

Recenze tématické studie

Hodnocení inforatického výzkumu, autor: Radim Bělohlávek

vypracované v rámci projektu SoSIReČR – „Sociální síť inforaticků v regionech České republiky, CZ.1.07/2.4.00/12.0039, OP Vzdělávání pro konkurenceschopnost“

zpracoval: Martin Štěpnička

Tématická studie se zabývá různými variantami hodnocení výzkumu v oblasti inforaticky jak z obecného hlediska, tak i demonstracemi na konkrétních případech. Velice pozitivním aspektem studie je, že se nesnaží pouze diskutovat celkovou koncepci, parametrické nastavení a dopady oficiální Metodiky hodnocení výsledků výzkumu a vývoje v roce 2009 dle RVVI, ale snaží se nejprve přistoupit k analýze specifčnosti inforatického výzkumu a k hledání odpovědi na otázku, co jsou prestižní výsledky v inforaticce. Zaměření studie na tzv. tvrdé výsledky nelze považovat za omezující, protože obecná studie na téma hodnocení včetně měkkých výsledků (neoponovaných a nerecenzovaných) by byla nad rámec rozsahu této studie.

Studie navíc adresně pojmenovává problémy dvou základních přístupů k hodnocení výzkumu bez ohledu na oborová specifika – na přístup založený na hodnocení pomocí odborných komisí a přístup scientometrický. U obou jsou jmenovány nezpochybnitelná negativa, a sice subjektivita a možná zaujatost v případě prvním a ztráta schopnosti relevantního měření výstupů v případě druhém, což je způsobeno snahou se plně přizpůsobit parametrům scientometrického měření na úkor opravdové kvality (Goodhartův zákon). Pasáž textu dokládá pokřivení scientometrického přístupu přímou finanční motivací např. na zvyšování IF lze ovšem analogicky použít i na jiné výstupy – uvedme např. finanční motivaci organizátorů konferencí (nárůst počtu účastníků) k zařazení sborníku do sledovaných databází. Tyto databáze pak nenesou žádnou kvalitativní informaci a jedná se čistě o kategorizační zařazení pro účely vyhledávání citací. Některé závěry studie mají tedy obecnou platnost i za hranicemi inforatického výzkumu.

Studie obsahuje srovnávací příklady: 1. srovnání podprůměrného českého inforaticka s třemi špičkovými vědci; 2. srovnání výše uvedených tří špičkových vědců s českým fyzikem. Příklady sice používají zřejmě ne příliš typické zástupce (český inforatic generuje 58 konf. příspěvků za tři roky, český fyzik 17 výsledků J_{imp} za stejné období), nicméně jedná se o existující osoby, tudíž metodika neselhává čistě v hypotetických příkladech, ale i v realitě. Podotkněme, že dopady druhého případu budou zřejmě částečně eliminovány novou metodikou rozdělovací bodové hodnoty oborově. Bohužel objemy bodů v jednotlivých oborech nebyly výsledkem sofistikované studie ani oprávněným strategickým rozhodnutím volené garnitury, ale vychází z objemů bodů generovaných jednotlivými obory v letech minulých, čili opět na základě vadné metodiky. První příklad ovšem zdá se upozorňuje na více závažný problém nekonzistence ve vnitrooborovém hodnocení. Zde se odráží i příliš simplifikující agregování výsledků obyčejným sčítáním:

Znamená opravdu 58 výsledků jednoho autora 58 krát více, než výsledek jediný? Potažmo, lze agregací mnoha neprestižních výsledků docílit vyššího hodnocení, než pomocí výsledků prestižních v nižším počtu?

