

Viitetiedolla vaikutusten jäljille

Ari Muhonen

Kirjastojen vaikuttavuus kiinnostaa niin kirjastoja itseään kuin niiden kehysorganisaatioita. Sitä vain on vaikea mitata. Yhden vaihtoehdon tähän tarjoaa tutkijoiden omien lähdeviitteiden analyysi. Samalla saadaan paljon muutakin hyödyllistä tietoa.

Kirjastot joutuvat entistä tarkemmin osoittamaan kehysorganisaatioille hyötynsä. Selvää on, että kirjastot vaikuttavat osaltaan tutkimuksen, opetuksen ja opiskelun laatuun ja tuloksellisuuteen. Tämän konkreettinen osoittaminen on kuitenkin hankalaa, tulos kun pitäisi mieluiten näyttää selkeinä numeroina. Vaikuttavuuden arviointi onkin nousemassa yhä tärkeämpään asemaan esimerkiksi yliopistokirjastojen neuvoston tavoitteissa.

Yhden näkökulman vaikuttavuuteen tarjoaa tieto, miten hyvin kirjaston kokoelmat tyydyttävät asiakkaiden tiedontarpeen. Mutta miten voimme tietää, mitä asiakkaat tarvitsevat? Tutkijat dokumentoivat käyttämänsä aineiston julkaisujensa lähdeluetteloihin. Näistä saamme selville, mitä materiaalia tutkijat ainakin ovat hyödyntäneet. Vertaamalla viitteitä kirjaston omiin kokoelmiin, saamme suhdeluvun, joka kertoo kirjaston palvelukyvystä.

Viiteanalyysi menetelmänä

Viiteanalyysi kehitettiin 1920-luvulla kokoelmien kehittämisen tarpeisiin (Gross ja Gross, 1927). Menetelmällä laadittiin kemian alan ydinlehtien lista hankinnan keskittämiseksi tärkeimpään materiaaliin, kun varat eivät enää riittäneet kaikkeen. 1930-luvulta lähtien tällaisia lehtilistoja on luotu viiteanalyysillä monellekin tieteenalalle.

Viiteanalyysiä on hyödynnetty myös moneen muuhun tarkoitukseen, esimerkiksi viime aikoina Suomessa Ahtola ja Juutistenaho (2003) sekä Pappila (2004) ovat käyttäneet sitä kokoelmien arviointiin. Erityisesti he ovat tarkastelleet teosten saatavuutta. Kortelainen (1999, 2000) puolestaan käytti väitöskirjassaan viiteanalyysia suomalaisen tieteellisen lehden diffuusion tutkimiseen. Teoksessa on selvitetty varsin perusteellisesti metodin hyviä ja huonoja puolia.

Viiteanalyysin käytettävyydestä ollaan montaa mieltä. Kovin kritiikki on kohdistunut viittauskäytäntöön yleensä.

Viittaus voi olla luonteeltaan negatiivinen, viitteet eivät kuvasta todellista aineiston käyttöä, ne eivät kerro, miten tutkijat ovat saaneet aineiston käsiinsä, eikä viitteiden perusteella voi ennustaa tulevaa aineistotarvetta (katso esimerkiksi Butkovich, 1996 sekä MacRoberts ja MacRoberts, 1996). Lisäksi voidaan olettaa, että tutkijat suosivat paikallisesti saatavilla olevaa aineistoa.

Kritiikki on varmasti oikeutettua silloin, kun analyysoidaan viitteiden merkittävyyttä ja käytettyjen lähteiden tärkeyttä. Vaikuttavuutta tutkittaessa viitteet kuitenkin muodostavat ”vain” listan, jota verrataan kokoelmätietoihin. Kortelaisen (1999, s. 14) mukaan ”jos kirjoittaja merkitsee julkaisun lähdeluetteloonsa, hän tietää sen olemassaolosta

riippumatta siitä, kuinka intensiivisesti hän on sitä käyttänyt tai onko hän ollenkaan lukenut sitä". Näin viitteet toimivat tieteellistä lehteä koskevan tietoisuuden indikaattorina.

