

Massadata-arkkitehtuuri sosiaali- ja terveydenhuollossa: Tiedonhallinnan yleislainsäädännöstä seuraavat vaatimukset henkilötietojen käsittelylle

Mikko Huovila

Sosiaali- ja terveysjohtamisen laitos, Itä-Suomen yliopisto, Kuopio

Mikko Huovila, Väitöskirjatutkija, YTM, Sosiaali- ja terveysjohtamisen laitos, Itä-Suomen yliopisto, PL 1627, 70211 Kuopio. Sähköposti: mhuovila@student.uef.fi

Tiivistelmä

Suurten tietomassojen käyttöä kohtaan on merkittäviä odotuksia sosiaali- ja terveydenhuollossa. Tiedon jalostamisen kautta pyritään tuottamaan uudenlaista ymmärrystä palveluiden kehittämiseksi mm. tutkimuksen, innovaatio toiminnan ja tietojohdamisen keinoin. Käsiteltäessä henkilötietoja on kuitenkin huomioitava rekisteröidyn oikeudet ja henkilötietojen käsittelyä koskevat säädökset. Tämä korostuu erityisesti tilanteissa, joissa käsittelyn kohteena on arkaluontoiset tiedot, kuten sosiaali- ja terveydenhuollossa syntyvät asiakastiedot.

Massadata-arkkitehtuureja suunniteltaessa on huomioitava sosiaali- ja terveydenhuollon tietojen käsittelyä ohjaavan lainsäädännön lisäksi tiedonhallintaa ohjaava yleislainsäädäntö eli EU:n yleinen tietosuojasetus, tiedonhallintalaki, julkisuuslaki, arkistolaki ja tietosuojalaki. Aiemmassa tutkimuksessa ei ole käsitelty näiden muodostamaa kokonaisuutta massadatatoteutusten näkökulmasta.

Tässä tutkimuksessa tavoitteena oli tunnistaa tiedonhallinnan yleislainsäädännöstä seuraavat vaatimukset massadata-arkkitehtuureille erityisesti henkilötietojen käsittelyn näkökulmasta. Tutkimuksen tuloksena lainsäädännöstä seuraavat vaatimukset ryhmiteltiin viiteen luokkaan: 1) hallinnolliset vaatimukset, 2) henkilötietojen suojaaminen, 3) tiedonhallinta ja tietoturva, 4) rekisteröidyn oikeudet ja läpinäkyvyys sekä 5) tiedonsiirrot. Lisäksi tunnistettiin lainsäädännön vaatimusten vaikutukset massadata-arkkitehtuurin toimijoihin.

Tyypillisten massadata-arkkitehtuurien ja lainsäädännön yhteensovittaminen ei ole ongelmatonta. Jännitteisyyttä liittyy erityisesti käyttötarkoitussidonnaisuuden huomiointiin, tietojen käsittelyn rajaamisen välttämättömään eli miniminointiperiaatteen toteutumiseen, tiedon oikeellisuuden varmistamiseen ja säilytyksen rajoittamiseen sekä tiedonkäsittelyn läpinäkyvyyden toteuttamiseen. Näissä kohdissa lainsäädännön soveltaminen vaatii erityistä huolellisuutta.

Sekä tietosuojasetus, että tiedonhallintalaki korostavat organisaatioiden johdon viime kätistä vastuuta rekisteröityjen oikeuksien suojaamisessa sekä hyvän hallinnon vaatimusten toteutumisessa viranomaisen tiedonhallinnassa. Riskien hallitsemiseksi, sisäänrakennetun ja oletusarvoisen tietosuojan sekä hyvän

Published under a CC BY 4.0 license (<https://creativecommons.org/licenses/by/4.0/>).

tiedonhallintatavan toteutumiseksi olennaisia välineitä ovat lakisäätteiset tietosuojavaikutusten arvioinnit sekä tiedonhallinnan muutosvaikutusten arvioinnit sekä tiedonhallintayksikön antamat ohjeet, koulutus sekä sisäinen valvonta. Sovellettavan lainsäädännön kokonaisuuden monimutkaisuuden vuoksi aihetta käsittelevä viranomaisohjeistus tai viranomaisille suunnattu kansallinen massadataviitearkkitehtuuri olisi hyödyllinen. Aihepiiriä koskevalle empiiriselle ja lainopilliselle lisätutkimukselle olisi tarvetta.

Avainsanat: tiedonhallinta, massadata, lakien toimeenpano, kokonaisarkkitehtuuri, henkilötiedot

Abstract

There are big expectations for the utilisation of data masses in health and social care. By processing data, there is an aim to create a new understanding of service development through research, innovation activities and knowledge management. When processing personal data, there are requirements to take into account citizens' rights and norms related to personal data processing. That is highlighted when processing is related to sensitive data, like health and social care data.

When designing and implementing Big Data architectures, there is a need to consider the general law of information management in addition to social and healthcare legislation (General Data Protection Regulation, Information Management Act, Act on the Openness of Government Activities, Data Protection Act and Archives Act). There has yet to be any earlier research on the overall effect of this legislation on big data implementations.

This article aims to recognise requirements from the general law of information management to big data architectures, especially from the personal data processing point of view. In this article, the main research results are five categories originating from the legislation: 1) Administrative requirements, 2) Personal data protection, 3) Information management and security, 4) Data subject's rights and transparency and 5) Data transfers. In addition, requirements from legislation are connected to big data architecture roles.

Integrating typical Big Data architectures and legislative requirements is not problem-free. There is tension, especially in purpose limitation, data minimisation, data accuracy and storage limitation and data processing transparency. When applying these requirements, there is a need for caution.

GDPR and Information Management Act emphasise management's last-hand responsibility for protecting data subjects' rights and good governance requirements in authorities' information management. Data protection impact assessments and information management change assessments are essential tools for risk management and data protection by default and design. Also, information management entities' instructions, education and internal auditing are important. Because applicable legislation is complex as a whole, official guidelines or mass data reference architecture could be helpful to authorities. More empirical and legal dogmatic research related to mass data use in health and social care is needed.

Keywords: information management, big data, legislation, enterprise architecture, personal data

Johdanto

Massadata ilmiönä

Massadatalta (eng. Big Data) tarkoitetaan toisaalta datan määrän valtavaa kasvua sekä toisaalta teknologioita, joilla kasvavasta ja laajasta datapohjasta pyritään saamaan jalostamalla uudenlaista arvokasta tietoa ja ymmärrystä. Käsitteen voidaan katsoa toisaalta viittaavaan yhteiskunnalliseen paradigmaan, jonka myötä datan kautta luodaan uudenlaista arvoa ja toisaalta tietotekniisiin ratkaisuihin, jotka mahdollistavat kasvavien datamäärien tehokkaan käsittelyn. Massadata on kiinteä osa uusia dataa hyödyntäviä teknologioita, kuten analytiikkaa ja tekoälyä sekä näitä hyödyntävää laajempaa alustataloutta [1–4].

Laajaa ihmiselämästä syntyvän tiedon keräämistä, tallentamista ja analysointia kutsutaan datafikaatioksi. Siinä keskeistä on ilmiöiden muuttaminen mitattavaan muotoon kvantifioimalla ja luokittelemalla niitä dataksi [5,6].

Suurten datamassojen hyödyntämiseen ja analysointiin on kohdistettu paljon toiveita myös julkishallinnossa niin Suomessa kuin kansainvälisestikin [1–3,7,8]. Datafikaatiolta ja massadatalta odotetaan paljon mm. kaupunkien kehittämisessä, liikennesektorilla, koulutuksessa, tutkimuksessa, sosiaali- ja terveyspalveluissa ja “tietosiilojen” purkamisessa [1–3,7–19].

Suomalaisessa julkishallinnossa on otettu käyttöön massadataratkaisuja viime vuosina erityisesti sosiaali- ja terveydenhuollossa. Käytännössä tämä on toteutunut esimerkiksi ns. tietoaallas-ratkaisujen leviämisenä [12,20,21].

