

Tieteenalojen välisiä eroja

- [Tieteenalaluokitukset](#)
- [Tekijöiden järjestys voi kertoa eri tieteenaloilla eri rooleista](#)
- [Tieteenala vaikuttaa viittausten määrään ja karttumisnopeuteen](#)
 - Käytettyjen lähteiden määrä vaihtelee aloittain
 - Eri tieteenaloilla viitataan eri ikäisiin lähteisiin
 - Viittaustietokannat eivät kata yhtäläisesti kaikkien tieteenalojen julkaisuja
 - Tekijöiden lukumäärä vaihtelee tieteenaloittain
 - Tieteenalojen keskimääräiset viittauskertymät vaihtelevat siis syistä, jotka eivät liity tutkimuksen laatuun tai vaikuttavuuteen
 - [Lähteet](#)



- Tieteenalaluokituksia on useita, julkaisun sijoittuminen tieteenalalle voi vaihdella riippuen käytetystä luokituksesta.
- Tieteenalojen julkaisukäytännöt poikkeavat toisistaan tavoilla, jotka heijastuvat sekä julkaisujen lukumääriin viittaustietokannoissa että näille karttuviin viittauksiin. Julkaisumetriikan tuottamisen kannalta eroja aiheuttavat erityisesti julkaisemisen kieli ja muoto (kirja, artikkeli, konferenssijulkaisut, jne.) sekä tieteenalalle tyypilliset viittauskäytännöt.

Tieteenalaluokitukset

Julkaisumetriikassa erilaiset tieteenalaluokitukset ovat keskeisiä työvälineitä. Erilaisia tieteenalaluokituksia on runsaasti ja niissä julkaisuja luokitellaan tieteenaloihin erilaisin perustein. Samat julkaisut voivat kuulua eri tieteen- tai tutkimusaloille eri tavoin riippuen siitä miten alat on määritelty.

Suomessa käytetään [tilastokeskuksen tieteenalaluokitusta](#) mm. Opetus- ja kulttuuriministeriön julkaisutiedonkeruussa. Kukin julkaisu sijoitetaan yhteen tai useampaan luokkaan joko julkaisun tekijän tai tutkimustietojärjestelmän pääkäyttäjän toimesta. Kyseessä on siis julkaisun asiasisältöön perustuva luokitus, jonka julkaisutiedon tallentaja määrittelee. Tätä luokitusta käytetään mm. [Vipusen](#) raporteissa ja [Tiedejatutkimus.fi](#) -sivustolla.

[Kansainvälisissä viittaustietokannoissa](#) julkaisuja ei yleensä sijoiteta tieteenaloille yksitellen, vaan tavallisesti joko **julkaisusarjan** tieteenalan perusteella tai artikkelin **viittaussuhteisiin** pohjautuen. Näiden lisäksi voidaan muodostaa tutkimusaloja **hakulauseen kaltaisilla algoritmeilla**, SciVal'in *Kestävän kehityksen tavoitteet (SDG Research areas)* on esimerkki tällaisesta luokittelusta.

Tieteenalan määrittely julkaisusarjan perusteella on suurimmissa viittaustietokannoissa toistaiseksi yleisintä. Kullekin sarjalle on tunnistettu yksi tai useampia tieteenaloja, ja kaikki sarjassa julkaistavat artikkelit sijoittuvat niihin automaattisesti. Käyttäjän kannalta järjestelmän hyvänä puolena on luokitteluperusteen läpinäkyvyys. Haittana on kuitenkin se, että luokittelutapa on hyvin karkea, ja sarjassa julkaistujen artikkeleiden tutkimusalojen kirjo on yleensä suurempi kuin sarjalle määriteltyt alat. Pienet tieteenalat tulevat usein pienuutensa vuoksi luokitelluksi osaksi suurempaa tieteenalaa, mikä aiheuttaa ongelmia normalisoitujen indikaattorien luotettavuudelle näiden kohdalla. Julkaisuja ilmestyy myös muissa kuin omalle tutkimusalalleen määritetyissä sarjoissa. Ongelmana ovat myös monitieteelliseksi luokitellut sarjat (multidisciplinary), joissa julkaistut artikkelit jäävät useimmissa sarjapohjaisissa luokitteluissa sijoittamatta omalle tieteenalalleen. Erityisen ongelmallista tämä on siksi, että moni suurinta huomiota saavista artikkeleista julkaistaan nimenomaan monitieteellisissä lehdissä (esim. Nature ja Science) (Szomszor 2021). Julkaisusarjoihin perustuvia luokitteluja on suuri määrä, ja viittaustietokannoissa on rakennettu niitä myös eräiden kansallisten tutkimuksen arviointien tarpeisiin.

