

Lajdenski manifest o vrednovanju istraživanja

Diana Hicks, Paul Wouters i saradnici pozivaju na poštovanje ovih deset principa prilikom vrednovanja naučnih istraživanja

Podaci se sve češće koriste za upravljanje naukom. Evaluacija naučnih istraživanja je nekada bila prilagođena pojedincu i sprovedena od strane kolega, a danas postaje stvar rutine i oslanja se na merenje¹. Problem je što se vrednovanje rukovodi podacima, a ne rasuđivanjem. Metrika postaje sve popularnija: obično se sprovodi u dobroj nameri, ne oslanja se uvek na pouzdane podatke, često se pogrešno primenjuje. Tako se rizikuje nanošenje štete sistemu upravo onim alatkama koje su namenjene njegovom poboljšanju, pošto evaluaciju sve češće sprovode organizacije koje nisu dovoljno upućene i ne razumeju dobru praksu evaluacije i pravilnu interpretaciju rezultata.

Pre 2000. godine, baza *Science Citation Index* je bila dostupna na kompaktnim diskovima *Instituta za naučne informacije* (ISI) iz Filadelfije, a koristili su je eksperti i obučeni analitičari. *Thomson Reuters* je 2002. godine pokrenuo integrisanu veb platformu, čime je baza podataka *Web of Science* postala široko dostupna. Nastali su i konkurentski citatni indeksi: Elsevierov *Scopus* (pokrenut 2004. godine) i *Google Scholar* (beta verzija pokrenuta 2004. godine). Uvedene su veb alatke za lako poređenje produktivnosti i uticajnosti istraživačkih ustanova, kao što su *InCites* (koji koristi *Web of Science*) i *SciVal* (koji koristi *Scopus*), kao i softver za analizu citiranosti pojedinaca u bazi *Google Scholar* (*Publish or Perish* objavljen 2007. godine).

Jorge Hirsch, fizičar sa Univerziteta u Kaliforniji u San Dijegu, predložio je 2005. godine *h-indeks* i na taj način popularizovao prebrojavanje citata pojedinačnih istraživača. Interesovanje za faktor uticajnosti časopisa (impakt faktor) stalno je raslo od 1995. godine (pogledajte sliku [Opčinjenost impakt faktorom](#) u ovom dokumentu).

U poslednje vreme, merenje bazirano na društvenoj upotrebi i internet komentarima postaje sve popularnije - *F1000Prime* je osnovan 2002. godine, *Mendely 2008*, a *Altmetric.com* (podržan od strane *Macmillan Science and Education*, vlasnika *Nature Publishing Group*) 2011. godine.

Kao scijentometričari, naučnici iz oblasti društvenih nauka i rukovodioci istraživanja, sa sve većom pažnjom smo posmatrali sveopštu pogrešnu primenu indikatora za vrednovanje naučnog učinka. Navodimo samo nekoliko od brojnih primera. Širom sveta, univerziteti postaju opsednuti svojom pozicijom na globalnim rang listama (kao što su *Šangajska lista univerziteta* i *Times Higher Education* lista), čak i ako se takve liste zasnivaju na, po našem mišljenju, nepotpunim podacima i proizvoljnim indikatorima.

