

Leidenmanifestet för bibliometri

Diana Hicks^a, Paul Wouters^b, Ludo Waltman^b, Sarah de Rijcke^b and Ismael Rafols^{d,e}

a School of Public Policy, Georgia Institute of Technology, Atlanta, USA

b Centre for Science and Technology Studies (CWTS), University of Leiden, The Netherlands

c Ingenio (CSIC-UPV), Universitat Politècnica de València, València, Spain

d Science Policy Research Unit (SPRU), University of Sussex, Brighton, UK

e Observatoire des Science et des Techniques (OST-HCERES), Paris, France

22 april 2015

Svensk översättning av de Hicks et al. (2015) The Leiden Manifesto for research metrics. *Nature*, 520, 429-431. <http://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>.

Översatt av översättare - informationssekreterare Ann-Sofie Källund, vetenskapliga biblioteket Tritonia på uppdrag av styrgruppen (ordf. Tua Hindersson-Söderholm, Mari Katvala, Johanna Lahikainen, Anne Lehto och Eva Tolonen) för Expertnätgruppen för forskningsservice, Rådet för Finlands Universitetsbibliotek. Tack riktas till Kimmo Tuominen, Jussi Piipponen, Margareta Danielsson, Eva Höglund och Diana Hicks.

Låt de här tio principerna vägleda utvärderingen av forskningen, uppmanar Diana Hicks, Paul Wouters och deras kolleger.

Data används allt mer för att styra forskningen. Forskningsutvärderingar som tidigare utfördes som referentgranskning görs numera rutinmässigt och baseras i allt högre grad på bibliometriska analyser¹. Problemet är att utvärderingarna styrs av data, inte av omdöme. De kvantitativa mätningarna har ökat snabbt och även om den bakomliggande tanken är god, klarar man inte alltid av att utföra arbetet professionellt då metoderna och resultaten haltar. Risken finns att vi skadar systemet med just de verktyg, vars uppgift var att förbättra det, då utvärderingarna i allt högre grad görs av organisationer som inte har tillräckligt med kunskap om eller som inte fått utbildning i goda förfaringssätt och analysmetoder.

Fram till år 2000 var databasen Science Citation Index (ISI) tillgänglig endast på CD-ROM och användes enbart av experter för specialanalyser. År 2002 publicerade Thomson Reuters ett användargränssnitt, som gjorde informationen tillgänglig för en större publik via databasen Web of Science. Strax därefter skapades konkurrerande citeringsdatabaser: Elseviers Scopus (2004) och Google Scholar (beta-version år 2004). För att jämföra institutioners forskningsproduktivitet och genomslag introducerades webbaserade verktyg såsom InCites (som använder data ur Web of Science) och Scival (som använder data ur Scopus) samt verktyg med vilka man kan analysera enskilda forskare s citeringsfrekvens baserat på data i Google Scholar (Publish or Perish, år 2007).

¹ 1 Wouters, P. in *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact* (eds Cronin, B. & Sugimoto, C.) 47–66 (MIT Press, 2014).

År 2005 presenterade fysikern Jorge Hirsch, University of California, San Diego, h-indexet och populariserade en citeringsanalys för enskilda forskare. Intresset för tidskrifternas impaktfaktorer har ökat kontinuerligt efter år 1995 (se ['Impaktfaktorfixering'](#)).

På senare tid har indikatorer baserade på användning i sociala medier och nätkommentarer fått fotfäste – F1000Prime skapades år 2002, Mendeley år 2008 och Altmetric.com (som upprätthålls av MacMillan Science och ägs av Education, Nature Publishing Group) år 2011.

Som experter inom scientometri, samhällsvetenskap och forskningsadministration har vi med växande oro följt med missbruket av de bibliometriska indikatorerna. Här följer bara ett litet axplock av många dåliga exempel. Universitet runt om i världen har börjat fixera sig vid sin placering i globala rankingar (såsom Shanghai och Times Higher Education), även om sådana listor enligt oss baserar sig på felaktigt material och godtyckliga indikatorer.

