

Litteraturhenvisninger, anseelse og magt

Citationer spiller en større og større rolle for, hvordan penge og indflydelse fordeles i forskningsverdenen. Men det er vigtigt at holde sig svaghederne ved dette "belønningssystem" for øje.

Det er i dag en selvfølge, at videnskabelig litteratur skal indeholde bibliografiske referencer, dvs. henvisninger til den litteratur, som er benyttet i forskningsarbejdet og ved udarbejdelsen af publikationen. Deres antal i hver videnskabelig artikel er voksende, og det er deres betydning i videnskabskommunikationen også. Det er nok ukendt for mange mennesker, at studiet af litteraturhenvisninger selv er blevet genstand for omfattende undersøgelser (citationsanalyser), som hvert år resulterer i tusindvis af videnskabelige publikationer. Det er en naturlig følge af den betydning, litteraturhenvisninger har fået i systemer til forskningsevaluering – og dermed til at afgøre, hvilke forskere der får prestige og penge. Videnskabssociologiske studier af bibliografiske referencer har givet megen ny viden.

Referencer, citationer og citationsindeks

Når en forfatter afgiver en reference til et andet dokument, modtager derved forfatteren til dette en citation. At modtage en citation er en anerkendelse fra en forskerkollega, en støtte i at man er på rette vej i sin forskning. At opnå mange citationer anses for at være et tegn på høj anseelse blandt kolleger. Derfor anvender man citationstal i forsknings-evalueringssystemer. Især i USA er det almindeligt at regne med, at fastansættelse og professorater ved universiteter forudsætter et rimeligt højt citationstal. Forskere i dag er i mindre eller større grad bevidste om, at de udøver symbolsk magt gennem deres referering eller ikke-referering til bestemte dokumenter og forfattere.

Overblik over hvilke forskere, der modtager citationer, kan man kun få i citationsindeks, hvoraf den vigtigste er *Science citation index*. Det er databaser over videnskabelige artikler (og andre dokumenter) og de bibliografiske referencer, der findes i artiklerne.

Citationsanalyser

Hele grundlaget for citationsanalyser er, at forskerne refererer til alle de dokumenter, som har været af stor betydning for det konkrete forskningsarbejde, og omvendt: At de ikke refererer til dokumenter, som ikke er benyttet. At det ikke altid er i skønneste orden har flere konstateret, bl.a. Anders Smith, der forsker i magnetisk køling. Han fortalte i *Aktuel Naturvidenskab* 5-2013, at næsten alle forskere han undersøgte, citerede en artikel, de ret sikkert ikke har læst og som ovenikøbet ikke var egentlig relevant. Det er langt fra et enestående eksempel.

Citationsforvrængninger der skaber ugrundet autoritet

I 2009 publicerede dr. med. Steven Greenberg en citationsanalyse med ovennævnte titel. I analysen indsamlede han alle udsagn i den medicinske litteratur om en bestemt videnskabelig påstand (at proteinet β -amyloid spiller en rolle i sygdommen IBM (Inklusionslegeme myositis)). I sin litteratursøgning fandt han 242 dokumenter, der behandlede emnet. De relevante af disse dokumenters referencer kategoriserede han som positiv, neutral eller negativ i forhold til påstanden om proteinet. Der var 636 positive, 18 neutrale og 21 negative citationer.

Af dette citationsnetværk fremgår også, at nogle dokumenter fremtræder som meget "autoritative", idet de modtager massevis af citationer. De 10 højest citerede dokumenter var alle positive, de gav altså udtryk for at påstanden var sand. En del referencer gav sig ud for at være positive i forhold til den videnskabelige påstand, men da Greenberg efterså de dokumenter, der refereredes til, viste det sig, at disse i nogle tilfælde:

- overhovedet ikke behandler emnet, eller
- at de forvrænger indholdet af teksten og i virke-

Forfatteren



Svend Bruhns er forhenværende ass. fagleder ved Danmarks Biblioteksskole (IVA) i Aalborg. bibliografica@c.dk

Science citation index (SCI)



Eugene Garfield
– manden bag Science
Citation Index.