Viime vuosina sarjapohjaisen luokittelun rinnalle on noussut julkaisujen keskinäisiin viittaussuhteisiin perustuva luokittelu. Suurimmilla analyysityökaluilla on omat viittaussuhteisiin perustuvat luokittelunsa, SciVal'in *Topic clusters* sekä InCites'in *Citation topics*, joka perustuu CWTS Leidenin kehittämään algoritmiin. Algoritmit tunnistavat toisiinsa viittaavista julkaisuista muodostuvia klustereita, joista muodostetaan tutkimusaloja. Viittaussuhteisiin perustuva luokittelu mahdollistaa hienosyisemmän luokittelun, joka huomioi paremmin artikkelin aihepiirin ja kontekstin. Käyttäjän kannalta haittapuolena on se, että luokitteluperusteet eivät ole avoimesti näkyvillä.

Tekijöiden järjestys voi kertoa eri tieteenaloilla eri rooleista

Julkaisun tekijöiden kirjaamisjärjestys kertoo [tekijöiden roolista](#) julkaisun kirjoittamisessa, ja sitä voidaan hyödyntää tarkastelun kohdentamisessa esimerkiksi julkaisuihin, joissa tarkasteltava taho on ollut johtavana tekijänä. Tieteenalojen käytännöt tekijöiden merkitsemisjärjestyksen suhteen vaihtelevat.

Useimmissa luonnontieteissä, lääketieteessä ja tekniikan aloilla kirjoittajat merkitään julkaisun tekemiseen annetun työpanoksen mukaisessa järjestyksessä, ensimmäiseksi merkitään pääkirjoittaja eli tekijä, jolla on ollut suurin rooli julkaisun kirjoittamisessa. Viimeiseksi merkitään tutkimushankkeen johtaja. Muista luonnontieteistä poiketen, matematiikan ja joidenkin fysiikan alan julkaisuissa kirjoittajat on tapana merkitä aakkosjärjestykseen kirjoittajien työpanoksesta riippumatta. Yhteiskuntatieteissä kirjoittajat voidaan merkitä joko työpanokseen perustuvaan järjestykseen tai aakkosjärjestykseen tieteenalasta riippuen. Esimerkiksi psykologian julkaisuissa kirjoittajat merkitään työpanoksen mukaiseen järjestykseen, mutta monilla muilla aloilla aakkosjärjestykseen. Humanistisilla aloilla kirjoittajat merkitään yleensä aakkosjärjestykseen. (Puuska and Miettinen 2008, s. 43-51).

Tieteenala vaikuttaa viittausten määrään ja karttumisnopeuteen

Käytettyjen lähteiden määrä vaihtelee aloittain

Teoriassa kullakin tieteenalalla on jaettavana se määrä viittauksia, jonka alan julkaisut tekevät. Tällöin viittausten määrä riippuisi siis siitä, kuinka moneen lähteeseen kukin alan julkaisu viittaa eli lähdeluettelon pituudesta (mm. Garfield 1979, Moed et al. 1985). Garfield kutsui tätä tieteenalan viittauspotentiaaliksi (*citation potential*). Käytettyjen lähteiden määrä vaihtelee tieteenaloittain ja voitaisiin olettaa, että tieteenaloilla, joilla lähdeluettelot ovat pitkiä, viitatuksi tulemisen todennäköisyys olisi suurempi kuin aloilla, joilla ne ovat lyhyitä. Tämä oletus pitäisi kuitenkin paikkansa vain, jos kaikilla tieteenaloilla olisi yhtä paljon viittauksia tekeviä julkaisuja, kaikki viittaukset tehtäisiin omalla alalla ilmestyneisiin julkaisuihin ja kaikki lähdeluetteloissa esiintyvät viittaukset olisivat laskettavissa. Näin ei ole.

