Haaga-Helia
ammattikorkeakoulu Oy

Lokien hallinta osana tiedonhallintalain vaatimuksia

Minna Ryyppö

Opinnäytetyö
Tietojenkäsittelyn koulutusohjelma
2020

Haaga-Helia
ammattikorkeakoulu Oy

Tiivistelmä

Tekijä(t)
Minna Ryyppö

Koulutusohjelma
Tietojenkäsittelyn koulutusohjelma

Raportin/Opinnäytetyön nimi
Lokien hallinta osana tiedonhallintalain vaatimuksia

Sivu- ja liitesivumäärä
28 + 2

Tässä tutkimuksessa selvitetään mitä vaatimuksia laki julkisen hallinnon tiedonhallinnasta (906/2019) asettaa lokien hallinnalle. Tarkoituksena on lain ja siihen liittyvien perustelujen ja suositusten valossa selvittää, mitä laissa esitetty lokien keräämisen vaatimus käytännössä tarkoittaa viranomaisorganisaatiolle. Lisäksi lokeja ja lokien hallintaa käydään läpi yleisellä tasolla. Tutkimuksen empiriaosassa käydään läpi Museoviraston lokien hallinnan nykytilaa sen kartoittamiseksi, mitä toimenpiteitä lain vaatimuksen täyttäminen vielä vaatii. Tutkimuksen perusteella tehdään toimenpidesuosituksia lokien hallinnan kehittämiseksi.

Lokeja on paljon erityyppisiä riippuen käyttötarkoituksesta. Esimerkkejä erilaisista lokeista ovat mm. muutoslokit, käyttölokit, pääsynhallintalokit, virhelokit, yhteyslokit jne. Kaikille lokeille yhteistä on se, että niiden tarkoituksena on varmistaa normaalia toimintaa ja havaita poikkeamia.

Lokien hallinta tarkoittaa suunnitelmallisesti toteutettua prosessia, joka etenee lokien keräämisestä lokitiedon käsittelyyn, säilyttämiseen ja lopulta lokitiedon poistamiseen. Lokien hallintaa organisaatiossa ohjaavat yhteisesti sovitut lokiperiaatteet sekä lainsäädäntö. Julkishallinnon organisaation kannalta keskeisimmät lait ovat tiedonhallintalain lisäksi EU:n tietosuoja-asetus sekä julkisuuslaki. Erityisaloilla voi myös olla omaa sääntelyä lokitiedoista, kuten sosiaali- ja terveysalalla.

Tiedonhallintalain 17 § mukaan lokeja tulee kerätä järjestelmän käytöstä ja tiedonluovutuksista aina, jos järjestelmä edellyttää kirjautumista. Tiedonhallintalautakunnan suositus, samoin kuin lain valmistelumateriaali, antaa kuitenkin ymmärtää, että lokien keräämisen ei ole syytä ryhtyä näin suoraviivaisesti vaan mieluummin tarpeeseen perustuen. Lisäksi mainituissa lähteissä korostetaan, että lokien keräämisen kannalta oleellisinta on, jos järjestelmässä on tai siitä luovutetaan henkilö- tai salassa pidettäviä tietoja.

Museovirastossa on käytössä noin 30 sellaista tietojärjestelmää, jotka edellyttävät kirjautumista. Tarkempi tarkastelu kuitenkin osoittaa, että vain osassa näistä on oleellista toteuttaa lokien keruuta tiedonhallintalain tarkoittamalla tavalla. Erilaisten lokien keräämiseen voi silti olla joitain muita tarpeita. Lokien hallinnan nykytilanne osoittaa, että lokeihin ei ole aikaisemmin järjestelmällisesti kiinnitetty huomiota. Tutkimuksen perusteella ehdotetaan, että Museovirasto laatisi lokiperiaatteet ja arvioisi lokien keräämisen tarpeen järjestelmäkohtaisesti.

Tiedonhallintalain voimaantulo varmastikin kiinnittää julkishallinnon organisaatioiden huomion lokien hallintaan aiempaa enemmän. Tämä on erittäin hyödyllinen kehitysaskel ajatellen koko yhteiskunnan tietojen tietosuojaa.

Asiasanat
lokitiedostot, tiedonhallinta, tietosuoja

Sisällys

1 Johdanto

Vuoden 2020 alussa astui voimaan laki julkisen hallinnon tiedonhallinnasta, joka velvoittaa viranomaiset järjestämään tiedonhallintansa ja kuvaamaan sen jatkuvasti ylläpidettävässä tiedonhallintamallissa. Laki antaa myös velvoitteita koskien tietoturvallisuuden huomioimista tiedonhallinnassa ja se velvoittaa mm. keräämään tarpeellisia lokitietoja silloin kun on kyseessä kirjautumista vaativa järjestelmä. Osa tiedonhallintalain vaatimuksista on täytettävä vuoden 2021 alkuun mennessä ja osa myöhemmin. Lokien keräämistä koskevan vaatimuksen takaraja on vuoden 2021 loppu.

Käyn tässä opinnäytetyössä läpi, mitä tiedonhallintalain 17 § vaatimus lokitietojen keräämisestä tarkoittaa tiedonhallintayksikön kannalta. Tiedonhallintayksikkö on tiedonhallintalaissa käytetty käsite, jolla tarkoitetaan sellaista viranomaista tai viranomaisen osaa, jonka tehtävänä on järjestää tiedonhallinta lain mukaisesti. Käytän lokien hallinnan arvioinnissa esimerkkinä omaa työpaikkaani Museovirastoa. Konkreettisena tavoitteena on selvittää missä tietojärjestelmissä tiedonhallintalain mukainen lokien kerääminen on toteutettava. Tutkimukseni tuloksena on tarkoitus esittää malli, miten lokien keräämisen tarvetta voisi määritellä.

Tutkimuksen tietoperustassa käyn läpi, mitä lokitiedot ja lokien hallinta yleisellä tasolla tarkoittavat. Lisäksi tarkastelen tiedonhallintalakia ja sen 17 § tarkemmin, muun muassa selvittämällä taustoja lain valmisteluaineiston kautta. Tärkeitä lähteitä työssä ovat Tiedonhallintalautakunnan suositus ”Suosituskokoelma tiettyjen tietoturvallisuussäädösten soveltamisesta” (Tiedonhallintalautakunta 2020), julkisen hallinnon tietohallinnon neuvottelukunta JUHTA:n koulutusmateriaalit (JUHTA 2017a, JUHTA 2017b) sekä hieman vanhemmat, mutta edelleen periaatteiltaan relevantit lokiohjeistukset, joita on tehty eri vuosina eri valtion virastossa (Valtiovarainministeriö 2009, Viestintävirasto 2016). Empiirisessä osassa olen selvittänyt mitä lain vaatimus tarkoittaa nimenomaan Museoviraston kannalta ja tehnyt analyysin Museoviraston järjestelmien nykytilanteesta ja mallin tarveanalyysille, jonka perusteella voi tehdä johtopäätöksiä tietojärjestelmäkohtaisista tarpeista lokien keräämiselle.

Koska kyseessä on vasta voimaan tullut laki, jonka lokien hallintaa koskevan pykälän siirtymäaika ei ole vielä kulunut loppuun, julkishallinnon organisaatioissa ei ole vielä tällä hetkellä täyttä selvyyttä mitä konkreettisia muutoksia organisaatioiden, tai oikeammin sanoen tiedonhallintayksiköiden, tulee lokien hallinnan osalta tehdä. Lokien hallintaa koskeva pykälä on melko suppea, mutta sitä täydentää tiedonhallintalautakunnan suositus. Olen

tässä tutkimuksessa pyrkinyt tekemään tulkintaa lain, suosituksen ja lain valmisteluaineistojen perusteella siitä, miten lokien hallinta konkreettisesti tulee ottaa huomioon, jotta lain vaatimus tulisi täytetyksi.

Joillain hallinnonaloilla, kuten sosiaali- ja terveyshallinnossa, on oman erityislainsäädännön nojalla jo huolehdittu lokitietojen hallinnasta, mutta monilla hallinnonaloilla asia tulee enemmän tai vähemmän uutena velvoitteena.

Lokien hallinta ja lokitietojen kerääminen varmistaa tietojen asianmukaista käsittelyä ja lokitietojen avulla voi esimerkiksi todentaa, että tietojen vastaanottajalla on ollut lainmukainen peruste saada tietoja tai että tietoja on katseltu vain siihen tarkoitukseen kuin on ollut tarkoituksenmukaista. Toisaalta lokit tarjoavat myös välineen tutkia järjestelmän virhetilanteita ja niiden avulla voidaan ajoittaa esimerkiksi vikatilanteen alkamisaika. Lokien avulla voidaan myös varmistaa tietoturvallisuutta ja havaita järjestelmiin kohdistuvat tunkeutumisyritykset.

Alla on lueteltu työssä esiintyvät termit ja lyhenteet.

GDPR	Eu:n tietosuoja-asetus, General Data Protection Regulation
Lokilähde	Tietojärjestelmä tai laite, joka tuottaa lokitietoa.
Lokitapahtuma	Tapahtuma, joka kirjataan lokiin. Esimerkiksi kirjautuminen, katselu, tiedon muuttaminen jne.
Lokitieto	Yksittäinen tieto, joka tallentuu jonkin tapahtuman yhteydessä, esimerkiksi käyttäjätunnus tai aikaleima.
Lokitiedosto	Tiedosto tai tietokanta, jonne tietyn järjestelmän lokitiedot tallentuvat.
SIEM	Keskitetty lokienhallintajärjestelmä, Security Information and Event Management
Sote-laki	Laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä 159/2007
Tiedonhallintayksikkö	Julkisen hallinnon viranomaisorganisaation osa tai monen organisaation muodostama kokonaisuus, jonka tulee järjestää tiedonhallintansa tiedonhallintalain mukaisesti
TiHL	Tiedonhallintalaki eli laki julkisen hallinnon tiedonhallinnasta
VM	Valtiovarainministeriö

2 Lokitiedot

2.1 Mitä lokitiedot ovat

Lokitiedot kertovat mitä, miksi ja milloin tapahtui (Viestintävirasto 2016). Useimmiten nykyaikana lokitiedoilla ja lokituksella viitataan tietojärjestelmien tapahtumien tallentamiseen, mutta periaatteessa lokitietoja saatetaan kerätä myös muunlaisista, ei digitaalisista tapahtumista. Esimerkiksi vieraskirjakin on eräänlainen loki. Tässä työssä lokitiedoilla tarkoitetaan tietojärjestelmien lokitietoja, ellei toisin mainita.

Loki-sana periytyy tietoteknistenkin tapahtumien kirjaamiseen laivojen lokikirjasta, johon on alun perin kirjattu laivan kulkema matka nopeuden arvioimiseksi. Laivan lokikirja puolestaan sai nimensä mittausvälineestä (chip log). Mittausvälineessä oli puinen esine sidottuna naruun, jossa puolestaan oli solmuja. Narua kerittiin tasaisella vauhdilla auki veteen ja samalla laskettiin solmujen määrää. Nämä lukemat kirjattiin säännönmukaisesti lokikirjaan. (Reaveley 2010). Tässä mielessä idea on edelleen sama: säännönmukaisuudet ja poikkeamat voidaan havaita, kun kirjaaminen on säännöllistä ja jatkuvaa.

Monenlaiset järjestelmät, laitteet, palvelut, reitittimet ja operaattorit keräävät lokeja tallentaakseen mitä järjestelmässä tapahtuu. Lokien avulla varmistetaan tarkoituksenmukainen toiminta sekä havaitaan poikkeamat normaalista toiminnasta (Viestintävirasto 2016). Vähimmillään lokiin tulisi tallentua aikaleima, tapahtuma, tekijä, tekijän käyttöoikeustaso, tapahtuman lähde sekä tapahtuman onnistuminen/status (Kyberturvallisuuskeskus 2020). Tiedonhallintalautakunnan suositus lokitietojen keräämisestä (Tiedonhallintalautakunta 2020, 40) antaa vielä yksityiskohtaisempia ohjeita kerättävistä tiedoista. Aikaleiman osalta tulee esimerkiksi kerätä sekä päivämäärä että kelloaika, ja mahdollisuuksien mukaan niin lokitiedolla kuin lokitettavalla tapahtumalla tulisi olla omat aikaleimansa. Lisäksi lokitiedolla tulee olla tunniste.

