

Viittausvaikuttavuusindikaattoreita

Tieteellistä vaikuttavuutta kuvaavat indikaattorit perustuvat pääasiallisesti julkaisujen saamiin viittausmääriin. Julkaisujen saama viittausmäärä, julkaisujoukon viittausten keskiarvo ja h-indeksi ovat esimerkkejä viittausvaikuttavuusindikaattoreista. Tieteenalojen väliset erot viittauskäytännöissä, julkaisujen ikä ja julkaisutyyppi on pyritty ottamaan huomioon normalisoiduissa viittausvaikuttavuusindikaattoreissa.

- [Viittausten lukumäärä](#)
 - [Vastuullinen käyttö](#)
- [Keskimääräinen viittausmäärä](#)
 - [Vastuullinen käyttö](#)
- [Normalisoidut indikaattorit](#)
 - [Tieteenalanormalisoidut Top x% -indikaattorit](#)
 - [Top x% -indeksit](#)
 - [Top x% -osuudet](#)
 - [Top x% -julkaisut](#)
 - [Vastuullinen käyttö](#)
 - [Tieteenalanormalisoidut viittausindikaattorit](#)
 - [Vastuullinen käyttö](#)
 - [Normalisoitujen indikaattoreiden toteutuksia tietokannoissa ja muissa lähteissä](#)
- [H-indeksi](#)
 - [Vastuullinen käyttö](#)
 - [H-indeksin toteutuksia tietokannoissa](#)

Viittausten lukumäärä

Julkaisuun tehtyjen viittausten määrä kuvastaa sitä, kuinka paljon sitä on hyödynnetty ilmestymisensä jälkeen tehdyssä tutkimuksessa. Viittausten voidaan ajatella kertovan julkaisussa esitetyn tutkimuksen hyödyllisyydestä tiedeyhteisölle, ja siten sen vaikuttavuudesta. Vaikuttavuus on kuitenkin käsitteenä laaja ja se voi ilmetä monin eri tavoin, joten käytämme tässä oppaassa käsitettä viittausvaikuttavuus. Lue lisää viittausten lukumäärästä vaikuttavuuden indikaattorina luvusta [Viittaukset](#).

Vastuullinen käyttö

- Ota huomioon tieteenalojen erilaiset julkaisu- ja viittauskäytännöt ja julkaisujen ilmestymisajankohta.
- Viittausmäärä on riippuvainen käytetystä tietokannasta. Eri tieteenalojen julkaisujen kattavuus eri viittaustietokannoissa voi vaihdella paljon.
- Katso myös vastuullisuuden arvioimiseen liittyvät kysymykset [Indikaattoreita](#)-sivulta.

Keskimääräinen viittausmäärä

Viittausten keskiarvo voidaan laskea erilaisille julkaisujoukoille. Keskiarvoja käytetään mm. normalisoiduissa indikaattoreissa. Haasteita keskiarvon käytölle aiheuttavat vinot viittausjakamat, jolloin keskiarvo ei välttämättä ole kuvaava indikaattori.

Vastuullinen käyttö

- Ota huomioon tieteenalojen erilaiset julkaisu- ja viittauskäytännöt ja julkaisujen ilmestymisajankohta.
- Viittausmäärä on riippuvainen käytetystä tietokannasta. Eri tieteenalojen julkaisujen kattavuus eri viittaustietokannoissa voi vaihdella paljon.
- Tutki kuinka hyvin keskiarvo kuvaa aineistoasi. Keskiarvo ei useinkaan kuvaa viittausten jakautumista, koska viittausjakamat ovat tavallisesti voimakkaasti vinoja, jolloin keskiarvo ja mediaani eroavat usein toisistaan paljon.
- Yksittäiset paljon viittauksia saaneet julkaisut voivat vaikuttaa keskiarvoon huomattavasti. Näitä julkaisuja esiintyy usein.
- Katso myös vastuullisuuden arvioimiseen liittyvät kysymykset [Indikaattoreita](#)-sivulta.

Normalisoidut indikaattorit

Normalisoiduilla indikaattoreilla tarkoitetaan indikaattoreita, joiden laskennassa on pyritty huomioimaan julkaisun tieteenala, ikä ja julkaisutyyppi. Normaalisoinnissa keskeistä on julkaisun tieteenalan määrittely. Julkaisut luokitellaan tietokannoissa erilaisiin tieteenalaluokituksiin erilaisin perustein ja menetelmin. Monissa kansainvälisissä tietokannoissa käytetään julkaisusarjojen tieteenalaluokitukseen perustuvaa luokitusta. Tällöin julkaisusarjassa (esim. lehdessä) ilmestynyt julkaisu (esim. tieteellinen artikkeli) katsotaan kuuluvaksi samaan luokkaan kuin mihin julkaisusarja kuuluu. Lisätietoa tieteenalaluokituksista on saatavilla luvusta [Tieteenalojen välisiä eroja](#).

