

Leideni manifest teaduse paremaks hindamiseks

Hicks, Diana, Paul Wouters, Ludo Waltman, Sarah de Rijcke, Ismael Rafols 2015. *Bibliometrics: The Leiden Manifesto for research metrics*. – *Nature* 520, 429–431.

<http://www.leidenmanifesto.org/>

Tõlkinud Hele Priimets, SA Eesti Teadusagentuur. Tartu 2019.

Diana Hicks, Paul Wouters ja nende kolleegid teevad üleskutse juhinduda teaduse hindamisel alljärgnevast kümnest põhimõttest.

Teadust on hakanud valitsema arvud. Kui algselt hindasid teadustööd eksperdid, kellele oli omane individuaalne lähenemine, siis nüüdseks on teaduse hindamisest saanud rutiinne tegevus, mis tugineb bibliomeetrialet.¹ Probleem seisneb selles, et juhtpositsioonil olnud eksperthinnang on asendunud arvandmete kultusega. Levib bibliomeetriausk: üldjuhul heatahtlik, mitte alati asjatundlik, sageli valesti rakendatud. Kuna üha enam hindavad teadustööd asutused, kus jääb vajaka tõlgendamisoskusest ja head hindamistava puudutavatest teadmistest ning nõuannetest, varitseb oht, et lõhume teaduse hindamise süsteemi, kasutades neid samu abivahendeid, mille eesmärk oli seda süsteemi parandada.

Enne 2000. aastat kasutasid eksperdid analüüside tegemisel CD-ROM-il olevat *Science Citation Index*'it, mille oli välja töötanud USA Teadusliku Informatsiooni Instituut (*Institute for Scientific Information; ISI*). 2002. aastal lõi Thomson Reuters integreeritud veebiplatvormi ja muutis *Web of Science*'i andmebaasi laialdaselt kättesaadavaks. Peagi loodi konkureerivad tsiteerimisindeksid: Elsevier'i andmebaas *Scopus* (avati 2004. aastal) ja *Google Scholar* (beetaversioon avati 2004. aastal). Võeti kasutusele ka veebipõhised analüütilised tööriistad nagu *InCites* (seotud andmebaasiga *Web of Science*) ja *SciVal* (seotud andmebaasiga *Scopus*), mida kasutades saab hõlpsalt võrrelda erinevate asutuste teadustöö produktiivsust ja mõju, ning tarkvara (*Publish or Perish*, mis avati 2007. aastal), mille abil saab analüüsida teadlaste tsiteeritavust ja mis põhineb *Google Scholar*'i andmestikul.

2005. aastal tutvustas San Diegos asuva California Ülikooli füüsik Jorge Hirsch h-indeksit, propageerides seeläbi teadlaste tsiteerimise mõõtmist. Huvi ajakirja mõjufaktori vastu on kasvanud stabiilselt pärast 1995. aastat (vt graafikut <https://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>).

Viimasel ajal on kogunud hoogu bibliomeetria suund, mille allikateks on internetikommentaariid ja postitused sotsiaalmeedias: 2001. aastal loodi *F1000Prime*, 2008. aastal *Mendeley* ja 2011. aastal *Altmetric.com* (koostöös kirjastusega *Macmillan Science and Education*, mille omanduses on *Nature Publishing Group*).

Teadusmeetria spetsialistidena, sotsiaalteadlaste ja teadusadministraatoritena oleme me üha suurema ärevusega jälginud teadustöö tulemuslikkuse näitajate ulatuslikku väärkasutust. Järgnevad näited on vaid mõned paljude seast. Kõikjal maailmas on ülikooli positsioon rahvusvahelistes pingeridades (näiteks Shanghai ja *Times Higher Education*'i edetabelites) muutunud kinnisideeks, vaatamata sellele, et meie arvates sisaldavad need pingeread ebatäpseid andmeid ja meelevaldseid näitajaid.

Mõned tööandjad küsivad kandidaatidelt nende h-indeksit. Mitmed ülikoolid langetavad edutamisoskuseid h-indeksile kehtestatud lävendi ja „kõrge mõjufaktoriga“ ajakirjades avaldatud artiklite arvu põhjal. Teadlase CV on muutunud nende näitajatega

hooplemise mängulavaks, eelkõige biomeditsiinis. Levinud on ka see, et doktoriõppe üliõpilaste juhendajad soovivad neil avaldada artikleid kõrge mõjufaktoriga ajakirjades ja hankida väliseid rahastamisallikaid veel enne seda, kui noored teadlased on selleks päriselt valmis.

