

## Možné nástrahy kvantitativního hodnocení

Pavel Mika

Knihovna Akademie věd ČR

[mika@knav.cz](mailto:mika@knav.cz)

INFORUM 2016: 22. ročník konference o profesionálních informačních zdrojích  
Praha, 24. - 25. 5. 2015

### Abstrakt

*Kvantitativní metody hodnocení vědeckých výstupů představují na první pohled objektivní, na tvrdých datech založený systém využívaný na různých úrovních vědeckých institucí. Příspěvek nabídne kritický pohled na praktické použití těchto bibliometrických metod. Postupně budou v příspěvku představeny úskalí a omezení, která snižují vypovídající hodnotu a zvyšují riziko chybné interpretace bibliometrických indikátorů. K tomu budou prodiskutovány možnosti řešení a obecná doporučení vedoucí ke správnému použití bibliometrie.*

*Konkrétně bude popsána problematika oborového členění, jehož použití je klíčovým faktorem k objektivnímu nahlížení na citační indikátory. Ovšem jaké členění použít a jak to ovlivní výsledné hodnoty je poměrně obsáhlá otázka. Samotné bibliometrické indikátory zahrnují široký okruh okolností, které je nutné pro jejich správné využití brát v potaz (např. extenzivní – intenzivní indikátory, skreslení průměrem, minimální počet výstupů). Zajímavou otázkou je také započítávání vědeckých výstupů mezi více autorů, tedy to, jestli využít tzv. plné počítání nebo frakcionalizaci a jak to může skreslit interpretaci indikátorů.*

*Cílem příspěvku není znehodnotit využití bibliometrie pro účely hodnocení vědy, ale naopak navrhnout taková řešení, která budou robustní a odolná všem možným nástrahám.*

---

### Sofistikované citační indikátory

Citovanost odráží míru vlivu dané vědecké práce a to obvykle pozitivního vlivu, ale v ojedinělých případech i negativního. Samotný počet citací nám může napovědět, jaká je míra vlivu dané práce, ale abychom mohli publikace různého zaměření navzájem porovnávat, je nutné využít sofistikovanějších citačních indikátorů. Tato nutnost je dána především odlišnými citačními zvyklostmi v rámci vědních oborů. Účelem tohoto příspěvku není představit širokou škálu a rozdíly v citačních indikátorech, proto budou v textu využity především tyto dva sofistikované indikátory: oborově normalizovaný citační vlivu (Mean Field Normalized Citation Impact – MFNCI, v praxi jsou používány i jiné názvy) a podílu na nejcitovanějších (obvykle 10% či 25%) pracích v oboru (proportion of top 10% publications – PP10%).

MFNCI je vypočítán jako podíl počtu citací publikace a průměrného počtu citací všech publikací ze stejného oboru, stejné roku vydání a stejného typu dokumentu (článek, review). V případě více oborů článku se vypočte podíl za každý obor a bere se jejich průměr. Pro určení hodnoty tohoto indikátoru u skupiny více publikací, obvykle představujících vědeckou produkci sledované jednotky, se bere průměrná hodnota za jednotlivé publikace. Tento indikátor je jednoduše interpretovatelný, jelikož

hodnota 1 představuje světový průměr (za předpokladu, že svět je vše co je obsaženo v databázi) a např. hodnota 1,45 znamená, že daná skupina prací překračuje průměrnou citovanost o 45 procent.

Oproti této výhodě je indikátor snadno ovlivnitelný jediným vysoce citovaným článkem a tak může být zkreslena jeho výpovědní hodnota. Vzhledem k tomu, že indikátor pracuje právě s průměrem, měl by být založen na dostatečném množství poměřovaných publikací.

PP10% (25%) vyjadřuje podíl prací, které se řadí mezi nejcitovanější práce v oboru za daný rok a typ dokumentu. Tento indikátor je vhodný ke sledování excelence a nedochází u něj ke zkreslení při zařazení jednoho vysoce citovaného článku do souboru méně citovaných.

