Joel Kataja

EU:N YLEINEN TIETOSUOJA-ASETUS: REKISTERÖIDYN OIKEUKSIEN TOTEUTUMINEN JA DATANHALLINNAN HYÖDYNTÄMINEN YRITYKSISSÄ

JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2021

TIIVISTELMÄ

Kataja, Joel
EU:n yleinen tietosuoja-asetus: Rekisteröidyn oikeuksien toteutuminen ja datanhallinnan hyödyntäminen yrityksissä.
Jyväskylä: Jyväskylän yliopisto, 2021, 91 s.
Tietojärjestelmätiede, pro gradu -tutkielma
Ohjaaja(t): Siponen, Mikko

Datan alentuneet käsittelykustannukset houkuttelevat yrityksiä keräämään dataa entistä enemmän, joka voi johtaa henkilötietojen käsittelyn väärinkäyttöön ja rekisteröidyn yksityisyyden loukkaamiseen. Euroopan unionin yleinen tietosuoja-asetus astui voimaan 25.5.2018, jonka myötä jokaisen, Euroopan unionin sisällä henkilötietoja käsittelevän yrityksen, on noudatettava asetuksessa määrättyjä vaatimuksia. Asetuksen tarkoituksena on yhtenäistää lainsäädäntö Euroopan unionin sisällä sekä lisätä oikeuksia yksilöille. Asetuksen vaatimusten rikkomisesta voi koitua yritykselle merkittäviä sanktiota – jopa 4 % yrityksen globaalista vuosittaisesta liikevaihdosta tai 20 miljoonaa euroa, riippuen kumpi näistä on suurempi. Vaatimusten toteuttamisen tekee haastavaksi konkretian puuttuminen asetuksen virallisesta tekstistä, joka myös vaikeuttaa yksilön tapaa käyttää oikeuksiaan sekä tarkastella käsittelyn lainmukaisuutta.

Tämän tutkimuksen tarkoituksena on selvittää, miten rekisteröidyn oikeuksia koskevat vaatimukset ovat toteutettu yrityksissä ja, kuinka datanhallintaa on hyödynnetty niiden toteuttamisessa. Tutkimus toteutettiin laadullisena tutkimuksena, jossa haastateltiin eri toimialojen tietosuojavastaavia tai yleisen tietosuoja-asetuksen parissa työskenteleviä henkilöitä. Tutkimukseen osallistui kahdeksan henkilöä. Tutkimuksessa havaittiin, että rekisteröidyn oikeuksia koskevat vaatimukset olivat toteutettu eri tavalla osallistuneissa yrityksissä. Tutkimuksessa myös havaittiin, että datanhallinnalla on ollut keskeinen rooli rekisteröidyn oikeuksien toteuttamisessa ja yleisesti asetuksen vaatimustenmukaisuudessa. Johtopäätöksenä voidaan todeta, että eri toimialojen lainsäädäntö ja henkilötietojen käsittelyn laajuus vaikuttavat osaltaan siihen, miten yrityksissä on toteutettu rekisteröidyn oikeuksia koskevat vaatimukset.

Asiasanat: Euroopan yleinen tietosuoja-asetus, rekisteröidyn oikeudet, datanhallinnan hyödyntäminen, master datan hallinta, datan laadun hallinta.

ABSTRACT

Kataja, Joel
The European General Data Protection Regulation: Implementing the data subject's rights and utilizing data management in companies.
Jyväskylä: University of Jyväskylä, 2021, 91 pp.
Information Systems, Master's Thesis
Supervisor(s): Siponen, Mikko

The decrease in costs of data processing is tempting companies to collect more data from their customers, which can lead to misconducts in personal data processing and violations of data subject's privacy. The European General Data Protection Regulation (later referred to as "Regulation") entered into force on 25 May 2018, which means that every company processing personal data within the European Union must comply with the requirements set out in the Regulation*. The purpose of the Regulation is to harmonize legislation around personal data processing within the European Union and to increase the rights of individuals as data subjects. Violation of the requirements set by the Regulation can result in significant sanctions for the company - up to 4% of the company's global annual turnover or 20 MEUR, whichever is higher. The implementation of the requirements is made complicated by the lack of specificity in the official text of the Regulation, which also makes it difficult for an individual to exercise his or her rights and to review the lawfulness of the data processing they are subject to.

The purpose of this study is to find out how the requirements regarding data subject's rights have been implemented in companies and how data management has been utilized in the implementations. The study was carried out as a qualitative study, in which data protection officers from different industries or people working on the general data protection regulation domain were interviewed. In total, eight people participated in the study. The study found out that the requirements regarding data subject's rights were implemented in multiple different ways within the participating companies. The study also discovered that data management has played a key role in enforcing the data subject's rights and, more generally, in complying with the Regulation. In conclusion, the legislation in different sectors and fields of industry as well as the scope of the processing of personal data contribute to the implementation of the requirements regarding data subject's rights in companies.

Keywords: General data protection regulation, data subject rights, utilization of data management, master data management, data quality management

KUVIOT

TAULUKOT

SISÄLLYS

1 JOHDANTO

Yrityksiin kohdistuu nykyään aiempaa enemmän tietosuojaan liittyvää sääntelyä, mutta samaan aikaan kuluttajat ovat tulleet Euroopan unionin asettaman yleisen tietosuoja-asetuksen myötä tietoisimmiksi omista yksityisyyteen liittyvistä oikeuksistaan (Forrester, 2019). Digitaalisella vallankumouksella ja henkilötietojen lisääntyneellä keräämisellä on huomattavia tietoturva haasteita ja riskejä. Alentuneet kulut käsitellä dataa houkuttelee yrityksiä keräämään tietoa enemmän kuin on tarpeen, joka altistaa henkilötietojen väärinkäytölle ja yksityisyyden loukkauksille. (Almeida Teixeira, Mira da Silva, & Pereira, 2019.)

Tutkimus käsittelee Euroopan unionin yleisen tietosuoja-asetuksen rekisteröidyn yksilön oikeuksia ja datanhallinnan hyödyntämistä niiden täyttämisessä. Yleisen tietosuoja-asetuksen keskeisiä tavoitteita on suojella luonnollisten henkilöiden yksityisyyttä, vapauksia ja oikeuksia, mutta samalla yhtenäistää lainsäädäntöä ja tietojen vapaata liikkuvuutta Euroopan sisällä (Team, 2017). Käyttäjien tietojen avulla yritykset pystyvät kehittämään palveluitaan ja kohdentamaan tarjontaa paremmin kuluttajille, mutta varjopuolena on yksilöiden yksityisyyden suojaaminen. Tutkimukset ovat osoittaneet, että kuluttajat ovat huolissaan yksityisyydestään. (Mikkonen, 2014; Eurobarometri, 2015.)

Aiken, Allen, Parker, ja Mattia, (2007.) määrittelevät datanhallinnan koko organisaationlaajuiseksi toiminnoksi, jonka tarkoituksena on ymmärtää nykyisiä ja tulevia datatarpeita sekä tukea organisaation liiketoimintaa. Datanhallinnan tarkoituksena on huolehtia, että sidosryhmien tarpeet täyttyvät datan saatavuuden, turvallisuuden ja laadun suhteen (Brackett, Earley, & Henderson, 2009). Väreen (2019) mukaan Euroopan unionin yleisen tietosuoja-asetuksen edellyttämät vaatimukset ovat osa hyvää datanhallintaa. Datanhallinta koostuu useasta päällekkäisestä osa-alueesta, joista tähän tutkimukseen on otettu aiheen kannalta relevanteimmat.

Datanhallinnan kontekstissa on syytä erotella data, informaatio ja tietämys. Data on merkityksetöntä raaka-ainetta, joka on tallennettuna numeroiden, graafien, kuvien, äänen tai videon muodossa. Informaatio syntyy, kun dataan lisätään konteksti. Kun informaatioon yhdistetään tietotoisuus, ymmärrys ja tilannetaju, niin siitä muodostuu tietämystä. (Rowley, 2007.)

Tutkimusaiheesta, jossa yhdistyy niin datanhallinta, kuin yleinen tietosuoja-asetus, ei itsessään ole juurikaan aiempia tutkimuksia löydettävissä. Tästä huolimatta yleisestä tietosuoja-asetuksesta ja datanhallinnasta on löydettävissä useita tutkimuksia. Euroopan unionin yleistä tietosuoja-asetusta koskevat artikkelit pohjautuvat pitkälti asetuksen lakitekstiin ja asetuksen tuomiin vaikutuksiin sekä asetuksen rekisteröidyn oikeuksien toteuttamiseen. Datanhallintaan liittyvät tutkimukset keskittyvät usein johonkin datanhallinnan osa-alueeseen, eikä niinkään siihen kokonaisuutena.

Euroopan yleisen tietosuoja-asetuksen voimaan astumisesta on kulunut tätä tutkimusta kirjoittaessa muutama vuosi. Tästä syystä on mielenkiintoista selvittää, miten yritykset ovat toteuttaneet rekisteröidyn oikeuksia koskevat vaatimukset ja, kuinka datanhallintaa on hyödynnetty niiden täyttämisessä. Aihe on ajankohtainen, sillä useampi yritys on saanut sanktioita asetuksen rikkomisesta ja lukuisat niistä ovat päätyneet eri medioiden etusivuille.

Yleisen tietosuoja-asetuksen vaatimusten rikkominen voi aiheuttaa yritykselle merkittäviä sanktioita. Sanktiot voivat olla suurimmillaan jopa 20 miljoonaa euroa tai 4 % yrityksen globaalista liikevaihdosta, riippuen kumpi näistä on suurempi (Tankard, 2016). Merkittävien sanktioiden ja mainehaittojen johdosta yritykset ovat mitä luultavimmin asetuksessa määrättyjen vaatimustenmukaisia jollakin tasolla. Tehokkaan datanhallinnan avulla yrityksien on mahdollista sekä täyttää asetuksen vaatimukset, mutta myös tehostaa, turvata ja hallita dataa voimavarana.

Tämän tutkimuksen tarkoitus on selvittää, miten Euroopan unionin yleisen tietosuoja-asetuksen rekisteröityä henkilöä koskevat vaatimukset ovat toteutettu yrityksissä ja, kuinka datanhallintaa on hyödynnetty niiden täyttämisessä. Tutkimuksen teoriaosuudessa selvitetään käsitteellisteoreettisen tutkimusmenetelmän avulla havaintoja aiheeseen liittyvästä kirjallisuudesta. Teoriaosuudessa pyritään vastaamaan seuraaviin tutkimusongelmiin:

- Euroopan unionin yleisen tietosuoja-asetuksen keskeisimmät vaatimukset rekisteröidyn henkilön näkökulmasta ja niiden toteuttaminen yrityksissä.
- Miten datanhallinta helpottaa täyttämään rekisteröidyn henkilön oikeuksia koskevat vaatimukset?

Empiirisessä tutkimuksessa pyritään selvittämään edellisten tutkimusongelmien pohjalta vastaus päätutkimusongelmaan:

- Miten yleisen tietosuoja-asetuksen rekisteröidyn oikeuksia koskevat vaatimukset ovat toteutettu yrityksissä ja, kuinka datanhallintaa on hyödynnetty niiden täyttämisessä?

Tutkimus koostuu johdannon lisäksi teoriaosuudesta, johon lukeutuu kolme sisältölukua ja yhteenveto. Ensimmäinen sisältöluku käsittelee Euroopan unionin yleisen tietosuoja-asetuksen taustaa, rekisteröidyn henkilön oikeuksia ja

asetuksen odotettuja ja todettuja vaikutuksia sekä niiden toteuttamista yrityksissä. Toinen luku koostuu tutkimuksen kannalta relevanteista datanhallinnan osa-alueista: data governancesta, datan laadun-, master datan-, datan tietoturvan- ja dokumenttien- sekä sisällönhallinnasta. Viimeinen sisältöluku keskittyy yleisen tietosuoja-asetuksen rekisteröidyn henkilön oikeuksien täyttämiseen datanhallinnan avulla. Teoriaosuuden päättää yhteenveto. Empiirinen osuus puolestaan koostuu tutkimuksen toteutuksesta, tuloksista, pohdinnasta sekä yhteenvedosta.

Kirjallisuutta tutkimuksen teoriaosioon haettiin hyödyntäen eri tietokantoja, kuten Google Scholaria, IEE:ta, ACM Digital Librarya. Keskeisimmät hakusanat olivat: genaral data protection regulation, data subject rights, data subject access requests, data management, data governance, data quality management, master data management ja data security. Kirjallisuuskatsaukseen on pyritty hyödyntää alan kirjallisuuden lisäksi myös kirjallisuutta kaupallisista lähteistä.

2 EUROOPAN UNION YLEINEN TIETOSUOJA-ASETUS

Tässä luvussa käsitellään Euroopan unionin yleisen tietosuoja-asetuksen taustaa sekä sen vaikutuksia yksilön henkilötietojen käsittelyyn. Aihetta tarkastellaan yksilön lisääntyneiden oikeuksien eli rekisteröidyn oikeuksien kautta, mikä on vastavuoroisesti kasvattanut yritysten velvollisuuksia käsitellä henkilötietoja asetuksen vaatimusten mukaisella tavalla. Tarkasteluun on otettu asetuksen keskeisimmät vaatimukset, jotka liittyvät yksilön oikeuksiin henkilötietojen käsittelyn osalta. Luvun lopussa esitetään kirjallisuudessa havaittuja yleisen tietosuoja-asetuksen odotettuja ja toteutuneita vaikutuksia sekä tutkimuksia rekisteröidyn oikeuksien toteuttamisesta yrityksissä.

2.1 Euroopan unionin yleisen tietosuoja-asetuksen tausta

Euroopan unionin yleinen tietosuoja-asetus on seurausta teknologian nopeasta kehittymisestä ja henkilötietojen käsittelyn lisääntymisestä yrityksissä. Eurobarometrin teettämän tutkimuksen mukaan lähes jokainen palveluntarjoaja kerää käyttäjiltä henkilötietoja ja 71 % osallistujista on sitä mieltä, että henkilötietojen antaminen on edellytys palvelujen käytölle (Eurobarometri, 2015). Lee (2016) lisää, että yritykset myös mahdollisesti käyttävät rekisteröidyn henkilötietoja tarkoituksiin, johon käyttäjä ei ole antanut alun perin suostumustaan. Jotta henkilötietojen käsittely olisi tiukemmin säädeltyä, asetti Euroopan unioni yleisen tietosuoja-asetuksen, jonka tarkoituksena on tiukentaa henkilötietojen käsittelyä sekä yhtenäistää Euroopan unionin jäsenvaltioiden lainsäädäntöä (Euroopan unionin virallinen lehti, 2016). Lainsäädännön yhtenäistämisen avulla on myös tarkoitus helpottaa tiedon vapaata liikkuvuutta Euroopan unionin sisällä sekä turvata yksilön oikeuksia, vapautta ja yksityisyyttä. Asetus astui voimaan 25.5.2018 ja se koskettaa kaikkia Euroopassa toimivia yrityksiä sekä myös muita yrityksiä, jotka käsittelevät Euroopan sisällä asuvien henkilöiden tietoja. (Team, 2017.)

Uusi tietosuoja-asetus korvasi vuonna 1995 laaditun henkilötietodirektiivin, joka ei enää soveltunut nykyajan tilanteeseen henkilötietojen käsittelyn osalta (Ryz & Grest, 2016). Vuoden 1995 henkilötietodirektiivi oli myös hajanainen, sillä jokainen EU:n jäsenvaltio asetti omia lakejaan tukemaan alkuperäisen direktiivin perusperiaatteita. Tämä vaikeutti liiketoiminnan harjoittamista EU:n alueella, sillä jokaisella jäsenvaltiolla oli omat lakinsa henkilötietojen käsittelyyn. Henkilötietodirektiivi tai DIR95 oli asetettu vastaamaan sen aikakauden henkilötietojen käsittelyyn, jolloin vain noin 1 % maailman väestöstä käytti internettiä. Henkilötietoja ei myöskään käsitelty tuolloin sähköisesti samalla mittakaavalla kuin nykyään, joka myös asetti henkilötietojen käsittelylle uudenlaiset haasteet, sillä sähköisesti tuotettua tietoa on vaikeampi suojella. (Tankard, 2016.)

Lainsäädännön yhtenäistäminen luotiin asetuksen kautta, josta tulee automaattisesti osa jokaisen jäsenmaan kansallista oikeusjärjestelmää. Asetuksen ja direktiivin ero on, että direktiivin tapauksessa jäsenvaltiot voivat itse asettaa omat lait vastaamaan direktiivin sisältöä. (Gilbert, 2012.)

Yleinen tietosuoja-asetus ei pelkästään aiheuta ongelmia tai haasteita yrityksille, sillä jotkut asiat helpottuivat asetuksen tulon myötä. Aikaisemmin yritysten oli ilmoitettava henkilötietojen käsittelystä sen jäsenvaltion tietosuojaviranomaiselle, jossa henkilötietojen käsittely tapahtuu. Yleisen tietosuoja-asetuksen mukaan yritys on velvollinen ilmoittamaan käsittelystä jäsenvaltiolle, jossa yrityksen pääkonttori on tai, jossa tapahtuu suurin osa sen käsittelystä. Tämän asetuksen tuoman muutoksen ansiosta yrityksien on mahdollista säästää resurssejaan ja vähentää kustannuksiaan. (Tankard, 2016.)

Asetuksen kontekstissa on syytä erottaa seuraavat käsitteet: rekisterinpitäjä (engl. data controller) ja datan käsittelijä (engl. data processor). Rekisterinpitäjä määritellään uudessa tietosuoja-asetuksessa tarkoittavan oikeushenkilöä, virastoa tai luonnollista henkilöä, joka päättää henkilötietojen käsittelyn tarkoitukset ja tavat käsitellä niitä. Henkilötietojen käsittelijä voi olla ulkoinen toimija, joka käsittelee rekisterinpitäjän henkilötietoja. (Euroopan unionin virallinen lehti, 2016.)

Henkilötietojen määritelmää on laajennettu uudessa tietosuoja-asetuksessa, jossa henkilötiedot määritellään luonnolliseen henkilöön suorasti tai epäsuorasti liittyviksi tunnistetiedoiksi. Tunnistetietoja voi olla yhden tai useamman fyysisen, fysiologisen, geneettisen, psyykkisen, kulttuurillisen tai sosiaalisen tiedon yhdistelmä. (Euroopan unionin virallinen lehti, 2016.) Jatkossa luonnollinen henkilö voidaan tunnistaa henkilötiedoista, joko suorasti tai epäsuorasti, jonka takia esimerkiksi myös IP-osoite ja evästeet lasketaan henkilötiedoiksi. (Tankard, 2016.) Henkilötietojen käsittelyllä viitataan taas asetuksessa toimintoihin, jotka liittyvät tiedon keräämiseen, tallentamiseen, varastointiin, muokkaamiseen tai muuttamiseen, hakemiseen, käyttöön, yhdistämiseen tai yhteensovittamiseen, poistamiseen tai tuhoamiseen, järjestämiseen, jäsentämiseen, säilyttämiseen, hakemiseen, kyselemiseen, luovuttamiseen, siirtämiseen, levittämiseen tai saataville asettamiseen tai rajoittamiseen (Ryz & Grest, 2016).

2.2 Henkilötietojen käsittelyyn liittyvät vaatimukset

Euroopan unionin yleinen tietosuoja-asetus linjaa henkilötietojen käsittelylle kuusi vaatimusta, joiden suhteen henkilötietoja on noudatettava. Ensimmäiseksi henkilötietoja on käsiteltävä lainmukaisesti, asianmukaisesti ja läpinäkyvästi. Käsittelylle on oltava aina jokin tarkoitus, eikä henkilötietoja saa käsitellä, mikäli alkuperäinen tarkoitus ei ole enää yhteensopiva myöhemmän tarkoituksen kanssa. Asetuksessa tähän viitataan käyttötarkoitussidonnaisuudella. Tietojen minimointi liittyy myös edellä mainittuun, mutta on erillinen periaate asetuksessa. Sen mukaan henkilötietojen on oltava rajattu käyttötarkoitukseen ja olennaisia niiden käytön kannalta. Täsmällisyys puolestaan viittaa, että kaikki henkilötiedot on oltava paikkansapitäviä ja tarvittaessa päivitettyjä. Virheelliset henkilötiedot on poistettava tai oikaistava mahdollisimman pian. Henkilötietojen säilyttämistä on myös rajoitettava uuden asetuksen myötä. Nykyään tietoja ei saa säilyttää tunnistettavassa muodossa kauempaa kuin niiden käsittelylle on tarpeen. Asetuksessa otetaan myös kantaa tietojen turvalliseen käsittelyyn ja eheyteen. Tietojen käsittely on turvattava tarvittavilla toimintatavoilla, jotta luvaton tai lainvastainen käsittely on estettävissä. Organisaatioiden on myös rakenteellisten tai teknisten toimien kautta pyrittävä estämään vahingossa tapahtuva tietojen häviäminen tai poistaminen. (Euroopan unionin virallinen lehti, 2016.)

2.3 Rekisteröidyn oikeudet

Yleisen tietosuoja-asetuksen keskeinen periaate on suojella yksilöiden yksityisyyttä, mikä on johtanut luonnollisten henkilöiden oikeuksien lisääntymiseen. Tietosuoja-asetus lisää rekisteröidyn oikeuksia kontrolloida henkilötietojaan entistä paremmin. (Team, 2017.) Tässä kappaleessa esitetään yleisen tietosuoja-asetuksen rekisteröidyn oikeuksia koskevat keskeisimmät vaatimukset. Tässä luvussa ei erikseen käsitellä artikloja 13, 14 ja 19.

2.3.1 Oikeus saada läpinäkyvää informaatiota henkilötietojen käsittelystä (Artikla 12)

Rekisteröidyillä henkilöillä on asetuksen mukaan oikeus saada tietoa henkilötietojen käsittelystä ja rekisteröidyn oikeuksista eli toisin sanoen toiminnan täytyy olla läpinäkyvää. Informaation vaaditaan olevan selkeästi ymmärrettävää ja helposti saatavilla olevaa. Myös standardoitujen kuvien avulla on mahdollista havainnollistaa henkilötietojen käsittelyprosessia. Tiedot on toimitettava pyydettäessä rekisteröidylle kuukauden sisällä pyynnön lähettämisestä. (Tikkinen-Piri, Rohunen, & Markkula, 2018.) Euroopan unionin virallinen lehti vielä täs-

mentää, että tiedot tulee esittää yksinkertaisella ja selkeällä kielellä etenkin, kun kyseessä on lapsi (Euroopan unionin virallinen lehti, 2016).

Läpinäkyvyys vaatii yrityksiltä tietosuojaselosteiden uudelleen kirjoittamista, jos niistä ei käy ilmi henkilötietojen käsittely riittävän yksinkertaisesti ja ymmärrettävästi. Yritysten tulee miettiä ja suunnitella kommunikointi, kun kyseessä oleva rekisteröity on lapsi.

2.3.2 Oikeus saada pääsy omiin tietoihin (Artikla 15)

Rekisteröidyllä on yleisen tietosuoja-asetuksen, kuten myös edellisen DIR95 mukaan oikeus saada pääsy omiin henkilötietoihin. Rekisteröidyllä on oikeus saada pyytämällä rekisterinpitäjältä vahvistus, mikäli häntä koskevia henkilötietoja käsitellään. Jos rekisteri sisältää rekisteröidyn henkilötietoja, on rekisterinpitäjällä velvollisuus toimittaa hänelle tiedot niitä pyydettäessä. Jo DIR95:ssä oli säädetty, että rekisterinpitäjän on toimitettava rekisteröidylle seuraavat tiedot omien henkilötietojen lisäksi: käsittelyn tarkoitus, prosessoidun tiedon kategoriat ja tiedon vastaanottajat/vastaanottajaryhmät, joille tiedot ovat tarkoitus toimittaa. Näiden lisäksi yleisessä tietosuoja-asetuksessa velvoitetaan rekisterinpitäjää lähettämään rekisteröidylle tiedot säilyttämisajasta ja tiedot sen määrittämisestä. Rekisteröidyllä on myös oikeus pyytää henkilötietojen oikaisemista, poistamista, tietojen käsittelyn rajoittamista tai sen vastustamista, tehdä valitus valvontaviranomaiselle, käsittelyyn liittyvästä logiikasta ja käsittelyn merkittävyydestä, seurauksista sekä tiedot automaattisen päätöksenteon olemassaolosta, johon luetaan mukaan myös profilointi. (Tikinen-Piri ym., 2018.)

2.3.3 Oikeus omien tietojen oikaisuun (Artikla 16)

Uuden tietosuoja-asetuksen myötä rekisteröidyllä on oikeus pyytää rekisterinpitäjää oikaisemaan virheelliset, epätarkat sekä puutteelliset tiedot ilman turhaa viivästystä. Täten rekisteröidyllä on oikeus täydentää omat henkilötietonsa toimittamalla rekisterinpitäjälle lisäselvityksen, mikäli tiedot ovat virheelliset tai puutteelliset. (Euroopan unionin virallinen lehti, 2016.)

2.3.4 Oikeus tietojen poistamiseen (”oikeus tulla unohdetuksi”) (Artikla 17)

Oikeus tietojen poistamiseen tarkoittaa, että yksilöllä on oikeus saada tietonsa poistetuksi yrityksen järjestelmästä. Yritykset ovat velvollisia poistamaan rekisteröidyn tiedot, kun käyttäjä niin pyytää. (Shao & Oinas-Kukkonen, 2019.) Oikeus tietojen poistamiseen on mahdollista, mikäli jokin seuraavista ehdoista täyttyy: henkilötietoja ei enää käytetä niihin tarkoituksiin, johon ne alun perin kerättiin, rekisteröity peruuttaa suostumuksensa henkilötietojen käsittelyyn, rekisteröity vastustaa käsittelyä, henkilötietoja käsitellään lainvastaisesti tai ne ovat kerätty tietoyhteiskunnan palvelujen tarjoamiseksi. Jos henkilötietojen käsittelylle on jokin perusteltu syy, kuten esimerkiksi henkilön verotiedot, ei re-

kisteröidyllä ole tällöin oikeutta tietojensa poistamiseen. Mikäli rekisterinpitäjä on tehnyt henkilötiedoista julkisia, niin tietojen poistaminen velvoittaa rekisterinpitäjää ilmoittamaan muille rekisterinpitäjille kyseisen henkilö tietojen, linkkien sekä mahdollisten kopioiden poistamisesta. (Tikkinen-Piri ym., 2018.)(Tikkinen-Piri ym., 2018)

Lee (2016) ottaa kantaa ”oikeuteen tulla unohdetuksi”, sillä internettiin ladattua tietoa voi olla jopa mahdotonta poistaa kokonaan, jos joku muu käyttäjä on ladannut tiedot ennen niiden poistamista. Tästä johtuen tietojen täydellinen poistaminen on käytännössä täysin mahdotonta, mutta asetuksen kontekstissa riittävää on, että tiedot poistuvat rekisterinpitäjien rekistereistä. (Lee, 2016.)

Teknisestä näkökulmasta katsottuna tietojen poistaminen tietovarastosta on monimutkaisempaa kuin asetuksesta voi ymmärtää. Nykyään melkein kaikki suurimmat tietovarastot noudattavat ACID vaatimuksia, jotka asettavat tietovarastojen toimintaan tiettyjä ominaisuuksia. Rekisteröidyn tietojen poistaminen vaatii näiden vaatimusten huomioimista, sillä tiedot voivat olla tallennettuina lokitietoihin, varmuuskopioihin tai eri paikkoihin sisäisten tietokantamekanismien sisälle. Tietojen täydellinen poistaminen vaatii näiden tilojen tunnistamisen ja tiedon ylikirjoittamisen sattumanvaraisella tiedolla. Tietovarastojen johdonmukaisuus kuitenkin vaarantuu tai jopa tuhoutuu kokonaan etenkin tapahtumalokien kohdalla, jos rekisteröidyn tiedot ylikirjoitetaan sattumanvaraisella tiedolla. (Villaronga, Kieseberg, & Li, 2018.) Shao ja Oinas-Kukkonen (2009) lisäävät, että onnistunut tietojen poistaminen vaatii tietojen riittävää dokumentointia, selvitystä siitä kuinka tietoja säilytetään sekä kelle osapuolille tiedot ovat jaettu.

Yleisesti tietopyyntöihin ja erityisesti rekisteröidyn poistopyyntöihin vastaaminen asetuksessa määrätyn aikaikkunan sisällä vaatii yrityksiltä valmiit ja riittävät prosessit sekä tekniset ratkaisut. Yritysten tulee myös suunnitella kommunikointikanava, jota kautta rekisteröityjä informoidaan aiheesta. (Tikkinen-Piri ym., 2018.)

2.3.5 Oikeus käsittelyn rajoittamiseen (Artikla 18)

Rekisteröidyllä on oikeus rajoittaa omien henkilötietojensa käsittelyä, mikäli yksi seuraavista ehdoista täyttyy: rekisteröity kiistää henkilötietojensa paikkansapitävyyden, käsittely ei ole lainmukaista, rekisterinpitäjä ei enää tarvitse henkilötietoja, mutta rekisteröity tarvitsee niitä esimerkiksi oikeudellisista syistä tai rekisteröity on vastustanut henkilötietojen käsittelyä, jolloin käsittelyä rajoitetaan tilanteen selvittämisen ajaksi. (Euroopan unionin virallinen lehti, 2016.)

2.3.6 Oikeus siirtää tiedot järjestelmästä toiseen (Artikla 20)

Rekisteröidyllä on oikeus siirtää tiedot järjestelmästä toiseen, jolloin rekisterinpitäjällä on velvollisuus toimittaa rekisteröidyn henkilötiedot toiselle rekisterinpitäjälle tai rekisteröidylle yleisesti hyväksytyssä ja koneellisesti luettavassa muodossa. Kuitenkin sillä ehdoin, että tietojen siirtäminen on teknisesti mah-

dollista. (Euroopan unionin virallinen lehti, 2016.) Yritysten on toimitettava rekisteröidyille kopio heidän antamistaan tiedoista, jos rekisteröity näin vaatii. Yritysten on myös tarkistettava, että tiedot lähetetään rekisteröidyille henkilöille yhtenäisessä muodossa, jotta rekisteröidyn henkilön on mahdollista vaivattomasti käyttää tietoja tai siirtää ne toiselle palveluntarjoajalle. (Shao & Oinas-Kukkonen, 2019).

Oikeuden siirtää tiedot järjestelmästä toiseen kannattaa pyrkiä tekemään automaattisesti, jolloin vuorovaikutukseen rekisteröidyn henkilön kanssa ei tarvitse kuluttaa yrityksen resursseja. Rekisteröidyn henkilön pyytäessä tietojaan rekisterin pitäjältä, järjestelmä automaattisesti luo kopion rekisteröidyn tiedoista, jonka jälkeen rekisteröity voi itse käydä lataamassa omat tietonsa. Tietojen muodossa voidaan hyödyntää yleisten standardien hyväksymiä tiedostomuotoja. (Shao & Oinas-Kukkonen, 2019).

2.3.7 Oikeus vastustaa tietojenkäsittelyä (Artikla 21)

Yksilöillä on myös oikeus vastustaa tietojenkäsittelyä. Kyseinen oikeus määritellään yleisessä tietosuoja-asetuksessa seuraavasti:

> Rekisteröidyllä on oikeus henkilökohtaiseen erityiseen tilanteeseensa liittyvällä perusteella milloin tahansa vastustaa häntä koskevien henkilötietojen käsittelyä, joka perustuu 6 artiklan 1 kohdan e tai f alakohtaan, kuten näihin säännöksiin perustuvaa profilointia. Rekisterinpitäjä ei saa enää käsitellä henkilötietoja, paitsi jos rekisterinpitäjä voi osoittaa, että käsittelyyn on olemassa huomattavan tärkeä ja perusteltu syy, joka syrjäyttää rekisteröidyn edut, oikeudet ja vapaudet tai jos se on tarpeen oikeusvaateen laatimiseksi, esittämiseksi tai puolustamiseksi. Jos henkilötietoja käsitellään suoramarkkinointia varten, rekisteröidyllä on oikeus milloin tahansa vastustaa häntä koskevien henkilötietojen käsittelyä tällaista markkinointia varten, mukaan lukien profilointia silloin kun se liittyy tällaiseen suoramarkkinointiin. (Euroopan unionin virallinen lehti. 2016.)

Tikkinen-Piri ja kumppanit (2018) korostaa, että rekisteröidyllä on oikeus vastustaa henkilötietojen käsittelyä, johon sisältyy myös profilointi. Asetuksen mukaan oikeus vastustaa henkilötietojen käsittelyä tulee esittää selkeästi ja erillään muusta informaatiosta. Jos rekisteröity henkilö vastustaa henkilötietojen käsittelyä, yrityksellä ei ole enää oikeutta käsitellä rekisteröidyn henkilötietoja, ellei rekisterinpitäjällä ole osoittaa käsittelyyn perusteltua ja tärkeää syytä. (Tikkinen-Piri ym., 2018.)

2.3.8 Automaattinen päätöksenteko ja profilointi (Artikla 22)

Euroopan unionin yleisessä tietosuoja-asetuksessa automaattisesta päätöksenteosta ja profiloinnista linjataan seuraavasti:

> Rekisteröidyllä on oikeus olla joutumatta sellaisen päätöksen kohteeksi, joka perustuu pelkästään automaattiseen käsittelyyn, kuten profilointiin, ja jolla on häntä kos-

kevia oikeusvaikutuksia tai joka vaikuttaa häneen vastaavalla tavalla merkittävästi. (Euroopan unionin virallinen lehti, 2016.)

Asetuksen mukaan rekisterinpitäjän on ilmoitettava rekisteröidyn oikeus vastustaa automaattista käsittelyä mahdollisimman selkeästi ja erotettuna muusta informaatiosta. Rekisterinpitäjän on lopetettava käsittely välittömästi, mikäli tällä ei ole esittää tarvittavia ja perusteltuja syitä käsittelylle. Esimerkiksi suoramarkkinointiin liittyvä profilointi on tämän kaltaista käsittelyä, johon rekisteröidyllä on oikeus olla joutumatta. (Tikkinen-Piri ym., 2018.)

2.4 Euroopan unionin yleisen tietosuoja-asetuksen vaikutukset

Euroopan unionin asettama yleinen tietosuoja-asetus on aiheuttanut yrityksille suuria haasteita täyttää asetuksen vaatimukset. Yleinen tietosuoja-asetus vaikuttaa voimakkaimmin yrityksiin, joiden ydintoimintana on henkilötietojen käsittely (Seo, Kim, Park, Park, & Lee, 2018). Kirjallisuudessa on myös pohdittu paljon asetuksen vaikutuksia yritysten toimintoihin ennen asetuksen voimaan astumista, joita käsitellään alempana luvussa ”Odotetut vaikutukset”. Kirjallisuudessa on löydettävissä artikkeleita, joissa on jo analysoitu yleisen tietosuoja-asetuksen todettuja vaikutuksia yritysten toimintoihin.

2.4.1 Odotetut vaikutukset

Monissa artikkeleissa spekuloitiin Euroopan unionin yleisen tietoturva-asetuksen vaikutuksia yritysten toimintoihin. Tankard (2016) esittää artikkelissaan, että erään tutkimuksen mukaan noin puolet organisaatioista uskoivat saavansa sanktioita yleisen tietosuoja-asetuksen laiminlyömisestä. Samaisessa tutkimuksessa enemmistö yrityksistä arveli asetuksen kasvattavan kustannuksia harjoittaa liiketoimintaa Euroopassa. Osa yrityksistä myös pohti budjettien kasvattamista noin 10 %:lla, jotta asetuksen seuraukset voitaisiin kattaa. (Tankard, 2016.)

