



doc. Mgr. Barbora Vidová Hladká, Ph.D.

hladka@ufal.mff.cuni.cz
<http://ufal.mff.cuni.cz/bvh>

V Praze 31. prosince 2020

Posudek na knihu *Citační analýza judikatury*

Vážené kolegyně, vážení kolegové,

byla jsem požádána vypracovat posudek na právní monografii *Citační analýza judikatury* autorského kolektivu vedeného dr. Jakubem Haraštou. Velmi ráda jsem se tohoto úkolu zhostila, protože počítačové zpracování dokumentů z legislativní domény je jedním z mých výzkumných témat. Vzhledem k tomu a vzhledem k mé kariéře v oblasti počítačové lingvistiky se cítím kompetentní hodnotit zejména ty kapitoly, které pojednávají o empirické studii nad soudními rozhodnutími.

Hlavním cílem knihy je popis citační praxe v české judikatuře prostřednictvím kvantitativní analýzy citací rozhodnutí, které se vyskytují v rozhodnutích vrcholných soudů České republiky, tj. analýzy odkazů na dřívější (minulá) rozhodnutí v rozhodnutích Nejvyššího (NS), Nejvyššího správního (NSS) a Ústavního soudu (ÚS). Jedná se o unikátní studii nad datovou sadou 237 tisíc soudních rozhodnutí z období let 1993-2018, která přináší nový kvalitativní pohled na citování v české judikatuře.

Knihy obsahuje 256 stran včetně seznamu použité literatury a zdrojů a je rozdělena do tří částí, ve kterých se prolínají teoretické a empirické aspekty vybraného tématu. Hlavní i dílčí témata jsou doprovázena podrobnou rešerší relevantních prací z oblasti angloamerického a kontinentálního právního systému.

Část I představuje teoretické a metodologické základy citační analýzy judikatury. Diskutuje význam pojmu judikatura a formuluje hypotézy, které se týkají role judikatury a motivace soudců citovat dřívější soudní rozhodnutí. Podrobně se zastavuje u vágních pojmů, zejména *důležitost* a *relevance*, jejichž chápání je značně subjektivní, v čase se měnící, dochází u nich k významovému překryvu a ne vždy jsou/mohou být definovány. Jsou-li k dispozici data, je nutné je vhodně reprezentovat mj. i proto, aby bylo možné vágní pojmy definovat pomocí kvantitativních měr. Kapitola 6 představuje soudní rozhodnutí jako síť, analogickou např. k železniční síti: rozhodnutí jsou stanice a citace jsou koleje, jednosměrné i obousměrné. Důležitost rozhodnutí pak může být definována jako počet rozhodnutí, které citují příslušné rozhodnutí (tj. počet kolejí, které vedou do stanice). V terminologii teorie grafů se jedná o orientované grafy.

Získání rozhodnutí jednotlivých vrcholných soudů pro robustní počítačové zpracování není bohužel otázka „jednoho kliknutí“. Rozhodnutí jsou sice zpřístupňována na různých platformách, ale pouze k vyhledávání. Část II popisuje technické aspekty zpřístupnění soudních rozhodnutí ve formě korpusu *Czech Court Decisions Corpus* (CzCDC), který obsahuje 460 524 867 slov ve 237 723 rozhodnutích s nejvyšším počtem rozhodnutí NS, dále pak ÚS a nakonec NSS. Celkový počet rozhodnutí CzCDC je natolik vysoký, že není reálné provádět například označení citací ručně. Proto autoři používají k anotování (obohacování rozhodnutí) metody řízeného strojového učení, které vyžadují existenci tzv. zlatých dat pro trénování (= část dat s ruční anotací). Konkrétně řeší dvě anotační úlohy, a sice segmentaci soudních rozhodnutí do významových celků (hlavička, shrnutí aj.) a detekci a klasifikaci referencí (identifikátor odkazovaného soudního rozhodnutí, soud, datum aj). Úspěšnost obou úloh je vyhodnocována standardními metrikami. Podobně jako autoři chápu tyto experimenty jako pilotní a zveřejnění dat jako správný krok k dalšímu experimentování mimo autorský kolektiv.

Část III popisuje kvantitativní analýzu sítě soudních rozhodnutí a její expertní interpretaci. Citace v počtu 668 375 byly ze soudních rozhodnutí extrahovány automaticky pomocí modelu popsaného v části II. Ze sítě citací byly získány hodnoty měř, které popisují jednak bezprostřední okolí soudních rozhodnutí, tj. rozhodnutí, které citují příslušné rozhodnutí (stanice, do kterých se dostaneme bez přestupu), jednak vzdálenější okolí (stanice, do kterých se dostaneme s přestupy). Hodnoty zvolených měř jsou interpretovány pro i přes jednotlivé soudy a srovnávány v čase. Výklady naznačují dílčí trendy například v tendenci jednotlivých soudů citovat specifická rozhodnutí. Sice nenaznačují zásadní trendy v citační praxi českých vrcholných soudů, ale rozhodně nabízejí celou řadu podnětů k diskusi a k dalšímu vytěžování soudních rozhodnutí. Domnívám se, že velký přínosem by byla vizualizace sítě rozhodnutí, která by zohledňovala i hodnoty vybraných měř a uživatelům by umožňovala přímo zobrazení soudních rozhodnutí

Právní monografii *Citační analýza judikatury* hodnotím vysoce jako původní příspěvek do diskuse o citační praxi v české judikatuře, který se opírá o významnou datovou základnu a o expertní interpretaci kvantitativní analýzy této základny. Zpřístupnění základny v repozitáři LINDAT/CLARIAH-CZ (<http://lindat.cz>) je vynikajícím krokem k otevření soudních rozhodnutí výzkumníkům z dalších společenských i přírodovědných oborů.

Se srdečným pozdravem