Tietokannassa saman ikäisten, samaa tieteenalaa ja samaa julkaisutyyppiä edustavien julkaisujen saamista viittaasmääristä muodostetaan viittausjakauma ja lasketaan keskiarvo. Normalisointi toteutetaan näiden viittausjakaumien ja keskiarvojen avulla. Tieteenalanormalisoiduilla indikaattoreilla pyritään mahdollistamaan eri alojen julkaisujen vertailu. Alla esitellään erilaisia tieteenalanormalisoituja indikaattoreita.

Tieteenalanormalisoidut Top x% -indikaattorit

Tieteenalanormalisoidut Top x%-indikaattorit perustuvat tietokannan tieteenalakohtaisiin viittausjakaumiin. Top x% -indikaattorit ilmaisevat, kuinka suuri suhteellinen osuus tarkasteltavasta julkaisujoukosta kuuluu saman tieteenalan, julkaisuvuoden ja -tyypin viitatuimpaan x prosenttiin. Niitä käytetään korkean viittausvaikuttavuuden indikaattoreina, esimerkiksi Top 10% - ja Top 1%-osuudet kertovat korkeaan (10%) että huippuvaikuttavuuteen (1%) yltävien julkaisujen suhteellisesta määrästä julkaisujoukossa.

Top x% -indeksit

Top x%-indeksit perustuvat tietokannan tieteenalakohtaisiin viittausjakaumiin. Esimerkiksi Top 1% -indeksin ykkösen ylittävä arvo tarkoittaa, että tarkasteltavan joukon julkaisuista enemmän kuin 1% kuuluu oman tieteenalansa viitatuimman 1% joukkoon.

Esimerkki kuinka Top 10% -indeksi lasketaan

- Muodostetaan viittausjakauma: Tieteenalalla Y on ilmestynyt 25000 julkaisua vuonna 2014. Viittausjakaumassa Top 10%:iin kuuluu siis 2500 julkaisua.
- Tarkasteltava julkaisujoukko: Samalla tieteenalalla toimii tiedekunta Z. Tiedekunnan tutkijat ovat julkaisseet vuonna 2014 500 julkaisua. Julkaisusta 100 kuuluu edellämainitun viittausjakauman top 10%:iin. Tilastollisesti odotusarvo tiedekunnan top 10%:ssä olevien julkaisuiden määrälle on 50 eli 10% tiedekunnan julkaisuista.
- Top 10% indeksin laskeminen: Top 10% indeksi lasketaan jakamalla laitoksen top 10% kuuluvien julkaisuiden lukumäärä odotusarvolla eli $100/50=2$.
- huomioi että esimerkki on yksinkertainen, todellisuudessa julkaisut kuuluvat eri tieteenaloille ja ovat ilmestyneet eri vuosina - tarkasteltavia julkaisuja verrataan siis moneen viittausjakaumaan. Lisäksi Top x% -luokka sisältää harvoin juuri täsmälleen x% julkaisuista, esimerkiksi Top 10% -luokka sisältää tuskin koskaan täsmälleen 10 % julkaisuista. Tämä johtuu siitä että viittaasmäärä on diskreetti muuttuja - se saa vain kokonaislukuarvoja.

Top x% -osuudet

Top x%-osuudet perustuvat tietokannan tieteenalakohtaisiin viittausjakaumiin. Top x% -osuus ilmaisee, kuinka suuri %-osuus tarkasteltavasta julkaisujoukosta kuuluu saman tieteenalan, julkaisuvuoden ja -tyypin viitatuimpaan x prosenttiin. Esimerkiksi mikäli julkaisujen Top 1% -osuus on 2% tarkoittaa se, että tarkasteltavan joukon julkaisuista 2% kuuluu oman tieteenalansa viitatuimman 1% joukkoon.