En del rekryterare ber om de arbetssökandes h-index. Flera universitet baserar sina befordringsbeslut på h-indexet och på antalet artiklar publicerade i tidskrifter med hög impaktfaktor. Forskares cv:n har blivit ett sätt att skryta med dessa mätetal, speciellt inom biomedicin. Överallt ber handledare forskarstuderande att publicera i tidskrifter med hög impaktfaktor och att skaffa extern finansiering innan de är redo för det.

I Skandinavien och Kina allokera en del universitet forskningsfinansiering eller bonusar baserat på en siffra, t.ex. genom att räkna ut individuella poäng för att fördela "prestationsresurser" eller för att ge forskare bonus för att ha publicerat i tidskrifter med en impaktfaktor högre än 15²

I många fall använder forskarna och utvärderare ännu ett balanserat omdöme. Missbruket av de bibliometriska metoderna har dock spritt sig så pass att det är svårt att ignorera det.

Därför presenterar vi detta Leidenmanifest, namngett efter den konferens där det utkristalliserats (se <http://sti2014.cwts.nl/>). Manifestets tio principer är inget nytt för experter inom scientometri, även om ingen av oss kunnat lista dem tidigare då kodifiering saknats. Förgrundspersonerna inom fältet, såsom t.ex. Eugene Garfield (ISI:s grundare), har framfört några av principerna^{3 4}. Men de är inte tillgängliga när utvärderare rapporterar till universitetets administratörer, vilka å sin sida inte är experter inom den relevanta metodologin. Forskare som söker information som skulle bevisa att forskningsutvärderingen är felaktig, märker att material är splittrat i tidskrifter som är obekanta för dem och som de saknar tillgång till.

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).

Vi erbjuder den här komprimerade modellen av bästa praxis så att forskarna ska kunna ställa utvärderarna till svars och så att utvärderarna ska kunna ställa sina indikatorer till svars.

TIO PRINCIPER

1) *En kvantitativ utvärdering bör stöda en kvalitativ expertutvärdering.* Kvantitativ bibliometri kan motarbeta partiska tendenser i referentgranskning och underlätta bedömningen. Detta borde stärka referentgranskningen, eftersom det är svårt att bedöma kolleger utan tillgång till relevant information. Utvärderarna får dock inte frestas att överlåta beslutsfattandet till siffrorna. Indikatorer kan inte ersätta en välgrundad bedömning. Ansvaret för bedömningen kvarstår hos utvärderarna.

2) *Mät prestationen i förhållande till institutionens, gruppens eller forskarens forskningsuppdrag.* Programmålen bör skrivas ut tydligt i början, och de indikatorer som används för att utvärdera prestationerna bör tydligt relateras till dessa mål. I valet av indikatorer och hur de används bör mera omfattande socioekonomiska och kulturella kontexter beaktas. Forskarnas forskningsuppdrag varierar. Forskning som närmar sig akademisk spetskunskap skiljer sig från forskning som fokuserar på att ge lösningar på samhällliga problem. En utvärdering kan vara baserad på meriter som har relevans för politik, industri eller allmänheten snarare än akademisk excellens. Ingen enskild modell för utvärdering passar för alla kontexter.

3) *Värna om forskningsexcellens i forskning som har lokal betydelse.* I många delar av världen likställs forskningsexcellens med publicering på engelska. Till exempel styr den spanska lagstiftningen forskare att publicera i tidskrifter med högt inflytande, s.k. high impact-tidskrifter. Impaktfaktorn beräknas för tidskrifter som indexerats i den USA-baserade databasen Web of Science, som till största delen innehåller engelskspråkigt material. Dessa tendenser är speciellt problematiska inom samhällsvetenskaper och humaniora, där forskningen är mera regionalt och nationellt förankrad. Många andra forskningsområden har en nationell eller regional dimension - t.ex. HIV-epidemiologin i Afrika söder om Sahara.