Massadata on tässä artikkelissa käännetty englanninkielisestä käsitteestä Big Data. Käsitteelle on esitetty hiukan toisistaan poikkeavia määritelmiä, mutta näissä yleisesti viitataan datan määrän

merkittävään kasvuun ja sen käsittelyyn ja analysointiin [1,3,22].

Massadataa kuvataan usein viittaamalla kolmeen V:hen. Näillä tarkoitetaan tiedon määrän merkittävää kasvua (Volume), tiedon kertymisen nopeutta (Velocity) sekä tietoaaineistojen monimuotoisuutta (Variety) [1,23].

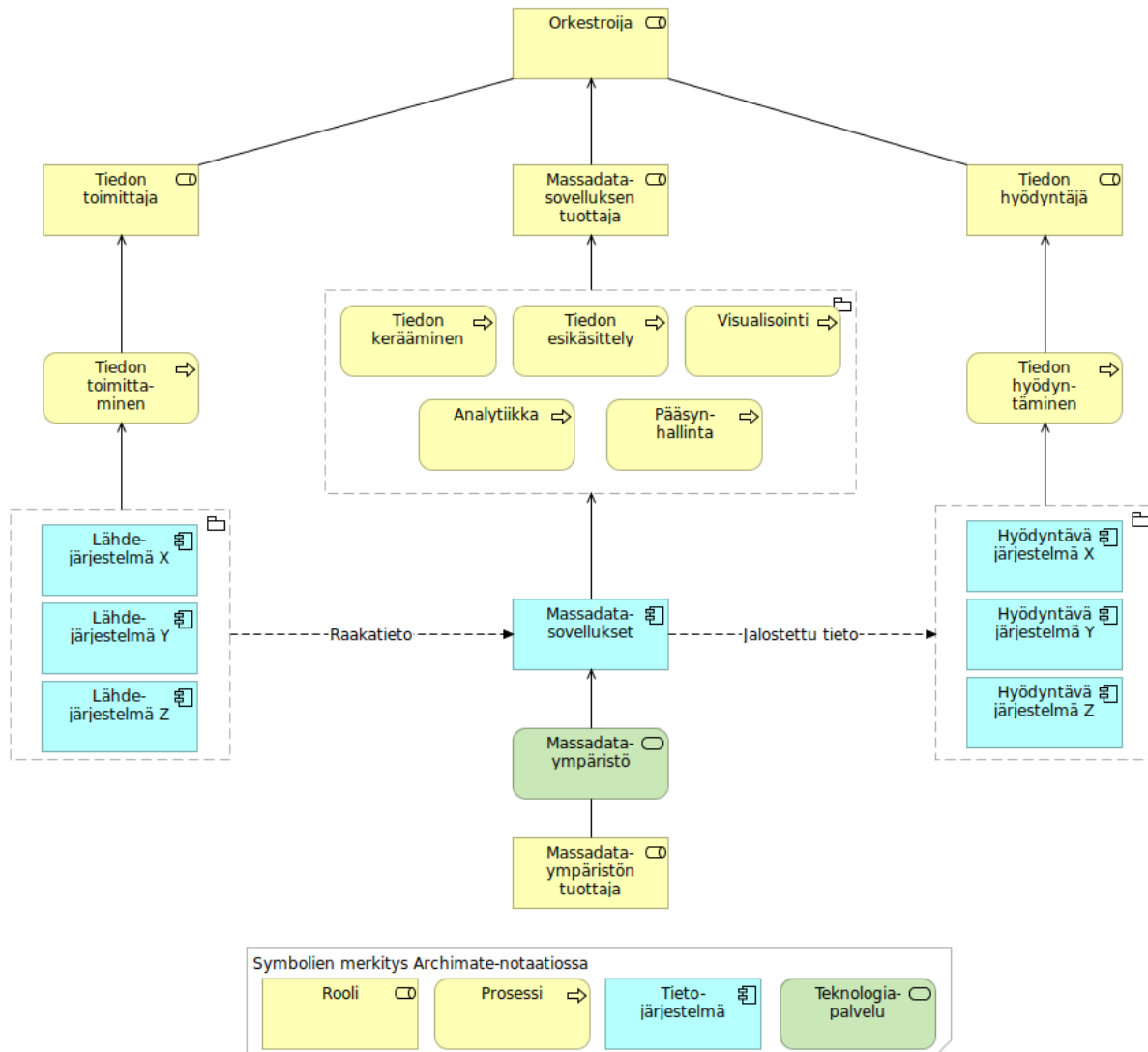
Kolmen V:n määritelmää on laajennettu useissa eri yhteyksissä lisäkäsittein. Näitä ovat mm. kattavuus (exhaustivity), erotuskyky ja indeksikaalisuus (resolution and indexicality), relationaalisuus (relationality) sekä laajennettavuus ja skaalautuvuus (extensionality and scalability) [1].

Yksi tapa hahmottaa massadataa on verrata sitä perinteisempään datan keräämiseen ja hyödyntämiseen. Kun halutaan korostaa massadatan eroja aiempaan, niin perinteisempää datan käsittelyä voidaan kutsua pieneksi dataksi (small data). Johtuen kustannuksista, resursseista, datan käsittelyn haasteista ja tallentamisen rajoituksista datan käsittely ja kokoaminen lähtökohtaisesti rajattiin tiettyyn kaapeaan käyttötarkoitukseen. Massadatalle taas tyypillistä on tiedon laaja kattavuus, jatkuva tuottaminen ja monimuotoisuus [1].

Massadata-arkkitehtuuri

Massadatalta on laadittu useita referenssiarkkitehtuureja, jotka voidaan jaotella akateemisiin ja käytännöllisiin kuvauksiin [24]. Tähän tutkimukseen valittiin analyysin tueksi yhdysvaltalaisen viranomaisen National Institute of Standards and Technologyn laatima ja ylläpitämä NIST Big Data Reference Architecture [25]. Kuvaus yhdistää sekä käytäntöä, että akateemisiä lähtökohtia. Useimmat referenssiarkkitehtuurit ovat laadittu vain toisesta lähtökohdasta [24]. Alkuperästäan huolimatta kuvauksen hyödyntämiselle Euroopassa ei ole esteitä.

Kuvioon 1 on mallinnettu tyypillinen massadata-arkkitehtuuri NIST Big Data referenssiarkkitehtuuria mukailien Archimate-notaatiolla [26].



Kuvio 1. Tyypillinen massadata –arkkitehtuuri.

NIST Big Data referenssiarkkitehtuurissa keskeistä on tiedonhallintaan liittyvien tehtävien roolittaminen. Kuhunkin rooliin liittyy tehtäviä, joiden toteuttaminen on olennaista datan käsittelyssä. [25]

Orkestroija on toimija, joka vastaa kokonaisuudesta ja asettaa sille vaatimukset. Tiedon toimittaja huolehtii siitä, että tieto on saatavilla sekä sen

toimittamisesta. Massadatasovelluksen tuottaja vastaa tiedonkäsittelyn toteuttamisesta orkestroijan vaatimusten mukaisesti. Massadataympäristön tuottaja taas vastaa tiedonkäsittelyn teknisestä ympäristöstä. Massadatasovelluksen jalostaman tiedon tulokset saatetaan lopuksi tiedon hyödyntäjälle [25]. Massadata-arkkitehtuuriin liittyviä rooleja on avattu tarkemmin taulukossa 1.

Taulukko 1. Massadata-arkkitehtuuriin liittyvät roolit ja tehtävät. Muokattu NIST Big Data referenssiarkkitehtuurin [25] pohjalta.