Greengardin (2018) mukaan yritykset ovat olleet huolissaan asetuksen vaikutuksista yritysten innovointiin, sillä asetus rajoittaa tiedon käsittelyä yrityksissä. Tiedon käsittely hankaloituu esimerkiksi applikaatioissa, tietovarastoissa ja erilaisissa online-palveluissa. Myös tiedon käyttö nähdään rajoittuvan asetuksen seurauksena etenkin mainonnassa, joka korostuu rekisteröidyn oikeudessa vastustaa suoramarkkinointia. Tiedon käsittelyn rajoittaminen saattaa myös vaikuttaa järjestelmiin, jotka ovat riippuvaisia tekoälystä, mutta myös esimerkiksi itseohjautuviin autoihin ja robotiikkaan. Organisaatiot voivat kyseisissä tilanteissa päätyä ratkaisuun, jossa Euroopan unionin kansalaisille on oma tietovarasto. (Greengard, 2018.)

Ennen asetuksen voimaan astumista pohdittiin myös kuluttajien reagoimista omiin oikeuksiinsa, kuten oikeuteen poistaa, oikaista tai saada pääsy omiin tietoihin. Yritysten on pohdittava uusia ratkaisuja, joiden avulla voidaan

hallita suuria määriä rekisteröityjen pyyntöjä koskien uudistuneita oikeuksiaan. Mahdollisia teknologisia ratkaisuja voisi olla esimerkiksi lohkoketjuteknologia tai tiedonhallintajärjestelmät. (Greengard, 2018.) Kuluttajien reagointia on haasteellista ennakoida, sillä myös yrityksen koko ja imago voi vaikuttaa siihen, kuinka kuluttajat aikovat hyödyntää oikeuksiaan. Ongelmaksi voi muodostua myös pk-yrityksien resurssit varautua äkilliseen pyyntöjen määrän kasvuun, joka luultavasti laantuu melko nopeasti.

Yleisen tietosuoja-asetuksen pohdittiin vaativan yrityksiltä olemassa olevien politiikoiden kehittämistä ja uudelleen kirjoittamista. Asetuksen vaatimusten täyttäminen ja yrityksien uusien toimintojen kehittäminen on nähty vaativan suuria investointeja ja vievän paljon resursseja yritykseltä. Vaikutukset on arveltu johtavan IT-osastojen budjettien kasvuun lisääntyneiden koulutusten, politiikoiden, menettelytapojen ja teknologioiden rahoittamiseksi. (Gilbert, 2011.)

Seo ja kollegat (2018) tutkivat yleisen tietosuoja-asetuksen taloudellisia vaikutuksia esineiden internetin toimialaan Gordon-Loeb -mallin avulla. Gordon-Loeb -malli on matemaattinen malli, jota käytetään tietoturva investointien optimaalisen tason selvittämiseen. Tutkimuksesta selvisi, että yleisen tietosuoja-asetuksen laadulliset vaikutukset vaikuttaisivat ehkäisevien ja oikeudellisten kulujen kasvamiseen. Määrällisten vaikutusten arviointi ennusti, että tietosuoja-asetuksen vaikutukset kasvattaisivat esineiden internettiin (engl. IoT, Internet of Things) keskittyneiden yritysten kustannuksia keskimäärin noin kolme-neljä kertaa. Enimmillään kustannukset voivat arvioin mukaan nousta 18 kertaiseksi. Tutkijat arvioivat, että yleinen tietosuoja-asetus vaikuttaa rajusti esineiden internetin valmistajiin, sillä laitteet hyödyntävät suuria määriä käyttäjien tietoja, jotka kulkevat laitteiden välillä. (Seo ym., 2018.)

Eräässä tutkimuksessa selvitettiin yleisen tietosuoja-asetuksen vaikutuksia uusiin ja kehittyviin teknologiayrityksiin kohdistuviin investointeihin. Tutkimuksessa arveltiin asetuksen nostavan kustannuksia tiedon pääsyn, keruun, vastuun sekä kaupallistamisen kautta, joilla on suora vaikutus liiketoimintamalleihin ja epäsuora vaikutus asiakkaiden hankintakustannuksiin. Toinen mahdollinen ja todennäköinen kustannuksien aiheuttaja on vaatimusten täyttäminen, jonka myös arvioidaan kasvattavan kustannuksia harjoittaa liiketoimintaa. (Jia, Jin & Wagman, 2018.)

2.4.2 Todetut vaikutukset

Yleisen tietosuoja-asetuksen voimaan astumisesta on jo useampi vuosi tätä tutkimusta kirjoittaessa ja tuona aikana on ehditty havaita asetuksen tuomia vaikutuksia. Kirjallisuudesta on löydettävissä artikkeleita ja tutkimuksia aiheesta. Tähän on koottu asetuksen todettuja vaikutuksia yritysten toimintoihin.

Yleisellä tietosuoja-asetuksella on havaittu olevan vaikutusta internetsivustojen tietosuojaselosteisiin. Degeling ja kollegat (2018) havaitsivat tutkimuksessaan, että monissa maissa lisättiin uusia ja päivitettiin olemassa olevia tietosuojaselosteita internetsivustoille ennen asetuksen voimaan astumista. Tämän

jälkeen selosteiden lisääminen ja päivittäminen hidastui voimakkaasti. Tutkimuksessa selvisi, että yleinen tietosuoja-asetus lisäsi internetsivustojen yleistä läpinäkyvyyttä. Myös monilla sivustoilla huomioitiin muut yleisen tietosuoja-asetuksen vaatimukset, kuten ilmoittaminen rekisteröidyn oikeuksista ja evästeiden käytöstä. Siitä huolimatta monilta sivustoilta puuttuu mekanismeja, joiden avulla käyttäjien on mahdollista hyväksyä tai kieltää tietojensa käsitteleminen. (Degeling ym., 2018.)

Jia, Jin ja Wagman (2019) havaitsivat tutkimuksessaan yleisen tietosuoja-asetuksen vaikuttavan negatiivisesti EU:n jäsenmaissa toimiviin uusiin ja kehittyviin yrityksiin verrattuna samaisiin yrityksiin, jotka toimivat EU:n ulkopuolella. Negatiiviset vaikutukset havaittiin EU:n jäsenmaiden yritysten rahoituskierrosten määrän vähentymisessä verrattuna samankaltaisiin yrityksiin Yhdysvalloissa. Samassa tutkimuksessa havaittiin odotetusti, että yritykset, joiden liiketoimintaan liittyy dataa ja sen käsittelyä, altistuvat asetuksen negatiivisille vaikutuksille herkemmin. Tutkimuksessa myös todettiin, että B2C-alan organisaatioihin (engl. business to consumers) kohdistuu enemmän negatiivisia vaikutuksia kuin B2B-alan toimijoihin (engl. business to businesses). Tämä luultavasti johtuu siitä, että B2C-alan yritykset altistuvat enemmän asetukselle ja niihin kohdistuu rekisteröityjen oikeuksia koskevat vaatimukset. (Jia ym., 2019.) Tutkimuksen tuloksia kannattaa tarkastella kriittisesti, sillä EU:n ulkopuolella toimivat yritykset voivat myös joutua asetuksen vaatimusten kohteeksi, jos ne käsittelevät EU:n jäsenmaiden kansalaisten tietoja. On myös syytä huomioida, että altistuminen asetuksen vaatimuksille ei aiheuta pysyviä negatiivisia vaikutuksia, vaikka yritys joutuisikin alkuun tekemään investointeja. Asetuksen vaatimusten täyttäminen saattaa tulevaisuudessa tuoda kustannussäästöjä yritykselle.

Shah, Banakar, Shastri, Wasserman ja Chidambaram (2019) tutkivat tutkimuksessaan tietosuoja-asetuksen vaikutuksia tiedon varastointijärjestelmiin (engl. storage systems). Tutkimuksessa selvitettiin kuusi ominaisuutta, joiden avulla varastointijärjestelmät täyttävät asetuksen vaatimukset. Näiden lisääminen kuitenkin hidastaa järjestelmän kykyä suorittaa toimintoja. Yleisen tietosuoja-asetuksen mukainen järjestelmä on noin 20 kertaa hitaampi kuin ei-modifioitu järjestelmä. (Shah ym., 2019.)

2.5 Aikaisempia tutkimuksia rekisteröidyn oikeuksista ja niiden toteuttamisesta

Seuraavaksi käsitellään aikaisempia tutkimuksia rekisteröidyn oikeuksien toteuttamisesta yrityksissä.

2.5.1 Tietosuojaseloste

Ausloos ja Dewitte (2018) tutkimuksessa havaittiin tietosuojaselosteiden osalta, että ne olivat vaikeasti löydettävissä yrityksen sivuilta. Syy tähän oli, että yritykset eivät noudattaneet nykyajan yleistä standardia, jossa linkki tietosuojaselosteeseen sisällytetään jokaisen sivun alaosaan. Suurimmassa osassa yrityksiä tietosuojapolitiikoista puuttui olennaisia tietoja, jotka oli määrätty henkilötietosuojadirektiivin ja yleisen tietosuoja-asetuksen vaatimuksissa. Puolet politiikoista oli rakenteeltaan ja kirjoitustyyliltään epäselviä ja haastavia. Useat yritykset myös tarjosivat yhden tietosuojaselosteen, joka kattoi kaikki heidän palvelunsa. Tutkijoiden mielestä tämä saattaa vaikuttaa rekisteröidylle selkeältä toimenpiteeltä, mutta todellisuudessa yksi keskitetty politiikka ei välttämättä huomioi eri palveluiden erityispiirteitä. (Ausloos & Dewitte, 2018.)

Tutkijoiden suositukset rekisterinpitäjille on, että tietosuojaselosteen tulisi olla helposti ymmärrettävä ja löydettävissä yrityksen sivuilta. Rekisterinpitäjien tulisi myös välttää yhden keskitetyn tietosuojaselosteen käyttämistä, jos sen toimintaan kuuluu useampi eri palvelu. Lisäksi tietosuojaselosteessa tulisi olla selkeästi esillä rekisteröidyn oikeudet. (Ausloos & Dewitte, 2018.)

Myös Hoiry ja Norris (2015) havaitsivat tutkimuksessaan puutteita yritysten tietosuojaselosteissa, tosin tutkimus oli toteutettu ennen yleisen tietosuoja-asetuksen voimaan astumista. Kuten edellisessä kappaleessa todettiin, niin myös Hoiry ja Norris linjaa, että tietosuojaosion tulisi olla selkeästi näkyvissä ja helposti saatavilla yrityksen sivuilla. He lisäävät myös, että tietosuojaselosteen tulisi sisältää tietopyynnöille tarkoitettu mallipohja ja ohjeistus tietopyynnön tekemiseen ja kenelle se tulisi ohjata. Tietosuojaseloste tulisi aina tarjota rekisteröidylle ennen palveluun rekisteröitymistä. (Hoiry & Norris, 2015.)

2.5.2 Rekisteröidyn tunnistaminen tietopyynnön yhteydessä

Yleisen tietosuoja-asetuksen rekisteröidyn oikeuksia koskevien tietopyyntöjen yhteydessä on tunnistettava pyyntöä tekevä henkilö, jotta tiedot eivät vuoda väärälle henkilölle. Tästä huolimatta yleisessä tietosuoja-asetuksessa ei ole kirjattu konkreettisia vaatimuksia, kuinka rekisteröity tulisi tunnistaa tietopyyntöjen yhteydessä. Konkretian puuttuminen asetuksen virallisesta tekstistä tuo haasteita muun muassa, kun tarkistetaan käsittelyn lainmukaisuutta ja vaikeuttaa rekisteröidyn oikeuksien käyttöä. (Boniface, Fouad, Bielova, Lauradoux, & Santos, 2019.)

Pyyntöä tekevä henkilö tulisi tunnistaa ennen kuin pyyntöön vastataan. Rekisterinpitäjällä on oikeus pyytää tarkentavia tietoja, mikäli rekisteröidyn henkilöllisyyttä ei voida tarpeeksi vahvasti varmistaa annetuista tiedoista. Sähköisesti rekisteröidyn olisi yleisesti pystyttävä käyttämään omia tunnistetietojaan henkilöllisyyden todentamiseksi ja todistamiseksi. Jos rekisteröity voi kirjautua palveluun kirjautumistietojen avulla, niin pyyntöjen täyttämisen yhteydessä tulisi käyttää vähintäänkin samaa suojaustasoa kuin kirjautumisprosessissa. Jos kyseessä on erittäin sensitiivisiä tietoja, niin on hyvä käyttää vahvem-

pia menetelmiä rekisteröidyn tunnistamiseen, kuten monivaiheista tunnistautumista. Tästä huolimatta rekisterinpitäjä voi käyttää vain kohtuullisia keinoja rekisteröidyn henkilön tunnistamiseksi. Esimerkiksi, jos rekisteröity voidaan jo tunnistaa olemassa olevista tiedoista, niin rekisterinpitäjän ei tulisi pyytää enää lisätietoja tunnistamista varten. Jossakin tapauksissa rekisterinpitäjän on myös hyvä varmistaa, että pyynnön tekijä oikeasti teki tietopyynnön omista tiedoistaan. Voidaan esimerkiksi lähettää rekisteröidylle palvelun rekisteröinnin yhteydessä annettuun sähköpostiin vahvistuslinkki, jolla voidaan välttää tilanne, jossa käyttäjän tunnukset ovat joutuneet vääriin käsiin. (Singh & Cobbe, 2019.)

Ausloos ja Bewitte (2018) havaitsivat myös tutkimuksessaan, että osa palveluntarjoajista vaativat kopion rekisteröidyn henkilötodistuksesta tunnistamista varten. Vaikka yrityksen on tunnistettava rekisteröity ennen tietopyynnön käsittelemistä, rekisteröity voi tuntea epämieluisaksi joutuakseen luovuttamaan vielä enemmän tietoja oikeuksiensa käyttämiseksi. Varsinkin kun tutkimuksessa havaittiin, että osa yrityksistä ei kelpuuttanut henkilötodistuksia, jossa osa epärelevanteista tiedoista oli peitetty. (Ausloos & Bewitte, 2018.)

Boniface ja kollegat (2019) etsivät tutkimuksessaan vastausta seuraaviin kysymyksiin: Onko rekisteröidyn turvallista tehdä tietopyyntö omista tiedoistaan? Milloin rekisterinpitäjällä on tarpeeksi tietoa rekisteröidyn todentamiseksi? He analysoivat tutkimuksessaan eri maiden tietosuojaviranomaisten suosituksia ja suosittujen verkkosivustojen ja kolmannen osapuolten seurantapalveluiden tunnistamiskäytäntöjä. Tutkimuksessa havaittiin, että osa rekisterinpitäjistä käyttivät turvattomia tai epäilyttäviä menetelmiä rekisteröidyn tunnistamiseen. Rekisterinpitäjien yleisin virhe oli tunnistaa rekisteröity kansallisen henkilötodistuksen kopion perusteella, joka toimitettiin suojaamattomana rekisterinpitäjälle. Vaarana on, että kopio henkilötodistuksesta päätyy vääriin käsiin ja hyökkääjä voi esiintyä sen avulla kyseisenä henkilönä toisille rekisterinpitäjille tietopyynnön yhteydessä. Tutkijat ehdottavatkin, että henkilötodistuksen kopioissa käytettäisiin vesi- ja aikaleimoja, jotka osoittavat kopion olevan vain kyseiselle rekisterinpitäjälle ja näin ollen ehkäisevät sen hyödyntämisen tulevaisuudessa. Tutkijat myös suosittelevat käyttämään verkkoportaalin kautta tapahtuvassa viestinnässä vähintään suojattua yhteyttä (https) ja salasanaa. Tutkimuksessa kuitenkin todetaan, että paras ratkaisu olisi käyttää monivaiheista tunnistautumista.

Kröger, Lindemann ja Herrmann (2020) tutkivat, kuinka mobiiliapplikaatioiden toimittajat vastaavat rekisteröidyn tietopyyntöihin. Tutkimus toteutettiin pitkittäistutkimuksena, jossa tehtiin kolme iteraatiota vuosien 2015–2019 aikana. Tutkimuksessa havaittiin, että yritykset ovat alkaneet ottamaan paremmin huomioon rekisteröidyn oikeuksia, mutta parannettavaa vielä löytyy. Etenkin henkilön tunnistaminen yhteydessä havaittiin, että suurin osa yrityksistä ei tarkastanut rekisteröidyn henkilöllisyyttä ennen tietojen lähettämistä. Tosin tutkimuksessa tietopyynnön tiedustelemiseen käytettiin rekisteröinnin yhteydessä ilmoitettua sähköpostia, joka itsessään tarjoaa minimaalisen tavan todennukseen. Tutkimuksessa havaittiin myös, että vaikka tunnistusmekanismit

olivat käytössä, niin ne kierrettiin huomaamatta yrityksen toimesta. Tutkijat saivat jopa vahingossa pääsyn toisen henkilön tietoihin. (Kröger ym., 2020.)

Singh ja Cobbe (2019) esittävät artikkelissaan organisaatioiden toteutettuja toimenpiteitä vastatakseen rekisteröidyn oikeuksia koskeviin vaatimuksiin. Ensinnäkin tietopyynnöt ja tietojen lataaminen tapahtuu yrityksen omalla alustalla, jolloin rekisteröity validoidaan käyttämällä samoja menetelmiä kuin palvelua käyttäessä. Toiseksi rekisterinpitäjän tulisi ilmoittaa tietopyynnöstä rekisteröidylle rekisteröinnin yhteydessä annettuun sähköpostiosoitteeseen ja mieluiten vielä pyytää vahvistamaan pyyntö. Näin vältetään tilanteet, jossa rekisteröidyn tili on vaarantunut. Viimeiseksi tiedot tulisi olla ladattavissa vain tietyn ajan yrityksen alustalta, jolloin vältetään tietojen mahdollinen vuotaminen tai jakelu tulevaisuudessa. (Singh & Cobbe, 2019.)

Martino ja kumppanit (2019) selvittivät tutkimuksessaan eri toimialoilla toimivien organisaatioiden käytäntöjä rekisteröidyn todentamiseen tietopyyntöjen yhteydessä. Tutkimukseen kuului empiirinen osuus, jossa tutkijat yrittivät saada pääsyn toisen henkilön tietoihin. Tässä onnistuttiin 15 yrityksen kohdalla, kun otos oli 55 organisaatiota. Tutkijat myös havaitsivat, että suurin osa organisaatioista käsitteli tietopyynnöt manuaalisesti, kun vain 14 organisaatiota käytti prosessointiin automatiikkaa. Tutkimuksessa myös havaittiin, että yrityksillä oli eri menetelmiä rekisteröidyn tunnistamiseen. Osa yrityksistä vaati ensiksi rekisteröidyn kirjautumisen palveluun, jonka jälkeen rekisteröidyn tuli kirjautua vielä erilliselle sivustolle, jotta tietopyyntö voitiin toteuttaa. Suurimmassa osassa organisaatioita tietopyyntö lähetettiin tietosuojavastaavan sähköpostiin, joka löytyi tietosuojaselosteesta. Ideaalissa tilanteessa rekisterinpitäjän tulisi toimittaa tiedot ainoastaan, jos tietopyynnön lähettäjän sähköpostiosoite on löydettävissä yrityksen järjestelmästä. Tutkittavista yrityksistä 12 noudatti tätä politiikkaa, mutta viisi yritystä oli valmis tunnistamaan rekisteröidyn myös muun käyttäjäkohtaisen tiedon perusteella, kun rekisteröidyllä ei ollut enää pääsyä omaan sähköpostiinsa. Osassa organisaatioista tunnistamiseen käytettiin myös henkilökortin kopiota, joka liitettiin sähköiseen lomakkeeseen tai sähköpostin liitetiedostoksi. Myös pienessä osassa yrityksiä rekisteröity tunnistettiin kotiosoitteen, puhelinsoiton tai muun käyttäjäkohtaisen tiedon perusteella. (Martino ym., 2019.)

2.5.3 Tietopyynnön lähettäminen ja käsitteleminen

Yrityksillä on käytössä eri tapoja ottaa vastaan rekisteröidyn tietopyyntö. Ausloos ja Dewitte (2018) selvittivät tutkimuksessaan rekisteröidyn oikeuksiin liittyvien vaatimusten toteuttamista yrityksissä. Tutkimus toteutettiin vuosi ennen yleisen tietosuoja-asetuksen voimaan astumista, jolloin yritykset noudattivat vielä henkilötietodirektiiviä – toiselta nimeltään DIR 95/46. Tästä huolimatta havaittuja tapoja lähettää tietopyyntö oli kirjepostilla yrityksen tarjoamaan postiosoitteeseen tai sähköisen lomakkeen kautta, joka oli mielekkäin tapa tehdä tietopyyntö. Osassa yrityksiä kuitenkin tietopyynnön postittaminen oli ainoa tapa toimittaa pyyntö.

Ausloos ja Dewitte (2018) suosittelevat, että yrityksissä hyödynnettäisiin yksinkertaisia pohjia tietopyyntöjen täyttämiseen. Sen avulla vähennettäisiin tietopyyntöjen muotoon ja sisältöön liittyviä epävarmuustekijöitä, joka havaittiin johtavan turhauttavaan viestintään rekisterinpitäjien kanssa. Rekisterinpitäjien tulisi myös hyödyntää enemmän teknologiaa tietopyyntöjen käsittelemisessä. Täysin automatisoidun ratkaisun kehittäminen vastatakseen rekisteröidyn pyyntöihin mahdollisesti vaatisi yrityksen koko järjestelmän uudistamista. Tutkimuksessa havaittiin, että suurimmalle osalle yrityksiä oli jopa haastavaa tunnistaa ja löytää rekisteröidyn tiedot. Kehittämällä järjestelmien tiedonhaku ominaisuutta voitaisiin ratkaista edellä mainittu ongelma. (Ausloos & Bewitte, 2018.)

2.5.4 Tietojen toimittaminen rekisteröidylle

Tietojen toimittaminen rekisteröidylle voidaan toteuttaa eri menetelmiä hyödyntäen. Kröger ja kollegat (2020) havaitsivat tutkimuksessaan, että tiedot tarjottiin rekisteröidylle sähköpostin, kirjepostin tai applikaation kautta. Useimmissa tapauksissa tiedot lähetettiin sähköpostin liitteenä vaihtelevissa tiedostomuodoissa, kuten PDF-, HTML-, JSON-, CSV-, JPG-, PNG-, DOXC-, tai TXT-tiedostona. Palvelun toimittajat olivat alkaneet kuitenkin kasvavin määrin mahdollistamaan tietojen lataamisen suoraan portaalin kautta salasanalla suojattuna arkistona. Tutkijat ehdottavat, että datan toimittamisen suhteen tulisi olla pakolliset standardin omaiset rajapinnat, jotta vähennettäisiin yleisen tietosuoja-asetukseen liittyvien pyyntöjen manuaalista käsittelemistä. (Kröger ym., 2020)

Myös Ausloos ja Dewitte (2018) havaitsivat tutkimuksessaan, että yrityksissä tiedot toimitettiin rekisteröidylle eri muodoissa, kuten sähköpostiviestin runko-osassa tai sähköpostin liitteenä PDF- tai XLSX-tiedostona. Osassa yrityksiä sähköpostin liitteenä olevat tiedot oli tuotu raakadatana yrityksen tietokannasta. Pienessä osassa yrityksiä tiedot oli mahdollista ladata URL-osoitteen kautta, joka vanheni tietyn ajan kuluessa. Ratkaisun koettiin ottavan hyvin huomioon tietoturvan tietojen toimittamisessa. Rekisteröidyn oli myös mahdollista ladata omat tiedot erillisen työkalun kautta, mutta tämä osoittautui monin tavoin puutteelliseksi. (Ausloos & Bewitte, 2018.)

Rekisterinpitäjän tulisi myös pohtia tietojen suojaamista, kun ne toimitetaan rekisteröidylle. Tietojen lähettäminen suojaamattomasti sähköpostilla voi olla jossain tilanteissa kätevää, mutta katsotaan usein riittämättömäksi tavaksi toimittaa tiedot rekisteröidylle. Parempi vaihtoehto – tosin vähemmän yleinen tapa – tietojen suojaamiseksi on lähettää ne salatussa arkistossa, jonka rekisteröity voi ladata itse määrittelemän salasanan avulla. Vaikka arkisto vuotaisikin vääriin käsiin, sen sisältö olisi ainakin jollakin tavalla suojattuna. (Singh & Cobbe, 2019.)

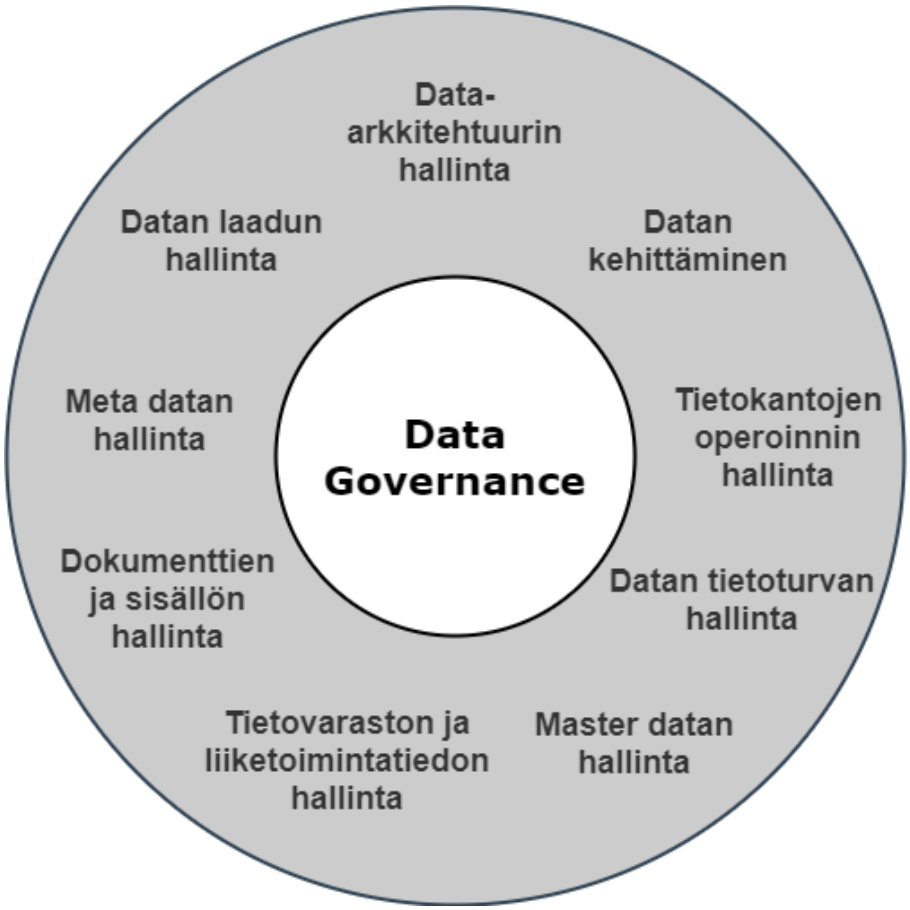
2.5.5 Tietojen siirtäminen järjestelmästä toiseen

Wong ja Henderson (2019) tutkivat tutkimuksessaan rekisteröidyn oikeutta siirtää tiedot järjestelmästä toiseen. Tutkimuksessa havaittiin, että toimitettu tiedostomuoto vaihteli hyvinkin paljon tutkittavien yritysten kesken ja osa niistä ei täyttänyt yleisen tietosuoja-asetuksen vaatimuksia. Asetuksen mukaan tiedot tulisi toimittaa rekisteröidylle yleisesti hyväksytyssä koneluettavassa muodossa, joka tutkittavien havaintojen perusteella vaatisi tarkempaa teknistä määrittelyä. Tutkimuksessa havaittu yleisin tiedostomuoto oli CSV, XLS ja XLSX. Tiedot tarjottiin rekisteröidylle myös muissa muodoissa, kuten esimerkiksi HTML-, PDF-, JSON-, XML-, TXT-, DOC-, DOCX-, RTF-muodossa sekä myös sähköpostin viestinä. Tutkijoiden mukaan PDF-, TXT-, DOC/DOCX- ja HTML-tiedostomuoto sekä sähköpostin viesti eivät täytä yleisen tietosuoja-asetuksen vaatimuksia tietojen siirtämisen osalta. (Wong & Henderson, 2019.)

3 DATANHALLINTA

Tässä luvussa perehdytään datanhallintaan ja sen eri osa-alueisiin. Datanhallinnan avulla voidaan parantaa yrityksen tehokkuutta ja päätöksentekoa, mutta myös täyttää Euroopan unionin yleisen tietosuoja-asetuksen vaatimukset. Tässä luvussa aluksi keskitytään data governanceen, jonka ympärille datanhallinnan muut toimet rakentuvat. Tämän jälkeen esitellään muut datanhallinnan kategoriat: datan laadun-, master datan-, datan tietoturvan- ja dokumenttien- sekä sisällönhallinta. Luvussa viitataan pitkälti ”The DAMA Guide to the data management body of knowledge” kirjaan, jossa esitetään laajasti datanhallinnan eri toimet. Teoriaa tuetaan aihealueeseen löydetyllä kirjallisuudella ja tieteellisillä artikkeleilla. Datan arkkitehtuurin-, datan kehittäminen, tietokanta toimintojen-, datavarastojen ja liiketoimintatiedon hyödyntämisen-, metadatan hallinta ovat sisällytetty osaksi datanhallintaa DAMA Internationalin teoksessa, mutta rajattu ulos tästä tutkimuksesta. Kuviossa 1 on esitetty datanhallinnan eri osa-alueet.

Käytän tutkielmassani osittain englanninkielistä ilmaisua datanhallinta, sillä datan suomennos on harhaanjohtava. Datan suomennos tieto voi tarkoittaa montaa asiaa, kuten jo työstettyä dataa eli informaatiota. Data on kuitenkin raaka-aine, josta muodostuu informaatiosta, kun siihen liitetään merkitys. Myös esimerkiksi tiedonhallinnalla voidaan viitata datanhallintaa, hankitun tiedonhallintaan (engl. knowledge management), tiedonhallinnan organisoinnin malliin (engl. data governance) jopa dokumenttien hallintaan. Sekaannuksien välttämiseksi käytän tiedosta englanninkielistä ilmaisua data. (Väre, 2019.) Data governance suomentuu tiedonhallinnan organisoinnin malliksi, mutta koin paremmaksi käyttää englanninkielistä termiä data governance, jota kuitenkin usein käytetään suomenkielisissä teksteissä.



KUVIO 1 Datanhallinnan osa-alueet (mukaillen Brackett ym., 2009)

3.1 Data governance

Data governance on vielä suhteellisen tuore suuntaus yritysmaailmassa, mutta sen avulla on mahdollista viedä datanhallinta seuraavalle tasolle. Data on nykypäivänä yksi yritysten tärkeimpiä omaisuuseriä ja strategisesti hallittuna se voi tuoda merkittävää kilpailuetua yritykselle. (Soares, 2010.) The Guardianin (2014) artikkelin mukaan jopa 80 % yrityksen datasta on suojaamatonta. Lisäksi vuonna 2012 noin kolmasosa yrityksen datasta tarvitsi suojauksen, mutta luvun arveltiin tuolloin kasvavan yli 40 % vuoteen 2020 mennessä johtuen henkilötietojen käsittelyn kasvusta (The Guardian, 2014).

3.1.1 Määritelmä ja tavoitteet

Oikein hallittuna datasta voi muodostua yrityksen arvokkain voimavara. Hyvän datanhallinnan avulla yrityksen on mahdollista pysyä kilpailukykyisen ja ketteränä, ennakoida asiakkaiden tarpeet sekä mahdollisuus pitää kulut hallinnassa. Yritykset, koosta riippumatta, ovat alkaneet nähdä datanhallinnan merkityksen liiketoiminnassaan ja hallita dataa nimenomaan yrityksen voimavara-

na jakamalla ja käyttämällä sitä organisaation laajuisesti eri järjestelmissä ja prosesseissa. Yritykset ovat havainneet, että datan käytölle, kehittämiselle ja hallinnalle on perustettava tarvittavat prosessit, politiikat ja standardit. On myös huomattu, että kriittistä datanhallinnassa on luoda oikeanlainen organisaatiorakenne ja kehittää teknologista infrastruktuuria, jotta voidaan tukea datanhallintaa riittävästi. (Panian, 2010.)

Datanhallinnan näkökulmasta data governance on datanhallinnan ydin. Sen tehtävänä on yrityksen tietovoimavarojen hallinta ja valvonta, kuten suunnittelu, monitorointi ja täytäntöönpano. Toisin sanoen se määrittelee ohjeet eri datanhallinnan toimintojen toteutukseen. (Brackett ym., 2009.) Data governance määritellään myös tarvittaviin prosesseihin, politiikoihin, standardeihin, organisointeihin sekä teknologioihin, joilla voidaan hallita ja turvata datan saatavuus, käytettävyys, laatu, muuttumattomuus, auditoitavuus ja tietoturva (Panian, 2010). Rifaie, Yorkshire, Ridley ja Yorkshire (2009) toteavat myös artikkelissaan data governancen prosessiksi, jonka tarkoituksena on edistää yritystä koskevien data investointien päätöksentekoa sekä hallita yrityksen tietovoimavaroja, jotta saavutettaisiin kilpailuetua markkinoilla. Hyvin käytettynä data governancen viitekehys tarjoaa ratkaisut seuraaviin kysymyksiin: Kuinka dataa koskevat päätökset on tehty? Kuka tekee kyseiset päätökset? Ketä pidetään vastuussa päätöksistä? Ja kuinka päätösten seurauksia mitataan ja valvotaan? (Rifaie ym., 2009.)

Väre (2019) lähestyy data governancea roolien ja vastuiden kautta. Se on malli, jonka pohjalta määrätään datan päätöksentekoa koskevat vastuut. Data governancen rooli on tehdä edellä mainitut vastuut selkeiksi ja tunnetuiksi koko organisaation laajuisesti. Sen tavoitteena on varmistaa laadukas päätöksenteko dataa koskevissa asioissa, jossa datan laatu, data prosessien toimivuus ja liiketoimintahyötyjen saavuttaminen varmistetaan hyvällä päätöksenteolla. (Väre, 2019.)

Data governance on osa laajempaa IT governancea, jolla taas on merkittävä rooli yrityksen hallinnossa. IT governance pohjautuu johtamis- ja organisaatiorakenteisiin sekä prosesseihin, joiden avulla pyritään varmistamaan IT:n rooli osana yrityksen strategiaa ja tavoitteita. (Rifaie ym., 2009.) Brackett ja Earley (2009) kuitenkin erottelevat teoksessaan data governancen irralliseksi osaksi IT governancea. IT governance keskittyy enemmänkin päätöksentekoon IT:n investoinneista, sovelluksista ja projekteista sekä linjaa IT:n strategian ja investoinnit yhteen yrityksen strategian ja tavoitteiden kanssa. Data governance taas keskittyy hallinnoimaan yrityksen tietovoimavaroja. (Brackett ym., 2009.)

Panin (2010) esittää data governancelle kolme keskeistä tavoitetta, jotka sen avulla tulisi saavuttaa. Ensimmäisenä tulisi varmistaa, että data täyttää liiketoiminnan tarpeet. Toisena dataa on suojeltava ja hallittava yrityksen voimavarana. Viimeisenä tavoitteena on alentaa datanhallinnan kustannuksia data governancen avulla. Myös Rifaie ja kollegat (2009) määrittelevät, että data governancen tarkoituksena on vaikuttaa IT:n pyrkimyksiin ja hallita dataintensiivisten toimintojen suuntaa, jotta voidaan saavuttaa sille asetetut tavoitteet. He jakavat data governancen tavoitteet neljään päätavoitteeseen, jotka organisaati-

on tulisi täyttää. Neljä päätavoitetta ovat datan arvo ja yhteenlinjaaminen, datan riskienhallinta, vastuullisuus ja toimintakyvyn mittaaminen.