Esimerkki kuinka Top 10% -osuus lasketaan

- Muodostetaan viittausjakauma: Tieteenalalla Y on ilmestynyt 25000 julkaisua vuonna 2014. Viittausjakaumassa Top 10%:iin kuuluu siis 2500 julkaisua.
- Tarkasteltava julkaisujoukko: Samalla tieteenalalla toimii tiedekunta Z. Tiedekunnan tutkijat ovat julkaisseet vuonna 2014 500 julkaisua. Julkaisuista 100 kuuluu edellämainitun viittausjakauman Top 10%:iin.
- Top 10% osuuden laskeminen: lasketaan tiedekunnan Top 10% kuuluvien julkaisuiden prosenttiosuus kaikista tiedekunnan julkaisuista: $100/500 * 100\% = 20\%$
- huomioi että esimerkki on yksinkertainen, todellisuudessa julkaisut kuuluvat eri tieteenaloille ja ovat ilmestyneet eri vuosina - tarkasteltavia julkaisuja verrataan siis moneen viittausjakaumaan. Lisäksi Top x% -luokka sisältää harvoin juuri täsmälleen x% julkaisuista, esimerkiksi Top 10% -luokka sisältää tuskin koskaan täsmälleen 10 % julkaisuista. Tämä johtuu siitä että viittaasmäärä on diskreetti muuttuja - se saa vain kokonaislukuarvoja.

Top x% -julkaisut

Essential Science Indicator - ja Web of Science -tietokannan Highly Cited Papers- ja Hot Papers-merkinnät ovat yksi versio Top x%-indikaattoreista. Ne perustuvat julkaisujen tieteenalakohtaisiin viittausjakaumiin. Highly Cited Papers- ja Hot papers- statuksen saavat julkaisut, jotka ylittävät tietokannassa määritellyt viittausten raja-arvot.

Highly Cited papers merkinnän saavat julkaisut, jotka ovat korkeintaan 10 vuotta vanhoja ja jotka ylittävät Top 1% raja-arvon samalla alalla samana vuonna ilmestyneiden julkaisujen viittausjakaumassa.

Hot Papers merkinnän saavat julkaisut, jotka ovat korkeintaan kaksi vuotta vanhoja ja jotka ylittävät Top 0,1% raja-arvon samalla alalla samana päivityskertana tietokantaan lisättyjen julkaisujen viittausjakaumassa. Päivitykset tehdään kahden kuukauden välein. Viittauksia laskettaessa otetaan huomioon ainoastaan viimeisimmässä päivityksessä lisättyjen julkaisujen tekemät viittaukset.

Vastuullinen käyttö

Top x% -indikaattorit pyrkivät ottamaan huomioon tieteenalojen väliset erot julkaisu- ja viittauskäytännöissä ja julkaisuajankohdan. Ne eivät ole herkkiä yksittäisille erittäin paljon viitatuille julkaisuille.

Ota huomioon, että

- indikaattorit arvottavat vain pienen osan julkaisuista, muut julkaisut jätetään huomiotta.
- rajanveto tiettyyn prosenttiin on mielivaltaisen, Top 1% ja Top 10% sijasta voisimme yhtä hyvin tarkastella Top 5% ja Top 15%.
- perustuvat tieteenalaluokituksiin, jotka eivät ole ongelmattomia (lisätietoja kohdasta [Tieteenalaluokitukset](#)).
- verrattain tuoreiden julkaisujen (alle 5 vuotta vanhat julkaisut) Top x%-indikaattoreita tarkasteltaessa on huomioitava, että joillain aloilla viittauskertymät ovat pieniä. Tällöin mielekästä jakaumaa, jonka perusteella Top x%-indikaattori voitaisiin laskea ei vielä ole olemassa.

- eri julkaisutyyppihin kertyy viittauksia eri tavoin, esimerkiksi konferenssijulkaisu voi päästä mukaan Top x% julkaisuihin huomattavasti pienemmällä viittausmäärällä kuin lehtiartikkeli.
- tutustu myös vastuullisuuden arvioimiseen liittyviin kysymyksiin [Indikaattoreita](#)-sivulta.

Tieteenalanormalisoidut viittausindikaattorit

Tieteenalanormalisoidut viittausindikaattorit kertovat julkaisujoukon viittausvaikuttavuudesta tieteenaloillaan. Tieteenalanormalisoiduilla indikaattoreilla pyritään mahdollistamaan eri alojen julkaisujen viittausvaikuttavuuden vertailu.

