

Riitta Kärki

Yhteisviittausanalyysin taustaa ja ongelmia

Kärki, Riitta, Yhteisviittausanalyysin taustaa ja ongelmia [On the background and some problems of cocitation analysis]. *Kirjastotiede ja informatiikka* 12 (2): 53–62, 1993.

The theoretical background and some methodological problems of citation studies are discussed. Although most citation studies are based on Robert K. Merton's work, also Thomas S. Kuhn's description of paradigms and some of the ideas developed within constructivist sociology are usable as frame of reference. Author cocitation analysis (ACA), is described as a special case of bibliometrics. Its limitations and validity are discussed. Usually the results of the ACA studies are represented in graphic displays, often called maps, produced by a multi dimensional scaling program. In the article, attention is paid to the benefits obtained when drawing the maps manually. All types of data (raw data, correlations, and data based on a factor or cluster analysis) can be considered simultaneously, enabling the researcher to construct a full-length portrait of the field covered by the study.

Address: University of Tampere, Department of Information Studies, P.O.Box 607, SF-33101 Tampere, Finland.

Informaatiotutkimuksessa¹ varsin laajalti harrastettu ja alan omimpana pidetty tutkimusmenetelmä on viiteanalyysi. Jo vuonna 1980 oli tiedossa yli 2000 viiteanalyyttistä tutkimusta eikä niiden määrä näytä olevan ollenkaan vähenemään päin (MacRoberts & MacRoberts 1989, 342). Viiteanalyysistä, kuten bibliometrisistä menetelmistä yleisemminkin, kehitellään jatkuvasti uusia muunnelmia. Maailmalla, joskaan ilmeisesti ei Suomessa², yksi erityisen kiinnostuksen kohde on yhteisviittausanalyysi (cocitation analysis). Sitä käyttäen laaditaan erityistieteistä tai tutkimusalueista karttoja, joiden on määrä kuvata näiden alojen rakennetta. Menetelmää on kehitelty 1970-luvulla harrastetusta julkaisuihin kohdistuvasta analyysistä 1980-luvulla yleistyneeksi kirjoittajiin kohdistuvaksi tutkimukseksi ja edelleen yhdistelmäksi, jossa tarkas-

tellaan tiettyjen avainsanojen ja viittausten esiintymistä samassa julkaisussa.

Tarkastelen aluksi yleisesti bibliometrisen tutkimuksen teoreettista taustaa sekä joitakin metodisia ja teknisiä ongelmia. Tämän jälkeen kuvailen yksityiskohtaisemmin yhteisviittausanalyysia. Sovelsin sitä työssä, jossa selvitetään, missä määrin ja millä tavoin tieteellisen viestinnän tutkimus yhdistää tieteen sosiologiaa ja informaatiotutkimusta. Tieteellisen viestinnän tutkimusta on luontevaa ajatella näiden erityistieteiden yhteiseksi alueeksi, mutta alan tutkijoiden yhdessä saamia viittauksia analysoimalla voidaan valaista sitä, miten intensiivinen suhde on kyseessä. Tässä artikkelissa en paljasta tuloksia, vaan tarkastelen yhteisviittausanalyysin periaatteita, toteutusta ja rajoituksia. Tavoitteena on tarjota aineksia menetelmän luotettavuuden ja

tarkoituksenmukaisuuden arvioinnille.

Perusta tieteesosiologiassa

Vaikka informaatiotieteilijät saattavat tulkita bibliometrian omaksi saavutukseksi, se on yhtä lailla tieteesosiologian menetelmä. Robert K. Mertonin työstä suuri osa käsittelee tieteellistä kirjallisuutta; mm. tutkimukset tieteellisestä julkaisu-järjestelmästä ja prioriteettikiistoista ovat tärkeä osa hänen työstään tieteesosiologina. Mertonin työ on ollut perustana lukuisille merkittävälle sosiologisille tutkimuksille, joissa on käytetty bibliometristä aineistoa. Esimerkkeinä mainittakoon Warren Hagstromin *The Scientific Community*, Jonathan ja Stephen Colen tutkimukset tieteen sosiaalisesta kerrostumisesta, Nicholas Mullinsin verkostanalyysi teoriaryhmistä sosiologiassa sekä Harriet Zuckermanin tutkimus Nobel-palkinnon saajista. (Pierce 1990, 54.)

Mertonin merkitys ei kuitenkaan rajoitu siihen, että hän on käyttänyt bibliometristä aineistoa ja analysoinut tieteellistä julkaisu-järjestelmää. Hänen luomansa malli sisältää myös käsityksen siitä, miten tieteellinen tutkimus toimii, so. kuuluisan tieteen normatiivisen rakenteen kuvauksen. Siinä tiede määritellään yhteiseksi ja julkiseksi. Tutkimustulosten julkaiseminen merkitsee toisaalta tutkijan antia yhteisen tiedon kasvattamiseen, toisaalta sen varmistamista, että kyseinen työ tunnustetaan tutkijan älylliseksi omaisuudeksi. Toisten tutkijoiden töihin viittaaminen on tärkeä osa tätä järjestelmää: siinä maksetaan älyllisiä velkoja, annetaan muille tunnustusta ja noudatetaan yhtä tieteellisen julkaisuuden etikettisääntöä. Mertonilaisen lähestymistavan puitteissa voidaan arvioida tieteellistä tuotteliaisuutta julkaisujen määrän perusteella ja julkaisujen vaikutusta viittausten määrän perusteella. (Mt., 54; Luukkonen 1990, 7.)

Mertonin sosiologia ei kuitenkaan ole ainoa mahdollinen viitekehys bibliometrialle, vaan se sopii yhteen myös Thomas S. Kuhnin paradigmallin kanssa. Viiteanalyysia voidaan käyttää tiedeyhteisöjen tai erityistieteiden identifioinnissa sekä tieteiden kognitiivisten rakenteiden analysissa. Tiedeyhteisöjen yleisiä tunnusmerkkejä ovat sen jäsenille yhteinen aihepiiri, tutkijoiden jäsenyys tieteellisissä seuroissa, tieteelliset aikakauslehdet, tiettyihin konferensseihin osallistuminen sekä formaalit ja informaalit viestintäverkostot, mukaan luettuina ne, jotka käyvät ilmi kirjeenvaihdosta ja

viittausverkostoista (Kuhn 1970, 177-178). Viiteanalyttisen tutkimuksen tulokset ovat yleensä tukeneet Kuhnin näkemystä, että tutkijat kommunikoivat suhteellisen pienen kollegajoukon kanssa, joka muodostaa tiedeyhteisön ja jonka mielipiteet ohjaavat tutkimusta. (Paisley 1990, 284.)

