

Aleksi Knuutila, Vesa Kokkonen, Heikki Sundquist, Ossi
Kuittinen & Salla Thure

MyData muutosvoimana: Julkishallinnon henkilötiedon ihmiskeskeisen hyödyntämisen mallit ja vaikutukset

Syyskuu 2017

Valtioneuvoston selvitys-
ja tutkimustoiminnan
julkaisusarja 61/2017

KUVAILULEHTI

Julkaisija ja julkaisu-aika	Valtioneuvoston kanslia, 1.9.2017		
Tekijät	Aleksi Knuutila, Vesa Kokkonen, Heikki Sundquist, Ossi Kuittinen, Salla Thure		
Julkaisun nimi	MyData muutosvoimana: Julkishallinnon henkilötiedon ihmiskeskeisen hyödyntämisen mallit ja vaikutukset		
Julkaisusarjan nimi ja numero	Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 61/2017		
Asiasanat	henkilötieto, digitalisaatio, yksityisyys, vaikuttavuus		
Julkaisun osat	Sähköinen pdf (lisenssi CC-BY-SA 4.0)		
Julkaisu-aika	Syyskuu, 2017	Sivuja 83	Kieli Suomi

Tiivistelmä

Valtio sekä kunnat ylläpitävät tehtäviään varten henkilötietorekistereitä, joista monet ovat sisällöltään ainutlaatuisia ja monikäyttöisiä. Joillakin hallinnonaloilla henkilöitä koskevan tiedon hyödyntämiseen liittyvät säännöt ovat sallivia, mutta henkilötiedon käyttöä rajaavat laissa määritellyt käyttötarkoitukset sekä tarvittaviin lupiin liittyvät monimutkaiset käytännöt.

MyData-periaatteet kuvaavat ajatusta, että ihmisillä on määräysvaltaa heitä koskevaan tietoon. Niihin liittyy toimintamalli, jossa ihmisille tarjotaan käytännön keinoja hallita omien tietojensa käyttöä ja jakaa tietojaan eri palveluihin antamiensa lupien perusteella. Tällaiset keinot hyödyntää tietoa voisivat pienentää uusien palveluiden kehityskustannuksia ja tukea nopeiden kokeilujen tekemistä sekä yksityisellä että julkisella sektorilla. Samalla ne vahvistaisivat julkisten palveluiden ymmärrettävyyttä ja yksilöiden oikeuksia omaan tietoonsa.

MyData-periaatteet voidaan toteuttaa useammalla erilaisella tavalla. Raportissa vaihtoehtoja jäsenetään luokittelulla, joka sisältää viisi eri vaatimustasoa, sekä käytännön ratkaisulla näiden saavuttamiseen. Selvityksessä ehdotetaan, että alustoja ja työkaluja tuotettaisiin niin, että rekisterinpitäjät voisivat valita tilanteeseensa sopivat keinot tehdä henkilötieto saatavaksi. Julkishallinnon tulisi pyrkiä myös synnyttämään luottamusta ja helpottamaan yksilöiden sekä palveluntarjoajien kanssakäymistä, esimerkiksi yleisillä malleilla käyttöehdoille sekä yhteisillä ymmärrettävillä keinoilla kuvata henkilötiedon käytön tapoja.

Raportissa kuvataan MyData-periaatteiden soveltamisen vaikutuksia julkishallintoon ja muihin sidosryhmiin kolmen tapaus tutkimuksen kautta. Tiedon parempi saatavuus mahdollistaa yksilöityjä palveluita ja uudenlaisia sovelluksia, muun muassa ennaltaehkäisevän terveydenhuollon saralla.

Liite 1 Rekisteröityjen oikeudet eri tietotyyppien kohdalla

Tämä julkaisu on toteutettu osana valtioneuvoston vuoden 2017 selvitys- ja tutkimussuunnitelman toimeenpanoa (tietokayttoon.fi).

Julkaisun sisällöstä vastaavat tiedon tuottajat, eikä tekstisisältö välttämättä edusta valtioneuvoston näkemystä.

PRESENTATIONSBLAD

Utgivare & utgivningsdatum	Statsrådets kansli, 1.9.2017		
Författare	Aleksi Knuutila, Vesa Kokkonen, Heikki Sundquist, Ossi Kuittinen, Salla Thure		
Publikationens namn	Modeller och effekter av MyData: Mänsklig centrerad hantering av personuppgifter från den offentliga sektorn		
Publikationsseriens namn och nummer	Publikationsserie för statsrådets utrednings- och forskningsverksamhet 61/2017		
Nyckelord	Personuppgifter, digitalisering, integritet, konsekvensbedömning		
Publikationens delar	PDF (lisens CC-BY-SA 4.0)		
Utgivningsdatum	September, 2017	Sidantal 83	Språk Finska

Sammandrag

För att sköta sina uppgifter upprätthåller staten och kommunerna personuppgiftsregister, som är både unika och användbara. Inom en del sektorer är reglerna för nyttjandet av informationen generösa, men bruket av personuppgifter begränsas av de användningsändamål som definieras i lag och av komplicerade tillståndsförfaranden.

MyData-principerna utgår från att en människa har bestämmanderätt över information som handlar om henne själv. MyData inkluderar också en modell som ger människor redskap att kontrollera informationen och förvalta användningstillstånd för olika tredje parter. Modellen kunde minska kostnaderna för utvecklingen av nya tjänster och stöda testning i tidigt skede både inom den offentliga och den privata sektorn. Samtidigt skulle den stärka de offentliga tjänsternas begriplighet och individens rätt till sin egen information.

MyData-principerna kan förverkligas på flera olika sätt. I rapporten struktureras alternativen med hjälp av en klassificering, som innehåller fem olika nivåer och konkreta lösningar för att uppnå dessa. I utredningen föreslår vi att plattformar och verktyg produceras, så att registeransvariga kan välja för situationen lämpliga metoder att nyttja personuppgifter. Dessutom behövs gemensamma verksamhetsmodeller, såsom gemensamma mallar för användarvillkor och beskrivningar av användningsändamål, som underlättar kommunikationen mellan registeransvariga, individer och serviceproducenter.

I rapporten beskriver vi också med hjälp av tre fallstudier hur implementering av MyData-principerna påverkar myndigheters och andra intressenters arbete. En bättre tillgång på information möjliggör anpassade tjänster och nya typer av applikationer, bland annat inom förebyggande hälsovård.

Bilaga 1 Registrerade personers rättigheter med olika typer av information

Den här publikation är en del i genomförandet av statsrådets utrednings- och forskningsplan för 2017 (tietokaytoon.fi/sv).

De som producerar informationen ansvarar för innehållet i publikationen. Textinnehållet återspeglar inte nödvändigtvis statsrådets ståndpunkt

DESCRIPTION

Publisher and release date	Prime Minister's Office, 1.9.2017		
Authors	Aleksi Knuutila, Vesa Kokkonen, Heikki Sundquist, Ossi Kuittinen, Salla Thure		
Title of publication	Models and impact of MyData: Human-centric management and processing of public sector personal data		
Name of series and number of publication	Publications of the Government's analysis, assessment and research activities 61/2017		
Keywords	personal data, digitalisation, privacy, impact assessment		
Parts of the publication	PDF (license CC-BY-SA 4.0)		
Release date	September, 2017	Pages 83	Language Finnish

Abstract

The state and municipal governments collect data to fulfil their various duties. Often the information they store is not available elsewhere and could be used for many purposes. For some branches of government, rules about data sharing are permissive, but the use of personal data is limited by the purposes defined in law and the complicated practices related to granting permissions to access.

The MyData principles start from the idea that people should have control over the data about them. They suggest that people should have practical tools for authorising the reuse of their data in other services and to make choices about its processing. Making data available through such means could reduce the costs of developing new services and experimentation in both the private and public sector. At the same time, they would make public services more understandable and strengthen the rights individuals to their own data.

The MyData principles can be implemented in several different ways. In the report, the options are described through a classification of five different levels and the practical solutions that could be used to achieve them. The report proposes new platforms and tools to be developed in a way that allows data controllers to choose means for data sharing according to their individual situation. Public bodies should also work to create trust and facilitate interaction between individuals and service providers, for instance through templates for terms and conditions and shared, understandable ways of expressing how data is used.

The report describes the impact of implementing MyData on the government and other stakeholders through three case studies. The improved availability of information enables personalised services and new kinds of applications, for instance moving health care to a preventative direction.

Appendix 1 The rights of data subjects with respect to different types of information

This publication is part of the implementation of the Government Plan for Analysis, Assessment and Research for 2017 (tietokayttoon.fi/en).

The content is the responsibility of the producers of the information and does not necessarily represent the view of the Government.



SISÄLLYS

Kiitokset	1
Esipuhe	2
1 Johdanto	3
1.1 MyData käsitteenä	4
MyData henkilötiedon osajoukkona	4
MyData-periaatteet	5
MyDataan liittyvää sanastoa	6
1.2 Julkishallinto ja henkilötieto	7
1.3 Kansallisia kehityskulkuja	9
MyData osana digitalisaatiota	9
Kansallinen arkkitehtuuri ja toimialakohtaiset alustat	9
1.4 Euroopan unionin tietosuojaa-asetus	10
2 Julkiset tietovarannot ja niiden hyödyntäminen	12
2.1 Julkiset henkilötiedon varannot	12
Katsaus julkishallinnon tietovarantoihin	13
Yksilön oikeudet eri tietovarannoissa oleviin henkilötietoihin	14
Tiedon uudelleenkäytön muodot ja määrät	15
Tiedon tyyppejä: Perusrekisterit ja alustojen keräämä tieto	16
2.2 Korkean lisäarvon MyData-tietovarannot	17
Perusrekisterien tiedot	17
Asiakkuuteen liittyvät tiedot	18
Nykyiset näkymät omiin tietoihin	19
3 MyDatan toteuttaminen: Tärkeimmät palvelut ja infrastruktuuri	20
3.1 MyDatan eri tasot	20
3.2 Tiedon jakaminen	21
3.3 Käytännön keinot tiedon hallintaan	22
3.4 Hyödyllisessä muodossa oleva tieto	23
3.5 Palveluiden toteuttamisen arkkitehtuurit	25

4 MyDatan taloudelliset vaikutukset ja rahoitusmallit	28
4.1 Palveluntarjoajien arvoketjut ja henkilötiedon kysyntä.....	28
4.2 Investointikustannukset ja juoksevat kustannukset	30
Operatiiviset kustannukset	31
Investointikustannukset	32
4.3 Tietosuoritteiden hinnoittelun lailliset perusteet	34
4.4 Virastojen tietosuoritteista saamat tulot	35
4.5 Vaihtoehtoja luvitetun henkilötiedon hinnoittelulle	38
5 MyDatan vaikuttavuus: kolme tapaustutkimusta	44
5.1 Väestötietojärjestelmä esimerkkinä.....	45
VTJ:n käyttötarkoitukset MyDatan näkökulmasta	45
VTJ:n kehityskohteet MyDatan näkökulmasta.....	46
MyDatan vaikutukset VTJ:n toimintaan, tietojärjestelmiin ja talouteen.....	47
5.2 MyData terveydenhuollossa ja hyvinvointipalveluissa.....	50
Terveysdatan tyypit.....	51
Datan toissijainen käyttö	54
Terveydenhuollon digitalisaatio	55
Terveydenhuollon ekosysteemit ja tulevaisuus	56
MyDataa rajoittavat tekijät terveydenhuollossa	58
5.3 MyData liikenteessä	59
5.4 Virkamiesten roolin muutos tapaustutkimusten pohjalta	61
6 Vastuullisen tiedon hyödyntämisen tukeminen	64
6.1 Sovellusten rekisteröityminen ja tarkistaminen	64
6.2 Selkeiden tiedon jakamiseen liittyvien valintojen tukeminen.....	65
6.3 Avoimet standardit ja yhteentoimivuus	69
7. Ehdotuksia toimenpiteiksi.....	71
7.1. Henkilötiedon hallinnan peruspalveluiden synnyttäminen.....	71
7.2. Yksittäisten julkishallinnon organisaatioiden toimenpiteet.....	72
7.3. Sääntely ja kansallinen ohjaus	72
7.4. Hallinnon sisäiset käytännöt tukemaan MyDataa	73
7.5. Suomen kansainvälisen edelläkävijyyden mahdollistaminen.....	74
Liite 1. Rekisteröityjen oikeudet eri tietotyyppien kohdalla.....	76



Lähteet.....81

KIITOKSET

Tässä selvityksessä esitetyt ajatukset ja ehdotukset nojaavat suurelta osin haastatteluaineistoon. Tutkimusta varten tehtiin teemahaastatteluita sekä henkilörekistereitä hallinnoivien julkisten rekisterinpitäjien että sitä käyttävien yritysten ja julkisten palvelujen parissa. Haastatteluja tehtiin helmikuusta toukokuuhun 2017. Samaan aikaan valmisteltiin myös vuonna 2014 Liikenne- ja viestintäministeriön julkaiseman *My Data - johdatus ihmiskeskeiseen henkilötiedon hyödyntämiseen* -raportin päivitystä, jota osa haastatteluista myös tuki. Järjestimme myös kaksi sidosryhmätyöpajaa, joihin osallistui yhteensä 18 henkilöä. Työpajojen tuloksia käytettiin etenkin tulevaisuuden julkishallinnon toimintamallien hahmottamiseen luvussa 6.

Kiitämme suuresti haastateltavia heidän ajastaan ja monipuolisesta osaamisestaan. Kiitos: Ulla Ahlbald-Bordi (STM), Antti Ailio (Sitra), Sirpa Arvonen (HUS), Annika Collin (VRK), Sampsa Hautaniemi (HY), Jani Heikkinen (VM), Mauri Honkanen (Nokia), Harri Honko (DHR), Konstantin Hyppönen (Kela), Hannu Hämäläinen (STM), Anneli Jätteenmäki (EU-parlamentti), Jari Kallela (VM), Joonas Kankaanpää (VRK), Anne Kauhanen-Simanainen (VM), Antti Kettunen (Tieto), Miikka Kiiski (IBM), Taito von Konow (Verohallinto), Jaakko Korhonen (Espoo), Juha-Pekka Koskinen (Nets Oy), Marko Latvanen (VRK), Torbjörn Lundqvist (Kela), Jukka Lähesmaa (STM), Kari-Pekka Mäki-Lohiluoma (Kela), Joonas Mäkinen (Opetushallitus), Hannu Parkkisenniemi (Trafi), Pauli Pekkanen (VRK), Mikko Pitkänen (PRH), Olli Pitkänen (HIIT), Kirsi Pulkamo (Trafi), Päivi Pösö (VRK), Pekka Rehn (VRK), Jani Ruuskanen (VRK), Antti Saarikoski (MML), Timo Salovaara (VRK), Matti Sillanmäki (OKM), Mia Sundblad (Kela), Teija Tarvainen (MML), Paula Vesterinen (Kela), Teemupekka Virtanen (STM), Janne Viskari (VRK) ja Tomi Voutilainen (VM).

Kiitos myös kaikille, joilta olemme saaneet asiantuntevaa apua raportin kirjoittamisessa. Näihin kuuluvat ohjausryhmämme jäsenet Riitta Autere (YM), Antti Eskola (TEM), Jani Heikkinen (VM), Jari Kallela (VM), Tomi Kytölä (OKM), Kimmo Mäkinen (VM), Markus Rahkola (VM), Taru Rastas (LVM) ja Suvi Remes (VM). Lisäksi haluaisimme kiittää seuraavia kirjoittamisessa auttaneita: Joonas Antikainen, Mika Honkanen, Eija Kalliala, Jaana Nevalainen, Juho-Matti Paavola, Jessica Parland-von Essen, Antti Poikola, Teemu Ropponen, Samuel Rinnetmäki ja Arttu Vainio. Kiitämme myös Kirmo Kivelää, joka on tehnyt aiempien MyData selvitysten grafiikat, ja näin vahvasti edesauttanut tämänkin selvityksen kuvituksen suunnittelussa, sekä Digital Health Revolution hankkeen tekijöitä vahvasta halusta kehittää MyDatan visuaalista maailmaa.

ESIPUHE

*MyData*n vaikutukset julkisen hallinnon organisaatioiden toimintaan on Valtioneuvoston rahoittama selvitys- ja tutkimushanke, jonka tarkoituksena on tukea ja edesauttaa hallitusohjelman *Digitalisoidaan julkiset palvelut* -kärkihankekokonaisuuden toteutusta. Yhtenä kärkihankekokonaisuuden toimenpiteenä valtiovarainministeriö on käynnistänyt *Yhteinen tiedon hallinta* -hankkeen (2016–2018), jossa tehtävän työn tavoitteena on vahvistaa periaatetta, että tietoa kysytään vain kerran ja hyödynnetään monipuolisesti. Taustalla on ollut myös tavoite vahvistaa kansalaisen oikeutta julkisen hallinnon hallussa oleviin tietoihin ja edesauttaa ihmiskeskeisen tiedonhallinnan, ns. *MyData*-lähestymistavan, kehittymistä Suomessa. *MyData*-lähestymistavassa kansalainen voi esimerkiksi jakaa tai myydä omia henkilötietojaan valtuuttamansa palvelun hyödynnettäväksi.

MyData-lähestymistavan soveltamisella julkisen sektorin hallussa oleviin henkilötietoihin ja muihin tietovarantoihin on julkisessa keskustelussa nähty olevan mahdollisuuksia edistää uusien liiketoimintamahdollisuuksien syntymistä myös yksityiselle sektorille, palveluille, joissa hyödynnetään tai jatkjalostetaan henkilöön liittyviä tietoja. Ajatusmalli, jossa kansalainen itse on aktiivinen toimija tiedonhallinnan keskiössä suhteessa yksityisen ja julkisen sektorin toimijoihin, nostaa väistämättä esille kysymyksiä viranomaisen vastuista ja roolista uudessa muuttuvassa toimintaympäristössä.

Tutkimushankkeen toimeksiannossa nostettiin esille erityisesti seuraavat tehtävät:

- Arvioida *MyData*-tyyppisen henkilötietojen hallinnan ja jakamisen vaikutuksia viranomaisten toimintaan julkisessa hallinnossa.
- Arvioida omien tietojen jakamisen mahdollistamisen vaikutuksia suhteessa viranomaisen nykyisiin resursseihin ja tietojärjestelmiin.
- Arvioida erilaisten skenaarioiden kautta julkiselle sektorille vaihtoehtoisia rahoitusmalleja *MyData*-tyyppisen henkilötiedon hallinnan järjestämiseksi.

Tutkimuksen aikana tutkimushankkeen ohjausryhmä päätti sisällyttää tutkimukseen kolme tapaustutkimusta, joissa pureuduttiin käytännön esimerkkien kautta nykyisiin ja mahdollisiin tuleviin ratkaisumalleihin. Osana tutkimusta järjestettiin myös kaksi avointa työpajaa sidosryhmille.

Tutkimushankkeen johtopäätöksiä ja suosituksia hyödynnetään julkisen hallinnon *MyData*-linjausten ja toimenpiteiden valmistelussa, ja ne voivat toimia myös oppaana julkishallinnon organisaatioille niiden kehittäessä henkilötiedon käsittelyään *MyData*-periaatteiden mukaiseksi. Tutkimustuloksia hyödynnetään soveltuvin osin myös mainitun *Yhteinen tiedon hallinta* -kärkihankkeen työpaketin *OmaData julkisessa hallinnossa* toteuttamisessa.

30.6.2017

Kimmo Mäkinen

Helsinki

Kehittämispäällikkö

Valtiovarainministeriö

1 JOHDANTO

Aleksi Knuutila – Salla Thure

Julkishallinto ylläpitää tehtäviään varten monia rekistereitä, joita vastaavaa tietosisältöä ei ole olemassa muualla. Sen tallettamalla tiedolla on usein myös erityisasema, sillä julkista tietoa pidetään luotettavana. Suomessa henkilötietorekisterit ovat kansainvälisesti katsoen kattavat ja niiden hyödyntäminen onkin verrattain kehittyneellä tasolla. Tiedon tallentamisen ja käsittelyn menetelmien kehittyessä jaetun tiedon määrä on kasvussa, ja niin yksityiset kuin julkisetkin palvelut yhdistelevät tietoja yhä suuremmasta määrästä rekistereitä.

Henkilötiedon hyödyntämistä koskevat lainsäädäntö ja käytännöt ovat usein sallivia, joskin hallinnonalojen välillä on merkittäviäkin eroja. Henkilötietoa julkisista rekistereistä saa käyttää laissa määriteltyihin käyttötarkoituksiin, niin viranomaisten tehtäviä suorittaessa kuin myös liiketoiminnassa ja rekistereihin pohjautuvassa tutkimuksessa. Pääsääntöisesti rekistereitä hallinnoivat viranomaiset ovat myös vastuussa sen valvomisesta, että tietoa käytetään asetettujen käyttötarkoitusten rajoissa ja turvallisesti.

Silti nykyisissä keinoissa hyödyntää henkilötietoa on omat rajoituksensa. Hyväksytyt käyttötarkoitukset tiedolle on määritelty ennakkoon laissa ja asetuksissa, jotka muuttuvat hitaasti eivätkä aina pysy mukana tarpeiden ja sovellusten kehityksessä. Pääsy viranomaisten hallinnoimaan tietoon edellyttää mahdollisesti monimutkaisia tietolupajärjestelyjä ja tarkistuksia, jotka pitää hoitaa jokaisen viranomaisen kanssa erikseen. Tiedon hinnoittelussa on lisäksi kirjavia käytäntöjä. Tulot kyllä tukevat viranomaisten toimintaa, mutta jo niiden periminen voi olla hallinnollisesti hankalaa ja kasvattaa tiedon käytön kynnystä.

MyData-periaatteet kuvaavat mallia henkilötiedon hyödyntämiseen, jossa ihmisillä on oikeus ja pääsy heitä koskevaan tietoon ja käytännön keinot hallita sen käyttöä. Periaatteet on kehitetty aikaisemmassa selvitys- ja tutkimustyössä, joka toimii tämän raportin lähtökohtana.¹ MyData-lähestymistapaan kuuluu ajatus tulevaisuudesta, jossa ihmiset voisivat antamiensa käyttöluvien perusteella monipuolisesti yhdistellä itseään koskevaa tietoa eri lähteistä, mikä auttaisi heitä hyötymään heistä kerätystä tiedosta parempien palvelujen muodossa. Henkilötietoa, joka on ihmisten itsensä saatavilla ja siirrettävissä tällä tavoin, kutsutaan MyDataksi.

MyData-periaatteet voivat toteutua useammalla erilaisella tavalla. Tässä raportissa hahmotellaan yksilöiden hallitseman henkilötiedon jakamiseen järjestelmä, joka tukisi sujuvaa palvelujen kehitystä ja nopeampia kokeiluja. MyDataa hyödyntävien palveluntarjoajien vastuulla olisi kuvailla toimintaansa niin, että heidän palveluitaan käyttävät henkilöt voivat tehdä perusteltuja ja merkityksellisiä valintoja omien tietojensa käytöstä. Tällaisten sovellusten tietoturva ja tietosuojat tulisi voida tarkistaa viranomaisten toimesta, mutta kevyellä tavalla, joka pitää kehittämisen kustannukset matalana. Ihmisten antama suostumus toimii (tietyissä rajoissa) yleisenä perusteena tiedon käsittelylle. Siksi hahmoteltu järjestelmä voisi tarjota yhtenäiset käytännöt tiedon hyödyntämiseen useasta julkisesta henkilörekisteristä. Tällainen tapa hyödyntää henkilötietoja pienentäisi laajasti palvelujen kehittämisen kustannuksia ja tukisi nopeiden, yllättävienkin kokeilujen tekemistä.

¹ Etenkin Poikola and Honko 2015

MyData-periaatteet tukisivat yksilöiden oikeuksia omaan tietoonsa ja käytännön mahdollisuuksia sen käytön hallintaan. Periaatteiden pohjalta henkilötieto nähdään uudelleen käytettävänä resurssina, jota ihmiset voivat hyödyntää haluamallaan tavalla. MyData-periaatteisiin kuuluu lisäksi ajatus avoimista ja yhteisistä toimintatavoista, jotka helpottaisivat käytännön työkalujen kehittämistä tiedon jakamisen hallitsemiseen.

Myös julkisten palvelujen suunta on kohti monipuolisempaa datan hyödyntämistä mm. tietojen yhdistelyssä, analytiikan käytössä ja automatisoidussa päätöksenteossa. Yksityisten ja yksilöiden itse hallinnoimien tietolähteiden hyödyntäminen julkisissa palveluissa voi lisääntyä. Tähän kehityskulkuun liittyy myös kasvavia ihmisten yksityisyyteen liittyviä riskejä, joihin MyData-periaatteet voivat osin vastata. Luottamus julkishallinnon toimintaan voi siis tulevaisuudessa vaatia nykyistä suurempaa tiedon käsittelyn läpinäkyvyyttä ja ihmisten osallistamista henkilötiedon käsittelyyn liittyviin valintoihin.

1.1 MyData käsitteenä

Tarkastellessamme henkilötiedon käsittelyä julkishallinnossa MyData-periaatteiden soveltamisen näkökulmasta lähtökohtanamme ovat toimineet aikaisemmat kirjoitukset aiheesta, etenkin liikenne- ja viestintäministeriön julkaisemat raportit.² Keskustelussa aiheesta sanaa MyData on käytetty viittaamaan useampaan asiaan: tiettyihin datan joukkoihin, henkilötiedon hallinnan järjestelmiin sekä näitä määrittäviin periaatteisiin.³ Tässä tekstissä pyrimme selkeyttämään sanastoa niin, että puhuessamme MyDatasta tarkoitamme nimenomaan sitä henkilötietoa, joka on ihmisten käytettävissä (katso tarkemmin alla). MyData-periaatteet puolestaan määrittelevät aikaisemmassa keskustelussa kiteytyneet tavoitteet ja ohjenuorat henkilötiedon hyödyntämiselle. Periaatteet voivat toteutua monissa tietojärjestelmissä, joihin viittaamme tapauskohtaisesti erilaisella sanastolla, kuten esimerkiksi kuvaillessamme jäljempänä MyData-tiliä.

MyData henkilötiedon osajoukkona

MyDatassa on kyse nimenomaan henkilötiedon hyödyntämisestä uudella tavalla. Tässä tekstissä tarkoitamme henkilötiedolla laajasti kaikkea henkilöön liittyvää tai henkilön toiminnan seurauksena syntyvää tietoa. Tapamme käyttää termiä on siis lähellä henkilötietolaissa annettua määritelmää.⁴

MyData on henkilötiedon osajoukko. Kaikki MyData on siis henkilötietoa, mutta kaikki henkilötieto ei ole MyDataa. Vähimmäiskriteeri sille, että jokin henkilötieto lasketaan MyDataksi, on se, että yksilöllä on laillinen oikeus ja tekninen mahdollisuus uudelleenkäyttää ja jakaa tietoaan haluamallaan tavalla.

² Poikola, Kuikkaniemi, & Kuitinen, 2014 ja Poikola & Honko, 2015

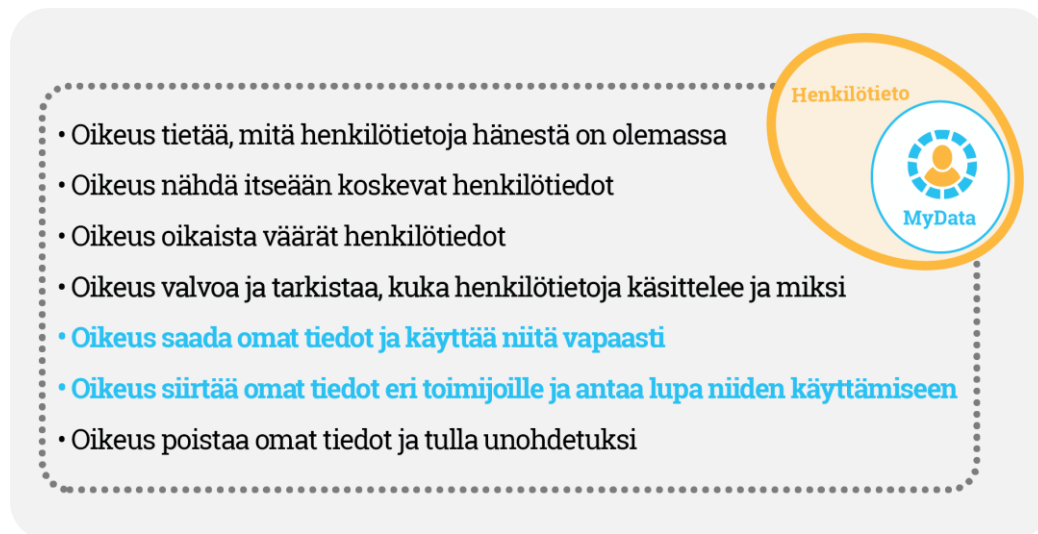
³ taan liittyvää sanastoa on pohdittu esimerkiksi Sanastokeskuksen toimesta, ks. <http://www.terminfo.fi/sisalto/my-data-mydata-oma-data-omadata-oma-tieto-omatieto-omat-tiedot--uudella-kasitteella-on-monta-nimea-368.html>

⁴ Tässä yksinkertaisuuden vuoksi MyData ymmärretään luonnollisen henkilön ja henkilötiedon välisen suhteen pohjalta. Tällaiseen määritelmään on ainakin kaksi huomautusta: Ensinnäkin joissain MyDataan liittyvissä ehdotuksissa samojen MyData-palveluiden kautta pystyisi hallitsemaan myös yrityksiin tai muihin organisaatioihin liittyvää dataa (Poikola ym., 2014). Esimerkiksi oikeilla valtuutuksilla yrityksen nimissä toimiva henkilö voisi päästä käsiksi yritykseen liittyviin lupa- tai tietoihin. Tämän järjestämiseksi voisi mahdollisesti käyttää samoja palveluita, rajapintoja ja käyttöliittymiä kuin luonnollisia henkilöitä koskevan MyDaten kohdalla. Jos henkilötiedon hyödyntämisen järjestelmä kehittyisi tähän suuntaan, sen voisi ymmärtää keinona, jolla mitä tahansa toimijaa koskeva data tulisi hallittavaksi ja luettavaksi. Toimijat voisivat olla myös yrityksiä tai yhdistyksiä, joiden nimissä yksittäiset käyttäjät voivat toimia, kun heidät on valtuutettu oikein.

Toinen huomautus koskee MyDataa henkilötiedon tyyppinä. MyData määritellään usein siksi henkilötiedon osajoukoksi, jonka kohdalla aktiiviset oikeudet toteutuvat (Poikola ym., 2014). Henkilötieto määritellään lain mukaan tiedon referentin pohjalta. Jos sen yksilön, jota tieto koskee, pystyy tunnistamaan, on kyse nimenomaan henkilötiedosta, joka koskee kyseistä henkilöä. Samalla on kuitenkin myös muita lainkohtia ja perusteita tiedonsaannille ja muille oikeuksille tiedon suhteen. Esimerkiksi Anna-Riitta Wallin esittelee luokittelun tiedonsaantioikeuksia koskevista määrittelyistä, jossa erotellaan "tiedon referenssin tarkastusoikeus" "rooliperusteisesta tiedonsaantioikeudesta" (Korhonen, 2003). Jälkimmäiseen luokkaan kuuluu tieto, johon henkilöillä voi olla oikeus esimerkiksi sillä perusteella, että he ovat perheenjäseniä tai työntekijöitä. Perheenjäsenet voivat joissakin tapauksissa asianomaisina selata esimerkiksi lastensa terveystietoja. Tämä on toinen mahdollinen poikkeus sille, että MyData tarkkaan ottaen rajoittuisi yksilön omaan henkilötietoon.

Henkilötiedon hallinnan toteutuminen MyData-periaatteiden mukaisesti ei ole pelkästään joko-tai-kysymys, vaan monimutkainen kokonaisuus. Kuva 1.1 listaa joukon mahdollisia oikeuksia, jotka voivat toteutua eri tilanteissa. Perustelluista syistä tiettyjen henkilötiedon lähteiden kohdalla joitakin näistä oikeuksista ei voi soveltaa. Erityisesti julkishallinto kerää tietoja kansalaisista lain määrittämässä rajoissa viranomaistehtäviä suorittaessaan. Kansalaisilla on usein mahdollisuus saada kopio tiedostaan, mutta harvoin oikeutta esimerkiksi kieltää viranomaisia käsittelemästä heidän henkilötietoiaan tai vaatia, että se poistettaisiin.

Kuva 1.1. Yksilön oikeudet itseään koskevaan tietoon



Erilaiset oikeudet, joita yksilöllä voi olla itseään koskevan tiedon hallinnassa. MyDatan minimiehto on pääsy tietoon ja mahdollisuus käyttää sitä uudelleen.

Henkilötiedon hallinnan tukemiseksi on MyDatan minimiehtoja laajemmin perusteltua pyrkiä siihen, että mahdollisimman moni yllä mainituista oikeuksista toteutuu. MyData-periaatteet toteutuisivat vahvasti julkishallinnossa esimerkiksi silloin, kun yksilöt voivat käyttää heistä kerättyä tietoa omiin tarkoituksiinsa ja seurata tiedon käyttöä julkishallinnon sisällä. Näin tulkittuna MyData-periaatteet eivät siis välttämättä edellytä, että tietoa käsitellään ihmisten suostumusten perusteella tai että he voisivat estää heitä koskevan tiedon käytön.

MyData-periaatteet

Edellä mainituissa lähteissä MyDatan perusteet on kiteytetty kolmeen periaatteeseen. Nämä ovat ihmiskeskeinen tiedonhallinta, käytettävä data sekä liiketoimintaympäristön avoimuus. Periaatteet täydentävät MyDatan määritelmää (henkilötiedon osajoukkona) kuvaamalla laajasti henkilötiedon hyödyntämisen edellytyksiä.

1. periaate: Ihmiskeskeinen tiedonhallinta ja yksityisyys

Ensimmäisen MyData-periaatteen mukaan yksilöt tulisi nähdä aktiivisina toimijoina, ei passiivisina toiminnan kohteina. Ihmisillä tulee olla paitsi oikeuksia suhteessa omaan tietoonsa, myös käytännön työkaluja tiedon hyödyntämisen ja yksityisyytensä hallintaan. Henkilön tulisi voida siirtää tai kopioida itseään koskeva tieto muiden käyttöön, ja myös lopettaa tällainen jakaminen helposti. Parhaimmillaan oman tiedon käyttöä voisi valvoa ja tehdä sitä koskevia yksityiskohtaisia valintoja.

Ihmiskeskeinen tiedonhallinta merkitsee myös, että yksilöä koskeva tieto eri lähteistä olisi kokonaisvaltaisesti ja helposti saatavissa. Yksilö on tietoja yhdistävä tekijä riippumatta siitä, mihin tiedot on talletettu. Ottaessaan käyttöön uusia palveluita käyttäjä voi valita, mitä itseään koskevia tiedon lähteitä hyödyntäisi.

2. periaate: Käytettävä data

Toinen periaate ehdottaa, että rekistereihin ihmisistä kerätty tieto tulisi nähdä uudelleen käytettävänä resurssina. Tällä hetkellä palvelut, jotka keräävät tai käsittelevät tietoa ihmisistä, tallentavat ja soveltavat tietoja ennalta määriteltyihin käyttötarkoituksiin. Yksilöt sen sijaan voivat hyötyä näistä tietovarannoista käyttämällä niitä joustavasti itse määrittelemiinsä tarkoituksiin. Datan käytettävyys edellyttää, että tietoon on helppo päästä käsiksi ja että se on esitetty muodossa, joka mahdollistaa uudelleenkäytön.

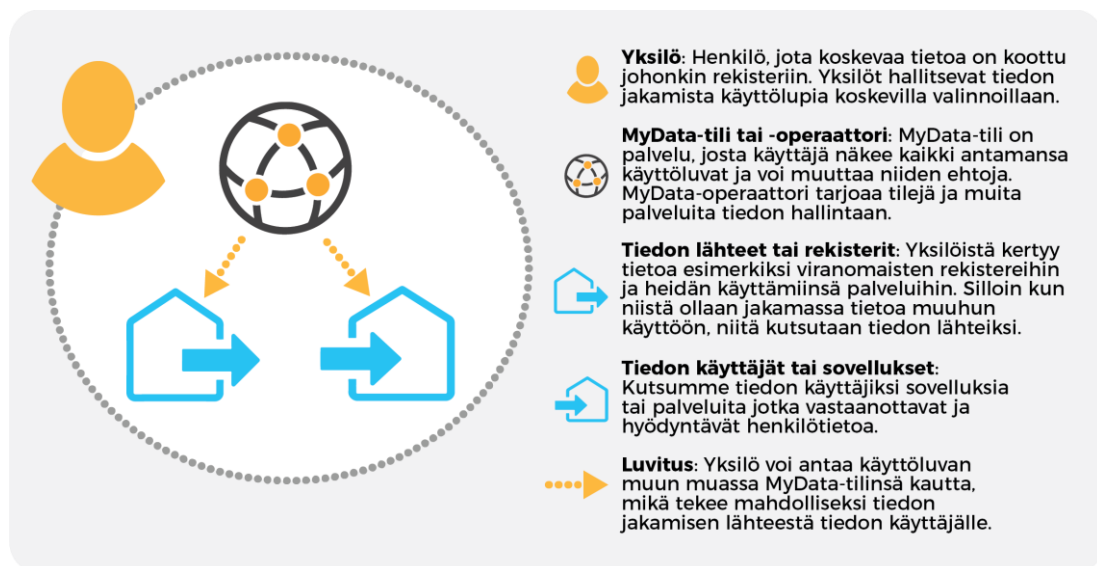
3. periaate: Liiketoimintaympäristön avoimuus

Ihmiskeskeinen tiedon hyödyntäminen voi tukea avoimia liiketoimintamalleja, joissa asiakkuuden vaihto on helppoa ja tieto seuraa ihmisten mukana. Julkishallinto ei luonnollisesti toimi liiketoiminnallisin perustein. Sen piirissä yksittäiset toimijat ovat usein ainoita, jotka keräävät jotakin tietoa viranomaistehtäviään varten. Tästä huolimatta julkinen sektori voi pyrkiä luomaan avoimen toimintaympäristön MyDataan pohjaavien palveluiden kehittäjille, mikä voi tukea niin yksityisen sektorin sovelluksia kuin julkishallinnon sisäistäkin toimintaa. Yhteisiä ja avoimia standardeja käytettäessä toimijoita ei suljeta pois käyttämästä tietoresursseja. Silloin kun palveluita syntyy monipuolisesti, eivät käyttäjätkään ole riippuvaisia yksittäisistä palveluntarjoajista tai käyttöliittymistä.

MyDataan liittyvää sanastoa

Henkilötiedon uudelleenkäytön mahdollistavassa järjestelmässä on useita toimijoita ja niiden välisiä suhteita. Joiltain osin näitä kuvaava sanasto on edelleen kehittymässä ja vakiintumassa. Käymme siksi selkeyden vuoksi lyhyesti läpi keskeisimmät tekstissä käyttämämme termit.

Kuva 1.2. Roolit henkilötietoja siirrettäessä



Neljä eri roolia silloin kun henkilötietoa siirretään yksilöiden antamien lupien perusteella

MyDataa koskevissa kirjoituksissa on eroteltu neljä erilaista roolia. Ne ovat seuraavat:

- **Yksilö:** Henkilö, jota koskevaa tietoa on koottu johonkin rekisteriin. Yksilöt hallitsevat tiedon jakamista käyttö lupia koskevilla valinnoillaan.
- **Tiedon lähteet tai rekisterit:** Yksilöistä kertyy tietoa esimerkiksi viranomaisten rekistereihin ja heidän käyttämiinsä palveluihin. Silloin kun niistä ollaan jakamassa tietoa muuhun käyttöön, niitä kutsutaan tiedon lähteiksi.
- **Tiedon käyttäjät tai sovellukset:** Kutsumme tiedon käyttäjiksi sovelluksia tai palveluita jotka vastaanottavat ja hyödyntävät henkilötietoa.
- **MyData-tili tai -operaattori:** MyData-tili on palvelu, josta käyttäjä näkee kaikki antamansa käyttöluvut ja voi muuttaa niiden ehtoja. MyData-operaattori tarjoaa tilejä ja muita palveluita tiedon hallintaan.

MyData on henkilötietoa, joka on jaettavissa ihmisten valintojen perusteella. Tiedon jakaminen tarkoittaa yksinkertaisesti siirtämistä tai kopioimista kahden toimijan välillä. Jakaminen ei siis edellytä, että tiedon lähteen kopio datasta poistetaan. Vaikka usein tiedon vastaanottaja pitää kopion saamastaan tiedosta, on myös mahdollista, että tietoa säilytetään vain sen alkuperäisessä lähteessä. Jakamisen ohella ihmiset voivat myös uudelleenkäyttää tietojään, eli soveltaa niitä uuteen käyttötarkoitukseen.

Ihmisten tekemiä valintoja kutsumme tässä yhteydessä luvan antamiseksi tai luvittamiseksi. Luvan antamiseen sisältyvät suostumukset, jotka ovat tietoisia ja vapaaehtoisia omien tietojen käytön hyväksymisiä. Suostumuksen antajien tulee tietää, mitä tiedon sisältöjä käsitellään ja mihin tarkoitukseen. Tässä raportissa puhumme luvan antamisesta myös silloin, kun henkilö ottaa palvelun tai sovelluksen käyttöönsä tai muulla esimerkiksi käyttöehdoissa mainitulla toimella sopii tietojensa käytöstä. Tällöin palvelun käyttöehtojen olisi hyvä selkeästi kuvata, miten tietoja käsitellään.⁵

MyData-periaatteiden mukainen järjestelmä voi toteutua monella erilaisella tavalla. Kuvassa 1.2 on kuvattu järjestely, jossa tieto siirtyy suoraan tiedon lähteestä sen käyttäjälle yksilön antamien käyttö lupien perusteella. Erilaisia malleja periaatteiden toteuttamiselle on kuvattu luvussa 3.

1.2 Julkishallinto ja henkilötieto

MyData on käsitteenä määritelty yleisesti periaatteiden tasolla. Silti henkilötiedon hyödyntämisen mahdollistavat järjestelmät kuten myös tiedosta saatava hyöty saattavat olla erilaisia yhteiskunnan eri alueilla. Tässä yhteydessä käymme lyhyesti läpi tärkeimmät julkishallinnon erityispiirteet, jotka vaikuttavat MyData-periaatteiden toteutumiseen.

Julkishallinnon tietoa pidetään luotettavana

Tietoa voi olla saatavilla eri lähteistä, mutta usein julkisten rekistereiden sisältöä pidetään ratkaisevana ja luotettavana. Julkisen tiedon erityisasema perustuu mahdollisesti myös lakiin. Jotkin rekisterit, kuten lainhuutorekisteri, määritellään laissa *julkisesti luotettavaksi*, eli

⁵ Suostumus ja asiakassuhde antavat kummatkin laillisen perustan henkilötiedon käsittelylle, mutta ovat laillisilta vaatimuksiltaan erilaisia. Tämän raportin yhteydessä ei oteta vahvasti kantaa esimerkiksi siihen, minkälaisia palvelujen käyttöehtoja pidetään hyväksyttävänä MyData:n jakamisessa. Sen sijaan raportissa käydään läpi merkityksellisten valintojen tukemisen yleisiä periaatteita.

yleisesti hyväksytyksi tiedonlähteeksi. Tällaisten rekistereiden sisältämän tiedon hyväksymistä ei lasketa rikkeeksi edes silloin, kun sen sisältö olisi jälkikäteen osoitettu virheelliseksi.⁶

Julkishallinnon erityisasemasta johtuen julkishallinnon tarjoaman tiedon arvoa voi vahvistaa se, että tiedon voi todistaa olevan peräisin julkisista lähteistä. Lähteiden varmentamisen mahdollisuutta ei aina priorisoida yksityisissä tiedon jakamisen järjestelmissä, mutta julkisten rekistereiden kohdalla se on tärkeää. Aikaisemmin asiakirjojen laatija varmennettiin allekirjoitusten tai leimasinten avulla. Sähköisiin tiedonjakojärjestelmiin on nykyisin rakennettu sisälle keinoja tiedon sähköisiin allekirjoituksiin. Esimerkiksi sähköinen reseptijärjestelmä mahdollistaa apteekissa asiointin todentamalla, että reseptitiedot ovat jonkin hyväksytyin tahon (kuten lääkärin) laatimia. Samankaltainen sähköisten allekirjoitusten järjestelmä voisi olla myös tärkeä osa henkilötiedon uudelleenkäyttöä tukevaa järjestelmää.

