

Лейденски манифест за научни метрики¹

Използвайте тези десет принципа, за да управлявате оценката на научните изследвания, съветват Диана Хикс, Пол Уотърс и техни колеги

Превод д-р **Биляна Яврукова** - Софийски университет „Св. Климент Охридски”

Все повече данни се използват за управление на науката. Оценката на научните изследвания, която някога е правена по поръчка и извършвана от колеги, сега е рутинна и разчита на метрики [1]. Проблемът е, че оценката днес се извършва на базата на данни, а не на реална преценка. Метриците стават все повече: обикновено създадени с добри намерения, не винаги добре обосновани, често зле прилагани. Рисуваме да разрушим системата с въвеждането на оценка на научните изследвания с голямо количество инструменти, създадени за нейното усъвършенстване, без институциите да познават достатъчно приложението и тълкуването им.

Преди 2000 г. Science Citation Index се разпространява на CD-ROM от Института за научна информация (Institute for Scientific Information – ISI) и е използвана от експерти за специализирани анализи. През 2002 г. Thomson Reuters създава интегрирана уеб платформа, правейки широко достъпна базата от данни Web of Science. Създават се конкурентни цитатни указатели: Scopus на Elsevier (пуснат през 2004 г.) и Google Scholar (бета версия пусната през 2004 г.). Появяват се уеб базирани инструменти с цел лесно сравнение на научната дейност и влиянието на институциите, като InCites (използвайки Web of Science) и SciVal (използвайки Scopus), както и софтуер за анализ на индивидуалните цитатни профили, използвайки Google Scholar (Publish or Perish, пуснат през 2007 г.).

През 2005 г. Хорхе Хирш, физик в Калифорнийския университет, Сан Диего, създава h-индекса, популяризирайки отчитане на цитиранията на отделни учени. Интересът към Journal Impact Factor нараства значително след 1995 г. (виж „Impact-factor obsession” [„манията импакт-фактор”]).

В последно време все по-популярни стават метриците, свързани със социалните мрежи и онлайн коментарите – през 2002 г. е създаден F1000Prime, през 2008 г. Mendeley, а през 2011 г. Altmetric.com (подкрепен от Macmillan Science and Education, който е собственик на Nature Publishing Group).

Като изследователи на науката, учени в областта на обществените науки и администратори на научни изследвания, ние наблюдавахме с нарастваща тревога повсеместното неправилно приложение на индикаторите за оценка на научните дейности. Следват малка част от многобройните примери. В световен мащаб университетите са обсебени от тяхната позиция в международните класации (като Shanghai Ranking и списъка на Times Higher Education), въпреки че подобни списъци са базирани на това, което сме от наша гледна точка, неточни данни и произволни индикатори.

Някои работодатели изискват от кандидатите за работа стойности на h-индекс. Някои университети вземат своите решения, свързани с кариерно израстване на сътрудниците си, на базата на техните h-индекси и броя на статиите, публикувани в списания с „висок импакт фактор“. Учени, най-вече в областта на биомедицината, подчертават тези стойности в своите автобиографии. Навсякъде научните ръководители

¹ Преводът е направен в изпълнение на проект „Проектиране и разработване на прототип на информационна система „Указател за цитируемост на публикации от български автори (обществени науки)“ [Договор № ДН 15/1 от 11.12.2017], финансиран от Фонд „Научни изследвания” към Министерство на образованието и науката

изискват от докторантите да публикуват в списания с висок импакт фактор и изискват външно финансиране.

В скандинавските страни и Китай някои университети разпределят заплащането на учените или бонусите на базата на цифри: например, чрез пресмятане на лични импакт фактори за разпределяне на „ресурси за научна дейност“ или чрез изплащане на бонуси на учени за публикация в списание с импакт фактор по-висок от 15 [2].

В много случаи учените и оценяващите правят балансираны оценки. Все пак злоупотребата с наукометрични показатели става все по-разпространена, за да бъде пренебрегвана.