Rooli	Yleiskuva	Tehtävät	Esimerkkejä toimijoista
Orkestroija	Asettaa vaatimukset kokonaisuudelle	Toiminnallisten vaatimusten asettaminen Hallinnasta vastaaminen Arkkitehtuurin määrittely Datan käsittelyn vaatimusten asettaminen Tietoturva- ja tietosuojat vaatimusten määrittely Laadun hallinta Palveluiden hallinta Auditointi vaatimusten määrittely	Johtajat Arkkitehdit Datatieteilijät
Tiedon toimittaja	Huolehtii tiedon saatavuudesta ja toimittamisesta	Tiedon kokoaminen Tiedon säilyttäminen Tiedon muokkaaminen toimitettavaan muotoon Metadatan laatiminen Tiedon käyttöoikeuksien määrittely Sopimusten tekeminen tiedon hyödyntäjien kanssa Rajapintojen määrittely ja toteutus Tiedon saatavuudesta informoiminen	Viranomaiset Yritykset Tietojärjestelmätoimittajat Tutkimusyhteisö Hakukoneet Loppukäyttäjät
Massadatasovelluksen tuottaja	Huolehtii tiedon käsittelystä orkestroijan vaatimusten mukaisesti	Tiedon kerääminen Tiedon esikäsittely Analytiikka Visualisointi Pääsynhallinta	Sovellusasiantuntijat Alusta-asiantuntijat
Massadataympäristön tuottaja	Huolehtii tiedon käsittelyn teknisestä ympäristöstä ja toteutuksesta	Infrastruktuurin hallinta Data-alustojen hallinta Datan prosessointi	Datakeskukset Pilvipalvelujen toimittajat
Tiedon hyödyntäjä	Hyödyntää massadatasovelluksen jalostamaa tietoa	Tiedon haku Tiedon lataaminen Tiedon paikallinen analysointi Tiedon raportointi Tiedon visualisointi Tiedon hyödyntäminen omissa prosesseissa	Loppukäyttäjät Tutkijat Ohjelmistot

Tutkimuksen tarkoitus ja tutkimuskysymys

Aiemmassa kansainvälisessä tutkimuksessa on käsitelty massadatan ja EU:n tietosuoja-asetuksen suhdetta niin yleisellä tasolla kuin terveydenhuollossakin [1,3,27–33]. Kotimaisessa tutkimuksessa on yksittäisiä viittauksia massadataan, lainsäädäntöön ja tietosuojaan [4,10,18,34–39], mutta julkisen hallinnon tiedonhallinnan yleislainsäädännöstä seuraavia vaatimuksia kokonaisuutena ei ole aiemmin käsitelty.

Tämän artikkelin tarkoituksena on tarkastella massadatan hyödyntämistä sosiaali- ja terveydenhuollossa tiedonhallintaa koskevan yleislainsäädännön näkökulmasta. Tutkimuskysymys on: “Millaisia vaatimuksia tiedonhallinnan yleislainsäädännöstä seuraa massadata-arkkitehtuurille erityisesti henkilötietojen käsittelyn näkökulmasta?”.

Sosiaali- ja terveydenhuoltoa koskeva tiedonhallinnan erityislainsäädäntö on rajattu tämän artikkelin tarkastelun ulkopuolelle. Artikkelissa ei oteta kantaa minkään yksittäisen massadatan käyttöä edistävän tietojärjestelmätoteutuksen lainmukaisuuteen vaan kootaan yhteen tiedonhallinnan yleislainsäädännöstä lainkohtia, niistä seuraavia vaatimuksia yleisellä tasolla ja arvioidaan niiden vaikutusta massadatan tiedonhallintaan.

Sosiaali- ja terveydenhuollon tiedonhallinnan tutkimus on monitieteistä [40]. Tässä artikkelissa tutkimuskohdetta lähestytään informaatio-oikeuden ja tietojenkäsittelytieteen näkökulmia yhdistelemällä.

Aineisto ja menetelmät

Tutkimusaineisto

Tutkimuksen aineistona on keskeisimmät julkisen hallinnon tiedonhallintaa koskevat yleislait eli Julkisuuslaki (myöhemmin JulkL), Tiedonhallintalaki

(TihL), EU:n yleinen tietosuoja-asetus (TsA), Tietosuojalaki (TsL) sekä Arkistolaki (ArkL). Lainsäädäntöä käsitellään tässä artikkelissa erityisesti henkilötietojen käsittelyn näkökulmasta.

EU:n yleinen tietosuoja-asetus säädettiin vuonna 2016. Se luo säännöt luonnollisten henkilöiden suojelle henkilötietojen käsittelyssä sekä henkilötietojen vapaalle liikkuvuudelle Euroopan Unionissa. Asetuksella suojellaan luonnollisten henkilöiden perusoikeuksia ja –vapauksia ja erityisesti oikeutta henkilötietojen suojaan. Asetus on suoraan jäsenvaltiossa sovellettavaa lainsäädäntöä ja se sisältää mm. henkilötietojen käsittelyä koskevat periaatteet sekä säännökset koskien rekisteröidyn oikeuksia, rekisterinpitäjän ja henkilötietojen käsittelijän velvollisuuksia, henkilötietojen siirtoa kolmansiin maihin sekä valvontaa ja oikeussuojakeinoja [41].

Laki julkisen hallinnon tiedonhallinnasta eli tiedonhallintalaki säädettiin vuonna 2019. Siinä säädetään julkisuusperiaatteen ja hyvän hallinnon vaatimusten toteuttamisesta viranomaisten tiedonhallinnassa. Laki sisältää koko julkista hallintoa koskevat säännökset tiedonhallinnan järjestämisestä ja kuvaamisesta, tietovarantojen yhteentoimivuudesta, tietojärjestelmien yhteentoimivuuden toteuttamisesta, teknisten rajapintojen ja katseluyhteyksien toteuttamisesta sekä tietoturvallisuuden toteuttamisesta [42].

Laki viranomaisten toiminnan julkisuudesta eli julkisuuslaki sääntelee viranomaisten asiakirjojen ja muiden tietoaineistojen julkisuutta ja salassapitoa sekä tiedon antamista asiakirjasta ja siihen liittyvää menettelyä. Lisäksi laissa on säännökset viranomaisten velvollisuudesta edistää tiedonsaantia. Julkisuuslaki astui voimaan vuonna 1999 [43].

Arkistolaki säädettiin vuonna 1994. Se pitää sisällään säännökset arkistotoimen järjestämisestä [44].

Tietosuojalaki täsmentää ja täydentää EU:n tietosuojasetusta kansallisen liikkumavaran osalta. Laki säädettiin vuonna 2018 [45].

Menetelmälliset valinnat

Tutkimusta tehdessä jokainen tutkimusaineistoon valittu laki käytiin läpi pykälittäin ja sieltä tunnistettiin lainkohdat, jotka ovat olennaisia massadatan hyödyntämisen näkökulmasta. Lainkohdat ryhmiteltiin luokkiin eli menetelmällisesti kyse on laadullisessa tutkimuksessa käytetystä koodaamisesta ja luokittelusta [46,47].

Valittu menetelmä on myös lähellä vaatimusten hankintaa, vaikka tarkempaa vaatimusmäärittelyä ei ole tehtykään. Tietojenkäsittelytieteeseen pohjautuvassa ohjelmistosuunnittelussa yksi vaatimusten hankintamenetelmä on dokumenttien tutkiminen vaatimusmäärittelyn kohteen kannalta olennaisia kohtia etsien [48,49].

Varsinaisten lakitekstien lisäksi säädösten tulkinnan apuna käytettiin myös lakien esitöitä, kotimaista ja kansainvälistä oikeuskirjallisuutta sekä aiheeseen liittyvää aiempaa tutkimusta.

Tutkimuksen tulokset

Lainkohtien luokittelu

Massadatan käsittelyn kannalta keskeisimmät lainkohdat ryhmiteltiin viiden luokan alle. Luokat on muodostettu tätä tutkimusta tehdessä ja poikkeavat jonkin verran yksittäisissä laeissa olevista tavoista jäsentää lainkohtia. Lisäksi yksittäiset säännökset voisivat kohdentua useampaankin luokkaan, mutta tässä yhteydessä ne on asian selkiyttämisen vuoksi sijoitettu vain yhteen luokkaan. Lainkohtien luokittelu on esitetty taulukossa 2.