Oikeudet vaikuttaa tiedon käsittelyyn julkishallinnossa

Julkisia rekistereitä hyödynnetään ennen kaikkea viranomaisten tehtävien suorittamiseen. Julkinen valta käyttää näitä tietoja myös sellaisissa tehtävissä, joita voidaan pitää yhteiskunnan toiminnan kannalta välttämättöminä tai perustavanlaatuisina, kuten esimerkiksi verotuksen toteuttamisessa, turvallisuuden ylläpitämisessä tai erilaisten julkisten palveluiden käyttöoikeuksien määrittämisessä. Viranomaisilla on myös vastuu rekistereiden kattavuuden ja luotettavuuden ylläpitämisestä, mikä voi samalla rajoittaa yksilöiden oikeuksia tietosisältöjen suhteen. Yksityisten palvelujen kohdalla ihmisillä on pääsääntöisesti mahdollisuus poistaa heitä koskevat merkinnät esimerkiksi sen jälkeen, kun he ovat päättäneet asiakassuhteensa palveluntarjoajan kanssa. Julkisten rekistereiden kohdalla yksilöiden mahdollisuudet kieltää heitä koskevan tiedon käsittelyä ovat sen sijaan rajatut. Näin ollen julkishallinnon MyDatan kohdalla on luontevaa puhua tiedon yhteiskäytön mahdollistamisesta, sen sijaan että tiedon katsottaisiin kuuluvan yksilölle vahvassa mielessä. Yksilön oikeudet käydään läpi tarkemmin liitteessä 1.

Lähtökohtana nykyiset tietojärjestelmät

Suomen julkishallinnon perusrekistereillä on jo pitkä historia tiedon hyödyntämisen ja hallinnan saralla. Monissa tapauksissa viranomaiset ovat kehittäneet hyvin suunniteltuja rajapintoja tiedon jakamiseen, ja niiden käyttö tiedon hakuun suorien kyselyiden muodossa onkin kasvussa. Myös esimerkiksi henkilötietojen korjaamiseen ja niiden käyttötapojen selvittämiseen on vakiintuneita käytäntöjä. MyData-periaatteiden toteuttaminen ei siis tapahdu tyhjältä pöydältä, vaan sen täytyy pohjata olemassa oleviin tietojärjestelmiin ja toimintatapoihin. Pyrkimykset ihmiskeskeiseen tiedonhallintaan olisi siis paras kytkeä osaksi näiden olemassa olevien järjestelmien uudistamista, joka on tapahtumassa joka tapauksessa. Joiltain osin jo olemassa olevat viranomaisiin kohdistuvat velvoitteet, kuten yhteisten sähköisten tukipalvelujen käyttö sekä yleinen tietosuoja-asetus, voivat myös tukea muutosta tähän suuntaan.

⁶ Esimerkiksi Laissa väestötietojärjestelmästä ja Väestörekisterikeskuksen varmennepalveluista 2009/661 todetaan, että "henkilötietoja pidetään julkisesti luotettavina tietoina, jollei osoiteta, että tieto on virheellinen tai puutteellinen".

1.3 Kansallisia kehityskulkuja

MyData osana digitalisaatiota

Sipilän hallituksen hallitusohjelman, *Ratkaisujen Suomi*, digitalisaatio-osuuden kaksi osiota, *digitalisoidaan julkiset palvelut ja rakennetaan digitaalisen liiketoiminnan kasvuympäristö*⁷ ovat MyDatan kannalta tärkeimmät. Hallitusohjelma viittaa MyDatan kaltaisiin periaatteisiin:

“Vahvistetaan kansalaisten oikeutta valvoa ja päättää itseään koskevien tietojen käytöstä, samalla varmistaen tietojen sujuva siirtyminen viranomaisten välillä.”

MyData on vahvasti esillä myös hallituksen kärkihankkeissa. Yksi merkittävä esimerkki on Yhteisen tiedon hallinnan kärkihanke (YTI-hanke), jonka tavoitteena on edesauttaa julkishallintoa esimerkillisenä MyData-toimijana sekä järjestelmän mahdollistajana. Hankkeen piirissä MyData ymmärretään osana julkishallinnon toimintatapojen uudistusta, uudenlaisten palvelukokemusten synnyttämistä sekä keinona tehostaa tietojen monipuolista käyttöä. YTI-hanke pyrkii luomaan yhteisiä periaatteita MyData-toiminnalle sekä edesauttamaan kansallisessa palveluarkkitehtuuriohjelmassa toteutettujen sähköisen asioinnin tukipalveluiden jatkokehittämistä MyData-periaatteiden toteuttamiseksi.⁸ Digitaalisen liiketoiminnan kasvuympäristö -kärkihanke (Digi2) sen sijaan tukee digitaalisen liiketoiminnan toimintaympäristön rakentumista. Se pyrkii lisäämään datan käyttöä digitaalisessa palvelutuotannossa, johon liittyy myös henkilötiedon hyödyntäminen kansalaisen ehdoilla. Digi2:n osana ollaan toteuttamassa kokeiluja, joissa myös julkishallinnon tietoja tehdään saataviksi MyData-periaatteita seuraten.⁹

Kansallinen arkkitehtuuri ja toimialakohtaiset alustat

Kansallisessa palveluarkkitehtuuriohjelmassa (KaPa) toteutetut sähköisen asioinnin tukipalvelut tukevat ja edesauttavat julkishallinnon organisaatioita sähköistämään palvelujansa. KaPassa toteutetut palvelut ovat Suomi.fi-perhettä. KaPA-laissa on veloitettu julkishallinnon organisaatioita käyttämään ensisijaisesti Suomi.fi-tukipalveluja. Suomi.fi-tukipalvelut tarjoavat yhteisiä, vakioituja tapoja järjestää näitä toimintoja hallinnon eri aloilla. Suomi.fi-tukipalveluiden avulla helpottuu mm. henkilön vahva tunnistus, tiedonsiirto sekä sähköisten papereiden lähetys ja toisen puolesta asiointi. Suomi.fi-palvelutietovarantoon on veloitettu julkishallintoa kuvaamaan palvelunsa, jotka näkyvät kansalaisille yhdestä paikasta, Suomi.fi-verkkopalvelusta. Suomi.fi:ssä kansalainen voi nähdä, mitä tietoja hänestä on julkishallinnon eri rekistereissä. Raportin kirjoittamisen aikana tiedon tyyppiä näkyy palvelussa vain muutama, mutta KaPA-laki velvoittaa tuomaan henkilötunnuksella ja Y-tunnuksella haettavat rekisteritiedot Suomi.fi-palveluun. Tämä vaatii organisaatiolta liittymistä Suomi.fi-palveluväylään. Toki organisaatio päättää, mitkä tiedot rekistereistä näkyvät sivustolla. Esimerkiksi VRK:n henkilörekisteriä ei täysin näytetä Suomi.fi:ssä. Lisäksi viranomaisilla on velvollisuus kytkeä merkittävimmät tietovarantonsa sekä sähköisissä palveluissaan tarjoamansa tiedot Suomi.fi-palveluväylään. Suomi.fi-sivuston kautta ihmisillä on jo valmiiksi melko kokonaisvaltainen näkymä omiin tietoihinsa, ja on mahdollista, että näihin sähköisiin tukipalveluihin kehitetään tulevaisuudessa MyData-operaattorin tapaista toiminnallisuutta, jonka kautta käyttäjät voisivat jakaa omaa tietoaan muihin palveluihin.

⁷ (Hallituksen julkaisusarja, 2015)

⁸ <http://vm.fi/yhteinen-tiedon-hallinta>; Hankesuunnitelma <https://wiki.julkict.fi/julkict/yti/vit-projektiryhma>

⁹ <https://www.lvm.fi/digitalisaatio>

Suomessa on kehitetty muutamia toimialakohtaisia järjestelmiä, joiden on tarkoitus helpottaa tiedon välitystä alan sisällä, ja tarjota myös loppukäyttäjälle pääsy omaan tietoon. Kanta-palvelut sisältävät valtakunnallisia terveystietojen ja sosiaalipalveluiden tukijärjestelmiä ja Koski on opintotietojen hallintaa varten suunniteltu alusta.

Omakanta on sähköinen terveystietojen ja reseptien hallinnan palvelu, joka tarjoaa näkymän Kanta-palvelujen tallettamaan tietoon. Sen avulla voi hallita ja tarkastella sähköisiä reseptejä ja Potilastiedon arkistoon tallennettuja tietoja. Käytännössä Omakannassa pääsee näkemään itseä tai alle 10-vuotiasta omaa lasta koskevat reseptit, joita palvelun kautta voi myös uusien hoitoon liittyvät kirjaukset kuten diagnoosit ja kuvaukset sekä laboratoriotutkimusten tiedot. Omakantaan voi myös tallentaa hoito- ja elinluovutustahdon. Terveystietojen siirtämisen eri terveydenhuollon palveluiden välillä voi luvittaa tai luvan kumota Omakannassa. Palvelun käyttöä varten asiakas tarvitsee suomalaisen henkilötunnuksen ja sähköisen tunnistautumisvälineen kuten pankkitunnukset tai sähköisen henkilökortin.¹⁰ Kehitteillä on myös Omatietovaranto, johon asiakkaat voivat tallettaa itse tuottamiaan terveystietojen ja hyvinvointitietoja.

Koski on opetushallituksen opintotietojen hallintaa selkeyttävä hanke. Tavoitteena on luoda palvelu, joka kattaa julkisen valvonnan alaisilta koulutuksen tarjoajilta opinto-oikeuksia, läsnäoloa, opintosuorituksia ja tutkintoja koskevat tiedot. Nämä tiedot tuodaan kansalaisten ja viranomaisten käyttöön keskitetysti ja yhdenmukaisessa muodossa. Sähköisen palvelun kautta tiedot ovat henkilön nähtävissä, mutta ne voitaisiin myös siirtää turvallisesti muiden palvelujen käyttöön. Koskeen tallennettuja tietoja voi käyttää esimerkiksi opintojen arviointiin, työnhakuun ja opintojen suunnitteluun, mutta myös koulutuksen rahoituksen suunnitteluun ja tilastojen laadintaan. Opintotiedot liikkuvat opintojen järjestäjältä opintojen suorittajille ja siitä eteenpäin niitä tarvitseville viranomaisille, työvoimatoimistoille, Kelalle ja toisille koulutusta järjestävälle tahoille.¹¹

1.4 Euroopan unionin tietosuoja-asetus

Henkilötiedon hallinta tulee muuttumaan koko Euroopan unionin alueella toukokuussa 2018, kun Euroopan unionin uutta tietosuoja-asetusta (englanniksi General Data Protection Regulation, GDPR)¹² aletaan soveltaa. Tietosuoja-asetus antaa yksilöille joitakin uusia oikeuksia heidän koskien heidän henkilötietoaan, ja etenkin yksityisellä sektorilla henkilötiedon käsittelyn edellytykset tiukentuvat ja niiden toimeenpanosta tulee vahvempaa. Euroopan unionin jäsenmailla on kansallista liikkumavaraa sen suhteen, kuinka asetuksen osa-alueita sovelletaan julkishallinnon kohdalla. Raportin kirjoittamisen ajankohtana lainsäädäntö liikkumavaran käytöstä ja uudesta tietosuojaa koskeva yleislaista on valmisteilla.¹³

Tietosuoja-asetuksen myötä ihmisten oikeudet heitä itseään koskevaan tietoon vahvistuvat. Tarkistuspyyntöihin, joiden kautta yksilöt voivat varmistaa, että rekistereihin talletetut tiedot ovat oikein, liittyy vastaisuudessa myös oikeus saada kopio omista tiedoista sähköisessä muodossa. Asetus vahvistaa myös ihmisten mahdollisuuksia saada tietää tietovuodoista sekä siitä, kuinka heidän henkilötietojensa on käsitelty. Tarkastusoikeuden lisäksi

¹⁰ Kela, 2016

¹¹ Mäkinen & Sassi, 2017

¹² <http://www.eugdpr.org/>

¹³ TATTI Työryhmä, 2017

asetuksessa määritellään uusi oikeus, nimittäin oikeus tietojen siirtämiselle. Tämä oikeus ei kuitenkaan sellaisenaan pääsääntöisesti päde julkishallinnon käsittelemään tietoon.¹⁴

Yritysten osalta vaatimukset henkilötiedon hallinnan suhteen tiukkenevat. Tämä tarkoittaa ennen kaikkea sitä, että asetuksen mukaan yritysten on kyettävä osoittamaan, että ne noudattavat lakia. Paljon henkilötietoa käsittelevät yritykset joutuvat nimeämään tietoturvasta vastaavan henkilön. Kokonaisuudessaan tietosuoja-asetus lisää yritysten dokumentointi- ja viestintätarvetta, kun niiden on todennettava tietoturvakäytäntönsä sekä viestittävä niistä asiakkailleen sekä tämän lisäksi tiedotettava mahdollisista tietovuodoista.

Yleinen tietosuoja-asetus siis tukee MyData-periaatteiden toteutumista monilta osin. Tietosuojan edellytysten vahvistuminen voi lisätä luottamusta ja vähentää tiedon uudelleenkäyttämisen liittyviä riskejä. Sääntöjen muutosten johdosta miltei kaikkien toimijoiden täytyy harkita uudelleen omia toimintatapojaan. Henkilötiedon saatavuus MyDatana voi olla keino rekisterinpitäjille varmistaa, että he toteuttavat asetuksen vaatimukset sen suhteen, että ihmisille annetaan pääsy heidän tietoonsa. Henkilötiedon hyödynnettävyyden ja sen käsittelyn läpinäkyvyyden kannalta olisikin toivottavaa, että uuden tietosuoja-asetuksen kuvaamat yksilön oikeudet tulisivat osaksi myös julkista sektoria ohjaavaa lainsäädäntöä.

¹⁴ Oikeus tiedon siirtämiseen pätee kuitenkin vain tietoihin, jotka on "toimitettu" rekisterinpitäjälle, ja joita käsitellään suostumuksen perusteella (Yleinen tietosuoja-asetus, Artikla 20). Miltei kaikissa julkishallinnon rekistereissä tiedon käsittely perustuu muuhun kuin yksilöiden suostumukseen (ks. Liite 1 lisätietoja varten).

2 JULKISET TIETOVARANNOT JA NIIDEN HYÖDYNTÄMINEN

Aleksi Knuutila

Tietosuojavaltuutetun mukaan Suomessa on yli miljoona henkilörekisteriä.¹⁵ Rekisteritutkimuksen tukikeskus on käynyt läpi rekisteriselosteet 21:sta merkittävimmästä tietoa tallentavasta virastosta, ja löytänyt yhteensä 177 rekisteriä.¹⁶ Kuntien tasolla erilaisia tietojärjestelmiä ja rekistereitä löytyy helposti satoja: Helsingin kaupungin julkaisemista selosteista löytyy 778 tietojärjestelmää ja 209 rekisteriä.¹⁷ Rekistereiden lukumäärä ei tietysti ole kovin osuva mitta kuvaamaan tallennetun tiedon määrää. Kuten mainituissa selvityksissä todetaan, useampi tietokokonaisuus katetaan ajoittain yhdellä rekisteriselosteella, joten erillisten tietojärjestelmien määrä voisi olla vielä paljon suurempi.

Henkilötietoa tallettuu siis lukuisiin tietojärjestelmiin entistä suurempia määriä. Tässä kappaleessa käymme lyhyesti läpi aikaisempia kokonaisvaltaisia selvityksiä siitä, mitä tietovarantoja julkishallinnon toimijoilla on, ja kuinka tietoa käytetään.

Henkilörekistereitä, jotka MyData-periaatteiden mukaan tulisi avata ihmisten hyödynnettäviksi, on lukuisia. Käytännössä viranomaisten ja muiden tahojen pitää tehdä valintoja siitä, mitä tietovarantoja he kehittävät ja kytkevät tietojärjestelmiin, jotka mahdollistavat niiden hallinnan. Siksi käymme myös läpi joitakin yleisiä periaatteita siitä, mitkä tietovarannot olisivat mahdollisesti hyödyllisimpiä MyDataan liittyville sovelluksille.

2.1 Julkiset henkilötiedon varannot

Ennen kuin käymme läpi yleisellä tasolla julkisia tietovarantoja, on aihetta lyhyesti määritellä, mitä tähän kategoriaan ylipäätään kuuluu. Julkishallinnon ja yhteiskunnan muiden osien rajanveto tehdään eri tavoin eri yhteyksissä. EU:n julkista tietoa koskevassa vihreässä kirjassa listataan kolme tapaa tehdä erottelu:

- Funktionaalinen lähestymistapa, jossa julkinen sektori sisältää valtion toimivaltaa käyttävät sekä julkisia palveluita tuottavat tahot.
- Lainsäädäntöön perustuva lähestymistapa, jonka mukaan asianomaisissa laeissa luetellut elimet lasketaan julkisiksi.
- Rahoitukseen perustuva lähestymistapa, jonka mukaan julkiseen sektoriin kuuluvat ne organisaatiot, jotka rahoitetaan pääsääntöisesti julkisin varoin.¹⁸

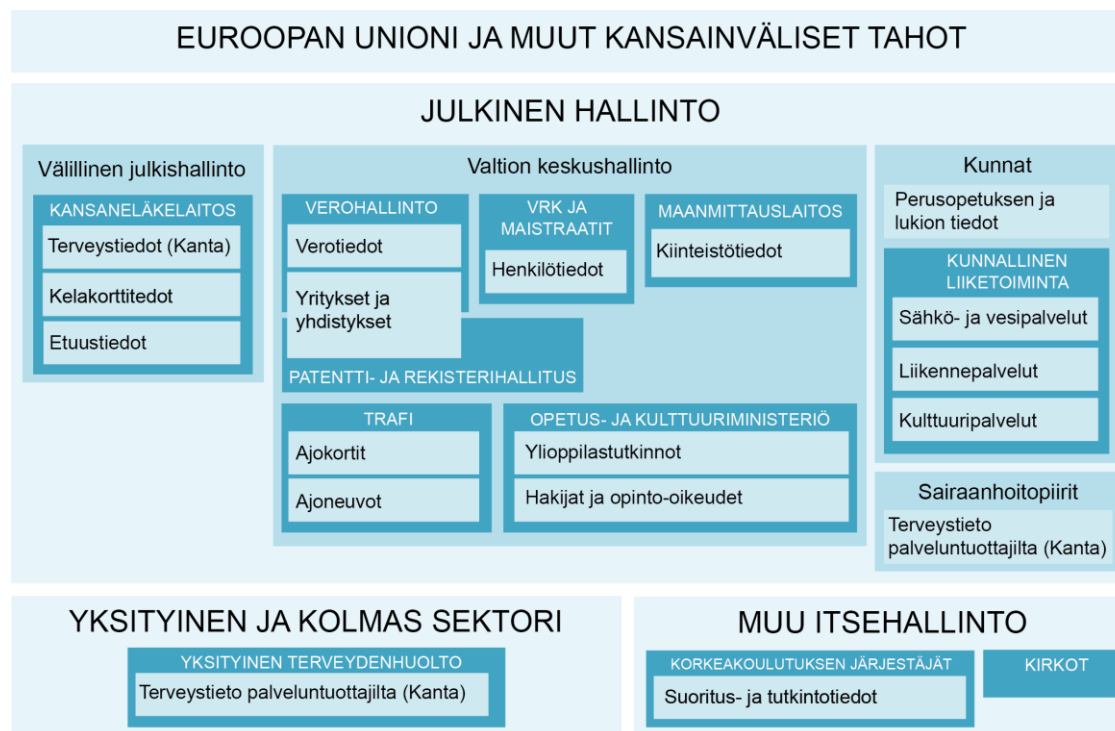
¹⁵ Salminen, 2011

¹⁶ Tukikeskus, 2017

¹⁷ Poikola, Hjelm, & Schildt, 2017

¹⁸ Euroopan komissio, 1998

Kuva 2.1 Tärkeimpiä julkishallinnon henkilötiedon varantoja



Kuvio esittää kuinka jotkin merkittävät henkilötiedon varannot ovat jakautuneet eri julkishallinnon osiin, ja mainitsee joitakin julkisiin palveluihin liittyviä julkishallinnon ulkopuolisia tietovarantoja.

Kerättyjen henkilötietojen kannalta esimerkiksi valtion ja kuntien liikelaitokset sekä kuntien omistuksessa olevat yhtiöt ovat merkittäviä rajatapauksia. Ne toteuttavat useita merkittäviä kuluttajille suunnattuja palveluita kuten julkista liikennettä tai energian tuotantoa. Sellaisina niiden haltuun voi kerääntyä merkittäviä mm. asiointiin ja kulutuskäytäntöihin pohjaavia tietovarantoja, kuten esimerkiksi liikkumiseen liittyviä tietoja matkustuskorteilta. Organisaatioiden tyypit myös osin määräävät tiedonhallintaan liittyviä sääntöjä. Esimerkiksi liikelaitokset ovat pääsääntöisesti kunnallisen hallinnon alla ja niihin pätee julkisuusperiaate, mutta eivät maksuperustelain rajoitukset tietotuotteiden hinnoittelusta. Tilanne voi edelleen muuttua, jos liikelaitoksia esimerkiksi yhtiöitetään. Joitakin perustavanlaatuisia tietovarantoja on jo yhtiötettyjen toimijoiden hallussa, kuten esimerkiksi Suomen Posti Oy:n osoiterekisteri.

Tässä selvityksessä keskitymme yksinkertaisuuden vuoksi suurimmaksi osaksi valtiotasolla kerättyihin rekistereihin, vaikka samalla pyrimme myös kuvaamaan kuntatasoa sekä liikelaitoksia. MyData-periaatteita tukevia järjestelmiä rakennettaessa kannattaisi luoda tekninen mahdollisuus sekä oikeat kannustimet sille, että myös liikelaitosten tietovarantoja voidaan hyödyntää.

Katsaus julkishallinnon tietovarantoihin

Valtiovarainministeriön Julkisen hallinnon tietoarkkitehtuuri -hanke on yksi kattavimmista julkishallinnon tietovarantoja kartoittaneista töistä. Hankkeen loppuraportissa ns. loogiset tietovarannot (siis tietosisältöjen kokonaisuudet, ei yksittäiset rekisterit) jaetaan neljään eri ryhmään. Yhteensä ryhmään kuuluu satoja eri varantoja. Käytyämme läpi tämän kartoituksen sekä aikaisemmin mainitut rekistereiden listaukset, teimme yhteenvedon viranomaisten hallussa olevasta tiedosta (taulukko 2.1.). Taulukko pyrkii olemaan suuntaa

antava, eikä pysty olemaan täysin kattava kuvaus julkishallinnon ja julkisten organisaatioiden rekistereistä. Käytetty jaottelu perustuu ensi sijassa tiedon sisältöön, eikä välttämättä hallinnon alaan tai rekisteriä pitäviin viranomaisiin. Taulukosta käy myös ilmi, kuinka tiettyyn aiheeseen (kuten esimerkiksi henkilön talouteen ja tuloihin) liittyvät tiedot voivat toistaiseksi olla jakautuneena lukuisten eri virastojen ja tietojärjestelmien välillä.¹⁹

Taulukko 2.1. Viranomaisten hallinnoimat henkilötiedon lajit

	Sisältö	Rekistereitä
Henkilötiedot	Nimi, henkilötunnus, asuintiedot, yms.	Väestörekisteri (VRK)
Terveystiedot	Potilastiedot, mittaukset, sähköiset reseptit, suostumukset ja tahdonilmaisut	Lukuisia terveyspalvelujen rekistereitä ja yhteinen Kanta-järjestelmä
Taloustiedot	Tulot eri vuosilta, maksetut verot, saadut etuudet	Verotietokanta (Verohallinto), Etuusrekisteri (Kela), Eläketapahtumarekisteri (Eläketurvakeskus), Toimeentulotukirekisteri (THL)
Kiinteistötiedot	Tiedot sijainnista, omistajista, viranomaispäätöksistä, yms.	Kiinteistörekisteri (MML), Rakennus – ja huoneistorekisteri (VRK)
Yritystiedot	Yrityksen yhteystiedot, omistajuus, toimiala, yms.	Kaupparekisteri, säätiörekisteri ja yhdistysrekisteri (PRH)
Liikenne- ja ajoneuvotiedot	Ajoneuvojen omistajuus ja vakuutukset, sekä liikkumistiedot julkisen liikenteen palveluista	Ajoneuvoliikennerekisteri (Trafi), liikennepalvelujen asiakasrekisteri
Oikeustiedot	Rikokset, sakot	Rikosrekisteri, sakkorekisteri (Oikeusrekisterikeskus)
Kulutus- ja asiakastiedot	Esimerkiksi sähkön tai veden kulutus, erilaisten kulttuuripalvelujen käyttö	Kunnallisten palvelujen, liikelaitosten ja yhtiöiden asiakasrekisterit
Koulutustiedot	Tiedot tutkinnoista, opinto-oikeuksista, käydyistä kursseista sekä hakemuksista	Hakija- ja opinto-oikeusrekisterit (Opetushallitus), opiskelijatietojärjestelmät (oppilaitokset)

Yhteenvedo viranomaisten hallinnoimista tietotyypeistä, niiden sisällöstä ja rekisterinpitäjistä.

Yksilön oikeudet eri tietovarannoissa oleviin henkilötietoihin

Ihmisten oikeuksista heitä koskevaan henkilötietoon määräävät yleislaait, kuten henkilötietolaki ja julkisuuslaki, sekä monet hallintoalakohtaiset erityislaait. Oikeudet koskien pääsyä tietoon, sen käytön kieltoa tai tiedon poistamista riippuvat siis tiedon tyypistä. Aiheeseen liittyvää lainsäädäntöä käymme läpi tarkemmin liitteessä 1. Tässä esitämme yhteenvedona neljä eri luokkaa, joihin julkiset henkilötiedot kuuluvat sen mukaan, millaisia rekisteröityjen oikeudet ovat käsitellyn tiedon suhteen. Luokat on ryhmitelty heikommista vahvempiin oikeuksiin. Huomattavaa on, että suurin osa julkishallinnon tiedoista kuuluu luokkaan 3.

- 1. Salaiset henkilötiedot.** Viranomaisten lain pohjalta käsittelemät tiedot, joihin yksilöllä ei ole pääsyä eikä sen käyttöä voi valvoa. Esimerkki: Turvallisuuden tai poliisin toimintaan liittyvät rekisterit.
- 2. Tieto, johon liittyy arvioitavia riskejä.** Tähän tietoon pääsy on ammatinharjoittajien arvioinnin varassa, ja sen siirto toisiin sovelluksiin on rajattua. Toisaalta myös salassapidettävän tiedon luovutus kolmannelle osapuolelle usein vaatii yksilön suostumusta. Esimerkki: Terveys- ja sosiaalitiedot.

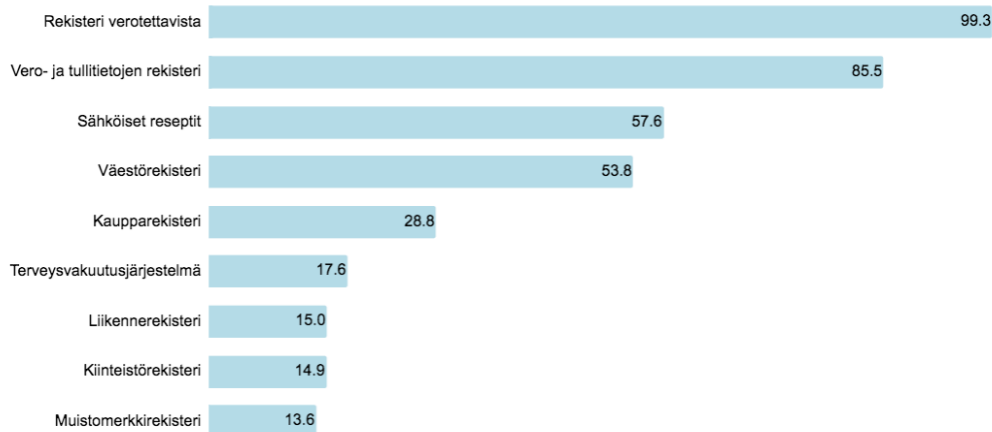
¹⁹ Tulevan kansallisen tulorekisterin on tarkoitus helpottaa ja keskittää tuloihin liittyvän tiedon keräämistä.

3. **Laissa määriteltyjen viranomaisten tehtävien tai palvelusuhteen pohjalta käsitelty tieto.** Tähän tietoon henkilöillä on pääsääntöisesti pääsy, sen käytön kieltäminen tai poistaminen ei kuitenkaan ole mahdollista. Esimerkki: Verotiedot.
4. **Suostumuksen pohjalta käsitelty tieto.** Tämän tiedon käsittelyä yksilö voi kieltää ja pyytää sen poistoa. Esimerkki: Terveystiedot, jotka on siirretty kolmansille osapuolille suostumuksen perusteella.

Tiedon uudelleenkäytön muodot ja määrät

Viron julkisten viranomaisten tiedonsiirtoväylä, X-Road, on ollut olemassa riittävän pitkään, jotta sen käytöstä on syntynyt kattavasti tilastotietoja. Kuva 2.2 esittää vuoden 2016 10 tärkeintä tietolähdettä käsiteltyjen tietokyselyiden mukaan. Merkittävimmiksi tiedon tyypeiksi lukeutuvat henkilötiedot, verotiedot, resepti- ja vakuutus tiedot terveysalalta, sekä kaupparekisteri ja ajoneuvotietoihin keskittyvä rekisteri. X-Road-järjestelmän asiakkaina ovat sekä viranomaiset että yksityiset toimijat. Huomattavaa on, että merkittävä osa palveluväylää pitkin kulkevasta tiedosta näyttäisi olevan henkilötietoa, vaikka tilastot eivät tätä erottelua tee.

Kuva 2.2. 10 merkittävintä tiedonlähdettä X-Road-järjestelmässä 2016



Kyselyitä vuonna 2016, tuhatta

Viron X-Road-järjestelmän 10 eniten tietokyselyitä vastaanottanutta tietojärjestelmää vuonna 2016. Lähde: Estonian Information System Authority²⁰

Vastaavaa tilastotietoa Suomen eri viranomaisiin kohdistuvista tietokyselyistä on vähemmän.²¹ Väestörekisterikeskus ja Trafi ovat tuottaneet "tietotilinpäätöksiä", joissa virastot käyvät läpi tietotuotteitaan, miten niitä on tuotettu ja kuinka paljon kysyntää niille on.²² Esimerkiksi Väestörekisterikeskus (VRK) ilmoittaa väestötietojärjestelmän jakaneen vuonna 2015 yhteensä noin 550 miljoona "tietoyksikköä". Tietokyselyt voi jakaa eri tyyppisiin niiden kysyjän tai käyttötarkoituksen mukaan. Esimerkiksi VRK:n tapauksessa "liiketaloudellisen sektorin" osuus kaikista luovutetuista tiedoista oli noin 400 miljoona yksikköä, joista noin 125 miljoona liittyi asiakasrekisterien päivityksiin.²³

²⁰ Viron tilastot saatavissa osoitteessa <https://www.ria.ee/en/statistics-about-x-road.html>. Vastaavia tilastoja Iso-Britanniasta, ks. <https://www.gov.uk/performance/government-transactional-services/transactions-overview>

²¹ Valtiovarainministeriö, 2012, s. 36

²² Trafi, 2017b; Väestörekisterikeskus, 2016

²³ Väestörekisterikeskus, 2016, s. 36

Suomessa rekisteritietoja on myös mahdollista käyttää laajasti tutkimusta varten. Kaiken kaikkiaan rekisteritutkimukseen liittyvät tiedonsiirron määrät ovat suhteellisen pieniä: vuonna 2011 merkittävät sosiaali- ja terveystietoisuudella luovuttivat tietoja tutkimukseen noin 350 kertaa.²⁴

Tietovarantoon kohdistuvan kysynnän määrä kertoo osaltaan sen arvosta käyttäjille. Myöhemmin tässä luvussa käymme läpi muita tapoja ennakoita, mitkä tietovarannot olisivat merkittävimpiä henkilötiedon uudelleenkäytön kannalta.

Tiedon tyyppiä: Perusrekisterit ja alustojen keräämä tieto

Julkishallinnon hallitsemista tiedoista puhuttaessa käytetään usein perusrekisterin tai perustietovarannon käsitteitä. Näiksi on myös laskettu henkilöitä, yhteisöjä tai kiinteistöjä yksilöiviä rekistereitä, jotka usein muodostetaan lakisääteisesti.²⁵ Esimerkiksi Rauno Korhonen listaa perusrekisterien välttämättömiksi ominaisuuksiksi niiden kattavuuden, monikäyttöisyyden, tietojen suojauksen ja tunnuseheyden.²⁶ Kattavuus tarkoittaa sitä, että perusrekisterin on tarkoitus listata kaikki rekisterin määritelmään sisältyvät yksiköt (kuten henkilöt tai yritykset). Perusrekisterien tietosisältö on siis kapeasti ja tarkasti määritelty, ja niiden tietosisältö voi olla hitaasti muuttuvaa. Tämän datan käytännön merkitys ei ole pelkästään siinä, että se kuvaa tiedon kohteesta tiettyjä ominaisuuksia (esimerkiksi väestörekisteri sisältää osoitteita ja puhelinnumeroita). Perusrekisterit viime kädessä määrittävät yksiselitteisen tunnuksen kullekin yksikölle, jota voidaan käyttää, kun muissa tietovarannoissa halutaan viitata samaan henkilöön. Perusrekisterit siis mahdollistavat tiedon yhdistelemisen ja prosessoinnin useasta eri lähteestä.

Näiltä osin perusrekisterit vertautuvat havainnollisella tavalla muun tyyppisiin tietokantoihin. Palveluihin kertyy jonkinlainen asiakasrekisteri ja kasvava määrä tietoa jokaisesta palvelutapahtumasta. Kirjastolle karttuu tiedot kirjoista, joita sen asiakkaat lainaavat. Yleisemmin voisi sanoa, että on erilaisten digitaalisten alustojen ominaisuus, että ne keräävät tietoa toiminnasta, jota ne tukevat.²⁷ Sähköinen infrastruktuuri yhdistettynä kasvavaan tallennuskapasiteettiin johtaa siihen, että erilaisten tapahtumien ja transaktioiden yksityiskohtien tallentamisen kynnys pienentyy entisestään. Ihmisten toiminnasta sivuvaikutuksena syntyy eräänlainen "datajalanjälki". Jotain toimintaa tai palvelun käyttöä seuraavat tietokannat poikkeavat perusrekistereistä sillä tavalla, että niiden sisältämien tietueiden määrä ei ole ennalta rajattu. Niin kauan kuin alusta on käytössä, tiedon määrä kasvaa jatkuvasti. Näitä tietovarantoja kutsumme tässä yhteydessä asiakastiedoiksi.

Perusrekisterien ja asiakastietojen erottelu voi joskus olla vaikeaa. Myös perusrekistereihin lukeutuvien tietojärjestelmien sisältö voi vaatia jatkuvaa päivytystä. Esimerkiksi kiinteistörekisterin omistustiedot voivat muuttua kauppohenkilöiden seurauksena. Tällä hetkellä monia perusrekistereitä ylläpitävät nimenomaan julkisen hallinnon toimijat johtuen niiden viranomaistehtävien hoitamiseen tarvitsemasta tiedosta. Samalla julkisella sektorilla on vain rajattu määrä alustoja, joihin kustakin yksilöstä kertyisi suuri määrä yksityiskohtaista tietoa. Julkisiin palveluihin, jotka tuottavat merkittäviä määriä käytöstään kertyvää henkilötietoa, kuuluvat ainakin infrastruktuuripalvelut kuten liikenne ja energia sekä julkisesti tuotetut palveluklusterit kuten terveyspalvelut. Näissä digitaaliset toimintatavat ovat kehittyneet jo

²⁴ Rekisteritutkimuksen Tukikeskus, 2012, s. 4

²⁵ Perusrekisterien määritelmät vaihtelevat jossain määrin. Esimerkiksi Kunnallishallinnon tietotekniikkaneuvottelukunta sisällyttää määritelmäänsä myös osoiterekisterin, saanto-, lainhuuto- ja vuokrasopimusrekisterin, suunnitelmarekisterin, kiinnitysrekisterin sekä toimipaikkarekisterin.

²⁶ Korhonen, 2003

²⁷ Ks. esimerkiksi Ailisto ym., 2016 s. 31-32

niin pitkälle, että dataa kerätään järjestelmällisesti. Sähköisen asioinnin kehittyessä julkishallinnon ylläpitämät digitaaliset alustat tulevat edelleen yleistymään.

Eri tyyppisistä datoista syntyy myös erilaista arvoa niiden käyttäjille. Suurten tietoaaineistojen pohjalta voi tehdä erilaisia profiloitteja tai analyysejä, visualisoinneista tilastollisiin laskelmiin, kun taas perusrekisterien käytännön merkitys liittyy usein tiedon yhdistelyn mahdollisuuteen tai rekisteriotteiden käyttöön todisteina. Näitä hyötyjä käydään tarkemmin läpi myöhemmin tässä luvussa.

2.2 Korkean lisäarvon MyData-tietovarannot

Viranomaisten hallussa olevan henkilötiedon määrä on suuri. Haastatteluissamme kävi ilmi, että joissakin yksittäisissä virastoissa on kymmeniä tuhansia taulukoita, jotka sisältävät tietoa yksittäisistä henkilöistä. Kaiken tämän tiedon muokkaaminen saatavaksi MyData-mallin mukaisena ei olisi käytännöllisesti mahdollista. Henkilötiedon saatavuuden parantaminen edellyttää priorisointia samaan tapaan kuin avoimen datan julkaisun kohdalla, jossa parhaat keinot priorisointiin ovat olleet pitkään keskustelun aiheena.²⁸ Toisaalta tiedon avaamisen suunnitelmissa on myös todettu, että aineiston täsmällinen valikointi etukäteen voi olla vaikeaa, ja voi olla vaikeaa tietää, mihin tarkoituksiin ja kenen toimesta tietoa tullaan käyttämään. On mahdollista, että datalle löytyy piilevää kysyntää, joka käy ilmi vasta siinä vaiheessa, kun se tulee saatavaksi ja tunnetummaksi.²⁹ Seuraavaksi käymme läpi joitakin perusteita sille, minkä tiedon avaaminen olisi hyödyllisintä tiedon käyttäjille tai erilaisten palveluiden kehityksen kannalta. Näkemykset perustuvat tekemiimme teemahaastatteluihin.³⁰

Perusrekisterien tiedot

Yllä kuvatun jaottelun mukaan kutsumme tässä perusrekistereiksi tietokantoja, jotka ovat kattavia luetteloita ja usein tietosisällöltään selkeästi rajattuja. Tällaisia ovat esimerkiksi rekisterit ajokortin omaavista ihmisistä tai ylioppilastutkinnon suorittaneista. Useissa tapauksissa näiden rekisterien tietoja tarvitaan myös todisteina tai virallisina asiakirjoina. Esimerkiksi ajokorttite tiedot pitää joskus esittää koulutusta tai työpaikkaa varten. Ajokorttiotteen, kuten monet muut asiakirjat, saa jo nyt sähköisesti haltuunsa tulosteena. Jos vastaavia virallisia tietoja olisi saatavilla myös rakenteisena datana, voitaisiin näitä tietoja käyttäviä palveluprosesseja automatisoida ja vähentää kansalaisten hallintotaakkaa. Kenties esimerkiksi oman ajokorttiotteen tiedot voisi lähettää usealle vakuutuksen tarjoajalle, kun niitä kilpailutetaan.³¹

Monien rekisterien sisältämä tieto on arvokasta myös siksi, että se mahdollistaa tietojen yhdistelemisen eri lähteistä. Tietojen yhdisteltävyys edellyttää rekisterien välisiä yhteyksiä niiden avaintiedoissa.³² Yhdistelemällä tietoja voidaan esittää datana entistä monimutkaisempia suhteita. Esimerkiksi yhdistelemällä kiinteistörekisterin, väestötietojärjestelmän ja mahdollisesti yritysrekisterin tietoja, voitaisiin katsoa, asuuko

²⁸ Esimerkiksi Koski, Honkanen, Luukkonen, Pajarinen, & Ropponen, 2017, s. 43-44.

²⁹ Poikola, Kola, & Hintikka, 2010, s. 38

³⁰ Asiantuntijoiden näkemyksiä siitä, mitkä tietovarannot perusrekistereistä ovat heidän toiminnalleen merkittävimpiä, on myös selvitetty mm. osana JUHTAn Pertiva-hanketta, joka vuonna 2012 tilasi kyselyn sidosryhmiltä perustietovarantojen kehittämistarpeista. Kyselyyn osallistui ihmisiä julkiselta sektorilta, liike-elämästä sekä tieteen ja tutkimuksen parissa työskentelevistä. Merkittävimmät kehitystarpeet liittyivät tälle joukolle mm. tietosuojan parantamiseen niin, että tiedon yhteiskäyttö helpottuu, sekä vakio-muotoisten tietopalveluiden saatavuuteen. Tärkeimmiksi tietotyypeiksi vastaajat nostivat erityisesti terveydenhuoltoon liittyvän tiedon sekä yritys- ja henkilötiedot (Jauhiainen & Kuusinen, 4.9. 2012).

³¹ Huomionarvoista on, että tiedon ajantasaisuus on haaste. Esimerkiksi ajokorttien kohdalla tilanne on yksinkertaisempi sikäli, että niiden voimassaoloaika on tarkkaan määritelty. Sen sijaan vaikka omistustietojen kohdalla pitäisi merkittävässä transaktioissa tarkistaa rekisteristä, ovatko esitetyt tiedot ajan tasalla.

³² Pertiva, 2016

henkilö omistamassaan asunnossa. Lisäksi samoja asioita täydentävästi kuvaavien tietolähteiden yhdistely helpottuu. Voi kuvitella esimerkiksi, että omistaja hankkii asunnolleen hinta-arvion sekä kuntotarkastuksen. Jos näiden arvioiden tarjoajat käyttäisivät samojen kansallisten perusrekisterien tunnuskeitä, olisi tietojen automaattinen yhdistäminen (esimerkiksi asunnon myyntiin käytetyssä palvelussa) yksiselitteisen tunnusten avulla helppoa.³³

Perusrekisterien kaltaiset tietovarannot tarjoavat siis lisäarvoa MyDatana etenkin silloin kun:

- ne kuvaavat myönnettyjä todistuksia tai lupia, joita ihmiset usein tarvitsevat toiminnassaan.
- vastaavia tietoja ei ole saatavilla muualla.
- kyseiset henkilötiedot eivät ole arkaluontoisia eikä niiden käsittelyyn liity poikkeuksellisen suuria riskejä.
- tietojen yhdisteleminen on mahdollista, esimerkiksi hyvin toteutettujen avaintietojen pohjalta.
- tietojen oikeellisuus ja korjaaminen on keskeistä; tietojen parempi saatavuus voi myös parantaa niiden sisältöä, jos ihmiset todennäköisemmin oikaisevat omia tietojaan.

Asiakkuuteen liittyvät tiedot

Iso-Britanniassa kansalaisen oikeuksien vahvistamisesta henkilötiedon hallinnassa puhuttiin jo vuonna 2011, usein käyttäen termiä "midata".³⁴ Tässä yhteydessä tiedon siirrettävyys ja käytettävyys ymmärrettiin keinoksi sille, että yritysten asiakkaat pystyisivät ymmärtämään omia kulutustottumuksiaan, vertailemaan palvelujen tarjoajia ja halutessaan vaihtamaan niiden välillä. Myös julkisen sektorin hallussa olevan henkilötiedon sujuvampi hallinta voi helpottaa asiakassuhteiden siirtymistä ja käyttäjien mahdollisuutta valita. Esimerkiksi toiseen kuntaan muutettaessa tai terveystietojen siirtäessä ihmisten tietojen kulkeminen mukana on tärkeää. Päälimmäinen ongelma on järjestelmien yhteensopivuuden puuttuminen, mutta joskus kyse voi olla myös puutteista viranomaisten lupajärjestelyissä. Joissakin tapauksissa arkaluontoisten tietojen jakaminen edellyttää suostumuksen saamista, mikä tällä hetkellä tapahtuu usein paperilomakkeilla. Kummassakin tapauksessa MyData-tilin kautta tapahtuva suostumusten hallinta voisi tehdä käsittelyä sujuvammaksi.

