Liikenne- ja viestintäministeriö

Antti Poikola
Kai Kuikkaniemi
Ossi Kuittinen

My Data

– johdatus ihmiskeskeiseen henkilötiedon hyödyntämiseen

Avoimesti lisensoitu opas

Selvitystyön on toteuttanut Open Knowledge Finland ry. liikenne- ja viestintäministeriön rahoituksella. Yhteistyössä selvityksen tekemistä on tukenut myös Aalto-yliopiston tietotekniikan tutkimuslaitos HIIT ja TEKESin iso strateginen avaus Digital Health Revolution.

Ulkoasu: **Kirmo Kivelä**

ISBN 978-952-243-417-3 (painotuote)

ISBN 978-952-243-418-0 (verkkojulkaisu http://www.lvm.fi/asiakirjat_ja_muut_julkaisut)

My Data

– johdatus ihmiskeskeiseen henkilötiedon hyödyntämiseen

Antti Poikola | Kai Kuikkaniemi | Ossi Kuittinen

Kiitokset

Selvittääksemme henkilötiedon käsittelyn tulevaisuutta ja My Datan teknisiä, juridisia ja liiketoiminnallisia piirteitä olemme haastatelleet asiantuntijoita eri toimialoilta ja julkishallinnon organisaatioista.

Sakari Vaelma, Gecko Systems
Jussi Muurikainen, Balancion
Ville Peltola, IBM
Anu Talus, Oikeusministeriö
Juha Kenraali, Trafi
Jouni Sintonen, Telia-Sonera Finland
Anssi Mikola, Wellbookers
Tuomas Teuri, Taltioni
Eero Toivanen, Ravintokoodi

Haastatteluiden lisäksi on keskusteltu lukuisien muiden henkilöiden kanssa, jotka tarkastelevat aihetta eri tulokulmista. Open Knowledge Finlandin My Data -työryhmässä joukko yritysten ja julkishallinnon edustajia, tutkijoita ja muita aiheesta kiinnostuneita on yhteisönä kasvattanut ymmärrystä henkilötiedon ihmiskeskeisen käsittelyn mahdollisuuksista. Työryhmän jäsenet ovat olleet mukana selvityksen suunnittelussa ja kommentoinnissa sekä erityiskysymysten asiantuntijoina. Työryhmä on avoin kaikille asiasta kiinnostuneille http://okf.fi/my-data

Kiitos kaikille, joilta olemme saaneet asiantuntevaa apua oppaan kirjoittamisessa. Erityisesti haluamme mainita seuraavat henkilöt: **Jouni Alanen, Emil Asp, Reuben Binns, Leif Beilinson, Myles Byrne, Antti Eskola, Konsta Hansson, William Heath, Kari A. Hintikka, Nina Honkela, Eija Kalliala, Matti Kinnunen, Otso Kivekäs, Miska Knapek, Jaakko Korhonen, Ismo Kosonen, Johanna Kotipelto, Krista Lagus, Mark Lizar, Sami Majaniemi, Markus Petteri Laine, Kiti Müller, Ville Oksanen, Mika Pantzar, Juuso Parkkinen, Olli Pitkänen, Olli-Pekka Pohjola, Elias Pöyry, Taru Rastas, Mikael Rinnetmäki, Samuel Rinnetmäki, Aleksi Rossi, Minna Ruckenstein, Jaakko Talvitie** ja **Laura Tarhonen**.

Tämä selvitys on kiteymä käynnissä olevasta keskustelusta. Julkaisu on jo ennen valmistumistaan ollut avoimesti kiinnostuneiden kommentoitavissa verkossa Google Docs -alustalla ja kommentointimahdollisuutta jatketaan. Toivotamme lukijat tervetulleeksi esittämään ajatuksia, kysyymyksiä, vastalauseita ja parannusehdotuksia osoitteesta http://bit.ly/mydata_selvitys löytyvään selvityksen elävään versioon.

Esipuhe

Meistä kaikista tallennetaan jatkuvasti valtavia määriä tietoa julkishallinnon rekistereihin, yritysten asiakasjärjestelmiin ja käyttämiemme verkkopalvelujen tietokantoihin. Henkilöistä kerätyllä tiedolla on suuri merkitys yksityisyydensuojalle. Samalla sillä on jo tänä päivänä merkittävää ja jatkuvasti kasvavaa yhteiskunnallista, taloudellista ja vaihdannallista arvoa. Yksityisyydensuojaa koskevien huolien lisäksi henkilöistä kerätyn tiedon hyödyntämistä vaikeuttaa nykyisin tiedon ja sen hallinnan pirstaleisuus.

