

## ANALYYSI

# Tekoäly ja massiiviset data-aineistot muovaavat rahoitusmarkkinoita

Rahoitusvakaus | 27.11.2019 | Lauri Jantunen, Kimmo Koskinen

### KIRJOITTAJAT



Lauri Jantunen  
Tietostrategiapäällikkö



Kimmo Koskinen  
Neuvonantaja

Tekoäly (artificial intelligence tai AI) ja massiiviset data-aineistot ovat nopeasti vakiinnuttamassa asemaansa osaksi rahoitusmarkkinoiden arkipäivää. Näitä edistyneitä teknologioita sovelletaan rahoitusprosessin kaikissa eri vaiheissa aina riskienhallinnasta portfolioiden optimointiin. Näillä uusilla menetelmillä ja teknologioilla on erinomaiset edellytykset tehostaa olemassa olevaa toimintaa ja luoda kokonaan uusia tuotteita, palveluita ja jakelukanavia. Näissä innovaatioissa on kuitenkin ominaisuuksia, jotka voivat altistaa markkinat uudentyyppisille haavoittuvuuksille. Siksi kehitystä tulee seurata tarkasti.



- Suomen Pankki järjestää Data Driven Financial Stability: Opportunities and Challenges in Big

Data -konferenssin Finlandia-talossa 2.–3.12.2019. Ks. lisätietoja [suomenpankki.fi](http://suomenpankki.fi).

- Seuraa konferenssin videolähetystä osoitteessa: <https://www.youtube.com/watch?v=uihJFBS1G9Q>
- Tutustu myös Lauri Jantusen englanninkieliseen blogiin: [Artificial intelligence and massive data are transforming financial markets](#).

## AI on osa suurempaa digitalisaation trendiä

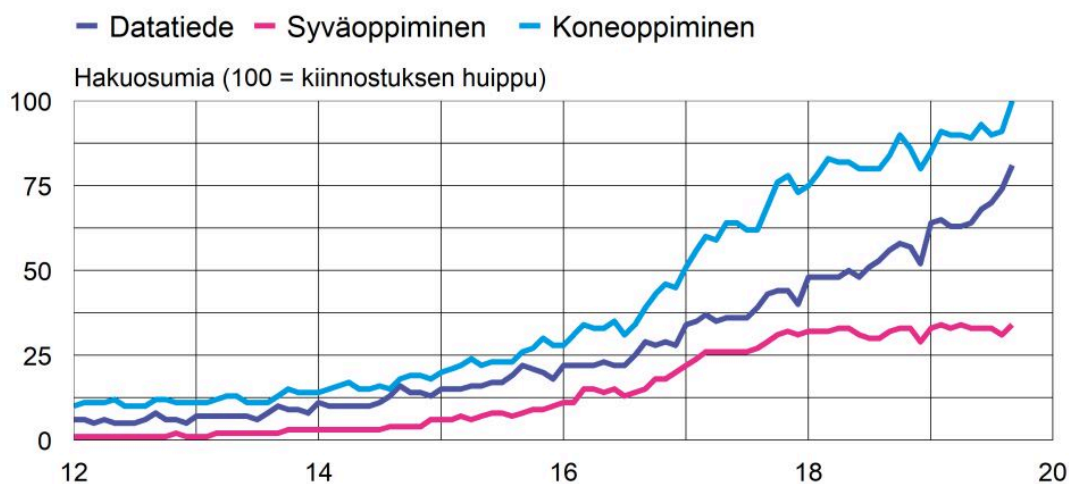
Tekoäly ei tapahdu tyhjiössä, vaan se on osa isompaa digitalisoitumisen murrosta, joka tuottaa yhä enemmän sellaista dataa, jota voidaan hyödyntää liiketoiminnan eri vaiheissa. Tätä ilmiötä kiihdyttävät innovaatiot edistyneessä analytiikassa ja datan prosessoimisessa, kuten mm. koneoppimisessa ja pilvipalveluissa. Nämä innovaatiot ovat kehityksen keskiössä, sillä AI tarvitsee datan lisäksi huomattavan edistynyttä osaamista, teknologiaa ja infrastruktuuria toimiakseen.

