉMŮ: ROZDÍLY MEZI SYSTÉMY PŘI VYUŽITÍ RŮZNÝCH VYHLEDÁVACÍCH STRATEGIÍ¹

JAKUB HARAŠTA²

ABSTRAKT

Vyhledávání právních informací v právních informačních systémech je běžnou součástí života právníka. V zásadě je možné ve dvou modech: jako vyhledávání zprostředkované a nezprostředkované. V rámci zprostředkovaného vyhledávání se spoléháme na souvislosti připravené tvůrcem systému. Jedná se například o přiřazení dokumentu do oblasti úpravy nebo, na specifitější úrovni, o označení soudních rozhodnutí jako souvisejících s určitým předpisem nebo ustanovením. Při nezprostředkovaném vyhledávání se uživatel naopak spoléhá sám na sebe. Sám formuluje plnotextové dotazy a vyhledává výrazy, které pro něj symbolizují rešeršovaný problém.

V tomto textu se zaměřuji na porovnání výsledků poskytnutých právními informačními systémy ASPI, Beck-online a Codexis v odpovědi na srovnatelné dotazy využívající zprostředkovaných i nezprostředkovaných právních informací.

Po úvodu a přehledu literatury následuje představení výzkumných otázek. Čtvrtá část pak popisuje metodu použitou pro srovnání systémů a data k tomuto srovnání využítá. Pátá část popisuje výsledky srovnání napříč právními infor-

¹ Část tohoto textu v rozsahu podobnosti výsledků jednotlivých právních informačních systémů v odpovědi na metadatové vyhledávání (zejména metoda v části IV.A a výsledky v části V.A) byla publikována jako konferenční příspěvek z Internationalen Rechtsinformatik Symposium IRIS 2019 v Salzburku.

² JUDr. Mgr. Jakub Harašta, Ph.D., Odborný asistent, Ústav práva a technologií, Právnická fakulta MU. Kontakt: jakub.harasta@law.muni.cz.

mačními systémy ASPI, Beck-online a Codexis. Šestá část se věnuje diskuzi těchto výsledků a jejich zasazení do kontextu. V poslední části pak předkládám závěr a formuluji doporučení ohledně volby vyhledávacích strategií při sběru právních informací (dokumentů) pro akademickou i praktickou činnost.

Závěrem tohoto textu je, že při využití různých systémů pro zprostředkované vyhledávání (identifikaci soudních rozhodnutí souvisejících s konkrétním ustanovením) pozorujeme výrazné rozdíly. Naopak při zařazení metod nezprostředkovaného vyhledávání (plnotextový dotaz) jsou si výsledky poskytnuté jednotlivými systémy podobnější, byť jsou celkově horší oproti vyhledávání zprostředkovanému.

KLÍČOVÁ SLOVA

vyhledávání právních informací, vyhledávací strategie, srovnání

ABSTRACT

Use of legal information retrieval systems is a common part of every lawyer's life. In principle, it is possible to use two approaches to searching: mediated and non-mediated. As a part of the mediated search, we rely on the context provided by the provider of the information system. This might entail assignment of document with particular topic or labelling court decision as related to particular provision. As a part of the non-mediated search, user relies on herself. She formulates full-text queries to search expressions representing the issue at hand.

In this text, I focus on comparison of results provided by legal information systems ASPI, Beck-online and Codexis in response to comparable search queries employing both mediated and non-mediated search.

The introduction and the literature review is followed by part dedicated to research questions. The fourth part described the method used to compare the systems and the data used for the comparison. In the fifth part, I describe the results and the sixth part is devoted to discussion of these results in wider context. The last part presents a conclusion and contains recommendations regarding the selection of search strategies.

In conclusion, use of different systems for mediated search leads to significant differences between systems. On the contrary, when including non-mediated

search (full-text queries), results provided by individual systems are more similar to each other, but generally less precise compared to mediated search.

KEYWORDS

legal information retrieval, search strategies, comparison

1. ÚVOD

Právníci byli brzkými uživateli systémů pro získávání informací (IR, *information retrieval*). V důsledku toho se problematika vyhledávání právních informací v rámci elektronických systémů stala předmětem diskuze na stránkách (zejména anglickojazyčných) odborných i profesních periodik,³ ale také předmětem výuky mladých právníků v průběhu studia i na počátku profesní kariéry. Otázky rozvoje strategií pro vyhledávání dokumentů, strategií pro zajištění a udržení uživatelských kompetencí⁴ nebo evaluace úspěšnosti vyhledávacích úkonů prováděných uživateli⁵ doprovázely rozvoj těchto systémů i jejich rozšiřování v praxi. Proces efektivní identifikace relevantních dokumentů a jejich získávání není, pochopitelně, pouze předmětem akademického zájmu. Má i nezpochybnitelný praktický rozměr. Svědomité vyhledání zdrojů je základem pro kompetentní odpověď klientovi nebo pro náležitou obranu jeho zájmů. Praktici jsou odpovědní za nesprávnou právní radu a její nesprávnost může být způsobena nedokonalostmi v prováděném sběru podkladů (rešerše). Znalost limitů konkrétních právních informačních systémů a obsahu, který se v nich nachází (nebo naopak nenachází), je důležitá. Stejně důležitá je i schopnost právníka pochopit rozdíly mezi systémy a praktické dopady, které tyto rozdíly mají na

³ Srov. GERSON, Kevin. Evaluating Legal Information Retrieval Systems: How Do the Ranked-Retrieval Methods of WESTLAW and LEXIS Measure Up? *Legal Reference Services Quarterly*, 1999, roč. 17, č. 4, s. 53-67. Viz také KNAPP, Melanie a Rob WILLEY. Comparison of Research Speed and Accuracy Using Westlaw Classic and WestlawNext. *Legal Reference Services Quarterly*, 2013, roč. 32, č. 1-2, s. 126-141. Viz také MART, Susan Nevelow. The Case for Curation: The Relevance of Digest and Citorator Results in Westlaw and Lexis. *Legal Reference Services Quarterly*, 2013, roč. 32, č. 1-2, s. 13-53.

⁴ Viz BING, Jon. Legal Information Retrieval Systems: The Need for and the Design of Extremely Simple Retrieval Strategies. *Computer/Law Journal*, 1978, roč. 1, č. 1, s. 379-400.

⁵ Viz BLAIR, David a M. E. MARON. An Evaluation of Retrieval Effectiveness for a Full-Text Document-Retrieval System. *Communications of the ACM*, 1985, roč. 28, č. 3, s. 289-299.

vyhledávání. Do třetice je pak nutné být si vědom různých vyhledávacích strategií a rozdílů mezi nimi, ať už s cílem dosáhnout větší kvality výsledků nebo nižší časové náročnosti vyhledávání.

Naneštěstí není literatura, která by se kriticky věnovala právním informačním systémům v České republice, respektive některým aspektům uživatelské práce s nimi, v podstatě vůbec k dispozici. Omezujeme se na výuku práce s nimi a rozvoj praktických dovedností vyhledávání v informačních systémech. Porovnání jednotlivých systémů nebo třeba diskuze limitů některých vyhledávacích postupů však v literatuře absentuje.

V České republice mají tradičně dominantní postavení tři právní informační systémy – ASPI (Wolters Kluwer), Beck-online (Nakladatelství C.H. Beck) a Codexis (Atlas Consulting). Součástí kánonu české praxe, který stále přetrvává, je přistupovat k těmto systémům jako k plně zaměnitelným, protože obsahují totožné právní předpisy a z velké části totožné kolekce judikatury. Skutečné rozdíly můžeme najít, jak se často uvádí (a jak i já sám při různých příležitostech studentům opakuji) až na úrovni dalšího obsahu – typicky komentářové literatury zařazené do jednotlivých systémů. Dle mých zjištění nebyla zatím žádná pozornost věnována kritické analýze výsledků poskytovaných těmito systémy v odpovědi na srovnatelné vyhledávací dotazy. Systémy nebyly v České republice podrobeny podobnému srovnání, jako je například možné vidět v některých jiných zemích.⁶ Podobným způsobem absentuje i diskuze na téma efektivity vyhledávacích strategií pro práci s právními informačními systémy. Pokud se například zamyslíme nad tím, jak vzniká kresba, jedná se o sled využití jednotlivých technik, které jsou různě časově náročné. Pokud poskytneme k vytvoření

⁶ Srov. GERSON, Kevin. Evaluating Legal Information Retrieval Systems: How Do the Ranked-Retrieval Methods of WESTLAW and LEXIS Measure Up? *Legal Reference Services Quarterly*, 1999, roč. 17, č. 4, s. 53-67. Také viz TAYLOR, William L. Comparing KeyCite and Shepard's for Completeness, Currency, and Accuracy. *Law Library Journal*, 2000, roč. 92, č. 2, s. 127-142. Dále KNAPP, Melanie a Rob WILLEY. Comparison of Research Speed and Accuracy Using Westlaw Classic and WestlawNext. *Legal Reference Services Quarterly*, 2013, roč. 32, č. 1-2, s. 126-141. Dále MART, Susan Nevelow. The Case for Curation: The Relevance of Digest and Citor Results in Westlaw and Lexis. *Legal Reference Services Quarterly*, 2013, roč. 32, č. 1-2, s. 13-53. Také viz HELLYER, Paul. Evaluating Shepard's, KeyCite, and BCite for Case Validation Accuracy. *Law Library Journal*, 2018, roč. 110, č. 4, s. 449-476.

kresby časovou dotaci deseti hodin, výsledek bude diametrálně odlišný od situace, kdy poskytneme pouze deset minut. Zkušený profesionál si podle dotace rozvrhne práci a při zpracování s nižší časovou dotací nebude takovou pozornost věnovat jednotlivým detailům. Rozdíl ve výsledku pak bude na první pohled patrný. Bude ale rozdíl na první pohled patrný, pokud se nebude jednat o kresbu, ale o vyhledávání právních informací? Bude si právník schopný poradit s odlišnou časovou dotací a zvolit pro vyhledání informací různé postupy tak, aby jeho vyhledávání nebylo perfektní, ale bylo kompetentní? Z mých (spíše anekdotických) zjištění formou neformálních dotazů plyne, že tomuto věnují právníci jen velmi málo pozornosti. Často mají pocit, že vyhledávání právních informací, jako jakési „pomocné vědy právnické“, je pod jejich úroveň a má na jejich práci je relativně malý vliv.

Série metodologických článků uveřejněných v časopise *Jurisprudence* mne přivedla k mnoha otázkám směřujícím právě tímto směrem. Pokud například Kosář s Petrovem uvádí, že „[p]odstatou případové studie je podrobná analýza jednoho anebo několika málo případů určitého fenoménu či kategorie“⁷, jak se vůbec dozvím o existenci konkrétních případů vhodných k analýze, pokud zpracovávám rozhodovací praxi? Je nějaký rozdíl mezi vyhledáváním vhodných případů formou plnotextových dotazů nebo formou metadatových vazeb mezi ustanovením předpisu a rozhodnutími? Pokud Bobek v poznámce pod čarou poznamenává, že „[d]obrý komentář totiž zasadí dané ustanovení do širšího kontextu, osvětlí smysl a původ, kriticky analyzuje problémy a navrhne řešení, vtáhne do sebe a kriticky zhodnotí existující judikaturu“⁸, jakým způsobem autor komentáře zajistí, že identifikoval existující judikaturu dostatečným způsobem? A konečně, pokud Urbániková a Smékal uvádí, že při nepopsání použité metody „jde o tzv. black-box výzkum, kdy závěry padají shůry, nevíme, jak k nim autor došel, a tím pádem

⁷ Viz KOSAŘ, David a Jan PETROV. Jak vybrat „případy“ do případové studie a pracovat s nimi v právu: poznatky z výzkumu na pomezí práva a politologie. *Jurisprudence*, 2016, roč. 25, č. 6, s. 24.

⁸ Viz BOBEK, Michal. Výzkum v právu: reklama na Nike anebo kvantová fyzika? *Jurisprudence*, 2016, roč. 25, č. 6, s. 5 (poznámka pod čarou 8).

*jsme odkázáni jen na slepou důvěru*⁹, nedlužíme čtenářům akademických článků vysvětlení, proč jsme do právní analýzy konkrétního ustanovení zařadili tři konkrétní rozhodnutí? Nemůže být fakt, že nám tři konkrétní rozhodnutí nabídne jako související s konkrétním ustanovením ASPI, ale Beck-online nebo Codexis nabídne jako relevantní jiná rozhodnutí, pro výzkumníka, čtenáře nebo praktikujícího právníka důležitý? Poskytuje srovnatelný dotaz do různých systémů stejné výsledky? Pokud ne, jak moc se liší a proč? Pokud se celou tuto záležitost pokusím převést do vyložené praktických dimenzí: pokud jsem v pozici advokáta připravujícího dovolání k Nejvyššímu soudu a nezohledním v argumentaci relevantní rozhodnutí, mohu poškodit klienta. Ale jak – jakými metodami, jakými postupy a jakým způsobem – zajistím, že při rešerši budu schopen identifikovat všechna relevantní rozhodnutí?