My Datalla, suomennettuna omadatalla, viitataan ihmiskeskeisiin henkilötiedon organisointitapoihin, jossa yksityisyydensuojan ja pirstaleisuuden haasteita pyritään ratkaisemaan asettamalla ihminen tiedon hallinnan keskiöön. Organisaatio tarjoaa keräämänsä yksilöä koskevan tiedon takaisin ihmiselle itselleen. Yksilö voi hyödyntää tietoaan sekä jakaa, vaihtaa tai myydä sitä edelleen muihin palveluihin. My Data -malli voisi samalla mahdollistaa muun muassa uudenlaisten palveluiden ja liiketoimintamallien kehityksen. Ihmiskeskeisyyden ansiosta kyseessä on perustavanlaatuinen systeeminen muutos, jolla on vahvoja, vakiintunutta toimintaa muuttavia yhteiskunnallisia ja sosiaalisia ulottuvuuksia.

Liikenne- ja viestintäministeriön tilaama My Data -selvitys on samalla laajempi keskustelunavaus, joka haastaa kansalaisia, yrityksiä ja hallintoa pohtimaan uudenlaisen henkilötietomallin mahdollisuuksia ja vaikutuksia. Selvitys on toteutettu ministeriön digitalisaatiosta ja digitaalisista palveluista kasvua hakevan Kide-ohjelman puitteissa. Ohjelma edistää digitaalisten palveluiden syntyä pilottihankkeiden ja kokeilevan toiminnan avulla. Selvityksen toteutuksesta ja sisällöstä vastaa Open Knowledge Finland ry.

Selvitys toimii yleiskatsauksena My Datan eri alueisiin ja luo osaltaan yhteistä kieltä, mikä selkeyttää keskustelua laajan aihepiirin ympärillä. Samalla selvitys luo pohjaa My Datasta kiinnostuneiden toimijoiden verkostolle sekä jatkotyölle.

Uutuutensa vuoksi My Dataan liittyy monia keskeneräisiä kysymyksiä, kuten henkilötietoon liittyvät eri osapuolten intressit, oikeudet sekä käytännön tekniset ratkaisut. My Data linkittyy vahvasti myös suurten tietoaineistojen, big datan, hyödyntämiseen. Kehitys My Datan ympärillä on avoimista kysymyksistä huolimatta maailmanlaajuisesti vauhdikasta ja näkyy osaltaan myös globaaleiden toimijoiden erilaisissa palveluissa. Suomessa My Data -ajattelu tulee yhä vahvemmin esiin esimerkiksi erilaisissa tutkimushankkeissa.

Selvitystä tehtäessä on kerätty ideoita ja ajatuksia lukuisista käynnissä olevista keskusteluista Suomessa ja kansainvälisesti. Kirjallisuustietoa aihepiiristä on vielä niukasti. Selvitys kannustaa kehittämään käytännön My Data- kokeiluja, joilla erilaisia malleja ja niiden toimivuutta voidaan testata ja levittää.

Emil Asp
Liikenne- ja viestintäministeriö

Sisältö

Infoboksit:

Johdanto

My Data on ihmiskeskeinen lähestymistapa henkilötiedon hallintaan ja käsittelyyn. Siinä ihmisille annetaan oikeus ja pääsy heistä kerättyyn dataan kuten ostotietoihin, liikennetietoihin, teletietoihin, terveystietoihin, taloustietoihin ja eri verkkopalveluihin kertyvään dataan. Keskeistä on ihmisten mahdollisuus siirtää tietojaan nykyistä uudelleenkäytettävämmässä muodossa itselleen tai valtuuttamaansa palveluun hyödynnettäväksi. My Data on maailmanlaajuisesti kehitysvaiheessa oleva ilmiö, malli ja tulevaisuusskenaario, jonka ympärille on kertymässä kasvavaa vauhtia teknologiaa ja liiketoimintaa.