По тази причина ние представяме Лейденския манифест, носещ името на конференцията, на която е разработен (виж <http://sti2014.cwts.nl>). Той се състои от десет принципа, които не са новост за изследователите на науката, въпреки че никой от нас не би могъл да ги изложи в тяхната пълнота, тъй като до този момент няма кодификация. Светилата в тази област, като Юджийн Гарфилд (основател на ISI), също са посочвали някои от тези принципи [3, 4]. Но на тях не им се обръща внимание, когато оценителите ги представят на администрациите на университетите, които не са експерти в съответната методология. Учените търсят литература, с която да оспорят оценката, но тя е разпръсната в неизвестни и недостъпни списания.

Ние предлагаме същността на най-добрата практика в оценката на научните изследвания, базирана на метрики, така че учените да могат да проверяват как са оценявани, а оценителите да могат да проверят своите индикатори.

Десет принципа

1. Количествената оценка трябва да се допълва от качествена експертна оценка. Количествените метрики могат да повлияят на тенденциите при рецензирането и да улеснят обсъждането. Те трябва да подкрепят рецензирането, защото е трудно без достатъчно информация да се оценява колега. Въпреки това оценителите не трябва да се изкушават да вземат решение на основата на числа. Индикаторите не могат да заместят информираното оценяване. Всеки трябва да носи отговорност за своята оценка.

2. Съпоставяйте научната дейност с изследователските задачи на институцията, групата или учения. Целите на програмата трябва да бъдат описани в началото и използваните индикатори за оценка на научната дейност трябва да съответстват ясно на тези цели. Изборът на индикаторите и начините, по които те се използват, трябва да отчитат широко социално-икономическия и културния контекст. Учените имат различни изследователски цели. Научни изследвания, които са в границите на академичното знание се различават от научните изследвания, които се фокусират върху намирането на решения на обществени проблеми. Оценката може да се основава на академичните идеи за научни постижения, но и да се вземат под внимание заслугите за политическите решения, индустрията или обществото. Не всеки модел за оценка отговаря на всеки контекст.

3. Подкрепяйте научните постижения в научни изследвания, свързани с региона. В много части на света качеството на научните постижения се измерва на базата на англоезични публикации. Например, испанският закон затвърждава препоръката испанските учени да публикуват в списания с висок импакт фактор. Импакт факторът се изчислява за списания, индексирани в американската и все още основно англоезична база от данни Web of Science. Тези тенденции се явяват основен проблем в обществените и хуманитарните науки, при които научните изследвания са свързани основно с регионални и национални въпроси. Много други научни области имат национално или регионално значение – напр. Епидемията от СПИН в Субсахарска Африка.

Този плурализъм и обществена значимост започват да се negliжират, за да се създават публикации, които представляват интерес за пазителите на високия импакт фактор – англоезичните списания. Испанските социолози, които имат висока цитируемост в Web of Science са работили с абстрактни модели или са използвали данни на САЩ. Загубена е характерната особеност на работата на социолозите във високо значими испаноезични публикации: теми като местно трудово право, семейно здравеопазване за възрастните или трудова заетост на имигрантите [5]. Метриците, прилагани за висококачествена неанглоезична литература, трябва да идентифицират и оценяват научните постижения на изследвания, значими за съответния регион.

4. Поддържайте процеса на събиране на данни и анализ открит, прозрачен и прост. Структурата на базите от данни, използвани за оценка, трябва да спазват определени правила, установени преди завършване на научното изследване. Това е общоприета практика сред академичните и търговските групи, които изработват в последните десетилетия методологията за библиометрична оценка. Тези групи публикуват свои протоколи в рецензирани издания. Подобна прозрачност позволява щателна проверка. Например, през 2010 г. обществено обсъждане на техническите качества на един от важните индикатори, използван от една от четирите групи (Centre for Science and Technology Studies at Leiden University in the Netherlands [Центъра за научни и технически изследвания в Лейденския университет в Нидерландия]), доведе до преразглеждане на изчисляването на този индикатор [6]. Споменатите търговски организации трябва да се придържат към същите стандарти; никой не трябва да прилага при оценяване машина тип „черна кутия”.

