

Tekoälyn huippuosaamisen kartoitus

Valtioneuvoston Tekoälyn kokonaiskuva ja osaamiskartoitus –selvityksen julkaisuanalyysi

Anssi Neuvonen, Heikki Ailisto (Teknologian tutkimuskeskus VTT Oy)

19.6.2018

Sisältö

Johdanto.....	1
1.1 Tutkimussuunnat	2
1.2 Näkökulmat ja menetelmät.....	3
1.3 Julkaisuanalyysit.....	5
1.3.1 Tekoäly yleisesti.....	5
1.3.2 Data-analyysi	15
1.3.3 Havainnointi ja tilannetietoisuus.....	22
1.3.4 Luonnollinen kieli ja kognitio	27
1.3.5 Vuorovaikutus ihmisen kanssa	32
1.3.6 Digitaaliset työelämässä, ongelmanratkaisu ja laskennallinen luovuus	35
1.3.7 Koneoppiminen	38
1.3.8 Järjestelmätaso ja systeemivaikutukset	44
1.3.9 Tekoälyn laskentaympäristöt, alustat ja palvelut, ekosysteemit	46
1.3.10 Robotiikka ja koneautomaatio - tekoälyn fyysinen ulottuvuus	48
1.3.11 Etiikka, moraali, regulaatio ja lainsäädäntö	53
1.4 Patentointi	57
1.4.1 Tekoäly	58
1.4.2 Koneoppiminen	60
1.4.3 Hermoverkot	63
1.5 Yhteenveto	66
1.6 LIITE: tiedonhaut	68

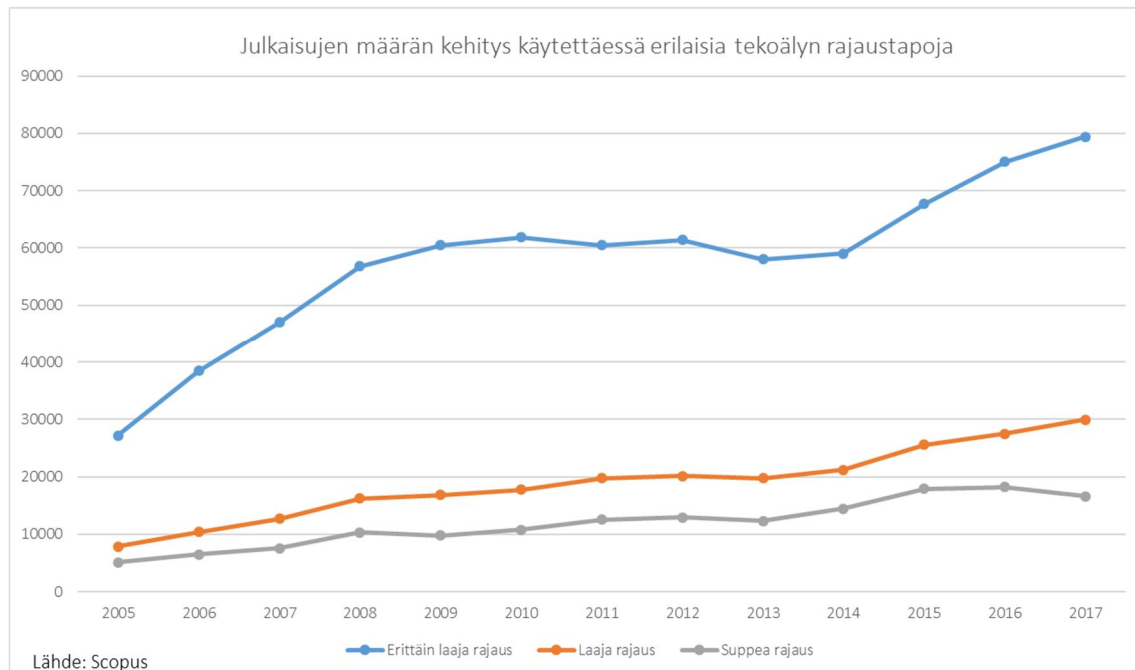
Johdanto

Teknisestä näkökulmasta tarkastellen tekoäly ei ole yksi teknologia, vaan joukko tekniikoita ja tutkimussuuntauksia.¹ Fei-Fei Li, Stanfordin yliopiston AI Labin johtaja, on verrannut tekoälyä salaattikulhoon, joka pitää sisällään eri osa-alueita ja tieteenaloja.² Tekoälyn tutkimuksen suuntauksilla on myös ollut erilainen asema eri aikoina kokonaisuuden sisällä. Kuten kaikessa tutkimustoiminnassa eri teemat ja suuntaukset voivat nousta esiin ja laskea eri aikoina. Näin on myös tekoälytutkimuksen piirissä.

¹ A Layered Model for AI Governance, Urs Gasser and Virgilio A.F. Almeida. Computing Edge, February 2018, Volume 4, Number 2, p. 15

² GrAI expectations, The Economist, vol 426, nr 9085, March 31 2018

Tekoälyn koostuessa erilaisista tutkimusaloista muuttuu myös tekoälyn ”ulkorajojen” määrittäminen haasteelliseksi. Mihin tekoäly päättyy? Rajanvedon ongelmista kertoo se, kuinka erilaisia julkaisumääriä eri määrittely- ja rajaustavoilla on mahdollista löytää. Maailmanlaajuisesti tekoälyä käsittelevien julkaisujen määrä vaihtelee tuhansista lähes sataan tuhanteen julkaisuun vuodessa riippuen pelkästään aiheen rajauksesta. Tämä vertautuu tilanteeseen, jossa yritettäisiin löytää globaalista tutkimuskentästä kaikki tietotekniikkaa käsittelevät tutkimusjulkaisut. Tekoälystä on vastaavalla tavalla muodostumassa vähitellen keskeinen mahdollistava teknologia laajasti koko tutkimuksen kentälle.



Kuva 1. Tekoälyä käsittelevien julkaisujen määrällinen kehitys 2003-2017 aiheen erilaisilla rajauksilla.

Tekoälylle löydetään jatkuvasti uusia sovelluskohteita. Yksi lähestymistapa tutkimuskirjallisuuteen onkin erottaa toisistaan sovellukset ja itse tekoälyn tutkimus. Johtuen tekoälymenetelmien ja -tekniikoiden luonteesta, eron tekeminen soveltavan ja puhtaan tutkimuksen välille on tekoälyn tapauksessa kuitenkin kenties vielä haastavampaa kuin tutkimusaloilla yleensä. Esimerkiksi tekoälyn kehittämisessä keskeisiä algoritmeja tyypillisesti muokataan ja sovelletaan uusien tutkimuskohteiden ollessa kyseessä.

Tekoälymenetelmien ja -tekniikoiden suuri määrä tarjoaa tutkimukselle jatkuvasti uusia mahdollisuuksia yhdistää ja soveltaa niitä eri tavoin. Pelkästään nopeasti kehittyvän koneoppimisen alueella on nykyisin erotettavissa jo lähes 150 eri algoritmista menetelmää.³ Tässä tilanteessa puhe tekoälyn perustutkimuksesta ja soveltavasta tutkimuksesta toisistaan erillisinä tutkimuskonteksteina, joilla on omat tutkimusparadigmansa ja -menetelmänsä, johtaa helposti ongelmallisiin tulkintoihin koko tekoälykäsitteen sisällöstä.

1.1 Tutkimussuunnat

Sillä, miten tekoäly rajataan ja mitä katsotaan kuuluvan sen piiriin, on ratkaiseva merkitys myös julkaisuanalyysin ja sovellettavan bibliometriikan tulosten kannalta.

Tekoälyn tutkimuskentän laajuus ja moninaisuus tarjoavatkin useita tapoja ryhmitellä julkaisuja erilaisiin alakokonaisuuksiin. Tässä selvityksessä on lähdetty siitä, että yhtä ainoaa ja yleisesti hyväksyttävissä olevaa tapaa rajata ja ryhmitellä tekoälyn tutkimusta ei ole järkevää tavoitella.

Tekoälyn julkaisukenttä on tässä katsauksessa ryhmitelty noudattamalla koko tekoälyselvityksessä käytettyä mallia. Sisältöperusteisen jaottelun lisäksi käytössä on myös yleinen geneerinen rajaus:

1. Tekoäly yleisesti tulkittuna
2. Data-analyysi

³ How Much to Trust Artificial Intelligence? George Hurlburt. Computing Edge, February 2018, Volume 4, Number 2, p. 26-27.

3. Havainnointi ja tilannetietoisuus
4. Luonnollinen kieli ja kognitio
5. Vuorovaikutus ihmisen kanssa
6. Digitaaliset työelämässä, ongelmanratkaisu ja laskennallinen luovuus
7. Koneoppiminen
8. Järjestelmätaso ja systeemivaikutukset
9. Tekoälyn laskentaympäristöt, alustat ja palvelut, ekosysteemit
10. Robotiikka ja koneautomaatio - tekoälyn fyysinen ulottuvuus
11. Etiikka, moraali, regulaatio ja lainsäädäntö

Selvityksessä hyödynnettiin Elsevier Scopus- ja SciVal-tietokantoja tekoälyjulkaisujen identifioinnissa, ryhmittelyssä sekä tieteellisen tason ja vaikuttavuuden arvioinnissa. Sisältöperusteisen ryhmittelyn apuna käytettiin Elsevierin Scopus –tietokannasta paikannettuja julkaisujen avainsanoja. Koska aiheen rajaamisessa käytettiin myös yleisempiä termejä (esimerkkeinä etiikka, moraali, lainsäädäntö, ennakointi, havaitseminen jne.), oli kyseiset termit liitettävä tekoälyn tutkimukseen sopivilla tekoälyn ”ankkuritermeillä”. Kyseiset tekoälyn nykyistä tutkimuskenttää mahdollisimman laajasti kuvaavat termit olivat tässä selvityksessä: *artificial intelligence, expert systems, cognitive systems, learning systems, machine learning ja neural networks*.

1.2 Näkökulmat ja menetelmät

Jokaisesta kymmenestä tutkimussuunnasta ja yleisesti rajatusta tekoäly-teemasta käytiin läpi Scopus-tietokannasta saatujen julkaisumäärien kehitys v. 2000-2017 (v. 2017 tilanne: 04/2018) sekä keskeiset kyseisen tutkimussuunnan aihepiirit ja teemojen tutkimusintensiteetin muutos (julkaisumäärien keskimääräinen vuosimuutos) v. 2008-2017.

Lisäksi seuraavista tutkimussuunnista käytiin läpi keskeiset maat, maiden tutkimusintensiteetin muutos v. 2015-2018 ja painotukset yleisimmissä tutkimusaiheissa (maiden vertailu). Johtavien maiden lisäksi tarkasteluun otettiin Pohjoismaat ml. Suomi:

- § Data-analyysi
- § Havainnointi ja tilannetietoisuus
- § Luonnollinen kieli ja kognitio
- § Koneoppiminen
- § Robotiikka ja koneautomaatio - tekoälyn fyysinen ulottuvuus
- § Etiikka, moraali, regulaatio ja lainsäädäntö

Lisäksi selvitettiin myös keskeiset tutkimusorganisaatiot ja organisaatioiden tutkimusintensiteetin muutos v. 2015-2018 ja painotukset yleisimmissä tutkimusaiheissa seuraavissa tutkimussuunnissa:

- § Koneoppiminen
- § Tekoäly yleisesti

Eri maiden ja tutkimuskentän toimijoiden tutkimusprofiilien eroja vertailtiin tässä selvityksessä vain yleisellä tasolla. Tarkastelussa olivat julkaisujen avaintermit sellaisenaan ilman termien sisältöihin liittyviä tulkintoja, termien yhdistämistä tai muualla tavoin tapahtuvaa muokkausta.

Julkaisujen tieteellisen vaikuttavuuden arviointi

Tieteellisen tutkimuksen vaikuttavuutta voidaan arvioida useista näkökulmista. Tässä selvityksessä on käytetty lähdeaineistona tutkimuksen tuottamia julkaisuja. Perinteisten bibliometristen mittarien – julkaisumäärien, viittausmäärien ja lehtien ns. impact factorien - rinnalle on viime vuosina noussut lukuisia niitä täydentäviä ja täsmentäviä tieteellisen vaikuttavuuden arvioinnin välineitä.

Mittarit voidaan jakaa karkeasti kolmeen ryhmään: a) lähdeviittausten lukumääriin, b) käytettyyn julkaisukanavaan ja nykyisin myös c) julkaisun käyttöön/näkyvyyteen/jakamiseen perustuviin mittareihin.

Näiden lisäksi tieteellistä vaikuttavuutta tarkastellaan usein tutkimusyhteistyön näkökulmasta: yhteistyön on tutkimuksissa todettu mm. lisäävän julkaisujen viittaasmääriä.⁴

Mikään yksittäinen mittari ei silti itsessään tarjoa kattavaa näkemystä tutkimuksen tieteellisestä vaikuttavuudesta, vaan eri mittareita on aina syytä tarkastella rinnakkain. Erityisen tärkeää on huomioida myös julkaisuutoiminnan volyymi silloin, kun tarkastellaan julkaisujen vaikuttavuutta. Pienet julkaisumäärät tuottavat suurta vaihtelua tieteellisen vaikuttavuuden arvioinnin mittareihin.

Tutkimustulosten vaikuttavuuden arvioinnin erityisenä haasteena voi pitää viivettä, joka syntyy julkaisun ilmestymisen ja julkaisun keräämien lähdeviittausten välille. Koska tämän selvityksen eräs keskeinen lähtökohta oli tarjota mahdollisimman ajantasainen kokonaiskuva - tilannekuva - tekoälystä, niin painopiste julkaisuaineistoissa oli myös vaikuttavuusselvitysten osalta uudemmassa aineistossa. Tällöin myös lähdeviittausten keräämisen aikaikkunaa jouduttiin joissakin kohdin supistamaan.

Tutkimuksen tieteellistä vaikuttavuutta tarkasteltiin selvityksessä viidellä mittarilla:

1. Suhteutettu viittausten lukumäärä (tutkimusala, julkaisuvuosi ja julkaisutyyppi huomioitu)
2. Julkaisujen määrä
3. Käytettyjen julkaisukanavien merkittävyys (top 10% -lehtien/julkaisukanavien osuus)
4. Kansainvälinen julkaisuyhteistyö
5. Yritys-yliopisto/tutkimuslaitos -yhteistyö
6. Normalisoitu julkaisutietojen katselumäärä (lähteenä Scopus-tietokanta)

Vaikka selvityksessä on painotettu nimenomaan suhteutettua viittausten lukumäärää, niin kaikkia biblio- ja muun metriikan mittareita on syytä tarkastella tässä yhteydessä rinnakkain. Näin erityisesti, koska selvityksen tarpeista johtuen painopiste oli uusimmassa julkaisuaineistossa.

Tutkimussuuntien analysointi

Tutkimussuuntia tarkasteltiin yllä mainituista näkökulmista seuraavasti:

Tutkimussuunta	Julkaisukehitys 2000-2017	Aiheet: volyymit ja muutokset 2008-2017	Maat: 1) volyymit ja muutokset 2003-2017 2) profiilivertailu 2015-2018	Tutkimusorganisaatiot: 1) volyymit ja muutokset 2003-2017 2) profiilivertailu 2015- 2018	Julkaisujen tieteellisen vaikuttavuuden arviointi: 1) maat ja 2) tutkimus- organisaatiot
Tekoäly yleisesti tulkittuna	x	x	x	x	x
Data-analyysi	x	x	x		x ⁵
Havainnointi ja tilannetietoisuus	x	x	x		x
Luonnollinen kieli ja kognitio	x	x	x		x
Vuorovaikutus ihmisen kanssa	x	x	x		
Digitaidot työelämässä, ongelmanratkaisu ja laskennallinen luovuus	x	x			
Koneoppiminen	x	x	x	x	x
Järjestelmätaso ja systeemi-vaikutukset	x	x			

⁴ Esim. Is collaboration among scientists related to the citation impact of papers because their quality increases with collaboration? Lutz Bornmann. Journal of Association for Information Science and Technology. Vol. 68, Issue 4, April 2017, 1036-1047. <https://doi.org/10.1002/asi.23728>. Do more distant collaborations have more citation impact. Nomalera, Ö, Frenkena, K., Heimeriks G. Journal of Informetrics, 7 (2013) 966– 971

⁵ Data-analyysi –teeman osalta tieteellisen vaikuttavuuden arviointi jaettiin kahtia: 1) big data, 2) data mining

Tekoälyn laskentaympäristöt, alustat ja palvelut, ekosysteemit	x	x			
Robotiikka ja koneautomaatio - tekoälyn fyysinen ulottuvuus	x	x	x		x
Etiikka, moraali, regulaatio ja lainsäädäntö	x	x	x		x

1.3 Julkaisuanalyysit

1.3.1 Tekoäly yleisesti

Julkaisumäärät ja aiheet

Tekoälyllä on tutkimusalan pitkä perinteet. Ensimmäiset tekoälyä ja sen mahdollisuuksia käsittelevät julkaisut ilmestyivät jo noin 60 vuotta sitten. Julkaisumäärien kasvu ei kuitenkaan ole ollut jatkuvaa, vaan väliin on mahtunut tasaisemman kehityksen ja jopa taantumisen kausia. Kuuluisin taantuma lienee ollut 1990-luvun ”tekoälyn talvi”⁶. Kokonaisuutena tekoälyn julkaisumäärät ovat kuitenkin kasvaneet 1960-2017 keskimäärin 19,2% vuodessa. Mutta kasvunopeus on siis vaihdellut tänä aikana paljon. Nopean kasvun vaiheet ovat osuneet erityisesti tämän vuosituhannen puolelle.

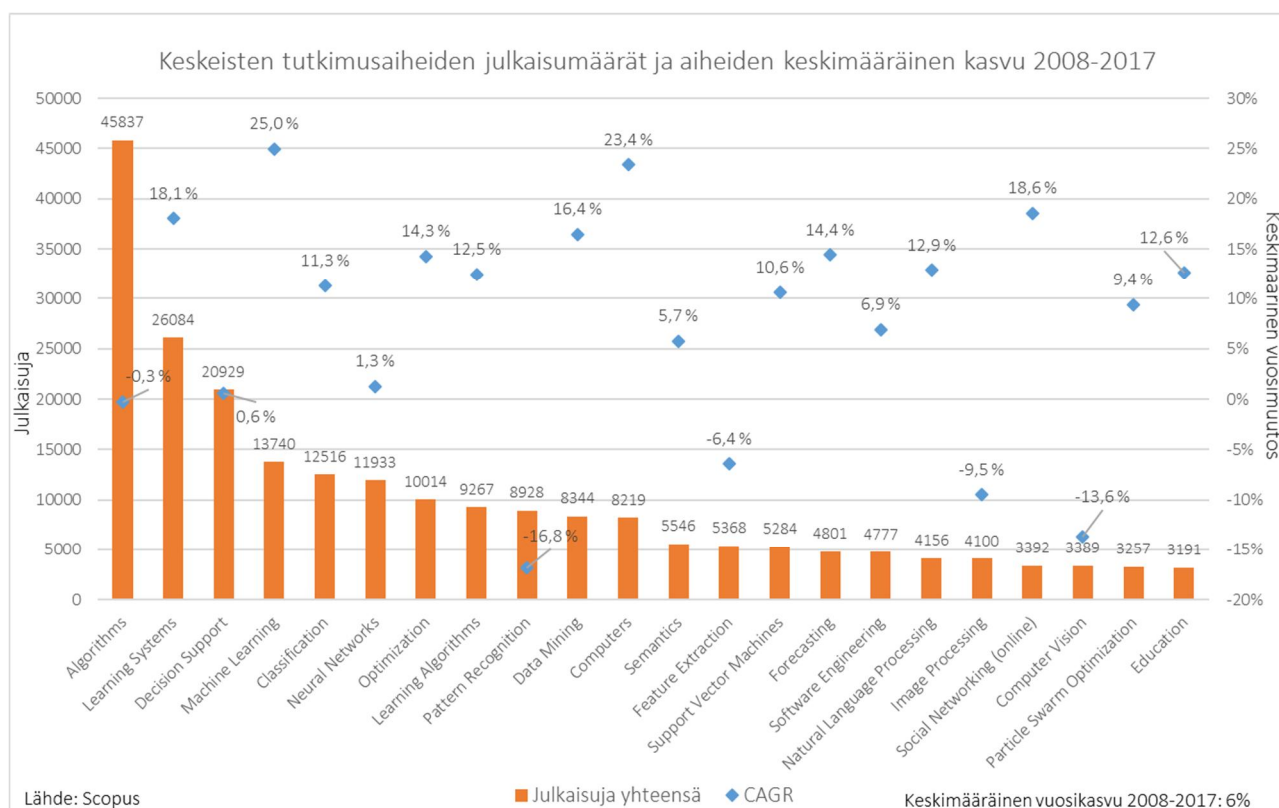


Kuva 2. Tekoäly⁷ käsittelevien julkaisujen määrä 1958-2017. Lähde: Scopus.

Viime vuosien julkaisumääriä tarkasteltaessa nousevat aiheissa esiin erityisesti koneoppimiseen liittyvät teknologiat ja tutkimussuunnat ml. erilaiset neuroverkkoteknologioita ja –algoritmeja hyödyntävät ratkaisut. Vaikka esimerkiksi ns. syväoppiminen ja syvät neuroverkot ovat aivan muutamana viime vuonna nousseet erääksi nopeimmin kasvavaksi yksittäiseksi tutkimusalaksi, on ns. oppivien järjestelmien nousu tekoälytutkimuksen sisällä jatkunut jo pitkään. Toisaalta myös perinteiset aiheet ovat edelleen vahvasti esillä, kun tarkastellaan aiheiden julkaisumääriä useamman vuoden aikaperspektiivistä. Alla on poimintoja isoimmista teemoista ja nopeimmin kasvavista aiheista tekoälyjulkaisuissa.

⁶ AI: The Tumultuous Search for Artificial Intelligence, Crevier, Daniel (1993)

⁷ Tulostajien rajoituksissa käytettiin seuraavia termejä: artificial intelligence, learning systems, learning algorithms, deep learning, image recognition, pattern recognition, machine learning, natural language processing, speech recognition, supervised learning, unsupervised learning, reinforcement learning, transfer learning, neural networks, feature extraction, decision support systems. Vuoden 2017 tilanne: 13.4.2018.



Kuva 3. Tekoälyn keskeisiä tutkimusaiheita: julkaisumäärät ja vuosikasvu 2008-2017. Kaikkien alan julkaisujen vuosikasvu ko. ajanjaksolla oli 6%. Lähde: Scopus.⁸

Maat

Kiina ja Yhdysvallat ovat tekoälyn suurvallat.⁹ Kiinan kasvavaa roolia kuvaa yleisemminkin se, että maa ohitti äskettäin Yhdysvallat maailman suurimpana tieteellisten julkaisujen tuottajana. Yhdessä kyseiset maat kattoivat v. 2016 n. 36% kaikista tieteellis-teknillisistä julkaisuista.¹⁰ Tekoälytutkimuksen piirissä Kiinan ja Yhdysvaltain osuus on vielä tätäkin merkittävämpi. Toisaalta osuus vaihtelee riippuen käytetystä tekoälykäsitteen rajauksesta. Esimerkiksi koneoppimisessa Kiinan ja Yhdysvaltojen osuus oli v. 2017 44% kaikista julkaisuista. Kommunikaatiossa, havainnoinnissa ja toiminnassa osuus oli samana vuonna n. 39%.

Tekoälytutkimuksen piirissä tilanne pysyi tarkastellulla ajanjaksolla suurimpien toimijoiden suhteen melko vakaana: kaikkien julkaisumäärät lisääntyivät¹¹. Suurinta vaihtelu oli Kiinan osalta. Vuosina 2014-2015 tapahtui selvä hyppäys julkaisumäärissä, mikä näyttäisi kuitenkin sittemmin tasaantuneen.

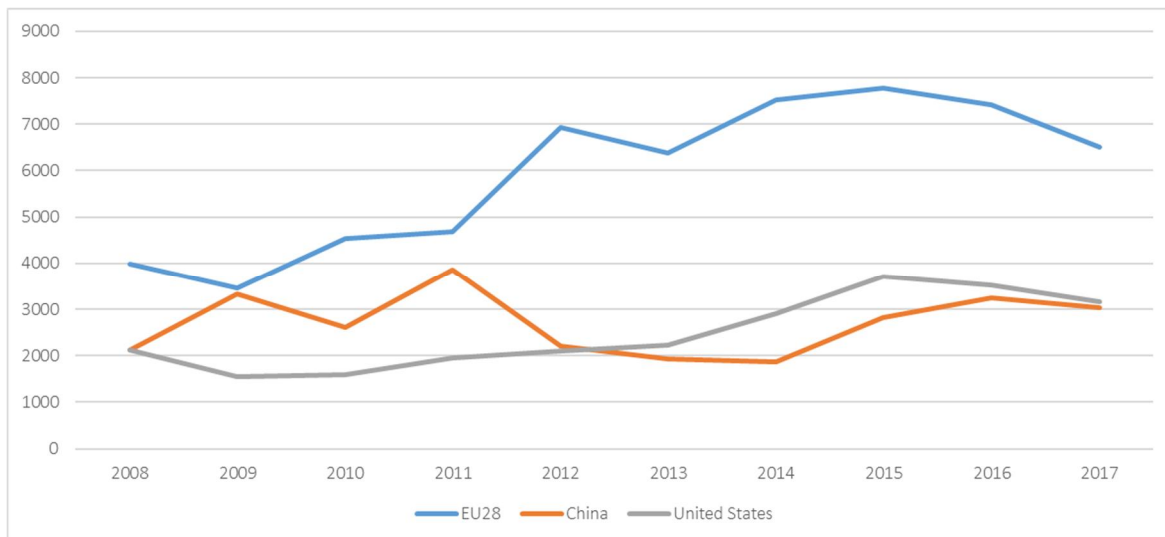
⁸ Haku: TITLE-ABS-KEY ("Artificial intelligence") AND (LIMIT-TO (SUBJAREA, "COMP")), status: 5.3.2018

⁹ Esim. Emmanuel Macron Talks To Wired About France's Ai Strategy, Wired 31.3.2018. "When you look at artificial intelligence today, the two leaders are the US and China."

¹⁰ US National Science Foundation (NSF), Science & Engineering Indicators 2018.

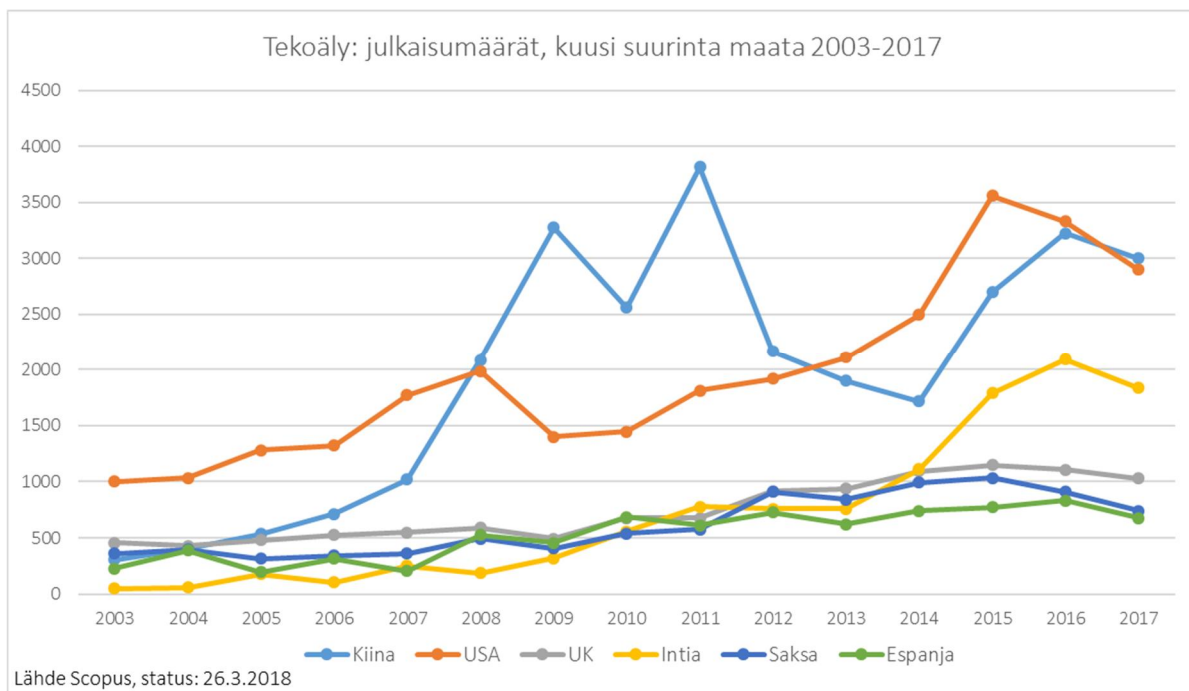
<https://www.nsf.gov/statistics/2018/nsb20181/figures>

¹¹ Rajaus Scopus-tietokannasta: KEY ("Artificial intelligence") OR TITLE ("Artificial intelligence") AND (LIMIT-TO (SUBJAREA, "COMP")). Yhteensä mukana julkaisuja v. 2015-2018 (Vuosi 2018: tilanne 27.3.2018): 54950.



Kuva 4. Julkaisumäärin kehitys: EU28, USA ja Kiina 2008-2017. EU28: yhteisjulkaisut laskettu kaikkien osallistujamaiden lukumääriin¹². Lähde: Scopus

Taustalla näkyy ennen kaikkea datapohjaisten ja oppivien menetelmien lisääntynyt tutkimus.¹³ Jo tällä hetkellä Kiinassa on eräs suurimmista tekoälytutkijoiden keskittymistä ja on ennustettu, että Kiina ohittaa Yhdysvallat tekoälyn kehittämisessä vuonna 2025.¹⁴ Tämä näkyy myös julkaisutoiminnan kehityksessä. Selvästi nopeinta kasvu on kuitenkin ollut Intiassa sekä tarkasteltaessa koko ajanjaksoa 2003-2017 että vuosina 2008-2017. Toisaalta Intian lähtötaso on ollut huomattavasti Yhdysvaltoja ja Kiinaa matalampi.

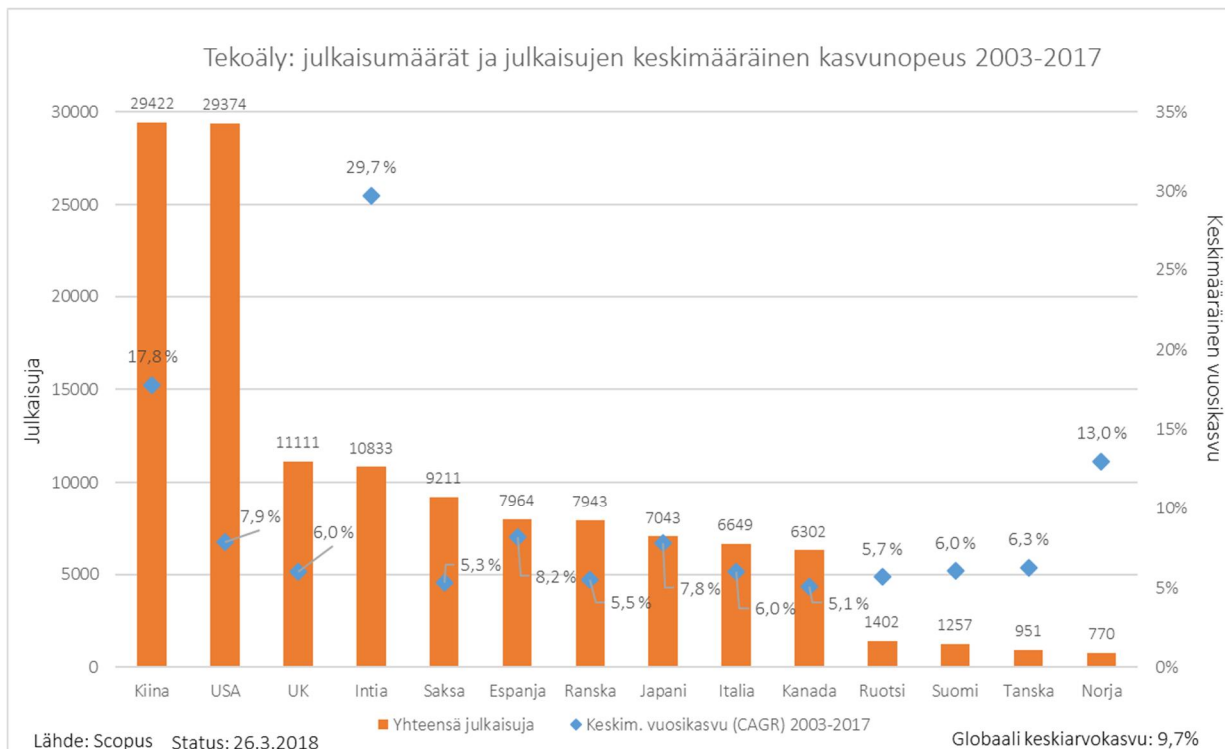


Kuva 5. Suurimmat maat julkaisumäärillä mitattuna 2003-2017. Lähde: Scopus

¹² Rajaus Scopus-tietokannassa: TITLE-ABS-KEY ("Artificial intelligence") AND (LIMIT-TO (SUBJAREA, "COMP")).

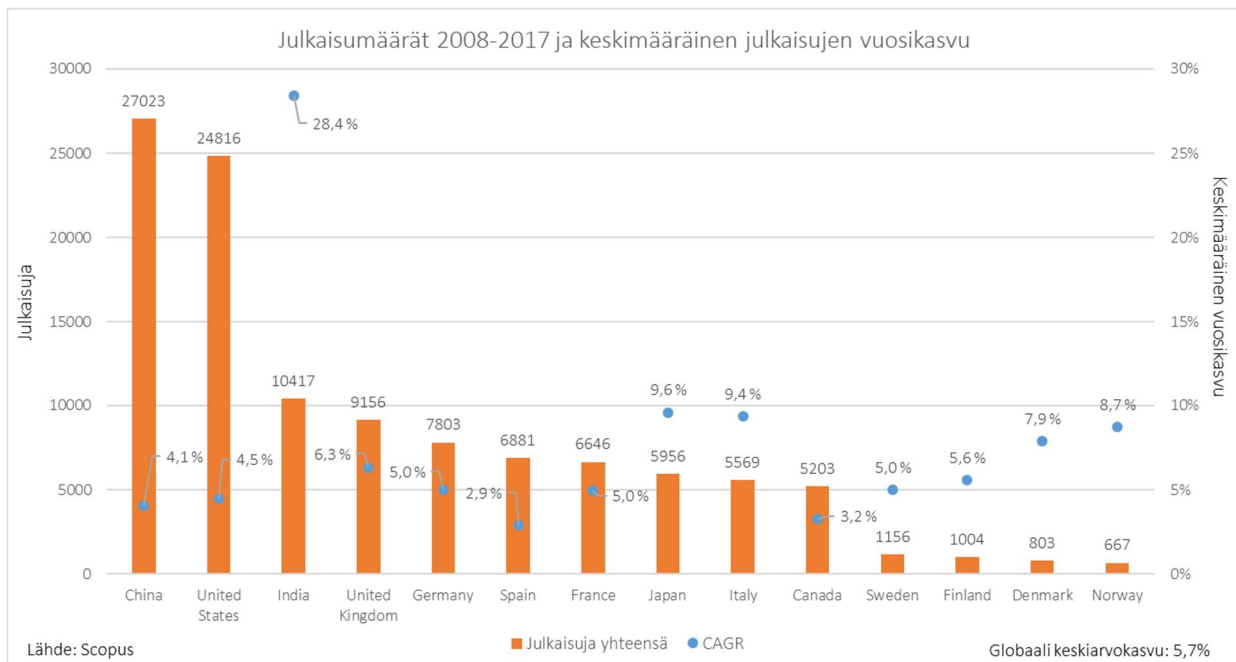
¹³ Ks. Kappale 2.3.7 Koneoppiminen

¹⁴ Esim. The Economist, Marche 17th 2018, s. 11



Kuva 6. Tekoäly: suurimmat maat Pohjoismaat 2003-2017, julkaisujen määrä ja julkaisutoiminnan kasvunopeus. Lähde: Scopus

Mikäli tarkastellaan erikseen vain vuosia 2008-2017 ja otetaan mukaan julkaisuja väljemmällä hakurajauksella, niin Kiinan ja USA:n kasvunopeudet ovat ko. ajanjaksolla pienemmät. Muista maista Espanjan, Japanin ja Italian julkaisujen kasvu oli puolestaan nopeampaa.



Kuva 7. Suurimmat maat sekä Pohjoismaat: julkaisumäärät ja vuosikasvu 2008-2017. Haku: title-abs-key ("artificial intelligence") and (limit-to (subjarea, "comp")). Lähde: Scopus

Paitsi julkaisumäärissä ja tutkimuksen kasvunopeudessa on maiden välillä eroja myös tutkimuksen painotuksissa. Ohessa aiheiden prosentiosuudet julkaisuista eri maissa vuosina 2015-2018.

Julkaisuja yhteensä:	10004	9213	5871	3403	2762	2388	2361	2301	2201	1847	1743	1250	468	372	295	248
Avainsana	USA	Kiina	Intia	UK	Saksa	Espanja	Ranska	Japani	Italia	Kanada	Australia	Etelä-Ko	Ruotsi	Suomi	Tanska	Norja
Algorithms	9.1 %	14.9 %	12.1 %	7.9 %	6.0 %	7.7 %	7.6 %	5.3 %	5.0 %	10.4 %	9.9 %	6.6 %	8.3 %	8.1 %	7.1 %	6.9 %
Big Data	5.0 %	4.7 %	4.7 %	3.7 %	3.1 %	3.2 %	2.5 %	3.3 %	4.5 %	3.8 %	4.8 %	4.3 %	3.6 %	4.6 %	2.0 %	5.2 %
Classification (of Information)	7.7 %	7.6 %	11.4 %	7.9 %	6.4 %	7.7 %	6.0 %	4.8 %	7.1 %	9.3 %	9.4 %	6.2 %	5.6 %	8.6 %	6.1 %	7.3 %
Computation Theory	3.3 %	4.5 %	6.3 %	3.7 %	2.3 %	2.8 %	3.4 %	2.5 %	4.3 %	2.9 %	3.0 %	1.3 %	2.6 %	1.3 %	2.7 %	4.4 %
Computer Science	4.1 %	4.3 %	1.4 %	7.7 %	13.1 %	4.5 %	12.1 %	10.0 %	7.7 %	5.6 %	4.6 %	2.5 %	10.3 %	11.0 %	16.6 %	11.7 %
Computers	5.9 %	5.5 %	1.9 %	10.7 %	18.6 %	6.0 %	17.6 %	12.9 %	10.7 %	8.0 %	6.4 %	2.9 %	14.5 %	15.1 %	26.1 %	15.7 %
Data Mining	8.5 %	7.0 %	17.1 %	8.1 %	7.7 %	8.0 %	6.8 %	6.6 %	7.5 %	8.8 %	10.7 %	6.9 %	7.3 %	8.6 %	1.7 %	4.8 %
Decision Making	3.9 %	2.7 %	3.5 %	4.9 %	3.9 %	5.8 %	5.3 %	2.3 %	5.9 %	4.3 %	5.9 %	3.0 %	4.9 %	4.8 %	2.0 %	5.6 %
Decision Support Systems	3.7 %	2.1 %	3.5 %	5.3 %	5.4 %	6.8 %	5.9 %	1.4 %	8.3 %	4.4 %	5.3 %	2.9 %	10.5 %	4.8 %	4.7 %	15.7 %
Decision Trees	3.6 %	2.4 %	4.7 %	4.1 %	2.6 %	3.8 %	2.9 %	3.1 %	3.2 %	4.5 %	3.8 %	3.5 %	5.1 %	2.7 %	1.0 %	3.6 %
Education	5.4 %	3.0 %	2.3 %	2.9 %	2.1 %	4.4 %	1.6 %	4.7 %	1.7 %	2.9 %	3.4 %	2.6 %	3.6 %	3.2 %	1.7 %	3.2 %
Feature Extraction	3.2 %	5.1 %	6.9 %	4.2 %	2.6 %	4.5 %	1.9 %	2.9 %	3.0 %	3.1 %	4.0 %	4.4 %	2.1 %	2.2 %	3.4 %	4.4 %
Forecasting	6.2 %	5.1 %	5.8 %	4.7 %	4.0 %	5.5 %	2.8 %	4.1 %	3.6 %	6.0 %	5.0 %	4.3 %	7.1 %	4.8 %	5.8 %	4.4 %
Learning Algorithms	14.6 %	10.4 %	12.6 %	10.0 %	9.5 %	8.1 %	8.0 %	8.2 %	8.4 %	13.0 %	10.4 %	10.2 %	9.2 %	7.5 %	8.5 %	8.1 %
Machine Learning	15.4 %	8.5 %	10.6 %	12.3 %	11.9 %	13.3 %	8.6 %	11.7 %	9.8 %	14.4 %	9.1 %	13.6 %	15.8 %	15.6 %	8.8 %	14.9 %
Neural Networks	5.0 %	5.8 %	6.8 %	7.2 %	6.8 %	11.6 %	5.3 %	6.3 %	6.0 %	6.2 %	5.6 %	6.5 %	3.6 %	7.8 %	3.7 %	3.6 %
Optimization	7.5 %	14.4 %	12.3 %	8.3 %	5.8 %	7.6 %	6.9 %	6.3 %	6.5 %	7.9 %	11.6 %	5.4 %	7.3 %	6.2 %	6.1 %	12.1 %
Pattern Recognition	2.5 %	3.0 %	4.8 %	3.6 %	2.6 %	3.3 %	2.5 %	2.5 %	3.0 %	3.2 %	3.2 %	2.7 %	2.6 %	6.2 %	2.0 %	
Semantics	4.0 %	5.2 %	2.6 %	4.4 %	5.4 %	3.9 %	5.8 %	3.8 %	7.3 %	3.5 %	4.5 %	2.3 %	5.8 %	5.1 %	4.4 %	2.8 %
Social Networking (online)	3.3 %	2.7 %	4.1 %	2.6 %	2.5 %	3.0 %	1.9 %	2.1 %	3.7 %	3.2 %	3.7 %	3.1 %	3.2 %	2.4 %	2.0 %	2.4 %
Support Vector Machines	3.4 %	4.6 %	6.4 %	3.0 %	2.4 %	2.7 %	1.8 %	3.7 %	2.9 %	4.0 %	3.0 %	3.4 %	1.7 %		2.0 %	2.8 %

Kuva 8. Tekoäly: suurimmat kymmenen maata ja pohjoismaat 2015-2018, julkaisujen painotukset. Yleisimmät sisältöavainsanat koko julkaisuaineistossa (yleisimmät kuivailutermit "artificial intelligence" ja "learning systems" eivät mukana vertailussa). Termien prosentuaalinen osuus on $\neq$ 100% johtuen eri aiheiden esiintymisestä samoissa julkaisuissa. Yhteensä analysoituja julkaisuja oli 54950 (tilanne: 27.3.2018). Lähde: Scopus

Tutkimuksen vaikuttavuus: maat

Yhdysvallat on perinteisesti ollut joko kärjessä tai hyvin lähellä sitä tarkasteltaessa viittausmääriä millä tahansa tutkimusalalla. Toisaalta viittauksia keräävien kiinalaisten tekoälyjulkaisujen määrä on nykyisin jo 89% amerikkalaisten julkaisujen määrästä¹⁵. Kiina on siten lähestymässä Yhdysvaltoja myös julkaisujen laadussa, mikäli julkaisun laadulla ymmärretään julkaisun keräämiä viittauksia muista julkaisuista. Julkaisun keskimäärin keräämien viitteiden määrässä Yhdysvallat on kuitenkin edelleen selvästi edellä Kiinaa.

Alla on listattuna tekoälytutkimuksen¹⁶ merkittävimmät maat normalisoidun (suhteutetun / skaalatun) viittausindeksin (FWCI) perusteella.¹⁷ Raja-arvona tässä on FWCI:n arvo 1,10 (eli jos FWCI-arvo on 1,10, niin maan tekoälyä käsitteleviin julkaisuihin on viitattu keskimäärin 10% enemmän kuin koko maailman keskiarvon mukaan olisi odotettavissa). Alle tuhat julkaisua v. 2013-2017 tuottaneet maat jätetty vertailusta pois.

Normalisoidulla viittausmäärällä mitaten tekoälytutkimuksen kolme kärkimaata ovat Singapore, Sveitsi ja Saudi-Arabia. Yksittäisiä huippujulkaisuja oli puolestaan suhteellisesti eniten Hong Kongilla, Singaporella ja Israelilla. Seuraavina tulivat Australia, Espanja ja Kanada. Kansainvälistä yhteistyötä oli eniten Hong Kongilla ja Saudi-Arabialla, molemmilla kaksinkertaisesti kansainväliseen keskiarvoon verrattuna. Yritysten ja yliopistojen välinen tutkimusyhteistyö oli kaikkein tiiveintä Irlannissa (ei kuvassa), Israelissa ja Yhdysvalloissa. Seuraavina tulivat Ruotsi, Sveitsi ja Saksa.

¹⁵ Chinese tech v American tech. The Economist, vol. 426 nr 9079, February 18th 2018, s. 61

¹⁶ Tekoälyn rajauksena käytettiin tässä yhteydessä Elsevier SciValin research area –määrittelyä: "artificial intelligence"

¹⁷ FWCI: maailman keskiarvo = 1. Esim. lukuarvo 1,45 tarkoittaa, että organisaation tekoälyjulkaisuihin on viitattu tarkastelujaksolla (2012-2017) 45% maailman keskiarvoa enemmän. Lähteenä on käytetty Scopus- ja SciVal-aineistoja.