Skandinaavia riikides ja Hiinas eraldavad mõned ülikoolid uurimistoetusi või lisatasusid arvulise näitaja põhjal: näiteks arvutatakse „tulemustasu“ eraldamiseks välja individuaalne mõjufaktor või määratakse teadlasele lisatasu selle eest, et ta on avaldanud artikli ajakirjas, mille mõjufaktor on suurem kui 15.²

Enamasti võtavad teadlased ja hindajad vastu tasakaalukaid otsuseid. Siiski on bibliomeetriliste andmete väärkasutamine muutunud nii laialdaseks, et seda pole enam võimalik ignoreerida.

Seega esitleme me Leideni manifesti, mis on saanud nime konverentsi järgi, mille käigus omandas see selge kuju (vt <http://sti2014.cwts.nl>). Manifesti kümme põhimõtet ei ole küll teadusmeetria spetsialistide jaoks midagi uut, kuid kuna neid ei ole siiani kodifitseeritud, ei suuda keegi neid tervikuna esitada. Bibliomeetria suurkujud, näiteks Eugene Garfield (*ISI* asutaja), on mõnda nendest põhimõtetest enda kirjutistes maininud.^{3 4} Siiski ei viibi nad ruumis siis, kui ekspert esitab tulemused ülikooli administratiivtöötajale, kes ei valda vastavat metoodikat. Teadlased, kes otsivad allikaid, millele tuginedes tulemus vaidlustada, leiavad vaid infokilde, mis on pärit (nende meelest) kummalistest ja nende jaoks piiratud ligipääsuga ajakirjadest.

Alljärgnevas kokkuvõttes pakume välja teaduse bibliomeetriapõhise hindamise hea tava, mille kohaselt peavad teadlased julgema kahelda hindajates ja hindajad julgema kahelda bibliomeetrilistes näitajates.

Kümme põhimõtet

1) Kvantitatiivne hindamine peab toetama kvalitatiivset eksperthinnangut. Kvantitatiivsed näitajad võivad aidata vältida erapoolikut suhtumist eksperthinnangu andmisel ja hõlbustada kaalutletud otsuse tegemist. See peaks lisama eksperthinnangule kaalukust, sest kui puudub mitmekülgne vajalik info, siis on keeruline langetada kolleege puudutavaid otsuseid. Siiski ei tohi eksperdid anda järele kiusatusele asendada enda otsustuspädevus arvudega. Arvnäitajad ei tohi muutuda kaalutletud otsuste aseaineks. Igaüks vastutab enda hinnangu eest ise.

2) Teadustulemusi tasub kõrvutada asutuse, rühma või teadlase uurimistöö eesmärkidega. Seatud sihid peavad olema selgelt sõnastatud algusest peale ja näitajad, mida kasutatakse teadustulemuste hindamiseks, peavad olema nende sihtidega otseses seoses. Näitajate valik ja nende kasutusviis peab võtma arvesse ka laiemat sotsiaalpoliitilist ja kultuurilist konteksti. Teadlased seavad enda uurimistööle mitmesuguseid eesmärke. Teadustöö, mis avardab teoreetilise mõtte piire, erineb teadustööst, mis on keskendunud sellele, et leida ühiskondlikele probleemidele lahendusi. Hinnang teadustööle võib lähtuda selle mõjust seadusloome, ettevõtluse või avalikkuse hüvangule, mitte sellest, et akadeemilises sfääris on tegemist silmapaistva ideega. Universaalset hindamismeetodit ei eksisteeri.

3) Kohaliku tähtsusega teadustööd on vaja väärtustada. Mitmel pool maailmas võrdsustatakse tippteadust ingliskeelsete publikatsioonidega. Näiteks sätestab üks Hispaania õigusakt, et erilist tunnustust väärivad need Hispaania teadlased, kes avaldavad artikleid

kõrge mõjufaktoriga ajakirjades. Mõjufaktorit arvutatakse aga Ameerika Ühendriikides indekseeritud ajakirjade ja peajasjalikult ingliskeelse *Web of Science*'i põhjal. Selline erapoolikus on suureks probleemiks iseäranis sotsiaal- ja humanitaarteadustes, sest nendes valdkondades on teadustöö seos piirkondliku ja riikliku tasandiga tuntavam. Ka paljudes teistes valdkondades eksisteerib riiklik või piirkondlik mõõde. Üheks näiteks on HIV epidemioloogia Saharast lõunasse jäävas Aafrikas.

Mitmekesisust ja ühiskondlikku olulisust surutakse aga tagaplaanile, et luua artikleid, mis pakuvad huvi kõrge mõjufaktori suunanäitajatele: ingliskeelsetele ajakirjadele. Hispaania sotsioloogid, kelle töid on *Web of Science*'i andmetel väga palju tsiteeritud, on keskendunud abstraktsetele mudelitele või Ühendriikide andmete uurimisele. Sotsioloogidele eriomased teemad nagu kohaliku tasandi tööõigus, eakate omaste hoolekanne ja sisserändajate tööhõive on kõrge mõjufaktoriga hispaaniakeelsetest artiklitest kadunud.⁵ Näitajad, mis võtavad arvesse kõrgetasemelisi mitte-ingliskeelseid kirjutisi, aitaksid ära tunda ja väärtustada silmapaistvalt head kohaliku tähtsusega teadustööd.