Väitöskirjat lähdemateriaalina

Vuonna 2003 Teknillisen korkeakoulun (TKK) julkaisurekisteriin kirjattiin 3 304 julkaisua, joista oli lehtiartikkeleita 1 148, konferenssiesitelmiä 1 468 ja raportteja 334. Tällaista massaa ei voida ottaa kokonaan viiteanalyysin lähdemateriaaliksi. Siitä pitää valita otos, joka on edustava mutta käsiteltävissä oleva määrä teoksia.

Vähänkin suuremman viitemäärän käsittely edellyttää, että ne saadaan valmiiksi sähköisessä muodossa. Tähän on kaksi vaihtoehtoa.

Useissa tutkimuksissa on hyödynnetty ISI:n (Institute of Scientific Information) tietokantaa. Täältä kuitenkin löytyy lähinnä angloamerikkalaisten lehtien artikkeleita. Kun tarkastellaan suomalaista tutkimusta, sellaisen aineiston edustavuus on hyvin kyseenalaista.

Elektroniset väitöskirjat tarjoavat toisenlaisen vaihtoehdon. TKK:n julkaisuarkistosta löytyy varsin kattava kokoelma uusimpia väitöskirjoja pdf-muodossa. Kyseinen aineisto edustaa hyvin TKK:n tutkimusta, tekevähän jatko-opiskelijat yleensä työnsä korkeakoulun omissa tutkimusprojekteissa.

22 eri alojen väitöskirjaa tutkittavana

Tutkimuksen lähdeaineiston muodostaa 22 TKK:ssa vuonna 2003 tehtyä väitöskirjaa (ks. taulukko 1). Määrä on 17% kaikista tuon vuoden väitöksistä. Teokset valittiin satunnaisesti kullakin osastolla tehtyjen väitöskirjojen joukosta siten, että arkkitehtiosastoa lukuunottamatta kaikista osastoista tuli vähintään yksi väitöskirja mukaan. Mukaan otettujen teosten yhteenlaskettu viitemäärä oli 2871.

Taulukko 1. Teknillisessä korkeakoulussa vuonna 2003 tehtyjen väitöskirjojen sekä tutkimukseen mukaan otettujen väitöskirjojen määrä osastoittain.

Osasto	Tehty 2003	Mukana tutkimuksessa	%
Sähkö- ja tietoliikennetekniikan osasto	31	5	16
Tietotekniikan osasto	10	2	20
Teknillisen fysiikan ja matematiikan osasto	27	4	15
Rakennus- ja ympäristötekniikan osasto	4	1	25
Konetekniikan osasto	7	1	14
Automaatio- ja systeemitekniikan osasto	2	1	50
Tuotantotalouden osasto	16	3	19
Kemian tekniikan osasto	14	2	14
Materiaali- ja kalliotekniikan osasto	3	1	33
Puunjalostustekniikan osasto	5	1	20
Maanmittausosasto	7	1	14
Arkkitehtiosasto	1	0	0
Yhteensä	127	22	17

Vaikka väitöskirjojen lähdeviitteet saatiin suoraan elektronisessa muodossa, niiden käsittely oli työlästä. Viitteet piti saattaa yhtenäiseen muotoon sekä erotella eri tyyppiset viitteet omiksi joukoikseen. Tämä tehtiin tekstinkäsittely- ja taulukkolaskentaohjelmilla.

Tulokseksi saatiin julkaisutyypeittäin lajiteltu viitteiden luettelo, jossa jokainen viitteen osa (tekijä, otsikko jne.) oli omassa sarakkeessaan. Tämä antoi mahdollisuuden monipuolisten ristiintaulukointien ja laskelmien suorittamiseen.

Perustiedot

Taulukot 2 ja 3 antavat taustatietoja materiaalista. Viitteiden kokonaismäärä vaihteli huomattavasti teoksittain, pienimmän ja suurimman viitemäärän ero oli kymmenkertainen. Keskimäärin viitteitä oli 131. TKK:lla suosituissa niin sanotuissa nippuväitöskirjoissa yhteenvedo-osan viitemäärä oli hieman monografiäväitöskirjan viitemäärää pienempi.

Taulukko 2. Analysoitujen lähdeluetteloiden tilastotietoja (N = 2871).