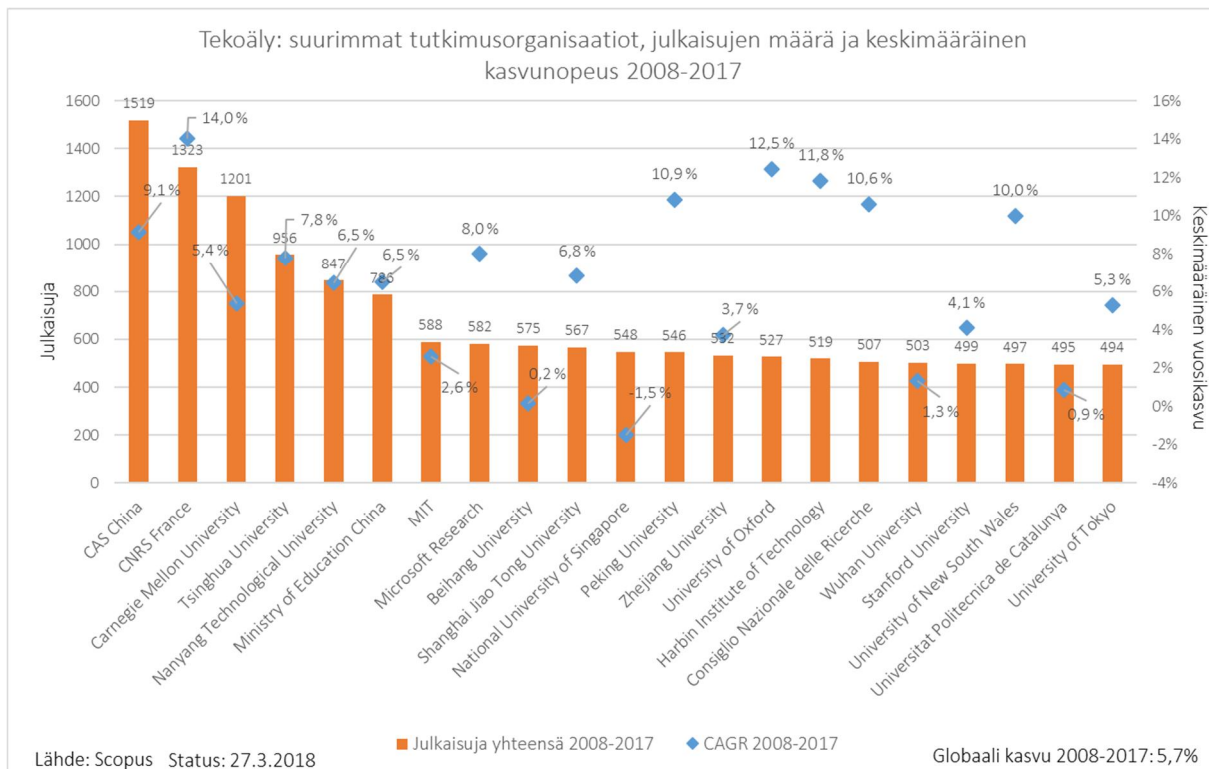
Maa	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys -yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisutietojen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Singapore	2,26	3141	34,8	1,64	3,7	1,35
Switzerland	2,07	1993	22,9	1,64	6,3	1,36
Saudi Arabia	2,01	1603	26,7	2,01	1,2	1,67
Hong Kong	1,94	2370	41,2	1,99	3,2	1,45
Australia	1,73	6132	30,6	1,51	5,2	1,43
United States	1,69	30792	26,4	1,02	6,9	1,1
United Kingdom	1,63	11877	25,3	1,53	3,8	1,3
Austria	1,5	1522	23	1,47	4,1	1,21
Israel	1,5	1506	31,4	1,28	9,9	0,84
Italy	1,49	7484	18,8	1,13	2,6	1,77
Canada	1,49	5933	27,4	1,42	2,6	1,2
Pakistan	1,48	1275	11,4	1,26	0,6	1,71
Netherlands	1,46	2648	20,6	1,46	3,7	1,25
Finland	1,45	1291	19	1,41	3,6	1,45
Germany	1,44	9699	17,3	1,12	5,3	1,06
Belgium	1,4	1637	26,1	1,59	3,7	1,27
Spain	1,33	7335	28,3	1,16	2,3	1,46
Greece	1,25	1804	17	1,11	0,8	1,57
France	1,23	7672	19,4	1,39	4,6	1,04
Malaysia	1,19	2964	17,6	0,98	0,9	2,44
Egypt	1,16	1001	9,6	1,32	1,5	1,06
Poland	1,14	4042	17,6	0,77	0,7	1,05
Sweden	1,14	1739	13,2	1,5	6,7	1,26
China	1,11	49044	21,9	0,64	1,7	0,93

Kuva 9. Tekoälytutkimuksen maat järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2013-2017. Viittausindeksin raja-arvo kuvassa: 1,10. Alle 1000 julkaisua tuottaneet maat eivät ole mukana taulukossa. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen yhteistyön ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa. (Tiedot Scopusesta: katselukerrat antavat tietoa julkaisun herättämästä kiinnostuksesta täydentämällä viittausmäärien välittämää informaatiota erityisesti uusimman aineiston osalta.) Normalisoiduissa mittareissa maailman keskiarvo = 1. Lähde: Scopus ja SciVal. Julkaisuja n. 209113. Tilanne: 6.6.2018

Tutkimusorganisaatiot: yliopistot, tutkimuslaitokset ja yritykset

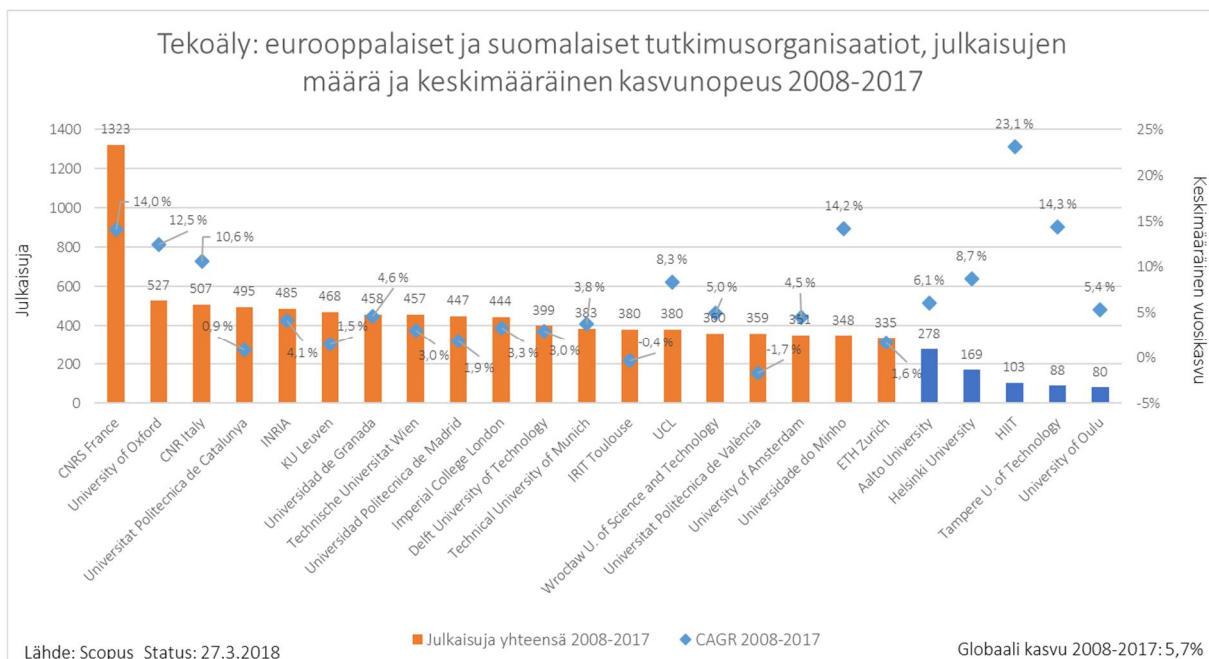
Kiinan ja Yhdysvaltojen lisäksi Ranskan ja Iso-Britannian yliopistot ja tutkimuslaitokset ovat keskeisiä toimijoita tekoälytutkimuksessa. Ranskan kansallinen tutkimuskeskus CNRS, Centre national de la recherche scientifique, on nopeimmin tekoälyn tutkimustoimintaansa julkaisujen määrällä mitattuna lisännyt tutkimuslaitos. Kuten useilla muillakin tieteen aloilla Kiinan tiedeakatemia CAS on myös tekoälytutkimuksen alueella johtava toimija.

Suurimmat tutkimusorganisaatiot olivat tarkasteltavana ajanjaksona Kiinan tiedeakatemia ja Ranskan kansallinen tutkimuskeskus CNRS.

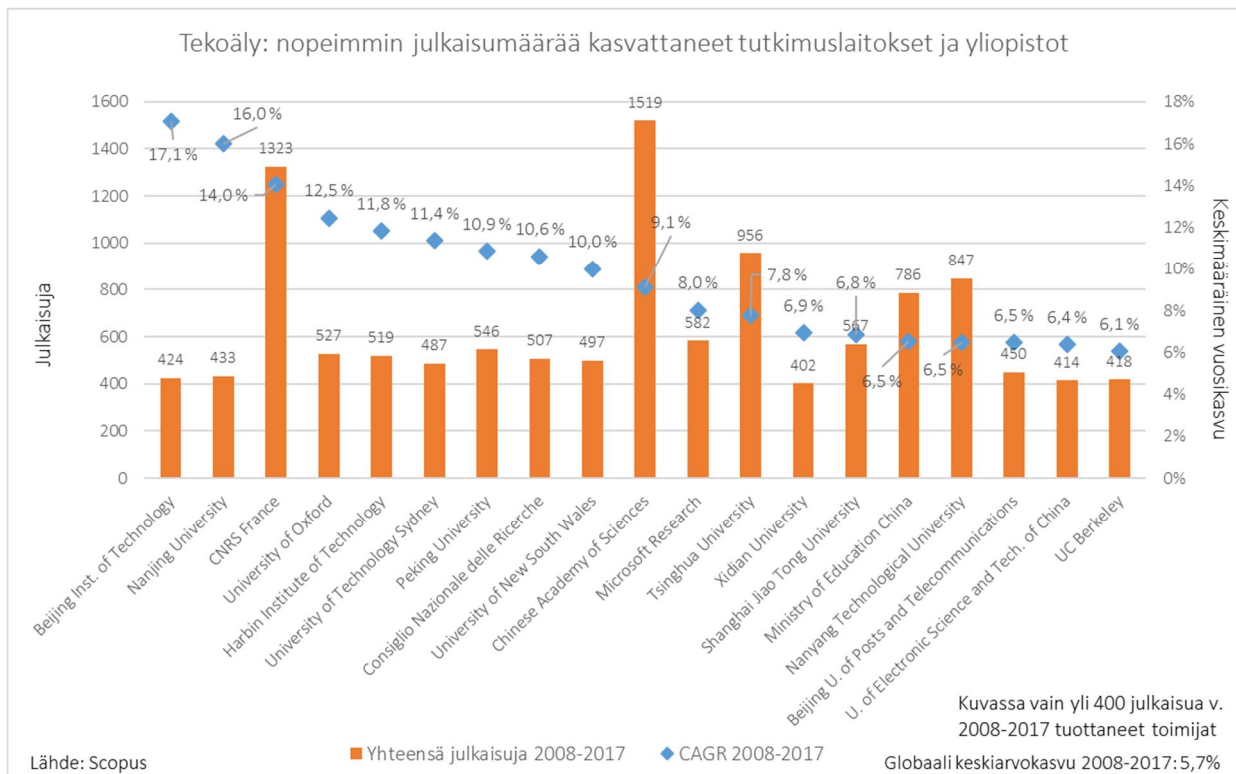


Kuva 10. Tekoäly yleisesti tulkittuna: suurimmat tutkimusorganisaatiot 2008-2017, julkaisujen määrä ja julkaisutoiminnan kasvunopeus. Rajaus Scopus-tietokannasta: key ("artificial intelligence") or title ("artificial intelligence") and (limit-to (subjarea, "comp")). Yhteensä mukana julkaisuja v. 2015-2018 (Vuosi 2018: tilanne 27.3.2018): 54950. Lähde: Scopus

Euroopan suurimmat yliopistot tekoälyn julkaisumäärillä mitaten ovat viimeisen kymmenen vuoden aikana olleet Oxfordin yliopisto, Katalonian teknillinen yliopisto ja Leuvenin kuninkaallinen yliopisto. Tutkimuskeskuksista merkittävimmät ovat olleet CNRS:n lisäksi Italian CNR ja Ranskan INRIA. Suomessa suurimmat yliopistot ovat v. 2008-2017 odotetusti olleet Aalto yliopisto ja Helsingin yliopisto.

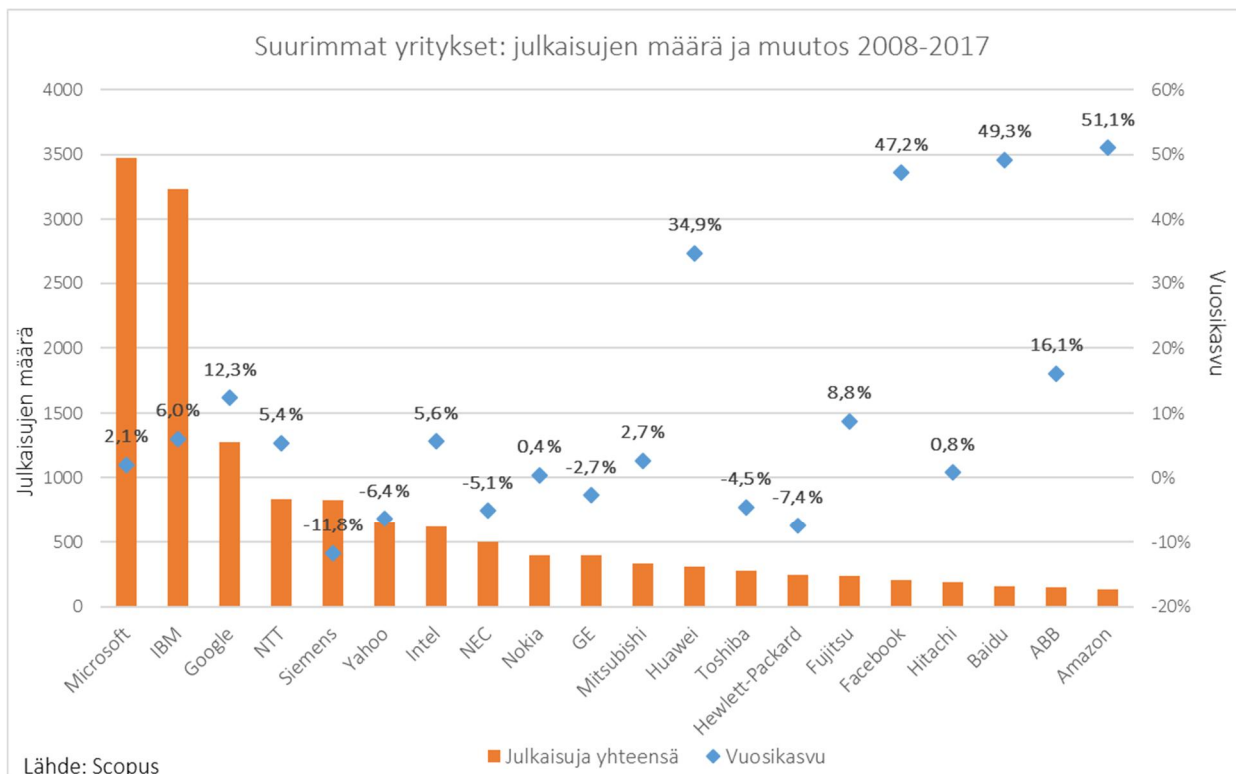


Kuva 11. Tekoäly yleisesti tulkittuna: suurimmat eurooppalaiset ja suomalaiset tutkimusorganisaatiot 2008-2017, julkaisujen määrä ja julkaisutoiminnan kasvunopeus. Lähde: Scopus



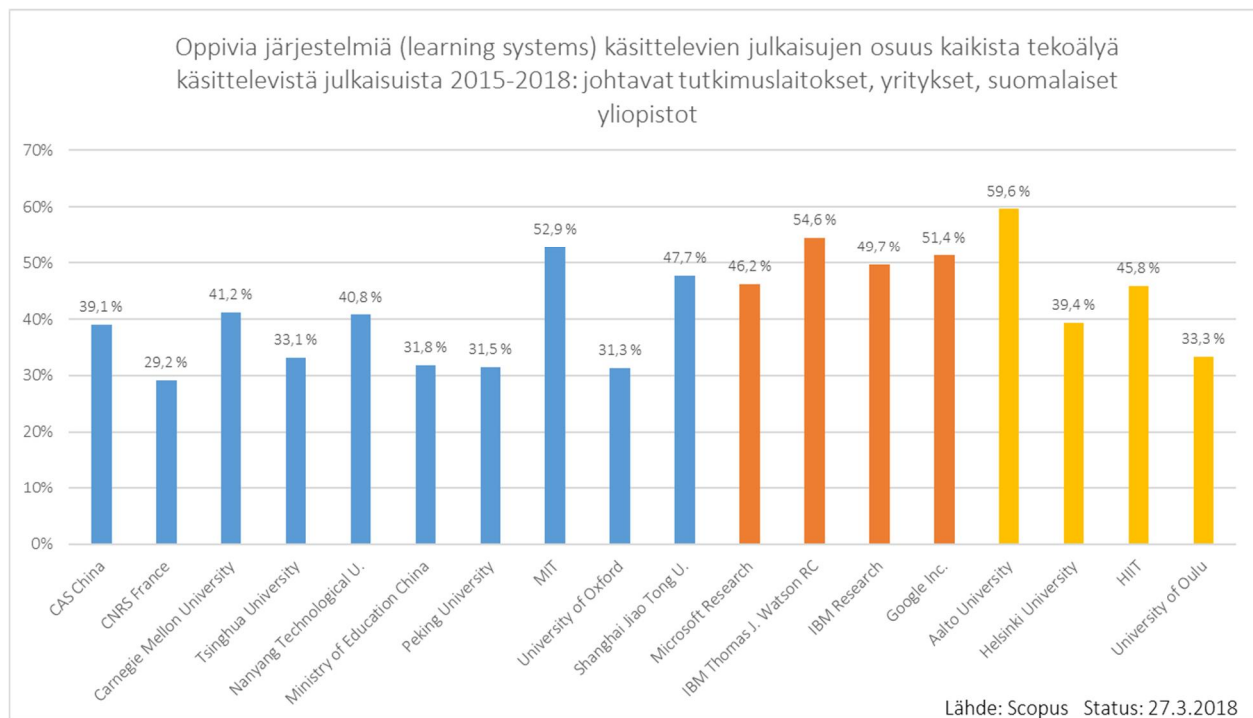
Kuva 12. Tekoäly yleisesti tulkittuna: nopeimmin julkaisujen määrää lisänneet tutkimuslaitokset ja yliopistot 2008-2017, julkaisujen määrä ja julkaisutoiminnan kasvunopeus. Mukana vain suuret tutkimusorganisaatiot. Lähde: Scopus

Selkeästi suurimmat tutkimusjulkaisuja tuottavat yritykset v. 2008-2017 olivat amerikkalaiset Microsoft ja IBM. Julkaisumäärissä mitattuna nopeimmin panostustaan julkiseen tutkimukseen ovat lisänneet kuitenkin kiinalaiset Tencent, Baidu ja Alibaba Group sekä yhdysvaltalaiset Amazon ja Facebook. Julkaisumäärät ovat näillä kaikilla kuitenkin vielä toistaiseksi suhteellisen pieniä.



Kuva 13. Tekoäly yleisesti tulkittuna¹⁸: suurimmat yritykset 2008-2017, julkaisujen määrä ja julkaisutoiminnan kasvunopeus. Lähde: Scopus

Koneoppimiseen, oppiviin algoritmeihin ja muihin oppiviin järjestelmiin liittyvän tutkimuksen osuus tekoälyn tutkimuksessa on nykyisin merkittävä yliopistoissa ja tutkimuslaitoksissa. Alla oleva tilasto kuvaa ennen kaikkea sitä, miten vahvasti oppivat järjestelmät nähdään osana tekoälyn kenttää. Suuri osuus kertoo lähinnä tekoälyn tutkimuksen perinteistä kyseisessä maassa: oppivat järjestelmät nähdään kyseisissä maissa todennäköisesti luontevammin osana tekoälyn tutkimuksen pitkää traditiota kuin siitä erillisenä tutkimussuuntana. Taulukossa on erotettu kymmenen johtavaa yliopistoa / tutkimuslaitosta, eniten julkaisevat yritykset ja eniten tekoälyjulkaisuja tuottaneet suomalaiset yliopistot v. 2015-2018.



Kuva 14. Oppivia järjestelmiä käsittelevien julkaisujen osuus kaikista tekoälyä käsittelevistä julkaisuista 2015-2018 (27.3.2018). Lähde: Scopus

Paitsi julkaisumäärien kasvunopeudessa on maiden välillä eroja myös tutkimuksen painotuksissa. Oppivat järjestelmät ja ohjelmistot ovat vahvasti edustettuna tutkimusaiheissa. Ohessa aiheiden prosenttiosuudet julkaisuista johtavilla organisaatiolla vuosina 2015-2018.

¹⁸ Tutkimussuunnan rajauksessa käytettiin seuraavia termejä: artificial intelligence, learning systems, learning algorithms, deep learning, image recognition, pattern recognition, machine learning, natural language processing, speech recognition, supervised learning, unsupervised learning, reinforcement learning, transfer learning, neural networks, feature extraction, decision support systems, cognitive systems, knowledge based systems, knowledge representation

Julkaisuja yhteensä 2015-2018	699	582	565	483	353	349	286	276	262	256	277	163	147	142	114	66	48	36
	CAS China	CNRS	Carnegie Mellon U.	Tsinghua University	Nanyang Technological U.	Ministry of Education China	Peking University	MIT	University of Oxford	Shanghai Jiao Tong U.	Microsoft Research	IBM Thomas J. Watson RC	IBM Research	Google Inc.	Aalto University	Helsinki University	HIIT	University of Oulu
Avainsana																		
Algorithms	9,6 %	6,7 %	6,4 %	9,5 %	10,8 %	20,9 %	9,8 %	9,8 %	7,3 %	9,8 %	7,2 %	9,8 %	11,6 %	9,2 %	9,6 %	6,1 %	8,3 %	11,1 %
Big Data	4,6 %	2,9 %	4,1 %	5,4 %	4,5 %	3,2 %	3,1 %	3,6 %	3,8 %	6,3 %	2,9 %	8,6 %	6,1 %	4,9 %	4,4 %	1,5 %		8,3 %
Classification (of Information)	8,3 %	5,8 %	4,1 %	7,0 %	10,2 %	9,5 %	6,3 %	5,4 %	5,3 %	9,0 %	5,8 %	6,1 %	4,8 %	1,4 %	10,5 %	1,5 %	8,3 %	13,9 %
Computation Theory	4,9 %	3,1 %	2,5 %	1,7 %	5,1 %	4,3 %		5,8 %	4,6 %	2,7 %	4,7 %			2,1 %				
Computer Science	5,3 %	16,7 %	3,0 %	6,4 %	3,1 %	5,2 %	11,5 %	5,4 %	7,6 %	3,5 %	6,5 %	3,7 %	2,0 %	2,1 %	6,1 %	13,6 %	10,4 %	5,6 %
Computer Vision	5,3 %	2,4 %	4,2 %	3,5 %	1,4 %	3,2 %	3,8 %	1,4 %	2,3 %	4,3 %	4,0 %			2,1 %			2,1 %	5,6 %
Computers	7,0 %	22,2 %	4,2 %	8,3 %	5,1 %	6,3 %	12,2 %	6,5 %	9,9 %	5,9 %	7,6 %	4,3 %	3,4 %	3,5 %	10,5 %	18,2 %	14,6 %	5,6 %
Data Mining	4,9 %	6,0 %	6,0 %	6,2 %	9,3 %	6,0 %	7,3 %	5,1 %	3,4 %	5,5 %	7,2 %	7,4 %	5,4 %	9,2 %	5,3 %	9,1 %	6,3 %	13,9 %
Decision Making	2,1 %	4,0 %	3,0 %	2,9 %	3,4 %	2,0 %	1,7 %	3,3 %	3,8 %	2,3 %	2,9 %	3,7 %	4,8 %	1,4 %	2,6 %			5,6 %
Feature Extraction	7,6 %	1,2 %	1,2 %	4,3 %	4,5 %	8,3 %	3,8 %	1,1 %	1,5 %	3,5 %		3,7 %	2,7 %	1,4 %	3,5 %			
Forecasting	6,9 %	1,4 %	4,6 %	5,4 %	4,0 %	2,6 %	4,9 %	8,0 %	1,9 %	8,6 %	7,2 %	9,2 %	13,6 %	2,8 %	7,9 %	3,0 %	4,2 %	8,3 %
Iterative Methods	4,0 %	2,4 %	4,2 %	2,9 %	2,8 %	2,9 %	4,5 %	5,4 %	1,1 %	6,6 %	2,9 %	3,1 %	2,7 %	3,5 %		3,0 %		
Learning Algorithms	13,6 %	5,7 %	13,1 %	9,7 %	17,3 %	11,7 %	9,8 %	13,8 %	6,9 %	11,3 %	13,4 %	18,4 %	11,6 %	8,5 %	8,8 %	9,1 %	10,4 %	8,3 %
Machine Learning	12,2 %	6,7 %	10,4 %	9,3 %	13,0 %	8,3 %	8,0 %	12,0 %	6,9 %	12,5 %	7,9 %	11,7 %	9,5 %	11,3 %	17,5 %	6,1 %	4,2 %	11,1 %
Matrix Algebra	2,9 %	2,9 %	2,8 %	2,5 %	2,3 %	2,3 %	3,1 %	4,7 %		4,3 %	4,0 %	3,7 %	4,1 %	2,1 %	4,4 %		4,2 %	
Natural Language Processing Sys.	4,0 %	1,4 %	2,1 %	2,9 %	5,4 %	2,3 %	7,0 %	2,2 %			4,7 %	4,3 %	3,4 %	4,9 %				
Neural Networks	11,6 %	2,9 %	1,2 %	5,6 %	7,1 %	4,9 %	4,9 %	4,0 %	2,7 %	5,5 %	4,3 %	3,7 %	4,1 %	4,2 %	7,0 %	4,5 %	4,2 %	5,6 %
Optimization	8,4 %	7,4 %	7,3 %	7,7 %	11,0 %	17,5 %	14,3 %	9,4 %	3,8 %	9,8 %	10,8 %	9,2 %	6,1 %	8,5 %	7,0 %	9,1 %	8,3 %	
Semantics	6,0 %	8,1 %	5,1 %	9,7 %	7,4 %	4,6 %	10,5 %	2,9 %	8,0 %	5,9 %	5,1 %	7,4 %	5,4 %	5,6 %		10,6 %	10,4 %	
Social Networking (online)	4,7 %	1,2 %	2,1 %	6,0 %	3,1 %	3,4 %	3,8 %	2,2 %	1,9 %	2,0 %	3,6 %	2,5 %			4,4 %		8,3 %	

Kuva 15. Tekoäly yleisesti: suurimmat kansainväliset tutkimuslaitokset ja yritykset sekä ja suomalaiset yliopistot 2015-2018. Yleisimmät avainsanat koko julkaisuaineistossa: osuus toimijoiden julkaisuista. (Yleisimmät termit: "artificial intelligence" ja "learning systems" eivät mukana vertailussa.) Osa julkaisutermeistä jätetty pois termin päällekkäisyyden tai yleisluonteisuuden perusteella. Termien prosentuaalinen osuus on $\neq$ 100% johtuen eri aiheiden esiintymisestä samoissa julkaisuissa. Lähde: Scopus. Vuoden 2018 status: 27.3.2018.

Tutkimuksen vaikuttavuus: tutkimusorganisaatiot

Yhdysvaltalaiset yliopistot ovat odotetusti vahvasti edustettuna, kun tarkastellaan tutkimuksen tieteellistä merkitystä. Eurooppalaisista toimijoista Zürichin ETH ja Oxfordin yliopisto ovat lähellä kärkeä käytettäessä mittarina julkaisujen keräämää keskimääräistä ja normalisoitua (odotettavissa olevaan viittausten keskiarvoon suhteutettua) viittausmäärää. Huippulehdissä julkaistujen artikkelien osuudella mitaten parhaiten pärjäsivät v. 2013-2017 Microsoft, Hong Kongin City University (ei kuvassa) ja Granadan yliopisto Espanjassa. Yritys-yhteistyötä oli yliopistoista suhteellisesti eniten Illinois'n, Berkeleyyn ja Stanfordin yliopistoissa, joissa kaikissa kyseisten yhteisjulkaisujen osuus oli yli 10%, mitä voi pitää poikkeuksellisen suurena koko tieteellisen tutkimuksen kenttää tarkasteltaessa.

Yrityksen (Microsoft) asema tämän tekoälyä koskevan vertailun kärjessä on poikkeuksellista tutkimusjulkaisujen merkittävyttä arvioitaessa. Osittain tämä tosin selittyy tiiviillä yhteistyöllä alan huippuyliopistojen kanssa. On silti hyvä todeta, että Microsoftin julkaisumäärät tekoälyn alalla ovat verrannolliset monien isojen yliopistojen julkaisumääriin.

Erilaisten tutkimuksen vaikuttavuutta kuvaavien mittareiden rinnalla on syytä aina tarkastella myös julkaisuvolyymejä. Vaikuttavuusmittarit sellaisenaan voivat antaa hyvinkin vinoutuneen kuvan toimijan (tai maan) tutkimuksen tasosta, mikäli tutkimussuunnan julkaisumäärä on pieni.

Tutkimuslaitos / yliopisto	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimuslalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankiingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys -yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisutietojen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Microsoft USA	4,55	713	45	1,39	87,8	1,1
Stanford University	3,61	619	37,3	0,9	10,5	1,24
King Abdulaziz University	3,43	586	36,7	2,24	0,9	1,41
University of California at Berkeley	3,29	594	34,2	1,03	10,9	1,2
MIT	3,17	1180	28,8	1,1	9,1	1,41
ETH Zurich	2,85	631	26,1	1,63	7,6	1,33
University of Oxford	2,79	674	28	1,57	3,4	1,18
U. of Illinois at Urbana-Champaign	2,5	534	38,2	1,03	11,4	1,2
University of Michigan	2,49	492	-	0,73	9,6	1,31
Nanyang Technological University	2,48	1526	31,2	1,59	4,1	1,42
University of Granada	2,4	575	44,5	1,2	0,9	1,93
Harvard University	2,35	511	35	1,26	8	1,48
Carnegie Mellon University	2,33	1493	25,3	0,97	8,2	1,12
National University of Singapore	2,31	997	40,8	1,6	3,6	1,26
University of Toronto	2,28	516	28,3	1,44	6,8	1,29
Chinese University of Hong Kong	2,24	572	39,1	1,85	4,5	1,1
University of Technology Sydney	2,19	856	35,6	1,72	1,8	1,39
Institute of Automation CAS China	2,18	874	30,3	0,6	2,4	1,03
Hong Kong Polytechnic University	2,13	584	42,8	1,79	1,7	1,69
University of Southern California	2,1	702	24,7	0,94	6	1,09
Istituto Italiano di Tecnologia	2,05	506	-	1,32	2	1,55
University of Maryland	1,98	458	36,2	0,98	8,3	1,19
Imperial College London	1,96	631	24,9	1,58	2,7	1,34
EPFL Lausanne	1,96	454	-	1,69	5,1	1,24
Peking University	1,92	775	29,7	0,96	7,4	1,06
INRIA France	1,9	909	26,1	1,34	4,4	1,11
Georgia Institute of Technology	1,88	870	20,8	0,93	9,4	1,2
Technical University of Munich	1,86	788	13,7	0,94	8,8	1,29
University of Rome La Sapienza	1,82	548	21,4	1,06	2	1,77
University of Texas at Austin	1,82	451	-	0,77	8,6	1,02
Huazhong U. of Science and Tech.	1,77	965	30,1	0,82	0,6	1,19
Xidian University	1,71	857	35	0,57	1,8	1,02
IBM	1,68	510	-	1,36	71,8	1,07

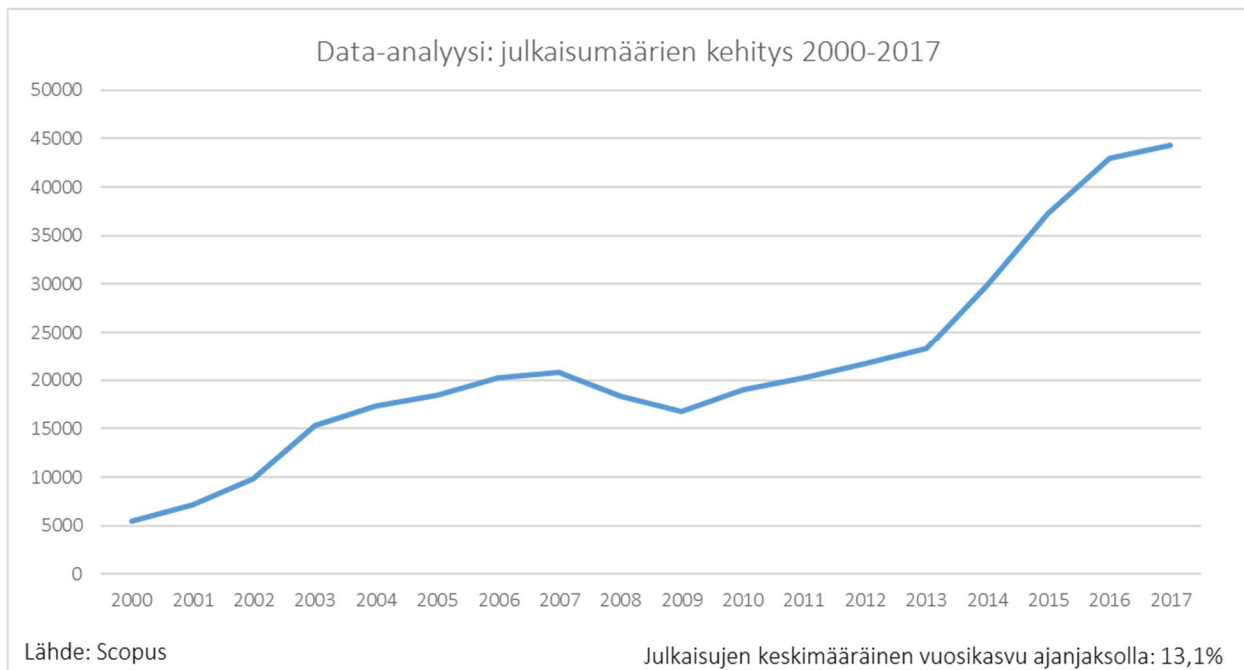
Kuva 16. Tekoäly: merkittävimmät yliopistot, tutkimuslaitokset ja yritykset järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2012-2017. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa (katselutiedot Scopusesta). Normalisoiduissa mittareissa maailman keskiarvo = 1. Lähde: Scopus ja SciVal. Julkaisuja 209 113. Tilanne: 6.6.2018

1.3.2 Data-analyysi¹⁹

Julkaisumäärät ja aiheet

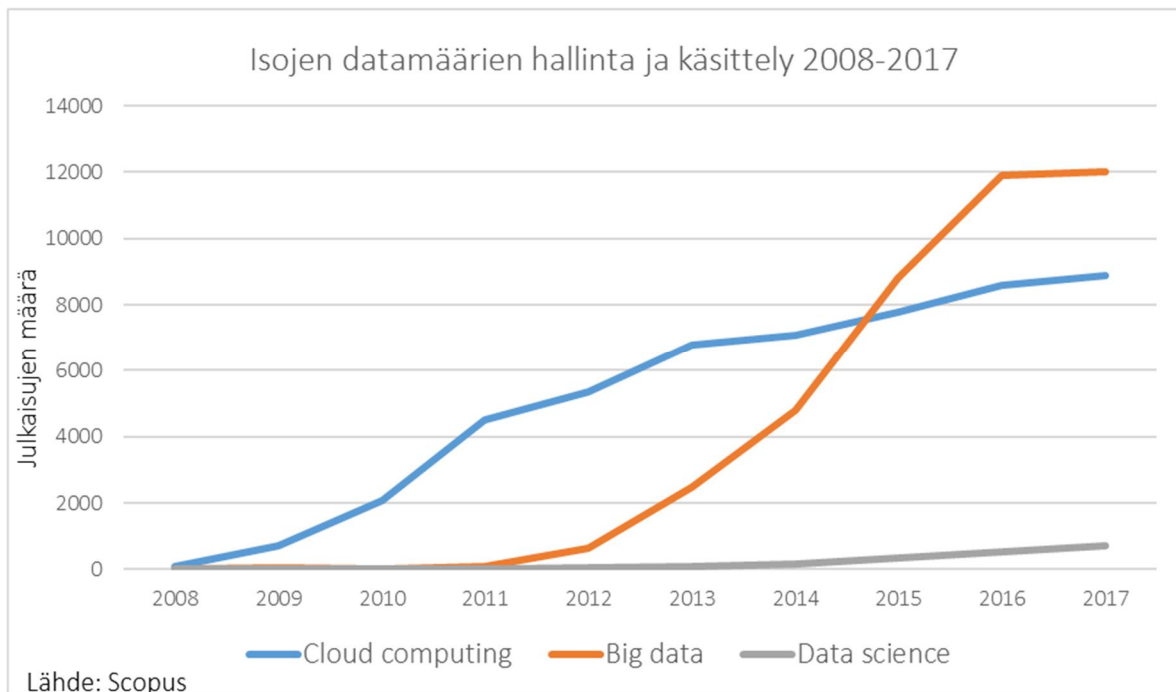
Isojen datamäärien käsittelystä, tallentamisesta ja hyödyntämisestä on tullut viime vuosina myös tekoälyn tutkimuksen valtavirtaa. Julkaisumäärät ovat kehittyneet ripeästi lukuun ottamatta vuosia 2008-2009, jolloin julkaisujen määrä jopa laski väliaikaisesti. Erityisen nopeaa kasvu oli vuosina 2014-2016.

¹⁹ Tutkimussuunnan rajoituksessa käytettiin seuraavia termejä: big data, classification (of information), clustering algorithms, data mining, data analytics, predictive analytics sekä yhdessä tekoälyn yleiskäsitteiden kanssa myös paljon dataa tuottavia sovelluspainotteisia termejä: bioinformatics, gene expression, genetics, internet of things, remote sensing



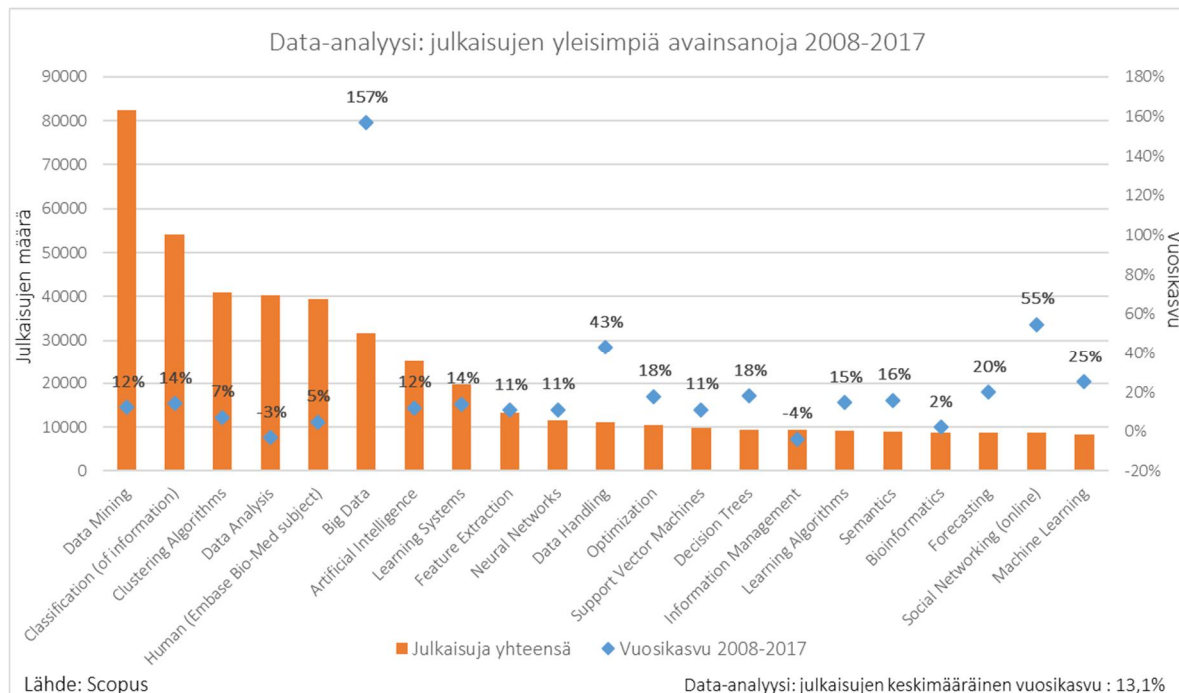
Kuva 17. Tieteellisten julkaisujen määrä 2000-2017. Lähde: Scopus

Datan analysointi edellyttää datan määrätietoista ja suunnitelmallista hallintaa. Tarkasteltavana ajanjaksona nopeimmin nouseva uusi termi julkaisuissa oli ymmärrettävästi "big data". Vielä vuonna 2010 sitä ei käytännössä esiintynyt tutkimusjulkaisuissa ollenkaan ja v. 2017 siihen liittyviä julkaisuja ilmestyi maailmalla jo lähes 10 000. Kasvuvauhti on hyvin poikkeuksellista, erityisesti ottaen huomioon, että itse asia, suuret aineisto- ja datamäärät sekä niiden ympärille rakentuva tutkimus, "data-tiede" eivät sinänsä ole mitenkään uusia teemoja tutkimuksessa. Isojen datamassojen hallinta käsittelyn tutkimusaiheet "cloud computing", "big data" (ja osittain myös "data science") ovat leimallisesti 2010-luvun tutkimusaiheita.



Kuva 18. Isojen datamäärien hallinta. Cloud computing, big data ja data science: julkaisumäärät. Lähde: Scopus

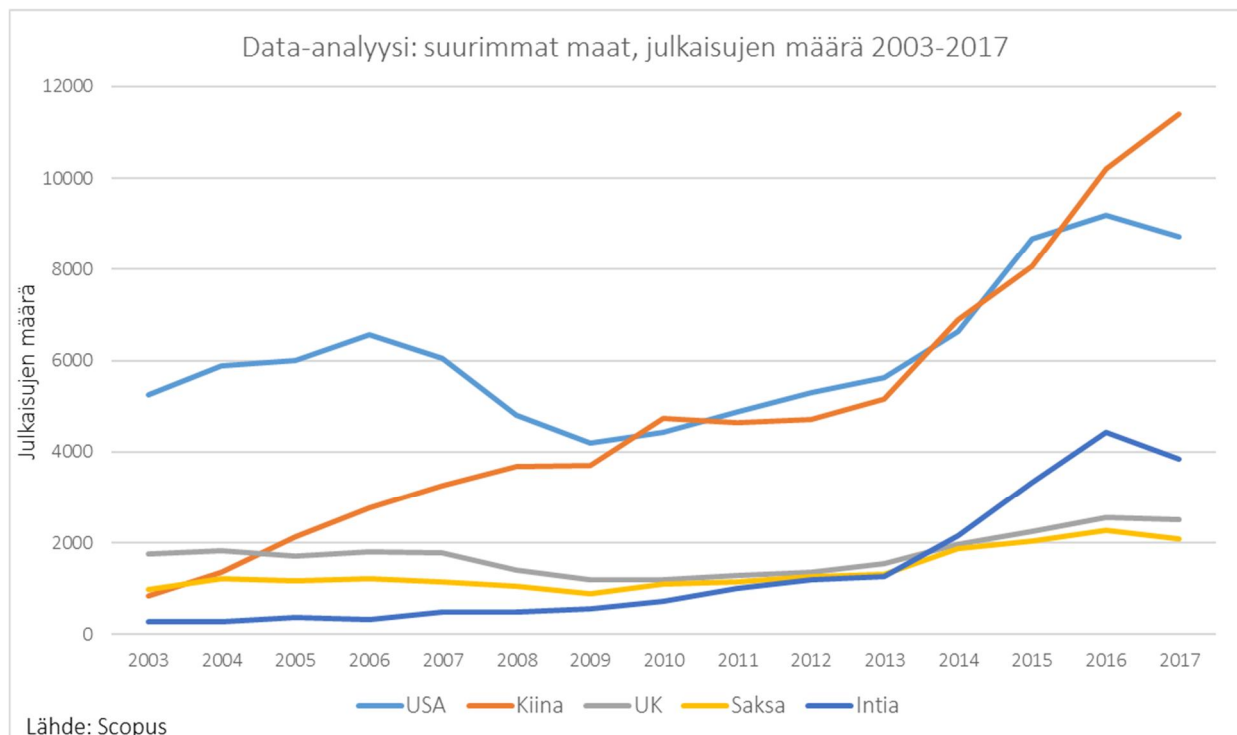
Tarkasteltaessa lähemmin datan analysointiin liittyviä teemoja nousevat datan louhinta sekä luokittelu ja ryhmittely volyymeiltaan keskeisiksi tutkimusaiheiksi.