Datan arvo ja yhteenlinjaaminen on yksi data governancen tärkeimmistä tavoitteista. Yhteenlinjaamisen myötä pyritään varmistamaan IT:n ja liiketoimintayksiköiden yhteistyö. Tarvittavien prosessien ja rakenteiden määrittäminen tietovoimavaroille auttaa johtoa varmistamaan, että yrityksen resurssit kohdistetaan strategian kannalta tärkeiden projektien edistämiseen. Datan arvon tuottamisella liiketoiminnalle viitataan yrityksen kykyyn pysyä lainsäädännön vaatimusten mukaisena, kasvattamalla yrityksen tuottoja, vähentämällä kustannuksia, sekä lisäämällä uusia tuotteita ja palveluja. (Rifaie ym., 2009.)

Datan riskienhallinta vastaa yrityksen liiketoiminnan riskeistä etenkin, jos yrityksen arvolupaus pohjautuu dataan liittyviin riskeihin. Dataan liittyvät riskit sisältävät tietoturvarikkomuksia, yksityisyyttä koskevia rikkomuksia, katastrofista palautumista, joustavuutta palvelukatkojen myötä sekä projektien epäonnistumisiin liittyviin riskeihin. Tietoturvarikkomuksia aiheutuvat hakkerien kohdistamista verkkohyökkäyksistä. Yksityisyyttä koskevat rikkomukset viittaavat esimerkiksi identiteettivarkauksiin. (Rifaie ym., 2009.)

Data governancen kolmas päätavoite on vastuullisuus. Data governance on vastuussa datanhallinnasta ja tietovoimavaroihin tehtyjen sijoituksien tuotoista sekä myös omien tietojen luotettavuudesta ja valvonnasta. Viimeinen tavoite liittyy toimintakyvyn mittaamiseen. Yrityksen on kyettävä mittaamaan data governancen toimintaa käyttämällä jotakin yleisesti suosittua mittaria, kuten esimerkiksi tasapainotettua tuloskorttia. Tasapainotettu tuloskortti sisältää kaksi data governancen kannalta keskeistä osa-aluetta: IT:n arvo ja riskienhallinta. IT:n arvo sisältää keskeisiä mittareita, jotka kohdistuvat IT:n ja liiketoiminnan yhteenlinjaamiseen ja IT:n arvon määrittelemiseen. Operatiivinen kyvykkyys taas sisältää keskeisiä mittareita IT:n riskien arvioimiseen. (Rifaie ym., 2009.)

Tietovastaavat (engl. data stewards) huolehtivat tietovaroista koko muun organisaation puolesta ja he ovat nimetty eri sidosryhmien datan omistajiksi. Tietovastaava vastaa yrityksen datan laadusta ja tehokkaasta käytöstä. (Brackett ym., 2009.) He ovat mukana siis kaikissa dataa ja sen käyttöön liittyvissä päätöksissä sekä datanhallinnan jokaisessa osa-alueessa.

3.1.2 Ajurit

Panian (2009) tunnistaa artikkelissaan yleisimmät ajurit, jotka ajavat yritykset hyödyntämään data governancea. Yksi yleisimmistä ajureista on kasvattaa yrityksen myyntiä ja parantaa olemassa olevien asiakkaiden säilyvyyttä ymmärtämällä asiakkaiden tarpeet syvällisemmin. Asiakastiedot voivat olla tallennettuina lukuisiin yrityksen eri järjestelmiin, joka tuo haasteita datan hallinnalle. Yrityksen ongelmia voivat olla esimerkiksi eri järjestelmien integraatiosta johtuva tiedonsiirron hitaus, joka aiheuttaa viivettä tiedon saatavuuteen. Tämän johdosta asiakastiedot eivät ole välttämättä aina ajan tasalla, joka taas voi aiheuttaa asiakaspalvelijalle ja asiakkaalle turhautumista. Data governance tarjo-

aa yritykselle tavan vastata näihin ongelmiin kehittämällä esimerkiksi asiakkaista yhden yhtenäisen näkymän ja parantamalla tiedon laatua. (Panian, 2009.)

Toinen tunnistettu ajuri on kustannusten alentaminen tehostamalla liiketoimintaprosesseja automatisoimalla ja poistamalla tarpeetonta dataa. Tähän liittyy vahvasti datanhallinnan osa-alue master data ja sen hallinta, jonka avulla voidaan yksinkertaistaa liiketoimintaprosesseja järjestelemällä, jakamalla ja siivoamalla yrityksen tärkeimmät ydintiedot. Data governance tarjoaa ohjeet, jotta yritys voi vastata master dataan liittyviin organisaatio- ja prosessikysymyksiin. (Panian, 2009.)

Yritykset siirtyvät data governancen pariin myös siitä syystä, että voidaan varmistaa ulkoisten sekä sisäisten lakien ja politiikoiden noudattaminen yksinkertaistamalla raportointitietojen keräämistä ja lisäämällä auditoitavuutta. Yrityksen raportointi voi olla toteutettu esimerkiksi taulukkolaskentaohjelman avulla, joka ei kuitenkaan täytä ulkoisten asetusten tai lakien vaatimuksia. Yrityksissä myös mahdollisesti pelätään käsitellä sensitiivistä tietoa, sillä väärinkäytettynä siitä voi koitua suuria sanktioita yritykselle. (Panian, 2009.)

3.1.3 Toiminnot

Data governancessa keskeisessä osassa on tehdä datastrategia, joka on koko datanhallintaohjelman suunnitelma datan ylläpitämiseksi ja parantamiseksi laadun, eheyden, turvallisuuden ja pääsyn osalta. Datastrategia voi myös sisältää liiketoiminnallisen näkökulman hyödyntää dataa kilpailukyvyn saavuttamiseen ja tukeakseen yrityksen tavoitteita. Se perustuu liiketoimintastrategiaan ja vaatii data tarpeiden ymmärtämisen liiketoimintastrategian pohjalta. Datastrategia voi sisältää esimerkiksi mission ja pitkänaikavälin datanhallinnan tavoitteet, datanhallinnan onnistumisen hallintatoimenpiteet sekä datanhallinnan roolit ja organisoinnit sisältäen vastuut ja päätöksenteko oikeuksien määrittämisen. (Brackett ym., 2009.)

Datastandardit ja ohjeistukset sisältävät muun muassa nimeämis-, vaatimusten määrittely-, datan mallinnus-, tietokantojen suunnittelu-, arkkitehtuurisekä proseduuri standardit jokaiselle datanhallinnan toiminnolle. Standardit vaihtelevat yritysten välillä, mutta niistä vastaa usein datanhallinnan asiantuntija ja ne arvioidaan ja hyväksytetään datanhallinta neuvoksella (engl. data governance council). (Brackett ym., 2009.)

Lainsäädännön vaatimustenmukaisuus sisältyy data governancen toimintoihin. Jokaista yritystä koskettaa hallituksen tai alan asettamat lait ja asetukset sekä useat näistä käsittelee dataa ja sen hallintaa. Useat yritykset ovat kiinnostuneet data governancen perustamisesta lakien vaatimusten täyttämisen kautta. Data governancen tehtävänä on ohjeistaa riittävien kontrollien käyttöönottaminen, jotta voidaan varmistaa, dokumentoida ja valvoa dataa koskevien lakien vaatimustenmukaisuus. (Brackett ym., 2009.)

Data governancen tehtäviin sisältyy myös ongelmien hallinta, jolla viitataan dataan liittyvien ongelmien tunnistamiseen, hallintaan ja ratkaisemiseen. Data ongelmat voivat liittyä esimerkiksi sen laatuun, nimeämiseen, tietoturvaan,

lain vaatimustenmukaisuuteen tai politiikoihin. Ongelmat voidaan myös ratkoa paikallisesti. Mikäli niitä ei saada ratkaistua, niin voidaan ne viedä aina datanhallinnan neuvokselle asti. (Brackett ym., 2009.)

Tietovoimavarojen arvioiminen on data governancen toiminto, joka on hyvä tehdä säännöllisin väliajoin lisätäkseen tietoisuutta organisaation halussa olevasta datasta. Organisaatiot voivat arvioida tietovoimavaroja muun muassa suorien ja epäsuorien hyötyjen kautta. Toinen tapa arvioida hallussa olevaa dataa on arvioida hinta datan häviämiselle ja tunnistaa vaikutukset sille, jos datan määrä tai laatu ei olisikaan sama verrattuna nykyiseen. Myös data yrityksen tietovarana voidaan arvioida kilpailijoiden kautta esimerkiksi pohtimalla, kuinka paljon kilpailijat maksaisivat datasta. (Brackett ym., 2009.)

3.2 Datan laadun hallinta

Tässä luvussa käydään läpi datan laadun hallinnan vaikutukset, ulottuvuudet, tehtävät ja sen mittaamista.

3.2.1 Vaikutukset

Heikkolaatuinen data saattaa vaikuttaa yritykseen niin strategisella, kuin myös operationaalisella tasolla. Datan laadun ongelmat ovat aiheuttaneet esimerkiksi Yhdysvalloissa vuosittain 611 miljardin dollarin posti-, paino- ja henkilökustannuksia. (Khatri & Brown, 2010) Gartnerin teettämän tutkimuksen (2018) mukaan tutkimukseen osallistuneet organisaatiot uskoivat, että pelkästään heikko datan laatu aiheuttaa 15 miljoonan dollarin kustannukset vuosittain. Heikosta datan laadusta saattaa aiheutua yritykselle myös riski täyttää eri asetusten ja lakien vaatimuksia, jos datan laatu ei täytä valvontaviranomaisten odotuksia. (Cichy & Rass, 2019.)

Datan laatua voidaan pitää synonyyminä informaation laadun kanssa, sillä heikko datan laatu vaikuttaa epätarkkaan informaatioon ja heikkoon liiketoiminnan toimintakykyyn (Brackett ym., 2009). Heikko datan laadun taso voi vaikuttaa myös yrityksen liiketoimintaan pitkällä aikavälillä esimerkiksi huonojen päätöksien ja menetettyjen liiketoimintamahdollisuuksien muodossa. Tämä johtuu usein siitä, että data ei anna tarpeeksi kattavaa kuvaa vallitsevista olosuhteista. (Cichy & Rass, 2019.) KPMG:n (2017) teettämän katselmuksen mukaan 45 % vastanneista kertoi huonon datan laadun vaikuttavan negatiivisesti asiakkaasta muodostettavaan kokonaiskuvaan. Vastanneita (56 %) huolestutti myös datan laadun eheys, joihin päätökset pohjautuivat. Katselmus pohjautui toimitusjohtajien kohtaamiin haasteisiin ja mahdollisuuksiin. Selvitykseen osallistui noin 1300 toimitusjohtajaa maailman suurimmista yrityksistä. (KPMG, 2017.)

3.2.2 Ulottuvuudet

Datan laatu voidaan määritellä sille tunnistettujen ulottuvuuksien kautta. Useassa datan laatua käsittelevässä tieteellisessä artikkelissa esiintyy kolme tyypillistä datan laadun ulottuvuutta, jotka ovat tarkkuus, ajankohtaisuus ja täydellisyys. Edellä mainittujen lisäksi on myös muita tunnistettuja datan laadun ulottuvuuksia, mutta ne voivat vaihdella julkaisusta toiseen. Datan ulottuvuudet ovat esitetty taulukossa 1. Khatri ja Brown (2010) toteavat osuvasti artikkelissaan, että ulottuvuudet ovat suhteellisia ja ne tulee määritellä vasta, kun tiedetään datan lopullinen käyttötarkoitus. Wangia ja Strongia (1996) voidaan pitää pioneereina datan laadun tutkimuksessa ja he ovat tutkimuksessaan löytäneet 118 datan laadun ominaisuutta. Heidän mukaansa parantaakseen datan laatua on ymmärrettävä käyttäjän tarpeita datan laadun suhteen. Datan laatu riippuu myös liiketoiminnan luonteesta ja vaatimuksista. Esimerkiksi 85 % tarkkuus nimen, numeron ja osoitteen kohdalla voi olla hyväksyttävää vakuutusyhtiölle, jotka esimerkiksi kohdistavat yhteydenotot potentiaalisille asiakkaille. Sama tarkkuusprosentti ei kuitenkaan ole hyväksyttävä esimerkiksi tilanteessa, jossa yrityksen on ilmoitettava lääkäreille lääkkeiden palauttamisesta. (Khatri & Brown, 2010) Myös datan laadun mittaamisessa usein käytetään apuna yrityksen ympäristössä tunnistettuja datan laadun ulottuvuuksia.

Brackett ym., (2009) määrittelevät datan laadun oikeellisuuden ulottuvuuden viittaavan tasoon, jossa data virheettömästi kuvaa oikean elämän entiteettejä. Khatri ja Brown (2010) mukaan oikeellisuus datan laadun ulottuvuutena kuvaa datan oikeellisuutta suhteessa todellisuuteen eli, että kirjattu arvo vastaa todellista arvoa huomioiden sen käyttötarkoituksen. Cichy ja Rass (2019) vielä lisäävät, että kun data täyttää oikeellisuuden vaatimukset, niin se on oikeaa, luotettavaa ja varmennettua.

Datan laadun ajanmukaisuuden määritelmissä on eroja eri julkaisuiden välillä. Perusajatuksena on kuitenkin, että kirjattu arvo on oikea-aikaisesti saatavilla ja ajan tasalla sitä käsiteltäessä (Khatri & Brown, 2010). Ajanmukaisuuden osa-alueeseen sisältyy myös, että data ei ole vanhentunutta, vaan soveltuu tehtävän suorittamiseen (Cichy & Rass, 2019).

Sisällön kattavuus datan laadun ulottuvuutena viittaa sen eheyteen. Datajoukosta löytyy tietyt nimetyt attribuutit ja rivit, jotka ovat aina olemassa kyseisessä datajoukossa (Brackett ym., 2009). Eli toisin sanoen kirjatut arvot eivät ole puutteellisia ja asetetut vaatimukset täyttyvät (Khatri & Brown, 2010). Cichy ja Rass (2019) lisäävät, että sisällön kattavuudella viitataan siihen, että data on riittävällä tasolla käsiteltävään tehtävään.

Edellä mainittujen lisäksi datan laadulle on tunnistettu muita ulottuvuuksia, kuten uskottavuus, jolla viitataan kirjatun arvon lähteen ja sisällön luotettavuuteen (Khatri & Brown, 2010). Datan laadun johdonmukaisuudella tarkoitetaan, että data on esitetty yhtenäisessä muodossa ja sopii yhteen yrityksen muun tiedon kanssa. Esteettömyydellä taas viitataan, että data on saatavilla, kun sitä tarvitaan tai, että se on nopeasti ja helposti haettavissa. (Cichy & Rass, 2019) Väre (2019) viittaa kirjassaan datan laadun vaatimuksen-

mukaisuuteen, jolla tarkoitetaan datan kykyä täyttää sille asetetut muodolliset ja sisällölliset vaatimukset. Muodollisten vaatimusten avulla data on teknisesti paremmin tulkittavampaa sekä mahdollistaa sen siirrettävyyden eri järjestelmien välillä. Esimerkiksi sähköpostin virheellinen muoto johtaa sähköpostin lähettämisen epäonnistumiseen, mutta myös automatisoidun prosessin keskeytymiseen. Sisällölliset vaatimukset puolestaan viittaavat siihen, että data täyttää sille kohdennetut liiketoiminnalliset vaatimukset. Esimerkiksi jossain virallisissa raporteissa voi olla vaatimuksena henkilön koko nimi, jolloin lempinimi tai pelkkä etunimi ei riitä. (Väre, 2019.)

Brackett ym. (2009) esittää edellä mainittujen lisäksi muita datan laadun ulottuvuuksia, kuten eheys, yksityisyys, kohtuullisuus sekä ainutlaatuisuus. Eheydellä tarkoitetaan, että eri järjestelmien keskinäiset viittaukset ovat keskenään oikein. Toisin sanoen yksi datan tietue on luotettavasti yhteydessä muihin järjestelmiin. Yksityisyyden ulottuvuus merkitsee, että johonkin tiettyyn datajoukkoon on pääsy vain valituilla henkilöillä, jolloin kulun- ja käytönvalvontaa tulee tarkkailla. Kohtuullisuudella viitataan tarkkaavaisuuteen johdonmukaisuusodotuksien kanssa tietyissä toimintaympäristöissä. Esimerkiksi voidaan olettaa, että transaktioiden määrä ei ylitä yhtenä päivänä 105 % edellisten 30 päivän keskimääräisestä tapahtumien määrästä. Ainutlaatuisuus datan laadun ulottuvuutena linjaa, että data entiteetti tulisi esiintyä vain kerran datajoukossa eli toisin sanoen datajoukossa ei saisi olla duplikaatteja. (Brackett ym., 2009.)

TAULUKKO 1 Kirjallisuudessa havaitut datan laadun ulottuvuudet

Datan laadun ulottuvuudet	**Kuvaus**	**Lähde**
Oikeellisuus (engl. accuracy)	Kuvaa datan oikeellisuutta eli missä määrin kirjattu arvo vastaa todellisuutta.	*Brackett ym., 2009; Khatri & Brown, 2010; Cichy & Rass, 2019*
Ajanmukaisuus (engl. timeliness)	Data on ajan tasalla ja saatavilla, kun sitä tarvitaan.	*Khatri & Brown, 2010, Cichy & Rass, 2019*
Sisällön kattavuus (engl. completeness)	Data ei ole puutteellista, vaan on riittävällä tasolla tehtävän suorittamiseen.	*Brackett ym., 2009; Khatri & Brown, 2010; Cichy & Rass, 2019*
Uskottavuus (engl. credibility)	Datan sisältö ja lähde on luotettavaa.	*Khatri & Brown, 2010*
Johdonmukaisuus (engl. consistency)	Data on muodoltaan yhtenäistä muun tiedon kanssa.	*Cichy & Rass, 2019*
Esteettömyys (engl. accessibility)	Data on esteettömästi saatavilla, kun sitä tarvitaan tai vähintään nopeasti ja helposti haettavissa.	*Cichy & Rass, 2019*
Vaatimustenmukaisuus (engl. validity)	Data on sekä muodollisesti, että sisällöllisesti vaatimustenmukaista.	*Brackett ym., 2009; Väre, 2019*
Eheys (engl. integrity)	Datan eri järjestelmien keskinäiset viittaukset ovat yhdistettävissä.	*Brackett ym., 2009; Väre, 2012*
Yksityisyys (engl. privacy)	Pääsyn- ja kulunvalvonta tiettyjen datajoukkojen osalta.	*Brackett ym., 2009*
Kohtuullisuus (engl. reasonableness)	Kohtuullisten odotuksien määrittäminen toimintaympäristön kontekstissa.	*Brackett ym., 2009*
Ainutlaatuisuus (engl. uniqueness)	Yksi data entiteetti tulisi esiintyä vain kerran datajoukossa.	*Brackett ym., 2009*

3.2.3 Hallinnan tehtävät

Lyhyesti sanottuna datan laadun hallinnan tehtäviin kuuluu suunnittelu, implementointi ja kontrollointi, joissa hyödynnetään laadunhallintatekniikoita mittaamalla, arvioimalla, parantamalla sekä varmistamalla datan sopivuus. Datan laadun hallinta ei siis sisällä pelkästään datan korjaamista, vaan siihen kuuluu myös datan elinkaaren hallinta datan luomisen, muuntamisen ja siirron osalta. Näissä datan elinkaaren vaiheissa pyritään varmistamaan, että datan laatu vastaa yrityksen datakuluttajien tarpeita. Datan laadun valvonnan, hallinnan ja parantamisen prosessien virallistaminen riippuu datan laadulle määritettyjen liiketoiminnan tarpeiden tunnistamisesta ja määrittelemällä parhaat tavat mitata, valvoa, kontrolloida sekä raportoida datan laadusta. Tietojenkäsittelyvirtojen ongelmien tunnistamisen jälkeen tietovastaava (engl. data steward)

tekee tarvittavat toimenpiteet korjatakseen dataan liittyvät ongelmat. (Brackett ym., 2009.)

Datan laadun hallinta on jatkuva prosessi, jossa pyritään jatkuvasti parantamaan datan laatua määrittelemällä hyväksyttyjä parametreja, jotta datan laadun taso vastaa liiketoiminnan vaatimuksia ja, että datan laatu täyttää kyseiset vaatimukset. Datan laadun hallinta sisältää datan laadun analysointia, data poikkeamien tunnistamista ja liiketoiminnan vaatimuksien määrittelemistä ja niiden liittämistä datan laadun vaatimuksiin. Lisäksi datan laadun hallinta sisältää prosessien luomista, kuten valvonta- ja tarkastusprosessien, jotta voidaan valvoa datan laadulle asetettujen sääntöjen noudattamista. Myös sääntöjen poikkeamille on luotava omat prosessit, jotta tarvittaessa voidaan jäsentää, standardoida, puhdistaa ja yhdistää dataa. Datan laadun hallinnalle kuuluu myös palvelutasosopimusten noudattamisen valvominen datan laadun osalta. (Brackett ym., 2009.)

Datan laadun hallintaa ja sen tehtäviä ohjaa ulkoiset ja sisäiset vaatimukset. Moni järjestelmä on riippuvainen datasta ja sen käytöstä tavoitteena täyttää liiketoiminnan prosessien vaatimukset. Liiketoiminnan prosessit määräytyvät liiketoiminnan politiikoiden pohjalta, joita säätelee niin ulkoiset kuin sisäiset vaatimukset. Ulkoisia vaatimuksia ovat esimerkiksi eri lainsäädäntöjen, alan standardien tai tiedonsiirtomuotojen noudattaminen. Sisäisiä vaatimuksia taas ovat yrityksen omat politiikat, jotka ohjaavat esimerkiksi markkinointia, myyntiä, palkkioita ja logistiikkaa. Datan laadun vaatimukset ovat usein piilotettuja ennalta määritettyihin liiketoimintaprosesseihin, joka hankaloittaa niiden tunnistamista. Liiketoimintaprosessien yksityiskohtainen arvioiminen ja asteittainen määrittely auttaa laatuvaatimusten tunnistamisessa. (Brackett ym., 2009.)

3.2.4 Mittaaminen

Datan laatua voidaan mitata joko subjektiivisesti kyselyiden tai objektiivisesti eri mittareiden avulla. Subjektiivinen mittaus perustuu dataa käyttävän työntekijän arvioon datan laadun ulottuvuuksista. Toinen tapa mitata datan laatua on määrittää mittarit, jotka voivat antaa arvioin datan laadusta. Viimeksi mainitussa on hyötynä, että se mittareiden avulla voidaan mitata datan laadun eri osia objektiivisesti. Toisaalta yksi mittari ei riitä mittaamaan kaikkia datan laadun ulottuvuuksia, vaan ideana on yhdistellä eri mittareita, jotta saadaan selkeä ja kokonaisvaltainen käsitys datan laadun tasosta. (Cichy & Rass, 2019.)

Mittarit usein mittaavat ennalta asetettujen raja-arvojen rikkomusten lukumäärää tai prosenttiosuutta. Laadullisesti tehdyissä mittauksissa mitataan esimerkiksi virheellisten päätösten lukumäärää pohjautuen heikkoon datan laatuun. Kaikkia datan laadun tasoja ei välttämättä voida mitata objektiivisesti mittareiden avulla, vaan näiden ulottuvuuksien mittaamiseen tarvitaan subjektiivisia arvioita datan laadusta. Datan laadun eri viitekehyksissä kuitenkin suositaan enemmän objektiivisia mittareita. (Cichy & Rass, 2019.)

Objektiivisia datan laadun mittareita ovat esimerkiksi prosenttiosuus virheellisistä arvoista, indikaattori datan päivityksestä, prosenttiosuus olematto-

mista käyttäjätileistä sekä viitteellistä eheyttä rikkovien tietueiden lukumäärä. Osa datan laadun viitekehyksistä suosii mittareiden yhdistelemistä. Subjektiivisilla kyselyillä saadaan datan käyttäjiltä näkemys datan laadusta, jolloin voidaan päättää mitä objektiivisia mittareita tarvitaan laadun mittaamiseen. (Cichy & Rass, 2019.)

3.3 Master datan hallinta

Tässä luvussa perehdytään master datan hallintaan sen määritelmän, tavoitteiden ja hyötyjen sekä arkkitehtuurin kautta.

3.3.1 Määritelmä

Nykyään datan käyttö nähdään jokaisen yrityksen toiminnan perustana niin operationaalisessa, taktisessa, kuin myös strategisessa päätöksenteossa. Tehokkuuteen pyrkivälle yritykselle on erityisen tärkeätä, että yrityksen tiedot ovat riittävällä tasolla eli toisin sanoen yrityksen menestyksen perustana voidaan pitää korkealaatuista tietoa. (Madnick, Wang & Xian, 2003.) Master data ja sen hallinnan rooli on kasvanut merkittävästi yrityksissä viime aikoina. Syynä tähän on yritysten halu luoda asiakkaasta kokonaisvaltainen kuva, valmius täyttää eri asetusten vaatimukset tai yhtenäistää ja yhdistää liiketoimintaprosesseja organisaation laajuisesti. (Otto, 2012.) Myös jatkuva paine lisätä organisaatioiden läpinäkyvyyttä ja vastuullisuutta sekä parantaa lainsäädännön noudattamista on vaikuttanut organisaatioiden kiinnostumiseen datanhallinnasta. Organisaatiot ovat ottaneet käyttöön master datan hallinnan, jotta voivat vastata edellä mainittuihin haasteisiin. (Sammon, Adam, Nagle, & Carlsson, 2010.)

Master datan – toiselta nimeltään ydintieto - määritellään olevan liiketoimintayksiköiden perusominaisuuksia. Master datan luonteelle on tyypillistä, että ne luodaan vain kerran, niitä käytetään usein ja pysyvät muuttumattomina ajan kuluessa. (Knolmayer & Röthlin, 2006) White, Newman, Logan ja Radcliffe (2006) lisäävät artikkelissaan ydintietojen olevan johdonmukaisia ja yhtenäisiä tunnisteita, jotka kuvaavat yrityksen ydinkokonaisuuksia. Väre (2019) liittää master datan määritelmään kaksi keskeistä tekijää. Ensimmäinen tekijä viittaa ydintiedon kriittisyyteen; ilman master dataa organisaation toiminta heikkenee. Toinen tekijä täsmentää, että master data on koko organisaation käytettävissä ja eri toiminnot hyödyntävät tätä samaa tietoa. Väre kuitenkin toteaa kirjassaan, että pääosassa on ensimmäinen tekijä, sillä asiakastiedot ovat usein master dataa, vaikkei niitä olisikaan jaettu koko organisaation kesken. Tiedon kriittisyys on suuressa roolissa, kun määritellään yrityksen master dataa. (Väre, 2019.) Yhteenvetona voidaan todeta, että ydintiedon määritelmä on moniulotteinen. Kiteytettynä master datan voidaan ajatella olevan yritykselle kriittistä, johdonmukaista ja yhtenäistä tietoa, jota käytetään eri yksiköiden kesken ja, joka pysyy suhteellisen muuttumattomana.

DAMA International teoksessa ”Data Management Body of Knowledge”, johon myös Väre (2019) viittaa kirjassaan, ryhmittelee master datan kolmeen ryhmään: osapuolet, asiat ja paikat. Osapuolet ovat yrityksen sidosryhmiä, jotka jollain tavalla osallistuvat yrityksen toimintaan. Osapuolet voivat olla yksittäisiä henkilöitä tai kokonaisia organisaatioita ja ne voidaan määrittää, joko asiakkaiksi tai toimittajaksi riippuen toiminnan luonteesta. Asioihin sisältyy organisaation ydintoiminnan kannalta merkittäviä asioita, kuten palvelut, tuotteet ja sopimukset. Sopimusten osalta on kuitenkin tärkeää tarkastella asiaa sen elinkaaren kautta. Jos sopimuksen elinkaari on lyhyt, luonnehditaan se ennemmin transaktio dataksi kuin ydintiedoksi. Paikat osio sisältää ydintietoa organisaation toimipaikoista, kiinteistöstä tai esimerkiksi geolokaatioista. (Väre, 2019.)

Edellisten lisäksi organisaatioissa on dataa, jota voidaan jaotella ajalliseksi sekä rakenteelliseksi master dataksi. Ajallinen master data sijoittuu varsinaisen master datan ja tapahtumatiedon välille, ja näin ollen on kriittistä tietoa yrityksen toiminnan kannalta, mutta sen elinkaari voi vaihdella hyvinkin paljon. Esimerkiksi sopimustiedot, hintalistat ja tuotekonfiguraatiot voidaan lukea ajalliseksi master dataksi. Rakenteelliseksi ja taloudelliseksi master dataksi lukeutuu tiedot, jotka liittyvät toimintoihin, kuten talouteen, raportointiin ja liiketoiminnan ohjaamiseen ja seurantaan. Esimerkiksi tiedot organisaatiokaavioista, hierarkioista, tilikartoista, tuotehierarkioista tai asiakassegmenteistä voivat olla rakenteellista ja taloudellista master dataa. On tärkeää tunnistaa organisaation kannalta merkittävät tiedot, jotka koskevat tämän tyyppistä tietoa ja liittää ne osaksi master datan hallintaa. (Väre, 2019.)

Master data voidaan jakaa eri luokkiin, joista useimmin käytettyjä ovat toimittajat, asiakkaat, materiaalit ja tuotteet. Master data ei kuitenkaan rajoitu näihin luokkiin, vaan siihen voi sisältyä myös tiedot organisaatiosta, kuten tiedot kustannuspaikoista, työntekijöistä tai jaettuja tietoja eri käyttäjistä. Se mikä määritellään master dataksi, riippuu pitkälti yrityksen toimialasta. Esimerkiksi sopimustiedot voidaan määritellä teollisuusalalla master dataksi, sillä ne pysyvät suhteellisen muuttumattomina ajan kuluessa. Teletoimialalla sopimustiedot nähdään ennemminkin transaktiotietona, sillä sopimuksia päivitetään ja puretaan jatkuvasti asiakkaiden toimesta. Vaikkakin master datan määritelmä voi tuntua hieman sekavalta, niin perustana on käsitys siitä, että tiedot koskevat yrityksen keskeisten liiketoimintatavoitteiden ominaisuuksia. Näille tiedoille on tyypillistä, että ne ovat yksiselitteisesti määritelty ja yksilöity organisaation laajuisesti. (Otto, 2012.) Master dataa myös käytetään organisaation eri liiketoiminnan prosesseissa eri toimijoiden toimesta. Samaa master dataa voidaan käyttää hankintaosastolla sekä ostoreskontrassa ja usein master dataa joko käytetään tai tallennetaan useaan eri yrityksen tietojärjestelmään. (Otto & Reichert, 2010.)

Master datalle on tiettyjä tyypillisiä piirteitä, jotka erottavat sen muusta tiedosta. Muuttumattomuuden lisäksi ydintiedolle on ominaista, että se ei ole sidoksissa aikaan. Esimerkiksi varastotietoihin, kuten varastossa säilytettäviin eriin, liitetään usein aikaleima, toisin kuin esimerkiksi master dataksi kuuluviin materiaalitietoihin. Materiaalitietoihin sisältyy materiaalin tunnusnumero, joka

säilyy koko sen elinkaaren ajan. Master datalle on myös tyypillistä, että sen volyymi pysyy suhteellisen vakiona toisin kuin transaktiodatalle (esimerkiksi tilaukset), jonka määrät vaihtelevat usein päivien ja kuukausien mukaan. Näiden lisäksi master dataa voi olla luotuna ilman riippuvuutta muuhun tietoon, kun taas transaktiodata on liitettävä aina johonkin tiettyyn ydintietoon. (Otto & Reichert, 2010.)

Smith ja McKeen (2008) määrittelevät master datan hallinnan (engl. master data management) järjestelmästä tai sovelluksesta riippumattomaksi prosessiksi, joka kuvaa, omistaa ja hallinnoi yrityksen liiketoiminnan kannalta tärkeitä tietoja eli toisin sanoen yrityksen master dataa. Master datan hallinta asettaa ohjeet, jonka avulla voidaan varmistaa tietojen johdonmukaisuus ja tarkkuus luoden samalla yhtenäisen kuvan yrityksen avaintiedoista. (Smith & McKeen, 2008) White ja kollegat (2006) kuvailee master datan hallinnan prosessiksi, jossa liiketoiminta- ja IT-yksiköt yhdessä pyrkivät yhtenäistämään, puhdistamaan, suojaamaan ja julkaisemaan yleisiä ja koko organisaation kesken jaettuja yrityksen voimavaroiksi kuuluvia tietoja. Master datan hallinnalla pyritään varmistamaan yrityksen master datan johdonmukaisuus, tarkkuus, hallinta ja vastuullisuus. (White ym., 2006.) Väre (2019) lisää kirjassaan, että master datan hallinnalla viitataan organisaation toimintatapoihin ja menetelmiin, joiden avulla turvataan master datan tarkoituksenmukaisuus. Kiteytettynä master datan hallintaa voidaan kuvailla järjestelmästä riippumattomaksi prosessiksi, jolla hallinnoidaan liiketoiminnan kannalta kriittisiä tietoja varmistaen niiden johdonmukaisuuden, tarkkuuden, hallinnan, tarkoituksenmukaisuuden sekä vastuullisuuden.

Otto (2012) esittelee artikkelissaan master datan hallinnalle kolme keskeistä asiaa, joiden kautta voidaan ymmärtää paremmin master datan hallinnan käsitettä. Ensimmäiseksi todetaan, että master datan hallinnassa ei ole kyse pelkästään sovelluksesta tai järjestelmästä, vaan se nähdään pikemminkin organisaation toimintona, johon sisältyy master datan omistajuus. Toiseksi artikkelissa todetaan, että master datan tehokkuuden mittarina käytetään tiedon laatua. Viimeiseksi selvennetään, että sen säilyttämiseen ja jakamiseen on erilaisia arkkitehtuurillisia lähestymistapoja. (Otto, 2012.)

3.3.2 Tavoitteet ja hyödyt

Master datan hallinnalla on Väreen (2019, s. 37) mukaan neljä päätavoitetta, jotka tulisi sovittaa organisaation liiketoiminnan tavoitteisiin ja tarpeisiin. Ensimmäinen tavoite on, että organisaation yhtä oikeita asioita sisältävää tietuetta vastaa yksi tosielämän kohde. Tietuella tarkoitetaan tietojärjestelmän itsenäistä loogista tietokokonaisuutta eli esimerkiksi relaatiotietokannassa yksi rivi kuvaa tietuetta. Toinen tavoite alleviivaa, että tiedot luodaan vain kerran ja käytetään useasti. Kolmas tavoite liittyy tiedon päivittämiseen. Tiedot on päivitettävä vain kerran, mutta on tärkeää varmistaa, että tiedot ovat päivitettynä kaikkialla. Viimeisellä tavoitteella viitataan tiedon saatavuuteen, ja tavoitteena on pyrkiä

siihen, että tiedot ovat aina saatavilla oikeassa paikassa oikeaan aikaan. (Väre, 2019.)

Master datan hallinnan tavoitteena on myös saavuttaa yrityksissä korkealaatuinen master data ja niin kutsuttu ”kultainen rekisteri”, johon on kuvattu ja määritelty koko organisaation kaikki tietokokonaisuudet. Master datan avulla on myös tarkoitus vähentää yrityksen kuluja ja helpottaa yrityksien turhaa monimutkaisuutta eri standardien kautta sekä tukea tiedon integroimista. Sen avulla voidaan myös tukea yritysten liiketoimintatavoitteita, kuten lisätä yrityksen tuloja, helpottaa mahdollisia yritysostoja ja sulautumisia, tehostamaa toimintaa, vähentää kustannuksia, parantaa riskienhallintaa, tukea liiketoimintatietoa (engl. business intelligence) ja helpottaa lainsäädännön noudattamista. Näiden saavuttamiseksi tulee kuitenkin ymmärtää master datan integroimisen tarpeet, tunnistaa sen lähteet ja myötävaikuttajat, määritellä ja ylläpitää arkkitehtuuria sekä muutosten hallintaa. (Otto, 2012; Otto & Reichert, 2010.)