Foto: Bringes med tilladelse af
Eugene Garfield

I et citationsindeks kan man søge på de enkelte referencer fra artiklerne og finde, hvilke nyere artikler der har citeret dem (dvs. i hvilke andre litteraturlister, de findes). Man kan for det første se, om et bestemt dokument er blevet citeret i nogle af kilde-dokumenterne. For det andet kan man se, hvem der har citeret de dokumenter, som faktisk er citeret. For det tredje kan man tælle, hvor mange gange et bestemt dokument er blevet citeret.

Det var informationsvidenskabsmanden Eugene Garfield (f.1925), der i 1960'erne på sit eget firma *Institute for Scientific Information (ISI)* udgav det første omfattende og multidisciplinære citationsindeks, kaldet *Science citation index (SCI)*. Citationsindekserne er blevet meget mere udbredt siden ISI indførte deres *Web of Science (WoS)* – internetversionen af deres citationsindekser.

ISI blev senere købt af Thomson-Reuters. Den naturvidenskabelige SCI suppleredes senere af citationsindekser for samfundsvidenskab og humaniora. Garfields citationsindekser har leveret de nødvendige hjælpemidler til udførelse af citations-søgninger og citationsanalyser, og der foretages nu hver dag utallige citationssøgninger og -analyser verden over. En næsten jævnbyrdig konkurrent til WoS er *Scopus*, som udgives af Elsevier, der er verdens største forlag for videnskabelige tidsskrifter.

Det har været et kæmpemæssigt og kostbart arbejde at udvikle og vedligeholde disse vældige databaser. Men profitten er også stor. Det er dyrt at få adgang til at søge i dem, så i Danmark er det i praksis stort set kun på universitetsbibliotekerne, at vi kan bruge citationsindekserne.

- ligheden ikke udsiger noget positivt om emnet, eller
- at de ikke selv bringer data om emnet, men blot refererer til andre dokumenter, eller
- at de fremstiller noget som fakta, skønt det blot er hypotese.

Dette er kun nogle af de mange typer af forvrængninger, som Greenberg fandt i dette netværk af referencer/citationer. Han tager ikke stilling til, om den videnskabelige påstand, han undersøgte, er sand eller falsk, men han viser, at den dominerende stilling påstanden har fået *gennem citationer* ikke er ægte autoritet. Om denne citationsbias ("skævhed") skyldes ond vilje, eller om de citerende forskere ikke ordentligt har læst de artikler, de refererer til, og derfor blot som lemminger følger strømmen af flertallet af citationer, er uvist.

Anders Smith og Greenberg er eksempler på forskere, som sætter spørgsmålstejn ved den måde citationstal ofte indgår i forskningsevalueringer på: hver citation tæller 1, alle citationer er lige meget værd. Men Greenberg tæller ikke kun, han *læser* også de dokumenter, som referencerne går til og hvilke de ikke går til og konstaterer derved bl.a., at de få kritiske artikler er blevet ignoreret (se figur).

Citationsproblemer

Der er indbygget skævhed i citationsantal af mange grunde, bl.a. er de i høj grad en funktion af, om publikationen er et kildetidsskrift i *Science Citation Index* eller ej. Der er en høj grad af selvreference i de fleste tidsskrifter. Der er også store forskelle disciplinerne imellem. Fx vil en artikel i molekylærbiologi i gennemsnit modtage 10 gange så mange citationer som en i matematik. Statistisk vil en artikel af flere forfattere opnå højere citationstal, hvis der

er en USA-forfatter med, end en ligeså god artikel, som ikke har en USA-forfatter med.

Selvom et højt citationstal må anses for at være en god indikator for god forskning, kan man ikke sætte lighedstegn mellem citationstal og kvalitet. Bl.a. af ovennævnte grunde samt af den grund, at der er forekommet (relativt sjældne) tilfælde af højtciteret svindel. Husk også, at man ud fra en citationsanalyse intet kan sige om kvaliteten af et ikke-citeret dokument.