Det finns en tendens att undertrycka denna pluralism och samhällliga betydelse för att skapa texter som intresserar bevakarna av high impact: de engelskspråkiga tidskrifterna. De spanska sociologer som citeras frekvent i Web of Science är sådana som har arbetat med abstrakta modeller eller studerar data från USA. Däremot saknas den specialkompetens som framkommer i högt rankade vetenskapliga publikationer på spanska: ämnen som lokal arbetslagstiftning, äldreomsorg inom familjen eller sysselsättning av immigranter⁵. Bibliometri baserad på högklassig icke-engelsk litteratur skulle göra det lättare att identifiera och belöna regionalt relevant forskning.

⁵ López Piñeiro, C. & Hicks, D. *Res. Eval.* 24, 78–89 (2015).

4) *Bevara datainsamlingen och de analytiska processerna öppna, transparenta och enkla.* De databaser som krävs för utvärdering bör konstrueras enligt tydliga regler som fastslagits innan forskningen har genomförts. Detta utgjorde sedvanlig praxis bland de akademiska och kommersiella grupper som byggde upp metodologin för bibliometrisk utvärdering under flera årtionden. Dessa grupper agerade i enlighet med

"Enkelhet är en dygd
vad gäller
indikatorer, eftersom
det ökar
transparensen."

protokoll publicerade i referentgranskad litteratur. Denna transparens möjliggjorde granskning. Till exempel ledde år 2010 en offentlig debatt om de tekniska egenskaperna för en viktig indikator som användes av en av våra grupper (the Centre for Science and Technology Studies at Leiden University in the Netherlands) till en revidering av hur denna indikator

beräknas⁶. Nya kommersiella aktörer borde följa samma standarder; ingen borde acceptera en utvärderingsapparat i form av en svart låda.

Enkelhet är en dygd vad gäller indikatorer, eftersom det ökar transparensen. Däremot kan förenklad bibliometri snedvrída resultatet (se princip 7). Utvärderare måste sträva efter balans - enkla indikatorer som beaktar forskningsprocessens komplexa natur.

5) *Ge de forskare som utvärderas möjligheten att verifiera data och analyser.* För att säkerställa datakvaliteten borde alla forskare som berörs av bibliometriska studier ha en möjlighet att granska att deras resultat har identifierats på ett korrekt sätt. Var och en som styr och leder en utvärderingsprocess borde säkerställa att all data är korrekt, via egenkontroll eller tredjemanskontroll. Universiteten kunde implementera detta i sina forskningsinformationssystem och det borde vara en ledande princip när man väljer systemleverantör. Insamling och behandling av exakt, högkvalitativ data kräver tid och resurser. Budgetera i enlighet med det.

6) *Ta i beaktande att det finns skillnader i praxis för publicering och citering mellan olika vetenskapsgrenar.* Det bästa tillvägagångssättet är att välja en uppsättning möjliga indikatorer och ge ämnesområdena själva möjlighet att välja bland dem. För några år sedan fick en grupp europeiska historiker en relativt låg ranking i en nationell referentgranskning. Orsaken var att de i huvudsak publicerade böcker, inte artiklar i tidskrifter som indexerades av Web of Science. Historikerna hade oturen att tillhöra en psykologisk institution. För historiker och samhällsvetare är det viktigt att böcker och litteratur på det inhemska språket räknas med i antalet publikationer; för datavetare ska antalet konferenspapper räknas⁷.

Citeringsgraden varierar från disciplin till disciplin: topprankade tidskrifter i matematik har impaktfaktorer på omkring 3; topprankade tidskrifter i cellbiologi har impaktfaktorer på omkring 30. Indikatorerna bör

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).

normaliseras, och den mest stabila normaliseringsmetoden baseras på percentiler: varje publikation vägs på basis av den percentil den tillhör inom citeringsfördelningen i disciplinen (topp-1 %, -10 % eller -20 %, till exempel). En enskild högt citerad publikation förbättrar i viss mån universitetets ranking om denna baseras på percentiler, men kan lyfta universitetet från mittskiktet till toppen om rankingen baseras på medeltal för citeringar.