V tomto textu představuji experiment, kterým (i) porovnávám výsledky poskytnuté právními informačními systémy ASPI, Beck-online a Codexis, v odpovědi na srovnatelné dotazy a (ii) porovnávám výsledky poskytnuté právními informačními systémy ASPI, Beck-online a Codexis při změně vyhledávací strategie. Po úvodu, kterým mimo jiné vysvětluji i motivaci pro přípravu tohoto textu, následuje přehled literatury věnované některým otázkám získávání relevantních informací a evaluace úspěšnosti těchto úkonů. Ve třetí části popisují idealizovaný model, pomocí kterého s kolegy vyučujeme práci s právními informačními systémy mimo jiné i na Právnické fakultě MU a na Justiční akademii, a který poslouží jako základní struktura pro představení tohoto experimentu. Model totiž stojí na některých předpokladech, které v tomto textu ve velmi zjednodušené formě ověřuji a přímo z něj tak vycházejí konkrétní výzkumné otázky, které se tímto textem snažím zodpovědět. Čtvrtá část popisuje metodu a využitá data, pátá prezentuje výsledky. V šesté části výsledky diskutuji a upozorňuji na některé (nezanedbatelné) limity tohoto textu. Poslední část je závěrečná – shrnuje tento text a také obsahuje základní úvahy o možném dalším postupu v této oblasti.

⁹ Viz URBÁNIKOVÁ, Marína a Hubert SMEKAL. Právní věda a právní psaní: postačí vždy jako výzkumná metoda "číst, přemýšlet a psát"? *Jurisprudence*, 2017, roč. 26, č. 4, s. 40.

2. LITERATURA

Pokud do jakéhokoli informačního systému zadáme konkrétní dotaz a v odpovědi na něj získáme seznam výsledků, ideálně bychom chtěli dokumenty v databázi rozdělit do 2 kategorií. První kategorií jsou tzv. pravá pozitivita (TP, *true positives*). Jedná se o dokumenty, které jsou pro nás relevantní a jsou obsaženy v seznamu výsledků, který jsme dostali v odpovědi na dotaz. Druhou kategorií jsou tzv. pravá negativa (TN, *true negatives*), což jsou dokumenty, které pro nás nejsou relevantní a v seznamu výsledků, který jsme dostali v odpovědi na dotaz, nejsou obsaženy. Takto jednoduchá však situace není, protože v reálném světě se nám objevují další kategorie dokumentů. Třetí kategorií tak jsou tzv. falešná pozitivita (FP, *false positives*). Jedná se o dokumenty, které pro nás nejsou relevantní, ale z nějakého důvodu došlo k jejich zařazení na seznam výsledků. Poslední kategorií jsou tzv. falešná negativa (FN, *false negatives*), které jsme sice chtěli získat (jsou pro naše vyhledávání relevantní), ale nedošlo k jejich zařazení na seznam výsledků. Nedostatky (tedy zařazení falešně pozitivních dokumentů nebo nezařazení falešně negativních) může být způsobené nedostatky na straně informačního systému nebo databáze (nízká kvalita dat a metadat, nízká uživatelská přívětivost při formulování dotazu) nebo na straně uživatele (nedostatečná kompetence při formulaci dotazů).

Obecně se úspěšnost získávání informací v právních informačních systémech měří pomocí přesnosti (P, *precision*), úplnosti (R, *recall*) a míry F1 (F1, *F1 measure*).¹⁰ Přesnost je poměr počtu relevantních dokumentů získaných v odpovědi na konkrétní dotaz vůči celkovému počtu dokumentů získaných v odpovědi. Úplnost je poměr mezi počtem relevantních dokumentů získaných v odpovědi na dotaz a všech relevantních dokumentů v databázi. Míra F1 je pak harmonickým průměrem přesnosti a úplnosti.¹¹ Tyto veličiny při evaluaci nabývají hodnoty mezi 0 a 1, kde hodnota 0 indikuje nejméně přesný (resp. úplný nebo správný) výsledek vyhledávání a hodnota 1 indikuje nejvíce přesný (resp. úplný nebo správný) výsledek

¹⁰ Podrobněji ASHLEY, Kevin. *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*. Cambridge: Cambridge University Press, 2017. S. 222.

¹¹ Tamtéž.

vyhledávání. Dosáhnout v rámci jednoho vyhledávání perfektní přesnosti i perfektní úplnosti (a tím perfektní míry F1) je v podstatě nemožné.¹²

Maximální přesnost ($P=1$) je možné představit si tak, že každý dokument zařazený na seznam výsledků je právě pozitivum (TP) bez ohledu na jejich celkový počet. Pokud mám v databázi sto relevantních dokumentů a v odpovědi na můj dotaz je jich do seznamu výsledků zařazeno pouze pět a seznam výsledků neobsahuje žádný falešně pozitivní dokument, dosáhl jsem maximální přesnosti. Celkový počet falešně negativních dokumentů (FN), tedy relevantních dokumentů, které nebyly zařazeny do seznamu výsledků, při výpočtu přesnosti nehraje roli. Maximální úplnost ($R=1$) je možné představit si tak, že na seznam výsledků se musí dostat všechny relevantní výsledky bez ohledu na celkový počet informačního šumu tvořeného falešně pozitivními (FP) výsledky. Pokud tedy mám v databázi sto relevantních dokumentů a do seznamu výsledků jich zařadím pět set, přičemž nezůstane žádný falešně negativní dokument (FN), který bych vyhledáváním nezískal, dosáhnou maximální úplnosti vyhledávání.

Stoprocentně přesné je vyhledávání, které identifikuje právě jeden relevantní dokument – takové vyhledávání je ale mimořádně neúplně. Tento vztah ale platí i opačně. Pokud na seznam výsledků zařadím všechny dokumenty z právního informačního systému, určitě jsem zachytil všechny relevantní dokumenty a dosáhl jsem tak stoprocentní úplnosti. Takové vyhledávání je ale mimořádně nepřesné. Jakýkoli proces získávání informací je tak balancováním mezi přesností a úplností tak, abych si akcentem na jednu metriku nepotlačil metriku druhou. Právě z toho důvodu existuje míra F1, která je, jak bylo uvedeno výše, harmonickým průměrem přesnosti P a úplnosti R .

$$\text{Výpočet přesnosti } P:^{13} P = (TP) / (TP + FP)$$

$$\text{Výpočet úplnosti } R:^{14} R = (TP) / (TP + FN)$$

¹² Viz BING, Jon. Performance of Legal Text Retrieval Systems: The Curse of Boole. *Law Library Journal*, 1987, roč. 79, č. 2, s. 196.

¹³ Převzato z Ashley 2017 op. cit. s. 222.

¹⁴ Převzato tamtéž.

$$\text{Výpočet míry } F1: \text{ } F1 = \frac{2 * TP}{2 * TP + FN + FP}$$

Jednou ze slavných studií věnovaných evaluaci procesu získávání právních informací je studie STAIRS publikovaná v roce 1985.¹⁶ Studie poukazovala na velmi nízkou schopnost právníků formulovat relevantní plnotextové vyhledávací dotazy do databáze právních informací. Blair a Maron požádali v rámci této studie právníky o vyhledávání v databázi s tím, že jakmile získají alespoň 75 % všech relevantních dokumentů (tedy dosáhnou $R > 0,75$), mají vyhledávání ukončit a požádat výzkumníky o evaluaci tohoto vyhledávání. Po evaluaci výzkumníci zjistili, že ač měli právníci za to, že dosáhli slušné úplnosti vyhledávání, ve skutečnosti byly jejich výsledky velmi špatné ($R < 0,2$). Jakkoli se tato studie soustředila zejména na proces identifikace skutkově relevantních dokumentů a nikoli dokumentů právně relevantních, naznačila, že vyhledávání založené na plnotextových dotazech je mimořádně náročné¹⁷ a nevede k optimálním výsledkům. Následné studie pak demonstrovaly, že tento problém se projevuje i v jiných právních informačních systémech¹⁸ a že je typický pro vyhledávání, při kterých uživatelé usilují o vysokou úplnost.¹⁹

Plnotextové vyhledávání není jediným způsobem, jak získat z informačních systémů relevantní dokumenty. V současné době získáváme právní informace ze systémů využívajících plnotextové vyhledávání, znalostní inženýrství nebo můžeme využít získávání informací založené na zpracování přirozeného jazyka.²⁰ Všechny tyto postupy mají své výhody a nevý-

¹⁵ Převzato tamtéž.

¹⁶ Viz Blair a Maron 1985 op. cit.

¹⁷ Dabney označuje kolekce psaných záznamů s odkazem na Platóna za nevyhléditelně ob- skurní reprezentace idejí. Viz DABNEY, Daniel. The Curse of Thamus: An Analysis of Full- Text Legal Document Retrieval, *Law Library Journal*, 1986, roč. 78, č. 5, s. 40

¹⁸ Tamtéž.

¹⁹ Viz SORMUNEN, Eero. Extensions to the STAIRS Study – Empirical Evidence for the Hypo- thesised Ineffectiveness of Boolean Queries in Large Full-Text Databases. *Information Retrieval*, 2001, roč. 4, č. 3-4, s. 257-273.

²⁰ Viz MAXWELL, Tamsin a Burkhard SCHAFER. Concept and Context in Legal Information Retrieval. *Proceedings of Jurix 2008*. S. 64.

hody, protože perfektní míra F1 (tedy kombinace perfektní přesnosti P a úplnosti R) zůstává nedosažitelná. Systémy založené na znalostním inženýrství se pokouší vystihnout způsob, jakým si právníci pamatují případy, a toto zachytit v rámci počítačových algoritmů a datových struktur.²¹ Hlavním problémem těchto systémů je jejich obtížné využití v praxi. Jak uvádí Maxwell a Schafer, systémy využívající znalostního inženýrství jsou často velmi specificky zaměřené a umožňují využívat pouze velmi malé a vysoce strukturované kolekce dokumentů.²² Získávání informací založené na zpracování přirozeného jazyka pak má značnou výhodu ve škálovatelnosti a také vykazuje od počátku lepší výsledky.²³ Zpracování přirozeného jazyka získává v procesu získávání informací stále prominentnější pozici. Konečným cílem se zdá být snaha zajistit vyhledávání informací prostřednictvím odpovědí na otázky²⁴ nebo prostřednictvím konceptuálního vyhledávání.²⁵

Vyhledávání relevantních právních informací je tak velkou výzvou z pohledu použitých technologií a postupů. Neméně výzvou je ale i samotný termín *relevance* – nikoli tedy samotný proces vyhledávání, ale vůbec způsob, jakým rozlišujeme mezi relevantními a nerelevantními dokumenty. Termín *relevance* je totiž velmi často chápán intuitivně. Konceptualizace termínu *relevance* specificky pro právní prostředí se objevila až v roce 2017, kdy Opijnen a Santos přišli s následujícími kategoriemi *relevance*: algoritmická, bibliografická, tematická, kognitivní, situační a doménová.²⁶ Tato kategorizace je dle mého názoru zcela zásadní pro pochopení procesu vyhledávání. Některé kategorie mají projevy, které jsou v rámci práva téměř

²¹ Viz HAFNER, Carole. Conceptual Organization of Case Law Knowledge Bases. *Proceedings of ICAIL 1987*. S. 35.

²² Viz Maxwell a Schafer 2008 op. cit., s. 66.

²³ Srov. TURTLE, Howard. Natural Language vs. Boolean Query Evaluation: A Comparison of Retrieval Performance. *Proceedings of SIGIR 1994*. S. 212-220.

²⁴ Viz Maxwell a Schafer 2008 op. cit., s. 68.

²⁵ Viz GRABMAIR, Matthias, Kevin ASHLEY, Ran CHEN, Preethi SURESHKUMAR, Chen WANG, Eric NYBERG a Vern WALKER. Introducing LUIMA: an Experiment in Legal Conceptual Retrieval of Vaccine Injury Decisions Using a UIMA Type System and TOOLS. *Proceedings of ICAIL 2015*. S. 69-78.

²⁶ Viz VAN OPIJNEN, Marc a Cristiana SANTOS. On the Concept of Relevance in Legal Information Retrieval. *Artificial Intelligence and Law*, 2017, roč. 25, č. 1, s. 73.

univerzální – z hlediska práva jako domény je relevantní judikatura publikovaná ve sbírkách, například ve Sbírce soudních rozhodnutí a stanovisek Nejvyššího soudu. V tomto případě se jedná o doménovou relevanci, kde jsou individuální dokumenty vnímány jako relevantní na základě příslušnosti ke třídě dokumentů, která je vnímána jako relevantní. Jiné kategorie jsou však velmi subjektivní a přímo závislé na osobě, která se systémem pracuje. Zkušenému uživateli s dvacetiletou praxí přinese pouze malé množství dokumentů nové informace, zatímco čerstvému absolventovi může novou informaci přinést každý druhý dokument, který nalezne. V tomto případě se jedná o kognitivní relevanci.

Dvě z kategorií, které definovali Opijnen a Santos, jsou využity v experimentu popisovaném v tomto textu. Je to zejména z důvodu, že pokud nezohledníme žádnou konkrétní formu relevance, ale budeme se spoléhat na intuitivní chápání termínu, bude to jako porovnávání jablek s hruškami. První je tzv. tematická relevance, což je vztah mezi tématem formulovaným v uživatelském vyhledávání a tématem dokumentů. Vytvoření vztahu mezi dokumentem a tématem může být explicitní nebo implicitní prostřednictvím automatického indexování slov, manuálního indexování dokumentů nebo poloautomatizované klasifikace.²⁷ Druhou je tzv. kognitivní relevance, která vystihuje, zda je dokument relevantní pro konkrétního uživatele. Tento typ relevance je, jako takový, subjektivní a silně závislý například na uživatelské apriorní znalosti tématu.²⁸

3. MODEL A VÝZKUMNÉ OTÁZKY

Právní informace, ať jsou obsažené v právních předpisech, soudních rozhodnutích nebo jakýchkoli jiných dokumentech, musí být možné v právních informačních systémech vyhledávat. Z toho důvodu musí být dokumenty, které tyto informace obsahují, nějakým způsobem organizovány. Možnost uživatele vyhledat konkrétní dokument vyžaduje určitou organizaci, kterou uživatel musí poznat a na kterou se při vyhledávání může spolehnout.

