

Datatalouden koon ja vaikuttavuuden arviointi Suomessa



Heli Koski, ETLA

Johannes Anttila, Demos

Anna Björk, Demos

Vera Djakonoff, Demos

Otto Kässi, ETLA

Hanna Niemi-Hugaerts, TIEKE

Mika Pajarinen, ETLA

Jussi-Pekka Parkkari, TIEKE

Suosittelava lähdeviittaus:

Koski, Heli, Anttila, Johannes, Björk, Anna, Djakonoff, Vera, Kässi, Otto, Niemi-Hugaerts, Hanna, Pajarinen, Mika & Parkkari, Jussi-Pekka (18.12.2024). ”Datatalouden koon ja vaikuttavuuden arviointi Suomessa”.

ETLA Muistio nro 150.

<https://pub.etla.fi/ETLA-Muistio-Brief-150.pdf>

Tiivistelmä

Käytössä olevien rekisteriaineistojen avulla datan roolia taloudessa voidaan parhaiten arvioida kustannusperusteisilla menetelmillä. Kokeellisten laskelmiemme mukaan vuonna 2021 dataperusteisen työn tuottaman arvonlisäyksen osuus oli kymmenisen prosenttia ja datainvestointien arvo vajaa kaksi prosenttia Suomen bruttokansantuotteesta. Suurin osa datatalouden arvosta syntyi tieto- ja viestintäalalla.

Datatalouden koon ja vaikuttavuuden tarkempi arviointi ja julkisten datataloutta edistävien toimien vaikuttavuuden arviointi edellyttäisivät panostuksia tiedonkeruuseen ja tilastokäytäntöihin.

- Tilinpäätöskäytäntöjen muutokset ja panostukset uuteen tiedonkeruuseen (kuten työntekijöiden ajankäyttö dataperusteiseen työhön) ovat välttämättömiä, jos datatalouden kokoa ja vaikuttavuutta halutaan arvioida tarkemmin.
- Kansainvälinen yhteistyö yhtenäisten mittauskäytäntöjen ja tiedonkeruun kehittämisessä on tärkeää, jotta saadaan vertailukelpoista tietoa Suomen datatalouden kehityksestä kansainvälisessä kontekstissa.
- Julkisten toimien, kuten datatalouteen liittyvien hankintojen ja tki-rahoituksen, vaikuttavuuden arvioimiseksi pitäisi parantaa datatalouteen liittyvien toimien tunnistamista kerättävistä aineistoista.
- Tiedon ja kokemusten jakamisella datataloutta tukevien toimenpiteiden parhaista käytännöistä voisi edistää datatalouden kasvua EU-alueella.

Abstract

Assessment of the Size and Impact of the Data Economy in Finland

Within the scope of the available register data, the role of data in the economy is best estimated using cost-based methods. Our experimental calculations indicate that, in 2021, data-driven work contributed approximately 10 percent of value added, while data investments represented slightly under two percent of Finland's GDP. The majority of the value in the data economy is generated within the information and communication sector.

To better assess the size and impact of the data economy and the effectiveness of public policies aimed at fostering its growth, investments in data collection and statistical methodologies are essential.

- Changes in accounting practices and investments in new data collection (such as measuring employees' time spent on data-driven work) are essential for a more accurate assessment of the size and impact of the data economy.
- International collaboration in developing consistent measurement practices and data collection is important to obtain comparable information on the development of Finland's data economy in an international context
- To evaluate public policy initiatives—such as data economy procurement and R&D funding—the identification of data economy-related activities within collected data must be improved.
- Additionally, sharing knowledge and best practices across the EU could further stimulate the growth of the data economy.