7) *Basera utvärderingen av enskilda forskare på en kvalitativ bedömning av deras portfolio.* Ju äldre du är, desto högre är ditt h-index, även i avsaknad av nya publikationer. H-indexen varierar från disciplin till disciplin: forskare inom biovetenskap kan uppnå ett index på 200; fysiker 100 och samhällsvetare 20-30⁸. H-indexet är beroende av databas: det finns forskare inom datavetenskap som har ett h-index på omkring 10 i Web of Science men 20–30 i Google Scholar⁹. Att läsa och bedöma en forskares arbete är mycket mer ändamålsenligt än att förlita sig på en siffra. Även då man jämför ett stort antal forskare är det bästa tillvägagångssättet att också beakta en individs expertis, erfarenhet, arbete och inflytande.

8) *Undvik missriktad konkretisering och falsk precision.* Indikatorer inom vetenskap och teknologi är känsliga för konceptuell oklarhet och osäkerhet och kräver starka antaganden som inte är accepterade universellt. Betydelsen av antalet citeringar, till exempel, har debatterats länge. Därför är det bästa sättet att använda flera indikatorer för att ge en mera stabil och pluralistisk bild. Om osäkerhet och felaktigheter kan uttryckas till exempel med felstaplar, borde denna information ingå i de indikatorvärden som publiceras. Om detta inte är möjligt borde de som producerar indikatorer åtminstone undvika falsk precision. Till exempel uttrycks impaktfaktorn för en tidskrift med tre decimaler för att undvika att flera tidskrifter får samma impaktfaktor. Med beaktande av den konceptuella oklarheten och slumpmässiga variationen i antalet citeringar är det dock inte vettigt att göra åtskillnad mellan tidskrifter på basis av mycket små differenser i impaktfaktorer. Undvik falsk precision: endast en decimal är motiverad.

9) *Inse systemeffekterna av utvärdering och indikatorer.* Indikatorer förändrar systemet genom de incitament de skapar. Dessa effekter borde förutses. Därmed är en serie indikatorer alltid att föredra - en enskild indikator kommer att inbjuda till spel och till att målet förskjuts (mätningen i sig blir det nya målet). I Australien finansierades på 1990-talet universitetsforskning med en formel som till stora delar baserades på antalet publikationer vid universiteten. Universiteten kunde beräkna 'värdet' på en artikel i en referentgranskad tidskrift; år 2000 var det 800 australiska dollar (ca 480 amerikanska dollar) inom forskningsfinansieringen. Som man kunde förutspå ökade antalet publikationer av australiska forskare, men de återfanns i mindre citerade tidskrifter, vilket tyder på att kvaliteten på artiklarna sjönk¹⁰.

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).

10) *Granska och uppdatera indikatorerna regelbundet.* Forskningsmål och utvärderingsmål förändras och forskningssystemet självt utvecklas med dem. Bibliometriska indikatorer som tidigare varit användbara blir otillräckliga, nya indikatorer utvecklas. Indikatorsystemen måste utvärderas och eventuellt modifieras. Efter att man i Australien insett effekterna av den förenklade formeln introducerade man år 2010 ett mångsidigare system, Excellence in Research for Australia, som betonar kvalitet.

FÖLJANDE STEG

Om dessa tio principer efterföljs kan forskningsutvärdering spela en viktig roll när det gäller att utveckla vetenskapen och dess interaktion med samhället. Bibliometri kan ge central forskningskunskap som skulle vara svår att fånga eller förstå genom individuell expertis. Denna kvantitativa information får dock inte ändras från ett verktyg till ett mål.

De bästa besluten fattas när gedigen statistik kombineras med känsla för den forskning som utvärderas, dess mål och natur. Både kvantitativa och kvalitativa belägg behövs; båda är objektiva på sitt eget sätt. Beslutsfattande om vetenskap måste baseras på högkvalitativa processer med tillgång till data av högsta kvalitet.

Översatt av translator-informatör Ann-Sofie Källund, Tritonia, på uppdrag av Expertnätgruppen för forskningsservice, Rådet för Finlands universitetsbibliotek.