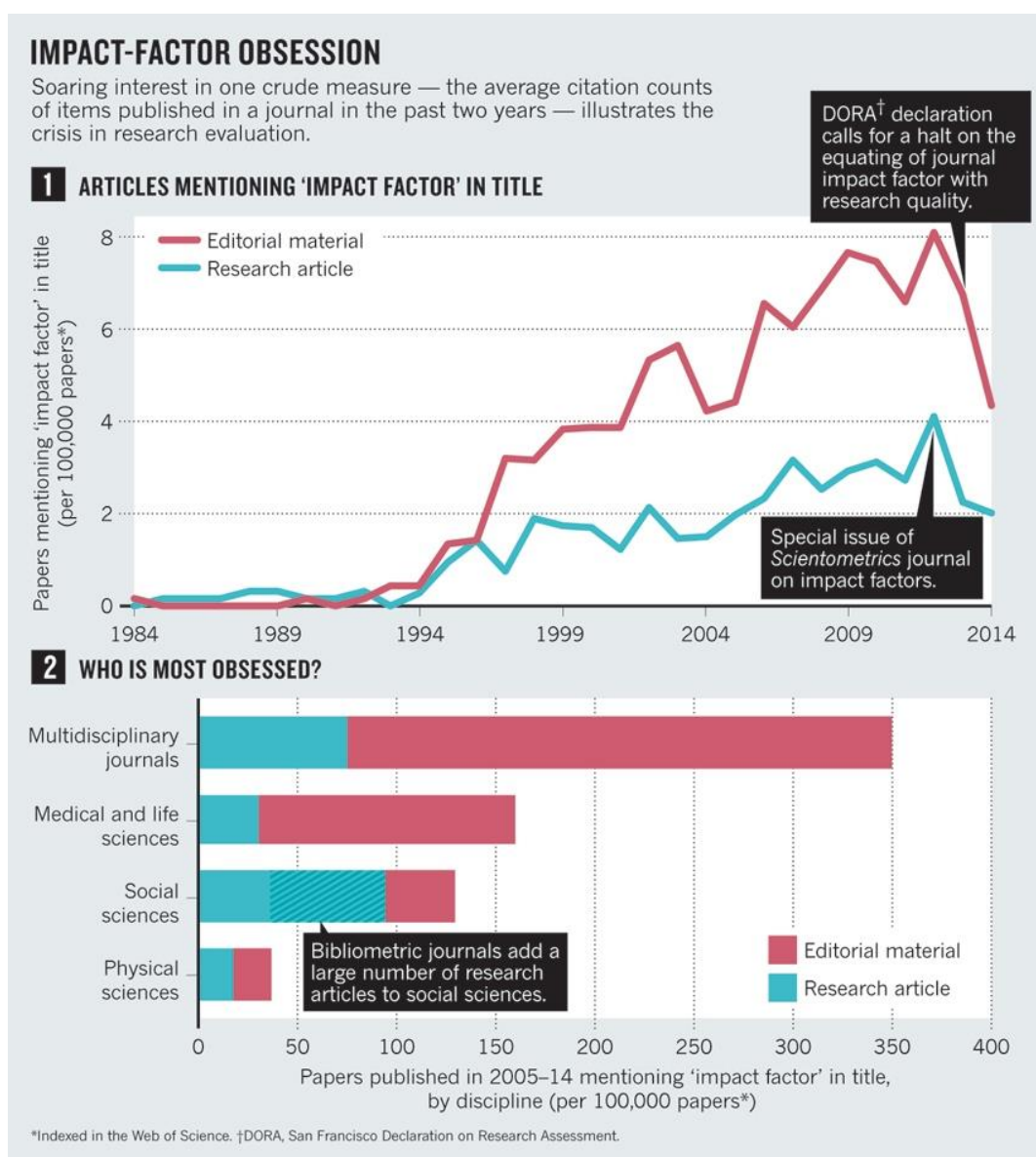
Neki poslodavci zahtevaju određene vrednosti h-indeksa kandidata prijavljenih na konkurse za radna mesta. Mnogi univerziteti donose odluke o izborima u zvanja na osnovu vrednosti h-indeksa i broja članaka u časopisima sa visokim impakt faktorom. Biografije istraživača postale su način da se razmeće ovim rezultatima, naročito u biomedicini. Mentori traže od studenata doktorskih studija da objavljuju radove u časopisima sa visokim impakt faktorom i obezbede spoljno finansiranje pre nego što su zaista spremni za to.

U Skandinaviji i Kini neki univerziteti izdvajaju sredstva za istraživanja ili dodeljuju bonuse na osnovu brojki, računajući, na primer, indikatore uticajnosti pojedinaca za raspodelu "sredstava baziranih na učinku" ili finansijski nagrađujući istraživače za publikovanje u časopisima s impakt faktorom većim od 15².

U mnogim slučajevima, istraživači i procenjivači još uvek primenjuju objektivnu i izbalansiranu procenu. Ipak, zloupotreba metrike istraživanja postaje previše zastupljena da bi se ignorisala.

Zbog toga predstavljamo Lajdenski manifest, nazvan po konferenciji na kojoj je uobličen (vidi <http://sti2014.cvts.nl>). Navedenih deset principa nisu novost za scijentometričare, mada nijedan od nas ne bi bio u stanju da ih izrecituje u celini, jer do sada nisu bili sistematizovani. Začetnici ove oblasti, kao što je Eugene Garfield (osnivač ISI), već su javno iznosili neka od ovih načela³⁴. Ali oni se ne oglašavaju kada evaluatori predaju izveštaje upravama univerziteta koje ne poznaju u dovoljnoj meri relevantnu metodologiju. Naučnici koji traže literaturu da bi osporili evaluaciju, pronalaze materijal raštrkan u onome što su za njih časopisi nejasnog statusa ili teško dostupni časopisi.