KTT **Heli Koski** on Elinkeinoelämän tutkimuslaitoksen tutkimusjohtaja. (heli.koski@etla.fi)

VTM **Johannes Anttila** on vanhempi policy-asiantuntija Demos Helsingissä. (johannes.anttila@demoshelsinki.fi)

YTT **Anna Björk** on tiiminvetäjä ja johtava tutkija Demos Helsingissä. (anna.bjork@demoshelsinki.fi)

VTM **Vera Djakonoff** on asiantuntija Demos Helsingissä. (vera.djakonoff@demoshelsinki.fi)

VTT **Otto Kässi** on Elinkeinoelämän tutkimuslaitoksen tutkija. (otto.kassi@etla.fi)

TaM Uusi media **Hanna Niemi-Hugaerts** on Tietoyhteiskunnan kehittämiskeskus (TIEKE) ry:n toiminnanjohtaja. (hanna.niemi-hugaerts@tieke.fi)

KTM **Mika Pajarinen** on Elinkeinoelämän tutkimuslaitoksen tutkija. (mika.pajarinen@etla.fi)

TkK **Jussi-Pekka Parkkari** on Tietoyhteiskunnan kehittämiskeskus (TIEKE) ry:n Dataosaaja. (jussi-pekka.parkkari@tieke.fi)

Ph.D. (Econ.) **Heli Koski** is a Research Director at ETLA Economic Research.

MSc **Johannes Anttila** is a Senior Policy Expert at Demos Helsinki.

D.Soc.Sci **Anna Björk** is a Team Lead and Leading Researcher at Demos Helsinki.

M.Soc.Sci **Vera Djakonoff** is an Expert at Demos Helsinki.

Dr. Soc. Sc. **Otto Kässi** is a researcher at ETLA Economic Research.

MA (New Media) **Hanna Niemi-Hugaerts** is the Executive Director of TIEKE Finnish Information Society Development Centre.

M.Sc. (Econ.) **Mika Pajarinen** is a Researcher at ETLA Economic Research.

B.Sc (Tech) **Jussi-Pekka Parkkari** is a Data Expert at TIEKE Finnish Information Society Development Centre.

Kiitokset: Kirjoittajat kiittävät *Datatalouden vaikuttavuus* -hankkeen ohjausryhmää arvokkaista kommentteista.

Acknowledgements: The authors thank the steering group of the *Impacts of the Data Economy* project for their valuable comments.

Avainsanat: Data, Datatalous, Taloudelliset vaikutukset

Key words: Data, Data economy, Economic impacts

JEL: D24, O3, O5

Johdanto

Digitaalinen data tarjoaa uusia mahdollisuuksia arvontuontiin ja tuottavuuden parantamiseen, ja sen hyödyntäminen synnyttää uusia liiketoimintamalleja, tuotteita ja palveluja. Datalähtöiset innovaatiot ovat jo muuttaneet useiden toimialojen toimintaa, ja datan rooli talouden veturina korostuu entisestään tulevaisuudessa. Tästä huolimatta meillä on vielä verrattain vähän tietoa siitä, miten nämä muutokset vaikuttavat kansantalouden tasolla ja millä tavoin datan hyödyntäminen tukee talouskasvua.

Olemassa olevien aineistojen puitteissa ei ole mahdollista seurata tai arvioida tarkasti datatalouden kokoa, tuottamaa arvonlisäystä tai sen vaikutuksia talouskasvuun. Tiedon ja yhteisesti sovitun tarkastelukehikon puute hankaloittaa myös datatalouden julkisten kehittämis- ja tukitoimien vaikuttavuuden arviointia. Datatalouden kokoa voidaan kuitenkin mitata karkeasti olemassa olevien rekisteriaineistojen perusteella.

Datatalouden koko Suomessa

Data poikkeaa perinteisistä hyödykkeistä ja tuotannontekijöistä. Se on kulumaton, samaa dataa voi käyttää usea taho samanaikaisesti ja sitä voi yhdistellä muuhun dataan luoden lukemattomia käyttömahdollisuuksia. Data on aineeton omaisuus, kun se tuotetaan kerran ja sitä käytetään toistuvasti taloudellista hyötyä tuottavana tuotantopanoksena.