Taulukko 2. Massadataan liittyvät lainkohdat tiedonhallinnan yleislainsäädännössä

<p>1. HALLINNOLLISET VAATIMUKSET</p> <p>Lainsäädännön tunnistaminen ja sen mukaan toimiminen (TsA 5 art.)</p> <p>Tiedonhallintayksikön ja rekisterinpitäjän vastuut (TihL 4§, TsA 24. art.)</p> <p>Tiedonhallintamallin ylläpito ja seloste käsittelytoimista (TihL 5 §, TsA 30. art)</p> <p>Muutosvaikutusten arviointi (TihL 5§)</p> <p>Tietoturvaluustoimenpiteiden varmistaminen hankinnoissa (TihL 13 §)</p> <p>Tietosuojavaikutusten arviointi (TsA 35 art)</p> <p>Ennakkokuuleminen (TsA 36 art)</p> <p>Tietosuojavastaava ja resurssien varmistaminen (TsA 37–39 art)</p> <p>Osoitusvelvollisuus (TsA. 5. art.)</p> <p>Henkilötietojen käsittelijän velvoittaminen ja ohjeistaminen (TsA 28. Art)</p>	<p>2. HENKILÖTIETOJEN SUOJAAMINEN</p> <p>Käyttötarkoitussidonnaisuus (TsA 5. ja 89 art, TsL 31 §)</p> <p>Minimointi ja välttämättömyys (TsA 5. art.)</p> <p>Käsittelyperusteen määrittely (TsA 6. art)</p> <p>Eriyisten henkilötietoryhmien ja muiden arkaluontoisten tietojen tunnistaminen ja suojaaminen (TsA 9. art TsL 6 §),</p> <p>Käyttöoikeuksien määrittäminen (TihL 16 §)</p> <p>Käyttöoikeuksien käytännöistä huolehtiminen (TihL 16 §)</p> <p>Lokitietojen keruu (TihL 17 §)</p> <p>Automaattisen päätöksenteon tunnistaminen (TsA 22. Art)</p> <p>Sisäänrakennetun ja oletusarvoisen tietosuojan toteutuminen (TsA 25. Art).</p>
<p>3. TIEDONHALLINTA JA TIETOTURVA</p> <p>Tietojen täsmällisyys (TsA 5. art.)</p> <p>Tietojen eheys ja luotettavuus (TsA 5. art)</p> <p>Luotettavuutta edellyttävien tehtävien tunnistaminen ja mahdolliset tietoturvaselvitykset (TihL 12 §)</p> <p>Riskien tunnistaminen (TihL 13 §, TsA 32 art.)</p> <p>Vikasietoisuuden ja käytettävyyden testaaminen (TihL 13 §)</p> <p>Tietojärjestelmien ja tietoliikennejärjestelyjen arviointi sitä koskevan lainsäädännön mukaisesti (TihL 13 §)</p> <p>Häiriötilanteisiin varautuminen (TihL 13a §)</p> <p>Saatavuus ja käyttökelpoisuuden varmistaminen (TihL 13 §)</p> <p>Alkuperäisyys, ajantasaisuus ja virheettömyys (TihL 15 §)</p> <p>Säilytyksen rajoittaminen, säilytystarpeen arviointi ja määrittely, arkistointi (TsA 5. art., TihL 21 §, ArKL 8 §)</p> <p>Asiakirjajulkisuuden toteuttamisvelvoite (TihL 13 § ja 27 §)</p> <p>Muuttumattomuuden varmistaminen (TihL 13§)</p> <p>Arkistointimahdollisuuden varmistaminen (TihL 15 §)</p> <p>Tietoturvaloukkausten menettelyt (TsA 33 ja 34. Art)</p>	<p>4. REKISTERÖIDYN OIKEUDET JA LÄPINÄKYVYYS</p> <p>Läpinäkyvyys ja informointi (TsA 12, 13 ja 14 art.)</p> <p>Tarkistusoikeus (TsA 15 art.)</p> <p>Oikaisuoikeus (TsA 16 art.)</p> <p>Käsittelyn rajoittaminen ja vastustaminen (TsA 18. Ja 20. Art),</p> <p>5. TIEDONSIIRROT</p> <p>Salaukset ja suojaukset (TihL 14 §)</p> <p>Vastaanottajan varmistaminen ja tunnistaminen (TihL 14 §)</p> <p>Keräämisen rajapinta (TihL 20 §)</p> <p>Luovuttamisen rajapinta (TihL 22 §)</p> <p>Luovuttamisen tapauskohtaisen tarpeellisuuden ja välttämättömyyden arviointi (TihL 22§)</p> <p>Katseluyhteys ja siihen liittyvät suojatoimet (TihL 23 §)</p> <p>Tiheästi päivittyvän tiedon saatavuuden varmistaminen (TihL 24 §)</p> <p>Henkilötietojen siirtäminen kolmansiin maihin (TsA 44-50 art)</p> <p>Salassa pidettävien tietojen luovuttaminen (Julkl 16 §, 26 § ja 29 §)</p>

Ensimmäinen luokka on hallinnolliset vaatimukset. Sillä tarkoitetaan säännöksiä, jotka liittyvät tiedonhallinnan organisointiin, johtamiseen ja erilaisiin hallinnollisiin arvioihin, sopimuksiin ja ohjeisiin. Näitä vaatimuksia tulee erityisesti tietosuoja-asetuksesta ja tiedonhallintalaista, joissa vastuutetaan tiedonhallintayksikön johtoa huolehtimaan

tiedonhallintaan liittyvistä velvoitteista. Massadataan käsittelyn kannalta keskeisiä hallinnollisia vaatimuksia ovat mm. kokonaisuuteen liittyvän lainsäädännön tunnistaminen, sen mukaan toimiminen ja osoitusvelvollisuus (TsA 5. art.), tietosuojavaikutusten arvioinnin tekeminen (TsA 35 art), tietoturvallisuudesta huolehtiminen hankinnoissa (TihL 13§)

sekä henkilötietojen käsittelijän velvoittaminen toimimaan vaatimusten mukaan sopimuksin ja ohjein. (TsA 28 art).

Toinen luokka on henkilötietojen suojaaminen. Se pitää sisällään erityisesti tietosuoja-asetuksesta nousevia velvoitteita. Näistä keskeisimpiä ovat käsittelyperusteen määrittely (TsA 6 art), tiedon käyttötarkoitusten tunnistaminen ja määrittely (TsA 5 art ja 89 Art, TSL, 31 §), tiedon käsittelyn minimointi (TsA 5. art) sekä erityisiin henkilötietoryhmiin ja muihin arkaluontoisiin tietoihin liittyvät suojatimet (TsA 9. art, TSL 6 §). Sisäänrakennetun ja oletusarvoisen tietosuojan varmistaminen on olennaista jo henkilötietojen käsittelyn suunnitteluvaiheessa sekä läpi koko tiedon elinkaaren. (TsA 25 art). Henkilötietojen suojaamisen kannalta keskeisiä vaatimuksia tulee myös tiedonhallintalaista, jossa edellytetään käyttöoikeuksista ja lokituksista huolehtimista (TihL 16–17 §).

Kolmas luokka on tiedonhallintaa ja tietoturva koskevat lainkohdat. Näillä tarkoitetaan tiedon elinkaaren aikaiseen saatavuuteen, löydettävyyteen, hyödynnettävyyteen, eheyteen ja luottamuksellisuuteen liittyviä vaatimuksia. Sekä tietosuoja-asetuksessa, että tiedonhallintalaissa painotetaan tiedon oikeellisuutta, ajantasaisuutta ja alkuperäisyyttä (TsA 5. art., TihL 13 § ja 15 §). Keskeisiä ovat myös säilyttämiseen liittyvät velvoitteet, kuten säilytyksen rajoittaminen (TsA 5. art.), säilytystarpeen arviointi ja määrittely (TihL 21 §) sekä arkistoinnin hoitaminen (ArKL).

Neljäntenä luokkana on rekisteröidyn oikeuksia ja läpinäkyvyyttä edistävät säännökset. Nämä liittyvät henkilötietojen käsittelyn avoimuuteen, informointiin sekä tietojenkäsittelyn kohteena olevien oikeuteen tarkistaa itseään koskevat tiedot (TsA 15 art.), pyytää niihin oikaisua (TsA 16 art) sekä tietyissä tilanteissa rajoittaa (TsA 18 art) tai vastustaa tietojen käsittelyä (TsA 20 art).

