

Úvodní komentář

Hodnocení kvality a výkonnosti základního výzkumu založené na publikačních výstupech a jejich citačním ohlasu je v oblasti věd o živé i neživé přírodě běžně používáno na úrovni jednotlivců, vědeckých týmů i celých výzkumných organizací (VO). Základem tohoto bibliometrického přístupu je předpoklad, že výstupem vědecké činnosti (zejména v základním výzkumu) je odborná publikace a ta může (ale nutně nemusí) být podkladem pro další výzkum a případně i praktické aplikace pro lidskou společnost. Dokladem toho je pak ohlas ve formě citací v odborných textech, patentech apod.

Jako podklad pro bibliometrické hodnocení se využívá řada tzv. indikátorů (např. IF nebo AIS časopisu, počet citací, typ citací aj.) a další parametry měřící ohlas výsledků vědecké činnosti. Pro zjednodušení se běžně hodnoty indikátorů rozdělují do skupin podle kvality (obvykle kvartilů a decilů).

Studie ukazují, že výsledky bibliometrického hodnocení se dobře shodují s hodnocení expertů, tzv. „peer review“. Přihlédneme-li k tomu, že podklady pro bibliometrické hodnocení jsou snadno (např. i automatizovaně) dostupné a vlastní hodnocení pak nezávislé na názoru jednotlivce a také levné (ve srovnání s „peer review“), je bibliometrie cenným nástrojem. Má však i své nedostatky.

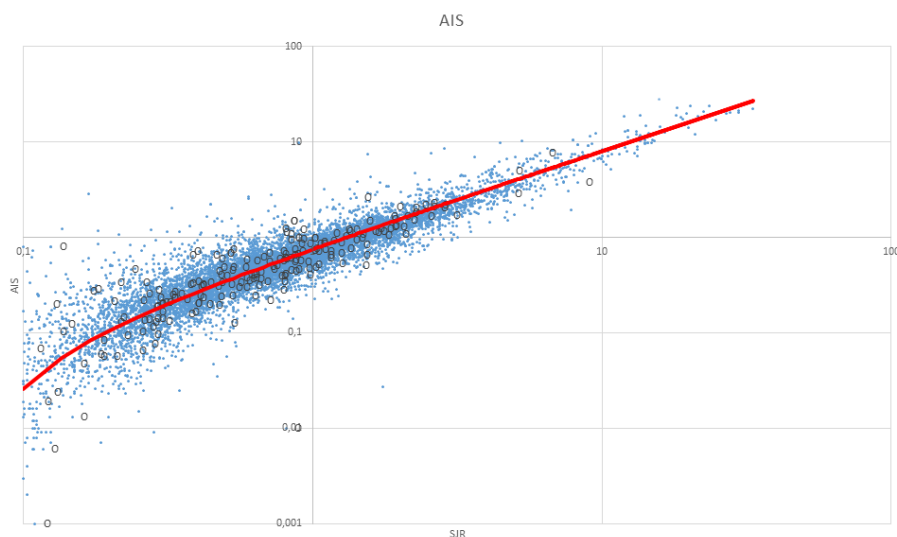
Nejspolehlivějším parametrem hodnocení jsou bezesporu citace. Každý citující vědec si práci prošel a našel v ní něco zajímavého (počty tzv. negativních citací jsou obvykle nízké). Pokud má tedy nějaká publikace vysoký počet citací, je pravděpodobně kvalitní a přínosná, přinejmenším z hlediska společenského užitku a aktuálních trendů v oboru. Počty citací však mají nevýhodu v tom, že se mění v čase. Je tedy nutno vždy uvádět datum, ke kterému byly stanoveny. U nedávných publikací budeme těžko tento indikátor používat, protože počty citací budou nízké a tudíž objektivně nehodnotitelné. Citační a publikační zvyklosti v jednotlivých oborech se navíc výrazně liší, což komplikuje porovnávání mezi jednotlivými disciplínami a jejich dílčími částmi (příkladem mohou být „Computer sciences“, ve kterých konferenční sdělení hraje mnohdy významnější roli než časopisecká publikace).

I proto se k hodnocení používají „náhradní indikátory“ a to jsou parametry časopisů, ve kterých byla práce publikována. IF (Impact Factor), odvozený index AIS (Article Influence Score) popř. také parametr SJR (Scientific Journal Ranking) více či méně odrážejí kvalitu časopisu a umožňují roztřídit časopisy do různých kategorií tak, jak tomu je v podkladech Úřadu vlády, jež byly poskytnuty pro hodnocení. AIS je zřejmě méně závislý na oborech (ve srovnání s IF) a není zatížen vlivem autocitací. Jako takový je pro hodnocení vhodnější (ve srovnání s IF). Při porovnání parametrů AIS pro jednotlivé podobory však vychází najevo, že tento parametr běžně nadhodnocuje časopisy specializované na v současnosti moderní (módní) tematiku a bohužel také referátové časopisy, které původní výsledky vůbec nepublikují.