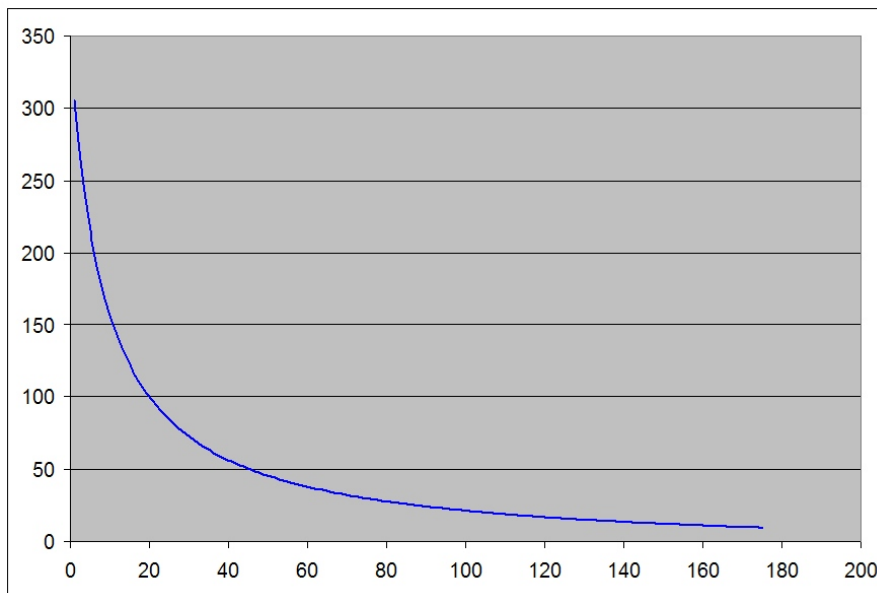
Na obě otázky se nabízí negativní odpověď. Nikdy nebylo oficiálně uvedeno, že účelem hodnocení není kvalita, ale kvantita!

Autor uvádí jako dva základní kameny úrazu v zavedeném hodnocení jednak nerozlišování mezi prestižními konferencemi sloužícími jako zásadní publikační fórum a konferencemi běžnými, a jednak využití pouze databáze Web of Science, která nejenže nemá dostatečný záběr, ale současně její vnitřní oborové dělení na subject categories je velmi hrubé, nepřesné a v mnoha ohledech stojí na

srovnávání nesrovnatelného. Nad těmito závěry nelze v žádném případě polemizovat, jedná se o nezpochybnitelná fakta opět důsledně doložená příklady. Opět bych si dovolil doplnit, že tyto závěry mají obecnou platnost i mimo informatiku. Pro doložení tohoto poznatku vezměme v úvahu např. časopisy vhodné pro publikace v oboru analýzy časových řad. Možná nejprestižnějším je International Journal of Forecasting s IF=1,685. Kdyby byl řazen do Statistics & Probability, byl by hodnocen jako sedmáctý z 93. Bohužel spadá do jiných kategorií, a sice do Economics a do Management, kde je až na 36. pozici z 89. Pro autory publikující v oblasti časových řad tak vlastně neexistuje možnost získat plný počet bodů, chtějí-li své výsledky uplatňovat především ve své komunitě, kam neoddiskutovatelně patří. Jedná se o přímý důsledek špatného zařazení do kategorií (umístění do kategorie Management) firmou Thomson Reuters a také nedostatečné jemnosti dělby do kategorií, ani zařazení do Statistics & Probability by totiž problém neřešilo.

Ještě vážnější jsou dopady v oblasti logiky, která je úzce spjatá s informatickým výzkumem. Prestižní časopisy tohoto oboru dosahují IF kolem 0.5 (Journal of Symbolic Logic IF=0,439), což je vzhledem k častému zařazení do kategorie Mathematics řadí do nízkých pozic (JSL je na pozici 159 z 219). Nejkrivější případem je pak nezařazení časopisu Studia Logica do JCR, a to se bez diskusí jedná o jeden z nejváženějších časopisů v logice vůbec. Vědečtí pracovníci jsou pak před volbou publikování v prestižních časopisech, sbírání bodů dle oficiálních metodik, nebo neetického chování (fragmentace výsledků, opakované publikování, MPI atd.). Zde je negativní vliv Goodhartova zákona naprosto zásadní.

Opět se tedy dostáváme do diskuse, zda není vhodnější nahradit scientometrický přístup hodnocením odborníky. Sám autor zmiňuje, že vedle vhodnosti hodnocení je třeba mít na zřeteli i proveditelnost hodnocení. Dostatečný počet „nezávislých“ (často zahraničních) komisí na všechny pracovní týmy v ČR je zřejmě nad rámec možností. Pokud by navíc daný přístup adoptovaly i ostatní státy, jedná se čistě o iluzi. Zde se tedy opět dá jen souhlasit jak s autorem, tak i s obecně známými názory prof. Tomáše Opatrného, člena RVVI. Krátce se jen pozastavme nad zmínkou, o schopnosti scientometrických přístupů oddělit zrna od plev. Obecně je tvrzení pravdivé. Je otázkou, na kolik toho je schopna schválená Metodika 2009.