### Oborová klasifikace

Oba tyto indikátory jsou založeny na porovnání vědeckých výstupů v rámci jejich vědního oboru, čímž se dostáváme k problematice samotného rozřazení do vědních oborů. Pro účely evidence a hodnocení vědeckých výstupů se používají různé oborové klasifikace, které mohou být vytvořeny národními orgány pro účely institucionálního hodnocení, mezinárodními organizacemi, komerčními subjekty jakými jsou například producenti citačních databází či výzkumnými institucemi zabývajícími se kvantitativními metrikami.

U oborové klasifikace můžeme rozlišit míru jejího členění (takzvanou granularitu), tedy to na z jak malých oborů respektive podoborů se klasifikace skládá. Z toho plyne i aktuálnost klasifikace a její schopnost postihnout nové vědní obory. Dalším hlediskem je způsob jejího použití, tedy to jestli jsou do oborů zařazovány časopisy společně s celým jejich obsahem nebo jestli je použita na úrovni jednotlivých článků. Porovnání čtyř oborových klasifikací je vidět v Tabulce 1.

Název klasifikace/zdrojové databáze	Počet oborů	Úroveň zařazení
IS VaVal	125	Články - ručně
WoS	252	Časopisy
Scopus	307	Časopisy
Leiden Ranking	3822	Články - algoritmicky

Tabulka 1.: Porovnání vybraných oborových klasifikací

Výběr oborové klasifikace může mít zásadní vliv na výpovědní hodnotu citačních indikátorů a to především z důvodu, že by mělo být dosaženo toho, aby publikace přiřazené do daného oboru byly opravdu vzájemně porovnatelné. Jedná se jednak o problém příliš širokých oborů, kde je možné vysledovat více skupin s rozdílnými průměrnými počty citací, které jsou si blízké. Obvykle se jedná o rozdělení samotného oboru na teoretickou a praktickou část.

Při výpočtu výše zmíněných oborově normalizovaných indikátorů se obvykle pracuje s klasifikací dané citační databáze -Web of Science (WoS) nebo Scopus, kde může docházet ke zkreslení výsledků ze dvou důvodů: publikace jsou přiřazovány do oboru na základě časopisu (může se tedy stát, že se týkají daného oboru pouze okrajově), časopis může být zařazen do více oborů (články z méně citovaného oboru budou pozitivně ovlivněny a naopak). Z těchto dvou důvodů může právě nastat situace, kdy publikace není porovnávána vůči publikacím s podobným citačními zvyklostmi.

Tento problém řeší například klasifikační schéma výzkumníků z Leidenské univerzity, které je uměle vytvořené na základě analýzy několika stovek milionů referencí mezi pracemi obsaženými v databázi WoS. Vzniklé kategorie představují klastry vymezené klíčovými slovy. Při vzniku této klasifikace se snažili najít optimální velikost klastrů, přičemž pro účely světového žebříčku univerzit Leiden Ranking byla použita klasifikace obsahující 3822 mikro oborů.

Kromě výběru oborové klasifikace je nutné být obezřetný k analýzám prací zařazených v oborech s nízkou průměrnou citovaností. V oborech, kde stačí několik málo (1-5) citací k zařazení práce mezi PP25%. Tento problém je možné řešit nastavením minimálních hodnot počtů citací pro málo citované obory (tento postup byl například použit při přípravě bibliometrických podkladů v rámci Hodnocení výzkumné a odborné činnosti pracovišť AV ČR za léta 2010-2014). Další možností, která je využita v již zmiňovaném Leiden Rankingu, je rozdělení publikací respektive časopisů na takzvané „core“. Časopisy tvořící toto jádro musejí splňovat dvě základní podmínky: časopis musí mít mezinárodní záběr (na základě autorů článků a původu citací) a musí mít dostatečné množství referencí na ostatní core časopisy, což právě dokazuje to, že se nachází v oboru, který je vhodný pro citační analýzu.

### **Ovlivňující aspekty**

Mezi dalšími aspekty, které mohou ovlivnit výsledky při použití citačních indikátorů, rozhodně patří zahrnutí pouze vybraných druhů dokumentů – v potaz by měli být brány pouze dokumenty typu article, review a případně proceeding paper. Zde jsme odkázáni na správnost indexačního procesu dané databáze, kdy dochází k určení druhu dokumentu. Je také nutné podotknout, že citační databáze WoS a Scopus mají odlišné politiky indexace a určení druhů dokumentu a publikace pak může být hodnocena rozdílně na základě dat z jedné či druhé databáze.