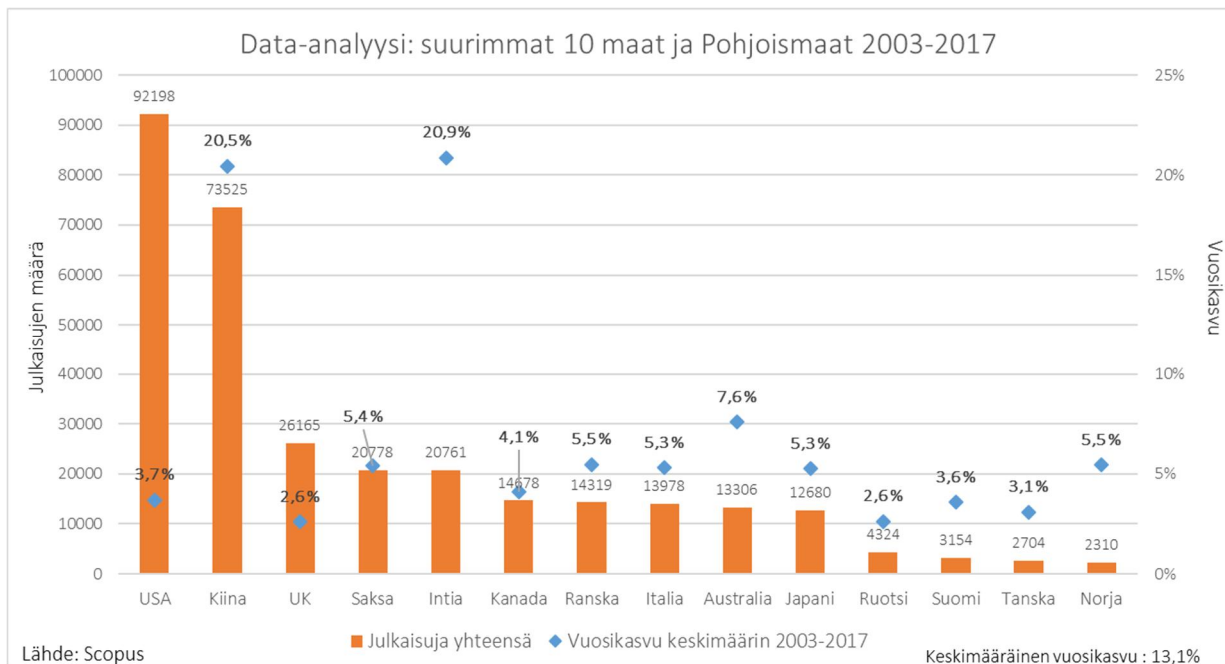
Kuva 19. Data-analyysin keskeisiä tutkimusaiheita: julkaisumäärät ja vuosikasvu. Kaikkien alan julkaisujen vuosikasvu ko. ajanjaksolla oli 13,1%. Osa sisällöltään yleisemmistä avainsanoista on jätetty kuvasta pois. Lähde: Scopus

Maat

Kiina on noussut Yhdysvaltojen ohi suurimmaksi julkaisijaksi. Kiinan kasvu näyttäisi myös edelleen jatkuvan suhteellisen voimakkaana. Intian kasvunopeus data-analyysissä vastaa Kiinan lukuja, mutta v. 2017 kasvu näyttäisi taittuneen. Vielä on tosin liian aikaista sanoa, onko käänne pysyvämpi.



Kuva 20. Data-analyysi, suurimmat maat, julkaisumäärät 2003-2017. Lähde: Scopus



Kuva 21. Data-analyysi: suurimmat kymmenen maata ja Pohjoismaat, julkaisujen määrä ja julkaisumäärien keskimääräinen vuosikasvu 2003-2017. Lähde: Scopus

Maiden julkaisuprofiileja tarkastellessa yleisimpien termien (data mining, big data) suhteellinen osuus oli suurin Kiinassa, Intiassa, Espanjassa ja Australiassa. Lääke- ja terveystieteelliset ("Human (Bio-Medical)") aiheet olivat vahvimmin esillä Iso-Britanniassa, Tanskassa ja Ruotsissa.

Julkaisuja yhteensä 2015-2018	32158	28453	12888	7977	6862	4948	4920	4820	4581	4216	1420	1018	823	801
Termi	Kiina	USA	Intia	UK	Saksa	Italia	Australia	Ranska	Kanada	Espanja	Ruotsi	Suomi	Tanska	Norja
Artificial Intelligence	7,7%	10,3%	16,3%	10,3%	8,9%	11,0%	10,6%	9,6%	11,7%	13,8%	7,4%	10,4%	5,0%	7,9%
Big Data	27,8%	22,2%	17,5%	19,6%	17,5%	19,8%	21,1%	16,2%	17,5%	17,8%	18,3%	20,5%	14,5%	24,0%
Classification (of Information)	25,4%	17,1%	27,7%	19,0%	19,2%	20,3%	22,4%	23,7%	22,4%	26,4%	15,9%	23,5%	13,6%	17,0%
Cloud Computing	4,0%	2,0%	3,7%	1,9%	1,2%	2,0%	2,9%	1,7%	2,4%	1,5%	1,9%	2,1%	1,3%	4,2%
Clustering Algorithms	19,9%	8,7%	15,5%	7,1%	7,9%	9,3%	10,1%	11,6%	10,8%	7,2%	6,5%	10,2%	5,6%	8,2%
Data Analysis	4,5%	15,8%	3,8%	20,6%	14,8%	11,3%	15,1%	11,1%	17,4%	10,5%	19,6%	12,7%	26,6%	19,4%
Data Handling	7,7%	5,6%	5,6%	4,9%	6,9%	6,7%	5,4%	6,0%	5,5%	6,2%	6,3%	6,5%	6,7%	9,6%
Data Mining	26,8%	26,2%	38,9%	21,6%	24,1%	25,8%	27,5%	25,4%	24,9%	27,2%	21,8%	24,5%	18,2%	18,7%
Decision Making	2,2%	2,9%	3,5%	3,2%	2,5%	2,7%	3,7%	2,8%	3,3%	3,0%	3,5%	2,4%	1,6%	2,5%
Decision Trees	3,3%	2,7%	5,6%	3,2%	2,8%	2,4%	3,7%	3,0%	4,2%	3,7%	3,2%	2,5%	1,8%	3,1%
Digital Storage	4,1%	3,4%	3,8%	2,1%	2,5%	2,8%	2,9%	2,2%	2,8%	2,5%	2,0%	3,3%		2,5%
Feature Extraction	6,9%	3,2%	10,0%	3,9%	3,0%	2,8%	5,5%	3,3%	4,5%	5,2%	2,2%	3,5%	2,8%	1,9%
Forecasting	4,3%	4,1%	4,8%	2,9%	2,3%	2,6%	3,7%	2,3%	3,7%	3,2%	2,8%	2,2%	3,6%	3,2%
Human (Bio-Med)	5,6%	21,5%	5,8%	25,2%	17,7%	14,6%	18,9%	13,2%	21,7%	14,0%	24,6%	13,4%	29,4%	22,1%
Image Classification	4,8%	2,0%	3,0%	1,7%	1,8%	2,2%	2,8%	3,0%	2,0%	2,2%		2,7%		
Image Processing	4,9%	2,9%	5,6%	2,5%	3,3%	3,3%	3,0%	4,6%	2,9%	3,5%	2,5%	3,1%	2,3%	1,9%
Information Management	3,7%	2,6%	2,0%	2,5%	3,3%	3,9%	2,4%	3,1%	2,4%	2,9%	3,1%	4,1%	2,7%	2,6%
Learning Algorithms	4,2%	3,9%	5,4%	3,1%	2,5%	2,8%	4,1%	2,8%	4,1%	4,0%	2,3%	2,4%	1,6%	2,4%
Learning Systems	8,8%	10,4%	11,1%	9,4%	8,7%	8,4%	10,4%	8,8%	10,7%	11,2%	6,6%	11,7%	5,7%	8,6%
Machine Learning	2,8%	5,4%	4,3%	5,1%	4,0%	3,5%	4,1%	3,3%	5,3%	5,8%	3,5%	6,0%	3,8%	4,5%
Neural Networks	5,0%	3,0%	5,7%	3,8%	2,8%	3,3%	3,9%	2,8%	3,2%	4,4%	2,3%	3,9%	1,5%	2,9%
Optimization	7,0%	3,3%	6,2%	3,1%	2,1%	2,5%	4,2%	3,5%	3,1%	3,4%	2,5%	3,3%	1,7%	2,9%
Pattern Recognition	3,2%	2,2%	3,4%	2,5%	2,2%	2,6%	3,0%	2,7%	3,3%	3,3%	2,2%	3,9%	1,9%	1,7%
Semantics	4,9%	3,3%	2,7%	3,0%	4,1%	4,9%	3,7%	4,1%	2,2%	3,6%	3,0%	3,0%		2,2%
Social Networking (online)	3,6%	4,2%	5,3%	3,4%	2,8%	4,7%	4,1%	3,2%	3,8%	3,8%	2,4%	3,6%	4,1%	3,5%
Support Vector Machines	5,2%	2,2%	6,7%	2,2%	1,8%	2,2%	3,4%	2,5%	3,4%	2,6%	1,2%	2,3%		2,2%

Kuva 22. Suurimmat kymmenen maata ja pohjoismaat 2015-2018, julkaisujen painotukset. Yleisimmät sisältöavainsanat koko julkaisuaineistossa (yleisimmät kuivailutermit "artificial intelligence" ja "learning systems" eivät mukana vertailussa). Osa julkaisutermeistä jätetty pois termin päällekkäisyyden tai yleisluonteisuuden perusteella. Termien prosentuaalinen osuus on $\neq$ 100% johtuen eri aiheiden esiintymisestä samoissa julkaisuissa. Vuoden 2018 tilanne: 24.4.2018. Lähde: Scopus

Tutkimuksen vaikuttavuus: maat, big data²⁰

Kiina on ylivoimaisesti suurin maa isojen datamäärien tutkimuksessa ja tutkimukseen soveltamisessa. Kun tarkasteluun otetaan julkaisumäärien rinnalle julkaisujen keräämä huomio tiedemaailmassa ja tutkimusyhteistyön intensiteetti, muuttuu kuva kuitenkin jossain määrin erilaiseksi. Käytettäessä mittarina normalisoituja julkaisujen viittausmääriä ovat kolme kärkimaata Saudi-Arabia, Alankomaat ja Ruotsi.

Huippujulkaisujen osuudella mitattuna kärjessä on Thaimaa, tosin maan julkaisumäärä on toistaiseksi suhteellisen pieni. Seuraavina tulevat Hong Kong ja Kanada. Kansainvälinen tutkimusyhteistyö oli v. 2016-2017 intensiivisintä Irlannissa, Hong-Kongissa ja Sveitsissä, kun tarkastellaan yli sata julkaisua tuottaneita maita. Yliopisto-yritys –yhteistyötä oli suhteellisesti eniten Sveitsissä ja Ruotsissa.

Maa	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankiingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys -yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisujen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Saudi Arabia	2,61	135	12,9	1,83	3	2,36
Netherlands	2,4	134	12,7	1,44	4,5	2,07
Sweden	2,31	117	14,5	1,55	11,1	2,52
Australia	2,18	492	12,8	1,63	2,6	2,4
Norway	2,11	103	17,1	1,76	2,9	2,17
United Kingdom	1,91	667	14,7	1,61	4,9	2,5
Austria	1,84	99	7,7	1,84	6,1	2,82
Spain	1,81	370	11,7	1,37	5,7	2,67
Singapore	1,77	140	13,8	1,76	3,6	1,89
Italy	1,73	496	9,4	1,03	3,4	2,98
Canada	1,72	390	18,1	1,48	4,4	2,04
Greece	1,71	174	7	1,23	2,9	2,04
Malaysia	1,69	146	4,3	1,26	0,7	3,94
Russian Federation	1,61	207	1,2	0,55	1,4	2,11
Hong Kong	1,61	170	19,5	1,87	2,4	2,2
Finland	1,58	90	10,5	1,63	7,8	2,78
United States	1,57	2772	15,1	1,02	8,4	1,76
Poland	1,57	121	5,5	0,72	0,8	2,23
Indonesia	1,52	121	9,1	0,69	0	3,32
Switzerland	1,44	123	12,3	1,85	15,4	2,29
Germany	1,32	526	5,1	1,15	8,4	2,21
South Korea	1,31	607	6	0,61	4	1,87
Portugal	1,23	121	0	1,16	1,7	2,72
Ireland	1,16	96	9,8	1,9	8,3	1,62
Taiwan	1,02	316	11,8	0,69	1,9	2,05
France	1,01	395	5,8	1,63	4,1	1,99
China	0,95	5046	7,9	0,61	2,2	1,55
Japan	0,95	479	7	0,73	4,8	1,36

Kuva 23. Big data: maat järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuosi 2016-2017 (viittausindeksin raja-arvo kuvassa: 0,95). Alle 90 julkaisua tuottaneet maat eivät ole mukana taulukossa. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen yhteistyön ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa. (Tiedot Scopusesta: katselukerrat antavat tietoa julkaisun herättämästä kiinnostuksesta täydentämällä viittausmäärien välittämää informaatiota erityisesti uusimman aineiston osalta.) Normalisoiduissa mittareissa maailman keskiarvo = 1. Lähde: Scopus ja SciVal. Analysoituja julkaisuja yhteensä 14817. Tilanne: 7.5.2018

²⁰ Julkaisut rajattiin tässä vertailussa Scopus-tietokannasta käyttämällä hakua: (key ("big data") or title ("big data")) and (subjarea (comp)) and (limit-to (pubyear , 2017) or limit-to (pubyear , 2016)) and (limit-to (doctype , "cp") or limit-to (doctype , "ar"))

Tutkimuksen vaikuttavuus: maat, data mining²¹

Suurimmat julkaisijat v. 2016-2017 olivat odotetusti Kiina, Yhdysvallat ja Intia. Normalisoiduilla viittausmäärillä tarkastellen maista parhaiten sijoittuivat Singapore, Hong Kong ja Alankomaat, joilla kaikilla viittausmäärät olivat yli kaksinkertaiset odotettavissa olevaan kansainväliseen keskiarvoon verrattuna. Suhteellisesti eniten yksittäisiä top 10%:n huippujulkaisuja oli samana ajanjaksona Hong Kongilla, Belgiassa ja Yhdysvalloilla. Kansainvälistä yhteistyötä oli puolestaan suhteellisesti eniten Hong Kongilla, Saudi Arabialla ja Iso-Britanniassa. Yritysten ja yliopistojen välistä yhteistyötä oli eniten Sveitsillä, Yhdysvalloilla ja Ruotsilla.

Maa	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankiingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys - yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisujen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Singapore	2,25	213	14,9	1,56	4,2	1,59
Hong Kong	2,24	196	29,6	1,9	6,6	2,08
Netherlands	2,08	232	12,6	1,44	7,3	1,61
United States	1,86	2932	20,5	1	9,2	1,49
Canada	1,86	478	12,5	1,42	3,1	1,55
Italy	1,85	576	9,1	1	2,8	2,23
Switzerland	1,83	131	17,7	1,49	10,7	1,65
United Kingdom	1,76	703	20,5	1,66	6,3	2,1
Australia	1,63	602	16,1	1,49	3,8	1,81
Malaysia	1,61	252	8,7	1,16	0,8	4,09
Belgium	1,6	117	21,1	1,63	4,3	1,41
Indonesia	1,57	196	3,7	0,7	0	2,06
Spain	1,56	503	19,9	1,35	2	2,3
Russian Federation	1,55	294	1,7	0,61	0,7	2,18
Finland	1,51	115	14	1,36	5,2	1,81
Austria	1,5	133	9,7	1,64	3	1,36
Poland	1,47	276	6,2	0,63	0,4	1,42
Greece	1,31	185	13	1,01	1,6	1,59
Viet Nam	1,31	136	10,1	1,14	0	1,29
Sweden	1,29	108	18,4	1,31	8,3	1,58
Germany	1,27	678	8,5	1,16	6,8	1,43
Iran	1,24	239	13	0,5	0	1,66
South Korea	1,2	428	17,5	0,81	4,2	1,85
Saudi Arabia	1,2	220	10,7	1,89	0,9	1,64
Portugal	1,19	232	7,7	1,11	2,2	1,97
France	1,12	541	13,4	1,46	3,5	1,23
China	1,08	3771	10,7	0,71	3,2	1,19

Kuva 24. Datan louhinta: maat järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2016-2017 (viittausindeksin raja-arvo kuvassa: 1,08). Alle 100 julkaisua tuottaneet maat eivät ole mukana taulukossa. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen yhteistyön ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa. (Tiedot Scopusesta: katselukerrat antavat tietoa julkaisun herättämästä kiinnostuksesta täydentämällä viittausmäärien välittämää informaatiota erityisesti uusimman aineiston osalta.) Normalisoiduissa mittareissa maailman keskiarvo = 1. Lähde: Scopus ja SciVal. Analysoituja julkaisuja yhteensä 17405. Tilanne v. 2017 osalta: 7.5.2018

Tutkimuksen vaikuttavuus: tutkimusorganisaatiot, big data

²¹ Julkaisut rajattiin tässä vertailussa Scopus-tietokannasta käyttämällä hakua: (key ("data mining") or title ("data mining")) and (subjarea (comp)) and (limit-to (pubyear , 2017) or limit-to (pubyear , 2016)) and (limit-to (doctype , "cp") or limit-to (doctype , "ar"))

Yliopistoista ja tutkimuslaitoksista eniten big dataan liittyviä julkaisuja tuottivat v. 2016-2017 Kiinan tiedeakatemia CAS ja Tsinghua-yliopisto. Tarkasteltaessa tutkimuksen vaikuttavuutta suhteutetuilla viittausmäärillä kärjessä oli Kiinan Xiamenin yliopisto. Suuremmista toimijoista paras tulos oli Shanghain Jiao Tong –yliopistolla. Huippujulkaisujen osuudella mitattuna kärkiyliopistot olivat Hong Kongin polytekninen yliopisto sekä Stanfordin ja Melbournen yliopistot. Pidettäessä mittarina kansainvälisen yhteistyön osuutta tutkimusjulkaisuissa oli kärjessä Aalto yliopisto. Yritysten ja yliopistojen yhteistyössä kärjessä olivat amerikkalaiset yliopistot: MIT ja Marylandin yliopisto (ei mukana kuvassa).

Tutkimuslaitos / yliopisto	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimuslalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys - yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisujen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Xiamen University	3,95	34	17,6	1,15	0	2,58
King Saud University	3,46	42	13,8	1,77	4,8	2,92
Kyungpook National University	3,31	35	11,8	0,81	2,9	3,22
University of Naples Federico II	3,06	40	11,8	0,7	2,5	4,44
Central South University	2,6	51	11,1	0,67	0	1,59
Deakin University	2,59	64	9,7	2,09	0	1,97
Massachusetts Institute of Technology	2,57	43	6,2	0,98	20,9	2,63
Stanford University	2,52	38	27,8	1,63	5,3	1,61
University of Southern California	2,43	38	17,6	1,04	10,5	1,82
Shenzhen University	2,34	43	23,1	1,16	0	1,54
Sun Yat-Sen University	2,24	55	14,8	1,03	0	2,57
South China University of Technology	2,24	68	8,1	0,65	0	2,99
University of Melbourne	2,19	43	27,3	1,27	4,7	2,57
Hong Kong Polytechnic University	2,18	47	29,6	2,15	2,1	2,09
Shanghai Jiao Tong University	2,17	124	11,6	0,98	7,3	1,8
CNR	2,11	68	11,8	0,63	0	3,31
UTS University of Technology Sydney	2,02	53	3,8	1,51	1,9	2,03
University of California at San Diego	2,01	40 -		0,6	10	1,67
Georgia State University	1,97	40 -		1,36	0	1,35
University of New South Wales	1,97	52	24,1	1,74	9,6	2,5
Wuhan University	1,97	103	5,9	0,76	1	1,92
Dalian University of Technology	1,94	84	10,6	0,69	1,2	1,62
Tongji University	1,93	86	12,8	1,08	9,3	1,69
Aalto University	1,92	33	10,5	2,21	9,1	1,82
New Jersey Institute of Technology	1,91	37	15,8	1,38	0	1,31
Intel	1,89	35 -		0,62	68,6	1,45

Kuva 25. Big data: merkittävimmät yliopistot, tutkimuslaitokset ja yritykset järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2016-2017. Viittausindeksin raja-arvo kuvassa: 1,89. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa (katselutiedot Scopusesta). Normalisoiduissa mittareissa maailman keskiarvo = 1. Analysoituja julkaisuja yhteensä 14817. Lähde: Scopus ja SciVal, tilanne: 7.5.2018

Tutkimuksen vaikuttavuus: tutkimusorganisaatiot, data mining

Suurimmat julkaisujen tuottajat olivat Kiinan Tiedeakatemia, Tsinghua-yliopisto ja Ranskan kansallinen tutkimuskeskus CNRS. Yhdysvaltalaiset yliopistot olivat puolestaan viittausmäärissä mitaten kärjessä. Yksittäisiä huippujulkaisuja oli eniten Hong Kongin City-yliopistolla ja polyteknisellä yliopistolla (45%). Kolmantena listalla on Granadan yliopisto Espanjassa. Eniten kansainvälistä yhteistyötä oli v. 2016-2017 King Abdulaziz –yliopistolla ja Lontoon UCL:llä, molemmilla yli kaksinkertaisesti kansainväliseen keskiarvoon verrattuna. Hong Kongin polyteknisellä yliopistolla oli suhteellisesti eniten yritys-yhteistyötä (14% julkaisuista tehty yhteistyössä yrityksen kanssa, ei mukana kuvassa).

Tutkimuslaitos / yliopisto	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankiingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisujen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
University of Washington	6,67	38	-	0,83	0	3,39
University of Illinois at Chicago	4,84	56	-	1,58	6,6	1,91
Stanford University	4,42	66	20,7	0,76	0	1,29
Queensland University of Technology	4,4	40	23,8	1,52	0	0,65
Universiti Tun Hussein Onn Malaysia	4,36	47	-	1,19	0	2,05
Eindhoven University of Technology	3,77	69	14	1,39	4,2	1,59
Yahoo Research Labs	3,65	43	-	0,81	0	1,98
Arizona State University	3,56	95	13,8	0,79	1,8	1,46
Hong Kong U. of Science and Technology	3,26	42	-	1,93	0	2,55
Nanyang Technological University	2,88	98	9,1	1,57	2	2,3
University of Rome La Sapienza	2,88	40	-	1,62	0	2,83
Microsoft USA	2,82	69	-	1,47	6,6	2,08
University of Malaya	2,75	48	12,5	1,85	0	1,95
Google Inc.	2,61	50	-	1,11	0	2,37
Tsinghua University	2,58	195	13,7	1,17	9,2	1,49
SUNY Buffalo	2,46	43	-	0,96	0	1,42
Rutgers - U. of New Jersey, New Brunswick	2,37	47	-	1,48	0	1,95
Carnegie Mellon University	2,32	108	20,6	1,15	3,8	1,81
University of Minnesota	2,32	44	-	0,48	0	1,02
Georgia Institute of Technology	2,26	52	21,7	0,99	4,2	1,96
University of Illinois at Urbana-Champaign	2,2	68	-	0,74	2,2	2,2
University of Granada	2,19	41	27,8	1,43	0	0,6
Harbin Institute of Technology	2,07	148	9,4	1,37	6,8	1,43
Universite Paris Saclay	2,05	81	6,7	1,29	0,4	1,42
CNR	2,04	84	3,9	0,55	0,7	2,18
King Abdulaziz University	2,03	54	24,3	2,15	7,2	1,17
Microsoft Research Asia	2	39	-	1,16	0	1,49
Northwestern Polytechnical University Xian	1,96	41	19,2	0,57	0	1,63
University of Michigan	1,95	51	-	0,89	3,6	1,24
Aalto University	1,89	38	-	1,76	0	0,7
Massachusetts Institute of Technology	1,89	66	25,8	1,01	0	1,24
UTS University of Technology Sydney	1,85	81	27,5	1,76	1,5	1,57
IBM	1,83	98	14,8	1,15	3,5	1,23

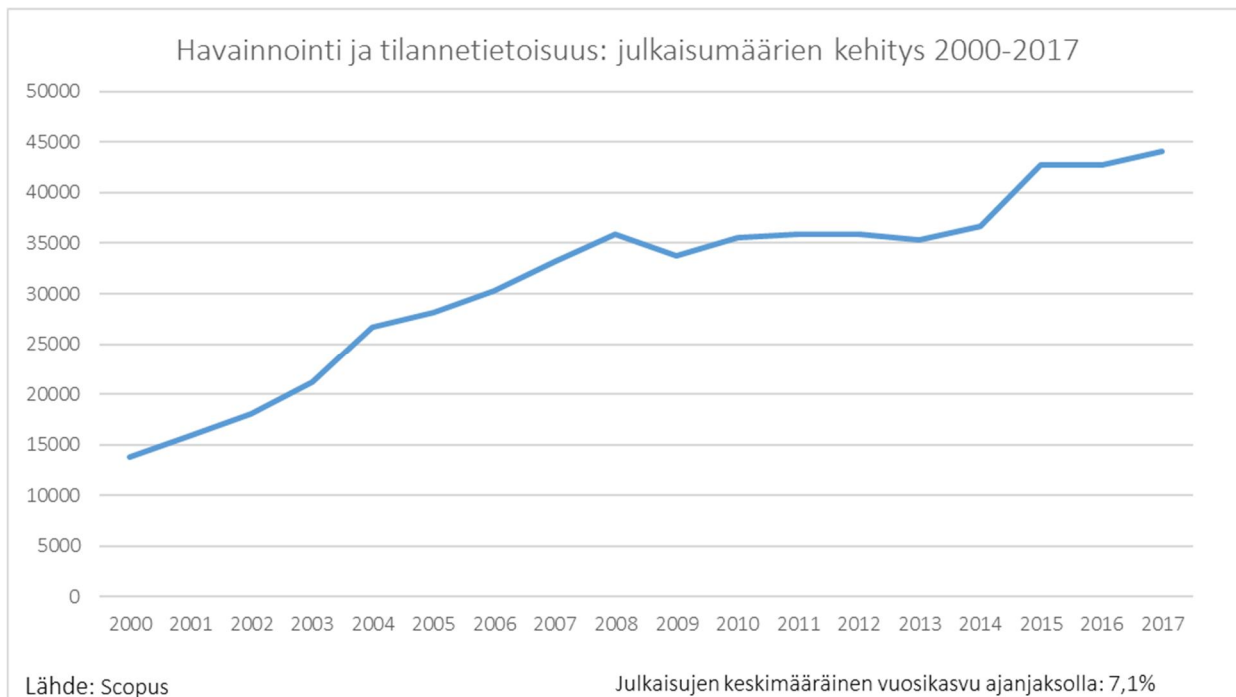
Kuva 26. Datan louhinta: merkittävimmät yliopistot, tutkimuslaitokset ja yritykset järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2016-2017. Viittausindeksin raja-arvo kuvassa: 1,83. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa (katselutiedot Scopusksesta). Normalisoiduissa mittareissa maailman keskiarvo = 1. Analysoituja julkaisuja yhteensä 17405. Lähde: Scopus ja SciVal, tilanne: 7.5.2018

1.3.3 Havainnointi ja tilannetietoisuus²²

Julkaisumäärät ja aiheet

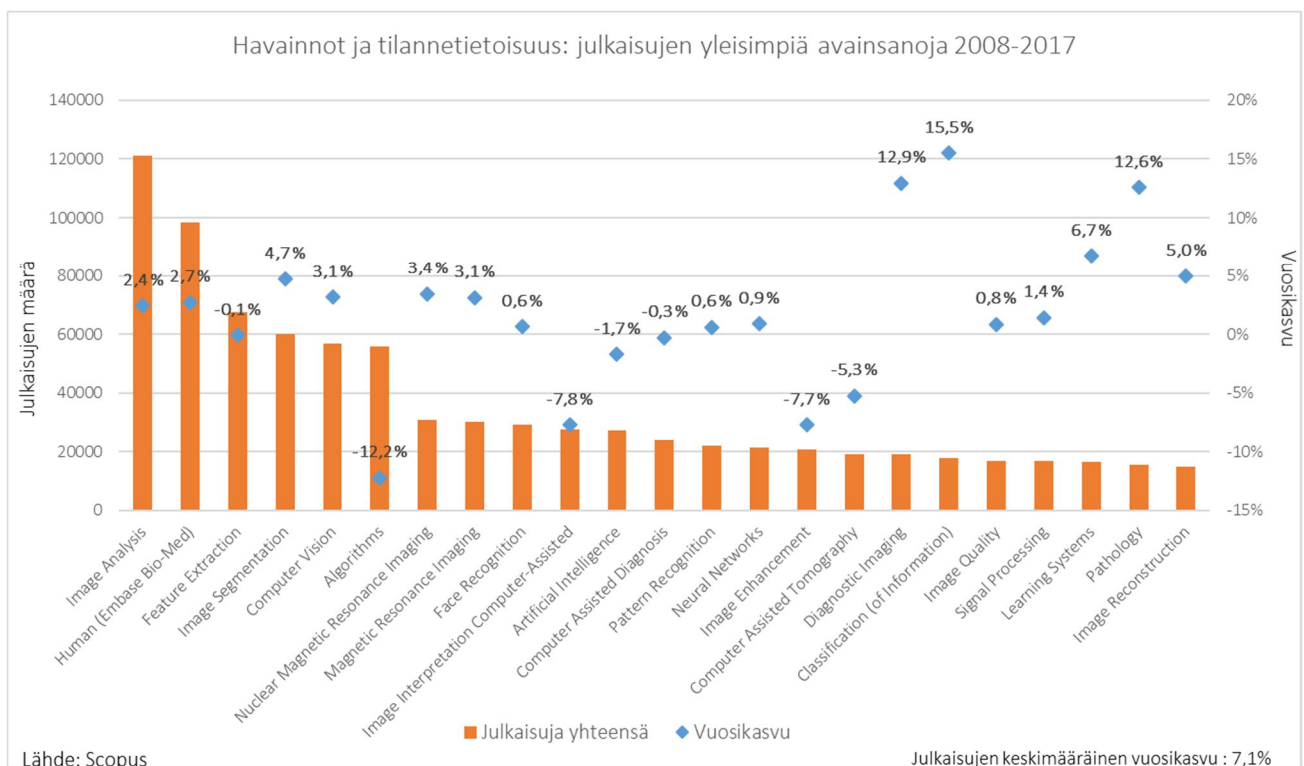
Havainnointia ja tilannetietoisuutta käsittelevien julkaisujen määrä on kehittynyt tarkastelujaksolla melko tasaisesti. Jaksoon on kuitenkin osunut myös useamman vuoden jakso v. 2009-2014 jolloin julkaisumäärät pysyivät käytännössä ennallaan. Viimeinen merkittävä julkaisujen määrällinen lisäys osui vuoteen 2015. Nopeat kasvupyrähdykset eivät kuitenkaan ole olleet leimallisia tutkimussuunnalle, toisin kuin eräillä muilla tekoälyn tutkimusaloilla.

²² Tutkimussuunnan rajauksessa käytettiin seuraavia termejä: ambient intelligence, automated pattern recognition, computer vision, face recognition, feature extraction, image analysis, image classification, image interpretation, image segmentation, machine vision, pattern recognition sekä osittain myös: behavioral research, medical imaging, quality control, remote sensing, virtual reality



Kuva 27. Tieteellisten julkaisujen määrä 2000-2017. Lähde: Scopus

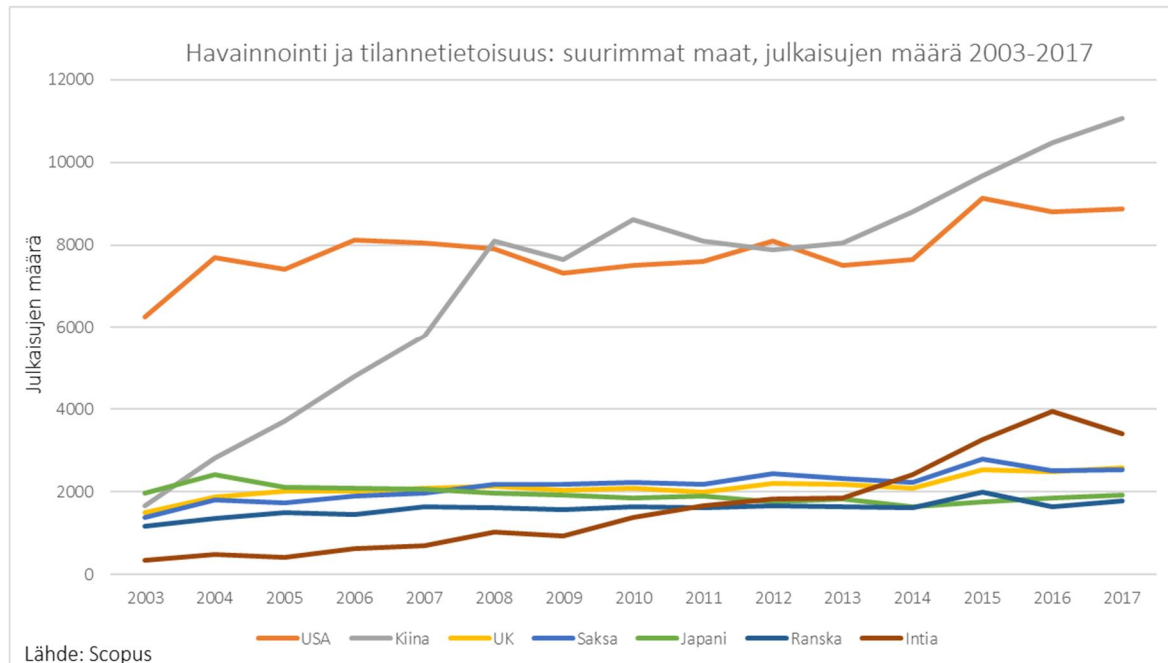
Koska erilaiset näkemiseen ja muuhun havaitsemisen ja tunnistamisen liittyvät tehtävät ovat perinteisesti kuuluneet keskeisiin tekoälyn sovelluskohteisiin, on ymmärrettävää, että myös kyseisen tutkimussuunnan julkaisumäärät muodostavat huomattavan suuren osan myös koko tekoälykentän julkaisuista. Suuret volyymit selittävät osaltaan myös tutkimussuunnan suhteellisen tasaisen kehityksen.



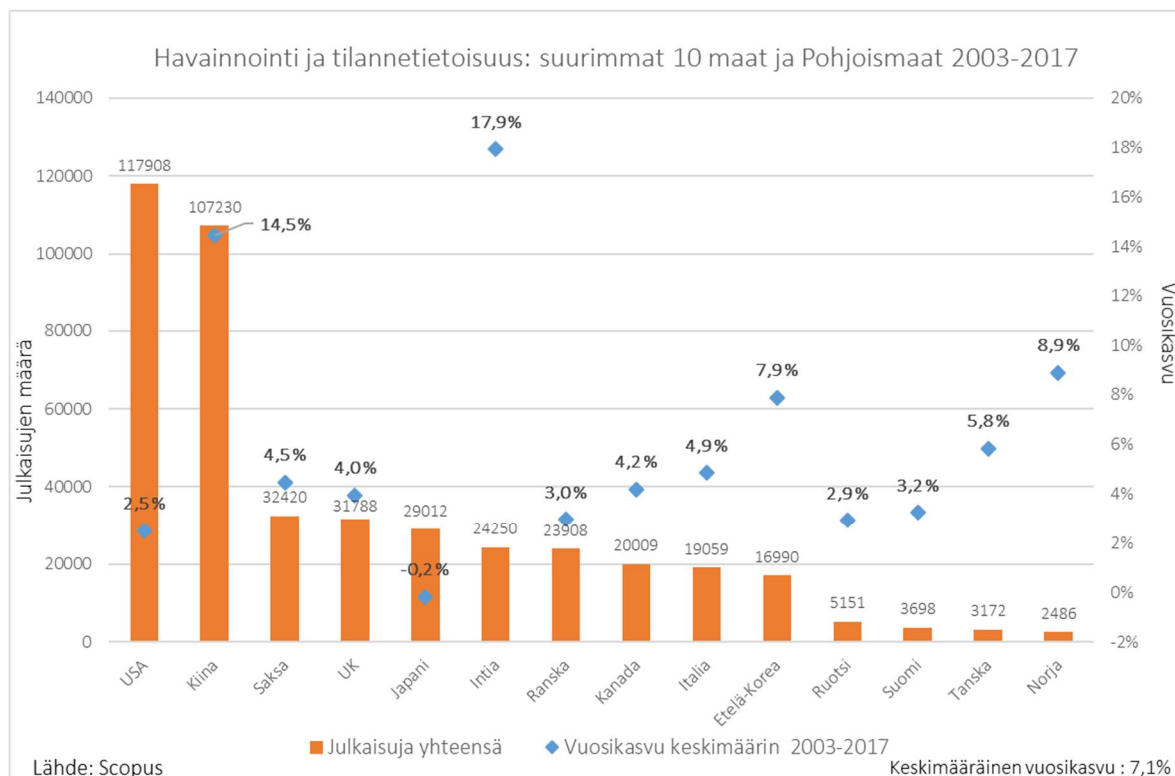
Kuva 28. Keskeisiä tutkimusaiheita: julkaisumäärät ja vuosikasvu. Kaikkien alan julkaisujen vuosikasvu ko. ajanjaksolla oli 7,1%. Osa sisällöltään yleisistä avainsanoista on jätetty kuvasta pois. Lähde: Scopus

Maat

Suurimmat julkaisujen tuottajat ovat odotetusti USA ja Kiina. Kuvien tunnistamiseen, konenäköön ja yleisemmin ympäristön ja tilan havainnoimiseen liittyvän tutkimuskirjallisuuden määrässä Kiina ohitti Yhdysvallat jo vuonna 2008, mikä on aikaisemmin kuin eräillä muilla tutkimusaloilla. Intian nopea nousu näkyy myös täällä, tosin nopeimman kasvun vaihe näyttäisi taittuneen viime vuonna.



Kuva 29. Havainnointi ja tilannetietoisuus, suurimmat maat, julkaisumäärät 2003-2017. Lähde: Scopus



Kuva 30. Havainnointi ja tilannetietoisuus: suurimmat kymmenen maata ja Pohjoismaat, julkaisujen määrä ja julkaisumäärien keskimääräinen vuosikasvu 2003-2017. Lähde: Scopus

Tarkasteltaessa julkaisujen painotuksia maittain huomio kiinnittyy biolääketieteen vahvaan asemaan myös data-analyysiin kohdistuvassa tutkimuksessa. Alan jossain määrin perinteistä tutkimuskohdetta,

tietokonenäköä, käsitteleviä tai siihen liittyviä julkaisuja oli suhteellisesti eniten Suomessa ja Kiinassa. Esimerkiksi NMRI-kuvantamiseen liittyviä julkaisuja ilmestyi puolestaan eniten Saksassa, Yhdysvalloissa ja Tanskassa.

Julkaisuja yhteensä 2015-2018	33814	28647	11585	8372	8215	5860	5828	5362	5048	5048	1378	980	934	739
Termi	Kiina	USA	Intia	Saksa	UK	Japani	Ranska	Italia	Kanada	Etelä-Korea	Ruotsi	Tanska	Suomi	Norja
Algorithms	12,7 %	11,1 %	7,9 %	10,9 %	9,1 %	6,9 %	9,1 %	8,2 %	10,9 %	10,2 %	10,7 %	10,0 %	9,5 %	7,4 %
Artificial Intelligence	6,1 %	6,9 %	11,2 %	5,9 %	8,3 %	6,6 %	6,4 %	8,2 %	6,9 %	9,0 %	4,6 %	4,6 %	7,9 %	5,0 %
Classification (of Information)	10,1 %	5,6 %	13,6 %	4,5 %	6,5 %	4,5 %	6,7 %	5,8 %	7,0 %	5,1 %	4,4 %	3,9 %	9,0 %	3,0 %
Computer Assisted Diagnosis	3,7 %	11,1 %	1,7 %	13,8 %	8,9 %	10,9 %	7,9 %	8,3 %	9,7 %	10,6 %	13,4 %	12,7 %	7,4 %	9,1 %
Computer Assisted Tomography	2,2 %	6,2 %	2,0 %	6,2 %	4,2 %	9,4 %	4,6 %	5,3 %	5,6 %	7,5 %	5,0 %	7,1 %	3,7 %	3,1 %
Computer Vision	16,0 %	15,7 %	12,5 %	15,4 %	13,9 %	14,0 %	15,5 %	13,2 %	13,6 %	14,6 %	13,5 %	11,6 %	16,1 %	11,4 %
Diagnostic Imaging	4,3 %	13,9 %	2,4 %	14,7 %	11,0 %	15,5 %	9,4 %	12,5 %	12,1 %	14,6 %	14,4 %	16,3 %	9,2 %	11,6 %
Face Recognition	9,6 %	4,8 %	11,0 %	3,0 %	6,2 %	5,0 %	4,4 %	4,0 %	5,4 %	7,8 %	3,0 %	3,5 %	10,2 %	6,4 %
Feature Extraction	24,2 %	9,2 %	30,9 %	7,6 %	12,0 %	11,1 %	8,9 %	9,0 %	11,9 %	13,5 %	8,9 %	7,6 %	11,3 %	10,7 %
Human (Bio-Medical)	13,1 %	40,4 %	10,8 %	41,7 %	37,8 %	39,9 %	28,8 %	36,7 %	37,1 %	35,9 %	43,0 %	47,8 %	27,1 %	35,3 %
Image Analysis	21,7 %	42,7 %	17,0 %	43,7 %	41,1 %	44,8 %	38,7 %	46,5 %	39,2 %	36,0 %	48,1 %	51,7 %	39,7 %	46,5 %
Image Classification	8,4 %	4,0 %	6,5 %	3,2 %	3,4 %	3,2 %	5,1 %	4,6 %	3,9 %	3,0 %	2,5 %		4,8 %	3,5 %
Image Interpretation Computer-Assisted	2,5 %	8,2 %		10,1 %	7,1 %	6,5 %	6,0 %	6,2 %	7,4 %	7,1 %	8,9 %	8,5 %	5,7 %	7,3 %
Image Processing	20,1 %	17,4 %	22,9 %	19,1 %	17,7 %	15,7 %	18,9 %	18,6 %	17,9 %	15,4 %	18,4 %	15,6 %	17,9 %	19,5 %
Image Segmentation	22,5 %	14,6 %	26,7 %	14,2 %	14,9 %	10,7 %	19,6 %	12,2 %	19,4 %	15,4 %	15,9 %	13,4 %	15,0 %	14,3 %
Learning Systems	7,1 %	6,8 %	7,6 %	5,3 %	7,7 %	5,6 %	5,4 %	5,3 %	6,9 %	4,7 %	3,6 %	4,0 %	8,7 %	4,7 %
Magnetic Resonance Imaging	4,1 %	12,3 %	5,3 %	13,3 %	11,2 %	9,1 %	8,5 %	9,3 %	12,7 %	11,3 %	11,7 %	12,6 %	9,0 %	11,2 %
Medical Imaging	2,8 %	5,6 %	6,9 %	4,1 %	5,7 %	3,1 %	4,3 %	2,7 %	5,6 %	3,1 %	3,6 %	4,3 %	2,6 %	3,5 %
Neural Networks	8,0 %	5,4 %	8,1 %	4,0 %	6,2 %	5,2 %	4,0 %	5,4 %	4,6 %	5,9 %	3,1 %	2,7 %	6,4 %	4,2 %
Nuclear Magnetic Resonance Imaging	4,1 %	15,1 %	3,0 %	16,1 %	12,6 %	12,7 %	9,7 %	11,4 %	14,3 %	13,2 %	14,7 %	15,1 %	10,6 %	14,2 %
Pathology	3,0 %	10,5 %	1,9 %	10,1 %	9,1 %	10,9 %	7,3 %	10,0 %	8,9 %	10,7 %	10,2 %	10,3 %	8,7 %	10,7 %
Pattern Recognition	7,3 %	7,5 %	8,8 %	6,0 %	6,9 %	5,6 %	6,5 %	6,3 %	6,4 %	5,5 %	6,0 %	3,6 %	8,8 %	3,8 %
Physiology	1,6 %	6,3 %		7,1 %	6,8 %	4,8 %	5,0 %	4,9 %	6,0 %	3,4 %	8,0 %	7,3 %	5,0 %	5,5 %
Signal Processing	4,4 %	3,2 %	6,8 %	3,3 %	3,6 %	2,6 %	4,2 %	2,9 %	4,0 %	2,7 %	3,3 %	3,6 %	4,3 %	2,3 %

Kuva 31. Suurimmat kymmenen maata ja pohjoismaat 2015-2018, julkaisujen painotukset. Yleisimmät sisältöavainsanat koko julkaisuaineistossa. Osa julkaisutermeistä jätetty pois termin päällekkäisyyden tai yleisluonteisuuden perusteella. Termien prosentuaalinen osuus on $\neq 100\%$ johtuen eri aiheiden esiintymisestä samoissa julkaisuissa. Vuoden 2018 tilanne: 24.4.2018. Lähde: Scopus

Tutkimuksen vaikuttavuus: maat

Yhdysvallat ja Kiina ovat ylivoimaisesti suurimmat maat tietokonenäössä ja hahmon tunnistuksessa²³, kun tarkastellaan pelkästään julkaisujen määriä vuosina 2013-2017. Kun tarkasteluun otetaan julkaisumäärien rinnalle julkaisujen keräämä huomio ja julkaisujen hyödyntäminen tiedemaailmassa sekä tutkimusyhteistyön intensiteetti, muuttuu kuva kuitenkin erilaiseksi.