Простотата е силата в един индикатор, защото гарантира прозрачност. Но опростените метрики могат да изкривят резултатите (виж принцип 7). Оценителите трябва да се стремят към баланс – прости индикатори, отговарящи на сложността на изследователския процес.

5. Позволявайте на оценяваните да проверяват данните и анализа. За да осигурите качеството на данните, всички учени, включени в библиометричното изследване, трябва да могат да проверят дали техните трудове са коректно идентифицирани. Всеки, който се занимава и ръководи процеса на оценка, трябва да е убеден в точността на данните чрез самопроверка или проверка от трета страна. Университетите трябва да го въведат в своите научни информационни системи и това трябва да е водещ принцип в избора на доставчик на тези системи. По правило висококачествените данни отнемат време и пари, за да се проверят и обработят. За това е нужен бюджет.

6. Научните области се отличават една от друга по практики на публикуване и цитиране. Най-добрата практика е да се подбере набор от възможни индикатори и да се позволи на научните области да избират от тях. Преди няколко години Европейската група на историците получи сравнително ниска оценка, защото те публикуват повече книги отколкото статии в списания, индексирани в Web of Science. Историците са имали лошия късмет да бъдат част от Департамента по психология. При отчитане на публикациите на историците и учените в областта на обществените науки трябва да бъдат отчитани книги и статии на националния език; за информатиците трябва да бъдат отчитани доклади от конференции.

Количеството цитати зависи от научната област: математическите списания с най-висок рейтинг имат импакт фактор около 3; списанията в областта на клетъчната биология с най-висок рейтинг имат импакт фактор около 30. Необходими са нормализирани индикатори и най-убедителният метод за нормализация е основан на проценти: всеки труд се измерва на основата на проценти в зависимост от разпределението на цитиранията в научната област, към която спада (например, в

първия/ите 1%, 10% или 20%). Една високоцитирана публикация малко подобрява позицията на университета в класация, която е основана на процентни индикатори, но може да изтласка университета от средата към върха на класация, разработена на осреднени показатели на цитиране [7].

7. Оценявайте отделните учени на основата на качествена оценка на тяхното портфолио. Колкото сте по-възрастни, толкова вашият h-индекс е по-висок, дори и да нямате нови публикации. H-индексът зависи от научната област: учените в областта на природните науки достигат 200; физиците достигат 100 и учените в областта на обществените науки до 20-30 [8]. Зависи и от базата от данни: има учени, работещи в областта на информатиката, които имат h-индекс около 10 в Web of Science, но 20-30 в Google Scholar [9]. Четенето и оценяването на труда на учения е много по-уместно, отколкото осланянето на едно число. Дори когато се сравняват голям брой учени, най-добрият подход е този, който разчита на повече информация за индивидуалните знания, опит, дейност и влияние.

8. Избягвайте неуместната конкретност и фалшивата точност. Научно-технологичните индикатори са склонни към концептуална двусмисленост и неопределеност и изискват ясни предположения, които не са приети от всички. Например, дълго време бе обсъждано значението на отчитането на цитиранията. Така най-добрата практика използва множество индикатори, за да предостави по-ясна и плуралистична картина. Ако несигурността и грешката могат да бъдат количествено определени, например с помощта на ленти за грешки, тази информация трябва да придружава публикуваните стойности на показателя. Например, за да се избегнат съвпадения импакт факторът на списанията се публикува до третия десетичен знак. Въпреки това концептуалната двусмисленост и случайната изменчивост на отчетените цитирания нямат значение при оценка на списанията на основата на много малки разлики в импакт фактора. Избягвайте фалшивата прецизност: необходим е само една цифра след десетичния знак.

9. Отчитайте системното въздействие на оценката и индикаторите. Индикаторите променят системата чрез стимулите, които създават. Тези ефекти трябва да бъдат предвидени. Това означава, че набор от индикатори винаги е за предпочитане – използването само на един измества целите (показателят става цел). Например през 90-те години на XX век Австралия финансира университетско научно изследване, използващо формула основаваща се основно на броя на публикациите от всеки институт. Университетите можели да изчисляват „стойността“ на дадена публикация в рецензирано списание; през 2000 г. тя се е равнявала на 800 австралийски долара (около US\$480 през 2000 г.). Както може и да се очаква броят на публикациите на австралийските учени нараснал, но те били в по-малко цитирани списания, което предполага намаляване на качеството на статиите [10].