Master datan hallinnan avulla on myös mahdollista poistaa aikaa vievä väittely tiedon paikkansapitävyydestä. Sen avulla voidaan pitää tiedot ajan tasalla ja yhdessä paikassa, joten oikea tieto on helposti löydettävissä. Tietojen huonolaatuisuus ajaa yritykset heikkoon päätöksentekoon. Master datan hallinnan onnistunut toteutus riippuu monesta tekijästä, kuten tiedon laadusta, hallinnasta, johtamisesta ja muutoksen hallinnasta. Se ei kuitenkaan perustu pelkästään tekniikkaan, vaikka osa sen toiminnoista onkin teknologian mahdollistamia. Yritysten tulisi aloittaa master datan hallinnan toteuttaminen varmistamalla koko organisaation sitoutuminen siihen. Liiketoimintayksiköt tulee sisältyä master datan hallinta -prosessin toteuttamiseen yhdessä IT-organisaation kanssa jakamalla vastuu koko projektista. (White ym., 2006.)

Master dataa käytetään organisaatiossa, jotta voitaisiin varmistaa prosessien, analyysien ja viestinnän johdonmukaisuus, yksinkertaisuus sekä yhdenmukaisuus. Kun master dataa käytetään laajasti koko organisaationtasolla, vähenee erillisten osittain ylläpidettyjen tietojen tai niistä muodostettujen versioiden tarve. Master data -ohjelman avulla yrityksien on mahdollista edistää yrityksen ketteryyttä sekä myös yksinkertaistamalla integraatioon liittyviä toimintoja. (White ym., 2006.)

Haug ja Arlbjørn (2011) toteavat tutkimuksessaan heikkolaatuisen tiedon olevan yksi master datan suurimmista ja kustannuksia aiheuttavista ongelmista. Korkealaatuisella tiedolla viitataan usein tiedon tarkkuuteen (Madnick, Wang, & Xian, 2003). Tästä huolimatta useat yritykset jättävät tiedon laadun kokonaan huomioimatta tai eivät ainakaan panosta siihen riittävällä tasolla. (Marsh, 2005.)

Master datan hallinnan avulla voidaan hakea yritykselle uutta liiketoimintaa. Hyvän hallinnan avulla on mahdollista tarjota nykyisille tai uusille asiakkaille henkilökohtaisesti kohdennettuja palveluita tai tarjouksia. Myös esimerkiksi teollisuuden yritykset ovat master datan hallinnan avulla tarjonneet asiakkaille parempaa huoltopalvelua ominaisuuksien kautta, jossa laite lähettää dataa toiminnastaan sen valmistaneelle yritykselle. Erinomaisen master datan hallinnan avulla yrityksen on mahdollista yhdistää kyseessä oleva tuote sitä vastaavaan asiakkuuteen. (Väre, 2019.)

Asiakasdataa hyödyntämällä organisaation on mahdollista parantaa asiakaspalvelua ja -tyytyväisyyttä. Tämän lisäksi yritys voi myös parantaa henkilöstön tyytyväisyyttä, joka johtaa hyvään asiakaspalveluun, mutta tätä kuitenkin harvoin tavoitellaan master datan hallinnan avulla. Yrityksen on mahdollista saavuttaa edellä mainitut hyödyt pitämällä asiakkaisiin, tuotteisiin ja organisaatioon liittyvä master datan ajan tasalla ja kaikkien saatavilla. Esimerkiksi asiakaspalvelun on helpompi palvella asiakasta, kun tunnistautumisen jälkeen asiakkaasta saadaan kokonaisvaltainen kuva master datan avulla. (Väre, 2019.)

Kun master dataa ei syötetä tai päivitetä useaan eri järjestelmään, on tarvittava data saatavilla sitä tarvittaessa. Se myös helpottaa ja nopeuttaa asioiden käsittelyä ja resursseja voidaan kohdentaa vaativampiin asioihin. Työvaiheita voidaan automatisoida ja sitä kautta tehostaa toimintaa. (Väre, 2019.) Tavoitteena on, että data on luotettavaa, ajan tasalla ja se on löydettävissä yhdestä paikasta.

3.3.3 Arkkitehtuuri

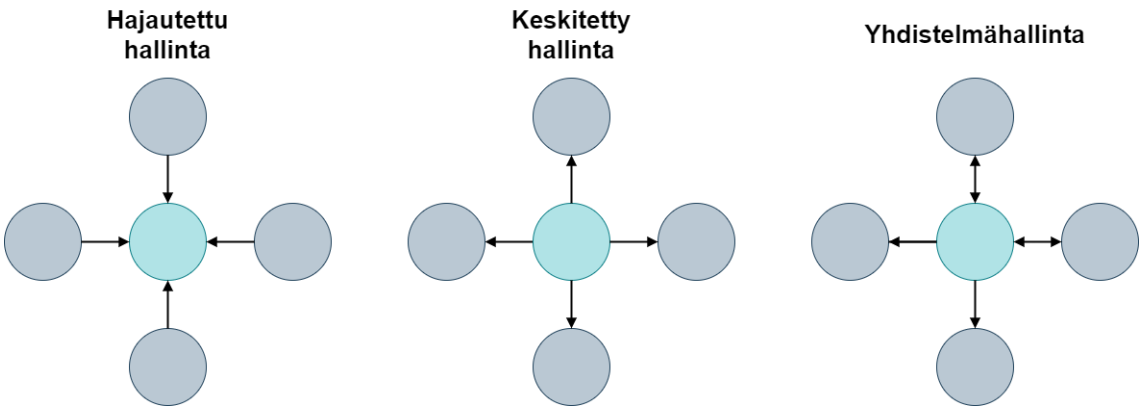
Master datan arkkitehtuurin suunnittelu ja ylläpito sisältyy master datan hallinnan tehtäviin (Otto, 2012). Arkkitehtuurin tehtävänä on varmistaa datan laatu kontrolloimalla pääsyä, replikointia, datan kulkua (Brackett ym., 2009). Jotta master dataa voidaan hyödyntää täysin, on data pystyttävä olla helposti hyödynnettävissä ja jaettavissa koko organisaation kesken. Tämän toteutus vaatii järjestelmien keskinäisen arkkitehtuurin suunnittelemista. Arkkitehtuurimalleja on useampia, jotka sopivat eri organisaatioihin ja niiden tarpeisiin.

Väre (2019) jakaa yrityksen järjestelmät operationaalisiin tietojärjestelmiin ja master data -järjestelmiin. Operationaalisissa järjestelmissä käsitellään master dataa päivittäisessä toiminnassa. Master data -järjestelmä voi olla joko master datan hallintaan tarkoitettu erillinen järjestelmä tai sovellus tai edellä mainittu operationaalinen järjestelmä, jossa käsitellään master dataa. Esimerkiksi operationaalisena master data -järjestelmänä voi toimia organisaation toiminnanohjaus- tai asiakkuudenhallintajärjestelmä.

Väre (2019) esittelee kirjassaan kolme erilaista master datan järjestelmäkarttaa, johon vaikuttaa organisaatioiden järjestelmien määrä ja monipuolisuus. Yksinkertaisessa järjestelmäkartassa yrityksen master data on yhdessä järjestelmässä, joka on integroitu korkeintaan muutamaan eri järjestelmään. Usean järjestelmän järjestelmäkartassa on domain kohtaisia järjestelmiä, jossa data luodaan, mutta sitä ylläpidetään muissa yrityksen järjestelmissä. Esimerkiksi asiakastiedot ovat tallennettuna yrityksen asiakkuudenhallintajärjestelmään, jonka lisäksi lisätietoja asiakkaista on myös toiminnanohjausjärjestelmissä. Data jaetaan vielä yrityksen muihin pienempiin järjestelmiin, kuten taloushallinnon- ja sisällönhallintajärjestelmiin. Monimutkaisessa järjestelmäkartassa on useita eri järjestelmiä ja samaa master dataa käsitellään monessa eri järjestelmissä. Usean eri master dataa käsittelevän järjestelmän lisäksi ne integroidaan moniin pienempiin järjestelmiin, jotka hyödyntävät master dataa. Haasteena voi olla

järjestelmien toimittajien ristiriitaisuus teknologioiden suhteen, joka tekee niiden integroinnista monimutkaista.

Master datan hallinnassa on tunnistettu myös kolme erilaista järjestelmäarkkitehtuuria: hajautettu, keskitetty ja yhdistelmä. Edellä mainitut järjestelmäarkkitehtuurit ovat esitetty kuviossa 2. Hajautetussa hallinnassa master data -järjestelmä toimii ainoastaan datan välittäjänä ja datan käsitteleminen tapahtuu hajautetusti eri operationaalisissa järjestelmissä. Master data -järjestelmän roolina on myös tarkistaa duplikaattien olemassaolo ja valvoa yleisesti datan laatua. Kaikki toimenpiteet datan suhteen, kuten duplikaattien poistaminen, tapahtuu operationaalisissa järjestelmissä, joka tekee mallista haastavan datan laadun hallinnan suhteen. Keskitetyssä hallinnassa kaikki datan käsittelyyn liittyvät toimenpiteet voidaan tehdä suoraan yhdestä järjestelmästä, josta data siirretään integraatioiden avulla muihin järjestelmiin. Muissa järjestelmissä ei kuitenkaan ole mahdollista poistaa tai muuttaa master dataa, mutta siihen voidaan lisätä muuta tietoa. Keskitetty malli on yksinkertaisin datan laadun hallinnan kannalta. Yhdistelmähallinnassa master dataa käsitellään master data -järjestelmissä sekä muissa operationaalissa järjestelmissä. Jotkut operationaaliset järjestelmät voivat hyödyntää pelkästään master dataa ja järjestelmien välillä voi olla sääntöjä liittyen master datan luontiin ja ylläpitämiseen. Tämä malli on teknisesti haastava toteuttaa, mutta voi hyvin toteutettuna mahdollistaa hyvän datan laadun hallinnan. (Väre, 2019.)



KUVIO 2 Master datan hallinnan arkkitehtuurimallit (mukaillen Väre, 2019)

3.4 Datan tietoturvan hallinta

Tässä alaluvussa käsitellään datan tietoturvan hallintaa. Aihetta käsitellään tietoturvan tarpeiden ja lainsäädännön vaatimuksien, -politiikkojen ja -standardien, käyttäjien- ja pääsynhallinnan, datan luokittelun sekä auditoinnin kautta.

3.4.1 Tietoturvan tarpeet ja lainsäädännön vaatimukset

Muuttuva ja globaali ympäristö on tuonut esiin uusia lakeja ja asetuksia, joita organisaatioiden on noudatettava. Useasti organisaatioiden kohtaamat eettiset ja oikeudelliset ongelmat johtavat uusien lakien ja standardien laatimiseen. Uudemmat asetukset, kuten Euroopan unionin yleinen tietosuoja-asetus ja Euroopan unionin Basel 2 huomioivat tietoturvan vaatimuksissaan tiukemmin. (Brackett ym., 2009.)

Tietoturvan implementointi yrityksessä alkaa tunnistamalla ja ymmärtämällä liiketoiminnan tarpeet. Liiketoiminnan tarpeet määrittelevät, kuinka korkealla tasolla tietoturva tulee huomioida. Tähän vaikuttaa myös yrityksen koko ja toimiala. Tietyt alat, kuten finanssiala tai terveydenhuolto, ovat tarkemmin säädeltyjä kuin muut alat riippumatta yrityksen koosta. Kun taas esimerkiksi vähittäiskaupan alalla suurten ja pienten yritysten tietoturvan tasossa saattaa olla suuriakin eroja. Liiketoiminnan säännöt ja prosessit määrittelevät tietoturvan kosketuspinnat ja työnkulun jokainen tapahtuma sisältää omat tietoturvaa koskevat vaatimukset. (Brackett ym., 2009.)

3.4.2 Tietoturvapolitiikat ja -standardit

Tietoturvapolitiikka määrittelee, mitä pitää tehdä organisaation tietovarojen suojelemiseksi. Hyvin rakennettu politiikka sisältää määritelmän ”mitä tehdä”, jotta myös ”miten” voidaan tunnistaa, mitata ja arvioida. Hyvin kirjoitettu ja rakennettu tietoturvapolitiikka voi parantaa organisaation yleisen tietoturvallisuuden tasoa. Usein tietoturvapolitiikat koostuvat yleisestä korkeamman tason politiikasta ja useista yksityiskohtaisemmista dokumenteista ja niiden tulisi yhdessä sisältää toimenpiteet, joilla organisaatiot aikovat noudattaa lakien ja asetusten vaatimuksia. Tietoturvapolitiikat ovat erittäin tärkeä myös muistaa päivittää ja usein päivitystarpeet johtuvat teknologioiden ja lakien muutoksista. (Tuyikeze & Pottas, 2010.)

Datan tietoturvapolitiikan määrittely perustuu sen tietoturvavaatimuksiin ja vaatii IT:n tietoturva järjestelmävalvojien, tietovastaavan, sisäisten ja ulkoisten auditoijien sekä lakiosaston yhteistyön. Tietoturvan ammattilaiset usein suhtautuvat tietoturvaan teknisestä perspektiivistä, joka saattaa tehdä politiikoista turhan vaikeaselkoisia datan käyttäjälle, jolla ei ole teknistä osaamista. Tietoturvapolitiikka on pyrittävä luomaan siten, että sen noudattaminen on helpompaa kuin noudattamatta jättäminen. Lopuksi datan tietoturvapolitiikka arvioidaan ja hyväksytään tietohallintoneuvoston toimesta (engl. data governance council). (Brackett ym., 2009.)

On yleistä, että yrityksissä on sekä IT:n, että datan tietoturvapolitiikat, mutta ne tulisi olla eroteltuna. Datan tietoturvapolitiikka keskittyy nimensä mukaisesti datan tietoturvaan, kuten esimerkiksi yksittäisten sovellusten, tietokanta roolien, käyttäjäryhmien ja salasanastandardien määrittelemiseen. Vastaavasti IT:n tietoturvapolitiikka voi liittyä hakemistorakenteiden ja identiteetinhallinta viitekehysten määrittelemiseen. (Brackett ym., 2009.)

Panhila, Siponen ja Mahmood (2007) selvittivät tutkimuksessaan, miksi yritysten työntekijät eivät noudata tietoturvaohjeita, ja mitkä ovat vaikuttavia tekijöitä tietoturvapolitiikoiden noudattamiselle. Tutkimuksessa saatiin selville, että informaation laadulla eli, kuinka selkeästi se on esitetty ja kuinka helposti tarvittava tieto löydetään, on merkittävä vaikutus tietoturvakäytäntöjen noudattamiseen. Tutkimuksessa havaittiin myös, että työntekijöiden aikomukseen noudattaa tietoturvapolitiikkaa vaikuttaa asenteet, normatiiviset uskomukset sekä tottumukset. Myös uhka-arviointien tekeminen ja olosuhteiden helpottaminen edistää politiikoiden noudattamista. Tämän takia on tärkeää informoida työntekijöitä organisaatiota koskevista tietoturvauhkista, niiden vakavuudesta ja mahdollisista vaikutuksista. Aiemmissa tutkimuksissa on havaittu, että sanktioilla on positiivinen vaikutus tietoturvapolitiikkojen noudattamiseen, joka kuitenkin kyseenalaistettiin kyseisessä tutkimuksessa. Myöskään palkkioilla ei havaittu olevan vaikutusta, joka kuitenkin voi johtua kunnollisen palkkiojärjestelmän olemassaolon puutteesta. (Panhila ym., 2007)

Microsoftin teettämän selvityksen mukaan yritykset näkevät lakien ja asetuksien vaikutuksen (28 %) vähemmän tehokkaana tapana parantaa heidän tietoturvaansa verrattuna vapaaehtoisiin tietoturvastandardeihin tai -viitekehyksiin (37%), kuten NIST tai ISO (Marsh & Microsoft, 2019). Standardisoinnin tarkoituksena on tuoda läpinäkyvyyttä sekä tukea parhaita käytäntöjä. Hyvien käytäntöjen avulla voidaan noudattaa tuotteille tai palveluille asetettuja vaatimuksia, jotka on täytettävä globaaleilla markkinoilla. Standardisoinnin etuna on myös kyky mitata ja arvioida, millä tasolla tuotteet tai palvelut täyttävät kansainvälisten standardien vaatimukset. Edellä mainitut standardien hyödyt pätevät myös tietojärjestelmien tietoturvaan, sillä standardien avulla edistetään yhteistä ymmärrystä tietoturvan vaatimuksista ja varmistetaan, että toteutetut tietoturvamekanismit noudattavat globaalisti hyväksyttyjä sääntöjä ja käytänteitä. Täten toteutetut tietojärjestelmät saavuttavat yleisesti hyväksytyn tietoturvatason ja toimivat yhdessä tehokkaasti sekä tietoturvallisesti. Tunnettuja ja yleisesti hyväksi todettuja tietoturvastandardeja ovat esimerkiksi International Organization for Standardization (ISO), Information Systems Audit and Control Association (ISACA), National Institute of Standards and Technology (NIST) ja Payment Card Industry Security Standards Council. (Tsohou, Kokolakis, Lambrinoudakis, & Gritzalis, 2010.)

Tietoturvastandardien tarkoituksena on varmistaa, että yrityksessä on otettu käyttöön tietyt tietoturvaprosessit ja -toiminnot. Ne eivät kuitenkaan huomioi, kuinka ne tulisi käytännössä toteuttaa. Tässä yhteydessä prosessilla viitataan tietoturvatoimintoihin tai -periaatteisiin, joilla voidaan turvata järjestelmät. Prosessi voi esimerkiksi koostua seuraavista periaatteista: ”suorita riskianalyysi” ja ”perusta tietoisuusohjelma”, joiden sisältöön ei oteta kantaa. Standardit siis pyrkivät huolehtimaan, että organisaatioissa noudatetaan tietoturvatoimia, mutta eivät ota kantaa siihen, kuinka hyvin ne toteutetaan. (Siponen, 2006.)

Yritykset voivat hakea tietoturvallisuuden toteutuksille hyväksyntää eli sertifikaattia, jolla voidaan osoittaa vaatimustenmukaisuus. Asiakkaat voivat

vaatia yritykseltä jonkun tietyn sertifikaatin olemassaoloa. Tosin standardit eivät vaadi sertifioitumista, vaan niitä voidaan noudattaa ilmankin. Sertifioitumalla voidaan kuitenkin osoittaa muille, että yrityksen tietoturvataso on standardin vaatimustenmukainen. Auditiointien avulla voidaan osoittaa ja todistaa, että standardin vaatimuksia noudatetaan. Organisaatio voi itse toteuttaa auditoinnin tai sen voi tehdä kolmas osapuoli, mikä on suositeltavaa luotettavuuden kannalta. (Traficom, 2019.)

Datastandardien ja lakien vaatimusten mukaiseen tietoturvan käyttöönottoon ei ole yhtä oikeaa tapaa. Lait ja asetukset yleisesti asettavat tietoturvalle lopputavoitteen, mutta eivät huomioi sen suuremmin merkitystä. Täten organisaatioiden on suunniteltava ja dokumentoitava itse omat tietoturvakontrollit, joiden kautta voidaan osoittaa lain ja asetusten vaatimustenmukaisuus. Tehokas datan tietoturvastrategia huomioi myös fyysisen turvallisuuden, kuten mobiililaitteet, kannettavien laitteiden datan tallennuksen ja laitteiden tietoturvallisen hävittämisen. (Brackett ym., 2009.)

3.4.3 Käyttäjien- ja pääsynhallinta

Yksittäisille käyttäjille voidaan myöntää ja päivittää pääsyoikeuksia, mutta se vaatii turhaa työtä. Käyttäjienhallinta tulisi toteuttaa roolien mukaan, jotta niiden hallinta olisi helpompaa ja loogisempaa. On myös vaikeampaa ymmärtää syytä sille, miksi yksittäiselle käyttäjälle on myönnetty jokin tietty käyttöoikeus ilman roolin mukaista käyttäjäoikeusryhmää. Aina, jos vain mahdollista, tulisi yrittää määrittää käyttäjälle vain yksi rooliryhmä. Datan tietoturvaan liittyy pääsyn estämisen lisäksi myös sen mahdollistaminen asianmukaisin keinoin yrityksen eri tietovaroihin. Varsinkin sensitiivisen datan käyttöoikeuksia tulisi hallita myöntämällä käyttöoikeuksia vain, kun niille on perusteltu tarve. Käyttöoikeuksia on suositeltavaa hallita ryhmätasolla etenkin suurissa yrityksissä, mutta esimerkiksi pienissä organisaatioissa voi olla hyväksyttävää yksilötasoinen hallinta. (Brackett ym., 2009.)

Roolipohjainen pääsynhallintamenetelmä (engl. role-based access control, RBAC) on yleinen tapa hallita pääsyä. Roolipohjaisessa pääsynvalvonnassa käyttöoikeudet ovat liitetty rooleihin ja käyttäjät saavat käyttöoikeudet eri roolien perusteella. RBAC tukee hierarkiaa eli ylemmän tason rooleille niin kutsutuille isäntärooleille (engl. parent roles) voidaan antaa perusoikeuksia, jotka periytyvät alemmille lapsirooleille (engl. child roles). (Ray & Toahchoodee, 2007.) Hierarkia kannattaa rakentaa liiketoimintayksiköiden pohjalta, jolloin lapsiroolit rajoittavat isäntäroolien käyttöoikeuksia (Brackett ym., 2009.)

Pääsynhallintamenetelmiä on useita muitakin edellä mainitun lisäksi, kuten attribuutteihin perustuva- (engl. attributes based access control, ABAC), pakotettu pääsynvalvonta (engl. mandatory access control, MAC) sekä ehdollinen pääsynvalvonta (engl. conditional access). ABAC perustuu roolien sijasta attribuutteihin eli ominaisuuksiin, jotka voivat kuulua käyttäjille, palveluille, operaatioille tai ympäristöille (Adda, Abdelaziz, McHeick, & Saad, 2015).

MAC pääsynhallintamenetelmä perustuu tietoturvaluokitusten käyttöön, eikä näin pääsyoikeuksia voi vapaasti antaa eri käyttäjille tai tiedostoihin. Kyseisessä mallissa tieto luokitellaan johonkin tietoturvaluokkaan ja vastaavasti käyttäjille annetaan pääsy eri tietoturvaluokkiin. Kun käyttäjällä on sama tietoturvaluokka, kun dokumentilla, hän saa pääsyn dokumenttiin. (Osborn, 1997.)

Ehdollinen pääsynhallinta (engl. conditional access) perustuu ehtoihin ja niiden täyttyessä se sallii käyttäjän pääsyn palveluun. Ehtoja, joihin ehdollinen pääsy perustuu, voi olla esimerkiksi sijainti, laite, palvelu ja tunnistautumismenetelmä tai jokin näiden yhdistelmä. Esimerkki ehdollisesta pääsynvalvonnasta on, että jotakin tiettyä palvelua voi käyttää ainoastaan yrityksen sisäverkosta. (Traficom, 2019.)

Pääsynhallinta heikkenee suuresti, jos jaettuja- tai palvelutilejä ei arvioida huolellisesti. Näillä tileillä on usein parannetut käyttöoikeudet järjestelmänvalvojien toiminnan parantamiseksi, joka saattaa aiheuttaa yritykselle riskin tietoturvaloukkauksille. Yritykset jopa jättävät valvomatta kyseisiä tilejä, joka vain suurentaa riskin määrää. Jaettuja- ja palvelutilejä ei tulisi käyttää säännöllisesti tai oletuksena. (Brackett ym., 2009.)

Käyttäjien oikeaksi todentaminen ja pääsynvalvonta on oleellinen osa yrityksen datan tietoturvaa, sillä sen avulla voidaan seurata käyttäjien pääsyä yrityksen tietovaroihin. Useimmat lait ja asetukset vaativat tämän tyyppistä valvontaa auditoinneissaan ja sen avulla voidaan muun muassa ennakoida ennalta arvaamattomia tilanteita. Valvonnan voi lisäksi toteuttaa sekä automaattisesti tai passiivisesti ja sen avulla voidaan havaita normaalista toiminnasta poikkeavia ja epänormaaleja tapahtumia, jotka voivat vaatia lisätutkimusta. Valvonnan suhteen on mietittävä mitä valvotaan, kuinka kauan ja mitkä tapahtumat aiheuttavat hälytyksen. Se kannattaa toteuttaa useammalla tasolla tai datan kosketuspinnalla. Se voidaan ottaa käyttöön myös eri osa-alueilla esimerkiksi järjestelmän, käyttäjän ja ryhmän, käyttöoikeuksien tai datan eheyden vahvistamisen osalta. (Brackett ym., 2009.)

3.4.4 Luottamuksellisen datan luokittelu

Datan luokitteleminen on prosessi, jossa määritetään datan ja sen sensitiivisyyden tasot. Datan luokitus määrää, millä tasolla data on suojattava ja kertoo niiden tärkeyden yrityksen liikevaroina. Dataa voidaan luokitella eri näkemyksien mukaan, kuten niiden paljastumiseen liittyvän riskin perusteella. Tällöin data voidaan luokitella esimerkiksi julkiseen, sisäiseen, luottamukselliseen, rajoitettuun, säädeltyyn tai erittäin salaiseen luokkaan. Vaihtoehtoisia tapoja luokitella dataa ovat esimerkiksi luontitavan, käyttäjän henkilökohtaisen datan tai käyttötapojen perusteella. Datan luokittelu vahventaa datan tietoturvaa ja suojaa dataa sen sisällön sekä saavutettavuuden perusteella. (Shaikh & Sasikumar, 2015.)

Yrityksen data luokitellaan sen sisällön luottamuksellisuuden mukaan. Datan luokittelussa käytetään tyypillisesti viittä eri tasoa perustuen siihen, kuinka luottamuksellisesta datasta on kyse. Edellä mainitut tasot ovat esitetty taulukossa 2. Luottamuksellisen datan erottaa etiketistä, joka tulisi merkitä do-

kumentin jokaisen sivun ylä- tai alaviitteeseen. Ensimmäistä tasoa voidaan pitää oletuksena kaikille dokumenteille, eikä niitä tarvitse erikseen merkitä etiketillä. Nämä dokumentit ovat tarkoitettu kaikkien saataville, eivätkä ne sisällä luottamuksellista tietoa. Seuraava taso sisältää dataa, jotka ovat rajattu vain sisäiseen käyttöön. Tämän tyyppistä dataa voidaan kuitenkin esittää organisaation ulkopuolelle, mutta ei kopioida. Kolmannen tason dataa ei saa jakaa missään muodossa organisaation ulkopuolelle. Neljäs taso on vieläkin rajoittuneempi ja sisältää dataa, jotka liittyvät jonkun tietyn roolin tai työtehtävän suorittamiseen. Tämän datan käsitteleminen voi vaatia turvallisuusselvityksen tekemisen. Viimeinen datan luottamuksellisuuden luokka on ääripää ja sisältää hyvin arkaluonteista dataa. Tähän pääsy vaatii oikeudellisen sopimuksen allekirjoittamisen ja salassapitovelvollisuuden noudattamista. (Brackett ym., 2009.)

Luokittelu ei saa rajoittua pelkästään dokumentteihin, vaan myös tietokannat, relaatiotaulut, sarakkeet ja näkymät tulisi olla luokiteltu datan turvallisuusluokkien mukaan. Tietovastaavat ovat vastuussa datan turvallisuusluokkien arvioimisesta ja määrittelemisestä. Dokumentin tekijä vastaa taas siitä, että turvaluokitus on kirjattu oikein. (Brackett ym., 2009.)

TAULUKKO 2 Datan turvallisuusluokat (Brackett ym., 2009)

Datan turvallisuusluokka	**Kuvaus**
Laajemmalle yleisölle	Oletustaso, joka kaikkien saatavilla.
Vain sisäiseen käyttöön	Data rajoitettu vain sisäiseen käyttöön, joka kuitenkin oikeuttaa siitä puhumisen tai sisällön näyttämisen ulkopuolisille, mutta ei kopioimista.
Luottamuksellinen	Data, jota ei saa jakaa organisaation ulkopuolelle.
Rajoitettu turvaluokitus	Data, joka on rajoitettu eri rooleihin liittyvien työtehtävien suorittamiseen. Voidaan vaatia turvallisuusselvityksen tekemistä.
Rekisteröity turvaluokitus	Arkaluonteista dataa, johon pääsy edellyttää oikeudellisen sopimuksen allekirjoittamisen ja salassapitosopimuksen noudattamisen.

3.4.5 Auditointi

Auditointi on keskeinen säännöllinen toimenpide, jonka avulla voidaan parantaa järjestelmien tietoturvaa. Tietoturvakontrollien asettaminen on aikaa vievää, eikä niiden olemassaoloa välttämättä huomata ennen kuin suoritetaan auditointi tai yritykseen tehdään tietomurto. Eri lokien säilyttäminen ja kerääminen on tärkeä yksittäinen tietoturvatoimenpide, jolla pystytään osoittamaan auditoijalle yksityiskohtaiset tapahtumat, kuten kuka teki mitä ja milloin sekä mahdolliset tietoturva rikkomukset. (Suduc, 2010.)

Suduc (2010, s. 45) määrittelee auditoinnin muodolliseksi tarkastukseksi, jossa tarkastetaan, noudatetaanko yrityksessä standardia tai ohjeita, ovatko dokumentit tarkkoja tai onko tehokkuus- ja vaikutustavoitteet saavutettu. Audi-

toinnin voi suorittaa sisäinen tai ulkoinen tarkastaja. Datan tietoturvan auditointi on toimenpide, jonka avulla analysoidaan, vahvistetaan, neuvotaan ja suositellaan tietoturvan hallintaan liittyviä käytäntöjä, standardeja ja toimintoja. Auditointi toteutetaan johtajien toimesta yhdessä työntekijöiden kanssa, jotka osaavat auttaa toteutukseen ja yksityiskohtiin liittyvissä kysymyksissä. Auditoijien on oltava riippumattomia tarkastukseen liittyvistä tiedoista tai prosesseista. Eheyden varmistamiseksi auditoijalla ei saa olla suoraa yhteyttä tarkastettavaan kohteeseen. Muuten auditoijat voivat olla joko yrityksen ulkopuolelta tai sisältä. (Brackett ym., 2009.)

Auditoinneissa on tavoitteena tarjota yritykselle objektiivisia arvioita ja käytännön suosituksia, eikä niinkään etsiä ja listata löytyneitä vikoja. Yleisimmin auditoinneissa tarkastetaan yrityksen datan tietoturva politiikat, standardien dokumentit, toteutusoppaat, muutospyynnöt, pääsynvalvontalokit, raportit ja muut tallenteet. Lisäksi voidaan tarkastaa esimerkiksi yrityksen lain tai asetusten vaatimustenmukaisuus, tietomurron ilmoittamiseen liittyvät toimenpiteet ja prosessit, datan jakamiseen liittyvien sopimusten arvioiminen tai vaatimustenmukaisuuden parantamisen suositteleminen. (Brackett ym., 2009.)

3.5 Dokumenttien- ja sisällönhallinta

Seuraavaksi käsitellään dokumenttien- ja sisällönhallintaa. Tässä alaluvussa ensiksi määritellään dokumenttien- ja sisällönhallinta, jonka jälkeen siirrytään datan elinkaaren kappaleeseen.

3.5.1 Määritelmä

Yrityksen tietovaroista suurin osa on strukturoimattomassa muodossa. Strukturoimaton data määritellään esimerkiksi dokumentiksi, tiedostoksi tai raportiksi, jota ei ole merkattu tai muuten jäsennelty riveiksi ja sarakkeiksi tai tietueiksi. Monien arviointien mukaan jopa 80 % kaikesta datasta on tallennettu tietokantojen ulkopuolelle. (Brackett ym., 2009)

Dokumenttien- ja sisällönhallinta keskittyy relaatiotietokantojen ulkopuolisen datan tallentamiseen, varastointiin, pääsyn- ja käytönhallintaan. Näiden tavoitteena on datan parempi eheys ja saatavuus. Dokumenttien- ja sisällönhallinta menee osittain päällekkäin muiden datanhallinnan toimintojen kanssa osoittaakseen tarpeen data governancelle, arkkitehtuurille, tietoturvalle, metadatan- ja datan laadun hallinnalle strukturoimattoman datan osalta. (Brackett ym., 2009.)

Dokumenttien- ja sisällönhallinta jakautuu loogisesti kahteen osaan: Dokumenttien hallintaan ja sisällönhallintaan. Dokumenttien hallinta sisältää sekä sähköisten, että paperisten dokumenttien varastoinnin, inventaarion ja valvonnan. Sisällönhallinnalla viitataan puolestaan prosesseihin, tekniikoihin ja teknologiaan, joiden avulla voidaan organisoida, luokitella ja jäsentää pääsyä datan

sisältöön tavoitteena tehostaa datan hakemista ja uudelleenkäyttöä. Sisällönhallintaan saatetaan myös viitata kirjallisuudessa yrityksen sisällönhallinnalla (engl. Enterprise content management, ECM), joka tarkoittaa sisällönhallintaa koko organisaation laajuisena toimintona. Yleisesti voidaan todeta, että dokumenttien hallinta keskittyy tiedostoihin ja kohtelee tiedostoa yhtenä entiteettinä, kun taas sisällönhallinta keskittyy enemmänkin tiedostojen sisältöön pyrkimyksenä tunnistaa ja käyttää sisällöstä löydettyjä käsitteitä. Dokumenttien hallintaan keskittynyt järjestelmä mahdollistaa sähköisten dokumenttien seuraamisen, varastoinnin, versionhallinnan, metadatan hallinnan, sisällön indeksoinnin ja hakemisen. (Brackett ym., 2009.)

3.5.2 Datan elinkaari

Niin kuin jokaisella muullakin yrityksen voimavaralla, myös datalla on oma elinkaarensa. Jotta yrityksen on mahdollista hallita dataa, on sen myös osattava hallita datan elinkaarta. Yleisellä tasolla datan elinkaareen kuuluu viisi vaihetta eli luonti tai hankinta, säilytys tai ylläpito ja lopulta sen poistaminen. Datan elinkaaren eri vaiheissa dataa voidaan purkaa, viedä, tuoda, siirtää, validoida, muokata, päivittää, puhdistaa, muuttaa, konvertoida, yhtenäistää, erotella, koota, viitata, tarkistaa, raportoida, analysoida, louhia, varmuuskopioida, arkistoida ja palauttaa ennen lopullista poistamista. (Brackett ym., 2009.)