Käytettäessä mittarina suhteutettua viittausindeksiä ovat kolme kärkimaata Hong Kong, Yhdysvallat ja Israel. Eniten yksittäisiä top 10% -huippujulkaisuja tuottivat Israel, Hong Kong ja Singapore. Eniten kansainvälistä yhteistyötä tutkimusjulkaisujen tuottamisessa oli Saudi-Arabiassa ja Hong Kongilla. Yritys-yliopisto – yhteistyötä oli puolestaan eniten Yhdysvalloilla, Alankomailla ja Sveitsillä. Seuraavina tulivat Suomi ja Israel, joten Suomessa oli yritysten ja yliopistojen / tutkimuslaitosten yhteistyö kansainvälisestikin katsoen tiivistä.

²³ Tutkimusteeman rajauksena käytettiin tässä vertailussa SciVal-tietokannan tutkimusalueita: computer vision and pattern recognition

Maa	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys -yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisujen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Hong Kong	2,69	1442	42,6	1,83	5,7	1,3
United States	2,02	20658	33	1	9,1	1,05
Israel	1,99	634	43,7	1,02	7,4	0,81
Singapore	1,88	2031	35,4	1,65	4	1,15
Switzerland	1,88	1800	34,6	1,59	7,7	1,11
Australia	1,66	3523	30,8	1,33	6,7	1,38
United Kingdom	1,63	6200	28,6	1,46	4,9	1,21
Sweden	1,55	1299	21,6	1,36	6,2	1,27
Netherlands	1,51	1642	25,9	1,52	7,9	1,28
Belgium	1,45	1042	27,9	1,69	4	1,03
Italy	1,41	3963	12,8	1,17	2,9	1,69
Germany	1,37	7344	22,7	1,07	6,1	1,03
Canada	1,34	3541	22,9	1,21	3,3	1,14
Czech Republic	1,33	623	29,6	1,03	3,9	1,32
Saudi Arabia	1,32	653	19,3	1,97	1,1	2,04
Norway	1,31	623	11,3	1,57	2,2	1,26
France	1,29	6186	21,5	1,28	4	0,9
Austria	1,27	1439	26,9	1,26	4,1	1,26
Finland	1,2	1159	14,1	1,34	7,5	1,32
Spain	1,11	3336	16,1	1,22	2,1	1,31
Denmark	1,11	779	17,4	1,39	4,7	1,19
China	1,04	22334	18	0,72	3,1	0,86
South Korea	0,98	3398	15,5	0,65	6,2	0,95
Brazil	0,92	1825	9	0,86	2	1,13
Pakistan	0,92	674	5,3	1,12	1	1,22
Portugal	0,91	1378	11,6	1,09	1,7	1,21

Kuva 32. Tietokonenäkö ja hahmon tunnistus: maat järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2013-2017 (viittausindeksin raja-arvo kuvassa: 1). Alle 600 julkaisua tuottaneet maat eivät ole mukana taulukossa. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen yhteistyön ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa. (Tiedot Scopusesta: katselukerrat antavat tietoa julkaisun herättämästä kiinnostuksesta täydentämällä viittausmäärien välittämää informaatiota erityisesti uusimman aineiston osalta.) Normalisoiduissa mittareissa maailman keskiarvo = 1. Julkaisujen määrä analyysissä: 115849. Lähde: Scopus ja SciVal. Tilanne: 9.5.2018

Tutkimuksen vaikuttavuus: tutkimusorganisaatiot

Yliopistoista ja tutkimuslaitoksista selvästi eniten tietokonenäköön ja hahmontunnistukseen liittyviä julkaisuja tuottivat v. 2013-2017 Kiinan tiedeakatemia CAS ja Ranskan kansallinen tutkimuskeskus CNRS.²⁴

Tarkasteltaessa tutkimuksen vaikuttavuutta suhteutetuilla viittausmäärillä, huippujulkaisujen osuudella, tutkimusyhteistyön määrällä ja julkaisutietojen katselukerroilla, niin kuva muuttuu erilaiseksi.

Suhteutettuja viittausmääriä tarkasteltaessa olivat kärjessä Berkeleyn yliopisto, Google ja Stanfordin yliopisto. Amerikkalaiset teknologiayritykset, Google ja Microsoft, olivat vertailun kärjessä myös käytettäessä mittarina huippujulkaisujen osuutta, kolmantena Australian kansallinen yliopisto, joka on korkealla myös käytettäessä mittarina yritys-yliopistoyhteistyön osuutta julkaisujen tuottamisessa. Suurista yrityksistä Microsoft teki enemmän yhteistyötä yliopistojen kanssa kuin Google. Kansainvälistä tutkimusyhteistyötä on suhteellisesti eniten Hong Kongin yliopistoissa, mutta tilastoon vaikuttanee tiivis yhteistyö Manner-Kiinan yliopistojen kanssa. Eurooppalaisista yliopistoista kansainvälistä yhteistyötä harjoittivat eniten ETH Zürich, Leuvenin kuninkaallinen yliopisto ja Lontoon UCL.

²⁴ Tutkimusteeman rajauksena käytettiin tässä vertailussa SciVal-tietokannan tutkimusaluetta: computer vision and pattern recognition

Tutkimuslaitos / yliopisto	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankiingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys - yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisujen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
University of California at Berkeley	7,34	446	57,8	0,92	15,7	1,25
Google Inc.	6,75	294	83,3	0,9	72,1	1,1
Stanford University	6,55	459	59,7	1,12	12,9	1,35
Microsoft USA	6,28	743	75,6	1,45	89,2	1,2
University of Oxford	5,03	355	54,6	1,52	7	1,38
Chinese University of Hong Kong	4,58	528	56,4	1,86	5,7	1,29
University of Michigan	4,12	325	27,4	0,99	8,9	1,23
University of Washington	3,05	319	42,8	0,73	13,8	1,08
Carnegie Mellon University	2,81	916	53,7	1,08	8,7	1,12
Imperial College London	2,74	484	43,1	1,52	4,3	1,48
ETH Zurich	2,7	630	49,2	1,64	9,8	1,11
Massachusetts Institute of Technology	2,69	671	35	1,11	9,2	1,29
University of Maryland	2,67	426	55,1	0,71	14,6	1,08
Australian National University	2,62	307	71,8	1,23	28,7	2
Columbia University	2,47	291	41,4	1,22	15,8	1,1
Royal Institute of Technology	2,44	271	10,2	1,16	5,5	1,14
UTS University of Technology Sydney	2,44	459	30,1	1,77	4,1	1,45
University of Illinois at Urbana-Champaign	2,34	483	52,4	0,91	10,6	1,29
Hong Kong Polytechnic University	2,33	272	44	1,69	0,7	1,26
University of California at San Diego	2,31	343	43,5	0,96	14,3	1,22
PSL Research University	2,3	281	40,7	1,27	5	1,06
Graz University of Technology	2,27	287	50,7	1,33	5,2	1,71
University of Texas at Austin	2,23	294	36,6	0,71	10,2	1,28
INRIA France	2,23	935	42,3	1,34	4,4	1
Queen Mary University of London	2,22	273	-	1,34	2,6	0,97
University of California at Los Angeles	2,2	275	58,3	1,18	7,6	1,08
National University of Singapore	2,12	765	46	1,71	4,8	1,14
Johns Hopkins University	2,07	309	35	1,18	8,7	1,07
University College London	2,06	343	39,9	1,55	5,5	1,5
Dalian University of Technology	2,06	294	31,7	0,95	5,1	0,8
KU Leuven	2,06	403	42	1,63	4,2	1,05

Kuva 33. Merkittävimmät yliopistot, tutkimuslaitokset ja yritykset järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2013-2017. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa (katselutiedot Scopusesta). Normalisoiduissa mittareissa maailman keskiarvo = 1. Julkaisujen määrä analyysissä: 115849. Lähde: Scopus ja SciVal, tilanne: 9.5.2018

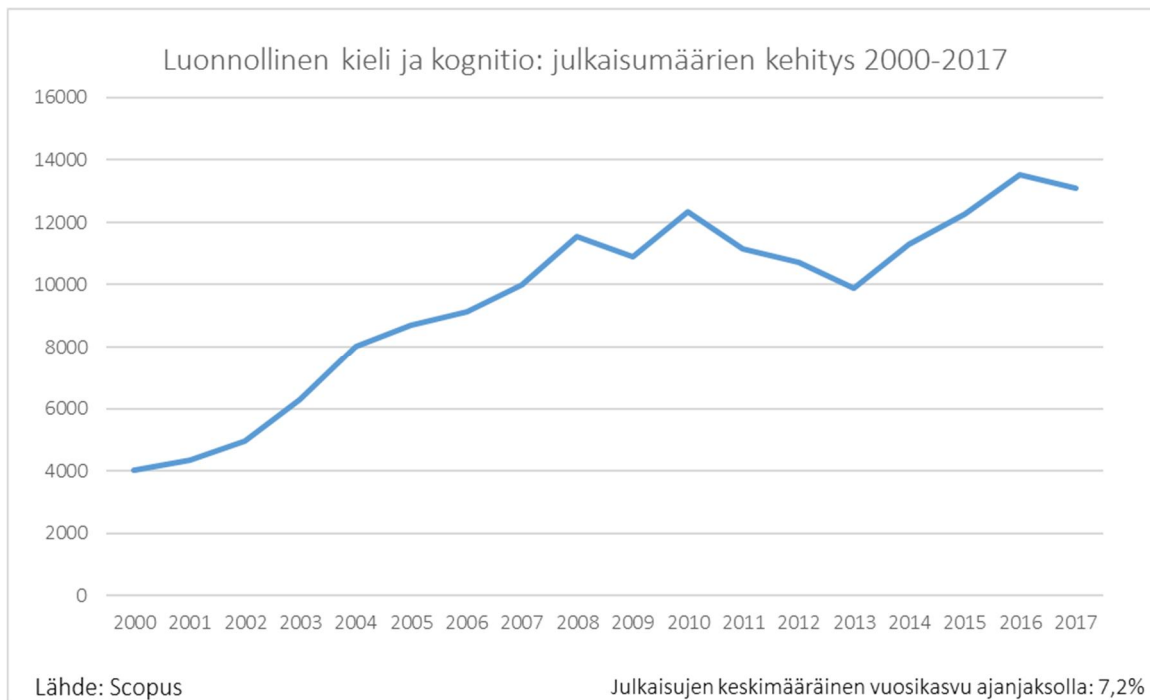
1.3.4 Luonnollinen kieli ja kognitio²⁵

Julkaisumäärät ja aiheet

Erilaiset kognitiiviset ja tietoperusteiset järjestelmät ja ratkaisut ovat alusta saakka olleet vahvasti osa tekoälyn tutkimusta. Koneavusteisella kielenkääntämisellä ja puheen tulkinalla on puolestaan ollut paljon yhtymäkohtia myös tekoälyn tutkimukseen.

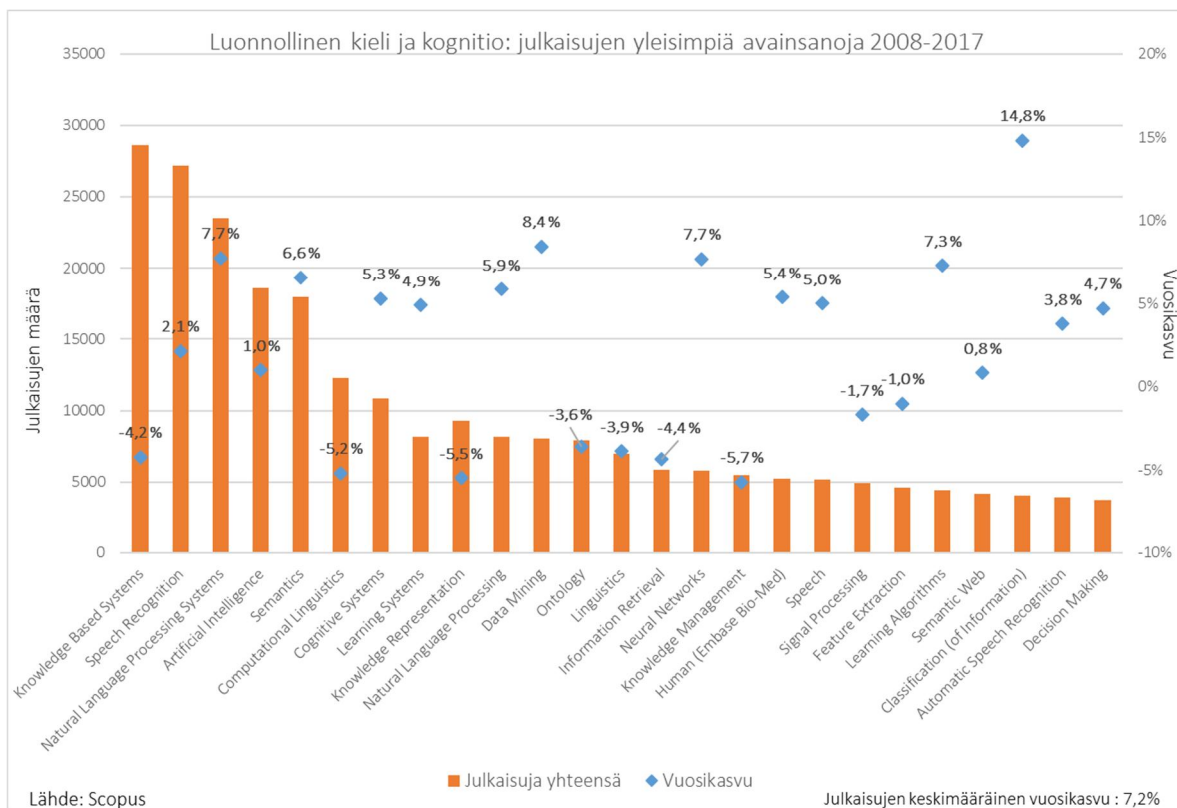
Luonnollista kieltä ja kognitiota käsittelevien julkaisujen määrä on tarkastelujaksolla kasvanut vuosittain keskimäärin yli 7%. Vuosina 2011-2013 tutkimusalan julkaisujen määrä kuitenkin laski selvästi. Selvää yksittäistä syytä tähän ei ole nähtävissä, mutta samaan aikaa myös eräissä muissa tutkimussuunnissa Uudelleen v. 2014 kasvaneen kiinnostuksen taustalla ovat todennäköisesti erilaiset koneoppimiseen tukeutuvat menetelmät, joilla myös puheen ja kirjoitetun tekstin tulkinassa on päästy nopeasti hyviin tuloksiin. Vuonna 2017 nousu näyttäisi ainakin tilapäisesti kuitenkin taittuneen.

²⁵ Tutkimussuunnan rajauksessa käytettiin seuraavia termejä: cognitive systems, heuristic algorithms, information retrieval, knowledge based systems, knowledge representation, natural language processing, speech recognition, sekä osittain myös: information retrieval, machine translation, cognition, reasoning, ontology, semantics



Kuva 34. Tieteellisten julkaisujen määrä 2000-2017. Lähde: Scopus

Isoista tutkimusaiheista nopeimmin v. 2008-2017 kasvoivat tiedon luokittamiseen, big dataan ja koneoppimiseen liittyvät teemat. Pienemmistä teemoista nopeimmin kasvoivat aiheet, joissa luonnollisen kielen ja kognition teemat yhdistyivät uusiin koneoppimisen alateknologioihin kuten *deep learning*, *convolutional neural network*, *recurrent neural networks*. Näiden lisäksi myös *sentiment analysis* oli nopeimmin kasvaneiden aiheiden joukossa (keskimääräinen kasvu 2008-2017 52%).



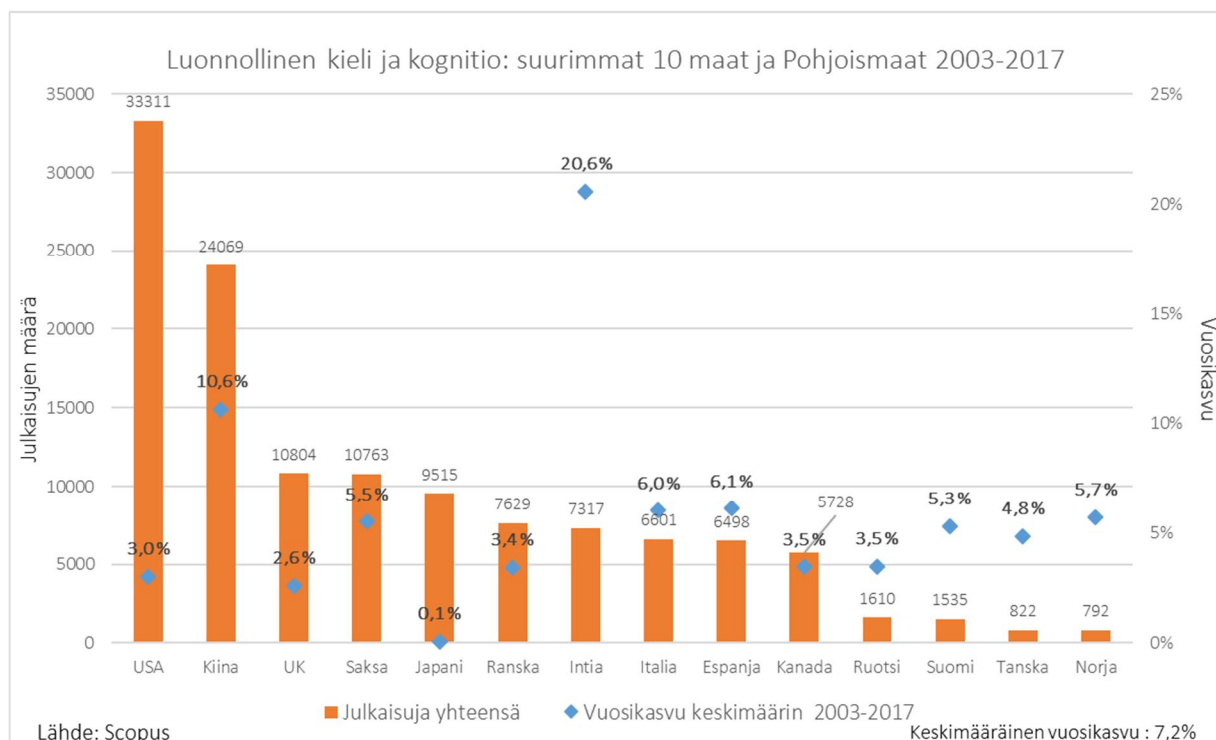
Kuva 35. Keskeisiä tutkimusaiheita: julkaisumäärät ja vuosikasvu. Kaikkien alan julkaisujen vuosikasvu ko. ajanjaksolla oli 7,2%. Osa sisällöltään yleisemmistä avainsanoista on jätetty kuvasta pois. Lähde: Scopus

Maat

Suurimmat julkaisijamaat ovat myös kommunikaation ja kognition tutkimussuunnalla Yhdysvallat ja Kiina, mutta erot suurimpien maiden välillä ovat pienemmät kuin esimerkiksi data-analyysissä tai tilannetietoisuuteen keskittyvissä julkaisuissa. Esimerkiksi kolmannen maan, Iso-Britannian, julkaisumäärä vastasi v. 2003-2017 45% Kiinan julkaisujen lukumäärästä, kun esimerkiksi tilannetietoisuudessa kolmanneksi aktiivisimman julkaisumaan, Saksan, julkaisujen määrä oli vain 30% toiseksi suurimman julkaisumaan, Kiinan, julkaisumäärästä.

Kognition ja kieliteknologioiden tutkimus on monissa Euroopan maissa vahvaa, vaikka kokonaisuudessaan puhutaankin selvästi pienemmästä tutkimussuunnasta kuin esimerkiksi data-analyysistä, koneoppimisesta puhumattakaan.

Intia kasvatti julkaisuvolyymiään tarkastelujaksolla selkeästi nopeimmin. Myös Kiinan julkaisujen kasvuvauhti ylitti koko tutkimussuunnan keskimääräisen kasvuvauhdin maailmassa.



Kuva 36. Luonnollinen kieli ja kognitio: suurimmat kymmenen maata ja Pohjoismaat, julkaisujen määrä ja julkaisumäärien keskimääräinen vuosikasvu 2003-2017. Lähde: Scopus

Tarkasteltaessa julkaisujen sisältöä maittain, huomio kiinnittyy laajojen avaintermien osalta

Julkaisuja yhteensä 2015-2018	8214	6846	3585	2717	2607	1878	1870	1865	1454	1246	391	384	218	184
Termi	USA	Kiina	Intia	Saksa	UK	Italia	Ranska	Japani	Espanja	Kanada	Ruotsi	Suomi	Tanska	Norja
Algorithms	5,0 %	6,3 %	4,2 %	3,0 %	3,3 %	3,2 %	3,7 %	3,2 %	3,6 %	5,3 %	4,6 %	3,1 %	2,3 %	2,7 %
Artificial Intelligence	17,9 %	16,5 %	17,6 %	18,1 %	20,9 %	23,9 %	19,8 %	14,6 %	22,8 %	21,2 %	17,1 %	11,2 %	15,1 %	16,3 %
Classification (of Information)	5,2 %	7,0 %	10,6 %	5,2 %	4,6 %	6,0 %	5,3 %	4,3 %	6,2 %	5,3 %	4,9 %	6,8 %	6,0 %	1,6 %
Cognitive Systems	16,1 %	8,5 %	6,2 %	11,5 %	15,0 %	10,5 %	7,0 %	6,2 %	8,9 %	18,9 %	15,9 %	15,9 %	7,8 %	12,0 %
Computational Linguistics	12,5 %	11,5 %	12,3 %	11,7 %	11,7 %	13,8 %	12,0 %	11,6 %	15,9 %	8,7 %	7,2 %	7,0 %	14,2 %	6,5 %
Data Mining	10,5 %	9,8 %	14,4 %	7,5 %	8,9 %	10,8 %	8,1 %	8,0 %	11,1 %	10,3 %	7,4 %	7,0 %	9,6 %	7,6 %
Decision Making	4,3 %	2,7 %	3,0 %	3,3 %	5,2 %	3,7 %	4,8 %	2,2 %	4,7 %	5,4 %	4,9 %	3,6 %	6,0 %	4,9 %
Deep Neural Networks	5,1 %	7,1 %	3,3 %	2,9 %	4,3 %	3,1 %	2,1 %	7,7 %	2,0 %	2,9 %	1,8 %	3,9 %	4,1 %	2,2 %
Human (Bio-Medical)	13,5 %	2,9 %	1,9 %	5,2 %	8,7 %	3,3 %	4,7 %	2,3 %	5,4 %	8,6 %	9,2 %	4,2 %	8,7 %	6,5 %
Information Retrieval	4,7 %	3,5 %	6,9 %	4,2 %	4,0 %	4,5 %	6,2 %	3,3 %	5,1 %	3,8 %	2,3 %	4,4 %	5,0 %	7,1 %
Knowledge Based Systems	15,4 %	20,0 %	13,2 %	26,8 %	19,3 %	26,3 %	26,6 %	25,9 %	21,3 %	21,0 %	28,9 %	18,0 %	20,2 %	41,8 %
Knowledge Representation	4,2 %	5,0 %	3,3 %	7,0 %	6,6 %	10,3 %	9,4 %	2,2 %	7,9 %	6,4 %	5,9 %	6,5 %	8,3 %	3,8 %
Learning Algorithms	6,3 %	5,8 %	7,1 %	3,7 %	4,5 %	5,2 %	3,6 %	3,4 %	5,0 %	5,8 %	3,8 %	2,1 %	3,2 %	2,2 %
Learning Systems	11,8 %	10,8 %	12,6 %	9,9 %	11,8 %	9,9 %	8,6 %	11,2 %	9,4 %	13,0 %	6,1 %	5,5 %	9,2 %	7,6 %
Linguistics	6,9 %	6,7 %	6,5 %	6,3 %	7,4 %	10,6 %	6,6 %	7,5 %	9,9 %	5,3 %	6,1 %	3,4 %	11,0 %	4,3 %
Machine Learning	5,9 %	3,0 %	4,8 %	3,3 %	4,7 %	3,0 %	2,9 %	3,1 %	5,0 %	5,0 %	4,6 %	3,4 %	3,2 %	3,3 %
Neural Networks	6,7 %	13,1 %	7,1 %	5,2 %	6,2 %	4,3 %	3,3 %	7,6 %	4,7 %	5,7 %	4,3 %	5,7 %	3,2 %	2,7 %
NLP Systems	23,7 %	25,8 %	29,7 %	22,4 %	20,8 %	27,6 %	20,2 %	21,2 %	28,5 %	19,8 %	16,1 %	11,5 %	25,2 %	16,8 %
Ontology	3,0 %	3,2 %	5,7 %	5,6 %	5,4 %	7,5 %	9,9 %	2,8 %	7,2 %	3,9 %	5,1 %	5,2 %	1,8 %	7,6 %
Semantic Web	2,2 %	2,5 %	3,5 %	6,5 %	3,5 %	7,7 %	7,3 %	2,6 %	4,6 %	2,9 %	4,3 %	3,6 %	3,2 %	6,5 %
Semantics	17,0 %	27,8 %	12,2 %	19,0 %	16,5 %	24,1 %	21,7 %	13,9 %	20,2 %	13,6 %	20,5 %	14,6 %	14,7 %	15,8 %
Social Networking (online)	3,7 %	3,3 %	6,0 %	2,5 %	3,6 %	5,1 %	3,1 %	3,3 %	5,1 %	4,4 %	1,8 %	3,1 %	5,0 %	3,8 %
Speech Recognition	21,9 %	23,0 %	37,1 %	18,8 %	20,4 %	11,5 %	19,9 %	31,9 %	16,9 %	18,6 %	11,8 %	35,7 %	23,4 %	9,8 %

Kuva 37. Suurimmat kymmenen maata ja pohjoismaat 2015-2018, julkaisujen painotukset. Yleisimmät sisältöavainsanat koko julkaisuaineistossa. Osa julkaisutermeistä jätetty pois termin päällekkäisyyden tai yleisluonteisuuden perusteella. Termien prosentuaalinen osuus on $\neq 100\%$ johtuen eri aiheiden esiintymisestä samoissa julkaisuissa. Vuoden 2018 tilanne: 24.4.2018. Lähde: Scopus

Tutkimuksen vaikuttavuus: maat

Normalisoidun (suhteutetun) viittausindeksin perusteella Singapore, Kanada ja Sveitsi tuottivat v. 2013-2017 tieteellisesti vaikuttavimpia luonnollista kieltä ja kognitiota käsitteleviä tekoälyn tutkimusjulkaisuja. Top 10% -huippujulkaisuissa ilmestyneitä artikkeleita tuottivat eniten Hong Kong, Israel, Yhdysvallat ja Singapore, joilla kaikilla yli 20% julkaisuista kuului kyseisiin huippujulkaisuihin. Yritys-yliopisto –yhteistyötä oli tutkimusalalla v. 2013-2017 suhteellisesti eniten Israelissa ja Irlannissa, molemmissa yli 10% julkaisuista tuotettiin yhteistyössä yritysten kanssa. Kansainvälistä yhteistyötä oli ajanjaksolla suhteellisesti eniten Saudi-Arabiassa ja Hong Kongissa. Euroopasta eniten kansainvälistä yhteistyötä oli Sveitsissä ja Irlannissa. Myös Suomessa sekä yritys-yliopisto että kansainvälistä yhteistyötä oli selvästi keskimääräistä enemmän.

Maa	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankiingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys -yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisutietojen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Singapore	2,52	939	20,7	1,62	4	1,34
Canada	2,29	1904	16	1,41	4,3	1,32
Switzerland	2,21	661	19,7	1,67	7,6	1,26
United States	2,15	12085	22,2	0,88	9,7	1,17
Denmark	2,14	332	15	1,6	6,6	1,46
Hong Kong	2,06	568	26,5	1,84	6,9	1,36
United Kingdom	1,97	3888	18,2	1,54	6,1	1,42
Finland	1,91	599	12,4	1,61	6,3	1,4
Israel	1,81	385	22,9	1,42	11,2	1,2
Italy	1,77	2706	12,8	1,11	3,1	1,75
Netherlands	1,76	1066	18,3	1,47	6,8	1,39
Germany	1,75	4179	11,7	1,11	5,1	1,14
Austria	1,75	781	9	1,41	4,7	1,34
Ireland	1,58	438	6,7	1,61	11,2	1,31
Australia	1,55	1662	18,6	1,52	6,1	1,5
France	1,53	3015	11,2	1,36	4,5	1,01
Norway	1,48	310	13,3	1,37	3,5	1,42
Saudi Arabia	1,46	477	15,5	1,85	2,7	1,82
Greece	1,43	718	12,3	1,19	2,2	1,62
Spain	1,36	2225	15,9	1,2	3	1,45
Poland	1,36	1146	4,8	0,65	0,9	1,19
Belgium	1,31	566	17,3	1,57	5,1	1,2
Sweden	1,27	621	11,8	1,48	7,7	1,44
Russian Federation	1,24	1099	2,1	0,57	1,2	1,79
Czech Republic	1,24	688	4,9	0,85	3,5	1,39
Mexico	1,2	408	8,8	1,27	0,7	1,57
Portugal	1,16	700	9,6	1,22	5,1	1,41
Turkey	1,11	684	15,7	0,74	6	1,22
China	1,09	9592	9,2	0,69	3,6	0,96

Kuva 38. Luonnollinen kieli ja kognitio: maat järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2013-2017. Viittausindeksin raja-arvo kuvassa: 1,09. Alle 300 julkaisua tuottaneet maat eivät ole mukana taulukossa. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen yhteistyön ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa. (Tiedot Scopusesta: katselukerrat antavat tietoa julkaisun herättämästä kiinnostuksesta täydentämällä viittausmäärien välittämää informaatiota erityisesti uusimman aineiston osalta.) Normalisoiduissa mittareissa maailman keskiarvo = 1. Lähde: Scopus ja SciVal. Analysoituja julkaisuja yhteensä 60240. Tilanne: 4.6.2018

Tutkimuksen vaikuttavuus: tutkimusorganisaatiot

Eniten julkaisuja v. 2013-2017 tuottivat Ranskan kansallinen tutkimuskeskus CNRS (1082), Kiinan tiedeakatemia CAS (861) ja Tsinghua-yliopisto Kiinassa (648). Mikäli mittarina pidetään normalisoitua viittausten määrää, niin Stanfordin yliopisto oli kärjessä, seuraavina Toronton ja Berkeleyn yliopistot. Yrityksistä Google ja Microsoft olivat myös vahvasti esillä. Huippulehdissä ilmestyneitä artikkeleita tuottivat Harvard, Pennsylvanian yliopisto, Lontoon UCL ja Stanford. Kansainvälistä yhteistyötä oli suhteellisesti eniten Sydneyn teknillisellä yliopistolla ja Hong Kongin kiinalaisella yliopistolla. Yritysyhteistyötä oli tutkimuslaitoksista eniten Australian CSIRO:lla, Columbian yliopistolla New Yorkissa ja Marylandin yliopistolla. Eurooppalaiset yliopistot pärjäsivät vertailussa myös hyvin.

Tutkimuslaitos / Yliopisto	Normalisoitu viittausedeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankiingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys - yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisutietojen katseluideksi (Scopus-katselukerrat), keskiarvo = 1
Stanford University	7,5	280	36,7	0,87	12,1	1,35
University of Toronto	6,92	190	22,4	1,35	12,6	1,94
University of California at Berkeley	6,09	176	35,7	1,07	9,1	1,24
Google Inc.	5,07	283	19	0,88	55,8	1,28
Microsoft USA	4,86	466	25,4	1,16	77,5	1,28
Johns Hopkins University	3,88	224	24	1,03	15,2	1,02
University of Cambridge	3,17	267	22,6	1,31	12,4	1,18
Carnegie Mellon University	3,15	580	17,5	1,04	13,1	1,2
Nanyang Technological University	3,11	390	20,7	1,53	3,8	1,6
Imperial College London	3,09	209	13,7	1,73	7,2	1,51
University of Oxford	3,05	202	26,2	1,15	5,9	1,1
University of Edinburgh	2,98	288	12,3	1,7	11,1	1,18
Microsoft Research Asia	2,95	232	11,1	0,98	54,7	1,08
Technical University of Munich	2,94	214	10	1,42	13,1	1,54
Chinese University of Hong Kong	2,7	154 -		1,75	7,1	1,15
MIT	2,64	333	30,2	1	10,5	1,37
The Ohio State University	2,59	169	17,8	0,7	7,7	1,13
University of Rome La Sapienza	2,52	186	14,8	1,15	3,8	1,57
IBM Research	2,5	179 -		1,58	50,8	1,23
Georgia Institute of Technology	2,48	274	15,8	1,22	13,1	1,53
University of Michigan	2,45	184	30,1	0,65	9,2	1,39
University of Washington	2,42	281	24,5	0,78	14,9	1,15
IBM	2,39	416	9,5	1,09	53,4	1,19
University of Sheffield	2,38	227	14,2	1,31	5,3	1,81
University of Trento	2,37	174 -		1,49	5,7	1,2
INRIA France	2,35	288	13,3	1,16	6,9	0,93
Technische Universität Dresden	2,34	148 -		1,24	1,4	0,86

Kuva 39. Luonnollinen kieli ja kognitio: merkittävimmät yliopistot, tutkimuslaitokset ja yritykset järjestettynä normalisoidun viittausedeksin mukaan, julkaisuvuosi 2013-2017. Viittausedeksin raja-arvo kuvassa: 2,34. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa (katselutiedot Scopusesta). Normalisoiduissa mittareissa maailman keskiarvo = 1. Analysoituja julkaisuja yhteensä 60240. Lähde: Scopus ja SciVal, tilanne: 4.6.2018

1.3.5 Vuorovaikutus ihmisen kanssa²⁶

Julkaisumäärät ja aiheet

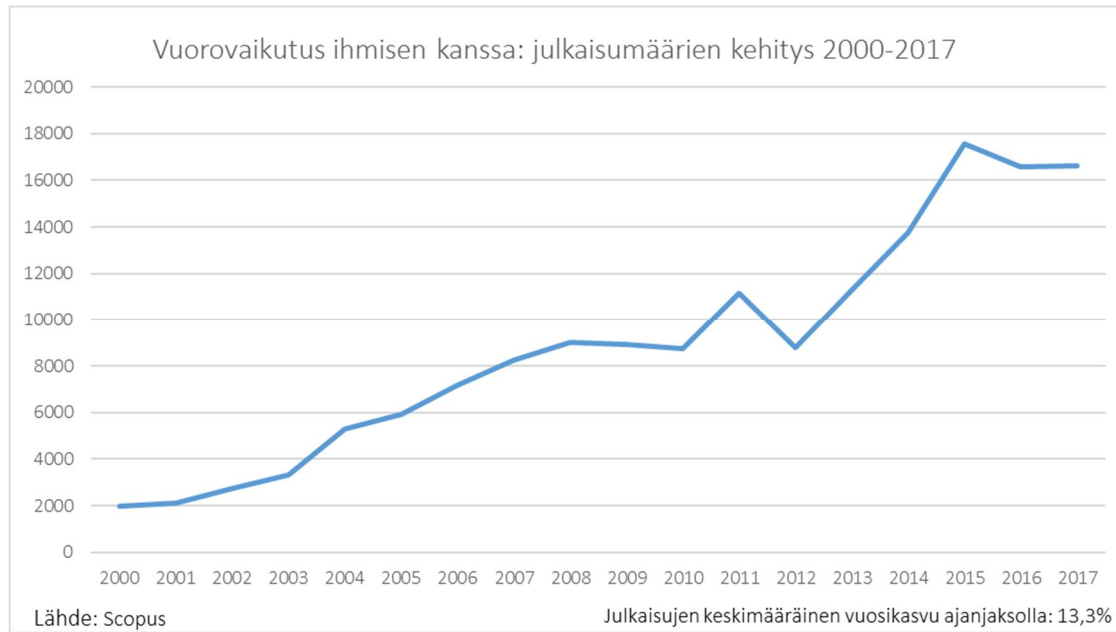
Vaikka tekoälyn kehittämisen eräänä tavoitteena on sovellusten mahdollisimman pitkälle viety autonomia, toimivat tekoälyn nykyiset sovellukset hyvin harvoin täysin autonomisesti. Päinvastoin: tulevaisuuden työelämän eräänä keskeisenä haasteena tulee olemaan robottien ja muiden tekoälyyn nojaavien ratkaisujen toiminta ihmisen rinnalla, työntekijän työparina.

Vuorovaikutus ihmisen kanssa voi tapahtua monilla tavoin ja eri tasoisena. Tässä selvityksessä lähdetään vuorovaikutuksen hyvin avoimesta tulkinnasta: kyse voi olla chatbot-tyyppisestä asiakaspalvelusta tai

²⁶ Tutkimussuunnan rajoituksessa käytettiin seuraavia termejä: affective computing, chatbots, computer assisted diagnosis, conversational agents, emotion recognition, game ai, game artificial intelligence, human computer interaction, human-robot interaction, interactive agent, machine emotional intelligence, recommendation algorithms, recommendation systems, recommender systems, sekä osittain myös: virtual reality

käyttäjilleen elämyksiä tuottavista musiikkialgoritmeista. Perinteisen robotin ja ihmisen vuorovaikutuksen tutkimuksen rinnalle ovat nousseet esimerkiksi käyttöliittymien pelillistäminen ja erilaiset suosittelevat järjestelmät, joilla vahvistetaan käyttäjien sitoutumista erilaisiin palveluihin. Ratkaisujen taustalta löytyy hyvin usein tekoälyyn tai koneoppimiseen liittyvää tutkimusta.

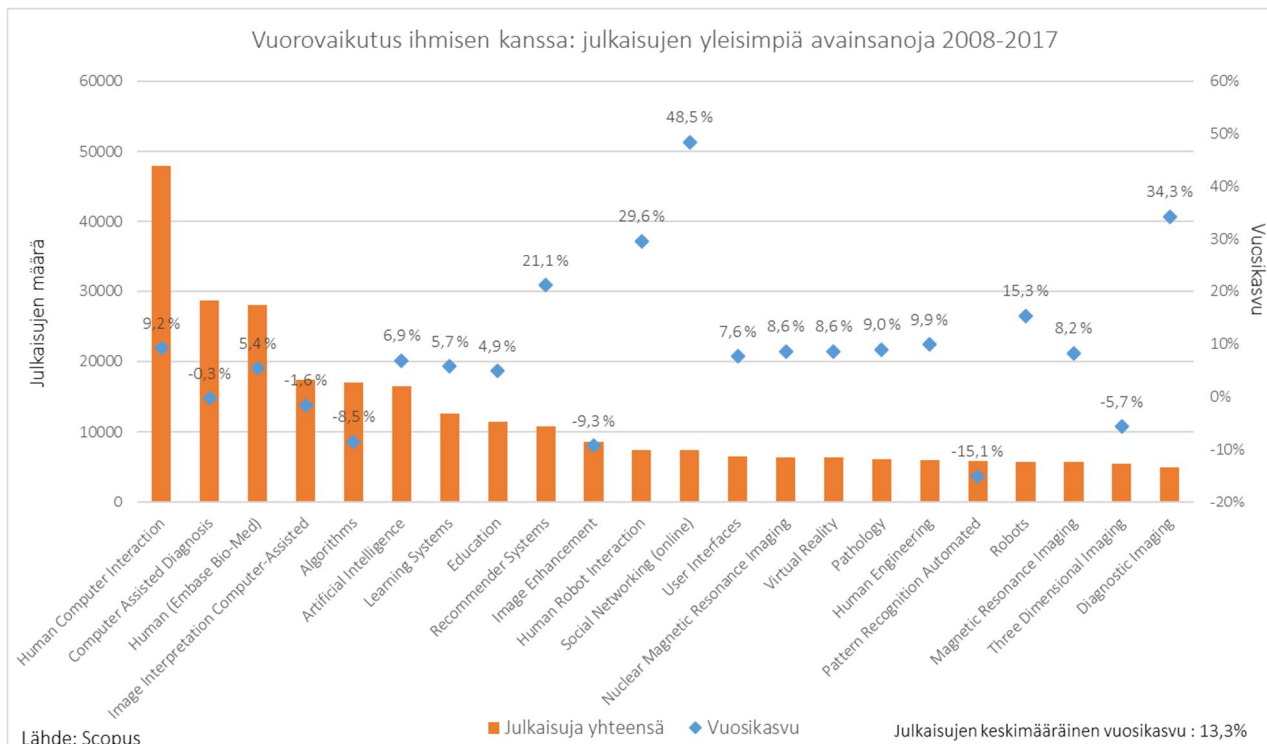
Vuorovaikutusta käsittelevien julkaisujen määrä on kasvanut koko 2000-luvun suhteellisen nopeasti. Erityisen vahvaa kasvu oli vuosina 2013-2015. Jaksolle on kuitenkin osunut myös tekoälylle tyypillisiä taantumajaksoja. Viimeisten kahden vuoden aikana tutkimusalan kasvu näyttäisi taittuneen, mutta vakiintuneen kuitenkin uudelle korkeammalle tasolle.



Kuva 40. Tieteellisten julkaisujen määrä 2000-2017. Lähde: Scopus

Tarkasteltaessa julkaisujen painotuksia, niin suurimmat aiheet liittyvät luonnollisesti tietokoneen (ja robotin) ja ihmisen vuorovaikutukseen, lisäksi erityisesti terveydenhoidon diagnostiikkaan liittyvät sovellukset ovat vahvasti edustettuna sekä volyymiltaan että julkaisumäärän kasvulla mitattuna. Nopeimmin suurista aiheista kasvoi ajanjaksolla sosiaaliseen mediaan liittyvä tutkimus ja siihen liittyen erilaisten suosittelevat järjestelmien ja -ratkaisujen tutkimus (kasvu: 21%).

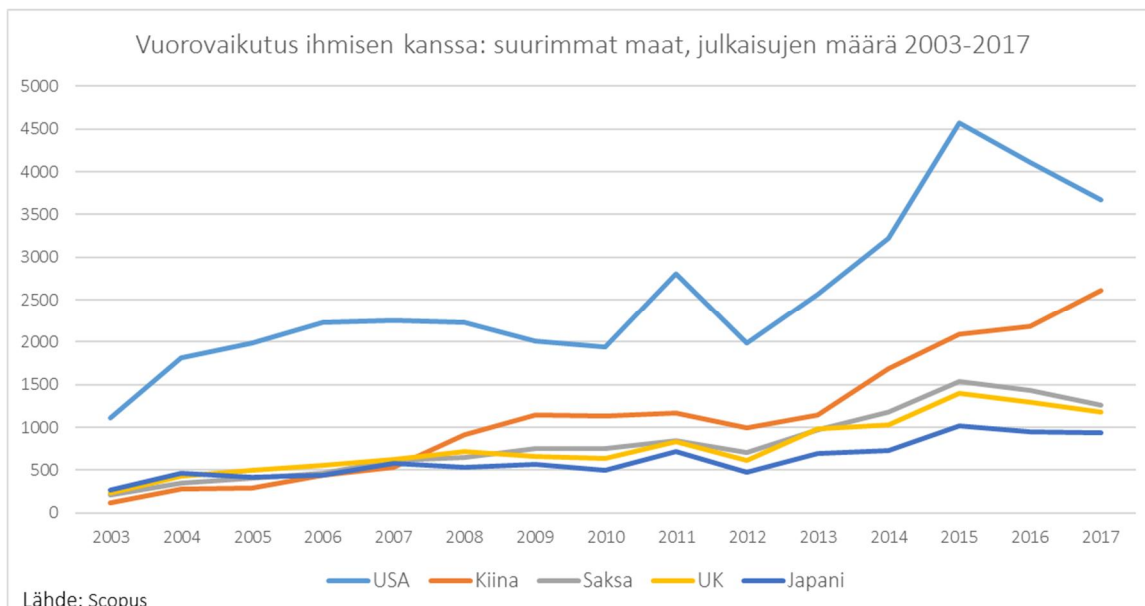
Pienemmistä aiheista voimakkaassa kasvussa olivat erityisesti neuroverkkoja ja syväoppimista hyödyntävä tutkimus (convolutional neural network, kasvu: 96%; deep learning, kasvu: 74%; deep neural networks, kasvu: 70%) sekä mieliala-/mielipideanalyysit (sentiment analysis, kasvu 56%), joista jälkimmäinen liittyi läheisesti (66% julkaisuista) sosiaalisen median (social networking) tutkimukseen. Myös ”tunteiden tunnistaminen” (emotion recognition) oli yksi suhteellisen suurista ja nopeasti kasvavista tutkimuskohteista (3400 julkaisua v. 2008-2017, kasvu 14 %).