Indikátor SJR používaný v databázi Scopus je vhodný při posuzování většího spektra časopisů (v databázi WOS je něco přes 11 tis. časopisů, zatímco v databázi Scopus jich je přes 20 tis.). Z toho plyne, že např. v horním decilu dle AIS bude mít daná VO méně výsledků než při hodnocení dle SJR. SJR je tak ze statistického pohledu robustnější ale zároveň „měkčím“ (rozuměj méně náročným) kritériem, neboť časopisy, které z pohledu SJR spadají do horního decilu, často bývají z hlediska AIS hodnoceny hůře (typicky v horním kvartilu). Jinými slovy: WOS obsahuje sice méně časopisů, ale kvalitnějších, a je tudíž vhodnější pro posuzování vyšší kvality, zatímco Scopus umožní přiřadit určitou míru kvality i méně kvalitním výsledkům (a bohužel i výsledkům publikovaným v časopisech s horší publikační praxí). Současné používání obou přístupů je podle názoru panelu nevhodné, neboť obě zmíněné databáze člení výsledky jiným způsobem a přiřazení výsledků do jednotlivých hodnocených oblastí není snadno

a hodnověrně přenositelné. Hodnocení podle obou kritérií (tedy AIS i SJR) se jeví nadbytečné i s ohledem na těsnou korelaci mezi těmito parametry (viz obrázek č. 1).

Obr. č. 1. Vztah mezi parametry AIS a SJR pro cca 10 tis. časopisů; černě jsou zobrazeny matematické časopisy.



Pádnou námitkou proti indikátorům časopisů může být skutečnost, že pro jednotlivé publikace jejich hodnoty nic neznamenají, tj. publikace ve vysoce hodnoceném časopise může mít po 10 letech 0 citací a naopak práce z časopisu s nízkým impaktním faktorem za stejnou dobu mnoho citací. To je však pravda jen částečně. Ve skutečnosti se vliv renomovaného časopisu na publikované práce projevuje v pravděpodobnosti citování (což je dílem dáno dobrým recenzním řízením, které horší práce nepropustí). Jinými slovy, nevíme, jak dopadne konkrétní publikace, ale pro statisticky robustnější soubor publikací (velikost souboru závisí na oboru) si můžeme být jisti, že součet citací bude blízký očekávání. Proto je hodnota IF nebo AIS pro soubor prací (týmu, VO i země) indikátorem, se kterým lze dále pro v jeho výpovědní hodnotu a robustnost zodpovědně pracovat.

Při porovnávání distribuce publikačních výstupů v jednotlivých kvartilech a decilech podle zvoleného parametru je vždy nutné všimnout si statistické významnosti posuzovaného souboru údajů (u malých souborů může hrát roli statistická chyba), brát v úvahu různou frekvenci publikování v jednotlivých oborech, které nevyhnutelně vyplývají mj. z náročnosti provedení experimentu, formulování a ověření hypotézy a managementu přístrojových nároků. Z těchto důvodů nelze srovnávat mezi sebou instituce podle počtu publikací v horním decilu (kvartilu), jestliže se liší oborově (např. mezi „buněčnou biologii“ a „fyzikální chemii“ je rozdíl přibližně dvojnásobný). Stejně tak nelze stanovit jednotný počet publikací, které by se měly odevzdat jako „nejlepší“ za určitou dobu jako určitý násobek počtu pracovníků týmu. Koeficient zde musí být proměnný, závislý na oboru.

Současná metodika hodnocení bohužel pomíjí zásadní parametr, kterým je status korespondujícího či hlavního autora. Ačkoli tato kategorie nehraje ve všech oborech zahrnutých do „Natural sciences“

klíčovou roli (např. v matematice), bývá často významným indikátorem místa, kde klíčový výsledek či myšlenka vznikly. U publikací uvádějících více afiliací na národní a zvláště na mezinárodní úrovni by proto tento aspekt měl být sledován, aby byl kredit za kvalitní publikační výstup dán jeho hlavnímu původci či původcům. V rámci hodnocení by měl tuto skutečnost vysvětlit vedoucí celého týmu (nikoliv nezávislý expert, protože takový se těžko najde).

V neposlední řadě panel nedoporučuje identifikovat nejkvalitnější výzkumná pracoviště na základě publikací v malé skupině vybraných časopisů (konkrétně *Nature* a *Science*), které adekvátně nereflektují stav výzkumu v ČR. Jako vhodnější k hodnocení se jeví publikace v prvním decilu nebo kvartilu s nezbytným přihlédnutím na velikost hodnocené jednotky (FTE) a oborovým specifikům.