Tiedonsiirtoja koskevat säännökset muodostavat viidennen luokan. Näistä keskeisimpiä ovat salassa pidettävien tietojen luovuttamisen perusteen säännökset (JulkL 16 §, 26 § ja 29 §) sekä tiedon siirtoihin liittyvien teknisten edellytysten varmistaminen (TihL 14 §, 20 §, 22 §, 23 § ja 24 §). Lisäksi tarvittaessa on huomioitava kolmansiin maihin eli EU- ja ETA-alueen ulkopuolelle tapahtuvien tiedonsiirtojen sääntely (TsA 44–50 §). Henkilötietojen ja salassa pidettävien tietojen osalta on tapauskohtaisesti varmistettava siirtojen välttämättömyys (TihL 22 §).

Lainsäädännön suhde referenssiarkkitehtuuriin rooleihin

Suhteessa aiemmin esitettyihin tyypillisen massa-data-arkkitehtuuriin rooleihin, orkestroijan rooli voidaan nähdä jaetun julkishallinnossa tiedonhallintalain tarkoittaman tiedonhallintayksikön ja mahdollisten tiedonhallintayksikön sisällä toimivien viranomaisten kesken. Jälkimmäiset toimivat myös tietosuoja-asetuksen mukaisena rekisterinpitäjänä. Orkestroijan roolin kannalta lainkohdista olennaisia ovat erityisesti tiedonhallintalain 4 § mukainen tiedonhallintayksikön johdon vastuu sekä tietosuoja-asetukseen 5 art. velvoite tunnistaa lainsäädäntö ja sen mukaan toimiminen. Lisäksi tietosuoja-asetuksen 24 artikla säättää rekisterinpitäjän vastuista. Orkestroija vastaa viime kädessä kokonaisuudesta ja sille asetetuista vaatimuksista. Näin ollen sen tulee huolehtia, että kaikessa tiedon käsittelyssä tulisi toteutua sisäänrakennettu ja oletusarvoinen tietosuoja (TsA 25 art.). Riskiperusteisen lähestymistavan mukaisesti erityisesti arkaluontoisten tietojen käsittelyssä keskeistä on huolehtia tietosuoja vaikutusten arvioinnista jo ennen tietojen käsittelyn aloittamista (TsA 35 art.) Lisäksi on määriteltävä säilyttämiseen, sen rajoittamiseen ja arkistointiin liittyvät vaatimukset.

Orkestroija vastaa myös rekisteröidyn oikeuksien ja läpinäkyvyyden toteutumisesta.

Tiedon toimittajan roolissa voi toimia sama viranomainen kuin joka vastaa massadatan käsittelystä. Se voi olla myös toinen viranomainen tai yksityinen toimija. Riippumatta mikä toimija roolissa toimii, niin keskeistä on huolehtia tiedon oikeellisuudesta (eheys, täsmällisyys, virheettömyys TsA 5. art, TihL, 15 §) sekä käsiteltävän tiedon rajaamisesta käsittelyn kannalta vain välttämättömään eli minimointiperiaatteen toteutumisesta. (TsA 5. art). Mikäli tiedot toimitetaan luovutuksena toiselle viranomaiselle, tulee tunnistaa millä perusteella luovutus voidaan tehdä (JulkL 16 §, 26 § ja 29§, TihL 22 §) sekä huolehtia myös muiden tiedonsiirtoja koskevien vaatimusten täyttymisestä (kts. Taulukko 2, luokka 5)

Massadatasovelluksen tuottajan ja massadataympäristön tuottajan toiminnan lähtökohtana voi pitää sitä, että henkilötietojen käsittelijän on pystyttävä toteuttamaan sellaiset suojoimet, joilla varmistetaan, että käsittely täyttää tietosuojasetuksen asettamat vaatimukset (TsA 28 art.) Toteutettaessa tiedon keräämiseen, esikäsittelyyn, analytiikkaan, visualisointiin ja pääsynhallintaan liittyviä tehtäviä sen tulee tapahtua orkestroijan määrittelmällä tavalla. Tässä korostuu erityisesti TihL 4 § tarkoittamat ohjeet ja valvonta, tiedonhallintayksikön hankinnoissa asettamat vaatimukset (TihL 13 §) sekä henkilötietojen käsittelijän kanssa tehdyt sopimukset ja ohjeet (TsA 28 art). Näissä ohjeissa ja sopimuksissa tulisi määritellä tarkkaan vaatimukset massadatasovelluksen tuottajalle ja massadataympäristön tuottajalle, jotta tiedonkäsittely tapahtuu koko sen elinkaaren ajan lainmukaisesti. Käytännössä tämä edellyttää lähes kaikkien viiteen luokkaan jaoteltujen lainkohtien läpikäyntiä ja niiden soveltamisen varmistamista. Näin ollen orkestroijalla on hyvin merkittävä vastuu

massadatasovelluksen tuottajan ja massadataympäristön tuottajan ohjeistamisesta ja valvonnasta.

Tiedon hyödyntäjä voi olla osa sama viranomainen tai sen ulkopuolinen toimija, kuten toinen viranomainen tai yksityinen toimija. Mikäli tiedon hyödyntämisvaiheessa käsitellään henkilötietoa, tulee siinä huomioida henkilötietojen suojaamiseen liittyvät lainkohdat. Näitä ovat mm. käyttöoikeuksiin ja lokittamiseen liittyvät vaatimukset (TihL 16–17 §). Mikäli tietoja siirretään toimijalta toiselle, tiedonsiirtoja koskevat lainkohdat ovat myös olennaisia.

Pohdinta ja yhteenveto

Julkisen hallinnon tiedonhallintaa koskevaa lainsäädäntöä on verrattain paljon, se on syntynyt eri aikoina eikä se ole käsitteellisesti yhtenäistä. Siksi sen soveltaminenkin on työlästä. Erityisesti kun uudet teknologiat kehittyvät ja niitä tuodaan myös julkishallintoon, voi olla haastavaa yhteensovittaa näitä olemassa olevaan lainsäädäntöön [3,50].

Artikkelin lopuksi nostetaan esiin neljä näkökulmaa, joissa jännitteisyyttä erityisesti on ja jossa lainsäädännön soveltaminen vaatii huolellisuutta. Käsiteltävät näkökulmat ovat käyttötarkoitussidonnaisuus, minimointiperiaate, tietojen oikeellisuus ja säilytyksen rajoittaminen sekä tiedon käsittelyn läpinäkyvyys [1,3,13,27,51–56].

Ensinnäkin käyttötarkoitussidonnaisuus edellyttää, että yhteen käyttötarkoitukseen kerättyä tietoa ei saa myöhemmin käyttää sen kanssa yhteensopimattomalla tavalla. Massadatan hyödyntämiseen sen sijaan liittyy olennaisesti tiedon yhdistely eri lähteistä. Käyttötarkoitussidonnaisuudesta voidaan poiketa vain, mikäli siitä on erikseen säädetty poikkeus tietosuojasetuksessa tai kansallisessa lainsäädännössä (esim. tilastointi ja tutkimus) [3,41,45,50]. Sosiaali- ja terveydenhuollossa ns.

toisiolaki mahdollistaa tiedon käytön toisiin käyttö-tarkoituksiin, joko luvalla tai muuten laissa tarkoin määritellyin tavoin [57].

Toiseksi käyttötarkoituksen määrittely liittyy läheisesti minimointiperiaatteeseen. Sen lähtökohta on käsiteltävän tiedon rajaaminen käyttötarkoituksen kannalta välttämättömään. Tietoja ei tule käsitellä enempää mitä käyttötarkoituksen kannalta on tarpeellista ja tämä tulee määritellä etukäteen ennen käsittelyn aloittamista [3,50,58]. Massadatasessa taas usein on olennaista tiedon suuri määrä sekä yhdistelyn kautta saatava uudenvuoroinen tieto, jota ei ole pystytty ennakoimaan tiedon käsittelyä aloitettaessa.