Kuhnin paradigmakäsitteellä on monta merkitystä. Yksi niistä on käsitys, jonka mukaan paradigma on esikuva, ts. tietty tai tietynlainen tutkimus, joka toimii mallina kaikelle myöhemmälle tutkimukselle. Tässä konkreettisessa merkityksessä merkityksessä paradigma tarkoittaa esim. oppikirjoja ja klassikkoja (Verronen 1986, 55). Samalla tavalla hyvin runsaasti viittauksia saaneiden julkaisujen voidaan ajatella edustavan alansa paradigmoja.

... Kuhnin ajatukset ovat vahvasti vaikuttaneet siihen tieteesosiologian suuntaukseen, jossa tieteellisen tiedon ajatellaan olevan lähinnä neuvonpidon tulos, tutkijoiden luoma sosiaalinen konstruktio (Mulkay 1977, 238; Ben-David 1978, 204-207). Konstruktivistisesta näkökulmasta dokumentit ovat pelkkää dataa ja pysyvätkin sellaisena, kunnes yksilö antaa niille merkityksen (Lievrouw 1990, 68). Ensi alkuun vaikuttaa mahdolltomalta yhdistää tätä ajattelutapaa viiteanalyttiseen tutkimukseen. Yhdistelmä ei kuitenkaan ole looginen mahdollisuus, sillä oleellista ei ole se, millaista tutkimusaineistoa käytetään, vaan se, missä viitekehyksessä tulokset tulkitaan:

"...John Van Maanen --- huomauttaa, että 'maasto ei ole yhtä kuin kartta'. Voimme lisätä, ettei tieteellisen viestinnäkään 'maastoa' löydetä koskemattomana bibliometrian kartoista. Kuitenkin tiedämme myös, ettei maastossa voida helposti suunnistaa ilman karttaa ja että kartta on merkityksetön, ellei matkaaja osaa tulkita sitä. Sekä bibliografinen kartta että sen etnografinen tulkinta on välttämätön liikuttaessa läpi tieteellisen viestinnän maaston." (Lievrouw 1990, 69; suom. RK.)

Konstruktivistisesta näkökulmasta viittaaminen ja viittamatta jättäminen voidaan katsoa osaksi tieteen retoriikkaa. Tieteen retorisen järjestelmän on todettu muodostavan perusedellytykset viittaamiselle: viitteillä on perustavanlaatuinen tehtävä argumentoinnissa ja ne liittyvät ideoiden tekemiseen 'relevanteiksi', 'loogisesti johdetuiksi' ja 'vakuuttaviksi' (Luukkonen 1990, 14). Tutkija saattaa esim. kirjoittaa "luotettavimmat tulokset saavutti Jones...", koska Jones sattuu kuulumaan samaan koulukuntaan, tai "yleisesti on tiedossa, että..." koska ei viitsinyt tarkistaa asiaa mistään

lähteestä (Gilbert & Mulkay 1984, 176-177). Oma erityistapauksensa on opinnäytejärjestelmä, jossa viittaamiseen vaikuttaa mm. työn ohjaaja sekä se, että opinnäytteen tekijän odotetaan osoittavan lukeneisuutta (Rautio & Suhonen 1981, 5-6). Tässä mielessä viittaaminen heijastaakin tieteen sosiaalisia rakenteita.

Viiteanalyysin kritiikkiä

Sitaatti-indekseihin perustuvan viiteanalyysin puutteista on keskusteltu siitä lähtien, kun sitä on harrastettu 1960-luvun alusta alkaen. Nykyiset menetelmän puoltajat väittävät tuntevansa sen rajoitukset, kun taas vastustajat väittävät, että puoltajat eivät yksistään vähättele viiteanalyysin ongelmia vaan suorastaan epäonnistuvat niiden tunnistamisessa. (MacRoberts & MacRoberts 1989, 342.)

Bibliometrinen menetelmien kritiikki ulottuu niiden epistemologiseen perustaan saakka. David Edge on kirjoittanut: "Tarjoamalla 'kokonaisobjektiivisen' kuvan tieteestä 'sellaisena kuin se on' nuo menetelmät sisältävät tausta-ajatuksena epätarkoituksenmukaisen positivistisen ja realistisen lähestymistavan ja kääntävät huomion pois tutkijoiden omasta dynaamisesta näkökulmasta: kvantitatiivisiin menetelmiin johdetut selitykset on tarkoitettu 'korjaamaan' asianosaisten 'puolueelliset' havainnot." (Edge 1979, 108-109, suom. RK.)

Edgen sävy sekä tapa käyttää lainausmerkkejä tiettyjen avainsanojen ympärillä saattaa johdattaa luulemaan, että viiteanalyytikot todella olisivat itse väittäneet jotain sellaista, mitä hän heihin yhdistää. Itse asiassa hänen omat lainauksensa viiteanalyytikoiden teksteistä eivät osoita mitään tämänkaltaista. Kukaan ei väitä korjaavansa virheellisiä käsityksiä, vaan pyrkimyksenä on rikastuttaa muilla menetelmillä saatua kuvaa ja esittää todisteita ilmiöistä, jotka ovat saattaneet jäädä huomaamatta esim. historiallisissa tutkimuksissa. (White 1990, 94.)

Koska Edge antaa harhaanjohtavan kuvan viiteanalyytikoiden lähtökohdista ja tavoitteista, ei ole liioin syytä hyväksyä hänen päätelmänsä viiteanalyytisen tutkimuksen tieteenteoreettisesta perustasta. Viiteanalyytikot saattavat olla realisteja, mutta se ei ole välttämätön seuraus siitä, että he tekevät viiteanalyysia. Normatiivinen teoria kylläkin olettaa tieteellisen tiedon kumulatiivisen kasautumisen todeksi ja siihen nojautuva viiteanalyysi

edustaa realismia, mutta kuten edellä tuli ilmi, viiteanalyyttistä tutkimusta voidaan tehdä muunkinlaisessa viitekehyksessä.

Kaikkein useimmin viiteanalyysin kritiikki (esim. Weinstock 1971) kohdistuu siihen, että viittaamisella on lukuisia erilaisia funktioita, joita pelkkä viittausten laskeminen ei voi paljastaa. Läheskään kaikenlaisessa viiteanalyytisessä tutkimuksessa ei kuitenkaan ole suurtakaan merkitystä sillä seikalla, että osa esim. Thomas S. Kuhniin kohdistuvista viittauksista tehdään vain muodon vuoksi, rituaalinomaisesti tai negatiivisessa tarkoituksessa. Merkitystä on sillä, että esim. tieteesosiologiassa julkaisuissa häneen yleensä on viitattu enemmän kuin kehenkään muuhun ja että hän on saanut varsin runsaasti yhteisviittauksia. Vaikka kaikki Kuhniin kohdistuvat viittaukset olisivat kriittisiä tai kielteisiä, hän osoittautuisi silti bibliometrisen evidenssin valossa erittäin tärkeäksi hahmoksi (White 1990, 90).