USA og engelsk sprog dominerer i høj grad indholdet af de mest udbredte citationsindekser (*WoS* og *Scopus*). Det er forståeligt, da USA er verdens førende forskningsnation, og engelsk er videnskabsens internationale sprog. Men det bevirker, at andre sprog og lande, også store forskningsnationer som Frankrig og Tyskland, bliver underrepræsenteret, også med hensyn til antal citationer.

Ikke desto mindre indgår citationstal i mange forskningsevalueringssystemer.

Citationer og forskningsevaluering

Der har i mange år været foretaget forskningsevaluering ved stillingsbesættelser, antagelse af manuskripter til publicering osv. Det relativt nye er, at evalueringen bliver et instrument i forskningsstyring og -strategier. Nutidens forskningsadministratorer ønsker sig helst kvantitative metoder, angiveligt fordi metoderne er "objektive". Måske også fordi de tror, at man ikke behøver at have noget personligt kendskab til forskningsområdet, som er under vurdering, hvis man blot kan aflæse nogle kurver og tal. Men der er mange metodologiske udfordringer og faldgruber ved citationsanalyser, fx hvordan man skal "normalisere" materialet for at kompensere for

skævheder. Det er egentlig kun professionelle bibliometrikere, der kan vurdere materialet.

I Danmark indførtes fra 2010 efter norsk mønster Den bibliometriske forskningsindikator (BFI), som udelukkende tæller publikationer i godkendte publikationskanaler, dvs. artikler i videnskabelige tidsskrifter eller bøger på videnskabelige forlag. Det er godt, at det er de danske forskere, der selv bestemmer hvilke publikationskanaler, der kan godkendes og hvilke af dem, der udgør den femtedel, der udløser højere pointtal. BFI er nok det bedste system, når alle forskningsdiscipliner skal dækkes i ét system. Der har dog hele tiden været røster fremme, som talte for, at citationstal også skulle tælle med i forskningsevaluering i Danmark. I de internationale "rankings" af forskningsinstitutioner indgår som regel også deres opnåede citationstal. Men det ville være ødelæggende for balancen i de danske forskningsbevillinger, hvis systemet skulle baseres på citationstal.

Citationsperspektiver

Citationsindeksering og -analyse breder sig mere og mere inden for og uden for forskningsverdenen. Forskerne selv bliver også mere og mere *citationsbevidste*. Bl.a. ved måltrettet at stræbe efter at placere deres bedste manuskripter i tidsskrifter, der gennemsnitligt opnår højere citationstal. Hvilket nogle gange bevirker, at de mest relevante fagtidsskrifter ikke modtager disse manuskripter. På længere sigt vil det formodentlig accelerere tidsskriftsdøden. Det er i forvejen op ad bakke for mange nordiske videnskabelige tidsskrifter.

Forholdet mellem antal videnskabelige publikationer og antal citationer er ikke fuldstændigt klarlagt. Mange undersøgelser har vist, at lande med stor produktion af artikler også har opnået mange citationer. Danmark er et af de lande i verden, der opnår størst antal citationer pr. artikel i gennemsnit. Det har imidlertid vist sig, at lande, der har givet forskerne incitamenter til at forøge produktionen af artikler, ofte må konstatere, at citationstallene pr. artikel falder, hvilket ofte tolkes som at kvaliteten falder. Dette har fx været tilfældet i Norge i de seneste år. Norge indførte allerede i 2004 den BFI-model, som Danmark næsten uændret kopierede fra 2010. Det bliver interessant at se, om det samme vil ske i Danmark.

Citationsindekserne udvikles og forbedres hele tiden. Indtil videre er der meget dårlig repræsentation af den ikke-vestlige verden i dem, men for få år siden er et kinesisk citationsindeks indlemmet i *Wos*, og flere andre naturvidenskabelige, bl.a. et russisk, er startet. Der er imidlertid meget lang vej igen, inden de nu kendte citationsindekser har en rimelig dækning af ikke-angloamerikansk videnskab. Jeg er bange for, at firmaerne bag citationsindekserne i virkeligheden

ikke er interesserede i at udvide dækningen. De har altid været stolte af, at de med et snævert udvalg af kildetidsskrifter har så stor procentdel af vigtige tidsskrifter og bekymrer sig næppe over, at de ikke medtager tidsskrifter, som man i et tilfældigt europæisk land finder meget vigtige.

