



Tekoäly, koneoppiminen ja datatoimijuus

Henriikka Vartiainen & Matti Tedre

Tekoäly ja koneoppiminen ovat monin tavoin läsnä arkipäiväisessä elämässämme. Koneoppiminen suodattaa roskaposteja, neuvoo meille reittivaihtoehtoja sekä auttaa löytämään meitä kiinnostavia uutisia, musiikkia tai elokuvia. Yhtälailla käytämme sosiaalista mediaa ja erilaisia pikaviestinpalveluja yhteydenpidon välineinä tai vaikkapa harrasteryhmien toiminnan organisointiin. Voimme myös mitata omaa hyvinvointia tai liikuntasuorituksia, valvoa kodin turvajärjestelmiä sekä seurata miten robotti-imuri huolehtii kodin puhtaudesta. Koneoppimisen ohjaamana saamme myös juuri meille suunnattuja mainoksia ja tiedotteita tuotteista, tapahtumista ja palveluista, joita emme aina edes tienneet tarvitsevamme (Valtonen ym., 2021).

Westin (2019) mukaan uusia sosiaalisen median palveluja luonnehditaan tyypillisesti henkilökohtaisen, sosiaalisen ja demokraattisten etujen näkökulmasta. Samaan tapaan älykkäiden tuotteiden myyntipuheissa esitetään niiden tarjoavan räätälöityä tukea oppimiseen, liikunnallisten elämäntapojen ja hyvinvoinnin edistämiseen, sairauksien ehkäisyyn ja ylipäättään elämänlaadun parantamiseen (Lupton, 2018). Samalla kun uusien teknologioiden ja palvelujen käyttäjistä on tullut osallistujia, sisällöntuottajia ja oman hyvinvoinnin mittaajia, henkilökohtaisista datajäljistä on tullut globaalia kauppatavaraa. Datan keruu on arkipäiväisessä elämässämme usein näkymätöntä, emmekä aina ole tietoisia kolikon kääntöpuolella olevasta datan keruun laajuudesta, ongelmista ja riskeistä.

Tässä kirjoituksessa tarkastelemme sosiaalisen median ansaintalogiikkaa sekä siihen sisältyviä potentiaalisia riskejä erityisesti yksityisyydensuojan sekä algoritmisen vaikutuksen näkökulmista. Tämä kirjoitus pohjautuu Itä-Suomen yliopistolla (UEF) toimivan monitieteisen tutkimusryhmän yhteistutkimuksiin ja -julkaisuihin, joissa olemme tarkastelleet miten lasten sekä nuorten toimijuutta voidaan tukea koneoppimisen muo-
vaamassa maailmassa.

Miksi dataa kerätään?

Zuboffin (2015) mukaan erityisesti Amazonin, Facebookin, Applen ja Googlen kaltaisten suuryritysten ansaintamallit perustuvat massiiviseen käyttäjätiedon keräämiseen sekä sen jälleenyymiseen mainostajien käyttöön. Palveluista eivät siten maksa niiden käyttäjät vaan mainostajat, jolloin myytävä tuote on palveluiden käyttäjistä kerätty tieto (Zuboff, 2015). Käytännössä laajat aineistot voivat olla esimerkiksi klikkauksia, tykkäyksiä, ostotapahtumia, päivityksiä, selaushistoriaa ja sijaintitietoja, joita syntyy palvelun käytön sivutuotteena. Kun ihmisen toimintaa tarkastellaan historiallisena kehitysprosessina, dataa yhdistellään esimerkiksi asiakas- ja luottokorttitietoihin sekä verataan toisten ihmisten tekemisiin, on myös mahdollista rakentaa erittäin tarkkoja profiileja sekä yksittäisistä ihmisistä että erilaisista kuluttajaryhmistä ja -segmenteistä (Chen ym., 2015; Valtonen ym., 2019). Yritykset voivat esimerkiksi analysoida eri käyttäjien profiilien samankaltaisuutta ja tehdä varsin osuvia ennustuksia siitä, mitä eri kuluttajasegmenteille kannattaa mainostaa tai millaisia sisältöjä heille näytetään, jotta he koukuttuvat palvelun käyttöön (Chen ym., 2015). Toisin sanoen, algoritmien ohjaamien profilointi- ja personointiteknologioiden seurauksena eri käyttäjät saavat erilaista sisältöä ja kohdennettuja mainoksia riippuen siitä, mitä palvelut tietävät käyttäjästä, hänen mieltymyksistään sekä hänen käyttöhistoriastaan (Kramer ym., 2014; Tufekci, 2015).

