

O kryteriach bibliometrycznych w ewaluacji – głos zza oceanu

W różnych środowiskach naukowych, m.in. na łamach „PAUzy Akademickiej”, toczy się ostatnio dyskusja na temat zasadności stosowania kryteriów bibliometrycznych, takich jak liczba publikacji w czołowych czasopismach w danej dziedzinie, liczba cytowań, h-indeks, itp., przy ewaluacji indywidualnych naukowców lub placówek naukowych. Liczne wypowiedzi, na ogół krytyczne, dotyczą bądź jakości poszczególnych mierników liczbowych, bądź w ogóle sensu stosowania takich kryteriów przy ocenie dorobku naukowego. Ta druga, bardziej radykalna grupa krytyków podaje argument, że tylko wnikliwa merytoryczna ocena dorobku badawczego przez ekspertów, czyli *peer review*, może być podstawą rzetelnej ewaluacji badacza lub placówki naukowej, a żadne wskaźniki liczbowe takiej oceny zastąpić nie mogą. Mój głos polemiczny dotyczy właśnie tej drugiej, bardziej radykalnej opinii i ogranicza się do problemu oceny indywidualnych badaczy. Jest to głos przedstawiciela nauk ścisłych, matematyka i informatyka, który ostatnie 30 lat spędził w kanadyjskim środowisku akademickim. Ta perspektywa z zewnątrz w stosunku do nauki polskiej pozwoli może rzucić nowe światło na dyskutowany problem.

Na początku zaznaczę, że trudno nie zgodzić się z wyżej zacytowanym argumentem krytycznym. W idealnej sytuacji byłoby oczywiście najlepiej dokonywać wszelkich ocen naukowych wyłącznie na podstawie lektury prac ocenianego badacza, dysponując głęboką wiedzą z zakresu tematyki tych prac. Niestety, jak postaram się wykazać, tego typu sytuacje ewaluacyjne nie zawsze są możliwe i wtedy właśnie kryteria bibliometryczne stanowią istotną pomoc.

Jako stary już naukowiec (w Kanadzie powiedziałyby się elegancko *very senior*) miałem wielokrotnie okazję być członkiem różnych gremiów, takich jak komitet European Research Council przyznający granty, czy jego kanadyjski odpowiednik National Science and Engineering Research Council, bądź uniwersyteckich komitetów przyznających promocje profesorskie lub nagrody naukowe. Takie gremia mają wspólną cechę: muszą dokonać oceny porównawczej badacza z różnych dziedzin lub przynajmniej równych poddziedzin, w sytuacji gdy żaden członek gremium nie jest ekspertem we wszystkich dziedzinach czy poddziedzinach reprezentowanych przez kandydatów. Oto komitet informacyjny ERC musi rozstrzygnąć, czy wielomilionowy grant należy przyznać znakomitemu aplikantowi z obliczeń rozproszonych czy świetnemu specjalistce z dziedziny sztucznej inteligencji. Oto ogólnouniwersytecki komitet musi zdecydować, czy doroczną (jedyną na cały uniwersytet) *Prix d'excellence en recherche* (czyli dosłownie nagrodę za znakomitość badawczą) powinien dostać wybitny informatyk czy wybitna psycholożka. Żaden z członków komitetu nie zna się jednocześnie na obliczeniach rozproszonych i na sztucznej inteligencji, a tym bardziej żaden z nich nie jest jednocześnie informatykiem i psychologiem. Oczywiście komitet dysponuje wnikliwymi opiniami tzw. recenzentów zewnętrznych, dokonany przez ekspertów w dziedzinie danego kandydata. Ci eksperci nie mają jednak niestety możliwości porównania różnych kandydatów. Taką możliwość mają tylko członkowie komitetu. Co powinni zrobić?

Własne doświadczenie i obserwacja zachowań kolegów pouczają, że członkowie komitetu stosują w takich sytuacjach na ogół zasadę, którą nazwałbym „zaufanie pośrednie”. Otóż próbują np. dowiedzieć się od kolegów z nieznanego sobie dziedziny (najlepiej spoza komitetu), jakie pisma w tej dziedzinie uważa się za czołowe. (Eksperci, jeśli nie są osobiście zainteresowani wynikiem danej decyzji i nie powoduje nimi solidarność dyscyplinarna, są na ogół zadziwiająco zgodni przy typowaniu takich pism). Jeśli okaże się, że kandydatka do nagrody, psycholożka ma 6 prac w czołowych pismach z psychologii osobowości, a jej konkurent informatyk ma tylko jedną pracę w liczącym się piśmie i nadrabia ten brak długą listą artykułów w pismach, o których wiem, że są drugorzędne, to (wbrew solidarności dyscyplinarnej) zagłosuję na psycholożkę, rozumując tak: coś musi być w jej pracach (których nie rozumiem), skoro przebiła się z nimi do tych renomowanych pism. Kryteria związane z liczbą cytowań nie bardzo nadają się do porównywania badaczy z różnych dziedzin, bo np. przeciętna liczba cytowań w historii i w biologii jest zupełnie inna. Ponadto w naukach humanistycznych, w których dla celów polemicznych cytuje się czasem artykuły zawierające pomysły nieudane, duża liczba cytowań może być niekiedy, choć pewnie rzadko, wskaźnikiem dwuznacznym. Natomiast wewnątrz tej samej dziedziny, w wypadku nauk ścisłych, gdzie kiepskich prac się raczej nie cytuje (nawet po to, by je skrytykować), tylko się je po prostu ignoruje, takie kryteria mogą być użyteczne. Otóż jeśli oceniam podania o grant dwóch informatyków o podobnym stażu pracy i kandydat A ma 5000 cytowań i h-indeks 35, a kandydat B ma 1500 cytowań i h-indeks 20, to prawdopodobieństwo, że A jest lepszym naukowcem niż B, jest ogromne; nie stuprocentowe, ale ogromne. Wysoki h-indeks kandydata A eliminuje np. anomalie, gdzie mierny badacz „wystukał” połowę swoich cytowań dzięki dwóm bardzo popularnym w dziedzinie artykułom przeglądowym, niemającym wartości innowacyjnej. Musiałbym mieć bardzo poważne przesłanki świadczące na korzyść kandydata B (w informatyce taką przesłanką mogłaby być np. Best Paper Award na poważnej konferencji), żeby mimo tych różnic w cytowaniach głosował za przyznaniem grantu kandydatowi B, a nie A.

W obu rozważanych przykładach zasadę zaufania pośredniego można wyrazić następująco:

Ponieważ w danej sytuacji decyzyjnej nie jestem w stanie dokonać merytorycznej analizy dorobku kandydatów, używam przy podjęciu decyzji pewnych wskaźników liczbowych, o których wiem, że w podobnych sytuacjach na ogół wiernie oddają różnice poziomów dorobku naukowego.

Tylko tyle i aż tyle.

Reasumując, nie potępiamy w czambuł kryteriów bibliometrycznych przy ocenie dorobku naukowego. Stanowią one nieraz użyteczne narzędzie, o ile są używane obok innych kryteriów i o ile, jak każdego narzędzia, używamy ich ostrożnie i z rozważą.

ANDRZEJ PELC

Kierownik Katedry Badawczej Obliczeń Rozproszonych
Université du Québec en Outaouais, Canada