Dataomaisuuden kirjaamiseen yrityksen taseeseen liittyy kuitenkin haasteita, eivätkä tilinpäätöskäytännöt edellytä raportointia datan osalta. Tällä hetkellä tietoja datan hyödyntämisestä ei kerätä kattavasti eikä olemassa olevien aineistojen perusteella saada tarkkaa kuvaa datatalouden koosta eikä siitä, miten se vaikuttaa talouskasvuun tai hyvinvointiin.

Kansainvälisiä mittareita, kuten EU:n DESI-indeksi ja IDC:n datatalouden mittaamiskehikko European Data Market Monitoring Tool, on jo kehitetty datatalouden

arvioimiseksi. IDC (2021) määrittelee **datatalouden arvonlisäyksenä tai hyvinvointina, jonka datan hyödyntäminen tuottaa koko taloudessa**. Tämä mittari on hyödyllinen, koska sen avulla datatalouden tuottamaa arvonlisäystä voidaan arvioida suhteessa bruttokansantuotteeseen. IDC arvioi datatalouden arvoa eri EU-maissa useiden tekijöiden pohjalta. Heidän laskelmiensa mukaan vuosina 2020 ja 2021 datatalouden arvo Suomessa (sisältäen suorat ja epäsuorat vaikutukset) oli noin 4,4 prosenttia bruttokansantuotteesta.

IDC:n mittauskehikko ei kuitenkaan ole luonteeltaan toistettava. IDC:n kehittämän työkalun käyttö edellyttäisi merkittäviä tutkimusresursseja ja pääsyn erilaisiin rekisteriaineistoihin sekä IDC:n hallussa oleviin tietokantoihin. Näiden aineistojen analyysin lisäksi tarvittaisiin lisätietojen keruuta, laadullisia analyyseja ja syvähaastatteluja.

Datatalouden kehityksen vuositason seurantaan sopivat paremmin pienemmillä panostuksilla käytettävissä olevat epäsuorat datatalouden arvon mittaamisen menetelmät. Arviomme mukaan näistä menetelmistä parhaiten Suomeen sovellettavissa on kustannusperusteinen lähestymistapa¹, joka perustuu datan luomisen, keräämisen, varastoinnin ja tietokantojen ylläpidon kustannuksiin. Tämä menetelmä on yhdenmukainen kansantalouden tilinpidon periaatteiden kanssa, ja sitä on hyödynnetty jo monissa maissa kokeellisesti arvioimaan datainvestointien ja -omaisuuden määrää.

Kanadan tilastoviranomaisen (Statistics Canada, 2019) kehittämässä menetelmässä ja kokeellisissa laskelmissa datan arvo yrityssektorilla arvioitiin hyödyntämällä rekisteriaineistoista kerättyjä työllisyys- ja palkkatietoja, jotka kuvastivat dataperusteiseen työhön käytettyjä työvoimapanoksia. Menetelmän käyttö edellyttää, että dataperusteiseen työhön osallistuvat ammattiryhmät ja heidän dataperusteiseen työhön käyttämänsä aika voidaan arvioida.

Goodridge ym. (2022) arvioivat kustannusperusteisella menetelmällä dataperusteisten investointien määrää EU-maissa ja datapääoman vaikutusta tuottavuuden kasvuun. He havaitsivat, että 57 prosenttia työvoimasta, joka osallistuu ohjelmisto- ja datapääoman muodostamiseen, sisältyy jo kansantalouden tilinpitoon, mutta merkittävä osa jää edelleen SNA:n ulkopuolelle. Laajennetun inves-

tointimääritelmän käyttö kasvulaskennassa vaikutti sekä panoksiin että tuotoksiin ja nosti työn tuottavuuden kasvua EU13-maissa 0,79 prosentista 0,83 prosenttiin vuodessa ja lisäsi tuotantoa 6,7 miljardilla eurolla vuonna 2016. Samalla ohjelmisto- ja datapääoman syventymisen vaikutus tuottavuuteen kolminkertaistui lisäen arvoa 9,4 miljardilla eurolla samana vuonna.