Koneoppimisen ja massadatan erityinen potentiaali piilee niiden kyvyssä paljastaa ihmisten käyttäytymisestä ja arvomaailmasta asioita, joista emme aina ole itsekään tietoisia. Henkilökohtaisen datan perusteella voidaan tehdä pitkälle meneviä päätelmiä esimerkiksi ihmisten arvoista, asenteista, verkostoista, sosiaalisesta statuksesta tai vaikkapa sukupuolisesta suuntautuneisuudesta (Kosinski ym., 2013). Esimerkiksi Youyou kollegoineen (2015) osoittivat tutkimuksessaan, kuinka Facebookin algoritmi ennusti 10 tykkäyksen perusteella henkilön persoonallisuutta tarkemmin kuin hänen kollegansa ja 300 tykkäyksen perusteella paremmin kuin heidän puolisonsa tai läheiset ystävät. Kosinkin ja kumppaneiden (2013) tutkimuksessa puolestaan osoitettiin, miten Facebookin tykkäyksistä voidaan ennustaa esimerkiksi seksuaalista suuntautumista, etnistä taustaa, onnellisuutta, uskonnollisia ja poliittisia näkemyksiä sekä riippuvuutta aiheuttavien aineiden käyttöä. Vaikuttamisen mahdollisuuksia on myös kuvattu Kramer ym. (2014) kohua herättäneessä tutkimuksessa, jossa Facebook ohjasi viikon ajan sitä, millaisia päivityksiä lähes 700 000 ihmistä näkivät uutisvirrassaan. Tutkimus osoitti, että käyttäjien tunnetiloja voidaan todellakin ohjata, sillä käyttäjät alkoivat käyttää päivityksissään tilastollisesti merkittävästi enemmän negatiivisia tai positiivisia sanoja sen mukaan, mille tunnetilalle heidät oli tietämättään altistettu (Kramer ym., 2014).

Näkymätöntä ohjailua hyödynnetään myös poliittisessa vaikuttamisessa. Esimerkiksi Bond ym. (2012) osoittivat tutkimuksessaan, miten sosiaalisen median välityksellä voidaan vaikuttaa kohderyhmän äänestyskäyttäytymiseen. Kenties tunnetuin esimerkki poliittisesta vaikuttamisesta on kuitenkin Cambridge Analytica -yhtiön rooli

Yhdysvaltain presidentinvaaleissa 2016. Tällöin kymmenien miljoonien äänestäjien käyttäytymistä ohjailtiin kohdentamalla personoituja mainoksia ja valeuutisia valituille kohderyhmille (Isaak & Hannah, 2018). Tiedon kerääminen, profilointi ja personoidut mediasisällöt ovat siten osa algoritmista vaikuttamista, jonka kautta voidaan ohjata ihmisten valintoja ja käyttäytymistä sekä kaupallisten, ideologisten että poliittisten intressien suuntaan. Vihman ja kumppaneiden (2018) mukaan elämmekin niin sanottua totuuden jälkeistä aikaa, jossa kybervaikuttaminen, tunteisiin vetoaminen, populismi, propaganda ja kohdennetut valeuutiset leviävät kuin tartuntataudit.