Tapahtuman osalta lokitietoihin tulee tallentua tapahtuman kohde, jotta käy mistä laitteesta, järjestelmän osasta tai tiedosta on kysymys. Tekijän osalta tulee tallentaa mahdollisimman tarkat käyttäjätiedot, mutta ei kuitenkaan henkilötunnuksia, salasanoja tai arkaluonteisia erityisiin henkilötietoryhmiin kuuluvia tietoja. Myöskään luottokorttinumeroita tai muitakaan lokitietoja korkeampaa turvallisuusluokitusta vaativaa tietoa ei tule tallentaa.

Lisäksi tapahtuman osalta tulisi tiedonhallintalautakunnan suosituksen mukaan tallentaa paitsi tapahtuman tyyppi ja onnistuminen, myös sen merkitys ja kuvaustietoja.

Lokitettavien tietojen määrä ja laatu riippuvat siitä, minkälaisesta lokista on kysymys ja mihin sitä aiotaan käyttää. Erilaisia lokeja ovat mm. virhelokit, käyttölokit, tapahtumalokit, viestintälokit, muutoslokit, pääsynhallintalokit tai haltijalokit (Viestintävirasto 2016).

2.2 Erilaisia lokityyppejä

Lokeja käytetään eri käyttötarkoituksiin ja siten kaikki lokit eivät ole samanlaisia. Alla esitetty lokien tyypittely on suuntaa antava ja sen tarkoituksena on antaa käsitys lokin moninaisista käyttötarkoituksista. Joskus lokissa voi olla kyse kahden tyyppisestä lokista samanaikaisesti, tai joissain yhteyksissä jotain tiettyä lokityyppinimitystä saatetaan käyttää eri merkityksessä kuin jossain toisessa yhteydessä. Karkeasti jaotellen kuitenkin suurin osa lokeista on jonkun alla luetellun tyypin mukainen.

A) Virhelokit

Virhelokeihin tallentuu tietoa järjestelmässä tapahtuvista virhetilanteista. Osa virheistä voi liittyä käyttäjiin, mutta virheitä voivat aiheuttaa myös käyttäjistä riippumattomat seikat, kuten ongelmat tietoliikenneyhteyksissä, ohjelmistokoodissa, muissa ohjelmissa jne. Toisaalta esimerkiksi myös tietomurrot voivat ilmetä virheinä virhelokin kautta. (Techopedia 2020).

B) Käyttölokit

Käyttölokeihin tallentuu tietoa nimensä mukaisesti käyttäjien toimista tietojärjestelmässä, esimerkiksi sisään- ja uloskirjautumisista samoin kuin epäonnistuneista kirjautumisista. Toisaalta myös käyttöoikeuksien muutokset kuuluvat käyttölokiin. Käyttölokin tallennus kohdistuu useimmiten tietojen käyttämiseen ja katseluun ja sen tarkoituksena on valvoa ja seurata järjestelmässä olevien tietojen käyttöä. Käyttöloki voi sisältää myös tiedot tallennuksista ja poistoista, vaikka toisinaan näiden tapahtumisen kirjaamisen yhteydessä puhutaan myös muutoslokista. (Kansallinen terveysarkisto 2013, Kyberturvallisuuskeskus 2020).

C) Muutoslokit

Muutosloki on joissain tapauksissa sama asia kuin käyttöloki. Muutoslokissa huomio on kiinnittynyt tietojen lisäämisen, muuttamisen ja poistamisen jäljittämiseen. Muutosten kanssa on tärkeää tallentaa myös tieto muuttuneen tiedon alkuperästä. (Kyberturvallisuus-

keskus 2020). Muutoslokeja saatetaan jossain yhteyksissä kutsua myös transaktiolokeiksi. Tällöin kyse on erityisesti tietokantojen muutostapahtumien tallentamisesta. Tietokannoissa muutoksista on tärkeää tallentaa lokitietoja tietojen eheydenkin kannalta ja yleensä lokitieto on tietokantaan erottamattomasti kuuluva osa. (Valtiovarainministeriö 2009, 37).

D) Luovutuslokit

Luovutuslokeihin tallennetaan tiedot luovutustapahtumista, eli tilanteista, joissa järjestelmästä on luovutettu tietoja jollekin ulkopuoliselle taholle. Luovutuslokit ovat tärkeitä etenkin silloin, jos luovutettavat tiedot sisältävät henkilötietoja, turvallisuusluokiteltuja tietoja tai muuten salassa pidettäviä tietoja. Esimerkiksi sähköinen asiakastietolaki, eli laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä on velvoittanut sosiaali- ja terveysalan toimijat ylläpitämään luovutuslokeja. (Kansallinen terveysarkisto 2013). Tietoluovutuksissa kerättävän lokitiedon tärkeyttä korostetaan myös tiedonhallintalautakunnan suosituksissa. Näiden suositusten mukaan luovutuslokin keskeinen tehtävä on varmistaa, että tietojen luovuttamiselle on laillinen peruste (Tiedonhallintalautakunta 2020).

E) Viestintäloki

Puhelimet, sähköposti ja muut viestintälaitteet tallentavat tietoja viestien lähettämisestä ja vastaanottamisesta. Ns. teletunnistetiedot ovat viestintälokia. (Kyberturvallisuuskeskus 2020). Viestintälokin tarkoituksena voi olla kyseisen järjestelmän vikatilanteiden selvittäminen, tietoturvallisuuden varmistaminen tai viestin perille menemisen todistaminen. (JUHTA 2017).

F) Haltijaloki

Haltijalokissa pidetään kirjaa, kenelle tietty laite, ohjelma, lisenssi, ip-osoite tai nettiosoite on ollut annettuna tiettynä ajankohtana. Haltijalokia voi käyttää myös yhdistettynä käyttölokiin, jolloin esimerkiksi tietyn ip-osoitteen tekemät käyttötapahtumat voidaan jäljittää pidemmälle (Tolonen 2017).

G) Pääsynhallintaloki

Pääsynhallintalokiin kirjautuu tiedot järjestelmän sisään- ja uloskirjautumisista. Käyttötarkoitus liittyy vahvasti tietoturvaan ja lokin avulla voidaan seurata esimerkiksi, jos järjestelmään yritetään kirjautua vanhentuneilla tai poistetuilla tunnuksilla. Järjestelmätason lisäksi

myös verkkotasolla tallennetaan päädynhallintalakia, joka toisaalta on melko kohdan I) yhteyslokia. Myös kiinteistöjen kulunvalvonta on eräänlainen pääsynhallintaloki. (Kyberturvallisuuskeskus 2020). Pääsynhallintalokille on tyypillistä useista muista lokeista poiketen, että se sisältää henkilötietoja.

H) Sovelluslokit

Sovellus-, järjestelmä- ja käyttöjärjestelmälokit tallentavat tietoa sovelluksen sisäisistä prosesseista, niiden käynnistymisistä, toimenpiteistä sekä mahdollisista virhetilanteista. (JUHTA 2017). Sovelluslokeiksi voidaan laskea myös verkkopalvelulokit, joihin tallentuu tietoa esimerkiksi www-käyttäjän selaimesta, sijainnista tai ip-osoitteesta. Käyttöjärjestelmälokien kerääminen halutulla tavalla edellyttää useimmiten, että käyttöjärjestelmässä on säädetty lokiasetuksia. Tämä koskee sekä Windows- että Unix-ympäristöjä. (Tolonen 2017). Käyttöjärjestelmät keräävät usein samalla useita eri lokeja. Esimerkiksi Windows-käyttöjärjestelmä kerää sovellus-, suojaus- ja järjestelmälokia (Valtiovarainministeriö 2009, 34).

I) Verkon yhteyslokit

Verkon yhteyslokeja tallentavat mm. reitittimet ja palomuurit. Näistä lokeista ilmenee mistä osoitteesta on mennyt liikennettä mihin osoitteeseen. Korkeamman protokollatason lokista näkyy myös mihin tietoliikenneporttiin liikenne on kohdistunut. Verkkotason yhteyslokien tarkoitus on ennen kaikkea virhetilanteiden ja tietoturvapoikkeamien selvittäminen. Yhteyslokin avulla voidaan luotettavimmin havaita esimerkiksi palvelimiin kohdistuneet tietomurrot, vaikka niistä ei vielä näekään mitä palvelimella on tapahtunut (Valtiovarainministeriö 2009, 36).

2.3 Miksi lokitietoja kerätään

Edellä luetelluista lokityypeistä käy ilmi, että lokeja on tarpeen kerätä useita erilaisia käyttötarkoituksia varten. Lokien kerääminen voi liittyä järjestelmän teknisten ongelmien selvittämiseen, jolloin lokeja tarvitaan joko jälkikäteen tai jopa reaaliaikaisesti. Keräämisen tarkoitus voi liittyä myös järjestelmiin kohdistuvien luvattomien tunkeutumisten estämiseen tai tutkimiseen. Lokeja voidaan myös käyttää sisäisten väärinkäytösten valvontaan ja niiden selvittämiseen. Lokien tallentuminen voi joskus olla myös virkavastuun todentamisen tai eri osapuolien oikeusturvan toteutumisen kannalta tärkeää. Näiden syiden takia sekä

automaattiset tapahtumat että käyttäjien tekemät toimet ovat lokituksen kohteena. (Kyberturvallisuuskeskus 2020).

Ennen kuin lokien keräämistä voi alkaa suunnitella, täytyy olla selvillä mihin tarkoitukseen niitä aiotaan käyttää. Lokien kerääminen ilman syytä ei ole järkevää eikä kustannustehokasta. Lokien keräämistä ohjaavat usean lait. Laeista ns. sotelaki, laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä 159/2017 velvoittaa lokien tallentamiseen aina tietojen luovuttamisen yhteydessä, mutta monissa muissa tapauksissa, esimerkiksi uusi tiedonhallintalaki, laki sisältää ajatuksen tarpeellisuuden arvioinnista. Tarve lokien keräämiselle määräytyy näin ollen organisaatio- ja tietojärjestelmäkohtaisesti.

Lokien keräämisestä on säädetty lukuisissa eri laeissa. Tämän työn aiheena oleva tiedonhallintalaki on osaltaan pyrkinyt yhtenäistämään julkishallinnon tiedonhallintaa koskevaa sääntelyä tältä osin ja tuomaan myös lokien hallinnan osaksi yleisiä tiedonhallinta-asioita. (Eduskunta 2020, Kivivasara 2019). Tiedonhallintalaissa lokien käyttötarkoitus rajataan teknisten virheiden selvittämiseen ja järjestelmien käytön ja niistä tehtävien tietoluovutusten seurantaan. Käytännössä lokeja keräävän ja hallinnoivan tahon on kuitenkin otettava lokien muutkin käyttötarkoitukset huomioon. Näin ollen tässä työssä huomioidaan lokien käyttötarkoitukset laajasti, vaikka samalla tarkastellaankin nimenomaan tiedonhallintalaista tulevien velvoitteiden toteuttamista julkishallinnon organisaatiossa.

Lokeja tarvitaan sekä normaalitilanteessa että poikkeusoloissa. Normaalitilanteessa lokitiedot auttavat seuraamaan toiminnan häiriötöntä jatkumista, tarjoavat tietoa tilastointiin tai esimerkiksi käytön perusteella tapahtuviin laskutuksiin. Poikkeusoloissa lokeja puolestaan tarvitaan, kun selvitetään vikaan johtaneita syitä, tapahtumia ja niihin liittyneitä osapuolia. (Tiedonhallintalautakunta 2020, 37-38).

3 Lokien hallinta

3.1 Lokien hallinnan hyödyt

Edellä on jo kuvattu syitä, miksi lokitietoja on hyödyllistä kerätä. Lokitiedot eivät useinkaan tallennu itsestään, vaan niiden kerääminen ja tallentaminen täytyy olla tietoisesti suunniteltua. Esimerkiksi käyttöjärjestelmissä täytyy erikseen asettaa lokitietojen kerääminen päälle (Tolonen 2017) ja erilaisten sovellusten suunnittelussa lokitietojen tallentaminen täytyy ottaa huomioon jo suunnitteluvaiheessa. Lokien kerääminen organisaatiossa tulee muutenkin tapahtua hallitusti lakien ja lokiperiaatteiden ja -suunnitelmien mukaisesti (Tiedonhallintalautakunta 2020, 38).