Kolmanneksi massadatasessa lähtökohtana usein on tietojen laajamittainen kopiointi alkuperäisistä lähdejärjestelmistä massadatasovelluksiin aiemmin esitetyn arkkitehtuurin mukaisesti [20,21,25]. Sekä tiedonhallintalaki että tietosuojalaki korostavat tiedon virheettömyyttä, eheyttä ja muuttumattomuutta, alkuperäisyyttä sekä säilytyksen rajoittamista. Näin olleen massadatan käsittelyjärjestelmissä tulisi erityisen huolellisesti kiinnittää huomiota näihin tiedon oikeellisuuden varmistamiseen. Lisäksi tulee huolehtia kopioitujen tiedon hävittämisestä, kun sen käsittelylle ei ole perustetta säilytyksen rajoittamisperiaatteen mukaisesti [3,50,58].

Neljänneksi rekisteröidyn oikeuksien sekä hyvän hallinnon kannalta keskeistä on tiedon käsittelyn läpinäkyvyys. Siksi rekisteröidyn informointeihin heidän tietojensa käsittelystä ja oikeuksistaan sekä asiakirjajulkisuuskuvauksiin tulisi kiinnittää erityistä huomiota [3,36,50,59].

Kuten tässä artikkelissa on todettu, niin massadatarakkitehtuurissa vastuut jakautuvat usealle eri taholle. Näin ollen tiedonhallintayksikön johdon vastuut tiedonkäsittelyn lainmukaisuudesta, sen

ohjeistamisesta, riittävästä osaamisesta ja valvonnasta ovat merkittäviä. Lakisääteiset tietosuojavaikeuksien arvioinnit ja tiedonhallintalain muutosten vaikutusten arvioinnit ovat keskeisiä välineitä varmistettaessa, että tiedonhallinnan toteutuksen kannalta olennaisimmat näkökulmat tulevat huomioitua riittävän aikaisessa vaiheessa [3,50,60–63].

Lainsäädännön laajuuden ja monimutkaisuuden vuoksi yhteinen kansallinen ohjeistus tai viitearkkitehtuuri voisi olla hyödyllinen massadataratkaisuja käyttäville ja sitä suunnitteleville viranomaisille. Tietosuoja-asetuksen soveltamisen tueksi on tehty tällaisia viitearkkitehtuureja [64–67]. Suomalaisen laajemman lainsäädäntökokonaisuuden vuoksi ohjeistuksen tulisi kuitenkin olla kattavampi.

Julkisten palveluiden datafikoituminen on laajuudeltaan merkittävää ja siihen liittyy monia eri näkökulmia. Mm. Petteri Orpon hallitusohjelmassa on tähän paljon odotuksia. Massadatan hyödyntämiseen liittyvää empiirinen jatkotutkimus aiheesta olisi tämän vuoksi tarpeellista: Kuinka hyvin lainsäädännön vaatimukset tunnetaan ja miten ne on huomioitu massadatatoteutuksissa? Millainen on oikeudellisen osaamisen taso? Millaisia aiheeseen liittyviä koulutustarpeita on? Kuinka ymmärrettävänä ja toimivana nykyistä sääntelyä pidetään? Miten massadataradigman ja tietosuoja-sääntelyn välisiä jännitteitä on ratkottu?

Näiden lisäksi olisi tarvetta lainopilliselle tutkimukselle, joka auttaisi kuvaamaan ja selkiyttämään erilaisia tulkintatilanteita ja oikeudellisten intressien välisiä jännitteitä sekä tuottamaan ehdotuksia systemaattisen ja soveltajan kannalta selkeän lainsäädännön kehittämiseksi.

Sidonnaisuudet

Kirjoittajalla ei ole sidonnaisuuksia.

Lähteet

[1] Kitchin R. The Data Revolution: A Critical Analysis of Big Data, Open Data and Data Infrastructures. 2. painos. SAGE Publications Ltd; 2022.

[2] Ooijen C van, Ubaldi B, Welby B. A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance [Internet]. OECD Working Papers on Public Governance, No. 33. Paris: OECD; 2019 touko [viitattu 20. huhtikuuta 2022]. Saatavissa: <https://doi.org/10.1787/09ab162c-en>

[3] Forgo N, Wagner MC. The Principle of Transparency in Medical Research. Applying Big Data Analytics to Electronic Health Records. Teoksessa: AI in EHealth: Human Autonomy, Data Governance and Privacy in Healthcare. Cambridge University Press; 2022. s. 209–250. Saatavissa: <https://doi.org/10.1017/9781108921923.013>

[4] Aalto E, Gustafsson R, Lipiäinen N. Alustatalouden innovoinnin ja kasvun haasteet [Internet]. Aalto-yliopiston julkaisusarja. Report No.: 15/2020. Aalto yliopisto; 2020 [viitattu 19. elokuuta 2023]. Saatavissa: https://www.aalto.fi/sites/g/files/flghsv161/files/2020-05/Taustadokumentti%20Alustatalouden%20innovoinnin%20ja%20kasvun%20haasteet_0.pdf

[5] Dijck J van. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveill Soc.* 9. toukokuuta 2014;12(2):197–208. Saatavissa: <https://doi.org/10.24908/ss.v12i2.4776>

[6] Mejias UA, Couldry N. Datafication. *Internet Policy Review.* 2019;8(4). Saatavissa: <https://doi.org/10.14763/2019.4.1428>

[7] Freeman Engstrom D, Ho DE. Artificially intelligent government: A review and agenda. Teoksessa: Vogl R (toim.). *Research Handbook on Big Data Law.*

Edward Elgar Publishing; 2021. s. 57-86. Saatavissa: <https://doi.org/10.4337/9781788972826.00009>

[8] Neittaanmäki P, Lehto M, Ruohonen T, Kaasalainen K, Karla T. Suomen terveystietojen hyödyntäminen. Loppuraportti Vol. 4 [Internet]. Jyväskylän yliopisto; 2019 [viitattu 31. heinäkuuta 2023]. Saatavissa: <http://urn.fi/URN:ISBN:978-951-39-7715-3>

[9] Antikainen J, Eskelinen J, Koski H, Niemi T, Pajarinen M, Pyykkönen S, ym. Massadatatista liiketoimintaa ja tehokkaita julkisia palveluja [Internet]. Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 16/2016. Valtioneuvoston kanslia; 2016 [viitattu 2. elokuuta 2023]. Saatavissa: <http://urn.fi/URN:ISBN:978-952-287-250-0>

[10] Tuomisto JT. Massadatat kansanterveyden edistämiseksi. *Lääketieteellinen Aikakauskirja Duodecim.* 2015;131(22):2179–87.

[11] Euroopan parlamentti. Massadatat: määritelmä, hyödyt, haasteet (infografiikka). [Internet]. Ajankohtaista, Euroopan parlamentti; 2021 [viitattu 31. heinäkuuta 2023]. Saatavissa: <https://www.europarl.europa.eu/news/fi/headlines/society/20210211STO97614/massadatat-maaritelma-hyodyt-haasteet-infografiikka>

[12] Virkki A. Tietoa avulla kohti laadukasta ja kustannus tehokasta lääkehoitoa [Internet]. Lääkealan turvallisuus- ja kehittämiskeskus Fimea; 2018 [viitattu 31. heinäkuuta 2023]. Saatavissa: <https://urn.fi/URN:NBN:fi-fe2018061926082>

[13] De Gregorio GD, Ranchordás S. Breaking down information silos. A legal analysis of data sharing. Teoksessa: Cannataci J, Falce V, Pollicino O, toim. *Legal challenges of big data.* Northampton: Edward Elgar Publishing; 2020. Saatavissa: <https://doi.org/10.4337/9781788976220.00014>

[14] Ouli J, Voutilainen T. Oppimisanalytiikka ja opiskelijatietojen käsittely yliopistoissa [Internet]. EDILEX 30. syyskuuta 2019 [viitattu 8. elokuuta

2023]. Saatavissa: <https://www.edilex.fi/artikkelit/20064>

[15] Nevaranta M, Lempinen K, Kaila ET. Insight on the Finnish Field of Learning Analytics: Applications and Ethics in Adaptive Education Models. Teoksessa: The Asian Conference on Education 2019: Official Conference Proceedings. Iafor; 2019. Saatavissa: https://papers.iafor.org/wp-content/uploads/papers/ace2019/ACE2019_53283.pdf