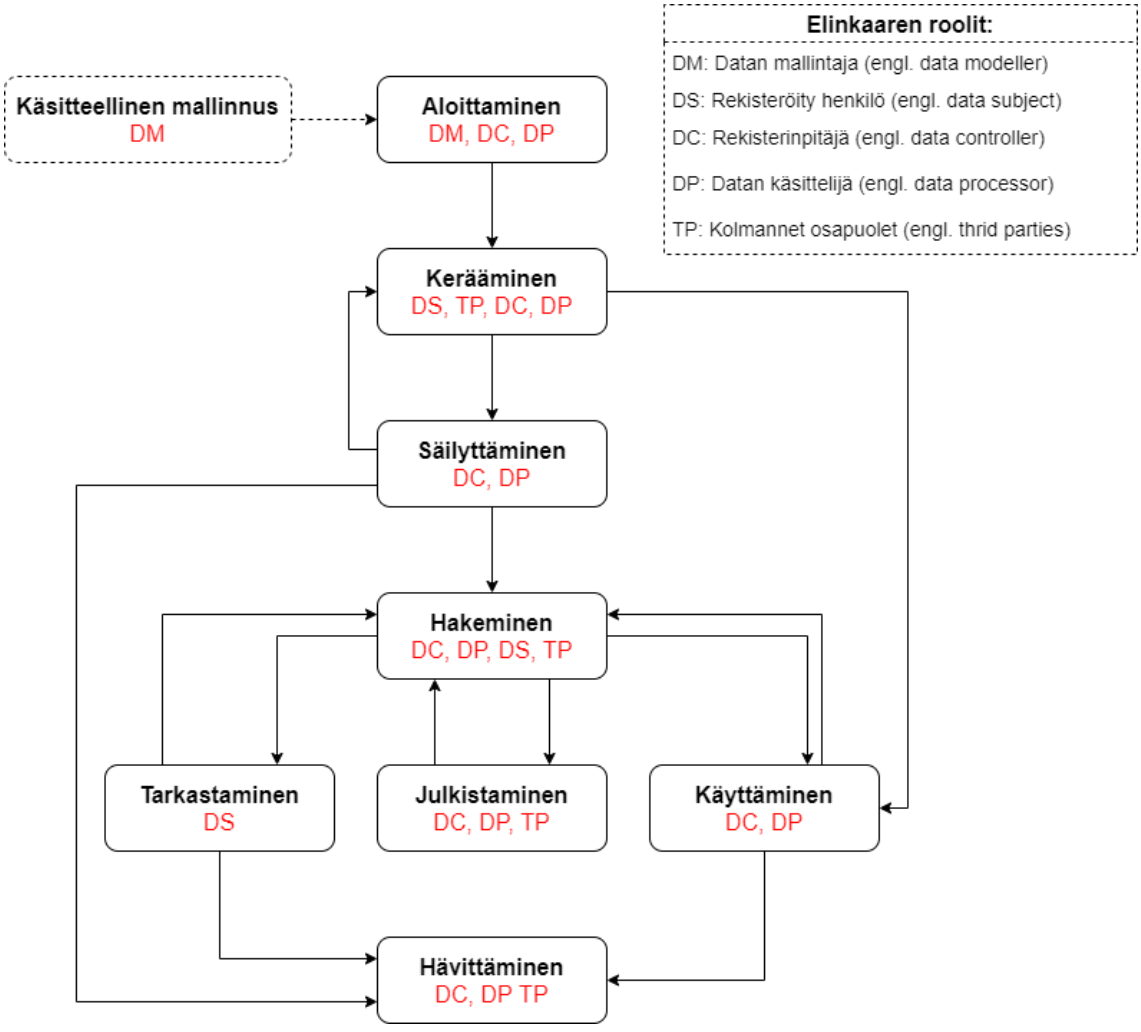
Yksityisyyteen, tietosuojaan ja identiteettivarkauksiin liittyvien tapausten johdosta huomio on kiinnittynyt entistä enemmän henkilötietojen käsittelyyn yrityksissä. Tästä syystä dokumenttien hallintaprosesseja ei tulisi säilyttää, eikä tiettyjä tietoja rekisteröidyistä henkilöistä siirtää kansainvälisten rajojen ulkopuolelle. Markkinoiden sekä sääntelyn, kuten Euroopan unionin yleisen tietosuoja-asetuksen, aiheuttama paine on johtanut yrityksiä keskittymään datan säilytysaikaan, sijaintiin, siirtämiseen ja hävittämiseen. Brackett ym. (2009) esittävät dokumenttien hallinnan elinkaari sisältää seuraavat aktiviteetit:

- Olemassa olevien ja uusien dokumenttien tunnistaminen
- Dokumentteja koskevien politiikoiden luominen, hyväksyminen ja käyttöönottaminen
- Dokumenttien luokittelu
- Varastointi: Fyysisten ja sähköisten dokumenttien lyhyen ja pitkän aikavälin varastointi.
- Hakeminen ja levittäminen: Dokumentteihin pääsyn ja niiden levittämisen salliminen politiikoiden, tietoturvan, standardien ja lakien sekä asetusten vaatimusten mukaisesti.
- Säilyttäminen ja hävittäminen: Dokumenttien arkistoiminen ja tuhoaminen organisaation tarpeiden, sääntöjen ja lakien/asetusten vaatimusten mukaisesti.

Dokumenttien säilytys- ja hävittämisohjelma määrittelee, kuinka kauan tiettyjä dokumentteja tai dataa on säilytettävä ja milloin ne eivät ole enää aktiivisia,

jolloin dokumentit voidaan siirtää toissijaiseen varastoon arkistoitavaksi. Ohjelma määrittelee prosessit vaatimusten noudattamiselle ja metodit sekä aikataulut dokumenttien hävittämiselle. Laki- ja asetusten vaatimukset on huomioitava, kun asetetaan dokumenttien säilyttämiselle aikataulua. Jotta organisaation datan varastointi kustannuksia saataisiin pienennettyä, on kaikki lisäarvoton tieto poistettava. Kuitenkin harvoin organisaatiot poistavat dataa, sillä varastointitila on halpaa ja sen ostaminen prosessina on helpompi, kuin arkistointi ja poistaminen. (Brackett ym., 2009.)

Alshammari ja Simpson (2018) esittelevät artikkelissaan henkilötietojen elinkaaren käsitemallin (engl. abstract personal data lifecycle model, APDL), joka kuvastaa henkilötietojen prosessointia ottaen huomioon sen vaiheet, toiminnot ja roolit. APDL-malli tunnistaa henkilötietojen elinkaaren vaiheet ja osoittaa niiden järjestyksen sekä syvyyden. Kuviossa 3 on kuvattu APDL-mallin vaiheet, toiminnot ja roolit.



KUVIO 3 Henkilötietojen elinkaaren käsitemalli (mukaillen Alshammar & Simpson, 2018)

Ensimmäistä vaihetta kutsutaan käsitteelliseksi mallinnukseksi, jossa kehitetään malli, joka kuvaa ongelman ja sen ratkaisun hyödyntäen alalle tyypillistä sanastoa. Tätä voidaan käyttää muun muassa helpottamaan kommunikointia useiden sidosryhmien kanssa. Seuraavaksi on aloitusvaihe, joka sisältää ”täydellisen käsittelysuunnitelman” määrittelyn, joka täsmentää henkilötietojen keräämisen ja prosessoinnin tarkoituksen suhteutettuna kontekstiin. Tämä on tietosuojaselosteen laatimisen perusta, joka välitetään rekisteröidyille. (Alshammari & Simpson, 2018.)

Keräysvaiheeseen henkilötietojen rekisteröinti, kerääminen ja tallentaminen riippumatta ovatko ne hankittu suoraan rekisteröidyltä vai ulkoisista lähteistä. Vaiheeseen liittyy neljä keskeistä arviointikriteeriä: henkilötietojen on oltava tarkoituksenmukaisia ja relevantteja suhteessa määriteltyyn tarkoitukseen; henkilötiedot ovat kerätty laillisin ja oikeudenmukaisin menetelmin; tietosuojailmoitus on toimitettu rekisteröidylle keräyshetkellä tai ennen sitä; ja rekisteröidyltä on saatu suostumus tietojen käsittelemiseen. Edellä mainittujen täyttyessä henkilötietojen kerääminen ja tallentaminen voidaan aloittaa. Jos vaatimukset eivät täyty, niin toiminnot on korjattava. (Alshammari & Simpson, 2018.)

Seuraava vaihe liittyy datan säilyttämiseen ja sisältää henkilötietojen järjestämisen, jäsentämisen ja tallentamisen tietyn ajanjakson ajan tietovarastossa. Hakemisenvaihe seuraa säilyttämisenvaihetta ja sisältää henkilötietojen määrittelemisen ja hakemisen. Tarkastusvaiheeseen sisältyy toimintoja, kuten käyttöoikeuksien toteuttaminen ja rekisteröityjen henkilötietojen tarkastaminen, jotta voidaan varmistaa tietojen tarkkuus, ajantasaisuus ja täydellisyys. Käyttövaiheessa käsitellään ja käytetään henkilötietoja ennalta määritetyn tarkoituksen mukaisesti. (Alshammari & Simpson, 2018.)

Julkistamisvaiheeseen liittyy toimintoja aiemmin käytettyjen ja haettujen henkilötietojen levittämiseksi, saataville asettamiseksi tai välittämiseksi kolmansien osapuolien käyttöön. Hävittäminen on mallin viimeinen vaihe ja siinä toiminnot keskittyvät henkilötietojen poistamiseen, tuhoamiseen, muokkaamiseen tai hävittämiseen olemassa olevien politiikoiden mukaisesti. (Alshammari & Simpson, 2018.)

Ymmärtämällä datan elinkaaren, yritysten on mahdollista säästää kokonaiskustannuksissa optimoimalla tallennustilaa datan käyttöön perustuvien mallien avulla. Useat yritykset eivät myöskään tiedä mitä dataa heillä on, kuinka kriittisiä ne ovat, niiden lähteitä tai, kuinka tarpeellisia ne ovat yritykselle. Datan elinkaaren hallinnassa on hyvä määrittää datan taksonomia, joka voidaan sisällyttää metadataan. Datan tallentaminen liiketoiminnan tarpeiden mukaan auttaa sen jakamista tehokkaammin useille resursseille. Tämä johtaa pienempiin datan varastointi kustannuksiin ja tallennustilan tehokkaampaan käyttöön. (Khatri & Brown, 2010.)

Datan elinkaaren usein määrittelee myös lain asettamat vaatimukset tietojen säilyttämisestä. Säilyttämisessä on syytä erottaa arkistoinnin ja varmuuskopioinnin ero. Kun data arkistoidaan, se poistetaan lähteestä ja korvataan metadatan osoittimella, joka sallii datan palauttamisen arkistosta. Varmuuskopio

tarkoittaa suuren data määrän tallentamista toiseen tietovarastoon, joka tarjoaa väliaikaisen suojauksen datalle. (Khatri & Brown, 2010.)

4 YLEISEN TIETOSUOJA-ASETUKSEN VAATIMUSTENMUKAISUUS DATANHALLINNAN AVULLA

Tässä osiossa tarkastellaan, kuinka datanhallinnan avulla voidaan täyttää Euroopan unionin yleisen tietosuoja-asetuksen määräämät rekisteröidyn oikeuksia koskevat vaatimukset. Luvussa pyritään tuomaan esiin hyvän datanhallinnan käytänteitä, jolla voidaan saavuttaa yleisen tietosuoja-asetuksen vaatimustenmukaisuus. Kuten Väre (2019) toteaa kirjassaan, Euroopan unionin yleisen tietosuoja-asetuksen edellyttämät vaatimukset ovat osa hyvää datanhallintaa.

4.1 Data governance keskiössä

Euroopan unionin yleisen tietosuoja-asetuksen vaatimusten täyttämisessä helpottaa tehokas ja hyvin dokumentoitu data governance -strategia, jossa selkeiden politiikkojen, proseduurien ja prosessien avulla hallitaan ja turvataan data sekä siihen lukeutuvat henkilötiedot. Yleisen tietosuoja-asetuksen vaatimukset ohjeistavat, kuinka henkilötietoja kerätään, käsitellään, käytetään ja varastoidaan. Microsoftin (2017) teettämässä julkaisussa vaatimukset voidaan jakaa neljään ryhmään, jotka myös muodostavat perustan tehokkaalle data governancen suunnitelmalle.

- Datan etsiminen: henkilötietojen tunnistaminen ja luokitteleminen
- Datanhallinta: sisältää rekisteröityjen pyyntöihin vastaamisen
- Datan suojaaminen: henkilötietojen kokonaisvaltainen suojaaminen
- Raportointi: henkilötietoihin liittyvien toimintojen ja ehtojen dokumentointi

Data governancen yksi tärkeimmistä toiminnoista on löytää data nopeasti ja hallita sitä tehokkaasti. Samalla se myös auttaa täyttämään asetuksessa määritettyjä rekisteröidyn oikeuksia, kuten datan poistamista, pääsyä omiin tietoihin ja oikeutta niiden oikaisemiseen. Data governancen tehtävänä on mahdollistaa

rekisteröidyn oikeuksien täyttyminen ja niiden suojeleminen yrityksessä. Sen tehtävänä on tarjota menetelmiä oikeuksien pyytämiselle, luoda siihen liittyvät prosessit, löytää ja tunnistaa henkilötietoja teknologian avustamana sekä hallita ja vastata niihin liittyviin pyyntöihin. Yksi tärkeä ja yhteinen tekijä data governancen suunnittelussa ja asetuksen vaatimusten täyttämisessä on tunnistaa organisaation hallussa olevat henkilötiedot, jotta niitä voidaan suojata riittävällä tasolla ja helpottaa vastaamaan rekisteröidyn pyyntöihin. (Microsoft, 2017.) Almeida, Teixeira, Mira da Silva ja Pereira (2019) toteavat artikkelissaan, että hyvän datanhallinnan avulla organisaatioiden on mahdollista vastata helpommin asiakkaiden pyyntöihin liittyen datan poistamiseen ja siihen pääsyyn, mutta myös osoittaa valvontaviranomaisille asetuksen vaatimustenmukaisuuden.

Datan luokittelupolitiikassa (engl. data classification policy) määritellään datalle tietoturvan- tai yksityisyydentaso ja kertoo siihen perustuen, kuinka dataa käsitellään. Kun pohditaan datalle asetettavaa suojausta, on ensimmäiseksi määritettävä datan luokitus. Datan luokituspolitiikat määrittelevät tasot, joissa dataa suojataan ja hallitaan sille asetetun tason mukaisesti. Henkilötiedot ja erityiset henkilötietoryhmät merkitään tai leimataan tunnistamisen ja löytämisen helpottamiseksi. (Microsoft, 2017.)

Datan luokittelu prosessina auttaa organisaatiota ymmärtämään, kuinka henkilötietoja käytetään tehokkaasti ja kuinka tietoturvatoimia sekä käyttöoikeuksia sovelletaan datan arkaluontoisuuden perusteella. Datan etsintä-, tunnistus- ja luokitteluprosessit hyödyntävät hakumenetelmiä datan paikantamiseen ja merkitsemiseen, jotta ne ovat löydettävissä helpommin ja niihin pystytään soveltamaan hallinta- ja suojaustoimintoja. (Microsoft, 2017.)

Jotta voidaan osoittaa, että rekisteröidyn käyttäjän pyyntöjä on hallinnoitu asianmukaisesti, on pyynnöt dokumentoitava. Dokumentoitaviin pyyntöihin tulee sisällyttää sen luonne, kuten tietojen tarkastus, muokkaus tai poistaminen sekä lopullinen päätös. Dokumentit on oltava sääntelyviranomaisten saatavilla. (Microsoft, 2017.) Organisaation datan auditointi on tärkeää datanhallinnan ja yleisen tietosuoja-asetuksen kontekstissa, jossa dokumenttien- ja sisällönhallinta on erittäin tärkeässä roolissa (Laybats & Davies, 2018.).

Teknisen toteutuksen näkökulmasta roolipohjaisen pääsynvalvonnan (RBAC) avulla voidaan säädellä eri rooleissa olevien käyttäjien mahdollisuutta suorittaa vain heille kuuluvia tehtäviä ja pääsyä dataan, joka on välttämätöntä työtehtävän suorittamisen kannalta. Yleisessä tietosuoja-asetuksessa määritellään rekisteröidyn oikeuksien lisäksi rekisterinpitäjille velvoite, jonka mukaan yrityksen tulisi toteuttaa tekniset ja organisatoriset toimenpiteet, jotta käsitellään vain tietyn tehtävän kannalta oleellisia henkilötietoja. Roolipohjaisen pääsynvalvonnan avulla yritys voi täyttää edellä mainitun vaatimuksen, mutta myös turvata ja hallita pääsyä henkilötietoihin. (Microsoft, 2017.)

4.2 Master datan- ja datan laadun hallinta yleisen tietosuoja-asetuksen vaatimusten täyttämisessä

Master datan hallintaan kuuluu tietovirtojen, tietomallien ja datan poistamisen prosessin kuvaaminen, jotka auttavat tietosuoja-asetuksen vaatimusten täyttämisessä. Yleisen tietosuoja-asetuksen kontekstissa – erityisesti omien tietojen poistamisen ja niihin pääsyn osalta – on kriittistä tietää, mitä dataa yrityksen eri järjestelmissä on ja mihin niitä käytetään. Edellä mainittujen toimenpiteiden avulla voidaan määritellä data, joka sisältyy tarkastusoikeuden ja poistettavien piiriin. (Väre, 2019.)

Hyvän datan laadun hallinnan avulla voidaan sekä täyttää osa yleisen tietosuoja-asetuksen vaatimuksista, mutta myös varmistaa datan oikeellisuus. Yrityksen vastuulla on varmistaa, että henkilötiedot ovat oikeita eli data on tarkkuudeltaan riittävän tasoista ja ajanmukaista (Väre, 2019). Parantamalla datan laatua ja ottamalla käyttöön dokumenttien hallinnan, yritykset saavat kattavan kuvan olemassa olevasta datasta. Edelliseen voidaan lisätä datan analytiikka, jolla on mahdollista maksimoida datan potentiaali ja arvo. (Garber, 2018.)

Strukturoimaton data tuottaa suuria ongelmia yleisen tietosuoja-asetuksen vaatimustenmukaisuudessa, sillä asetuksen mukaan henkilötietojen paikkansapitävyys on suojattava ja turvattava. Aina yritykset eivät edes tiedä strukturoimattoman datan sisältöä tai kenellä kaikilla on niihin pääsy. Strukturoimattoman datan vaikea hallinta tekee siitä alttiimpaa erilaisille verkkohyökkäyksille. Tästä voi koitua yritykselle suuria kustannuksia sanktioiden muodossa, perustuen yleisen tietosuoja-asetuksen vaatimusten rikkomiseen. Yritysten on tarjottava työntekijöille työkalut datan analysointiin, etsimiseen, säilyttämiseen ja jakamiseen, jotta datan käyttö pysyisi yrityksen hallinnassa. (Bisnode, 2021.)

5 EMPIIRISEN TUTKIMUKSEN TOTEUTUS

Luvussa käsitellään empiirisen tutkimuksen toteutusta. Luvussa ensimmäiseksi kuvataan tutkimuksen tavoite, kohde ja rajaus, jonka jälkeen esitetään tutkimusmenetelmät ja tutkittavien valinta sekä motivointi. Lopuksi esitellään haastatteluiden suunnittelu ja toteutus sekä kerätyn aineiston käsittely ja analyysi.

5.1 Tutkimuksen tavoite, kohde ja rajaus

Tutkimuksen tavoitteet voidaan jakaa kahteen kategoriaan: tutkijan henkilökohtaisiin- ja tieteellisen kontribuution tavoitteisiin. Vaikka tavoitteet ovat jaettu kategorioihin, ne eivät silti poissulje toisiaan. Tutkimuksen tavoitteena on selvittää, miten yrityksissä on toteutettu yleisen tietosuoja-asetuksen rekisteröidyn henkilön oikeuksia koskevat vaatimukset ja, kuinka datanhallintaa on hyödynnetty niiden täyttämisessä. Toisin sanoen haastattelemalla eri yritysten tietosuojavastaavia tai asetuksen parissa työskenteleviä henkilöitä pyritään selvittämään, millaisia ratkaisuja yritykset ovat tehneet asetuksen suhteen rekisteröidyn henkilön oikeuksien näkökulmasta. Datanhallinnan hyödyntäminen vaatimusten täyttämisessä viittaa kestäviin ratkaisuihin ja yrityksen haluun tehostaa omaa toimintaansa datan paremman hallinnan kautta. Ajatuksena kuitenkin on, että useassa yrityksessä vaatimukset ovat täytetty tiukan aikarajan painostamana, eikä kestäviä ratkaisuja niiden täyttämisen suhteen ole ehditty välttämättä miettiä. Tieteellisen kontribuution tavoitteena on pyrkiä muodostamaan kattava selvitys yleisen tietosuoja-asetuksen tämänhetkisestä tilanteesta ja ymmärtämään, miten yritykset ovat toteuttaneet rekisteröidyn henkilön oikeuksia koskevat vaatimukset.

Tutkielmassa vastataan kolmeen tutkimusongelmaan, josta kahteen ensimmäiseen vastattiin tutkielman teoriaosiossa ja viimeiseen etsitään vastausta empiirisen tutkimuksen kautta. Teoriaosiossa vastattiin seuraaviin tutkimusongelmiin:

- Euroopan unionin yleisen tietosuoja-asetuksen keskeisimmät vaatimukset rekisteröidyn henkilön näkökulmasta ja niiden toteuttaminen yrityksissä.
- Miten datanhallinta helpottaa täyttämään rekisteröidyn henkilön oikeuksia koskevat vaatimukset?

Edellisten pohjalta tarkoitus on vastata päätutkimusongelmaan, joka on seuraava:

- Miten yleisen tietosuoja-asetuksen rekisteröidyn oikeuksia koskevat vaatimukset ovat toteutettu yrityksissä ja, kuinka datanhallintaa on hyödynnetty niiden täyttämisessä?

Tutkimuksen kohteeksi valikoitui eri yritysten tietosuojavastaavat tai henkilöt, joilla oli kokemusta asetuksen parissa työskentelemisestä. Tutkimukseen valittiin yrityksiä eri toimialoilta, jotta saataisiin mahdollisimman kattava ja kokonaisvaltainen kuva rekisteröidyn oikeuksia koskevien vaatimusten toteuttamisesta. Lisäksi tutkimukseen valittiin yrityksiä, jossa käsitellään rekisteröidyn henkilötietoja. Kaikissa yrityksissä käsiteltiin laajasti rekisteröidyn henkilötietoja, paitsi yhdessä yrityksessä, jossa rekisteröidyn tietojen käsittely perustui rekrytointitilanteissa kerättyihin tietoihin.

Tutkimuksen strategia on kartoittava tutkimus, jolla pyritään selvittämään vähän tunnettua ilmiötä ja löytää uusia näkökulmia (Hirsjärvi, Remes & Sajavaara, 2004). Tutkimuksen tarkoituksena on selvittää, miten yrityksissä on toteutettu yleisen tietosuoja-asetuksen rekisteröidyn oikeuksia koskevat vaatimukset. Näin ollen tutkimuksen strategiaksi valikoitui kartoittava tutkimusstrategia.

5.2 Tutkimusmenetelmät

Tässä alaluvussa esitetään tutkimuksen empiirisessä osuudessa käytettyjä tutkimusmenetelmiä.

5.2.1 Laadullinen tutkimus

Tutkimus toteutetaan kvalitatiivisena tutkimuksena, jonka ominaispiirteisiin kuuluvat ei-numeeriset analyysit ja haastava mittaaminen. Kvalitatiivisessa tutkimuksen mittaamisen haasteisiin kuuluu sen tulkinnanvaraisuus, sillä tutkijan on aina käytettävä omaa näkemystä mittaustulosten tulkitsemiseen. Kvalitatiivisen tutkimuksen tulokset ovat useimmiten monimuotoisia, eivätkä yksiselitteisiä, kuten määrällisille tutkimuksille on ominaista. (Thornhill, Saunders & Lewis, 2009.) Hirsjärvi ym. (2004) vielä lisää, että kvalitatiivisessa tutkimuksessa selvitetään uutta ilmiötä ja tavoitteena on tutkimuskohde sekä sen ymmärtäminen. Kvalitatiivinen tutkimus valikoitui tutkimusmenetelmäksi, sillä tut-

kimuksessa pyritään kartoittamaan tutkittavaa aihetta ja löytää uusia ilmiöitä siitä.

Tutkimuksessa käytetään niin kutsuttua mono-metodia eli aineistonkeruussa sekä analyysissä käytetään samaa menettelytapaa – tässä tapauksessa siis laadullista menetelmää (Thornhill ym., 2009.) Aineistonkeruumenetelmänä käytetään laadulliselle tutkimukselle tyypillistä haastattelua, jolla pyritään tuottamaan syvempää tietoa haastateltavista yrityksistä. Haastattelujen haasteena on niiden tulkinnanvaraisuus ja, että ne vievät aikaa. Tämän takia tutkimukseen osallistuvia yrityksiä on rajoitettava määrällisesti, jotta tutkimus ei venyisi aikataulullisesti.

5.2.2 Teemahaastattelu

Tutkimuksen tiedonkeruumenetelmänä käytetään haastattelua, sillä tutkimuksen alue on vähän kartoitettu, eikä vastausten suuntaa pysty arvioimaan etukäteen (Hirsjärvi ym., 2004). Sarajärvi ja Tuomi (2009) korostavat, että haastattelun etuna on joustavuus. Haastateltavan kanssa on mahdollista käydä keskustelua, korjata väärinymmärryksiä sekä tarpeen tullen toistamaan kysymys. Joustavuutta lisää myös se, että kysymykset ovat mahdollista esittää tutkijan itse valitsemassa järjestyksessä. Tarkoituksena ei ole saada oikeita vastauksia kysymyksiin, vaan saada mahdollisimman paljon tietoa tutkimuksen aiheesta. Kysymykset tulisikin antaa haastateltavalle etukäteen, jotta tämä voi tutustua niihin ennakkoon. (Sarajärvi & Tuomi, 2009.)

Tarkemmin ottaen haastattelumenetelmäksi valikoitui teemahaastattelu eli puolistrukturoitu haastattelu. Teemahaastattelussa käsitellään ennalta valittuja teemoja ja niihin kohdistettuja tarkentavia kysymyksiä. Tästä huolimatta haastattelussa ei voi kysellä mistä tahansa aiheesta, vaan vastaukset on oltava merkityksellisiä tutkimuksen tarkoituksen ja -ongelman kannalta. Etukäteen valitut teemat perustuvat teoriaosiossa esitettyyn aineistoon. (Sarajärvi & Tuomi, 2009.)

Haastattelun luotettavuuteen voi mahdollisesti vaikuttaa haastateltavan pyrkimys antaa sosiaalisesti suotavia vastauksia haastattelutilanteessa ja näin ollen antaa itsestään ja yrityksestään todellisuutta paremman kuvan (Hirsjärvi ym., 2004). Etenkin, kun tutkimuksen aihe koskettaa lainsäädäntöä, on mahdollista, että haastateltavat pyrkivät antamaan suotavampia vastauksia haastatteluissa esitettyihin kysymyksiin.

Haastattelujen heikkouksena on, että ne vievät aikaa. Hirsjärvi ym. (2004) toteaa kirjassaan, että teemahaastattelut kestävät yleisesti tunnista kahteen tuntiin ja, että haastattelun suunnittelu ja rooliin paneutuminen itsessään vievät jo aikaa. Haastattelussa on varauduttava sekä vähemmän puheliaisiin, mutta myös puheliaisiin haastateltaviin (Hirsjärvi ym., 2004). Tutkimuksen haastattelun kestivät keskimäärin 30–90 minuuttia.

Ennen varsinaisia haastatteluja, on hyvä testata omaa haastattelumenetelmää ja -tilannetta testihenkilöillä. Samalla voi myös kokeilla teemojen toimivuutta ja tarvittaessa muokata niitä. (Hirsjärvi ym., 2004) Tutkimuksessa toteutettiin kaksi esihaastattelua ennen varsinaisten haastatteluiden toteuttamista.

5.3 Tutkittavien valinta ja motivointi

Haastattelun eduksi katsotaan, että tutkimukseen voidaan valita henkilöt, jotka omaavat kokemusta tai tietoa tutkittavasta aiheesta (Sarajärvi & Tuomi, 2009). Tutkimuksen kohderyhmäksi valikoitui eri toimialalla toimivien yritysten tietosuojavastaavat tai henkilöt, jotka ovat työskennelleet asetuksen parissa. Tutkimukseen osallistuneet yritykset pyrittiin valitsemaan eri toimialoilta, jotta saataisiin mahdollisimman laaja kuva siitä, miten rekisteröidyn oikeuksia koskevat vaatimukset ovat toteutettu eri toimialojen yrityksissä. Tutkittavilta myös varmistettiin ennen haastattelua, että he tietävät aiheesta ja ovat kykeneviä vastaamaan haastatteluissa esitettäviin kysymyksiin.

Tutkittavia motivoitiin raportilla keskeisimmistä tutkimustuloksista, jotka kaikki osallistuneet yritykset saavat tutkimuksen valmistuttua. Raportin avulla osallistuneet yritykset voivat havaita, miten rekisteröidyn henkilön oikeuksia koskevat vaatimukset ovat toteutettu eri yrityksissä. Yrityksen on mahdollista peilata raportin tuloksia omaan toimintaa ja pohtia eri ratkaisuita rekisteröidyn oikeuksia koskevien vaatimusten toteuttamiseen.

Tutkittavien määrä koettiin olevan tarpeeksi suuri, sillä tutkimuksessa ei pyritty yleistämään tutkimuksen tuloksia, vaan kuvaamaan ja kartoittamaan nykyistä tilannetta. Aineistoa havaittiin olevan tarpeeksi, sillä viimeiset haastattelut eivät enää tuoneet tutkimusongelman kannalta uutta tietoa, vaan enemmänkin toistivat aiemmin havaittuja tuloksia aiheesta. Tätä kutsutaan aineiston kyllääntymiseksi eli saturaatioksi, ja se on yksi tapa ratkaista aineiston riittävyys (Eskola & Suoranta, 1998).

5.4 Haastatteluiden suunnittelu ja toteutus

Tutkimukseen osallistuneet tutkittavat löydettiin LinkedIn-sivuston ja omaa verkostoa hyödyntäen. Tutkimuksen onnistumisen kannalta on tärkeää, että haastateltavat voivat tutustua aiheeseen ja kysymyksiin hyvissä ajoin ennen haastattelua. On myös eettisesti korrektia kertoa tutkittaville tutkimuksen tarkoituksesta ja haastattelun aiheesta ennen haastattelua. (Sarajärvi & Tuomi, 2009.) Keskustelu tutkittavien kanssa käytiin sähköpostin ja LinkedIn -sivuston välityksellä. Tutkittaville kerrottiin tutkimuksen aihe ja tarkoitus sekä haastatteluiden luonne. Haastattelun kysymykset lähetettiin sähköpostin liitteenä ennen haastattelua. Vallitsevan koronatilanteen vuoksi kaikki haastattelut käytiin etänä Teams -ohjelman välityksellä.

Haastatteluiden kesto sijoittui 30–90 minuutin välille ja haastattelija käytti apunaan Excel-ohjelmaa muistiinpanoja varten sekä esillä oli myös etukäteen suunnitellut haastattelukysymykset (liite 1). Kaikilta haastateltavilta kysyttiin lupa haastatteluiden nauhoittamisesta myöhempää litterointia ja analysointia varten. Haastatteluiden nauhoittaminen sopi kaikille tutkittaville ja hyödynnettiin Teams-ohjelman nauhoitus ominaisuutta. Haastattelut litteroitiin nauhoi-

tusten avulla. Haastatteluita ei litteroitu kokonaisuudessaan, vaan niistä pyrittiin poimimaan aiheen kannalta keskeisimmät asiat. Litteroitu aineisto vielä koottiin yhteen Excel-ohjelman avulla ja jaoteltiin teemoittain eri kategorioihin.

Haastattelutilanteessa ennen haastatteluiden aloittamista tutkittaville kerrottiin vielä tutkimuksen tarkoitus ja varmistettiin, että he ovat työskennelleet omassa yrityksessä yleisen tietosuoja-asetuksen parissa ja, että yrityksen toimintaan liittyy henkilötietojen käsitteleminen. Tutkittavilta kerättiin myös taustatietoina työnkuva ja -nimike.

Tutkimuksen teemat määräytyivät tutkimuksen teoriaosion pohjalta. Kaikki edellä mainitut teemat esiintyivät jokaisessa haastattelussa suhteellisen saman laajuisena. Osassa haastatteluja esitettiin tarkentavia lisäkysymyksiä liittyen johonkin tiettyyn kysymykseen ja teemaan. Esihaastatteluiden pohjalta teemoja muokattiin vielä sopivammaksi tutkimuksen tarkoitukseen nähden.

Haastatteluissa teemat voitiin jakaa kolmeen eri osioon: tutkittavan taustatiedot, rekisteröidyn oikeudet sekä datanhallinta. Taustatietojen kohdalla tutkittavilta selvitettiin työnimike, yrityksen toimiala, kokemus tietosuojasta ja tietosuoja-asetuksesta sekä mitä tietoja yrityksessä kerätään rekisteröidystä henkilöstä. Toinen teema keskittyi rekisteröidyn oikeuksiin ja niiden toteuttamiseen. Yrityksiltä selvitettiin rekisteröidyn informointiin, tunnistamiseen, tietojen toimittamiseen, poistamiseen ja siirtämiseen sekä pyyntöjen käsittelemiseen liittyviä asioita. Datanhallinnassa yrityksiltä tiedusteltiin data governancen, master datan, datan laadun, tietoturvastandardien, pääsynhallinnan ja strukturoimattoman datan hallinnan hyödyntämistä asetuksen vaatimustenmukaisuudessa. Haastattelun lopussa tutkittavilta kysyttiin mielipidettä datanhallinnan tavoista, joilla tulevaisuudessa voitaisiin vastata tietosuoja-asetusten vaatimuksiin. (Liite 1, Teemahaastattelun runko).

5.5 Haastatteluaineiston käsittely ja analyysi

Haastatteluiden aikana tehtiin muistiinpanoja Excel -ohjelmaa apuna käyttäen. Excel -ohjelmaan oli tehty valmis runko muistiinpanoille, joihin oli helppo lisätä oleellisimmat muistiinpanot haastatteluiden aikana. Muistiinpanot täydennettiin heti haastatteluiden jälkeen hyödyntäen Teams-ohjelman nauhoitusta. Haastatteluiden litterointi suoritettiin huhti-toukokuussa 2021.

Litteroitu aineisto yhdistettiin erilliseen Excel-tiedostoon, johon koottiin kaikkien haastateltavien vastaukset teemoittain. Tiedoston pohjalta oli helppo havaita eroavaisuuksia ja yhteneväisyyksiä eri teemojen väliltä. Yksittäisten kysymysten kohdalla vielä koottiin yhteen tutkittavien vastaukset ja luokiteltiin ne avainsanojen mukaan.

6 EMPIIRISEN TUTKIMUKSEN TULOKSET

Tämä luku keskittyy tutkimuksen tulosten käsittelyyn. Luvussa ensimmäiseksi esitellään tutkittavien taustatiedot, kuten työnimike, yrityksen toimiala sekä kokemus tietosuoja-asetuksen parissa työskentelemisestä. Tämän jälkeen käsitellään haastatteluista kerättyjä havaintoja liittyen Euroopan unionin yleisen tietosuoja-asetuksen rekisteröidyn oikeuksien toteuttamiseen. Lopuksi vielä esitetään tutkimuksessa löydettyjä havaintoja siitä, kuinka rekisteröidyn oikeuksien toteuttamisessa on hyödynnetty datanhallintaa.

6.1 Tutkittavien taustatiedot

Tutkimukseen valittiin yrityksiä eri toimialoilta tavoitteena saada mahdollisimman laaja käsitys, miten eri toimialoilla on toteutettu rekisteröidyn oikeuksia koskevat vaatimukset. Taulukkoon 3 on koottu tutkimukseen osallistuneiden yritysten toimialat ja tutkittavien lukumäärä.

TAULUKKO 3 Tutkimukseen osallistuneiden yritysten toimialat

Toimiala	Tutkittavien määrä
Julkishallinto	1
Teollinen valmistus	1
Tietoliikenne- ja digitaaliset palvelut	1
Tietoliikenne	1
Finanssiala	1
Sosiaali- ja terveydenhuolto	1
Energiateollisuus	1
IT konsultointi ja -palvelut	1
Yhteensä	8

Haastateltaviksi valittiin ensisijaisesti yritysten tietosuojavastaavia, mutta jokaisen haastateltavan kohdalla vielä varmistettiin, että hän on oikea henkilö vastaamaan tutkimuksessa esitettyihin kysymyksiin. Erityisesti haluttiin varmistaa,

että tutkittava tietää, kuinka yrityksessä on toteutettu rekisteröidyn oikeuksia koskevat vaatimukset sekä datanhallinta. Yhden yrityksen kohdalla osoittautui relevantiksi haastatella kehitysinsinööriä, joka oli ollut mukana tietosuoja-asetuksen rekisteröidyn henkilön oikeuksia koskevien vaatimusten toteuttamisessa. Kyseinen haastateltava oli toiminut yrityksessä myös privacy championina. Tutkimuksen kannalta nähtiin hyödylliseksi haastatella tietosuoja-asetuksen parissa työskentelevää konsulttia, jolla oli laaja näkemys, kuinka eri yrityksissä on toteutettu rekisteröidyn oikeuksia koskevat vaatimukset. Taulukkoon 4 on koottu tutkittavien henkilöiden määrä ja työnimikkeet.