²⁷ Tamtéž, s. 74.

²⁸ Tamtéž, s. 81.

Aby bylo možné od sebe odlišit jednotlivé způsoby vyhledávání – vyhledávací strategie, chceme-li – a dostatečně je studentům vysvětlit, vytvořili jsme si s kolegy Jakubem Míškem a Matějem Myškou zjednodušený model, který jsme před několika lety začali využívat při výuce na PrF MU. Pomocí tohoto zjednodušeného modelu studentům popisujeme a vysvětlujeme některé situace, se kterými se mohou při vyhledávání setkat. Právní informační systémy totiž, jak si nejspíše všiml každý, kdo je při své činnosti intenzivně využívá, neposkytují při srovnatelných dotazech totožné výsledky. V některých situacích mohou být dokonce rozdíly mezi systémy při využití srovnatelného dotazu větší než shody. Ve výuce používaný model, se kterým budu dále pracovat i v tomto textu, umožňuje tyto odlišnosti, když ne přímo vysvětlit, tak alespoň ukázat, jak vznikají.

Zmíněný model, který využíváme ve výuce, má dvě základní složky: data a metadata. V případě metadat pak rozlišujeme mezi objektivními a subjektivními metadaty.

Uživatelé v právních informačních systémech zajímají primárně data – tedy texty právních předpisů, soudních rozhodnutí, komentářů nebo jakýchkoli jiných zdrojů právních informací. Textem ustanovení jsme informováni o existenci určitého pravidla, které je textem tohoto ustanovení přímo konstituováno. Často si však nevystačíme pouze s textem ustanovení, ale při snaze toto ustanovení pochopit nám může významně pomoci například jeho zařazení do systematiky předpisu (například do konkrétní hlavy, části a podobně). Data samotná nepřináší v rámci českých právních informačních systémů výrazné problémy. Je zcela legitimní očekávat, že § 12 občanského soudního řádu nalezneme v právním informačním systému ASPI, Beck-online i Codexis, případně i v jakémkoli jiném. Stejně tak je zcela legitimní očekávat, že text tohoto ustanovení bude ve všech právních informačních systémech totožný. V případě soudních rozhodnutí, které se v systémech vyskytují, je situace poněkud komplikovanější. Editorские zásahy nebo přejímání rozhodnutí z různých zdrojů (zákonem předepsané sbírky jednotlivých vrcholných soudů, anotace v časopisech, webové databáze vrcholných soudů a podobně) mohou vést k rozdílům v textech těchto

rozhodnutí.²⁹ Obecně ale platí, že i v případě soudních rozhodnutí Ústavního soudu, Nejvyššího soudu a Nejvyššího správního soudu tak nějak očekáváme, že texty – tedy data – budou stejné (nebo aspoň velmi podobné) napříč systémy.

Organizace dokumentů výhradně v datové podobě by však znamenala, že se jako uživatelé musíme při jejich vyhledávání spolehnout na plnotextové dotazy. Plnotextové vyhledávání v právních informačních systémech nám umožňuje pomocí různých operátorů specifikovat, zda nás zajímá slovní spojení, zda se mají v textu vyskytovat všechna zadaná slova nebo stačí pouze jedno z nich, případně zda se zadávaná slova mají vyhledat v určité vzdálenosti.³⁰ Plnotextové vyhledávání s sebou ale nese nevýhody – uživatel sám musí definovat, jaké výrazy vyskytující se v textu dokumentu signalizují, že je pro něj dokument relevantní. Jakkoli přímočaře tento požadavek zní, rozhodně se nejedná o triviální záležitost. Abychom totiž dokázali konkrétní dokument nebo ustanovení nalézt plnotextovým dotazem, musíme do určité míry tušit, co je v tomto dokumentu napsáno. Alespoň do určité míry tedy musíme vědět, jak je například dané ustanovení naformulováno, jaké obsahuje výrazy, slovní spojení a podobně. U relativně strukturovaných a terminologicky pevně vymezených zdrojů je to často otázka cviku a zkušenosti – právník, který si více pamatuje a „více toho viděl“, má zjednodušenou pozici. U soudních rozhodnutí však takové vyhledávání může být náročné i pro velmi zkušeného uživatele. Při vyhledávání rozsáhlých databází například důkazního materiálu, které budou relativně nestrukturované, je to pak ještě horší, jak ostatně ukázala i shora citovaná STAIRS studie z 80. let minulého století.³¹ V rámci plnotextového vyhledávání hraje roli také to, jak daný uživatel umí využívat různé možnosti nabízené informačním systémem. Schopnost práce se systémem se relativně jednoduše nabývá, ale udržet uživatelskou kompetenci (tedy schopnost aktivně používat všechny nebo většinu dostupných funkcí

²⁹ Srov. HARAŠTA, Jakub. Nejednoznačnost odkazů k soudním rozhodnutím a možnosti řešení. *Revue pro právo a technologie*, roč. 6, č. 11, s. 15-28.

³⁰ Ne každý právní informační systém nabízí všechny tyto možnosti. Jaké operátory zařadit je i otázkou rozhodnutí tvůrců systému na základě třeba výzkumu uživatelských preferencí.

³¹ Viz Blair a Maron 1985 op. cit.

v podobě jednotlivých plnotextových operátorů) je poměrně náročné.³² I z těchto důvodů hrají při vyhledávání značnou roli metadata, tedy data o datech.

Metadata nám umožňují vyhledat data (ustanovení, soudní rozhodnutí atd.) nikoli nutně pomocí jejich obsahu, ale podle nějaké vlastnosti. Je tak možné omezit vyhledávání soudních rozhodnutí na rozhodnutí vydaná například Nejvyšším soudem v období od 1. ledna 2010 do 31. prosince 2011. Uživatel sám musí přijít s kritérii, která mu pomohou vyhledat relevantní informace. Část kritérií plyne z možností konkrétního informačního systému (co nabízí a co naopak nenabízí), část pak plyne z vědomostí o právním řádu, v jehož rámci uživatel relevantní dokumenty vyhledává. Jak bylo uvedeno výše, v našem modelu rozeznáváme metadata dvojího typu – objektivní metadata a subjektivní metadata. Objektivní metadata jsou specifická pro daný dokumenty a vyskytují se napříč všemi systémy, zatímco subjektivní metadata jsou často specifická pro konkrétní systém. Mohou se tak u jednotlivých iterací dokumentu zařazených v různých informačních systémech lišit. Příkladem objektivního metadatového záznamu je například spisová značka rozhodnutí, datum jeho vydání nebo soud původu. Pomocí těchto kritérií jsme schopni omezit vyhledávání a zároveň jako uživatelé legitimně očekáváme, že dokument bude pomocí těchto metadatových záznamů popsán ve všech systémech. Se subjektivními metadaty je situace opět o něco složitější, ale pro tento text zcela zásadní.

Subjektivní metadatové záznamy se mohou mezi systémy lišit. Mohou se lišit například přítomností příslušného metadatového záznamu. Například ASPI využívá metadatovou kategorii „Oblasti úpravy“ pro popis dokumentů, kterou ale Codexis nemá ekvivalentním způsobem zařazenou. Z mého pohledu jsou ale mnohem zajímavější odlišnosti v metadatových kategoriích, které jsou jako kritérium zařazené do více informačních systémů. Příkladem tohoto stavu může být metadatová kategorie identifikující rozhodnutí jako vztahující se k určitému ustanovení. Tato kategorie je nejen

³² Příkladem může být vyhledávací funkce umožňující nastavování vzdáleností mezi zadanými výrazy při plnotextovém v právním informačním systému ASPI. Tato funkce je z mých zkušeností uživateli hodnocena velmi pozitivně, ale zároveň na ni uživatelé často zapomínají a často si musejí její existenci připomínat.

přítomná v určité podobě v systémech ASPI, Beck-online i Codexis, ale je také mimořádně často využívána k vyhledávání relevantních dokumentů (a to nejen v České republice³³). V českém právu neexistuje žádný centrální seznam rozhodnutí, která by se vztahovala k nějakému ustanovení a byla tedy relevantní pro jeho výklad. Než si však povzdechne nad tím, jak je to české právo zpátečnické, musím poznamenat, že takovýto seznam nebude s největší pravděpodobností existovat v autoritativní podobě v žádné zemi kontinentální právní tradice. A zatímco spisová značka konkrétního rozhodnutí bude totožná ve všech třech zmíněných systémech (byť chyby a překlepy se samozřejmě mohou objevit), vytvoření vazby mezi ustanovením a rozhodnutím může být, při absenci nějaké centrální autority, vytvořeno s důrazem na různé aspekty.³⁴ Bude tedy přímo záležet na definici situace, kdy je nutné rozhodnutí a ustanovení provázat metadatovým záznamem. Pokud budeme například vztah definovat na základě zmínky daného ustanovení v soudním rozhodnutí, budeme vytvářet metadatové vazby na jiných místech, než pokud budeme za kritérium k vytvoření vazby považovat nejen zmínku daného ustanovení v soudním rozhodnutí, ale i jeho historických předchůdců. Abych uvedl praktický příklad, seznam výsledků rozhodnutí souvisejících s § 14 zákona č. 89/2012 Sb. může zahrnovat (i) rozhodnutí zmiňující § 14 zákona č. 89/2012 Sb. nebo (ii) rozhodnutí zmiňující § 14 zákona č. 89/2012 Sb. a zároveň § 6 zákona č. 40/1964 Sb., protože se jedná o „ideového“ předchůdce rešeršovaného ustanovení. Podobně můžeme zvolit variantu, která bude pro tvůrce právního informačního systému výrazně náročnější, kdy budeme vazbu mezi ustanovením a rozhodnutím vytvářet pouze tehdy, pokud je v soudním rozhodnutí vyhledávané ustanovení nejenom zmíněno, ale substantivně diskutováno (soud například rozebírá limity jeho aplikace). Takovéto výsledky jsou pro uživatele mimořádně přínosné, protože snižují jeho informační zahlcení. Přináší s sebou ale další palčivé otázky. Jednou z nich je například

³³ Viz Van Opijnen a Santos 2017 op. cit., s. 74-75.

³⁴ Srov. ARREDONDO, Pablo. Shepard for a Day: A Novel Class Exercise for Teaching Citators. *Legal Reference Service Quarterly*, 2015, roč. 34, č. 3, s. 244, kde autor volá po zdravém skepticizmu směrem k rozhodnutím označeným v rámci amerických informačních systémů červenými, žlutými nebo zelenými vlaječkami podle jejich vzájemného vztahu.

zajištění kontinuity tvorby těchto záznamů při personální nebo jiné změně na straně tvůrce informačních systému. Mělo by nás zajímat, kdo rozhoduje o tom, že je dané ustanovení v soudním rozhodnutí substantivně diskutováno. Jsou rozhodnutí zpracovávána v roce 2020 se stejnou péčí, s jakou byla zpracovávána v roce 2000? Jsou vazby vytvářeny podle stále stejného standardu?

Zdráhám se tento model nazvat teorií organizace dokumentů v právních informačních systémech. S kolegy jej nicméně používáme ve výuce, kde se nám celkem osvědčil. Rozdělení obsahu systémů na data a metadata a rozdělení metadat v systémech na metadata objektivní a subjektivní není složité a pomáhá ilustrovat některé problémy související s vyhledáváním v systémech. Jedná se samozřejmě o model idealizovaný – data nejsou zcela totožná, protože právní informační systémy neobsahují zcela stejné dokumenty. V důsledku odlišného přístupu při budování konkrétní databáze a tvorbě jejího uživatelského rozhraní nemusí být možné je ani totožným způsobem prohledávat. Drobné rozdíly plynou z odlišného postupu zvoleného pro plnotextovou indexaci při zařazení obsahu do systému³⁵, mohou se vyskytovat chyby (šotci) a podobně.

Podle toho, na jaké části tohoto modelu se uživatel při vyhledávání spolehne pak můžeme rozlišovat vyhledávací strategie zprostředkované, nezprostředkované a kombinované. Volbou zprostředkované strategie se uživatel spoléhá na tvůrce informačního systému, respektive na subjektivní metadata v systému obsažená. Typickým příkladem je všemi uživateli hojně využívaná funkce hledání judikatury související s konkrétním ustanovením. Volbou nezprostředkované strategie se uživatel spoléhá primárně sám na sebe a formuluje plnotextové dotazy nebo využívá objektivní metadata. Plnotextovým vyhledáváním nebo využitím objektivních metadat se sám snaží nadefinovat relevantní dokumenty, které hledá a které pro svoji práci potřebuje. Každý z těchto postupů má svá pro a proti. Zprostředkovanou strategií budeme, pokud shora uvedený model odpovídá realitě, dostávat výsledky závislé na použití toho kterého konkrétního systému. Na druhou

³⁵ Více viz ŠAVELKA, Jaromír, Matěj MYŠKA, Adam PTAŠNIK a Danuše SPÁČILOVÁ. *Právní informační systémy*. Brno: Tribun EU, 2011. S. 64-72.

stranu ale bude vyhledávání zřejmě časově i kognitivně méně náročné. Ne-zprostředkovanou strategií strávíme větší množství času a budeme u jejího použití muset jako uživatelé i více přemýšlet. Na druhou stranu se ale oprostíme od závislosti seznamu výsledků na využití konkrétního systému. Kombinované vyhledávací strategie pak využívají kombinaci jak nezprostředkovaných (data, objektivní metadata), tak zprostředkovaných (subjektivní metadata) informací obsažených v systémech. Například onen shora uvedený příklad s dotazem na judikaturu související s § 14 zákona č. 89/2012 Sb. můžeme doprovodit plnotextovým dotazem ve snaze zajistit si menší množství výsledků k dalšímu zpracování.