Predstavljamo ovaj sažetak dobre prakse vrednovanja istraživanja zasnovanog na merenju kako bi istraživači mogli da imaju poverenja u evaluatore, a evaluatori u svoje indikatore.



Deset principa

1 Kvantitativna evaluacija treba da bude dopuna kvalitativnoj proceni stručnjaka. Kvantitativno vrednovanje može da umanjí pristrasnost recenzija i da olakša donošenje odluke. Ono bi trebalo da osnaži proces recenziranja, jer je donošenje suda o kolegama teško bez dovoljne količine pouzdanih informacija. Međutim, procenjivači ne smeju da budu u iskušenju da prepuste donošenje odluke brojka. Kvantitativni indikatori ne smeju da zamene informisanu procenu. Svaki pojedinac je odgovoran za ocene koje je dao.

2 Vrednujte naučni učinak u skladu sa istraživačkim ciljevima institucije, grupe ili istraživača. Programske ciljeve treba navesti na početku, a indikatore koji se koriste za procenu naučnog učinka treba jasno povezati s tim ciljevima. Prilikom izbora pokazatelja i načina na koji se koriste, treba uzeti u obzir širi socio-ekonomski i kulturni kontekst. Naučnici imaju različite istraživačke misije. Bazična istraživanja koja unapređuju granice akademskog znanja razlikuju se od istraživanja kojima je cilj da ponude rešenja društvenih problema. Recenzija može da bude zasnovana na doprinosima relevantnim za razvoj strategija, industriju ili širu javnost, a ne na akademskim idejama izvrsnosti. Ne postoji samo jedan model evaluacije koji je primenjiv u svim kontekstima.

3 Zaštitite specifičnosti lokalno važnih istraživanja. U mnogim delovima sveta, izvrsnost istraživanja se izjednačava s objavljivanjem na engleskom jeziku. Španski zakon, na primer, naglašava potrebu da španski naučnici objavljuju u časopisima sa visokim impakt faktorom. Impakt faktor se izračunava za časopise indeksirane u bazi Web of Science koja se objavljuje u SAD i još uvek najvećim delom na engleskom jeziku. Ova pristrasnost je naročito problematična u društvenim i humanističkim naukama u kojima su istraživanja u većoj meri usmerena na regionalni i nacionalni nivo. I mnoge druge oblasti imaju nacionalnu ili regionalnu dimenziju - na primer, epidemiologija HIV-a u Podсахarskoj Africi.

Ova raznovrsnost i specifičan društveni značaj često se potiskuju da bi se objavili radovi od interesa za čuvare visokog uticaja: časopise na engleskom jeziku. Španski sociolozi koji su visoko citirani u bazi Web of Science radili su na hipotetičkim modelima ili istraživali američke podatke. Izgubljena je specifičnost sociologa u časopisima s visokim impakt faktorom na španskom jeziku: teme kao što su lokalno radno pravo, porodična zdravstvena zaštita za starije i zapošljavanje imigranata⁵. Merenje bazirano na visokokvalitetnoj neengleskoj literaturi trebalo bi da služi za identifikaciju i nagrađivanje izvrsnosti u lokalno relevantnim istraživanjima.

4 Učinite prikupljanje podataka i analitičke procedure javnim, transparentnim i jednostavnim. Kreiranje baza podataka neophodnih za evaluaciju treba da se sprovodi u skladu s jasnim pravilima, postavljenim pre nego što je istraživanje završeno. Ovo je bila uobičajena praksa među akademskim i komercijalnim grupama koje su tokom nekoliko decenija izgradile metodologiju bibliometrijskog vrednovanja. Te grupe su se pozivale na procedure opisane u recenziranoj literaturi. Ovakva transparentnost je omogućavala kontrolu i preispitivanje. Na primer, 2010. godine je javna rasprava o tehničkim karakteristikama važnog indikatora koju je koristila jedna od naših grupa (Centar za nauku i tehnološke studije Univerziteta u Leidenu, Holandija) dovela do revizije u računanju ovog indikatora⁶. I novi komercijalni učesnici treba da se drže istih standarda; niko ne treba da prihvati crnu kutiju kao mašinu za evaluaciju.