TAULUKKO 4 Haastateltavien työnimikkeet

Haastateltavan työnimike	**Tutkittavien määrä**
Tietosuojavastaava	6
Konsultti	1
Kehitysinsinööri (Privacy champion)	1
Yhteensä	8

Kaikissa (8) tutkimukseen osallistuneista yrityksissä käsiteltiin laajasti eri henkilötietoja rekisteröidyistä henkilöstä. Yksi osallistuneista yrityksistä ei keskittynyt kuluttaja liiketoimintaan. Kyseisessä yrityksessä käsiteltiin henkilötietoja omasta henkilöstöstä ja rekrytoinnin yhteydessä rekisteröityneistä henkilöistä. Kahdessa yrityksessä käsiteltiin myös erityisiä henkilötietoryhmiä, joihon sisältyy terveyttä koskevat tiedot, kuten potilastiedot.

Kaikki kahdeksan haastateltavaa olivat työskennellyt Euroopan unionin yleisen tietosuoja-asetuksen parissa ja suurimmalla osalla (7) oli monen vuoden (6–16) kokemus tietosuoja-asioiden kanssa työskentelemisestä. Yhteenvetona voitiin todeta, että yleisesti tutkittavilla oli merkittävä kokemus tietosuojasta ja tietosuoja-asetuksesta.

6.2 Rekisteröidyn oikeuksien toteutus yrityksissä

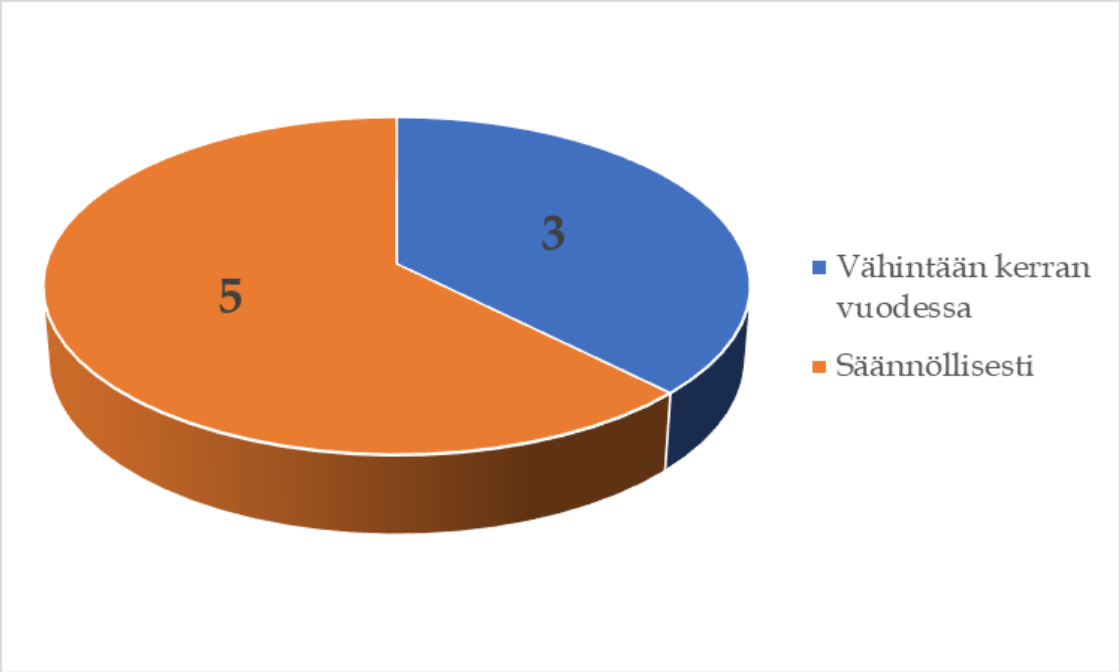
Jokaisessa kahdeksassa yrityksessä oli huomioitu ja täytetty rekisteröidyn henkilön oikeuksia koskevat vaatimukset. Tämä tosin oli olettamus, sillä asetus on ollut voimassa jo useamman vuoden. Rekisteröidyn oikeudet oli kuitenkin toteutettu eri tavalla tutkimukseen osallistuneissa yrityksissä. Tähän muun muassa vaikutti yrityksen toimiala, sillä esimerkiksi finanssiala sekä sosiaali- ja terveydenhuolto ovat toimialoina tiukasti säädeltyä eri lakien toimesta, mikä myös asettaa tiedon käsittelylle erityispiirteitä.

Kaikissa kahdeksassa yrityksessä kattavampi kuvaus tietojen käsittelystä oli sisällytetty tietosuojaselosteeseen tai vastaavaan. Jokaisessa kahdeksassa yrityksessä tietosuojaselosteet olivat sisällytetty yrityksen verkkosivuille erilliseen tietosuoja osioon. Kahden yrityksen kohdalla tuli ilmi, että olemassa oli myös paperiversio tietosuojaselosteesta. Kahdessa yrityksistä tietosuojaseloste

tarjottiin PDF-dokumenttina yrityksen sivuilta, kun taas neljässä yrityksistä se tarjottiin internetsivuna. Kahden yrityksen kohdalla tietosuojaseloste oli sekä PDF-dokumenttina, että internetsivuna. Verkkopohjaisissa tietosuojaselosteissa informaatio oli jaettu eri otsikoiden alle helpottamaan tiedon löytämistä. Osassa (3) yrityksistä oli hyödynnetty informointiin videota (2) tai kuvaa (1) havainnollistamaan, kuinka tietoja käsitellään yrityksessä.

Tutkimuksessa havaittiin, että yrityksissä (8) viittaus tietosuojaselosteeseen tai vastaavaan tarjottiin rekisteröidylle asiakaskohtaamisen, kuten rekisteröinnin, sopimuksen solmimisen, suostumuksen pyytämisen tai asiakasviestinnän yhteydessä. Tutkimuksessa havaittiin, että kahdessa yrityksistä käytettiin lähestymistapaa, jossa rekisteröidylle tarjottiin tilanteeseen mahdollisimman relevanttia informaatiota tietojen käsittelystä. Viestiin sisällytettiin viite, joka ohjasi rekisteröidyn yrityksen sivuille laajempaan tietosuojaselosteeseen. Tässä oli ajatuksena, että rekisteröidylle tarjotaan tärkeimmät tiedot ensiksi ja mahdollisuus tutustua laajemmin tietojen käsittelyyn kiinnostuksen mukaan.

Suurimmassa osassa (5) yrityksiä tietosuojaselosteet arvioitiin vähintäänkin säännöllisesti, mutta ei ollut asetettu tarkkaa aikaväliä arvioinneille. Kolmessa yrityksessä oli asetettu tietosuojaselosteiden päivittämiselle vuosittainen tarkastus. Jokaisen kahdeksan yrityksen kohdalla tuli ilmi, että mikäli päivityksen tarve havaittiin, niin tietosuojaseloste oli päivitetty mahdollisimman nopeasti. Päivitys tarpeet johtuivat usein esimerkiksi yrityksien nimien, palvelujen, oikeuskäytäntöjen ja tulkintojen muutoksista. Myös yhdessä yrityksessä vaikutustenarviointien yhteydessä havaitut muutokset johtivat tietosuojaselosteiden arviointiin ja päivittämiseen. Kuviossa 4 on esitetty yritysten aikataulu tietosuojaselosteiden arvioimiselle.



KUVIO 4 Tietosuojaselosteiden arvioimisen aikataulu yrityksissä

Osallistuneissa yrityksissä rekisteröidyillä oli vaihtoehtoisia tapoja tehdä tietopyyntö omista tiedoistaan. Rekisteröidyn tietojen pyytäminen tapahtui seitsemässä yrityksessä erillisen sähköisen lomakkeen kautta yrityksen sivuilta. Sähköinen lomake oli tyypillisesti PDF-muodossa ja muokattavissa suoraan selaimen kautta. Neljässä yrityksessä sähköinen lomake postitettiin ja kolmessa yrityksessä se toimitettiin turvasähköpostin kautta. Osassa yrityksistä (4) oli myös vaihtoehtona toimittaa sähköinen PDF-lomake yrityksen fyysiseen toimipisteeseen tai täyttää paikan päällä. Viidessä yrityksessä rekisteröity pystyi itse sähköisen asiointipalvelun kautta tarkastella omia tietojaan ja tekemään tietopyynnön. Lisäksi yhdessä yrityksessä oli käytössä sähköisesti allekirjoitettava verkkolomake, jossa varmistettiin rekisteröidyn henkilön henkilöllisyys käyttämällä vahvaa sähköistä tunnistautumista. Lisäksi kahdessa yrityksessä tietopyyntö oli mahdollista tehdä myös asiakaspalvelun kautta. Taulukossa 5 ovat koottuna tutkimuksessa havaitut tavat, joita kautta rekisteröidyn oli mahdollista tehdä tietopyyntö.

TAULUKKO 5 Rekisteröidyn tavat tehdä tietopyyntö

Rekisteröidyn tavat tehdä tietopyyntö	**Havaintojen lukumäärä**
Sähköinen asiointipalvelu	4
Sähköisen PDF-lomakkeen postittaminen	4
Fyysisessä toimipisteessä asioiminen	4
Sähköisen PDF-lomake turvasähköpostilla	3
Asiakaspalvelun kautta	2
Sähköisesti allekirjoitettava verkkolomake	1

Suurimmassa osassa (5) yrityksistä oli käytössä yksi keskitetty lomake kaikille rekisteröidyn pyynnöille. Tutkimuksessa kuitenkin havaittiin kahden yrityksen kohdalla, että eri toiminnoille ja palveluille oli omat lomakkeensa. Yksi tähän vaikuttava tekijä oli eri palveluille kohdistuva lainsäädäntö, joka asetti henkilötietojen käsittelylle erityispiirteitä. Yhdessä yrityksessä esimerkiksi tarkastuspyynnöille ja oikaisupyynnöille oli omat lomakkeensa. Kahdessa sähköistä asiointipalvelua hyödyntävässä yrityksessä pystyi tarkastelemaan ja lataamaan omat tiedot, mutta oikeus poistaa omat tiedot ja pyytää käsittelyn rajoittamista vaati erillisen lomakkeen täyttämistä. Lomaketta oli mahdollista pyytää yrityksen asiakaspalvelusta. Yhdessä yrityksessä myös oikeus saada pääsy omiin tietoihin ja tietojen siirtäminen vaati erilliset tietopyynnöt.

Rekisteröidyn henkilön tunnistamiseen oli käytössä useampi menetelmä yrityksissä. Eniten käytössä (5) ollut menetelmä tunnistaa rekisteröity oli vahva sähköinen tunnistaminen. Sähköisessä vahvassa tunnistautumisessa rekisteröity tunnistautui verkkopankkitunnuksilla tai mobiilivarmenteen avulla. Neljässä yrityksessä rekisteröidyn oli mahdollista myös tunnistautua henkilötodistuksen avulla toimittamalla kopion henkilötodistuksesta tai tunnistautumalla fyysisesti toimipisteessä. Yhdessä yrityksessä rekisteröity henkilö voitiin myös tunnistaa rekrytoinnin yhteydessä tehtyjen tunnusten kautta. Yhdessä yrityksessä myös allekirjoitus katsottiin riittäväksi tunnistamaan rekisteröity henkilö, kun ky-

seessä oli kirjallinen lomake. Taulukossa 6 on koottu eri tavat, joiden avulla tutkittavissa yrityksissä voitiin tunnistaa rekisteröity henkilö.

TAULUKKO 6 Tavat tunnistaa rekisteröity tietopyyntöjä tehdessään

Rekisteröidyn tunnistamiseen käytetyt tavat	**Havaintojen lukumäärä**
Sähköinen vahva tunnistautuminen (verkkopankkitunnukset tai mobiilivarmenne)	6
Henkilötodistus	4
Kirjautumistiedot	1
Allekirjoitus	1

Tietopyyntöjen käsitteleminen oli toteutettu eri tavalla tutkittavissa yrityksissä. Kahdella yrityksistä oli käytössä asianhallintajärjestelmä, jota kautta pyynnöt kohdistettiin oikealle henkilölle käsiteltäviksi. Asianhallintajärjestelmässä oli myös mahdollisuus seurata tietopyyntöjen etenemistä. Kahdessa yrityksessä tietopyyntöjen käsittelemiseen oli koulutettu erillinen asiakaspalvelutiimi, joka käsitteli rekisteröidyn tietopyynnöt, ja yhdessä yrityksessä tietopyyntöjen käsitteleminen oli osa back office -toimintoja. Yhdessä yrityksessä tiedot haettiin osittain automaattisesti sähköisen asiointipalvelun kautta. Tutkimuksessa havaittiin myös, että yhdessä yrityksessä oli käytössä nimetty lista henkilöitä, joille tietopyynnöt osoitettiin. Tietosuojavaltuutettu avusti tietopyyntöjen käsittelyssä. Lisäksi yhdessä yrityksessä tietosuojavastaava käsitteli rekisteröidyn pyynnöt sovellusten omistajien avustuksella.

Tutkimuksessa selvitettiin, pääseekö rekisteröity käsiksi suurimpaan osaan omista tiedoistaan yrityksen sivujen kautta. Suurimmassa osassa (7) yrityksessä selvisi, että rekisteröity näki yrityksen asiointipalvelusta omat oleelliset perustiedot, kuten yhteys- ja laskutustiedot tai rekrytoinnin yhteydessä kerätyt tiedot. Yhden edellä mainittujen yritysten kohdalla tämä oli toteutettu ainoastaan yhden palvelun osalta. Suurimmassa osassa (7) yrityksistä rekisteröidyn oli mahdollista muokata osaa omista tiedoistaan, kuten sähköpostia ja puhelinnumeroa. Kolmessa yrityksessä osoitetiedot haettiin Digi- ja väestötietoviraston kautta, joten niitä rekisteröity ei pystynyt muokkaamaan. Yhden yrityksen kohdalla rekisteröidyn ei ollut mahdollista päästä käsiksi omiin tietoihin yrityksen asiointipalvelusta, eikä myöskään muokata niitä. Viidessä yrityksessä myös rekisteröidyn muokkaamat tiedot päivittyivät yrityksen muihin järjestelmiin.

Tutkimuksessa selvitettiin, poistetaanko rekisteröidyn tiedot tietopyynnön seurauksena automaattisesti vai manuaalisesti. Yhdessä yrityksessä rekisteröidyn ei ollut mahdollista esittää poistopyyntöä, sillä yrityksen toimialalla käsiteltiin tietoja, joihin liittyi muita käsittelyperusteita ja pitkiä säilytysaikoja – osaa tiedoista tuli säilyttää jopa 12 vuotta rekisteröidyn kuoleman jälkeen. Lisäksi yhdessä yrityksessä poistopyyntö koski vain erittäin rajallista osaa rekisteröidyn tiedoista. Yrityksissä (6), joissa poistopyyntö oli mahdollista tehdä, se arvioitiin aina erikseen manuaalisesti. Tähän vaikutti muun muassa toimialakohtaiset tiukat lakisääteiset säilytysajat tiettyjen tietojen osalta sekä jokin muu

mahdollinen tietoihin liittyvä käsittelyperuste. Kolmen yrityksen kohdalla selvisi, että kun arviointi poistopyynnöstä oli tehty, tiedot poistettiin järjestelmistä automaattisesti. Yhden yrityksen kohdalla havaittiin, että rekisteröidyn oli mahdollista itse poistaa omat rekrytointitilanteessa antamansa tiedot. Tutkimuksessa ei saatu vastausta siihen, miten yritykset varmistavat, että automaattiset tiedot ovat varmasti poistuneet yrityksen eri järjestelmistä.

Osallistuneissa yrityksissä tiedot voitiin toimittaa rekisteröidylle käyttäen useampaa eri kanavaa. Viidessä yrityksessä tiedot oli mahdollista toimittaa rekisteröidylle kirjepostina. Kahdessa yrityksessä tuli ilmi, että tiedot lähetettiin Digi- ja väestötietoviraston mukaiseen osoitteeseen. Näin pienennettiin riskiä lähettää tiedot väärään osoitteeseen. Kahdessa yrityksessä tiedot lähetettiin kirjattuna kirjeenä, joka vaati henkilön tunnistamisen kirjettä noudettaessa. Viidessä yrityksessä tiedot voitiin lähettää myös salatun sähköpostin kautta. Kahden haastateltavan kohdalla tuli ilmi, että tämä ei ollut suositeltava tapa toimittaa tietoja siihen liittyvien riskien, kuten rekisteröidyn antaman väärän sähköpostiosoitteen takia. Neljässä yrityksessä rekisteröidyn oli mahdollista ladata omat tiedot sähköisen asiointipalvelun kautta itselleen. Kahden haastateltavan näkemys oli, että kaikkien yritysten tulisi pyrkiä toimittamaan tiedot sähköisen asiointipalvelun kautta. Näin pienennettäisiin tarvetta siirrellä rekisteröidyn tietoja Internetin yli ja, ehkäistäisiin siihen liittyviä riskejä. Taulukkoon 7 on koottuna eri kanavat, joiden kautta tiedot voitiin toimittaa rekisteröidylle.

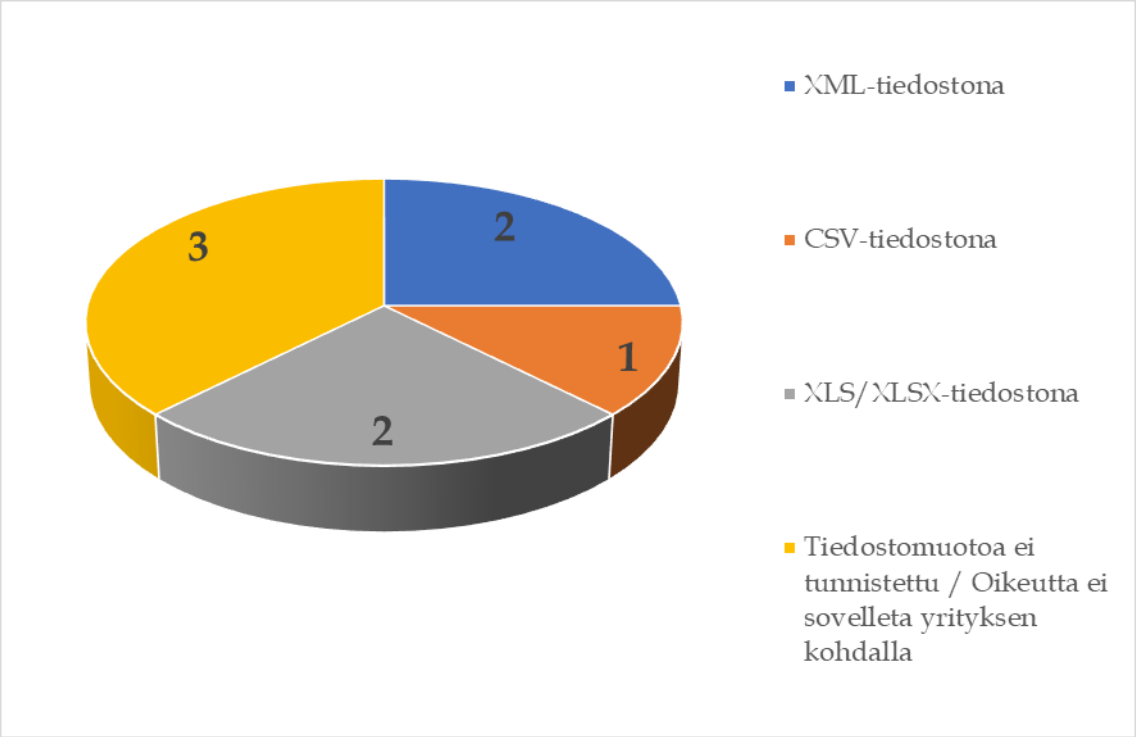
TAULUKKO 7 Omien tietojen toimittaminen rekisteröidylle tutkimukseen osallistuneissa yrityksissä

Tietojen toimittaminen rekisteröidylle	**Havaintojen lukumäärä**
Sähköinen asiointipalvelu	4
Salattu sähköposti	4
Kirjeposti	5

Jäljennös rekisteröidyn tiedoista toimitettiin eri muodoissa tutkimukseen osallistuneissa yrityksissä. Kahdessa yrityksessä tiedot oli mahdollista ladata PDF-formaatissa sähköisen asiointipalvelun kautta ja yhdessä yrityksessä ne lähetettiin edellä mainitussa muodossa salatun sähköpostin liitteenä. Kahdessa yrityksessä jäljennös rekisteröidyn tiedoista voitiin lähettää yleisesti tunnistettuna sähköisenä dokumenttina, kuten Word-, Excel- tai vastaavassa muodossa olevana tiedostona salatun sähköpostin liitteenä. Lisäksi osassa (5) yrityksistä rekisteröidyn oli mahdollista saada omat tiedot, joko pelkästään, tai vaihtoehtoisesti paperisena kopiona.

Kolmessa yrityksessä rekisteröidyn ei ollut mahdollista pyytää tietojen siirtämistä järjestelmästä toiseen, sillä yritykset toimivat toimialalla, jossa tätä oikeutta ei voitu soveltaa. Viidessä yrityksessä rekisteröidyn oli mahdollisuus käyttää oikeuttaan siirtää tiedot järjestelmästä toiseen. Käytännössä tämä tarkoitti, että tiedot toimitettiin rekisteröidylle yleisesti käytetyssä ja koneellisesti luettavassa muodossa. Kolmessa yrityksessä rekisteröidyn oli mahdollista ladata tiedot yrityksen sähköisen asiointipalvelun kautta CSV- (1) ja XML-

tiedostomuodossa (2). Kahdessa yrityksessä tiedostomuodoksi oli tunnistettu XLS/XLSX-tiedostomuoto. Kuitenkin toisessa näistä tiedostomuoto tuli sopia erikseen asiakkaan kanssa. Kuvioon 5 on koottu yritysten tunnistamat tiedostomuodot koskien rekisteröidyn oikeutta siirtää tiedot järjestelmästä toiseen.



KUVIO 5 Tunnistetut tiedostomuodot koskien rekisteröidyn siirto-oikeutta

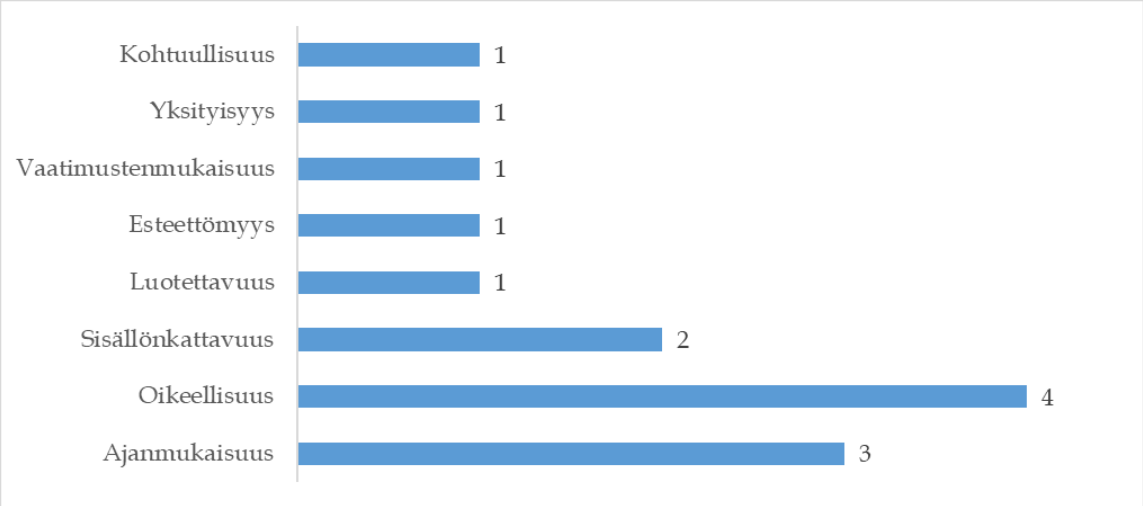
6.3 Datanhallinnan rooli rekisteröidyn oikeuksia koskevien vaatimusten täyttämisessä

Tutkimukseen osallistuneilta yrityksiltä tiedusteltiin, millainen rooli datanhallinnalla on tietosuoja-asetuksen vaatimusten täyttämisessä. Suurimmassa (6) osassa yrityksiä datanhallinta oli todella merkittävässä ja keskeisessä roolissa asetuksen vaatimusten täyttämisessä. Haastatteluissa tuli ilmi muun muassa, että datanhallinta muodostaa perustan tietosuojan toteuttamiselle ja sen avulla luodaan kokonaisvaltaisen kuva yrityksen datasta ja sen käyttötarkoituksista. Yhdessä yrityksessä todettiin, että datanhallinta oli erillinen organisaatio tietosuojan hallinnasta, eikä rooli korostunut niin merkittävästi asetuksen vaatimusten täyttämisessä. Lisäksi yhden yrityksen kohdalla havaittiin, että datanhallinnan rooli näyttäytyi asianhallintajärjestelmän kautta, johon oli kuvattu vastuuhenkilöt eri prosesseille.

Yhtä yritystä lukuun ottamatta kaikissa (7) yrityksissä data governance - toiselta nimeltään datan hallintamalli - oli otettu käyttöön ja yrityksissä sen tehtävänä oli muun muassa määrätä vastuut dataa koskevista päätöksistä. Kahden julkishallinnon toimijan kohdalla tuli ilmi, että tiedonhallintalaki on määrännyt

data governancen käyttöönoton. Tutkimuksessa myös selvitettiin, miten osallistuneissa yrityksissä data governance on auttanut yleisen tietosuoja-asetuksen vaatimusten täyttämisessä. Suurin osa (6) haastateltavista totesi, että data governancen kautta on määritetty ja dokumentoitu datan omistajat ja selkeät vastuuhenkilöt yrityksen eri rekistereistä. Yhden yrityksen kohdalla tuli ilmi, että tietosuojaprosessit olivat sisäänrakennettuina osaksi data governancea jo ennen yleisen tietosuoja-asetuksen voimaan astumista. Kyseisessä yrityksessä varsinkin dataa hyödyntävien uusien teknologioiden käyttöönoton osalta on ollut erittäin kriittistä, että tietosuojavaatimukset ovat olleet sisäänrakennettuina osaksi data governancea ja datanhallintaa.

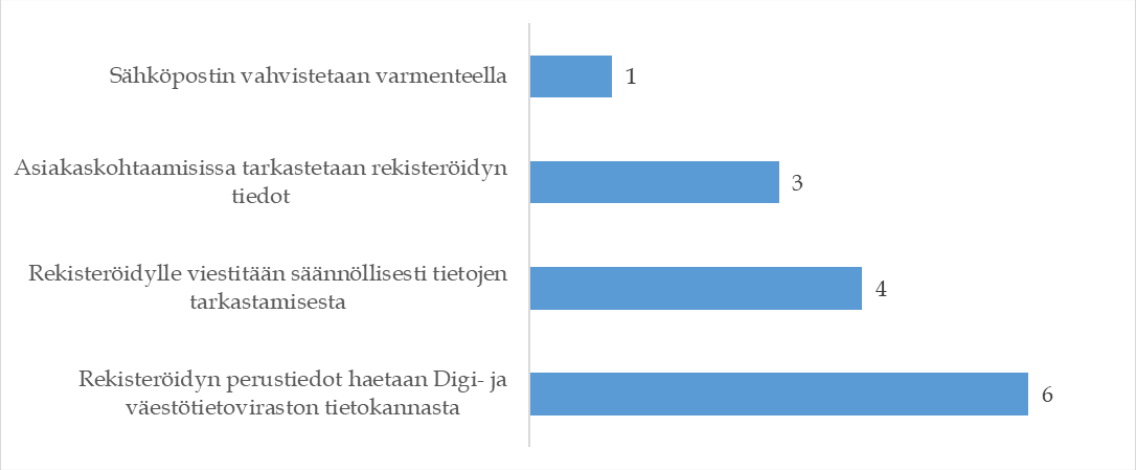
Tutkimuksessa selvitettiin, mitkä datan laadun ulottuvuudet ovat erityisesti korostuneet yleisen tietosuoja-asetuksen vaatimusten täyttämisessä. Kysymys oli haasteellinen, sillä yrityksissä datan laatu nähtiin enemmänkin kokonaisuutena, eikä niinkään yksittäisinä ulottuvuuksina. Yrityksissä muun muassa todettiin, että datan eheyttä oli pyritty parantamaan ja tarpeetonta dataa oli siivottu säännöllisesti. Yhden yrityksen kohdalla havaittiin, että datan laadun ulottuvuudet katetaan tiedonhallinnan, tietosuojan ja tietoturvan kautta. Viidessä yrityksessä kuitenkin korostui seuraavat datan laadun ulottuvuudet: ajanmukaisuus, oikeellisuus, kattavuus, luotettavuus, sisällönkattavuus, esteettömyys, vaatimustenmukaisuus, yksityisyys ja kohtuullisuus. Yhden yrityksen kohdalla tuli ilmi, että erityisesti ajanmukaisuus asiakkuudenhallintajärjestelmissä koettiin haasteeksi. Lisäksi yhdessä yrityksessä rekisteröidyn itse ilmoittamien tietojen oikeellisuus ja sen varmistaminen on korostunut erityisesti asetuksen kontekstissa. Kuvioon 6 on koottu datan laadun ulottuvuudet, jotka ovat korostuneet yrityksissä yleisen tietosuoja-asetuksen kontekstissa.



KUVIO 6 Yrityksissä korostuneet datan laadun ulottuvuudet

Yrityksissä oli varmistettu rekisteröidyn tietojen laatu usealla eri tavalla. Suurin osa (6) yrityksiä käytti ainakin yhtenä menetelmänä Digi- ja väestötietoviraston tietokantaa, josta rekisteröidyn tiedot haettiin ja päivitettiin säännöllisesti. Neljässä yrityksessä lisäksi rekisteröidylle viestittiin säännöllisesti tietojen tarkas-

tamisesta ja päivittämisestä. Kolmen yrityksen kohdalla asiakaskohtaamisissa tarkastettiin ja varmistettiin tietojen oikeellisuus. Lisäksi yhdessä yrityksessä rekisteröidyn sähköposti varmistettiin vahvistusviestillä. Kuvioon 7 on koottuna eri menetelmät, joilla osallistuneissa yrityksissä varmistettiin rekisteröidyn tietojen laatu.



KUVIO 7 Yrityksien menetelmät varmistaa rekisteröidyn tietojen laatu

Yrityksiltä selvitettiin, miten henkilötietoihin liittyvä master data on järjestetty. Neljässä yrityksessä oli käytössä yksi keskitetty master järjestelmä, joka oli integroitu muihin yrityksen järjestelmiin. Yhdessä yrityksessä havaittiin, että master datan hallinnalla on pyritty varmistamaan tiedon siirtyminen muihin järjestelmiin muuttumattomana. Lisäksi osassa (2) edellä mainituista yrityksiä oli myös muita palvelukohtaisia järjestelmiä, jotka eivät olleet integroituna master järjestelmään. Yhdessä yrityksessä rekisteröidyn tiedot haettiin henkilötunnuksen avulla eri järjestelmistä, sillä käyttöoikeuksien takia tietojen täytyi olla eriytetyissä järjestelmissä. Yhden yrityksen kohdalla master datan hallinta oltiin ottamassa käyttöön ja tällä hetkellä käytettiin tiedon kategorisointia, joka oli edellytys rekisteröidyn oikeuksien toteuttamiselle. Tietojen kategorisoinnilla pystyttiin hakemaan eri järjestelmistä saman tyyppiset tiedot, kuten esimerkiksi sopimustiedot. Lopuissa yrityksistä master datan hallinta ei ollut käytössä (1) tai kysymykseen ei voitu vastata (1).

Neljän yrityksen mukaan master datan hallinta oli auttanut vaatimustenmukaisuudessa, kun rekisteröidyn tiedot olivat haettavissa ja päivitettävissä yhdestä järjestelmästä. Näistä lisäksi yhden yrityksen kohdalla nousi esiin datan määritteleminen ja mallinnus, kun esimerkiksi palvelu on useamman rekisteröidyn yhteiskäytössä. Yhden yrityksen kohdalla tietojen kategorisoinnin avulla pystyttiin kohdentamaan data tiettyyn rekisteröidyn oikeuteen. Kategorisoinnilla voitiin myös rajata sellainen data rekisteröidyn oikeuksien ulkopuolelle, johon liittyi jokin muu säilytystä vaativa laillinen velvoite. Yhdessä yrityksessä oli havaittu, että kaikkein palveluiden kohdalla master datan hallinnan kypsyystaso ei ollut riittävällä tasolla. Rekisteröidyn tietoja oli eri järjestelmissä, jonka myötä tietojen hakeminen vaati kohtuuttoman paljon resursseja. Kahden

yrityksen kohdalla master datan hallinnan vaikutuksia vaatimustenmukaisuuteen ei voitu todeta, sillä sitä ei ollut otettu käyttöön yrityksissä.

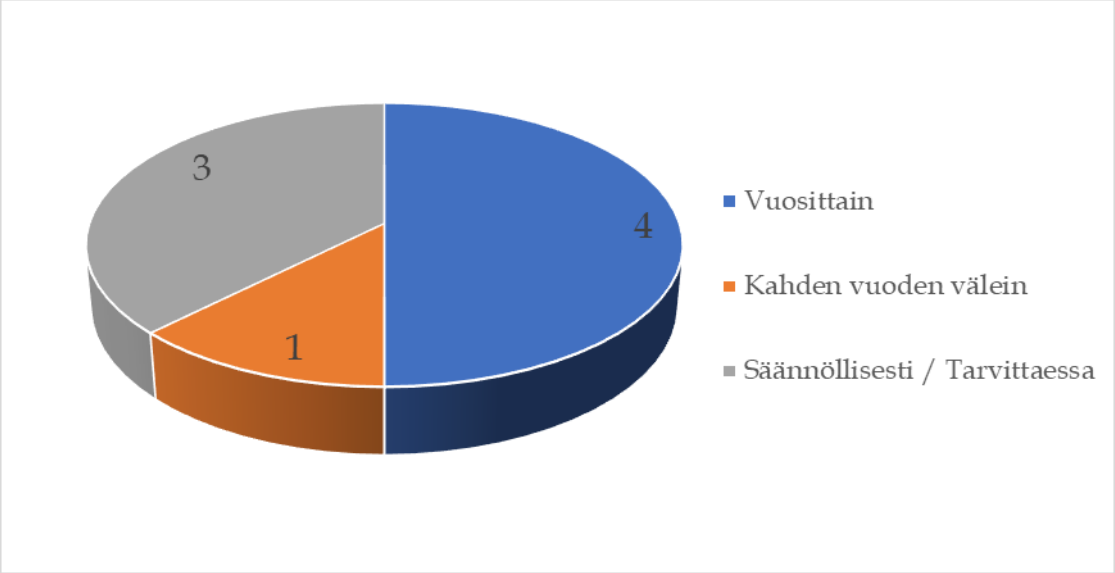
Tutkimuksessa havaittiin, että ainoastaan kahdessa yrityksessä oli käytössä tietoturvastandardi. Kummatkin yritykset olivat sertifioituneet ISO/IEC 27001 mukaisiksi sekä lisäksi toinen yrityksistä oli myös ISO/IEC 27002 mukainen. Osassa (2) yrityksiä oli sertifioitu ISO/IEC 27001 mukaisiksi vain osa liiketoiminta yksiköistä. Kolmessa yrityksessä ei ollut käytössä mitään tietoturvastandardia, mutta toimintaa oli mitattu ISO/IEC 27001 ja VAHTI-ohjeita vasten. Lisäksi kolmen yrityksen kohdalla tuli ilmi, että toiminnassa on seurattu tietoturvan viitekehyksiä, kuten information security forumin tietoturvan hyviä käytäntöjä (engl. standard of good practice for information security) ja NIST:n tietoturvan viitekehystä (engl. cyber security framework, CSF). Lisäksi muutamalla toimialalla lainsäädännön asettamat vaatimukset säätelevät toimintaa tiukasti ja osittain ISO/IEC 27001 mukaisesti, joten standardin sertifioitumiselle ei nähty tarvetta. Taulukossa 8 on esitetty yritysten tietoturvastandardien käyttö.