Tekoäly on määritelmänä kuitenkin haastava, sillä se ei yleisesti ottaen viittaa mihinkään yksittäiseen tekniseen toimintoon<sup>1</sup>. Termiä AI käytetään pikemminkin kuvaamaan useampien erilaisten toimintojen lopputuloksena syntyvää kokonaisuutta. World Economic Forum (WEF) esimerkiksi määrittelee<sup>2</sup> tekoälyn sellaisiksi teknologioiksi tai teknologioiden yhdistelmäksi, joita voidaan käyttää säännönmukaisuuksien tunnistamiseen, ennustamiseen, päätöksentekoon, interaktiivisuuteen tai palvelujen ja tuotteiden räätälöimiseen.

Tällaiset edistyneet teknologiaratkaisut ovat jo arkipäivää ja niitä sovelletaan yhä laajemmin rahoitusmarkkinoiden eri osa-alueilla. Tällä hetkellä yleisimmät tekoälyn käyttökohteet ovat mm. sääntelyssä ja asiakaspalvelussa, mutta näitä teknologioita käytetään myös luottoriskien hallintaan, sekä arvopaperikauppaan<sup>3,4</sup>. Tekoälyratkaisujen laajempi käyttöönotto tulee vaikuttamaan liiketoiminnan periaatteisiin ja markkinoiden toimintaan. Selkein muutos on siinä, että yritysten kilpailuetu tulee olemaan yhä tiiviimmin sidottuna laadukkaaseen dataan, datan määrään ja yrityksen kykyyn tuottaa näistä datamassoista liiketoiminnalle arvokasta informaatiota.

Kuvio 1.

## Yleisimmillä hakutermeillä mitattuna\* kiinnostus tekoälyn menetelmiin on kehittynyt nopeasti



\*) Asteikossa hakuosumat osoittavat hakutermin suosiota suhteessa valitun ajanjakson huippulukemaan. Asteikon arvo 100 on hakutermin huippuarvo.

Lähde: Google Trends.  
29.10.2019  
eurojatalous.fi

## Suurimmat haasteet edistyneissä teknologiaratkaisuissa liittyvät laatuun ja läpinäkyvyyteen

Data on tämän neljänneksi teolliseksi vallankumoukseksi nimetyn teknologiavetoisen murroksen polttoainetta. Laadukkaan datan avulla yritykset voivat optimoida koko arvoketjuna.

Datavetoiset yritykset voivat tehostaa liittymäpintojaan arvoketjun muiden toimittajien kanssa esimerkiksi tehokkaiden rajapintojen avulla. Data luo yrityksille myös edellytykset ymmärtää asiakkaita paremmin ja syventää näitä asiakassuhteita. He voivat tarjota kuluttajien henkilökohtaisille preferensseille paremmin soveltuvia digitaalisia palveluita, mutta he voivat myös oppia milloin kuluttajille kannattaa tarjota näitä tuotteita.

Datavetoisen liiketoiminnan etu perustuu vahvasti kustannustehokkuuteen, sillä ohjelmistoratkaisujen rajakustannukset ovat hyvin matalat. Alkuinvestointien<sup>5</sup> jälkeen nämä ohjelmistotuotteet ovat usein monistettavissa laajoille joukoille digitaalisten jakelukanavien

välityksellä. Prosessi on teoriassa itseään vahvistava, eli mitä enemmän dataa on<sup>6</sup>, sitä älykkäämpiä ja paremmin asiakkaan yksilöllisiin tarpeisiin räätälöityjä tuotteita voidaan kehittää. Älykkäämmät ja paremmin räätälöidyt tuotteet houkuttelevat lisää asiakkaita, joilta saadaan kerättyä enemmän dataa, jonka avulla voidaan kehittää tuotteita vieläkin paremmiksi. AI-projekteissa onnistujat pystyvät siten nopeasti kasvattamaan edelläkävijän etuaan (nk. first-mover advantage) ja saavuttamaan markkinaosuus.