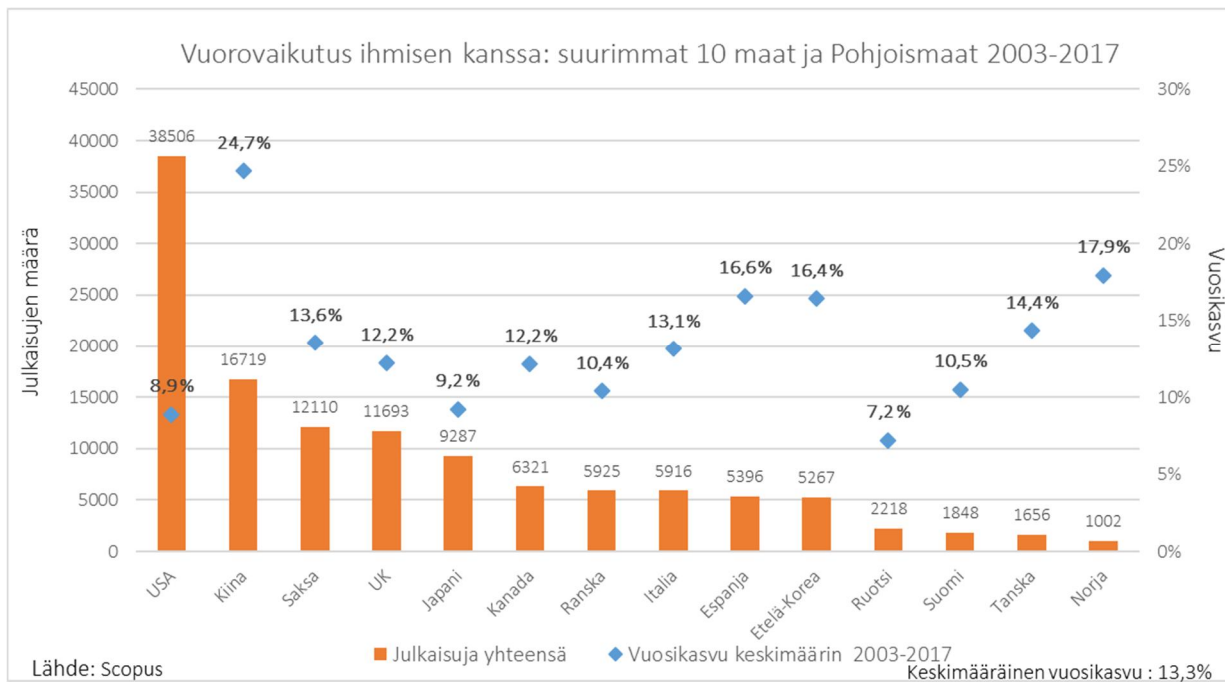
Kuva 41. Keskeisiä tutkimusaiheita: julkaisumäärät ja vuosikasvu. Kaikkien alan julkaisujen vuosikasvu ko. ajanjaksolla oli 13,3%. Osa sisällöltään yleisemmistä avainsanoista on jätetty kuvasta pois. Lähde: Scopus

Maat

Yhdysvallat oli selvästi suurin julkaisujen tuottaja 2003-2017. Kiina lisäsi julkaisumääriään nopeimmin eikä julkaisujen määrän kasvu näytä taittuneen 2016-2017 niin kuin muilla suurimmilla mailla. Pohjoismaista Norja kasvatti julkaisumääriään nopeimmin. Suurista maista kuvan ulkopuolelta voi mainita lisäksi Intian, jonka julkaisutuotanto kasvoi ajanjaksolla kaikkein nopeimmin, keskimäärin 33,2% vuodessa.



Kuva 42. Vuorovaikutus ihmisen kanssa: suurimmat maat, julkaisumäärät 2003-2017. Lähde: Scopus



Kuva 43. Vuorovaikutus ihmisen kanssa: suurimmat kymmenen maata ja Pohjoismaat, julkaisujen määrä ja julkaisumäärien keskimääräinen vuosikasvu 2003-2017. Lähde: Scopus

Maiden tutkimuspainotuksissa oli havaittavia eroja 2015-2018. Intian (ja osittain Kiinan) profiili poikkeaa muista suurista maista eniten. Muissa maissa painotus oli vahvasti lääketieteen ja terveydenhoidon ratkaisujen kehittämisessä, kun taas Intiassa algoritmit, oppivat järjestelmät ja muut tekoälyn teknologiat olivat vahvasti esillä.

Yhteensä 2015-2018	12800	7332	4371	4032	3008	2692	2196	2088	1907	1904	790	634	572	424
Termi	USA	Kiina	Saksa	UK	Japani	Intia	Italia	Kanada	Ranska	Australia	Ruotsi	Tanska	Suomi	Norja
Algorithms	11,8 %	13,8 %	11,1 %	8,8 %	6,5 %	6,9 %	8,0 %	11,1 %	11,4 %	7,8 %	10,5 %	8,8 %	6,5 %	6,1 %
Artificial Intelligence	14,7 %	15,6 %	9,0 %	12,7 %	13,0 %	23,0 %	13,1 %	12,2 %	14,1 %	13,3 %	8,5 %	8,0 %	8,2 %	8,3 %
Behavioral Research	4,9 %	4,3 %	4,6 %	4,8 %	5,7 %	3,8 %	4,6 %	4,2 %	3,6 %	4,3 %	3,7 %	2,5 %	7,3 %	3,1 %
Computer Assisted Diagnosis	28,6 %	18,6 %	30,1 %	21,4 %	23,1 %	9,5 %	23,8 %	26,4 %	27,1 %	17,4 %	28,5 %	22,6 %	13,5 %	19,3 %
Data Mining	3,7 %	5,4 %	2,3 %	2,8 %	2,8 %	12,6 %	3,2 %	3,8 %	3,5 %	5,3 %	2,0 %		2,1 %	1,9 %
Diagnostic Imaging	11,2 %	8,3 %	11,4 %	7,7 %	11,1 %	2,9 %	10,2 %	10,8 %	10,4 %	6,6 %	10,1 %	9,8 %	5,2 %	6,8 %
Education	10,3 %	10,4 %	4,5 %	8,4 %	10,3 %	10,7 %	5,4 %	7,7 %	5,6 %	10,2 %	5,7 %	6,3 %	7,9 %	11,8 %
Human (Bio-Med)	31,6 %	18,7 %	31,9 %	26,5 %	23,7 %	10,5 %	26,4 %	30,0 %	28,8 %	23,8 %	30,6 %	24,6 %	15,2 %	21,9 %
Human Engineering	10,7 %		7,1 %	12,2 %	4,0 %	1,5 %	2,3 %	12,0 %	5,3 %	5,8 %	9,9 %	13,2 %	11,5 %	5,0 %
Human Robot Interaction	8,7 %	4,5 %	8,6 %	7,2 %	14,3 %	2,2 %	11,2 %	5,5 %	9,9 %	4,3 %	12,0 %	7,4 %	3,8 %	3,5 %
Image Interpretation	18,4 %	11,8 %	19,3 %	14,4 %	12,7 %	3,7 %	15,1 %	18,0 %	18,2 %	10,3 %	15,4 %	13,1 %	9,3 %	12,7 %
Learning Systems	10,4 %	13,7 %	7,7 %	10,5 %	11,4 %	19,8 %	8,0 %	10,0 %	10,0 %	11,2 %	5,6 %	5,2 %	9,1 %	8,3 %
Machine Learning	4,9 %	4,3 %	2,9 %	4,3 %	2,7 %	7,5 %	3,7 %	4,1 %	3,8 %	3,6 %	2,8 %	2,1 %	3,7 %	2,4 %
NMR Imaging	9,6 %	4,9 %	10,3 %	6,9 %	6,2 %	1,8 %	5,6 %	9,7 %	8,0 %	4,8 %	8,1 %	5,2 %	4,5 %	5,9 %
Pathology	7,9 %	5,2 %	7,8 %	6,7 %	6,9 %	2,4 %	7,7 %	7,2 %	8,1 %	4,8 %	6,7 %	4,1 %	5,6 %	7,1 %
Recommender Systems	6,2 %	17,2 %	6,5 %	3,8 %	5,4 %	18,8 %	14,2 %	6,7 %	8,8 %	10,2 %	3,2 %	2,8 %	5,4 %	10,4 %
Robots	8,0 %	4,7 %	8,6 %	7,1 %	14,5 %	2,0 %	11,4 %	5,1 %	9,8 %	4,2 %	12,4 %	7,7 %	4,2 %	3,5 %
Social Networking (online)	7,0 %	10,3 %	3,8 %	5,6 %	4,4 %	16,3 %	9,2 %	6,5 %	5,0 %	8,0 %	4,1 %	3,5 %	5,6 %	8,5 %
User Interfaces	3,8 %	2,3 %	6,6 %	5,0 %	4,1 %	3,5 %	6,0 %	4,5 %	4,6 %	4,6 %	6,1 %	5,7 %	8,9 %	8,3 %
Virtual Reality	5,1 %	5,5 %	5,4 %	4,8 %	5,2 %	3,8 %	5,6 %	5,0 %	6,6 %	5,4 %	2,9 %	2,5 %	5,4 %	5,4 %

Kuva 44. Suurimmat kymmenen maata ja pohjoismaat 2015-2018, julkaisujen painotukset. Yleisimmät sisältöavainsanat koko julkaisuaineistossa (yleisin kuivailutermin "human-computer interaction" ei mukana vertailussa). Osa julkaisutermeistä jätetty pois termin päällekkäisyyden tai yleisluonteisuuden perusteella. Termien prosentuaalinen osuus on $\neq$ 100% johtuen eri aiheiden esiintymisestä samoissa julkaisuissa. Vuoden 2018 tilanne: 24.4.2018. Lähde: Scopus

1.3.6 Digitaidot työelämässä, ongelmanratkaisu ja laskennallinen luovuus²⁷

Julkaisumäärät ja aiheet

²⁷ Tutkimussuunnan rajoituksessa käytettiin seuraavia termejä: algorithmic composition, artificial emotional intelligence, computational creativity, creative systems, decision support systems, decision trees, digital

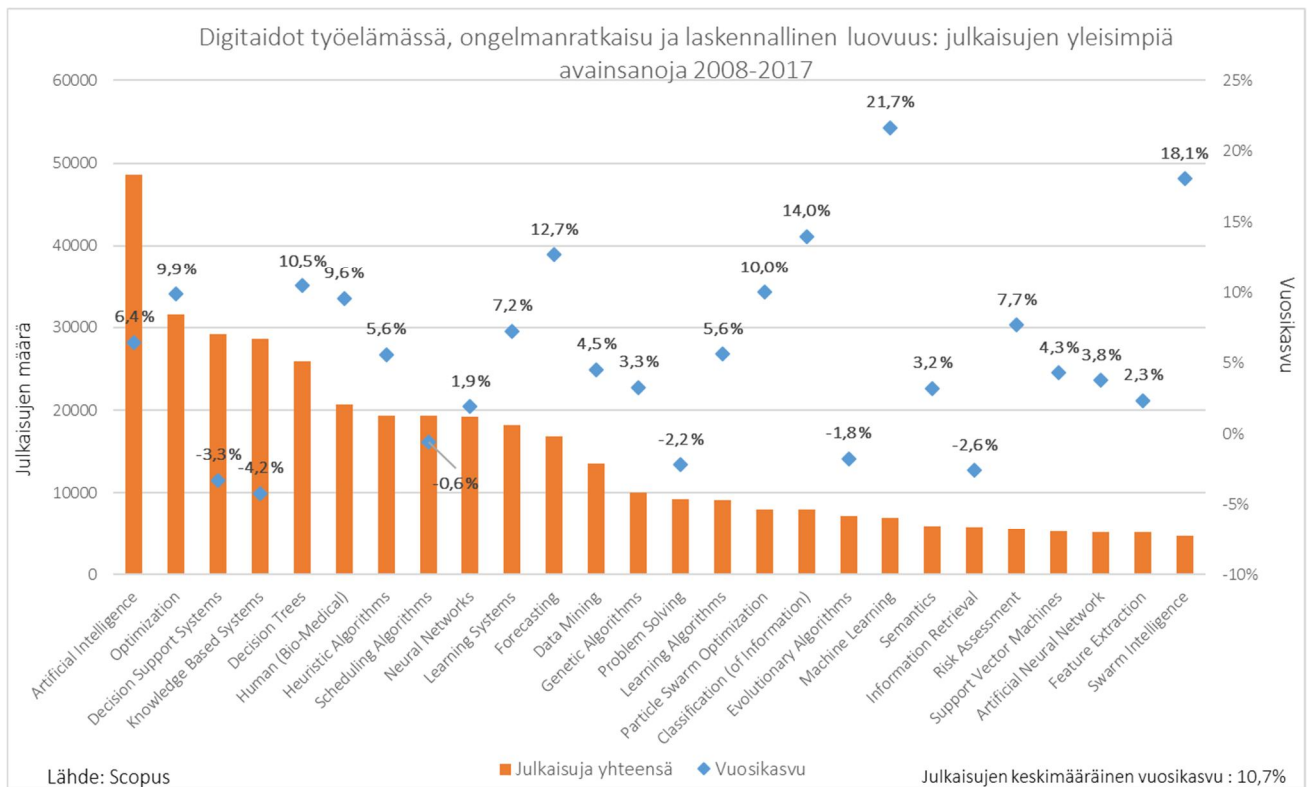
Ongelmanratkaisulla on pitkään ollut keskeinen sija tekoälyn kehittämisessä. Laskennallinen luovuus ja ”luovat järjestelmät” ovat sen sijaan hyvin uusia teemoja tekoälyn tutkimuskirjallisuudessa. Jälkimmäinen esiintyi julkaisuissa ensimmäistä kertaa v. 2007 eikä kumpaakaan esiinny julkaisuissa vieläkään muutamaa kymmentä julkaisua enempää vuosittain. Digitaidot ovat sen sijaan olleet jo pidempään esillä erityisesti digitaalinen lukutaito –käsitteen myötä.



Kuva 45. Tieteellisten julkaisujen määrä 2000-2017. Lähde: Scopus

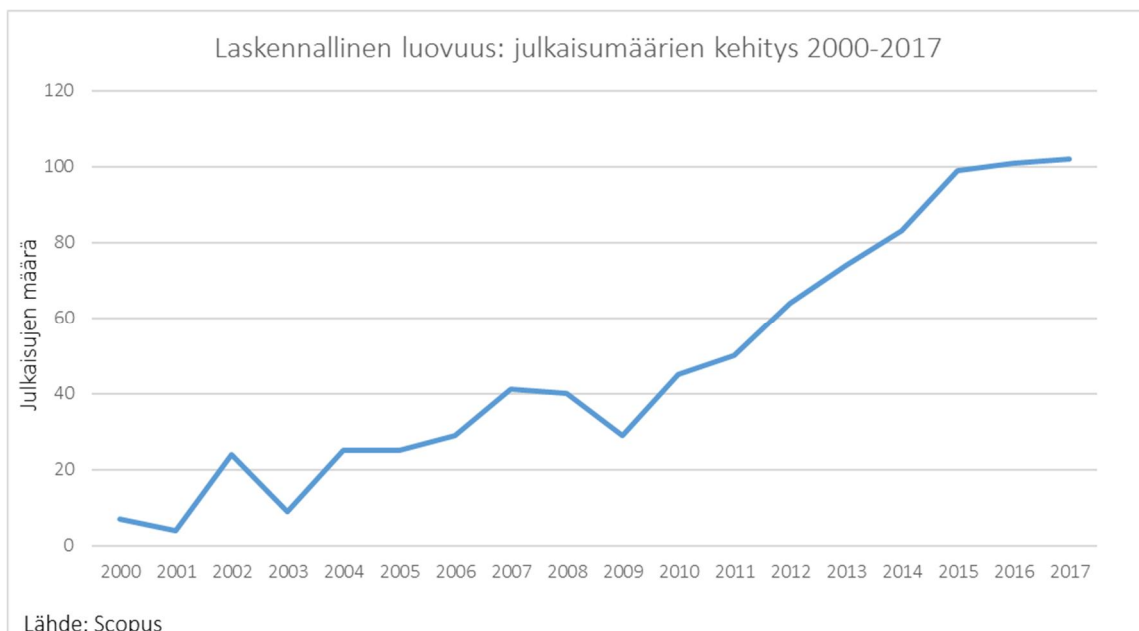
Tutkimussuunnan keskeisiä aiheita ovat v. 2008-2017 olleet optimointiin, päätöksentekoon ja tietämykseen liittyvät teemat. Näiden lisäksi erilaiset algoritmratkaisut ovat näkyvästi esillä tutkimuskirjallisuudessa. Nopeimmin kasvavat aiheet ovat pitkälti samoja kuin muillakin tekoälyä hyödyntävillä tutkimusaloilla. Toisaalta koneoppimisen ja tiedon luokittamisen lisäksi parviälyyn liittyvä tutkimus on ollut vahvasti nousussa. Pienemmissä aiheissa erityisesti syväoppiminen, CNN- ja RNN-neuroverkot sekä big data ovat voimakkaassa kasvussa.

competences, digital literacy, digital skills, digital storytelling, heuristic algorithms, scheduling algorithms, swarm intelligence, sekä osittain myös: decision making, forecasting, information retrieval, optimization, problem solving, reasoning



Kuva 46. Keskeisiä tutkimusaiheita: julkaisumäärät ja vuosikasvu. Kaikkien alan julkaisujen vuosikasvu ko. ajanjaksolla oli 13,3%. Osa sisällöltään yleisistä avainsanoista on jätetty kuvasta pois. Lähde: Scopus

Laskennallinen luovuus ja ylipäättään tekoälyn hyödyntäminen luovissa prosesseissa on edelleen suhteellinen pieni tutkimusala, vaikkakin jatkuvasti kasvussa. Tällä hetkellä julkaisuja ilmestyy n. 100 vuodessa.

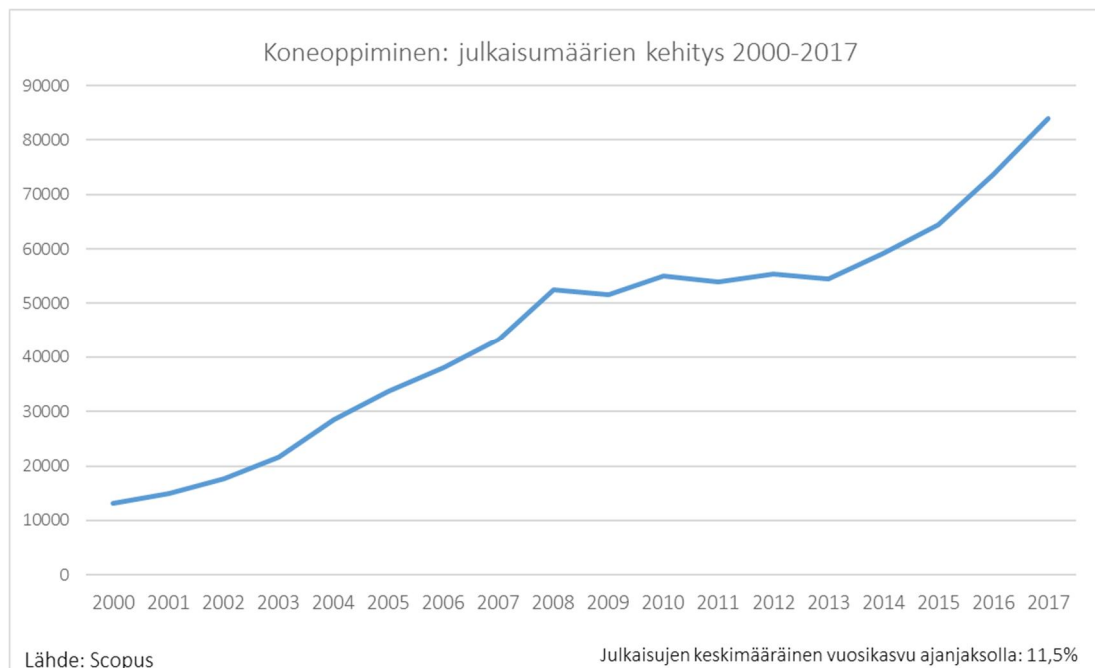


Kuva 47. Laskennallinen luovuus, julkaisujen määrä 2000-2017. Rajaavina termeinä: algorithmic composition, computational creativity, sekä tekoälyä käsittelevien aiheiden yhteydessä: creative process, creativity, creative systems. Lähde: Scopus

1.3.7 Koneoppiminen²⁸

Julkaisumäärät ja aiheet

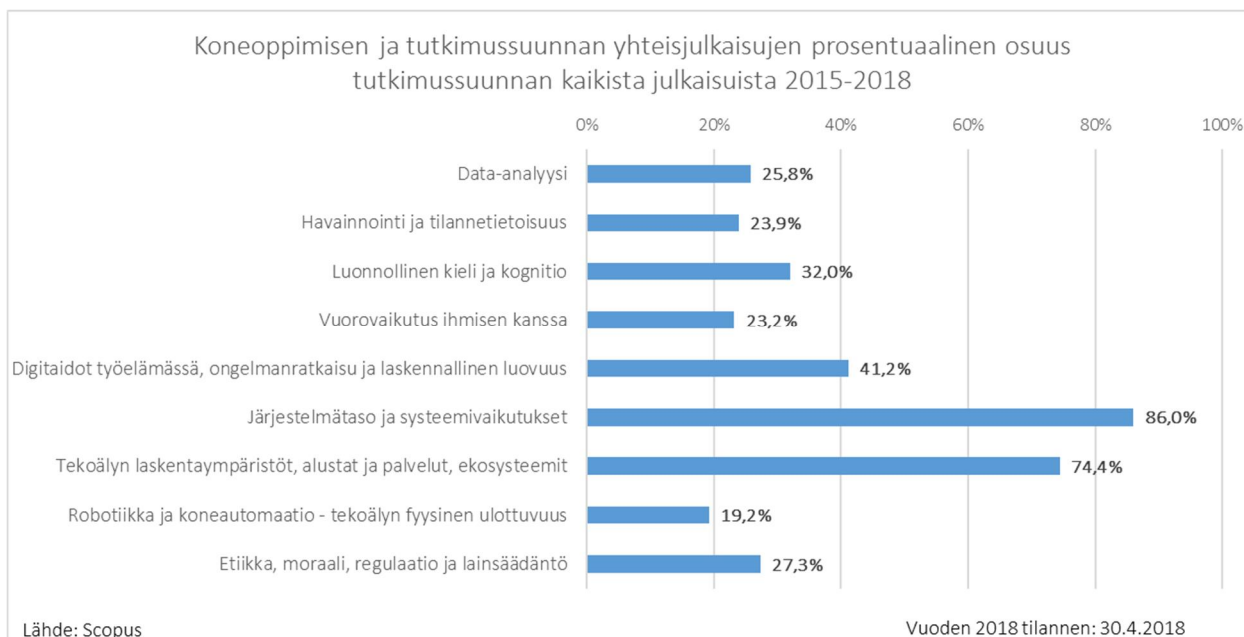
Koneoppiminen muodostaa nykyisin tekoälyn selvästi suurimman alateknologian ja samalla tutkimussuunnan. Koneoppimisen julkaisuja ilmestyi v. 2015-2017 n. 40% enemmän kuin toiseksi suurimman tutkimussuunnan, havainnointi ja tilannetietoisuus, julkaisuja. Toisaalta samoin kuin muillakin tekoälyn tutkimusalueilla, myös koneoppimisen tutkimuksessa on ollut sekä nopean kasvun että hitaan kehityksen vaiheita. Viime vuosina kasvu on ollut nopeaa, mutta v. 2009-2013 julkaisumäärät pysyivät käytännössä ennallaan. Tässä suhteessa koneoppiminen ei juuri poikennut muista tässä selvityksessä tarkastelluista tutkimussuunnista.



Kuva 48. Tieteellisten julkaisujen määrä 2000-2017. Lähde: Scopus

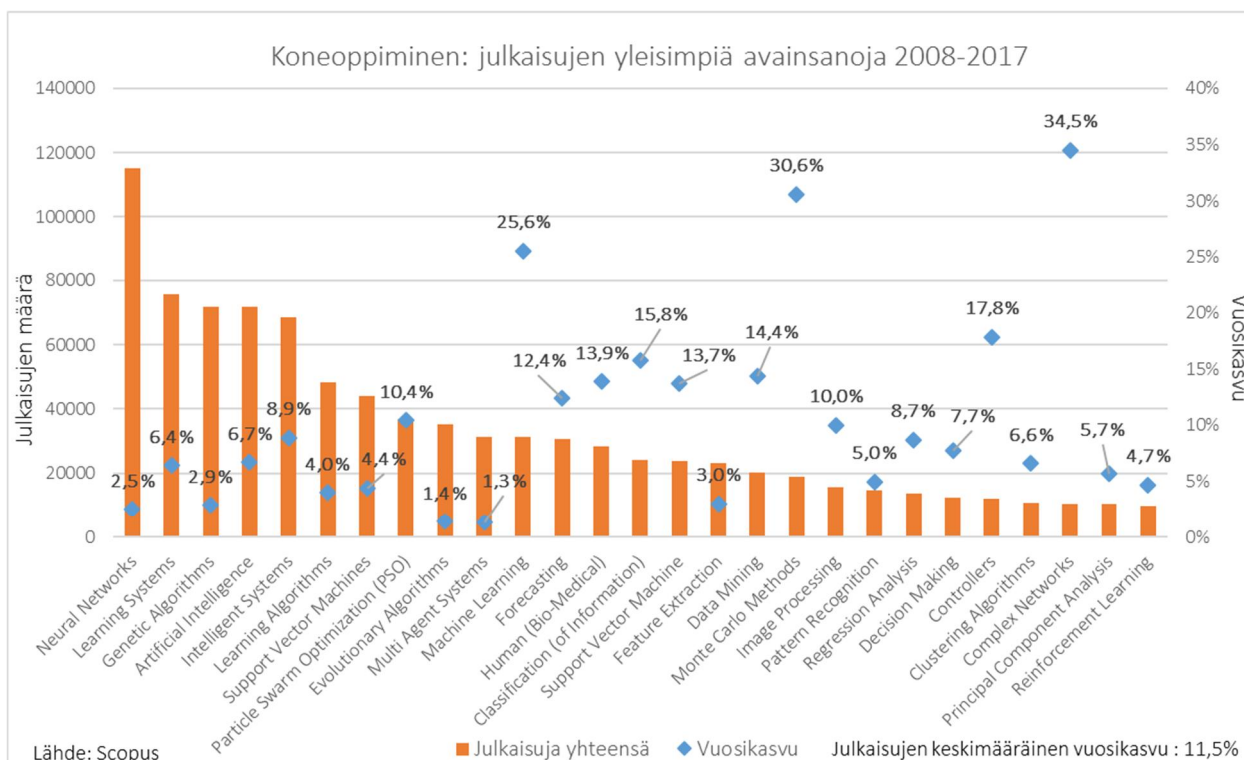
Koneoppimiseen, oppiviin algoritmeihin ja muihin oppiviin järjestelmiin liittyvän tutkimuksen osuus tekoälyn tutkimuksessa on nykyisin merkittävä. Koneoppimista käsittelevien julkaisujen osuus muiden tutkimusalan julkaisuista kuitenkin vaihtelee tutkimussuunnasta riippuen. Suurimmillaan se on järjestelmätasoa ja systeemivaikutuksia käsittelevissä julkaisuissa, joista peräti 86% ovat yhteisiä koneoppimisen julkaisujen kanssa. Pienin osuus on robotiikan ja koneautomaation julkaisuissa, joissa yhteisjulkaisujen osuus on alle viidennes.

²⁸ Tutkimussuunnan rajauksessa käytettiin seuraavia termejä: ant colony optimization, artificial bee colony, artificial neural networks, deep learning, deep neural networks, evolution algorithms, evolutionary algorithms, genetic algorithms, intelligent agents, intelligent control, intelligent systems, learning algorithms, learning systems, machine learning, multi agent systems, neural networks, particle swarm optimization, reinforced learning, reinforcement learning, statistical learning, supervised learning, support vector machines, unsupervised learning sekä seuraavia termejä yhdessä tekoälyaiheiden: bayesian networks, fuzzy systems, iterative methods, principal component analysis, random forests, regression analysis, soft computing, stochastic systems



Kuva 49. Koneoppimisen ja tutkimussuunnan yhteisjulkaisujen prosentuaalinen osuus tutkimussuunnan kaikista julkaisuista 2015-2018.

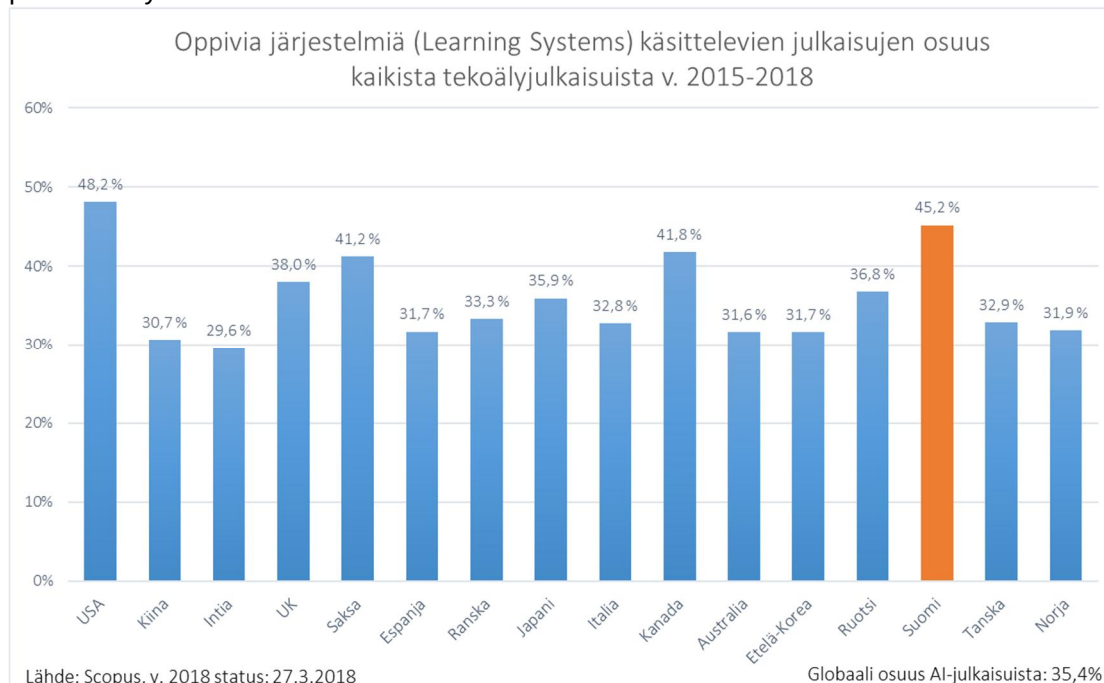
Yksittäisistä aiheista nopeimmin kasvoi syviin neuroverkkoihin (ml. "convolutional neural networks", "deep learning") ja isoihin datamääriin liittyvät soveltava koneoppimisen tutkimus. Neuroverkkoja käsittelevät julkaisut olivat selkeästi suurin aihe v. 2008-2017.



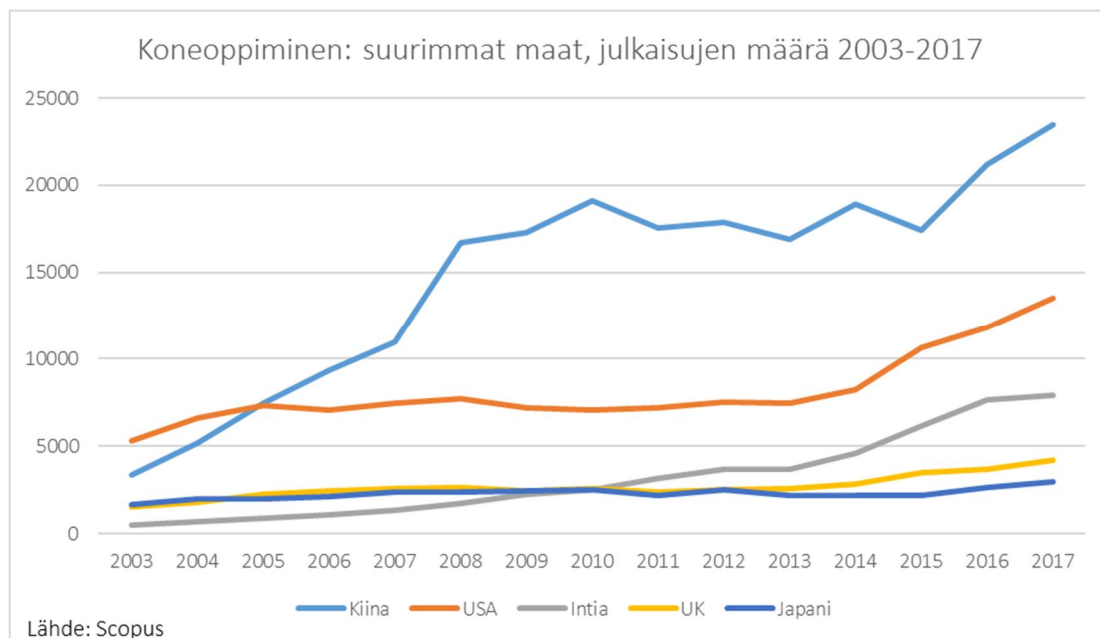
Kuva 50. Keskeisiä tutkimusaiheita: julkaisumäärät ja vuosikasvu. Kaikkien alan julkaisujen vuosikasvu ko. ajanjaksolla oli 11,5%. Osa sisällöltään yleisistä avainsanoista on jätetty kuvasta pois. Lähde: Scopus

Maat

Koneoppimista käsittelevien julkaisujen osuus kaikista tekoälyä käsittelevistä julkaisuista vaihtelee maittain, alle kolmasosasta lähes puoleen kaikista tekoälyä käsittelevistä julkaisuista. Alla kyseisten julkaisujen osuudet suurimmissa tutkimusmaissa ja Pohjoismaissa. Taulukko kuvaa ennen kaikkea sitä, miten vahvasti oppivat järjestelmät nähdään osana tekoälyn kenttää. Suuri osuus kertoo siten lähinnä tekoälyn tutkimuksen perinteistä kyseisessä maassa: oppivat järjestelmät nähdään kyseisissä maissa (myös) osana tekoälyn pitkää kehityshistoriaa



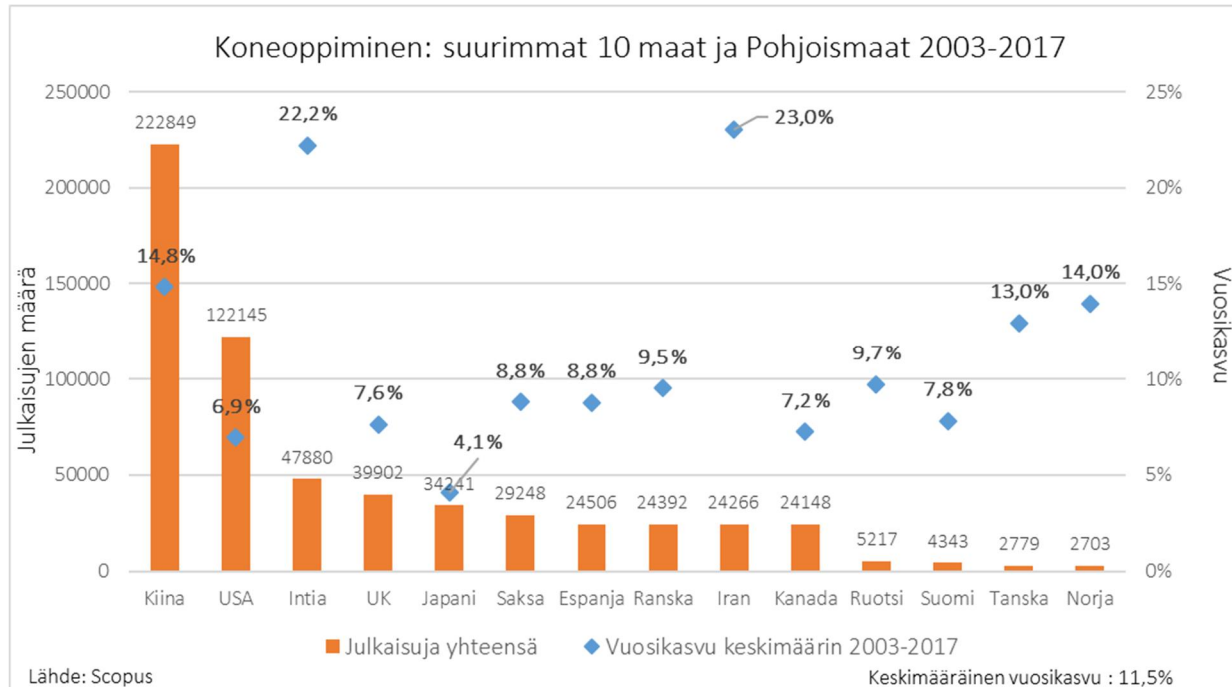
Kuva 51. Oppivia järjestelmiä käsittelevien julkaisujen osuus kaikista tekoälyä käsittelevistä julkaisuista 2015-2018 (27.3.2018). Lähde: Scopus



Kuva 52. Koneoppiminen, suurimmat maat, julkaisumäärät 2003-2017. Lähde: Scopus

Koneoppimisen tutkimusta nopeimmin lisänneet maat olivat tarkastellulla ajanjaksolla Iran, Intia ja Kiina. Kiinan ja Intian vahva rooli koneoppimisen tutkimuksen kasvu ei ole yllätys – sama trendi nähtävillä mm. data-analyyssissä sekä havainnoinnissa ja tilannetietoisuudessa – mutta Iranin nousu viime vuosina 10 eniten

julkaisevan maan joukkoon on poikkeus muista tekoälyn tutkimussuunnista. Mainitut kolme maata pitävät myös yllä koneoppimisen kasvua. Tosin koko tutkimussuunnan kasvuvauhti on sinänsä melko hyvin linjassa koko tekoälyn julkaisumäärien kasvun kanssa. Poikkeuksen tekee tutkimussuunnan julkaisuvolyymien koko: kyseessä on selvästi suurin yksittäinen tekoälyn tutkimussuunta maailmassa. Yksin Kiina julkaisi koneoppimiseen liittyviä julkaisuja vuonna 2017 kaksi kertaa niin paljon kuin kymmenen suurinta maata julkaisi robotiikan ja koneautomaation julkaisuja yhteen laskien samana vuonna.



Kuva 53. Koneoppiminen: suurimmat kymmenen maata ja Pohjoismaat, julkaisujen määrä ja julkaisumäärien keskimääräinen vuosikasvu 2003-2017. Lähde: Scopus

Julkaisujen painotukset vaihtelevat maittain. Osa eroista selittyy paitsi tutkimusalojen suuntauksilla myös käytetyllä terminologialla. Esimerkiksi Kiinassa ja Iranissa ei koneoppimisen julkaisuja välttämättä tulkita osana myös "tekoälyä".

Julkaisuja yhteensä 2015-2018	69057	39367	24561	12514	9466	9154	8456	7748	7377	7308	1858	1280	1088	1025
Termi	Kiina	USA	Intia	UK	Saksa	Iran	Japani	Ranska	Kanada	Italia	Ruotsi	Suomi	Tanska	Norja
Artificial Intelligence	9,9 %	20,6 %	15,5 %	19,4 %	18,8 %	8,8 %	17,8 %	17,7 %	19,0 %	17,6 %	16,3 %	19,2 %	16,4 %	17,8 %
Classification (of Information)	6,0 %	6,0 %	8,1 %	6,0 %	5,5 %	3,9 %	5,0 %	6,0 %	6,7 %	5,5 %	3,6 %	8,3 %	4,0 %	5,2 %
Complex Networks	3,4 %	3,0 %	2,5 %	3,1 %	3,0 %	2,9 %	2,6 %	3,2 %	4,0 %	4,3 %	2,2 %	3,0 %	3,3 %	2,4 %
Convolutional Neural Network	4,8 %	5,1 %	1,2 %	3,8 %	4,3 %	0,5 %	5,2 %	3,0 %	3,1 %	3,4 %	3,0 %	5,2 %	2,4 %	2,8 %
Data Mining	3,3 %	4,9 %	6,5 %	4,1 %	3,8 %	2,8 %	3,8 %	3,6 %	4,3 %	4,0 %	3,8 %	4,5 %	2,6 %	4,6 %
Decision Making	1,9 %	3,1 %	2,1 %	3,4 %	2,8 %	3,6 %	1,6 %	3,3 %	2,9 %	2,8 %	3,4 %	2,8 %	3,5 %	3,2 %
Deep Learning	4,5 %	5,7 %	1,8 %	4,6 %	4,2 %	0,5 %	5,0 %	3,4 %	4,7 %	3,6 %	3,6 %	4,0 %	4,1 %	4,8 %
Evolutionary Algorithms	6,4 %	3,2 %	8,3 %	7,5 %	4,7 %	7,9 %	5,6 %	5,4 %	5,5 %	5,0 %	4,0 %	6,2 %	6,6 %	4,6 %
Feature Extraction	4,5 %	2,8 %	6,6 %	3,4 %	2,3 %	3,2 %	3,1 %	2,3 %	3,2 %	2,3 %	2,2 %	3,1 %	2,3 %	3,0 %
Forecasting	7,0 %	5,9 %	5,5 %	5,0 %	4,2 %	9,0 %	4,3 %	4,0 %	6,6 %	4,6 %	5,3 %	4,8 %	5,6 %	6,9 %
Genetic Algorithms	12,3 %	6,7 %	13,1 %	9,3 %	5,6 %	23,5 %	8,6 %	9,4 %	10,2 %	10,2 %	6,8 %	6,3 %	8,7 %	10,6 %
Human (Bio-Medical)	3,7 %	13,0 %	3,4 %	12,0 %	11,7 %	4,4 %	5,7 %	7,4 %	10,3 %	8,6 %	12,8 %	8,7 %	13,0 %	8,9 %
Image Processing	4,2 %	3,7 %	4,6 %	3,5 %	3,7 %	2,5 %	3,5 %	3,7 %	3,1 %	3,4 %	2,5 %	4,6 %	3,0 %	3,6 %
Intelligent Systems	9,9 %	12,7 %	11,2 %	11,9 %	17,7 %	9,7 %	14,4 %	17,4 %	13,0 %	16,6 %	22,0 %	13,5 %	12,4 %	17,3 %
Learning Algorithms	8,1 %	10,0 %	8,2 %	7,8 %	7,4 %	6,2 %	6,7 %	7,3 %	9,1 %	6,8 %	7,4 %	6,4 %	6,8 %	9,1 %
Learning Systems	12,5 %	24,4 %	13,1 %	20,9 %	22,6 %	5,4 %	22,0 %	18,6 %	20,9 %	18,9 %	18,9 %	26,3 %	20,4 %	21,7 %
Machine Learning	4,6 %	15,5 %	6,2 %	13,2 %	13,0 %	3,1 %	8,5 %	9,0 %	13,0 %	9,8 %	13,6 %	13,7 %	15,0 %	13,3 %
Monte Carlo Methods	5,4 %	8,5 %	2,9 %	8,4 %	10,9 %	6,3 %	8,2 %	11,6 %	9,2 %	10,5 %	14,1 %	8,7 %	9,7 %	11,8 %
Multi Agent Systems	4,1 %	4,5 %	1,4 %	5,4 %	5,4 %	2,0 %	5,1 %	8,9 %	4,2 %	7,0 %	8,8 %	3,3 %	7,2 %	5,4 %
Neural Networks	17,0 %	14,0 %	16,1 %	14,9 %	13,5 %	21,8 %	15,8 %	12,7 %	14,9 %	16,6 %	9,5 %	15,2 %	10,8 %	12,4 %
Particle Swarm Optimization	8,8 %	2,5 %	10,7 %	3,5 %	1,6 %	12,6 %	3,8 %	2,9 %	4,7 %	2,5 %	1,8 %	2,2 %	3,8 %	2,8 %
Pattern Recognition	2,6 %	3,3 %	2,6 %	3,0 %	2,9 %	1,6 %	2,7 %	2,7 %	3,1 %	3,3 %	2,0 %	3,8 %	1,7 %	2,0 %
Signal Processing	2,6 %	2,5 %	4,1 %	2,6 %	2,1 %	2,4 %	2,4 %	2,5 %	2,6 %	2,1 %	2,4 %	5,1 %	3,8 %	2,0 %
Support Vector Machines	8,9 %	4,7 %	9,1 %	4,3 %	3,9 %	6,4 %	5,9 %	5,1 %	5,8 %	4,5 %	2,5 %	4,1 %	3,2 %	6,4 %

Kuva 54. Suurimmat kymmenen maata ja pohjoismaat 2015-2018, julkaisujen painotukset. Yleisimmät sisältöavainsanat koko julkaisuaineistossa. Osa julkaisutermeistä jätetty pois termin päällekkäisyyden tai yleisluonteisuuden perusteella. Termien prosentuaalinen osuus on $\neq$ 100% johtuen eri aiheiden esiintymisestä samoissa julkaisuissa. Vuoden 2018 tilanne: 2.5.2018. Lähde: Scopus

Tutkimusorganisaatiot: yliopistot, tutkimuslaitokset ja yritykset

Kiinalaiset yliopistot ja tutkimuslaitokset ovat erittäin vahvasti edustettuna koneoppimisen tutkimuskentässä. Ainoat ei-kiinalaiset tutkimusorganisaatiot 20 eniten julkaisseeseen yliopistoon tai tutkimuslaitokseen joukossa olivat vuosina 2008-2017 Ranskan CNRS (9.), Nanyangin teknillinen yliopisto Singaporessa (11.) ja Teheranin yliopisto (19.). Mikäli Kiinan tiedeakatemiaa ja opetusministeriötä ei huomioida, niin viisi aktiivisinta kiinalaista yliopistoa olivat Tsinghua-yliopisto, Harbinin teknologinen instituutti, Zhejiangin yliopisto, Beihang-yliopisto ja Shanghai Jiao-Tong –yliopisto.