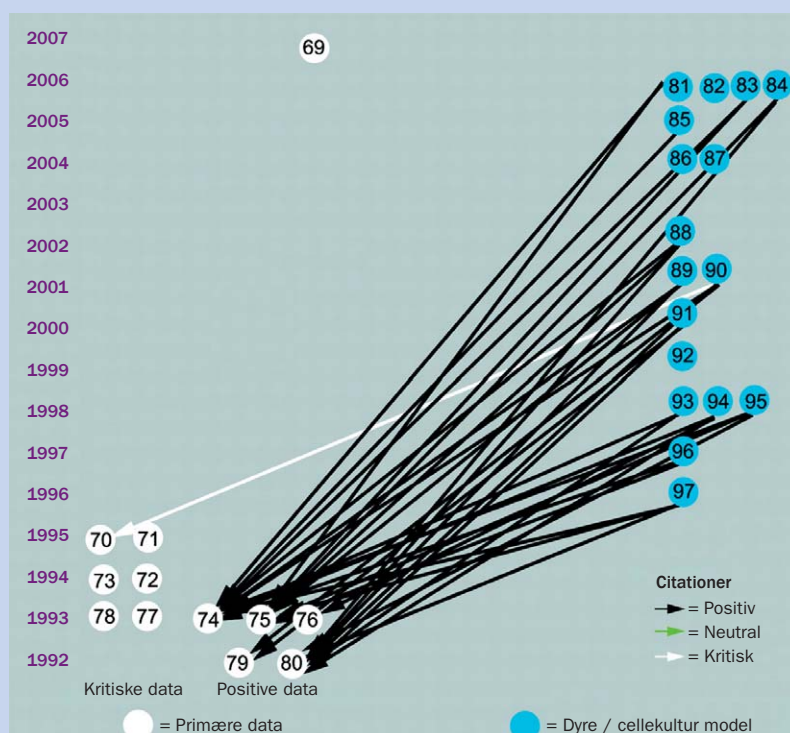
Det er især inden for humaniora og samfundsvidenskab, at dækningen af ikke-angloamerikansk videnskab er – for at sige det rent ud – nærmest til grin. The European Science Foundation (ESF) har i nogle år arbejdet på at skabe et europæisk "referenceindeks" for humaniora, som også skulle give mulighed for citationssøgning. Jeg håber, det lykkes, alene af den grund, at citationssøgning er et vigtigt supplement til (men ikke erstatning for) almindelig bibliografisk søgning.

Der er sikkert også en lysende fremtid for ikke-kommercielle citationsindekser som *Google Scholar*, der udnytter Internettet til citationssøgning. Alt tyder på at interessen for citationer kun vil vokse i de kommende år. Men vi får også brug for flere undersøgelser à la Anders Smiths og Greenbergs, der dokumenterer nogle af de svagheder, der er ved citationssøgning. ■

Videre læsning

Bruhns S. (2010). Citationsindeksering. Med et kapitel: Kortlægning og visualisering af Jesper W. Schneider. Bibliografica, Romdrup.

Greenberg, S. A. (2009), "How citation distortions create unfounded authority: analysis of a citation network", *bmj* : Online first, vol. 339, p.1-14 + 121 supplementary pp online. www.bmj.com/content/339/bmj.b2680



Figuren stammer fra Steven Greenbergs citationsanalyse omkring den videnskabelige påstand, at proteinet β -amyloid spiller en rolle i sygdommen IBM (Inklusionslegeme myositis). Her betragter Greenberg en gruppe af de citerende artikler (nr. 81-97) og konstaterer, at alle deres 32 referencer (pilene) i netværket går til positive dokumenter (nr. 74 m.fl.) bortset fra en enkelt, der går til et kritisk dokument (nr. 70).

En relevant tolkning af figuren er, at den viser en blanding af bevidst strategisk adfærd og en "lemming-effekt": en del af forskerne refererer simpelthen til de dokumenter, som i forvejen har opnået mange citationer. Herved opnår disse dokumenter – især nr. 74 og 80 – uberettiget mange citationer.