Z tohoto modelu pak přímo plynou i výzkumné otázky, na které v tomto textu hledám odpověď.

Za prvé: liší se výsledky vyhledávání podle subjektivních metadat mezi jednotlivými systémy? A pokud ano, jak moc? Pokud přijmeme z představeného modelu vycházející premisu, mezi systémy ASPI, Beck-online a Codexis budou při využití srovnatelné zprostředkované vyhledávací strategie seznamy výsledků odlišné. V rámci této otázky tak vlastně budu ověřovat, (a) zda shora představený model odpovídá realitě a (b) jak velké jsou rozdíly mezi výsledky poskytnutými systémy v odpovědi na srovnatelné dotazy.

Za druhé: povede zařazení plnotextového kritéria do vyhledávání podle subjektivních metadat k dosažení výrazně přesnějších výsledků v rámci jednotlivých právních informačních systémů? Tato otázka přímo navazuje na první výzkumnou otázku, kde dochází k využití zprostředkované vyhledávací strategie. Při odpovědi na druhou otázku dojde k využití kombinované vyhledávací strategie (subjektivní metadata plus plnotextový dotaz) a pokusím se zjistit, zda budou výsledky oproti využití zprostředkované vyhledávací strategie přesnější ve smyslu vyšší dosažené úplnosti R, přesnosti P a tím i míry F1.

Za třetí: povede zařazení plnotextového kritéria do vyhledávání podle subjektivních metadat k dosažení podobnějších výsledků napříč jednotlivými informačními systémy? Tato otázka přímo navazuje na obě předcházející výzkumné otázky. Pokud přijmeme z představeného modelu vycházející

premisu, mezi systémy ASPI, Beck-online a Codexis budou při využití srovnatelné kombinované vyhledávací strategie výsledky sice odlišné, ale budou si navzájem podobnější než při využití zprostředkované vyhledávací strategie.

4. METODA A DATA

4.1 ZPROSTŘEDKOVANÉ ÚDAJE: METADATOVÉ VYHLEDÁVÁNÍ

Pochopitelně zde není možné porovnat výsledky poskytované právními informačními systémy ASPI, Beck-online a Codexis kompletně a komplexně, tedy porovnat všechny v těchto systémech existující vazby. Proto jsem pro tento text zvolil porovnání judikatury související se zákonem č. 121/2000 Sb., autorský zákon (dále jen AutZ), a s dnes již neúčinným zákonem č. 101/2000 Sb., o ochraně osobních údajů (dále jen ZOOÚ). K jednotlivým ustanovením těchto dvou předpisů byla mezi 8. a 10. říjnem 2018 vytěžena související rozhodnutí českých soudů pomocí metadatové vazby označující tzv. související judikaturu. Tuto metadatovou vazbu obsahují pod různými názvy všechny tři právní informační systémy a jedná se tak o dotaz, který je srovnatelný svým významem při vyhledávání.

Soudní rozhodnutí získaná popsáním způsobem jsem zorganizoval do databáze, kde bylo každé z rozhodnutí uspořádáno do páru obsahujícího ustanovení *U* a rozhodnutí *R* a bylo indikováno, ve kterém systému je možné tento pár zachytit. Tím byla vytvořena databáze získaných rozhodnutí, jejíž výňatek je možné vidět v Tab. 1. Zápis v tabulce na prvním řádku indikuje, že pár *U-R* zahrnující ustanovení § 2 AutZ a rozhodnutí č.j. 10 A 186/2013-49 je možné získat ze systému ASPI, ale nikoli ze systémů Beck-online a Codexis. Uvedené rozhodnutí se sice v těchto databázích nachází, ale nevrací se jako výsledek k vyhledávání judikatury související s § 2 AutZ. Podobně pak druhý řádek ukazuje, že pár *U-R* zahrnující ustanovení § 21 ZOOÚ a rozhodnutí sp. zn. I. ÚS 1783/10 je možné získat z právních informačních systémů ASPI a Codexis, ale nikoli ze systému Beck-online. Opět nám tento záznam neříká nic o tom, zda se dané rozhodnutí v právním informačním systému nachází, ale pouze indikuje, zda je

rozhodnutí metadatově provázáno s konkrétním ustanovením jako související.

Ustanovení <i>U</i>	Rozhodnutí <i>R</i>	ASPI	Beck-online	Codexis
§ 2 AutZ	10 A 186/2013-49	1	0	0
§ 21 ZOOÚ	I. ÚS 1783/10	1	0	1

Tab. 1 Příklad dvou řádků z databáze párů *U-R* s indikací systémů, ve kterých je pár možné identifikovat pomocí přítomností metadatové vazby

Tímto způsobem tak vznikla databáze, která obsahuje každý pár *U-R*, který se v systémech ASPI, Beck-online a Codexis vyskytl alespoň jednou – tedy obsahuje rozhodnutí *R*, které bylo při metadatovém vyhledávání judikatury související s ustanovením *U* zachyceno alespoň v jednom z těchto tří systémů. Do vyhledávání byla zahrnuta všechna ustanovení AutZ a ZOOÚ na úrovni jednotlivých paragrafů. Nižší jednotky, tedy odstavce a písmena, nebyla zohledňována.

Na tomto místě je naprosto zásadní zopakovat, že neexistuje žádný autoritativní nebo centrální seznam rozhodnutí, která by se měla objevit v seznamu výsledků v odpovědi na konkrétní dotaz. Shora uvedená databáze tak nahrazuje tento seznam. Porovnávání výsledků poskytnutých jednotlivými systémy se tak děje srovnáním s jakousi idealizovanou databází, která je představována sjednocením množin výsledků poskytnutých jednotlivými systémy. Tato základní databáze (de facto seznam rozhodnutí uvedených jako související s jednotlivými ustanoveními) bude dále obohacena o zavedení plnotextového vyhledávacího parametru a o zohlednění kognitivní relevance jednotlivých rozhodnutí pro expertního a neexpertního uživatele.

4.2 NEZPROSTŘEDKOVANÉ ÚDAJE: PLNOTEXTOVÉ VYHLEDÁVÁNÍ

Do výše uvedené databáze jsem ve druhém kroku zanesl plnotextový parametr, který jsem využil jako druhé kritérium pro vyhledávání. Použitý plnotextový parametr jsem stanovil jako výskyt tří obecných termínů souvisejících s AutZ a tří obecných pojmů souvisejících se ZOOÚ v textu soudního

rozhodnutí. Tyto pojmy se v textu mohou vyskytovat kdekoli a nemusí tak mezi nimi být žádná vzdálenostní či jiná souvislost. Pro AutZ byl plnotextový parametr zvolen jako současný výskyt tří následujících slov: *autor*, *dílo*, *duševní*. Pro ZOOÚ byl plnotextový parametr zvolen jako současný výskyt tří následujících slov: *údaj*, *zpracování*, *správce*. Takováto vyhledávací strategie je mimořádně jednoduchá a je zvolena s přihlédnutím k předpisu, ale nikoli s přihlédnutím ke konkrétnímu ustanovení.

Do databáze, jejíž dva řádky jsou uvedeny v Tab. 1 tak přibyly další hodnoty. Výňatek databáze se zařazeným plnotextovým parametrem je možné vidět v Tab. 2. První řádek indikuje, že pár *U-R* obsahující ustanovení § 2 AutZ a rozhodnutí č.j. 10 A 186/2013-49 v sobě obsahuje dokument, který splňuje plnotextovou podmínku stanovenou jako související s ustanovením AutZ (tedy současný výskyt slov *autor*, *dílo* a *duševní* v textu rozhodnutí). Druhý řádek pak indikuje, že pár *U-R* obsahující ustanovení § 21 ZOOÚ a rozhodnutí sp. zn. I. ÚS 1783/10 v sobě obsahuje dokument, který splňuje plnotextovou podmínku stanovenou pro rozhodnutí související s ustanovením ZOOÚ (tedy současný výskyt slov *údaj*, *zpracování* a *správce*).

Ustanovení <i>U</i>	Rozhodnutí <i>R</i>	Plnotextové (AutZ)	Plnotextové (ZOOÚ)
§ 2 AutZ	10 A 186/2013-49	1	-
§ 21 ZOOÚ	I. ÚS 1783/10	-	1

Tab. 2 Příklad dvou řádků z databáze párů *U-R* s indikací splnění plnotextového parametru pro AutZ/ZOOÚ

4.3 KOGNITIVNÍ RELEVANCE ROZHODNUTÍ

V návaznosti na vytvoření databáze obsahující jak výsledky metadatového vyhledávání, tak výsledky vyhledávání s plnotextovým parametrem, jsem pozval tři osoby s právním vzděláním k dalšímu zpracování výsledků. Jejich úkolem bylo po předložení páru *U-R* rozhodnout, zda dané rozhodnutí *R* přispívá k pochopení ustanovení *U* či nikoli. Jeden z přizvaných právníků byl odborníkem na problematiku ochrany osobních údajů a dostal

tak za úkol vytvořit expertní anotace pro ZOOÚ. Druhý byl odborníkem na problematiku autorského práva a dostal tak za úkol vytvořit expertní anotace pro AutZ. Třetí přizvaný nebyl odborníkem ani v jedné této oblasti a dostal tak za úkol vytvořit neexpertní anotace pro oba předpisy. Do databáze, jejíž dva řádky jsou v Tab. 1 (metadatové vyhledávání) a následně i v Tab. 2 (plnotextový parametr), tak přibyly další hodnoty. Výňatek je možné vidět v Tab. 3. První řádek této tabulky indikuje, že pár *U-R* tvořený ustanovením § 2 AutZ a rozhodnutím č.j. 10 A 186/2013-49 byl v rámci expertních anotací hodnocen jako relevantní – rozhodnutí tedy přispívá odborníkovi k pochopení daného ustanovení (je kognitivně relevantní) – zatímco v rámci neexpertních anotací byl hodnocen jako nerelevantní. Druhý řádek pak indikuje, že pár *U-R* tvořený ustanovením § 21 ZOOÚ a rozhodnutím sp. zn. I. ÚS 1783/10 byl hodnocen jako kognitivně relevantní (přispívající k pochopení ustanovení) v rámci expertních i neexpertních anotací.

Ustanovení <i>U</i>	Rozhodnutí <i>R</i>	Expertní anotace	Neexpertní anotace
§ 2 AutZ	10 A 186/2013-49	1	0
§ 21 ZOOÚ	I. ÚS 1783/10	1	1

Tab. 3 Příklad dvou řádků z databáze párů *U-R* s indikací kognitivní relevance z pohledu expertního a neexpertního anotátora

Považuji za nutné na tomto místě vysvětlit, proč jsem do experimentu zahrnul i kognitivní relevanci. Začnu apelem na uživatelskou zkušenost každého, kdo s nějakým informačním systémem musel pracovat. Je poměrně častým jevem, že informační systém zařadí na seznam výsledků dokument, který nám v konečném důsledku k ničemu není. Splňuje totiž nějaké metriky, pro které ho tvůrce systému považoval za důležitý (tematicky relevantní, doménově relevantní a podobně), ale v situaci konkrétního uživatele (z hlediska kognitivní relevance) nám k ničemu není – nepřináší nám novou informaci. Toto je klasický problém všech informačních systémů a databází.

Pokud nám hypotetický systém X na dotaz vrátí seznam výsledků o 20 rozhodnutích a hypotetický systém Y nám na srovnatelný dotaz vrátí 2 rozhodnutí, na první pohled může být lepší využít systému X, který poskytne větší množství relevantních výsledků (v tomto případě mluvím o tematické relevanci v podobě provázání ustanovení a soudního rozhodnutí). Pokud ale do vyhodnocování zařadíme jinou formu relevance (například právě kognitivní relevanci), můžeme zjistit, že z dvaceti rozhodnutí identifikovaných použitím systému X nám k vyřešení problému, před který jsme postaveni, pomáhají pouze tři. Oproti tomu, ze dvou rozhodnutí identifikovaných jako relevantní systémem Y jsou kognitivně relevantní obě dvě. Na první pohled jednoduché srovnání systémů je najednou náročnější a méně přímočaré. Systém X sice identifikoval více kognitivně relevantních rozhodnutí, ale bylo nutné je odfiltrovat dalším zpracováním (přečtením) rozhodnutí, která jako relevantní (související s ustanovením) identifikoval tvůrce příslušného systému. S použitím systému X se tak váže větší časová náročnost a někteří uživatelé tak budou inklinovat k využití systému Y. Pro různé typy uživatelů pak bude sada kognitivních rozhodnutí vypadat jinak v závislosti na odbornosti v konkrétní oblasti a na zkušenostech.

4.4 METODA POROVNÁNÍ VÝSLEDKŮ

Shora uvedeným postupem jsem vytvořil idealizovanou databázi, která mi poslouží pro porovnání jednotlivých systémů a jednotlivých vyhledávacích strategií a tím k zodpovězení stanovených výzkumných otázek. Na tomto místě je nutné připomenout, že cílem tohoto textu není určit, jestli jsou v systémech ASPI, Beck-online a Codexis správně evidované relevantní dokumenty. Vzhledem k chybějícímu autoritativnímu seznamu rozhodnutí souvisejících s konkrétním ustanovením je takový úkon zcela jistě nemožný. Cílem je zjistit, jaká proporce párů *U-R* získaných ze všech tří systémů se vyskytuje v jednotlivých systémech (první výzkumná otázka), jestli vede zařazení plnotextového parametru do vyhledávání k lepším výsledkům (druhá výzkumná otázka) a konečně, zda jsou si výsledky napříč systémy po zařazení plnotextového parametru vzájemně podobnější (třetí výzkumná otázka). Kde je to vhodné, jsou dále v textu uváděny konkrétní

hodnoty přesnosti P , úplnosti R a míry $F1$, kterých bylo dosaženo porovnáním jednotlivých právních informačních systémů (ASPI, Beck-online, Codexis) s popsanou idealizovanou databází.