Hyvin suunniteltu lokien kerääminen mahdollistaa sen, että organisaatio pystyy tehokkaasti hyödyntämään hallussaan olevia lokitietoja. Lokitiedot voivat parhaimmillaan mahdollistaa säännöllisen seurannan ja analysoinnin toteuttamisen. Tiedonhallintalautakunnan suosituksessa puhutaan havainnointikyvykkyyden kehittymisestä, joka edesauttaa esimerkiksi kriittisten kohteiden osalta reaaliaikaisen poikkeamien havainnoimisen. Jotta tämä onnistuisi, jokaisen järjestelmän osalta on täytynyt muodostua kuva, kuinka järjestelmä toimii lokien näkökulmasta normaaliaikana. Näin saadaan vertailukohta poikkeavalle toiminnalle. (Tiedonhallintalautakunta 2020, 45).

Lokien hallintaan kuuluu koko lokitiedon elinkaaren suunnittelu keräämisestä tiedon säilyttämiseen, analysointiin ja poistamiseen.

3.2 Lokien hallinnan suunnittelu ja toteutus

Kun on tunnistettu, että lokitietoa tarvitaan, on syytä ryhtyä systemaattisesti suunnittelemaan lokitietojen hallintaa. Julkisen hallinnon digitaalisen turvallisuuden johtoryhmä (VAHTI) sekä julkisen hallinnon tietohallinnon neuvottelukunta (JUHTA) ovat vuosina 2017-2018 järjestämissään yhteishankkeissa ohjeistaneet lokien hallinnan suunnittelusta (JUHTA 2017b, 21). Lokien hallinnan suunnittelua ohjataan lokiperiaatteilla, jotka organisaation johto on hyväksynyt lokien keräämistä, käyttämistä ja säilyttämistä ohjaaviksi säännöiksi. Lokiperiaatteet ottavat kantaa lokien keräämisen tavoitteisiin, hallinnan prosesseihin, rooleihin, vastuisiin, elinkaareen ja tekniseen toteutukseen (JUHTA 2017b, 17). Lokiperiaatteet tekevät selväksi niin organisaation eri toimijoille kuin sidosryhmillekin, minkälaisten suuntaviivojen puitteissa lokitietoja kerätään. Periaatedokumentti antaa selkänojan toteutukselle. Lokiperiaatteisiin pitäisi myös sisällyttää kuvaus eri rooleista lokien käsittelyssä ja siitä, mikä näiden roolien vastuut ovat. Lisäksi myös sidosryhmien mahdol-

liset vastuut tulee kuvata (Ahonen, Seppänen & Pärssinen 2019, liite A). Ahosen, Seppäsen ja Pärssisen energia-alan kyberturvallisuutta koskevassa raportissa suositellaan lokiperiaatteiden liitteeksi lisäksi esimerkinomaista listausta erilaisista lokityypeistä. Lokityyppilistausta ei kuitenkaan tule nähdä kattavana tai määräävänä vaan siitä voi poiketa sen mukaisesti minkälaisia lokeja järjestelmä voi tuottaa (Ahonen ym. 2019, liite A).

Käytännön työ suunnittelussa tulisi Juhdan ohjeen mukaisesti aloittaa lokilähteiden tunnistamisesta, lokien käyttötarkoituksen tunnistamisesta sekä lokitietojen luokittelusta. Lokilähteillä viitataan yleensä lokeja tuottaviin järjestelmiin etenkin silloin, kun eri järjestelmien lokitietoja on tarkoitus tuoda yhteen keskitettyyn hallintajärjestelmään. Eri lähteet on tunnistettava, jotta voidaan suunnitella esimerkiksi tarpeelliset integraatiot. Vaikka kyseessä ei olisikaan keskitetty lokien hallinta, kokonaiskuvan muodostaminen on silti aiheellista.

Kun lokeja keräävät järjestelmät on tunnistettu, edetään Juhdan ohjeen mukaan vaiheeseen, jossa näiden eri lokien käyttötarkoitukset analysoidaan ja niiden sisältämät tiedot luokitellaan. Eri lokityypit voivat vaatia erilaista käsittelyä, esimerkiksi erilaiset säilytysajat, joten on tärkeää tunnistaa minkä tyyppisestä lokista on kyse ja mikä tietojen käyttötarkoitus on. Esimerkiksi teknisten virheiden havainnointiin riittää useimmiten lokien säilyttäminen joitakin kuukausia, kun taas erilaisten väärinkäytösten selvittämisessä lokeja on hyvä säilyttää joko se aika, mikä on syyteoikeuden vanhenemisaika tai vahingonkorvaustapauksissa maksimirangaistusaika, joka lasketaan vahinkoon johtaneesta tapahtumasta. Mikäli lokeja käytetään tämän tapaisten väärinkäytösten selvittelyyn, säilytysajat vaihtelevat yleensä 3 ja 10 vuoden välillä (JUHTA 2017b, 19).

Lokitietojen luokittelu myös auttaa tunnistamaan lokeille tarvittavan oikean tietosisällön. Tästä on hyötyä, jottei kerätä turhia tietoja, joiden käsittelystä tai säilyttämisestä saattaisi aiheutua lisätyötä tai -kustannuksia. Esimerkiksi liian tarkkoja henkilötietoja ei kannata kerätä, ellei siihen ole välttämätöntä syytä. Lokitietojen luokittelu vaikuttaa myös siihen, mikä lokitietojen suojaustaso on. Suojaustasoon vaikuttaa kuitenkin lokitietojen sisällön lisäksi myös lokilähde sinänsä (Juhta 2017b, 14).

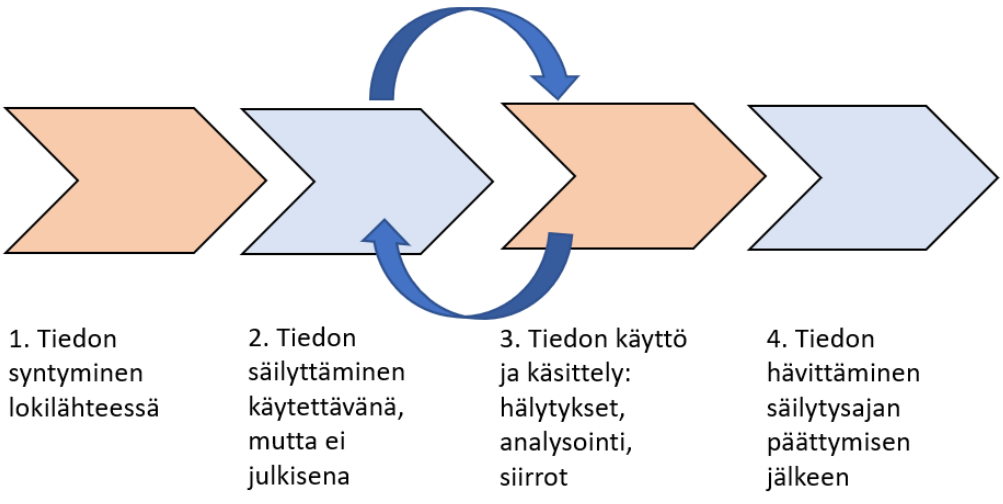
Lokien hallinnan suunnitteluun kuuluu myös se, että huomioidaan mahdollisuudet yhdistää eri järjestelmien lokitietoja. Lokien tulisi näin ollen olla mahdollisimman yhteensopivia. Esimerkiksi jotta aikaleimoja voisi vertailla, lokilähteiden kellojen tulisi olla samassa ajassa. Lisäksi hyvällä lokien hallinnan suunnittelulla varmistetaan, että lokitiedot pysyvät tallessa niin kauan kuin on tarpeen (Tiedonhallintalautakunta 2020, 43).

Lokit on tärkeää pystyä suojaamaan asiattomalta manipulaatiolta. Lokien tehtävänä on varmistaa tietojen ja tietokantojen eheys ja oikeellisuus sekä asianmukainen pääsy ja toiminta järjestelmistä. Lokit eivät täytä tehtäväänsä, mikäli niitä on mahdollista päästä muokkaamaan oikeudettomasti. Esimerkiksi kirjoitusoikeus lokiin tulisi olla vain sillä prosessilla, jonka tarkoituksena lokin tuottaminen on. Varma suojaus on se, että lokeja kirjoitetaan medialle, jonne on olemassa vain kertakirjoitusoikeus. Samoin lokien muuttumattomuutta suojaa niiden siirtäminen keskitettyyn järjestelmään, jonne lähdejärjestelmien ylläpitäjilläkään ei ole pääsyä. Etenkin silloin, jos lokitiedot sisältävät henkilötietoja tai salassa pidettäviä tietoja, myös lokitietojen katselu tulisi lokittaa. (Valtiovarainministeriö 2009, 62).

3.3 Lokien elinkaari

Lokitiedon elinkaari muodostuu pitkälti samoin kuin minkä tahansa datan elinkaari. Lokeihin liittyy kuitenkin tiettyjä erityispiirteitä elinkaaren kaikissa vaiheissa. Kansallisarkisto, joka visionsa mukaisesti pyrkii olemaan julkisen tiedon elinkaaren johtava asiantuntija, on esittänyt tiedon elinkaarelle seuraavia vaiheita: tiedon syntyminen ja käsittelyn alkaminen, julkisuuden arvioiminen, käyttöoikeuksien hallinta, tiedon rikastaminen, elinkaaren aikaiset tapahtumat (kuten siirrot tai migraatiot), säilyttäminen ja viimeisenä hävittäminen tai pysyvä säilyttäminen (Eräkaski 2017).

Lokitiedon kannalta elinkaari yllä mainittua soveltaen voisi olla seuraavanlainen:



Kuva 1. Lokitiedon elinkaari (Eräkasken kirjallisesta kuvauksesta kuvaksi piirretty ja lokien hallintaan sovellettu)

Lokitiedon elinkaari alkaa, kun lokitapahtuman seurauksena syntyy tieto jossakin lokilähteessä. Tieto säilyy ja pysyy käytettävänä lokityypin määrittämää käyttötarkoitusta varten. Kansallisarkiston luettelemiin elinkaarivaiheisiin kuuluu mm. julkisuuden/salassapidon määrittäminen. Lokitiedot usein eivät ole julkisia, mutta niihin voi liittyä tiedonsaantioikeus joko erityislakiin tai julkisuuslakiin perustuen (ks. kpl 4.2). Pääsy lokitietoihin tulee olla rajattu tarkoituksen mukaisesti. Lokitietojen säilytysratkaisu tulee olla sellainen, että tiedot ovat käytettävissä mm. analysointiin ja tapahtumien tarkistamiseen. Elinkaaren aikaisiin tapahtumiin voi liittyä myös tietojen siirtämistä lokilähteestä keskitettyyn tietovarantoon ja tiedon konvertointia tätä tarkoitusta varten.

Etenkin keskitetyn lokienhallinnan tilanteessa lokitiedon elinkaari voi olla pidempikin kuin lokilähteen elinkaari. Toki tilanne voi olla myös päinvastoin. Kansallisarkiston elinkaarivaiheissa tiedon elinkaari päättyy joko hävittämiseen tai pysyvään säilyttämiseen (tiedonhallintalaissa pysyvästä säilytyksestä käytetään termiä arkistointi). Lokien kohdalla on kuitenkin erittäin harvinaista, että lokitietoa tulisi säilyttää pysyvästi. Lokitiedon säilytysaika riippuu lokin käyttötarkoituksesta, eli esimerkiksi siitä, miten pitkän aikajänteen ajalta ajatellaan virhetilanteita selvitettävän. Jos lokien keräämisellä pyritään valvomaan mm. väärinkäytöksiä, ratkaisevaa on mikä on syyteoikeuden vanhenemisaika tai vahingonkorvausvastuun maksimipituus. Esimerkiksi törkeän viestintäsalaisuuden loukkauksesta, törkeän tietojärjestelmän häirinnästä ja törkeästä tietomurrosta syyteoikeus vanhenee 10 vuodessa. Näin ollen näitä tilanteita valvovien lokien säilytysaika tulisi olla vähintään 10 vuotta. Myös vahingonkorvausvaatimuksen voimassa olo on 10 vuotta vahinkoon johtaneesta tapahtumasta. Ns. sotelaki määrää sosiaali- ja terveysalan asiakastietojärjestelmien lokien säilytysajaksi 12 vuotta. (Juhta 2017b, 19). Sen sijaan joissain vikatilojen selvittämiseen tarkoitetuissa lokeissa säilytysaika voi olla vain joitain kuukausia tai viikkoja.