Internetin pioneereihin kuuluvan Jaron Lanierin (2018) mukaan sosiaalisen median kautta tapahtuva käyttäytymisen muokkaaminen perustuu pitkälti behaviorismista tuttuihin periaatteisiin. Sosiaalisen median sovellukset suunnitellaan addiktoiviksi ja käyttäjille syötetään jatkuvasti tunnereaktioita aiheuttavia ärsykyksiä, joiden perusteella saadaan tietoa siitä, miten niihin reagoimme. Näin järjestelmät myös oppivat reaktioitamme suhteessa erilaisiin ärsykkeisiin, jolloin myös käyttäytymisen ohjaaminen on entistä tehokkaampaa (Lanier, 2018). Huomionarvoista myös on, neuroverkkoihin perustuvat menetelmät toimivat pitkälti niin sanottuina ”mustina laatikoina” (engl. black box), sisältäen miljoonia tai jopa miljardeja parametreja, joiden merkitystä edes kehittäjät itse eivät pysty riittävän näkyvästi analysoimaan. Sosiaalisen median algoritmien toimintalogiikat ovat tyypillisesti myös tarkkaan varjeltuja liikesalaisuuksia, joiden kehitystä johtavat kansainväliset suuryritykset. Algoritmien on myös osoitettu vahvistavan sen käyttämässä datassa olevia virheitä, vinoutumia sekä syrjinnän mekanismeja, jotka voivat liittyä esimerkiksi sukupuoleen tai ihonväriin (Ollila, 2019). Yhdeksi datayhteiskunnan polttavaksi eettiseksi kysymykseksi onkin noussut algoritmien diskriminaatio, koska algoritmisia päätöksentekojärjestelmiä hyödynnetään joko itsenäisesti tai yhdessä ihmisten kanssa esimerkiksi rekrytointiprosesseissa, vakuutuspäätöksissä, lainapäätöksissä sekä osana tuomioistuinten toimintaa.

Lapsuuden datafikaatio

Viime vuosina tutkijat ovat myös nostaneet esiin lukuisia huolia siitä, miten arjen datafikaatio vaikuttaa lasten elämään ja toimijuuteen. Luptonin ja Williamsonin (2017) mukaan tilanne on erityisen huolestuttava lasten ja nuorten näkökulmasta, sillä jo hyvin pienistä lapsista jaetaan valtavasti tietoa kansainvälisten suuryritysten palveluihin. Dataa tuottavat sekä lapset itse että heidän läheisensä, jotka eivät useinkaan ymmärrä miten, missä tai miksi tietoja kerätään ja mitä riskejä tiedonkeruuseen liittyy (Pangrazio & Selwyn, 2019). Vaikka aikuisten oletetaan turvaavan lasten yksityisyyttä, datan keruu voi myös tapahtua ilman huoltajien tai kasvattajien tietoisuutta (Livingstone & Yoo, 2018; Mertala, 2020).

Lasten henkilökohtaisista tiedoista voi kuitenkin rakentua pitkiä elämänkaarua, alkaen sosiaalisen median palveluissa jaetuista ultraäänikuvista aina videoihin ja valokuviiin vauvan ensiaskelista, lapsuuden harrastuksista, koulutodistuksista, seurustelusuhteista sekä perhe-elämästä (van Dijck, 2013). Tänä päivänä lapsista kerätään myös hyvin intiimiä dataa erilaisten puettavien vaatteiden, valvontalaitteiden, älylelujen ja kaupallisten sovellusten kautta, joita voidaan käyttää esimerkiksi lasten kasvun, kehityksen, liikkeiden, hengityksen, sykkeen, terveyden ja nukkumis- ja ruokintatottumusten seurantaan (Holloway, 2019; Lupton & Williamson, 2017; Mertala, 2020). Huolta ovat myös herättäneet erilaiset kotiautomaatiojärjestelmät, jotka yhtälailla voivat tuottaa dataa lasten arjesta kaupallisille toimijoille (Holloway, 2019). Aikalaiskriitikoiden mukaan meidän tulisi olla huolissamme siitä, miten lasten kasvuympäristö, liikkeet, sijainti, leikki, sosiaaliset suhteet, tunnetilat ja keholliset reaktiot ovat enenevässä määrin kietoutumassa osaksi kaupallisten toimijoiden kudosta. Meillä ei kuitenkaan ole tietoa siitä, miten lapsista kerättyä dataa käytetään, mihin tarkoitukseen, ketkä kaikki sitä hyödyntävät ja mitä datalle tapahtuu tulevaisuudessa (Holloway, 2019; Leaver, 2017).