Joissakin tapauksissa julkiset palvelut ovat yksi osa laajempaa jatkumoa palveluita, joihin kuuluu myös paljon yksityisiä tarjoajia. Näin on etenkin terveystietojen palveluissa, mutta myös esimerkiksi liikenne- ja koulutuspalveluissa. Erityisesti silloin kun palveluntarjoajien hallussa olevat henkilötiedot sisältävät suuria tietomääriä tai päivittyvät nopeasti, voi niiden siirrettävyys muodostua esteeksi toisien palvelujen käyttämiselle tai aiheuttaa hallintotaakkaa yksilöille, jotka ovat vaihtamassa palvelua. On myös tapauksia, joissa julkiset liikelaitokset tai julkisessa omistuksessa olevat yhtiöt tarjoavat samoja palveluja kuin yritykset ja muut toimijat, kuten sähkömarkkinoilla. Näissä tapauksissa omien tietojen

³³ Samalla on huomattava, että joissakin tapauksissa tietojen yhdistelemisen mahdollisuus ei välttämättä ole yksilön etujen mukaista. Monet henkilöt saattavat esimerkiksi käyttää useita vaihtoehtoisia henkilöllisyyksiä eri palveluissa turvatakseen omaa yksityisyyttään ja estääkseen sen, että eri yhteyksissä kerääntyvät tiedot olisi yhdistettävissä toisiinsa.

³⁴ Department for Business, Innovation & Skills, 2012

tarkastelu voisi auttaa asiakkaita ymmärtämään paremmin omien kulutustottumuksiensa kustannuksia ja siten edesauttaa parempien valintojen tekemistä palveluntarjoajien suhteen.

Asiakastietojen kaltaiset tietovarannot ovat siis erityisen arvokkaita MyDatana silloin kun:

- ne sisältävät usein muuttuvia tai suuria tietomääriä, joiden siirtäminen ilman tätä tukevia järjestelmiä olisi vaikeaa.
- kyse on palvelujen tyypistä, jota toteuttavat useammat eri palveluntarjoajat (kuten esimerkiksi terveystietojen yksityiset ja julkiset tarjoajat.)
- tiedot on esitetty muodossa, joka on yhteensopiva eri palveluntarjoajien kesken.
- tiedot auttavat eri palveluntarjoajien välisessä vertailussa, kuten esimerkiksi sähkön tarjoajien kanssa.
- asiakastiedoilla on laaja joukko vaihtoehtoisia käyttötarkoituksia, kuten esimerkiksi ihmisten taloudellista tilannetta ja omistussuhteita kuvaavilla tiedoilla.

Nykyiset näkymät omiin tietoihin

Monilla viranomaisilla on jo heidän omaan alaansa kohdistuvia sähköisiä palveluita, joiden kautta käyttäjät saavat omat tietonsa näkyville. Esimerkiksi Kanta-palvelun käyttöliittymän kautta voi katsoa omia resepti- ja potilastietoja. Joissakin tapauksissa tällaiselle palvelulle on lakisääteiset vaatimukset. Lakiin asetetut vaatimukset voivat myös asettaa rajoituksia sille, millä lailla tietoihin pääsy on järjestetty. Kantaan tallennettuja tietoja ei tätä kirjoitettaessa ole voitu tehdä saataviksi koneluettavassa muodossa. Joissakin tapauksissa myös viranomaiset itse on määriteltä tahoksi, jonka on toteutettava palvelu kansalaisen omiin tietoihin pääsemiseksi.

Näkymät omiin tietoihin eivät pääsääntöisesti näytä kaikkea viranomaisten hallinnoimaa tietoa vaan ainoastaan valikoidut kentät. Yksi merkittävä peruste näytettävien kenttien valinnalle on tavoite näyttää ne viranomaistoimintaan perustuvat tiedot, joita ihmiset mahdollisesti haluaisivat korjata tietokannoista. Mahdollisuus korjata omia tietoja ja tehdä rekisterien sisältöjä kattavammiksi on keskeinen motiivi sähköisten näkymien luomiselle. Ensimmäinen askel siihen, että kansalaiset saadaan osallistumaan omien tietojensa tarkastamiseen ja oikaisemaan virheelliset tiedot, on tehdä selvemmäksi, mitä tietoja heistä ylipäätään on tallennettu.

Kynnys sille, että nyt sähköisissä palveluissa näkyvät tiedot kytkettäisiin myös MyDatana mahdollistamaan järjestelmään, on pääasiassa pieni. Lisäksi julkisen sektorin toimijoilla on valmiiksi velvollisuus tuoda suomi.fi-palveluväylään ne tiedot, jotka ovat jo valmiiksi saatavilla tietoverkon välityksellä.³⁵ Näiltä osin sähköisissä palveluissa jo näytetyn tiedon avaaminen MyDataksi voitaisiin mahdollisesti toteuttaa osana jo tarpeellista tietojärjestelmämuutosta.

³⁵ Lähde: <https://esuomi.fi/palveluntarjoajille/palvelunakymat/#toggle-id-1>

3 MYDATAN TOTEUTTAMINEN: TÄRKEIMMÄT PALVELUT JA INFRASTRUKTUURI

Aleksi Knuutila

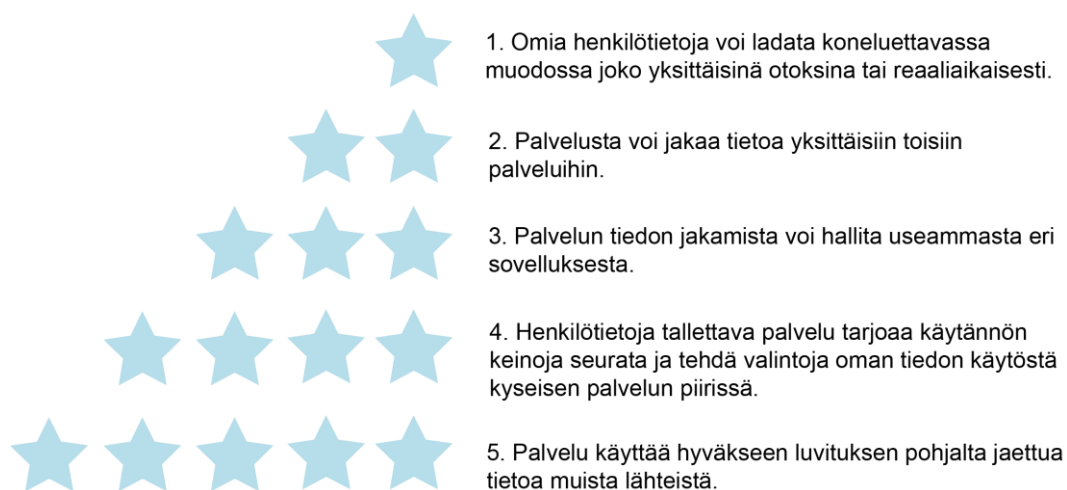
Ensimmäisessä luvussa MyData määriteltiin henkilötiedon osajoukoksi, jonka yksilö voi vähintäänkin ladata itselleen ja jakaa toisiin tietojärjestelmiin. Joidenkin tietovarantojen kohdalla on lisäksi käytännön keinoja seurata ja tehdä valintoja oman tiedon käytöstä. Näin ollen MyData-periaatteita voikin lähteä toteuttamaan usealla erilaisella tavalla.

Selvennämme erilaisten toteutusten eroja käymällä ensin läpi ajatuksen MyDatan useasta toinen toistaan vaativammasta tasosta. Sen jälkeen kuvaamme erilaisia sähköisiä tiedon jakamisen ja hallinnan palveluja, joiden pohjalta eri tasot toteutuisivat. Lopuksi kuvaamme myös kolme vaihtoehtoista arkkitehtuuria, jotka ovat erilaisia tapoja järjestää palvelujen tuottaminen.

3.1 MyDatan eri tasot

Samalla lailla kuin monia avoimen datan luokitteluja ja laadun arviointimalleja³⁶, seuraavaa viittä MyData-periaatteiden toteutumisen tasoa voidaan ajatella toinen toistaan vaativampina kriteereinä sen suhteen, lasketaanko jokin henkilötiedon varanto MyDataksi. Kriteerien esittämisen tarkoitus on selkeyttää erilaisten MyData-periaatteiden toteutusten eroja ja helpottaa niiden vertailua. Yksittäiset tietovarannot tai palvelut voisivat myös aloittaa yksinkertaisimmista MyDatan askeleista ja pyrkiä kohti haastavampia luokituksia.³⁷ Kunkin tason kohdalla viittaamme tämän tason osalta tärkeimpiin palveluihin, joista kirjoitamme myöhemmin tässä luvussa.

Kuva 3.1 MyData-tasot



Viisi toinen toistaan vaativampaa MyData-periaatteiden toteutumisen tasoa

³⁶ Esimerkiksi Tim Berners-Lee on kehittänyt viiden tähden avoimen datan laadun arviointimallin kriteereineen, ks. <http://5stardata.info/>

³⁷ Huomattavaa on, että kriteerit on kirjoitettu niin, että niitä voi soveltaa yksittäiseen tietoa tallettavaan palveluun, kuten vaikkapa julkisen liikenteen tarjoajaan. Kriteerien toteutuminen ei kuitenkaan riipu pelkästään yksittäisen tahon toimista tiedon avaamiseksi vaan myös datan käyttöä tukevan laajemman infrastruktuurin olemassaolosta. Yksittäisen luokittelun sisällä voi olla merkittäviä eroja, esimerkiksi sen suhteen, kuinka hyvin pääsy omiin tietoihin on järjestetty. Ihanteellisesti MyDatan eri tasot voisi nähdä toinen toistaan seuraavina, niin että jokainen luokitus vaatii myös edellisten toteuttamista. Näin ei käytännössä aina ole, koska MyDatan eri osa-alueet voivat toteutua toisistaan riippumatta.

1. Omia henkilötietoja voi ladata koneluettavassa muodossa joko yksittäisinä otoksina tai reaaliaikaisesti. Tämä on MyDatán minimiehto. Esimerkki: Yhdysvaltojen Blue- ja Green Button -järjestelmät. Katso alla: "Napit" tiedon lataamiseen.
2. Palvelusta voi jakaa tietoa yksittäisiin toisiin palveluihin. Näin tehdessään voi valita, mitä tietoa siirretään, mihin tarkoitukseen ja kuinka pitkäksi aikaa. Esimerkki: Useat suuret pilvipalvelut, joihin voi kytkeä muita sovelluksia, kuten Google tai Facebook. Katso alla: Dynaaminen sopiminen.
3. Palvelun tiedon jakamista voi hallita useasta eri sovelluksesta. Tämä on mahdollista, jos palvelu tukee avoimia standardeja käyttöilupien järjestämisessä. Tällöin käyttäjällä on valinnanvaraa sen suhteen, mitä ohjelmia ja käyttöliittymää hän käyttää oman tietonsa hallintaan. Katso alla: Ohjelmoitu sopiminen.
4. Henkilötietoja tallentava palvelu tarjoaa käytännön keinoja seurata ja tehdä valintoja oman tiedon käytöstä kyseisen palvelun piirissä. Julkishallinnon kohdalla tämä voi tarkoittaa lokitietojen sekä lain sallimien valintojen (kuten kiellon tietojen käytöstä markkinoinnissa) tarjoamista. Esimerkki: Privacy Dashboard -sovellukset. Katso alla: Lokitiedot ja yksityisyysasetukset.
5. Palvelu käyttää hyväkseen luvituksen pohjalta jaettua tietoa muista lähteistä joko tekemällä mahdolliseksi omien tietojen päivityksen jaetun tiedon pohjalta tai tuottamalla useiden tietojen lähteiden yhdistelyyn perustuvia palveluita. Tämä taso ei ole kaikille palveluille olennainen. Katso alla: Tehokas datakatalogi.

3.2 Tiedon jakaminen

Tässä esittelemme lyhyesti kolme tapaa järjestää henkilötiedon jakaminen yksilöiden antamien lupien pohjalta siten, että jokainen näistä tavoista toteuttaa kolme ensimmäistä tasoa yllä kuvatusta luokittelusta. Linkki kotisivulla on tapa toteuttaa ensimmäisen tason edellytykset, kun taas dynaaminen sopiminen vastaa toista tasoa ja ohjelmoitavan sopimisen pohjalta toteutuisi kolmas taso.

"Napit" tietojen lataamiseen

Yksinkertaisimmillaan omien tietojen hakeminen voi tapahtua niin kuin tiedoston lataaminen. Esimerkiksi Yhdysvalloissa on käynnissä *Green Button*- ja *Blue Button Initiative* -hankkeet, joissa sähkö- sekä terveyspalveluiden tarjoajat kautta maan tarjoavat samanlaisen käyttöliittymän omien tietojen hakemiseen. Monien terveyspalvelujen verkkosivuilta voi painaa sinistä nappia, jolloin tunnistautumisen jälkeen omat tiedot saa tiedostona.³⁸

Tiedon käyttäjien kannalta hyödyllisin tapa saada tietoa on ajantasaisten tietojen hankkiminen joustavien rajapintojen kautta. Silti joidenkin tietovarantojen kohdalla helpoin ja ensimmäinen ratkaisu voi olla tietojen tarjoaminen ladattavina otoksina. Green Buttonin kaltaiset ratkaisut ovat myös hyvä keino toteuttaa tietosuojasetuksen edellyttämä ihmisten pääsy sähköiseen kopioon tiedoistaan, ilman että rekisterinpitäjän tarvitsee vastata jokaiseen tietopyyntöön erikseen. Julkishallinto voisi siis tukea yhteisten työkalujen, käytäntöjen sekä käyttöliittymien kehittämistä, jotta "nappeja" voisi toteuttaa pienillä kustannuksilla eri organisaatioissa. Ranskassa ollaankin kehitetty ns. Rainbow Button -

³⁸ Green- ja Blue Button -hankkeet ovat osa Yhdysvaltojen MyData-aloitteita. Järjestelmien myöhempiin versioihin on kehitetty tiedoston latausta monipuolisempiakin toimintoja, kuten esimerkiksi sovellusten kytkentää jatkuvia tiedonsiirtoa varten. Lisätietoja varten ks. <https://obamawhitehouse.archives.gov/blog/2016/03/15/my-data-empowering-all-americans-personal-data-access>

hanketta, jonka piirissä pyritään Euroopan tasolla yhteiseen ohjeistukseen oman tiedon lataamisen toimintamalleista.³⁹

MyData-tilit ja dynaaminen sopiminen

Tällä hetkellä ihmisten käytännön keinot vaikuttaa heistä kerätyn tiedon käyttöön ovat rajatut. Yksittäisten palvelujen kohdalla käyttäjillä on mahdollisuus hyväksyä tai olla hyväksymättä käyttöehdot. Kutsumme MyData-tileiksi palveluita, jotka tarjoavat yksilöille kattavan näkymän heidän antamiinsa käyttö lupiin ja mahdollisuuden muuttaa niitä. MyData-tilit mahdollistaisivat siten ns. dynaamisen sopimisen, joka laajentaa valinnan mahdollisuuksia. Sen piirissä ihmisen suostumus tai sopimuksen hyväksyminen ei perustu pysyvään valintaan yhtenä hetkenä, vaan omaa suostumustaan voi muuttaa ajan kuluessa. MyData-tilin kautta käyttö lupia voisi kumota myöhemmin ja niiden yksityiskohtia voisi muuttaa.⁴⁰ Dynaamista sopimista tukevia palveluita ei toistaiseksi ole monia.

Ottaessaan sovelluksen tai palvelun käyttöönsä käyttäjät voisivat liittää sovelluksen MyData-tiliinsä. Ennen liittämistä käyttäjä tunnistautuu ja näkee palvelukuvauksen, joka selittää, kuinka käyttäjän tietoa tullaan käyttämään ja mahdollisesti tarjoaa valintoja tämän suhteen. Ihanteellisesti tämän prosessi ei olisi sidottu verkkopalveluihin vaan se toimisi yhtä hyvin esimerkiksi toimipisteessä asioitaessa. Liittymisen jälkeen käyttäjä on antanut sovelluksille luvan lukea tietoa hänestä rajapinnan kautta, ilman ylimääräistä käyttäjän osallistumista.

Ohjelmoitava sopiminen

Tässä yhteydessä kutsumme yksityisyysasetuksiksi kaikkia oman tiedon jakamista koskevia valintoja. Yksinkertaisessa mallissa henkilötiedon jakamisesta käyttäjä tekee ja muuttaa asetuksiaan ottaessaan palveluita käyttöön tai myöhemmin kirjautumalla MyData-tiliinsä. On olemassa myös teknisiä ratkaisuja, joilla tiettyjen sovellusten tai palveluiden yksityisyysasetuksia pystyisi muokkaamaan muista sovelluksista käsin rajapinnan kautta.⁴¹ Tulevaisuudessa tällainen ohjelmoitava sopiminen voisi tehdä järjestelmästä käyttäjäystävällisemmän.

Ohjelmoitavan sopimisen kanssa käyttö lupia voisi hallinnoida usean käyttö liittymän kautta, ei pelkästään MyData-tilin kautta. Vaihtoehtoisten hallintasovellusten kautta voisi samalla järjestää tiedon jakaminen paitsi julkishallinnon rekistereistä myös muista lähteistä, jos nämä tukevat samoja standardeja. Yksi etu useassa käyttö liittymässä olisi myös sovellusten saavutettavuuden varmistamisen helpottuminen, eli se, että yksityisyyden hallintaan voisi kehittää keinoja myös esimerkiksi heikosti näkeville ihmisille.

3.3 Käytännön keinot tiedon hallintaan

Seuraavassa esittelemme kaksi käytännön keinoa, joiden avulla tietojen käyttöä yksittäisen palvelun sisällä voisi seurata ja hallita. Tällaisten keinojen tarjoaminen on luokittelussamme neljännen tason edellytys.

³⁹ MesInfos-järjestön ja sen kumppaneiden Rainbow Button -hankkeesta, ks. <http://mesinfos.fing.org/rainbowbutton/>

⁴⁰ Tällaisen suostumuksen luonnetta koskien ks. (Whitley, 2009/8). Parhaimmillaan dynaamisessa sopimisessa voi myös hallinnoida tiedon jakamista hyvinkin yksityiskohtaisella tasolla. Joissakin palveluissa on esimerkiksi luotu mahdollisuus valikoida kenttä- tai resurssikohtaisesti, mitä tietoja jaetaan, mihin käyttö tarkoitukseen ja kuinka pitkäksi aikaa. Esimerkkejä suostumuksen hallinnan mallintamisesta ovat mm. MyData:n referenssi-implemtaatio (<https://github.com/HILT/mydata-stack>) ja Kantara Initiativen Consent Receiptit. Erilaisten mallintamisen ratkaisujen vertailusta, ks. Stizftung Datenschutz, 2017. Käytännössä suostumuksen dynaaminen hallinta vaatisi teknologiaa ja standardeja, jotka eivät vielä ole laajasti käytössä, vaikka niitä on toteutettu useissa ns. henkilökohtaisissa datavarastoissa.

⁴¹ Esimerkiksi Kantaran Initiativen hallinnoima User-Managed Access -standardi tukee ohjelmoitavaa sopimista.

Lokit tietojen käytöstä

MyData-tili toteutuessaan keräisi lokitietoja annetuista käyttöluvista ja tiedon jakamisesta. Tämän tiedon tallettaminen voi helpottaa esimerkiksi kiistojen ratkaisemista. Lisäksi tilin yhteydessä voisi olla palvelu, joka kokoaa lokitietoja myös niistä sovelluksista, joihin tilistä jaetaan tietoa. Näiden lokitiedot voisivat kuvata esimerkiksi tiedon käyttöön liittyviä tapahtumia sekä tiedon mahdollisia tiedon luovutuksia kolmansille osapuolille. Myös ne voisi antaa MyData-tilin käyttäjien nähtäväksi.

Jos MyData-tilin toteuttaja tarjoaa alustan lokitietojen keräämiseen ja keskitetyn näkymän niihin, voisivat myös julkiset rekisterit käyttää sen toimintoja. Monet julkishallinnon rekisterit tallentavat lokitietoja sisältönsä käytöstä ja luovutuksista. Pääasiallisesti tähän tietoon on pääsy vain pyynnöstä, usein silloin, kun väärinkäytöstä on perusteltu epäily. Rekistereiden lokitietojen käytettävyyden ongelmana on myös se, että ne ovat hajautuneet moneen eri paikkaan. Yksi alusta, josta voisi tarkastella useisiin paikkoihin kerääntyneitä lokitietoja, voisi olla käyttäjille selkein. Yksittäisiä käyttötapauksia kuvaavien lokien jakamisen lisäksi viranomaiset voisivat myös julkaista tietotilinpäätösten tapaisia yleisiä selvityksiä tiedon käytön laajuudesta.

Lokitietojen keräämistä perustellaan usein väärinkäytösten estämisellä. Tietojen urkinta on vähemmän todennäköistä, kun mahdollisimman moni tarkistaa, että tietojen käyttö on oikeutettua. Lokitietojen saatavuuden parantamiseen on myös myönteisempiä perusteluja. Läpinäkyvyys voi lisätä yleistä luottamusta järjestelmään, ja toisaalta esimerkiksi tiedon tutkimuskäytön näyttäminen voisi osoittaa nykyjärjestelyn luomaa lisäarvoa. Kattavassa järjestelmässä, jossa yksilöt voisivat seurata, mitä tietoja heitä henkilökohtaisesti koskeviin päätöksiin on käytetty, tiedonkäytön läpinäkyvyys myös lisäisi ymmärrystä siitä, kuinka julkiset palvelut toimivat, ja helpottaisi virheiden korjausta.

Yksityisyysasetukset

Julkisten rekistereiden tietoja voidaan käyttää laissa määriteltyihin tarkoituksiin ilman suostumusta, mutta rekisteröidyllä on mahdollisuus pyytää, ettei tietoja luovutettaisi esimerkiksi suoramarkkinointiin. Myös valmisteilla olevaan lakiin sosiaali- ja terveydenhuollon asiakastietojen käsittelystä saattaa sisältyä periaate, jonka mukaan yksilöillä on mahdollisuus valita, käytetäänkö heidän terveystietojaan tutkimuksessa.⁴²

Tietojen käyttöön liittyvien valintojen helpottamiseksi olisi hyvä, jos valintoja pystyisi tekemään kattavasti yhdestä paikasta. Ensimmäinen askel järjestelmälle voisi olla, että se listaa saatavilla olevat valinnat, niille olennaisimmat rekisterit sekä keinot rajoittaa tiedon luovuttamista. Tulevaisuudessa ihmiset voisivat kenties saman järjestelmän kautta näyttää, hyväksyvätkö he henkilötietojensa käytön joukkoon yleishyödyllisiä tarkoituksia, jotka eivät välttämättä kuulu lain piiriin, kuten esimerkiksi kolmannen sektorin tekemiin tutkimuksiin.

3.4 Hyödyllisessä muodossa oleva tieto

Tietovarantojen kuvaus ja tehokas datakatalogi

Jos tiedon lähteitä ja tietosisältöjä kuvataan rikkaalla metatiedolla, on varantojen käyttöönotto sovelluksissa nykyistä helpompaa. Sovellukset voisivat esimerkiksi tarkistaa, mihin eri lähteisiin samankaltaista tietoa ihmisistä on tallennettuna. Jos ihminen käyttää

⁴² Lähde: <http://www.mediouutiset.fi/uutisarkisto/tietosuojalaki-tutisee-potilastietoja-voi-kohta-jakaa-laajalti-hyotykayttoon-6633798>

vaikka useaa kirjastoa, voisi näiden tiedot helposti yhdistää. MyData-tilien yhteyteen voisi kerätä datakatalogin, eli listan kaikista mahdollisista tai kunkin yksilön tiedon lähteistä.

Julkishallinnossa on kehitetty useita tietovarantojen kuvauksen malleja. Esimerkiksi Julkisen hallinnon tietohallinnon neuvottelukunnan (JUHTAn) JHS-jaosto on luomassa suositusta rekisteritiedon metatiedoista, joka valmistuu syksyllä 2017.⁴³ Myös Suomi.fi-palvelutietovarannossa voisi mahdollisesti viitata katalogiin datan lähteistä, ja esimerkiksi näyttää, mitä tiedon lähteitä palvelu pystyy käyttämään yksilön antamien lupien pohjalta. Tietovarantojen yhteensopivuuden varmistamiseksi tietomalleja ja sanastoa tulisi myös kehittää yhdessä eri toimijoiden kanssa, käyttäen tätä tarkoitusta varten luotua välineistöä.⁴⁴

Tietovarantojen kuvaukset ja datakatalogit voisivat sisältää seuraavia tietoja:

- Skeema, joka kuvaa tietovarannon sisältöä ja tietotyyppejä.
- Rekisteritietoja yhdistellään usein, joten metatietojen tulisi viitata niihin toisiin rekistereihin, joihin rekisteritietoja voidaan yhdistää yhteisten tunnistetietojen pohjalta.
- Tieto säilyy usein rekistereissä pitkään, joten metatiedon versiointi on hyödyllistä tietueisiin ja tiedon mallinnukseen liittyvien muutosten tunnistamiseksi.
- Koska samankaltaista tietoa on saatavilla useista eri rekistereistä, metatiedot voisivat viitata sanastoon ja muihin yhteentoimivuustyökaluihin. Näin tiedon käyttäjä voi selvittää eroja esimerkiksi tulotietojen määrittelyistä riippuen tiedon lähteistä.

Sähköiset allekirjoitukset ja tiedon lähteen varmentaminen

Rekistereistä saatu tieto on hyödyllisempää, jos se hyväksytään todisteeksi, koska sen katsotaan olevan luotettavasta lähteestä. Lähteen todistaminen on mahdollista sähköisten allekirjoitusten avulla. MyDatakin olisi siis erityisen hyödyllistä, jos sen saa allekirjoitettuna ja varmenteiden tarkistus on helppoa. Allekirjoituksen voisi joko hoitaa tiedon lähteestä, tiedon luovuttavan organisaation toimesta tai keskitetympin esimerkiksi usean rekisterin tiedon jakamista hoitavan rajapinnan kautta.

Sähköisesti allekirjoitetun tiedon laillinen asema voi riippua hallinnon alojen erityislainsäädännöstä. Yleiset sähköisiä allekirjoituksia koskevat säädökset ovat tiukkoja. Esimerkiksi Viron X-Road-järjestelmässä ollaan siirtymässä kohti laitteistolla (eikä ohjelmistoilla) luotuja sähköisiä varmenteita, jotta järjestelmä olisi lähempänä Euroopan Unionin direktiivien vaatimuksia.⁴⁵

Koneluettavan sähköisen allekirjoituksen lisäksi tietojen välityksen voi suunnitella niin, että se tekee lähteen varmentamisen helpoksi ihmisille, ilman että nämä tarvitsevat monia ylimääräisiä työkaluja. Yksi keino tähän voisi olla mahdollisuus luoda väliaikaisia linkkejä, jotka viittaavat julkishallinnon yleiseen tietoa välittävään palveluun (kuten alla kuvattuun MyData-operaattoriin). Käyttäjät voivat jakaa linkit tiedon vastaanottajan kanssa, joka näkee rekisterin sisällön selkeässä muodossa. Tiedon vastaanottajan selain myös tarkistaa, että tieto on peräisin julkishallinnon yhteisestä järjestelmästä.⁴⁶

⁴³ JHS-hanke Rekisteritiedon metatiedot, <http://www.jhs-suositukset.fi/web/quest/jhs/projects/rekisteritiedon-metatiedot>

⁴⁴ Esimerkiksi YTI-kärkihankkeen yhteydessä ollaan kehittämässä kansallista koodistopalvelua ja tietomallieditoria, ks. <https://wiki.julkict.fi/julkict/yti/>

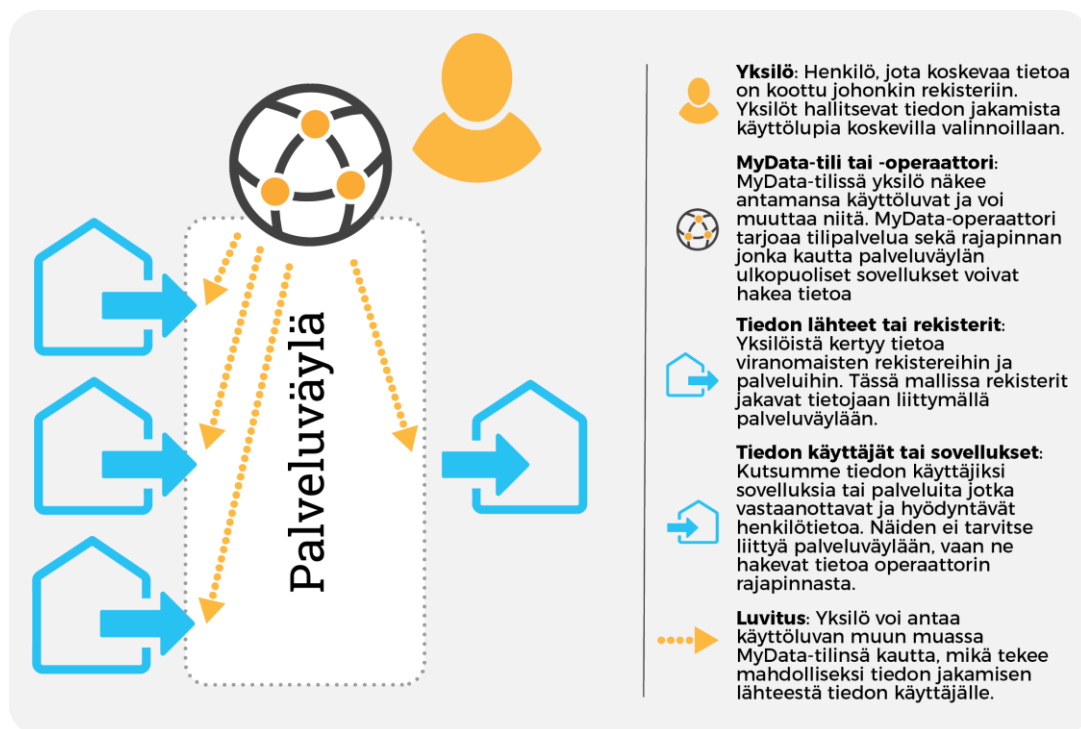
⁴⁵ Ansper, Buldas, Freudenthal, & Willemson, 2013

⁴⁶ Yksinkertaisin ratkaisu tätä varten olisi käyttää SSL-varmenteita, joita yleisesti käytössä olevat selaimet tukevat.

3.5 Palveluiden toteuttamisen arkkitehtuurit

Luvun loppuksi kuvaamme erilaisia arkkitehtuureja eli tapoja järjestää tiedon tallettaminen ja tietopalvelujen toteuttaminen. Jotkin malleista perustuvat keskitettyihin, yhteisiin ratkaisuihin, kun taas toiset toimivat hajautetummalla tavalla. Esitetyistä arkkitehtuurista etenkin hajautetuimmat edellyttävät yhteisiä standardeja ja toimintamalleja, jotka ovat toistaiseksi vielä kehittymässä. Arkkitehtuurit eivät välttämättä sulje toisiaan pois, vaan ne voisivat kehittyä samanaikaisesti tai toisiaan seuraten. Ensimmäisiä ratkaisuja kokeiltaessa kannattaa siis pitää mielessä myös vaihtoehtoiset tavat järjestää henkilötiedon hyödyntämisen verkostoja sekä niiden mahdolliset edut.

Kuva 3.2. Suomi.fi-palveluväylä ja yhteinen operaattori

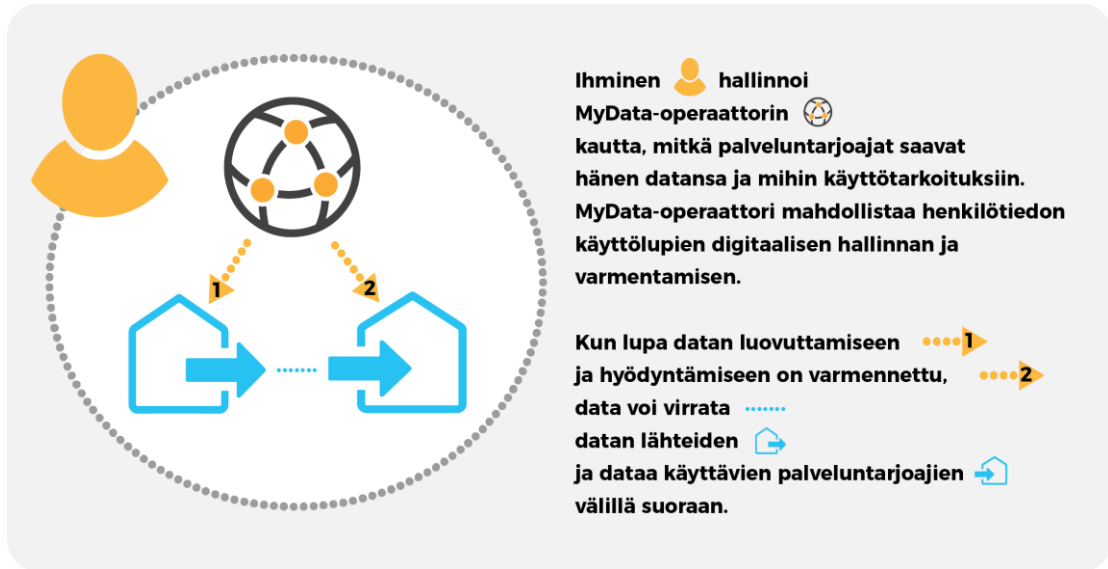


Palveluväylään perustuva yhteinen MyData-operaattori

Suomi.fi-palveluväylä tarjoaa vakioitun tavan tietojen jakamiseen organisaatioiden välillä. Lisäksi julkishallinnon organisaatioilla on velvollisuus kytkeä tärkeimmät tietovarantonsa palveluväylään. Monien rekisterinpitäjien kannalta helpoin tapa toteuttaa luvitettu tiedon jakaminen olisi siis toteuttaa palveluväylään yhteinen MyData-periaatteiden pohjalta toimiva ulostulo. Tällaista yhteistä palvelua kutsumme tässä yhteydessä MyData-operaattoriksi. Koska rekisterinpitäjät toteuttavat joka tapauksessa yhteentoimivuuden palveluväylän kanssa, tämä arkkitehtuuri aiheuttaisi vähemmän ylimääräisiä kustannuksia julkisille rekisterinpitäjille.

Sovellukset voisivat rekisteröityä operaattorin palveluun, joka tarkistaisi sovellusten asiallisuuden jollakin kevyellä tavalla. Myös palveluväylään kytkeytyneet rekisterit sopivat operaattorin kanssa tietonsa välittämisestä ja siihen liittyvistä rajoista. MyData-operaattori täydentäisi (eikä korvaisi) rekisterinpitäjien muita rajapintoja tai luvanvaraisia tietosuoritteita. Palveluväylään pohjaavan operaattorin kautta sovellukset voisivat mahdollisesti käyttää suurta määrää erilaisia tiedon lähteitä yhtenäisen rajapinnan sekä yhtenäisten käyttöehtojen kautta.

Kuva 3.3. Pohjoismainen malli henkilötiedon hallinnasta



Datan jakaminen pohjoismaisessa mallissa

Aikaisemmissa MyDataan liittyvissä kirjoituksissa on ehdotettu, että monien yhteentoimivien operaattoreiden mallia voisi kutsua pohjoismaiseksi malliksi henkilötiedon hallintaan.⁴⁷ Tähän malliin liittyy ensinnäkin arkkitehtuuri, jossa tieto on tallennettuna hajautetusti lukuisiin eri palveluihin ja organisaatioihin. Nämä eri rekisterinpitäjät ovat sopineet yhteisistä standardeista mm. tunnistautumiseen ja käyttölupien esittämiseen. Tämä mahdollistaa sen, että käyttäjä voi MyData-operaattorinsa kautta hallita omia yksityisyysasetuksiaan. Samalla tieto siirtyy suoraan tiedon lähteiden ja hyödyntäjien välillä kulkematta operaattorin kautta. Käyttäjän antaessa luvan operaattori avaa teknisen mahdollisuuden tiedon luovuttamiseen. Samalla operaattorille sekä tiedon siirron osapuolille jää tietosuojalainsäädännön mukainen merkintä annetusta suostumuksesta.

Lisäksi pohjoismaiseen malliin kuuluu ajatus mahdollisuudesta vaihtaa operaattoreita. Vertauskuvana toimivat teleoperaattorit, jotka mahdollistavat viestinnän verkon sisällä, niin että yhden operaattorin kautta saa myös yhteyden muiden operaattoreiden asiakkaisiin. Samaan verkkoon voi liittyä usean operaattorin kautta. Samaan tapaan keskenään yhteensopivat MyData-operaattorit olisivat väylä verkostoon, jonka piirissä henkilötietoa jaetaan luotettavalla tavalla. Halutessaan ihmiset voisivat alkaa hoitamaan asioitaan toisen operaattorin kautta, ja ihanteellisesti viedä käyttölupia koskevat asetukset mukanaan.

Pohjoismainen malli vastaa sikäli Suomen julkishallinnon tietojärjestelmien arkkitehtuuria, että siinäkin tavoitteena on säilyttää tietoa useissa paikallisissa tietojärjestelmissä, sen sijaan että valikoituja perustietovarantoja koottaisiin kattavasti yhteen järjestelmään.⁴⁸ Lähestymistavan etuna on mm. useiden tietovarantojen saatavuus ja tiedon ajantasaisuus. Pohjoismainen malli pyrkii myös muuttamaan asiakassuhteiden laatua, niin että kynnys palveluiden käyttöönottoon helpottuu ja yritysten välillä on helppo vaihtaa. Onnistuessaan tämä lisäisi palveluntarjoajiin kohdistuvaa luottamusta ja kannustimia uusien palvelujen luomiseen.

⁴⁷ Poikola & Honko, 2015

⁴⁸ Esimerkiksi Pertiva, 2016 vertailee Suomen lähestymistapaa Tanskan järjestelmään.

Hajautetut tietojärjestelmät

Hajautetut tilikirjat (distributed ledger) ovat viime aikoina kehittynyt uusi tiedon tallettamisen tapa. Tilikirjat ovat tietokantoja, jotka sisältävät ajan myötä kasvavia listoja tapahtumista. Lohkoketjut (blockchain) ovat yksi tämän teknologian muoto. Hajautetut järjestelmät mahdollistavat tiedon tallentamisen sellaisella tavalla, että monien ihmisten on mahdollista nähdä kaikki tietokantaan tallennetut tapahtumat sekä itse lisätä niihin tietoa sovittujen sääntöjen puitteissa. Tietokanta on kopioitu verkoston sisällä useisiin eri paikkoihin, joilla kullakin on kopio kaikista tallennetuista tapahtumista. Kun tieto on kerran tallennettu, sitä on vaikea poistaa kenenkään toimesta. Etenkin tämä viimeinen ominaisuus vahvistaa luottamusta tietokannan sisältöön.

Hajautetut järjestelmät voivat myös tukea MyDatan edellyttämän infrastruktuurin syntymistä tarjoamalla uusia malleja yhteiseen toimintaan useiden toimijoiden välillä. Hajautettuihin tilikirjoihin voisi esimerkiksi tallentaa merkintöjä siitä, millaisia henkilötietoja henkilö on jakanut muille toimijoille ja millaisilla ehdoilla. Sopimuksilla olisi sähköiset allekirjoitukset ja ne olisi tallennettu tavalla, joka turvaa ne peukaloinnilta. Lohkoketjujen avulla voisi myös hallita ihmisten identiteettiä koskevaa tietoa sekä tunnistautumista. Sovrin on yksi hajautetuista järjestelmistä, jotka pyrkivät tarjoamaan keinoja tunnistautumiseen ja jotka olisivat yksilöiden itsensä hallussa ja hallinnassa. Sen ehdottamassa viitekehyksessä ihmiset voivat liittää omaan profiiliinsa ominaisuuksia useilta identiteetin tarjoajilta, ja tunnistautua eri palveluihin jakamalla näitä tietoja valintojensa mukaan.⁴⁹

Lohkoketjujen tapaiselle arkkitehtuurille on siis kehitteillä useita erilaisia sovelluksia. Tässä vaiheessa on vielä epäselvää, missä tarkoituksissa tämä järjestämisen tapa erottuu edukseen. Hajautetut järjestelmät tarjoavat kuitenkin yhden uuden ja lupaavan reitin kehittää yhteisiä käytäntöjä ja standardeja, joita datan entistä sujuvampi siirrettävyys vaatii. Sellaisenaan ne voivat myös tukea pohjoismaisen mallin toteutumista. Onnistuessaan hajautettujen tilikirjojen tapaiset järjestelmät voisivat toteuttaa joitakin henkilötiedon jakamiseen liittyviä palveluita, jotka muuten edellyttäisivät erikoistuneita palvelimia. Esimerkiksi tiedon luvittaminen voisi tapahtua suoraan tiedon lähteen ja sen käyttäjän vuorovaikutuksen pohjalta silloin, kun käyttäjän identiteetti ja sopimus tiedonsiirrosta löytyvät hajautetusta tilikirjasta. Tällöin MyData-operaattoria ei välttämättä tarvittaisi erikoistuneen tietoa välittävän palvelimen muodossa.⁵⁰

⁴⁹ <https://sovrin.org/>

⁵⁰ Silti esimerkiksi Sovrin-järjestelmässä on sovelluksia, joita kutsutaan agenteiksi. Näiden tarkoitus on toimia palvelimien tavoin ja tarjota pysyvä osoite, johon käyttäjän identiteettiin liittyvää viestintää voi kohdistaa.

4 MYDATAN TALOUDELLISET VAIKUTUKSET JA RAHOITUSMALLIT

Heikki Sundquist - Aleksi Knuutila - Vesa Kokkonen (Kustannusten arviointi)

Tässä osiossa käymme läpi MyDataan taloudellisia vaikutuksia, erityisesti MyDataan liittyvien infrastruktuuripalveluiden kehittämiseen ja toimintaan liittyviä kustannuksia sekä järjestelmän vaikutuksia virastojen tuloihin. Luvussa tarkastelumme kohteena on tilanne, jossa julkishallinto kehittää MyData-operaattorin kaltaisen palvelun henkilötiedon jakamiseen. Tämän operaattorin oletetaan täydentävän virastojen olemassaolevia rajapintoja ja tietotuotteita, niiden korvaamisen sijaan.

Ymmärtääksemme paremmin tietoon liittyvää kysyntää kuvaamme ensin datan hyödyntämiseen liittyviä arvoketjuja ja tietotuotteiden maksullisuuden vaikutusta niiden käyttöön. Hinnoittelun laillisia perusteita ja virastojen nykyisiä tuloja käsittelevien osioiden jälkeen käymme läpi mahdolliset rahoitusmallit MyDataan liittyvän toiminnan kustannusten kattamiseksi.

4.1 Palveluntarjoajien arvoketjut ja henkilötiedon kysyntä

Henkilötiedon hyödyntämisestä kiinnostuneiden joukko on mahdollisesti laaja. Jo olemassa olevat palveluntarjoajat voivat sen avulla kehittää tuotteitaan entistä sujuvammiksi, ja pienet startup-yritykset voivat luoda kokonaan uusia sovelluksia. Kolmannen sektorin toimijat voivat kerätä tietoa esimerkiksi jäsenistään, ja joitakin esimerkkejä on myös henkilötiedon keräämisestä tutkimuskäyttöön yksilöiden kautta.⁵¹ Myös julkiset palvelut voivat mahdollisesti kokeilla uusia toimintatapoja, jos henkilötiedon hyödyntäminen helpottuu. Kasvavassa määrin ihmiset voivat myös käsitellä omaa tietoaan itse, esimerkiksi terveystietoa lukevien sovellusten avulla.