Vakavamman ongelman muodostaa se bibliometrian perustavanlaatuinen olettamus, itse asiassa uskomus, että tutkijat todella viittaavat niihin teoksiin, joilla on ollut merkitystä heidän työlleen. Erityisesti klassikoihin saatetaan jättää viittaamatta, koska teosten sisällöstä on tullut omalla alallaan yleissivistystä. Michael ja Barbara MacRobertsin suorittamassa testissä ilmeni, että kirjoittajat viittaavat vain noin kolmannekseen niistä julkaisuista, jotka heihin ovat vaikuttaneet. Suurin osa kirjoittajista yksinkertaisesti ei viittaa auktoriteetteihinsa, perusteellisinkin kirjoittaja MacRobertsien otoksessa viittasi vain 64 prosenttiin vaikuttajistaan. (MacRoberts & MacRoberts 1989, 343.)

Se, että bibliometrinen tutkimusten aineistot yleensä kerätään Science Citation Indexistä (SCI) tai Social Sciences Citation Indexistä (SSCI), aiheuttaa sekä teknisiä ongelmia että tulosten systemaattista vinoutumista. Kun teoksella on monta kirjoittajaa, haussa löytyy vain ensimmäinen tekijä. Tietoa ei ole siitä, kuinka paljon tämä vaikuttaa tutkimustuloksiin. Tietokannoissa esiintyy saman henkilön nimi monissa muodoissa ja henkilö saattaa myös vaihtaa sukunimeään. Toisaalta kahdella eri henkilöllä voi olla sama nimi tai vähintään sama sukunimi ja etunimien alkukirjaimet. Nämä seikat on yritetty ottaa viiteanalyysissa huomioon, mutta mitään testituloksia ei ole siitä, miten perinpohjaisesti se on onnistuttu tekemään. SCI ja SSCI kattavat noin 10 % edustamistaan tieteenaloista. Englanninkieliset lehdet ja länsimainen tiede ovat yliedustettuina, kun taas pienet maat ja ei-länsi-

maat sekä muilla kuin latinalaisilla aakkosilla kirjoitetut tekstit ovat aliedustettuina. Myös eri tieteenhaarat on katettu erisuuruisesti. (Mt., 346.)

Yhteisviittausanalyysin periaatteet

Yhteisviittausanalyysin idea on lähtöisin Eugene Garfieldilta ja sittemmin tätä menetelmää ovat soveltaneet ja kehittäneet ennen muuta Henry Small, Bolver Griffith ja Howard White. Yhteisviittausanalyysin peruserä on, että mitä useammin tutkijat viittaavat kirjoituksissaan kahteen samaan julkaisuun (tai samoihin kahteen tekijään), sitä todennäköisempää on, että näiden kahden viitattun julkaisun välillä vallitsee jokin yhteys. Tämän yhteyden katsotaan merkitsevän useimmiten sitä, että viitattut julkaisut ovat aihepiiriltään samankaltaisia. Mitä enemmän julkaisuihin viitataan yhdessä, sitä läheisempi niiden suhde on. (Bellardo 1980, 231; White 1990, 84.) Tutkimalla runsaasti yhdessä viitattujen julkaisujen verkostoja voidaan luoda rakenteellisia esityksiä laajoista tieteen sektoreista, vaikka analyysiyksiköitä on melko vähän (Hargens 1978, 129).

Aluksi yhteisviittausanalyyssejä tehtiin tietyistä julkaisuista, mutta näissä tutkimuksissa oli enemmän teknisiä ongelmia ja suurempia aineiston kokoon liittyviä rajoituksia kuin sittemmin yleistyneessä kirjoittajien yhteisviittausanalyysissä (author cocitation analysis). Tutkimalla kirjoittajiin kohdistuneita yhteisviittauksia saadaan mukaan ko. kirjoittajien koko tuotanto eikä vain ennalta valittuja dokumentteja.

Kirjoittajien yhteisviittausanalyysissä ei tutkita nimiä tai henkilöitä, vaan sitä, mitä nimet edustavat jollakin alalla. Analyysi ei välttämättä tuo esiin yksittäisen tutkijan työn kaikkia puolia, koska tutkija on saattanut toimia erilaisilla tutkimus- tai tieteenaloilla. Esim. Derek J. de Solla Priceen tieteenhistorialliset kirjoitukset tuskin saavat yhteisviittauksia jonkun informaatiotieteilijän kirjoitusten kanssa, mutta hänen informaatiotieteelliset kirjoituksensa saavat. Kun nimistä tehdään nimipareja, saadaan aikaan kahden kontribuution leikkaus, joka edustaa tiettyä tutkimushaaraa. On tärkeää huomata, että nimenomaan kahden nimen yhdistelmä muodostaa tuon tutkimushaaran: kukin tutkija on kirjoittanut monesta aiheesta, joten jokainen nimi yksinään edustaa jotain laajempaa kuin kahden nimen yhdistelmä. (White & Griffith 1981, 163-164.)

Kirjoittajien väliset yhteydet, joita analyysissä löydetään, eivät välttämättä ole erityisen suoria. Esim. viiteanalyysia tarkastelevassa metodisessa julkaisussa saattavat saada yhteisviittauksia Robert K. Merton, Thomas S. Kuhn, Howard D. White sekä Bolver C. Griffith. Kaikki neljä eivät suinkaan ole kirjoittaneet samasta aiheesta, sillä kaksi ensimmäistä ovat tietensosiologian keskeisiä teoreetikkoja ja jälkimmäiset bibliometrinen menetelmien kehittäjiä. Sisällöllinen yhteys näiden kahden parin välille syntyy siitä, että viiteanalyysia arvioidaan tieteen tutkimuksen perusasettamusten valossa. Yhteisviittausanalyysin tulosten luotettava tulkinta edellyttää tämän seikan tietämistä; analyytikon on siis tunnettava tutkittava ala.