Obrázek 1: Funkce přiřazující body na základě pořadí v kategorii dle IF.

Při pohledu na funkci přiřazující body od 10 do 305 (Obrázek 1) dle pořadí časopisu v kategorii na základě IF je zřejmé, že s plevou je odděleno i mnoho zrn. Zvláště vezmeme-li v úvahu výše uvedené problémy s kategorizací časopisů do oborů, vede prudce klesající konvexní funkce na nízké obodování i mnoha prestižních časopisů a efekt je podobný efektu ignorace prestižních selektivních konferencí

v informatice. Mediánový časopis je hodnocen cca 25body, což činí v intervalu <10,305> neuvěřitelný pokles o 91,5% z možného maxima. Lineární pokles by vedl na snížení o 50% na 157,5 bodu. Naopak konkávní funkce s mírným poklesem v první části a prudkým poklesem ve svém závěru by vedla na vyrovnání průměrných časopisů se špičkovými a nevýznamné hodnoty jen u nevýznamných periodik – strategie, kdy je odděleno zrno od nejhrubějších plev, ale nemalá část plev zůstává mezi zrnem.

Všech šest závěrů studie je tedy dle názoru recenzenta dostatečně podloženo a plně se s nimi ztotožňuje. K námětům na pokračování si recenzent dovoluje doporučit přidat i pokračování v úvahách nad oběma základními přístupy hodnocení (scientometrie vs. komise) s cílem zachovat jejich výhody a eliminovat jejich nevýhody. Jednou z možných aplikací těchto úvah je využití nezávislých odborných komisí na tvorbu oborově úzce specializovaných hodnotících kritérií scientometrické podstaty. Komise by tak nehodnotily přímo pracoviště, čímž by se znemožnila zaujatost a neobjektivita a zároveň by se snížila náročnost (počet i velmi úzce specializovaných oborů je jednoznačně nižší než počet výzkumných týmů – zvláště v celosvětovém měřítku). Naopak by práce takových komisí umožnila nepodléhat obecně nevhodným databázím (Web of Science, Scopus, CPCI či jiné), pracovat s prestižními konferencemi v oblasti informatiky stejně jako s prestižními časopisy, vysoce hodnotit prestižní časopisy i přes jejich nízký IF nebo dokonce přes jejich nezařazení do JCR. Komisionální hodnocení, které by stále mělo mít své místo v kariérních (habilitačních a profesorských) řízeních, akreditačních řízeních a projektovém financování VaV by tak své prvky odrazilo i v nastavení scientometrického přístupu. Ten by se naopak posléze mohl stát opěrným bodem ve výše uvedených hodnoceních odbornými komisemi. Nezanedbatelnou výhodou by měla být vyšší stabilita, která se při dnešních prudkých změnách IF mezi jednotlivými roky stává absentujícím aspektem v hodnocení VaV – v informatice obzvlášť.

- [1] R. Bělohlávek, Hodnocení infromatického výzkumu, studie v rámci projektu SoSIREČR CZ.1.07/2.4.00/12.0039, <http://www.cs.cas.cz/hsi/clanky/belohlavek.pdf>
- [2] P. Mikulecký, Infromatický výzkum v regionech – přednosti a úskalí, studie v rámci projektu SoSIREČR CZ.1.07/2.4.00/12.0039 <http://www.cs.cas.cz/hsi/clanky/mikulecky.pdf>
- [3] T. Opatrný, blog iHned, <http://blog.ihned.cz/?p=YOpatr>
- [4] Metodika hodnocení výsledků výzkumu a vývoje v roce 2009, Úřad vlády ČR, č.j. 08724/09-RVV, http://www.vyzkum.cz/storage/att/CDDC542199F1640B59A7D1E841B7151C/Metodika%202009_schváleno.pdf