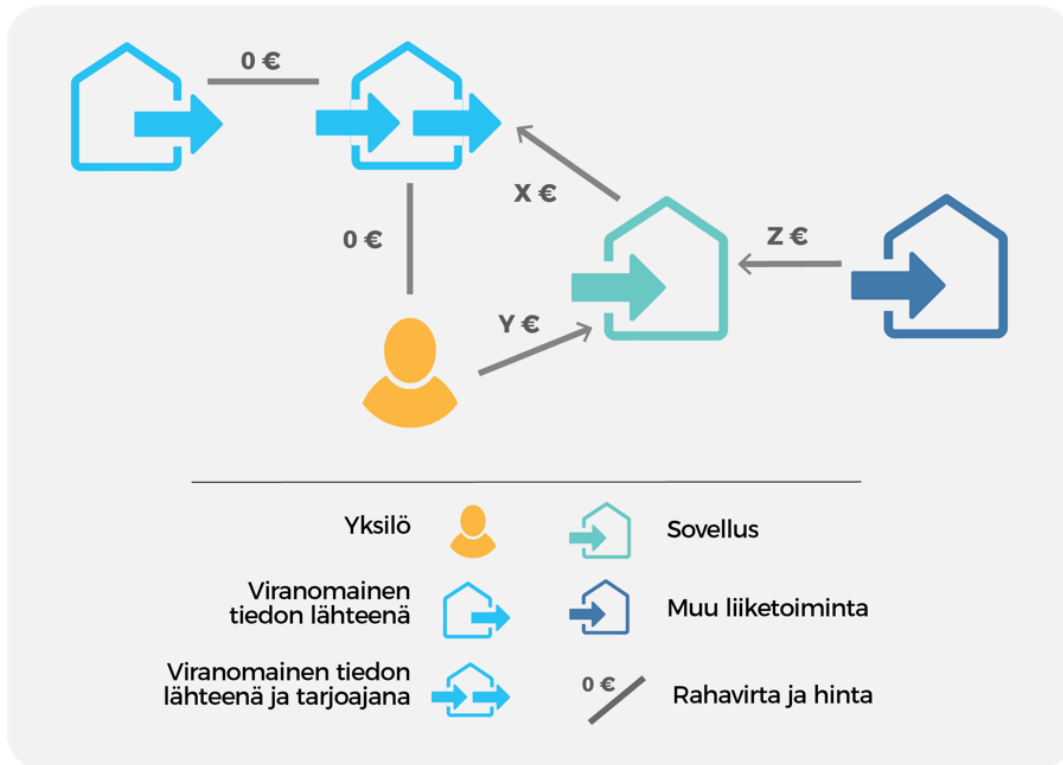
Kuvassa 4.1. esitetään yksinkertaistettu arvoketju, johon sisältyy julkishallinnon henkilötietoa käyttävä yksityinen palveluntarjoaja. Tämä luku käsittelee muun muassa viranomaisten ja palveluntarjoajan välisen vaihdon maksullisuutta. Huomattavaa on, että palveluntarjoaja voi saada myös korvauksen toiminnastaan paitsi tuloina asiakkailta myös kolmansilta osapuolilta tai muun oman liiketoimintansa edistämisen kautta.

Kokonaiskustannukset ja datan käytettävyys

Data ei ole niukka hyödyke, eli sen käyttö yhdessä yhteydessä ei sulje muita käyttötapoja pois. Näin ollen datan hyödyntäminen kannattaa järjestää niin, että mahdollisimman monet käyttötavat ja -tarkoitukset ovat mahdollisia ja tuettuja. Keinot henkilötiedon tehokkaaseen siirtämiseen tukevat myös muita tietopolitiikan tavoitteita, kuten päällekkäisten rekistereiden välttämistä ja palveluntarjoajien tallentaman datan määrän minimointia.

⁵¹ Tutkimuskäyttöä suunnitellaan esimerkiksi Decode-hankkeen yhteyteen. Lähde: <https://capssi.eu/data-sovereignty-for-the-sharing-economy-decode-project-kickoff/>

Kuva 4.1. MyDatan hyödyntämisen arvoketjut



Yksinkertainen MyDatan hyödyntämisen arvoketju. Tiedon siirto on pääsääntöisesti maksutonta sekä viranomaisten välillä että yksilöiden heidän omaa tietoaan. Tiedon käyttäminen yksityisissä sovelluksissa voi olla maksullista. Sovellusten tuottajat voivat saada tuloja paitsi palveluita käyttäviltä yksilöiltä myös muusta liiketoiminnasta.

Olemassa olevat keinot hakea tietoa julkisista rekistereistä ilman yksilöiden antamia käyttöluvia ovat useissa tapauksissa toimivia tiedon käyttäjille. Joissakin tapauksissa MyDatan käyttö saattaa vähentää toimijoiden kokonaiskustannuksia. Riippuen siitä, kuinka MyData-operaattorin tapaiset palvelut toteutetaan, niihin liittyminen voi olla nopeaa ja kustannuksiltaan kevyttä. Lisäksi operaattori voisi vähentää monimutkaisuutta tarjoamalla yhden rajapintojen kokonaisuuden, jonka kautta tietoa voi hakea useasta lähteestä.

Etenkin vakiintuneet yritykset arvioivat datan käyttöä sen kokonaiskustannusten ja järjestelyn ennustettavuuden mukaan. Kokonaiskustannuksiin kuuluu paitsi maksu tietopalvelusta, myös aika ja vaiva, joka kuluu tietopalveluun liittymiseen. Datan käyttämiseen liittyy myös kehityskustannuksia, jotka riippuvat niiden julkaisun laadusta ja saatavilla olevasta tuesta. Vakiintuneiden palveluntarjoajien näkökulmasta ei siis haetun henkilötiedon hinnalla välttämättä ole suurtakaan merkitystä. Tärkeintä ovat selkeät sopimuskäytännöt ja rajapintojen korkea palvelutaso. Näin on erityisesti, koska tiedonhauk viranomaisilta ovat yleisesti varsin edullisiksi hinnoiteltuja. Kaikista tiedoista kustannusvastaavuus oli kokonaisuudessaan vuonna 2015 94 %.⁵²

Toisaalta erilaisia toimintamalleja kokeileville yrityksille ero halvan tietoaaineiston ja sen maksuttomuuden välillä voi olla suuri. Silloin kun halutaan joustavasti kokeilla erilaisia datan lähteitä, voivat maksujärjestelyt, olla merkittävä kynnys. Sama pätee myös henkilötiedon käyttöön esimerkiksi kolmannella sektorilla tai tutkimustoiminnassa.

⁵² Valtiokonttori, 2016, s. 4

MyDatan ja muiden tietotuotteiden suhde

Tiedon maksuttomuuden taloudellisista seurauksista on käyty pitkään keskustelua avoimen datan kohdalla. Viranomaisten kannalta yksi avoimen datan haasteista oli, että se saattoi merkitä joidenkin maksullisten tietotuotteiden muuttumista maksuttomiksi, ja siten muutosta virastojen tuloissa. Toisaalta datan julkaiseminen avoimena edellyttää vain rajoitetun määrän lisäkustannuksia niissä tapauksissa, joissa tietovarannot ovat jo valmiiksi digitaalisessa muodossa ja niiden käsittelyn käytännöt ovat hyvässä mallissa.

MyDataan liittyvät taloudelliset haasteet ovat jossain määrin erilaisia. Tässä luvussa lähtökohtamme on, että MyData täydentäisi viranomaisten olemassa olevia rajapintoja ja tietotuotteita. Tällaisena se ei ole täydellinen korvike tietojen luovutuksille muilla perusteilla ja järjestelyillä. Toisaalta MyDatan toteutuminen edellyttää uudenlaisten infrastruktuuripalvelujen kehittämistä esimerkiksi tiedon käyttölupien hallintaan. Näihin liittyvien investointikustannusten kattaminen on keskeinen taloudellinen haaste.

Monet viranomaisten nykyisistä tietotuotteista ovat saatavilla määritelyihin käyttötarkoituksiin tietolupien hoitamisen jälkeen. Koska käytännöt tiedon hankkimiseen yksilön antamien lupien pohjalta ovat erilaiset, arvioimme siis, että MyDatan saatavuus vaikuttaa muiden tietotuotteiden käyttöön vain maltillisesti. Osa viranomaisten tiedon luovutuksista sisältää myös esimerkiksi massahakuja, jolloin vastaavan tiedon hankkiminen yksittäisten lupien kautta ei olisi käytännössä mahdollista.

Seuraavissa tapauksissa MyDatan käyttö voisi vaikuttaa olemassa olevien tietotuotteiden kysyntään:

- Monista viranomaisten rajapinnoista tehdään yksilöitä koskevia kyselyitä tilanteissa, joissa kyseinen yksilö on palveluntarjoajan kanssa asiakassuhteessa tai jollain muulla tavalla asioinnin alullepanijana. Tällaisissa tilanteissa olisi mahdollista, että käyttäjä voisi antaa luvan MyDatan käyttöön sen sijaan, että tiedot haettaisiin muilla keinoilla (kuten tietolupiin perustuvien rajapintojen kautta). Yksilöiden antamiin lupiin pohjaavat käytännöt todennäköisesti korvaavat muita tiedon haun tapoja vain silloin, kun ne ovat sujuvampia ja kustannuksiltaan pienempiä tai ne avaavat uusia mahdollisuuksia palveluntarjontaan laajemman tiedon käsittelyn pohjalta. On myös kuviteltavissa, että uudet palvelut käyttävät useita tiedonhaun tapoja, jolloin eri väylät henkilötietoon tosiasiaa olisivat toisiaan täydentäviä sen sijaan, että ne sulkisivat toisensa pois.
- Joissakin tapauksissa ihmisille myydään heitä itseään koskevaa tietoa esimerkiksi virallisten otteiden muodossa. Näiden otteiden arvo niiden tietosisällön lisäksi on usein siinä, että ne ovat luotettuja tai lainmukaisia todisteita. Erilaiset MyData-palvelut voivat korvata tätä myyntiä siinä määrin kuin niiden kautta siirrettyä, mahdollisesti sähköisesti varmennettua tietoa pidetään riittävän luotettavana.

4.2 Investointikustannukset ja juoksevat kustannukset

MyData-periaatteen mukaiseen tietojärjestelmään siirtyminen aiheuttaisi sekä operatiivisia- että investointikustannuksia. Lopulliset kustannukset riippuvat siitä, missä muodossa MyData-ratkaisuihin siirrytään.

Operatiiviset kustannukset

MyDataan liittyviä operatiivisia kustannuksia syntyisi siinäkin tapauksessa, että henkilölle tarjottaisiin häntä koskeva data aina ilmaiseksi tai henkilön luvalla kolmas taho voisi saada sen ilmaiseksi. Datan maksuton saatavuus lisäisi sen kysyntää, mikä osaltaan kuormittaisi tietojärjestelmää ja vaatisi lisää resursseja ylläpitoon ja neuvontaan. Nämä operatiiviset kustannukset tulisivat nykyisistä maksutuloista luopumista seuraavien tulonmenetysten päälle. MyData tuskin kasvattaa kyselyjen määrää ainakaan aluksi merkittävästi. Dataliikenne on kuitenkin ollut kasvussa jo ilman uudistuksiakin luoden kustannuspaineita.

Keskeinen operatiivinen kustannus syntyy dataoperaattorista. Dataoperaattorin lopullinen kustannus riippuu siitä, millaisia tehtäviä sille lopulta jää. Oleellista kuitenkin on se, että operaattorilla tulisi olemaan vakituista henkilökuntaa ja infrastruktuurin ylläpitoon liittyviä kuluja. Hyvän vertailukohdan kustannuksille antaa Hyvis-ICT Oy, joka vastaa Taltioni-terveystilin ylläpidosta. Vuonna 2016 Hyvis-ICT:llä oli 28 työntekijää ja se teki 2,3 miljoonan euron liikevaihdolla 908 000 euron tappiot⁵³. Kokonaiskustannukset olivat siis n. 3,2 miljoonaa euroa. Tämä vastaa jokseenkin vuotuisia operatiivisia kustannuksia. Taltioni ei vastaa täysin MyData-operaattoria, mutta toimintaperiaatteissa on paljon samaa. Oleellista kuitenkin on se, että toimiva MyData-operaattori edellyttää miljoonaluokan kustannuksia toimiakseen. Tämä ei välttämättä jäisi kokonaan julkishallinnon rasiitteeksi, jos operaattorille rakennetaan oma ansaintamalli. Tämä voi tarkoittaa esimerkiksi sitä, että korkeamman palvelutason tuotteet olisivat maksullisia tietoa hyödyntäville yrityksille.

Pienempiä juoksevia kustannuksia syntynee tunnistaumisesta lupien antamisen yhteydessä sekä tiedon käyttäjien valvonnasta. Tunnistaumisesta aiheutuu jo nykyisin kustannuksia, mutta luvitus tulee kasvattamaan tämän tarvetta.

Taulukko 4.1. Operatiivisia kustannuksia

Ominaisuus/ kustannustekijä	Laajuus	Nykytila	Merkitys MyDatalle	Kustannusarvio/ Arvio kustannus- vaikutuksesta
Lisääntynyt dataliikenne	Koskee kaikkia MyData-tietovarantoja	Dataliikenne aiheuttaa resurssipulaa jo nyt	Edellytys tiedon saatavuudelle	Maltillinen, riippuu lopullisesta hinnoitteluratkaisusta
Tunnistauminen luvan antamisen yhteydessä	Koskee kaikkia MyData-tietovarantoja	Luvitusta tässä muodossa ei ole	Luvitus on edellytys MyDatan käytön toteutumiselle	Alle euro per tunnistauminen
Tiedonkäytön ja tietoturvan valvonta	Koskee kaikkia tiedostotyyppejä	Lokit olemassa, mutta ei avoimia yksilöille. Tietoturvaa valvotaan tietolupien yhteydessä	Lisää luottamusta, jossain laajuudessa välttämätön	
Operaattori	Yhdistää eri tietolähteitä	Ei olemassa	Laajan käytettävyyden kannalta oleellinen	Miljoonaluokkaa

Operatiiviset kustannukset ja kuvaus niiden tarpeen laajuudesta, kyseisen kustannustekijän nykytilanteesta, merkityksestä MyDatan kannalta ja arvio kustannuksen suuruudesta.

⁵³ <https://www.finder.fi/IT-konsultointia+IT-palveluja/HYVIS-ICT+Oy/Kotka/yhteystiedot/3053241>

Investointikustannukset

Investointikustannusten määrittäminen on huomattavasti haastavampaa. Kustannusten määrittäminen tulee ajankohtaiseksi siinä vaiheessa, kun on esillä selkeitä vaihtoehtoisia suunnitelmia. Tämän takia arviointi tehdään skaalautuvasti niin, että voidaan arvioida eri toteutuksien kustannuksia. Suppealla tasolla muutokset voidaan toteuttaa osana esimerkiksi Suomi.fi- ja Kanta-hankkeita, niin että ihmiset saavat käyttöön omat datansa. Vastaavasti MyData-näkökulmasta Omakantapalvelua tulisi kehittää niin, että data olisi koneluettavassa muodossa siirrettävissä toiseen järjestelmään.

Täydellinen MyData-arkkitehtuuri voidaan toteuttaa lukemattomilla eri tavoilla huomioiden MyDatan eri tasot. Tässä yhteydessä nostetaan esiin muutamia kustannuksia, joita MyData voi aiheuttaa. Kustannukset eivät ole täysin erotettavissa yleisistä tietojärjestelmäkustannuksista. Merkittävä esimerkki tästä on datan yhteensopivuus. Tämä on merkittävä haaste tietojohdalliselle yleisesti, mutta myös MyDatalle. Keskeiset investointikustannukset syntyvät erityisesti siitä, miten tiedon saatavuus varmennetaan. Tämä edellyttää datakatalogia eli kuvausta siitä, mitä tietoa ja millä tavalla on saatavissa. Tämän jälkeen täytyy päättää tekniset ratkaisut siitä, miten data annetaan käytettäväksi. Tämä voidaan tehdä luomalla API:t eri tietolähteille. Tämän kaltainen hajautettu järjestelmä lienee toimivampi ratkaisu kuin mikään keskitetty tietovaranto. Hajautettu ratkaisu sopii myös paremmin MyData-hengen pohjaksi, koska tavoitteena ei ole keskittää henkilötietoa millekään tietylle taholle.

Jos ihmisille luotaisiin mahdollisuus seurata sitä, miten heidän tietojaan käytetään, pitäisi julkisten tietovarantojen lokijärjestelmät avata myös kansalaisten käyttöön. Kattavan lokijärjestelmän rakentamisen kustannukset olisivat merkittävät ja toisaalta siihen liittyy myös hallinnollisesti kaksi ongelmaa. Toinen syntyy siitä, että on tilanteita, joissa ihmisen ei kuulukaan saada tietoa viranomaisten toiminnasta. Tämä voi tarkoittaa esimerkiksi poliisin esitutkintaa tai vaikka valmisteilla olevaa lapsen huostaanottoa. Tämä haaste voidaan hallita esimerkiksi niin, että lokitietojen saatavuus on joko viivästytetty tai tiettyjen viranomaisten toiminnan synnyttämät merkinnät eivät olisi vapaasti selailtavissa. Toinen haaste liittyy siihen, että lokimerkinnät ovat jonkin toisen henkilön MyDataa – eli yksityisyyden suoja vaarantuu, jos lokitiedoissa merkinnät näkyvät henkilötasolla. Tämä ongelma on kuitenkin helposti kierrettävissä, jos lokimerkinnät syntyisivät tai näkyisivät vain tietoja katsoneen organisaation tasolla – eli lokitiedoissa lukisi KELA, Kannelmäen terveyskeskus jne. Tarkempien lokitietojen avaaminen edellyttäisi perusteltua epäilyä tietojen väärinkäytöstä.

Väestörekisterikeskukseen rakennettava lokijärjestelmä maksaisi haastattelujen pohjalta arviolta ½ - 1 miljoonaa euroa. Tämä ei kuitenkaan tarkoittaisi vielä kattavaa lokijärjestelmää, koska tällä hetkellä henkilötietoa välittyy useissa eri järjestelmissä. Näin ollen kattavan lokitietojärjestelmän luominen edellyttäisi sitä, että kaikki reitit henkilötiedon katseluun kartoitettaisiin, niin että näistä syntyvät lokitiedot voitaisiin näyttää kootusti tai linkitetysti.

Taulukko 4.2. Investointikustannukset

Ominaisuus	Laajuus	Nykytila	Merkitys MyDatalle	Kustannukset
Lupien antaminen tietojen käyttöön muille osapuolille/ muiden yksityisyyden tietojen hallinta	Kaikki henkilötiedot	Toteutus osana valtuutusrekisteriä	Mahdollistaa tiedon käytön luvittamisen	Osa YTI:n toteutusbudjettia
Tietojen todentaminen/ sähköinen allekirjoitus/	Viranomaisten välittämät henkilötiedot	Perustuu laajasti paperisiin, maksullisiin todistuksiin	Datan arvo riippuu sen luotettavuudesta	Vaihtelee ⁵⁴
Lokitiedostot	Kapan kautta liikkuva data	Saatavissa, mutta ei helposti	Mahdollistaa henkilön valvomaan, miten tietoja käytetään	Alle 1 milj. euroa
API:t tietojen saatavuuden parantamiseksi	Tarkemmin määritettävät tietovarannot/ rekisterit	Ei olemassa	Mahdollistaa datan käytettävyyden sovelluksissa	Huomattava kustannus – vaatii myös ylläpitoa
Pilotointi	Kaikki avattavat tietovarannot/ rekisterit	Ensimmäiset pilotit käynnissä	Haetaan käyttökohteita sekä testataan rajapintojen yms. toimivuutta	Voidaan toteuttaa eri laajuisena. Kustannukset osittain palveluntuottajien maksettavina
Rekistereille merkittävien henkilötietojen liittäminen palveluväylään	Kaikki henkilötietorekisterit	Kehitteillä KaPa-hankkeessa	Mahdollisuus toteuttaa MyDatan saatavuus	Osana meneillään olevia kehityshankkeita
Yhteentoimivuus Kanta/ Koski kanssa	Eri järjestelmien yhteen tuominen	Kannan osalta rakennetaan yhteyttä KaPaan ja mm. yhteiseen palvelunäkymään	Luo mahdollisuuden eri tietojen yhdistämiselle	
Datakatalogi	Mahdollisimman kattavasti kaikki tietovarannot	Koostuu useista erillistä listauksista, ei keskitettyä katalogia	Tiedon hyödynnettävyys edellyttää riittävää metatietoa.	

Investointikustannukset ja kuvaus niiden tarpeen laajuudesta, kyseisen kustannustekijän nykytilanteesta, merkityksestä MyDatan kannalta ja arvio kustannuksen suuruudesta.

Kokonaiskustannukset nousevat merkittäviksi, jos MyData pyritään toteuttamaan laajalla rintamalla yhdellä kertaa. Käytännössä toimivampi ratkaisu lienee se, että MyData-periaatteet otetaan mukaan julkisten tiedonhallintahankkeiden suunnitteluun. Tämä tarkoittaa sitä, että tiedon uudelleenkäytettävyys, lupien antaminen jne. otetaan huomioon aina, kun järjestelmiä kehitetään. Jatkuvalle suunnittelulle voidaan pitää huoli siitä, etteivät kustannukset nouse kohtuuttoman suuriksi samalla kun tavoitellut hyödyt saavutetaan. Merkittävää on myös se, että MyDatalla on yhtymäkohtia moniin muihin tietohallinnon tavoitteisiin, kuten datan uudelleenkäytettävyyteen ja yhdisteltävyyteen. Näillä voidaan hakea tehokkuutta julkishallintoon. Näin ollen mahdolliset kustannukset eivät liity pelkästään MyDataan.

Näiden kustannusten jälkeen jää vielä tiettyjä, määrittelemättömiä alueita. Esimerkiksi osa ihmisten tiedoista voi olla arkaluontoisia ja tietyt tietueet saattavat sisältää viittauksia muihin ihmisiin. Esimerkiksi väestötietojärjestelmässä on tietoa mm. alaikäisten lasten

⁵⁴ Sähköisen allekirjoituksen voi toteuttaa eri teknologioilla. Yksi tapa on käyttää PKI (public key infrastructure) -teknologiaa. Tätä on käytetty mm. Kanta-hankkeessa. Haasteena tälle on se, että myös vastaanottavan tahon tietojärjestelmän on tuettava tätä. Distributed ledgerin kaltaiset ratkaisut tarjoavat oman vaihtoehdon, mutta teknologian uutuuden takia referenssihintaa on vaikea löytää.

sosiaaliturvatunnuksista, minkä ei pitäisi olla vapaasti siirrettävää tietoa. Vastaavasti dokumenttimuotoisessa (esim. KELA:n päätökset) tiedossa saattaa olla viitteitä muihin ihmisiin. Tällä hetkellä tämän tiedon automaattinen suodattaminen ei ole mahdollista.

4.3 Tietosuoritteiden hinnoittelun lailliset perusteet

Viranomaisten tiedon luovuttamiseen liittyvien palvelujen maksullisuudesta ja maksujen suuruuden yleistä perusteista säädetään lailla. Maksuperustelakia sovelletaan, ellei muussa erityislaissa tai laissa olevien valtuuksien nojalla ole annettu siitä poikkeavia säännöksiä. Erityislakeja on sitemmin säädetty monien eri virastojen toimialoille, kuten esimerkiksi Patentti- ja rekisterihallituksen tai Väestörekisterikeskuksen kohdalla.

Maksuperustelain periaatteiden mukaan maksut voidaan jakaa neljään ryhmään:

- Julkisoikeudelliset suoritteet, joista perittävän maksun tulee vastata suoritteen tuottamisen kokonaiskustannuksia (omakustannusarvo).
- Muut kuin julkisoikeudelliset suoritteet, joita kutsutaan markkinasuoritteiksi tai liiketaloudellisesti hinnoitelluiksi suoritteiksi.
- Monopolisuoritteet, joiden tuottamiseen on valtiolla monopoliasema ja joiden hinta voidaan määrätä omakustannusarvon mukaiseksi.
- Maksuttomat suoritteet, jotka ovat siis maksuttomia, ellei maksullisuudelle ole erityistä syytä.⁵⁵

Kokonaisuudessaan julkistaloudellisten suoritteiden kustannusvastaavuus oli vuonna 2015 94 %.⁵⁶ Maksullisten suoritteiden tuotot ovat kasvaneet 2007 vuoden 1,25 miljardista eurosta vuoden 2015 1,5 miljardiin euroon. Samaan aikaan liiketaloudellisesti laskutettujen tuottojen osuus pieneni noin 700 miljoonasta 360 miljoonaan. Pienentyminen johtuu siitä, että valtion toimintoja on siirretty budjettitalouden ulkopuolelle liikelaitostamisen ja yhtiöittämisen myötä, jolloin viranomaisten liiketaloudellinen laskutus on vähentynyt.⁵⁷

Yksityishenkilön omien henkilötietojen saanti viranomaisilta on pääsääntöisesti maksutonta. Mikäli tietojen luovutukseen liittyy muita palveluita kuten esimerkiksi tiedon oikeaksi todistamista tai muita toimenpiteitä, joiden toteuttamisen kustannukset ovat todennettavissa, niistä voidaan periä valtion maksuperustelain mukainen korvaus.

Vuonna 2015 valtiovarainministeriön taustamuistiossa selvitettiin tietojen luovutusten maksullisuutta. Selvityksessä tarkasteltiin MML:n, Trafín, PRH:n, VRK:n, Veroviranomaisten, THL:n ja Kelan toimintaa.⁵⁸ Raportin mukaan kaikissa näissä vakiomuotoiset tietoluovutukset kunnan ja valtion viranomaisille olivat jo maksuttomia yleisen maksupoliittisen linjauksen mukaisesti.

⁵⁵ Valtion maksuperustelaki, 21.2.1992/150

⁵⁶ Valtiokonttori, 2016, s. 4

⁵⁷ Männikkö, 2015

⁵⁸ Kauhanen-Simanen, 2013

4.4 Virastojen tietosuoritteista saamat tulot

Tässä osiossa käsitellään Väestörekisterikeskuksen, Maanmittauslaitoksen, Patentti- ja rekisterihallituksen sekä Trafín tietosuoritteista saamia tuloja. Nämä valittiin, koska niiden katsottiin olevan merkittävimmät henkilötiedon myynnistä tuloja saavat valtiotason viranomaiset. Tarkastelun ulkopuolelle jätettiin markkinasuoritteista suuria tuloja saavia rekisterinpitäjiä (raja-arvo 2,9 miljoonaa euroa), joilla on merkittävästi salassa pidettäviä tietoja tai sitten laskutusta tiedoista, joihin ei sisälly henkilötietoja.⁵⁹

Tarkastelussa on pyritty arvioimaan itse tietoluovutuksista saatavia maksuja jättämällä pois palvelumaksuja siltä osin kuin se on ollut mahdollista. Silloin kun se on tietojen pohjalta mahdollista, pyrimme erittelemään tuloista ne, jotka keskittyvät yksilötason henkilötietoon. Nämä tiedon luovutukset ovat sisältönsä puolesta samankaltaisia kuin MyData-operaattorin kaltaisten palvelujen kautta saatava tieto.

Väestörekisterikeskus

Taulukko 4.3. VRK:n maksulliset tiedonsiirrot ja niiden tulot

TIETOPALVELUSUORITTEET, VRK	Tulot 2016	Milj. € 2015	Kasvuprosentti 2015-16	Kustannusvastaavuus 2016
Julkiset palvelut yhteensä	1,15	0,97	19 %	106 %
Julkiset kyselypalvelut	0,15	0,13	12 %	106 %
Julkinen muutostietopalvelu	0,20	0,21	-4 %	86 %
Julkiset muut eräkäyttöiset palvelut	0,81	0,63	29 %	112 %
Liiketaloudelliset palvelut yhteensä	9,30	8,41	11 %	301 %
Kyselypalvelut	5,41	4,60	17 %	313 %
Asiakasrekisterin päivitys	1,90	1,53	24 %	313 %
Poimintapalvelut	1,91	2,19	-13 %	277 %
Nimi ja osoitepalvelut	0,09	0,09	4 %	127 %

Lähde: Väestörekisterikeskuksen tilinpäätös 2016

Ylläolevassa taulukossa on kerätty tiedot maksullisista tiedonsiirroista, joihin ei sisälly merkittävästi palvelua. Tiedot on kerätty VRK:n tietotilinpäätöksen tiivistelmästä 2015⁶⁰ ja tilinpäätöksestä 2016⁶¹. Pääosa henkilötietojen hakujen tuloista saadaan palveluyrityksiltä, joko palvelun suorakäytöllä (58%) tai eräkäyttöhouilla (42%). Näiden lisäksi VRK laskuttaa merkittävästi varmennepalveluista (8,5 miljoonaa euroa), jotka jätetään tämän tarkastelun ulkopuolelle, koska niihin sisältyy henkilötiedon siirron lisäksi merkittävässä määrin muita palveluita.

Suorakäytöksi luotujen tietohakujen kohdalla haetaan yksilötason tietoa. Suorakäyttö on siis tiedon sisällön puolesta lähimpänä tietoa, jota voisi hakea MyData-operaattorista. Lisäksi

⁵⁹ Näitä olivat: Poliisihallitus, Puolustusvoimat, Puolustushallinnon rakennuslaitos, Tilastokeskus, Luonnonvarakeskus, Liikennevirasto, Ilmatieteen laitos, Geologinen tutkimuskeskus, Suomenlinnan hoitokunta, Opetushallitus ja Ympäristökeskus. (Lähde: (Valtiokonttori, 2016))

⁶⁰ Väestörekisterikeskus, 2016

⁶¹ Väestörekisterikeskus, 2017

suorahakujen käyttötarkoituksia voi tulkita niin, että yksilö on aina alullepanija sille toiminnalle, jonka osana tietoa käsitellään.

Yritysten tekemien suorakäyttöjen kasvu on ollut edellisestä vuodesta 18 % ja koko henkilötietojen hakujen tulot ovat kasvaneet 12 %. Yksityishenkilöiden tekemien hakujen tulot ovat olleet huomattavasti pienempiä ja niiden kasvukin on ollut vain noin 3 %. Kasvu ja kokonaisvolyymi ovat siten toteutuneet erityisesti valtion rekisterinpitäjien ulkopuolisten palveluntarjoajien kautta tehtävissä hauissa.

Maanmittauslaitos

Maanmittauslaitoksen (MML) rekistereihin tallennetaan kiinteistöjen tunnistetietoja sekä niiden omistuksiin ja muutoksiin liittyviä tietoja. Lisäksi virasto hallinnoi lainhuutorekisteriä sekä rekisteriä kiinteistökaupoista ja kauppahinnoista.

MML:ssä ei pääasiassa tuoteta palveluja loppukäyttäjille vaan rajapintoja, joiden pohjalta muut tahot tuottavat lisäarvopalveluja. Palveluntarjoajille liiketoiminnalliseen käyttöön on tarjolla kiinteistötietopalvelu. Palvelussa voi selata valtakunnallisen kiinteistötietojärjestelmän (KTJ) tietoja. Tietoja voi hakea myös kiinteistökaupoista, rakennuksista sekä kiintopisteistä. Palveluntarjoajia ovat mm. pankit ja kiinteistönvälittäjät, joilla on keskeinen rooli loppuasiakkaiden kaupanteon palveluissa ja rahoituskysymyksissä.

Julkisoikeudellisten tietopalvelujen rajapintahakujen määrä kasvoi tavoitteen mukaisesti 3 %. KTJ:n tietolupia myönnettiin yhteensä 626 käyttäjälle, mikä on 17 % vähemmän kuin edellisenä vuonna. Tästä voi päätellä, että lähes kaikilla kiinteistötietojen lainmukaisen käyttötarpeen täyttävillä asiakkaila alkaa olla mahdollisuus käyttää sähköisiä palveluja.⁶²

Tietopalveluiden tuotot vuonna 2016 olivat 9,8 miljoonaa euroa. Pääosa tuloista syntyy myymällä otteita, joita voidaan käyttää virallisina todisteina. Ne toimitetaan joko paperimuodossa tai PDF:nä. Otteiden käyttöön liittyvä palvelukokonaisuus on asiakkaalle monimutkainen. Asiakas joutuu ensin hakemaan erilaisia todistuksia erikseen ja sitten liittämään ne varsinaiseen asiaa koskevaan hakemukseen. Toisaalta yksittäisten kansalaisten kiinteistöjen omistukseen tehtävät muutokset hoidetaan laajasti kiinteistönvälityspalvelujen kautta, jolloin loppuasiakkaan kokema palvelurajapinta voidaan toteuttaa asiakaslähtöisesti.

Liikenteen turvallisuusvirasto, Trafi

Trafin rekistereissä on tieliikenteeseen, meriliikenteeseen, ilmailuun sekä rautatieliikenteeseen liittyviä tietoja. Kaikissa on sekä ajoneuvotietoja että henkilötietoja. Yksityisten ihmisten kannalta merkittävimmät tiedot koskevat tieliikennettä. Rekisterissä on 5 miljoonaa henkilöön liittyviä tietoja. Tieliikenteen tietojen hyödyntäjiä ovat viranomaiset, sopimusrekisteröijät (katsastuskonttorit, vakuutusyhtiöt, autoliikkeet, romuttajat) sekä erilaiset sopimuskumppanit (tietopalvelujen tuottajat). Toiseksi yleisin rekisteri liittyy merenkulkuun, joka käsittää 170 000 henkilön tietoja sekä 203 tuhannen vesikulkuneuvon tietoja.

Vuonna 2016 Trafin liiketaloudelliset tuotot eli tuotot tietopalveluiden hinnastosuoritteiden myynnistä olivat 11,4 miljoonaa euroa. Kasvua edellisestä vuodesta oli 19 %. Tuottojen kustannusvastaavuus oli 210 %. Merkittävä osuus tuloista liittyy ns. suorakyselyihin ja sovelluskyselyihin, joita vuonna 2016 tehtiin noin 11 miljoonaa ja 27 miljoonaa kappaletta.

⁶² Maanmittauslaitos, 2017

Suorakyselyillä haetaan yksittäisen yksilön tietoja, kun taas sovelluskyselyt voivat sisältää myös laajempia hakuja.⁶³

Patentti- ja rekisterihallitus

Patentti- ja rekisterihallitus (PRH) ylläpitää laaja-alaisesti rekistereitä kahdella isolla tulosalueella. Se ylläpitää ensinnäkin patenttirekisteriä, tavaramerkkirekisteriä ja mallirekisteriä. Lisäksi se hallinnoi yrityksiä ja yhteisöjä kuvaavaa tietoa, jota on muun muassa kauppa-, yhdistys- ja säätiörekisterissä. Lisäksi PRH:n toimialueelle kuuluu tilintarkastusvalvonta vuoden 2016 alusta ja siihen liittyvä tilintarkastusvalvontarekisteri. Henkilötietoa kuuluu etenkin yrityksiä ja yhteisöjä kuvaaviin rekistereihin.

PRH:n hallintoimien rekistereiden osalta omien tietojen tarkistaminen on keskitetty Suomi.fi:n Kansalaisen palvelunäkymän yhteyteen. Yritysten tilaa voidaan tarkastella maksutta YTJ.fi-palvelusta. Merkittävä osa PRH:n palveluista toteutetaan yksityisten palveluyritysten kautta, joita kutsutaan tietojen jakeluyrityksiksi. Kaupparekisterin suhteen suurimmat jakelijat ovat Asiakastieto Oy ja Kauppalehti. Niiden palvelut perustuvat erilaisten viranomaistietojen sekä muista tietolähteistä saatavien tietojen yhdistelyyn. Patenttiasioiden suhteen erilaiset patenttiasiamiesyhtiöt ovat tietojen tuojia ja käyttäjiä yksityisten henkilöiden tai yritysten palvelijoina. Yritysten tilinpitoa toteuttavat tilitoimistot hakevat ja syöttävät yritysten tietoja rekistereihin.

PRH:n julkisoikeudellisten palvelujen tuotot olivat vuonna 2015 46 miljoonaa euroa ja markkinasuoritteiset 3,6 miljoonaa euroa.⁶⁴ Tulot jälkimmäisestä luokasta koostuvat pääosin Virre-tietopalvelujen käytöstä. Virre-tietopalvelussa on tietoja kaupparekisteristä, säätiörekisteristä ja yritys kiinnitysrekisteristä. Perustiedot ovat maksuttomia, mutta rekisteriotteet, todistukset ja muut asiakirjat sekä yritysten laajemmat tiedot ovat maksullisia. Erityisesti on huomioitava, että tietojen tuottaminen ja muuttaminen esimerkiksi kaupparekisteriin on maksullista.

Kustannusvastaavuus kummassakin luokassa oli 94 %. Tuloja ei ole eritelty tiedonsiirtojen ja muiden palvelujen suhteen. On huomioitava, että markkinasuoritteisten tulojen kustannusvastaavuus esimerkiksi VRK:n ja Trafín osalta on ollut yli 200 % siinä missä PRH:n vastaava luku jää merkittävän alhaiseksi.

Yhteenveto

Näiden virastojen osalta yhteenvetona markkinasuoritteisista tuloista saadaan alla oleva taulukko. On huomioitava, että luvut eivät ole tarkkoja tiedonluovutusten suhteen, sillä eri virastojen luokittelu markkinasuoritteisten tulojen suhteen ei ole yhdenmukainen ja toisaalta toisiin tuloihin sisältyy myös palveluja tiedonluovutusten yhteydessä.

⁶³ Trafi, 2017a

⁶⁴ Patentti- ja Rekisterihallitus, 2017

Taulukko 4.4. Virastojen markkinasuoritteiset tulot ja niiden kustannusvastaavuusprosentit

Virasto	Tulot/ M€	Kustannusvastaavuus/%
VRK	9,3	301
Trafi	11,4	210
MML	2,9	105
PRH	3,6	94

Lähde: Virastojen tilinpäätökset vuodelta 2016

On oletettavissa, että pelkkien tiedonluovutusten suhteen volyymit kasvavat merkittävästi lähivuosina yksityisten palveluntarjoajien toteuttaessa yhä suuremman osuuden palvelujen arvoketjusta. Mitä se sitten toisaalta merkitsee talouden suhteen, riippuu täysin valtion maksupolitiikasta ja sen kehittymisestä lähivuosina. Keskeinen kysymys on periaatepäätös yksityisen palveluntarjoajan tiedonsaannin maksullisuudesta tai maksuttomuudesta henkilön suostumuksen perusteella saatavan tiedon suhteen.

4.5 Vaihtoehtoja luvitetun henkilötiedon hinnoittelulle

Esityksessä lähdetään siitä, että omien tietojen hakeminen on henkilöille maksutonta. Lisäksi viranomaisten keskinäinen tiedonsiirto on myös maksutonta. Osiossa käymme läpi erilaisia ansaintamalleja MyData-palveluihin liittyvien kustannusten kattamiseksi. Yksi vaihtoehto on, että yksityiset palveluntarjoajat maksaisivat tiedon hauista joissakin tapauksissa. Itse tiedon haun tuotteistamisen lisäksi viranomaisten on myös mahdollista myydä tiedon käyttämiseen liittyviä tukipalveluita.

Tässä osiossa kuvaamme ensin lyhyesti tulonjakoon liittyviä kysymyksiä. Esitämme sitten kattavasti kaikki merkittävät vaihtoehdot henkilötiedon jakamisen hinnoittelusta ja muista siihen liittyvistä palveluista. Vaihtoehtojen kuvailun jälkeen käymme lyhyesti läpi niiden seurauksia. Vaihtoehtoisia hinnoitteluperiaatteita kuvatessa olemme käyttäneet vaihtoehtojen nimeämiseen kirjaimia (2A, 2B, 2C..). Tehdäksemme kuvauksesta yksinkertaisemman, puhumme tässä yhteydestä MyData-operaattorin kaltaisesta palvelusta, johon palveluntarjoajat voisivat liittyä voidakseen hyödyntää henkilötietoa. Kuten kappaleessa 4 totesimme, operaattoreita voi tosiasiallisesti olla useita ja niiden toteuttamiseen on useita vaihtoehtoisia malleja.

Tulojen jakautumisesta ja tavoitteellisesta tulotasosta

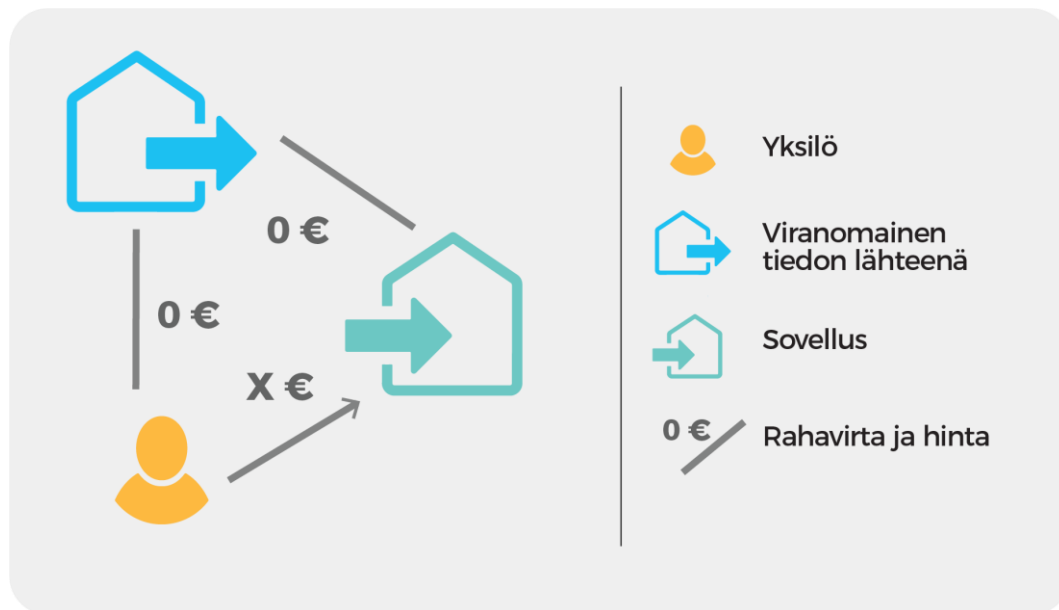
Tietojen hakeminen MyData-operaattorin kautta aiheuttaa kustannuksia usealle toimijalle. MyData-operaattorin lisäksi myös rekisterinpitäjä, joka on vastuussa tietovarannon keräämisestä ja tietojen luovuttamisesta, joutuu hallinnoimaan ja ylläpitämään omaa tietojärjestelmäänsä. Operaattori voisi tehdä tulonjakosopimuksen kaltaisen järjestelyn rekisterinpitäjien kanssa, jossa se sitoutuu vähintäänkin kattamaan rekisterinpitäjien ylimääräiset kustannukset. Silloin kun rekisterinpitäjät tarjoavat rajapintaa suorakyselyyn liiketoiminnallisen hinnoittelun pohjalta, voisi tämä olla vaihtoehtoinen pohja tulojen jakamiselle.

MyData-operaattorista vastaavien tahojen on syytä pyrkiä kattamaan palvelun tuloilla paitsi osuus juoksevista kustannuksistaan myös tarpeet jatkokehitykselle. MyData-operaattorin palvelut pitäisi hinnoitella niin, että muuttuvien kustannusten lisäksi ne kattavat myöhemmät investoinnit järjestelmän kehitykseen ainakin osittain.

Vaihtoehto 1: MyData-operaattorin käyttö on maksutonta

Ensimmäisessä vaihtoehdossa käyttäjät voivat käyttöluvien kautta tehdä omaa henkilötietoaan saatavaksi erilaisille sovelluksille. Käyttöluvien antamisen edellyttämä tunnistautuminen sekä tietojen jakaminen rajapinnasta tapahtuvat ilman maksuja käyttäjälle tai tiedot lataavan sovelluksen tarjoajalle. Näiltä osin skenaario vastaa aikaisempia kirjoituksia MyDatasta, joissa lähtökohtana on ollut tiedon avoin saatavuus ja siirrettävyys.

Kuva 4.2. Skenaarion 1 mukainen palveluprosessi



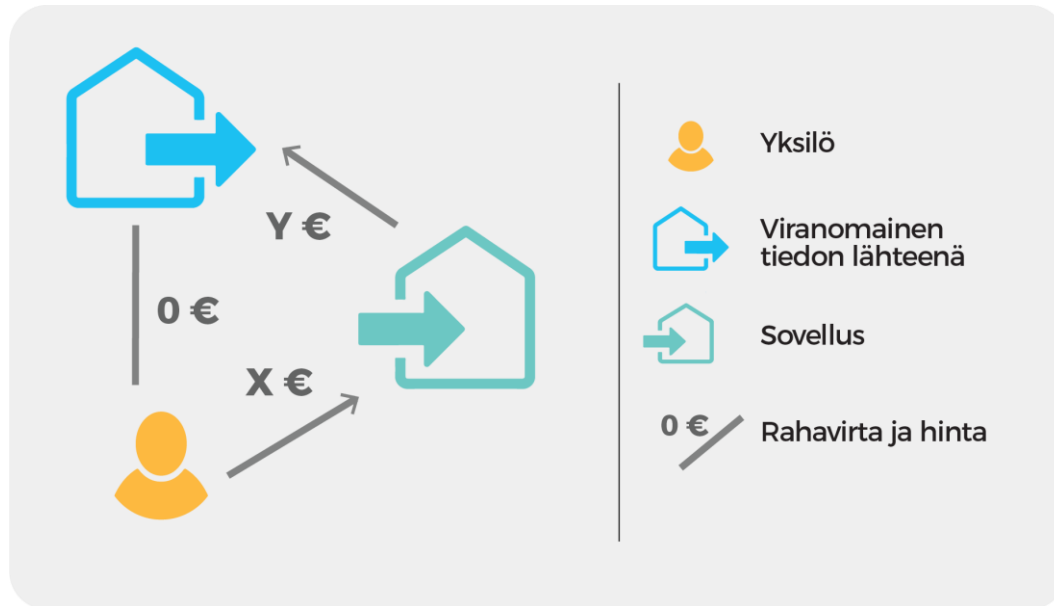
Ensimmäisessä skenaariossa henkilötieto on yksilön antaman käyttöluvan pohjalta saatavilla maksutta erilaisille sovelluksille. Sovellusten käyttö voi olla yksilölle joko maksullista tai maksutonta.