Kahteen henkilöön voidaan yhdessä julkaisussa tietenkin viitata monesta eri syystä. Esim. yhteisviittauksia saaneet tutkijat ovat voineet osallistua saman teorian kehittelyyn tai he saattavat edustaa täysin vastakkaisia teoreettisia näkemyksiä. Joka tapauksessa viittausanalyysi perustuu siihen, että kun tutkijat julkaisuissaan viittaavat toisiin tutkijoihin, he samalla toistavat vanhoja ja luovat uusia linkejä aiempien tieteellisten saavutusten välille. (Bayer & Smart & McLaughlin, 1990, 444-445.)

Yhteisviittausanalyysi on tietysti täsmälleen yhtä onnistunut kuin on tutkimukseen mukaan otettavien kirjoittajien valinta. Kirjoittajalistoilta on taipumus vinoutua asemansa vakiinnuttaneiden kirjoittajien suuntaan eivätkä ne koskaan voi sisältää kokonaista tieteen- tai tutkimusalaa, mutta niihin pyritään sisällyttämään kaikki tunnetut, relevantit nimet. (Esim. White & Griffith 1981, 164.)

Yhteisviittausanalyysin toteutus

Kirjoittajien valinta

Yhteisviittausanalyysi aloitetaan keräämällä joukko sopivia nimiä, jotka edustavat tutkittavaa aihepiiriä. Omassa tutkimuksessani oli tarkoitus löytää keskeisimmät tietensosiologit ja informaatiotieteilijät, jotka joko ovat tutkineet tieteellistä viestintää tai ovat alallaan sikäli merkittäviä, että heidän voi olettaa vaikuttaneen myös tieteellisen viestinnän tutkimukseen. Tieteellisen viestinnän tutkimus määritettiin seuraavien aihepiirien tutkimukseksi: tieteellisen julkaisutoiminnan ja kirjallisuuden rakenteen tutkimus, tieteellisen informaation leviämisen ja kirjallisuuden käytön tut-

kimus, viestintäverkot (communication networks), tutkijoiden epäviralliset mutta suhteellisen vakiintuneet kontaktiverkot (invisible college) sekä epävirallinen keskinäisviestintä.

Analyysiin mukaan otettavien kirjoittajien valintakriteerit voidaan tiivistää seuraavasti: 1. Tieteellistä viestintää tutkineet informaatiotieteilijät. 2. Tieteellistä viestintää tutkineet tieteesosiologit. 3. Muita tieteellisen viestinnän tutkimukseen keskeisesti vaikuttaneita henkilöitä (esim. John Bernal). 4. Henkilöitä, jotka ovat vaikuttaneet tieteen tutkimukseen niin vahvasti, ettei heitä voi sivuuttaa, vaikka he eivät suoranaisesti kuulu mihinkään kolmesta edellisestä ryhmästä, esim. Thomas S. Kuhn. Viimeksi mainittu valintaperuste saattaa ensisilmäyksellä vaikuttaa liian epämääräiseltä, mutta kyseessä olevan alueen kirjallisuutta tuntevan ei ole erityisen vaikea sitä soveltaa.

Tutkijajoukon etsintää edelsi melko laaja tutkiminen alan kirjallisuuteen, joten lähtökohtana oli oma alan tuntemus. Lisäksi käytiin ensyklopedioita, käsikirjoja ja katsauskirjallisuutta. Tein myös informaatiotutkimuksen ja sosiologian Dialog-tietokannoista viitehakuja; niiden ensisijainen tarkoitus oli varmistaa, että mukaan saadaan mahdolliset muissa lähteissä mainitsematta jääneet tieteellisen viestinnän tutkijat. Viimeiset lisäykset listaan tein käyttämällä Whiten ja Griffithin vastaavia yhteisviittausanalyyssejä informaatiotieteestä (1981) ja tieteen tutkimuksesta (1982). Heidän tutkimuksensa informaatiotieteilijät jakautuivat neljään ryhmään. Valitsin omalle listalleni kaikki ryhmään 'tieteellinen viestintä' kuuluvat sekä keskushahmot ryhmistä 'bibliometria' ja 'generalistit'. Tieteen tutkijoista otin mukaan kaikki viestinnän tutkijat sekä sellaiset erityistieteiden rakenteiden ja tieteen palkkiojärjestelmien tutkijat, joilla osoittautui olevan selvä vaikutus myös viestinnän tutkijoiden ryhmään³.

Nimiä kertyi aluksi 55. Joukossa oli kirjoittajia -muitakin kuin kaikkein tunnetuimpia ja merkittävimpinä pidettyjä - eri maista ja he edustivat erilaisia suuntauksia. Tässä vaiheessa oli jo karsiutunut pois sellaisia, jotka aihepiiriin kannalta olisivat kuuluneet joukkoon, mutta joita oli turha ottaa mukaan SSCI:n aineistoon perustuvaan analyysiin, koska heitä ei ko. indeksissä juuri näy. Tästä syystä mukaan ei esim. otettu yhtään suomalaista.

Tein SSCI:stä hakuja, joissa jokainen nimi oli parina jokaisen muun kanssa ja kone laski näiden pariin saamat yhteisviittaukset. Tässä vaiheessa listalta putosi pois muutamia henkilöitä, koska

heihin oli yleensä viitattu niin vähän, etteivät olisi saaneet monia yhteisviittauksiakaan. Kirjoittajat, joihin yksinään oli viitattu vähemmän kuin 45 kertaa, jätettiin analyysin ulkopuolelle. Tämä valinta oli aineistolähtöinen: kukaan ei ollut saanut 20-44 viittausta ja alle 20 viittausta saaneet luultavasti olisivat joka tapauksessa jääneet pois jatkossa käyttämäni karsintaperusteiden vuoksi.

Kun kaikkien kirjoittajaparien keskinäiset viittaukset on saatu lasketuksi, kaikkia ei vieläkään kannata ottaa jatkoanalyysin kohteeksi, koska liian pienet yhteisviittausmäärät antavat vaikeasti tulkittavia tuloksia. Karsintakriteerinä on käytetty esim. arvon yhdeksän ylittävää yhteisviittauksen keskiarvoa tai yhteisviittauksia vähintään kolmasosan kanssa kaikista mukana olevista kirjoittajista. Erään testin mukaan tyydyttävä validisuus ja tulkittavuus saavutetaan, jos kirjoittajia on 50 ja neljän tai viiden vuoden periodin kattavassa analyysissä jokaisen kirjoittajan saamien yhteisviittauksen keskiarvo on vähintään neljä. (McCain 1990, 435.) Omassa tutkimuksessani karsintarajat olivat seuraavat: yhteisviittauksia kaikkien muiden kanssa piti olla keskimäärin vähintään yhdeksän ja vähintään kolmasosan kanssa piti olla vähintään yksi yhteisviittaus. Jos toinen kriteereistä jäi täytymättä, kirjoittaja karsittiin lopullisesta aineistosta. Tämän karsinnan tuloksena syntynyt lopullinen lista sisälsi 41 nimeä.