3.4 Lokitapahtumien analysointi

Jotta lokien keräämisestä saisi suurimman mahdollisen hyödyn, lokitapahtumien tuottamia tietoja tulisi pystyä analysoimaan. Analysointivaihe voi olla työläs, mutta toisaalta sitä on mahdollista myös automatisoida. Etukäteen on kuitenkin syytä suunnitella, miten analysointia tehdään. Samalla on pidettävä mielessä, että analysointi myös kehittyy sen mukaan, minkälaisia tietoja lokiin kertyy ja mitä niistä opitaan. (Valtiovarainministeriö 2009, 47).

Valtiovarainministeriön lokiohje (2009, 47-48) mainitsee seuraavat asiat, jotka ovat tärkeitä lokitietojen analysoinnissa:

1. Säännöllisyys
2. Ymmärrys normaalitilanteen tapahtumista, joihin poikkeamia voidaan verrata
3. Analysointiin tulevien tietojen suodattaminen massasta priorisointiin pohjautuen.

Kolmantena kohtana mainittu priorisointi tulisi suunnitella VM:n lokiohjeen mukaan seuraavia näkökohtia mielessä pitäen:

1. Lokimerkinnän tyyppi (onko kyse esimerkiksi tiedon muuttamisesta vai tiedon katselusta)
2. Lokimerkinnän harvinaisuus tai uutuus (onko merkintä kokonaan uuden tyyppinen?)
3. Lokilähde (kriittiset järjestelmät)
4. Lähde- ja kohde-ip (onko esim. ns. mustalla listalla)
5. Kellonaika ja viikonpäivä (tapahtuuko tapahtuma väärään tai huomiota herättävään aikaan?)
6. Lokitapahtuman esiintymisen tiheys (esimerkiksi hyvin tiheästi ilmenevät tapahtumat)

3.5 Keskitetty lokien hallinta

Jos lokilähteitä on runsaasti ja tietoa kertyy paljon, ja jos samanaikaisesti lokitietojen tehokkaalle hyödyntämiselle on suuri tarve, organisaation kannattaa harkita keskitettyä lokien hallintaa. Keskitetty lokien hallinta tarkoittaa, että eri lokilähteissä syntyvät lokitiedot analysoidaan ja säilytetään yhdessä, nimensä mukaisesti keskitetyssä paikassa. Yleisimmin keskitetyn lokien hallinnan ratkaisuna esitetään ns. SIEM-ratkaisut (Security Information and Event Management). Kuitenkin esimerkiksi Kyberturvallisuuskeskus neuvoo aloittamaan pienemmillä askelilla, esimerkiksi oman lokipalvelimen pystyttämisellä, ennen varsinaisen SIEM-järjestelmän käyttöönottoa (Viestintävirasto 2016, 8). Samaisessa ohjeessa tähdennetään, että SIEMiä ei tule ajatella järjestelmänä, joka ratkaisee lokien hallinnan, vaan se on ennemminkin prosessi, joka vaatii organisaation toimintojen sovittamista keskitettyyn lokien hallintaan sopivaksi.

SIEM ei ole synonyymi keskitetylle lokien hallinnalle. Keskitettyä lokien hallintaa voi tehdä ilman SIEM-ratkaisua ja toisaalta SIEMin avulla voi kontrolloida vaikka vain yhdenkin järjestelmän turvallisuutta. SIEM kuitenkin tarjoaa pitkälle automatisoidut työkalut poikkeamien havainnointiin ja niihin reagoimiseen. (Viestintävirasto 2016, 11).

Keskitettyyn lokien hallintaan saattaa helposti alkaa kertyä todella paljon tietoa, mikä vaikeuttaa sen analysointia. Näin ollen jo suunnitteluvaiheessa tulee pyrkiä riittävän minimoituun tietosisältöön. Lisäksi integroinnissa voi tulla vastaan hankaluuksia mm. sen kanssa,

että eri lokilähteiden tapahtumat eivät ole verrannollisia keskenään. Toisaalta sen etuna on se, että lokeja voidaan tarkastella yhtäaikaisesti ja esimerkiksi erilaisten tietoturvapoikkeamien havaitseminen tehostuu. Keskitetyssä lokien hallinnassa voidaan esimerkiksi hyökkäys huomata aiemmin, kun signaaleja saadaan monesta lähteestä. (Viestintävirasto 2016, 8-9). Tietomurtotapauksissa murtaja voi päästä käsiksi myös lokitietoihin, mikäli niitä ei ole siirretty toiseen sijaintiin. Tämäkin tukee keskitetyn lokien hallinnan hyödyllisyyttä. Vaikka esimerkiksi verkkolaitteiden lokit ovat yleensä paremmassa turvassa kuin palvelin- tai järjestelmätason lokitiedot, myös niiden lokit on turvallisinta siirtää säilöön keskitettyyn järjestelmään. (Valtiovarainministeriö 2009, 36).

Keskitettyyn, tai muuten lokilähteen tiedoista eriytettyyn lokitietojen säilytykseen, voi olla aihetta myös silloin, jos lokien säilytysajaksi on määritelty pidempi aika kuin mitä lokeja tuottavan järjestelmän muiden tietojen säilytysaika on. Tällöin lokitiedot kannattaa siirtää erilliselle lokipalvelimelle talteen (Tiedonhallintalautakunta 2020, 44). Tiedonhallintalautakunnan suosituksessa suositellaan

4 Lainsäädäntö

4.1 Tiedonhallintalaki

Laki julkisen hallinnon tiedonhallinnasta (906/2019) astui voimaan 1.1.2020. Lain tavoitteena on varmistaa viranomaisten palvelujen ja tietoaineistojen hyödynnettävyys henkilöiden tietosuojaa ja muuta tietoturvallisuutta unohtamatta. Samalla lain tarkoituksena on ohjata viranomaisia yhdenmukaiseen ja laadukkaaseen tiedonhallintaan sekä yhä enemmän kohti aineistojen ja järjestelmien yhteentoimivuutta.

Laki säätää tiedonhallinnan järjestämisestä tiedonhallintayksikössä. Tiedonhallintayksiköt on määritelty lain 4 §:ssa: käytännössä niitä ovat esimerkiksi valtion virastot, laitokset, kunnat, kuntayhtymät ja yliopistot. Kaikkien tiedonhallintayksiköiden on lain mukaan ylläpidettävä tiedonhallintamallia, joka määrittelee ja kuvaa organisaation tiedonhallinnan. Mallin tulee sisältää tiedot prosesseista, tietojärjestelmistä ja tietovarannoista sekä näiden välisistä kytkennöistä. Tiedonhallintamallit on laadittava vuoden 2020 loppuun mennessä.

Tiedonhallintamallin lisäksi tiedonhallintayksiköiden tulee lain mukaisesti huolehtia tietoaineistojen sähköiseksi muuttamisesta ja mm. niiden säilytysaikojen määrittämisestä. Laki myös velvoittaa avaamaan teknisiä rajapintayhteyksiä tietojen luovuttamiseksi tai katseluyhteyden avaamiseksi eri viranomaisten välillä. Samoin laki määrää myös viranomaisen asioiden rekisteröimisestä asiarekisteriin.

Tiedonhallintalaki säätää myös valtionhallinnon yleisestä tiedonohjauksesta ja sen nojalla on perustettu tiedonhallintalautakunta, jonka tehtävänä on arvioida lain kohtien toteutumista ja noudattamista sekä neuvoa ja ohjeistaa tiedonhallintayksiköitä lain vaatimusten täyttämisessä.

Tämän työn aihe, lokitietojen keräämisestä säätäminen, löytyy tiedonhallintalain neljännestä luvusta, joka käsittelee tietoturvallisuustoimenpiteitä. Lokien keräämisen ohella muita tietoturvaa koskevia velvoitteita ovat, että viranomaisten tulee tunnistaa luottamuksellisuutta edellyttävät tehtävät, huolehtia tietojärjestelmien ja tietoaineistojen turvallisuudesta, varmistaa tietojen turvallinen siirtäminen tietoverkossa, varmistaa tietoaineistojen turvallisuus, järjestää tietojärjestelmien käyttöoikeuksien hallinta sekä luokitella asiakirjansa turvallisuusluokkien mukaisesti.

4.1.1 Tiedonhallintalain 17 § - Lokitietojen kerääminen

Lokien keräämistä koskeva pykälä tiedonhallintalaissa on seuraavanlainen:

”Viranomaisen on huolehdittava, että sen tietojärjestelmien käytöstä ja niistä tehtävistä tietojen luovutuksista kerätään tarpeelliset lokitiedot, jos tietojärjestelmän käyttö edellyttää tunnistautumista tai muuta kirjautumista. Lokitietojen käyttötarkoituksena on tietojärjestelmissä olevien tietojen käytön ja luovutuksen seuranta sekä tietojärjestelmän teknisten virheiden selvittäminen.”

Lain kohta on vain kahden virkkeen mittainen, mutta sillä on suuria vaikutuksia viranomaisten tarvitsemaan järjestelmäkehitykseen, koska vaatimus ei todennäköisesti toteudu tällä hetkellä läheskään kaikissa tietojärjestelmissä. Toisaalta pykälä ei määrittele vaatimusta kovin tarkasti. Mitä ovat tarpeelliset lokitiedot missäkin yhteydessä? Missä muodossa lokitiedot tulee kerätä? Miksi käyttötarkoituksiin on määritelty käytön ja luovutuksen seuranta ja tekniset virheet, mutta ei esimerkiksi tietomurtojen tai hyökkäysten estämistä?

Lokien hallintaa koskeva pykälä on kuitenkin hyvin linjassa lain hengen ja tavoitteen kanssa: myös muissa pykälissä tavoitellaan sitä, että tietoja luovutettaisiin viranomaiselta kansalaisille ja muille viranomaisille, mutta sen pitää tapahtua tietosuoja ja tietoturvallisuus huomioiden. Viranomaisten on täytynyt huomioida lokien keräämistä koskeva vaatimus 1.1.2020 alkaen uusissa hankittavissa järjestelmissä, mutta vanhojen järjestelmien on oltava lain vaatimuksen mukaisia 24 kk siirtymäajan jälkeen, eli vuoden 2022 alussa.

Jo ennestään voimassa olleessa lainsäädännössä, mm. laissa henkilötietojen käsittelystä rikosasioissa ja kansallisen turvallisuuden ylläpitämisen yhteydessä on säädetty lokitietojen keräämisestä silloin kun tietojen käsittely, luovutus tai muu toiminta on kohdistunut henkilötietoihin. Uusi tiedonhallintalain vaatimus kuitenkin laajentaa lokien keräämisen vaatimuksen myös muita kuin henkilötietoja koskeviin tietoihin ja tietojärjestelmiin.

4.1.2 Tiedonhallintalain 17 § tausta

Lain valmistelumateriaaleista (hallituksen esitys, valiokuntakäsittelyt ja eduskunnan vastaus) käy ilmi, että 17 § ei ole muuttunut eduskuntakäsittelyn aikana ehdotetusta. (Eduskunta 2020). Asiantuntijalausunnoissa pykälää on kommentoitu ja kritisoitukin, mutta eduskunnan valiokuntakäsittelyssä sen sisältöä ei ole katsottu aiheelliseksi muuttaa. Huolta lausunnonantajien keskuudessa on aiheuttanut mm. pelko kustannusten ja työmäärän lisääntymisestä, kun vaatimus on vietävä käytäntöön sekä hankittavissa että 24

kk kuluessa myös olemassa olevissa järjestelmissä (mm. Kela, Suomen kuntaliitto). Samoin oikeusministeriön lausunnossa on otettu kantaa laissa säännellyistä teknisistä katseluyhteyksistä (22-24 §) ja niiden lokitiedoista. Lausunnossa esitetään huomio, että lokien kerääminen näistä uusista vaadituista toiminnoista tulee viranomaisille mietittäväksi ja toteutettavaksi (Korhonen 2019). Valiokunnat ovat kuitenkin katsoneet valtiovarainministeriön tavoin, että lain vaatimus ei ole kokonaan uusi jo ennestään olleisiin vaatimuksiin nähden eikä siten aiheuta järjestelmiin suuria muutoksia (Kivivasara 2019).