Sekä käyttöluviin liittyvään tunnistautumiseen että kyselyihin vastaamiseen liittyy luonnollisesti juoksevia kustannuksia. Lisäksi on mahdollista, että MyData-operaattorin kautta tapahtuvat kyselyt korvaavat joitakin virastojen henkilötietoihin liittyviä tietosuoritteita, jotka on hinnoiteltu kustannusten mukaisesti tai liiketaloudellisin perustein. Kustannusten kasvun lisäksi on siis mahdollista, että viranomaisten tulot laskevat. Yllä esitetyn arviomme mukaan muihin tietosuoritteisiin kohdistuvan kysynnän muutos olisi kuitenkin pieni.

Vaihtoehto 2: Tiedon hinnoittelu yksityisille palveluntarjoajille

Toisessa vaihtoehdossa MyData-operaattorin käytölle voi olla hinta niissä tapauksissa, joissa tietoa hakee yksityinen palveluntarjoaja jonkin yksilön antamien käyttöluvien perusteella. Hinnan määrittelyyn on useita erilaisia vaihtoehtoja, jotka käymme läpi. Oletuksena on edelleen, että yksilöiden pääsy omaan tietoonsa sekä muiden viranomaisten tietojen käyttö on maksutonta. Tämäkin vaikuttaa hinnoittelun seurauksiin, kuten selitämme alla. Maksujen periminen tiedon käytöstä sulkee pois joitakin tiedon mahdollisia käyttötapauksia.

Kuva 4.3. Skenaariossa 2 mukainen palveluprosessi



Toisessa skenaariossa henkilötieto on yksilön antaman käyttöluvan pohjalta saatavilla sovelluksille maksua vastaan. Yksilölle itselleen hänen oma tietonsa on saatavilla maksutta. Sovellusten käyttö voi olla yksilölle joko maksullista tai maksutonta.

Vaihtoehto 2A: Kaikki tiedonsiirto yksityisille palveluntarjoajille on maksullista

Tällä hetkellä henkilötietoa koskevien suorakyselyiden teko virastoista on pääasiassa maksullista ja hinnoiteltu liiketoiminnallisin perustein. Olisi mahdollista jatkaa tätä käytäntöä myös MyData-operaattorin osalta. Yksittäiset tiedon käyttäjät tekisivät operaattorin kanssa sopimuksen. Käytettävä henkilötieto olisi saatavilla joko yleisin hinnoitteluperusteiden tai operaattorin yksittäisten rekisterinpitäjien kanssa tekemien sopimusten perusteella. Palveluntarjoajille MyData-operaattorin käyttö voisi edelleen olla houkuttelevaa, esimerkiksi jos sen käyttöönotto on helpompaa kuin muiden rajapintojen tai jos se mahdollistaisi tiedon soveltamista (annettujen lupien pohjalta) käyttötarkoituksiin, jotka muut rajapinnat sulkevat pois.

Vaihtoehto 2B: Korkeamman palvelutason rajapinnat ovat maksullisia

Yksi yleinen datan tuottajien ansaintamalli on, että tietoa saa yksinkertaisemmassa muodossa ilmaiseksi. Korkeampaa palvelutasoa edellyttävät käyttäjät maksavat siitä. Jälkimmäisestä käyttötavasta saadut tuotot tukevat järjestelmän toimintaa kokonaisuutena.

Lisäpalvelut on tuotteistettu niin, että niistä maksetaan joko käytön mukaan tai tilauksena esimerkiksi kuukausi tai vuosi kerrallaan. Mahdollisiin lisäpalveluihin kuuluu:

- Useampien tiedon esitysmuotojen saatavuus (esimerkiksi muitakin talletusmuotoja tavallisen JSON-formaatin lisäksi)
- Mahdollisuus tehdä monipuolisempia kyselyitä tietokannasta (esimerkiksi mahdollisuus pyytää tietoa muodossa, jossa useampi yksilöä koskeva tietue on valmiiksi yhdistelty)
- Reaaliaikainen syöte muutoksista henkilöä koskevissa tiedoissa

- Useamman tiedon lähteen tai tyyppien yhdistäminen (esimerkiksi asuintietojen yhdistäminen kotialuetta koskevaan avoimeen dataan)
- Tarjottujen tietueiden rikastaminen lisätiedolla (esimerkiksi yksityiskohtia siitä, kuinka yksilöä koskeva tieto on kerätty, mikä voi helpottaa tiedon oikeellisuuden arvioinnissa)
- Joustavat keinot tiedon luotettavuuden ja ajankohtaisuuden tarkistamiseksi
- Palvelut, jotka tukevat ihmisten yksityisyyden säilyttämistä (esimerkiksi rajapinta, joka tarkistaa, onko ihminen täysi-ikäinen, mutta ei kerro tämän tarkkaa ikää)

Kehittämällä uusia rajapintatuotteita yhdessä tiedon käyttäjien kanssa syntyy myös tuotteita, jotka ovat hyödyllisiä palveluntarjoajille ja tukevat taloudellisen toiminnan kehittymistä. On myös mahdollista, että jokin yksityisten tahojen ryhmittymä tukisi tietyn rajapinnan rakentamista silloin kun sille on selkeä tarve. Toisaalta myös hyvin kehitettyjen avoimien rajapintojen pohjalta yksityinen sektori itse voi tuottaa niihin pohjaavia lisäarvopalveluita.

Tietojärjestelmiä voi rakentaa tavalla, joka helpottaa rajapintojen muotoilua joustavasti käyttäjän tilauksen ja tarpeen mukaan. Tämä mahdollisuus olisi hyvä ottaa huomioon järjestelmän teknistä toteutusta suunniteltaessa, jos lisäarvopalveluiden tuottaminen rajapintojen avulla nähdään tärkeäksi rahoituksen lähteeksi.

Vaihtoehto 2C: Hinnoittelu käytön määrän tai käyttömuodon mukaan

Useissa dataan liittyvissä tuotteissa on asetettu organisaatiokohtainen määrärajoitus, joka voi olla esimerkiksi sata hakua päivässä. Rajoituksen ylittävät tahot joutuvat maksamaan käytöstään joko tilauksen muodossa tai suhteessa käytön määrään. Määrärajoituksen etuna on, että se tekee järjestelmään kohdistuvan kuormituksen ennustettavammaksi ja voi jopa ehkäistä palvelunestohyökkäyksiä. Suurempia juoksevia kustannuksia aiheuttavat tahot tällöin myös tukevat järjestelmän toimintaa.

Toisinaan avointa dataa on myös tehty saatavaksi niin, että sen lisenssi riippuu käyttötarkoituksesta. Tietoja on saatavilla vain muuhun kuin kaupalliseen käyttöön, joka puolestaan edellyttäisi maksullisen sopimuksen tekemistä. Rajausten tekemisessä on lukuisia vaihtoehtoja. Aloitteleville yrityksille, joille lisenssointimaksut ovat hankalia, voisi myöntää poikkeuksia. Suomen julkishallinnossa joillakin tietoaineistoilla on ollut käyttäjäryhmä-kohtaisia rajoituksia, aineistot ovat esimerkiksi olleet maksuttomia vain tutkimuskäyttöön.⁶⁵ Monimutkaiset menettelyt hinnoittelussa luonnollisesti lisäävät kustannuksia ja epävarmuutta sekä tiedon käyttäjän että tarjoajan puolella.

⁶⁵ Poikola ym., 2010



Tulevaisuuden vaihtoehto: Käyttäjä tiedonvälittäjänä

Hyvin kehittyneessä henkilötiedon jakamisen järjestelmässä käyttäjä voisi hakea ja tallettaa omaa tietoaan itse hallinnoimaansa järjestelmään, esimerkiksi niin sanottuun henkilökohtaiseen datavarastoon (personal data storage). Parhaimmillaan tällaisen datavaraston käyttäminen tiedon lähteenä olisi sovelluksille yhtä helppoa kuin julkisen MyData-operaattorin.

Silti käyttäjän oman kopion käyttämiselle tiedon lähteenä on rajoituksensa. Riippuen käytetystä teknologiasta, tietoa ei välttämättä ole varmennettu eikä sen ajankohtaisuudesta ole takeita. Joihinkin käyttötarkoituksiin vähemmän luetettava tieto riittää hyvin, mutta joskus sovelluksen tarvitsee hakea ajankohtaisin ja varmennettu tieto suoraan viranomaiselta, MyData-operaattorin kautta.

Tällainen järjestely, jossa sovellus voisi saada tietoa joko käyttäjän kautta tai suoraan viranomaiselta, voisi olla toinen tapa hinnoitella kaksi eri palvelutasoa eri lailla. Käyttäjä voi hakea kopion tiedostaan ilmaiseksi ja jakaa sitä ilman kustannuksia. Tällöin myös julkishallinnon palvelimiin kohdistuu vähemmän rasitetta. Jos sovellus haluaa käydä tarkistamassa tuoreimman tiedon operaattorilta, tästä lisäpalvelusta voitaisiin periä maksu.

Vaihtoehto 3: Tukipalveluiden tarjoaminen

MyData-operaattorin käytön hinnoittelun lisäksi tuloja voi saada tukipalveluilla, jotka helpottaisivat tiedon käyttämistä. Tämä vaihtoehto ei siis sulje pois kahta ensimmäistä vaihtoehtoa vaan voi olla niitä täydentävä. Olisi esimerkiksi mahdollista yhdistää korkeamman palvelutason rajapinnat niihin liittyviin tukipalveluihin. Tukipalvelut voivat olla edellytys monimutkaisten rajapintoja käyttävien sovellusten kehittämiselle ja ne voivat lisätä rajapintojen käyttöä ylipäättään.

Mahdollisiin tukipalveluihin kuuluvat:

- Sovitut takuut tietopalvelujen saatavuudesta, mahdollisesti mukaan lukien myös rajapintojen aikaisempien versioiden jatkaminen
- Oikea-aikainen tuki rajapintojen käyttämiselle ja virheiden korjaamisen priorisointi
- Uusien sovellusten kehittäjien koulutus
- Ohjelmistojen kehityspakkausten (software development kit) lisensointi

Vaihtoehtoisten ansaintamallien vertailu

Avoimeen dataan liittyvissä tutkimuksissa tiedon maksuttomuus on vakuuttavasti näytetty kokonaisuuden kannalta edullisimmaksi vaihtoehdoksi.⁶⁶ Useissa tutkimuksissa myös henkilötiedon arvoa on pidetty kertaluokkaa avoimen tiedon varantoja suurempana.⁶⁷ Luonnollisesti kuitenkin tiedon keräämiseen ja yhteiskäyttöön sekä järjestelmien kehitykseen liittyvien kustannusten kattaminen on todellinen ongelma.

⁶⁶ Ks. esimerkiksi Stott, 2014, myös Valtiovarainministeriö, 2012 s. 53-60

⁶⁷ Schwab, Marcus, Oyola, Hoffman, & Luzi, 2011

On toistaiseksi epäselvää, millaisia käyttötarkoituksia julkiseen henkilötietoon liittyy, ja millaisissa palveluketjuissa yksilöiden lupien perusteella toimiminen erottuu edukseen. Varhaisiin innovaatioihin liittyviä mahdollisuuksia ei kannattaisi hidastaa monimutkaisilla lupa- tai hinnoittelukäytännöillä, vaan etenkin alkuvaiheessa MyData-operaattorin käytön tulisi olla maksutonta. Kokeiluvaiheessa, jossa suurin osa kustannuksista liittyy uusien tietojärjestelmien ja rajapintojen luomiseen, on mahdollista, että kehitystä johtava toimija voisi saada kohdennettua rahoitusta järjestelmän kehittämiseen. Samalla pitäisi kuitenkin varmistaa, että uusien palveluiden juoksevien kustannusten kattaminen sekä jatkokehityksen rahoitus ovat kestäväällä pohjalla.

Kun joitakin henkilötiedon varantoja on ensimmäisenä askeleena avattu käyttöön, voivat MyData-operaattorin kehittäjät myös pyrkiä ymmärtämään eri palvelumuotojen, toimialojen ja käyttötarkoitusten tarpeita. Toimiminen yhteistyössä yksityisten toimijoiden kanssa paitsi auttaa ymmärtämään, kuinka palvelujen kehitystä voidaan tukea, myös auttaa tunnistamaan mahdollisia lisäarvopalveluita, jotka voivat liittyä rajapintojen räätälöintiin ja muuhun tukeen. Kokeilujen kautta voidaan siis toimia asiakaslähtöisesti, mutta myös pyrkiä tilanteeseen, joissa merkittävimmät tiedon hyödyntäjät voisivat osallistua uuden infrastruktuurin kehityskustannuksien kattamiseen heille edullisten järjestelyjen kautta.

5 MYDATAN VAIKUTTAVUUS: KOLME TAPAUSTUTKIMUSTA

Vesa Kokkonen – Ossi Kuittinen

Älykkäisiin järjestelmiin perustuvat ratkaisut edellyttävät datan saatavuutta. Useat palveluntarjoajat pystyvät tuottamaan personoituja palveluita asiakkailleen tuntemalla näiden historian. Yksinkertaisia esimerkkejä tästä ovat mm. Netflixin ja Amazonin suositukset, jotka perustuvat käyttäjän omaan toimintahistoriaan. Näissä palveluissa yksittäiset organisaatiot käyttävät itse keräämäänsä henkilödataa omien tuotteidensa räätälöintiin.

Markkinoille on ilmestymässä myös palveluita, joissa hyödynnetään muiden toimijoiden tuottamaa henkilödataa. Esimerkiksi vakuutusyhtiö Turva tarjoaa liikennevakuutuksesta bonusta uusillekin asiakkaille. Tämä edellyttää, että henkilö antaa Turvalle luvan käyttää ajokortti- ja vakuutustietoja. Näin ollen Turva saa tietää hakijan henkilöhistorian ja voi tämän pohjalta myöntää tai olla myöntämättä bonusta. Järjestelmä on rakennettu niin, että Turva ei näe lähtödataa, vaan se analysoidaan koneellisesti ja tämän pohjalta tehdään päätös bonusten tasosta. Tämä on esimerkki uudesta palvelusta, joka on lähellä MyData-periaatteita. Oleellista on, että käyttöluvien kautta tieto on saatavilla myös uusiin palveluihin. MyDatan vaikuttavuudesta puhuttaessa kannattaa huomioida kiinnittää juuri tällaisiin uusiin palveluihin.

MyDatan hyödyntäminen yhteiskunnassa edellyttää uudenlaisia toimintamalleja ja uutta infrastruktuuria, joiden luominen aiheuttaa kustannuksia. Tapaustutkimuksissa selvitetään laajemmin MyDatan vaikuttavuutta talouteen ja yhteiskuntaan. Tapaustutkimukset on kuitenkin valittu huomioiden julkinen sektori eli pyritty nostamaan esiin vaikutuksia julkissektorin hallinnollisiin käytänteisiin. Yksittäisten virastojen tai rekistereiden näkökulmasta on riskinä, että MyData näyttäytyy etenkin taloudellisena rasitteena. Tämä johtuu mm. siitä, että maksuttomuuden korostaminen luo tarvetta joko uusille ansaintamalleille tai budjettirahoituksen kasvattamiselle. Tämän lisäksi tulevat vielä tietojärjestelmiin liittyvät investoinnit. Näiden panostusten vastapainona on saavutettavat hyödyt. Hallinnollisia muutoksia sekä tästä johtuvia kustannuksia arvioidaan väestötietojärjestelmän näkökulmasta.

MyData on tällä hetkellä kehityksen alkuvaiheessa. Tämän takia ollaan tilanteessa, jossa MyDataan pohjautuvia palveluita ja ratkaisuja ei vielä ole tai palvelut ovat pilotointivaiheessa. Vaikuttavuuden arvioiminen perustuu siis ennen kaikkea olemassa olevan trendin kuvaamiseen ja MyDatan merkityksen määrittämiseen trendin vauhdittamisessa. Tämä korostuu etenkin terveydenhuollossa, joka on tällä hetkellä yksi nopeimmin digitalisoituvia toimialoja. Liikenteessä digitalisaatio on myös kasvussa. Molemmilla toimialoilla merkittävänä kysymyksenä on se, miten ihminen – oli hän sitten asiakas, potilas tai kansalainen – saataisiin paremmin palveluiden keskiöön. Etenkin terveydenhuollossa kiinnostuksen kohteena on vaikuttavuuden parantaminen ja kustannusten hillitseminen. Näihin tavoitteisiin MyData tarjoaa kiinnostavan näkökulman.

5.1 Väestötietojärjestelmä esimerkkinä

Väestötietojärjestelmä (VTJ) on keskeinen hallinnollinen rekisteri. Sen sisältämä data muodostaa viranomaistoiminnan perustan silloin, kun viranomaisilla on tarve käsitellä henkilöön liittyviä tietoja. Datalla on kuitenkin myös muita kuin viranomaiskäyttötarkoituksia. Varsinainen viranomaiskäyttö rajataan tämän case-esimerkin ulkopuolelle lukuun ottamatta henkilön oikeutta tarkastella hänestä kerätyn tiedon käyttöä myös viranomaistoiminnassa. Tämän rajauksen lähtöoletuksena on se, että MyData ei oleellisesti tule muuttamaan viranomaisten oikeutta saada tietoa ja tarvittaessa myös niin, ettei henkilö itse saa tietää tästä. Periaatteessa MyData-lähestymistapa voi muuttaa viranomaisten tapaa saada ja hyödyntää tietoja. VTJ:n osalta tämä muutos on tuskin ensi alkuun kovin oleellinen. VTJ pysynee luotettavimpana henkilötietojen lähteenä ja VTJ:n tapoihin päivittää tietoja tuskin tulee suuria muutoksia. Esimerkiksi jo tällä hetkellä on mahdollisuus tehdä muuttoilmoitus myös maistraateille ja näin VTJ:hin Postin järjestelmän kautta. Vastaava tiedon jaettavuus saattaa toki laajentua muihinkin kanaviin.

VTJ:n käyttötarkoitukset MyDatan näkökulmasta

Väestötietorekisterin käyttötarkoitukset on määritelty Väestörekisterikeskuksen varmennepalveluissa annetussa laissa (661/2009) luvussa 4. Näitä käyttötarkoituksia ovat:

1. Yritysten tekemä yksilötiedon haku esimerkiksi henkilötietojen oikeellisuuden tarkistamiseksi. Tätä voidaan tarvita asiakkaan tunnistamisessa tai esimerkiksi tuotteen tai laskun toimittamisessa.
2. Yritysten asiakasrekisterien ylläpito. VTJ:n pohjalta tehdään massa-ajoja, joilla päivitetään automaattisesti eri yritysten asiakasrekisterit mm. yhteystietojen osalta, jos yritys tai muu organisaatio on saanut tähän luvan.
3. Otokset joukkopostituksia varten (esimerkiksi suoramarkkinointiin liittyen).
4. VTJ:n tietoja hyödynnetään myös pankki- ja vakuutustoiminnassa esimerkiksi asiakkaan tunnistamiseen, rahanpesun ja terrorismin rahoittamisen estämiseen sekä erilaisten verotukseen liittyvien velvollisuuksien todentamiseen. Tietoja käytetään myös perintätoiminnassa, yksityisessä terveydenhuollossa, isännöintitoimistojen taloyhtiöiden asukastietojen ylläpidossa sekä teleoperaattorien mobiilivarmenteiden luomisessa.
5. VTJ toimii lähdeaineistona maistraatin tuottamille virka- ja asuinpaikkatodistuksille.

MyDatan lähtökohdista eri käyttötapaukset avautuvat hieman eri tavoin. Suosituksia varten on myös punnittava, miten tiukat MyDatan määrittelyt sopivat yhteen asioiden sujuvuuden kanssa. Tähän liittyy mm. se, käytetäänkö tiedon käyttöilupien hallintaan opt in- vai opt out -ratkaisua. Yksilöiden antamien lupien kerääminen kaikkiin mahdollisiin käyttötapauksiin on työlästä. Jossain tapauksessa on sujuvaa, että esimerkiksi yritysten asiakasrekisterien massapäivityksiä voidaan tehdä, ellei tätä ole erikseen kielletty. Tällöin käytetään opt out -periaatetta.

Keskeinen haaste MyDatan näkökulmasta VTJ:n aineistoissa on, ettei VTJ:n sisältö ole vapaasti henkilöiden saatavissa ja siirrettävissä koneluettavassa muodossa. Teknisesti ratkaisu ei olisi monimutkainen, koska rekisteriaineisto on siirrettävissä olemassa oleviin koneluettaviin formaatteihin. Yritysten tietokyselyissä käytetään jo nyt koneluettavaa XML-

skeemaa. Käytännössä tietopyynnöt tapahtuvat VTJ-rajapinnan kautta vakimuotoisina kyselyinä. Kyselyjen sisältö on rajoitettu käyttötarkoituksen mukaan. Rajoite tietojen jakamiseen on siis ensisijaisesti hallinnollinen – eli kaikkea tietoa ei voi luovuttaa edes henkilön itse antamalla luvalla.

Jos tiedonsiirron formaattia tai väylää muutetaan, tämä vaatii kuitenkin tietojärjestelmään investoimista. Nämä muokkaukset eivät rajoitu pelkästään VRK:n järjestelmiin vaan myös tiedon hyödyntäjien järjestelmiä on muutettava, jos toiminnan logiikka muuttuu. Tämän lisäksi on ratkaistava tiedon oikeellisuuden varmistamiseen liittyvät haasteet. Jos ihminen itse välittää dataansa, on siitä käytävä ilmi datan lähde (sähköisen allekirjoituksen avulla) ja aikaleima. Tämä korostuu etenkin VTJ:n aineistoissa, koska niiden käytettävyydessä merkittävintä on, että VTJ tiedonlähteenä on luotettavampi kuin ihminen itse.

Toinen keskeinen tekijä VTJ:n käytölle on helppous ja ajantasaisuus. Olemassa oleva rekisteri ja siihen rakennettu kysely-yhteys mahdollistavat tietopyyntöjen automatisoinnin ja tekevät asioinnin helpoksi. VTJ toimii väylänä tiedonjakamiseen. Eri organisaatiot voivat luvan saatuaan päivittää oman asiakasrekisterinsä VTJ:n kautta. Näin ollen VTJ toimii väylänä datan joustavalle päivittämiselle eri käyttäjille. Tämä säästää myös henkilön vaivaa, kun hänen ei erikseen tarvitse huolehtia näistä kaikista päivityksistä.

VTJ:n kehityskohteet MyDatan näkökulmasta

Näistä lähtökohdista VTJ:n toiminnassa olisi seuraavia kehityskohteita MyDatan näkökulmasta:

1. Nykyisin tietojen luovuttaminen yritykselle massapäivityksiin ja poimintoihin tapahtuu opt out -periaatteella. Tämä ei ole suoranainen ongelma MyDatan kannalta. Suurempi haaste on kieltojen kategorisuus. Jos tietoja ei haluta luovuttaa jollekin taholle, täytyy sulkea pois kaikki muutkin. VRK on listannut ne organisaatiot, joiden asiakasrekisteriä se päivittää⁶⁸. Ihmisen täytyy kuitenkin itse tietää, minkä organisaatioiden asiakasrekisterissä hän on. Huomattavaa on, että tietojen luovuttamista ei kuitenkaan voi kieltää silloin, kun se perustuu lakiin. Tällaisia tapauksia on mm. viranomaiset, oppilaitokset, pankit, vakuutusyhtiöt, perintä ja Veikkaus Oy.
2. MyDatan hengessä henkilön pitäisi pystyä luovuttamaan datansa käyttöluvan perusteella kolmannelle osapuolelle. Tiukimman MyData-tulkinnan mukaan tämän pitäisi olla maksutonta. Tällä hetkellä tietokyselyt maksavat yrityksille 20–30 senttiä. Useissa tietokyselytilanteessa henkilö on jossain muodossa antanut suostumuksen tietojen käyttöön. Maksullisiksi jäisivät tällöin lähinnä yritysten lakisääteiset kyselyt. Näiden luvittaminen on kuitenkin periaatteessa mahdollista, jos henkilö haluaa tämän tehdä. Esimerkiksi pankkien palvelusopimuksissa voidaan antaa lupa kaikkiin pankkitoiminnan hoitamisen kannalta välttämättömien tiedonhakujen suorittamiseen, niidenkin, jotka perustuvat lakiin.
3. Ihmisillä ei ole mahdollisuutta seurata omia lokitietojaan. Kaikilta osin lokitietoja ei edes tuoteta, sillä esimerkiksi massapäivityksistä ja -poiminnoista ei jää selkeää lokia. Suorahauista jää lokimerkinnot, mutta nämä eivät ole avoimesti henkilön itsensä seurattavissa, vaan saatavissa vain väärinkäytösepäilyjen yhteydessä. Täydellisiä lokitietoja ei voi avata ihmisille kahdesta syystä. Yhtäältä nämä saattavat

⁶⁸ <https://eevertti.vrk.fi/osoitetietoja-vaestotietojarjestelmasta-paivittavat-asiakkaat>

häiritä viranomaisten toimintaa esimerkiksi poliisitutkinnan yhteydessä. Toisaalta henkilötietoja paljastava lokimerkintä voi itsessään loukata hauntekijän yksityisyyden suojaa. Tämä voidaan kiertää niin, että lokimerkinnässä henkilölle näkyy vain tietoa hakenut organisaatio.

4. Neljäs kehityskohde liittyy virkatodistuksiin. Todistus maksaa asiakkaalle 9 euroa. Käytännössä se on vain oikeaksi todistettu ote väestötietojärjestelmästä. Tässä yhteydessä ihminen maksaa oikeaksi todistamisesta sekä tietojen rajauksesta. Virkatodistukseen tulevat tiedot määräytyvät todistuksen käyttötarkoituksen mukaan, eikä siinä saa ilmetä tietoja, jotka eivät lain mukaan saa vaikuttaa käyttötarkoituksen kohteena olevaan päätökseen. Tyypillisesti ulos rajattavia tietoja ovat kieli ja uskonto ja varsinkin näiden yhdistelmä, jonka pohjalta voidaan arvata henkilön etnisuus. Tietojen luovutuksen kontrollointi on MyDatan hengen vastaista. Toisaalta järjestelmä, jossa ihmiset voivat jakaa virkatodistusta vastaavia oikeaksi todistettuja tietoja itse, vaikeuttaa luovutettavien tietojen rajausta. Tämä luo riskin, että eri toimijoiden päätöksiin alkavat vaikuttaa nekin tiedot, joita lain mukaan ei saa käyttää. Nykyiset tietojen luovuttamisen käytännöt perustuvat olemassa olevaan lainsäädäntöön.

MyDatan vaikutukset VTJ:n toimintaan, tietojärjestelmiin ja talouteen

MyData-periaatteen mukaiseen tiedonjakamiseen siirtymisen vaikutukset Väestörekisterin toimintaan ilmenevät maksutuloissa, toimintamalleissa sekä tarvittavassa infrastruktuurissa. Muutokset virkatodistuksia koskevilla käytännöillä vaikuttaisivat lisäksi maistraatteihin. Tiedonluovutuksen käytäntöjen muuttuminen vaikuttaisi myös tietoa käyttäviin yrityksiin, jotka saattaisivat joutua uusimaan tietojärjestelmiään.

Jos henkilöä koskevat kyselyt muutetaan maksuttomiksi silloin kun niihin on yksilön antama lupa, tarkastusten tuottama tulovirta pienenee. Väestötietojärjestelmän liiketaloudelliset palvelut ovat tuottaneet yhteensä 8,41 miljoonaa euroa vuodessa. Näistä 4,6 miljoonaa syntyy kyselypalveluista.⁶⁹ Tuottoja on tarkemmin käsitelty luvussa 4. Kyselypalvelun käytön takana on usein jokin toiminto, joka riippuu ihmisen itse käynnistämästä prosessista eli esimerkiksi kaupallinen palveluntarjoaja haluaa tarkistaa jonkin tiedon. Tämä voi olla esimerkiksi palvelun käyttöön vaikuttava ikä tai osoite. Näin ollen ihmisellä on insentiivi tämän tiedon luovuttamiseen. Toisaalta kyselyn taustalla voi olla myös jokin yrityksen omaan prosessiin liittyvä velvoite tai esimerkiksi perintätoimi. Myös asiakasrekisterien ylläpito on mahdollista hoitaa yksilöiden luvituksella, jolloin sen 1,53 miljoonan euron vuotuiset tuotot voitaisiin menettää. Yhteensä kyselypalvelut ja asiakasrekisterit ovat tuottaneet vuodessa 6,13 miljoonaa euroa VTJ:lle kohdistettuja tuloja. Tämä on teoreettinen katto tulonmenetykselle, jos MyData toteutetaan maksuttomuuden periaatteella. Periaatteessa tulonmenetykset voivat nousta tästäkin, kun huomioidaan kysynnän mahdollinen kasvu. Lopullinen tulonmenetykset riippuu siitä, millä toimintalogiikalla MyDataan siirtyminen toteutetaan ja miten yritykset muuttavat käyttäytymistään. Periaatteessa yritykselle voi olla edullisempaa toteuttaa kyselyt edelleen ilman yksilöiden hallitsemia lupia kuin hakea säästöä näiden luvituksen kautta.

Luvussa 4 on käsitelty erilaisia hinnoittelumalleja datan käytölle. Yhtenä vaihtoehtona on esitetty kattoa ilmaisen datan määrälle sekä maksullisuuden säilyttämistä yksityisille palveluntarjoajille. Merkitys VTJ:n tulonmuodostukseen riippuu myös MyData-operaattorin ja VTJ:n suhteesta ja MyData-operaattorin hinnoittelupolitiikasta. Tämän lisäksi oletuksena on,

⁶⁹ Väestörekisterikeskus, 2017

että MyData-operaattori toimisi rinnakkain olemassa olevien rekisterien tietotuotteiden myynnin kautta. Jos VTJ säilyttäisi omat tuotteet maksullisina ja vain operaattori toimisi maksuttomasti, niin on mahdollista, että VTJ säilyttäisi maksutulonsa ainakin jonkin aikaa. Tämä johtuu etenkin siitä, että yritykset ovat jo rakentaneet kysely-yhteyden VTJ:n järjestelmiin, eikä niillä välttämättä ole halua uusia järjestelmiään.

Jos MyDatan kehittämisen lähtökohtana halutaan pitää, että luvitetut tietopyynnöt ovat maksuttomia, niin on kuitenkin varauduttava siihen, että myyntitulojen menetykset voivat nousta miljoonaluokkaan. Muutokset myyntituloissa riippuvat siitä kuinka helppoa yksilöiltä on pyytää lupa tiedon käyttöön erilaisissa käyttötilanteissa. Tätä tutkimusta varten ei ollut saatavilla yksityiskohtaista tietoa VTJ:n luovuttaman henkilötiedon käyttötarkoitusten jakautumisesta, joka olisi mahdollistanut tarkemman arvion myyntitulojen muutoksesta. Huomattavaa on, että hinnoittelun muutokset tulevat todennäköisesti lisäämään VTJ:n kysyntää. Tämä johtuu etenkin siitä, että osittain vastaavia palveluita tarjoavat myös kaupalliset toimijat. Yritykset saattavat siirtyä näiden tarjoamien palveluiden piiristä käyttämään VTJ:tä, jos sen hinnoittelulogiikka muuttuu.

Virkatodistuksen muuttuessa maksuttomaksi ja datapohjaiseksi maistraatit menettäisivät niiden myynnistä tulevat tulot. Virkatodistusten hinnoittelu on kuitenkin lähellä kustannusperusteisuutta, joten oletettavasti myös manuaalisen työn väheneminen laskisi kustannuksia. Myyntitulot eivät siis suoraan aiheuttaisi tappiota julkiselle sektorille. Jäljelle jääisivät kuitenkin ne kustannukset, jotka syntyisivät sähköisen järjestelmän ylläpidosta.

Käytänteiden muutos: MyData-lähestymistapa muuttaisi jonkin verran hallinnollisia käytänteitä. Hallinnollisten käytänteiden muutos heijastuu myös tiedon käyttäjiin kuten viranomaisiin ja yrityksiin. Oleellista on se, minkä roolin Väestörekisterikeskus ottaa tiedon käytön hallinnassa.

Väestörekisterikeskus antaa tietojen hyödyntäjälle hallinnollisen päätöksen, ns. tietoluvan. Tämä pitää sisällään tiedon hyödyntäjän oikeudet, velvollisuudet ja mahdolliset maksut. MyData muuttaisi nämä tietoluvat yksilöiden hallitsemiksi käyttöluviksi. Sopimuksen tietojen käytöstä tekisi se henkilö, jonka tiedoista on kysymys. Tämä vähentää tarvetta luvitukselle ja siihen liittyvälle hallinnolle. Kokonaan tietoluvat eivät häviä, mikäli halutaan edelleen mahdollistaa tietojen luovutus maksullisesti ilman yksilötason käyttölupia. Kiinnostava lisä kokonaisuuteen on se, että VRK:lla on oikeus tarkastaa tiedonluovutuksen kohteena olevien organisaatioiden tietoturvakäytänteet, jotka ovat luottamuksellisia. Jos tiedon jakaminen siirtyy ihmisen itsensä vastuulle, Väestörekisterikeskus ei voi enää tarkastaa tietoturvaa. Tämä jälkeen täytyy pohtia, miten tietoturvan tasoa valvotaan. Oleellinen kysymys on mm. se, voidaanko MyDataa hyödyttäviltä yrityksiltä vaatia esimerkiksi auditointia. Tämäkään ei riitä, jos henkilöt jakavat datansa koneluettavassa muodossa itse, jolloin viranomaisilla ei ole keinoja valvoa, minne data joutuu. Vastaanottava yritys voi toimia myös ulkomailla.

Mahdollinen tapa riskien hallintaan olisi se, että laadittaisiin standardisopimus pohjia osaksi MyDataa hyödyntävien palvelujen käyttäjäehtoja. Näin julkinen valta voi vaikuttaa siihen, miten henkilötietoa käytetään myös MyDataan perustuvissa ratkaisuisa. Sopimus tehtäisiin kuitenkin aina palveluntarjoajan ja henkilön välillä, ja sopimus pohja olisi ennen kaikkea suositus käyttöehdoiksi. Standardisopimukset olisivat siis tapa edistää hyviä käytänteitä avoimessa tiedon jakamisessa sekä hallita ongelmaa, joka syntyy siitä, etteivät henkilöt käytännössä lue sopimusehtoja ja että heidän mahdollisuutensa vaikuttaa niiden sisältöön on muutenkin rajallinen. Tietojen väärinkäytön riskiä voidaan pienentää, jos käyttäjäehdoissa heti alussa viitataan siihen, että henkilötiedon käyttö on linjassa yleisten, standardoitujen sopimusehtojen välillä. Tämän valvonta on luonnollisesti haasteellisempi

asia. Standardisopimusten luomisessa täytyy myös tarkastella viranomaisten työnjakoa. Käytännöllistä voisi olla esimerkiksi se, että rekisteridatoilla luotaisiin yhteisiä sopimuspohjia. Kehitystyön tavoitteena tulisi olla osoittaa tämä työ oikealle viranomaiselle.

MyDatan avulla olisi mahdollista osallistaa ihmiset tietosisältöjen oikeellisuuden huoltamiseen. VTJ:n tietosisällöissä ilmenee puutteellisuuksia ja virheitä etenkin niissä sisällöissä, jotka eivät ole pakollisia. Vapaaehtoinen tieto on mm. sähköpostiosoite. Tämän lisäksi virheitä on mm. syntymäpaikkakunnissa varsinkin kirjoitusasun kohdalla. MyData-hengessä ihmisillä tulisi olla suurempi oikeus vaikuttaa itseään koskevien tietojen sisältöön. Ihmisten mahdollisuus hyödyntää VTJ:tä laajemmin saattaisi motivoida heitä myös tarkistamaan tietojensa oikeellisuuden. VTJ:n osalta tietosisällöt eivät kuitenkaan ole keskeinen ongelma, joten tältä osin ero nykytilanteeseen jää pieneksi. Tilanne saattaa muuttua, jos VTJ:n tietoja hyödyntävät tahot haluavat kommunikoida henkilön kanssa sähköpostilla. Tästä voi muodostua merkittävä ongelma, jos tiedot eivät ole ajantasaisia.

Inframuutokset: MyDatan toteuttaminen vaatisi useita tietojärjestelmätason muutoksia. Näitä ovat mm.

Yksityisyysasetusten hallinta: Valtuutusten hallintaa rakennetaan osaksi Suomi.fi-palvelua. VTJ:hin kohdistuu laajasti kaupallisten toimijoiden kyselyitä. Nykyisin suostumukset tiedonhakuun haetaan joka palvelutilanteessa kysymällä, osana käyttöehtoja tai suullisesti palvelutilanteessa. Suomi.fi-palveluun voidaan kuitenkin luoda mahdollisuus näiden käyttöilupien hallintaan. Tällöin henkilölle voidaan tarjota nykyistä kattavampi mahdollisuus seurata sitä, kenelle lupia on annettu sekä perua niitä tarpeen mukaan.

Lokitiedot: VTJ tuottaa jo nyt lokitietoja, mutta nämä eivät ole käyttäjien saatavilla. Omien tietojen käytön valvomisen mahdollistamiseksi lokitiedot tulisi avata myös ihmiselle itselleen. Tämän tosin täytyy tapahtua edellä esitetyin rajoituksin. Lokitietojärjestelmää tulisi kehittää niin, että se mahdollistaa myös massapäivitysten ja -poimintojen seurannan. Tämä ei kuitenkaan ole kovin kriittistä.

Kapasiteetin kasvattaminen: Tietopyyntöjen määrä kasvaa jo nyt useita kymmeniä prosentteja vuodessa. Tietoliikenteen kasvaminen vaatii myös lisää teknistä kapasiteettia. Jos MyData mahdollistaa edullisemman tiedonhaun, ei kysyntä ainakaan vähene. Tämän takia on varmistettava, että tiedonsiirtokapasiteetti on riittävällä tasolla. Teknisesti on myös huomioitava se, halutaanko tiedon siirtäminen mahdollistaa myös muilla tekniikoilla kuin niillä, joilla se on nykyisin mahdollista. Nyt tietöformaattina ovat XML-skeemat, ja tiedonsiirto edellyttää kysely-yhteyttä VTJ:n kanssa. Tämä ei vielä mahdollista tiedon jakamista kaikille mahdollisille hyödyntäjille.

Yhteensopivuus muun MyData-infran kanssa. Kappaleessa 4 esitetään yhtenä mahdollisena arkkitehtuurina MyDatan toteuttamiseen yhteistä kansallista operaattoria. Sitä voi hyvin soveltaa myös väestötietojärjestelmään. Arkkitehtuurissa datan ja sovellusten hallinta tapahtuu palveluväylän yhteyteen rakennetun MyData-tilin kautta. Lähtökohtana on se, että VTJ on yksi niistä rekistereistä, jotka liittyvät palveluväylään. Tämä arkkitehtuuri on toteutettavissa VTJ:n osalta. Huomattavaa on se, että arkkitehtuuri käsittelee ensisijaisesti tiedon jakamista. Tämän lisäksi tulee vielä lokitiedonhallinta. Päätettäväksi jää, toteutetaanko tämä rekisterikohtaisesti vai rakennetaanko keskitetty järjestelmä, jossa ihmiset näkevät yhdessä paikassa kaikkien julkisten tietolähteiden lokimerkinnot.

Muut kustannusvaikutukset: Muuttuvat käytänteet saattavat vaikuttaa jonkin verran hallinnollisiin kustannuksiin, mutta pääasiassa VTJ:n hallinto kevenisi, jos esimerkiksi käyttöilupien hallinta siirtyisi operaattorille. VTJ:ään kohdistuvien kyselyiden määrä on jo nyt

kasvussa, eikä MyData tule tätä vähentämään. Nykyisin VTJ:n palvelussa käsitellään 300 miljoonaa tietoyksikköä vuodessa ja käsittelyn hinta on laskennallisesti 4,6 senttiä. Näin ollen tietoyksiköiden käsittelymäärän lisäys yhdellä prosentilla maksaisi laskennallisesti 13 800 euroa. Myös järjestelmien peruskapasiteettiin pitäisi investoida. Osittain järjestelmäinvestoinnit voidaan toteuttaa osana Suomi.fi- ja YTI-hankkeita. Kokonaisuudessaan investointitarve nousee VTJ:n osalta haastatteluissa annettujen arvioiden perusteella miljoonaluokkaan, jos kaikki MyDatan asettamat vaatimukset toteutetaan. Tässä oletuksena on se, että olemassa olevat kyselyrajapinnat palvelevat myös MyData-käyttöä tai vaativat vain hieman kehitystyötä.

MyData-malli toisi myös hieman säästöjä. Tietolupahallinto on jo siinä määrin tehostettu, ettei tästä ole saavutettavissa mainittavia säästöjä. MyData-periaatteeseen siirryttäessä säästöjä syntyy kuitenkin siitä, että tietojen luovutukseen liittyvä valvonta sekä tietoturvaan liittyvät tarkastukset vähenisivät. Käytännössä säästöt syntyisivät siis siitä, että valvontavastuu siirtyisi VRK:lta yksittäisille henkilöille itselleen. Toisaalta jos tietoturvaa halutaan kehittää, on tietoturvalle löydettävä uusi valvontamalli, joka saattaa lisätä jonkin toisen viranomaisen kustannuksia.

Kustannussäästöt eivät riitä kattamaan menetettyjä maksutuloja. Näin ollen tulonmenetykset on kompensoitava VRK:lle muilla tavoilla. Vaihtoehtoiset ansaintamallit ovat rajoitetut, koska lähtökohtaisesti viranomaistoiminta on ilmaista mm. neuvonnan osalta. Uusien palvelujen tuottaminen viranomaistoimintana saattaisi johtaa kilpailutilanteeseen markkinoilla. Tämäkään ei kuulu viranomaistoimintaan.

5.2 MyData terveydenhuollossa ja hyvinvointipalveluissa

Suomessa kehitetään vahvasti sekä terveydenhuollon digitalisaatiota⁷⁰ että terveystietojen⁷¹ toissijaista käyttöä⁷¹. Nämä molemmat kehityssuunnat sivuavat MyDataa. Kokonaiskuvan saavuttamiseksi kannattaa tarkastella myös näitä kumpaakin kehityssuuntaa. Terveydenhuollon digitalisointi edellyttää datan siirrettävyyttä ja sen saatavuuden parantamista. Datan toissijainen käyttö taas kiinnittää huomiota niihin periaatteellisiin keskusteluihin, joita datan hyödynnettävyydestä juuri nyt käydään.

Terveydenhuollon näkökulma on valittu esimerkiksi siitä syystä, että terveydenhuollossa henkilötietoja syntyy enemmän kuin useilla muilla julkishallinnon aloilla. Terveystietojen⁷¹ lisäksi on myös juridinen erikoisasema. Esimerkiksi yksityisyyden merkitys korostuu vielä muita datatyyppisiä enemmän. Terveyssektorilla on myös laajasti erilaisia julkisia ja yksityisiä toimijoita, jotka voivat tuottaa ja hyödyntää dataa eri tavoilla. Myös ihmisen oma aktiivinen rooli muuttaa datan luonnetta. Uudet laitteet mahdollistavat ihmisten oman hyvinvoinnin tilan seuraamista, joko osana sairauden hoitoa tai pelkästään kiinnostuksesta omaan hyvinvointiin.⁷² Nämä kaikki kehityssuunnat luovat tarvetta arvioida sitä, miten myös julkisen terveydenhuollon luomaa dataa voitaisiin käyttää laajemmin MyData-periaatteiden mukaisesti. Ennakointi, ihmisten vastuuttaminen ja big datan hyödyntäminen ovat keskeisiä kehityssuuntia, jotka sivuavat myös MyDataa.