Nämä indikaattorit kuvaavat julkaisun tai julkaisujoukon saaman viittausmäärän suhdetta samalla tutkimusalalla, samana vuonna ilmestyneiden samaa julkaisutyyppiä olevien julkaisujen viittausten keskiarvoon. Keskimääräistä viittausvaikuttavuutta vastaava arvo on 1. Sen alittavat arvot kertovat keskimääräistä pienemmästä viittausmäärästä ja sen ylittävät keskimääräistä suuremmasta viittausmäärästä. Kun yksikkö saa esimerkiksi arvon 1,78, merkitsee se 78 prosenttia maailman keskiarvoa suurempaa viittausmäärää.

Esim. kuinka Tieteenalanormalisoidut viittausindikaattorit lasketaan

Esimerkki 1. Katsausartikkeli on saanut 15 viittausta. Saman aihepiirin lehdissä, samana vuonna ilmestyneet katsausartikkelit ovat saaneet keskimäärin 12 viittausta. Katsausartikkelin normalisoitu viittausindeksi lasketaan: $15/12 = 1,25$.

Esimerkki 2. Kun lasketaan normalisoitua viittausindeksiä julkaisujoukolle, yksittäisten julkaisujen viittausindeksit täytyy yhdistää. Lasketaan tieteenalanormalisoitu viittausindeksi kolmen julkaisun joukolle.

Julkaisu	artikkelin saamat viittaukset	samanlaisten artikkelien viittausten keskiarvo	Artikkelin viittausindeksi
1.	15	12	$15/12 = 1,25$
2.	100	85	$100/85 = 1,17$
3.	60	30	$60/30 = 2$
Julkaisujoukon viittausindeksi on $(1,25 + 1,17 + 2)/3 = 1,47 = 1,5$			

Vastuullinen käyttö

Tieteenalanormalisoidut viittausindikaattorit pyrkivät ottamaan huomioon tieteenalojen väliset erot julkaisu- ja viittauskäytännöissä ja julkaisuajankohdan.

Ota huomioon, että

- perustuvat keskiarvoihin, jolloin yksittäiset paljon viittauksia saaneet julkaisut voivat vaikuttaa indikaattorin arvoon huomattavasti.
- verrattain tuoreiden julkaisujen tarkasteltaessa on huomioitava, että joillain aloilla viittauskertymät ovat pieniä, jolloin pienet muutokset viittausmäärissä johtavat suuriin muutoksiin indikaattorin arvossa.
- perustuvat tieteenalaluokituksiin, jotka eivät ole ongelmattomia (lisätietoja kappaleesta [Tieteenalaluokitukset](#))
- katso myös vastuullisuuden arvioimiseen liittyvät kysymykset [Indikaattoreita](#)-sivulta

Normalisoitujen indikaattoreiden toteutuksia tietokannoissa ja muissa lähteissä

Top x% indeksit	Tietokanta /lähde	Huomioita menetelmästä	Linkki menetelmän kuvaukseen
PP(top x%)	Leiden ranking		Indicators -The CWTS Leiden Ranking 2021
top 10-indeksi	Tieteen tila		Concepts related to bibliometric analyses Scroll down the page for version in english.
Top x% osuudet			
% Documents in Top 1% % Documents in Top 10%	InCites		Katso InCites: Understanding the Metrics välilehti Percentiles Myös: Documents in Top 1% and 10%
Outputs in Top Citation Percentiles	SciVal	Tieteenalanormalisoimaton Outputs in Top Citation Percentile - indikaattori perustuu viittauksiin. Tieteenalanormalisoitu Outputs in Top Citation Percentile määritetään julkaisujen FWCI-arvojen perusteella. FWCI-arvojen viittausikunnan pituus on julkaisuvuosi ja tätä seuraavat kolme täyttä vuotta.	katso SciVal Metric: Outputs in Top Citation Percentiles Yrjö Leino & Marianne Gauffriaun. Guest Post: Understanding SciVal's calculation of field-weighted percentile indicators – The Bibliomagician
Top % julkaisut			

Highly cited papers	Essential Science Indicators, Web of Science, InCites		Essential Science Indicators - Highly Cited Papers Highly Cited Thresholds
Hot papers	Essential Science Indicators, Web of Science, InCites.		Hot Paper Thresholds
Normalisoidut viittausindikaattorit	Tietokanta		
Category Normalized Citation Impact (CNCI)	InCites		InCites: Understanding the Metrics See tab "Normalized"
Field-Weighted Citation Impact (FWCI)	SciVal	FWCI-arvojen viittausikkunan pituus on julkaisuvuosi ja tätä seuraavat kolme täyttä vuotta.	Research Metrics Guidebook
Mean Normalized Citation Score (MNCS)	Leiden ranking		Indicators The CWTS Leiden Ranking 2021
The Field Citation Ratio (FCR)	Dimensions		What is the FCR?
Co-citation Percentile Rank (CPR)	JYUcite		JYUcite: What is CPR?