Eurooppalaisista yliopistoista ja tutkimuslaitoksista eniten julkaisuja tuottivat CNRS:n lisäksi Lontoon Imperial College (54.), Delftin yliopisto Hollannissa (61.), Lontoon UCL (67.) sekä Münchenin teknillinen yliopisto (70.) ja ETH Zürichissä (71.).

Pohjoismaisista yliopistoista aktiivisin v. 2008-2017 oli Aalto yliopisto (852 julkaisua), Ruotsin KTH (825 julkaisua) ja Tanskan teknillinen yliopisto (594 julkaisua). Suomessa Tampereen teknillinen yliopisto tuotti 362 ja Helsingin yliopisto 309 julkaisua kyseisenä ajanjaksona. Seuraavaksi aktiivisimmat suomalaiset tutkimusorganisaatiot olivat Oulun ja Jyväskylän yliopistot ja Teknologian tutkimuskeskus VTT.

Tutkimuksen vaikuttavuus: maat

Normalisoidun (suhteutetun) viittausindeksin perusteella Israelin, Alankomaiden ja Kanadan julkaisut olivat v. 2013-2017 viitatuimpia tutkimusjulkaisuja koneoppimisen alueella.²⁹ Top 10% -huippujulkaisuissa ilmestyneitä artikkeleita tuotettiin eniten Hong Kongissa. Myös Singaporessa, Taiwanissa ja Australiassa huippujulkaisujen suhteellinen osuus oli korkea. Euroopassa huippujulkaisuja syntyi erityisen paljon Belgiassa ja Sveitsissä. Molemmissa yli 20% artikkeleista ilmestyi top 10 –huippujulkaisuissa. Suomen luku oli 13,5%, mikä oli lähellä Pohjoismaista keskitasoa.

Yritys-yliopisto –tutkimusyhteistyötä oli v. 2013-2017 koneoppimisessa eniten Israelissa, Ruotsissa, USA:ssa ja Sveitsissä. Suomessa yritys-yliopisto –yhteistyötä oli kuudenneksi eniten vertailussa mukana olevista maista. Enemmän kuin esimerkiksi seuraavina tulevilla mailla, Saksassa ja Australiassa.

Kansainvälisessä tutkimusyhteistyössä olivat kärjessä Saudi-Arabia, Hong-Kong, Sveitsi ja Singapore. Kansainvälinen tutkimusyhteistyö lisää yleensä julkaisuviittausten määrää, joten usein samat maat ja yliopistot ovat näkyvästi esillä sekä viittausmääriä että tutkimusyhteistyön laajuutta kuvaavissa indekseissä.³⁰

²⁹ Julkaisujoukko rajattiin Scopus-tietokannasta avainsanoilla "machine learning", "learning systems" ja "deep learning". Lisäksi haku kohdistettiin computer science -tutkimusalaan.

³⁰ "International co-authorship has a positive contribution to the FWCI of the institution", Influence of international co-authorship on the research citation impact of young universities, Khor K, Yu L. Scientometrics. 2016, 107, 1095-1110

Maa	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankiingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys - yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisutietojen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Israel	3,04	432	23,2	1,21	11,8	1,35
Netherlands	2,93	680	19,4	1,61	6,3	1,87
Canada	2,6	1863	17,1	1,37	4,6	1,66
Hong Kong	2,55	677	31,4	1,79	4,6	1,8
Switzerland	2,54	640	21,1	1,76	10,3	1,63
Singapore	2,47	996	25,6	1,72	4,5	1,66
United States	2,36	11559	21,6	0,93	10,6	1,5
United Kingdom	2,14	3224	19,7	1,49	5,9	1,75
Italy	2,04	1647	15,2	1,11	4,3	2,14
Austria	2,01	423	14,5	1,46	3,3	1,76
Australia	1,8	1621	23,7	1,35	6,4	1,81
Saudi Arabia	1,76	447	18,1	1,8	2,9	2,31
France	1,68	1859	18,1	1,37	6	1,41
Finland	1,65	523	13,5	1,42	8,4	1,58
Spain	1,64	1764	17,5	1,15	4,3	2,04
Belgium	1,63	501	23,6	1,47	5,8	1,62
Germany	1,6	2647	14,4	1,12	7,3	1,44
Sweden	1,56	422	14,8	1,45	10,7	1,85
Greece	1,53	578	10	1,09	1,4	1,97
Russian Federation	1,5	571	6,2	0,66	1,2	2,07
China	1,49	9980	14,8	0,81	2,9	1,23
Portugal	1,35	570	11,9	1,12	2,1	1,91
Iran	1,33	510	14,8	0,85	0	1,65
South Korea	1,32	1153	15,1	0,79	4,9	1,59
Poland	1,32	679	14	0,81	1,2	1,62
Brazil	1,18	1179	12,6	0,83	1,7	1,6
Malaysia	1,1	676	12,6	0,99	0,6	2,53
India	1,06	3483	9,6	0,38	3,2	1,69
Taiwan	1,05	1230	24,8	0,49	2,1	1,63
Turkey	0,98	832	19,3	0,52	2,3	1,45
Japan	0,9	2355	8	0,62	6,2	1,1

Kuva 55. Koneoppiminen ja oppivat järjestelmät³¹: maat järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2013-2017. Alle 400 julkaisua tuottaneet maat eivät ole mukana taulukossa. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen yhteistyön ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa. (Tiedot Scopusesta: katselukerrat antavat tietoa julkaisun herättämästä kiinnostuksesta täydentämällä viittausmäärien välittämää informaatiota erityisesti uusimman aineiston osalta.) Normalisoiduissa mittareissa maailman keskiarvo = 1. Analysoituja julkaisuja yhteensä 50273. Lähde: Scopus ja SciVal. Tilanne: 10.5.2018

Tutkimuksen vaikuttavuus: tutkimusorganisaatiot

Amerikkalaiset yliopistot ja yritykset ovat vahvasti edustettuna, kun näkökulmana on julkaisujen keräämät lähdeviittaukset. Yrityksen (Google) sijoittumista viittausvertailun kärkeen voi pitää melko poikkeuksellisenä tieteellisessä tutkimuksessa. Jos toisaalta arvioinnin kriteerinä käytetään huippujulkaisujen määrää, siirtyy painopiste yliopistoihin. Eniten huippujulkaisuja suhteessa muuhun julkaisutuotantoon oli Kalifornian San Diegon yliopistolla ja Hong Kongin kiinalaisella yliopistolla, molemmilla osuus oli yli 40%. Kansainvälisen julkaisuyhteistyön osuus oli suurin Zürichin ETH:lla, Hong Kongin Polyteknisellä Yliopistolla ja Sydneyn Teknisellä Yliopistolla. Suurista yrityksistä eniten yhteistyötä yliopistojen kanssa oli Microsoftilla. Yliopistoista eniten yhteistyötä yritysten kanssa v. 2013-2017 oli puolestaan Cambridgen yliopistolla, Georgia Techilla ja

³¹ Haku Scopus-tietokannasta avainsanoilla "machine learning", "learning systems" ja "deep learning". Lisäksi haku kohdistettiin tutkimusalaan computer science.

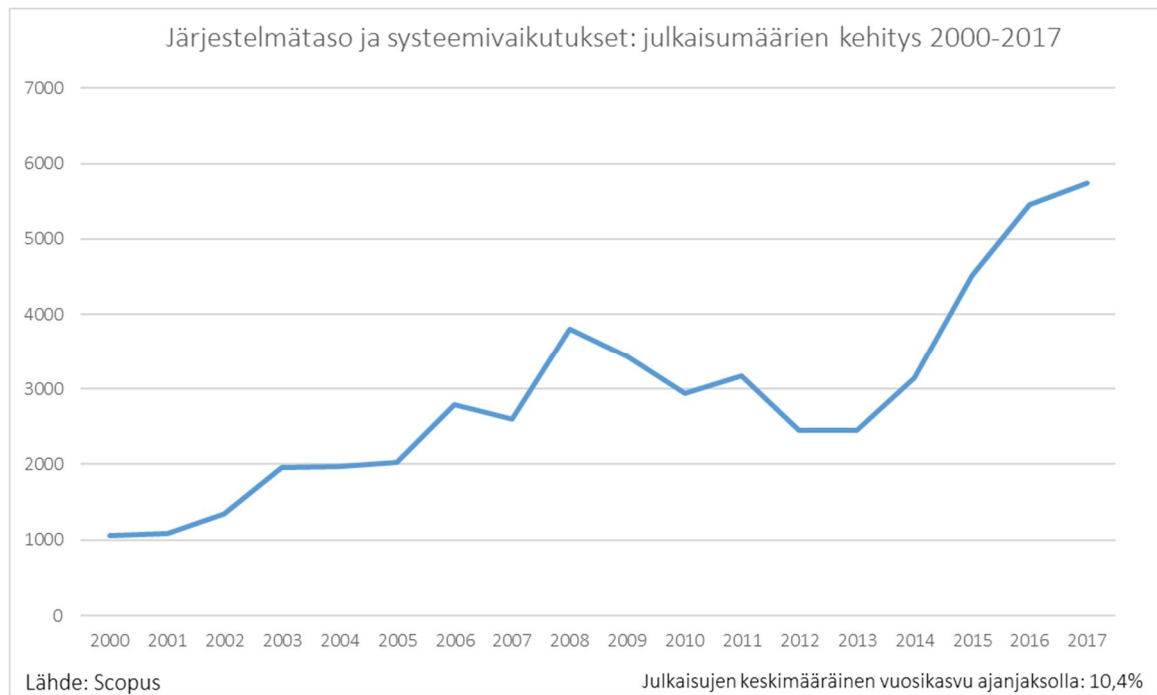
Eurooppalaista yliopistoista yritysyehtistyötä oli eniten Lausanne EPFL:llä, Edinburghin yliopistolla ja Münchenin teknisellä yliopistolla. Myös Aalto-yliopisto oli tässä vertailussa korkealla, viidentenä eurooppalaisista yliopistoista (julkaisuista 10% oli laadittu yhteistyössä yritysten kanssa). Viittausmääriä vertailtaessa oli eurooppalaisista yliopistoista kärjessä Cambridgen yliopisto ja Lontoon Imperial College. Eurooppalaisista yliopistoista suhteessa eniten huippujulkaisuja oli Oxfordin yliopistolla ja Lausannen EPFL:llä.

Tutkimuslaitos / yliopisto	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimuslalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys - yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisujen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Google Inc.	6,76	207	20,2	1,09	67,6	1,56
University of Toronto	6,69	172	20,9	1,39	14	1,97
University of California at Berkeley	6,42	260	29,8	0,91	17,3	1,6
Cornell University	4,33	124	-	0,98	14,5	1,27
Stanford University	4,31	314	31,9	0,84	17,2	1,54
University of Oxford	4,19	204	28,4	1,42	7,8	1,61
Microsoft USA	4,11	401	28,7	1,27	82,8	1,54
Chinese University of Hong Kong	3,9	163	43,2	1,66	4,3	1,64
New York University	3,85	149	28	1,1	16,8	1,67
University of Cambridge	3,73	167	-	1,61	19,8	1,45
Imperial College London	3,48	250	23,9	1,53	4	1,98
Purdue University	3,41	130	-	0,66	16,9	1,58
University of Trento	3,29	125	-	1,55	8,8	2,13
Massachusetts Institute of Technology	3,24	365	24,3	0,97	13,7	1,9
University of British Columbia	3,09	150	21,1	1,46	14,7	1,86
University of Washington	3,01	220	26	0,57	15	1,34
University of Wisconsin	2,98	140	27,5	0,79	8,6	1,25
Pennsylvania State University	2,95	135	-	0,79	10,4	1,53
University of California at Los Angeles	2,93	141	29,2	0,87	15,6	1,62
ETH Zurich	2,93	170	23,2	1,92	11,8	1,47
Technical University of Munich	2,93	207	16,9	1,43	12,6	1,73
Intel	2,85	122	-	0,87	71,3	2,45
University of Michigan	2,79	193	33,7	0,85	8,3	1,68
EPFL Lausanne	2,71	153	27,5	1,7	13,1	1,56
Nanyang Technological University	2,67	514	27,3	1,61	5,1	1,88
University of Pittsburgh	2,67	137	24,1	0,8	5,8	1,52
UTS University of Technology Sydney	2,64	234	37,6	1,79	3,4	1,61
University of California at San Diego	2,62	161	44,2	0,79	16,1	1,44
University of Illinois at Urbana-Champaign	2,62	216	21	0,72	17,1	1,39
IBM	2,59	282	12,9	1,13	64,5	1,61
Columbia University	2,53	186	19,2	0,93	17,7	1,62
University of Maryland	2,49	150	26	0,85	14,7	1,45
Tsinghua University	2,46	536	24,2	1,11	7,3	1,31
Institute of Computing Technology Chinese Academy of Sciences	2,45	175	15,2	0,57	5,1	1,3
Microsoft Research Asia	2,43	134	-	1,14	65,7	1,5
National University of Singapore	2,43	256	32,3	1,75	3,1	1,42
Sun Yat-Sen University	2,43	146	27,3	0,92	2,1	1,18
Carnegie Mellon University	2,35	560	26	1	11,6	1,37
Chinese Academy of Sciences	2,33	1001	19,2	0,84	3,2	1,31

1.3.8 Järjestelmätaso ja systeemivaikutukset³²

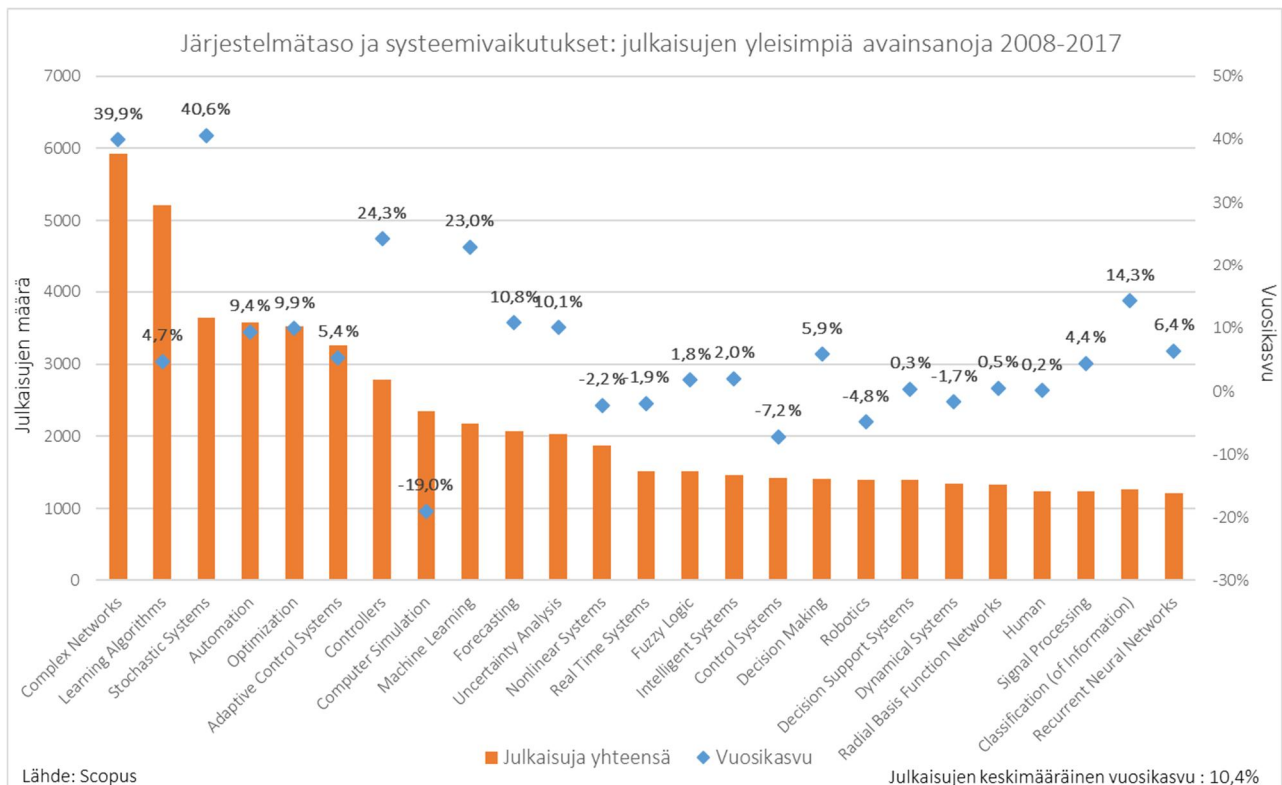
³² Tutkimussuunnan rajoituksissa käytettiin seuraavia termejä yhdessä tekoälyaiheiden kanssa: adaptive control systems, complex networks, control systems, dynamical systems, electric power systems, electric power transmission networks, large scale systems, linear systems, manufacture, multi-body system dynamics, nonlinear control systems, nonlinear systems, operational analysis , power system dynamics, real

Erilaisten systeemien hallinta sekä systeemiset lähestymistavat ja ratkaisut mm. prosessien ja verkostojen asettamiin haasteisiin ovat tyypillisiä valmistamassa teollisuudessa, energiasektorilla sekä esimerkiksi logistiikassa ja liikenteessä. Näillä aloilla myös rajatusti tekoälyä hyödyntäviä sovelluksia ja ratkaisuja on käytetty jo pitkään. Tämä näkyy myös tutkimussuunnan julkaisumäärissä. Pitkän nousuvaiheen voi katsoa päättyneen vuoden 2008 talouskriisiin, minkä jälkeen tutkimusjulkaisujen määrä laski vuoteen 2013 saakka. Vaikka monilla tekoälyn tutkimussuunnilla vuodet 2008 ja 2009 muodostivat tietynlaisen käännekohtan, oli muutos ja julkaisumäärien lasku järjestelmätasoa ja systeemien hallintaa käsittelevissä julkaisuissa kuitenkin poikkeuksellisen voimakas. Uusi ja aiempaa selvästi vahvempi nousu alkoi vuonna 2014 ja jatkuu yhä, vaikkakin aiempaa hieman hitaampana. Taustalla on nähtävissä mm. oppiviin järjestelmiin kohdistunut lisääntynyt kiinnostus.



Kuva 57. Tieteellisten julkaisujen määrä 2000-2017. Lähde: Scopus

Tutkimussuunnan aiheita tarkasteltaessa huomio kiinnittyy nopeimmin kasvaneisiin avaintermeihin. Isoista aiheista erityisen nopeasti julkaisumääriä ovat lisänneet complex networks ja stochastic systems. Koko ajanjaksoa tarkastellessa pienemmistä aiheista ovat huomattavan nopeasti kasvaneet mm. deep learning (230% v. 2012-2017), convolutional neural networks (79% v. 2008-2017) ja deep neural networks (71% v. 2008-2017)



Kuva 58. Keskeisiä tutkimusaiheita: julkaisumäärät ja vuosikasvu. Kaikkien alan julkaisujen vuosikasvu ko. ajanjaksolla oli 10,4%. Tilastosta puuttuvat yleisimmät avainsanat: artificial intelligence, neural networks ja learning systems. Lisäksi muutamia muita sisällöltään yleisiä avainsanoja on jätetty kuvasta pois. Lähde: Scopus

1.3.9 Tekoälyn laskentaympäristöt, alustat ja palvelut, ekosysteemit³³

Julkaisumäärät ja aiheet

Tutkimussuunnan julkaisumäärien kehityksen voi jakaa 2000 luvulla karkeasti kolmeen vaiheeseen. Julkaisumäärät kasvoivat vuoteen 2006 asti melko tasaisesti. Tämän jälkeen seurasi vaihe, jolloin julkaisumäärät pysyivät käytännössä paikoillaan vuoteen 2014 asti. Valtaosa tutkimusalueen kasvusta onkin tapahtunut aivan viime vuosina. Kun koko ajanjaksolla keskimääräinen vuosittainen kasvunopeus on ollut n. 15%, niin v. 2015-2017 julkaisumäärä kasvoi vuosittain lähes 40%, mitä voi pitää jo hyvin harvinaisena kasvulukuna tutkimusjulkaisuissa.

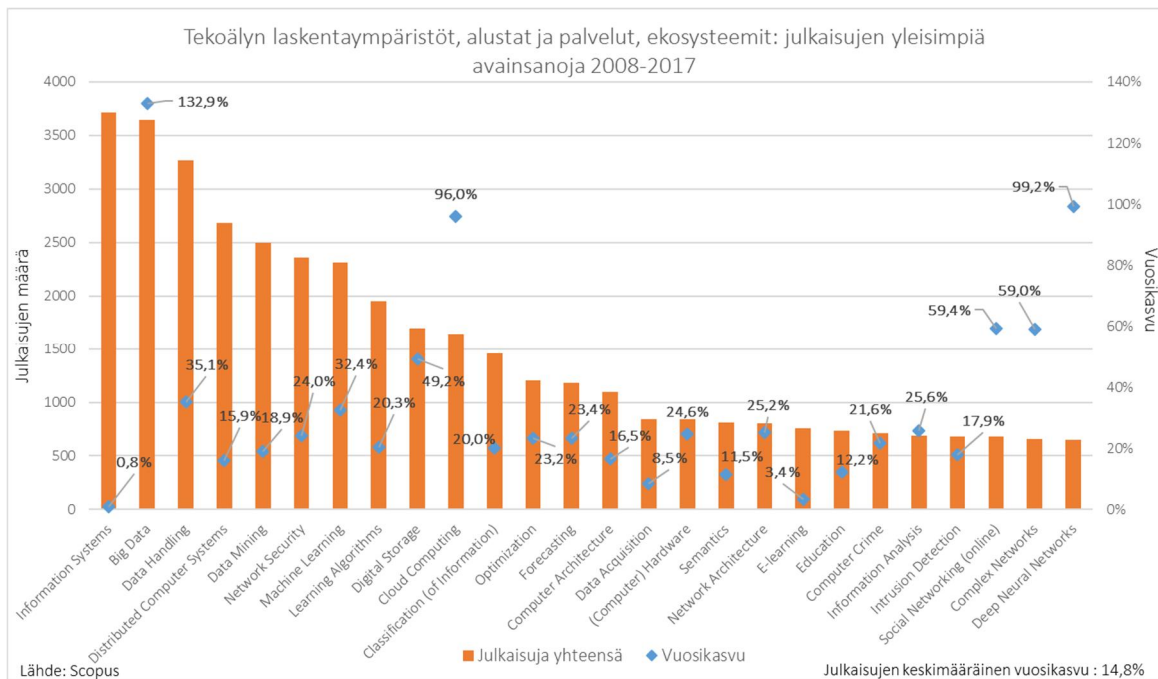
³³ Tutkimussuunnan rajauksessa käytettiin seuraavia termejä yhdessä tekoälyaiheiden kanssa: cloud computing, cloud environments, computer architecture, computer hardware, data acquisition, data analysis software, data handling, digital storage, distributed computer systems, hadoop, high-performance computing, information systems, mapreduce, network security



Kuva 59. Tieteellisten julkaisujen määrä 2000-2017. Lähde: Scopus

Sisällöllisesti tarkastelen isojen datamassojen hallintaan ja hyödyntämiseen liittyvä tutkimus selittää pääosin julkaisumäärien voimakkaan kasvun viime vuosina. Suuret aineistomäärät asettavat luonnolliset myös alustoille ja laskentaympäristöille uusia vaatimuksia. Toisaalta esimerkiksi hajautetut järjestelmät ovat olleet koko tarkastelujakson yksi merkittävä tutkimuskohde, vaikka myös sitä käsittelevien julkaisujen määrä on kasvanut koko tutkimusala nopeammin.

Isoista tutkimusaiheista nopeimmin kasvoivat big data, pilvilaskenta ja digitaalinen tallentaminen ylipäättään. Pienemmistä teemoista julkaisumäärät lisääntyivät nopeimmin koneoppimisen alateknologioissa myös tässä tutkimussuunnassa. Näitä olivat mm. CNN- ja syvät neuroverkot sekä ns. syväoppiminen. Lisäksi nk. kompleksiset verkot ja sosiaalisen median alustoihin liittyvä tutkimus olivat vahvasti nousussa. Laiteympäristöjen turvallisuus on keskeisessä roolissa kehitettäessä esimerkiksi pilvilaskentaa ja niihin tukeutuvia ekosysteemejä. Kyberrikollisuuden torjuntaan liittyvät julkaisut olivatkin selvästi näkyvissä nousevien aiheiden joukossa.

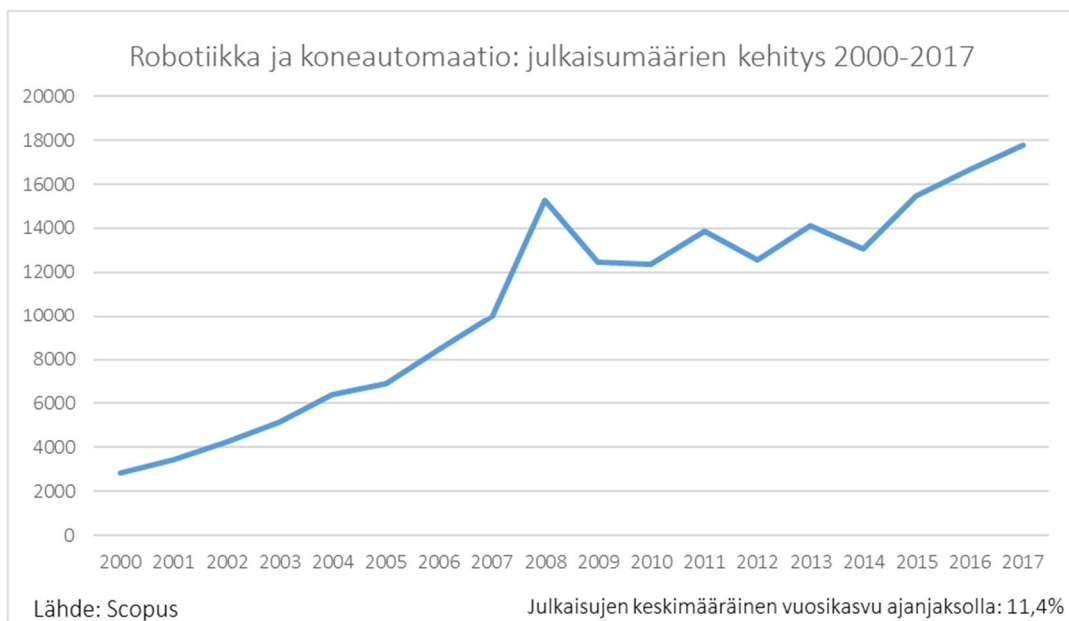


Kuva 60. Keskeisiä tutkimusaiheita: julkaisumäärät ja vuosikasvu. Kaikkien alan julkaisujen vuosikasvu ko. ajanjaksolla oli 14,8%. Tilastosta puuttuvat yleisimmät avainsanat: artificial intelligence, neural networks ja learning systems. Lisäksi muutamia muita sisällöltään yleisiä avainsanoja on jätetty kuvasta pois. Lähde: Scopus

1.3.10 Robotiikka ja koneautomaatio - tekoälyn fyysinen ulottuvuus³⁴

Julkaisumäärät ja aiheet

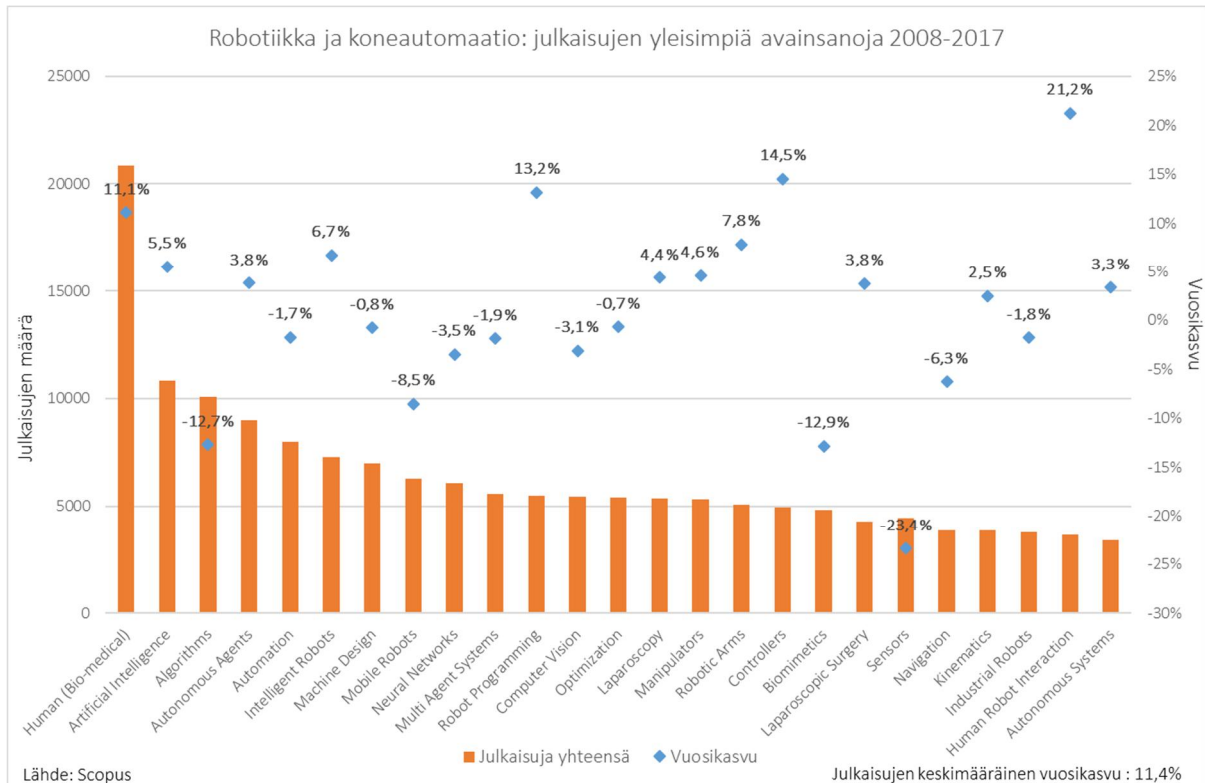
Robotiikan ja koneautomaation julkaisumäärät kasvoivat verrattain nopeasti 2000-luvun alussa. Nousu kuitenkin päättyi vuoteen 2008. Finanssikriisin ja sitä seuranneen kansainvälisen taantuman seurauksena julkaisumäärät palasivat vuoden 2008 tasolle vasta vuonna 2015. Viime vuosina julkaisujen määrässä ollaan palattu 2000-luvun alun kasvulukuihin.



³⁴ Tutkimussuunnan rajauksessa käytettiin seuraavia termejä: autonomous agents, autonomous systems, industry 4.0, robotic systems, robot learning, robotics sekä seuraavia termejä yhdessä tekoälyaiheiden: internet of things, robots, automation, process control, factory automation

Kuva 61. Tieteellisten julkaisujen määrä 2000-2017. Lähde: Scopus

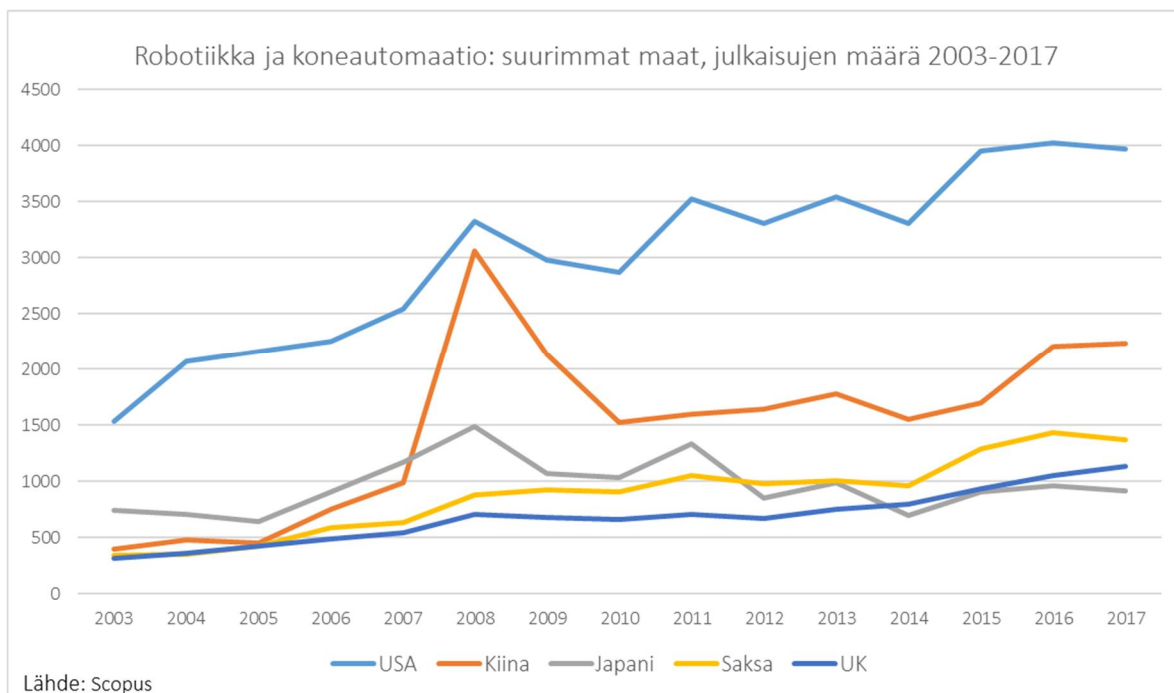
Robottiikan ja koneautomaation aiheita lähemmin tarkastellessa kiinnittyy huomio erityisesti kirurgian ja muun sairaanhoidon robotiikkatutkimuksen vahvaan osuuteen. Isommista aiheista nopeimmin on kirurgiarobotiikan lisäksi kasvanut ihmisen ja koneen vuorovaikutuksen tutkimus. Pienemmistä teemoista julkaisumäärät lisääntyivät koneoppimisen alateknologioiden lisäksi "Industry 4.0"-aihepiirin julkaisut (kasvu peräti 250% v. 2012-2017). Muita nopeasti kasvaneita aiheita olivat ihmisen ulkoiseen tukirankaan (exoskeleton, kasvu 102% v. 2008-2017), antropomorfisiin robotteihin ja mm. toimintaympäristön kartoittamiseen (mapping) keskittyvät julkaisut.



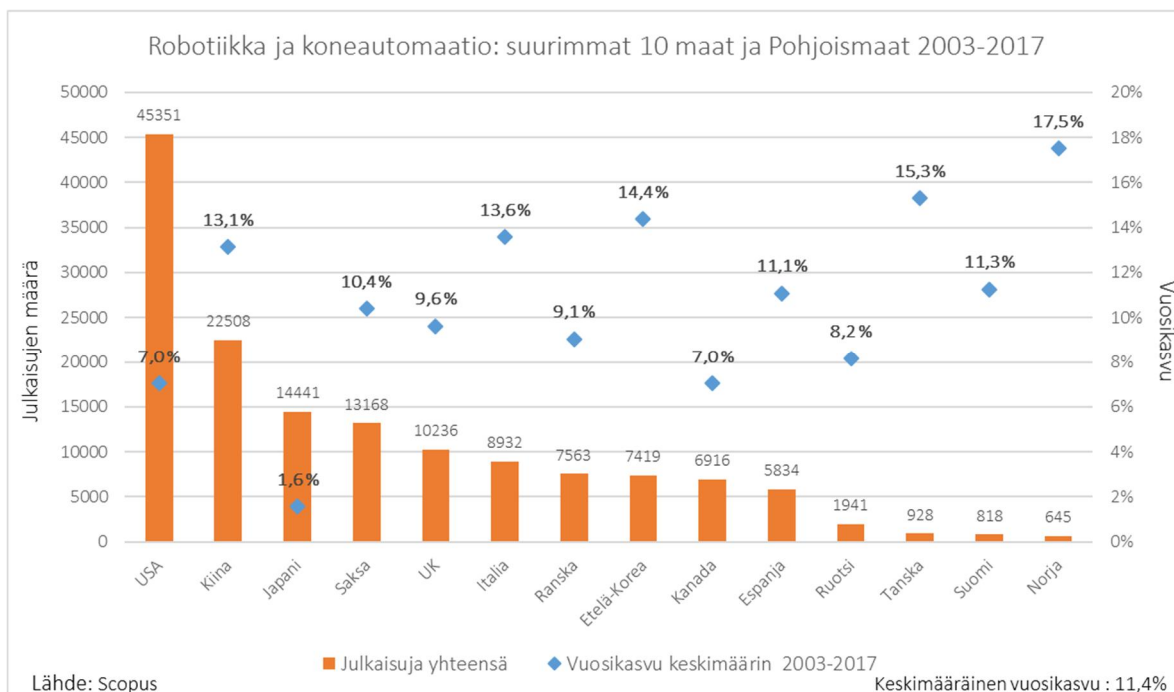
Kuva 62. Keskeisiä tutkimusaiheita: julkaisumäärät ja vuosikasvu. Kaikkien alan julkaisujen vuosikasvu ko. ajanjaksolla oli 11,4%. Tilastosta puuttuvat tutkimussuunnan yleisimmät avainsanat: robots ja robotics. Lisäksi muutamia muita sisällöltään yleisiä avainsanoja on jätetty kuvasta pois. Lähde: Scopus

Maat

Toisin kuin monissa muissa tekoälyn tutkimussuunnissa (esim. Data-analyysi tai Havainnointi ja tilannetietoisuus) Yhdysvallat hyvin selkeästi suurin julkaisumaa, vaikka 2. suurin maa, Kiina, kasvattaakin julkaisumääriään Yhdysvaltoja nopeammin. Japani on perinteisesti ollut robotiikan edelläkävijämaita, näin on myös tutkimusjulkaisuissa. Tosin Japani suhteellinen osuus on pienenemässä, koska julkaisujen määrä ei näytä enää Japanissa kasvavan. Muista suuremmista julkaisijoista Etelä-Korea ja Italia ovat lisänneet nopeimmin julkaisuvolyymiaan. Pohjoismaista Norja ja Tanska kasvattavat tutkimustaan suhteellisesti nopeimmin.

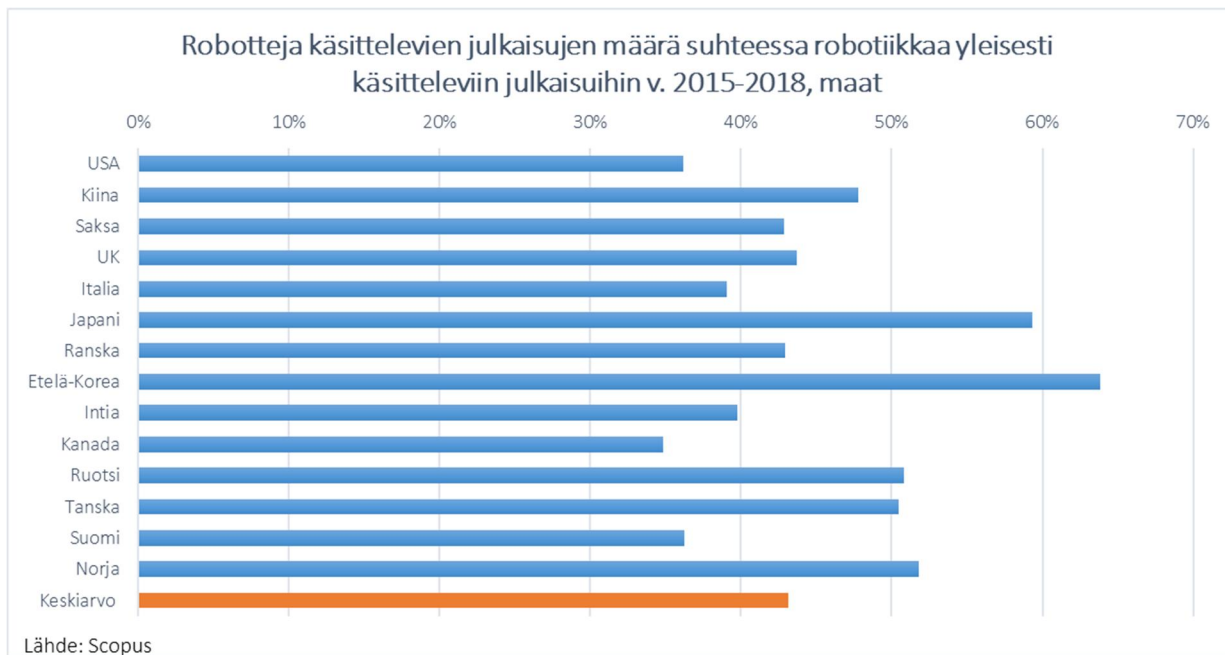


Kuva 63. Robottiikka ja koneautomaatio, suurimmat maat, julkaisumäärät 2003-2017. Lähde: Scopus



Kuva 64. Robottiikka ja koneautomaatio: suurimmat kymmenen maata ja Pohjoismaat, julkaisujen määrä ja julkaisumäärien keskimääräinen vuosikasvu 2003-2017. Lähde: Scopus

Kun tarkastellaan maiden välisiä eroja julkaisujen sisällössä, yksi tapa on katsoa maiden panostusta robotteja käsitteleviin julkaisuihin ja verrata kyseisten julkaisujen osuutta kaikkiin maan robottiikka-aiheisiin julkaisuihin. Japani ja Etelä-Korea erottuvat tässä vertailussa muista enemmän robottitutkimukseen panostavina maina.



Kuva 65. Robotteja ("robots") käsittelevien julkaisujen määrä suhteessa robotiikkaa ("robotics") yleisesti käsitteleviin julkaisuihin v. 2015-2018. Lähde: Scopus.

Myös tarkasteltaessa lähemmin painotuksia julkaisuissa on tiettyjä eroja nähtävissä maiden välillä. Robotiikan hyödyntäminen terveydenhoidossa ja lääketieteen apuna on erityisen vahvasti esillä Yhdysvalloissa ja Italiassa, joissa noin neljännes julkaisuista liittyy robotiikan ja automaation hyödyntämiseen terveyssektorilla. Etelä-Koreassa älykkäisiin robotteihin liittyvät julkaisut kattavat lähes neljänneksen koko maan julkaisuista v. 2015-2018.