Tätä tarkoitusta varten dataa kerätään nykyään valtavia määriä. Tutkimusyhtiö IDC ennustaakin datan määrän kasvavan 175 tsettatavuun<sup>7</sup> vuoteen 2025 mennessä. Datan hienojakoistuminen ja frekvenssin kasvaminen luovat kuitenkin kasvavaa tarvetta automaatiolle ja järjestelmien kehittämiseksi, koska käsiteltävät datamassat ovat niin massiivisia, että perinteisempien analytiikka- ja infrastruktuuriratkaisujen kapasiteetti ylitetään moninkertaisesti. Nämä uudet tietojärjestelmät on yleensä rakennettu tiedon yhteiskäyttöisyyden periaatteelle, eli sille ajatukselle että data tulisi olla mahdollisimman helposti hyödynnettävissä myös muissa järjestelmissä ja analyysissä. Lisäksi näissä kehittyneissä tietojärjestelmissä on täysin uudentyyppisiä ominaisuuksia. AI-menetelmiä voidaan esimerkiksi nykyään soveltaa yhä laajemmassa määrin myös puolistrukturoidulle ja strukturoimattomalle datalle, kuten sosiaalisen median viestien tai sopimustekstien analysoimiseen.

Tietomassoissa korostuu laadun ja tietoturvan merkitys. Suurissa tietotaltaisissa (nk. data lake) tiedon suojaus ja arkaluontoisen datan anonymisointi voi olla erityisen monimutkainen toimenpide, koska tiedon strukturoimaton luonne vaikeuttaa tietohallintoa ja turvallisuusauditoiteja. Tällaisten massiivisten data-aineistojen kohdalla on siksi haastavaa varmistua datan alkuperästä ja aitoudesta, sekä vaikeampaa valvoa sitä että kaikki arkaluonteinen materiaali on suojattu riittävän hyvin ja sitä ei käytetä väärin.

Massiivisten data-aineistojen kanssa työskentely tuottaa haasteita myös analyysille, sillä pieniltäkin tuntuvat ongelmat datan laadussa voivat tuottaa vääristymiä tuloksissa. AI-menetelmillä saatavat tulokset riippuvat siten täysin niille syötettävän datan laadusta (nk. garbage in, garbage out -ilmiö). Ongelmaksi voi tällöin muodostua se, että malli voi systemaattisesti tuottaa virheellisiä johtopäätöksiä esimerkiksi luottokelpoisuuden arvioimisessa.

Tällaisia datan laatuun liittyviä ongelmia, voi vahvistaa se, että AI-ratkaisuissa yleisesti käytettävät algoritmit eivät ole yksinkertaisia lineaarisia malleja. Esimerkiksi koneoppimisen avulla tuotettujen päätöspuiden syvyys ja monimutkaisuus voi tehdä näiden mallien toimintalogiikan ymmärtämisen äärimmäisen vaikeaksi. Tämä nk. black box -ilmiö on sitä vahvempi, mitä heikommin pystymme erittelemään yksittäisten parametrien vaikutuksen mallin käyttäytymiseen. Mikäli näiden mallien toimintalogiikkaa ei tunneta kunnolla, niin mallin tuottamia systemaattisia

virheitä ei välttämättä havaita ajoissa.

Puutteellinen mallintaminen, data ja algoritmit voivat pahimmillaan johtaa korvausvaatimuksiin, tai tilanteeseen jossa riskit ovat joko pahasti aliarvioitu ja/tai hyvin keskittyneet. Ymmärrys näiden teknologioiden puutteista näyttäisi kuitenkin parantuvan ja tekoälyn sovellukset ovat usein liitetty liiketoiminnan riskienhallintaan. Englannin keskuspankin äskettäin teettämässä kyselyssä<sup>8</sup> vastaajat näyttävät olevan hyvin perillä näiden välineiden ongelmista ja siksi koneoppimista soveltavissa AI-prosesseissa käytetään pääsääntöisesti ihmisiä vahvistamaan päätökset tai ongelmallisiksi merkityt tapaukset.