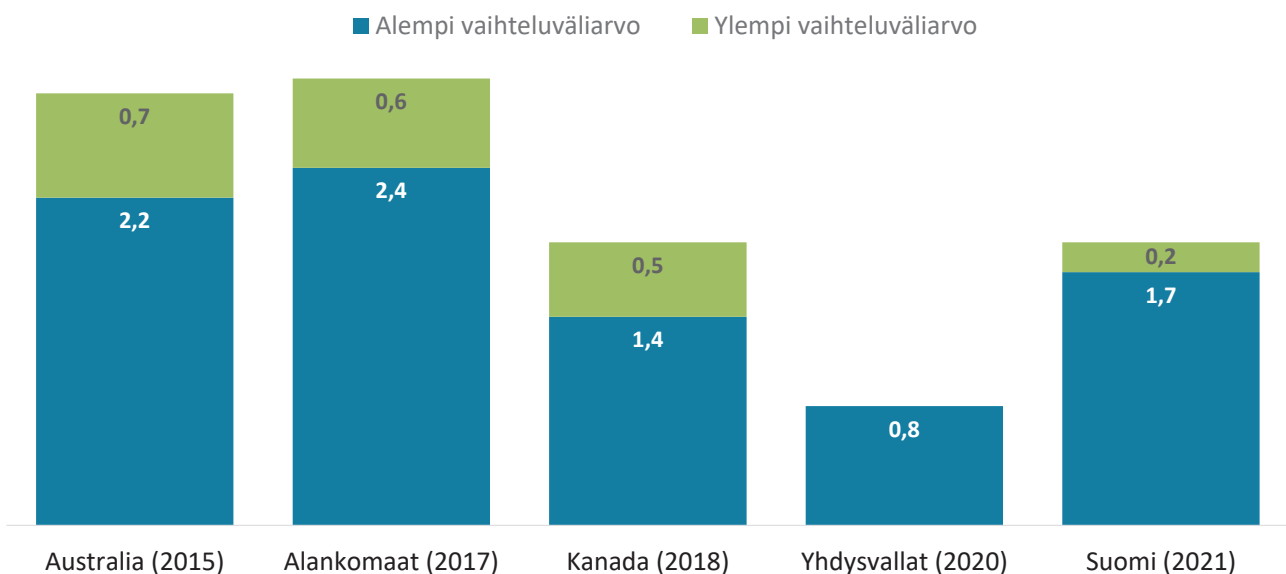
Hyödynnämme Kanadan tilastoviranomaisen kehittämää viitekehystä, jossa dataomaisuus luokitellaan kolmeen kategoriaan: data, tietokannat ja datatiede. Mittaamme dataperusteisen arvonluonnin tuotantokustannuksia keskeisten ammattiryhmien palkkatietojen perusteella sekä arvioimme erikseen ohjelmisto- ja tietokantainvestointien kustannuksia Goodridgen ym. (2022) lähestymistapaa mukaillen. Haasteena on se, että datatyöhön käytettyä työaikaa ei ole suoraan saatavilla, joten käytämme Statistics Canadan arvioita työajan jakautumisesta datatyöhön ja otamme laskelmissa huomioon vaihteluvälillä näille työaikaosuuksille. Tarkempi kuvaus käytetyistä menetelmistä ja aineistoista löytyy Kosken ja Pajarisen (2024) ja Kosken ym. (2024) raporteista.