5. VÝSLEDKY

5.1 ZÁKLADNÍ PŘEHLED

Idealizovaná databáze, která byla vytvořena sjednocením množin párů $U-R$ získaných z právních informačních systémů ASPI, Beck-online a Codexis obsahuje 403 párů pro AutZ a 262 párů pro ZOOÚ. V případě AutZ je 307 těchto párů možné získat z právního informačního systému ASPI, 135 ze systému Beck-online a 248 ze systému Codexis. V případě ZOOÚ je 176 párů možné získat z právního informačního systému ASPI, 81 z právního informačního systému Beck-online a 140 ze systému Codexis. Rozdíly mezi systémy tedy existují a jsou zcela evidentní již z prosté sumy nalezených párů $U-R$.

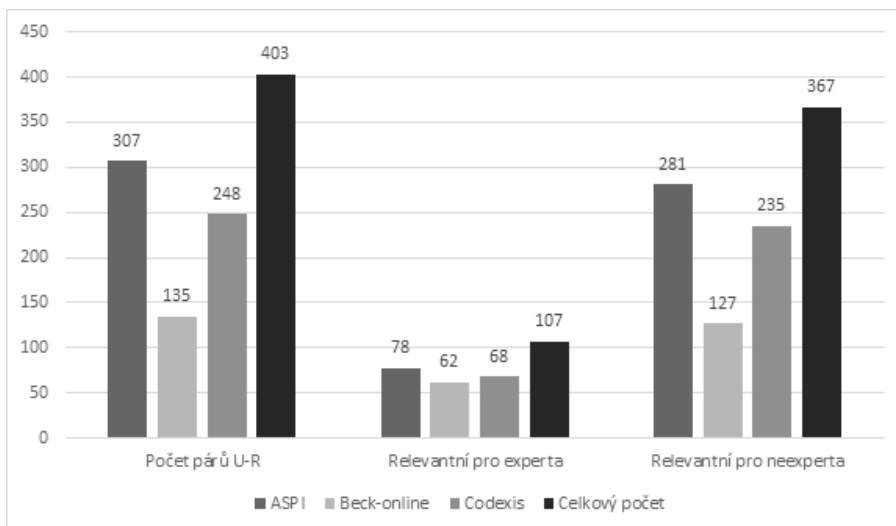
Logickou námitkou je, že právní informační systémy obsahují jiná rozhodnutí. Pokud ale prohlédneme jednotlivé databáze, tak zjistíme, že ze 403 párů $U-R$, které se objevují v idealizované databázi pro AutZ, jich může ASPI obsahovat 380 (94,29 %), Beck-online 368 (91,32 %) a Codexis dokonce 401 (99,51 %). Tato čísla znamenají, že z páru $U-R$ se rozhodnutí R v dané databázi vyskytuje a vazbu mezi ustanovením a rozhodnutím by tedy bylo možné hypoteticky vytvořit. V případě ZOOÚ je z celkového počtu 262 párů možné v ASPI i v Codexisu možné vytvořit 257 z nich (98,1 %). V případě Beck-online je možné jich vytvořit 247 (94,27 %). Rozdíly mezi systémy co do dokumentů v nich obsažených tedy existují. Tímto ale není možné vysvětlit rozdíly v párech $U-R$ (tedy rozhodnutích indexovaných jako souvisejících s určitým ustanovením) objevujících se v jednotlivých databázích. Zjednodušeně řečeno: je jen velmi malý počet rozhodnutí, které by se nenacházely ve všech třech systémech. Konkrétně se jedná o 19 rozhodnutí z celkového počtu 193 pro AutZ (obsah jednotlivých systémů je tedy shodný na 90,16 %) a o 9 rozhodnutí z celkových 154 pro ZOOÚ (obsah jednotlivých systémů je tedy shodný na 94,16 %).

Rozdíly mezi systémy tak primárně nejsou tvořeny tím, jestli se konkrétní rozhodnutí v příslušném systému nachází, ale opravdu tím, zda je dané rozhodnutí indexováno.

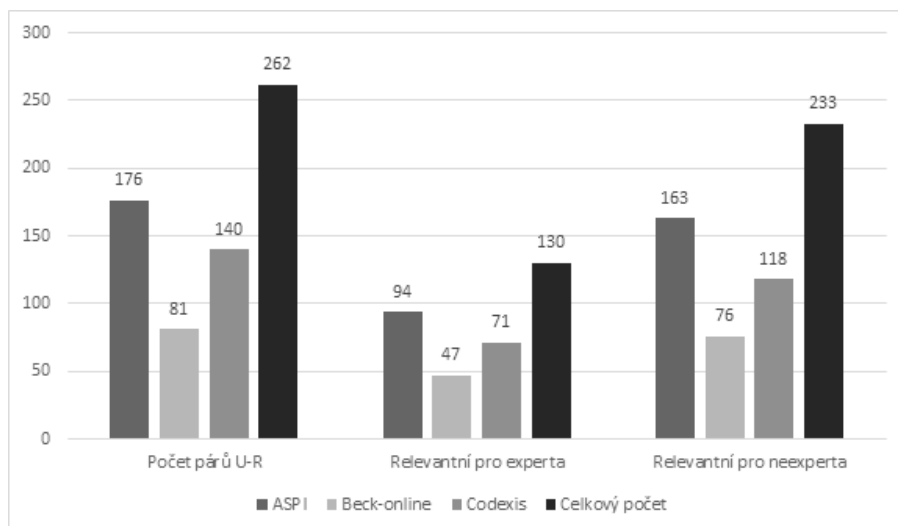
Ze 403 párů *U-R* tematicky relevantních (metadatově provázaných) pro AutZ jich expertní anotátor označil 107 (cca 26 %) jako kognitivně relevantní a neexpertní anotátor pak označil za relevantní 367 párů (cca 91 %). Z počtu 262 párů *U-R* tematicky relevantních pro ZOOÚ jich expertní anotátor označil za kognitivně relevantní 130 (cca 49 %) a neexpertní anotátor pak 233 (cca 90 %).

Společným základem, na kterém se tvůrci všech tří systémů shodli, je přítomnost 74 párů *U-R* v případě AutZ (cca 18 % celkového počtu) a 44 párů v případě ZOOÚ (cca 17 % celkového počtu).

Přehledněji jsou tyto statistiky znázorněné v rámci Obr. 1 a Obr. 2.



Obr. 1 Statistiky párů *U-R* pro ustanovení AutZ a jejich distribuce napříč systémy s přihlédnutím ke kognitivní relevanci pro expertního a neexpertního uživatele



Obr. 2 Statistiky párů U-R pro ustanovení ZOOÚ a jejich distribuce napříč systémy s přihlédnutím ke kognitivní relevanci pro expertního a neexpertního uživatele

Čísla, tedy počty párů *U-R* vyskytujících se v jednotlivých systémech, jsou tedy odlišná. Základní předpoklad, že při využití subjektivních metadata (tedy vazeb mezi ustanovením a soudním rozhodnutím) pro vyhledávání odlišná budou, se tak potvrdil.

V další části do analýzy zahrnu i kognitivní relevanci, čímž bude možné získat zajímavější vhled do existujících rozdílů mezi jednotlivými systémy. Kromě sumy, tedy prostého konstatování, že konkrétní systém obsahuje více či méně indexovaných párů *U-R* je po zařazení kognitivní relevance možné rozdíly kvantifikovat. Sjednocením množin rozhodnutí získaných z jednotlivých informačních systémů dojde k vytvoření onoho chybějícího autoritativního seznamu souvisejících rozhodnutí, byť pouze pro účely tohoto textu. Tento autoritativní seznam je pak možné porovnat s vyhledáváním v jednotlivých systémech a vypočítat přesnost *P*, úplnost *R* a míru *F1*. Pak můžeme srovnat výsledky poskytnuté jednotlivými systémy a diskutovat vhodnost použití různých vyhledávacích strategií z pozice různých typů uživatelů.

5.2 PODOBNOST VÝSLEDKŮ POSKYTOVANÝCH JEDNOTLIVÝMI SYSTÉMY

V této fázi je možné přikročit ke srovnání systémů. Za tímto účelem z idealizované databáze vzniklé sjednocením množin rozhodnutí získaných z právních informačních systémů ASPI, Beck-online a Codexis získáme výsledky pro jednotlivé právní informační systémy. Cílem je určit, jaká část párů *U-R* z idealizované databáze, se nachází v právním informačním systému ASPI, Beck-online a Codexis a jaká část výsledků v jednotlivých systémech je kognitivně relevantní z pohledu expertního a neexpertního uživatele. V Tab. 4 jsou uvedeny hodnoty přesnosti *P*, úplnosti *R* a míry *F1* vypočtené pro jednotlivé systémy a oba předpisy z pohledu expertního i neexpertního uživatele.

		ASPI			Beck-online			Codexis		
		P	R	F1	P	R	F1	P	R	F1
AutZ	Expertní uživatel	0,254	0,729	0,377	0,459	0,579	0,512	0,274	0,635	0,383
	Neexpertní uživatel	0,915	0,766	0,834	0,941	0,346	0,506	0,948	0,640	0,764
ZOOÚ	Expertní uživatel	0,534	0,723	0,614	0,580	0,362	0,445	0,507	0,546	0,526
	Neexpertní uživatel	0,926	0,700	0,797	0,938	0,326	0,484	0,843	0,506	0,633

Tab. 4 Evaluace výsledků vyhledávání jednotlivými systémy v idealizované databázi judikatury související s AutZ a ZOOÚ z pohledu expertního a neexpertního uživatele

Jakým způsobem tyto výsledky chápat vysvětlíme na příkladu výsledků ASPI. První řádek nám pro právní informační systém ASPI říká, že ze všech párů *U-R* nacházejících se v idealizované databázi, kde expertní anotátor označil rozhodnutí *R* jako relevantní pro pochopení ustanovení *U*, jsme vyhledáváním v ASPI dosáhli přesnosti 25,4 % a úplnosti 72,9 %. Vyhledáváním v ASPI nám tedy poskytlo 72,9 % ze všech párů kognitivně relevantních pro expertního uživatele, které se vyskytují ve všech třech informačních systémech zařazených do tohoto textu. Přesnost 25,4 % pak ale znamená, že všechny tyto relevantní páry byly skryty v množství falešně

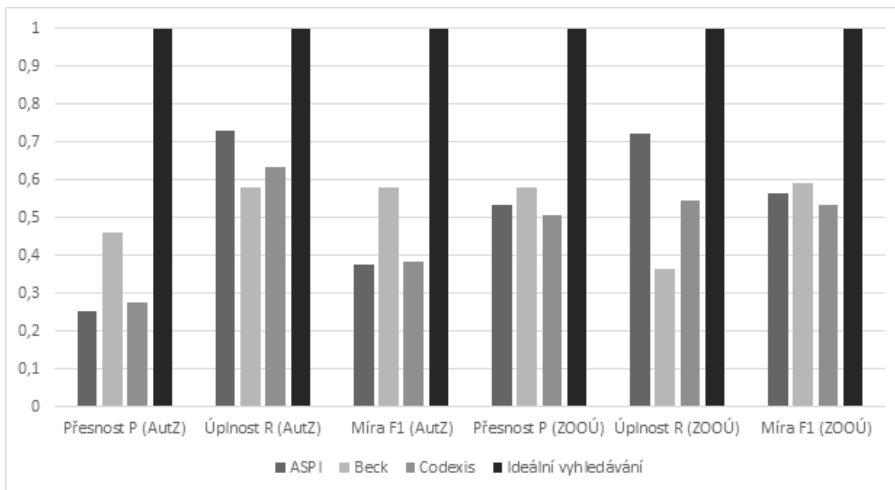
pozitivních výsledků, tedy výsledků, které expertní anotátor nepovažoval za relevantní, ale právní informační systém ASPI je zahrnul do seznamu výsledků.

Poněkud zjednodušeně tedy: pokud si manuálně zpracuji a ohodnotím všechna rozhodnutí, která jsou v právních informačních systémech navázána k AutZ, vykazuje vyhledávání v ASPI pokrytí tří z každé čtveřice relevantních rozhodnutí (to čtvrté bylo expertem posouzeno jako relevantní a zároveň bylo indexováno jiným systémem, než je ASPI). Toto je ale vykoupeno tím, že z každých čtyř rozhodnutí indexovaných v právním informačním systému ASPI vyhodnotil expert jako relevantní jenom jedno. Relativně malé procento relevantních rozhodnutí tedy zůstalo právním informačním systémem ASPI nezachycené (malý počet falešně negativních dokumentů), ale zároveň se jako uživatel budu muset při vyhodnocování výsledků probrat poměrně velkým množstvím dokumentů, které mi pro mé expertní chápání konkrétního ustanovení nepřinesou žádné informace (velký počet falešně pozitivních dokumentů).