## Uusilla toimijoilla on paremmat edellytykset soveltaa toiminnassaan edistynyttä teknologiaa

Laadukkaat tekoälyn sovellukset vaativat organisaatiolta hyvin korkeaa analyttisen kypsyyden tasoa ja hyvin erilaisten asiantuntijoiden saumatonta yhteistyötä. AI-ratkaisuissa korostuvat järjestelmien, ohjelmoinnin, tilastollisten menetelmien ja liiketoiminnan substanssiosaamisen merkitys. Monien AI-projektien epäonnistumisen<sup>9</sup> taustalla on mm. liian kapea osaamis pohja.

Perinteisten pankkien ja luottolaitosten suurin kompastuskivi on ollut siinä, että heidän tietovarastonsa ja tietoarkkitehtuurinsa perustuvat usein vanhentuneelle teknologialle. Näihin vanhentuneisiin ratkaisuihin on sisältynyt paljon manuaalisia prosessiratkaisuja, jotka eivät ole perustuneet jakamisen ajatukselle ja joita ei ole rakennettu tukemaan uusia edistyneitä teknologiaratkaisuja. Data on tällaisissa järjestelmissä usein siiloutunutta ja sitä on hankala soveltaa analyysikäyttöön. Tämä on vaikeuttanut sellaisten uusien innovaatioiden jalkauttamista, joilla olisi voitu parantaa esimerkiksi luottoriskien arvioimista tai tehostaa viranomaisraportoinnin tuotantoa. Uusilla toimijoilla ei ole ollut vastaavanlaisia vanhoja järjestelmiä ylläpidettävänä, vaan he ovat voineet rakentaa tietoarkkitehtuurinsa suoraan ketterien periaatteiden mukaan tukemaan datavetoisen liiketoiminnan erityistarpeita.

Kansainvälinen järjestelypankki (BIS) on viime aikoina kiinnittänyt erityistä huomiota siihen miten suurten teknologiajättien<sup>10</sup> (nk. big tech) tuleminen vaikuttaa rahoitusmarkkinoiden toimintaan<sup>11</sup>. Huolena on erityisesti se, että näillä suurilla teknologiayhtiöillä on erinomaiset edellytykset hyödyntää liiketoiminnassaan dataa ja hyvät valmiudet monistaa toimivaksi todettua liiketoimintaa globaalisti eri maanosissa. Lisäksi nämä suuret teknologiayhtiöt pystyvät tukemaan tappiollistakin liiketoimintaa muiden liiketoimien tuottamalla kassavirroilla markkinaosuuksien saavuttamiseksi.

Mikäli innovaatiot ja kilpailuetu ovat tulevaisuudessa pääosin sidottu dataan ja nämä

teknologiajätit hallitsevat suurinta osaa tästä datan keruusta, ei ole selvää, että heillä on kannustin jakaa tällaisia tietovarastoja kilpailijoidensa kanssa. Uhkana on, että rahoitusmarkkinoiden syvenevä datavetoisuus johtaa uudentyypisiin kilpailurajoitteisiin, datamonopoleihin, markkinoiden vahvempaan jakautumiseen markkinoita dominoiviin massatuotteisiin ja hyvin kapeaa markkinasegmenttiä hyödyntäviin erikoistuotteisiin.

Markkinoilla on kuitenkin näkyvissä paljon aloitteita ja yhteistyöprojekteja, jotka perustuvat datan ja lähdekoodin avoimelle jakamiselle. Erityisesti julkisyhteisöt ovat kunnostautuneet tarjoamalla dataa maksuttomasti koneluettavassa muodossa. Avoimen datan ja lähdekoodin aloitteilla voidaan ehkäistä datan monopolisoitumista, mutta myös vähentää kehitystyön kustannuksia, lisätä päätöksenteon läpinäkyvyyttä ja avata ovia uusille datavetoisille innovaatioille.