Ennakoivan terveydenhuollon tavoitteena on ennaltaehkäistä sairauksia ennen niiden ilmenemistä. Tähän liittyy terveydentilaan vaikuttavien tekijöiden kuten verenpaineen, kolesterolin jne. seuraaminen sekä elintapoja koskevan datan yhdistäminen tähän tietoon.

⁷⁰ Hyppönen & Ilmarinen, 2016

⁷¹ Raula & Savolainen, 2016

⁷² Swan, 2013

Ennakoitavuutta pidetään keskeisenä tapana hallita vanhenevien länsimaiden nousevia terveydenhoidon kustannuksia. Ennakoitiin liittyä osittain myös ihmisten oma vastuuttaminen, mikä tarkoittaa kaikkia niitä menetelmiä, joilla tuetaan ihmisiä ottamaan laajemmin vastuuta omasta terveydestään. Tähän tavoitteeseen päästään vain, jos ihmisille voidaan tuottaa heidän terveyttään koskevaa tietoa ymmärrettävässä muodossa.^{7374 75}

Big datan hyödyntäminen terveydenhuollossa ja lääketieteellisessä tutkimuksessa kasvaa samalla kuin datan määrä on lisääntynyt. Terveydenhuollon tuottamat datamäärät ovat olleet todella nopeassa kasvussa ja dataformaatit ovat heterogeenisiä. Big datan potentiaalina nähdään erityisesti kustannustehokkuuden paraneminen, hoidon oikeellisuuden varmistaminen sekä kansanterveyden seuranta. Laajat datamassat mahdollistavat myös entistä paremman lääketieteellisen tutkimuksen.⁷⁶ Tämä käytötapaus vastaa Suomessakin edistettävää datan toissijaista käyttöä.

Ennakoivan terveydenhuollon ja big datan haasteena on molemmissa datan saatavuus. Big datan osalta haaste nousee esiin varsinkin, kun tarvitaan henkilötietoja. Osa haasteista on selkeämmin teknisiä, kuten tietojärjestelmien ja dataformaattien yhteensopimattomuus. Oleellista kuitenkin on se, että eri datalähteitä voidaan yhdistää mahdollisimman tehokkaasti. Tämä tarkoittaa myös sitä, että terveydenhuollossa olisi käytössä muitakin kuin sen itsensä tuottamaa dataa. Tätä voidaan edistää MyData-tyylisellä lähestymistavalla, koska tietolähteiden käyttöä rajoittaa yksityisyyden suoja. Tämän lisäksi ihmisten oma vastuuttaminen hoidosta vaatii myös uudenlaista lähestymistapaa. Oleellista on tukea sitä, että ihminen ymmärtää omien valintojensa ja terveydentilansa välisen yhteyden. Tämän vuoksi pitäisi kehittää sellaisia ihmisen toimintaa tukevia applikaatioita, jotka voivat hyödyntää ihmistä koskevaa dataa monipuolisesti.

Valtio on laatinut Sote-tieto hyötykäyttöön -strategian 2020.⁷⁷ Se määrittää useampia toimenpiteitä, joissa kehitetään sote-tiedon hyödynnettävyyttä. Esimerkiksi raportissa esitellyt ODA- ja virtuaalisairaalahankkeet ovat strategian mukaisia toimenpiteitä. Myös datan toissijainen käyttö on osa strategiaa. Strategia on moniulotteinen ja pitää sisällään niin teknologisia kuin lainsäädännöllisiä osioita. Oleellisena tavoitteena on parantaa terveydenhuollon vaikuttavuutta. Strategian tavoitteena on toiminnallisesti integroitua terveydenhuollon ekosysteemi. Sen keskiössä on lisäarvon tuottaminen yhteisen tiedon jalostamisesta.

Terveysdatan tyypit

Ihmistä koskeva terveysdata voidaan jaotella useilla tavoilla. Tässä yhteydessä jaottelu on tehty etenkin siitä näkökulmasta, mikä datan hallittavuus on MyDatan näkökulmasta. Oleellista tällöin on se, kuka dataa hallinnoi ja minkälaisen lainsäädännön alla se on. Tehty jaottelu perustuu haastatteluihin. Datatyyppit ovat:

- Genomidata
- Henkilön itsensä mittaama data
- Terveyspalveluiden tuottajien tuottama data

⁷³ Grossglauser & Saner, 2014

⁷⁴ Chawla & Davis 2013

⁷⁵ Lupton, 2013

⁷⁶ Raghupathi & Raghupathi, 2014

⁷⁷ https://www.julkari.fi/bitstream/handle/10024/125500/URN_ISBN_978-952-00-3548-8.pdf?sequence=1

- Muu data

Genomidatan hyödyntäminen on vahvasti kasvussa, ja huomattava osa yksilöityä terveydenhuoltoa käsittelevää kirjallisuutta koskee juuri genomidatan hyödyntämistä. Genomin selvittämisen kustannukset ovat laskeneet radikaalisti, mikä on lisännyt siihen liittyvien kaupallisten palveluiden tarjontaa. Genomidata mahdollistaa yksilöidyn terveydenhuollon ja riskiarvioinnit. Genomidata voi olla MyDataa tai muuta dataa eli periaatteessa se voi olla vain terveystietopalvelun tuottajan hallussa, jolloin henkilö itse saa vain riskiarvion. Muusta datasta poiketen genomidata tarvitsee tuottaa vain kerran. Näin ollen ihminen voi luovuttaa kerran tuottamansa genomidatan haluamilleen tahoille. Kansallisesti laaditussa genomistrategiassa on ehdotettu, että Kanta-palveluun lisättäisiin tiedot ihmisen perimästä⁷⁸.

Ihmiset tuottavat itse yhä enemmän terveyteen ja hyvinvointiin liittyvää dataa. Trendiä ovat lisänneet mm. sykemittarit ja niitä seuranneet aktiivisuusrannekkeet. Yhä useammin ihmiset mittaavat itse myös suoraan terveydenhoitoon liittyvää dataa kuten verenpainetta, verensokeria tai vaikkapa astman hoitoon liittyviä puhallusarvoja. Tästä itseseurannasta on tullut oleellinen osa terveydenhuoltoa kroonisten sairauksien hoidossa. Periaatteessa kannettavat älylaitteet tuottavat ihmisestä huomattavan määrän sellaista dataa, jota ei tällä hetkellä tallenneta ja tai hyödynnetä. Itse mitattu data jää luonnollisesti myös ihmiselle itselleen. Tämä mahdollistaa myös uudet datapohjaiset palvelut. Itse tuotetun datan ja sen yhdistelyn mahdollisuudet ovat vielä pitkälti kartoittamatta, mutta datan vaikuttavuuden kannalta tämä on kuitenkin erittäin kiinnostava alue⁷⁹.

Terveystietopalveluiden tuottama data koostuu kaikista niistä aineistoista, joita terveydenhuollon palveluiden tuottajat keräävät. Ne muodostuvat mm. laboratoriokokeista, röntgenkuvista, lääkärinlausunnoista, lääkeresepteistä ja potilaskertomuksista. Tämän datan hallintaan Suomessa ollaan kehittämässä Kanta- ja Omakanta-palveluita. Omakanta antaa ihmiselle mahdollisuuden selata omia tietojaan. MyData-ratkaisuna tätä ei kuitenkaan voida pitää, koska tietojen siirto muihin järjestelmiin koneluettavassa muodossa ei onnistu. Kanta-järjestelmällä on kuitenkin merkittävä rooli siinä, että sen avulla ihmisen hoitohistoria seuraa häntä palveluntuottajalta toiselle. Järjestelmää voivat hyödyntää lähtökohtaisesti kaikki Suomessa toimivat viralliset terveystietopalveluiden tuottajat. Muilla tahoilla ei ole mahdollisuutta saada tietoa Kanta-järjestelmästä edes henkilön itsensä valtuuttamana. Kaikki terveydenhuollon toimijat eivät ole siirtyneet Kanta-järjestelmän käyttöön. Kanta-arkkitehtuuri on ristiriidassa MyData-ajattelun kanssa myös siinä, että se on nimenomaan keskittävä arkkitehtuuri eikä hajautettu, kuten MyData-arkkitehtuurin yleensä ajatellaan olevan. Pragmaattisesti ajatellen Kanta on kuitenkin merkittävä kehitysaskel hoitoketjussa tapahtuvan tiedonsiirron kehittämisessä ja se parantaa tiedon uudellenkäyttöä mm. perus- ja erikoissairaanhoidon välillä sekä auttaa tiedon siirtymistä ihmisen muuttaessa paikkakunnalta toiselle. Tiedon liikkumisen merkitys korostuu sote-uudistuksen ja siihen mahdollisesti liittyvän valinnanvapauden myötä.

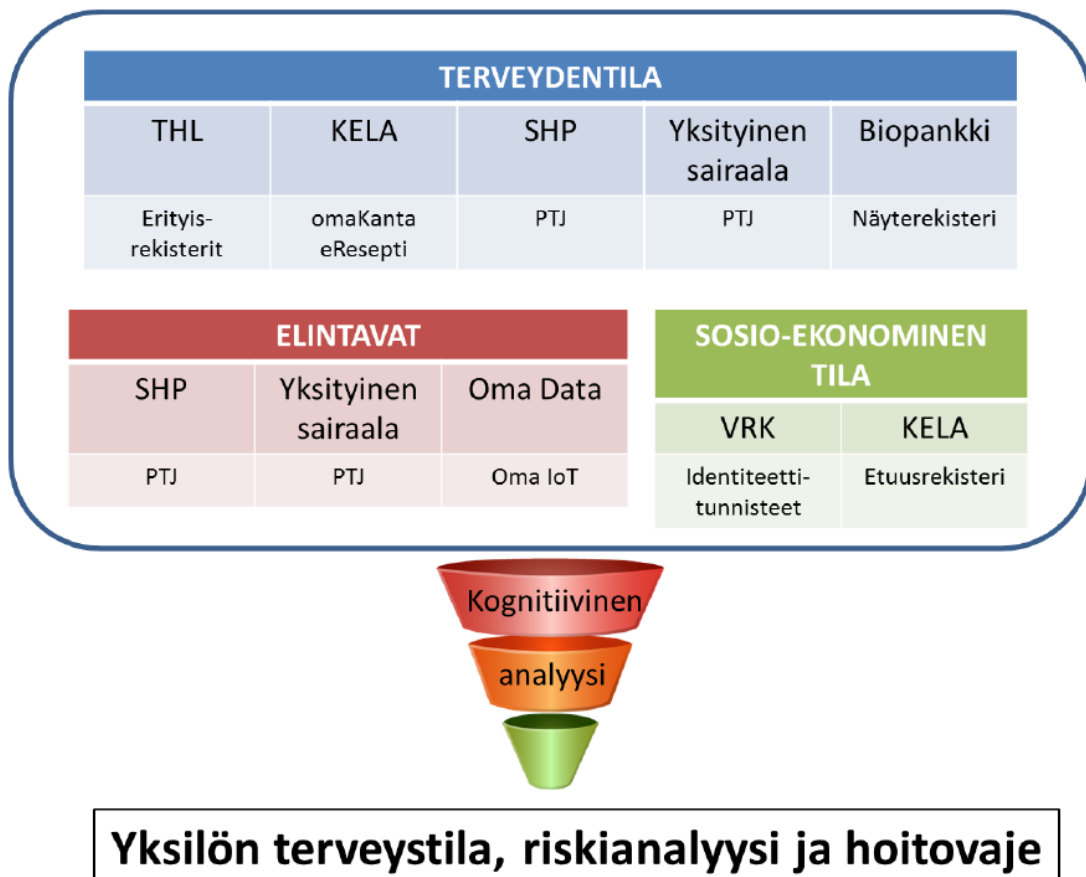
Laajempi haaste terveystiedon hallittavuudelle on palveluntuottajien omiin tietojärjestelmiin tallennetun tiedon laajempi hyödyntäminen. Terveystoimijoiden käytössä on useita eri järjestelmiä, joihin sisältyy potilas- ja asiakastietoa. Tämän tiedon siirrettävyys ja yhdisteltävyys ei aina ole helppoa edes sisäisesti, vaikka järjestelmien integraatiota on kehitetty. Näiden tietovarantojen hyödynnettävyyden parantaminen vaatii kehitystä myös tekniseltä integraatiolta.

⁷⁸ Kansallisen Genomistrategian työryhmä, 2015

⁷⁹ Swan 2013

Muuta kuin edellä määritettyä dataa voidaan hyödyntää myös terveydenhuollossa. Toteutuneita esimerkkejä on tunnistettu vasta muutama. Esimerkkinä tarpeesta on aivoinfarktista kotona toipuvan henkilön kuntoa seuraava järjestelmä, jota kehitetään yrityksen ja HUS:n välisenä yhteistyönä. Tämä järjestelmä voisi hyötyä teletiedoista, koska muihin ihmisiin suuntautuvien yhteydenpitotapojen muuttuminen voi kertoa ongelmista. Toisena esimerkkinä haastatteluissa nousi esiin se, että poikkeava käyttäytyminen sosiaalisessa mediassa saattaa ennakoida esimerkiksi itsetuhoisuutta masennuspotilailla. Nämä esimerkit ovat vielä teoreettisia sovelluskohteita, mutta osoittavat sen, että terveydenhuollon järjestelmät saattavat hyötyä datasta, jota ei alunperin ole tuotettu terveydenhoitoa varten.

Kuva 5.1 Yksilön terveydentila



Erilaisten tiedonlähteiden hyödyllisyys hyvinvointiin liittyvien sovellusten kannalta. Lähde: Lehto & Neittaanmäki, 2017⁸⁰

Kuva 5.1 tiivistää datan hyödyllisyyden yksilön hyvinvoinnin kannalta. Dataa on kertynyt useista lähteistä useisiin eri tietovarantoihin. Näitä yhdistämällä saadaan tietoa niin yksilön terveydentilasta, riskeistä kuin myös hoidosta ja sen vaikuttavuudesta. Tiedon yhdistäminen vaatii kuitenkin muutakin kuin saatavuuden. Avuksi tarvitaan kehittyneitä analytiikkaa. Kuvan 5.1 mukaisesti analytiikan apuna voidaan käyttää kognitiivista analyysia tai tietojenkäsittelyä. Tämä tarkoittaa massadatan, tekoälyn ja näitä tukevan viestinnän yhdistämistä. Tietoisuus tästä on levinnyt sekä käytännön terveystyötä tekevien pariin että lääketieteellisiin tiedekuntiin. Tällä hetkellä tutkitaan, miten kognitiiviseen tai koneoppimiseen perustuvia sovelluksia voitaisiin käyttää hyväksi hoitotyössä.

⁸⁰ Kuvan tekijät (Lehto & Neittaanmäki, 2017) antoivat luvan julkaista kuvan tässä yhteydessä CC-BY 4.0 lisenssin pohjalta. Tekijänoikeudet kuuluvat heille.

Datan toissijainen käyttö

Sosiaali- ja terveysministeriö on valmistellut lakia sosiaali- ja terveyshuollon asiakas- ja potilastiedon toissijaista käyttöä varten. Tiedon jatkohyödyntäminen on tarkoitus mahdollistaa tutkimuksessa, johtamisessa sekä toiminnan kehittämisessä ja valvonnassa. Lainsäädännön on taattava tietojen teknologiariippumaton hyödyntäminen sekä oltava linjassa EU:n tietoturva-asetuksen kanssa.

Datan toissijainen käyttö tarkoittaa sitä, että dataa käytetään muuhun kuin siihen tarkoitukseen, jota varten se on alunperin tuotettu. Terveysdatan tapauksessa ihmisestä tuotetaan paljon dataa, mm. laboratoriotestejä, kuvia, lausuntoja jne. Tämän datan ensisijainen käyttötarkoitus on henkilön terveydentilan hoitaminen ja seuranta. Toissijaisia käyttötarkoituksia ovat kaikki muut, kuten esimerkiksi tutkimus tai johtaminen.

Datan toissijaisesta käytöstä ollaan tekemässä Suomelle kilpailutekijää, koska terveysteknologia on toimialana vahvassa kasvussa⁸¹. Tätä tuetaan sekä lainsäädännöllä että käytännön kehityshankkeilla. Toissijainen käyttö ei ole MyDataa, mutta senkin tavoitteena on parantaa datan käytettävyyttä. Terveysdatan toissijainen käyttö tarkoittaa kerätyn datan käyttöä muuhun kuin alkuperäiseen tarkoitukseen eli lähtökohtaisesti ihmisen hoitamiseen. Tyypillisesti toissijainen käyttö on potilastietojen käyttämistä tieteelliseen tutkimukseen. Valmisteltavana olevassa lakiluonnoksessa ehdotetaan, että tutkimuskäyttö laajennetaan myös innovaatiotoimintaan niin, että data on myös tietyin ehdoin yritysten käytettävissä.⁸²

Toissijainen käyttö haastaa yksityisyydensuojaa. Tätä ylläpidetään datan anonymisoinnilla tai pseudonymisoinnilla, joten henkilön yksityisyyden ei pitäisi vaarantua. Lisäksi tietosuojakriteerit ovat vahvat. Tästä on tosin myös eriäviä mielipiteitä, koska riittävän suurista yhdistellyistä aineistoista on mahdollista tunnistaa myös yksilöitä. Tätä riskiä pyritään kuitenkin minimoimaan. Läpinäkyvämpi lähestymistapa saattaisi myös edistää data-aineistojen kerryttämistä. Tällä hetkellä biopankit⁸³ noudattavat jo aika hyvin MyData-periaatteita. Ihmisellä on mm. oikeus tietää, missä hänestä otettu näyte on, mihin sitä on käytetty ja mitä tuloksia tutkimus on tuottanut. Näytteen antaminen biopankkiin perustuu suostumukseen. Näytteen siirrettävyys ei luonnollisesti ole mahdollista, mutta fyysistä näytettä käsitellään luonnollisesti eri tavalla kuin dataa. Näytteen analyysistä syntyneen datan siirrettävyys on luonnollisesti toinen asia. Merkittävä ihmiskeskeinen ulottuvuus biopankeissa on myös se, että näytteen antajalla on oikeus saada tietoa näytteisiin perustuvasta tutkimuksesta, jos sen voidaan olettaa auttavan henkilön omassa hoidossa.

Käytännössä toissijaisen datan hyödynnettävyyden parantamiseksi Sitra toteuttaa Isaacus-hanketta. Hankkeen tavoitteena on luoda yhden luukun palveluoperaattori hyvinvointidatan hyödyntämiseen. Palveluun on tarkoitus saada mm. luvat tutkimusta varten. Käynnissä olevat pilottihankkeet kattavat mm. tietojohdamisen hankkeita terveydenhuollossa. Toissijaisen tiedon merkitys näkyy esimerkiksi siinä, että IBM on sijoittanut Watsonin kehitysyksikön Suomeen juuri terveydenhuollon alalla. Watson on IBM:n tekoälyalusta, jota IBM kehittää mm. Suomessa.

Toissijaisella hyödyntämisellä haetaan vaikuttavuutta tutkimiseen ja johtamiseen ja pyritään luomaan Suomesta kiinnostavampi toimintaympäristö kansainvälisille investoinneille. Tämä on eri tulokulma datan hyödyntämiseen kuin MyData, mutta nämä voivat täydentää toisiaan.

⁸¹ http://healthtech.teknologiateollisuus.fi/sites/healthtech/files/htf_vientitilasto_2016.pdf

⁸² Raula & Savolainen, 2016

⁸³ <http://www.biopankki.fi/>

Toissijainen käyttö mahdollistaa tieteellisen tutkimuksen ja mm. tekoälyn kehittämisen, koska nykyiset tekoälyteknologiat edellyttävät laajaa aineistoa toimintansa pohjaksi. MyData taas mahdollistaa sen, että yksilö voi täysimääräisesti hyödyntää näitä tekoälysovelluksia omassa hoidossaan. Tekoäly voi arvioida ihmisen hyvinvointia vain niin tarkasti kuin sillä on käytettävissä ihmisen terveyttä koskevaa dataa, jota se vertaa laajoista aineistoista (toissijainen käyttö) luotuihin malleihin. Haastatteluaineistojen perusteella kehitystyö eri aineistojen yhdistelyyn ja tämän pohjalta tehtäviin suosituksiin on vasta alkamassa. Ensimmäiset pilotit eivät suoraan koske terveydenhuoltoa vaan esimerkiksi yksilöllisten kunto-ohjelmien tukemista. MyData on kuitenkin yksi katalyytti sille, että kehitys yksilöllisempään terveydenhuoltoon ja hyvinvointipalveluihin voi jatkua. Tällä hetkellä tekoälyä käytetään lääkärin apuna diagnoosien teossa jne.

Datan hyödynnettävyys kilpailutekijänä nostattaa hyvin merkittäviä kysymyksiä tietosuojan luonteesta. Näistä merkittävin on se, että kansainvälisessä liiketoimintaympäristössä toimijat eivät ole vain suomalaisia. Tämän takia datanhallinta täytyy voida toteuttaa kansainvälisessä ympäristössä.

Terveydenhuollon digitalisaatio

Terveydenhuollon digitalisaatio on laaja käsite, jota käsittelee erittäin laaja eHealth- ja mHealth-kirjallisuus. Koko digitalisoitumisen kenttää ei tässä yhteydessä kuvata, eikä se välttämättä ole relevanttia MyDatan näkökulmasta. Keskeisiä trendejä ovat mm. etäkonsultoinnin lisääntyminen sekä terveydentilan seurannan tehostaminen teknologiaa hyödyntäen. Teknologia mahdollistaa mm. kotihoidon lisäämisen, kun esimerkiksi kuntoutettavan tilaa voidaan valvoa luotettavammin. Diabeteksen kaltaisissa kroonisissa sairauksissa taas verensokerin yms. valvonta tehdään jo nyt laajamittaisesti kotona. Kehittyvät algoritmit ja laitteiden kehityksen myötä tarkentuvan datan avulla voidaan hoidon seurantaa osittain automatisoida ja vaikuttavuutta seurata paremmin.

Digitalisaatio on yhdistelmä uutta teknologiaa, palvelumalleja ja teknologiaa. Tavoitteena on vaikuttavuuden lisääminen, kustannussäästöt sekä ihmisten itsensä osallistaminen. MyDatan näkökulmasta on kiinnostavaa, että käytännössä kaikki palvelut tuottavat tai hyödyntävät henkilödataa jollain tavalla. Tällä hetkellä palveluja kuitenkin kehitetään niin, että julkisen terveydenhuollon järjestelmät eivät tarjoa dataa ihmisille. Kehitys ei näin ollen tukeudu MyData-periaatteeseen.

Suomessa on meneillään useita digitaaliseen terveydenhuoltoon liittyviä hankkeita. Esimerkkejä näistä ovat ODA (Omahoito ja Digitaaliset Arvopalvelut) ja Virtuaalisairaala. ODA keskittyy perusterveydenhoitoon ja Virtuaalisairaala erikoissairaanhoidon. Molemmilla hankkeilla on laajoja tavoitteita kehittää terveydenhuoltoa kustannustehokkaammaksi ja vaikuttavammaksi teknologiaa hyödyntäen.

ODA-hankkeen toteutuksesta vastaa yhdessä 14 kuntaa ja sairaanhoitopiiriä. Sen tarkoituksena on tukea asiakkaiden itse- ja omahoitoa. Keskeisenä tavoitteena on muuttaa hoitoketjuja niin, että asiointia sairaanhoidossa voidaan vähentää ja näin ollen yhteiskunta saa säästöjä. Virtuaalisairaalahanketta vetää Helsingin ja Uudenmaan sairaanhoitopiiri. Hankkeella on vastaavia tavoitteita kuin ODA:lla. Lähtökohtana on se, että tiedon parempi hyödyntäminen mahdollistaa mm. oma- ja itsehoidon. Tämä luo kustannussäästöjä.

Molemmille hankkeille on yhteistä se, että ne integroivat terveyspalveluiden itsensä tuottamaan dataan asiakkaiden tuottamaa dataa. MyData-näkökulmasta on kiinnostavaa, että hankkeet hakevat tehokkuutta yhdistelemällä eri henkilödatan tyyppejä – eli ennen

kaikkea ihmisen itsensä mittaamaa tai ihmisen itsensä antamia oirekuvauksia. Tätä dataa yhdistetään terveydenhuoltojärjestelmän tuottamaan dataan. Datavirta on kuitenkin ihmisestä järjestelmään – ei toisin päin. Näiden hankkeiden pohjalta terveystoimijoille syntyy yhä suuremmat tietovarannot ihmisestä.

MyDatan avulla tälle tietomassalle voisi löytää uusia käyttökohteita. Dataa voisi käyttää mm. nykyistä laajemmin eri terveystalveluiden tarjoajien välillä ml. ulkomaiset palveluntarjoajat, muissa hyvinvointipalveluissa sekä esimerkiksi sovelluksissa, jotka auttavat ihmistä ymmärtämään omaa terveyttään ja tukemaan hoitoprosessia. Nämä sovellukset voivat olla pelejä tai muita mobiilisovelluksia. Virtuaalisairaalanhanke⁸⁴ mm. tutkii opetuspelien käyttöä ihmisten motivoinnissa.

Terveydenhuollon ekosysteemit ja tulevaisuus

Terveysdatan ympärille on syntyemässä useita ekosysteemejä, jotka tukevat terveydenhuoltoa. MyDatan vaikuttavuus syntyy tämän ekosysteemikehityksen kautta. Käytännössä tämä tarkoittaa sitä, että ekosysteemi luo eri palveluntarjoajien verkoston. Yksi toimija voi tarjota tekoälyalustan, toinen tätä alustaa hyödyntävän applikaation ja tämä applikaatio voi hyödyntää MyDataa tai muita datalähteitä. Vaikutus syntyy, kun MyData luo mahdollisuuden kehittää sovelluksia tai muita palveluita. Nämä palvelut taas luovat pohjaa liiketoiminnan kasvulle sekä terveydenhuollon tehostumiselle.

Haastattelut ja aiheesta tehdyt selvitykset^{85 86} ovat nostaneet esille sen, että terveydenhuoltoon on eHealthin, mHealthin ja digitaalisen terveydenhuollon ympärille syntyemässä kasvavaa liiketoimintaa. Yhtenä merkittävänä tekijänä tässä kasvavassa liiketoiminnassa on henkilötiedon käytettävyyden parantuminen. Kansainväliset selvitykset eivät käytä termiä MyData, mutta ongelman määrittely on sama. Palvelujen tuottama lisäarvo edellyttää uusia tapoja tuottaa ja jakaa dataa. Tämä täytyy kuitenkin tehdä tietosuojaa kunnioittavalla tavalla. Kansainvälisessä keskustelussa on luotu useita eri MyDataa sivuavia konsepteja. Esimerkkinä tästä on mm. MIT:n⁸⁷ luoma dataratkaisu.

Suomessa on meneillään useita kiinnostavia kehityshankkeita ja niiden ympärille syntyviä ekosysteemi-aihoita. Esimerkki näistä on mm. selvitykseen haastateltu Virtuaalisairaalanhanke⁸⁸, joka testaa suurten ICT-yritysten koneoppimislustoja ja tekee yhteistyötä analytiikkayhtiöiden kanssa. ODA-hanke⁸⁹ on järjestänyt oirehackathoneja uusien sovellusten löytämiseksi. Oleellista suomalaisissa hankkeissa on se, että niissä tuotetaan dataa (mittaukset) julkisen terveydenhuollon tueksi, kehitetään analytiikkaa sekä koulutetaan lääkäreitä käyttämään tätä analytiikkaa. Tämän lisäksi Suomessa toteutetaan Oulun yliopiston johdolla Tekes-rahoitteista Digital Health Revolution -hanketta⁹⁰, joka on nostanut MyData-lähestymistavan hankkeen keskiöön.

Tällä hetkellä terveysdata virtaa ihmisten omista sovelluksista julkisen terveydenhuollon järjestelmiin. Kotona tehtävässä terveydentilan seurannassa ja oireiden arvioinnissa on menossa useita hankkeita. Ekosysteemimäisen ajattelun mukaisesti näissä hankkeissa yritykset tekevät yhteistyötä julkisen sektorin kanssa.

⁸⁴ Lähde: Virtuaalisairaala 2.0 Hankesuunnitelma

⁸⁵ Taylor, Ronte, & Hammett, 2014

⁸⁶ Levy, Wasden, DiFilippo, & Sur, 2012

⁸⁷ <http://openpds.media.mit.edu/>

⁸⁸ Virtuaalisairaala 2.0 Hankesuunnitelma

⁸⁹ <https://oda-projekti.blogspot.fi/>

⁹⁰ <http://www.digitalhealthrevolution.fi/>

Ekosysteemin kehityksen mahdollistajana olisi merkittävää, että julkisen terveystietojen datan käytettävyys paranisi. Kehitystä tämänkin asian osalta on tapahtumassa. Haastattelujen perusteella terveystietojen potentiaalini ymmärtävät niin viranomaiset, yritykset kuin tutkijatkin. Merkittävänä kehityshankkeena on se, että THL:llä on meneillään Omätietovarannon päälle rakennettavien sovellusten hyväksymismenettelyiden kommentointi⁹¹. Tämä avaa tietä sovelluksille, jotka hyödyntävät julkisia tietovarantoja. Sovelluksille asetettaneen huomattavat turvallisuuskriteerit, mutta samalla kynnys liittymiselle ekosysteemin pyritään pitämään mahdollisimman alhaalla.

Teknologiayhtiöt kokoavat omia ekosysteemejään oman teknologiansa ympärille. Suomessa tätä työtä tekee IBM Watson-teknologioidensa kanssa. Toinen esimerkki on Google, joka kehittää omaa DeepMind- teknologiaansa.⁹² Nämä ekosysteemit eivät kuitenkaan nojaa vielä siihen, että niihin perustuvat palvelut jakaisivat dataa keskenään. Julkiset hankkeet voivat kuitenkin tukea tätä kehitystä toimimalla MyData-hengessä tietopalustoina erilaisille palveluille. Tässä vaiheessa merkitystä on enemmän datan toissijaisella käytöllä, jota hyödynnetään tekoälyn kehityksessä. Tästä toissijaisesta käytöstä haetaan kilpailuetua Suomelle. Ensimmäisenä onnistumisena voidaan pitää IBM:n Watsonin kehityskeskusten sijoittumista Suomeen. Toimiva lainsäädäntö ja -infrastruktuuri lisäävät tätä kiinnostusta entisestään.

Digital Health Revolution -hanke⁹³ samoin kuin Demosin selvitys⁹⁴ ovat nostaneet MyDatan vaikuttavuuden vahvuudeksi sen, että se muuttaa terveydenhuollon logiikan enemmän ennaltaehkäiseväksi ja terveyttä edistäväksi. Tämä edellyttää kahta asiaa. Ensiksikin ihmisen terveydentilaa ja hyvinvointia koskevan datan on oltava ihmisen itsensä käytössä. Tämä mahdollistaa sen, että ihminen voi ottaa paremmin vastuun omasta terveydestään. Data on tarjottava ihmiselle ymmärrettävässä muodossa ja mielellään ymmärrettävien suositusten kanssa. Tämä edellyttää sitä, että datan ja ihmisen väliin luodaan sovelluksia, jotka auttavat datan tulkitsemisessa ja datan muuttamisessa toimintasuosituksiksi. Toinen kehityssuunta on avata MyData niille toimijoille, jotka tuottavat hyvinvointipalveluja varsinaisen terveydenhuoltojärjestelmän ulkopuolella. Näitä ovat mm. ravitsemusneuvojat, personal trainerit jne. Periaatteessa dataa voi jakaa laajemmin yhteisöille, joilla on merkitystä yksilön hyvinvoinnille. Tämä voi tarkoittaa esimerkiksi työpaikkaa, hoitokotia ja koulua. Arvokeskustelua kaivataan sille, mitä dataa ja miten voidaan jakaa. Vaatii tasapainoilua löytää oikea suhde yksilöllisyyden suojan ja vaikuttavuuden välille.

Ekosysteemissä data voi ohjata moneen suuntaan. Hyvinvointipalveluiden tuottama data voi olla myös lääkäreiden käytössä. Työnseurantaan liittyvää dataa voidaan hyödyntää arvioitaessa työn kuormittavuuden merkitystä terveydelle. Yhdistelemällä eri ihmisten tietoja voidaan saada tietoa myös siitä, onko esimerkiksi huono sisäilma sairastavuuteen vaikuttava tekijä.

Maailma, jossa kaikki hyvinvointi- ja terveystietopalvelut olisivat täydellisesti yksilöitävissä henkilötietojen avulla, avaa futuristisia mahdollisuuksia. Mallien kehittyminen mahdollistaa yhä tehokkaamman ennakoinnin, jolloin ilmeneviin terveystietoihin voidaan puuttua paljon aikaisemmassa vaiheessa, mikä lisää terveiden elinvuosien määrää. Kustannukset voidaan pitää kurissa, kun palvelut voidaan digitalisoida ja ainakin osittain automatisoida. Alkuvaiheessa automatisointi koskee enemmän terveydentilan seuranta, mutta pidemmällä

⁹¹ <https://www.thl.fi/documents/10531/1895262/Hyvinvointisovellusten+kansalliset+hyvy%C3%A4ksymiskriteerit-02022017.pdf/c31a56d5-3d46-41f8-9c73-a0b7d438c899>

⁹² <https://techcrunch.com/2016/11/22/patient-data-api-pivotal-to-deepminds-push-into-uks-nhs/>

⁹³ <http://www.digitalhealthrevolution.fi/>

⁹⁴ Koponen et al., 2015

aikavälillä myös diagnosointia ja esimerkiksi lääkemääräyksiä. Datan pohjalta jokaiselle ihmiselle voidaan koostaa esimerkiksi ruokavalio, liikuntasuunnitelma jne., joka tukee terveyttä, mutta samalla huomioi ihmisen luontaiset preferenssit ja motivaation.

MyDatan taloudellinen vaikutus syntyy kahta eri reittiä. Ensimmäinen on taloudellinen vaikutus, joka syntyy, kun MyData kiihdyttää ekosysteemin kasvua ja tätä kautta luo liiketoimintamahdollisuuksia. Toinen vaikutus syntyy kansanterveyden paranemisen kautta. Vaikutusten arvioinnin suhteen kannattaa olla varovainen. Tämä johtuu ennen kaikkea siitä, että MyData on vain yksi osatekijä vaikutusten syntyemisessä. Mittavimmillaan arviot ovat siinä, miten aiempaa ennakoivampi ja yksilöllisempi terveydenhuolto säästää julkisia terveystuloja.

Global Market Insight⁹⁵ on arvioinut digitaalisen terveydenhuollon markkinan kasvavan 379 miljardiin Yhdysvaltain dollariin vuoteen 2020 mennessä. Kasvua syntyisi yhteensä 25 % vuodesta 2016. 60 % tästä markkinasta oli vuonna 2015 digitaalisilla terveystileillä (electronic health records). Kasvua vauhdittavat etenkin ratkaisut, jotka tukevat yhä älykkäämpää päätöksentekoa sekä potilaiden etävalvontaa. Suomen elinkeinopolitiikan kannalta oleellinen kysymys on se, millä politiikan toimenpiteillä voidaan tukea sitä, että Suomi olisi mahdollisimman kiinnostava kehitysympäristö tulevaisuuden ratkaisuille. Yksi mahdollinen keino on MyDatan edistäminen niin, että Suomi saavuttaisi edelläkävijyyden terveysdatan yhdistelemisessä ja sitä kautta luotavissa uusissa palveluissa.

Ennakoivan terveydenhuollon merkitys korostuu kustannussäästöjen kautta. Meneillään olevissa kehityshankkeissa – samoin kuin useissa edellä viitatuissa raporteissa – säästöt syntyvät etenkin sen kautta, että etäasiointi vähentää fyysisen asioinnin tarvetta. Fyysisen asioinnin tarpeen väheneminen edellyttää datan siirtymistä sekä terveystietojen tuottajien että asiakkaan välillä. Tavoiteltuja säästöjä on arvioitu tapauskohtaisesti. Suomessa vertailukohtana antaa Virtuaalisairaala-hanke. Tällä haetaan vuositasolla 40 miljoonan euron säästöjä. Kaikki säästöt eivät perustu datapohjaisiin ratkaisuihin, mutta mm. big data ja tähän perustuva koneälypohjainen neuvonta ovat osa kokonaisuutta. Tämän lisäksi asiakashyötyjä tuotetaan arjen toiminnassa auttavilla sovelluksilla. Hyödyt syntyisivät 12 miljoonan euron kertapanostuksella ja tämän jälkeen arviolta miljoonan euron vuosittaisilla ylläpitokustannuksilla. Vastaavan tason säästöjä esitetään myös kansainvälisissä selvityksissä. Konkreettisenä esimerkkinä PwC (2013) esittelee WellDocin, diabateksen hoitoon tarkoitetun sovelluksen, joka auttaa verensokerin hallitsemisessa. Sovellus toimii myös kommunikaatioväylänä lääkärin ja potilaan välillä. Potilasta kohden järjestelmän odotetaan säästävän 10 000 dollaria vuodessa. MyData mahdollistaa uusien, tämänkaltaisten sovellusten kehittämisen ja niiden käytettävyyden parantamisen.

MyDataa rajoittavat tekijät terveydenhuollossa

MyData-ekosysteemin ja hyödynnettävyyden lisääntymistä hidastavat useat tekijät. Teknisesti merkittävimpiä ovat datan yhteensopivuuteen ja siirrettävyyteen liittyvät haasteet. Nämä korostuvat terveydenhuollossa, jossa järjestelmiä ja datatyyppejä on useita. Esimerkiksi klinikkejärjestelmiä on useita samoin kuin kuvantamisjärjestelmiä. Eri järjestelmien integroiminen ei välttämättä toimi edes yhden organisaation sisällä, vaikka työtä tiedon käytettävyyden parantamiseen tehdään kaiken aikaa.

Toinen rajoite on periaatteellisempi. Kanta-järjestelmää ei kokonaisuudessaan haluta avata tai mahdollistaa tietojen jakamista muille kuin terveydenhuollon ammattilaisille. Tämän

⁹⁵ <https://www.gminsights.com/pressrelease/digital-health-market>

taustalla on osittain professionaalinen⁹⁶ lähtökohta siitä, että terveysdataa voi hyödyntää vain terveydenhuollon ammattilainen. Tämä lähtökohta on kirjoitettu myös osittain nykyiseen lainsäädäntöön. Edellä esitettyjen hyötyjen lisäksi terveysdatan vapaaseen siirrettävyyteen liittyy riskejä. Onkin tehtävä merkittävä poliittinen päätös siitä, halutaanko nämä riskit ottaa. Selvityksen yhteydessä järjestetyissä sidosryhmätyöpajoissa nousi esiin mm. seuraavia haasteita:

- Asiakkaan omalla luvalla terveysdata voi päätyä kenelle tahansa. Tämä voi johtua joko siitä, että asiakas ymmärtämättömyyttään jakaa ja antaa käyttöluvan datalle tilanteissa, joissa se ei ole hänen etunsa mukaista. Luonnollisesti tiedon leviäminen voi tapahtua myös tietovuotojen kautta, joiden riski kasvaa, kun sama data on useampien toimijoiden hallussa.
- Asiakas voi terveysdatansa avulla pyytää konsultaatiota ei-professionaaliselta taholta. Tämä siirtää riskin mahdollisista hidoista asiakkaalle itselleen. Haastattelujen valossa tätä ei pidetä haluttavana. Asia lienee noussut esiin useissa taustakeskusteluissa, vaikka haastattelujen pohjalta ei selkeästi voida osoittaa, kuka tätä kehitystä vastustaa.

Riskien suhteen nousi kuitenkin esiin se, että itse tuotettava data muodostaa reguloimattoman dataekosysteemin. Datan kasvaessa myös siihen liittyvien palveluiden määrä kasvaa. Tätä ei voida tällä hetkellä kuitenkaan säädellä.

Tutkijahaastatteluissa MyDatan ongelmaksi nähtiin se, että verifoidun, saatavilla olevan terveysdatan määrä ei vielä riitä big data -malleihin, mikä jarruttaa sovellusten luomista. Esimerkiksi ennakoivien mallien verifioimiseen menisi 30–40 vuotta. Tutkimusrahoituksessa tulisikin kohdentaa resursseja laajempiin ja pitkäkestoisempiin hankkeisiin, jotka mahdollistavat kokonaisvaltaisempien terveyttä edistävien mallien luomisen.

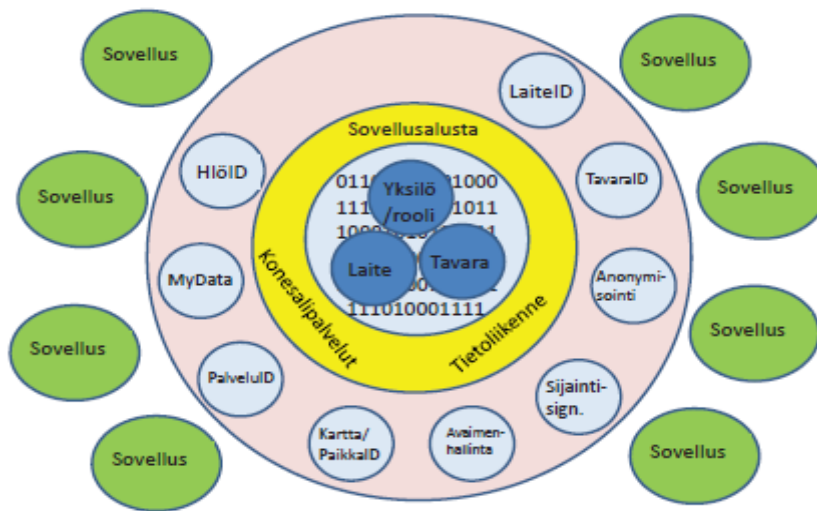
5.3 MyData liikenteessä

Liikenne on vahvasti digitalisoitumissa. Talouden uudet toimintamallit, kuten jakamis- ja alustatalous vaikuttavat myös liikenteeseen. Über on tunnetuimpia esimerkkejä alustaloudesta. Jakamistaloutta taas edustavat esimerkiksi erilaiset kimppekyyti- ja autonjakamispalvelut. Näitä on olemassa sekä sovelluksiin perustuvia että esimerkiksi sosiaalisessa mediassa toimivia ryhmiä. Tämän lisäksi markkinoille on tullut huomattava määrä palveluita – kuten Waze – jotka keräävät käyttäjiltä dataa ja luovat sitä hyödyntäen käyttäjille reittisuosituksia ja ottavat huomioon ruuhkat, tietyöt yms. muut liikennettä haittaavat tekijät. Vastaavat palvelut hyödyntävät myös avointa dataa. Digitalisaation seurauksena maksaminen, liput yms. ovat muuttuneet sähköisiksi. Tämän seurauksena ihmisestä jää yhä laajempi digitaalinen jalanjälki.

Nämä eri kehityssuunnat ovat johtaneet siihen, että datan määrä on liikenteen osalta eksponentiaalisessa kasvussa, samoin kuin sen hyödyntäminen. MyDatan näkökulmasta kehitys on kuitenkin vielä aluillaan. Oleellista on nimenomaan tietojen siirrettävyys järjestelmien välillä, mikä mahdollistaa yhä kohdennetummat ja yksilöllisemmät palvelut. Siirrettävyys mahdollistaa kuluttajaa myös kilpailuttamaan nykyistä tehokkaammin palveluntarjoajat.

⁹⁶ Professionit ovat työ sosiologian käsite siitä, kenellä on oikeus harjoittaa tiettyä professiota/ammattia tai toimia asiantuntijana. Lääkäri on malliesimerkki vahvasta professiosta, koska on suuri määrä asioita, jotka vain lääkäri saa tehdä ja lääkärin pätevyysvaatimukset on tiukasti määritelty. Ks. esim. Pirttilä & Eriksson, 2015

Kuva 5.2 Liikennetiedon ekosysteemi



Lähde: Linturi ja Kuittinen 2014⁹⁷

Liikenneviraston selvityksessä Liikennetiedon visiot⁹⁸ on arvioitu eri datatyyppien merkitystä liikenne-ekosysteemille. MyDatan arvo ekosysteemille syntyy yhdessä muiden datatyyppien kanssa ja näitä yhdistelemällä. Datan hyödyt syntyvät sitä yhdistelmällä sekä rakentamalla sen päälle sovelluksia. Näiden sovellusten määrä on kasvussa; käytetyimpiä ovat reittioppaat, ruuhkatiedotteet jne. Liikenteessä henkilödatan arvo korostuu tällä hetkellä mm. vakuutuksissa, joita kehitetään yhä enemmän älyvakuutustyyliksi, eli vakuutuksen hintaan vaikuttaa henkilön oma käyttäytyminen. Liikennevakuuttamisella on tässä pitkät perinteet, koska ajohistoria on vaikuttanut hinnoitteluun tähänkin asti. Datan siirrettävyys ja yhdisteltävyys mahdollistaa kuitenkin uusien tuotteiden luomisen.