Matriisit

Yhteisviittauksen määrät syötetään symmetriseen matriisiin, jossa jokainen kirjoittaja on sekä rivillä että palstalla ja täten myös jokainen luku esiintyy siinä kahteen kertaan. Keskellä, diagonaalisesti ylhäältä vasemmalta alas oikealle, on puuttuvien arvojen jono. (Kuva 1.)

	Allen	Andrews	Barber	Ben-David
Allen		26	8	14
Andrews	26		12	11
Barber	8	12		77
Ben-David	14	11	77	

Kuva 1. Esimerkki kirjoittajien saamien yhteisviittauksen määriä kuvaavasta matriisista.

Yhteisviittausanalyysissä yleensä katsotaan, miten eri henkilöiden saamien yhteisviittauksen

määrät korreloivat keskenään. Korrelaatioiden laskemisen perusteena on ensinnäkin se, että saadaan tasoitetuksi mittakaavaerot. Joku on voinut saada yleisesti paljon yhteisviittauksia ja joku toinen vain vähän, joten näitä kahta on vaikea vertailla. Toiseksi korrelaatioilla ilmaistaan tutkijoiden profiilien samankaltaisuutta ja erilaisuutta: korkea positiivinen korrelaatio merkitsee, että kun tutkija A saa paljon yhteisviittauksia tutkijan 1 kanssa, niin samoin saa tutkija B, ja kun tutkija B saa vähän yhteisviittauksia tutkijan 2 kanssa, niin samoin saa tutkija A. Mitä yhdenmukaisemmat kahden viitatus tutkijan korrelaatioarvot ovat, sitä yhteneväisemmät ovat myös heidän profiilinsa ja sitä läheisemmin he kuuluvat yhteen. (McCain 1990, 435-436.)

Matriisin diagonaaliset puuttuvat arvot ovat tuotaneet jonkin verran päänvaivaa korrelaatioita laskettaessa. Tyhjään paikkaan on kokeiltu sijoittaa kunkin kirjoittajan yksin saamien viittausten yhteismäärä, mutta luku saattaa olla huomattavan paljon suurempi kuin yksikään yhteisviittausten määrä ja aiheuttaa tuloksiin liian suuren virheen. Myös nollan on osoitettu tuottavan liikaa vääritymää. White ja Griffith ratkaisivat pulman jakamalla yhteisviittausten kolme korkeinta arvoa kahdella ja sijoittamalla näin saadut luvut tyhjiin paikkoihin. Heidän mukaansa luku indikoi yleisellä tasolla kunkin kirjoittajan suhteellista tärkeyttä. (White & Griffith 1981, 165.) Noudatin tätä laskentatapaa omassa tutkimuksessani.

Korrelaatioiden laskeminen tuottaa myös matriisin, joka on muodoltaan samanlainen kuin absoluuttisista luvuista koostuvakin, mutta joka tuloksia esittäessä on yleensä riisuttu malli: mukana eivät ole molemmat symmetriset puoliskot vaan ainoastaan toinen niistä. Myös desimaalipisteet sekä diagonaaliset arvot (kirjoittajan korrelaatiokerroin itsensä kanssa, joka on tietenkin aina 1) jätetään pois. (Kuva 2.)

	Allen	Andrews	Barber	Ben-David
Andrews	528			
Barber	-008	288		
Ben-David	046	247	754	
Bernal	015	122	532	602

Kuva 2. Esimerkki korrelaatioarvojen esittämisestä.

Analysointi

Aiemmissa tutkimuksissa ei tavallisesti ole tehty johtopäätöksiä suoraan korrelaatiokertoimista, vaan niitä on käytetty vain datana jatkoanalyysissa. Tällaiselle menettelylle ei kuitenkaan ole erityisen vahvoja perusteluja, samoin kuin ei ole erityisiä esteitä päinvastaiselle valinnalle. Kun matriisista karsitaan kaikkein pienimmät arvot, saadaan joukko lukuja, jotka osoittavat, keiden kirjoittajien välillä on vahvoja yhteyksiä ja keiden välillä puolestaan ei ole mainittavaa yhteyttä. Kun poistin ne arvot, jotka olivat alle 0,6, ja järjestin matriisin uudelleen siten, että jäljelle jääneet luvut kasautuivat ylös ja tyhjät paikat alas, sain kuvan 3 kaltaisia ryhmittymiä. Siitä nähdään esim., että Smallin profiili on hyvin samankaltainen kuin Griffithin ja Garfieldin, mutta Meadowsin profiilista se jo poikkeaa. Välillinen yhteys on silti olemassa myös Smallilla ja Meadowsilla, sillä molemmat korreloivat voimakkaasti Garfieldin ja Narinin kanssa.

	Griffith	Small	Garfield	Narin
Small	821			
Garfield	696	871		
Narin		732	870	
Meadows			768	629

Kuva 3. Esimerkki valikoitujen korrelaatioarvojen esittämisestä.

Koska pienimpien arvojen karsimisen tuloksena jää jäljelle suhteellisen vähän lukuja, niitä voidaan tarkastella sellaisinaan, ts. kokonaisuus on hahmotettavissa ilman sofistikoituneita tietokoneohjelmia, jollaisia korrelaatioiden jatkoanalyysiin yleensä käytetään. Omassa aineistossani syntyi melko selkeitä ryhmiä, joiden jäsenistä miltei kaikilla on korkeat korrelaatioarvot keskenään ja vastaavasti matalat arvot toisten ryhmien jäsenten kanssa. Lisäksi korrelaatioihin perustuva analyysi on yksittäisten kirjoittajien tasolla tarkempi kuin jollakin tietokoneohjelmalla laskettu ryhmittely. Ohjelmat esim. tunnistavat kyllä ne nimet, joilla on aivan muista poikkeava profiili, mutta eivät osaa määrittellä poikkeaman laatua. Tällöin jää selvittämättä, millä tavoin eri ryhmien keskinäiset välitysketjut rakentuvat ja ketkä ovat ketjun missäkin kohdassa avainasemassa.