Laskelmiemme mukaan datainvestointien osuus suhteessa bruttokansantuotteeseen oli Suomessa vuosina 2012–

2021 keskimäärin hieman alle kaksi prosenttia. Samana ajanjaksona OECD:n tietojen mukaan tilinpäätöstiedoista johdetut ohjelmisto- ja tietokantainvestoinnit vastasivat noin 1,4 prosenttia bruttokansantuotteesta. Tämä viittaa siihen, että laajennetun dataomaisuuden määritelmän käyttö nostaa datainvestointien osuutta bruttokansantuotteesta noin 0,5 prosenttiyksiköllä.

Vuonna 2021 dataperusteisen työn tuottama arvonlisäys vastasi noin kymmentä prosenttia bruttokansantuotteesta, ja datainvestointien osuus bruttokansantuotteesta Suomessa oli arviolta 1,7–1,9 prosenttia. Kuviossa 1 esitetään vertailu Suomen ja muiden maiden tuloksista, joissa on hyödynnetty vertailukelpoisia laskentamenetelmiä. Suomen osuus on samaa suuruusluokkaa kuin Kanadan, jossa datainvestointien osuus bruttokansantuotteesta oli vuonna 2018 1,4–1,9 prosenttia (Statistics Canada, 2019). Muita vertailukohtia ovat Australia, jossa datainvestointien osuus oli vuonna 2016 hieman suurempi, 2,2–2,9 prosenttia (Smedes ym., 2022), sekä Alankomaat, jossa osuus vuonna 2017 oli 2,4–3,0 prosenttia (de Bondt & Mushkudiani, 2021). Yhdysvalloissa datainvestointien osuus oli kuitenkin selvästi alhaisempi, vain 0,8 prosenttia vuonna 2020 (Calderón & Rassier, 2022).

Kuvio 1 Datainvestointien arvo suhteessa bruttokansantuotteeseen, %



Julkisen sektorin mahdollisuudet edistää datataloutta ja sen vaikuttavuuden arviointia

Mittaamisen ja vaikuttavuuden arvioinnin kehittäminen

Kansainvälinen työ datan sisällyttämiseksi osaksi kansantalouden tilinpitoa on käynnissä osana SNA 2008:n (System of National Accounts) uudistusta. Uudistuksen yhtenä tavoitteena on kehittää yhtenäiset kansainväliset tilinpäätöskäytännöt datan käsittelemiseksi varallisuuseränä ja ratkaista siihen liittyviä kirjaamishaasteita. Eurostatin ja IMF:n yhteinen työryhmä tukee vuonna 2025 valmistuvaa uudistusta laatimalla käsikirjan, joka ohjeistaa datan käsittelyä varallisuuseränä kansantalouden tilinpidossa. Tämä kehitys edistää yhtenäisten arvioiden laatimista datainvestoinneista ja -varannoista sekä parantaa edellytyksiä kansainvälisten vertailujen tekemiseen.

Panostukset datatalouden arvon ja kehityksen mittaamiseen Suomessa ovat erityisen hyödyllisiä, jos mittaamisen kehitystyötä tehdään rinnakkain kansainvälisellä tasolla yhteisten mittareiden luomiseksi ja vertailukelpoisen tiedon tuottamiseksi. Kansainväliset vertailut tarjoavat arvokasta näkökulmaa Suomen datatalouden asemaan ja kehitykseen suhteessa muihin maihin. Tiedonkeruuta datatyöhön käytetystä ajasta voitaisiin toteuttaa esimerkiksi lisäämällä kysymyksiä Eurostatin kaikissa EU-maissa toteutettaviin harmonisoiutuihin kyselytutkimuksiin. Tämä mahdollistaisi yhteisten mittareiden luomisen ja vertailukelpoisen tiedon tuottamisen.