Vm:n edustajat kirjoittavat: ”*Voimassa oleva lainsäädäntö edellyttää luovuttamisen laillisen perusteen varmistamiseksi luovutuslokitietojen keräämistä, jos järjestelmästä luovutetaan rajapintojen tai katseluyhteyden avulla salassa pidettäviä tietoja tai henkilötietoja. Lisäksi käyttölokitiedot tulisi kerätä ainakin tietojärjestelmistä, joissa käsitellään salassa pidettäviä tietoja.”* Voimassa olevalla lainsäädännöllä tarkoitetaan ilmeisesti lakia henkilötietojen käsittelystä rikosasioissa ja kansallisen turvallisuuden ylläpitämisen yhteydessä, jonka pykälä 19 koskee lokitietojen keräämistä ja säilyttämistä silloin kun tietojärjestelmässä käsitellään tai luovutetaan henkilötietoja. Vaikka jo ennestään on ollut vaatimus kerätä lokitietoja henkilötietoja sisältävistä järjestelmistä, uusi vaatimus laajentaa keräämisen myös muihin kuin henkilötietoihin. Tiedonhallintalain 17 §:ään kirjattu ainoa rajaus kohdistuu siihen, käytetäänkö tietojärjestelmässä kirjautumista vai ei. Viranomaisten käytössä on paljon tietojärjestelmiä, rajapintoja ja rekistereitä, joissa käytetään kirjautumista, mutta jotka eivät sisällä henkilötietoja. Tällaisiin järjestelmiin lokitietojen kerääminen kohdistuu uutena vaatimuksena. Tilastokeskus onkin tuonut esiin, että viranomaisella voi olla hallinnassaan useita kirjautumista vaativia järjestelmiä, joissa ei välttämättä ole ennestään lokien keräämistä toteutettu lain vaatimalla tavalla. Tilastokeskus on esittänyt arvion, että vaatimuksen täyttäminen edellyttää virastolta yhteensä 5 htv työpanoksen (Mikkelä 2019).

Edelleen Vm:n lausunnossa todetaan: ”*Ehdotetun säännöksen perusteella käyttölokitietojen keräämisen tarve tulisi arvioida puolestaan sillä perusteella, tarvitaanko niitä virheselvittelyjä varten tai yksilön etujen, oikeuksien ja velvollisuuksien sekä oikeusturvan toteuttamiseksi taikka virkavastuun todentamiseksi.”* Vaikka lakitekstissä mainitaankin sana ”tarpeellinen” lokitietojen tallennuksen yhteydessä, laista ei selkeästi käy ilmi missä määrin lokien kerääminen kirjautumista vaativista järjestelmistä on tarveharkintaista. Myös oikeusasiamies on ottanut lain tähän kohtaan kantaa omassa kannanotossaan vuodelta 2018. Oikeusasiamiehen mukaan on ongelmallista ja epäasianmukaista, jos osa sääntelystä on lakitekstin ulkopuolella sen perusteluissa (Pölönen 2018).

Koska säädösteksti ja sen perustelut ovat ristiriitaisia, käytännön soveltamisessa tullaan todennäköisesti nojautumaan pitkälti tiedonhallintalautakunnan antamiin suosituksiin.

Oikeustieteen kandidaatti Anna-Riitta Wallin on myös tuonut lausunnossaan esille säännöksen niukkuuden. Wallin kiinnittää huomiota etenkin siihen, että lokitietojen keräämisessä kerätään henkilötietoa ja täten olisi pitänyt myös säätää rekisteröidyn oikeuksista tarkastaa tietonsa. Pykälän perusteluissa on mainintoja lokeja koskevista tiedonsaantioikeuksista, mutta ne eivät käy ilmi lakipykälästä (Wallin 2019). Tähän VM on vastineessaan todennut, että niin tiedonsaantioikeudesta kuin rekisteröidyn oikeuksista henkilörekisterissä on säädetty toisaalla (Kivivasara 2019). Rekisteröidyn oikeudet varmistuvat myös lokitietojen osalta tietosuoja-asetuksen ja tietosuojalain kautta ja tiedonsaantioikeuteen pätee mitä julkisuuslaissa sanotaan.

Liikenne- ja viestintäministeriö on todennut lokien hallinnan kuuluvan yleisiin tietojärjestelmien tietoturvallisuustoimenpiteisiin. Lain edellyttäessä sitä vain kirjautumista vaativilta järjestelmiltä, tietoturvallisuutta itse asiassa heikennetään (Luomajärvi 2019). Lain valmistelusta vastaava valtiovarainministeriö on kuitenkin perustellut säännöksen muotoilua sillä, että myös muissa tilanteissa lokitietoja saa kerätä, tiedonhallintalaki määrää sen milloin niitä täytyy kerätä pykälässä mainittuja käyttötarkoituksia ajatellen (Kivivasara 2019).

Kela on asiantuntijalausunnossaan puolestaan huomauttanut, ettei lokitietojen käsitettä ole määritelty laissa ollenkaan. Kelan lausunnossa ihmetellään, miksi laissa säädetään vain käyttö- ja luovutuslokeista, eikä esimerkiksi tietojärjestelmien virhelokeista säädetä mitään. (Vähä-Erkkilä 2019). Lausunto liittyy samaan asiaan kuin yllä mainittu Liikenneviraston lausunto ja sen osalta esitetty vastine: tiedonhallintalaki ei estä lokien keräämistä muistakin tarpeellisista tilanteista, mutta se antaa minimivaatimuksen, milloin lokia on kerättävä tiedonhallintalain tarkoittamien asioiden näkökulmasta.

4.2 Tietosuojalainsäädäntö ja julkisuuslaki

Henkilötietojen suojasta ja rekisteröidyn oikeuksista omiin tietoihinsa säädetään EU:n GDPR-asetuksessa (General Data Protection Regulation) sekä kansallisessa tietosuojalaissa 1050/2018.

Koska monissa lokeissa (mm. käyttölokit, pääsynhallintalokit, haltijalokit) kerätään tietoa myös käyttäjistä, ne muodostavat henkilörekisterin. Rekisterinpitäjä on velvollinen laatimaan selosteen henkilötietojen käsittelytoimista sekä informoimaan rekisteröityjä tietojen keräämisestä, käytöstä ja poistamisesta. Henkilörekisteriä koskevat säädökset koskevat suurimmaksi osaksi myös lokirekisteriä. Esimerkiksi tietojen säilytysajoista on ilmoitettava

rekisteröidyille. Sen sijaan rekisteröity ei voi pyytää tietojensa poistamista lokirekisteristä kuten muista henkilörekistereistä (JUHTA 2017b).

Jos lokitietoja käytetään henkilöstön valvontaan, työnantajan on kerrottava tästä henkilöstölle ja asiasta on käynnistettävä yhteistoimintamenettely (Viestintävirasto 2016, 6.).

Rekisteröidyllä on aina oikeus saada tietoonsa itseään koskevat järjestelmään tallennetut tiedot. Lisäksi joissain tapauksissa henkilöllä on oikeus pyytää tietojärjestelmän käyttäjiä koskevia lokitietoja saadakseen tietää, kuka hänen tietojaan on katsellut tai kenelle niitä on luovutettu. Näin on esimerkiksi potilastietojärjestelmien lokien kanssa, jossa noudatetaan lakia sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä 159/2017. Muista tietojärjestelmistä tiedonsaantioikeus lokitietojen osalta perustuu julkisuuslakiin. Julkisuuslain 11§ 1. momentin perusteella tiedonsaanti on mahdollista myös salassa pidettävistä asiakirjoista, jos henkilö on asianosainen asiassa, johon tietojen katselu tai muu käsittely järjestelmässä on vaikuttanut. Julkisuuslain tiedonsaantioikeus koskee viranomaisen asiakirjoja. Korkeimman hallinto-oikeuden vuosikirjapäätöksestä KHO 2014:69 käy ilmi, että lokitiedot katsotaan asiakirjaksi silloin kun ne ovat osa viranomaisen tietojärjestelmää (Korkein hallinto-oikeus 2020). Samaisesta vuosikirjapäätöksestä käy ilmi, että kynnys salassa pidettävään lokitietoon kohdistuvasta tiedonsaantioikeudesta on kuitenkin melko korkea. Tiedon saannin perusteeksi ei esimerkiksi riitä, että epäilee virkamiesten katselleen itseä koskevia tietoja epäasianmukaisesti.

Juhta on kiteyttänyt lokien hallinnan ja tietosuojavaatimusten välisen yhteyden seuraavan muistilistan avulla (JUHTA 2017b, 27):

- Lokien keräämiselle tulee olla oikeudellinen peruste
- Käyttötarkoitus täytyy olla määrittely etukäteen
- Kerättävä tieto on minimoitu ja luokiteltu
- Lokitietojen käsittely lokitetaan myös
- Tiedonantovelvollisuudet tulevat täytetyiksi
- Huolehditaan YT-menettelystä silloin kun se on tarpeen

On kuitenkin huomattava, että Juhdan muistilista ei perustu tiedonhallintalakiin eikä sen asema suosituksena ole yhtä vahva kuin tiedonhallintalautakunnan suositus lokitietojen keräämisestä. Esimerkiksi lokitietojen käsittelyn lokittaminen on realistista vain silloin kun kyseessä on keskitetty lokien hallintajärjestelmä, ja sellaiseen tiedonhallintalaki ei velvoita.

5 Lokien hallinta Museovirastossa

5.1 Nykytilanteen kartoitus

Museovirasto on viranomaistehtäviä hoitava valtion virasto, joka muodostaa tiedonhallintalaissa tarkoitetun tiedonhallintayksikön. Näin ollen myös lain 17§ vaatimus lokien keräämisestä koskee Museovirastoa.

Museovirastossa on vuonna 2020 laaditun tietojärjestelmäkartan mukaan käytössä n. 50 tietojärjestelmää. Näistä n. 30 järjestelmää vaatii kirjautumisen, eli niissä toteutuu 17§ mukainen kohta *Viranomaisen on huolehdittava, että sen tietojärjestelmien käytöstä ja niistä tehtävistä tietojen luovutuksista kerätään tarpeelliset lokitiedot, jos tietojärjestelmän käyttö edellyttää tunnistautumista tai muuta kirjautumista.* (Museovirasto 2020). Yllä mainitussa lukumäärässä on mukana sekä järjestelmiä, jotka vaativat kaikilta käyttäjiltä kirjautumisen, että järjestelmiä, joiden ylläpito- tai hallintasovellus vaatii kirjautumisen.

Nykytilanteen kartoittamiseksi tässä tutkimuksessa on selvitetty Museoviraston tietojärjestelmistä seuraavat asiat:

1. Vaatiiko järjestelmä kirjautumista tai muuta tunnistautumista?
2. Onko Museovirasto järjestelmän omistaja?
3. Onko Museovirasto järjestelmässä ylläpidettävän henkilörekisterin rekisterinpitäjä?
4. Kerätäänkö tietojärjestelmästä lokia järjestelmän käytöstä?
5. Luovutetaanko tietojärjestelmästä tietoja?
6. Kerätäänkö tietojärjestelmästä lokia, jota voi käyttää virheiden selvittämiseen?
7. Kerätäänkö tietojärjestelmästä lokia tietojen luovuttamisesta?
8. Missä loki sijaitsee?

Kysymysten avulla on tarkoitus saada rajattua tietojärjestelmien joukkoa, joita tiedonhallintalain lokien keräämisen vaatimus koskee ja joissa Museovirasto on vastuussa lokien keräämisen järjestämisestä.