Viitteiden määrä	Kaikki	Monografiat	Nippuväitöskirjat
Keskiarvo	131	158	97
Mediaani	123	141	95
Pienin	47	47	60
Suurin	416	416	168

Taulukko 3. Viitteiden jakauma aineistotyypeittäin (N = 2871).

Aineistotyyppi	Viitteiden määrä	%	Pienin %	Suurin %
Lehti	1661	58	23	86
Monografia	703	24	2	65
Konferenssi	263	9	1	38
Kokoomajulkaisu	124	4	0	8
Opinnäyte	89	3	1	17
Elektroninen	22	1	NA	NA
Muu	9	0	NA	NA

Eniten lähdemateriaalina on odotetusti käytetty lehtiartikkeleita, seuraavaksi tulevat monografiat ja konferenssijulkaisut. Tutkimuksessa laskettiin viitelajien prosenttiosuudet myös osastoittain. Saaduista luvuista pienimmät ja suurimmat on kirjattu taulukon 3 kahteen viimeiseen sarakkeeseen. Näistä nähdään, että aineistotyyppien tärkeys tieteenaloittain vaihtelee suuresti.

Kaikki tutkimuksen väitöskirjat oli kirjoitettu englannin kielellä. Kieleltään tunnistettavista viitteistä englanninkielisiä oli 94%. Muita käytettyjä kieliä olivat suomi (4%), saksa (2%) ja ruotsi (1%). Muita kieliä oli mukana vain muutama hajatapaus.

Viitteiden keskimääräinen ikä oli 13 ja mediaani 8 vuotta. Edellistä lukua rasittaa muutama hyvin vanha viite, vanhin niistä oli vuodelta 1772! Keskimäärin uusimpia viitteitä olivat opinnäytteet (mediaani 5 vuotta) ja konferenssit (6 vuotta), vanhimmat viitteet taas monografioita (11 vuotta).

Elektronisessa muodossa saatavilla olleiden lehtiartikkeleiden suurimmat kustantajat on lueteltu taulukossa 4. Elsevier on odotetusti suurin sekä viitteiden että nimekkeiden

määrässä. American Physical Society on viitteissä toinen mutta nimekkeissä vasta kolmastoista, sillä sen viitteet keskittyvät vain viiteen lehteen. Tekniikan alalla tärkeät kustantajat IEEE (Institute of Electrical and Electronics Engineers), Wiley ja Kluwer (tässä vielä itsenäisenä) ovat myös mukana taulukossa, samoin taloustieteiden alalla tärkeä kokotekstitietokanta Abi/Inform.

Taulukko 4. Elektronisessa muodossa saatavilla olleiden lehtiartikkeleiden löytyvyys kustantajittain / kokotekstitietokannoittain (n = 1208).

Kustantaja / Kokotekstitietokanta	Osuus viitteistä, %	Osuus nimekkeistä, %
Elsevier	24	31
American Physical Society	16	2
Abi/Inform	11	10
IEEE	9	11
Wiley	7	4
Kluwer	3	5
Muut	30	37