Tutkittavan alan kokokuvan hahmottamisessa käytetään yleensä joko klusteri- tai faktorianalyysia. Kummassakin kirjoittajat sijoittuvat ryhmiin, joiden lukumäärä ja koostumus riippuu paitsi kirjoittajien keskinäisistä korrelaatioista, myös siitä, millaisiksi tietyt katkaisusäännöt määrätään käytettävällä ohjelmalla. Esim. ennen faktorianalyysin suorittamista voidaan valita vaihtoehto "eigenvalue < 1"⁴. Faktorianalyysi on mainituista laskentatavoista parempi vaihtoehto sikäli, että ensinnäkin se antaa moniulotteisempaa tietoa ja hukkaa vähemmän informaatiota, ja toiseksi se mahdollistaa kunkin nimen esiintymisen monessa paikassa eikä vain yhdessä. Lisäksi analyysin tuloksena muodostuneiden faktoreiden interkorrelaatio tuo esiin ryhmien keskinäiset yhteydet. (McCain 1990, 440; White & Griffith 1982, 259-260.)

Sen, mihin ryhmään kukin kirjoittaja faktorianalyysin perusteella kuuluu, määrää se, miten suuri lataus hänellä on kullakin faktorilla. Mitä suurempi latauksen arvo on välillä -1+1, sitä voimakkaammin kirjoittaja on vaikuttanut kyseisen faktorin rakentumiseen. Kaikkein pienimmät luvut on syytä jättää kokonaan huomiotta, joten vain tietyn rajan ylittävät arvot raportoidaan. Esim. White & Griffith (1981) käyttivät katkaisukohtana arvoa 0.40, mutta tulosten tulkinnan kannalta käyttökelpoisena kynnyksenä on pidetty niinkin korkeaa arvoa kuin 0.70 (McCain 1990, 440). Omassa tutkimuksessani katkaisukohta 0.50 tuotti ryhmittelyn, jossa yhtä lukuunottamatta kaikilla kirjoittajilla oli rajan ylittävää painoarvo ainakin yhteen faktoriin nähden ja toisaalta niin moni nimi ei latautunut usealla faktorilla, etteikö selkeää rakennetta olisi syntynyt.

Yleensä tutkittavasta alueesta tehdään jonkin multidimensional scaling -ohjelman avulla kaksi- tai kolmiulotteinen kartta, jossa on kunkin kirjoittajan sijaintia kuvaavia pisteitä sekä jälkikäteen piirrettyjä, eri ryhmiä osoittavia rajoja. Pidän kuitenkin tällaisista automaattisesti tuotetuista kartoista saatavaa hyötyä hieman kyseenalaisena. Ensinnäkin ohjelmat tuottavat vain kirjoittajien laskennallisen sijainnin suhteessa kaikkiin muihin, mutta eivät ota huomioon tällä tavoin syntyviä ryhmiä, jotka siis analytykon on joka tapauksessa konstruoitava itse. Toiseksi näitä ohjelmia käytettäessä ei voida yhdistää eri laskentatavoilla saatua dataa, vaan esim. korrelaatioarvojen perusteella piirrettyä karttaa ja faktorianalyysin tuottamia ryhmiä on tarkasteltava toisistaan erillään. Kolmanneksi yhteisviittausanalyysin tavanomainen aineisto

on sen kokoinen, että siitä pystyy mainiosti piirtämään kaksiulotteisen kartan manuaalisestikin. Kolmiulotteiset kuvat (esim. White & Griffith 1982, 267) puolestaan ovat paperilla esitettyinä sen näköisiä, ettei kolmas ulottuvuus helposti hahmotu muille kuin kuvan tekijälle.

Vaihtoehto automaattisesti tuotetulle kartalle on siis piirtää kuva itse. Piirtäminen ei kylläkään ole yhtä nopea ja vaivaton tapa kuin tietokoneen työllistäminen. Jotkut saattavat pitää tällaista menettelyä myös jotenkin vähemmän objektiivisena kuin valmiiseen ohjelmaan luottamista. Tutkija on kuitenkin jo ennen tätä vaihetta tehnyt niin monta valintaa, että tulos on joka tapauksessa hänen luomuksensa eikä ns. objektiivinen kuva todellisuudesta. Sitä paitsi tuloksen luotettavuuden arviointi perustuu käytettäviin kriteereihin: onko parempi uskoa tietokoneen tulostamaan graafiseen esitykseen tietyistä numeroista vai tutkijan rakentamaan, monia aspekteja huomioon ottavaan kokonaiskuvaan?

Tutkija voi kuvaa piirtäessään pitää mielessään ensinnäkin varsinaiset yhteisviittausten absoluuttiset määrät: koko tutkittavan alan keskushenkilöitä ovat ne, joilla on yleisesti ottaen paljon yhteisviittauksia toisten kanssa ja yksittäisten ryhmien keulahahmoja ne, joiden viittausmäärät ryhmän sisällä ovat suuret. Toiseksi kirjoittajien keskinäiset korrelaatiot osoittavat, kuinka lähelle toisiaan he kartalla sijoittuvat. Faktorianalyysin tuottamat ryhmät kertovat, missä kohtaa kartassa on rajaviivoja, ja samaisesta analyysistä ilmenevät faktorien keskinäiset korrelaatiot ilmaisevat, mitkä joukot leikkaavat toisiaan, mitkä ovat muuten vain lähellä ja mitkä kaukana toisistaan. Nyt voidaan vielä palata kirjoittajien keskinäiskorrelaatioiden tasolle ja katsoa, ketkä sijoittuvat mainittuihin leikkauksiin. Lopuksi asetetaan ei-kenenkään-maalille leijumaan ne mahdolliset tapaukset, jotka ovat välittäjiä ryhmien kesken, mutta eivät kunnolla minkään joukon jäseniä.

Sopivilla typografisilla tehosteilla ja niiden selityksillä varustettuna tuloksena syntyvä kuva kertoo oikeastaan yhteisviittausanalyysin kaikki tulokset. Siitä voidaan lukea analysoidun alueen intertekstuaalinen rakenne kokonaisuutena, alan sisäisten ryhmien koostumus ja keskinäissuhteet sekä yksittäisten kirjoittajien asema kyseessä olevassa joukossa. Se, mitä tämän luennan jälkeen on pääteltävissä, edellyttää tietenkin, että lukija tuntee sekä tutkittavan aihepiirin että yhteisviittausanalyysin rajoitukset.

Yhteisviittausanalyysin rajoitukset

Yhteisviittausanalyysin luotettavuudesta ja validiteetista ei olla yksimielisiä. On väitetty, että paljon yhteisviittauksia saaneet tutkijat eivät muodosta koherentteja sosiaalisia ryhmiä, mutta kun asiaa on yritetty testata, he näyttävät kuitenkin muodostavan niitä. Missään tapauksessa yhteisviittausanalyysi ei yksinään ole riittävä keino tutkia tieteenharjoittajien sosiaalisia suhteita eikä sen avulla voida muodostaa täydellistä kuvausta tietyistä erityisalasta. (Hargens 1978, 130.)