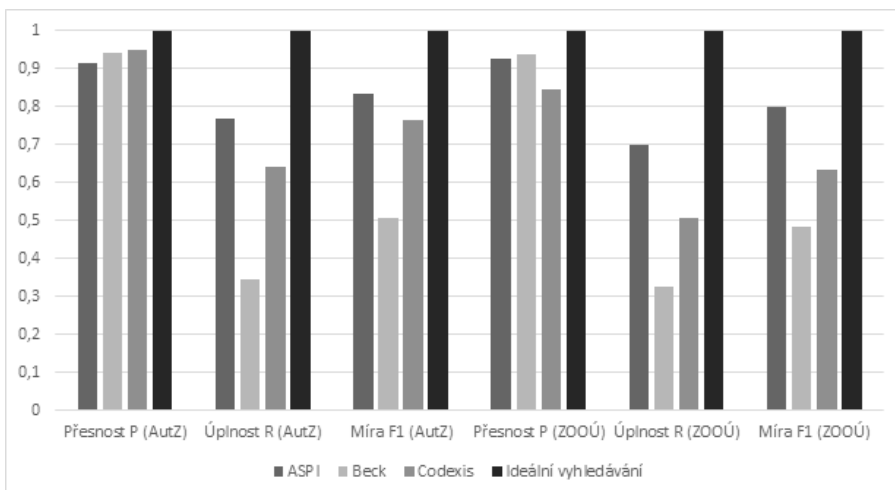
Situace je pak, celkem pochopitelně, odlišná pro neexpertního uživatele. Neexpertní anotátor totiž označil výrazně větší množství soudních rozhodnutí jako kognitivně relevantní. V důsledku nižší znalosti dané domény takovému uživateli přináší více soudních rozhodnutí nové informace. Pro takového uživatele bude mít metadatové vyhledávání v právním informačním systému ASPI, jak je evidentní z Tab. 4, přesnost P 91,5 % a úplnost R 76,6 %. Neexpertní uživatel tedy nedostane všechny dokumenty (zhruba každé čtvrté rozhodnutí označené anotátorem jako důležité mine). Na druhé straně mu ale více než devět z deseti rozhodnutí, které dostane v seznamu výsledků, přinese nové informace.

Srovnání jednotlivých metrik pro právní informační systémy ASPI, Beckonline a Codexis z pozice expertního uživatele je graficky znázorněno na Obr. 3 a z pozice neexpertního uživatele na Obr. 4. Ideálním vyhledáváním (míra $FI = 1,00$), se kterým jsou hodnoty srovnávány, je onen perfektní stav, po kterém všichni při práci s právními informačními systémy toužíme. Tedy stav, kdy na svůj dotaz dostaneme (i) všechna rozhodnutí, která nám pomohou pochopit nebo aplikovat konkrétní ustanovení (úplnost $R = 1,00$)

a (ii) všechna rozhodnutí, která dostaneme, jsou pro nás relevantní (přesnost $P = 1,00$).



Obr. 3 Porovnání úspěšnosti vyhledávání mezi systémy s ideálním vyhledáváním z pozice expertního uživatele



Obr. 4 Porovnání úspěšnosti vyhledávání mezi systémy s ideálním vyhledáváním z pozice neexpertního uživatele

5.3 ZMĚNA V ÚSPĚŠNOSTI VYHLEDÁVÁNÍ PO ZAVEDENÍ PLNOTEXTOVÉHO PARAMETRU

V této fázi je možné přikročit ke srovnání systémů po zavedení plnotextového parametru. Za tímto účelem z idealizované databáze, vzniklé sjednocením množin rozhodnutí, získaných z právních informačních systémů ASPI, Beck-online a Codexis, získáme výsledky pro jednotlivé právní informační systémy. Cílem je určit, jaká část párů *U-R* z idealizované databáze se nachází v právním informačním systému ASPI, Beck-online a Codexis a jaká část výsledků v jednotlivých systémech je kognitivně relevantní z pohledu expertního a neexpertního uživatele, po zařazení plnotextového parametru nad rámec vyhledávání souvislostí. V Tab. 5 jsou uvedeny hodnoty přesnosti *P*, úplnosti *R* a míry *F1* vypočtené pro jednotlivé systémy a pro oba předpisy z pohledu expertního a neexpertního uživatele. V Tab. 6 je pak znázorněno, jakým způsobem se hodnoty změnilы zařazením plnotextového parametru do vyhledávání rozhodnutí souvisejících s konkrétním ustanovením proti vyhledávání bez plnotextového parametru (jedná se tedy o rozdíl hodnot uvedených v Tab. 4 a hodnot uvedených v Tab. 5).

		ASPI			Beck-online			Codexis		
		P	R	F1	P	R	F1	P	R	F1
AutZ	Expertní uživatel	0,333	0,271	0,299	0,517	0,290	0,371	0,366	0,280	0,318
	Neexpertní uživatel	0,954	0,266	0,366	0,950	0,156	0,267	0,927	0,207	0,339
ZOOÚ	Expertní uživatel	0,602	0,500	0,546	0,612	0,315	0,416	0,606	0,485	0,539
	Neexpertní uživatel	0,963	0,446	0,610	0,955	0,275	0,426	0,913	0,408	0,564

Tab. 5 Evaluace výsledků vyhledávání jednotlivými systémy v idealizované databázi judikatury související s AutZ a ZOOÚ z pohledu expertního a neexpertního uživatele po zařazení plnotextového parametru

		ASPI			Beck-online			Codexis		
		P	R	F1	P	R	F1	P	R	F1
AutZ	Expertní uživatel	0,079 ↑	0,458 ↓	0,077 ↓	0,057 ↑	0,290 ↓	0,141 ↓	0,091 ↑	0,356 ↓	0,066 ↓
	Neexpertní uživatel	0,039 ↑	0,540 ↓	0,469 ↓	0,001 ↑	0,191 ↓	0,239 ↓	0,021 ↓	0,433 ↓	0,426 ↓
ZOOÚ	Expertní uživatel	0,068 ↑	0,223 ↓	0,069 ↓	0,032 ↑	0,046 ↓	0,029 ↓	0,010 ↑	0,061 ↓	0,013 ↑
	Neexpertní uživatel	0,037 ↑	0,253 ↓	0,187 ↓	0,017 ↑	0,051 ↓	0,057 ↓	0,071 ↑	0,099 ↓	0,069 ↓

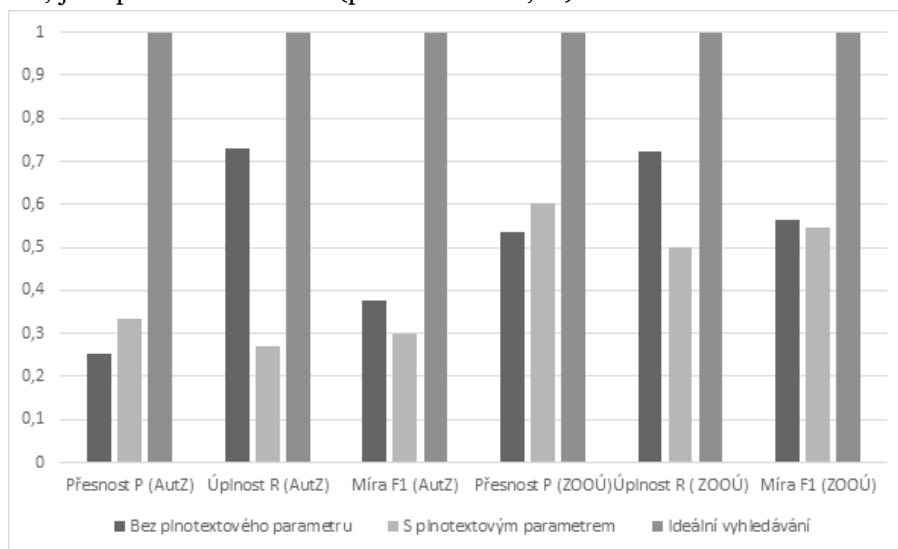
Tab. 6 Změna výsledků vyhledávání po zavedení plnotextového parametru proti vyhledávání bez plnotextového parametru

První řádek Tab. 5 nám (opět na příkladu výsledků ASPI) říká, že ze všech párů *U-R* nacházejících se v idealizované databázi, kde expertní anotátor označil rozhodnutí *R* jako relevantní pro pochopení ustanovení *U*, jsme vyhledáváním v ASPI po zařazení plnotextového parametru dosáhli přesnosti 33,3 % (zlepšení o 7,9 procentního bodu) a úplnosti 27,1 % (zhoršení o 45,8 procentních bodu). Vyhledávání v ASPI nám tedy poskytlo 27,1 % ze všech párů kognitivně relevantních pro expertního uživatele, které se vyskytují ve všech třech informačních systémech zařazených do tohoto textu. Přesnost 33,3 % pak ale znamená, že všechny tyto relevantní páry byly skryty v množství falešně pozitivních výsledků. Tedy výsledků, které expertní anotátor nepovažoval za relevantní, ale právní informační systém ASPI je zahrnul do seznamu výsledků.

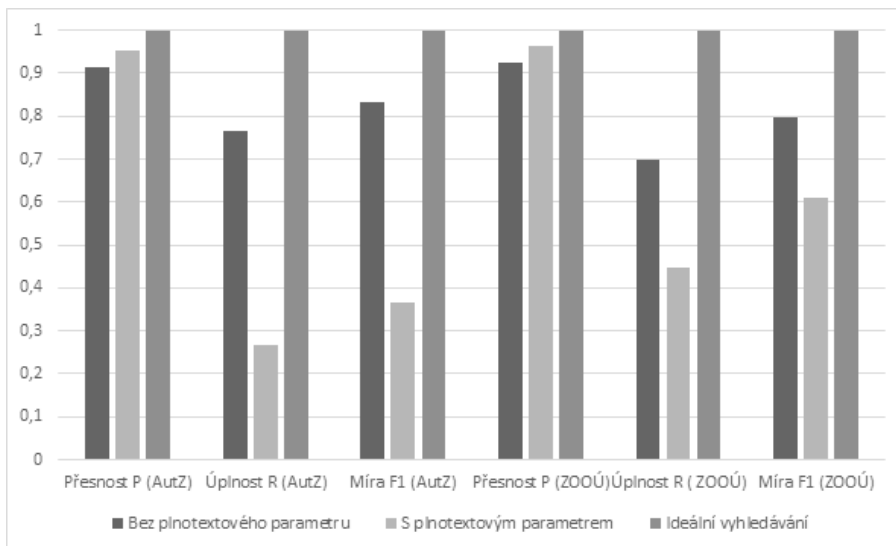
Poněkud zjednodušeně tedy: pokud si manuálně zpracuji a ohodnotím všechna rozhodnutí, která jsou v právních informačních systémech navázána k AutZ, vykazuje vyhledávání v ASPI po zařazení plnotextového parametru pokrytí jednoho z každé čtveřice relevantních rozhodnutí. Toto je pak doprovázeno tím, že z každých čtyř rozhodnutí indexovaných v právním informačním systému ASPI vyhodnotil expert jako relevantní jenom jedno. Část relevantních rozhodnutí tedy zůstala právním informačním systémem ASPI nezachycená (počet falešně negativních dokumentů). Zároveň se jako uživatel budu muset při vyhodnocování výsledků probrat poměrně velkým množstvím dokumentů, které mi pro mé expertní chápání konkrétního ustanovení nepřinesou žádné informace (počet falešně pozitivních dokumentů).

Stejně jako před zařazením plnotextového parametru je situace i v tuto chvíli odlišná pro neexpertního uživatele. Pro takového uživatele bude mít metadatové vyhledávání v právním informačním systému ASPI, jak je evidentní z Tab. 5 a Tab. 6, přesnost 95,4 % (zlepšení o 3,9 procentního bodu) a úplnost 26,6 % (zhoršení o 54 procentního bodu). Neexpertní uživatel tedy ani zdaleka nedostane všechny dokumenty (zhruba tři ze čtyř rozhodnutí označené anotátorem jako důležité mine). Na druhé straně mu ale více než devět z deseti rozhodnutí, které dostane v seznamu výsledků, přinese nové informace.

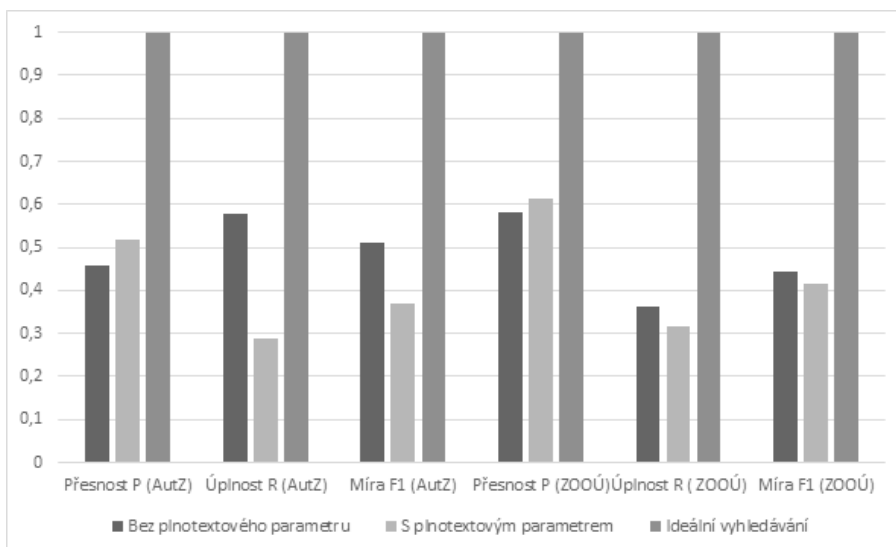
Srovnání jednotlivých metrik je tentokrát rozděleno na samostatné grafy pro jednotlivé právní informační systémy – Obr. 5 a 6 pro ASPI, Obr. 7 a 8 pro Beck-online a Obr. 9 a 10 pro Codexis. Toto zobrazení výsledků umožňuje jednoznačné srovnání úspěšnosti výsledků v jednotlivých systémech po zavedení plnotextového parametru. Dosažené hodnoty jsou opět srovnávány s ideálním vyhledáváním (míra $F1 = 1,00$), kdy dostáváme (i) všechna rozhodnutí, která nám pomohou pochopit nebo aplikovat konkrétní ustanovení (úplnost $R = 1,00$) a (ii) všechna rozhodnutí, která dostaneme, jsou pro nás relevantní (přesnost $P = 1,00$).



