

Výzkum citovanosti publikací na různých institucích podle počtu autorů

Bc. Petr Urban¹, Bc. Petr Štěpánek², Bc. Václav Hrabík³

1 Úvod

Spolupráce s ostatními institucemi je úzce spojena s popularitou vydaných publikací. Rowlands, Ian (2018) v rámci své prezentace ukázal, že publikace vytvořené ve spolupráci s autory z různých mezinárodních institucí jsou populárnější než publikace, na kterých pracuje pouze jeden autor v jedné instituci. Popularita je zde měřena počtem citací bez ohledu na samocitace. Dále Rowlands uvádí, že publikace publikované v mezinárodním prostředí mají vyšší impakt než publikace publikované v rámci jedné instituce nebo jednoho státu.

Tato práce se zaměřuje na ověření hypotézy, zda počet citací publikací skutečně roste s počtem autorů a počtem různých institucí, do kterých patří. Pro ověření této hypotézy byly využity vizualizační techniky v nástroji PowerBI a datové zdroje DBLP pro získání náhodně vybraných autorů z institucí Harvard a Oxford. Kromě toho byl využit datový zdroj OBD, který obsahuje jména pracovníků Katedry informatiky a výpočetní techniky (KIV) Fakulty aplikovaných věd (FAV) Západočeské univerzity v Plzni (ZČU), pro porovnání se zahraničními institucemi. Pro získání počtu citací daných publikací bylo využito API z aplikace Scopus.

2 Metody a výsledky

Pro ověření hypotézy oxfordské instituce byly použity tři vizualizační metody znázorněné na obrázku 1a. Metoda se spojnicovým grafem nedokázala prokázat souvislost mezi počtem autorů, institucí a citovaností článků. Metoda se sloupcovým grafem ukázala vyšší průměrnou citovanost při více autorech a institucích, ale vzácně byly publikace s jedním autorem z jedné instituce. To naznačuje mezinárodní spolupráci vědců oxfordské instituce, což může zvyšovat popularitu publikací, ale neověřuje hypotézu. Poslední metoda s grafovou vizualizací potvrdila nezávislost mezi počtem autorů, institucí a citovaností. Z těchto vizualizací vyplývá, že v oxfordské instituci neexistuje korelace mezi počtem autorů, institucí a citovaností článků.

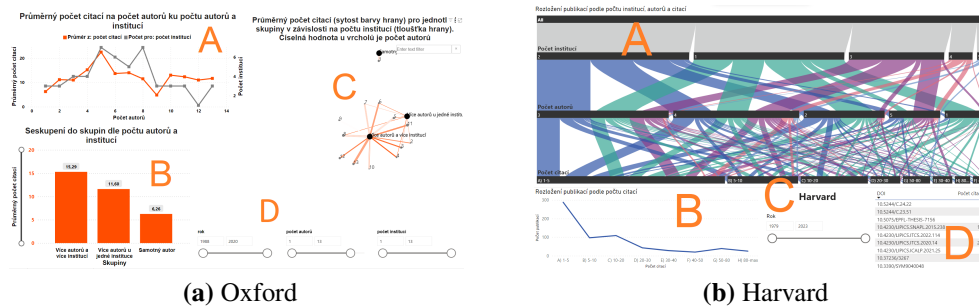
K prozkoumání publikací v instituci Harvard je využit především Sankeyův diagram s paralelními souřadnicemi znázorněný na obrázku 1b. Sankeyův diagram zobrazuje tok dat nebo informací mezi různými kategoriemi. Tok v tomto diagramu představuje počet publikací. Spojnicový graf pak ukazuje rozložení publikací mezi kategorie počtu citací a tvar křivky v něm se mění podle rozbalených kategorií v Sankeyově diagramu. Výsledky ukazují, že tvar křivky rozložení publikací v kategoriích počtu citací se téměř nemění při různých kombinacích autorů a institucí, ani při filtrování podle roku vydání. To naznačuje, že počet autorů a institucí nemá vliv na počet citací publikací.

Pro ověření hypotézy využívající data z Katedry informatiky a výpočetní techniky bylo

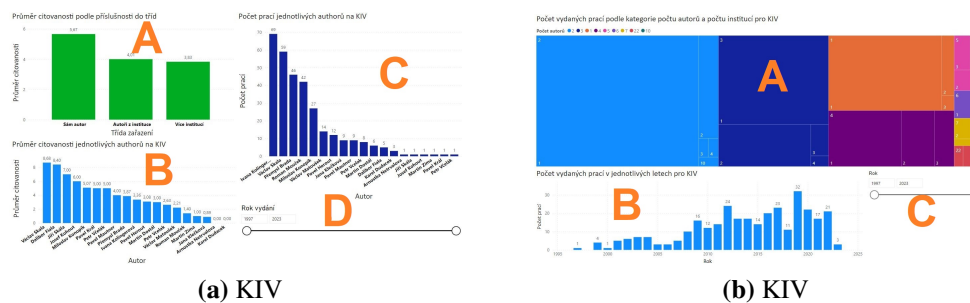
¹ student navazujícího studijního programu Softwarové a informační systémy, e-mail: urbanp@students.zcu.cz

² student navazujícího studijního programu Softwarové a informační systémy, e-mail: petrs1@students.zcu.cz

³ student navazujícího studijního programu Softwarové a informační systémy, e-mail: hrabikv@students.zcu.cz



Obrázek 1: Vizualizace pro zahraniční instituce



Obrázek 2: Vizualizace pro instituci KIV ZČU

vytvořeno pět různých vizualizací s interaktivními prvky, které jsou znázorněny na obrázku 2. Veškeré vizualizace zobrazují vztah mezi počtem autorů, institucí, publikací a citovaností za pomoci sloupcových grafů nebo mapy. Publikace jsou v zásadě rozděleny do tří tříd: publikace s jedním autorem, publikace s více autory z jedné instituce a publikace s více autory z různých institucí. Žádná z použitých vizualizací nepotvrdila hypotézu. Ukazuje se, že na KIV nejsou stejné podmínky jako v původním článku a v ostatních institucích, protože na KIV je velké množství autorů, kteří často publikují sami.

3 Závěr

Tato studie má několik omezení, která brání jednoznačnému potvrzení nebo vyvrácení hypotézy. Náhodný výběr jmen vědeckých pracovníků a omezení použité databáze DBLP na oblast počítačových věd ovlivňují reprezentativnost vzorku. Použití Scopusu pro zjištění počtu citací může být neaktuální a nemusí obsahovat veškerá relevantní data.

Z výsledků vizualizací vyplývá, že počet citací není přímo závislý na počtu autorů a institucí, ale spíše na konkrétním tématu publikace.

Publikace z Oxfordu a Harvardu se vyznačují větším počtem autorů z různých institucí, často mezinárodních. Naopak Katedra informatiky a výpočetní techniky na ZČU v Plzni projevuje pokles mezinárodní spolupráce a nárůst publikací s jedním autorem nebo s více autory z této instituce, což může mít negativní vliv na prestiž fakulty.

Budoucí výzkum by měl rozšířit datový soubor o další instituce a využít další zdroje. Dále by mohlo být zajímavé zkoumat jednotlivá témata publikací a jejich popularitu.

Literatura

Rowlands, Ian (2018) Benchmarking: How to Compare Research Outputs. *Prezentace v rámci školení European Academy for Taxes, Economics & Law v Berlíně*