Datatalouteen liittyvien julkisten toimien, kuten tki-rahoituksen ja julkisten hankintojen, vaikuttavuuden arviointia voitaisiin parantaa lisäämällä julkisrahoitteisiin hankkeisiin datatalous-tunniste. Nykyisellään datataloushankkeita ei ole mahdollista tunnistaa suoraan aineistoista, mikä vaikeuttaa niiden seuranta. Tunniste helpottaisi hankkeiden tunnistamista ja arviointia, vaikka datatalous ei olisikaan hankkeen ensisijainen teema. Lisäksi strategisten tavoitteiden ja mittarien määrittäminen jo tki-rahoitushakujen valmisteluvaiheessa voisi edistää

hankkeiden tavoitteiden saavuttamista ja parantaa niiden tulosten arvioitavuutta.

Julkisten toimien kausaalivaikutusten arviointi edellyttäisi, erityisesti satunnaistettujen koeasetelmien puuttuessa, vähintäänkin aineistoja, jotka mahdollistavat huolellisesti suunniteltujen näennäisten koeasetelmien käytön. Tällöin esimerkiksi datataloushankkeeseen tukea saaneita yrityksiä voitaisiin verrata samankaltaisiin, mutta tukea saamatta jääneisiin yrityksiin, mikä mahdollistaisi verrattain luotettavan analyysin tukitoimenpiteiden vaikutuksista.

Datatalouden edistäminen

Julkinen sektori voi tukea datatalouden kehittämistä paitsi huolehtimalla markkinoiden toimivuudesta edistämällä työvoiman saatavuutta ja yritysten toimintaedellytyksiä, myös kohdennetun tki-rahoituksen ja julkisten hankintojen avulla. Tki-rahoituksella voidaan tukea datatalouden kasvua edistäviä tutkimus- ja kehityshankkeita, kun taas julkiset hankinnat voivat luoda kysyntää datatalouden ratkaisuja hyödyntäville innovaatioille. Lisäksi julkinen sektori voi tukea datatalouden kehittymistä edistämällä pehmeän infran ratkaisuja, kuten tunnistautumisarjelmia ja datalompakoita.

Näiden lisäksi paikallisten toimenpiteiden suunnittelussa on tärkeää varmistaa niiden yhteensopivuus EU-tason sääntelyn kanssa. Säädösten velvoitteiden ja oikeuksien selkeyttäminen voisi erityisesti pk-yritysten näkökulmasta vähentää markkinoilla toimimiseen liittyvää epävarmuutta ja tukea datalähtöisen liiketoiminnan kehittämistä. Tämä on olennaista, sillä EU:n digisäädösten laaja kokonaisuus (esim. DMA, DSA, Data Act, AI Act, GDPR, PSD2) on monimutkainen ja osin tulkinnanvarainen, minkä vuoksi sen täytäntöönpanon vaikutukset tartentuvat osittain vasta tulevien oikeustapausten kautta.

Parhaiden käytäntöjen ja kokemusten jakaminen sekä kansallisesti että EU-alueella voisi osaltaan vauhdittaa datatalouden kasvua ja tukea jäsenmaita sääntelyn täytäntöönpanossa. Näin julkiset toimet voivat paitsi edistää datatalouden kehittymistä myös luoda pohjan pitkäjänteiselle kasvulle ja innovaatioille.

Suomi sijoittuu kuitenkin kansainvälisessä vertailussa datatalouden osalta keskitasoa heikommalle tasolle eri-

tyisesti data-startupien määrässä, rahoituksessa ja datalähtöisissä innovaatioissa (Koski ym., 2024). Tässä kontekstissa sääntelyn selkeyttäminen ja pk-yritysten toimintaympäristön parantaminen ovat keskeisiä edellytyksiä datatalouden kilpailukyvyyn vahvistamiselle.