Käytännössä viittaukset lasketaan viittaustietokannoista, niiden ulkopuolelle jääviin julkaisuihin tehdyt viittaukset jäävät laskematta. Niinpä lähdeluetteloiden pituutta merkittävämpi vaikutus tutkimusalueelle tyypillisiin viittausmääriin onkin sillä, kuinka paljon kyseisellä alalla julkaistaan viittaustietokantojen indeksoimissa sarjoissa ja kirjoissa ja kuinka suuri osa viittauksista tehdään näihin (Dorta-González and Dorta-González 2013, Patience et al. 2017).

Eri tieteenaloilla viitataan eri ikäisiin lähteisiin

Eri tieteenalojen viittauskäytännöissä on eroa sen suhteen minkä ikäisiä lähteitä käytetään. Kulloinkin jaossa olevaan viittausten määrään vaikuttaa se kuinka tuoreisiin artikkeleihin tieteenalalla on tapana viitata, ja toisaalta kuinka pitkään ilmestymisensä jälkeen artikkeleihin edelleen viitataan – myös tämä vaihtelee tieteenaloittain ja vaikuttaa alalle tyypillisiin viittauskertymiin (Moed et al. 1985, Dorta-González and Dorta-González 2013, Patience et al. 2017). Tuoreisiin artikkeleihin viittaaminen näyttää lisäävän tieteenalalla kattuvia viittauksia (Patience et al. 2017).

Jos analyyseissä käytetään hyvin lyhyttä [viittausikkunaa](#), saattavat laskematta jäävät viittaukset painottua tietyille tieteenaloille (Abramo 2011).

Viittaustietokannat eivät kata yhtäläisesti kaikkien tieteenalojen julkaisuja

Viittaustietokannat kattavat eri tieteenalojen julkaisut hyvin vaihtelevasti. Yleisesti ottaen luonnontieteelliset ja lääketieteelliset julkaisut löytyvät suurimmista viittaustietokannoista (Web of Science, Scopus, Dimensions) selvästi kattavammin kuin yhteiskuntatieteelliset tai humanistiset julkaisut. Luonnontieteeseen, lääketieteeseen ja yhteiskuntatieteeseen kuuluu kuitenkin sekä hyvin että heikommin viittaustietokantojen kattamia tutkimusaloja. Tietokantojen kattavuuksista löytyy lisää tietoa oppaan luvusta [Käytetyimpiä monitieteisiä viittaustietokantoja](#).

Erot kattavuudessa johtuvat mm. siitä että viittaustietokannat (Web of Science ja Scopus) indeksoivat **tieteellisiä artikkeleita** julkaisevia sarjoja **kattavammin kuin muita julkaisukanavia**. Osalla tieteenaloista merkittävä osa tutkimuksesta julkaistaan kuitenkin artikkeleiden sijasta kirjoissa (esim. monet yhteiskunta- ja humanistiset tieteet) (Larivière et al. 2006) tai konferenssijulkaisuissa (tietojenkäsittelytiede) (Freyne et al. 2010). Suuri osa näiden tutkimusalojen julkaisuista jää viittaustietokantojen ulkopuolelle.