## Nopea kasvu ja uusien tulokkaiden kokemattomuus voivat lisätä haavoittuvuuksia

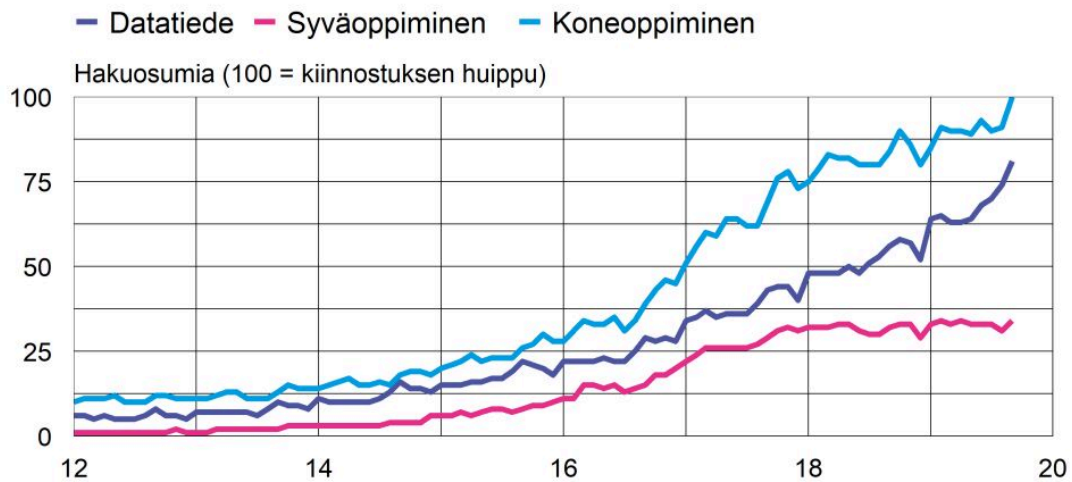
Vaikka tekoälyyn ja massiivisiin data-aineistoihin perustuvat tuotteet voivat osaltaan tehostaa rahoituksen välittymistä, niin kova kilpailu AI-teeman ympärillä voi kannustaa toimijoita lanseeraamaan tekoälyyn perustuvia ratkaisuja kovalla aikataululla ja tiukalla budjetilla.

Uudentyyppisillä fintech-toimijoilla ei välttämättä kuitenkaan ole samanlaista teknistä osaamista ja kokemusta rahoitustoiminnasta perinteisiin rahoituslaitoksiin verrattuna, vaikka ne usein tarjoavatkin pankkimaisia palveluita. Ei ole selvää kuinka hyvin aloittavien yritysten liiketoimintamalleissa ja innovatiivisissa tuotteissa on otettu huomioon luottosyklkien eri vaiheet ja kuinka hyvin nämä kestäisivät kireämmän rahapolitiikan tai poikkeuksellisen volatiilien markkinaolosuhteiden ajanjaksot. Tällaiset puutteet voivat korostua asiakasdatassa, jonka historia ei välttämättä ulotu kovin pitkälle menneisyyteen, ei sisällä samanlaisia suhdannehuippuja ja taantumia, tai jonka laadusta ei ole varmuutta. Mikäli tuotteita ja palveluita ei ole suunniteltu ja testattu riittävän hyvin, tai jos näiden tuotteiden riskienhallinnassa esiintyy puutteita, niin näihin voi sisältyä rakenteellisia haavoittuvuuksia.