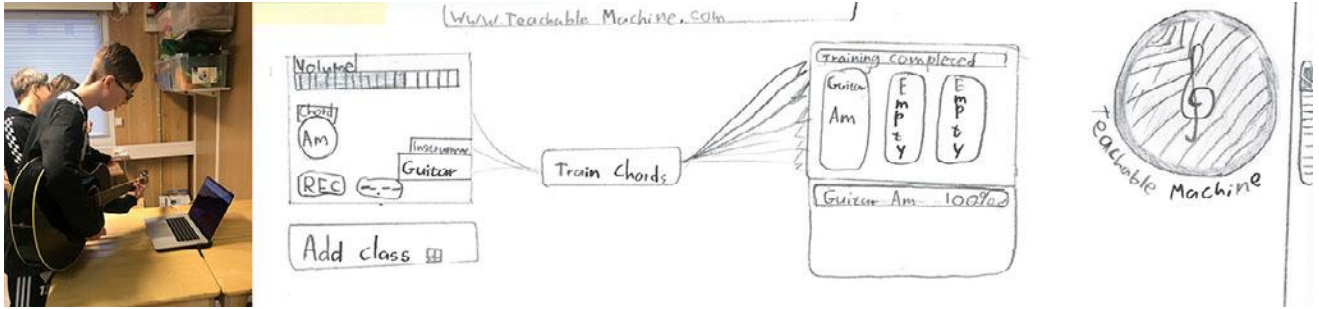
Lapsiin kohdistuvan datan luovuttamista on kuitenkin pyritty säätelemään esimerkiksi tietosuojasetuksen (GDPR) sekä palvelujen tarjoajien asettamien ikärajoitusten kautta. Toisaalta tietosuojasetus jättää vastuuta päätöksestä paljolti huoltajille ja nuorille itselleen. Tällöin heidän myös epäsuorasti odotetaan tekevän rationaalisia päätöksiä suhteessa palvelujen tarjoajien esittämiin pitkiin ja usein kovin vaikeaselkoiisiin lakiteksteihin (Keen, 2020). Tutkimuksissa on myös arvioitu, että monet palveluntarjoajien ikärajoitukset luovat lähinnä illuusiota lapsiin potentiaalisesti kohdistuvien riskien huomioimisesta, sillä ikärajoitukset ovat helposti kierrettävissä esimerkiksi antamalla vääriä tietoja tai käyttämällä lahjakortteja tai vanhempien omia käyttäjätunnuksia (Smirnova ym., 2021).

Tutkimusten mukaan lapset ja nuoret kuitenkin välittävät yksityisyydestään, mutta kokevan itsensä usein voimattomaksi vaikuttaa siihen, millä tavoin heistä kerättyä dataa kerätään, prosessoidaan ja hyödynnetään kaupallisiin tarkoituksiin (Selwyn & Pangrazio, 2018). Hargittain ja Marwickin (2016) mukaan tiedonkeruun normalisaatio voi myös synnyttää apatiaa, kyynisyyttä ja uskomuksia siitä, että yksityisyyden loukkaukset ovat väistämättömiä. Hälyttävää myös on, että lapset saavat hyvin vaihtelevasti opastusta vanhemmilta sekä opettajilta yksityisyyden suojaan liittyvissä kysymyksissä, joka luo kasvavaa eriarvoisuutta sekä asettaa osan lapsista haavoittuvampaan asemaan (Stoilova et ym., 2020).

Onko datalukutaito ratkaisu?

Halusimme tai emme, olemme jo siirtyneet uudenlaiseen datatalouden aikakauteen, jossa ihmisen ajattelu, toiminta ja vuorovaikutus on entistä enemmän koneoppimisen ohjaamaa (Hallinan & Striphos, 2016). Vastauksena edellä kuvattuihin haasteisiin ovat monet tutkijat peräänkuuluttaneet datalukutaidon sisällyttämistä kouluopetukseen. Datalukutaidosta on erilaisia määritelmiä, mutta tyypillisesti siihen sisältyy kyky tunnistaa, miten ja missä henkilökohtaista dataa luodaan ja käsitellään, kyky hallita henkilökohtaista dataa ja datakäytänteitä sekä kyky osallistua datayhteiskunnan käytänteisiin (Data-Pop Alliance and Internews, 2015; McCosker, 2017; Pangrazio & Selwyn, 2019). Tavoitteena on siten tarjota välineitä tietoisempiin valintoihin erilaisten palveluiden käytön sekä niiden arjen valintojen suhteen, joihin palveluilla pyritään vaikuttamaan (Valtonen ym., 2021). Lehtiniemen ja Haapojan (2020) mukaan datalukutaidon ytimessä onkin tukea ihmisiä olemaan aktiivisia ja vastuullisia datatoimijoita, eikä vain passiivisia datankeruun kohteita. Datalukutaito nähdään myös tärkeässä roolissa kansalaisaktiivisuuden tukemisessa sekä sosiaalisessa inkluusiossa, jossa kansalaisilla on sekä mahdollisuus, halu että kyky vaikuttaa siihen, miten ja kenen toimesta henkilökohtaisia tietoja kerätään, mihin niitä käytetään ja kenen toimesta (Data-Pop Alliance and Internews, 2015).