4) Andmete kogumine ja analüüsimine peab olema avatud, läbipaistev ja lihtne. Hindamiseks vajalike andmebaaside loomisel on oluline järgida reegleid, mis on eelnevalt sõnastatud. Selline praktika oli valdav teadlas- ja äriühingutes, kus bibliomeetriapõhise hindamise meetodika aastakümnete jooksul välja töötati. See inimeste ring tugines eelretsenseeritavates ajakirjades avaldatud allikatele ja seesugune läbipaistvus rajas teed üksikasjalikumale analüüsile. Näiteks toimus 2010. aastal avalik arutelu ühe olulise bibliomeetrilise näitaja tehniliste aspektide teemal. Selle tulemusena arvutatakse antud näitajat, mida kasutab ka üks osa meie endi seast (nimelt Hollandis asuva Leideni Ülikooli teadus- ja tehnoloogiauringute keskuse töötajad) nüüd teisiti.⁶ Uued turule sisenejad peavad samuti nendest standarditest lähtuma. Mitte keegi ei tohiks olla päri sellega, et teadustöö hindamiseks kasutatakse n-ö musta kasti meetodit.

Iga näitaja suurimaks vooreseks on lihtsus, sest see parandab läbipaistvust. Ülemäära lihtsustatud näitajad võivad aga tegelikkust moonutada (vt 7. põhimõtet). Hindajad peavad seega leidma tasakaalu ja tuginema lihtsatele näitajatele, mis arvestavad ka hindamisprotsessi keerukusega.

5) Teadlastel, keda hinnatakse, peab olema võimalik tutvuda andmete ja analüüsiga. Andmete kvaliteedi tagamiseks peab kõikidel bibliomeetrilisse analüüsi kaasatud teadlastel olema võimalik kontrollida, kas nende teadustöö väljundid on korrektselt identifitseeritud. Kõik, kes suunavad ja juhivad hindamisprotsessi, peavad tagama andmete õigsuse, olgu selleks vajalik enesekohaste andmete kontroll või kolmanda osapoole sõltumatu hinnang. Ülikoolidel on võimalik tagada andmete õigsus teadusinfosüsteemis. Lisaks on ülikooli ülesanne valida taoliste süsteemide pakkujate seast välja need, kes juhivad samadest põhimõtetest. Täpse ja kvaliteetse andmestiku kogumine ja töötlemine nõuab aega ja raha. Selleks on vaja eraldada piisavalt rahalisi vahendeid.

6) Valdkondlikke erinevusi publikatsioonide arvus ja tsiteeritavuses tuleb austada. On hea tava pakkuda välja teatud hulk bibliomeetrilisi näitajaid ja võimaldada eri valdkondadel nende seast sobivaimad välja valida. Paar aastat tagasi hindasid kohalikud eksperdid üsna kesisel Euroopa ajaloolastest koosneva uurimiserühma taset, sest nad ei avaldanud mitte artikleid *Web of Science*'is indekseeritud ajakirjades, vaid hoopis monograafiaid. Õnnetul kombel töötasid need ajaloolased psühholoogia osakonnas. Ajaloolaste ja sotsiaalteadlaste puhul

tuleb nende publikatsioonide hulka arvestada ka monograafiad ja emakeelsed kirjutised, arvutiteadlaste puhul aga ei tohi jätta kõrvale konverentsikogumike artikleid.

Tsiteeritavuse hulk on valdkonniti väga erinev: matemaatika-alaste tippajakirjade mõjufaktor on ligikaudu 3, rakubioloogia tippajakirjade puhul on see aga umbes 30. Seetõttu on vajalik lähtuda võrreldavatest näitajatest. Kõige usaldusväärsem on selleks kasutada protsentiili: iga artikli hindamise aluseks on protsentiil, mis põhineb vastava valdkonna tsiteeritavuse hulgal (näiteks kuulumine 1%, 10% või 20% sekka). Kui tsiteeritavuse keskmiste näitajate pingereas võib üks korduvalt tsiteeritud publikatsioon lennutada tabeli keskosas paikneva ülikooli lausa tippu, siis pingereas, mis põhineb protsentiilidel, parandab see publikatsioon ülikooli positsiooni minimaalselt.⁷