Julkaisuja yhteensä 2015-2018	12777	6732	4366	3351	3281	2913	2244	2217	2078	1814	587	362	281	263
Termi	USA	Kiina	Saksa	UK	Italia	Japani	Ranska	Etelä-Korea	Intia	Kanada	Ruotsi	Tanska	Suomi	Norja
Algorithms	5,2 %	5,7 %	3,6 %	3,8 %	3,5 %	3,1 %	4,5 %	5,2 %	4,7 %	5,7 %	4,1 %	2,5 %	2,1 %	5,7 %
Anthropomorphic Robots	2,0 %	1,9 %	4,1 %	2,9 %	4,4 %	6,1 %	4,9 %	2,9 %			2,6 %	3,0 %	2,8 %	
Artificial Intelligence	7,3 %	8,9 %	6,9 %	9,6 %	7,1 %	8,3 %	8,4 %	23,1 %	17,2 %	8,3 %	12,9 %	7,7 %	10,3 %	13,7 %
Automation	5,1 %	5,7 %	7,6 %	7,5 %	3,8 %	3,1 %	4,4 %	3,7 %	11,6 %	7,9 %	8,2 %	3,9 %	8,9 %	9,9 %
Autonomous Agents	8,9 %	5,1 %	6,4 %	12,4 %	6,8 %	4,9 %	10,3 %	2,6 %	4,9 %	9,0 %	8,3 %	6,1 %	11,7 %	10,3 %
Biomimetics	1,6 %	9,2 %	1,8 %	2,1 %	1,8 %	6,0 %	2,0 %	1,7 %		1,5 %				2,3 %
Computational Methods	4,4 %	2,8 %	2,9 %	5,1 %	3,5 %	2,0 %	4,8 %	1,7 %	1,5 %	5,0 %	3,2 %	3,0 %	6,8 %	7,2 %
Computer Vision	2,0 %	4,5 %	2,4 %	2,7 %	1,4 %	2,5 %	3,6 %	3,0 %	3,4 %	3,7 %	3,6 %	4,1 %	3,6 %	4,6 %
Controllers	2,7 %	6,0 %	2,7 %	2,5 %	2,6 %	2,8 %	3,9 %	3,4 %	8,0 %	4,9 %	1,2 %	3,6 %	2,8 %	2,3 %
Human (Bio-Medical)	26,4 %	7,6 %	12,7 %	19,5 %	24,6 %	14,1 %	17,9 %	21,0 %	9,2 %	19,8 %	16,4 %	20,2 %	8,2 %	14,4 %
Human Robot Interaction	4,0 %	2,5 %	4,8 %	5,0 %	4,8 %	6,2 %	4,6 %	3,6 %	1,6 %	3,2 %	7,5 %	6,6 %	3,6 %	2,3 %
Intelligent Robots	5,4 %	6,5 %	4,8 %	5,2 %	4,4 %	6,9 %	4,4 %	23,2 %	3,8 %	4,2 %	5,1 %	5,5 %	2,5 %	3,4 %
Laparoscopy	5,9 %	1,6 %	2,4 %	3,6 %	5,7 %	3,1 %	4,8 %	6,7 %	2,0 %	2,1 %	2,7 %	5,8 %	1,8 %	5,7 %
Learning Systems	4,1 %	4,1 %	4,5 %	5,0 %	3,1 %	4,2 %	4,2 %	3,5 %	8,4 %	4,7 %	6,3 %	5,0 %	7,5 %	5,7 %
Machine Design	4,8 %	6,6 %	4,0 %	4,7 %	5,1 %	5,6 %	4,1 %	5,6 %	4,8 %	3,8 %	3,4 %	8,3 %	4,3 %	3,8 %
Manipulators	3,2 %	8,7 %	3,0 %	2,8 %	5,0 %	4,2 %	4,0 %	5,8 %	7,4 %	7,2 %	3,2 %	5,0 %	7,5 %	6,5 %
Mobile Robots	1,4 %	4,3 %	2,8 %	2,4 %	1,8 %	3,9 %	2,6 %	3,5 %	4,1 %	2,4 %	2,7 %	2,2 %	4,6 %	1,9 %
Motion Planning	3,7 %	3,8 %	2,7 %	1,7 %	1,6 %	1,8 %	2,1 %	3,2 %	5,1 %	2,6 %	4,3 %	2,5 %	2,8 %	4,9 %
Multi Agent Systems	3,8 %	3,0 %	3,4 %	6,0 %	3,9 %	2,4 %	6,8 %		3,3 %	3,4 %	5,1 %	2,8 %	4,6 %	3,4 %
Neural Networks	1,9 %	6,0 %	2,8 %	3,8 %	2,2 %	3,9 %	2,5 %	3,0 %	6,5 %	3,8 %	1,4 %	2,8 %	2,5 %	2,7 %
Optimization	2,9 %	5,4 %	3,0 %	2,5 %	1,7 %	1,8 %	3,5 %	1,4 %	6,4 %	3,4 %	2,9 %	4,4 %	2,1 %	6,1 %
Procedures	13,8 %	3,9 %	7,2 %	9,0 %	13,7 %	7,0 %	9,6 %	12,9 %	3,5 %	9,1 %	9,2 %	11,0 %	3,2 %	9,1 %
Robot Programming	4,9 %	6,1 %	5,5 %	3,6 %	3,8 %	3,4 %	4,1 %	2,7 %	6,0 %	3,6 %	7,8 %	7,7 %	4,6 %	6,8 %
Robotic Arms	2,4 %	6,5 %	2,9 %	2,7 %	3,1 %	6,8 %	2,8 %	4,5 %	5,2 %	3,0 %	1,9 %	3,3 %		1,9 %
Robotic Surgical Procedure	12,0 %	2,4 %	4,7 %	7,3 %	10,2 %	4,8 %	7,8 %	11,0 %	4,0 %	6,0 %	7,0 %	7,2 %	2,5 %	6,8 %

Kuva 66. Suurimmat kymmenen maata ja pohjoismaat 2015-2018, julkaisujen painotukset. Yleisimmät sisältöavainsanat koko julkaisuaineistossa. Osa julkaisutermeistä jätetty pois termin päällekkäisyyden tai yleisluonteisuuden perusteella. Termien prosentuaalinen osuus on $\neq 100\%$ johtuen eri aiheiden esiintymisestä samoissa julkaisuissa. Vuoden 2018 tilanne: 26.4.2018. Lähde: Scopus

Tutkimuksen vaikuttavuus: maat

Normalisoidun (suhteutetun) viittausindeksin perusteella Sveitsi ja Belgia tuottivat v. 2013-2017 tieteellisesti vaikuttavimpia robotiikkaan ja koneautomaatioon liittyviä tekoälyn tutkimusjulkaisuja. Myös esim. Ruotsissa

julkaisujen viittausmäärät olivat suhteellisesti korkealla tasolla. Yksittäisiä top 10% -huippujulkaisuissa ilmestyneitä artikkeleita tuottivat suhteellisesti eniten Belgia, Alankomaat, Hong-Kong ja Ruotsi (maista joilta ilmestyi vähintään 300 julkaisua).

Kansainvälinen yhteistyö tutkimusjulkaisuissa oli intensiivisintä Hong Kongissa ja Belgiassa. Hong Kongin korkeaa lukua toisaalta selittänee osaltaan tiivis yhteistyö Kiinan kanssa. Yritys-yliopisto –yhteistyötä oli tutkimusalalla v. 2013-2017 selvästi eniten Ruotsissa ja Saksassa, molemmissa vähintään 5% julkaisuista tuotettiin yhteistyössä yrityksen kanssa. Seuraavina tulivat Japani, Alankomaat, Sveitsi ja Yhdysvallat. Myös Suomessa yritysten ja yliopistojen / tutkimuslaitosten välinen yhteistyö oli selvästi keskimääräistä tiiviimpää: Suomi oli kyseisessä vertailussa yhdeksäntenä, heti Israelin ja Ranskan jälkeen.

Maa	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys -yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisutietojen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Switzerland	2,2	1464	22,7	1,59	3,2	1,7
Belgium	1,88	728	28,2	1,82	2,5	1,69
Sweden	1,69	882	24	1,59	5,1	1,65
Italy	1,67	4526	16	1,22	1,5	2,11
Germany	1,65	6038	15,1	1,03	5	1,7
United States	1,63	18823	21,5	0,73	3,2	1,3
Norway	1,62	332	23,3	1,4	2,4	1,73
Netherlands	1,58	1345	25,4	1,56	3,7	1,61
Greece	1,58	686	14,9	1,35	1,2	1,66
Singapore	1,56	1209	18,6	1,52	2,2	1,41
Austria	1,55	734	13,1	1,59	2,5	1,74
United Kingdom	1,54	4686	18,1	1,44	2,8	1,55
Australia	1,49	1986	19,7	1,23	2,4	1,64
Israel	1,49	485	22,3	1,3	3,1	1,28
Denmark	1,49	450	12,2	1,4	2,2	1,58
Hong Kong	1,48	710	25	2,03	1	1,52
Portugal	1,44	1050	8,6	1,2	1,6	1,76
Spain	1,43	2515	13,5	1,21	2,6	1,69
Finland	1,35	380	12,9	1,3	2,9	1,69
France	1,33	3223	15,3	1,3	2,9	1,26
Canada	1,29	2736	19	1,2	2,5	1,32
Poland	1,29	1500	10,2	0,64	0,5	1,11
Czech Republic	1,27	509	7,2	0,76	0,4	2,16
Malaysia	1,24	1064	4,5	0,78	0,6	2,58
New Zealand	1,19	383	14,4	1,28	0,8	2,05
Egypt	1,18	302	13,7	1,76	2,3	1,32
Russian Federation	1,16	1008	3,1	0,7	1,1	1,98
Slovakia	1,09	334	0	0,58	0	2,4
Pakistan	1,01	354	6	1,19	0,8	1,58
Iran	0,98	1046	8	0,61	0	1,34
Turkey	0,98	955	12,6	0,9	2,1	1,33
South Korea	0,97	3412	17,3	0,61	2,5	1,17
Brazil	0,94	1358	10,4	1,03	1,5	1,4
Romania	0,92	1009	2,5	0,76	0	1,24
Mexico	0,9	641	5,4	1,02	0,2	1,69
China	0,88	9611	8,5	0,8	1,2	1,07
Japan	0,85	4499	9,1	0,84	3,9	1,03
India	0,79	2647	8,5	0,47	1,7	1,36
Taiwan	0,76	1217	9,8	0,45	0,4	1,23

Kuva 67. Robotiikka ja koneautomaatio: maat järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2013-2017. Alle 300 julkaisua tuottaneet maat eivät ole mukana taulukossa. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen yhteistyön ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa. (Tiedot Scopuksesta: katselukerrat antavat tietoa julkaisun herättämästä kiinnostuksesta täydentämällä viittausmäärien välittämää informaatiota erityisesti uusimman aineiston osalta.) Normalisoiduissa mittareissa maailman keskiarvo = 1. Lähde: Scopus ja SciVal. Analysoituja julkaisuja yhteensä 77098. Tilanne: 30.5.2018

Tutkimuksen vaikuttavuus: tutkimusorganisaatiot

Eniten julkaisuja v. 2013-2017 tuottivat Ranskan kansallinen tutkimuskeskus CNRS ja Kiinan tiedeakatemia CAS. Mikäli mittarina pidetään normalisoitua viittausten määrää, niin julkaisujen korkein tieteellinen vaikuttavuus saavutettiin Harvardin yliopistossa ja MIT:ssä. Suhteellisesti eniten top 10% -huippujulkaisuja tuottivat Harvardin ja Cornellin yliopistot.

Eniten kansainvälistä yhteistyötä oli Hong Kongin kiinalaisella yliopistolla, Lausannen EPFL:llä ja Universitat Politecnica de Catalunyalla. Yritys-yliopisto –yhteistyötä oli eniten Berkeleyn ja Washingtonin yliopistoilla Yhdysvalloissa. Euroopassa eniten yritys-yhteistyötä oli Münchenin teknillisellä yliopistolla.

Tutkimuslaitos / Yliopisto	Normalisoitu viittaussindeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankiingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys -yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisutietojen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Harvard University	3,05	659	39	1,09	3,2	1,96
Massachusetts Institute of Technology	2,75	727	20,8	0,96	5,2	1,64
University of California at Berkeley	2,63	299	21,8	0,83	9,4	1,51
University of Washington	2,48	246	27,3	0,63	8,5	1,61
Carnegie Mellon University	2,46	685	15,4	0,87	4,2	1,35
ETH Zurich	2,45	553	21,7	1,41	3,4	1,7
Cornell University	2,39	276	38,9	0,92	2,5	1,59
EPFL Lausanne	2,33	352	25,2	1,79	0,9	1,73
University of Southern California	2,31	394	28,2	1,01	2,8	1,47
University of Pennsylvania	2,28	393	18,2	0,68	2	1,81
University of Pisa	2,27	278	20,8	1,12	1,4	2,39
University of Zurich	2,27	214	24,7	1,47	2,3	1,74
Stanford University	2,22	409	27	0,97	6,6	1,51
University of Michigan	2,19	429	32,2	0,55	2,8	1,51
Johns Hopkins University	2,1	523	29,7	0,83	5,2	1,51
South China University of Technology	2,1	286	16	0,94	0,7	2,11
Technical University of Munich	2,05	565	14,5	1,08	6	1,52
University of Oxford	2,01	225	24,2	1,47	1,3	1,27
California Institute of Technology	2	205	19,3	0,61	2,4	1,15
University of Illinois at Urbana-Champaign	2	195	24,1	0,5	4,6	1,33
Scuola Superiore Sant'Anna, Pisa	1,99	362	16,7	1,48	0,6	2,34
Queensland University of Technology	1,95	188	15,1	0,93	0,5	1,41
King's College London	1,94	276	33,2	1,61	1,1	1,61
Universitat Stuttgart	1,94	189	-	0,87	6,9	2,03
German Aerospace Center (DLR)	1,92	299	10,2	1,02	1,7	1,22
Universidade de Lisboa	1,88	314	7,1	1,47	2,9	1,65
University of Pittsburgh	1,88	270	30,6	0,88	0	1,45
Istituto Italiano di Tecnologia	1,82	548	16,8	1,32	1,3	1,71
Royal Institute of Technology	1,79	194	20,3	1,65	2,6	1,56
University of California at Los Angeles	1,77	196	31,8	0,91	6,6	1,74

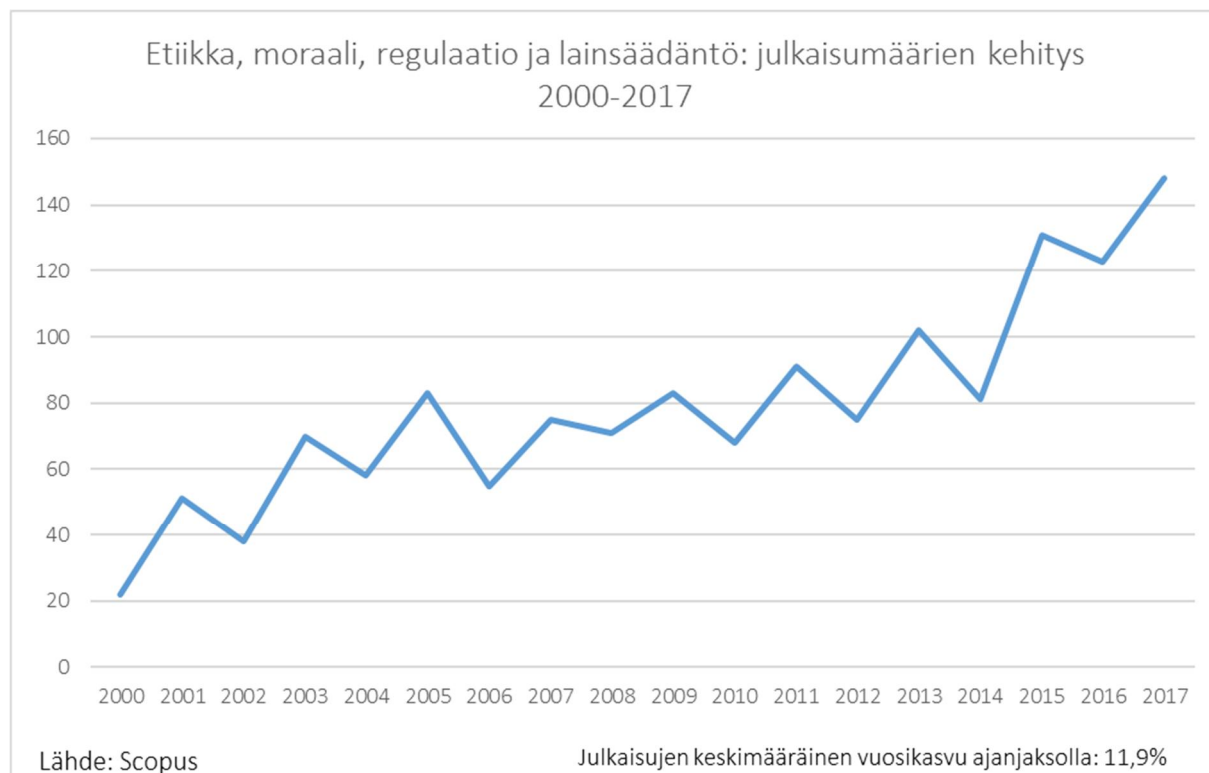
Kuva 68. Robotiikka ja koneautomaatio: merkittävimmät yliopistot, tutkimuslaitokset ja yritykset järjestettynä normalisoidun viittaussindeksin mukaan, julkaisuvuodet 2013-2017. Viittaussindeksin raja-arvo kuvassa: 1,77. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen ja tutkimuslaitos/yliopisto-yritysyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa (katselutiedot Scopusesta). Normalisoiduissa mittareissa maailman keskiarvo = 1. Analysoituja julkaisuja yhteensä 77098. Tilanne: 30.5.2018

1.3.11 Etiikka, moraali, regulaatio ja lainsäädäntö³⁵

Julkaisumäärät ja aiheet

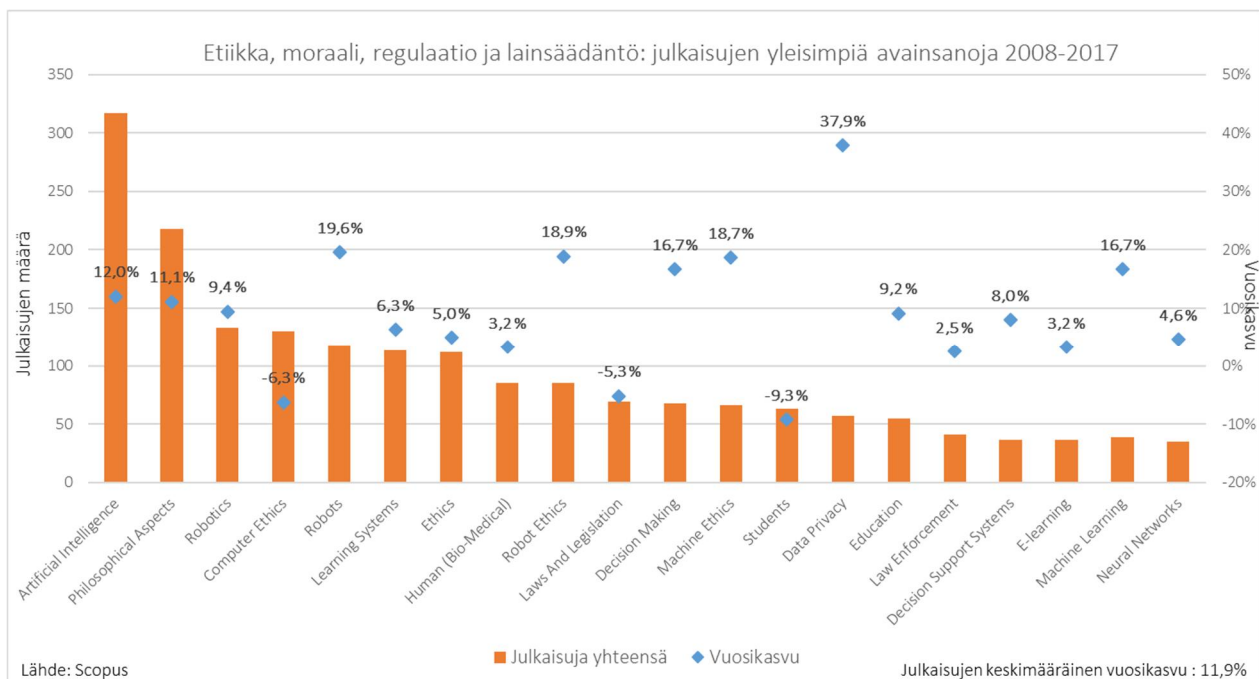
³⁵ Tutkimussuunnan rajauksessa käytettiin seuraavia termejä: computer ethics, machine ethics, robot ethics, ethical robots, machine morality, ethical machines, artificial morality, data ethics sekä seuraavia termejä yhdessä tekoälyaiheiden: ethics, moral obligations, morality, morals, regulation, legislation, laws, compliance, information privacy, data privacy, transparency, liability, legal, politics, public policy, ethical issues, data protection, GDPR, general data protection regulation, ethical aspects, ethical system, moral agency

Tekoälyn eettiset ja lainsäädännölliset reunaehdot ovat olleet viime aikoina paljon julkisuudessa. Tutkimuskirjallisuuden puolella tämä on näkynyt julkaisumäärien kasvuna. Toisaalta aiheeseen liittyvien julkaisujen tuotanto on kokonaisuudessaankin edelleen hyvin pientä.



Kuva 69. Tieteellisten julkaisujen määrä 2000-2017. Lähde: Scopus

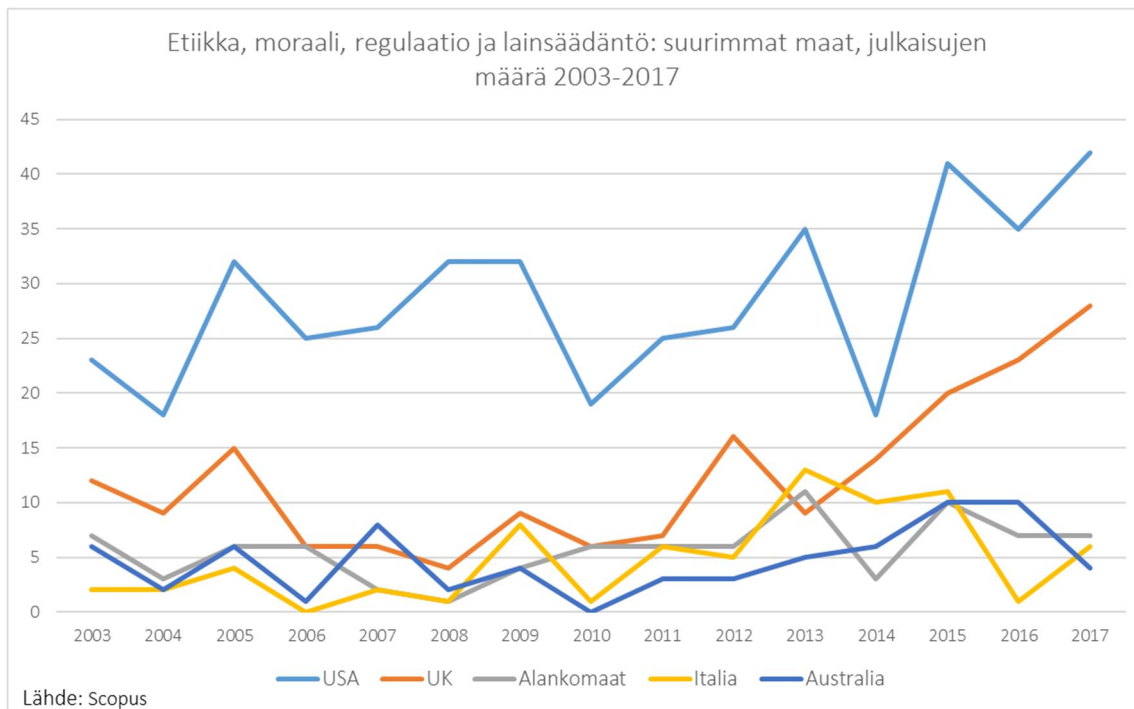
Tutkimussuunnan aiheista nousevat vuosina 2008-2017 esiin toisaalta robotiikkaan liittyvät eettiset ja lainsäädännölliset teemat ja toisaalta datan hallintaan ja (henkilödatan) suojaamiseen liittyvät kysymykset (data privacy, big data). Nopeimmin kasvavia aiheita olivat edellä mainittujen lisäksi mm. ihmisen ja robotin vuorovaikutus, koneoppiminen sekä robotti- ja kone-etiikka. Aivan viime vuosina on myös näistä erillinen "dataetiikka" noussut uutena aiheena tutkimuksen kohteeksi. On kuitenkin hyvä todeta, että kaikki kyseisten aiheiden julkaisumäärät ovat edelleen erittäin pieniä. Tästä johtuen myös julkaisumäärien vaihtelu on voimakkaampaa kuin suuremmilla tekoälyn tutkimussuunnilla tai sovellusalueilla.



Kuva 70. Keskeisiä tutkimusaiheita: julkaisumäärät ja vuosikasvu. Kaikkien alan julkaisujen vuosikasvu ko. ajanjaksolla oli 12,7%. Joukko sisällöltään yleisiä avainsanoja on jätetty kuvasta pois. Lähde: Scopus

Maat

Tekoälyyn, robotiikkaan ja isojen datajoukkojen hallintaan liittyvät eettiset kysymykset ovat nousseet esille lähinnä länsimaissa. Julkaisumäärät ovat kuitenkin niin pieniä, että kovin pitkälle meneviä johtopäätöksiä ei tästä voida vielä tehdä. Nopeinta kasvu viime aikoina on ollut suurimmista maista Iso-Britanniassa ja Yhdysvalloissa. Muista maista selkeää kasvua on ollut Saksassa ja Kiinassa. Kaikissa Pohjoismaissa on julkaistu alle 20 kyseisen tutkimussuunnan julkaisua tarkasteltavana ajanjaksona.



Kuva 71. Etiikka, moraali, regulaatio ja lainsäädäntö: suurimmat maat, julkaisumäärät 2003-2017. Lähde: Scopus

Tutkimusorganisaatiot: yliopistot, tutkimuslaitokset ja yritykset

Yliopistoista eniten tekoälyn, robotiikan ja datan hallinnan eettisiin ja lainsäädännöllisiin kysymyksiin keskittyviä julkaisuja tuottivat v. 2015-2018 (30.4.2018 mennessä) Oxfordin ja De Montfortin yliopistot Englannissa sekä Twenten, Amsterdamin ja Delftin yliopistot Alankomaissa. Yhdysvalloissa aktiivisin oli Carnegie-Mellonin yliopisto. Pohjoismaissa tekoälyn etiikkaan, moraaliin ja regulaatioon liittyviä julkaisuja tuottivat Uppsalan ja Oslon yliopistot, mutta määrät olivat molemmissa hyvin pieniä: vain kolme julkaisua v. 2015-2018 (v. 2018 tilanne: 04/2018).

Yleisesti voidaan sanoa, että tarkasteltaessa sekä maittain että toimijoittain ovat tekoälyn etiikkaan, moraaliin ja lainsäädäntöön liittyvien julkaisujen määrät hyvin pieniä. Käytännössä ei voida edes puhua varsinaisista osaamiskeskittymistä tässä yhteydessä.

Tutkimuksen vaikuttavuus: maat

Normalisoidun (suhteutetun) viittausindeksin perusteella Sveitsi, Belgia ja Saksa tuottivat v. 2012-2017 viitatuimpia tutkimusjulkaisuja tekoälyn etiikan, moraali, lainsäädännön ja regulaation tutkimuksessa. Top 10% -huippujulkaisuissa ilmestyneitä artikkeleita tuotti eniten Suomi. On kuitenkin samalla todettava, että julkaisumäärät olivat melko pieniä muualla kuin Yhdysvalloissa ja Iso-Britanniassa. Kansainvälistä tutkimusyhteistyötä tehtiin eniten Indonesiassa, Portugalissa ja Belgiassa.

Maa	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Top 10 -indeksi: huippujulkaisujen %-osuus kaikista julkaisuista (perustuen SCImago lehti-rankiingiin)	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Tutkimuslaitos-yritys - yhteistyön %-osuus kaikista julkaisuista	Normalisoitu julkaisujen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
Switzerland	2,48	12	0	1,69	8,3	3,2
Belgium	2,26	7	20	2,65	14,3	1,7
Germany	2,25	29	5,6	1,39	10,3	2,13
Italy	2,05	47	6,5	1,13	6,4	2,71
Austria	2	6	16,7	1,41	16,7	1,93
Netherlands	1,9	45	0	0,99	4,4	1,91
Sweden	1,88	10	11,1	1,17	10	1,5
India	1,85	6	0	0	0	4,69
Australia	1,71	39	4,8	1,17	7,7	2,2
Taiwan	1,69	7	0	0,5	0	1,87
Indonesia	1,64	6	0	3,94	0	1,52
United Kingdom	1,61	111	6	1,13	1,8	3,05
United States	1,56	191	15,4	0,83	3,7	1,98
Portugal	1,39	9	0	3,42	0	2,19
Denmark	1,35	8	0	0,46	0	1,72
Norway	1,35	7	20	1,23	0	1,81
New Zealand	1,24	8	0	2,03	0	2,39
Ireland	1,18	8	0	1,26	12,5	2,61
Canada	1,12	32	15,4	1,31	3,1	2,33
Greece	1,09	9	0	1,68	0	1,59
Iran	0,96	6	20	1,86	0	2,71
Israel	0,9	7	0	1,59	0	1,38
Spain	0,86	10	0	1,44	0	2,08
Japan	0,85	19	14,3	1,15	0	1,15
France	0,82	20	20	1,34	10	1,51
Finland	0,67	6	25	1,45	0	1,87
Turkey	0,6	7	20	0,95	0	2,36
China	0,51	22	7,7	0,88	0	5,13

Kuva 72. Etiikka, moraali, regulaatio ja lainsäädäntö: maat järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2012-2017 (viittausindeksin raja-arvo kuvassa: 1). Alle 6 julkaisua tuottaneet maat eivät ole mukana taulukossa. Muut mittarit: julkaisujen määrä, huippujulkaisujen (top 10%) osuus, kansainvälisen yhteistyön ja tutkimuslaitos/yliopisto-yritys-yhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa. (Tiedot Scopusesta: katselukerrat antavat tietoa julkaisun herättämästä kiinnostuksesta täydentämällä viittausmäärien välittämää informaatiota erityisesti uusimman aineiston osalta.)

Normalisoiduissa mittareissa maailman keskiarvo = 1. Lähde: Scopus ja SciVal. Julkaisuja yhteensä 662. Tilanne: 4.5.2018

Tutkimuksen vaikuttavuus: tutkimusorganisaatiot

Amerikkalaiset yliopistot hallitsevat tilastoa, kun tarkastellaan julkaisujen viittausmääriä. Eurooppalaista yliopistoista kärjessä olivat 2012-2017 Liverpoolin ja Oxfordin yliopistot, joilla oli myös keskimääräistä enemmän julkaisuja tutkimusteemassa. On kuitenkin samalla todettava, että julkaisumäärät ovat kaikilla toimijoilla liian pieniä, jotta pidemmälle meneviä johtopäätöksiä olisi mahdollista tehdä.

Tutkimuslaitos / yliopisto	Normalisoitu viittausindeksi (FWCI), maailman keskiarvo tutkimusalalla = 1	Julkaisujen määrä	Normalisoitu kansainvälisen yhteistyön indeksi (maailman keskiarvo = 1)	Normalisoitu julkaisujen katseluindeksi (Scopus-katselukerrat), keskiarvo = 1
University of Maryland	6,3	5	0,85	1,68
Brown University	4,49	5	0	2,37
Tufts University	3,73	6	0	1,87
Rensselaer Polytechnic Institute	3,64	6	0	0,67
Northern Illinois University	3,16	5	1,66	3,19
University of Liverpool	2,6	11	1,15	1,63
University of Oxford	2,52	13	1,72	3,27
Delft University of Technology	2,42	6	1,84	3,61
Utrecht University	2,37	5	1,69	1,84
University of Louisville	2,36	5	0	2,44
Carnegie Mellon University	2,35	7	1,45	1,9
NICTA	2,22	6	1,96	1,84
University of Groningen	2,07	6	1,38	1,83
Georgia Institute of Technology	1,95	10	0	4,24
Universidade Nova de Lisboa	1,89	5	4,73	1,39
University of Indonesia	1,89	5	4,73	1,39
University of Warwick	1,88	5	0,44	1,46
University of Pittsburgh	1,73	6	1,31	2,12
University of Sheffield	1,71	5	0	1,84
University of Turin	1,61	10	1,15	3,99
University of Washington	1,56	6	1,64	2,65
De Montfort University	1,48	13	0,82	1,46
University of Twente	1,42	15	0,77	1,95

Kuva 73. Etiikka, moraali, regulaatio ja lainsäädäntö: merkittävimmät yliopistot, tutkimuslaitokset ja yritykset järjestettynä normalisoidun viittausindeksin mukaan, julkaisuvuodet 2012-2017. Taulukossa mukana vain toimijoita, joilla ollut vähintään viisi julkaisua tarkastelujaksona. Muut mittarit: julkaisujen määrä, kansainvälisen tutkimusyhteistyön osuus sekä julkaisutietojen katselukerrat tietokannassa (katselutiedot Scopusesta). Normalisoiduissa mittareissa maailman keskiarvo = 1. Lähde: Scopus ja SciVal. Julkaisuja yhteensä 662. Tilanne: 7.5.2018

1.4 Patentointi

Kuten tutkimusjulkaisut ovat myös tekoälyyn keskittyvät patentit mahdollista ryhmitellä eri tavoin. Alla on tarkasteltu patentointia kolmesta näkökulmasta: tekoälyyn yleisesti tulkittuna liittyvää patentointia ja erikseen koneoppimisen ja hermoverkkojen patentointia. Koneoppiminen ja hermoverkkoteknologiat ovat valittu tarkasteluun niihin viime vuosina kohdistuneen voimakkaan kiinnostuksen takia.

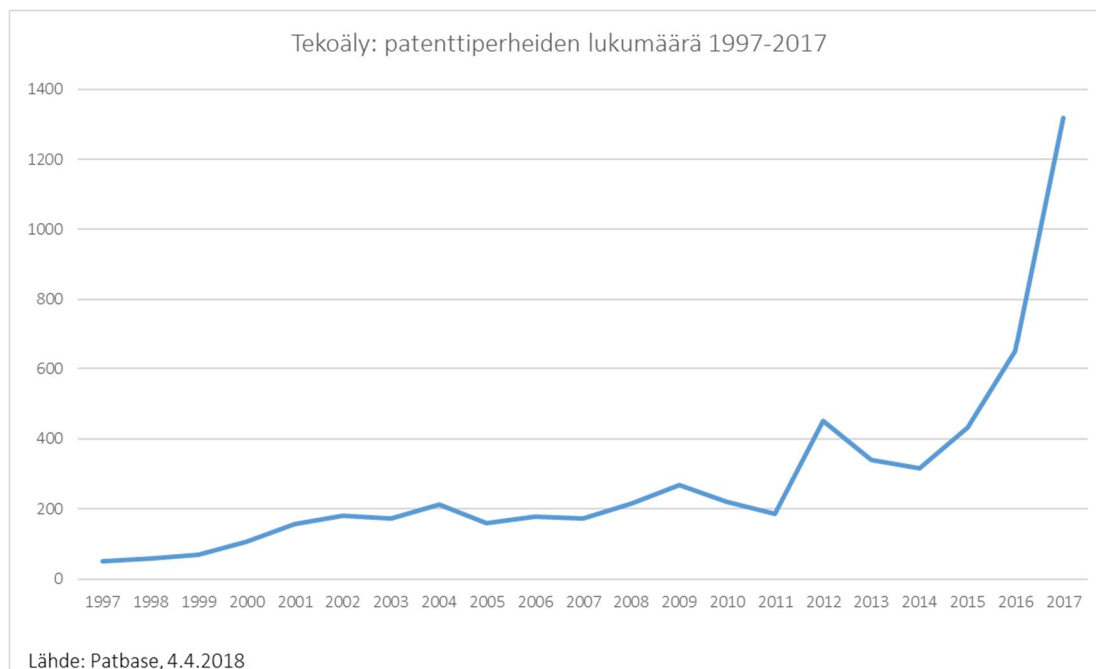
Tilastoinnissa kiinnitetään ohessa erityistä huomiota patenttiperheiden lukumääriin. Yksi patenttiperhe liittyy yhteen ja samaan *keksintöön* kuuluvia julkaisuja, joita syntyy mm. hakemalla patenteja eri maissa ja juridisissa toimialueissa tai täydentämällä alkuperäistä patenttihakemusta uusilla asiakirjoilla.

Kiinnostavaa on huomata, että tutkimusjulkaisujen tavoin myös tekoälyyn liittyvien patenttien määrä kääntyi selvään nousuun vuoden 2014 vaiheilla. Tärkein yksittäinen tekijä lienee juuri oppiviin järjestelmiin ja menetelmiin liittyvien patenttihakemusten määrän nopea lisääntyminen.

1.4.1 Tekoäly

Tekoälyyn yleisesti liittyvän patentoinnin erottaminen omaksi patenttiryhmäksi ei ole olemassa yksinkertaista ratkaisua. Seuraavassa on patenttien rajaamisessa käytetty pelkästään tekoälytermin esiintymistä patenttien otsikoissa tai tiivistelmissä. Suomalaisia toimijoita ei tässä katsauksessa erikseen selvitetty.

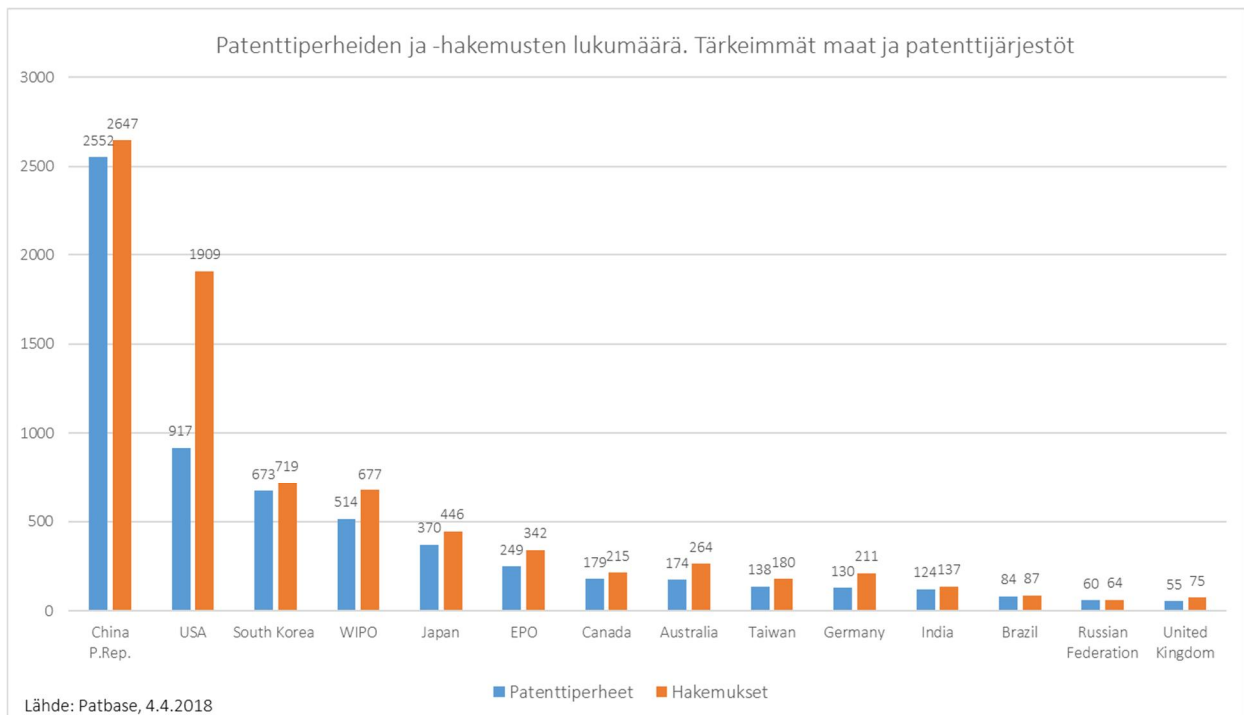
Kasvu patenttiperheiden julkaisuissa on ollut erittäin nopeaa viime vuosina. Keskimääräinen vuosikasvu on ollut vuoden 2011 jälkeen 38,4%. Erityisen nopeaa kasvu on ollut kahden viime vuoden aikana. Kasvu näyttää jatkuvan: vuoden 2018 kolmen ensimmäisen kuukauden aikana julkaistiin 540 uutta patenttiperhettä, kun vuonna 2017 niitä julkaistiin yhteensä 1317.



Kuva 74. Tekoäly yleisesti tulkittuna: patenttiperheet 1997-2017. Lähde: PatBase³⁶

Maista suurimmat patenttoijat olivat Kiina ja Yhdysvallat. Myös Etelä-Korea on merkittävä toimija omassa maassa haettujen patenttien (patenttiperheiden) määrällä mitaten. Eniten hakemuksia suhteessa voimassa olevien patenttiperheiden määrään oli suurista maista Yhdysvalloilla, mikä kertoo mm. Yhdysvaltojen muista poikkeavasta patentoinnin kulttuurista ja lainsäädännöstä.

³⁶ Tiedot haettu PatBase-tietokannasta hakemalla patentin otsikosta ja tiivistelmästä termillä "artificial intelligence"



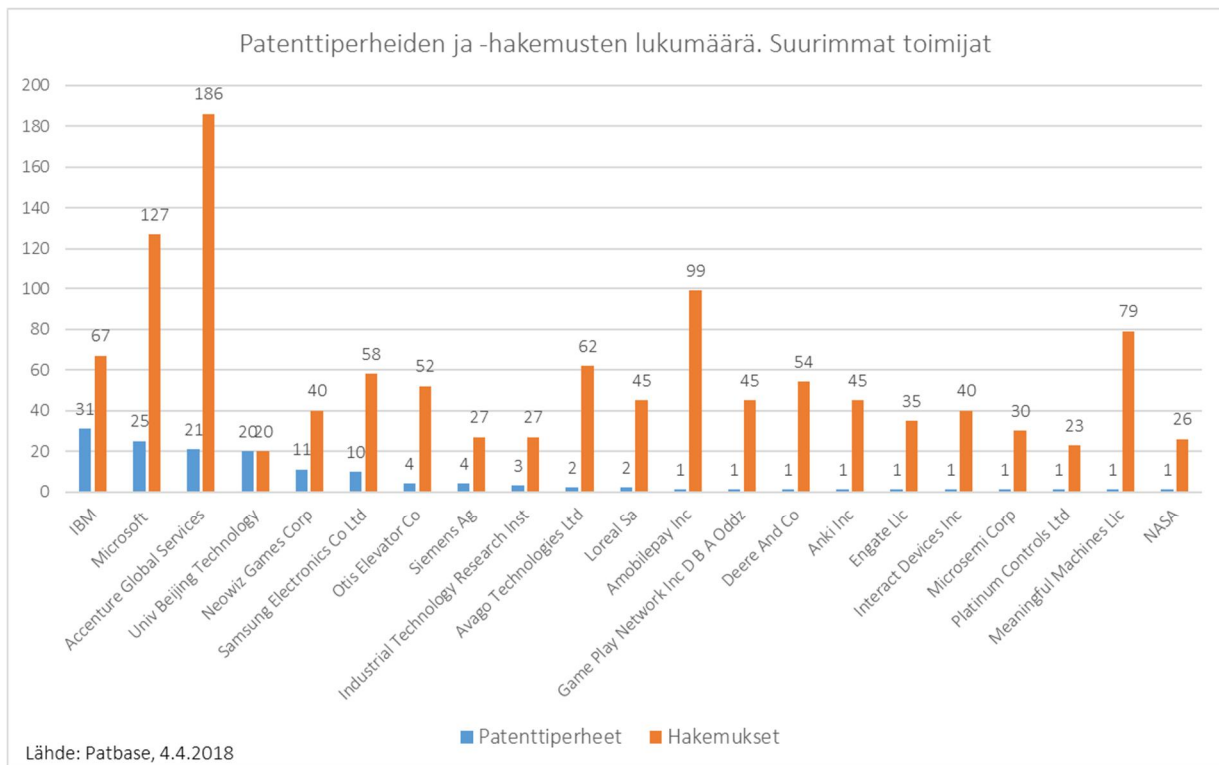
Kuva 75. Tekoäly yleisesti. Aktiivisimmat maat patentoinnissa. Lähde: PatBase

Mikäli tarkastelujaksoksi valitaan vain aivan viimeiset vuodet, niin Kiinan määrällisesti johtava asema korostuu patentoinnissa entisestään:

Maa / lainkäyttöalue	Patenttiperheet	Hakemukset
China P. Rep	1914	1980
United States of America	355	942
WIPO	238	345
South Korea	189	194
Japan	100	121
India	63	71
European Patent Office	57	105
Canada	47	66
Australia	37	56
Taiwan	35	46
Turkey	24	24
Germany	22	35
United Kingdom	20	29

Taulukko 1. Eniten uusia patenteja (patenttiperheitä) patentin prioriteettipäivän mukaan: maat tai toimialueet 2013-2018. Status 5.4.2018. Lähde: Patbase

Aktiivisimmat tutkimusorganisaatiot ovat patenttiperheiden osalta IBM ja Microsoft. Suurimmista yrityksistä Accenturella oli hakemusten lukumäärän suhde patenttiperheiden määrään poikkeuksellisen suuri.