Artikkeleita julkaisevien tieteellisten lehtienkään kattavuus viittaustietokannoissa ei ole sama kaikille tieteenaloille. Eniten käytetyt viittaustietokannat (Web of Science ja Scopus) indeksoivat **vain lehtiä, joissa vähintään otsikko ja tiivistelmä ovat englanninkielisiä**. Siksi pääosa esimerkiksi suomenkielisistä lehdistä jää niiden ulkopuolelle. Englanniksi julkaiseminen on tavallisempaa mm. luonnontieteissä ja lääketieteessä kuin monissa yhteiskunta- ja humanistisissa tieteissä, joissa tehdään paikka- ja kulttuurisidonnaisempaa tutkimusta. **Julkaisukieli** siis selittää osaltaan yhteiskuntatieteiden sekä taiteiden ja humanististen tieteiden heikompaa kattavuutta viittaustietokannoissa.

Näistä syistä viittausanalyysien tekeminen sellaisille tieteenaloille, jotka julkaisevat pääosan tutkimustuloksistaan artikkeleina kansainvälisissä (englanninkielisissä) tieteellisissä sarjoissa, on yksinkertaisempaa kuin tieteenaloille, joilla julkaistaan lähinnä suomeksi tai muissa muodoissa kuin tieteellisinä artikkeleina (esim. kirjat tai konferenssijulkaisut). Viittaustietokantojen heikommin kattamien tieteenalojen julkaisuanalyyseissä voidaan käyttää tutkimusorganisaatioiden omia [tutkimustietojärjestelmiä](#). Tällöin tarkastelu ei perustu viittaustietoihin, vaan julkaisujen muihin ominaisuuksiin kuten julkaisukieli tai kansallisten ja kansainvälisten yhteisjulkaisujen määrät ja osuudet.

Tekijöiden lukumäärä vaihtelee tieteenaloittain

Artikkelien kirjoittajien lukumäärässä on tutkimusalaakohtaisia eroja. Esimerkiksi tähtitieteen, astro-, ydin- ja hiukkasfysiikan artikkelilla voi usein olla yli 100 kirjoittajaa, kun taas historian ja kirjallisuudentutkimuksen tutkimusjulkaisuilla on tyypillisesti vain yksi kirjoittaja (Patience et al. 2017). Suuria kirjoittajajoukkoja esiintyy erityisesti tieteenaloilla, joilla tutkimus edellyttää suuria laboratorioita, tai muuta infrastruktuuria tai laajaa yhteistyötä esim. aineistojen jakamisessa. Kirjoittajamäärien erot heijastuvat tieteenalan keskimääräisiin artikkelikohtaisiin viittauksiin, sillä tekijöiden lukumäärän kasvun on havaittu lisäävän artikkelin saamia viittauksia (Adams et al. 2019, Patience et al. 2017).

Tieteenalojen keskimääräiset viittauskertymät vaihtelevat siis syistä, jotka eivät liity tutkimuksen laatuun tai vaikuttavuuteen

Viittausten kokonaismäärien vaihtelusta seuraa, että eri tieteenaloja tai niillä toimivia organisaatioita tai tutkijoita ei voi viittausten perusteella verrata toisiinsa tasapuolisesti (mm. Garfield 1979, Moed et al. 1985, Dorta-González and Dorta-González 2013, Mongeon and Paul-Hus 2016) muuten kuin [normaalisoiduilla](#) indikaattoreilla. Esimerkiksi puhtaasti yhteiskunnallista tutkimusta tekevän tutkijan [h-indeksi](#) ei käytännössä voi nousta läheskään yhtä korkeaksi kuin vastaavassa uravaiheessa olevan genetiikan tutkimusta tekevän tutkijan.