Yllä mainitut tiedot esitetään taulukkona liitteessä 1. Liitteen tietojärjestelmäkohtaiset tiedot ovat salassa pidettävä ja siksi opinnäytetyön julkisessa versiossa on mukana liite, josta järjestelmät on esitetty numerokoodeina. Taulukossa järjestelmät on lisäksi ryhmitelty tietojärjestelmäkartan mukaisesti. Ryhmät ovat esimerkiksi ”taloushallinnon järjestelmät”, henkilöstöhallinnon järjestelmät” jne, mutta opinnäytetyössä ryhmät ilmaistaan kirjainkoodeina.

Ensimmäinen kysymys koskien järjestelmään kirjautumista on oleellinen, koska se on tiedonhallintalain asettama selkeä kriteeri, että lokia vaaditaan. Osa järjestelmistä on sellaisia, joita osa käyttäjistä käyttää kirjautumatta ja osa, vähintään ylläpitäjä, käyttävät kirjautuen. Tällaisissa tapauksissa kohtaan on kirjattu ”kyllä”.

Kysymykset 2 ja 3 liittyvät siihen, että osa järjestelmistä on jonkun muun tahon ylläpitämiä. Tällaisia ovat mm. Palkeiden (Valtion talous- ja henkilöstöhallinnon palveluskeskus) ylläpitämät järjestelmät, Museoliiton ylläpitämä kokoelmahallintajärjestelmä tai opetus- ja kulttuuriministeriön ylläpitämä työryhmäalustapalvelu eDuuni. Tiedonhallintalain vaatimus lokittamisesta velvoittaa sitä tiedonhallintayksikköä, jonka omistama järjestelmä on. Toisaalta tieto järjestelmissä kuuluu järjestelmää tiedon tallentamiseen käyttävälle organisaatiolle. Tiedontallentajaorganisaatio voi olla esimerkiksi henkilötietojen osalta rekisterin pitäjä yksinään tai yhdessä muiden järjestelmän käyttäjätahojen kanssa. Tiedonhallintalautakunnan suositus tietoturvallisuusnäkökohtien huomioimisesta muistuttaa, että järjestelmän tai siinä olevien tietojen omistajan tulee varmistua, että toimittaja on huolehtinut kaikista teknisistä tietoturvallisuustoimista. Samoin täytyy olla sovittuna, miten ja missä tilanteissa lokitietoja saadaan nähtäville järjestelmästä, jota pitää yllä joku ulkopuolinen taho.

Kysymys tietojen luovuttamisesta tarkoittaa kaikkia mahdollisia tapoja, joilla järjestelmässä olevia tietoja saatetaan luovuttaa joko kansalaisille tai toisille viranomaisille. Luovutuksen tapa voi olla järjestelmän sisällä tiedon jakaminen, rajapinnan kautta jakaminen tai erillinen järjestelmän ulkopuolella tapahtuva tiedonluovutus. Tässä kohdassa ei kuitenkaan huomioida avoimena datana jaettavaa tietoa tai julkisilla www-sivuilla julkaistavaa tietoa. Tiedon luovuttamisen tilannetta tarkastellaan, koska tiedonhallintalaki mainitsee tietojen luovutuksen yhdeksi lokien käyttötarkoitukseksi. Useissa tapauksissa tietoa luovutetaan tietopyyntöjen kautta, joiden tiedot kirjautuvat asianhallintajärjestelmään sopimuksina tai asiakaspalvelun rekisteriin tietopyyntöinä. Jos järjestelmästä ei luovuteta tietoja, viimeiseen kohtaan luovutuslokin tallentamisesta ei ole otettu kantaa, vaan kohta on merkitty taulukkoon harmaalla.

Viimeinen kysymys liittyy lokin sijaintiin suhteessa järjestelmään. Kysymyksen tarkoituksena on selvittää, kuinka monessa tapauksessa loki tallentuu järjestelmään ja milloin lokitiedot puolestaan tallentuvat tai siirtyvät muuhun sijaintiin. Nykytilanteen tarkastelu osoitti, että muutoslokit tallentuvat kaikissa tapauksissa tietokantaan eikä tietoja siirretä ulkoiseen sijaintiin. Muiden lokien osalta tilanne on samankaltainen, muutamia isoimpia järjestelmiä lukuun ottamatta.

Nykytilanteessa osa Museoviraston tietojärjestelmistä täyttää tiedonhallintalain vaatimuksen jo valmiiksi. Osassa puolestaan on tilanne, että lokia kerätään jostain asiasta, mutta ei välttämättä laissa mainituista kaikista käyttötarkoituksista, mm. tietojen luovuttamisesta. Joukossa on myös järjestelmiä, joista ei kerätä lokitietoja lainkaan. Valtaosassa järjestelmiä syntyy lokitietoa tietojen tallentamisesta tai poistamisesta, mutta pelkästä tietojen katselusta ei välttämättä kerry lokia. Yllättävänkin monessa järjestelmässä tilanne on, ettei lokien tallentumisen tilanne ole aivan selvillä. Tietojen luovuttamisesta saatetaan pitää kirjaa järjestelmän ulkopuolella ja varsinaista järjestelmälokia luovutuksista ei synny. Tämä kuitenkin riittää tiedonhallintalain vaatimuksen täyttämiseksi, sillä laki ei edellytä, että loki olisi digitaalinen.

Museovirasto käyttää lukuisia järjestelmiä, joiden ylläpitäjänä ja omistajana on joku muu viranomainen tai muu taho. Nykytilanteen tarkastelu osoitti, että näiden järjestelmien osalta tiedot lokien keräämisestä ja tallentamisesta ovat huonot. Kuitenkin järjestelmän käyttäjän ja sinne tietoja tallentavan (eli Museoviraston) tulisi varmistua tietoturvallisuustoimenpiteiden toteutumisesta, vaikka toteutusvastuu sinänsä kuuluukin järjestelmän omistajalle.

5.2 Tarveanalyysi

Koska lokien hallinta ja sen suunnittelu on työlästä, lokitietoja ei kannata kerätä turhan takia. Myös tiedonhallintalain pykälässä on mainittu tarveharkinta: --- *tietojärjestelmien käytöstä ja niistä tehtävistä tietojen luovutuksista kerätään* ***tarpeelliset*** *lokitiedot.* Lakipykälää voinee tulkita siten, että voi olla myös tilanteita, joissa lokin kerääminen ei ole lainkaan tarpeellista tai joissa lokitietojen ei tarvitse olla kovin laajoja. Tätä tukee tiedonhallintalautakunnan antama suositus lokien hallinnasta. Suosituksessa todetaan, että lokien kerääminen on sidottu tarpeellisuuteen, jota puolestaan tulee arvioida suhteessa riskeihin (Tiedonhallintalautakunta 2020, 38).

Lokien keräämisen tarpeellisuutta arvioitaessa painotettavia asioita ovat tiedonhallintalautakunnan suosituksen perusteella se, onko järjestelmässä salassa pidettäviä tietoja, luovutetaanko järjestelmästä salassa pidettäviä tietoja tai henkilötietoja, tarvitaanko lokeja virheiden selvittämiseen, onko tarpeena varmistaa yksilön oikeuksien, etujen tai velvollisuuksien toteutumista tai oikeusturvan tai virkavastuun toteutumista. (Tiedonhallintalautakunta 2020, 38).

Tässä tutkimuksessa on tehty tarpeen arvioimista varten matriisi (liite 2), jonka avulla yllä mainittuja seikkoja voi arvioida järjestelmä kerrallaan. Laista ja suosituksesta on poimitettu

niissä mainitut kriteerit tai muuten painotetut kohdat ja annettu niille pistemäärä. Pistemäärän on tarkoitus toimia nopeana karkean tason työkaluna järjestelmien välillä mahdollisesti tehtävää priorisointia ajatellen. Luonnollisesti järjestelmäkohtaisesti on kuitenkin pohdittava tilanne pistemäärästä huolimatta erikseen. Taulukossa lasketaan järjestelmille pisteitä seuraavasti:

- vaatii kirjautumisen 1 p
- sisältää henkilötietoja 2 p
- sisältää turvallisuusluokiteltuja tietoja 2 p
- sisältää muuten salassa pidettäviä tietoja 2 p
- järjestelmästä luovutetaan tietoja katselu- tai rajapintayhteyden kautta (muuta kuin avointa dataa) 1 p
- järjestelmästä luovutetaan katselu- tai rajapintayhteyden kautta henkilö- tai salassa pidettäviä tietoja 2 p

Koska tiedonhallintalautakunnan suositus ja tiedonhallintalain valmisteluaineistot antavat ymmärtää, että henkilötietojen ja muiden luottamuksellisten tietojen suojaaminen on lokien hallinnan päätarkoituksia, pisteytysmallissa annetaan näistä seikoista enemmän pisteitä (2 p). Kirjautumisen vaatimus, tai se että järjestelmästä luovutetaan tietoja, pisteytetään yhden pisteen arvoisesti.

Tarveanalyysiin on valittu nykytilan kartoitustaulukosta ne järjestelmät, jotka edellyttävät kirjautumista tai muuta tunnistautumista ja ovat joko Museoviraston omistamia tai joissa Museovirasto toimii henkilörekisterin pitäjänä.

Taulukossa arvioidaan ensisijaisesti, minkä järjestelmien kohdalla tiedonhallintalain vaatimus lokien keräämisestä toteutuu vahvimmin. Tämän lisäksi on syytä muistaa, että järjestelmää hallinnoivalla taholla voi olla muita syitä kerätä lokitietoja. Esimerkiksi lähes kaikissa tietokannoissa muutosloki on oleellinen ja tärkeä mm. tietojen eheyden varmistamisen kannalta. Tietokantojen muutoslokit eivät kuitenkaan kuulu tiedonhallintalaissa painottuviin asioihin, joten muutoslokeja ei huomioida taulukon ”TiHL”-pisteissä. Tiedonhallintalaissa mainitaan, että lokitietoja voidaan hyödyntää virheiden selvittämiseen, mutta tiedonhallintalain kokonaisuutta ajatellen järjestelmävirheiden selvittäminen ei ole lakiin keskeisesti kuuluva seikka. Näin ollen myös virhelokit on muutoslokien ohella mainittu taulukossa ”Muut tarpeet” osassa. Samoin on pääsynvalvontalokien laita. Organisaation arvioidessa lokien hallinnan tarpeita, on kaikki mahdolliset tarpeet otettava huomioon tiedonhallintalain vaatimusten lisäksi. Esimerkiksi liitteessä 2 näkyvässä esimerkissä Museoviraston järjestelmissä 5 ja 19 on vain vähäinen tarve tiedonhallintalain mukaiselle lokien hallinnalle (vain 1 piste/ max 8 pistettä), mutta muunlainen lokien hallinta on tunnistettu tärkeäksi asiaksi (5 p/ max 6 p).

Tiedonhallintalaissa painotetut seikat viittaavat erityisesti siihen, että tietojen katselusta ja käytöstä tulisi tallentua lokitietoja, samoin kuin tiedon luovutuksista. Nämä liittyvät siihen, että voidaan todentaa tietojen vastaanottajan oikeudellinen peruste saada tietoja ja varmistaa virkavastuun ja oikeusturvan toteutuminen. Tarpeenarviointitaulukossa on ajatuksena, että jos järjestelmä saa korkeat ”TiHL-pisteet”, siihen tulisi todennäköisesti liittyä käyttöloki ja/tai luovutusloki.

5.3 Toimenpidesuositukset

Tiedonhallintalain vaatimusten, tiedonhallintalautakunnan suositusten ja nykytilanteen valossa Museoviraston kannattaisi ensimmäisenä toimenpiteenä aloittaa lokiperiaatteiden laatimisesta. Nykytilan kartoitus osoittaa, että Museovirastossa lokien hallintaa ei ole järjestetty suunnitelmallisesti. Lokiperiaatteissa tulisi mm. määrittää, onko tiedonhallintalaki yksinään lokien hallintaa määrittävä tekijä, vai kerätäänkö lokeja myös muita tarkoituksia varten.

Järjestelmien osalta tulisi suorittaa kattava tarveanalyysi, jossa käydään järjestelmät läpi ja huomioidaan painotettavat asiat (onko järjestelmässä salassa pidettäviä tietoja, luovutetaanko niitä jne). Tarveanalyysin voi tehdä esimerkiksi tässä työssä esitetyn excel-mallin mukaisesti.