Kuva 76. Tekoäly yleisesti: aktiivisimmat tutkimusorganisaatiot patentoinnissa. Lähde: PatBase

Jos tarkasteluajanjaksoksi otetaan vain vuodet viimeiset vuodet, niin kuva muuttuu radikaalisti. Listalla on IBM:n lisäksi enää lähinnä kiinalaisia yrityksiä ja tutkimuslaitoksia:

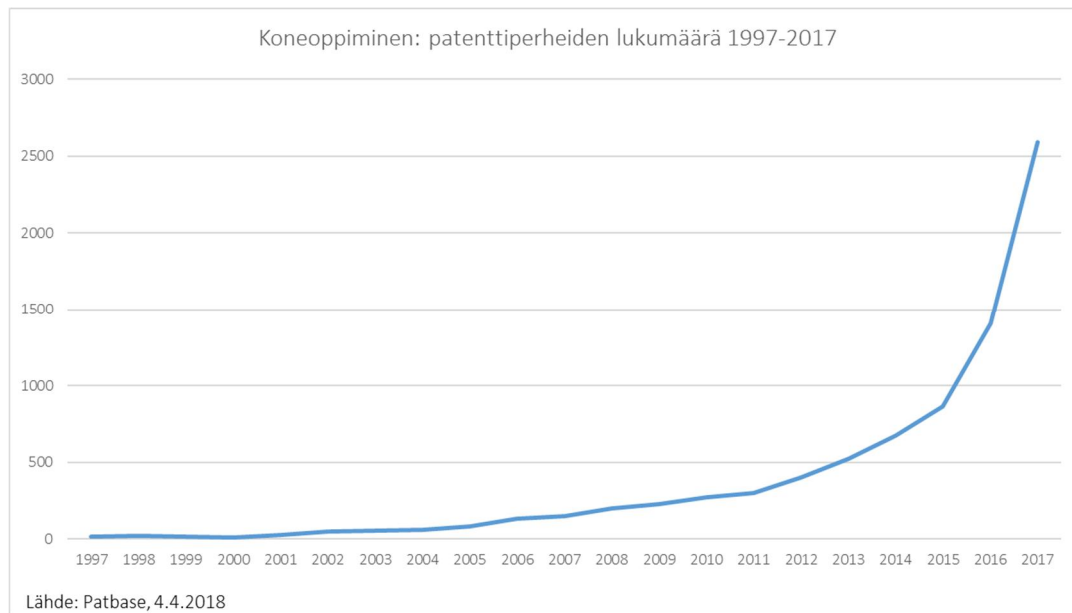
Yritys / tutkimuslaitos	Patenttiperheet	Hakemukset
Beijing Baidu Netcom Sci And Tec	133	142
Baidu Online Network Tech Beijing Co Ltd	89	114
Hu Ming Jian	33	33
Dalian Roiland Tech Co Ltd	26	26
Miyazaki Hiroaki	16	24
IBM	13	25
State Grid Corp China	13	13
Microblog Consulting Co Ltd Zhuji	12	12
Cangnan Boya Tech Co Ltd	11	11
Chen Yong	11	13
Guangdong Oppo Mobile Telecom.	10	10
Iteris Inc	9	25
Lmeca Co Ltd	9	25
Nanjing Automation System Co Ltd Baina	9	9
Shenzhen Gowild Robotics Co Ltd	9	17
Univ Beijing Technology	9	9
Univ Dalian Tech	9	9

Taulukko 2. Eniten uusia patenteja (patenttiperheitä) 2013-2018 patentin prioriteettipäivän mukaan. Status 5.4.2018. Lähde: Patbase

1.4.2 Koneoppiminen

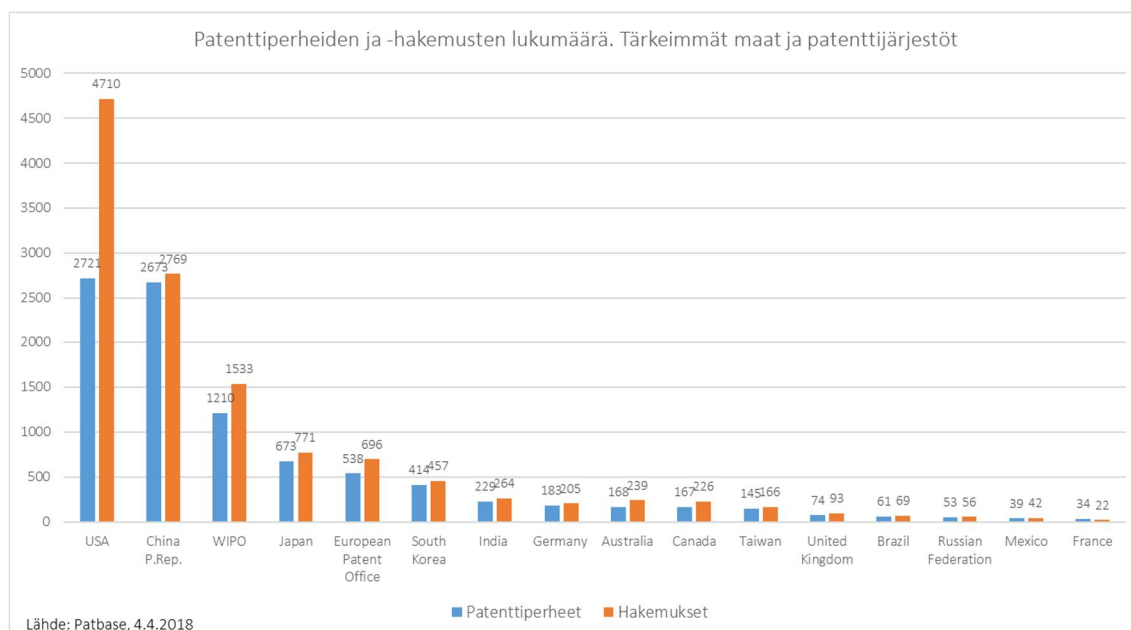
Koneoppimiseen liittyvä patentointi on kasvanut erittäin nopeasti viime vuosina. Keskimääräinen vuosikasvu on ollut vuoden 2011 jälkeen peräti 43,2%. Sama trendi näyttäisi myös jatkuvan. Ensimmäisten kolmen

kuukauden aikana v. 2018 julkaistiin yli tuhat uutta patenttiperhettä koneoppimisen alalla, kun v. 2017 niitä julkaistiin yhteensä 2590.



Kuva 77. Koneoppimiseen liittyvät patenttiperheet 1997-2017. Lähde: PatBase ³⁷

Kuten yleensäkin tekoälyn alueella suurimmat maat koneoppimiseen liittyvässä patentoinnissa olivat Yhdysvallat ja Kiina. Eniten hakemuksia suhteessa patenttiperheiden määrään oli suurista maista Yhdysvalloilla.



Kuva 78. Koneoppiminen: aktiivisimmat maat patentoinnissa. Lähde: PatBase

Mikäli tarkastellaan vain viimeisiä vuosia patentin prioriteetti- eli etusijaisuuspäivän mukaan, niin Kiina on ohittanut Yhdysvallat patenttiperheiden lukumäärässä. Sen sijaan hakemuksia on jätetty Yhdysvalloissa selvästi eniten.

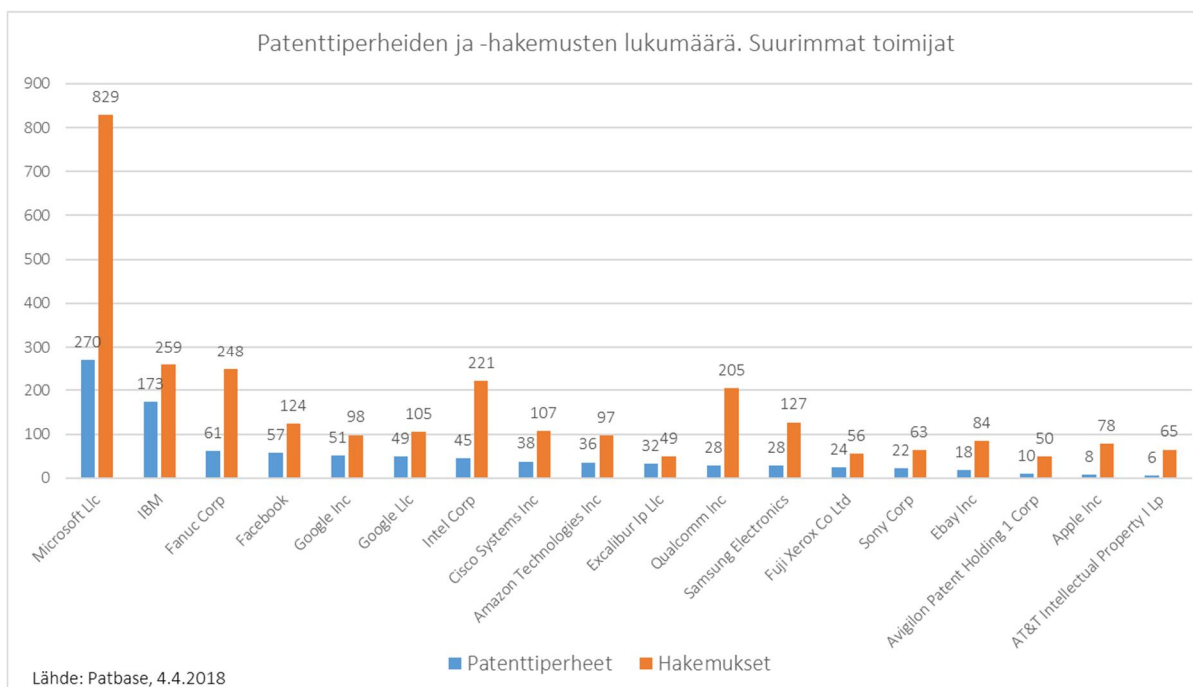
Maa / lainkäyttöalue	Patenttiperheet	Hakemukset
----------------------	-----------------	------------

³⁷ Tiedot haettu PatBase-tietokannasta hakemalla patentin otsikosta ja tiivistelmästä termillä "machine learning"

China P.Rep.	2296	2376
United States of America	1904	3610
WIPO	952	1238
Japan	416	485
European Patent Office	363	497
South Korea	271	305
India	173	205
Germany	128	138
Canada	116	160
Australia	105	154
Taiwan	98	111
United Kingdom	57	70
Brazil	40	47
Russian Federation	34	37

Taulukko 3. Eniten uusia patentteja (patenttiperheitä) patentin prioriteettipäivän mukaan: maat tai toimialueet 2013-2018. Status 5.4.2018. Lähde: Patbase

Aktiivisimmat tutkimusorganisaatiot ovat osittain samat suuret amerikkalaiset yritykset kuin julkaisujenkin puolella Microsoft, IBM ja Google. Eniten hakemuksia suhteessa patenttiperheiden määrään oli suurista yrityksistä Qualcommilla.



Kuva 79. Koneoppiminen: aktiivisimmat tutkimusorganisaatiot patentoinnissa. Lähde: PatBase

Jos tarkasteluajanjaksoksi otetaan vain viimeiset vuodet patentin prioriteettipäivän perustella, niin kuva patenttinhaltijoista pysyy pääpiirteissään samana. Uusina yrityksinä kymmenen eniten patentteja hakeneen joukkoon nousevat Alibaba, LinkedIn ja Adobe Systems.

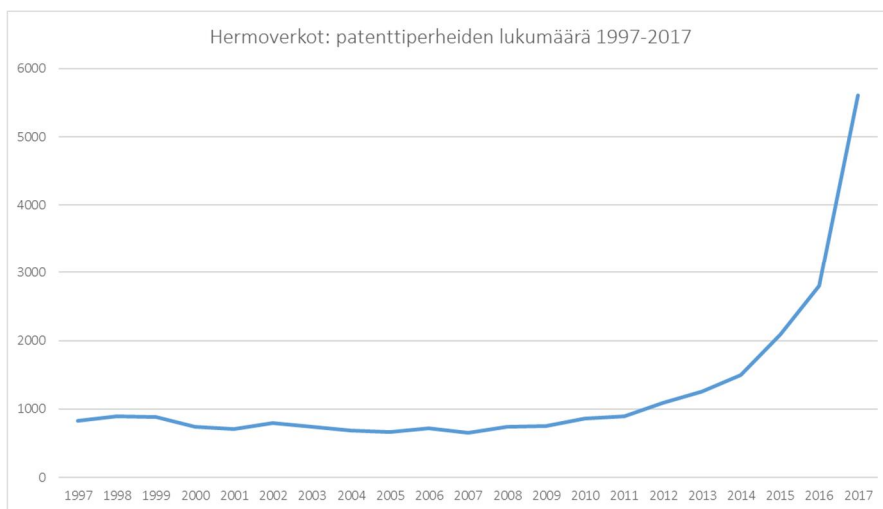
Yritys / tutkimuslaitos	Patenttiperheet	Hakemukset
IBM	134	191
Microsoft	133	470
Google	69	150
Fanuc Corp	61	248

Facebook Inc	47	91
Cisco Systems Inc	38	107
Intel Corp	38	204
Alibaba Group Holding Ltd	36	63
Linkedin Corp	33	50
Adobe Systems Inc	32	72
Amazon Technologies Inc	28	88
Univ Zhejiang	28	28
Dalian Roiland Tech Co Ltd	27	27
Gree Electric Appliances Inc Of Zhuhai	27	27
Qualcomm Inc	26	190
Univ Tsinghua	26	30
Fujitsu Ltd	24	47
Siemens Healthcare Gmbh	23	63
Samsung Electronics Co Ltd	22	116

Taulukko 4. Eniten uusia patenteja (patenttiperheitä) 2013-2018 patentin prioriteettipäivän mukaan. Status 5.4.2018. Lähde: Patbase

1.4.3 Hermoverkot

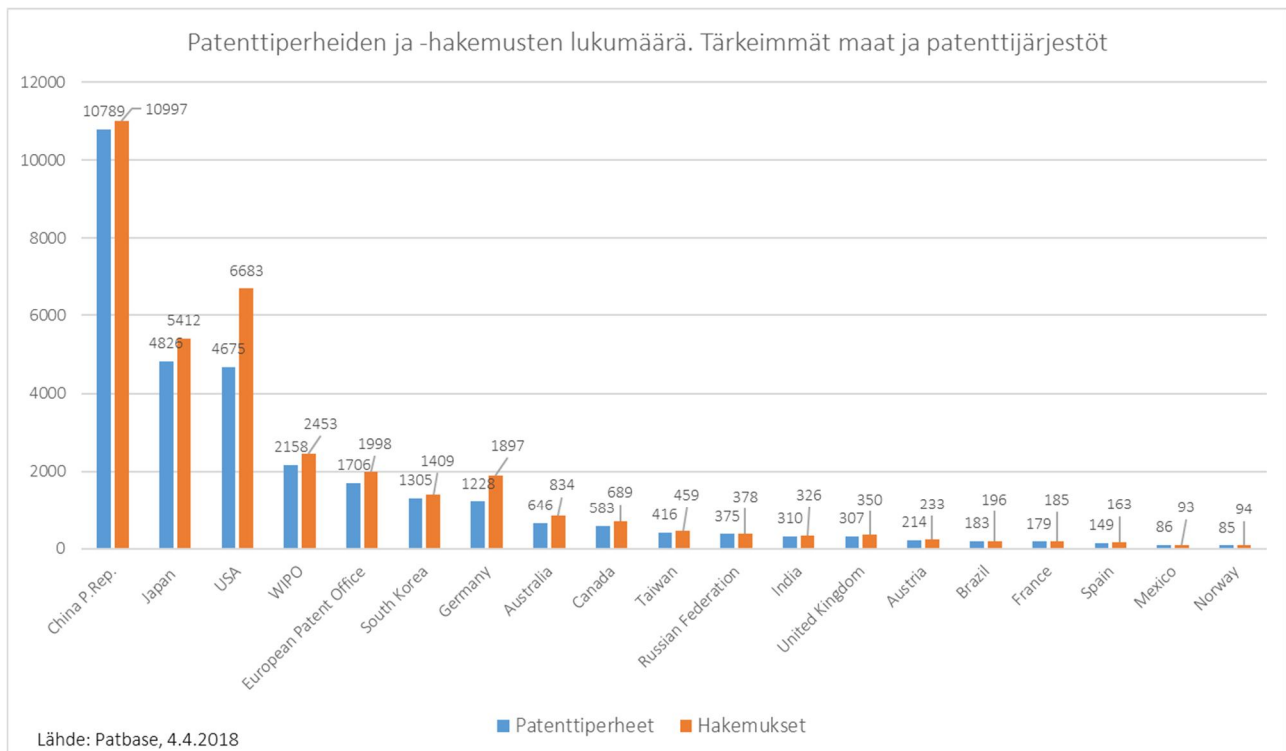
Mm. koneoppimisen taustalla usein käytettyjen hermoverkkoteknologioiden ja -ratkaisujen patentointi on lisääntynyt koko 2010-luvun, erityisesti muutaman viime vuoden aikana. Keskimääräinen vuosikasvu on ollut vuoden 2011 jälkeen 36%. Nopea kasvu näyttäisi myös jatkuvan, sillä vuoden 2018 ensimmäisellä neljänneksellä julkaistiin yli 2000 patenttiperhettä, kun v. 2017 niitä julkaistiin yhteensä 5607. Myös patenttien kokonaismäärässä hermoverkot on merkittävä teknologia-alue. Patenttiperheitä julkaistiin v. 2017 noin kaksikertaisesti verrattuna esimerkiksi koneoppimiseen liittyviin patenttiperheisiin.



Kuva 80. Koneoppimiseen liittyvät patenttiperheet 1997-2017. Lähde: PatBase ³⁸

Tosin kuin koneoppimisen tai yleisesti tekoälyn patentoinnissa Japani oli Yhdysvaltojen sijasta kahden yleisimmän patenttihakemusten jättömaan joukossa. Patenttiperheiden lukumäärillä mitattuna Kiinan ero muihin maihin oli poikkeuksellisen suuri. Eniten hakemuksia suhteessa voimassa olevien patenttiperheiden määrään oli suurista maista Saksalla ja Yhdysvalloilla.

³⁸ Tiedot haettu PatBase-tietokannasta hakemalla patentin otsikosta ja tiivistelmästä termillä "neural network**".



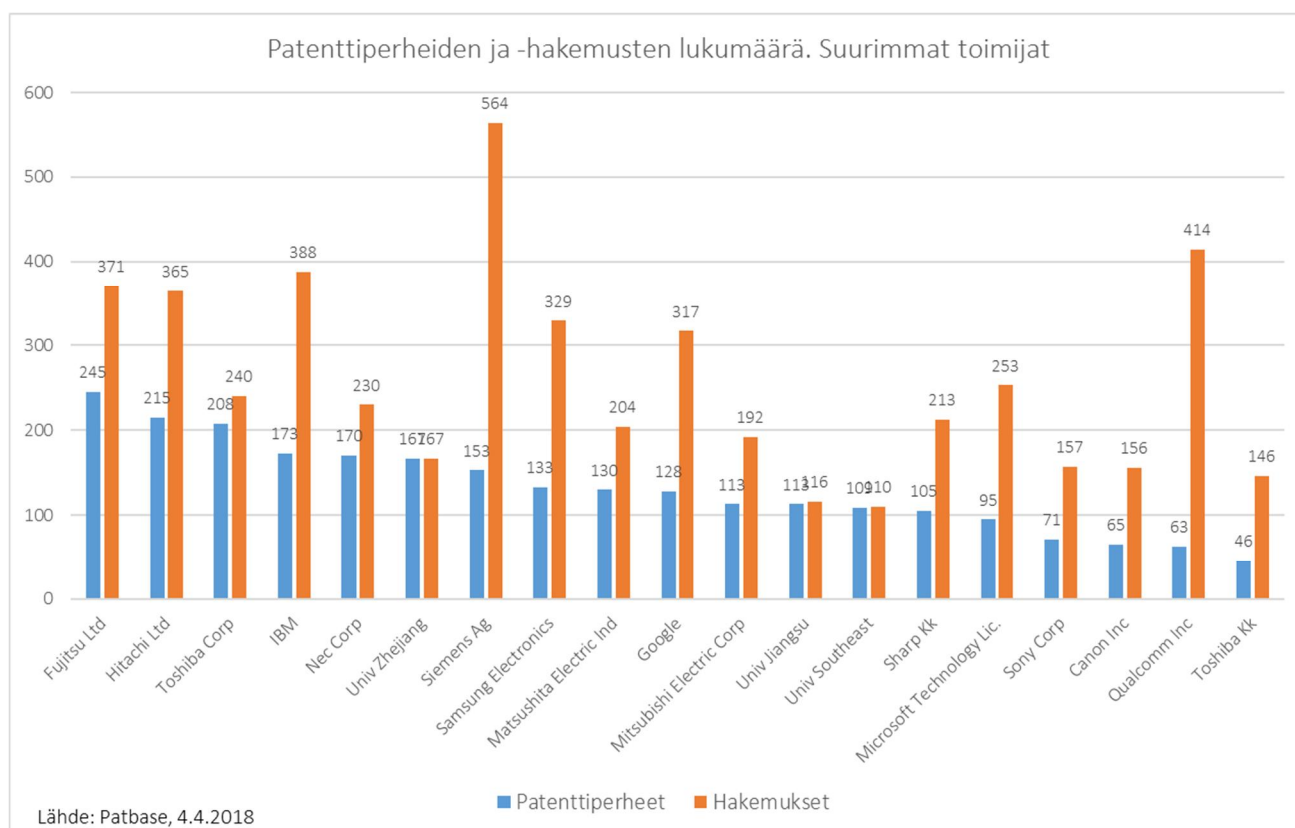
Kuva 81. Hermoverkot: aktiivisimmat maat patentoinnissa. Lähde: PatBase

Kun tarkastellaan vuosia 2013-, niin Kiinan välimatka Yhdysvaltoihin on patenttiperheiden määrällä mitaten kasvanut entisestään. Japani on tässä tarkastelussa pudonnut Yhdysvaltojen ja Etelä-Korean taakse.

Maa / lainkäyttöalue	Patenttiperheet	Hakemukset
China P. Rep.	8814	8967
USA	1462	2233
WIPO	915	1041
South Korea	384	404
European Patent Office	367	458
Japan	330	405
India	168	176
Russian Federation	135	135
Germany	127	144
Taiwan	98	115
Canada	88	116
United Kingdom	72	84

Taulukko 5. Eniten uusia patenteja patentin prioriteettipäivän mukaan: maat tai toimialueet 2013-2018. Status 5.4.2018. Lähde: Patbase

Ilman aikarajausta tarkasteltaessa olivat aktiivisimmat yritykset japanilaiset Fujitsu, Hitachi ja Toshiba. Eniten hakemuksia suhteessa voimassa olevien patenttiperheiden määrään oli suurista yrityksistä Siemensillä ja Qualcommilla.



Kuva 82. Hermoverkot: aktiivisimmat tutkimusorganisaatiot patentoinnissa. Lähde: PatBase

Mikäli tarkastelujaksoksi valitaan vain viimeiset vuodet, niin kuva patenttien omistajista muuttuu huomattavasti. Google on tässä tarkastelussa suurin toimija. Googlen, IBM:n ja Microsoftin lisäksi listalle ei mahdu juuri muita kuin kiinalaisia yliopistoja. Qualcommilla on selkeästi eniten patenttihakemuksia suhteessa varsinaisiin patenttiperheisiin, mikä kertonee myös kyseisen yrityksen aktiivisuudesta aihepiirin ympärillä.

Yritys / tutkimuslaitos	Patenttiperheet	Hakemukset
Google	175	443
State Grid Corp China	127	133
Univ Zhejiang	123	123
Univ Tianjin	106	107
Univ South China Tech	105	107
Univ Tsinghua	104	108
IBM	95	186
Univ Beijing Technology	94	95
Univ Xidian	90	90
Univ Southeast	80	81
Univ Jiangsu	69	71
Univ Zhejiang Technology	69	69
Univ Beihang	64	65
Microsoft	58	136
Univ Chongqing Sci And Tech	58	58
Guangdong Univ Of Technology	57	57

Taulukko 6. Eniten uusia patenteja (patenttiperheitä) 2013-2018 patentin prioriteettipäivän mukaan. Status 5.4.2018. Lähde: Patbase

1.5 Yhteenveto

Tässä selvityksessä on tarkasteltu tekoälyn julkaisukenttää eri näkökulmista. Tutkimussuunnat poikkeavat huomattavasti toisistaan sekä julkaisumäärien että niissä tapahtuvien muutosten osalta. Pienimmän - etiikkaan ja regulaatioon liittyvän - tutkimusalan julkaisumäärä oli vuosina 2000-2017 vain murto-osa suurimman, koneoppimisen, julkaisumäärästä.

Sovellettaessa laajaa tulkintaa tekoälyn sisällöstä, voidaan sitä käsittelevien julkaisujen määrän arvioida kasvaneen 2000-luvulla n. 9%:n vuosivauhtia. Tämä on hyvin lähellä tutkimusjulkaisujen vuosikasvua koko tiedekentässä.³⁹ Tekoälyn osalta kasvu on kuitenkin vaihdellut havainnoinnin ja tilannetietoisuuden 7%:n ja tekoälyn laskentaympäristöjen, alustojen ja palveluiden n. 15%:n keskimääräisen vuosikasvun välillä.

Tekoälyn tutkimussuunta	Julkaisuja yhteensä 2000-2017	Osuus	Vuosittainen kasvu 2000-2017
Koneoppiminen	815257	30,8 %	11,5 %
Havainnointi ja tilannetietoisuus	565819	21,4 %	7,1 %
Data-analyysi	388298	14,7 %	13,1 %
Digitaidot työelämässä, ongelmanratkaisu ja laskennallinen luovuus	272157	10,3 %	10,8 %
Robottiikka ja koneautomaatio	190918	7,2 %	11,4 %
Luonnollinen kieli ja kognitio	172268	6,5 %	7,2 %
Vuorovaikutus ihmisen kanssa	159287	6,0 %	13,3 %
Järjestelmätaso ja systeemivaikutukset	51933	2,0 %	10,4 %
Tekoälyn laskentaympäristöt, alustat ja palvelut, ekosysteemit	28710	1,1 %	14,8 %
Etiikka, moraali, regulaatio ja lainsäädäntö	1425	0,1 %	11,9 %

Tutkimussuunnan julkaisuvolymit, osuus ja julkaisumäärän vuosittainen kasvu vuosina 2000-2017. Lähde: Scopus

Erityisen nopeaa julkaisumäärien kasvu on ollut aivan viime vuosina, 2014-2017. Nopeaa kasvua selittää erityisesti koneoppimiseen ml. ns. syväoppimiseen ja keinotekoisin hermoverkkoihin liittyvä - erityisesti soveltava - tutkimus. Joitain merkkejä on kuitenkin siitä, että nopeimman kasvun vaihe saattaa olla taitumassa.

Tekoälyn tutkimukselle on tyypillistä soveltavan tutkimuksen vahva asema. Kysymykseen tekoälyn perustutkimuksen ja soveltavan tutkimuksen erottamisen hyödyllisyydestä ei tässä selvityksessä otettu kantaa. Esimerkiksi nopeimmin kasvavat tutkimusjulkaisujen aiheet viime aikoina, kuten edellä mainitut CNN-, DNN- ja RNN-hermoverkot, liittyvät usein valtaosaltaan pikemminkin tekoälyn soveltamiseen (lääke- ja terveystiede, liikenne, logistiikka, kiertotalous, ilmastotutkimus, ruokaketjun hallinta jne.) kuin tekoälyn tutkimukseen sinänsä. Tämä ei kuitenkaan tarkoita, etteikö myös kyseinen soveltava tutkimus silti voisi viedä myös tekoälyn perustutkimusta eteenpäin.

Kun tarkastellaan julkaisutuotannon kehitystä maittain ja toimijoittain, niin mikään yllätys ei ole, että kiinalaisten yliopistojen ja tutkimuslaitosten julkaisumäärät ovat lisääntyneet tarkasteltavana ajanjaksona nopeasti. Kiina ei kuitenkaan ole ainut tekoälytutkimustaan nopeasti kasvattanut maa. Nopeimmin tekoälyn tutkimuspanostuksia lisänneitä maita julkaisumäärillä mitaten löytyy etenkin Etelä- ja Kaakkois-Aasiasta: Intia, Bangladesh, Indonesia, Thaimaa, Vietnam ja Pakistan. Muita nopeasti julkista tutkimustoimintaansa vuoden 2012 jälkeen lisänneitä ja samalla julkaisuvolyymltaan merkittäviä maita olivat mm. Venäjä, Etelä-

³⁹ Ks. <http://blogs.nature.com/news/2014/05/global-scientific-output-doubles-every-nine-years.html> ja <https://blog.frontiersin.org/2017/03/06/160-million-papers-and-counting-the-worlds-information-deluge/>

Afrikka, Saudi-Arabia ja Turkki. Euroopan Unionin maista esimerkiksi Puola on nykyisin jo merkittävä toimija sekä julkaisuvolyymilla että julkaisumäärien kasvunopeudella mitaten.

Yleensä maiden tutkimuspanostukset painottuvat eri lailla eri tekoälyn tutkimussuunnissa. Esimerkiksi Intia on noussut USA:n ja Kiinan jälkeen kolmanneksi suurimmaksi tekoälyn julkaisumaaksi erityisesti koneoppimiseen, data-analyysiin ja havainnointiin liittyvän tutkimustyön myötä. Useissa tekoälyn eri tutkimussuunnissa on havaittavissa vastaavaa uusien maiden esiin nousua.

Tutkimusorganisaatioiden tasolla Kiinan Tiedeakatemia CAS ja Ranskan kansallinen tutkimuskeskus CNRS ovat tekoälytutkimuksessa omassa kokoluokassaan. CAS ja CNRS toimivat kuitenkin omissa maissaan eräänlaisina tutkimuksen sateenvarjo-organisaatioina ja osittain myös tutkimuksen rahoittajina, joten niiden - samoin kuin esimerkiksi Italian CNR:n - vertaaminen yksittäisiin yliopistoihin, ei ole mielekäästä.

Eräitä Kiinan ja Yhdysvaltojen suuria yliopistoja ja tutkimuslaitoksia lukuun ottamatta tutkimusorganisaatiot yleensä erikoistuvat omalle tutkimusalalleen. Näin on myös tekoälyn ollessa kyseessä. Jos kuitenkin katsotaan tekoälytutkimusta kokonaisuudessaan, niin kiinalaisten Tsinghua- ja Beihang-yliopistojen ja Harbinin Teknologiainstituutin lisäksi kannattaa nostaa esiin USA:n Carnegie Mellon –yliopisto ja MIT sekä Nanyangin teknillinen yliopisto Singaporessa. Euroopassa Universite Paris Saclay, Ranskan INRIA-tutkimuskeskus, Technische Universität München ja Oxfordin yliopisto on syytä mainita erityisen merkittävänä toimijoina omissa maissaan. Pohjoismaissa Aalto-yliopisto ja Teknologian tutkimuskeskus VTT olivat vuosina 2013-2017 korkeimmalla olevat tutkimusorganisaatiot, kun kriteerinä käytettiin huippujulkaisujen osuutta kaikista vertaisarvioituista tutkimusjulkaisuista.⁴⁰

Julkaisumäärien lisäksi on tärkeää pyrkiä arvioimaan myös tutkimustulosten tieteellisiä ansioita. Julkaisujen tieteellistä vaikuttavuutta arvioitaessa on olennaista tarkastella rinnakkain eri mittareita. Käytettäessä mittarina suhteutettuja viittausmääriä olivat maista kärjessä useilla tutkimusaloilla mm. Singapore, Sveitsi ja Hong Kong. Uusiin nouseviin maihin kuului Saudi-Arabia, jolla on julkaisumäärien lisäksi nykyisin näyttöä myös tutkimuksen tasosta. Suurista julkaisujen tuottajamaista Yhdysvallat, Iso-Britannia, Australia ja Kanada ovat myös julkaisujen vaikuttavuuden näkökulmasta (edelleen) merkittäviä toimijoita. Euroopassa Espanja on nykyään julkaisumääriltään Ranskan ja Italian kokoluokassa ja ohittanut ne esimerkiksi huippujulkaisujen osuudella mitaten.

Myöskään Kiina ei enää ole vahvasti esillä pelkästään julkaisumäärien takia, vaan yhä enemmän myös julkaisujen tieteellisen tason ansiosta. Tosin laatu näkyy vielä toistaiseksi ehkä enemmän yksittäisinä huippuina kuin tason kokonaisvaltaisena nousuna. Kiinan voimakas panostaminen tekoälytutkimukseen ja samalla maan lisääntyvä tutkimusyhteistyö selittänevät osaltaan myös Hong Kongin ja Singaporen vahvan aseman tekoälyn huippututkimuksessa.

Eräs Suomen vahvuus tutkimuksen vaikuttavuuden näkökulmasta on ollut kansainvälisen yhteistyön suhteellisen suuri rooli tutkimusjulkaisujen tuottamisessa. Kansainvälisen tutkimusyhteistyön osuus oli merkittävämmistä tekoälyjulkaisujen tuottajamaista suurin Saudi-Arabiassa ja Hong Kongissa, joissa molemmissa kansainvälisten yhteisjulkaisujen osuus oli kaksinkertainen verrattuna yhteisjulkaisujen kansainväliseen keskitasoon (Saudi Arabiassa 75% ja Hong Kongissa 73%). Maiden erona oli yhteistyön jakautuminen kumppanimaiden välillä: Hong Kong teki yhteistyötä ennen kaikkea Kiinan kanssa, Saudi-Arabian yhteisjulkaisut jakaantuivat puolestaan tasaisemmin useiden maiden kesken.

Kansainvälisen tutkimusyhteistyön osuutta v. 2013-2017 tarkastellessa erottuivat muista maista erityisesti Singapore ja Sveitsi. Käytännössä kyse on siten hyvin pitkälle samoista maista, joiden julkaisut keräävät paljon viittauksia muista julkaisuista. Kansainvälisen tutkimusyhteistyön määrän ja erityisesti juuri viittausmäärien onkin havaittu korreloivan toisiaan.

EU-maista kansainvälistä julkaisuyhteistyötä oli tarkastelujaksona suhteellisesti eniten Tanskassa, Belgiassa, Irlannissa ja Iso-Britanniassa. Myös Suomi ja muut Pohjoismaat olivat tässä vertailussa hyvin esillä, kansainvälisten yhteisjulkaisujen osuuden ollessa n. 50% sekä Ruotsissa, Norjassa että Suomessa.

⁴⁰ Haku Scopuksesta avainsanalla "artificial intelligence" ja julkaisut rajattu tietotekniikkaan. Data ajettu SciVal-työkalun läpi. Julkaisuja vertailussa mukana 78889.

Myös yritys-yliopisto (tutkimuslaitos) –yhteistyö on Suomessa keskimääräistä korkeammalla tasolla, vaikka ei ylläkään aivan huippumaiden tasolle. Kyseisiä maita olivat tarkasteluajanjaksona Israel, Sveitsi ja Yhdysvallat. Euroopan Unioni -maista erityisesti Ruotsilla ja Irlannilla on vahvaa näyttöä yritys-yliopisto – yhteistyöstä eri tekoälyn tutkimussuunnissa. Suurista EU-maista erityisesti Saksassa on yritysten ja yliopistojen tutkimusyhteistyöllä vakiintunut rooli. On kuitenkin hyvä todeta, että yritysten ja tutkimusorganisaatioiden yhteisjulkaisujen osuus ei huippumaissakaan ylitä 10%:tä kaikista maan julkaisuista.

Tutkimusorganisaatioiden vertaaminen keskenään tai listaaminen tutkimuksen vaikuttavuuden perusteella ei ollut tämän selvityksen ensisijaisena tavoitteena. Tekoälyn tutkimuksen tekee kuitenkin poikkeukselliseksi muutamien suurten yritysten, erityisesti Microsoftin, IBM:n ja osittain myös Googlen vahva panostus patentoinnin ohella myös julkiseen tutkimukseen. Esimerkiksi Microsoftin julkaisumäärät olivat v. 2008-2017 samaa luokkaa keskisuuren EU-maan kanssa, joten julkaisuja ja niihin liittyvää julkista tutkimusta ja tutkimusyhteistyötä ei voi pitää pelkkänä yrityksen t&k-toiminnan sivuvirtana, ainakaan suurimpien yritysten osalta. Microsoft sijoittui korkealle myös tarkasteltaessa julkaisujen tieteellistä vaikuttavuutta koko tekoälyn julkaisukentässä. Vastaavasti Google oli kärjessä koneoppimista käsittelevien julkaisujen viittausmäärissä. Jos lääketieteellistä tutkimusta ei tässä yhteydessä huomioida, niin yritysten sijoittumista korkealle myös julkaisujen vaikuttavuusvertailuissa voi pitää tieteellisen tutkimuksen piirissä poikkeuksellisenä. Yritysten tiivis yhteistyö ja kumppanuus huippuyliopistojen kanssa selittää tästä vain osan.

1.6 LIITE: tiedonhaut

Tutkimussuuntien rajauksessa käytetyt hakulauseet Elsevier-Scopus-tietokannassa.

Laaja tekoälyhaku:

KEY ("Artificial intelligence") OR KEY ("Learning systems") OR KEY ("Learning algorithms") OR KEY ("Deep Learning") OR KEY ("Image recognition") OR KEY ("Pattern recognition") OR KEY ("Machine Learning") OR KEY ("Natural Language Processing") OR KEY (nlp) OR KEY ("Speech Recognition") OR KEY ("Supervised learning") OR KEY ("Unsupervised Learning") OR KEY ("Reinforce* Learning") OR KEY ("Transfer Learning") OR KEY ("Neural network*") OR KEY ("Feature extraction") OR KEY ("Decision support system*") OR KEY ("Cognitive System*") OR KEY ("Knowledge Based System*") OR KEY ("Knowledge Representation") OR TITLE ("Artificial intelligence") OR TITLE ("Learning systems") OR TITLE ("Learning algorithms") OR TITLE ("Deep Learning") OR TITLE ("Image recognition") OR TITLE ("Pattern recognition") OR TITLE ("Machine Learning") OR TITLE ("Natural Language Processing") OR TITLE (nlp) OR TITLE ("Speech Recognition") OR TITLE ("Supervised learning") OR TITLE ("Unsupervised Learning") OR TITLE ("Reinforce* Learning") OR TITLE ("Transfer Learning") OR TITLE ("Neural network*") OR TITLE ("Feature extraction") OR TITLE ("Decision support system*") OR TITLE ("Cognitive System*") OR TITLE ("Knowledge Based System*") OR TITLE ("Knowledge Representation")

Data-analyysi:

KEY("Big Data" OR "Classification (of Information)" OR "Clustering Algorithm*" OR "Data Analy*" OR "Data Mining" OR "Predictive Analy*") OR TITLE("Big Data" OR "Classification (of Information)" OR "Clustering Algorithm*" OR "Data Analysis" OR "Data Mining" OR "Predictive Analy*") OR (KEY("Artificial Intelligence" OR "Expert System*" OR "Cognitive System*" OR "Learning System*" OR "Machine Learning" OR "Neural network*") AND KEY("Internet Of Things" OR IoT OR Bioinformatics OR Genetics OR "Gene Expression" OR "Remote sensing") AND (SUBJAREA(COMP)))

Havainnointi ja tilannetietoisuus:

KEY("Ambient Intelligence" OR "Automated Pattern Recognition" OR "Computer Vision" OR "Face Recognition" OR "Feature Extraction" OR "Image Analysis" OR "Image Classification" OR "Image Interpretation" OR "Image Segmentation" OR "Machine vision") OR TITLE("Ambient Intelligence" OR "Automated Pattern Recognition" OR "Computer Vision" OR "Face Recognition" OR "Feature Extraction" OR "Image Analysis" OR "Image Classification" OR "Image Interpretation" OR "Image Segmentation" OR "Machine vision") OR (KEY("Artificial Intelligence" OR "Expert System*" OR "Cognitive System*" OR "Learning System*" OR "Machine Learning" OR "Neural network*") AND KEY("Behavioral Research" OR "Medical Imaging" OR "Pattern Recognition" OR "Quality Control" OR Perception OR "Remote Sensing" OR "Virtual Reality") AND (SUBJAREA(COMP)))

Luonnollinen kieli ja kognitio:

KEY("Cognitive Systems" OR "Knowledge Based Systems" OR "Knowledge Representation" OR "Natural Language Processing " OR "Speech Recognition") OR TITLE("Cognitive Systems" OR "Heuristic Algorithms" OR "Knowledge Based Systems" OR "Knowledge Representation" OR "Natural Language Processing " OR "Speech Recognition") OR (KEY("Artificial Intelligence" OR "Expert System*" OR "Cognitive System*" OR "Learning System*" OR "Machine Learning" OR "Neural network*") AND KEY("Machine Translation" OR Cognition OR Reasoning OR Ontology OR Semantics) AND (SUBJAREA(COMP)))

Vuorovaikutus ihmisen kanssa:

KEY("Affective Computing" OR "Chatbots" OR "Computer Assisted Diagnosis" OR "Conversational Agents" OR "Emotion Recognition" OR "Game AI" OR "Game Artificial Intelligence" OR "Human Computer Interaction" OR "Human-Robot Interaction" OR "Machine Emotional Intelligence" OR "Recommendation Algorithms" OR "Recommendation Systems" OR "Recommender Systems") OR TITLE("Affective Computing" OR "Chatbots" OR "Computer Assisted Diagnosis" OR "Conversational Agents" OR "Game AI" OR "Game Artificial Intelligence" OR "Human Computer Interaction" OR "Human-Robot Interaction" OR "Machine Emotional Intelligence" OR "Recommendation Algorithms" OR "Recommendation Systems" OR "Recommender Systems") OR (KEY("Artificial Intelligence" OR "Expert System*" OR "Cognitive System*" OR "Learning System*" OR "Machine Learning" OR "Neural network*") AND KEY("Active Learning" OR Education OR "Interactive Agent*" OR "Social Networking " OR "Virtual Reality") AND (SUBJAREA(COMP)))

Digitaidot työelämässä, ongelmanratkaisu ja laskennallinen luovuus:

KEY("Algorithmic Composition" OR "Artificial Emotional Intelligence" OR "Computational Creativity" OR "Creative System*" OR "Decision Support System*" OR "Decision Tree*" OR "Digital Competences" OR "Digital Literacy" OR "Digital Skill*" OR "Digital Storytelling" OR "Scheduling Algorithm*" OR "Heuristic Algorithm*" OR "Knowledge Based Systems" OR "Swarm Intelligence") OR TITLE("Algorithmic Composition" OR "Artificial Emotional Intelligence" OR "Computational Creativity" OR "Creative System*" OR "Decision Support System*" OR "Decision Tree*" OR "Digital Competences" OR "Digital Literacy" OR "Digital Skill*" OR "Digital Storytelling" OR "Scheduling Algorithm*" OR "Knowledge Based Systems" OR "Swarm Intelligence") OR (KEY("Artificial Intelligence" OR "Expert System*" OR "Cognitive System*" OR "Learning System*" OR "Machine Learning" OR "Neural network*") AND KEY("Decision making" OR Forecasting OR Optimization OR "Problem Solving" OR Reasoning OR "Information Retrieval") AND (SUBJAREA(COMP)))

Koneoppiminen:

KEY("Ant Colony Optimization" OR "Artificial Bee Colony" OR "Artificial Neural Network*" OR "Deep Learning" OR "Deep Neural Networks" OR "Evolution Algorithms" OR "Evolutionary Algorithms" OR "Genetic Algorithms" OR "Intelligent Agents" OR "Intelligent Control" OR "Intelligent Systems" OR "Learning Algorithms" OR "Learning Systems" OR "Machine Learning" OR "Multi Agent Systems" OR "Neural Networks" OR "Particle Swarm Optimization" OR "Reinforced learning" OR "Reinforcement learning" OR "Statistical Learning" OR "Supervised learning" OR "Support Vector Machines" OR "Unsupervised Learning") OR TITLE("Ant Colony Optimization" OR "Artificial Bee Colony" OR "Artificial Neural Network*" OR "Deep Learning" OR "Deep Neural Networks" OR "Evolution Algorithms" OR "Evolutionary Algorithms" OR "Genetic Algorithms" OR "Intelligent Agents" OR "Intelligent Control" OR "Intelligent Systems" OR "Learning Algorithms" OR "Learning Systems" OR "Machine Learning" OR "Multi Agent Systems" OR "Neural Networks" OR "Particle Swarm Optimization" OR "Reinforced learning" OR "Reinforcement learning" OR "Statistical Learning" OR "Supervised learning" OR "Support Vector Machines" OR "Unsupervised Learning") OR (KEY("Artificial Intelligence" OR "Expert System*" OR "Cognitive System*" OR "Learning System*" OR "Machine Learning" OR "Neural network*") AND KEY("Bayesian Networks" OR "Fuzzy Systems" OR "Iterative Methods" OR "Principal Component Analysis" OR "Random Forests" OR "Regression Analysis" OR "Soft Computing" OR "Stochastic Systems") AND (SUBJAREA(COMP)))

Järjestelmätaso ja systeemivaikutukset:

KEY("Artificial Intelligence" OR "Expert System*" OR "Cognitive System*" OR "Learning System*" OR "Learning Algorithm*" OR "Machine Learning" OR "Neural network*") AND KEY("Adaptive Control Systems" OR "Automation" OR "Complex Networks" OR "Control Systems" OR "Dynamical Systems" OR "Electric Power Systems" OR "Electric Power Transmission Networks" OR "Large Scale Systems" OR "Linear Systems" OR "Manufacture" OR "Multi-body System Dynamics" OR "Nonlinear Control Systems" OR "Nonlinear Systems" OR "operation* analysis" OR "Power System Dynamics" OR "Real Time Systems" OR "Stochastic Systems" OR "Supply Chains" OR "System Dynamics" OR "System Of Systems" OR "System Stability" OR "System Theory" OR "Systems Analysis" OR "Systems Engineering" OR "Uncertainty Analysis") AND (SUBJAREA(COMP))

Tekoälyn laskentaympäristöt, alustat ja palvelut, ekosysteemit:

KEY("Artificial Intelligence" OR "Expert System*" OR "Cognitive System*" OR "Learning System*" OR "Machine Learning" OR "Neural network*") AND KEY("Big Data" OR "Cloud Computing" OR "Computer Architecture" OR "Data

Analysis Software" OR "Data Handling" OR "Cloud Environments" OR "Digital Storage" OR "Distributed Computer Systems" OR "High Performance Computing" OR "Information Systems" OR "Network Security" OR "Computer hardware" OR Hadoop OR MapReduce OR Map-Reduce OR "Data Acquisition") AND (SUBJAREA(COMP))

Robotiikka ja koneautomaatio - tekoälyn fyysinen ulottuvuus:

KEY("Autonomous Agent*" OR "Autonomous System*" OR "Industr* 4.0" OR "Robotic System*" OR "Robot Learning" OR "Robotics") OR TITLE("Autonomous Agent*" OR "Autonomous System*" OR "Industr* 4.0" OR "Robotic System*" OR "Robot Learning" OR "Robotics") OR (KEY("Artificial Intelligence" OR "Expert System*" OR "Cognitive System*" OR "Learning System*" OR "Machine Learning" OR "Neural network*") AND KEY("Internet Of Things" OR "Robot*" OR "Automation" OR "Process Control" OR "Factory Automation") AND (SUBJAREA(COMP)))

Etiikka, moraali, regulaatio ja lainsäädäntö:

KEY("Computer ethics" OR "Machine Ethics" OR "Robot Ethics" OR "Ethical Robots" OR "Machine Morality" OR "Ethical Machines" OR "Artificial Morality" OR "Data Ethics") OR TITLE("Computer ethics" OR "Machine Ethics" OR "Robot Ethics" OR "Ethical Robots" OR "Machine Morality" OR "Ethical Machines" OR "Artificial Morality" OR "Data Ethics") OR (KEY("Artificial Intelligence" OR "Cognitive System*" OR "Learning System*" OR "Machine Learning" OR "Neural network*" OR Robots OR Robotics) AND KEY ("Ethics" OR "Moral Obligations" OR Morality OR Morals OR Regulation OR Legislation OR Laws OR Compliance OR "Information Privacy" OR "Data Privacy" OR Transparency OR Liability OR Legal OR Politics OR "Public Policy" OR "Ethical Issues" OR "Data Protection" OR GDPR OR "General Data Protection Regulation" OR "Ethical aspects" OR "Ethical System" OR "Moral Agency") AND (SUBJAREA(SOCI)))