ESI Research Field	minimum number of citations received by the top 1% of papers from 2018 (2022)	ESI Research Field	minimum number of citations received by the top 1% of papers from 2018 (2022)
MOLECULAR BIOLOGY & GENETICS	164	COMPUTER SCIENCE	86
MULTIDISCIPLINARY	143	ENGINEERING	81
MATERIALS SCIENCE	139	PHYSICS	81

IMMUNOLOGY	113	PHARMACOLOGY & TOXICOLOGY	75
CHEMISTRY	104	GEOSCIENCES	69
SPACE SCIENCE	100	AGRICULTURAL SCIENCES	64
BIOLOGY & BIOCHEMISTRY	98	PSYCHIATRY /PSYCOLOGY	64
ENVIRONMENT /ECOLOGY	96	ECONOMICS & BUSINESS	63
MICROBIOLOGY	96	PLANT & ANIMAL SCIENCE	52
NEUROSCIENCE & BEHAVIOR	94	SOCIAL SCIENCES, GENERAL	49
CLINICAL MEDICINE	86	MATHEMATICS	36

Taulukko 1. Vuonna 2022 käytettävät kynnysarvot kullakin Web of Sciencen [ESI \(Essential Science Indicators\)](#) tieteenalalla neljä vuotta aikaisemmin (2018) ilmestyneille [Highly Cited Paper](#) artikkeleille. Highly Cited Paper merkinnän saavat artikkelit, jotka kuuluvat oman alansa ilmestymisvuonnaan eniten viittauksia saaneen 1 prosentin joukkoon. Luvuista voidaan päätellä, että jaettavien viittausten määrä vaihtelee tutkimusaloittain. Esimerkiksi Molekyylibiologian ja genetiikan lehdissä v. 2018 ilmestyneen artikkelin on saatava vähintään 164 viittausta ollakseen eniten viitatus prosentin joukossa vuonna 2022. Kun taas matematiikan alalla artikkelit jotka ovat saaneet yli 36 viittausta kuuluvat viitatuimpaan prosenttiin.

Tiedot poimittu Clarivaten palvelusta 9.3.2022 <https://esi.clarivate.com/ThresholdsAction.action>

Lähteet

Adams, J., Pendlebury, D., Potter, R. and Szomszor, M. (2019) *Global Research Report – Multi-authorship and research analytics*. Institute for Scientific Information, Clarivate, London and Philadelphia.

Dorta-González, P. and Dorta-González, M.I. (2013) Comparing journals from different fields of science and social science through a JCR subject categories normalized impact factor. *Scientometrics* 95, pp. 645–672. Saatavilla: <https://doi.org/10.1007/s11192-012-0929-9>

Freyne, J., Coyle, L., Smyth, B. and Cunningham, P. (2010) A quantitative evaluation of the relative status of journal and conference publications in computer science. *Communications of the ACM*, 53(11), pp. 124–132. Saatavilla: <http://dx.doi.org/10.1145/1839676.1839701>

Garfield, E. (1979) Is citation analysis a legitimate evaluation tool? *Scientometrics*, 1(4), pp. 359–375.

Larivière, V., Archambault, E., Gingras, Y. and Vignola-Gagné, E. (2006) The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology*, 57(8), pp. 997–1004. Saatavilla: <https://doi.org/10.1002/asi.20349>

Moed, H.F., Burger, W.J.M., Frankfort, J.G. and van Raan, A.F.J. (1985) The application of bibliometric indicators: important field- and time-dependent factors to be considered. *Scientometrics*, 8(3-4), pp. 177–203.

Mongeon, P. and Paul-Hus, A. (2016) The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics* 106, pp. 213–228. Saatavilla: <https://doi.org/10.1007/s11192-015-1765-5>

Patience, G., Patience, C., Blais, B. and Bertrand, F. (2017) Citation analysis of scientific categories. *Heliyon*, 3(5), e00300. Saatavilla: <https://doi.org/10.1016/j.heliyon.2017.e00300>

Puuska, H-M. and Miettinen, M. (2008) Julkaisukäytännöt eri tieteenaloilla. *Opetusministeriön julkaisuja*, 33, 111 p. Saatavilla: <https://core.ac.uk/download/pdf/198192223.pdf>

Szomszor, M., Adams, J., Pendlebury, D. and Rogers, G. (2021) *Global Research Report – Data categorization: Understanding choices and outcomes*. Institute for Scientific Information, Clarivate, London and Philadelphia.