TAULUKKO 8 Tietoturvastandardien käyttö yrityksissä

Onko yrityksessä käytössä tietoturvastandardia?	**Tietoturvastandardi**	
	ISO/IEC 27001	ISO/IEC 27002
Kyllä	2	1
Osittain	2	0
Ei	3	3

Tutkimuksessa selvitettiin, mitä hyötyä tietoturvastandardista on ollut yleisen tietosuoja-asetuksen suhteen. Kahdessa yrityksessä, joissa tietoturvastandardia noudatettiin, havaittiin standardin luovan perustan asetuksen vaatimusten täyttämiselle etenkin tiedon luokittelun ja prosessoinnin osalta. Myös asetuksen vaatimuksiin pystyttiin vastaamaan paremmin, kun tietoturvastandardi oli käytössä. Tietoturvan viitekehysten ja oman toiminnan mittaaminen standardien vaatimuksia vasten oli auttanut yrityksiä tunnistamaan tietojen käsittelyyn liittyviä riskejä, helpottanut tietosuoja-asetuksen vaatimusten täyttämisessä sekä auttanut havaitsemaan omien prosessien asianmukaisuutta.

Kaikissa (8) yrityksissä oli tietoturvapolitiikka, jossa ohjeistettiin henkilötietojen käsittelystä. Tosin osassa (2) yrityksistä oli erillinen tietosuojapolitiikka ja yhdessä yrityksessä tietosuojaperiaatteet. Kuviossa 8 on kuvattu tietoturva- tai suojapolitiikkojen arvioimisen aikataulu yrityksissä. Tietoturva- tai suojapolitiikkoja arvioitiin suurimmassa osassa (4) yrityksistä vuosittain. Yhdessä yrityksessä tietoturvapolitiikka arvioitiin kahden vuoden välein. Kolmen yrityksen kohdalla niitä arvioitiin vähintäänkin säännöllisesti tai tarvittaessa.



KUVIO 8 Tietoturvapolitiikkojen arvioimisen aikataulu yrityksissä

Kaikissa (8) yrityksissä työntekijöiden pääsyä rekisteröidyn tietoihin rajoitettiin roolipohjaisen pääsynhallinnan avulla. Yhdessä yrityksen edustaja toi ilmi, että roolipohjaisen pääsynhallinnan toteutus tulee miettiä erittäin tarkasti uusia järjestelmiä hankkiessa. Lisäksi yhdessä yrityksessä henkilöstön täytyi suorittaa koulutus, ennen kuin pystyivät käsittelemään henkilötietoja.

Tutkimuksessa selvitettiin, miten yrityksissä hallitaan strukturoimatonta dataa. Haastatteluiden perusteella selvisi, että suurimmassa osassa (5) yrityksistä strukturoimatonta dataa hallittiin ohjeistuksen kautta. Yksi haastateltava korosti, että yrityksen lähtökohtana on tallentaa henkilötietoja sisältävä data ainoastaan strukturoituihin tietokantoihin ja perusteltuna voitiin tallentaa tietoja strukturoimattomasti. Tämän lisäksi osassa (3) yrityksiä oli myös teknisiä ratkaisuita rakenteettoman datan hallintaan. Teknisiä ratkaisuita ei voitu nimetä, mutta käytännössä niiden avulla pystyttiin havaitsemaan henkilötietoja sisältäviä dokumentteja esimerkiksi yrityksen verkkolevyiltä tai pilvipalveluista. Kahdessa yrityksessä oli tehty manuaalinen tarkastus verkkolevyille henkilötietoja sisältävien dokumenttien osalta. Lisäksi yhdessä yrityksessä käytettiin datan luokittelua, jonka avulla voitiin hallita strukturoimatonta dataa. Taulukkoon 9 on koottu eri tavat, joilla yrityksissä hallittiin strukturoimatonta dataa.

TAULUKKO 9 Strukturoimattoman datan hallinnan menetelmät yrityksissä

Tutkimuksessa havaitut menetelmät	**Havaintojen lukumäärä**
Ohjeistus	5
Tekninen ratkaisu strukturoimattoman datan analysointiin	3
Manuaaliset tarkastukset	2
Datan luokitteleminen	1

Suurimmassa osassa (7) yrityksistä ei ollut käytössä omaa turvaluokitusta henkilötietoja sisältäville dokumenteille. Yrityksissä oli käytössä yleiset turvaluokitukset, kuten salassa pidettävä, sisäinen ja julkinen luokka, joihin henkilötiedot sisältyivät riippuen niiden arkaluontoisuudesta. Yksi haastateltava totesi, että henkilötiedoille ei ole relevanttia tehdä omaa luokaa, sillä useasti dokumentit sisältävät paljon henkilötiedoiksi luokiteltua tietoa. Tällöin kaikki dokumentit tulisi luokitella henkilötietoa sisältävään turvaluokkaan. Kahdessa yrityksessä henkilötiedot tyypillisesti kuuluivat salassa pidettävään turvaluokkaan. Lisäksi yhdessä yrityksessä havaittiin, että turvaluokitusten sijaan dokumentit luokitellaan niihin sisältyvän tiedon perusteella. Luokan avulla voidaan hakea ja tunnistaa henkilötietoja sisältävät dokumentit. Yhdessä yrityksessä henkilötiedoille oli oma turvaluokka ja dokumentin metatiedoissa oli merkintä, mikäli se sisälsi henkilötietoja.

Kuudessa yrityksessä henkilötiedot pystyttiin tunnistamaan strukturoimattomasta datasta. Yrityksissä oli käytössä erilaisia teknisiä ratkaisuja henkilötietojen tunnistamiseen strukturoimattomasta datasta, joita kuitenkaan suurin osa yrityksistä ei voinut paljastaa. Kolmen yrityksen kohdalla tuli ilmi, että käytössä oli data loss prevention (DLP) -ratkaisuja, joilla voitiin tunnistaa dokumenttien sekä myös kuvien tietosisältö. Näitä ratkaisuita käytettiin tunnistamaan henkilötietoja yhteisiltä verkkolevyiltä, pilvipalveluista ja työntekijöiden omilta päätelaitteilta. Yhdessä yrityksessä henkilötiedot tunnistettiin strukturoimattomasta datasta verkkolevyjen skannauksen ja manuaalisen tarkastuksen sekä erillisen auditoinnin kautta. Kahdessa yrityksessä henkilötietoja ei voitu tunnistaa strukturoimattomasta datasta.

6.4 Datanhallinta tulevaisuudessa

Haastattelun lopuksi selvitettiin tutkittavien mielipidettä datanhallinnan tapoihin, joilla voidaan vastata henkilötietojen käsittelyyn liittyvien asetusten vaatimuksiin tulevaisuudessa. Taulukkoon 10 on koottu tutkittavien näkemykset datanhallinnan tavoista tulevaisuudessa. Kolme haastateltava toi esille pääsynhallintaan liittyvät haasteet, jotka tulisi huomioida tulevaisuudessa paremmin. Etenkin huolenaiheita herätti pääsynhallinta työ- tai asiakassuhteiden päättyessä ja erityisesti huolehtiminen, että pääsy loppuu edellä mainituissa tilanteissa. Toinen tutkittava lisäsi vielä lokienhallinnan merkityksen, kun tarkastellaan jälkeenpäin, kuka on tarkastellut mitäkin tietoja.

Kaksi haastateltavaa mainitsi master datan hallinnan olevan keskeinen datanhallinnan tapa, jolla tulevaisuudessa voidaan vastata henkilötietojen käsittelyyn liittyvien asetusten vaatimuksiin. Tutkittavat toivat ilmi, että datan tallentaminen ja päivittäminen yhteen järjestelmään, josta se päivittyisi muihin järjestelmiin, olisi suuri hyöty eri vaatimusten täyttämiselle. Toinen tutkittava lisäsi myös asetusten vaatimusten määrittelemisen osaksi metadataa helpottavan vastaamaan tietosuoja-asetusten vaatimuksiin. Esimerkiksi metadatan attribuuttei-

hin voisi määritellä tiedon säilytysajan, jolloin tiedettäisiin, mitä tietoja joudutaan säilyttämään lain velvoittamana pidempään.

Kolme tutkittavaa korosti tietosuojan ja datan käsittelyn suunnittelua ennalta. Yksi tutkittava mainitsi datan käsittelyn suunnittelemisen ja riskiarviointien tekemisen ennakolta. Kriittistä olisi pohtia jo ennen datan keräämistä sen käyttötarkoitusta, kuten mitä, miksi ja miten kerätään, sekä miten dataa käsitellään. Nämä tulisi yhdistää lakisääteisiin vaatimuksiin ja osaksi yrityksen sisäisiä politiikoita ja data governancea. Toinen tutkittava keskittyi evästeiden ja käytönseurannan haasteisiin, jossa tulisi hyödyntää PDCA-mallia (Plan, Do, Check, Act). Tutkittava korosti, että suunnittelun lisäksi myös varmistetaan esimerkiksi sisäisten auditointien kautta, että ohjeita todellakin noudatetaan. Kolmas tutkittava vastasi, että sisäänrakennettu tietosuoja (engl. privacy by design) on otettava huomioon jo suunnitteluvaiheessa kaikessa kehityksessä. Lisäksi yhden yrityksen edustaja mainitsi datan luokittelemisen olevan erittäin keskeisessä roolissa tulevaisuudessa.

TAULUKKO 10 Tutkittavien näkemys datanhallinnan tavoista tulevaisuudessa tietosuoja-asetusten kontekstissa

Tutkittavien näkemykset datanhallinnan tavoista tulevaisuudessa	**Havaintojen lukumäärä**
Pääsynhallinta	3
Tietosuojan ja datan käsittelyn suunnittelu ennalta	3
Master datan hallinta	2
Datan luokitteleminen ja kategorisointi	1

6.5 Yhteenveto tutkimuksen tuloksista

Kaikissa yrityksissä oli vastattu yleisen tietosuoja-asetuksen rekisteröidyn oikeuksiin liittyviin vaatimuksiin. Tämä oli tosin olettamus, sillä asetus on ollut jo voimassa useamman vuoden. Rekisteröidyn oikeuksiin liittyvät vaatimukset olivat toteutettu eri tavalla osallistuneissa yrityksissä. Tähän havaittiin vaikuttavan muun muassa yrityksen toimiala ja henkilötietojen käsittelyn laajuus.

Kaikissa yrityksissä kuvaus rekisteröidyn henkilötietojen käsittelystä oli sisällytetty tietosuojaselosteeseen, joka oli tyypillisesti saatavilla ja luettavissa yrityksen verkkosivuilta tai ladattavissa sieltä PDF-muodossa. Tietosuojaseloste tarjottiin rekisteröidylle yleisesti asiakaskohtaamisen yhteydessä. Suurimmassa osassa ne arvioitiin säännöllisesti, mutta tarkkaa aikaväliä päivittämiselle ei ollut määritelty.

Rekisteröidyllä oli vaihtoehtoisia tapoja tehdä tietopyyntö omista tiedoistaan yritykselle. Pääasialliset tavat tehdä tietopyyntö olivat sähköisen asiointipalvelun, sähköisen PDF-lomakkeen postittamisen tai fyysisessä toimipisteessä asioimisen kautta. Tutkimuksessa havaittiin, että rekisteröidyn tunnistamiseen oli useampi eri menetelmä osallistuneissa yrityksissä. Havaitut tavat tunnistaa

rekisteröity tietopyynnön yhteydessä olivat sähköinen vahva tunnistautuminen, henkilötodistus, kirjautumistiedot sekä allekirjoitus. Sähköinen vahva tunnistautuminen oli yleisin tapa tunnistaa rekisteröity henkilö. Lisäksi toiseksi yleisin tapa tunnistaa rekisteröity oli henkilötodistuksen perusteella, josta lähetettiin kopio tietopyynnön yhteydessä tai tunnistauduttiin sen avulla yrityksen fyysisessä toimipisteessä. Tietopyyntöihin vaadittavan lomakkeen osalta, suurimmassa osassa yrityksistä oli käytössä yksi keskitetty lomake. Tietopyynnöt käsiteltiin useimmissa yrityksissä manuaalisesti, mutta yhdessä yrityksessä oli hyödynnetty sähköisessä asiointipalvelussa automatiikkaa tietojen hakemiseen. Suurimmassa osassa yrityksistä rekisteröidyn oli mahdollista tarkastella ja muokata omia perustietoja, kuten yhteys- ja laskutustietoja, kirjautumalla yrityksen sivuille. Lisäksi osassa yrityksiä muokatut tiedot päivittyivät yrityksen muihin järjestelmiin. Jokaisessa yrityksessä rekisteröidyn poistopyynnöt käsiteltiin aina erikseen ja manuaalisesti. Tähän liittyi usein muiden lakien asettamat käsittelyperusteet tai säilytysajat tietojen säilyttämiselle, jolloin tiedot tuli erikseen tarkistaa ennen niiden poistamista. Tiedot toimitettiin rekisteröidylle hyödyntäen eri kanavia ja useimmissa yrityksissä oli vaihtoehtoina lähettää tiedot kirjepostin, sähköisen asiointipalvelun tai salatun sähköpostin kautta. Rekisteröidyn oli mahdollista pyytää jäljennös omista tiedoistaan ja suurimmassa osassa yrityksiä ne tarjottiin sähköisesti PDF- tai XLS/XLSX-tiedostomuodossa, mutta myös osalla oli vaihtoehtoisesti tai pelkästään tarjota tiedot paperisena kopiona. Useimmissa yrityksissä rekisteröity pystyi käyttämään oikeuttaan siirtää tiedot järjestelmästä toiseen, jossa asetuksen vaatimaksi yleisesti hyväksytyksi koneluettavaksi muodoksi tunnistettiin XML-, CSV- ja XLS/XLSX-tiedostomuoto.

Suurin osa yrityksistä totesi, että datanhallinnalla on keskeinen rooli yleisen tietosuoja-asetuksen vaatimusten täyttämisessä. Myös useimmissa yrityksissä oli hyödynnetty data governancea ja sen havaittiin hyödyttävän asetuksen vaatimustenmukaisuudessa etenkin vastuiden ja datan omistajuuden kautta. Datan laadun ulottuvuudet, jotka olivat korostuneet yleisen tietosuoja-asetuksen kontekstissa, nähtiin yrityksissä enemmänkin kokonaisuutena kuin yksittäisinä ulottuvuuksina. Kuitenkin tutkimuksessa havaittiin korostuneen datan laadun ulottuvuuksista asetuksen kontekstissa etenkin oikeellisuus, ajanmukaisuus ja sisällönkattavuus. Suurimassa osassa yrityksistä hyödynnettiin Digi- ja väestötietoviraston tietokantaa rekisteröidyn perustietojen varmistamiseen, jolla parannettiin tietojen laatua. Myös muita havaittuja tapoja varmistaa rekisteröidyn tietojen laatu olivat säännöllinen viestiminen tietojen tarkastamisesta sekä asiakaskohtaamiset, jossa tiedot tarkastettiin. Puolessa yrityksiä henkilötietoihin liittyvää master dataa hallittiin yhden master järjestelmän kautta, joka oli integroitu muihin yritysten järjestelmiin. Yrityksissä koettiin sen auttavan asetuksen vaatimustenmukaisuudessa, kun rekisteröidyn tiedot olivat haettavissa ja päivitettävissä yhdestä paikasta. Tutkimuksessa havaittiin, että suurimmassa osassa yrityksistä ei ollut tai oli osittain käytössä tietoturvastandardi. Ainoastaan kahden yrityksen kohdalla havaittiin, että yritys oli sertifioitunut ISO/IEC 27001 mukaiseksi. Tietoturvastandardin koettiin hyödyt-

tävän vaatimusten täyttämisessä muun muassa luomalla perustan tietosuoja-asetuksen vaatimustenmukaisuudelle, vaatimuksiin pystyttiin vastaamaan paremmin sekä sen avulla tunnistettiin käsittelyyn liittyviä riskejä.

Kaikissa yrityksissä oli käytössä tietoturvapolitiikka ja osassa yrityksiä oli vielä erillinen tietosuojapolitiikka. Suurimmassa osassa tietoturvapolitiikoita arvioitiin vuosittain. Kaikissa yrityksissä työntekijöiden pääsyä rekisteröidyn tietoihin rajoitettiin roolipohjaisen pääsynhallinnan avulla. Strukturoimatonta dataa hallittiin useimmissa yrityksistä ohjeistuksen avulla, mutta osassa yrityksistä oli myös käytössä erilaisia teknisiä ratkaisuita datan analysointiin. Tutkimuksessa havaittiin, että suurimmassa osassa yrityksistä henkilötiedoille ei ollut omaa turvaluokitusta, vaan ne kuuluivat osaksi laajempia turvaluokituksia. Useassa yrityksessä henkilötiedot pystyttiin tunnistamaan strukturoimattomasta datasta erilaisten teknisten ratkaisuiden avulla, joita ei tutkimukseen osallistuneet voineet paljastaa. Osassa yrityksistä tuli ilmi, että käytössä oli data loss prevention -ratkaisuita, joilla voitiin tunnistaa dokumenttien sekä kuvien tietosisältö.

Tutkimuksessa nousi esille haastateltavien näkemyksiin pohjautuen kolme keskeisintä tulevaisuuden datanhallinnan tapaa, joilla pystytään vastaamaan tietosuoja-asetusten vaatimuksiin: pääsynhallinta, tietosuojan ja datan käsittelyn suunnittelu ennalta sekä master datan hallinta. Pääsynhallinnan osalta haasteeksi korostui pääsyn loppuminen työ- ja asiakassuhteiden päättyessä. Osa tutkittavista korosti tietosuojan ja datan käsittelyn suunnittelua ennalta, jossa tulisi pohtia jo ennen datan keräämistä, miksi, mitä ja miten kerätään, sekä miten dataa käsitellään. Master datan hallinnasta oli yhteinen näkemys, että sen avulla pystytään vastaamaan paremmin asetuksen vaatimuksiin, kun tiedot ovat haettavissa ja päivitettävissä yhdestä järjestelmästä.

7 POHDINTA

Tutkielmassa tarkasteltiin rekisteröidyn oikeuksien toteuttamista ja datanhallintaa yrityksissä. Tutkimuksessa tavoitteena oli selvittää, kuinka yrityksissä on toteutettu Euroopan unionin yleisen tietosuoja-asetuksen rekisteröidyn oikeuksia koskevat vaatimukset ja, kuinka datanhallintaa on hyödynnetty niiden täyttämisessä. Tutkimuksen kohderyhmäksi valikoitui eri yritysten tietosuojavastaavat tai henkilöt, joilla oli kokemusta asetuksen soveltamisesta. Tutkimus toteutettiin teemahaastatteluna. Tutkimukseen osallistui kahdeksan henkilöä, joka koettiin olevan riittävän kattava määrä tulosten analysointia ja johtopäätöksien tekemistä varten. Jokaiselta haastateltavalta varmistettiin ennalta, että heillä on kokemusta asetuksen parissa työskentelemisestä. Teemahaastattelu mahdollisti, että osallistujat ymmärsivät kysymykset, mutta myös tarkentavien kysymysten esittämisen. Tutkimuksessa etsittiin vastausta seuraavaan tutkimusongelmaan: Miten yleisen tietosuoja-asetuksen rekisteröidyn oikeuksia koskevat vaatimukset ovat toteutettu yrityksissä ja kuinka datanhallintaa on hyödynnetty niiden täyttämisessä? Tutkimuksen pohjalta voidaan todeta, että rekisteröidyn oikeudet olivat toteutettu yrityksissä eri tavalla. Toteutuksessa oli hyödynnetty datanhallintaa niin teknisin kuin myös hallinnollisin keinoin.

Tämä luku sisältää tutkimuksen keskeisimpien tulosten ja niiden johtopäätöksien esittämisen. Luvussa myös arvioidaan tutkimuksen onnistumista reliabiliteetin ja validiteetin kautta sekä esitetään tulosten hyödyntämismahdollisuuksia ja jatkotutkimusaiheita.

7.1 Keskeiset tulokset ja johtopäätökset

Yleisessä tietosuoja-asetuksessa ei oteta kantaa, miten vaatimukset tulisi toteuttaa, mikä on johtanut erilaisiin toteutustapoihin yrityksissä. Tutkimuksen perusteella voidaan todeta, että rekisteröidyn oikeuksia koskevat vaatimukset olivat toteutettu eri toimialoilla toimivissa yrityksissä eri tavoilla. Myös henkilötie-

tojen käsittelyn laajuuden havaittiin vaikuttavan siihen, millaisia tapoja yrityksissä on vastata rekisteröidyn oikeuksiin.

Tutkimuksessa havaittiin, että rekisteröidyn tietojen käsittelystä informoitiin tietosuojaselosteessa, joka oli saatavilla yrityksen internetsivuilla ja siihen viitattiin asiakaskohtaamisissa ennen rekisteröidyn tietojen käsittelemistä. Myös Ausloos ja Dewitte (2018) suosittelevat, että tietosuojaselosteet ovat helposti löydettävissä yrityksen internetsivuilta ja viite selosteeseen tulisi olla jokaisen sivun alareunassa. Hoiryn ja Norrisin (2015) tutkimuksen mukaan tietosuojaseloste tulisi tarjota rekisteröidylle ennen palveluun rekisteröitymistä, joka havaittiin toteutuvan tässä tutkimuksessa. Tutkimuksessa myös havaittiin, että tietosuojaselosteet arvioitiin säännöllisesti, mutta tarkkaa aikaväliä arvioinneille ei ollut määritelty. Tietosuojaselosteet olisi hyvä arvioida esimerkiksi vuosittain ja merkittävien muutosten yhteydessä.

Rekisteröity on aina tunnistettava ennen tietopyynnön käsittelemistä. Tutkimuksessa havaittiin, että yleisin tapa tunnistaa rekisteröity oli sähköisen vahvan tunnistautumisen kautta. Voidaan todeta, että sähköinen vahvatunnistautuminen on turvallinen ja hyvä tapa tunnistaa rekisteröity. Singhin ja Cobben (2019) mukaan tietopyyntöjen tekeminen tulisi tapahtua yrityksen omalla alustalla, jolloin rekisteröity voitaisiin tunnistaa hyödyntämällä samoja menetelmiä kuin palvelua käyttäessä. Erittäin sensitiivisten tietojen kohdalla olisi hyvä käyttää vahvaa monivaiheista tunnistautumista. Tässä tutkimuksessa havaittiin, että osa yrityksistä hyödynsi omaa alustaa tietopyyntöjen tekemiseen. Tietopyynnön tekemisen yhteydessä rekisteröity tunnistettiin sähköisen vahvan tunnistautumisen avulla palveluun kirjautumisen lisäksi. Tämän havaitsivat myös Martino ja kollegat (2019) omassa tutkimuksessaan. Osassa yrityksiä rekisteröity tunnistettiin henkilötodistuksen avulla, josta lähetettiin kopio tietopyynnön yhteydessä. Tämä on linjassa Ausloosin ja Bewitten (2018) tutkimuksen kanssa, jossa osa yrityksistä vaati kopion henkilötodistuksesta, jotta rekisteröity voitiin tunnistaa. Boniface ja hänen työryhmänsä (2019) esittivät tutkimuksessaan, että henkilötodistuksen kopion lähettäminen voi altistaa rekisteröidyn identiteettivarkauden uhriksi, mikäli se vuotaa vääriin käsiin. He suosittelevat yrityksien käyttävän monivaiheista tunnistautumista rekisteröidyn tunnistamiseen. Mikäli henkilötodistuksen kopiota käytetään rekisteröidyn tunnistamiseen, niin tulisi siinä käyttää vesi- ja aikaleimoja, jotka osoittavat kopion olevan vain kyseiselle rekisterinpitäjälle. (Boniface ym., 2019.) Johtopäätöksenä voidaan todeta, että yrityksien tulisi tunnistaa rekisteröity, joko sähköisen vahvan- tai monivaiheisen tunnistautumisen avulla ja hyödyntää tähän yrityksen omaa alustaa. Yrityksien tulisi myös välttää pyytämästä lisätietoja rekisteröidyn tunnistamiseksi, mikäli rekisteröity voidaan jo tunnistaa olemassa olevista tiedoista (Singh & Cobbe, 2019). Tietosuojavaltuutetun toimisto antoi tuoreimmassa päätöksessään 75 000 euron sanktiot yritykselle, joka laiminlöi tietojen minimointia pyytäessään rekisteröidyiltä tunnistamiseen liittyviä tietoja, joita sillä ei ollut alun perin hallussaan. Yrityksen rikkomukset lisäksi koskivat rekisteröidyn oikeuksien puutteellista toteuttamista ja tietojen säilytysaikojen rajoittamista. (Tietosuojavaltuutetun toimisto, 2021.)

Rekisteröidyn oli mahdollista tehdä tietopyyntö yrityksille useammalla eri tavalla. Tietopyyntö voitiin tehdä sähköisen asiointipalvelun, sähköisen PDF-lomakkeen postittamisen tai fyysisessä toimipisteessä asioimisen kautta. Tutkimuksessa havaittiin myös, että osassa yrityksiä tietopyyntöjen postittaminen oli ainoa vaihtoehto tietopyynnön lähettämiseen. Myös Ausloos ja Dewitte (2018) havaitsivat tutkimuksessaan, että osassa yrityksiä ei ollut muuta vaihtoehtoa kuin postittaa tietopyyntö yritykselle. Heidän tutkimuksessaan myös todettiin, että sähköinen lomake oli mielekkäin tapa tehdä tietopyyntö. Tässä tutkimuksessa sähköinen asiointipalvelu koettiin parhaaksi ja helpoimmaksi ratkaisuksi toteuttaa rekisteröidyn tietopyyntö. Sähköiseen asiointipalveluun on myös mahdollista rakentaa automaattinen tietojenhakuominaisuus, jolloin kummatkin osapuolet säästävät vaivaa tietopyyntöjen suhteen. Tutkimuksessa havaittiin, että vain yhdessä yrityksessä oli automatisoitu ratkaisu tietopyyntöjen käsittelemiseen ja tietojen hakemiseen. Ausloosin ja Dewitten (2018) mukaan automatisoidun ratkaisun kehittäminen voi vaatia jopa yrityksen koko järjestelmän uudistamista. Yrityksissä on mahdollisesti havaittu automaation tekeminen liian suurena investointina varsinkin, jos rekisteröidyn pyyntöjä on tullut suhteellisen vähän. Osa syy automatisoidun ratkaisun puuttumiseen voi olla datan luokittelun heikkous. Tutkimuksessa havaittiin, että yrityksissä tietopyynnöt käsiteltiin manuaalisesti, jotta voitiin tarkistaa, kohdistuuko rekisteröidyn tietoihin jokin muu käsittelyperuste. Datan luokittelu yhdistettynä automaattiseen tietojenhakuun mahdollistaisi vain sellaisten tietojen hakemisen, joihin ei kohdistu muuta käsittelyperustetta tai laillista säilytysvelvollisuutta.

Tutkimuksen perusteella yrityksissä on vaihtoehtoisia tapoja toimittaa tiedot rekisteröidylle. Tiedot voitiin toimittaa kirjepostilla, sähköisen asiointipalvelun tai salatun sähköpostin kautta. Havainto on linjassa Krögerin ja kollegoiden (2020) tutkimustulosten kanssa, sillä poikkeuksella, että heidän tutkimustulosten perusteella tiedot lähetettiin suojaamattoman sähköpostin kautta. Myös sähköisten tietojen tiedostomuoto vaihteli yrityksien kesken, mutta useimmissa yrityksissä tiedot toimitettiin PDF-muodossa. Saman havainnon olivat tehneet Kröger työryhmineen (2020) sekä Ausloos ja Dewitte (2018) omissa tutkimuksissaan. Tietojen lataaminen sähköisen asiointipalvelun kautta havaittiin olevan rekisteröidylle helpoin ja turvallisin ratkaisu, mutta tietojen suojaamisessa voisi olla vielä parannettavaa. Singh ja Cobbe (2019) ehdottavat tutkimuksessaan, että tiedot lähetettäisiin salatussa arkistossa, joka on suojattu rekisteröidyn itse määräämällä salasanalla. Tätä voitaisiin myös hyödyntää yrityksissä, jossa tiedot lähetettiin sähköisesti rekisteröidylle, jolloin rekisteröity voisi itse määrätä salasanan tietopyynnön yhteydessä. Näin myös estettäisiin henkilökohtaisten tietojen vuotaminen vääriin käsiin ja mahdollisesti vähennettäisiin identiteettivarkauksien riskiä. Ausloosin ja Dewitten (2018) tutkimuksessa myös havaittiin, että rekisteröidyn tiedot olivat ladattavissa vain tietyn aikaa tietopyynnön tekemisen jälkeen. Edellä mainitun yhdistäminen aiemmin mainittuihin suosituksiin lisäisi entisestään tietojen suojaa.

Tutkimuksessa selvisi, että rekisteröidyn tietojen siirto-oikeuden suhteen yleisesti hyväksytyksi ja koneluettavaksi muodoksi tunnistettiin XML-, CSV- ja

XLSX-tiedostomuoto. Myös Wong ja Henderson (2019) havaitsivat tutkimuksessaan, että yleisimpiin tiedostomuotoihin rekisteröidyn siirto-oikeuden kohdalla kuuluivat tässä tutkimuksessa havaitut tiedostomuodot. Wongin ja Hendersonin (2019) tutkimuksen perusteella voidaan myös todeta, että tässä tutkimuksessa havaitut tiedostomuodot täyttävät yleisen tietosuoja-asetuksen vaatimukset yleisesti hyväksyttyinä ja koneluettavina muotona.

Datanhallinnan todettiin olevan keskeisessä roolissa rekisteröidyn oikeuksia koskevien vaatimusten toteuttamisessa. Data governance auttoi asetuksen vaatimustenmukaisuudessa selkeyttämällä vastuita ja datan omistajuutta. Tutkimuksessa havaitut tietoturvastandardit (ISO/IEC27001 ja 27002) loivat pohjan asetuksen vaatimuksille. Lisäksi tässä tutkimuksessa havaittiin, että harva yritys oli sertifioitunut tietoturvastandardin mukaiseksi. Tämä havaittiin johtuvan toimialakohtaisten lakien vaatimuksista, jotka menivät osittain päällekkäin tietoturvastandardien vaatimusten kanssa, joten yrityksissä ei nähty kannattavaksi sertifioitua tietoturvastandardin mukaiseksi. Tietoturva- ja erillisten tietosuojapolitiikoiden avulla ohjeistettiin työntekijöitä henkilötietojen käsittelyyn liittyvistä asioista. Rekisteröidyn tietojen laadun varmistamiseen käytettiin Digi- ja väestöviraston tietokantaa sekä lisäksi rekisteröidyille viestittiin säännöllisesti tietojen tarkastamisesta. Tutkimuksessa perusteella havaittiin, että oikeellisuus, ajanmukaisuus ja sisällönkattavuus datan laadun ulottuvuuksina korostui yksittäisinä ulottuvuuksina asetuksen vaatimusten täyttämisessä. Edellä mainitut ulottuvuudet mahdollisesti korostuvat esimerkiksi tietopyyntöjen yhteydessä, sillä rekisteröidyn heikko laatuinen data voi johtaa tietojen toimittamisen väärälle henkilölle. Myös rekisteröidyn tunnistaminen voi olla haasteellista, jos tietojen laatu on heikko tasoista. Tutkimuksessa myös havaittiin, että keskitetyn master datan hallinta -järjestelmän avulla pystyttiin helpommin hakemaan ja päivittämään rekisteröidyn tietoja. Oletettavasti myös tietopyyntöihin vastaaminen on nopeampaa yhden keskitetyn master datan hallinta -järjestelmän avulla. Lisäksi tutkimuksessa havaittiin, että osassa yrityksiä oli erillisiä järjestelmiä, joita ei ollut integroitu yrityksen master datan hallinta -järjestelmään. Tähän havaittiin vaikuttavan toimialakohtaiset lait, jonka takia osaa rekisteröidyn tiedoista käsiteltiin erillisissä järjestelmissä.

Asetuksen mukaan työntekijöiden tulisi käsitellä vain tehtävän kannalta oleellisia henkilötietoja, johon yrityksissä oli vastattu roolipohjaisen pääsynhallinnan avulla. Roolipohjaisen pääsynhallinnan havaittiin olevan yleinen pääsynhallinnan ratkaisu yrityksissä. Suurin osa yrityksen datasta sijaitsee relaatiotietokantojen ulkopuolella, jonka takia strukturoimattoman datan hallinta on erityisen tärkeää. Asetuksen kontekstissa yritysten on tiedettävä, missä data sijaitsee ja mitä se pitää sisällään – etenkin, kun kyseessä on henkilötiedoiksi luokiteltavaa dataa. Tutkimuksessa havaittiin, että henkilötiedoille ei ollut relevanttia tehdä omaa turvaluokitusta, sillä monet dokumentit sisältävät henkilötietoja. Tällöin suurin osa dokumenteista luokiteltaisiin henkilötietojen turvaluokitukseen. Myöskään henkilötiedot turvaluokituksena ei ole tarpeeksi kuvaava, joten käyttäjän on haastavaa tietää, saako kyseisen turvaluokituksen dokumentteja jakaa esimerkiksi sisäisesti. Tutkimuksen perusteella voidaan todeta,

että strukturoimatonta dataa pyritään hallitsemaan ensisijaisesti ohjeistuksen avulla, mutta myös erilaisten teknisten ratkaisuiden avulla. Teknisillä ratkaisuilla voitiin muun muassa tunnistaa dokumenttien tietosisältö ja luokitella niitä sen mukaan. Koen, että teknisten ratkaisuiden avulla on mahdollista saada parempi kuva yrityksen hallussa olevasta datasta. Tietovarat on tunnistettava ennen kuin niitä voidaan hallita. Strukturoimattoman datan hallintaan voisi olla ratkaisuna kirjata dokumenttien metatietoihin, että kyseinen dokumentti sisältää henkilötietoa. Ratkaisu mahdollistaisi myös henkilötietojen etsimisen strukturoimattomasta datasta metatietojen perusteella.

Tutkimuksentulokset voidaan kokea olevan maltillisia, sillä samoja havaintoja oli tehty aiemmissa tutkimuksissa. Tuloksista voidaan kuitenkin havaita, miten rekisteröidyn oikeudet ovat toteutettu suomalaisissa tai Suomen maaperällä toimivissa yrityksissä. Mielenkiintoista oli huomata, että tutkimuksessa oli hyödynnetty datanhallinnan keinoja rekisteröidyn oikeuksien toteuttamisessa. Tutkimuksen perusteella voidaan todeta, että rekisteröidyn oikeuksien hyödyntäminen on otettu vakavasti ja pyritty tekemään mahdollisimman selkeäksi sekä helpoksi rekisteröidylle.

7.2 Reliabiliteetti ja validiteetti

Vaikka tutkimuksissa pyritään virheettömästi tutkimaan tutkittavaa ilmiötä, niin tästä huolimatta luotettavuus ja pätevyys vaihtelevat. Juuri tästä syystä tutkijan on aina arvioitava oman tutkimuksen luotettavuutta. (Hirsjärvi ym., 2004.)