Monografioiden ja lehtiartikkeleiden löytyvyys

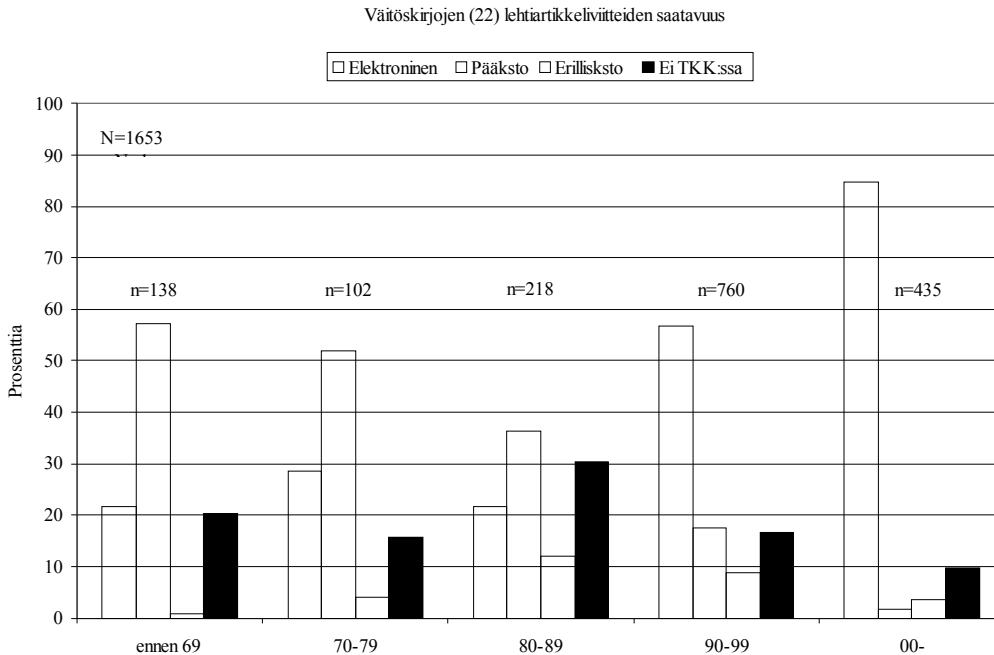
Monografioiden löytyvyys TKK:n kirjastoista on esitetty taulukossa 5. Kukin teos on mukana vain yhteen kertaan, jotta prosenttiosuuksien summa olisi 100. Sekä pää- että osastokirjastosta löytyneet teokset on kirjattu vain pääkirjaston kohdalle ja siksi osastokirjastoista löytyvien teosten määrä on tässä todellista pienempi. Taulukossa on myös ilmoitettu varsinaisten monografioiden, tutkimusraporttien ja erilaisten ohjelmistojen ym. manuaalien jakautumat.

Kaikkiaan TKK:n kirjastoista löytyy 49% käytetyistä monografioista. Varsinaisista monografioista löytyy 58%. Osuus on hyvin saman suuruinen kuin Pappilan (2004) tekemät havainnot koskien humanististen alojen väitöskirjojen lähdemateriaalin löytyvyyttä Helsingin yliopiston kirjastosta (55%).

Tutkimusraporttien ja manuaalien kohdalla löytyvyys oli hyvin huono. Nämä ovatkin tyypillisesti aineistoa, jota tutkijat hankkivat suoraan itselleen tai omaan laboratoriokirjastoonsa.

Taulukko 5. Monografioiden löytyvyys prosentteina (n = 698).

Mistä löytyy	Kaikki	Monografiat	Raportit	Manuaalit
Pääkirjasto	23	27	12	
Osastokirjasto	26	31	10	
Ei TKK:ssa	51	42	78	100



Kuva 1. Väitöskirjojen lehtiartikkeliviitteiden saatavuus. Elektronista aineistoa on preferoitu paperiseen verrattuna.

Kuvassa 1 näkyy lehtiartikkeliviitteiden saatavuus elektronisena sekä paperisena joko pää- tai osastokirjastoissa. Viitteet ovat mukana vain kertaalleen siten, että analyysissä on suosittu elektronista muotoa ja paperilehdistä pääkirjaston kappaletta. Kuvasta näkyy selvästi elektronisen aineiston osuuden kasvu uusimmassa aineistossa ja vastaavasti painetun aineiston osuuden pieneneminen. Kaikista lehtiviitteistä on elektronisesti saatavilla 55%, mutta uusimmasta, 2000-luvun aineistosta jo 85%.

TKK:n kirjastoista löytymättömän materiaalin osuus on uusimmassa aineistossa alle kymmenen prosenttia. Koko aineistosta TKK:sta löytyy 83%. Näistä on vain elektronisena 13%, elektronisena ja painettuna 42% ja vain painettuna 28%. Päällekkäisyyttä on siis vielä paljon. Toisaalta päällekkäisyyttä on nimenomaan vanhemmassa aineistossa.

Mielenkiintoinen havainto on se, että pahimmat kokoelmien puutteet ovat 80-luvun aineistossa. 80- ja 90-luvuilla tehtiin paljon leikkauksia lehtitilauksiin. Jälkimmäisiä on saatu paikattua elektronisten aineistojen lisensseillä. Nämä eivät kuitenkaan yleensä ulotu 80-luvulle. Siksi olisikin järkevää miettiä kustantajien takautuvien elektronisten kokoelmien hankintaa.