Jednostavnost je vrлина indikatora jer povećava transparentnost. Ali pojednostavljene metrike mogu da iskrive sliku (vidite 7. princip). Evaluatori treba da pronađu ravnotežu - jednostavne indikatore koji verno odražavaju složenost istraživačkog procesa.

5 Dozvolite onima koji se procenjuju da provere podatke i analizu. Da bi se obezbedio kvalitet podataka, svim istraživačima uključenim u bibliometrijske studije treba omogućiti da provere da li su njihovi rezultati pravilno prepoznati. Svi koji rukovode i sprovode evaluaciju treba da osiguraju tačnost podataka putem samo-verifikacije ili eksterne revizije. Univerziteti bi mogli da uvedu ovu praksu u svoje istraživačke informacione sisteme i to bi trebalo da bude osnovni princip pri izboru dobavljača ovih sistema. Potrebno je vreme i novac da bi se prikupili i obradili pouzdani i kvalitetni podaci. Predvidite budžet za to.

6 Uzmite u obzir razlike među oblastima u načinima objavljivanja i citiranja. Najbolja praksa je odabrati niz potencijalnih indikatora i dozvoliti istraživačima iz različitih oblasti da se opredele. Pre nekoliko godina, grupa istoričara iz Evrope je dobila relativno nizak rejting u nacionalnoj recenzentskoj proceni jer su pisali knjige, a ne članke u časopisima indeksiranim u bazi Web of Science. Istoričari su imali nesreću da budu deo odeljenja za psihologiju. Istoričari i naučnici iz oblasti društvenih nauka zahtevaju da knjige i literatura na nacionalnom jeziku budu uzeti u obzir prilikom vrednovanja njihove produktivnosti, dok naučnici iz oblasti informacionih tehnologija zahtevaju da se računaju saopštenja sa naučnih skupova.

Citiranost varira u zavisnosti od naučne oblasti: najbolje rangirani časopisi u matematici imaju impakt faktore od oko 3; vrhunski časopisi u mikrobiologiji imaju impakt faktore od oko 30. Neophodni su normalizovani indikatori, a najrobusniji metod normalizacije zasniva se na centilima: svaki članak se ponderiše na osnovu centila kome pripada u distribuciji citata u odgovarajućoj oblasti (npr. prvih 1%, 10% ili 20%). Jedna visoko citirana publikacija neznatno poboljšava poziciju univerziteta na rang listi koja se zasniva na centilima, ali može da pomeri poziciju univerziteta od sredine do vrha na rang listi formiranoj na osnovu prosečne citiranosti⁷.

7 Bazirajte procenu istraživača na kvalitativnoj oceni njihovog portfolija. Što ste stariji, to je veći vaš h-indeks, čak i ako ne objavljujete nove naučne radove. H-indeks varira u zavisnosti od naučne oblasti: najviši h-indeks u biologiji i biohemiji je oko 200, u fizici oko 100, a u društvenim naukama od 20 do 30⁸. On zavisi i od baze podataka koja se koristi: postoje istraživači u oblasti informacionih tehnologija koji imaju h-indeks od oko 10 u bazi Web of Science, a 20-30 u bazi Google Scholar⁹. Čitanje i ocenjivanje rada istraživača je prikladnije nego oslanjanje na jedan broj. Čak i kada se upoređuje veliki broj istraživača, najbolji je pristup koji uzima u obzir više informacija o stručnosti, iskustvu, aktivnostima i uticajnosti pojedinca.

8 Izbegavajte "pogrešno smeštenu" preciznost. Naučni i tehnološki indikatori su osetljivi na konceptualne dvosmislenosti i nejasnoće i zahtevaju jaku osnovu koja nije opšte prihvaćena. Na primer, dugo se raspravljalo o značenju citiranosti. Stoga je najbolja praksa upotreba većeg broja indikatora kako bi se obezbedila stabilnija i sveobuhvatnija slika. Ako neizvesnost i greška mogu da se kvantifikuju, na primer korišćenjem raspona greške, ove informacije treba da se dodaju objavljenim vrednostima indikatora. Ako to nije moguće, korisnici indikatora bi barem trebalo da izbegnu lažnu ili "pogrešno smeštenu" preciznost. Na primer, impakt faktor časopisa objavljuje se sa tri decimalna mesta, kako bi se omogućila preciznost. Međutim, s obzirom na konceptualne nejasnoće i slučajnu varijabilnost broja citata, nema smisla razlikovati časopise na osnovu minornih razlika u impakt faktoru. Izbegavajte lažnu preciznost: samo jedno decimalno mesto je opravdano.