Datalukutaitojen kehittämistä onkin pidetty erityisen tärkeänä lasten ja nuorten toimijuuden näkökulmasta (Tedre ym., 2020). Aktiivinen toimijuus, osallistuminen yhteiskunnan toimintoihin sekä kyky ymmärtää teknologioiden merkitys arjessa nousevat myös vahvasti esiin opetussuunnitelman tavoitteenasettelussa. Toisaalta tutkimusten mukaan datakapitalismin käytänteet sekä niihin kytkeytyvät koneoppimisen mekanismit ovat yleensä "mustia laatikoita" sekä oppilaille itselleen että heidän opettajilleen ja huoltajilleen (Livingstone ym., 2019; Vartiainen ym., painossa). Viime aikoina monet tutkijat ja kansainväliset työryhmät ovatkin tuottaneet erilaisia verkkokursseja, oppimateriaaleja sekä työvälineitä, joiden tuella voidaan tarkastella koneoppimisen peruskäsitteitä, mekanismeja sekä yhteiskunnallisia vaikutuksia (Tedre ym., 2021). Esimerkiksi Reaktor ja Helsingin yliopisto ovat yhdessä tuottaneet kaikille avoimia ja ilmaisia verkkokursseja teköälyn perusteista (<https://www.elementsofai.com/fi/>). Itä-Suomen yliopistossa on puolestaan kehitetty pedagogisia malleja sekä opetusteknologioita, joiden tuella eri-ikäiset oppijat ovat tutustuneet koneoppimisen mahdollisuuksiin ja mekanismeihin. Koneoppimisen kouluprojekteissa on käytetty Joensuussa kehitettyä sovellusta (Mariescu-Istodor & Jormanainen, 2019) sekä Googlen Teachable Machine (GTM) oppimisympäristöä (Vartiainen ym., 2020; Vartiainen ym., 2021). Koneoppimisen kouluprojekteissa oppilaat ovat suunnitelleet, toteuttaneet ja testanneet omia teköälysovelluksia. Samalla heitä on myös haastettu tutkimaan datan laatua sekä sitä, miten järjestelmään syntyy virheitä ja vinoutumia.



Alakoululaiset suunnittelemassa, toteuttamassa ja testaamassa koneoppimisen sovelluksia.

Koneoppimista sekä sen roolia arjessa voidaan myös lähestyä tarkastelemalla nuorten käyttämien sovellusten toimintaa, käyttöehtoja sekä omia datakäytänteitä (Valtonen ym., 2019). Yhtälailla oppijoiden kanssa voidaan tutkia myös sitä, miten koneoppimisen järjestelmät vaikuttavat yhteiskunnassa, kuinka koneoppimisen järjestelmät voivat vahingoittaa tai auttaa ihmisiä sekä pohtia kysymyksiä siitä, mitä ei pitäisi automatisoida ja kenen siitä pitäisi saada päättää? Tekoälyyn, koneoppimiseen sekä yleisemmin digitalisaatioon ja datafikaatioon liittyvien tietojen ja taitojen oppimisen tueksi olemme myös parhaillaan tuottamassa uudenlaista oppimiskokonaisuutta Helsingin yliopiston ja Itä-Suomen yliopiston yhteisessä Tekoäly ja lapset -hankkeessa.