[16] Sosiaali- ja terveysministeriö. Tieto hyvinvoinnin ja uudistuvien palvelujen tukena - Sote-tieto hyötykäyttöön -strategia 2020 [Internet]. Sosiaali- ja terveysministeriö; 2014 [viitattu 4. huhtikuuta 2022]. Saatavissa: <http://urn.fi/URN:ISBN:978-952-00-3548-8>

[17] Valtioneuvosto. Kestävää kasvua ja hyvinvointia – Tiekartta 2020–2023: Terveysalan tutkimus- ja innovaatiotoiminnan kasvustrategia [Internet]. Valtioneuvoston julkaisuja 2020:33. Valtioneuvosto; 2020 joulou [viitattu 19. elokuuta 2023]. Saatavissa: <http://urn.fi/URN:ISBN:978-952-383-520-7>

[18] Lipiäinen N, Gustafsson R, Lavinto S. Kuinka alustataloutta voidaan vauhdittaa terveysalalla? [Internet]. Poliitikkasuositus 7.6.2021. Aalto yliopisto; 2021 [viitattu 19. elokuuta 2023]. Saatavissa: https://www.aalto.fi/sites/g/files/flghsv161/files/2021-06/Politiikkasuositus_Miten%20alustataloutta%20voidaan%20vauhdittaa%20terveysalalla.pdf

[19] Larsio A. Datasta voimaa sote-järjestelmään [Internet]. Sitra; 2023 touko [viitattu 19. elokuuta 2023]. Saatavissa: <https://www.sitra.fi/julkaisut/datasta-voimaa-sote-jarjestelmaan/>

[20] Tamminen T. Tietoallas estää dataa hukummasta. Suomidigi; 2019 [viitattu 14. joulukuuta

2022]. Saatavissa: <https://www.suomidigi.fi/artikkelit/tietoallas-estaa-dataa-hukummasta>

[21] Oraskari J. Mikä on datajärvi? Tivi; 4.1.2021 [viitattu 31. heinäkuuta 2023]. Saatavissa: <https://www.tivi.fi/uutiset/mika-on-data-jarvi/a59711aa-acb1-4faa-a30c-fcc698b6d825>

[22] Finto. Tietotermit: massadata [Internet]. Finto suomalainen asiasanasto ja ontologiapalvelu [viitattu 2. elokuuta 2023]. Saatavissa: <https://finto.fi/tt/fi/page/t6>

[23] Zikopoulos PC, Eaton C, deRoos D, Deutsch T, Lapis G. Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data. McGraw-Hill; 2012.

[24] Ataei P, Litchfield AT. Big Data Reference Architectures, a systematic literature review [Internet] 2020. ACIS 2020 Proceedings. 30 [viitattu 2. tammikuuta 2023]. Saatavissa: <https://aisel.aisnet.org/acis2020/30>

[25] NIST Big Data Public Working Group. NIST Big Data Interoperability Framework: volume 6, reference architecture, version 2 [Internet]. Gaithersburg, MD: National Institute of Standards and Technology; 2018 [viitattu 2. tammikuuta 2023]. Saatavissa: <https://doi.org/10.6028/NIST.SP.1500-6r1>

[26] The Open Group. Introduction: ArchiMate® 3.2 Specification [Internet]. The Open Group; 2023 [viitattu 31. heinäkuuta 2023]. Saatavissa: <https://pubs.opengroup.org/architecture/archimate3-doc/index.html>

[27] Zarsky T. Incompatible: The GDPR in the Age of Big Data. Seton Hall Law Review. 2017;47(4)2. [Internet]. Saatavissa: <https://papers.ssrn.com/abstract=3022646>

[28] Yang YP. Data protection in the big data era: The broken informed consent regime and the way forward. Teoksessa: Lee J, Darbellay A (toim.). Data

Governance in AI, FinTech and LegalTech. Edward Elgar Publishing; 2022. s. 59–78. <https://doi.org/10.4337/9781800379954.00011>

[29] Cohen IG, Lynch HF, Vayena E, Gasser U. Big Data, Health Law, and Bioethics. Cambridge: Cambridge University Press; 2018. <https://doi.org/10.1017/9781108147972>

[30] Hulsén T. Challenges and solutions for big data in personalized healthcare. Teoksessa: Moustafa A. Big Data in Psychiatry and Neurology. Elsevier; 2021. s. 69–94. <https://doi.org/10.1016/B978-0-12-822884-5.00016-7>

[31] Cheung S. Disambiguating the benefits and risks from public health data in the digital economy. Big Data Soc. 1. tammikuuta 2020;7(1):2053951720933924. <https://doi.org/10.1177/2053951720933924>

[32] Forgó N, Johnson E, Kazeeva I, Steindl E. Big Data, AI, and Health Data: between National, European, and International Legal Frameworks [Internet]. Forthcoming in 'Handbook on the Politics and Governance of Big Data and Artificial Intelligence - Elgar Handbooks in Political Science. Rochester, NY; 2022 [viitattu 9. lokakuuta 2022]. Saatavissa: <https://papers.ssrn.com/abstract=4184395>. <https://doi.org/10.4337/9781800887374.00024>

[33] Marelli L, Testa G, Van Hoyweghen I. Big Tech platforms in health research: Re-purposing big data governance in light of the General Data Protection Regulation's research exemption. Big Data Soc. 1. tammikuuta 2021;8(1):20539517211018784. <https://doi.org/10.1177/20539517211018783>

[34] Jylhä V, Mykkänen J, Värri A, Kinnunen UM, Peltonen LM. Yhteinen sote-tieto? - Tiedon toisioikäyttö ja digitalisaation vaikutukset. FinJeHeW. 10. marraskuuta 2021;13(4):329–31. <https://doi.org/10.23996/fjhw.112086>

[35] Salovaara S, Leinonen J, Silén M. Tietojärjestelmien avulla kerätyn tiedon hyödyntämisen esteet sosiaalialan organisaatioiden tiedolla johtamisessa. FinJeHeW. 10. marraskuuta 2021;13(4):372–87. <https://doi.org/10.23996/fjhw.109930>

[36] Salokannel M, Ignatius E, Vesikukka T. Läpinäkyvä SOTE-tietojen käsittely: kuvakkeilla tiedottaminen. White paper 3/2023 [Internet]. DataLit; 2023 [viitattu 18. heinäkuuta 2023]. Saatavissa: <https://www.datalit.fi/wp-content/uploads/2023/04/GDPR-White-paper-Tietosuojakuvakkeet-3-2023.pdf>

[37] Häyrinen K. Big data sosiaali- ja terveydenhuollon tietohallinnon ja tietojenkäsittelyn tutkimuksessa. FinJeHeW. 2014;6(2–3):40.

[38] Häyrinen K. Tietosuojat – aina ajankohtainen aihe. FinJeHeW. 2015;7(1):1.

[39] Sorvettula J. Tietosuojat. FinJeHeW. 2015;7(1):36–45.

[40] Bichel-Findlay J, Koch S, Mantas J, Abdul SS, Al-Shorbaji N, Ammenwerth E, ym. Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics: Second Revision. Int J Med Inf. 1. helmikuuta 2023;170:104908. <https://doi.org/10.1016/j.ijmedinf.2022.104908>

[41] EUR-Lex. Euroopan parlamentin ja neuvoston asetus (EU) 2016/679, annettu 27 päivänä huhtikuuta 2016, luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta (yleinen tietosuojat-asetus) [Internet]. OJ L huhti 27, 2016. [viitattu 18. heinäkuuta 2023]. Saatavissa: <http://data.europa.eu/eli/reg/2016/679/oj/fin>

[42] Finlex. Laki julkisen hallinnon tiedonhallinnasta 906/2019 [Internet]. Valtiovarainministeriö; 2020 [viitattu 18. heinäkuuta 2023]. Saatavissa:

<https://www.finlex.fi/fi/laki/ajantasa/2019/20190906>

[43] Finlex. Laki viranomaisten toiminnan julkisuudesta 621/1999 [Internet]. Oikeusministeriö; 1999 [viitattu 31. heinäkuuta 2023]. Saatavissa: <https://www.finlex.fi/fi/laki/ajantasa/1999/19990621>

[44] Finlex. Arkistolaki 831/1994 [Internet]. Opetusministeriö; 1994 [viitattu 18 heinäkuuta 2023]. Saatavissa: <https://www.finlex.fi/fi/laki/ajantasa/1994/19940831>

[45] Finlex. Tietosuojalaki 1050/2018 [Internet]. Oikeusministeriö; 2019 [viitattu 18 heinäkuuta 2023]. Saatavissa: <https://www.finlex.fi/fi/laki/ajantasa/2018/20181050>

[46] Juhila K. Koodaaminen. Teoksessa: Vuori J (toim.). Laadullisen tutkimuksen verkkokäsikirja [Internet]. Tampere: Yhteiskuntatieteellinen tietoaristo [viitattu 31. heinäkuuta 2023]. Saatavissa: <https://www.fsd.tuni.fi/fi/palvelut/menetelma-opetus/kvali/analyysitavan-valinta-ja-yleiset-analyysitavat/koodaaminen/>

[47] Puusa A. Näkökulmia laadullisen aineiston analyysiin. Teoksessa: Puusa A, Juuti P, toim. Laadullisen tutkimuksen näkökulmat ja menetelmät. Gaudemus; 2020.