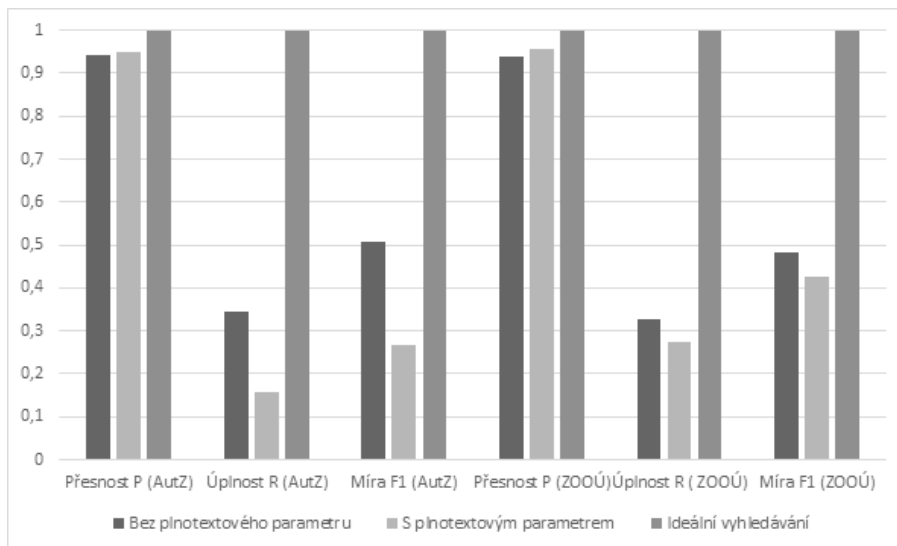
Obr. 5 Porovnání úspěšnosti vyhledávání v ASPI z pozice expertního uživatele před zadáním plnotextového kritéria a po jeho zadání



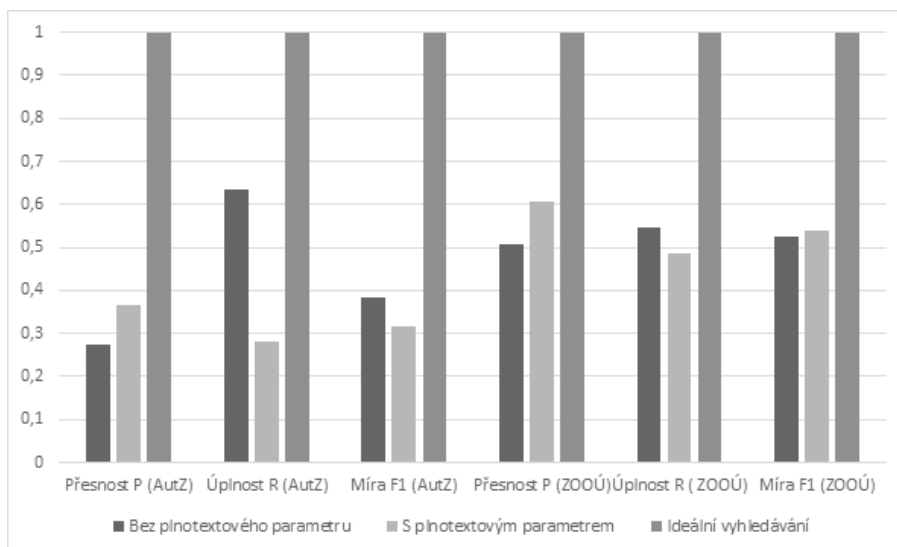
Obr. 6 Porovnání úspěšnosti vyhledávání v ASPI z pozice neexpertního uživatele před zadáním plnotextového kritéria a po jeho zadání



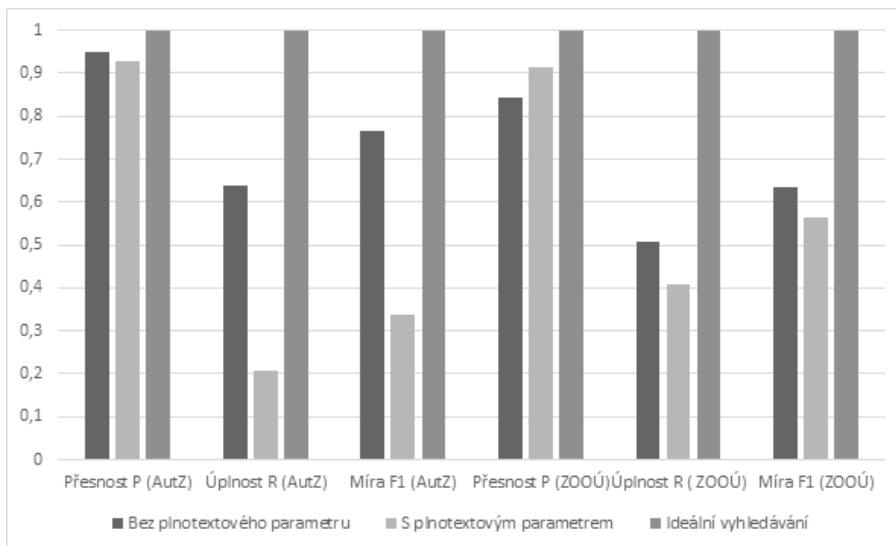
Obr. 7 Porovnání úspěšnosti vyhledávání v Beck-online z pozice expertního uživatele před zadáním plnotextového kritéria a po jeho zadání



Obr. 8 Porovnání úspěšnosti vyhledávání v Beck-online z pozice neexpertního uživatele před zadáním plnotextového kritéria a po jeho zadání



Obr. 9 Porovnání úspěšnosti vyhledávání v Codexis z pozice expertního uživatele před zadáním plnotextového kritéria a po jeho zadání



Obr. 10 Porovnání úspěšnosti vyhledávání v Codexis z pozice neexpertního uživatele před zadáním plnotextového kritéria a po jeho zadání

5.4 PODOBNOST MEZI SYSTÉMY PO ZAVEDENÍ PLNOTEXTOVÉHO PARAMETRU

Nyní je konečně nutné zjistit, zda zavedení plnotextového parametru vedlo k menším rozdílům mezi výsledky poskytnutými jednotlivými právními informačními systémy. Za tímto účelem je nutné zjistit variabilitu mezi hodnotami výsledků jednotlivých systémů při vyhledávání skrze zprostředkované informace (metadatová vazba mezi ustanovením a soudním rozhodnutím) a při vyhledávání skrze kombinaci zprostředkovaných a nezprostředkovaných informací (metadatová vazba mezi ustanovením a soudním rozhodnutím doprovázená plnotextovým parametrem). Jak je vidět z předcházejícího textu, zavedení plnotextového kritéria do vyhledání nevedlo ke zlepšení výsledků – zlepšila se přesnost pro většinu vyhledávání, ale došlo k výraznému snížení úplnosti. Směrodatné odchylky mezi hodnotami jsou v Tab. 7. Na prvním řádku je ve druhém sloupci směrodatná odchylka vypočtená z hodnot přesnosti výsledků poskytnutých právními informačními systémy ASPI, Beck-online a Codexis při vyhledávání expertním uživatelem před zavedením plnotextového parametru. Po zavedení plnotextového para-

metru je hodnota téhož uvedena ve druhém sloupci. Jak je vidět, je nižší – a toto platí pro všechna vyhledávání, ať už expertní či neexpertní a pro AutZ a ZOOÚ. Zavedení plnotextového kritéria jakožto kritéria, které umožňuje nezprostředkované vyhledávání (ve všech systémech je stejné), a jako doplňku ke zprostředkovanému vyhledávání (které je v každém systému jiné), vede ke statisticky podobnějším výsledkům. Ve všech případech je směrodatná odchylka výsledků vyhledávání v systémech bez plnotextového kritéria větší než s plnotextovým kritériem.

		Přesnost <i>P</i> (bez pln.)	Přesnost <i>P</i> (pln.)	Úplnost <i>R</i> (bez pln.)	Úplnost <i>R</i> (pln.)	Míra <i>F1</i> (bez pln.)	Míra <i>F1</i> (pln.)
AutZ	Expertní uživatel	0,092	0,080	0,062	0,008	0,062	0,031
	Neexpertní uživatel	0,014	0,012	0,176	0,030	0,141	0,042
ZOOÚ	Expertní uživatel	0,030	0,004	0,148	0,084	0,069	0,060
	Neexpertní uživatel	0,042	0,022	0,152	0,074	0,128	0,078

Tab. 7 Směrodatné odchylky výsledků dosažených při vyhledávání bez plnotextového parametru (bez pln.) a po zařazení plnotextového parametru (pln.)

6. DISKUZE A LIMITY

V této části se pokusím některé výsledky dále interpretovat a zároveň předstříit limity výzkumu prezentovaného v tomto textu. Těchto limitů je, jak bude za chvíli zřejmé, nemálo a jsou pro pochopení tohoto textu zcela zásadní.

Jakákoli interpretace výsledků je za prvé ovlivněna tím, že srovnání vyhledávání soudních rozhodnutí souvisejících pouze se dvěma předpisy (AutZ, ZOOÚ) poskytlo málo dat na zcela jasné a prokazatelné závěry. Jakkoli bylo manuální zpracování dat ze tří zařazených systémů časově náročné, stejně jako bylo náročné i jejich anotování, jednalo se spíše o práci mravenčí než o práci po stránce dat impozantní svým rozsahem. Tím je některé závěry nutné považovat automaticky za slabší, než by mohly být při zpracování výrazně většího datasetu. Podobně by mohlo být srovnání průkaznější při zpracování rozsáhlejšího předpisu, který má za sebou dlouhou historii (např. zákon č. 141/1961 Sb.) nebo s jehož praktickou aplikací jsou

spojeny těžkosti ohledně kontinuity judikatury (např. zákon č. 89/2012 Sb.).

Výsledky výrazně ovlivňuje i volba plnotextového dotazu, který byl zařazen jako druhý parametr do zprostředkovaného vyhledávání založeného na subjektivních metadatech. Plnotextový dotaz byl zvolen s přihlédnutím ke slově, která jsou typická pro celý předpis a nikoli pro konkrétní ustanovení. Skrze tento limit je nutné chápat výrazný pokles v úplnosti vyhledávání po zařazení plnotextového parametru. Na první pohled tedy zavedení plnotextového parametru vede k výrazně horším celkovým výsledkům při zcela marginálním zvýšení přesnosti. To je ale způsobeno spíše výběrem konkrétního plnotextového dotazu a výsledek je tak dle mého názoru spíše ilustrativní.

Posledním výrazným limitem tohoto textu je proces vyhodnocování kognitivní důležitosti jednotlivých soudních rozhodnutí z párů *U-R*. Využil jsem jednoho expertního anotátora pro AutZ, jednoho expertního anotátora pro ZOOÚ a jednoho neexpertního anotátora pro oba předpisy. Anotace tímto způsobem je velmi subjektivní. Zvýšení kvality při hodnocení kognitivní relevance jednotlivých rozhodnutí by bylo možné při zařazení většího množství anotátorů v rámci paralelních anotací (kognitivní relevance by pak byla hodnocena pro stejný pár *U-R* více než jedním anotátorem). Při čtení tohoto textu je nutné shora uvedené limity vést v patrnosti.

Co je tedy možné ze shora popsané evaluace vyvozovat? Zjednodušený model, který s kolegy používáme jako názornou pomůcku při výuce, je očividně poměrně přesný. Mezi výsledky poskytovanými jednotlivými právními informačními systémy v rámci dotazu využívajícího subjektivní metadata (rozhodnutí související s konkrétním ustanovením) jsou opravdu poměrně velké rozdíly. Množství párů *U-R*, které byly ze systémů ASPI, Beck-online a Codexis vytěženy, je odlišné. Z toho ale nelze automaticky usuzovat, že by systémy, nabízející více párů (typicky ASPI a Codexis) byly nutně lepší než ty, které jich nabízejí méně (typicky Beck-online). Zde je na místě jistá opatrnost při interpretaci. Mám za to, že tento rozdíl spíše ukazuje na odlišný způsob vytváření vazeb mezi ustanoveními a soudními rozhodnutími. Při vytváření těchto vazeb v systému Beck-online je nejspíše

postupováno konzervativněji, než v případě vytváření vazeb v systémech ASPI a Codexis.

Vyhodnocení výsledků (přesnosti P , úplnosti R a míry $F1$) po zařazení kognitivní relevance tomuto závěru nasvědčuje. Expertní anotátoři jak v případě AutZ, tak ZOOÚ vyhodnotili některá rozhodnutí jako relevantní pro jejich chápání konkrétního ustanovení a Beck-online v tomto hodnocení vykazuje vyšší hodnotu přesnosti, ale nižší úplnost. Výsledky poskytnuté systémy ASPI a Codexis pak dávají více možností, jak zachytit kognitivně relevantní dokumenty. Tyto rozdíly ale nejsou dle mého nijak zásadní. V případě obou odborníků se nicméně při hodnocení vyhledávání dostal dopředu Beck-online.

Neodborník, celkem pochopitelně, označil jako kognitivně relevantní velkou část všech párů $U-R$, protože jeho doménová znalost je nižší. Tímto se dostaly ve výsledcích dopředu právě systémy, u kterých je možné konstatovat tendenci označovat vazbou mezi ustanovením a rozhodnutím větší množství rozhodnutí (ASPI a Codexis). Pokud uživatel o dané problematice téměř nic neví, bude většina nalezených dokumentů kognitivně relevantní (tedy bude automaticky dosaženo vysoké přesnosti). Takový uživatel pak může těžit z toho, že při použití systému ASPI a Codexis dosáhne větší úplnosti svého vyhledávání.