7) Iga teadlast tuleb hinnata kvalitatiivselt ja vastavalt tema saavutuste loendile. Mida vanem on teadlane, seda kõrgem on tema h-indeks, ja seda isegi uute artiklite puudumisel. Valdkondlikud erinevused on h-indeksi puhul suured: loodusteadlaste haripunktiks on 200, füüsikute tipuks 100 ja sotsiaalteadlaste omaks 20–30.⁸ Samuti esineb erinevusi andmebaaside lõikes: arvutiteaduste valdkonnas on teadlasi, kelle h-indeks on *Web of Science*'is ligikaudu 10, kuid Google Scholaris 20–30.⁹ Selle asemel, et hinnata teadlast ainult ühe arvu põhjal, tasub lugeda ja analüüsida tema teadustööd laiemalt. Isegi siis, kui on vaja võrrelda suurt hulka teadlasi, sobib kõige paremini selline meetod, mis pöörab rohkem tähelepanu teadlase pädevuse, kogemuste, tegevuste ja mõjuga seotud infole.

8) Oluline on vältida ebavajalikku konkreetset ja eksitavat täpsust. Teaduse ja tehnoloogia näitajatele on iseloomulik kontseptuaalne ambivalentlus, ebakindlus ja tuginemine optimistlikele oletustele, mis ei ole laiemalt aktsepteeritud. Näiteks on juba ammu seatud küsimärgi alla tsiteeringute arvu tähendus. Seega on hea tava kasutada mitmeid näitajaid, et saada usaldusväärsem ja mitmekesisem ülevaade. Kui ebakindlust ja vigade ohtu on võimalik esitada arvandmetena, kasutades näiteks veatulba funktsiooni, siis tuleb see informatsioon märkida ära näitajate avaldamisel. Kui seda teha ei saa, siis vähemalt peavad näitajate esitajad vältima eksitavat täpsust. Näiteks seda, et viigiseisu välistamiseks avaldatakse ajakirja mõjufaktor kolme komakohaga. Kontseptuaalse ambivalentluse ja tsiteeringute arvu juhusliku varieeruvuse tõttu ei ole aga mõistlik teha ajakirjadel vahet üliväikeste mõjufaktori erinevuste põhjal. Eksitavat täpsust tuleb vältida: ühest komakohast piisab.

9) Hindamisprotsessi ja näitajate süsteemset mõju ei tohi jätta tähelepanuta. Bibliomeetrilised näitajad muudavad süsteemi nende endi loodud stiimulite abil. Võimalikku mõju peaks suutma ette näha. Seetõttu tasub alati eelistada mitmete näitajate kasutamist. Üksainus näitaja toob endaga kaasa mänguhimu ja eesmärgi silmist kaotamise, mille puhul võib meetmest endast saada eesmärk. Näiteks rahastati 1990. aastatel Austraalia ülikoolides tehtavat teadustööd suuresti lähtuvalt sellest, kui palju artikleid instituudi töötajad avaldasid. Ülikoolidel oli seega võimalik eelretsenseeritava artikli „väärus“ välja arvutada. 2000. aastal oli selleks 800 Austraalia dollarit. Austraalia teadlaste avaldatud artiklite arv hakkas mõistagi kasvama, kuid need ilmusid madala tsiteeritavusega ajakirjades, sest tõenäoliselt polnud artiklite kvaliteet kuigi hea.¹⁰

10) Bibliomeetrilisi näitajaid tasub regulaarselt ajakohastada ja põhjalikult kontrollida. Teaduse ja teaduse hindamise eesmärgid on pidevas muutumises ning ühes nendega areneb ka teadussüsteem. Kunagised kasulikud näitajad muutuvad ebasobivaks ja kerkivad esile

ued. Näitajate süsteemi on vaja kontrollida ja vajadusel ka modifitseerida. 2010. aastal, pärast seda, kui oli mõistetud lihtsustatud valemi kasutamise tagajärgi, pandi Austraalias alus Austraalia tippteaduse edendamise algatusele (*Excellence in Research for Australia*), mis väärtustab teadustöö kvaliteeti.

Edasised sammud

Kui järgida neid kümmet põhimõtet, siis võib teaduse hindamise protsess hakata mängima olulist rolli nii teaduse arengu kui selle ühiskondliku mõju seisukohast. Teadusmeetria andmed annavad olulist teavet, mida oleks eksperthinnangutele tuginedes keeruline koguda või mõista. Siiski ei saa lubada juhtuda sellel, et taoline kvantitatiivne teave muutub abinõust eesmärgiks iseeneses.

Parimaid otsuseid on võimalik langetada siis, kui usaldusväärne statistika kombineeritakse arusaamaga eesmärgist ja selle teadustöö olemusest, mida parajasti hinnatakse. Oluline on nii kvantitatiivne kui kvalitatiivne tõendusmaterjal – mõlemad on omal moel objektiivsed. Teadust puudutavate otsuste tegemine peab põhinema kõrgetasemelistel protsessidel, mis võtavad arvesse kõige kvaliteetsemaid andmeid.

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).