Tässä mielessä moni perinteisempi toimija voi itse asiassa olla uusia toimijoita selvästi paremmassa asemassa soveltamaan kehittyneitä analyysimenetelmiä, sillä näiden perinteisten rahoituslaitoksien vahvuus on juuri tekoälyn kannalta tärkeissä pitkissä aikasarjoissa ja erilaisten riskien arvioinnissa. Ongelmana on mukauttaa olemassa olevat järjestelmät uusiin menetelmiin ja työskentelytapoihin soveltuviksi. Tällaiset yhteensovittamishankkeet ovat kuitenkin kalliita ja voivat siten lisätä kustannuspaineita lyhyellä aikavälillä.

Kuvio 2.

## Yleisimmillä hakutermeillä mitattuna\* kiinnostus tekoälyn menetelmiin on kehittynyt nopeasti



\*) Asteikossa hakuosumat osoittavat hakutermin suosiota suhteessa valitun ajanjakson huippulukemaan. Asteikon arvo 100 on hakutermin huippuarvo.

Lähde: Google Trends.  
29.10.2019  
eurojatalous.fi

## Teknologia voi luoda myös uudentyyppisiä kytköksiä eri toimijoiden välillä

Hyvät ja laadukkaat tekoälyratkaisut eivät kuitenkaan synny tyhjästä, vaan ne edellyttävät edistynyttä tietoarkkitehtuuria ja osaavia asiantuntijoita. Tätä ongelmaa on osin pyritty ratkaisemaan panostamalla erilaisiin avoimen lähdekoodin ratkaisuihin<sup>12</sup>, mutta viime vuosina on yleistynyt ilmiö, jossa näitä järjestelmäkehityksen ja -testaamisen kuluja on pyritty laskemaan tuotteistamalla kehitettävä järjestelmä omaksi palveluksi. Esimerkiksi operationaalisista varainhallinnan back-office -järjestelmistä on tehty tuotteita, joita voidaan tarjota muille toimijoille maksua vastaan.

Liiketoiminnan kannalta tällaisten palveluiden<sup>13</sup> hyödyt ja kustannussäästöt voivat olla huomattavia. Pilvialustoilla toimivien palveluntarjoajien yleistyminen on tarkoittanut sitä, että

näitä järjestelmiä voidaan ottaa nopeasti käyttöön. Työasemaksi riittää oma tietokone, sillä kaikki laskenta ja tallennus voidaan tehdä joustavasti ja kustannustehokkaasti pilvipalvelun toimittajan laskentajärjestelmässä ilman suuria käyttöönottoprojekteja ja huomattavia investointeja järjestelmäarkkitehtuuriin ja ylläpitoon.

Näiden palvelujen laadun parantuessa, yhä suuremmalla osalla toimijoista on kannustin siirtynyt tällaisten palvelujen kuluttajaksi, kuin ryhtyä tuottamaan näitä toimintoja itsenäisesti. Tämä voi parantaa yksittäisten yritysten kilpailukykyä, mutta voi myös luoda uuden tyyppisiä haavoittuvuuksia rahoitusjärjestelmään, mikäli osat markkinoista tulevat hyvin riippuvaisiksi muutamien avaintoimijoiden tuotteista. Ongelmat, toimintahäiriöt ja haavoittuvuudet tällaisissa markkinoiden kannalta kriittisissä palveluissa voivat nopeasti heijastua kaikkien markkinaosapuolten toimintaan.

## Markkinoiden kehitystä tulee seurata tarkasti

Oikein käytettynä tekoäly ja massiiviset data-aineistot voivat edistää monella tapaa parempaa maailmaa. World Economic Forum listaa<sup>14</sup> useita erilaisia kanavia, joita pitkin nämä ratkaisut voivat tukea innovaatioita ja resurssien tehokkaampaa hyödyntämistä. Näihin teknologioihin liittyy kuitenkin monia teknisiä, eettisiä ja markkinoiden luonnetta koskevia haasteita, joita tulee seurata tarkasti.

Syvenevän digitalisoitumisen myötä syntyvät uudet tuotteet ja palvelut eivät välttämättä ole rajattuja maantieteellisiin alueisiin. Niiden rajat ylittävä luonne voi luoda haasteita valvoa mm. sitä, että kuluttajansuoja ja yksityisyys tulee huomioitua näissä palveluissa. Tämä voi luoda uudenlaisia tarpeita tiiviimmälle kansainväliselle koordinaatiolle näiden teknologioiden valvonnassa.