Odborné hodnocení

Srovnání rozložení publikačních výstupů vzniklých na národní úrovni a ve světě ukazuje, že u většiny oborů hodnocených v rámci panelu Přírodních věd je podíl prací autorů z ČR spadajících do prestižního prvního decilu nižší než odpovídá celosvětovému průměru. Čestnou výjimkou jsou „Computer and information sciences“ a „Other natural sciences“, kde je tento podíl nepatrně vyšší. V případech, kdy je frakce publikací v prvním decilu nižší oproti mezinárodnímu průměru, pak nacházíme značné rozdíly mezi obory. Největší „propad“ publikací v horním decilu lze nalézt v matematických a pak i chemických oborech. V ostatních hodnocených oborech jsou rozdíly marginální.

Z porovnání rozdělení publikačních výstupů mezi kvartily podle AIS, které je z důvodů uvedených výše pro hodnocení vhodnější, je zjevné, že publikace spadající do oborů „Mathematics“ a „Earth and related environmental sciences“ v podstatě kopírují mezinárodní parametry v nadprůměrném prvním a druhém kvartilu. Vyššího zastoupení prací v těchto kategoriích dosahují „Computer and information sciences“ a v menším rozsahu také „Other natural sciences“. Podíl publikací za chemické a biologické obory v prvním kvartilu je o málo nižší ve srovnání s mezinárodním průměrem, ale podíl prací v druhém kvartilu se od mezinárodního zastoupení neliší. Podstatné je, že podíl publikací v nejnižším čtvrtém kvartilu je pro většinu oborů (s výjimkou biologických oborů) nižší nebo nanejvýš stejný jako podíl těchto prací v mezinárodním měřítku. S tím koresponduje srovnání národních a mezinárodních oborových mediánů, které jsou v podstatě shodné pro „Mathematics“, „Earth and related environmental sciences“ a „Other natural sciences“. Výrazně vyššího mediánu dosahují „Physical sciences“, zatímco národní mediány „Chemical sciences“ and „Biological sciences“ jsou velmi mírně resp. výrazně nižší.

Významné navýšení podílu publikací v prvním kvartilu oboru „Physical sciences“ oproti mezinárodnímu průměru, které představuje nejvýznamnější odchylku mezi národními a mezinárodními výsledky v tomto hodnocení, lze vysvětlit specifikou oboru. Značná část „českých“ publikací je totiž vytvořena ve velkých mezinárodních kolektivech čítajících až několik tisíc autorů (konkrétně jde o 310 publikací z celkového počtu 789 prací spadajících do prvního kvartilu podle indexu AIS!). Jak už jsme psali výše, tyto publikace mají určité zvláštnosti – a) jsou většinou zveřejněny ve kvalitních časopisech, b) jako spoluautoři se často uvádějí všichni členové konsorcia, c) organizace, které se účastní těchto konsorcií, mohou mít zásadní podíl svých publikací tohoto typu, d) příspěvek jednotlivých autorů k originalitě výsledku je problematicky stanovitelný. Z těchto důvodů nelze rozdíl v hodnotě mediánu o 40% hodnotit jako vybočující z obecného trendu, že úroveň české vědy je velmi dobře srovnatelná s mezinárodní úrovní. Pokud například srovnáme českou výkonnost ve fyzice s okolními zeměmi, zjistíme totiž, že podle rozdělení výsledků podle AIS jsme srovnatelní s Maďarskem nebo Slovinskem (absolutní počty publikací se ale liší).

V širší perspektivě lze konstatovat, že rozdíly mezi ČR a mezinárodní úrovní jsou téměř ve všech případech malé (do 10-15%) se zmíněnou výjimkou fyziky. Tento závěr (lichotivý pro chronicky podfinancovanou českou vědu) však má určitý nedostatek, který tkví v tom, že mezinárodní úroveň není úrovní rozvinutých zemí. Počty publikací na hlavu rostou přibližně úměrně HDP dané země (s některými výjimkami) a počty kvalitních publikací (např. s IF nad 20 s korespondujícím autorem z dané země) dokonce rostou téměř kvadraticky s HDP. I rozdíl v počtu publikací na obyvatele se liší o několik řádů (mezi Indonésií – cca 60 prací za 5 let na 1 mil. obyvatel až po 20 000 prací za 5 let na 1 mil. obyvatel ve Švýcarsku). Ještě výraznější rozdíly jsou v počtu vysoce kvalitních publikací (např. v *Nature* či *Science* s korespondujícím autorem dané země). Pokud uvážíme publikace v *Nature* nebo *Science* uvedené v analýze Úřadu vlády, nalezneme jen 6 prací s korespondujícím autorem z ČR! Za stejné období má Švýcarsko (při nižším počtu obyvatel) 104 publikací v těchto špičkových časopisech s domácím korespondujícím autorem. Ze zemí jako je Bulharsko, Rumunsko nebo Indonésie (přes velký

počet obyvatel) jsou to jednotky a korespondující autoři jsou z jiných zemí. Mezinárodní úroveň je tedy velmi špatně definovaná veličina, protože zahrnuje země všech možných úrovní rozvoje, historického vývoje a způsobu podpory výzkumu a inovací. Proto je daleko vhodnější vztahovat českou úroveň k vybranému souboru srovnatelných a/nebo pokročilých zemí.