Validiteetilla arvioidaan, kuinka hyvin mittaus- tai tutkimusmenetelmä mittaa sitä ilmiötä, jota sen on tarkoituskin mitata. Tutkimus ei aina mittaa sitä ilmiötä, jota tutkija olettaa. Esimerkiksi kyselytutkimuksessa vastaaja on saattanut käsittää kysymykset toisin kuin tutkija on ne laatiessaan ajatellut. Edellisessä esimerkissä tulokset eivät ole päteviä, mikäli tutkija on käsitellyt vastauksia alkuperäisen ajatusmallinsa mukaisesti. (Hirsjärvi ym., 2004)

Reliabiliteetilla tarkoitetaan kykyä tuottaa samoja tuloksia eli toisin sanoen tutkimus tai mittaustulokset ovat toistettavissa. Reliabiliteetti voidaan havaita usealla tavalla, mutta ehkä yleisin tapa on toistaa tutkimus eri tutkijoiden toimesta. Jos tutkijat päätyvät samaan tulokseen, voidaan tuloksia pitää luotettavana. (Hirsijärvi ym., 2004)

Kvalitatiivisessa tutkimuksessa luotettavuutta voidaan parantaa kertomalla tarkkaan tutkimuksen toteutus. Tämä sisältää tutkimuksen kaikki vaiheet. Tutkijan on kerrottava selkeästi aineiston keruu ja siihen liittyvät toimenpiteet, kuten haastattelun olosuhteet sekä paikat. Tämän lisäksi on tuotava ilmi haastatteluun käytetty aika, häiriötekijät, virhetulkinnat ja tutkijan omakohtainen arvio tilanteesta. Lukijalle on kerrottava selkeästi luokittelun taustat ja miten niihin on päädytty. Tutkijan tulee myös tulkita saatuja tuloksia ja esittää lukijalle mihin päätelmät perustuvat. Validiutta voidaan parantaa käyttämällä use-

ampaa tutkimusmenetelmää, jota kutsutaan triangulaatioksi. (Hirsijärvi ym., 2004.)

Tärkeintä tutkimuksessa on kohdentaa se oikeille henkilöille eli tässä tapauksessa tietosuojavastaaville tai asetuksen parissa työskenteleville henkilöille. Tutkimuksessa selvitettiin, miten rekisteröidyn oikeuksia koskevat vaatimukset ovat toteutettu yrityksissä ja, kuinka datanhallintaa on hyödynnetty niiden täyttämisessä. Tutkimuksen kannalta voidaan todeta, että tutkimuskohteet olivat oikeita henkilöitä vastaamaan tutkimuskysymykseen. Tutkittavilta vielä varmistettiin ennen haastatteluita, että ovatko he oikeita henkilöitä vastaamaan tutkittavaan aiheeseen. Olennaista validiteetin kannalta on, että haastateltavat kuuluvat perusjoukkoon eli tässä tapauksessa henkilöihin, jotka vastaavat tietoturva-asetuksen vaatimusten toteuttamisesta yrityksissä.

Tutkimuksen reliabiliteetin voidaan olettaa olevan hyvä, mikäli tutkimus on toistettavissa. Tutkimuksessa on kuitenkin huomioitava sen olevan sidoksissa aikaan eli esimerkiksi kahden vuoden päästä tilanne yleisen tietosuoja-asetuksen suhteen voi olla eri kuin tällä hetkellä. Tutkimukseen osallistuneet yritykset olivat kooltaan keskikokoisia tai suuria yrityksiä, joten tulokset eivät välttämättä päde pieniin yrityksiin. Pienissä yrityksissä ei välttämättä ole yhtä paljon resursseja toteuttaa rekisteröidyn oikeuksia yhtä tehokkaasti kuin suuremmissa yrityksissä. Tutkimus käsittelee laajasti tutkittavaa aihetta, joten kaikkiin tekijöihin ei voitu paneutua yksityiskohtaisesti. Luotettavimpia ja tarkempia tuloksia voitaisiin saada, kun keskityttäisiin yksityiskohtaisemmin esimerkiksi yhteen rekisteröidyn oikeuteen ja sen toteuttamiseen yrityksissä. Kuitenkin tarpeelliseksi nähtiin tutkia aihetta laajalla mittakaavalla, sillä aiempaa tutkimusta ei ole tehty rekisteröidyn oikeuksien toteuttamisesta ja datanhallinnasta hyödyntämisestä niissä. Tutkimuksessa on hyvä huomioida, että otoskoko ei ole suuri, joka saattaa heikentää tutkimuksen luotettavuutta. Tästä huolimatta sen koettiin olevan riittävän suuri tulosten analysoinnin ja johtopäätösten tekemiselle. Tutkimuksen tarkoitus oli kartoitta, miten rekisteröidyn oikeudet on toteutettu yrityksissä, jolloin otoskoko voidaan todeta riittäväksi tässä kontekstissa. Suurempi otos luultavasti antaisi luotettavampia tuloksia aiheesta.

Tutkimuksen validiteetti varmistettiin sillä, että haastattelutilanteessa osallistujat ymmärsivät kysymykset oikein ja, että kysymykset olivat esitetty selkeästi ja ymmärrettävästi. Myös lyhyellä johdatuksella voitiin ennen itse tutkimusta varmistaa, että osallistujat ymmärsivät aiheen. Tutkittaville annettiin haastattelurunko ennalta, jotta osallistujat pystyivät halutessaan tutustumaan kysymyksiin ennen haastattelua. Kuitenkin on mahdollista, että haastateltavat ovat antaneet suotuisampia vastauksia, kuin todellisuudessa, joka vääristää tutkimuksen tuloksia. Ottamalla huomioon edellä mainitut seikat voidaan olettaa, että tutkimus on onnistunut ja olevan myös validi ja reliaabeli. Tutkimuksessa saatiin vastaus myös tutkimusongelmaan ja sen odotetaan tuovan arvoa niin yrityksille kuin tieteenalallekin.

7.3 Tulosten hyödyntäminen ja jatkotutkimus

Vaikka tutkimuksen otos ei ole kovin suuri, niin tulokset ovat silti hyödynnettävissä. Tulokset antavat hyvän kuvan, miten rekisteröidyn oikeudet on toteutettu Suomessa toimivissa yrityksissä. Tuloksia voidaan hyödyntää erityisesti uusissa yrityksissä, jotka eivät vielä ole toteuttaneet asetuksen vaatimuksia. Toisaalta tutkimustuloksia voidaan myös hyödyntää yrityksissä, jotka haluavat parantaa käytänteitään koskien rekisteröidyn oikeuksien toteuttamista. Organisaatioiden on mahdollista hyödyntää tutkimukseen osallistuneiden yritysten tapoja toteuttaa rekisteröidyn oikeudet sekä johtopäätöksistä johdettuja näkemyksiä oikeuksien toteuttamiseen. Tutkimus luo hyvän käsityksen datanhallinnan osa-alueista, jotka on hyvä huomioida asetuksen kontekstissa. Organisaatioiden on näin ollen myös mahdollista pohtia datanhallinnan ratkaisuita, joita ottaa käyttöön valmistautuessa asetuksen vaatimuksiin. Lisäksi tutkimustulokset lisäävät tieteellistä kontribuutiota aiheesta. Tutkimusta, jossa yhdistyy yleinen tietosuoja-asetus ja datanhallinta, on suhteellisen vähän, joten tulokset tuovat uutta tutkimusta aihealueesta. Aihealue on kuitenkin otollinen jatkotutkimukselle.

Tutkimus on aihealueeltaan laaja, joten jatkotutkimuksena olisi mielenkiintoista tutkia yksittäisten rekisteröidyn oikeuksien tai tietopyyntöjen toteuttamista yrityksissä. Tutkimukseen osallistuneet yritykset ovat keskimäärin suuria yrityksiä, joten olisi kiinnostavaa selvittää, miten rekisteröidyn oikeudet on toteutettu pienemmissä yrityksissä, jossa resursseja ei ole yhtä paljon käytettävissä kuin suurissa yrityksissä. Jatkotutkimusaiheita voidaan tarkastella myös datanhallinnan näkökulmasta, jossa olisi mielenkiintoista selvittää tarkemmin, miten master datan hallinnalla voidaan helpottaa rekisteröidyn oikeuksien täyttämistä. Lisäksi jatkossa olisi mielekästä myös selvittää metatietojen hyödyntämistä asetuksen kontekstissa. Toisaalta tutkimus olisi myös mielekästä toistaa tulevaisuudessa ja selvittää, onko rekisteröidyn oikeudet toteutettu eri tavalla ja onko datanhallinta korostunut entisestään ratkaisuiden toteutuksessa.

8 YHTEENVETO

Yrityksien toimintaa ohjaa entistä enemmän tietosuojaan liittyvät lait ja asetukset, mutta myös kuluttajat ovat tietoisempia omista oikeuksistaan. Euroopan unionin yleisen tietosuoja-asetuksen tavoitteena on lisätä rekisteröidyn oikeuksia ja yhtenäistää lainsäädäntöä sekä mahdollistaa tietojen vapaata liikkuvuutta Euroopan sisällä. Rekisteröidyn on mahdollista hallita yksityisyyttään ja omia tietojaan paremmin lisääntyneiden oikeuksien kautta. Toisaalta asetuksen vaatimusten noudattaminen voi kuitenkin olla haastavaa yrityksille, sillä sen viralllisessa tekstissä ei oteta kantaa, miten vaatimukset tulisi käytännössä toteuttaa.

Tämän pro gradu -tutkielman teoriaosuudessa vastattiin seuraaviin tutkimusongelmiin:

- Euroopan unionin yleisen tietosuoja-asetuksen keskeisimmät vaatimukset rekisteröidyn henkilön näkökulmasta ja niiden toteuttaminen yrityksissä.
- Miten datanhallinta helpottaa täyttämään rekisteröidyn henkilön oikeuksia koskevat vaatimukset?

Teoriaosion toisessa luvussa etsittiin vastausta ensimmäiseen tutkimusongelmaan liittyen yleisen tietosuoja-asetuksen keskeisimpiin vaatimuksiin rekisteröidyn näkökulmasta ja niiden toteuttamiseen yrityksissä. Euroopan unionin yleisen tietosuoja-asetuksen mukaan rekisteröidyn oikeuksiin sisältyy oikeus saada läpinäkyvää informaatiota henkilötietojen käsittelystä ja saada pääsy omiin tietoihin, oikeus tietojen oikaisuun, poistamiseen, tietojen siirtämiseen järjestelmästä toiseen sekä käsittelyn rajoittamiseen ja vastustamiseen. Rekisteröidyllä on myös oikeus olla joutumatta automaattisen päätöksenteon ja profiloinnin kohteeksi. Rekisteröidyn oikeuksia koskevat vaatimukset olivat toteutettu monellakin eri tapaa yrityksissä. Aiempien tutkimusten mukaan tietosuojaselosteet sisälsivät informaatiota rekisteröidyn tietojen käsittelystä, mutta olivat useassa yrityksessä puutteellisia, epäselkeitä ja vaikeasti löydettävissä yrityksen sivuilta. Tutkimuksissa myös selvisi, että rekisteröity voitiin tunnistaa henkilötodistuksen kopion, kirjautumistietojen, sähköpostin, kotiosoitteen, puhelinsoiton tai muun käyttäjäkohtaisen tiedon perusteella. Tietopyyntöjen lä-

hettäminen tapahtui joko kirjepostilla tai sähköisen lomakkeen kautta. Tiedot toimitettiin rekisteröidylle sähköpostin, kirjepostin tai applikaation kautta. Siirto-oikeutta koskevaksi yleisesti hyväksytyksi ja koneluettavaksi muodoksi oli tunnistettu CSV-, XLS-, XLSX-, HTML-, PDF-, JSON-, XML-, TXT-, DOC-, DOCX-, RTF-tiedostomuodot.

Toiseen tutkimusongelmaan vastattiin tutkimuksen kolmannessa luvussa ja neljännessä luvussa. Tutkimuksessa katsottiin data governancen, datan laadun-, master datan-, datan tietoturvan- ja dokumenttien sekä sisällönhallinnan keskeisimmiksi datanhallinnan tavoiksi vastata rekisteröidyn oikeuksiin koskeviin vaatimuksiin. Yleisen tietosuoja-asetuksen vaatimukset nähtiin olevan osa hyvää datanhallintaa. Yritykset myös ottavat käyttöön datanhallinnan tapoja vastatakseen eri asetusten vaatimuksiin. Data governancen katsottiin olevan keskeinen toiminto, jolla voidaan edesauttaa asetuksen vaatimustenmukaisuutta luomalla selkeät vastuut dataa koskevista päätöksistä. Master datan hallinnan tunnistettiin olevan merkittävässä roolissa rekisteröidyn oikeuksien, kuten tietojen poistamisen ja niihin pääsyn kannalta. Hyvän datan laadun hallinnalla taas pystyttiin täyttämään asetuksen vaatimukset rekisteröidyn tietojen oikeellisuudesta. Datan tietoturvan hallinnan avulla voitiin suojata yrityksen tietovoimavarat, mutta myös turvata henkilötietojen käsittely. Yhdeksi haasteeksi tietosuoja-asetuksen vaatimustenmukaisuudessa havaittiin strukturoimaton data ja sen hallinta, sillä valtaosa yrityksen datasta sijaitsee relaatiotietokantojen ulkopuolella.

Luvussa kuusi etsittiin vastausta empiirisen tutkimuksen kautta päätutkimusongelmaan: Miten yleisen tietosuoja-asetuksen rekisteröidyn oikeuksia koskevat vaatimukset ovat toteutettu yrityksissä ja, kuinka datanhallintaa on hyödynnetty niiden täyttämisessä? Empiirisessä tutkimuksessa havaittiin, että yrityksissä on toteutettu rekisteröidyn oikeuksia koskevat vaatimukset eri tavoilla. Yleisesti kuvaus henkilötietojen käsittelystä oli sisällytetty yrityksen tietosuojapolitiikkaan, joka oli löydettävissä yrityksen sivuilta. Tietopyyntö oli mahdollista tehdä sähköisen asiointipalvelun, sähköisen PDF-lomakkeen postittamisen tai fyysisessä toimipisteessä asioimisen kautta. Rekisteröity voitiin tunnistaa tietopyyntöjen yhteydessä sähköisen vahvan tunnistautumisen, henkilötodistuksen, kirjautumistietojen sekä allekirjoituksen perusteella. Suurimmassa osassa yrityksiä rekisteröidyn tietopyynnöt käsiteltiin manuaalisesti. Tiedot toimitettiin rekisteröidylle kirjepostilla, sähköisen asiointipalvelun tai salatun sähköpostin kautta. Jäljennös rekisteröidyn tiedoista tarjottiin PDF- tai XLSX-muodossa ja siirto-oikeutta koskevissa tiedoissa XML-, CSV- ja XLSX-muodossa. Datanhallinnalla nähtiin olevan keskeinen rooli yleisen tietosuoja-asetuksen vaatimusten täyttämisessä, jossa data governance selkeytti vastuita ja datan omistajuutta. Rekisteröidyn tietoja hallittiin osassa yrityksiä yhden keskitetyn master data -järjestelmän kautta, joka helpotti rekisteröityjen henkilöiden tietojen hakemista ja päivittämistä. Datan laadun ulottuvuudet eivät korostuneet selkeästi yksittäisinä ulottuvuuksina vaan näkyivät ennemmin kokonaisuutena. Tutkimuksessa havaittiin, että yrityksissä käytettiin Digi- ja väestötietoviraston tietokantaa rekisteröidyn perustietojen varmistamiseen. Tutkittavissa yrityksis-

sä harvalla oli käytössä tietoturvastandardia. Tietoturvapolitiikat taas vastaavasti olivat käytössä jokaisessa yrityksessä, ja lisäksi osassa oli erillinen tietosuojapolitiikka. Yrityksissä käytettiin roolipohjaista pääsynhallintaa rajoittamaan tietojen käsittelyä vain tehtävän kannalta oleellisimpiin tietoihin. Struktuuroimattoman dataa hallittiin ensisijaisesti ohjeistuksen avulla, mutta osassa yrityksiä oli käytössä myös teknisiä ratkaisuita siihen. Kuitenkaan henkilötiedoille ei ollut omaa turvaluokitusta vaan ne sisältyivät arkaluontoisuutensa mukaan osaksi laajempia turvaluokituksia. Johtopäätöksenä voidaan todeta, että datanhallintaa on hyödynnetty asetuksen vaatimusten täyttämisessä ja se mahdollistaa tehokkaan tavan vastata niihin.

Tutkimuksen aiheesta, jossa yhdistyy rekisteröidyn oikeudet ja datanhallinta, ei ole aikaisempaa tutkimusta. Kuitenkin tutkimustulokset rekisteröidyn oikeuksien osalta ovat linjassa aiempien tutkimusten kanssa. Tutkimukseen liittyy rajoitteita, jotka on huomioitava arvioitaessa tutkimuksen tuloksia. Ensinnäkin tutkimuksen otoskoko ei ollut kovin suuri, joten tulokset eivät ole välttämättä yleistettävissä. Lisäksi tutkimus käsittelee aihetta hyvin laajalta näkökulmalta ja näin ollen vaatisi tarkempaa tarkastelua yksittäisten osa-alueiden kautta.

Tutkimuksen pohjalta jatkotutkimusaiheiksi esitetään yksityiskohtaisempaa tarkastelua jonkun tietyn rekisteröidyn oikeuden toteuttamisesta. Mielenkiintoista olisi myös tutkia, miten rekisteröidyn oikeudet on toteutettu pienissä yrityksissä. Jatkotutkimuksessa voitaisiin myös keskittyä johonkin tiettyyn datanhallinnan osa-alueeseen, kuten master datan tai metatietojen hallintaan ja selvittää, kuinka niiden avulla voidaan vastata asetuksen vaatimuksiin tehokkaammin. Toisaalta tutkimus olisi myös mielenkiintoista toistaa tulevaisuudessa, jotta voitaisiin havaita, onko yritykset toteuttaneet rekisteröidyn oikeudet eri tavalla aikaisempaan.

LÄHTEET

Adda, M., Abdelaziz, J., McHeick, H., & Saad, R. (2015). Toward an access control model for IOTCollab. *Procedia Computer Science*, *52*(1), 428–435.

Aiken, P., Allen, M. D., Parker, B., & Mattia, A. (2007). Measuring data management practice maturity: A community's self-assessment. *Computer*, *40*(4), 42–50.

Almeida Teixeira, G., Mira da Silva, M., & Pereira, R. (2019). The critical success factors of GDPR implementation: a systematic literature review. *Digital Policy, Regulation and Governance* , *21*(4), 402–418.

Alshammari, M., & Simpson, A. (2018). Personal data management: An abstract personal data lifecycle model. *Lecture Notes in Business Information Processing*, *308*, 685–697.

Ausloos, J., & Dewitte, P. (2018). Shattering one-way mirrors - data subject access rights in practice. *International Data Privacy Law*, *8*(1), 4–28.

Bisnode. Tietosuoja-asetus ja datan laatu. Haettu 15.2.2021 osoitteesta: https://finland.bisnode.fi/syvenna-osaamistasi/ajatuksiamme/datan-laadun-vaikutus-gdpr-vaatimustenmukaisuuteen/

Boniface, C., Fouad, I., Bielova, N., Lauradoux, C., & Santos, C. (2019). Security Analysis of Subject Access Request Procedures: How to Authenticate Data Subjects Safely When They Request for Their Data. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11498 LNCS*, 182–209.

Cichy, C., & Rass, S. (2019). An Overview of Data Quality Frameworks. *IEEE Access*, *7*, 24634–24648.

Data Governance for GDPR Compliance: Principles , Processes, and Practices. (2017). *Microsoft*.

Degeling, M., Utz, C., Lentzsch, C., Hosseini, H., Schaub, F., & Holz, T. (2018). We Value Your Privacy ... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy. *arXiv*, (February).

Eskola, J. & Suoranta, J. (1998). Johdatus laadulliseen tutkimukseen. *Vastapaino.*

Eurobarometri. Haettu 6.4.2020 osoitteesta: http://ec.europa.eu/public_opinion/archives/ebs/ebs_431_en.pdf

Euroopan unionin virallinen lehti. Haettu 07.4.2020 osoitteesta: https://eur-lex.europa.eu/legal-content/FI/TXT/PDF/?uri=OJ:L:2016:119:FULL&from=FI

Forrester. (2019). *Data Privacy Is The New Strategic Priority*.

Garber, J. (2018). GDPR – compliance nightmare or business opportunity? *Computer Fraud and Security*, *2018*(6), 14–15.

Gartner. (2018). How to Create a Business Case for Data Quality Improvement Haettu 13.12.2020 osoitteesta: https://www.gartner.com/smarterwithgartner/how-to-create-a-business-case-for-data-quality-improvement/

Gilbert, F. (2012). European Data Protection 2.0: New Compliance Requirements in Sight-What the Proposed EU Data Protection Regulation Means for US Companies. *Santa Clara Computer & High Tech. LJ*, *28*(4), 815–863.

Greengard, S. (2018). Weighing the impact of GDPR. *Communications of the ACM*, *61*(11), 16–18.

Haug, A., & Arlbjørn, J. S. (2011). Barriers to master data quality. *Journal of Enterprise Information Management*, *24*(3), 288–303.

Hirsjärvi, S., Remes, P., & Sajavaara, P. (2004). Tutki ja kirjoita. 10. uusittu painos. *Tammi: Helsinki*.

Hoiry, L., & Norris, C. (2015). *The Honest Data Protection Officer ' s Guide to Enable Citizens to exercise their Subject Access Rights – Lessons from a Ten Country European Study*. (October 1995), 31–50.

Jia, J., Jin, G. Z., & Wagman, L. (2019). The Short-Run Effects of GDPR on Technology Venture Investment. *SSRN Electronic Journal*, (November).

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, *53*(1), 148–152. https://doi.org/10.1145/1629175.1629210

Knolmayer, G. F., & Röthlin, M. (2006, November). Quality of material master data and its effect on the usefulness of distributed ERP systems. In *International Conference on Conceptual Modeling* (pp. 362-371). Springer, Berlin, Heidelberg.

KPMG. (2017). Disrupt and grow. *Kpmg*. Noudettu osoitteesta https://home.kpmg.com/content/dam/kpmg/uk/pdf/2017/06/2017-ceo-outlook.pdf

Kröger, J. L., Lindemann, J., & Herrmann, D. (2020). How do app vendors respond to subject access requests?: A longitudinal privacy study on iOS

and Android Apps. *ACM International Conference Proceeding Series*.

Laybats, C., & Davies, J. (2018). GDPR: Implementing the regulations. *Business Information Review*, *35*(2), 81–83.

Lee, J. (2016). What the Right to be Forgotten Means to Companies: Threat or Opportunity? *Procedia Computer Science*, *91*(Itqm), 542–546.

Madnick, S., Wang, R., & Xian, X. (2003). The Design and Implementation of a Corporate Householding Knowledge Processor to Improve Data Quality. *Journal of Management Information Systems*, *20*(3), 41–70.

Marsh, & Microsoft. (2019). 2019 Global Cyber Risk Perception Survey. *Microsoft Insights*, (September), 1–36. Haettu 1.3.2021 osoitteesta https://www.microsoft.com/security/blog/wp-content/uploads/2019/09/Marsh-Microsoft-2019-Global-Cyber-Risk-Perception-Survey.pdf

Marsh, R. (2005). Drowning in dirty data? It's time to sink or swim: A four-stage methodology for total data quality management. *Journal of Database Marketing & Customer Strategy Management*, *12*(2), 105–112.

Martino, M. Di, Robyns, P., Weyts, W., Quax, P., Lamotte, W., & Andries, K. (2019). Personal information leakage by abusing the GDPR "right of access". *Proceedings of the 15th Symposium on Usable Privacy and Security, SOUPS 2019*, 371–386.

Microsoft. (2017). Data Governance for GDPR Compliance: Principles, Processes, and Practices Haettu 15.2.2021 osoitteesta: https://info.microsoft.com/DataGovernanceforGDPRCompliancePrinciplesProcessesandPractices-Registration.html

Mosley, M., Brackett, M., Earley, S., & Henderson, D. (2009). *The DAMA Guide to the Data Management Body of Knowledge (DAMA-DMBOK) Spanish Edition*.

Osborn, S. (1997). Mandatory access control and role-based access control revisited. *Proceedings of the ACM Workshop on Role-Based Access Control*, 31–40.

Otto, B. (2012). How to design the master data architecture: Findings from a case study at Bosch. *International Journal of Information Management*, *32*(4), 337–346.

Otto, B., & Reichert, A. (2010). Organizing master data management: Findings from an expert survey. *Proceedings of the ACM Symposium on Applied Computing*, 106–110.

Panhila, S., Siponen, M., & Adam Mahmood. (2007). Employees' Behavior

towards IS Security Policy Compliance. *Proceedings of the 40th Hawaii International Conference on System Sciences*.

Panian, Z. (2010). Some practical experiences in data governance. *World Academy of Science, Engineering and Technology*, *38*, 150–157.

Ray, I., & Toahchoodee, M. (2007). *A Spatio-temporal Role-Based Access Control Model*. 211–226.

Rifaie, M., Yorkshire, W., Ridley, M., & Yorkshire, W. (2009). *Data Governance Strategy : A Key Issue in Building Enterprise Data Warehouse This paper articulates data governance as one of the key is- sue in building Enterprise Data Warehouse . The key goals of this document are to : de ne the strategy for Data Gov-* . 587–591.

Rowley, J. (2007). The wisdom hierarchy: Representations of the DIKW hierarchy. *Journal of Information Science*, *33*(2), 163–180.

Ryz, L., & Grest, L. (2016). A new era in data protection. *Computer Fraud and Security*, *2016*(3), 18–20.

Sammon, D., Adam, F., Nagle, T., & Carlsson, S. (2010). Making sense of the master data management (MDM) concept: Old wine in new bottles or new wine in old bottles? *Frontiers in Artificial Intelligence and Applications*, *212*, 175–186.

Sarajärvi, A., & Tuomi, J. (2009). Laadullinen tutkimus ja sisällönanalyysi. 2009. *Vantaa: Tammi.*

Seo, J., Kim, K., Park, M., Park, M., & Lee, K. (2018). An analysis of economic impact on IoT industry under GDPR. *Mobile Information Systems*, *2018*.

Shaikh, R., & Sasikumar, M. (2015). Data classification for achieving security in cloud computing. *Procedia Computer Science*, *45*(C), 493–498.

Shao, X., & Oinas-Kukkonen, H. (2019). How Does GDPR (General Data Protection Regulation) Affect Persuasive System Design: Design Requirements and Cost Implications. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11433 LNCS*, 168–173.

Singh, J., & Cobbe, J. (2019). The Security Implications of Data Subject Rights. *IEEE Security and Privacy*, *17*(6), 21–30.

Siponen, M. (2006). Information security standards focus on the existence of process, not its content. Communications of the ACM, 49(8), 97-100.

Smith, H. A., & McKeen, J. D. (2008). Developments in Practice XXX: Master

Data Management: Salvation Or Snake Oil? *Communications of the Association for Information Systems*, *23*.

Soares, S. (2010*). The IBM data governance unified process: driving business value with IBM software and best practices*. MC Press, LLC.

Suduc, A.-M. (2010). *Audit for Information Systems Security | Paper | Microsoft Academic*. *14*(1), 43–48.

Tankard, C. (2016). What the GDPR means for businesses. *Network Security*, *2016*(6), 5–8.

Team, I. P. (2017). *EU general data protection regulation (GDPR): An implementation and compliance guide* IT Governance Ltd.

The Guardian. (2014). New Digital Universe Study Reveals Big Data Gap. Haettu 5.11.2020 osoitteesta: https://www.theguardian.com/news/datablog/2012/dec/19/big-data-study-digital-universe-global-volume

Tietosuojavaltuutetun toimisto. (2021). Rekisteröidyn yleisen tietosuoja-asetuksen mukaiset oikeudet ym.. Rekisterinpitäjä ParkkiPate Oy. Haettu 17.5.2021 osoitteesta https://tietosuoja.fi/-/yritykselle-seuraamusmaksu-tietosuojarikkomuksista-pysakoinninvalvontamaksujen-yhteydessa

Tikkinen-Piri, C., Rohunen, A., & Markkula, J. (2018). EU General Data Protection Regulation: Changes and implications for personal data collecting companies. *Computer Law and Security Review*, *34*(1), 134–153.

Traficom. (2019). Luottamuksen lähteillä: Näkökulmia tietoturvan standardointiin ja sertifiointiin. Haettu 20.5.2021 osoitteesta: https://www.kyberturvallisuuskeskus.fi/sites/default/files/media/publication/Luottamuksen_lahteilla.pdf

Traficom. (2019). Suojautuminen Microsoft Office 365 -tunnusten kalastelulta ja tietomurroilta. Haettu 13.3.2021 osoitteesta: https://www.kyberturvallisuuskeskus.fi/sites/default/files/media/publication/Suojautuminen%20Microsoft%20Office%20365%20-tunnusten%20kalastelulta%20ja%20tietomurroilta%20web.pdf

Tsohou, A., Kokolakis, S., Lambrinoudakis, C., & Gritzalis, S. (2010). A security standards' framework to facilitate best practices' awareness and conformity. *Information Management and Computer Security*, *18*(5), 350–365.

Tuyikeze, T., & Pottas, D. (2010). An information security policy development life cycle. *Proceedings of the South African Information Security Multi-Conference, SAISMC 2010*, (January 2020), 165–176.

Villaronga, E. F., Kieseberg, P., & Li, T. (2018). Humans forget, machines remember: Artificial intelligence and the Right to Be Forgotten. *Computer Law and Security Review*, *34*(2), 304–313.

Väre, T. (2019). Master data. Helsinki: Alma Talent.

Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, *12*(4), 5–33.

White, A., Newman, D., Logan, D., & Radcliffe, J. (2006). Mastering master data management. *Gartner Group, Stamford*.

Wong, J., & Henderson, T. (2019). The right to data portability in practice: Exploring the implications of the technologically neutral GDPR. *International Data Privacy Law*, *9*(3), 173–191.

LIITE 1 TEEMAHAASTATTELUN RUNKO

Taustatiedot

1. Mikä on työnimikkeesi ja -tehtäväsi?
2. Kuinka monta vuotta sinulla on kokemusta tietosuojan ja yleisen tietosuoja-asetuksen parissa työskentelemisestä?
3. Millä toimialalla yrityksesi toimii?
4. Mitä tietoja yrityksessäsi kerätään rekisteröidyistä henkilöistä?

Rekisteröidyn oikeudet

5. Millä tavalla rekisteröidylle kerrotaan henkilötietojen käsittelystä?
   - Arvioidaanko tietosuojaselostetta/-ilmoitusta säännöllisesti?
   - Missä kohtaa tietosuojaseloste/-ilmoitus tarjotaan rekisteröidylle?
   - Missä muodossa se tarjotaan?
6. Onko rekisteröidyn pyyntöjä varten olemassa lomaketta yrityksen sivuilla?
   - Mikäli ei ole, niin mitä kautta rekisteröidyn pyyntöjä käsitellään?
   - Onko kaikille pyynnöille oma lomakkeensa?
7. Miten henkilö tunnistetaan pyyntöjä esittäessä?
8. Miten ja kuka käsittelee rekisteröidyn pyynnöt?
9. Miten tiedot toimitetaan rekisteröidylle?
10. Pääseekö rekisteröity käsiksi kaikkeen hänestä kerättyihin tietoihin kirjautumalla yrityksen sivuille?
    - Voiko omien sivujen kautta muokata ja oikaista virheelliset tiedot?
    - Päivittyykö muutetut tiedot automaattisesti muihin järjestelmiin?
11. Jos rekisteröity pyytää tietojen poistamista, poistetaanko ne automaattisesti vai manuaalisesti?
    - Jos poistetaan automaattisesti, niin miten varmistetaan, että tiedot ovat poistuneet?
12. Onko rekisteröidyn mahdollista pyytää tietojen siirtoa järjestelmästä toiseen?
    - Miten se on toteutettu?
    - Missä muodossa tiedot siirretään toiseen järjestelmään?

Datanhallinta

13. Millainen rooli datanhallinnalla on yrityksessäsi yleisen tietosuoja-asetuksen vaatimusten täyttämisessä?
14. Onko yrityksessäsi datan hallintamallia tai -ohjelmaa (engl. data governance), joka määrää vastuut dataa koskevista päätöksistä?
    - Miten se on auttanut asetuksen vaatimustenmukaisuudessa?

15. Mikä/Mitkä datan laadun ulottuvuuksista ovat erityisesti korostuneet yrityksessäsi yleisen tietosuoja-asetuksen vaatimusten täyttämisessä? (ks. taulukko)
    - Miten ja mitkä datan laadun ulottuvuudet korostuvat erityisesti asetuksen kontekstissa?
    - Miten yrityksessäsi varmistetaan rekisteröidyn tietojen laatu?

TAULUKKO 11 Tutkimuksen teoriaosiossa käsitellyt datan laadun ulottuvuudet

Datan laadun ulottuvuudet	**Kuvaus**	**Lähde**
Oikeellisuus (engl. accuracy)	Kuvaa datan oikeellisuutta eli missä määrin kirjattu arvo vastaa todellisuutta.	*Brackett ym., 2009; Khatri & Brown, 2010; Cichy & Rass, 2019*
Ajanmukaisuus (engl. timeliness)	Data on ajan tasalla ja saatavilla, kun sitä tarvitaan.	*Khatri & Brown, 2010, Cichy & Rass, 2019*
Sisällön kattavuus (engl. completeness)	Data ei ole puutteellista, vaan on riittävällä tasolla tehtävän suorittamiseen.	*Brackett ym., 2009; Khatri & Brown, 2010; Cichy & Rass, 2019*
Uskottavuus (engl. credibility)	Datan sisältö ja lähde on luotettavaa.	*Khatri & Brown, 2010*
Johdonmukaisuus (engl. consistency)	Data on muodoltaan yhtenäistä muun tiedon kanssa.	*Cichy & Rass, 2019*
Esteettömyys (engl. accessibility)	Data on esteettömästi saatavilla, kun sitä tarvitaan tai vähintään nopeasti ja helposti haettavissa.	*Cichy & Rass, 2019*
Vaatimustenmukaisuus (engl. validity)	Data on sekä muodollisesti, että sisällöllisesti vaatimustenmukaista.	*Brackett ym., 2009; Väre, 2019*
Eheys (engl. integrity)	Datan eri järjestelmien keskinäiset viittaukset ovat yhdistettävissä.	*Brackett ym., 2009; Väre, 2012*
Yksityisyys (engl. privacy)	Pääsyn- ja kulunvalvonta tiettyjen datajoukkojen osalta.	*Brackett ym., 2009*
Kohtuullisuus (engl. reasonableness)	Kohtuullisten odotuksien määrittäminen toimintaympäristön kontekstissa.	*Brackett ym., 2009*
Ainutlaatuisuus (engl. uniqueness)	Yksi data entiteetti tulisi esiintyä vain kerran datajoukossa.	*Brackett ym., 2009*

16. Miten yrityksessäsi on järjestetty henkilötietoihin liittyvä ydintieto? (engl. master data management)?
    - Miten ydintiedon hallinta on auttanut yleisen tietosuoja-asetuksen vaatimusten täyttämisessä?
17. Onko yrityksessä tietoturvapolitiikka, jossa ohjeistetaan henkilötietojen käsittelystä?
    - Arvioidaanko ja päivitetäänkö sitä säännöllisesti?
18. Noudattaako yrityksen toiminta jotakin tietoturvastandardia?

- Onko standardista ollut hyötyä yleisen tietosuoja-asetuksen suhteen?

19. Onko henkilötietoja sisältäville dokumenteille oma turvaluokitus?
20. Tunnistetaanko rekisteröidyn henkilötiedot strukturoimattomasta datasta? Miten?
21. Rajoitetaanko työntekijöiden pääsyä rekisteröidyn henkilön tietoihin roolipohjaisen pääsynhallinnan avulla?
22. Mitkä ovat mielestäsi tärkeimmät datanhallinnan tavat, joilla voidaan yrityksissä vastata henkilötietojen käsittelyyn liittyvien asetusten vaatimuksiin tulevaisuudessa?