9 **Prepoznajte sistematske posledice vrednovanja i primene indikatora.** Indikatori menjaju sistem putem podsticaja koje stvaraju. Ove efekte treba predvideti. To znači da je uvek poželjno koristiti skup indikatora - samo jednim se može manipulirati i zameniti cilj (tako da merenje postaje cilj). Na primer, Australija je tokom devedesetih finansirala univerzitetsko istraživanje koristeći formulu zasnovanu najvećim delom na broju radova objavljenih od strane instituta. Univerziteti su mogli da računaju "vrednost" rada u recenziranom časopisu: 2000. godine svaki rad je obezbeđivao 800 australijskih dolara (oko 480 američkih) za finansiranje istraživanja. Kao što se moglo predvideti, broj radova australijskih istraživača je porastao, ali su objavljivani u slabije citiranim časopisima, što je ukazivalo na to da je kvalitet naučnih radova opao¹⁰.

10 **Redovno preispitujte indikatore i ažurirajte ih.** Istraživačke misije i ciljevi evaluacije se menjaju, a sami istraživački sistemi evoluiraju u skladu s tim. Metrika korisna u jednom momentu, postaje neadekvatna; pojavljuje se nova. Sistem indikatora mora da se preispita, a možda i modifikuje. Shvatajući efekte svoje pojednostavljene formule, Australija je 2010. godine uvela kompleksniju *Excellence in Research for Australia* inicijativu, u kojoj se naglasak stavlja na kvalitet naučnog rada.

Sledeći koraci

Uz poštovanje ovih deset principa, evaluacija istraživanja može da odigra važnu ulogu u razvoju nauke i njenoj interakciji s društvom. Merenje rezultata istraživanja može da obezbedi ključne informacije koje bi bilo teško prikupljati ili razumeti iz perspektive individualne ekspertize. Ali ne sme se dozvoliti da kvantitativne informacije od alatke postanu cilj.

Najbolje odluke se donose kombinacijom robusne statistike i poznavanja cilja i prirode naučnog istraživanja koje se evaluira. Neophodno je koristiti i kvantitativne i kvalitativne pokazatelje; svaki je objektivan na svoj način. Odlučivanje o nauci mora se zasnivati na visokokvalitetnim postupcima koji pružaju informacije na osnovu najkvalitetnijih podataka.

1. Wouters, P. in *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact* (eds Cronin, B. & Sugimoto, C.) 47–66 (MIT Press, 2014).
2. Shao, J. & Shen, H. *Learned Publ.* **24**, 95–97 (2011).
3. Seglen, P. O. *Br. Med. J.* **314**, 498–502 (1997).
4. Garfield, E. *J. Am. Med. Assoc.* **295**, 90–93 (2006).
5. López Piñeiro, C. & Hicks, D. *Res. Eval.* **24**, 78–89 (2015).
6. van Raan, A. F. J., van Leeuwen, T. N., Visser, M. S., van Eck, N. J. & Waltman, L. *J. Informetrics* **4**, 431–435 (2010).
7. Waltman, L. *et al. J. Am. Soc. Inf. Sci. Technol.* **63**, 2419–2432 (2012).
8. Hirsch, J. E. *Proc. Natl Acad. Sci. USA* **102**, 16569–16572 (2005).
9. Bar-Ilan, J. *Scientometrics* **74**, 257–271 (2008).
10. Butler, L. *Res. Policy* **32**, 143–155 (2003).

Izvor originala:

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429. doi:10.1038/520429a

O autorima:

Diana Hicks je profesor javne politike na Tehnološkom institutu u Džordžiji (*Georgia Institute of Technology, Atlanta, Georgia, USA*). Paul Wouters is profesor scijentometrije i upravnik Centra za naučne u tehnološke studije Univerziteta u Lajdenu (*Centre for Science and Technology Studies, Leiden University, the Netherlands*), Ludo Waltman je istraživač, a Sarah de Rijcke docent u istom centru. Ismael Rafols je istraživač koji se bavi naučnom politikom (*Spanish National Research Councili Polytechnic University of Valencia, Spain*).

e-mail: diana.hicks@pubpolicy.gatech.edu

Prevod / Translated by:

dr Snežana Smederevac i dr Dejan Pajić

Odsek za psihologiju Filozofskog fakulteta Univerziteta u Novom Sadu