Také problematika frakcionalizace neboli podílové započítávání prací autorům, vědeckým institucím či zemím v případě spoluautorství patří mezi důležité aspekty. Oproti takzvanému plnému započítávání, kdy je každá publikace započtena jednou každému autorovi (organizaci, zemi), při frakcionalizaci započteme sledované jednotce pouze podíl na vědecké práci (obvykle rovnoměrně rozdělený, případně lze váhově bonifikovat prvního autora). Frakcionalizace může být provedena na několika úrovních: autor, adresa, organizace, země. V praxi bývá nejčastěji použita právě na úrovni autora, což je případ i současné Metodiky hodnocení <sup>1</sup>.

V případě porovnávání na základě počtu vědeckých výstupů není použití frakcionalizace ničím podmíněno. Naopak u normalizovaných citačních indikátorů, konkrétně u oborově normalizovaného citačního vlivu, je pro správnou interpretaci toto použití záhodné, jak ukazuje nedávný článek Ludo Waltmana <sup>2</sup>. Zkreslení, ke kterému dochází v případě plného započítávání, je způsobeno několikanásobným započítáváním hodně citovaných článků. Pozitivně ovlivněny jsou tak obory, kde je možné sledovat korelaci mezi mírou citovanosti a počtem autorů.

Zapojení do velkých výzkumných kolaborací může také zkreslit výsledky citační analýzy a to především kvůli velkému počtu spoluautorů a výrazně vyšším průměrným počtem citací. Na druhou

---

<sup>1</sup> <http://vyzkum.cz/FrontClanek.aspx?idsekce=754005>

<sup>2</sup> Field-normalized citation impact indicators and the choice of an appropriate counting method. Journal of Informetrics. Journal of Informetrics. 2015, 9(4), 872-894. DOI: doi:10.1016/j.joi.2015.08.001.

stranu není možné autory takových publikací penalizovat vyřazením těchto prací. Řešením je tedy sledování podílu prací s velkým počtem autorů na všech pracích autora/pracoviště/země a při vyšší míře tohoto podílu brát výsledky citačních indikátorů s větší obezřetností.

V roce 2014 bylo v České republice vydáváno 51 časopisů indexovaných v databázi Journal Citation Report (neboli takzvaných impaktovaných časopisů – mající Impact factor). V rámci publikační analýzy je možné sledovat podíl prací ať už v domácích nebo přímo v časopisech vydávaných institucí, ve které autor působí. V případě vysokého podílu takovýchto prací bychom měli být znovu obezřetní při interpretaci výsledků.

### **„Metodický“ materiál**

Potřeba standardů pro použití bibliometrie v rámci hodnocení vědy vyústila v iniciativu Leiden Manifesto a stejnojmenný článek, který vyšel v časopise Nature<sup>3</sup>. V loňském roce také vyšel materiál The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management<sup>4</sup>, který mimo jiné čerpá z britského hodnocení Research Excellence Framework (REF).

Jednotlivá závěrečná doporučení jsou podpořena konceptem „Odpovědných metrik“, který v několika bodech poukazuje na adekvátní použití kvantitativních indikátorů, přičemž zde můžeme sledovat velkou míru shody s deseti doporučeními, které vycházejí z Leiden Manifesta. Těmito body jsou: robustnost, pokora, transparentnost, rozmanitost a reflexivita.

Zavádění těchto obecných doporučení do praxe společně s konkrétními opatřeními popsanými v příspěvku mohou vést k důvěryhodnému použití kvantitativního hodnocení vědy.

---

<sup>3</sup> HICKS, Diana, Paul WOUTERS, Ludo WALTMAN, Sarah de RIJCKE a Ismael RAFOLS. Bibliometrics: The Leiden Manifesto for research metrics. Nature. 2015, roč. 520, č. 7548, s. 429-431 [cit. 2015-04-28]. DOI: doi:10.1038/520429a. Dostupné z: <http://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>

<sup>4</sup> Wilsdon, J., et al. (2015). The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. DOI: 10.13140/RG.2.1.4929.1363