Kerätyn ja tallennetun tiedon määrä on digitalisaation seurauksena jatkuvassa ja nopeassa kasvussa, samoin sen liiketaloudellinen ja muu hyödyntäminen. Suuri osa tästä tiedosta on yksilöihin liittyvää henkilötietoa. Puhutaankin ihmisten kasvavasta digitaalisesta jalanjäljestä. World Economic Forum on arvioinut henkilötiedon yhdeksi merkittävimmistä tulevaisuuden liiketoimintakenttää muokkaavista voimista (World Economic Forum 2013). Henkilötiedon avulla voidaan vaikuttaa muun muassa terveydenhoitoon, kehittää sovelluksia oman elämän hallintaan ja itsestä oppimiseen sekä toteuttaa palvelujen räätälöintiä ja kohdistettua markkinointia yksilöiden ehdoilla ja yhteiskunnan kehitystä toivottuun suuntaan ohjaavasti. Esimerkiksi ostosdatan pohjalta voidaan tarjota palveluita, jotka auttavat ihmisiä muokkaamaan kulutustottumuksiaan nykyistä säästävämmiksi, ekologisemmiksi, eettisemmiksi tai terveellisemmiksi.

Henkilötiedon laajempaa hyödyntämistä varjostavat yksityisyyden katoamiseen liittyvät tulevaisuudenkuvat. Varsinkin Yhdysvaltain turvallisuusviranomaisten (NSA) massiivisen tiedonkeräysjärjestelmän paljastumisen jälkeen monessa maassa on käynnistynyt liikkeitä, joiden tavoitteena on vahvistaa yksityisyyden suojaa. On yleistä, että yksityishenkilöt kokevat yritysten ja valtioiden tietävän heistä liikaa, mikä on heistä epämiellyttävää. Toisaalta ihmiset eivät ymmärrä tapoja, joilla henkilötietoa hyödynnetään esimerkiksi sosiaalisen median suosittelupalveluissa tai kohdistetussa verkkomainonnassa, eivätkä hahmota mitä heistä kerättyä tietoa milläkin organisaatiolla on. Yhdysvaltain kuluttajasuojaviranomaisen Federal Trade Comissionin mukaan yrityksillä, jotka keräävät ja myyvät tietoa kuluttajista, on hallussaan tietovarannot, jotka kattavat lähes jokaisen yhdysvaltalaisen kuluttajan, ja jotka on kerätty pääosin ilman, että kuluttajat itse tietävät (FTC 2014).

Perinteinen yksityisyyden suojaamisen lähtökohta on, että ”mitä vähemmän henkilötietoa kerätään ja jaetaan, sen parempi”. Tämä jättää kuitenkin huomiotta datasta yksilölle itselleen kertyvän arvon ja on vastakkainen henkilötiedon määrän ja käytön lisääntymisen megatrendille. My Data -ajattelun tavoitteena on henkilötietoon liittyvän liiketoiminnan kehitys ihmiskeskeisesti niin, etteivät monopolistiset skenaariot (Newman 2013) tietoyhteiskunnan tulevaisuudesta toteutuisi. Se mahdollistaisi henkilötiedon keräämisen ja käytön niin, että hyödyt maksimoidaan ja yksityisyydensuojan heikkeneminen minimoidaan. Keskeinen keino henkilöihin liittyvän datan hyödyntämisen ja tietosuojan yhteensovittamisessa on vahvistaa yksilöiden asemaa, oikeuksia ja käytännön mahdollisuuksia heitä koskevien tietojen hallintaan.

Fyysisessä maailmassa yksityisyys ja julkisuus tunnistetaan selkeästi; meillä on globaali malli kodin (yksityisyys) ja kadun (julkisuus) käsitteille ja lainsäädäntö, joka suojaa kotirauhaa. Vaikka henkilötietolainsäädäntö pyrkiikin suojaamaan ihmisten yksityisyyttä digitaalisessa maailmassa, niin siellä ei ole vastaavaa käsitteellistä selkeyttä. My Data selkeyttää yksityisyyden suojaamisen ja henkilötiedon hyödyntämisen välisiä suhteita digitaalisessa maailmassa.