Závěrem lze říci, že hodnocení na základě bibliometrických analýz má ve skupině oborů „Přírodní vědy“ nepochybně velký význam. Mělo by však doznat určitých zlepšení tak, aby mohla být použita v roce 2020 pro první plnohodnotné hodnocení vědy v ČR. Konkrétní návrhy jsou následující.

- 1) Hodnocení různých vědních oblastí musí brát ohled na oborové zvyklosti, jakými jsou například různá publikační frekvence, preferované typy výstupů, počty citací atp. Co je možná důležitější, obdobný přístup by měl být aplikován i v rámci podoborů, aby byly spravedlivě hodnoceny i tzv. malé obory. Bibliometrické hodnocení musí být prováděno pouze porovnáním v rámci jednotlivých oborů, protože nelze vytvořit univerzální metodologii, která by byla aplikovatelná na všechny oblasti a umožňovala srovnání napříč vědními obory. Při hodnocení by měla být vždy uvažována velikost hodnoceného celku a to jak s ohledem na možné statistické ovlivnění (malé celky), tak i na výzkumnou kapacitu (FTE), která hodnocené výsledky vytvořila. Ve specifických případech, by měl být v úvahu brány i další funkce institucí, např. pro vysoké školy jejich pedagogické působení.
- 2) Bibliometrické hodnocení by i nadále mělo být založeno výhradně na dostupných informacích (databázích) a přístupech, které umožní automatizovanou přípravu robustních a objektivních podkladů. Jakékoli další zatěžování hodnocených subjektů (formou přípravy materiálů pro hodnocení) nepovažujeme za vhodné, neboť ve výsledku vede k plýtvání časem i finančními prostředky.
- 3) Hodnocení na základě parametru AIS (databáze Web of Science) je obecně akceptovatelné a vhodnější pro obory sdružené do skupiny „Natural sciences“. Použití parametru SJR (z databáze SCOPUS) v průběžném hodnocení nedoporučujeme a to jak na úrovni oborů, tak i jejich skupin. Důvodem je podobnost obou přístupů (tedy ve výsledku nadbytečnost), různá kvalita hodnocení (přístup založený na SJR je měkčí) a zejména obtížná srovnatelnost dílčích údajů plynoucích z odlišného oborového třídění v obou databázích.
- 4) Do hodnocení je důležité zahrnout více indikátorů kvality publikací, tedy nejen kvartil a decil dle AIS časopisu, ale například i analýzu počtu a kvality citací (aby se předešlo nadhodnocení u citačních spolků) a – kde je to možné – také váhu publikačního výstupu s ohledem na afilaci hlavního/korespondujícího autora. Poslední aspekt lze snadno řešit rozdělením publikací do kategorií: vytvořeno výhradně danou VO, vytvořeno ve spolupráci v rámci ČR, vytvořeno v rámci mezinárodní spolupráce a vytvořeno ve velkém konsorciu. Hodnoceno by mělo být širší časové období (navrhujeme 5 let), které zvýší statistickou výpovědní hodnotu dat omezením možných fluktuací.
- 5) Hodnocení vysoce kvalitních publikací (*Nature* a *Science*) vzhledem k jejich malému počtu a oborové nevyváženosti (v roce hodnocení) nedoporučujeme. V prvním kroku navrhujeme bibliometrické hodnocení publikací v prvním kvartilu a decilu s ohledem na velikost instituce a původnost výsledků v hodnocené entitě (autorství). Nejlepší výsledky týmů nebo VO je vhodné zaslat expertům v 1. fázi hodnocení (počet těchto výsledků musí být určen podle velikosti hodnocené organizace v závislosti na oborech).
- 6) Na bibliometrické hodnocení publikačních výstupů musí navazovat posouzení kolektivním orgánem (komisí), pro který je však potřebná i jiná forma bibliometrie, jež umožní analyzovat výsledek prvotního hodnocení. Při bibliometrickém hodnocení VO je vhodné omezit se na

nejdůležitější indikátory, jakými jsou kvalita prací dle časopisu, národního a mezinárodního srovnání dle časopisů, dle počtu citací pro pětiletý interval a také, zda jsou kvalitní výsledky skutečně produkovány v dané instituci, nebo zda za nimi stojí spolupráce.