10. Редовно преглеждайте индикаторите и ги актуализирайте. Изследователските задачи и целите на оценяването се променят и самата изследователска система се развива. Използваните вече индикатори стават неадекватни; появяват се нови. Системата от индикатори трябва да се преразглежда и вероятно да се променя. Осъзнавайки ефектите от опростената формула, през 2010 г. Австралия представи нова по-сложна формула за оценка, която гарантира качество – инициативата Excellence in Research in Australia [Отлични постижения на научните изследвания в Австралия].

Следващи стъпки

Спазвайки тези десет принципа, оценката на научните изследвания може да играе важна роля в развитието на науката и нейните взаимодействия с обществото. Наукометрията може да предостави значима информация, която би било трудно да бъде

събрана или разбрана чрез индивидуална експертиза. Но не бива да бъде позволявано тази количествена информация да се превърне от инструмент в цел.

Най-добрите решения се вземат чрез съчетаване на надеждна статистика с внимание към целта и същността на научното изследване, което се оценява. Необходими са както количествени, така и качествени данни; всички те са обективни по свой собствен начин. Вземането на решения за науката трябва да се основава на висококачествени процеси, които са осигурени с висококачествени данни.

Диана Хикс (Diana Hicks) е професор по публична политика в Джорджийския технологичен институт, Атланта, Джорджия, САЩ.

Пол Уотърс (Paul Wouters) е професор по наукометрия и директор на Центъра за научни и технологични изследвания към Университета в Лейден, Нидерландия.

Лудо Уолтман (Ludo Waltman) е изследовател в Центъра за научни и технологични изследвания към Университета в Лейден, Нидерландия.

Сара де Рийке (Sarah de Rijcke) е доцент в Центъра за научни и технологични изследвания към Университета в Лейден, Нидерландия.

Исмаил Рафолс (Ismael Rafols) е изследовател в областта на научната политика в Испанския национален научноизследователски университет и Политехническият университет във Валенсия, Испания.

Имейл: diana.hicks@pubpolicy.gatech.edu

1. Wouters, P. (2014). *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*. Cambridge. MIT Press.
2. Shao, J., Shen, H. (2011). The Outflow of Academic Papers from China: Why is it Happening and can it be stemmed? *Learned Publishing*. XXIV, (2), pp. 95-97.
3. Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *Br. Med. J.* № 314, pp. 498-502
4. Garfield, E. (2006). The history and meaning of the journal impact factor. *Journal of the American Medical Association*. Vol. 295, (1), pp. 90-93.
5. Lopez Pineiro, C., Hicks, D. (2015). Reception of Spanish Sociology by Domestic and Foreign Audiences Differs and Has Consequences for Evaluation. *Research Evaluation*. Vol. 24, (1), pp. 78-89.
6. van Raan, A. F. J. et al. (2010). Rivals for the Crown: Reply to Opthof and Leydesdorf. *Journal of Informetrics*. Vol. 4, (3), pp. 431-435.
7. Waltman, L. et al. (2012). The Leiden Ranking 2011/2012: Data Collection, Indicators, and Interpretation. *Journal of the American Society for Information Science and Technology*. Vol. 63, (12), pp. 2419-2432.
8. Hirsch, J. E. (2005). An Index to Quantify an Individual's Scientific Research Output. *Proceedings of the National Academy of Sciences of the United States of America*. Vol. 102 (46), pp.16569-16572.
9. Bar-Ilan, J. (2008). Which h-index – A Comparison of Web of Science, Scopus and Google Scholar. *Scientometrics*. Vol. 74, (2), pp. 257-271.
10. Butler, L. (2003). Explaining Australia's Increased Share of ISI Publications – The Effects of a Funding Formula Based on Publication Counts. *Research Policy*. Vol. 32, (1), pp. 143-155.