Hyvin voimakas kasvu teknologiapainoisilla toimialoilla vaikuttaa työvoiman kysyntään. Se paitsi lisää teknologiaosaajien kysyntää näillä aloilla, mahdollisesti muuttaa laajemminkin niitä tieto- ja taitoja, joita vaaditaan yhä suuremmalta osalta työvoimaa digitalisoituvassa taloudessa. Tällaisiin ilmiöihin vastaaminen vaatii aktiivista sopeutumista pitkällä aikavälillä mm. koulutuspolitiikan osalta.

Mitä laajemmin nämä teknologiat määrittelevät rahoituksen välittymistä, niin sitä enemmän ne myös muovaavat eri toimijoiden rooleja ja näiden välisiä kytköksiä. Nämä muutokset voivat muuttaa markkinoiden rakenteita ja toimijoiden rooleja. Muutokset vaikuttavat myös siihen, millä tekijöillä on vaikutusta rahoitusmarkkinoiden vakauteen ja vakauriskeihin. On siksi tärkeää varmistaa, että datavetoinen liiketoiminta ja innovaatiot tapahtuvat sellaisessa kehikossa, jossa kaikilla toimijoilla on yhtäläiset ja reilut toimintaedellytykset.

## Viitteet

1. Elements of AI, Helsingin Yliopisto ja Reaktor, 2018. ↑
2. The new physics of financial services, WEF, elokuu 2018. ↑
3. Artificial intelligence and machine learning in financial services, FSB, 2017. ↑
4. Monitoring of fast-paced electronic markets, BIS, 2018. ↑
5. Vaadittavat alkuinvestoinnit IT-infrastruktuuriin ja ylläpitoon voivat olla huomattavia, mutta esimerkiksi pilvipalveluita voidaan hyödyntää näiden investointitarpeiden kustannusten alentamiseen. ↑
6. On tärkeää huomioida, että laadukkaan informaation jalostaminen raakadatasta ei ole mikään yksinkertainen operaatio, vaan vaatii useita erilaisia työvaiheita, jotka eivät varsinaisesti liity tekoälyyn. ↑
7. Yksi tsettatavu on miljardi teratavua. ↑
8. Machine learning in UK financial services, Englannin keskuspankki ja FCA, 2019. ↑
9. Tutkimusyhtiö IDC arvioi, että neljänneksellä organisaatioista, jopa puolet AI-projekteista epäonnistuvat (Artificial Intelligence Global Adoption Trends & Strategies, IDC, 2019). ↑
10. Teknologiayhtiöt kuten Alibaba, Amazon, Facebook, Google ja Tencent. ↑
11. Annual Economic Report: Big tech in finance, BIS, 2019. ↑
12. Useat markkinoiden johtavista koneoppimisen ohjelmistokirjastoista (esim. Pytorch ja Tensorflow -kirjastot) perustuvat alun pitäen suurten teknologiayhtiöiden, kuten Facebook ja Google kehittämiin sisäisiin ohjelmistoratkaisuihin, jotka on päätetty julkaista avoimen lähdekoodin lisenssillä. ↑
13. Näistä palveluista käytetään yleensä nimitystä Software-as-a-Service (SaaS), tai Platform-as-a-Service (PaaS). ↑
14. Harnessing artificial intelligence for the earth, WEF yhteistyössä PwC ja Stanford Woods Institute for the Environment, 2018. ↑

Tässä artikkelissa esitetyt mielipiteet ovat kirjoittajien omia eivätkä välttämättä edusta Suomen Pankin näkemystä.

## Asiasanat

AI, digitalisaatio, fintech, innovaatio, koneoppiminen, rahoitusmarkkinat, rahoitusvakaus, teknologiajätit, tekoäly