Viime aikoina lasten datatoimijuutta koskeva keskustelu ja siihen liittyvät vaatimukset ovat myös ulottuneet opetuksen kehittämistä laajemmaksi. Kiistämättä opetuksen kehittämisen tärkeyttä, Sonia Livingstone on esittänyt aiheellisen kysymyksen siitä, onko opetuksen kehittäminen riittävä ratkaisu ongelmiin ja riskeihin, joita hallitsevien teknologiarytysten toiminta tuo lasten ulottuville? Tutkijat ovatkin peräänkuuluttaneet teknologiarytysten vastuunkantoa sekä omaehtoisia toimia lasten oikeuksien kunnioittamiseksi jo siinä vaiheessa, kun uudenlaisia tuotteita ja palveluja kehitetään (Lupton & Williamson, 2017; Stoilova ym., 2020; UNICEF, 2018). On selvää, että datakeruun kaikkiallisuus ja tekoälyteknologioiden kehittyminen tuo myös väistämättä pohdittavaksi lukuisia muita filosofisesti, oikeudellisesti ja yhteiskunnallisesti tärkeitä kysymyksiä, joihin ei ole olemassa yhtä ainoaa totuutta (Ollila, 2019). E erityisen tärkeää kuitenkin on, että uusien mahdollisuuksien ja riskien ymmärtämisestä käydään mahdollisimman laajaa-alaista ja avointa keskustelua tavalla, jossa myös lasten ääni tulee kuulluksi. Maailmassa, jossa datasta on tullut kaupankäynnin, kontrollin ja vaikuttamisen väline, aikamme keskeisimmät kysymykset liittyvät esimerkiksi lasten oikeuksien ja hyvinvoinnin turvaamiseen. Keskustelua on käytävä ja kriittisiä kysymyksiä on esitettävä, vaikkei vastauksia ei vielä olisikaan.

Kiitämme Tammikuun ryhmää.

Tekijät:

Vartiainen, H. & Tedre, M. (2022). *Tekoäly, koneoppiminen ja datatoimijuus*. LUKILOKI-hanke 2018–2022. Jyväskylän yliopisto.

Lähteet:

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415), 295–298.

Chen, F., Deng, P., Wan, J., Zhang, D., Vasilakos, A. V., & Rong, X. (2015). Data mining for the internet of things: Literature review and challenges. *International Journal of Distributed Sensor Networks*, 2015.

Data-Pop Alliance and Internews. (2015). Beyond data literacy: reinventing community engagement and empowerment in the age of data. <https://datapopalliance.org/wp-content/uploads/2015/11/Beyond-Data-Literacy-2015.pdf>

Hallinan, B., & Striphas, T. (2016). Recommended for you: The Netflix Prize and the production of algorithmic culture. *New Media and Society*, 18(1), 117–137.

Hargittai, E., & Marwick, A. (2016). “What can i really do?” Explaining the privacy paradox with online apathy. *International Journal of Communication*, 10, 3737–3757.

Holloway, D. (2019). Surveillance capitalism and children’s data: the Internet of toys and things for children. *Media International Australia*, 170(1), 27–36.

Isaak, J., & Hannah, M. J. (2018). User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection. *Computer*, 51(8), 56-59.

Keen, C. (2020). Apathy, convenience or irrelevance? Identifying conceptual barriers to safeguarding children’s data privacy. *New Media and Society*, 24(1),50-69.

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Digital records of behavior expose personal traits. *Proceedings of the National Academy of Sciences*, 110(15), 5802–5805.

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 111(24), 8788–8790.

Lanier, J. (2018). *Ten arguments for deleting your social media accounts right now*. London: The Bodley Head.

Leaver, T. (2017). Intimate surveillance: normalizing parental monitoring and mediation of infants online. *Social Media and Society*, 3(2).

Lehtiniemi, T., & Haapoja, J. (2020). Data agency at stake: MyData activism and alternative frames of equal participation. *New Media and Society*, 22(1), 87–104.

Livingstone, S, & Yoo, D. (2018). [More clarity brings more confusion debating what the European General Data Protection Regulation means for children in the UK.](#)

Livingstone, Sonia, Stoilova, M., & Nandagiri, R. (2019). Talking to children about data and privacy online: research methodology. <https://www.lse.ac.uk/media-and-communications/assets/documents/research/projects/childrens-privacy-online/Talking-to-children-about-data-and-privacy-online-methodology-final.pdf>

Lupton, D. (2018). 'Better understanding about what's going on': young Australians' use of digital technologies for health and fitness. *Sport, Education and Society*, 25(1), 1–13.

Lupton, D., & Williamson, B. (2017). The datafied child: The dataveillance of children and implications for their rights. *New Media and Society*, 19(5), 780–794.

Mariescu-Istodor, R., & Jormanainen, I. (2019). Machine learning for high school students. *ACM International Conference Proceeding Series*.

McCosker, A. (2017). Data literacies for the postdemographic social media self. *First Monday*, 22(10).