Toisaalta yhteisviittausanalyysin tekijät eivät ole koskaan väittäneetkään, että B:llä ja C:llä on välttämättä jokin yhteys silloin, kun A viittaa samassa työssä molempiin. Yhteisviittauksen *kasautuminen* on se seikka, joka indikoi keskinäistä suhdetta. Analyytikot eivät liioin yksinkertaisesti oletta, että yhteisviittauksia saaneilla kirjoittajilla on keskenään viestintäyhteys, vaikka se usein onkin olemassa. Suhde, jota viittauksen kasautumat yleensä kuvastavat, on viittaajien kirjallisuudessa havaitsema aihepiiriin tai metodologisen lähestymistavan samankaltaisuus. Yhteisviittausanalyysin avulla konstruoidut erityistieteiden kartat paljastavat kirjoittajien muodostaman kognitiivisen rakenteen ja ne kertovat ennen muuta intertekstuaalisista suhteista. (White 1990, 96-98.)

Yhteisviittausanalyysija on kritisoitu myös sen suhteen, ettei niissä ole riittävästi otettu huomioon aikaulottuvuutta. Klassikoihin ei välttämättä yhteisviitata siksi, että ne edustavat samaa paradigmaa, vaan siksi, että ne edustavat jotain varhaisempaa aikakautta. Myös aktiivitutkijoiden osalta voisi olla parempi, jos analyysissa tarkasteltaisiin heidän uransa eri kausia. Tämä menettely tarjoaisi empiirisen välineen analysoida viitattujen tutkijoiden kehitystä, paradigmasiirtymä ja antia erilaisille koulukunnille. Toinen tapa tarkastella viittauksia aikaulottuvuudella on jakaa aineisto eri periodeihin viittaavan kirjallisuuden ilmestymisajankohtien mukaan, mikä tuo esiin viittaajien näkemyksissä tapahtuvan muutoksen. (Bayer & Smart & McLaughlin 1990, 451-452.)

Aika-aspekti ei ole mikään ylitsepääsemätön ongelma – se pitää vain ottaa huomioon. Kirjoittajien yhteisviittausanalyysissa ei kylläkään pystytä tarkkaan määrittelemään viitattun julkaisun ilmestymisajankohtaa, ellei jokaista viittausta tutkita erikseen, mutta se tiedetään, minä ajanjaksona kukin viitattu henkilö on julkaissut. Viittaavan kirjalli-

suuden ilmestymisaikojen huomioon ottaminen on puolestaan suhteellisen yksinkertaista, ja jopa SSCI on ollut olemassa niin kauan, että siitä kerätyllä aineistolla voidaan tutkia viittaajien käsitysten kehitystä. Ellei aineistoa kuitenkaan sen keruuvaiheessa mitenkään jaotella eri aikakausien mukaisesti, ajallinen ulottuvuus voidaan silti pitää mukana tulosten tulkinnassa, kunhan analyytikko tuntee aineistonsa. Esim. tieteesosiologien saamia yhteisviittauksia tutkittaessa ei ole pahitteeksi tietää, että Karl Mannheimin keskeinen tuotanto tällä alueella on syntynyt 1930-luvulla, kun sen sijaan Michael Mulkay aloitti 1960-70-luvun vaihteessa ja on julkaissut uusimman monografiansa 1990-luvulla.

Kaikkein vakavin pulma, joka koskee kaikkea viiteanalyttistä tutkimusta, on sellainen yhteiskunnallinen kehitys, joka muuttaa tieteellistä käytäntöä ja saattaa heijastua myös viittaamisessa. Bibliometrinen karttojen tarkoitus on etsiä niitä kohtia, joihin tieteellinen huomio keskittyy, ja esittää näiden älyllisten keskittymisten keskinäisiä suhteita. Viittauksia ja yhteisviittauksia tutkimalla selvitetään kirjallisuuteen kohdistuvaa kollektiivista huomiota. Kirjallisuuteen suuntautuminen puolestaan on riippuvaista mm. niistä yhteiskunnallisista olosuhteista, joissa tutkimusta harjoitetaan. (Braam 1991, 267-268.)

Nykyoloissa kirjallisuus ei ole keskeisen tärkeä seikka monienkaan tieteenalojen ongelmanasetteluissa toisin kuin vielä 1800-luvulla ja 1900-luvun alussa. Mm. tieteen tuottama kasvava painolasti valtioiden budjeteissa on johtanut julkisen kontrollonin lisääntymiseen ja voimavaroja suunnataan tutkimukseen, joka ratkoo ajankohtaisia ongelmia. Varsinkin nuoria erityistieteitä kiinnostavat teemat valikoituvat yhä enemmän yhteiskunnan vaatimusten mukaan eikä niinkään aiemman tutkimuksen pohjalta. (Mt., 268-274.) Toiseksi tutkimusjärjestelmän painopiste on siirtynyt tieteestä teknologiaan, yliopistojen ulkopuolelle, ja lyhytjänteinen sopimustutkimus on lisääntynyt (Kaukonen 1988, 167). Varsinkin sopimustutkimukseen liittyvä aikaraja näyttää johtavan siihen, että tutkimusviestintä pohjautuu yhä vähemmän kirjallisuuteen samalla kun yhä laajemmin luotetaan nopeisiin informaaleihin viestintäkanaviin.

Edellä todettu merkitsee Braamin mukaan sitä, että yhteisviittausanalyysilla ei voida luoda esitystä nykyisten tieteenalojen temaattisesta struktuurista, joskin se soveltuu varhaisempien aikakausien kuvaamiseen (Braam 1991, 268). Ainakin toistaiseksi tieteenharjoituksen normit näyttävät kuitenkin

kin olevan voimassa, mitä tulee julkaisemiseen ja viittaamiseen, eli vaikka kirjallisuuden käyttö ei olisikaan tutkimuksessa keskeistä, kirjallisia raporteja julkaistaan ja niissä osoitetaan lukeneisuutta, annetaan tunnustusta aiemmille töille, kritisoidaan toisten koulukuntien edustajia jne. Meritoituminen edellyttää kirjallisuuteen suuntautumista, sekä julkaisemista että olevan kirjallisuuden tuntemuksen osoittamista. Niinpä yhteisviittausanalyysillä voidaan kuvata tieteellisten tekstien suhteita, vaikkakaan ei kokonaisten tieteenalojen rakenteita muuten kuin julkaisuissa ilmenevien virallisten yhteyksien osalta. Silti looginen johtopäätös on, että mitä marginaalisemmaksi perinteinen tieteellinen kirjallisuus käy, sitä hyödyttömämpää on nykymuotoinen viiteanalyttinen tutkimus.