Liikkuminen palveluna (Mobility as a Service, MaaS) on uusi toimintatapa, joka on kaikki liikennemuodot kattava sujuva ja saumaton liikennejärjestelmä ja mahdollistaa ”ovelta ovelle” -palvelutason ilman pakottavaa tarvetta omistaa yksityisautoa. MaaS on Suomessa kehitetty ja kansallisen liikennepolitiikan keskiössä. Niin yritysten kuin viranomaisten kehitystyö ja toimet MaaS:n saralla ovat globaalissakin mittakaavassa edelläkävijöitä.

Käyttäjänäkökulmasta asiakkaat hankkivat liikkumispalvelunsa joko yhden mobiilialustan ja palveluntarjoajan kautta tai erinäisiä helppokäyttöisiä palvelusovelluksia yhdistelemällä (esimerkiksi helpot mobiilivaraus- ja maksujärjestelmät sekä ajantasaiset matkatietopalvelut). Liikennemuotoja hyödynnetään ja yhdistetään liikkumispalveluiksi yli perinteisten raja-aitojen. Tieto liikkuu liikkujan ja logistiikan matkassa yli liikennemuotovaihdosten. Yritysten logistiikkapalvelut tehostuvat ja paranevat.

MaaS-palveluiden tuottaminen perustuu pitkälti yhä laajenevaan käyttäjätiedon hyödyntämiseen. Jo nyt kulkuneuvojen ja käyttäjien reaaliaikaiset sijaintitiedot ovat oletusarvona liikkumispalveluiden tuottamiseen. MaaS-konseptin osalta MyDatan tarve tulee kasvamaan yhä enemmän etenkin personoidumpien palveluiden tuottamisessa (esim. jaetut, kutsupohjaiset kuljetuspalvelut) kuin myös henkilön historiatietojen siirrettävyydessä

⁹⁷ Kuvan tekijät (Linturi ja Kuittinen, 2014) antoivat luvan julkaista kuvan tässä yhteydessä CC-BY 4.0 lisenssin pohjalta. Tekijänoikeudet kuuluvat heille.

⁹⁸ Linturi & Kuittinen, 2014

(esim. palveluoperaattorien kilpailuttaminen). Liikennesektorilla useat lipputuotteet ovat myös kytköksissä henkilötietoihin. Eri alennusliput joukkoliikenteessä sekä yhteiskunnan tukemat kuljetukset (esim. sairaskuljetukset) edellyttävät, että MyData-periaatteita sovelletaan yhä enenevässä määrin myös liikennesektorilla.

MyDataan perustuvilla ratkaisuilla voidaan mahdollistaa esimerkiksi liikkumishistoriaan, palvelujen käyttöön, asiointiin tai transaktioihin liittyvien tietojen siirtämistä järjestelmästä toiseen tai pääsyä näihin tietoihin erilaisten matkaketjujen toteuttamiseksi. Tämän lisäksi henkilötietojen joustava jakaminen mahdollistaa mm. kyydin jakamisen (eri ihmisten reittejä voidaan verrata automaattisesti) tai esimerkiksi ajo-oikeuksien hallinnan (ihminen voi vuokrata vain ajoneuvon, johon hänellä on ajo-oikeus) jne.

Suomessa varautuminen liikenteen murrokseen näkyy lainsäädännön valmistelussa. [Liikenteen palveluista annettu laki](#) (320/2017, hallituksen esityksessä laki oli nimeltään liikennekaari) on lakihanke, jolla pyritään edistämään tiedon hyödyntämistä uusien liikennepalveluiden synnyttämisessä. Sen ensimmäinen vaihe astuu voimaan 1.7.2018, mutta tiedon hyödyntämistä koskevat säännökset tulevat voimaan jo 1.1.2018. Lakihankkeen käynnissä olevaan toiseen vaiheeseen sisältyy ehdotuksena säännöksiä, jotka pyrkivät edistämään henkilön omien tietojen käyttöä. Lain toisessa vaiheessa ehdotettaneen muun muassa liikenteen rekistereiden kokonaisuudistusta, jonka avulla mahdollistettaisiin rekistereissä olevan tietopääoman hyödyntäminen uudella tavalla.

MyDatassa liikenne- ja viestintäministeriön rooli on ollut edistää kehitystä ja yhteistoimintaa mm. toteuttamalla MyData-konseptipaperin vuonna 2014 ([linkki](#)), käynnistämällä Mydata-allianssin toimintaa, osallistumalla kansainvälisesti mittavan MyData-konferenssitoiminnan käynnistymiseen, toteuttamalla MyData-pilotteja, kuten Maankoodauskurssin tuloksena toteutetun MyData-operaattorimallin ja sovelluksen liikenteen hiilijalanjäljen seurantaan henkilötasolla sekä asettamalla MyDatan kehittämisen tavoitteita virastoilleen seuraaviksi vuosiksi. Esimerkiksi vuonna 2016 Mydata-konferenssin yhteydessä pidetyn hackaton-kilpailun voittajaksi nousi Mobility Profile sovellus, joka yhdistää ihmisen kalenteri- ja liikkuvuustietoja ja antaa tämän pohjalta reittisuosituksia.

5.4 Virkamiesten roolin muutos tapaustutkimusten pohjalta

Tapaustutkimusten pohjalta ilmenee tekijöitä, jotka muuttavat virkamiesten ja julkisen sektorin roolia datan hallinnassa ja hyödyntämisessä. Todennäköistä on, että uudesta roolista tulee vuorovaikutteisempi. Esimerkiksi osa viranomaisten rooleista tulee siirtymään kansalaisille tai yrityksille. Tätä siirtymistä tapahtuu joka tapauksessa – esimerkiksi ihminen voi jakaa omaa henkilötietoaan erilaisten digitaalisten alustojen kautta ilman, että viranomaiset pystyvät tähän vaikuttamaan. Samaan aikaan kuitenkin mm. VTJ:n tietopyyntöjä rajoitetaan lainsäädännöllä. Näin ollen herää kysymys, onko VTJ:n rajoitteilla enää käytännön merkitystä ja toisaalta, pitäisikö viranomaisilla olla muitakin keinoja yksityisyyden suojan tukemiseen.

Yksityisyyden hallinta ja tietoturva: MyDatan yhtenä keskeisenä haasteena on noussut esiin se, että henkilötiedon jakaminen vaarantaa ihmisen yksityisyyden turvan – vapaasti levitetynä tietoja esimerkiksi terveydentilasta tai etnisyydestä saatetaan käyttää ihmistä haittaavalla tavalla. Haastetta lisää se, että väärinkäytös voi johtua myös tiedon hallinnoijan tietoturvan puutteellisuudesta. Näin ollen data voi levitä myös tietovuotojen seurauksensa ja tästä voi seurata henkilölle haitallista tiedon käyttöä.

Ongelmaa hallitaan nyt rajoittamalla tiedon siirrettävyyttä sekä sillä, että VRK:lla on valtuudet tarkastaa tiedon vastaanottajan tietoturvajärjestelmät. Uudessa roolissa viranomaisen tehtävänä tulisi olla enemmän valistus sekä toimijoiden sertifiointi; yritykset voisivat hakea julkisia sertifikaatteja, joilla ne todentavat tietoturvansa täyttävän minimivaatimukset. Sertifioija voi olla yksityinen toimija ja todennäköisesti sertifikaatille täytyisi luoda kansainväliset standardit. Viranomaisten rooliksi jäisi valvoa sertifiointia.

Henkilön heikko neuvotteluvoima nähdään myös syyksi rajoittaa MyDataa. Jos tiedon siirtäminen on mahdollista, tätä voidaan alkaa edellyttää tilanteissa, joissa näin ei saisi tapahtua. Esimerkiksi rekryointitapahtumissa voidaan alkaa vaatia tietoa terveystietojensa. Teoriassa tätä voidaan perustella sillä, että työhöntulotarkastus voidaan korvata olemassa olevilla terveystiedoilla, jos ne ovat ajantasaisia. Riskinä on kuitenkin se, että lopulliseen päätökseen rekryoinnista vaikuttavat irrelevantit tiedot – kuten jo hoidetut terveysongelmat tai ongelmat, jotka eivät ole haettavan työn kannalta aidosti merkityksellisiä.

Tätä riskiä voidaan hallinnoida luomalla standardisopimuksia siitä, mitä tietoa voidaan missäkin tilanteessa pyytää ja mihin sitä saa käyttää. Haasteena on toki määrittää se, mitä seurauksia sopimuksen noudattamatta jättämisestä olisi ja miten tätä käytännössä valvottaisiin. Riskinä olisi, että valvonta vaatisi kohtuuttomasti resursseja.

Rekisterien ylläpitäminen on tällä hetkellä viranomaistoiminnan kannalta keskeinen tukitehtävä. Ilman kattavia tietoaaineistoja viranomaiset eivät voisi hoitaa muita tehtäviään. MyData ja avoin data lisäävät kuitenkin mahdollisuuksia, että rekisterin ylläpitäjä voisi olla jokin muukin taho kuin viranomainen. Viranomaisten ei tarvitse muodostaa rekisteriä, jos julkisissa palveluissa tarvittavat tiedot on luotettavasti saatavilla muista lähteistä. Näköpiirissä ei tällä hetkellä ole se, että keskeiset rekisterit yksityistettäisiin. Sen sijaan olisi esimerkiksi pohdittava sitä, voisiko yksityisille rekisteritoimijoille kuten vaikka Postille antaa valtuudet tarjota samoja palveluita kuin VTJ tai laajentaa rekisteriaineistojaan samaan suuntaan.

Toinen kysymys on esimerkiksi se, voidaanko julkisia rekistereitä tehokkaammin täydentää yksityisillä aineistolla ja mitä tämä vaatisi. Esimerkiksi yksityisille kursseille ja esimerkiksi ICT-yhtiöiden sertifikaateille voisi rakentaa datapankin, joka täydentäisi Koskea, tai vakuutusyhtiöiden dataa voidaan käyttää ajo-oikeuksia myönnettäessä jne. Tämä edellyttäisi tietoturvan kaltaista sertifiointikäytäntöä, jolla viranomaiset tunnistaisivat luotettavia rekistereitä.

Konsultatiivisen roolin merkitys viranomaistyössä kasvaa. Tämä korostuu etenkin MyDatassa. MyData helpottaa henkilötiedon saatavuutta, mutta samalla laindäädäntö kiristyy sen suhteen, miten henkilötietoja saa käyttää ja tallentaa. Se, että ihminen luovuttaa tietonsa ei automaattisesti tarkoita sitä, että yritys saa tallentaa tiedon. Jo tällä hetkellä datan uudelleenkäyttö aiheuttaa yrityksissä vaikeuksia⁹⁹. Tämän takia viranomaisten tulisi ottaa aktiivisempi rooli tiedon hyödyntämisen mahdollistamisessa ja reunaehdoissa. Neuvonta voisi myös rohkaista yrityksiä hyödyntämään uusia mahdollisuuksia ja vastaavasti jakamaan omaa dataansa.

Viranomaisten mahdollistavasta roolista on olemassa jo useita malleja. Tätä edustaa edellä mainittu Liikennelabradori. Toinen hyvä esimerkki on Finanssivalvonnan innovaatio help-desk.

¹⁰⁰ Tämä on Pohjoismaiden ensimmäinen help-desk, mutta muualla maailmassa vastaavaa toimintaa on ollut aikaisemmin. Neuvonnan tarpeen on luonut finanssialalla FinTech-

⁹⁹ Nissilä, Kokkonen, Kuittinen, & Others, 2016

¹⁰⁰ Finanssivalvonta, 2016

ekosysteemin synty. Toimialalla on paljon sääntelyä eikä ole itsestään selvää, minkä sääntelyn piiriin uudet liiketoimintamallit kuuluvat ja mitä lupia niihin tarvitaan. Näihin haasteisiin help-desk pyrkii vastaamaan.

Viranomaistoiminnassa on huomioitava myös liiketoiminnan kansainvälistyminen ja ihmisten liikkuvuuden lisääntyminen. Näin ollen datan hallittavuudessa ja mahdollisessa sääntelyssä pitäisi huomioida kansainvälinen aspekti koko ajan. Esimerkiksi terveysdatan osalta pyritään Kanta-järjestelmän kautta mahdollistamaan datan käytettävyys Suomen sisällä kaikille terveystoimijoille. Tämä ei vielä huomioi mm. terveysturismia tai sitä, että ihminen saattaa joko muuttaa toiseen maahan tai asua useassa maassa. Vastaavasti kaupankäynnin kansainvälistyminen edellyttää kansainvälisesti yhteensopivia henkilötietojärjestelmiä.

Vastaavasti ohjelmistokehitys on kansainvälistä ja ohjelmistokehittäjät tuskin pitävät yksin Suomea kiinnostavana markkina-alueena. Näin ollen on rakennettava pelisäännöt ja toimintamalli kansainväliselle yhteistyölle.

Viranomaisten työnjakoa kannattaa kehittää. Datan hallinta ja siihen liittyvät kysymykset tulevat muuttamaan viranomaisten työnjakoa. Esimerkiksi datan vapaampi jakaminen luo tarvetta luoda uusia ratkaisuja tietoturvan valvontaan. Vastaavasti standardisopimukset kannattaa luoda keskitetysti jo senkin takia, että ne eivät olisi keskenään ristiriitaisia. Tietosuojakysymyksissä tietosuojavaltuutetun rooli on merkittävä, kuluttajan oikeuksien valvonnassa taas kilpailu- ja kuluttajavirastolla on keskeinen rooli. Oleellista on kuitenkin löytää roolitus, joka vastaa muuttuvaan tilanteeseen ja mahdollistaa koordinoitun toiminnan.

6 VASTUULLISEN TIEDON HYÖDYNTÄMISEN TUKEMINEN

Aleksi Knuutila

Ihmiskeskeinen henkilötiedon hyödyntäminen edellyttää uudenlaisten suhteiden syntymistä tietoa hallinnoivien viranomaisten, yksilöiden sekä tietoa käyttävien yksityisten ja julkisten palveluntarjoajien välille. Se vaatii uudenlaisia tehtäviä julkishallinnolle, jotta tämä voisi edesauttaa luottamuksen syntymistä, vastuullista toimintaa sekä käyttäjien kannalta selkeän ja avoimen järjestelmän syntymistä.

Tässä kappaleessa esitämme periaatteiden tasolla joitakin toimintamalleja, jotka vastaavat tiedon luovutuksen yhteydessä usein esiintyviin haasteisiin. Ensimmäisen osio kuvaa viranomaisten ja sovelluksia tuottavien palveluntarjoajien välistä suhdetta sekä siihen liittyviä tarkistamisen ja valvonnan käytäntöjä. Sen jälkeen kuvaamme tapoja, joilla käyttäjän ja sovelluksen välisen tiedon käyttöä koskevan sopimisen voi tehdä selkeäksi ja ymmärrettäväksi. Lopulta käsittelemme lyhyesti yhteentoimivuutta ja avoimia standardeja.

6.1 Sovellusten rekisteröityminen ja tarkistaminen

Yleiset mallit käyttöehdoille

Henkilötiedon jakamisessa MyData-operaattorin kaltaisen palvelun kautta on useita osapuolia: rekisterinpitäjä sekä operaattori, sovelluksen kehittänyt palveluntarjoaja sekä käyttäjä. Aloitteen tiedon siirrossa tekee käyttäjä, joka haluaa esimerkiksi tehdä sopimuksen ottaakseen palvelun käyttöönsä tai haluaa lähettää sille tietoa. Kaikkien osapuolien välisille suhteille voisi kehittää mallit ehdoista, joiden puitteissa tiedon jakaminen tapahtuu. Hyvä pohja käyttöehdoille voi selkeyttää käyttäjien valintoja, helpottaa uusien sovellusten kehittämistä sekä tukea vastuullisia tiedon käyttötapoja.

Julkishallinnossa on aikaisemminkin hahmoteltu yleisiä käyttöehtoja esimerkiksi perusrekisterien sisältöjä varten.¹⁰¹ Toimiva järjestely voisi olla, että MyData:n välittämistä kehittävä taho sopii yksittäisten rekisterinpitäjien kanssa sellaisella tavalla, että se pystyy tarjoamaan sovelluksille enemmän tai vähemmän yhdenmukaisten käyttöehtojen pohjalta siirtoja useammista rekistereistä. Lisäksi käyttäjän pitäisi mahdollisesti hyväksyä käyttöehdot käyttäessään MyData-tilin kaltaista palvelua ensimmäistä kertaa. Järjestelyn tulisi tehdä selkeäksi, kuinka vastuut jakautuvat ja mahdollistaa se, että käyttäjä voi päättää tiedon jakamisesta ilman merkittäviä julkishallintoon osuvia velvoitteita tai riskejä.

Käyttäjän ja sovelluksen välisen sopimisen sisältö voi vaihdella, mutta MyData-operaattori voisi tarjota tätä varten muuteltavia malleja. Yhdet käyttöehdot voisivat toimia pohjana sovelluksille, jotka tarvitsevat vain ajantasaiset yhteystiedot, kun taas omistuksiin liittyvien tietojen käsittely olisi rajoitetumpaa. Käyttöehdoissa voisi myös olla joitakin yleisiä edellytyksiä, kuten esimerkiksi tarkka rajausta sille, milloin tietoa jaetaan kolmansille osapuolille ja vaatimus sopimuksen helposta päättämisestä. Tiedon käsittely voi pohjata joko käyttöehdoilla määriteltyyn sopimukseen tai suostumukseen, joilla on erilaiset lailliset edellytykset.

¹⁰¹ Valtiovarainministeriö, 2011

Riskit ja sovellusten tarkistamisen tarve

Monille toimialoille osallistuvia valvotaan turvallisuuden ja yhteentoimivuuden varmistamiseksi. Valvonnan läpäisemiseen liittyvät kustannukset voivat olla pienille toimijoille merkittävä kynnys, sillä esimerkiksi yksin tietoturvan auditoinnin kustannukset voivat olla kymmenen tuhannen euron luokkaa.¹⁰² Luonnollisesti järjestelmän hyödyllisyyden kannalta kynnys kokeiluille ja käyttöönololle on hyvä pitää mahdollisimman pienenä. Valvonta voi olla vaativaa myös viranomaisille.

Valvonnan ja tarkistusten määrän on parasta vastata tiedon jakamiseen liittyviä riskejä. Vakavat riskit, joiden toteutumisella on merkittävä todennäköisyys, eivät ole hyväksyttäviä. Esimerkiksi THL on hahmotellut hyvinvointisovellusten hyväksymiskriteerejä myös tapauksille, joissa sovelluksen käyttäisivät Omakanta-palvelun asiakas- ja potilastietoja.¹⁰³ Terveyspalveluiden riskeihin lukeutuvat mahdolliset vääränlaiset neuvot. Yleisempiä riskejä ovat tietosuojaan liittyvät riskit, käyttötarkoituksen suhteen tiedon liian suuren määrän kerääminen, tietojen kulkeutuminen väärille osapuolille sekä tietoturvasuhteeseen liittyvät riskit kuten tietovuodot.

Vähemmän arkaluontoisen tiedon kohdalla yleiset riskit ovat vähemmän merkittäviä. MyDataan liittyvien kokeilujen olisikin hyvä lähteä liikkeelle tiedon tyypeistä, kuten esimerkiksi ajokorttirekisteristä, jotka eivät ole arkaluonteisia. Näiden kohdalla tiedon käyttäjien kevyempi valvonta ja sovellusten tarkistus ovat perusteltua.

Kevyet keinot sovellusten tarkistamiseen

Joillakin ohjelmistoja tarjoavilla alustoilla sovelluksia ei tarkisteta vahvasti etukäteen. Sen sijaan ohjelmien käyttöön liittyvät riskit minimoidaan toimintatapojen kautta ja käyttäjille annetaan tietoa, joka auttaa heitä tekemään hyviä valintoja. Jotkin henkilötietoja jakavat alustat, kuten esimerkiksi Android-puhelinten Google Play -alusta, toimivat tällä tavalla.

Vähemmän arkaluontoisia henkilötietoja käyttäville sovelluksille kevyt tarkastus voisi riittää myös silloin, kun nämä liittyvät MyData-operaattorin kaltaiseen palveluun. Palveluntarjoajilta edellytettäisiin palvelukuvausta, johon kuuluu pieni määrä välttämätöntä tietoa rakenteellisessa muodossa. Tähän voisi sisältyä paitsi tarvittavan tiedon kuvaus, myös kuvaus siitä, mihin tietoa käytetään, kuinka sitä talletetaan ja jaetaan ja kuinka palvelusta voi erota ja tiedot poistaa. Operaattoriin liittyminen voisi tapahtua nopeasti, kun viranomaiset ovat käyneet läpi palvelukuvauksen asiallisuuden.

Sovellukset, jotka vapaaehtoisesti kuvaavat toimintaansa tarkemmin ja esimerkiksi tarjoavat lokitietoja omasta järjestelmästä, voisivat saada laatumerkinnän (esimerkiksi "reilun datan toimijan" sertifikaatin), joka lisäisi käyttäjien luottamusta. Operaattorin vastuulla olisi myös poistaa sovellukset, joiden kohdalla väärinkäytökset ovat olleet ilmeisiä. Laatumerkinnät ja muuta asiayhteyteen liittyvää tietoa voi selkeästi välittää käyttäjille, kuten alla selitämme.

6.2 Selkeiden tiedon jakamiseen liittyvien valintojen tukeminen

Tietojen jakamista koskevia valintoja tehdään, kun jokin palvelu otetaan käyttöön tai tietoa halutaan jakaa muusta syystä. Lisäksi käyttäjillä on mahdollisuus jälkikäteen muuttaa

¹⁰² Lähteenä haastattelumme.

¹⁰³ <https://www.thl.fi/web/tiedonhallinta-sosiaali-ja-terveysalalla/ajankohtaista/lausuttavat-materiaalit/omatietovarannon-vaatimusten-komentointi>

käyttölupiin liittyviä asetuksiaan MyData-tilin kautta. Tässä osiossa käymme läpi joitakin palvelumuotoiluun ja standardoituun sopimiseen liittyviä periaatteita, joiden pohjalta henkilötiedon luvittaminen ja jakaminen voi tapahtua selkeällä ja käyttäjäystävällisellä tavalla.

Keskitetyn näkymän edut

Henkilötiedon hallinnan pohjoismaisessa mallissa (ks. luku 4) ihmisiä koskeva tieto sijaitsee hajautetusti useassa lähteessä. Hajautettu järjestely ei sulje pois sitä mahdollisuutta, että tietojen jakamista voisi hallita keskitetysti yhdestä paikasta. Käytettävyyden kannalta yksittäinen palvelu, joka tukee kattavasti oman tiedon hallintaa, olisi paras vaihtoehto.¹⁰⁴

Julkishallinnon kohdalla Suomi.fi:n sähköisen asioinnin palvelut voisivat olla palveluperhe, johon julkishallinnon tiedon jakamisen yleisnäkymän voisi sijoittaa. Palveluun voisi liittyä muun muassa seuraavia ominaisuuksia:

- Sen kautta voisi käydä läpi antamiaan käyttölupia, muuttaa niiden yksityiskohtia tai kumota niitä.
- Omia tietoja voisi tarkastella ja korjata.
- Palvelussa voisi olla "tieto-tiliotteen" kaltainen yleiskatsaus siihen, kuinka omaa tietoa on käytetty. Mahdollisuuksien mukaan tämä näkymä voisi sisältää myös kuvauksen siitä, kuinka omaa tietoa on käytetty julkishallinnon sisällä sekä toisiokäytössä esimerkiksi tutkimuksessa.
- Lähitulevaisuudessa todennäköisesti syntyy useampi järjestelmä tiedon jakamiseen, mukaan lukien julkisia toimialakohtaisia järjestelmiä kuten Kanta ja Koski. Jos suostumusten mallintamisen järjestelmät olisivat riittävän kehittyneitä, voisivat nämäkin järjestelmät keskustella keskenään. Tällöin keskitetystä Suomi.fi-näkymästä voisi myös nähdä esimerkiksi omien terveystietojen jakamiseen liittyvät valinnat.

MyData osana käyttäjäpolkuja ja palveluketjuja

Käytettävyyden kannalta on paras, jos tunnistautuminen ja tiedon siirrosta sopiminen voi tapahtua osana käyttäjien muuta asiointia ja askeleena muiden palveluiden suunnittelemissa käyttäjäpoluissa. MyData-tiliin itseensä tarvitsisi kirjautua vain, jos ihminen haluaa tarkistaa päätöksensä tiedon jakamisesta tai esimerkiksi hakea omia tietojaan omaan käyttöönsä.

Esimerkiksi moniin sovelluksiin, joko matkapuhelimien tai verkkosivujen kautta toimiviin, voi kirjautua käyttäen suurten pilvipalveluiden tunnuksia (esimerkiksi Google ja Facebook). Pilvipalvelut toimivat tällöin identiteetin sekä muun tunnuksiin liittyvän tiedon tarjoajina. Käyttäjien ei tarvitse erikseen siirtyä identiteetin tarjoajien palveluihin mahdollistaakseen tiedon vaihdon. Osana toisten sovellusten käyttöönottoa avautuu sivu, jonka kautta voi tunnistautua ja liittää identiteetin tarjoajan toisiin palveluihin. Ihanteellisesti MyDatan saisi luvitettua palveluiden käyttöön myös muiden kuin sähköisesti toimivien palveluiden kohdalla. Tällöin tietojen käyttöön liittyvien valintojen ilmaisuus voisi tapahtua osana niitä tavanomaisia keinoja, joilla ihmiset liittyvät asiakkaiksi (kuten esimerkiksi lomakkeiden täyttäminen).

¹⁰⁴ Tästä huolimatta joissakin tapauksissa käyttäjät saattavat itse valita esimerkiksi useamman profiilin ylläpitämisen. Monet tekevät näin suojatakseen omaa yksityisyyttään profiileilla, joita ei ole yhdistetty heidän julkiseen henkilöllisyyteensä.

Tulevaisuudessa myös esimerkiksi omien tietojen korjaaminen tai viestien vastaanottaminen voisi tapahtua osana muuta asiointia. Liittäessään omia rekisteritietojaan osaksi palvelua, käyttäjä voisi samalla tarkistaa, että ne ovat oikein. Jos esimerkiksi passi on vanhentumassa, tästäkin voisi samalla saada varoituksen.

Tunnistautuminen ja käyttöoikeuksien jakaminen osana muita palveluja luo uudenlaisia riskejä tietojen kalasteluun. Toteutusta suunniteltaessa pitäisi varmistaa, että käyttäjä voisi olla varma siitä, että hän tunnistautuu oikeaan järjestelmään ja ettei siinä, mitä tietoja siirretään, ole epäselvyyksiä.¹⁰⁵

Ymmärrettävät valinnat kun sovelluksia otetaan käyttöön

Tällä hetkellä uusia palveluita käyttöönottaessaan ihmisille jää usein epäselväksi, mitä tietoa heistä kerätään ja mihin tarkoitukseen. Parhaimmassa tapauksessa käyttäjien tulisi pystyä ilmaisemaan aikomuksensa selkeästi ja halutessaan valita yksityiskohtaisesta joukosta vaihtoehtoja. Suostumus sekä sopimuksien hyväksyminen edellyttävät jo lain mukaankin tiedon perusteella toimimista. Toisaalta liiallinen kuormitus tiedolla ja valinnoilla johtaa siihen, että käyttäjät hyväksyvät tiedon jakamisen perehtymättä siihen, kuten usein käy nykyisten käyttöehtojen kanssa.

MyData-operaattorin kaltaisten palvelujen toiminta kannattaa suunnitella alusta asti niin, että käyttäjät ymmärtävät omat valintansa ja niiden seuraukset. Mahdollisia ratkaisuja ovat yhteisesti sovitut muodot, joilla tiedon käyttöön liittyviä sopimuksia ja ehtoja voi kuvata sekä ymmärrettävät, mahdollisesti visuaaliset yhteenvedot niistä.

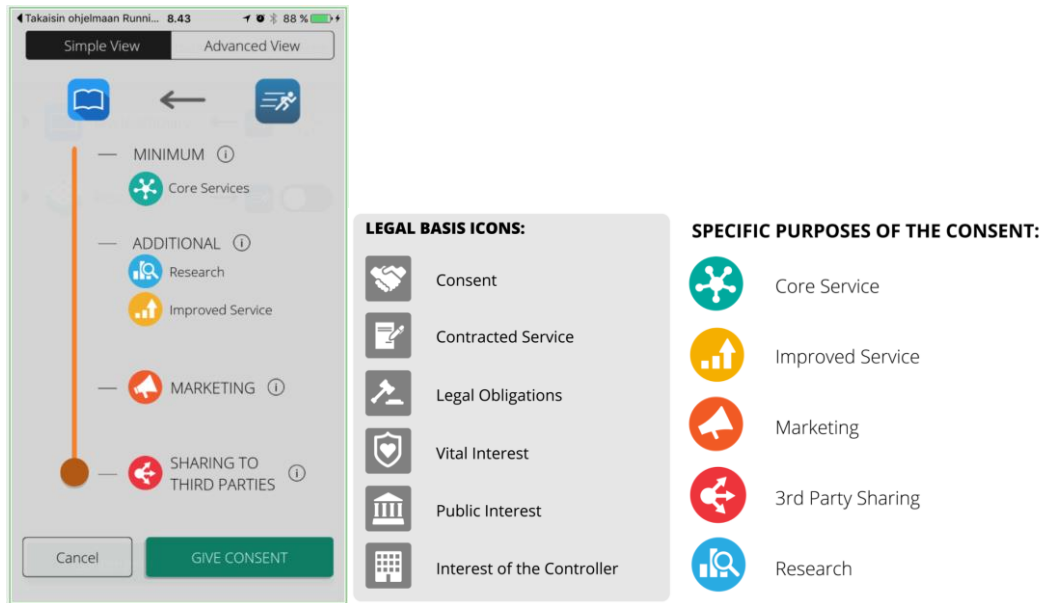
Standardisoituja malleja sopimuksista on jo käytetty esimerkiksi rahoitusallalla, jossa lainan tarjoajan on kerrottava tietyt tiedot.¹⁰⁶ Ne ovat keino suojata kuluttajia epäreiluilta sopimusten ehdoilta vähentämällä tilanteisiin usein liittyvää tiedon ja ymmärryksen epätasapainoa.

Kun sopimusten yleisistä muodoista on sovittu, ne on myös helpompi esittää selkeinä yhteenvedoina. Tähän tarkoitukseen on kehitetty myös useita ikonografisia keinoja, jotka mahdollistavat valintojen esittämisen ytimekkäästi ja visuaalisesti. Tutkimusten mukaan ihmiset ymmärtävät kuvakkeiden muodossa esitettyjä valintoja hyvin. Toisaalta tällaisilla yhteenvedoilla on välttämättä omat rajansa sen suhteen, millaisia sopimuksen sisältöjä ne voivat kuvata. Tiivistelmän yhteyteen on hyvä myös tarjota mahdollisuus nähdä palvelun laajempi kuvaus.

¹⁰⁵ Esimerkiksi Googlen käyttäjien tietoja on kalasteltu luomalla palveluja, joiden nimet näyttävät täsmälleen samalta kuin Googlen omien palveluiden. Tällöin käyttäjät eivät ole ymmärtäneet, mihin he ovat siirtämässä omia tietojaan. Lähde: <http://www.pcworld.com/article/3194816/security/google-docs-phishing-attack-undermines-oauth-security-risks.html>

¹⁰⁶ Edwards & Abel, 2014

Kuva 6.1. Ikoneita suostumuksen antamiseen



Suomalaisen tutkimusryhmän kehittämä joukko ikoneita ja käyttöliittymä, jonka pohjalta suostumuksen antamista voi tehdä ymmärrettävämmäksi. Lähde: Helsinki Institute for Information Technology¹⁰⁷

Tiedon jakamisen alueella standardisointia ja kuvakkeita on kehitetty jo pitkään.¹⁰⁸ Toistaiseksi yritysmaailman yritykset löytää yhteisiä malleja eivät ole onnistuneet. Tämä on siis alue, jossa julkishallinnon aloite voisi auttaa, mahdollisesti myös Euroopan Unionin tasolla.

Palvelun tietosuojan ja tietoturvan laadun kuvaus

Joidenkin asiantuntijoiden mielestä käyttäjän kannalta merkityksellisin tieto liittyisi nimenomaan tiedon jakamisen seuraukseen ja riskeihin, eikä pelkästään itse tiedon sisältöön ja käyttötarkoituksiin.¹⁰⁹ Liitettäessä sovelluksia osaksi MyData-tiliä tämä voisi tarjota monipuolisesti arvioita palvelujen laadusta ja vertailutietoa. Joitakin mahdollisuuksia ovat:

- **Laatumerkinnät.** Sovelluksille voi tarjota joukon kriteerejä, joiden suhteen niitä arvioidaan, joko skaalalla (esimerkiksi yhdestä viiteen tähteä) tai minimivaatimuksilla. Osallistuminen tällaiseen arviointiin olisi vapaaehtoista. Kriteerit voivat sisältää muun muassa palvelukuvauksen selkeyttä, turvallisuutta, tietosuojan mukaista datan minimointia. Sertifikaatit tai tähtiluokitus ja niiden selitys näytetään selkeästi käyttäjälle tämän kytkiessä sovellusta MyData-tiliin.
- **Mainejärjestelmä.** Sovellusten käyttäjät voisivat yhdessä arvioida sovelluksia niiden käyttökokemuksen, selkeyden ja turvallisuuden suhteen. Uuden käyttäjän valitessa sovellusta tämä näkisi esimerkiksi keskiarvon muiden käyttäjien antamasta arvosanasta sekä näiden jättämää palautetta.

¹⁰⁷ <https://github.com/HIIT/mydata-sdk/tree/master/graphics>

¹⁰⁸ Esimerkiksi vaikutusvaltaisen World Wide Web Consortiumin Platform for Privacy Preferences -hanke aloitettiin vuonna 2002.

¹⁰⁹ Ks. esimerkiksi Weigend, 2017

- Sovellusten välinen vertailu ja varoitukset. Sovelluksia voi luokitella esimerkiksi sen mukaan, mitä tarkoitusta varten ne hakevat tietoja (kuten yhteystietojen päivittämiseen tai käyttäjien profilointiin). Palveluja kuvattaessa voisi näyttää, vaatiiko sovellus enemmän tietoa kuin muut saman tyyppiset sovellukset. Silloin kun palvelun tiedon tarve näyttää tarpeettoman suurelta, tästä voisi antaa selkeän varoituksen, ja joissain tapauksissa tiedon jakamisen sovellukseen voisi estää.

Helppojen valintojen tukeminen: Yksityisyysapulaiset ja sovellukset tiedon hallintaan

Usein hyvä periaate sovellusten suunnittelussa on, että käyttäjille esitetään vain oleellisin tieto ja valinnat. Tarkemmat asetukset hoituvat usein oletusarvojen pohjalta. On myös mahdollista käyttää hyväkseen erilaisia ohjelmia tai jopa tekoälyjä helpottamaan omia valintoja.

Niin sanottuja yksityisyysapulaisia, ohjelmia, jotka helpottavat yksityisyyteen liittyviä valintoja, on kehitetty esimerkiksi matkapuhelimille. Carnegie Mellon -yliopistossa kehitetty Privacy Assistant kysyy joukon kysymyksiä siitä, millaisen tiedon jakamiseen käyttäjä on valmis, ja muokkaa sovellusten yksityisyysasetuksia vastaavasti. Sovellukset voivat siis käyttää erilaisia keinoja ihmisten mieltymysten mallintamiseen ja ehdottaa muutoksia tiedon jakamiseen. Järjestelmien kehittyttyä voikin olla, että valtuutamme erilaisia ohjelmia hallitsemaan tietovirtaamme ja käytämme niitä muokkaamaan tiedon siirtymistä hetkestä toiseen tarpeittemme mukaan.¹¹⁰

6.3 Avoimet standardit ja yhteentoimivuus

Avoimet standardit

Yleisesti käytössä olevat ja avoimet standardit ovat edellytys helposti toteutettavalle palvelujen väliselle yhteentoimivuudelle sekä sille, että palveluverkosto toimii suuremmalla skaalalla. Avoimilla standardeilla tarkoitetaan sellaisia, joiden käyttöön kaikilla on yhtäläiset mahdollisuudet, joiden käytöstä ei peritä maksua ja joiden kehitystyö on avointa eikä minkään yksittäisen yrityksen hallitsemaa. Yksi tapa yhdistää erilaisia standardeja on käyty läpi Helsinki Institute for Information Technologyn ns. MyData-viitearkkitehtuurissa, joka on hahmotelma siitä, millä malleilla MyData-tilin toteutuminen voisi tapahtua.¹¹¹

Tiedon ja tietojärjestelmien yhteentoimivuus

Julkishallinnon sisällä on yhteentoimivuuden eteen tehty paljon työtä, mm. tietovarantojen yhteisen arkkitehtuurin ja sanastotyön kautta.¹¹² Kun henkilötietojen käyttäminen ja yhdistely helpottuvat, tulee yksityisistä tahoista entistä tärkeämpi sidosryhmä yhteentoimivuustyöhön. Näin on erityisesti silloin, kun yksityiset toimijat myös tuottavat tietoa, jonka hyödyllisyys edellyttää yhteisiä standardeja ja käytäntöjä. Sekä yksityisiä että julkisia toimijoita sisältävien konsortioiden kautta voi luoda jaettuja malleja. Tulevaisuudessa myös yhteentoimivuus maiden välillä on entistä tärkeämpää. Koneoppiminen ja tekoäly saattavat tarjota uusia työkaluja tiedon harmonisointiin.

¹¹⁰ Yksi haaste tällä alueella voi olla sen tulkitseminen, mikä lasketaan lainvoimaiseksi suostumukseksi tai sopimuksen hyväksymiseksi. Esimerkiksi Tietosuoja-asetuksen mukaan suostumus on annettava "selkeästi suostumusta ilmaisevalla toimella" ja sen täytyy olla "tietoon perustuva". Lähde: Tietosuoja-asetus Recital 32, <http://www.privacy-regulation.eu/fi/r32.htm>

¹¹¹ <http://hiit.github.io/mydata-stack/>

¹¹² Esimerkiksi PerustA. Perustietovarantojen viitearkkitehtuuri, <https://www.avoindata.fi/data/fi/dataset/perusta-perustietovarantojen-viitearkkitehtuuri-1-0>

Henkilötiedon käytön saralla yksi haaste on, että lähitulevaisuudessa syntyy todennäköisesti sekä yleisiä että toimialakohtaisia järjestelmiä. Palveluväylän kautta siirtyy monien rekistereiden sisältö, mutta esimerkiksi terveystiedot siirtyvät Kanta-palveluiden kautta. Kunhan erilaisissa alustoissa tiedon jakamista varten käytetään avoimia standardeja, tulee olemaan mahdollista luoda sovelluksia, jotka yhdistävät tietoja useista järjestelmistä. Jokin yksittäinen julkinen toimija voisi olla vastuussa henkilötiedon hyödyntämisen järjestelmien kehityksen koordinoinnista, samaan tapaan kuin julkishallinnon kokonaisarkkitehtuurin määrittelyssä ehdotetaan tehtävän muidenkin ekosysteemien kokonaisuuksille.¹¹³

¹¹³ Ks. Ekosysteemimalli Julkisen hallinnon kokonaisarkkitehtuurin kuvauksista, <https://wiki.julkict.fi/julkict/juhta/juhta-tyoryhmat-2016/jhka-tyoryhma>

7. EHDOTUKSIA TOIMENPITEIKSI

Aleksi Knuutila – Vesa Kokkonen

Raporttia varten tehdyissä haastatteluissa ja keskusteluissa nousi esiin useita henkilötiedon uudelleenkäyttöön liittyviä haasteita sekä hyviä toimintamalleja. Ne on tässä luvussa esitetty koosteena toimintaehdotuksia.

Ihmiskeskeinen lähestymistapa henkilötietoon on paradigman muutos, jonka täysi toteutuminen julkishallinnossa vaatii pitkäjänteistä työtä, teknisen infrastruktuurin sekä laillisten puitteiden muutosta. Tavoitteena muutokselle ei ole parempi tietojärjestelmä vaan toimintatapojen uudistaminen sekä paremmat palvelut. Infrastruktuuria ja yhteisiä käytäntöjä muutettaessa harvoin syntyy helppoja ja nopeita voittoja. Silti tänään tehtävien ratkaisujen kanssa on eletävä pitkään, joten hallinnosta pitäisi löytyä tahtoa laaja-alaiseen muutokseen avoimuuden ja ihmiskeskeisen tiedon hallinnan periaatteisiin nojautuen.

7.1. Henkilötiedon hallinnan peruspalveluiden synnyttäminen

Yhteiset työkalut ja alustat datan hyödyntämiseksi. Henkilötiedon hyödyntämistä varten tulisi kehittää yhteisiä työkaluja ja alustoja, jotka myös mahdollistavat tiedon saatavuuden järjestämisen askelittain. Julkishallinnon tulisi kehittää ensinnäkin "Green Button"-tyyppinen ratkaisu, jota voi soveltaa hyvin laajasti omien tietojen lataamiseen niin valtio- kuin kuntatasolla. Tämän lisäksi tulisi kehittää MyData-tilin kaltainen alusta, joka sallii tiedon käyttöön annettujen lupien muuttamisen ja hallinnan yhdestä paikasta. MyData-tilin kehitys voi edetä kokeilevin keinoin, ja pitäen mielessä myös vaativammat tavoitteet, kuten ohjelmoidun sopimisen ja oman yksityisyyden hajautettu hallinta.

Käytännön keinot tiedon hallintaan. Luottamus julkisiin palveluihin sekä niiden ymmärrettävyys paranevat, jos kansalaisilla on helpot keinot seurata ja joissain tapauksissa myös ohjata tietojensa käyttöä. Julkishallinnon tulisi käyttäjälähtöisesti tuottaa palvelut tiedon käytön lokitietojen lukemiseen, rekisterien sisällön korjaamiseen sekä yksityisyysasetuksiin, joiden kautta ihmiset voivat tehdä (lain rajoissa) valintoja tietojensa luovutuksesta ja käytöstä. Lisäksi julkishallinnon tulisi kehittää yhteisiä toimintatapoja, jotka helpottavat ihmisiä ymmärtämään tiedon jakamisen seurauksia, siihen liittyviä merkityksellisiä valintoja sekä luottamusta järjestelmään. Sen tulisi luoda malli tietoa käyttävien sovellusten tietosuojan ja turvallisuuden arviointiin, johon kuuluisivat sekä käyttäjien antama palaute että kolmansien osapuolten tekemät kevyet arvioinnit.

Pidetään uusien sovellusten luomisen kynnyks matalana. Uudet sovellukset voivat päästä nopeasti käyttämään MyData-rajapintoja, jos julkishallinto kehittää yhteisiä malleja tarvittaville sopimuksille sekä käyttöehdoille, joita sovelletaan yksilöiden ja tietoa käyttävien palveluiden välisessä sopimisessa. Tavoitteina ovat selkeät säännöt ja vastuut, rekisterinpitäjien hallintotaakan vähentäminen, kokeilujen mahdollistaminen sekä yksilön merkityksellisten valintojen tukeminen. Kynnystä kokeiluihin ja innovaatioiden luomiseen voidaan madaltaa tuottamalla rajapinnat tiedon hyödyntäjien tarpeet mielessä. Jo alusta asti tietotuotteita suunniteltaessa voidaan palveluiden kehittäjiä ottaa mukaan ja kuulla. Myös laadukkaat yhteiset työkalut, kuten ohjelmistojen kehityspakkaukset, pienentävät uusien ratkaisujen kustannuksia.

Hyvät palvelut kyvyistä tai taidoista riippumatta. Uusia palveluita tulee luoda niiden käytön haasteet mielessä ja saavutettavuus lähtökohtana. Ihmiskeskeinen tiedon hyödyntäminen voi auttaa näissä tavoitteissa helpottamalla sitä, että palveluille luodaan useita vaihtoehtoisia käyttöliittymiä ja asiointin tapoja.