Kun lokien keräämisen tarpeet on selvitetty, lokien keruu ja lokitietojen hallinta tulee toteuttaa järjestelmiin, mikäli niin ei vielä ole. Lokitietojen käsittelyn osalta tulee päättää käsittelyn perusteista ja analysoinnin tarpeesta. Jos kerättävät lokitiedot sisältävät henkilötietoja, rekisteröityjä tulee informoida ja käsittelyn säännöistä tulee laatia seloste. Kun kokonaisuus lokien keräämisen osalta on selvillä, tulee pohtia miten ja missä lokitietoja säilytetään. Keskitetty lokijärjestelmä ei ole välttämättömyys, mutta sen mahdollisuus on hyvä selvittää, mikäli lokeja kerätään paljon.

Huomioitava on, että jatkossa uusissa tietojärjestelmähankinnoissa tulee jo hankintavaiheessa arvioida mikä lokien keräämisen tarve kyseisessä järjestelmässä tulee olemaan. Tietynlaiseen välitilaan jää järjestelmiä, joiden osalta Museovirasto ei ole järjestelmän omistaja tai sen teknisestä ylläpidosta vastaava taho. Silti järjestelmässä olevat tiedot ovat Museoviraston omistamia ja järjestelmässä saatetaan esimerkiksi ylläpitää henkilötietoja. Näiden järjestelmien osalta tiedonhallintalain tarkoittaman lokitietojen keräämisen järjestäminen ei ole Museoviraston vastuulla, mutta Museoviraston tulisi silti omien tietojensa turvallisuuden vuoksi varmistua siitä, kuinka lokien kerääminen on järjestetty ja miten tiedot tarvittaessa on saatavissa käyttöön.

6 Pohdinta

Tiedonhallintalain 17§ tekstin niukkuus ja pienehkö ristiriitaisuus lain perusteluiden kanssa aiheuttanee tiedonhallintayksiköille mietittävää. Lakitekstin mukaan ainoa lokien keräämiseen velvoittava tekijä on se, että järjestelmään täytyy kirjautua. Perustelu- ja suositusteksteissä sen sijaan painotetaan tarveharkintaa, henkilötietoja ja salassa pidettävien tietojen luovuttamista. Ristiriitaisuuden vuoksi painoarvoa tullee saamaan tiedonhallintalautakunnan suositukset. Suosituksissa onkin otettu paljolti lain perusteluita tukeva linja ja suosituksissa painotetaan mm. lokien keräämisen tarpeenmukaisuutta. Kuten suosituksessakin todetaan, kenenkään ei ole syytä käynnistää lokien keruuta suunnittelemattomasti ilman tarpeeseen perustuvaa suunnittelua (Tiedonhallintalautakunta 2020, 38). Tämän painotuksen, joka lakitekstin taustalla on, esiin tuominen ja havaitseminen on yksi tämän tutkimuksen tärkeimpiä havaintoja.

Toisena tutkimuksen tärkeänä johtopäätöksenä voi todeta, että lokien hallinta edellyttää ennen kaikkea tarpeellisuuden selvittämistä ja hyvää suunnittelua. Tarpeellisuuden arvioimista varten täytyy olla tiedossa syyt, miksi lokia kerätään ja tunnettava erilaiset lokityypit. Lokien keräämisen suunnittelu sinänsä vastaa lopulta melkein minkä tahansa tiedon keräämisen suunnittelua ja siinä pätevät samat periaatteet kuin yleisestikin koskien esimerkiksi tiedon elinkaarta.

Lisäksi tutkimus toi näkyville sen seikan, että tiedonhallintalain määrittelemät tarpeet eivät ole ainoita syitä lokien keräämiselle. Vaikka tiedonhallintalain tarkoittamat seikat eivät olisi oleellisia jossain tietojärjestelmässä, järjestelmässä voi olla silti muita syitä lokien keräämiselle. Tiedonhallintalaki ei esimerkiksi ota kantaa muutoslokien keräämiseen, vaikka ne ovat erityisen tärkeitä monissa tietovarannoissa.

Aiemmin mainituista ristiriitaisuuksista huolimatta tiedonhallintalain lokien keräämistä koskeva pykälä tukee lain yleistavoitteita mm. tietojen saamisesta uudelleen käytettäväksi tietoturvallisesti. Lokien hallinnan ottamisella osaksi tiedonhallintalain vaatimuksia on varmasti hyviä seurauksia sen kautta, että julkishallinnon organisaatioiden on pakko miettiä omaa lokipolitiikkaansa. Näin ei luultavasti tapahtuisi ilman lain vaatimusta. Myös tämän työn empiirinen osuus osoittaa, että lokien hallinta jää helposti toteuttamatta tai puolitiehen.

Henkilötietojen suojaus ja tietovuodoilta suojautuminen ovat hyvin ajankohtaisia asioita ja oletettavaa on, että julkishallinnon organisaatiotkin suhtautuvat näihin asioihin entistä tarkemmin. Tietoturvallisuuden huomioiminen tiedonhallinnassa tulee varmasti olemaan yksi

lähitulevaisuuden tärkeimmistä asioista, joihin tiedonhallinnan kehittämisessä suuntaudutaan. Lokien hallintaa koskeva pykälä tiedonhallintalaissa on siten hyvin paikallaan.

Tämän tutkimuksen myötä oma ymmärrykseni tutkimuksen kohteena olevasta aiheesta on luonnollisesti lisääntynyt. Lisäksi olen löytänyt julkishallinnon tiedonhallintaa koskevia ohjeistuksia, joita en ole aiemmin tuntenut. Nämä tulevat jäämään tärkeiksi lähteiksi myös tulevissa töissä. Lakitekstin tutkiminen oli uusi kokemus, ja myös sen osalta erilaisten lähteiden löytäminen oli hyödyllistä ja antoisaa. Opinnäytetyöni on ollut tutkimustyyppinen ja mielestäni olen ainakin itse saanut sen myötä paljon lisätietoa lokien hallinnasta ja ymmärtänyt sen kokonaisuuden, johon tiedonhallintalain lokien keräämistä koskeva pykälä liittyy. Toivon mukaan tieto hyödyttää myös tutkimuksen muita mahdollisia lukijoita. Seuraavana haasteena on viedä saatu tieto käytäntöön. Käytäntöön viemisen työtä riittää vielä paljon, eikä siihen ole tämän tutkimustyön yhteydessä ollut mahdollisuuksia. Tavoitteena on vuoden 2021 aikana saada hyödynnettyä tutkimuksessa selvinneitä asioita käynnössä.

Jälkikäteen ajatellen tutkimuksen empiirinen osa olisi voinut rajautua tarkemmin vain joihin järjestelmiin, jolloin niistä olisi voinut selvittää yksityiskohtaisempia asioita ja tehdä myös konkreettisempia toimenpidesuosituksia. Toisaalta koko kokonaisuuden karkea läpikäyminen on luonut hyödyllistä kokonaiskuvaa asiasta.

7 Lähteet

Ahonen P., Seppänen J. & Pärssinen J. 2019. KYBER-ENE Energia-alan kyberturvaaminen 1-2. Julkisten tulosten kooste. VTT Technology 353. Luettavissa: https://www.vttresearch.com/sites/default/files/pdf/technology/2019/T353.pdf. Luettu: 8.11.2020.

Eduskunta 2020. Asian käsittelytiedot HE 284/2018 vp. Www-sivusto. Luettavissa: https://www.eduskunta.fi/FI/vaski/KasittelytiedotValtiopaivaasia/Sivut/HE_284+2018.aspx. Luettu: 4.9.2020.

Eräkaski, M. (2017). Tiedon elinkaaren hallinta. Esitysmateriaali 18.5.2017. Luettavissa: https://avoinhallinto.fi/assets/files/2017/07/Tiedon-elinkaarenhallinta-18.5.2017-Mikko-Er%C3%A4kaski-kansallisarkisto.pdf. Luettu: 24.11.2020.

JUHTA 2017a. Julkisen hallinnon tietohallinnon neuvottelukunta. Lokittaminen ja riskienhallinta, työpajatilaisuuden materiaali 18.8.2017. Luettavissa: https://vm.fi/documents/10623/4914009/JUHTA+tietosuoja+lokitus+180817/14baba1f-98e5-4887-b3e4-65a98f465e75. Luettu 29.9.2020.

JUHTA 2017b. Julkisen hallinnon tietohallinnon neuvottelukunta. Tietosuojan osoitusvelvollisuutta edistävät työpajatilaisuudet 18.8.2018. Luettavissa: https://vm.fi/juhta-vahti-yhteishankkeiden-materiaalit##2.%20Lokittaminen%20ja%20riskienhallinta%20(18.8.2017). Luettu: 11.10.2020.

Kansallinen terveysarkisto (2013). Potilastietojärjestelmien käyttötapaukset. Liite 5. Vaatimukset potilastietojärjestelmien käyttölokeille. Luettavissa: https://www.kanta.fi/documents/20143/107839/Potilastietoj%C3%A4rjestelmien+k%C3%A4ytt%C3%B6tapaukset+Liite5+Vaatimukset+potilastietoj%C3%A4rjestelmien+k%C3%A4ytt%C3%B6lokeille.pdf/4e5675ec-1cf3-7baf-a5ad-aeee66a10950. Luettu: 29.9.2020.

Kivivasara S. (2019). Vastine hallituksen esityksestä 284/2018 vp laiksi julkisen hallinnon tiedonhallinnasta sekä eräiksi siihen liittyviksi laeiksi saaduista asiantuntijalausunnoista. 6.2.2019. Luettavissa: https://www.eduskunta.fi/FI/vaski/JulkaisuMetatieto/Documents/EDK-2019-AK-244888.pdf. Luettu 4.9.2020.

Korhonen V. (2019). Lausunto hallituksen esityksestä laiksi julkisen hallinnon tiedonhallinnasta ja eräiksi siihen liittyviksi laeiksi. Lausunto 5.2.2019. Luettavissa: https://www.eduskunta.fi/FI/vaski/JulkaisuMetatieto/Documents/EDK-2019-AK-243666.pdf. Luettu 4.9.2020.

Korkein hallinto-oikeus 2020. Vuosikirjapäätökset. KHO 2014:69. Luettavissa: https://kho.fi/fi/index/paatoksia/vuosikirjapaatokset/vuosikirjapaatos/1399372335852.html. Luettu: 29.10.2020.

Kyberturvallisuuskeskus 2020. Näin keräät ja käytät lokitietoja. Luettavissa: https://www.kyberturvallisuuskeskus.fi/fi/ajankohtaista/ohjeet-ja-oppaat/nain-keraat-ja-kaytat-lokitietoja. Luettu: 22.6.2020.

Luomajärvi J. 2019. Eduskunnan hallintovaliokuntapyyntö kirjallisesta asiantuntijalausunnosta HE 284/2018 vp. Lausunto 28.1.2019. Luettavissa: https://www.eduskunta.fi/FI/vaski/JulkaisuMetatieto/Documents/EDK-2019-AK-241659.pdf. Luettu: 4.9.2020.

Mikkelä H. (2019). Hallituksen esitys eduskunnalle laiksi julkisen hallinnon tiedonhallinnasta sekä eräiksi siihen liittyviksi laeksi. Lausunto 6.2.2019. Luettavissa: https://www.eduskunta.fi/FI/vaski/JulkaisuMetatieto/Documents/EDK-2019-AK-243976.pdf. Luettu 4.9.2020.

Museovirasto (2020). Tietojärjestelmäkartta. Ei julkaistu.

Pölönen P. (2018). Lausunto luonnoksesta hallituksen esitykseksi eduskunnalle laiksi julkisen hallinnon tiedonhallinnasta sekä eräiksi siihen liittyviksi laeiksi. Luettavissa: https://www.oikeusasiamies.fi/r/fi/ratkaisut/-/eoar/4254/2018. Luettu: 4.9.2020.

Reaveley P. (2010). Navigation and Logbooks in the Age of Sail. Www-sivu. Luettavissa: https://www.usna.edu/Users/oceano/pguth/website/shipwrecks/logbooks_lesson/logbooks_lesson.htm. Luettu 7.11.2020.