[48] Suomidigi. Julkisen hallinnon tietohallinnon neuvottelukunta. JHS 173 ICT-palvelujen kehittäminen: Vaatimusmäärittely [Internet]. Suomidigi; 2018. [viitattu 18 heinäkuuta 2023] Saatavissa: <https://www.suomidigi.fi/ohjeet-ja-tuki/jhs-suositukset/jhs-173-ict-palvelujen-kehittaminen-vaatimusmaarittely-vanhentunut>

[49] Tiihonen T, Mykkänen J, Savolainen S, Itälä T, Hiekkänen K, Luukkonen I, ym. Tarpeiden ja vaatimusten hallinta kokonaisarkkitehtuurissa [Internet]. Itä-Suomen yliopisto ja Aalto yliopisto; 2012

[viitattu 7. elokuuta 2023]. Saatavissa: <http://urn.fi/URN:ISBN:978-952-61-0722-6>

[50] Voutilainen T. Oikeus tietoon - informaatio-oikeuden perusteet (uud.) Digikirja [Internet]. Edita; 2019 [viitattu 31. heinäkuuta 2023]. Saatavissa: <http://shop.edita.fi/tuote/oikeus-tietoon-informaatio-oikeuden-perusteet-uud-digi>

[51] Mantelero A. Big data and data protection. Teoksessa: González-Fuster G, Van Brakel R, De Hert P (toim.). Research handbook on privacy and data protection law: values, norms and global politics. Research Handbooks in Information Law series. Edward Elgar Publishing; 2022.

[52] Ursic H, Custers B. Legal Barriers and Enablers to Big Data Reuse. Eur Data Prot Law Rev EDPL. 2016;2:209–21. <https://doi.org/10.21552/EDPL/2016/2/10>

[53] Jain P, Gyanchandani M, Khare N. Big data privacy: a technological perspective and review. J Big Data. 26. marraskuuta 2016;3(1):25. <https://doi.org/10.1186/s40537-016-0059-y>

[54] European Union Agency for Fundamental Rights (FRA). Council of Europe, Guidelines on the protection of individuals with regard to the processing of personal data in a world of big data [Internet]. European Union Agency for Fundamental Rights; 2020 [viitattu 17. heinäkuuta 2023]. Saatavissa: <http://fra.europa.eu/en/law-reference/council-europe-guidelines-protection-individuals-regard-processing-personal-data>

[55] Kiseleva A, de Hert P. Creating a European Health Data Space: Obstacles in Four Key Legal Area. Eur Pharm Law Rev EPLR. 2021;5(1):21–36. <https://doi.org/10.21552/eplr/2021/1/5>

[56] Ionescu S. EU Data Protection, Obstacle in the Way of Medical Scientific Research? Acad J Law Gov AJLG. 2020;8:124–8.

- [57] Finlex. Laki sosiaali- ja terveystietojen toissijaisesta käytöstä 552/2019 [Internet]. Sosiaali- ja terveysministeriö; 2019 [viitattu 7. elokuuta 2023]. Saatavissa: <https://www.finlex.fi/fi/laki/ajantasa/2019/20190552>
- [58] Euroopan tietosuojaneuvosto. Ohjeet 4/2019 25 artiklan mukaisesta sisäänrakennetusta ja oletusarvoisesta tietosuojasta. Versio 2.0 [Internet]. Euroopan tietosuojaneuvosto; 2020 [viitattu 18. heinäkuuta 2023]. Saatavissa: https://edpb.europa.eu/system/files/2021-04/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_fi.pdf
- [59] Tiedonhallintalautakunta. Suositus asiakirjajulkisuuskuvauksen laatimisesta [Internet]. Valtiovarainministeriön julkaisuja 2020:22. Valtiovarainministeriö; 2020 [viitattu 7. elokuuta 2023]. Saatavissa: <http://urn.fi/URN:ISBN:978-952-367-304-5>
- [60] Tiedonhallintalautakunta. Suositus johdon vastuiden toteuttamisesta tiedonhallinnassa [Internet]. Valtiovarainministeriön julkaisuja 2020:18. Valtiovarainministeriö; 2020 [viitattu 7. elokuuta 2023]. Saatavissa: <http://urn.fi/URN:ISBN:978-952-367-288-8>
- [61] Tiedonhallintalautakunta. Suositus tiedonhallinnan muutosvaikutusten arvioinnista [Internet]. Valtiovarainministeriön julkaisuja 2020:53. Valtiovarainministeriö; 2020 [viitattu 7. elokuuta 2023]. Saatavissa: <http://urn.fi/URN:ISBN:978-952-367-318-2>
- [62] Tietosuojatyöryhmä. Ohjeet tietosuojaa koskevasta vaikutustenarvioinnista ja keinoista selvittää ”liittyykö käsittelyyn todennäköisesti” asetuksessa (EU) 2016/679 tarkoitettu ”korkea riski” [Internet]. Tietosuojaryhmä; 2017 [viitattu 7. elokuuta 2023]. Saatavissa: <https://tietosuojafi.fi/documents/6927448/8316711/Vaikutustenarviointi+fi.pdf/af51e999-5326-4223-9deb-e21bdd2e0a63/Vaikutustenarviointi+fi.pdf?t=1527059635000>
- [63] Tietosuojavaltuutetun toimisto. Tietosuojan vaikutustenarvioinnin ohje [Internet]. Helsinki: Tietosuojavaltuutetun toimisto; 2021 [viitattu 7. elokuuta 2023]. Saatavissa: <https://tietosuojafi.fi/documents/6927448/66036250/TVA+ohje.pdf/ff0b6e1b-5b89-e85e-a2e5-6c4bd4c0ccfc/TVA+ohje.pdf?t=1639729535787>
- [64] Georgiadis G, Poels G. Enterprise architecture management as a solution for addressing general data protection regulation requirements in a big data context: a systematic mapping study. *Inf Syst E-Bus Manag.* 1. maaliskuuta 2021;19(1):313–62. <https://doi.org/10.1007/s10257-020-00500-5>
- [65] Burmeister F, Drews P, Schirmer I. A Privacy-driven Enterprise Architecture Meta-Model for Supporting Compliance with the General Data Protection Regulation. Teoksessa: *Proceedings of the 52nd Hawaii International Conference on System Sciences.* 2019. s. 6052–61. <https://doi.org/10.24251/HICSS.2019.729>
- [66] Blanco-Lainé G, Sottet JS, Dupuy-Chessa S. Using an enterprise architecture model for GDPR compliance principles. Teoksessa: Gordijn J, Guédria W, Proper H (toim). *The Practice of Enterprise Modeling. PoEM 2019. Lecture Notes in Business Information Processing, vol 369.* Cham, Springer; 2019. https://doi.org/10.1007/978-3-030-35151-9_13
- [67] Rhahla M, Allegue S, Abdellatif T. Guidelines for GDPR compliance in Big Data systems. *J Inf Secur Appl.* 1. syyskuuta 2021;61:102896. <https://doi.org/10.1016/j.jisa.2021.102896>