Hyväksytty julkaistavaksi 21.4.1993.

Viitteet

1. Tämän tekstin syntyä aikana Tampereen yliopiston kirjastotieteen ja informatiikan laitoksella on esitetty laitoksen nimen muuttamista informaatiotutkimuksen laitokseksi. Johdonmukaista olisi, että myös tieteenalasta käytettäisiin nimeä, joka esiintyy laitoksen nimessä. Nimenmuutoksen toteuttaminen on luultavaa, mutta se ei ole vielä tapahtunut. Alan englanninkielisessä terminologiassa 'information science' on ilmeisesti yhtä yleinen kuin 'information studies'. Niinpä voidaan sanoa, että tässä tilanteessa suomenkielisten termien käytön suhteen vallitsee jonkinasteinen valinnanvapaus. Itse olen tähän saakka pitänyt kiinni siitä, että 'informaatiotiede' olisi parempi termi kuin 'informaatiotutkimus', mutta alistun ryhmäpaineen edessä: paitsi, että Kalervo Järvelin ehti jo lehden viime numerossa lanseerata julkisuuteen informaatiotutkimus-termin, myös Tampereen yliopiston kirjastotieteen ja informatiikan laitoksen työntekijöiden enemmistö kannattaa sitä laitoksen nimeksi. Silti on joitakin kielellisiä ilmauksia, joihin 'informaatiotutkimus' ei mielestäni sujuvasti taivu ja siksi käytän esim. ilmaisua 'informaatiotieteellinen' sen sijaan, että puhuisin 'informaatiotutkimuksellisesta'.
2. Mainittakoon kuitenkin, että vuonna 1978 Pirjo Suutarinen käytti ko. menetelmää lisensiaatintutkimuksessaan "Viittausanalyysi suomalaisten lääketieteellisten julkaisujen kansainvälisestä käytöstä".
3. Tieteentutkimuksen tarkastelussaan White ja Griffith käyttivät faktorianalyysejä. Sen tuloksena syntyi seitsemän muuttujaryhmää. Monet muuttujat latautuivat voimakkaasti myös yhdellä tai useammalla muulla faktorilla kuin sillä, johon ne ensisijaisesti kuuluivat.
4. Eigenvalue = latausten neliöiden summa faktorilla (McCain 1990, 440).

Kirjallisuus

- Bayer, Alan E. & Smart, John C. & McLaughlin, Gerald W. - Mapping Intellectual Structure of a Scientific Subfield through Author Cocitations. *Journal of the American Society for Information Science* 41 (1990) 6, 444-452.
- Bellardo, T. - The Use of Co-citations to Study Science. *Library Research* 2 (1980), 231-237.
- Ben-David, Joseph - Emergence of national traditions in the sociology of science : the United States and Great Britain. San Francisco 1978, s. 197-218. Teoksessa: *Sociology of Science / toim. Jerry Gaston*.
- Braam, Robert R. - Mapping of Science: Foci of Intellectual Interest in Scientific Literature. Leiden 1991.
- Edge, David - Quantitative Measures of Communication in Science: A Critical Review. *History of Science* 17 (1979), 102-134.
- Gilbert, G. Nigel & Mulkay, Michael - Opening Pandora's Box : a sociological analysis of scientists' discourse. Cambridge 1984.
- Hargens, Lowell - Theory and Method in the Sociology of Science. San Francisco 1978, s. 121-139. Teoksessa: *Sociology of Science / toim. Jerry Gaston*.
- Kaukonen, Erkki - Tiedejärjestelmän murros ja tieteen itsearvioinnin tarpeellisuus. Tampere 1988, s. 167-174. Teoksessa: *Tiede, kriittisyys, yhteiskunta: näkökulmia ja taustoja tieteen käytäntösuhteseen / toim. Jorma Kuitunen*. - (Tampereen yliopisto, aluetieteen laitos, Sarja A, 9/1988).
- Kuhn, Thomas S. - The structure of scientific revolutions. 2nd ed. Chicago 1970.
- Luukkonen, Terttu - Citations in the rhetorical, reward, and communication systems of science. Tampere 1990. - (Acta Universitatis Tamperensis, ser A, vol 285).
- McCain, Katherine - Mapping Authors in Intellectual Space: A Technical Overview. *Journal of the American Society for Information Science* 41 (1990) 6, 433-443

- MacRoberts, Michael & MacRoberts, Barbara - Problems of Citation Analysis: A Critical Review. *Journal of the American Society for Information Science* 40 (1989) 5, 342-349.
- Mulkay, Michael - The sociology of science in Britain. Carbondale 1977, s. 224-257. Teoksessa: *The Sociology of Science in Europe* / toim. Robert K. Merton & Jerry Gaston.
- Paisley, William - The Future of Bibliometrics. Newbury Park 1990. Teoksessa: *Scholarly Communication and Bibliometrics* / ed. by Christine L. Borgman, s.281-299.
- Pierce, Sydney J. - Disciplinary Work and Interdisciplinary Areas: Sociology and Bibliometrics. Newbury Park 1990. Teoksessa: *Scholarly Communication and Bibliometrics* / ed. by Christine L. Borgman, s. 46-58.
- Rautio, Pertti & Suhonen, Pertti - Yhteiskuntatieteiden tietovirrat ja suomalainen tutkija : raportti sosiologian ja valtio-opin väitöskirjojen lähteistöstä. Helsinki 1981. - (Suomen Akatemia 2/1981).
- Verronen, Veli - The growth of knowledge : an inquiry into the Kuhnian theory. Jyväskylä 1986. - (Jyväskylän yliopiston filosofian laitoksen julkaisuja, No 35).
- Weinstock, N. - Citation Index. Teoksessa: *Encyclopedia of Library and Information Science* / ed. by A. Kent. New York 1971.
- White, Howard D. - Author Co-Citation Analysis: Overview and Defense. Newbury Park 1990. Teoksessa: *Scholarly Communication and Bibliometrics* / ed. by Christine L. Borgman, s. 84-106.
- White, Howard D. & Griffith, Belver C. - Author Cocitation: A Literature Measure of Intellectual Structure. *Journal of the American Society for Information Science* May 1981, 163-171
- White, Howard D. & Griffith, Belver C. - Authors as Markers of Intellectual Space: Cocitation in studies of Science, Technology and Society. *Journal of Documentation* 38 (1982), 255-272