7.2. Yksittäisten julkishallinnon organisaatioiden toimenpiteet

Hyödyllisimmät tietovarannot saatavaksi yksilöille. Kun henkilötietoja tehdään siirrettäväksi ihmisten antamien lupien perusteella, julkisten toimijoiden kannattaa aloittaa tiedosta, jonka katsotaan olevan hyödyllistä yksilöille, ja tietovarannoista, joiden avaamisen kustannukset ovat pienet. Hyödyllisiä tietoja ovat etenkin tunnistetiedot ja tietovarannot, joita ei ole saatavilla muualla, sekä usein muuttuvat ja määrältään suuret asiakastiedot. Joissakin sähköisen asiointin palveluissa on jo valmiiksi näkyviä omaan tietoon, jotka voivat olla muutettavissa MyDataksi suhteellisen pienellä panostuksella.

Tieto saataville käyttäjälähtöisesti. Hyödyllisimmät keinot henkilötiedon jakamiseen syntyvät julkishallinnon sekä dataa hyödyntävien yritysten, kolmannen sektorin sekä kansalaisten vuorovaikutuksessa. Pieniäkin datamääriä kannattaa tehdä saatavaksi kokeiluluonteisesti, esimerkiksi mahdollistamalla sen lataaminen viranomaisien omien sähköisten palvelujen kautta. Yhteiset alustat, kuten mahdollinen palveluväylään liittyvä operaattori, tulevat myös pienentämään kokeilujen kustannuksia tulevaisuudessa. Kun tietoa tehdään saatavaksi, tulisi luoda toimivat palautemekanismit sekä yhteistoimintaa kehittäjien kanssa, esimerkiksi kilpailujen ja muiden tapahtumien muodossa. Tämän vuorovaikutuksen pohjalta rekistereitä hallinnoivat tahot voivat paremmin ymmärtää, mitkä tiedon kokonaisuudet luovat eniten lisäarvoa ja millaiset tekniset ratkaisut parhaiten tukevat tiedon hyödyntämistä.

Panostetaan datan käytettävyyteen. Nykyiset tietovarantojen kuvaukset, esimerkiksi rekisteriselosteissa, ovat epäyhtenäisiä eivätkä erityisen hyödyllisiä. Viranomaiset voivat varautua MyDatan mahdollistamaan tietojen entistä laajempaan yhteiskäyttöön varmistamalla, että käytetyt tietomallit sekä sanasto on selkeästi määritelty. Tässä työssä voi soveltaa nykyisten kärkihankkeiden luomia yhteentoimivuuden välineitä sekä tulevaa JUHTAn ohjausta henkilötiedon metatiedoista. Julkishallinnon pitäisi lisäksi kehittää datakatalogien kaltaisia sovelluksia myös henkilötiedolle, jotta tämän löydettävyyden paranisi. Esimerkiksi Suomi.fi-palvelutietovaranto voisi sisältää viittauksia niihin tiedon lähteisiin, joita yksittäiset palvelut voivat hyödyntää.

7.3. Sääntely ja kansallinen ohjaus

Yhteisten sääntöjen selkeyttäminen. EU:n tietosuojasetus (GDPR) määrittää henkilötietojen käyttöehdot yrityksille. Asetus ei ole täysin sovellettavissa julkiselle sektorille – esimerkiksi ihmisillä ei ole oikeutta poistaa tietojaan rekistereistä. Julkishallinnon tulee kehittää kuitenkin omaa toimintaansa niin, että datanhallinta vastaa mahdollisimman hyvin GDPR:n henkeä. Tämä edellyttää mm. koneluettavaa datan saatavuutta, mahdollisuutta valvoa tietojen käyttö (lokiteidot) ja mahdollisuutta korjata väärää tietoa.

Ekosysteemin kehityksen koordinointi. Julkishallinnossa tulisi olla valtuutettu taho, jonka tehtävänä on neuvoa, tehdä mahdolliseksi ja koordinoida ihmiskeskeisen henkilötiedon ekosysteemin kehittämistä. Toimijalla olisi riittävät resurssit ja osaaminen, jotta se voisi olla

vastuussa yhteentoimivuudesta ja toisiaan täydentävien palveluiden syntymisestä julkishallinnossa sekä yhteistyössä yksityisen sektorin kanssa. Merkittävää on huomioida yhteys alustatalouteen. MyDataan ja muiden datalähteiden hyödyntäminen edellyttää toimivien alustojen kehittämistä niin, että eri toimijat pystyvät rakentamaan omia toimintojaan sen päälle. MyData-operaattori voi olla yksi alustoista, mutta oleellista on saada yksityinen sektori myös mukaan toimintaan ja mahdollistaa useampien eri tiedon tyyppien yhdistely suuremmissa mittakaavaissa. Tämä tarkoittaa yksityisten datavarantojen yhteensovittamista julkisten kanssa sekä myös muun kuin henkilötiedon yhdistämistä kokonaisuuteen.

Yhteisten standardien ja toimintamallien kehittäminen. MyData edellyttää yhteisiä standardeja paitsi tiedon muodossa ja sanastoissa, myös esimerkiksi käyttöehtojen esittämisessä, suostumusten hallinnassa sekä tunnistautumisessa. Yhteentoimivuus on entistä suurempi haaste, sillä sen täytyy toteutua yksityisen ja julkisen sektorin välillä sekä kansainvälisesti. Julkishallinnon tulisi resursoida toimialakohtaisia konsortioita sekä suomalaisten osallistumista kansainväliseen standardien kehitykseen.

7.4. Hallinnon sisäiset käytännöt tukemaan MyDataa

Datan liittyvien riskiarvioiden tekeminen. Hallinnollisesti yhtenä MyDataan keskeisenä haasteena nousi esille dataan liittyvät riskit. Henkilötieto on toisinaan niin arkaluontoista, ettei sen jakamista haluta jättää yksilön vastuulle. Ongelman ennakoinniseksi eri datalähteille tulisi laatia riskiarviot ja suunnitelma näiden riskien hallitsemiseen. Riskien hallintaan tulee luoda myös työkalut sekä seuranta. Riskiarvioiden tulokset voivat ohjata niiden käyttöehtojen muodostamista, joita tarjotaan sopimusmalleiksi tiedon hyödyntämistä varten. Tuloksia voidaan myös esittää selkeässä muodossa käyttäjille näiden tehdessä valintoja tiedon jakamisesta. Tämän lisäksi riskien toteutumista tulee valvoa. Jos ongelmia ilmenee, täytyy olla mahdollisuus tehdä myös päätös datan jakamisen keskeyttämisestä.

Kokeilut suostumusten pohjalla toimivista julkisista palveluista. Jotkin julkiset palvelut, kuten esimerkiksi tarpeita ennakoiva palveluista tiedottaminen, voisivat toimia suostumuksella saadun tiedon pohjalta. Julkishallinto voisi kokeiluna kehittää monipuolisiakin palveluja, jotka käyttävät MyData-rajapintoja ja joissa kansalaisilla olisi enemmän valinnanvaraa oman tietonsa käytöstä. Jos useampi tiedon lähde on käytettävissä yksilöiden antamien lupien pohjalta, avautuu myös uusia mahdollisuuksia analytiikan sekä tekoälyn soveltamiseen julkisten palvelujen tuottamisessa. Omien tietojen kulkeminen ihmisten mukana voi myös mahdollistaa kolmansien osapuolten tuottamia vaihtoehtoisia käyttöliittymiä julkisiin palveluihin. Tulevaisuudessa yritykset voisivat esimerkiksi kilpailla siitä kuka tuottaa sujuvammat keinot erilaisten lupa-asioiden hoitamiseen. Myös tällä saralla julkishallinto voisi tukea uusien kokeilujen syntymistä.

Valtio esimerkilliseksi toimijaksi henkilötiedon käytössä. Julkishallinto voisi näyttää tietä siinä, kuinka palveluita rakennetaan niin, että ihmiset luottavat niihin ja ymmärtävät kuinka heitä koskevaa tietoa käytetään. Valtion tulisi kehittää yleisiä suunnittelumalleja, joita voisi soveltaa asiain käyttöliittymien sekä tiedon esittämisen muotojen suunnittelussa. Suunnittelumallit sisältävät ohjeita sekä uudelleenkäytettäviä osia (esimerkiksi visuaalisia elementtejä), joita voi soveltaa laajasti uusia julkisia palveluita kehitettäessä. Näiden mallien tavoitteena olisi tehdä käyttäjille selväksi mm. miten heidän tietojensa käytetään palveluissa, ja kuinka he voivat saada apua jos heitä koskeva tieto tai sen perusteella tehty päätös ovat virheellisiä.

Virkamiesten dataosaamisen kehittäminen. Tiedon määrän ja sen arvon kasvaessa on harvoja tehtäviä, joihin datan ja henkilötiedon käsittely ei vaikuttaisi. Virkamiesten pitäisi siis paitsi ymmärtää tiedon hallinnan ja tietosuojan periaatteet myös tuntea tiedon julkisuuteen ja avoimuuteen liittyvät käytännöt. Julkinen hallinto voisi tuottaa aiheesta verkkokurssin sekä yhteisiä työkaluja tiedon hallinnan hyvien toimintatapojen tukemiseksi.

7.5. Suomen kansainvälisen edelläkävijyyden mahdollistaminen

Koulutusta ja tiedotusta datakansalaisille. Ihmiskeskeinen tiedon hallinta edellyttää kyvykkyyttä osallistuvilta ihmisiltä, jotta heillä olisi datan lukutaitoja ja he ymmärtäisivät tiedon jakamiseen liittyviä valintoja ja niihin liittyviä riskejä. Esimerkiksi tietosuojavaltuutettu voisi kehittää osuvaa tiedotusta aiheesta, ja datakansalaisten verkkokurssia voisi käyttää kouluopetuksessa esimerkiksi mediakasvatuksen yhteydessä.

Ratkaistaan ongelmia kokoamalla tietoa uudella tavalla. Usein yhteiskunnallisiin haasteisiin, kuten terveyspalveluiden tai liikenteen suunnitteluun, tarvittava tieto on hajautuneena julkishallinnon, yritysten ja yksilöiden hallussa. Tulevaisuudessa tarvitaan kestäviä ja yksityisyyden huomioivia keinoja kerätä tätä tietoa tutkimuskäyttöön ja muuhun ongelmanratkaisuun. Niihin voi kuulua myös yksilöiden osallistuminen, esimerkiksi suostumuksen antamisen tai jopa kansalaisten hallinnoimien organisaatioiden kautta. Suomen biopankit ovat esimerkki järjestelystä, jossa yksilöillä on mahdollisuus luovuttaa tietoa tutkimuskäyttöön ja nähdä millaiseen tutkimukseen sitä hyödynnetään. Sveitsistä löytyy terveystiedon osuuskunta, johon osallistuvat päättävät yhdessä kuinka syntyvää taloudellista arvoa käytetään.¹¹⁴ Tiedon kokoamista voisi tukea yhteisillä ratkaisulla ja pelisäännöillä tiedon anonymisointia varten. Esimerkiksi suostumusten hallinnan yhteydessä ihmisille voisi tarjota mahdollisuuden luovuttaa omaa dataa anonymisoina tutkimuskäyttöön.

Suomi viemään ihmiskeskeistä dataa Euroopan unioniin. Dataa hyödyntävä liiketoiminta ei ole kansallista toimintaa. Ekosysteemien synnyttämisen kannalta riskinä on se, että MyData-aloitteet jäävät suomalaisiksi kansallisiksi hankkeiksi, jotka eivät tue kansainvälisesti skaalautuvia liiketoimintamalleja. Matkapuhelinstandardit (ensiksi NMT ja sitten GSM) ovat loistavia historiallisia esimerkkejä siitä, miten suomalainen elinkeinoelämä on saanut globaalia kilpailuetua olemalla mukana luomassa kansainvälisiä standardeja. Suomen kannalta luontainen areena standardien luomiselle olisi Euroopan unioni ja sen digitaalisten sisämarkkinoiden kehittäminen.¹¹⁵ GDPR tähtää jo EU:n sisäisen digitaalisen markkinan kehittämiseen. Tätä voitaisiin viedä vielä pidemmälle luomalla alustoja ja standardeja, jotka mahdollistavat eri lähteistä tuotetun datan yhdistämisen.

Datan hyödyntämisen rahoitus. Liiketoiminnan käynnistäminen ja ekosysteeminen kehittäminen edellyttävät rahoitusta. MyData ei välttämättä tarvitse omaa ohjelmaa, mutta sen tuomia mahdollisuuksia voidaan nostaa esille digitaalista liiketoimintaa tukevissa ohjelmissa. Strategisia avauksia tehtäessä näitä kaikkia rahoitusinstrumentteja kannattaa suunnata niin, että ne tukevat henkilötiedon hyödyntämisen ekosysteemien kehitystä ja toisaalta eivät sulje pieniä toimijoita ulos. Rahoitus voi tukeutua esimerkiksi Tekesin ja

¹¹⁴ <https://midata.coop/>

¹¹⁵ Tällä hetkellä EU:n digitaalisen sisämarkkinoiden strategiassa on mainittu erikseen sekä liikenne että terveys. Näissä molemmissa MyDatalla on oleellinen merkitys ja näin ollen konsepti kannattaa tuoda osaksi tätä strategiatyötä. Esimerkiksi terveyspalveluiden aito kilpailuttaminen edellyttää sitä, että hoitohistoria on siirrettävissä uudelle palveluntarjoajalle. Tällä hetkellä se ei ole mahdollista kuin Suomen sisällä.

Suomen Akatemian ohjelmiin. Tämän lisäksi kannattaa hyödyntää innovatiivisia hankintoja, joissa julkinen sektori ja yritykset yhdessä kehittävät palveluita.

LIITE 1. REKISTERÖITYJEN OIKEUDET ERI TIETOTYYPPIEN KOHDALLA

Tämä osio kuvaa henkilötietoon liittyvien oikeuksien toteutumista tiedon eri tyyppien kohdalla. Tarkastelu keskittyy nykyiseen lainsäädäntöön sekä osin tulevaan yleiseen tietosuoja-asetukseen. Lainsäädäntö edesauttaa MyData-periaatteiden toteutumista vahvistamalla yksilön oikeuksia sekä lisäämällä luottamusta tietosuojaan kolmansien osapuolten hyödyntäessä tietoa. Luonnollisesti lainsäädännössä määritellyt oikeudet eivät yksinään riitä ihmiskeskeisen henkilötiedon hallinnan mahdollistamiseksi.

Erilaisten tietotyyppien luovuttamista ja käsittelyä koskeva lainsäädäntö on liian monimutkainen kokonaisuus, jotta sitä voisi tyhjentävästi käsitellä tässä yhteydessä. Tavoitteenamme on esitellä yleislainsäädännön pääpiirteet sekä esimerkkien avulla erityislakien poikkeuksien tyyppisiä. Erityislakien osalta viittaamme usein terveystietoihin ja väestörekisteriin kohdistuviin säännöksiin.

Tiedon luovuttaminen rekisteröidylle: Tarkastusoikeus ja oikeus tiedon siirtoon

Tarkastusoikeus on yksi henkilötietolain keskeisistä oikeuksista. Sen puitteissa yksilöille annetaan oikeus saada kopio rekisterinpitäjällä olevista itseään koskevista tiedoista. Tavoitteena on antaa rekisteröidyille mahdollisuus valvoa itseään koskevan tiedon oikeellisuutta. Virheellisen tiedon seuraukset vaikuttaisivat nimittäin usein juuri rekisteröityyn eikä rekisterinpitäjään.

Tarkastuspyynnön kautta ihmisillä on pääsy yksittäisiin otteisiin tiedoistaan. Pyyntö esitetään kirjallisina ja vastauksissa kestää usein pitkään. Tarkastuspyynnöt eivät siis helposti sovellu tiedon uudelleenkäyttöön. Tästä huolimatta tarkastusoikeuden rajoitukset voivat antaa suuntaa sille, minkä tyyppisiä tietovarantoja yksilöiden ei todennäköisesti ole mahdollista saada käyttöönsä lähitulevaisuudessa. On myös huomattava, että nykyisessä lainsäädännössä määritellään useita muita oikeuksia tietoon, kuten julkisuuslain tiedonsaantioikeus ja asianosaisuusjulkisuus, sekä potilaslaissa määritelty oikeus selvitykseen terveydentilasta.¹¹⁶

Taulukko A.1. Suomalaiset rajoitetun tarkastusoikeuden rekisterit

Hallinnonala/Toimiala	Rekisterit
Poliisi	Muun muassa epäiltyjen tietojärjestelmä; suojelupoliisin toiminnan tiedot; salaista tarkkailua ja erityistarkastuksia koskevat tiedot; tekoja, tarkkailuja tai tietolähteitä koskevat tiedot poliisin henkilörekisteristä
Rajavartiolaitos	Rikoksesta epäiltyjen rekisteri; turvallisuustietojen rekisteri; havaintotiedot; luokittelu, tarkkailu- tai tekotapatiedot henkilörekisteristä
Tulli	Valvontatietorekisterin tiedot; Euroopan unionin tullitietojärjestelmän tiedot; rekisterikilpien ja konttien kuvaus- ja tunnistusjärjestelmän tiedot; luokittelu, tarkkailu- tai tekotapatiedot henkilörekisteristä
Puolustusministeriö	Turvallisuustietorekisterin tiedot
Rikosseuraamuslaitos	Turvallisuustietorekisterin tiedot ja järjestystietorekisterin tiedot

Rekisterit, joissa tarkastusoikeus on rajoitettu. Näissä tapauksissa tietosuojavaltuutettu käyttää tarkastusoikeutta rekisteröidyn puolesta. Lähde: Tietosuojavaltuutetun toimisto 2014

¹¹⁶ Laki viranomaisten toiminnan julkisuudesta 621/1999, 9. momentti ja Laki potilaan asemasta ja oikeuksista 785/1992, 5. momentti

Tarkastusoikeuden tilannekohtaiset rajoitukset

Tarkastusoikeus ei ole ehdoton. Sitä rajoittavia säädöksiä sisältyy sekä henkilötietolakiin että erikoislakeihin, jotka määrittelevät käsittelyjen rajoja tietyillä toiminta-alueilla. Henkilötietolain mukaan tarkastusoikeutta ei ole, jos tiedon luovuttaminen voisi haitata mm. valtion turvallisuutta, yleistä järjestystä tai rikosten selvittämistä. Lisäksi oikeutta ei voi soveltaa, jos tiedon käytöstä voisi seurata vaaraa rekisteröidyn terveydelle tai jonkun muun oikeuksille.¹¹⁷ Tämän rajoituksen toteutuminen voi olla oleellista useiden julkisten palveluiden keräämien tietojen suhteen, kuten esimerkiksi henkilön hoitotietoja, sairauksia tai mielenterveyttä koskevien rekistereiden kohdalla.

Esimerkiksi silloin kun pyyntö tiedoista kohdistuu terveydenhuollon viranomaisen tai muun alan ammattihenkilön pitämään rekisteriin, jossa on terveydentilaa tai sairautta koskevia henkilötietoja, tarkastusoikeus on selkeästi rajattu niin, että se toteutuu ammatinharjoittajan välityksellä. Tarkastuspyyntö pitää osoittaa lääkärille tai jollekulle muulle alan ammattihenkilölle, ei esimerkiksi tietosuojasta vastaaville toimistovirkailijoille. Tiedon antamisesta myös huolehtivat kyseiset ammattihenkilöt, joiden kautta tieto luovutetaan, ja jotka voivat selvittää, toteutuvatko tarkastusoikeuden yleiset ehdot kyseisessä tilanteessa.¹¹⁸ Tarkastusoikeus ei siis näissä tapauksissa pelkisty pelkästään tietoon pääsyyn tai sen näkemiseen, vaan nykyisen lainsäädännön pohjalta edellyttää ammatinharjoittajan osallistumista tiedon välittämiseen ja hänen arviointiaan rekisteröityyn sekä muihin ihmisiin kohdistuvista riskeistä.

Oikeus tietojen siirtoon

Yleisessä tietosuoja-asetuksessa määritellään tarkastusoikeuden lisäksi oikeus tietojen siirtoon järjestelmästä toiseen. Artikla 20 asettaa velvollisuuden rekisterinpitäjille toimittaa tieto jäsennellyssä, koneluettavassa ja yleisesti käytettävässä muodossa. Jos vain se on teknisesti mahdollista, rekisteröidyllä on lisäksi oikeus saada tiedot siirrettyä suoraan toiseen rekisteriin. Henkilötietojen uudelleenkäytön kannalta tämä oikeus on merkittävästi vahvempi kuin tarkastusoikeus, ja myös lähempänä sellaista tiedonkäyttöä, jota MyData-periaatteet edellyttävät.¹¹⁹

Tapa, jolla oikeutta siirtoon tulkitaan, on edelleen kehittymässä. Tietosuoja-asetuksen tekstissä oikeus rajataan koskemaan nimenomaan tietoja, jotka rekisteröity on toimittanut rekisterille, ja joiden käsittelyn perusteena on joko rekisteröidyn antama suostumus tai sopimus. Koska artikla koskee vain toimitettuja tietoja, oikeus ei ilmeisesti pätsi tietoihin, jotka on esimerkiksi siirretty jostain toisesta tietojärjestelmästä ilman rekisteröidyn myötävaikutusta. Lisäksi on epäselvää, ulottuuko säännös myös tietovarantoihin, joissa on yhdistetty sekä rekisteröidyn toimittamia että muista lähteistä peräisin olevia tietoja. Suuri osa julkishallinnon henkilötiedon käsittelystä ei myöskään nojaa suostumukseen tai sopimukseen. Tältä pohjalta vaikuttaisi siltä, että oikeutta tiedon siirtoon ei välttämättä sovellettaisi suureen osaan viranomaisten hallussa olevasta henkilötiedosta.

Tiedon luovuttaminen rekisteriltä kolmannelle osapuolelle

Oikeus sekä käytännön keinot tehdä valintoja omien tietojen käytöstä on keskeinen osa MyData-periaatteita. Tässä raportissa kuvatuissa MyData-tileissä tiedonkäytön luvittamisella voidaan avata yhteys datan jakamiseen tiedon lähteen sekä sen käyttäjän välillä. Koska

¹¹⁷ Henkilötietolaki 523/1999, momentti 27.

¹¹⁸ Henkilötietolaki 523/1999, momentti 28.

¹¹⁹ Euroopan unionin yleinen tietosuoja-asetus, artikla 20: "Oikeus siirtää tiedot järjestelmästä toiseen"

samalla käyttäjä määrittelee suostumuksen oman datansa käyttöön, ovat myös jotkin henkilötiedon käsittelijän laillisista vaatimuksista katettu. Tässä olemme tulkinneet laillisia mahdollisuuksia luvan antamiseen jakamalla sen ensinnäkin mahdollisuuden järjestää tiedon luovuttaminen kolmannelle osapuolelle oman suostumuksen avulla, ja toiseksi estää tiedon luovutus kieltäytymällä antamasta omaa suostumusta.

Nykyisessä lainsäädännössä luovuttamisen yleiset reunaehdot viranomaisten rekistereistä on määritelty kahdessa yleislaissa, julkisuuslaissa ja henkilötietolaissa sekä lukuisissa erityislaissa. Julkisuuslaki sisältää julkisuusperiaatteen sekä salassapitomääräyksiä. Yksinkertaistaen, julkisuuslain mukaan tietoa voidaan luovuttaa henkilörekisteristä, jos tiedot vastaanottavalla taholla on oikeus käsitellä tietoa henkilötietolain sekä erityislakien rajaamalla tavalla. Tietoa pyytävän on määriteltävä tietojen käyttötarkoitus sekä muut tietosuojaan kannalta merkittävät tiedot¹²⁰. Arkaluontoiselle ja salassapidettävälle tiedolle vaatimukset ovat tiukempia.

Luovutukset ilman suostumusta ja käyttötarkoituksiin sidotut kiello-oikeudet

Väestörekisteri on perusrekisterinä laajasti niin julkisten kuin yksityistenkin toimijoiden käytössä. Se on esimerkki rekisteristä, jolle luovutuksen reunaehdot on määritelty erityislaissa. Lain mukaan väestörekisteristä voidaan luovuttaa tietoa erinäisiin käyttötarkoituksiin, joihin kuuluvat viranomaistehtävien hoitaminen, sekä rahoitus- ja vakuutustoimintaan, johon rekisteröity on osallisena.¹²¹ Lisäksi vähemmän arkaluontoisia tietoja voidaan luovuttaa myös muuhun yksityiseen käyttöön ilman käyttötarkoituksen määrittelyä. Osa väestörekisterille määritellyistä luovutuksen edellytyksistä, kuten laissa määriteltyjen viranomaistehtävien hoitaminen, pätevät pääsääntöisesti myös muille viranomaisten rekistereille. Näiltä osin tietojen luovutukset voivat siis tapahtua ilman rekisteröidyn suostumusta.

Sekä yleis- että erityislaeissa on säännöksiä, joiden perusteella yksilön on mahdollista kieltää tietojensa käsittely ja luovutus erityisiin tarkoituksiin. Tällaisia kieltoja ovat estot tietojen käytöstä suoramarkkinointiin (markkinointikielto) ja estot osoitetietojen luovutuksesta muille kuin viranomaisille (osoitteen luovutuskielto). Jos henkilöllä on perusteltu syy epäillä uhkaa itselleen tai perheelleen, hän voi kieltää yhteystietojensa käytön määrääjäksi (turvaamiskiello).¹²² Kielto-oikeus pätee siis vain poikkeuksellisissa olosuhteissa tai viranomaisen hallitseman tiedon käyttöä tiettyihin tarkoituksiin.

Luovuttaminen suostumuksen perusteella

Yksilön antama yksiselitteinen suostumus kattaa henkilötietolaissa määritellyn ns. tiedon käsittelyn yleisen edellytyksen.¹²³ Laissa määritellään myös salassapidettävään tai arkaluontoiseen (esim. vakaumusta tai rikoksia kuvaavaan) tietoon kohdistuvia rajoituksia, joista rekisteröity voi omalla suostumuksellaan luopua.¹²⁴ Suostumusta vaaditaankin usein silloin, kun yksityiset tahot käyttävät julkishallinnon rekistereistä tietoa. Koska tietosuoja on perus- ja ihmisoikeus, ei suostumuksen perusteella voi kuitenkaan luopua muista laissa

¹²⁰ Laki viranomaisten toiminnan julkisuudesta 621/1999, 13. momentti.

¹²¹ Laki väestötietojärjestelmästä ja Väestörekisterikeskuksen varmennepalveluista 2009/661, momentit 28-34.

¹²² Tietosuojaavaltuutetun toimisto, 2010: Rekisteröidyn kiello-oikeus, http://www.tietosuoja.fi/material/attachments/tietosuojaavaltuutettu/tietosuojaavaltuutetuntoimisto/oppaat/6.JfqC7x0Z/Rekisteroidyn_kiellooikeus.pdf

¹²³ Henkilötietolaki 523/1999, momentti 8.

¹²⁴ Laki viranomaisten toiminnan julkisuudesta 621/1999, momentti 26.

säädetyistä henkilötiedon suojusta.¹²⁵ Tiedon käsittelyn on edelleen oltava tarpeellista ja käyttötarkoituksen rajaamaa.

Erityislainsäädännössä tiedon luovuttamisen edellytyksiä sekä suostumukselta vaadittua muotoa on usein tarkennettu. Esimerkiksi tämänhetkinen lainsäädäntö potilastietojen sähköisestä käsittelystä sallii tietojen luovutuksen yksilön suostumuksella “ainoastaan toiselle terveydenhuollon palvelujen antajalle” kansallista tietopalvelua käytettäessä. Lisäksi suostumus täytyy esittää potilaan allekirjoittaman asiakirjan pohjalta, jossa potilas myös näyttää, että hänelle on selitetty, kuinka potilastieto on suojattu ja sen jakaminen järjestetty.¹²⁶ Terveydenhuollon asiakastietojen lainsäädäntöä ollaan tosin kirjoitushetkellä uudistamassa.

Tiedon poistaminen

Uuden yleisen tietosuoja-asetuksen myötä henkilötietoihin liittyviin oikeuksiin on myös kirjattu täsmällisesti rekisterinpitäjän velvollisuus poistaa tietoja tietyissä yhteyksissä. Tietosuoja-asetuksen artikla 17 määrittelee niin sanotun “oikeuden tulla unohdetuksi”. Rekisteröidyllä on oikeus pyytää poistamaan tiedot, jotka ovat vanhentuneita tai virheellisiä, tai sellaiset tiedot, jotka eivät ole enää tarpeellisia niiden käyttötarkoitusta varten. Lisäksi rekisteröidyn on mahdollista pyytää henkilötietoja poistettavaksi, jos niiden käsittelylle ei enää ole laillista perustaa.

Kuten yllä mainitsimme, henkilötiedon käsittelyn peruste on julkishallinnossa usein erikoislaissa säädetty tehtävä tai palvelusuhde henkilön kanssa. Toisin kuin suostumuksen kohdalla, useissa tapauksissa rekisteröidyllä ei ole mahdollista vaikuttaa näihin perusteisiin omilla valinnoillaan. Samalla myös oikeus tietojen poistamiseen voi olla rajoitettu. Esimerkiksi Tietosuojavaltuutettu päätöksessään vuodelta 2000 otti kannan, että henkilöillä ei ollut oikeutta pyytää omien potilasasiakirjojensa poistamista¹²⁷. Perusteluna päätökselle oli, että sosiaali- ja terveysministeriön asetuksen mukaan terveydenhuollon toimintayksiköillä oli laillinen velvollisuus säilyttää potilasasiakirjoja. Pääsääntöisesti siis aktiivinen oikeus poistaa omia henkilötietojaan pätee niihin henkilötietoihin, joita käsitellään rekisteröidyn suostumuksen tai asiakassuhteiden perusteella.

Tiedon käytön valvominen

Osana tarkastusoikeutta, henkilötietolain pohjalta, rekisteröidyllä on oikeus tietää tiedonkäsittelyn yleisistä piirteistä, kuten tietojen käyttötarkoituksesta, tietolähteistä ja mihin tietoja rekisteristä luovutetaan säännönmukaisesti.¹²⁸ EU:n yleinen tietosuoja-asetus tulee lisäämään muutaman tiedon lajin tähän listaan, mm. sen, kuinka kauan henkilötietoa säilytetään.¹²⁹ Tarkastusoikeuden pohjalta saa kuitenkin vain tietoa käsittelyn yleisistä linjoista ja suunnitelluista menettelytavoista, ei oman henkilötiedon käsittelyn yksityiskohdista.

Moniin tietojärjestelmiin liittyy myös ns. käyttölokeja, joihin kerääntyy tietoa henkilöihin liittyvien tietueiden käyttäjistä, käyttötarkoituksesta ja ajankohdista. Viranomaisten rekistereistä lokitietoja kertyy esimerkiksi väestörekisteristä sekä sosiaali- ja terveydenhuollon asiakastiedoista. Yksilöillä on pääsääntöisesti oikeus ja pääsy heitä

¹²⁵ Warma & Nieminen, 2016

¹²⁶ Laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä

¹²⁷ Tietosuojavaltuutetun kannanotto 6.11.2000, Dnro 1035/45/2000, <http://www.tietosuoja.fi/fi/index/ratkaisut/potilasasiakirjoissaolevientietojenpoist.html>.

¹²⁸ Henkilötietolaki 523/1999, momentti 26.

¹²⁹ Yleinen tietosuoja-asetus, artikla 15.

koskeviin lokitietoihin, joiden perusteella omien tietojen käyttöä voisi valvoa.¹³⁰ Jälleen kerran, erityislaeissa rajoitetaan luovutettujen lokitietojen käyttötarkoitusta. Esimerkiksi sosiaali- ja terveydenhuollon lokirekisterin kohdalla tietojen luovutus mm. riippuu ammatinharjoittajan arviosta, ja lokitietojen käyttö on rajattu asiakkaan “oikeuksien selvittämiseen tai toteuttamiseen”.¹³¹

¹³⁰ Tietosuojavaltuutetun toimisto, Lokitiedot henkilötietojen suojaamisen välineinä, http://www.tietosuojavaltuutettu.fi/material/attachments/tietosuojavaltuutettu/tietosuojavaltuutetuntoimisto/oppaat/kaytonvalvonnanhieet/wPJBn9UGn/Kaytonvalvonta_Lokitiedot_henkilotietojen_suojaamisen_valineina_paiv_10.06.2014.pdf Toisaalta korkeimman hallinto-oikeuden linjauksen perusteella on evätty pääsy vankitietojärjestelmän lokeihin, ks. (Voutilainen, 2014).

¹³¹ Laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä, 159/2007, momentti 18.

LÄHTEET

- Ailisto, Heikki, Jari Collin, Jari Juhanko, Martti Mäntylä, Sampsa Ruutu, Seppälä, Marco Halén, et al. 2016. "Onko Suomi jäämässä alustatalouden junasta." 16/2016. Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja.
- Ansper, Arne, Ahto Buldas, Margus Freudenthal ja Jan Willemson. 2013. "High-Performance Qualified Digital Signatures for X-Road." *Secure IT Systems*, 123–38. Springer, Berlin, Heidelberg.
- Department for Business, Innovation & Skills. 2012. "Better Choices: Better Deals. Consumers Powering Growth." Department for Business, Innovation and Skills.
- Edwards, L., ja W. Abel. 2014. "The Use of Privacy Icons and Standard Contract Terms for Generating Consumer Trust and Confidence in Digital Services." CREATE Working Paper 2014/15. CREATE.
- Euroopan komissio. 1998. "Julkisen sektorin tieto: Euroopan avainresurssi. Vihreä kirja julkisen sektorin tiedon käytöstä tietoyhteiskunnassa." KOM (1998) 585. Euroopan komissio.
- Finanssivalvonta. 2016. "Toimintakertomus 2016." Finanssivalvonta.
- Goodall, Warwick, Tiffany Dovey Fishman, Justine Bornstein, ja Brett Bonthron. 2017. "The Rise of Mobility as a Service--Reshaping How Urbanities Get around." *Deloitte Review* 20: 112–29.
- Grossglauser, M., ja H. Saner. 2014. "Data-Driven Healthcare: From Patterns to Actions." *European Journal of Preventive Cardiology* 21 (2 Suppl): 14–17.
- Hyppönen, Hannele, ja Katja Ilmarinen. 2016. "Sosiaali- ja terveydenhuollon digitalisaatio." *Suomen sosiaalinen tila: 2/2016*. THL.
- Jauhiainen, J., ja J. Kuusinen. 4.9. 2012. "Perustietovarantojen tehokkaaseen käyttöön: Verkkohaastattelun tulosityhteenveto." Pertiva.
- Kansallisen Genomistrategian työryhmä. 2015. "Parempaa terveyttä genomitiedon avulla." Sosiaali- ja terveysministeriö.
- Kauhanen-Simanan, Anne. 2013. "Perustietovarantojen maksuttomuuden toteutuminen ja vaikutukset." VM/041:00/2013 885/00.01.00.01/2013. Valtiovarainministeriön taustamuistio.
- Kavén, S. 2016. "Kanta-palvelut: Omakanta-palvelun kuvaus." Kela.
http://www.kanta.fi/documents/10180/4242542/Omakanta_palvelunkuvaus_TU_kantafi_v1.2/0f94df4a-857d-44f7-bb13-329559a2ecda
- Koponen, Johannes, Satu Korhonen, Veikka Lahtinen, Juha Leppänen, Alekski Neuvonen, Maari Parkkinen, Marjukka Parkkinen, ja Iris Sandelin. 2015. "Terveys 2050: Neljä skenaariota ihmislähtöisestä terveydestä ja valinnanvapaudesta." Terveiden tulevaisuudet. Tekes.
- Korhonen, Rauno. 2003. *Perusrekisterit ja tietosuoja*. Edita.
- Koski, H., M. Honkanen, J. Luukkonen, M. Pajarinen, ja T. Ropponen. 2017. "Avoimen datan hyödyntäminen ja vaikuttavuus." 40/2017. Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja.
- Levy, D., C. Wasden, D. DiFilippo, ja P. Sur. 2012. "Emerging mHealth: Paths for Growth." *PwC M-Health*, 1–44.
- Linturi, Risto, ja Ossi Kuittinen. 2014. "Liikennetiedon visiot." *Liikenneviraston Tutkimuksia Ja Selvityksiä: 42/2014*. Liikennevirasto.
- Lupton, Deborah. 2013. "The Digitally Engaged Patient: Self-Monitoring and Self-Care in the Digital Health Era." *Social Theory & Health: STH* 11 (3). Palgrave Macmillan UK: 256–70.
- Maanmittauslaitos. 2017. "Maanmittauslaitoksen tilinpäätös vuodelta 2016." MML 5565/02 00 02/2016. Maanmittauslaitos.

- Mäkinen, Joonas, and Päivi Sassi. 2017. "Opintosuoritukset ja opiskeluoikeudet kootusti yhdessä palvelussa." Accessed April 17 2017. <https://confluence.csc.fi/display/OPHPALV/Koski>
- Marko Männikkö, Klaus Naumanen. 2015. "Valtion maksupolitiikka, tuloksellisuustarkastuskertomus." Valtiontalouden tarkastusvirasto.
- Miettinen, Anne. 2017. "Liikenne- ja viestintäministeriön hallinnonalan datan avaaminen ja hyödyntäminen liiketoiminnassa ja päätöksenteossa." Liikenne- ja viestintäministeriö. <http://julkaisut.valtioneuvosto.fi/handle/10024/79150>
- Neittaanmäki, Pekka, ja Martti Lehto. 2017. "Suomen terveystietoympäristö." Informaatioteknologian tiedekunnan julkaisuja 35/2017. Jyväskylän yliopisto.
- Nissilä, Jussi, Vesa Kokkonen, Ossi Kuittinen, ja Others. 2016. "Kokemuksia massadatan, omadatan sekä älykkään robotiikan ja automaation osaamistarpeista ja tarjonnasta." Liikenne ja viestintäministeriö.
- Patentti- ja Rekisterihallitus. 2017. "Tilinpäätös vuodelta 2016." PRH/443/20/2017. Patentti- ja rekisterihallitus.
- Pirttilä, Ilkka, ja Susan Eriksson. 2015. "Asiantuntijoiden areenat." *SoPhi 66*. Jyväskylän yliopisto. https://jyx.jyu.fi/dspace/bitstream/handle/123456789/47925/SoPhi66_978-951-39-6501-3.pdf?sequence=3
- Poikola, A., E. Hjelm, ja D. Schildt. 2017. "Sähköinen asiointi ja henkilötieto: Selvitys ja kokeiluprojekti MyDatan hyödyntämisestä kaupungin palvelujen kehittämisessä." Helsingin kaupunki.
- Poikola, A., K. Kuikkaniemi, ja O. Kuittinen. 2014. "My Data - Johdatus ihmiskeskiseen henkilötiedon hyödyntämiseen." Liikenne- ja viestintäministeriö.
- Poikola, Antti, ja H. Honko. 2015. "Mydata: A Nordic Model for Human-centered Personal Data Management and Processing." Liikenne- ja viestintäministeriö.
- Poikola, Antti, Petri Kola, ja Kari A. Hintikka. 2010. "Julkinen data - Johdatus tietovarantojen avaamiseen." Liikenne- ja viestintäministeriö.
- Raghupathi, Wullianallur, ja Viju Raghupathi. 2014. "Big Data Analytics in Healthcare: Promise and Potential." *Health Information Science and Systems 2* (February). biomedcentral.com: 3.
- Raula, Helena, ja Tuija Savolainen. 2016. "Sosiaali- ja terveydenhuollon asiakas- ja potilastiedon toissijaista käyttöä koskevaa lainsäädäntöä valmisteleva työryhmä – Väliraportti työryhmän työskentelystä." Sosiaali- ja terveysministeriö.
- Rekisteritutkimuksen Tukikeskus. 2012. "Rekisteritutkimuksen tukikeskus toimintakertomus 2011." Rekisteritutkimuksen Tukikeskus.
- Salminen, Juho. 2011. "Suomessa miljoona henkilörekisteriä: "isoilta ruumiilta vältytty". *Suomen Kuvalehti*. November 9.
- Schwab, Klaus, Alan Marcus, J. O. Oyola, William Hoffman, ja M. Luzi. 2011. "Personal Data: The Emergence of a New Asset Class." World Economic Forum.
- Stiftung Datenschutz. 2017. "Neue Wege bei der Einwilligung im Datenschutz – Technische, Rechtliche Und ökonomische Herausforderungen." Stiftung Datenschutz.
- Stott, Andrew. 2014. "Open Data for Economic Growth." *Washington DC: World Bank*.
- Swan, Melanie. 2013. "The Quantified Self: Fundamental Disruption in Big Data Science ja Biological Discovery." *Big Data 1* (2). online.liebertpub.com: 85–99.
- TATTI Työryhmä. 2017. "EU:n yleisen tietosuoja-asetuksen täytäntöönpanotyöryhmän (TATTI) mietintö". Oikeusministeriön julkaisu 35/2017. Oikeusministeriö.
- Taylor, Karen, Hanno Ronte, ja S. Hammett. 2014. "Healthcare and Life Sciences Predictions 2020--a Bold Future." Deloitte, London.

- Tietäväinen, Jari, Pasi Lahtonen, ja Anne Kauhanen-Simanainen. 2013 "PerustA - Perustietovarantojen viitearkkitehtuuri, Liite 1 Nykytilan kuvaus." Valtiovarainministeriö.
- Tietosuojavaltuutetun toimisto. 2014. "Henkilörekisteriin talletettujen tietojen tarkastaminen." http://www.tietosuojafi.fi/material/attachments/tietosuojavaltuutettu/tietosuojavaltuutetuntoimisto/oppaat/dxJufFmWB/Henkilorekisteriin_talletettujen_tietojen_tarkastaminen_24.11.2014.pdf
- Trafi. 2016. "Trafin tietotilinpäätös 2016." Trafin julkaisuja 14/2016. Trafi.
- . 2017. "Liikenteen turvallisuusviraston tilinpäätös 2016." Trafin julkaisuja, 3/2017 . Trafi.
- Tukikeskus, Retki –. Rekisteritutkimuksen. 28.2. 2017. "Rekisterit rekisterinpitäjän mukaan." *ReTKi – Rekisteritutkimuksen Tukikeskus*. <https://rekisteritutkimus.wordpress.com/rekisterit/rekisterit-rekisterinpitajan-mukaan/>
- Väestörekisterikeskus. 2016. "Väestörekisterikeskuksen tietotilinpäätös 2015, Tiivistelmä." 1650/030/11. Väestörekisterikeskus.
- . 2017. "Väestörekisterikeskuksen tilinpäätös 2016." VRK/439/2017. Väestörekisterikeskus.
- Valtiokonttori. 2016. "Tietoja valtion maksullisesta toiminnasta 2015." Valtiokonttori.
- Valtioneuvosto. 2015. "Ratkaisujen Suomi: Pääministeri Juha Sipilän hallituksen strateginen ohjelma 29.5.2015." Hallituksen julkaisusarja.
- Valtiovarainministeriö. 2011. "Perusrekisterien tietojen käsittelyä koskevat yleiset käyttöehdot." Julkishallinnon tietovarantojen saatavuuden ja käytön edistämistä selvittävä työryhmä. Valtiovarainministeriö.
- . 2012. "Julkishallinnon tietoluovutusten periaatteet ja käytännöt." ICT-toiminta 2/2012. Valtiovarainministeriö.
- Viljanen, K., A. Poikola, ja P. Koponen. 30.4. 2012. "Tietosuunnistusta kaupungissa." Helsingin Kaupunki.
- Voutilainen, T. 2014. "Oikeus tietojärjestelmän lokitietoihin." *Lakimies*, no. 5: 741.
- Warmma, E., ja J. Nieminen. 2016. "Tietosuojaja kilpailuoikeus – Määrävässä markkina-asemassa olevan yrityksen toimitusvelvollisuudesta ja tietosuojalainsäädännöstä." *Defensor Legis* 4: 549.
- Weigend, Andreas. 2017. *Data for the People: How to Make Our Post-Privacy Economy Work for You*. Hachette UK.
- Whitley, Edgar A. 2009/8. "Informational Privacy, Consent and the 'control' of Personal Data." *Information Security Technical Report* 14 (3). Elsevier: 154–59.
- Wiiala, Niina, ja Joonas Kankaanrinne. 2016. "PERTTI Perustietovarantojen tietomallien yhteensovittaminen." PERTIVA raportti. Väestörekisterikeskus.



VALTIONEUVOSTON
SELVITYS- JA TUTKIMUSTOIMINTA

tietokaytoon.fi

ISSN 2342-6799 (pdf)
ISBN 978-952-287-446-7 (pdf)