Lopuksi

Dataomaisuuden arvoa voidaan mitata karkeasti suomalaisten rekisteriaineistojen perusteella käyttämällä kustannusperusteista menetelmää. Kokeellisten laskelmien mukaan datainvestointien osuus Suomen bruttokansantuotteesta oli vuosina 2012–2021 keskimäärin hieman alle kaksi prosenttia, mikä on suuruusluokaltaan samankaltainen kuin muissa maissa, joissa on käytetty vertailukelpoisia menetelmiä. On kuitenkin huomattava, että laskelmamme perustuvat tiettyihin oletuksiin, kuten kansainvälisissä tutkimuksissa käytettyihin dataperusteisen työn työaikaosuuksiin ja datan vuotuisiin arvonnmuutosprosentteihin, joiden todenmukaisuutta ei ole voitu testata. Datatalouden koon tarkempi arviointi edellyttäisi julkisen sektorin panostuksia uuden tiedonkeruun

kehittämiseksi, esimerkiksi työntekijöiden ajankäytön mittaamiseen dataperusteisessa työssä.

Vuonna 2025 valmistuva SNA-uudistus sekä Eurostatin ja IMF:n yhteisen työryhmän valmisteleva uudistusta tukeva käsikirja luovat pohjaa yhtenäisille kansainvälisille tilinpäätöskäytännöille datan käsittelemiseksi varallisuuseränä. Suomi voisi aktiivisesti edistää EU-tason harmonisoitujen kyselyjen ja mittareiden kehittämistä, jotta eri maista saataisiin vertailukelpoista tietoa. Kansainväliset vertailut tarjoaisivat tärkeää näkökulmaa Suomen datatalouden asemaan ja kehitykseen suhteessa muihin maihin.

Datatalouden kehitystä tukevien julkisten toimien, kuten hankintojen ja tki-rahoituksen, onnistumisen arviointia voitaisiin parantaa kehittämällä aineistoja ja mittaristoja, jotka mahdollistavat datataloushankkeiden selkeämmän tunnistamisen ja seurannan. Lisäksi pk-yritysten näkökulmasta säädösten velvoitteiden ja oikeuksien selkeyttäminen voisi vahvistaa datalähtöisen liiketoiminnan edellytyksiä. Parhaiden käytäntöjen ja kokemusten jakaminen datataloutta edistävästä toimenpiteistä voisi puolestaan edistää datatalouden kasvua EU-alueella.

Viite

- ¹ Kosken ym. (2024) raportissa keskustellaan muista epäsuorista menetelmistä, kuten tuloperusteiset ja markkina-arvoperusteiset menetelmät.

Kirjallisuus

de Bondt, H. & Muskhudiani, H. (2021). Estimating the Value of Data in the Netherlands. Paper prepared for the IARIW-ESCoE Conference November 11–12, 2021.

Calderón, J. & Rassier, D. (2022). Valuing the US data economy using machine learning and online job postings, paper prepared for the 37th IARIW General Conference.

Goodridge, P., Haskel, J. & Edquist, H. (2022). We see data everywhere except in the productivity statistics. *Review of Income and Wealth* 68, 4.

IDC (2021). The European Data Market Study 2021–2023. <https://digital-strategy.ec.europa.eu/en/library/results-european-data-market-study-2021-2023>

Koski, H., Anttila, J., Björk, A., Djakonoff, V., Kässi, O., Niemi-Hugaerts, H., Pajarinen, M. & Parkkari, J-P. (2024). Datatalouden mittaaminen ja vaikuttavuus. *Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja* 2024/44.

Koski, H. & Pajarinen, M. (2024). *Datasäädös – Tietopohja markkinoista*. Etlä raportti nro 148.

Smedes, M., Nguyen, T. & Tenburren, B. (2022). Valuing data as an asset, implications for economic measurement, paper prepared for Economic Implications of the Digital Economy Conference, 9–10 March, 2022.

Statistics Canada (2019). The value of data in Canada: Experimental estimates.

ETLA



Elinkeinoelämän tutkimuslaitos

ETLA Economic Research

ISSN-L 2323-2463
ISSN 2323-2463

Kustantaja: Taloustieto Oy

Puh. 09-609 900
www.etla.fi
etunimi.sukunimi@etla.fi

Arkadiankatu 23 B
00100 Helsinki