Mertala, P. (2020). [Data \(il\)literacy education as a hidden curriculum of the datafication of education](https://doi.org/10.23860/JMLE-2020-12-3-4). *Journal of Media Literacy Education*, 12(3), 30–42. <https://doi.org/10.23860/JMLE-2020-12-3-4>

Ollila, M.-R. (2019). *Tekoälyn etiikkaa*. Otava.

Pangrazio, L., & Selwyn, N. (2019). 'Personal data literacies': A critical literacies approach to enhancing understandings of personal digital data. *New Media and Society*, 21(2), 419–437.

Selwyn, N., & Pangrazio, L. (2018). Doing data differently? Developing personal data tactics and strategies amongst young mobile media users. *Big Data and Society*, 5(1).

Smirnova, S., Livingstone, S., & Stoilova, M. (2021). Understanding of user needs and problems: a rapid evidence review of age assurance and parental controls. <https://euconsent.eu/download/understanding-of-user-needs-and-problems-a-rapid-evidence-review-of-age-assurance-and-parental-controls/>

Stoilova, M., Livingstone, S., & Nandagiri, R. (2020). Digital by default: Children's capacity to understand and manage online data and privacy. *Media and Communication*, 8(4), 197–207.

Tedre, M., Toivonen, T., Kahila, J., Vartiainen, H., Valtonen, T., Jormanainen, I., & Pears, A. (2021). [Teaching machine learning in K-12 Classroom: Pedagogical and technological trajectories for artificial intelligence education](https://doi.org/10.1109/ACCESS.2021.3075558). *IEEE Access*, 9, 110558–110572.

Tedre, M., Vartiainen, H., Kahila, J., Toivonen, T., Jormanainen, I., & Valtonen, T. (2020). Machine Learning Introduces New Perspectives to Data Agency in K - 12 Computing Education. *Proceedings - Frontiers in Education Conference, FIE*, 2020-October.

Tufekci, Z. (2015). Algorithmic Harms beyond Facebook and Google: Emergent Challenges of Computational Agency. *Journal on Telecommunications & High Tech Law*, 13(23), 203–216.

UNICEF. (2018). [Children's online privacy and freedom of expression](https://www.unicef.org/childrens-rights-index/childrens-online-privacy-and-freedom-of-expression).

Valtonen, T., Tedre, M., Mäkitalo, K., & Vartiainen, H. (2019). [Media Literacy Education in the Age of Machine Learning](#). *Journal of Media Literacy Education*, 11(2).

Valtonen, T., Vartiainen, H., Tedre, M., Toivonen, T., Kokko, A., Kahila, J., Jormanainen, I., & Piispa-Hakala, S. (2021). Oppilaat digitalisoituneen yhteiskunnan toimijoina. *Kasvatus*, 52(1), 108–113.

van Dijck, J. (2013). “You have one identity”: Performing the self on Facebook and LinkedIn. *Media, Culture and Society*, 35(2), 199–215.

Vartiainen, H., Pellas, L., Kahila, J., Valtonen, T., & Tedre, M. (n.d.). Pre-service teachers' insights on data agency. *New Media and Society*.

Vartiainen, H., Tedre, M., Jormanainen, I., Kahila, J., Valtonen, T., & Toivonen, T. (2021). [Teko-äly, koneoppiminen ja teknologinen murros:Kohti datatoimijuutta ja tulevaisuuden design-taiteja](#). *Ainedidaktiikka*, 5(2), 03–120.

Vartiainen, H., Tedre, M., & Valtonen, T. (2020). Learning machine learning with very young children: Who is teaching whom? *International Journal of Child-Computer Interaction*, 25.

Vartiainen, H., Toivonen, T., Jormanainen, I., Kahila, J., Tedre, M., & Valtonen, T. (2021). [Machine learning for middle schoolers: Learning through data-driven design](#). *International Journal of Child-Computer Interaction*, 29, 100281.

Vihma, A., Hartikainen, J., Ikäheimo, H.-P., & Seuri, O. (2018). *Totuuden jälkeen : miten media selviää algoritmien ja paskapuheen aikana*. Helsinki: Teos.

West, S. M. (2019). Data Capitalism: Redefining the Logics of Surveillance and Privacy. *Business and Society*, 58(1), 20–41.

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences of the United States of America*, 112(4), 1036–1040.

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1), 75–89.