Techopedia 2020. Error log. www-sivu. Luettavissa: https://www.techopedia.com/definition/26306/error-log. Luettu: 14.9.2020.

Tiedonhallintalautakunta 2020. Suosituskokoelma tiettyjen tietoturvallisuussäädösten soveltamisesta. Valtiovarainministeriön julkaisuja 2020:21. Helsinki. Luettavissa: http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/162150/VM_2020_21.pdf?sequence=1&isAllowed=y. Luettu 22.6.2020.

Tolonen J. 2017. Lokitiedon merkitys käyttöjärjestelmän asennuksessa SCCM-ympäristössä. Opinnäytetyö. Kajaanin ammattikorkeakoulu. Luettavissa: https://www.theseus.fi/bitstream/handle/10024/127672/Tolonen_Jarkko.pdf?sequence=1&isAllowed=y. Luettu: 29.9.2020.

Valtiovarainministeriö 2009. Lokiohje. Valtiovarainministeriön julkaisuja 3/2009. Luettavissa: https://vm.fi/o/dms-portlet/document/0/371107. Luettu: 29.11.2020.

Viestintävirasto 2016. Lokien keräys ja käyttö. Opas lokitietojen keräämiseen ja hyödyntämiseen. Ohje 4/2016. Helsinki. Luettavissa: https://www.kyberturvallisuuskeskus.fi/sites/default/files/media/file/Lokitusohje.pdf. Luettu: 22.6.2020.

Vähä-Erkkilä A. (2019). Hallintovaliokunta ke 30.01.2019 / HE 284/2018 vp / Asiantuntijapyyntö. Lausunto 29.1.2019. Luettavissa: https://www.eduskunta.fi/FI/vaski/JulkaisuMetatieto/Documents/EDK-2019-AK-241940.pdf. Luettu: 4.9.2020.

Wallin A. (2019). Hallituksen esitys eduskunnalle laiksi julkisen hallinnon tiedonhallinnasta sekä eräiksi siihen liittyviksi laeiksi (284/2018 vp); lausunto eduskunnan perustuslakivaliokunnalle. Lausunto 6.2.2019. Luettavissa: https://www.eduskunta.fi/FI/vaski/JulkaisuMetatieto/Documents/EDK-2019-AK-243850.pdf. Luettu: 4.9.2020.

8 Liitteet

Liite 1. Lokituksen nykytilanteen kartoitus järjestelmäkohtaisesti. Julkaistavasta versiosta on poistettu järjestelmien nimet.

Tietojärjestelmä kartan mukainen ryhmittely	*Tietojärjestelmä*	*Vaatii kirjautumisen*	*Onko Museovirasto järjestelmän omistaja*	*Onko Museovirasto järjestelmässä pidettävän henkilörekisterin ylläpitäjä*	*Lokityyppi*	*Loki mahdollistaa virheiden selvittämisen (Kyllä/Ei /Osittain)*	*Luovutetaan tietoa*	*Loki tietojen luovuttamisesta*	*Lokin sijainti*
RYHMÄ A									
	Järjestelmä 1	Kyllä	Kyllä		Ei		Ei		
	Järjestelmä 2	Kyllä	Kyllä	Kyllä	Selvitettävä	Selvitettävä	Ei		
	Järjestelmä 3	Kyllä	Kyllä		Selvitettävä	Selvitettävä	Ei		
	Järjestelmä 4	Kyllä	Kyllä	Kyllä	Ei		Ei		
	Järjestelmä 5	Kyllä	Kyllä		Yhteysloki	Kyllä	Ei		Järjestelmässä
RYHMÄ B									
	Järjestelmä 6	Kyllä	Ei	Kyllä	Käyttöloki ja muutosloki	Kyllä	Kyllä	Järjestelmän ulkopuolella	Järjestelmässä
RYHMÄ C									
	Järjestelmä 7	Kyllä	Ei	Kyllä	Muutosloki	Kyllä	Rajapinnan kautta	Kyllä	Järjestelmässä
	Järjestelmä 8	Kyllä/Ei	Kyllä	Kyllä	Muutosloki	Osittain	Tietopyyntöihin perustuen	Järjestelmän ulkopuolella	Järjestelmässä
	Järjestelmä 9	Kyllä	Kyllä	Kyllä	Muutosloki	Osittain	Ei		Järjestelmässä
	Järjestelmä 10	Kyllä	Kyllä		Muutosloki	Osittain	Ei	Järjestelmän ulkopuolella	Järjestelmässä
	Järjestelmä 11	Kyllä	Kyllä	Kyllä	Muutosloki	Osittain	Järjestelmän käyttäjille	Ei	Järjestelmässä
	Järjestelmä 12	Kyllä	Kyllä	Kyllä	Muutosloki	Osittain	Ei		Järjestelmässä
	Järjestelmä 13		Kyllä		Selvitettävä	Selvitettävä	Ei		
	Järjestelmä 14	Kyllä	Kyllä		Ei		Ei		
	Järjestelmä 15	Kyllä/Ei	Ei	Kyllä	Selvitettävä	Selvitettävä	Ei		
RYHMÄ D									
	Järjestelmä 16	Kyllä/Ei	Kyllä		Ei		Ei		
	Järjestelmä 17	Kyllä	Kyllä	Kyllä	Ei	Ei	Ei		
	Järjestelmä 18	Kyllä	Kyllä	Kyllä	Muutosloki	Osittain	Järjestelmän käyttäjille	Ei	Järjestelmässä
	Järjestelmä 19	Kyllä/Ei	Kyllä		Muutosloki, muut selvitettävä	Selvitettävä	Ei		Järjestelmässä
	Järjestelmä 20	Kyllä/Ei	Kyllä		Muutosloki	Osittain	Tietopyyntöihin perustuen	Järjestelmän ulkopuolella	Järjestelmässä
	Järjestelmä 21	Kyllä/Ei	Kyllä	Kyllä	Muutosloki	Osittain	Tietopyyntöihin perustuen	Järjestelmän ulkopuolella	Järjestelmässä
	Järjestelmä 22	Ei	Kyllä		Virheloki	Kyllä	Rajapinnan kautta	Ei	Järjestelmässä
RYHMÄ E									
	Järjestelmä 23	Kyllä	Kyllä		Muutosloki		Ei		Järjestelmässä
	Järjestelmä 24	Ei	Ei		Selvitettävä	Selvitettävä	Ei		Selvitettävä
	Järjestelmä 25	Kyllä	Kyllä	Kyllä	Ei		Ei		
RYHMÄ F									
	Järjestelmä 26	Kyllä	Ei		Selvitettävä	Selvitettävä	Ei		Selvitettävä
	Järjestelmä 27	Kyllä	Kyllä		Selvitettävä	Selvitettävä	Ei		Selvitettävä
	Järjestelmä 28	Kyllä	Kyllä		Selvitettävä	Selvitettävä	Ei		Selvitettävä
	Järjestelmä 29	Kyllä	Kyllä	Kyllä	Selvitettävä	Selvitettävä	Ei		Selvitettävä
	Järjestelmä 30	Kyllä	Ei	Kyllä	Selvitettävä	Selvitettävä	Ei		Selvitettävä
RYHMÄ G									
	Järjestelmä 31	Kyllä	Kyllä	Kyllä	Selvitettävä	Selvitettävä	Ei		Selvitettävä
	Järjestelmä 32	Kyllä	Ei	Kyllä	Selvitettävä	Selvitettävä	Järjestelmän käyttäjille	Kyllä	Selvitettävä
RYHMÄ H									
	Järjestelmä 33	Kyllä	Ei	Kyllä	Selvitettävä	Selvitettävä	Järjestelmän käyttäjille	Kyllä	Selvitettävä
	Järjestelmä 34	Kyllä/Ei	Kyllä	Kyllä	Pääsynhallinta- ja muutoslokit	Kyllä	Ei		Järjestelmässä

Liite 2. Taulukko lokituksen tarpeen arvioimista varten.

		TiHL tarpeet							Muut tarpeet organisaatiossa			
RYHMÄ	**Järjestelmän nimi**	**Järjestelmä edellyttää kirjautumista, Kyllä = 1 p**	**Järjestelmä sisältää henkilötietoja, Kyllä = 2 p**	**Järjestelmä sisältää turvaluokiteltuja tietoja, Kyllä = 2 p**	**Järjestelmä sisältää muita salassa pidettäviä tietoja, Kyllä = 2 p**	**Järjestelmästä luovutetaan katselu- tai rajapintayhteyden kautta tietoa joka ei ole avointa dataa, Kyllä = 1 p**	**Järjestelmästä luovutetaan katselu- tai rajapintayhteyden kautta tietoa joka sisältää henkilö- tai salassapidettäviä tietoja, Kyllä = 2 p**	**TiHL mukaiset pisteet, yht**	**Järjestelmävirheiden selvittäminen oleellista, 0-2 (0=ei tarvetta, 1= mahdollinen tarve, 2 = tärkeää)**	**Tietokannan muutosloki oleellinen, 0-2 (0=ei tarvetta, 1= mahdollinen tarve, 2 = tärkeää)**	**Pääsynvalvontaan liittyvä valvonta oleellista, 0-2 (0=ei tarvetta, 1= mahdollinen tarve, 2 = tärkeää)**	**Oma tarve yht**
A	Järjestelmä 1	Kyllä	Kyllä	Ei	Ei	Ei	Ei	3	2	2	2	6
A	Järjestelmä 2	Kyllä	Kyllä	Ei	Ei	Ei	Ei	3	1	1	2	4
A	Järjestelmä 3	Kyllä	Kyllä	Ei	Ei	Ei	Ei	3	1	1	1	3
A	Järjestelmä 4	Kyllä	Kyllä	Ei	Ei	Ei	Ei	3	2	2	2	6
A	Järjestelmä 5	Kyllä	Ei	Ei	Ei	Ei	Ei	1	2	1	2	5
C	Järjestelmä 8	Kyllä	Kyllä	Ei	Ei	Ei	Ei	3	1	2	0	3
C	Järjestelmä 9	Kyllä	Ei	Ei	Ei	Ei	Ei	1	1	2	1	4
C	Järjestelmä 10	Kyllä	Kyllä	Ei	Ei	Ei	Ei	3	1	2	1	4
C	Järjestelmä 11	Kyllä	Kyllä	Ei	Ei	Kyllä	Kyllä	6	1	2	1	4
C	Järjestelmä 12	Kyllä	Kyllä	Ei	Kyllä	Ei	Ei	5	0	1	1	2
C	Järjestelmä 13	Kyllä	Ei	Ei	Ei	Kyllä	Ei	2	2	0	0	2
C	Järjestelmä 14	Kyllä	Ei	Ei	Ei	Ei	Ei	1	1	2	0	3
D	Järjestelmä 17	Kyllä	Kyllä	Ei	Ei	Ei	Ei	3	1	2	1	4
D	Järjestelmä 18	Kyllä	Kyllä	Ei	Ei	Ei	Ei	3	1	2	1	4
D	Järjestelmä 19	Kyllä	Ei	Ei	Ei	Ei	Ei	1	2	2	1	5
D	Järjestelmä 20	Kyllä	Kyllä	Ei	Ei	Kyllä	Kyllä	6	1	2	0	3
D	Järjestelmä 21	Kyllä	Kyllä	Ei	Ei	Ei	Ei	3	1	2	0	3
E	Järjestelmä 23	Kyllä	Ei	Ei	Ei	Ei	Ei	1	1	1	1	3
E	Järjestelmä 25	Kyllä	Kyllä	Ei	Ei	Ei	Ei	3	1	1	0	2
F	Järjestelmä 27	Kyllä	Kyllä	Ei	Ei	Ei	Ei	3	1	0	2	3
F	Järjestelmä 28	Kyllä	Kyllä	Ei	Kyllä	Ei	Ei	5	1	0	2	3
F	Järjestelmä 29	Kyllä	Kyllä	Ei	Ei	Kyllä	Ei	4	1	0	1	2
												0
G	Järjestelmä 31	Kyllä	Kyllä	Ei	Kyllä	Kyllä	Kyllä	8	1	2	2	5
H	Järjestelmä 34	Kyllä	Kyllä	Ei	Ei	Ei	Kyllä	5	2	2	2	6