H-indeksi

Yksi tunnetuimmista bibliometrisistä indikaattoreista on h-indeksi, jota kutsutaan myös Hirsch-indeksiksi. Sen kehitti Jorge E. Hirsch vuonna 2005. Lue lisää alkuperäisestä [artikkelista](#).

Tutkijan h-indeksi määritellään seuraavasti: tutkijan h-indeksin arvo on h, jos h kappaletta hänen julkaisuistaan ovat jokainen saaneet vähintään h viittausta ja loput julkaisut ovat saaneet korkeintaan h viittausta.

Esimerkki.

Tutkijan h-indeksin arvo 10 tarkoittaa, että tutkijalla on 10 julkaisua, joihin kaikkiin on viitattu vähintään 10 kertaa. Hänen mahdollisiin muihin julkaisuihinsa on viitattu korkeintaan kymmenen kertaa tai ei lainkaan.

Esimerkki.

Tutkijoiden h-indeksin arvo voi olla sama vaikka heidän julkaisuprofiiliinsa (julkaisuidensa määrä ja niiden saamien viittausten määrä) poikkeaisivat toisistaan paljon.

- Tutkija 1 on julkaissut viisi artikkelia joihin on viitattu seuraavasti: 9,9,7,6,5
- Tutkija 2 on julkaissut 11 artikkelia, joihin on viitattu seuraavasti: 200, 150, 99, 70, 5, 4,4,3,2,1,1
- Kummankin tutkijan h-indeksin arvo on 5

Vastuullinen käyttö

- Ota huomioon tieteenalojen erilaiset julkaisu- ja viittauskäytännöt ja julkaisujen ilmestymisajankohta.
- H-indeksi on riippuvainen käytetystä tietokannasta. Eri tieteenalojen julkaisujen kattavuus eri viittaustietokannoissa voi vaihdella paljon.
- H-indeksi ei voi olla suurempi kuin julkaisujen lukumäärä. Esim. Jos tutkija on julkaissut 10 julkaisua, h-indeksin maksimiarvo on 10 vaikka kaikkiin kyseisen tutkijan julkaisuihin olisi viitattu satoja kertoja.
- H-indeksi on kumuloituva mittari, se ei voi pienentyä. Se voidaan määritellä myös lyhyemmälle tarkastelujaksolle (esim. 5 vuotta) kuin koko uralle.
- Julkaisujen tekijöiden lukumäärä vaikuttaa h-indeksiin. Se suosii aloja, joissa on suuret tutkimusryhmät ja tutkijoilla siten paljon julkaisuja.
- H-indeksissä ei huomioida kirjoittajien roolia julkaisussa. Keskeinen rooli kirjoitustyössä ja muut roolit vaikuttavat yhtä paljon.
- Tutkijan uran pituus vaikuttaa h-indeksiin. Pidempään julkaisseella tutkijalla on aloittelevaa tutkijaa suurempi todennäköisyys saada korkea h-indeksi, koska hänellä on ollut enemmän aikaa julkaista ja julkaisuilla on ollut enemmän aikaa saada viittauksia.
- H-indeksissä ei huomioida milloin tekijä on julkaissut tai onko hän lopettanut julkaisemisen. Lopettamisen jälkeen h-indeksi voi jatkaa kasvuaan viittausten kertymisen myötä.
- Tutkijan merkittäviä yksittäisiä julkaisuja ei tule esille h-indeksissä
- Katso myös vastuullisuuden arvioimiseen liittyvät kysymykset [Indikaattoreita](#)-sivulta.

H-indeksin toteutuksia tietokannoissa

Tietokanta	Menetelmän kuvaus
Web of Science	The h-index value is based on a list of publications ranked in descending order by the Times Cited count. An index of h means that there are h papers that have each been cited at least h times. Lue lisää Web of Science: h-index information
Scopus	A scientist has an index h if h of his/her Np papers has at least h citations each, and the other (Np - h) papers have no more than h citations each. Lue lisää The Scopus h-index, what's it all about?
Google Scholar	The h-index of a publication is the largest number h such that at least h articles in that publication were cited at least h times each. For example, a publication with five articles cited by, respectively, 17, 9, 6, 3, and 2, has the h-index of 3. Lue lisää Google Scholar metrics