Při komplexním pohledu na výsledky tak opatrně uzavírám, že při zpracování dat pro Beck-online dochází k vyznačení spíše menšího množství vazeb – hodnocení tématické relevance při tvorbě systému je tedy spíše konzervativnější. Důsledkem je, že je větší šance, že budou takto vyznačené vazby kognitivně relevantní pro expertního uživatele. Naopak v případě systémů ASPI a Codexis to vypadá, že jsou rozhodnutí provazována k jednotlivým ustanovením po překročení nižší hranice. To vede k většímu množství falešných pozitiv při vyhledávání, ale zároveň je oproti Beck-online menší šance, že „něco uteče“.

A konečně, zařazení plnotextového parametru do vyhledávání vede k výsledkům, které jsou si napříč systémy statisticky podobnější. Kombinace zprostředkovaného (subjektivní metadata) a nezprostředkovaného (data) vyhledávání sice v tomto konkrétním případě výrazně zhoršilo metriky, po-

mocí kterých jsem vyhledávání hodnotil, ale potvrdilo domněnku, že zařazení dat a objektivních metadat do vyhledávání povede k „objektivnějším“ (napříč systémy podobnějším) výsledkům. Výrazné zhoršení výsledků po zařazení plnotextového parametru není dle mého názoru příliš znepokojivé – hovoří spíše o nevhodné zvoleném parametru v rámci tohoto výzkumu než o negativním vlivu plnotextových parametrů na vyhledávání obecně.

7. ZÁVĚR

V odpovědi na první výzkumnou otázku je tedy nutné uvést, že rozdíly mezi výsledky poskytnutými jednotlivými právními informačními systémy v odpovědi na dotaz obsahující subjektivní metadata opravdu existují. A to poměrně výrazné. Srovnatelné dotazy vyhledávající soudní rozhodnutí související s jednotlivými ustanoveními vedly k identifikaci odlišných dokumentů navázaných na jednotlivá ustanovení AutZ a ZOOÚ. Průnik množin dokumentů provázaných v jednotlivých systémech dosahoval pouze cca 18 % z celkového počtu vazeb u AutZ a pouze cca 17 % z celkového počtu vazeb u ZOOÚ. Po zohlednění kognitivní relevance pro vyhledávání z pohledu expertního a neexpertního uživatele jsou rozdíly mezi systémy značné. Za velmi důležité považuji uvést, že nízký poměr párů *U-R*, které se vyskytují ve všech systémech, má velmi praktický dopad – v podstatě máme čísla, která podporují argument o vhodnosti paralelního využití více právních informačních systémů pro zpracování rešerše. Ideální vyhledávací strategií se jeví využití alespoň dvou informačních systémů. To je sice časově i finančně náročnější, ale jsme pak při vyhledávání schopni identifikovat více relevantních párů a tím dosáhnout větší úplnosti vyhledávání. Jakkoli je přesnost důležitou metrikou, obecně považuji za menší problém získat v seznamu výsledků dokument, který pro mě není relevantní a musím ho vyřadit, než minout relevantní dokument.

V odpovědi na druhou výzkumnou otázku uzavírám, že zavedení plnotextového parametru do vyhledávání obecně zvýšilo přesnost vyhledávání, ale snížilo jeho úplnost. Dosažení vyšší přesnosti a nižší úplnosti přímo vychází z logiky věci. Pokud do vyhledávání doplníme plnotextový parametr,

vyhledávání se pak vlastně skládá ze dvou parametrů. Může tedy dojít k vyřazení některých dokumentů, které byly v seznamu výsledků zařazeny po metadatovém vyhledávání, ale již nemůže dojít k zařazení těch, které první parametr nenaplnily. Logicky tedy dojde ke zvýšení přesnosti (pokud se nám podaří plnotextovým dotazem vyřadit některé původně falešně pozitivní dokumenty), ale ke snížení úplnosti (pokud plnotextovým dotazem vyřadíme dokumenty, které byly původně pravými pozitivy). Podstatné v tomto případě bylo, že téměř vždy došlo k drobnému navýšení přesnosti. Podobně došlo vždy k poklesu úplnosti, často výraznému. To naznačuje, že plnotextový parametr nebyl zvolen zcela optimálně, což je jeden z výrazných limitů tohoto textu.

V odpovědi na třetí výzkumnou otázku je pak nutné uzavřít, že zavedení plnotextového kritéria vedlo k dosažení statisticky podobnějších výsledků napříč jednotlivými informačními systémy. Vyhledávání prostřednictvím zprostředkované strategie (využití subjektivních metadat) přineslo po dodatečném zařazení parametrů z nezprostředkované strategie (využití dat) výsledky, které jsou si napříč systémy podobnější.

Tato oblast – v užším smyslu evaluace výsledků získaných vyhledáváním v právních informačních systémech, v širším smyslu pak kritický přístup k vyhledávání v právních informačních systémech a k volbě vyhledávacích strategií – má dle mého názoru poměrně hodně potenciálu pro další akademický i praktický zájem. Zejména je dle mého názoru nutné ubírat se třemi směry. Za prvé pokusit se o komparaci na rozsáhlejších datech. Jak jsem již uvedl v tomto textu, vhodným se jeví zařazení některého rozsáhlého kodexu, zejména pak zákona č. 89/2012 Sb. Za druhé považuji za zajímavý směr pokusit se určit, zda mnou popsané rozdíly ve vyhledávání mají (a případně jak velký) praktický dopad. Je možné na základě sady dokumentů identifikované v jednom informačním systému dojít k odlišnému závěru než při zohlednění dokumentů získaných vyhledáváním v jiném informačním systému? Zjištění, zda nám sady dokumentů získané z různých systémů přinášejí odlišné informace, by mělo velký praktický význam. Za třetí by pak bylo dle mého názoru zajímavé věnovat se volbě optimální vyhledávací strategie v návaznosti na použité informační systémy a časovou do-

taci. Vyčíslení nákladů na vyhledávání by, opět dle mého názoru, mohlo vést k optimalizaci některých procesů v praxi, ale i ve výuce.³⁶

V tomto textu jsem se pokusil představit kritický přístup k právním informačním systémům. Evaluace vyhledávání v jednotlivých systémech s přihlédnutím k využití různých vyhledávacích strategií, jakož i evaluace vyhledávání napříč jednotlivými systémy, by měla být běžnou součástí úvah o práci s právními informačními systémy. Tyto úvahy navíc nejsou pouze teoretické, ale mají velmi silné implikace v rovině zajištění efektivního vyhledávání právních informací. Kritický přístup k vyhledávání v právních informačních systémech může pomoci rozvoji informační gramotnosti právníků.

PODĚKOVÁNÍ

Vznik tohoto textu byl podpořen Grantovou agenturou ČR v rámci projektu GA17-20645S.

Rád bych poděkoval Tereze Novotné, Janu Zibnerovi a Jakubu Míškovi za asistenci se zpracováním dat. Tereza Novotná si zaslouží ještě druhé poděkování, protože její komentáře k první verzi tohoto textu mi jej umožnily výrazně vylepšit.

Dřívější verze tohoto textu a různé části popsaného evaluačního experimentu byly prezentovány na konferencích *Qualitative Research in Law Conference* (název příspěvku *Collecting Case Law for Qualitative Content Analysis: What is the Correct ‘Population‘?*) v říjnu 2018 v Brně, *Internationales Rechtsinformatik Symposium IRIS* (název příspěvku *Case Law Retrieval: Critical Evaluation of Czech Legal Information Retrieval Systems*) v únoru 2019 v Salzburku a na konferenci *COFOLA* (název příspěvku *Efektivnost vyhledávání v právních informačních systémech: Porovnání strategií využívajících zprostředkované a nezprostředkované údaje k vyhledání relevantních soudních rozhodnutí*) v dubnu 2019 v Telči. Měl jsem také skvělou možnost představit východiska jako na *Work in Progress* semináři na Ústavu práva a tech-

³⁶ Srov. WILENSKY, Beth. When Should We Teach Our Students to Pay Attention to the Costs of Legal Research? *Perspectives: Teaching Legal Research and Writing*. 2016, roč. 26, č.1-2, s. 41-46.

nologií PrF MU v říjnu 2018. Na všech těchto akcích se mi dostalo cenných kritických připomínek, za které velmi děkuji.

8. LITERATURA

- [1] ARREDONDO, Pablo. Shepard for a Day: A Novel Class Exercise for Teaching Citators. *Legal Reference Service Quarterly*, 2015, roč. 34, č. 3, s. 239-244.
- [2] ASHLEY, Kevin. *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*. Cambridge: Cambridge University Press, 2017.
- [3] BING, Jon. Legal Information Retrieval Systems: The Need for and the Design of Extremely Simple Retrieval Strategies. *Computer/Law Journal*, 1978, roč. 1, č. 1, s. 379-400.
- [4] BING, Jon. Performance of Legal Text Retrieval Systems: The Curse of Boole. *Law Library Journal*, 1987, roč. 79, č. 2, s. 187-202.
- [5] BLAIR, David a M. E. MARON. An Evaluation of Retrieval Effectiveness for a Full-Text Document-Retrieval System. *Communications of the ACM*, 1985, roč. 28, č. 3, s. 289-299.
- [6] BOBEK, Michal. Výzkum v právu: reklama na Nike anebo kvantová fyzika? *Jurisprudence*, 2016, roč. 25, č. 6, s. 3-10.
- [7] DABNEY, Daniel. The Curse of Thamus: An Analysis of Full-Text Legal Document Retrieval. *Law Library Journal*, 1986, roč. 78, č. 5, s. 5-40.
- [8] GERSON, Kevin. Evaluating Legal Information Retrieval Systems: How Do the Ranked-Retrieval Methods of WESTLAW and LEXIS Measure Up? *Legal Reference Services Quarterly*, 1999, roč. 17, č. 4, s. 53-67.
- [9] GRABMAIR, Matthias, Kevin ASHLEY, Ran CHEN, Preethi SURESHKUMAR, Chen WANG, Eric NYBERG a Vern WALKER. Introducing LUIMA: an Experiment in Legal Conceptual Retrieval of Vaccine Injury Decisions Using a UIMA Type System and TOOLS. *Proceedings of ICAIL 2015*. S. 69-78.
- [10] HAFNER, Carole. Conceptual Organization of Case Law Knowledge Bases. *Proceedings of ICAIL 1987*. S. 35-42.
- [11] HARAŠTA, Jakub. Nejednoznačnost odkazů k soudním rozhodnutím a možnosti řešení. *Revue pro právo a technologie*, roč. 6, č. 11, s. 15-28.
- [12] HELLYER, Paul. Evaluating Shepard's, KeyCite, and BCite for Case Validation Accuracy. *Law Library Journal*, 2018, roč. 110, č. 4, s. 449-476.
- [13] KNAPP, Melanie a Rob WILLEY. Comparison of Research Speed and Accuracy Using Westlaw Classic and WestlawNext. *Legal Reference Services Quarterly*, 2013, roč. 32, č. 1-2, s. 126-141.
- [14] KOSAŘ, David a Jan PETROV. Jak vybrat „případy“ do případové studie a pracovat s nimi v právu: poznatky z výzkumu na pomezí práva a politologie. *Jurisprudence*, 2016, roč. 25, č. 6, s. 21-30.

- [15] MART, Susan Nevelow. The Case for Curation: The Relevance of Digest and Citator Results in Westlaw and Lexis. *Legal Reference Services Quarterly*, 2013, roč. 32, č. 1-2, s. 13-53.
- [16] MAXWELL, Tamsin a Burkhard SCHAFER. Concept and Context in Legal Information Retrieval. *Proceedings of Jurix 2008*. S. 63-72.
- [17] SORMUNEN, Eero. Extensions to the STAIRS Study – Empirical Evidence for the Hypothesised Ineffectiveness of Boolean Queries in Large Full-Text Databases. *Information Retrieval*, 2001, roč. 4, č. 3-4, s. 257-273.
- [18] ŠAVELKA, Jaromír, Matěj MYŠKA, Adam PTAŠNIK a Danuše SPÁČILOVÁ. *Právní informační systémy*. Brno: Tribun EU, 2011.
- [19] TAYLOR, William L. Comparing KeyCite and Shepard's for Completeness, Currency, and Accuracy. *Law Library Journal*, 2000, roč. 92, č. 2, s. 127-142.
- [20] TURTLE, Howard. Natural Language vs. Boolean Query Evaluation: A Comparison of Retrieval Performance. *Proceedings of SIGIR 1994*. S. 212-220.
- [21] URBÁNIKOVÁ, Marína a Hubert SMEKAL. Právní věda a právní psaní: postačí vždy jako výzkumná metoda "číst, přemýšlet a psát"? *Jurisprudence*, 2017, roč. 26, č. 4, s. 37-41.
- [22] VAN OPIJNEN, Marc a Cristiana SANTOS. On the Concept of Relevance in Legal Information Retrieval. *Artificial Intelligence and Law*, 2017, roč. 25, č. 1, s. 65-87.
- [23] WILENSKY, Beth. When Should We Teach Our Students to Pay Attention to the Costs of Legal Research? *Perspectives: Teaching Legal Research and Writing*. 2016, roč. 26, č.1-2, s. 41-46.

Toto dílo lze užít v souladu s licenčními podmínkami Creative Commons BY-SA 4.0 International (<http://creativecommons.org/licenses/by-sa/4.0/legalcode>).