Johtopäätöksiä

TKK:n kirjastolaitoksen kokoelmista löytyi 83% tutkittujen väitöskirjojen lehtiartikkeli- ja 49% monografiaviitteistä. Luvut ovat helposti ymmärrettäviä, ja ne antavat yhden näkymän kirjaston vaikuttavuudesta tutkimustyöhön. Toistamalla tutkimus vuosittain saataisiin aikasarja, joka näyttäisi kirjaston kyvyn kehittää kokoelmiaan. Lukuja voidaankin pitää sopivina vaikuttavuuden mittariksi.

Menetelmän luotettavuus on puolustettavissa. Tutkimusaineisto koostuu tutkijoiden itsensä mainitsemisen arvoiseksi katsomasta lähdemateriaalista. Toisaalta viittausten ”vanhuus” on ongelma.

Väitöskirjan teko on usein pitkä prosessi, ja lähteitä on tällä tavalla ajateltuna käytetty jopa vuosia ennen väitöskirjan valmistumista. Mutta vastaavassa tutkimuksessa todennäköisesti tarvittaisiin pääosaa näistä taustamateriaaliksi edelleen. Siksi viitejoukko voidaan katsoa soveltuvaksi myös tämän hetken kokoelmien arviointiin.

Tutkijoiden voidaan ajatella käyttävän helposti sitä aineistoa, mikä on paikallisesti saatavilla. Vaikuttavuuden arviointia voitaisiinkin parantaa ottamalla vertailuaineistoksi muiden vastaavan alan yliopistojen väitöskirjojen lähdeluetteloita, ovathan huippuyliopistojen väitökirjat nykyisin pääsääntöisesti elektronisessa muodossa saatavilla.

Viiteanalyysin ongelma on lähdeluetteloiden tarvitsema runsas käsityö viitteiden muokkauksessa analyysiin sopivaan muotoon. Pienellä otoksella tämä on kuitenkin hallittavissa. Siksi olisikin selvítettävä, millainen otos riittäisi luotettavan lopputuloksen aikaansaamiseen. Laajemmissa tutkimuksissa tarvittaisiin automaattista tai puoliautomaattista muokkausohjelmaa. Sellaisen tekeminen olisi kuitenkin hankalaa, koska viitteiden muoto vaihtelee suuresti lähdeluettelosta toiseen.

Viiteanalyysi antaa laajoja mahdollisuuksia tehdä erilaisia kirjastoa hyödyttäviä tutkimuksia. Tätä menetelmää kannattaa siis vakavasti harkita kokoelmien laadun arviointiin.

Lähteet

Ahtola, A. Anneli ja Juutistenaho, Riitta: Kokoelmien vahvuudet esiin - evaluointimenetelmät vertailtavina Tampereen ja Oulun yliopiston kirjastoissa. Signum 36(2003)3, ss. 48-52.

Butkovich, Nancy J.: Use studies: a selective review. Library resources & technical services 40(1996)4, ss. 359-368.

Gross, P.L.K. and Gross, E.M.: College libraries and chemical education. Science 66(1927)1713, ss. 385-389.

Kortelainen, Terttu: Kansallisesta kansainväliseksi : tutkimus suomalaisen tieteellisen lehden kansainvälisestä diffuusiosta. Oulu: Oulun yliopisto, Informaatiotutkimuksen laitos, 1999. Väitöskirja. Acta Universitatis Ouluensis. Series B, Humaniora, n:o 31. 152 s. Saatavissa elektronisena osoitteesta herkules oulu.fi/isbn9514253361/

Kortelainen, Terttu: Lehti leviää maailmalle - väitöskirja viittausanalyysiä käyttäen. Signum 33(2000)3, ss. 64-66.

MacRoberts M.H. and MacRoberts, Barbara R.: Problems of citation analysis. Scientometrics 36(1996)3, ss. 435-444.

Pappila, Jyrki: Humanistisen kokoelman evaluointi - tutkimuskirjallisuuden saatavuus Helsingin yliopiston kirjastossa. Kansalliskirjasto - Helsingin yliopiston kirjaston tiedotuslehti (2004)5-6, ss. 7-10.

Ari Muhonen toimii Helsingin teknillisen Korkeakoulun kirjaston johtajana. Email Ari.muhonen@hut.fi