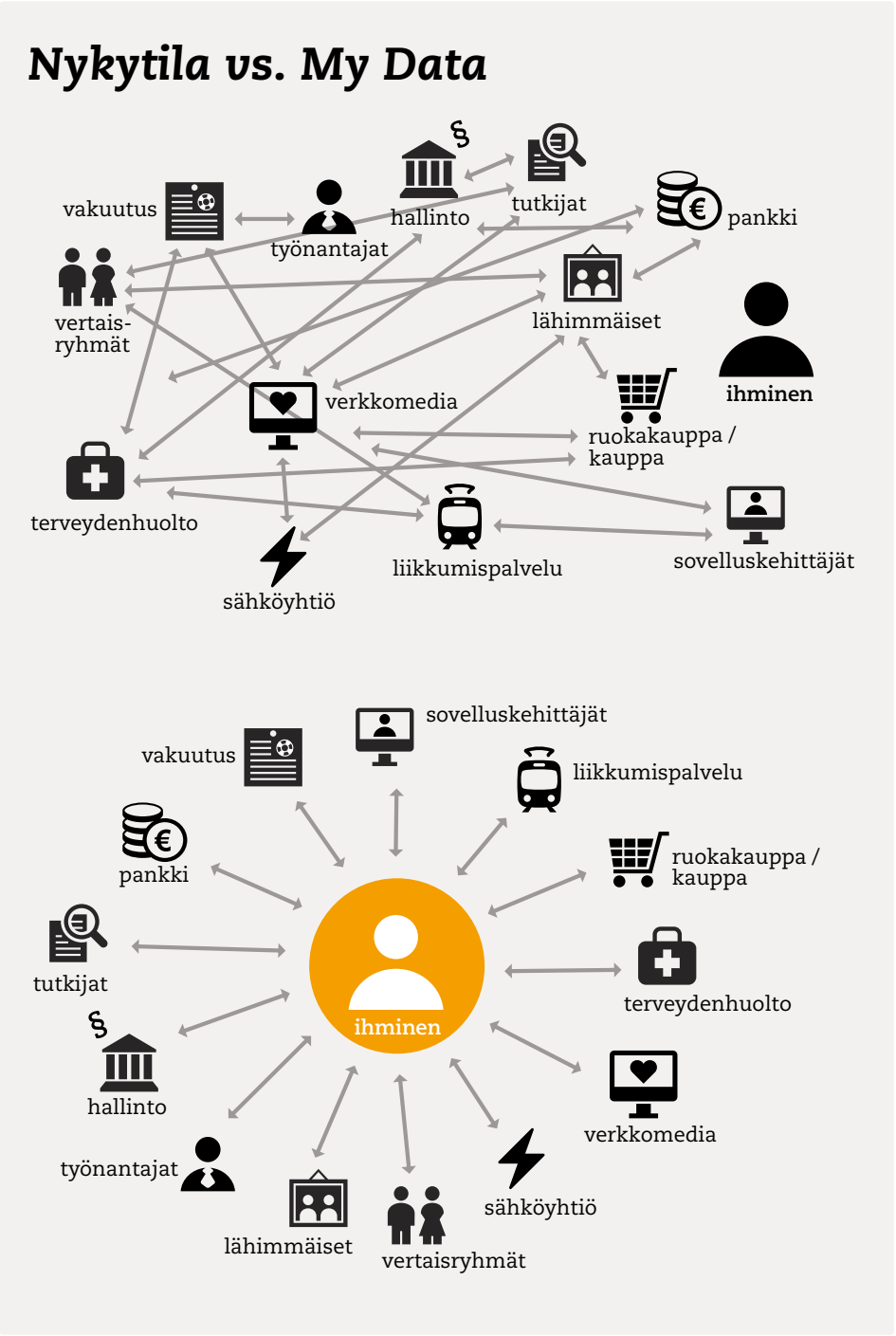
My Data -ajattelun lähtökohdat

My Datassa yhdistyy ihmiskeskeinen ajattelu, teknologia- ja data-ajattelu sekä liiketoimintanäkökulma. Lukija varmasti tunnistaa näistä itselleen tutuimman ja toisaalta joku muu suuntaus saattaa tuntua jopa häiritsevältä; *miksi korostetaan ihmisten oikeuksia, miksi puhutaan näin paljon datasta ja teknologiasta tai miksi yritysten hyötyjä alleviivataan*. Yhdessä ihmiskeskeinen lähestymistapa, tiedon hyödynnettävyys ja liiketoimintamallien avautuminen mahdollistavat kuitenkin systeemisen muutoksen tavassa, jolla henkilötietoa ja sen hyödyntämistä organisoidaan. Näitä eri lähtökohtia on pyritty selvityksessä käsittelemään mahdollisimman tasapainoisesti. Seuraavassa luvussa esiteltävät *My Data -periaatteet* konkretisoivat lähtökohtia ja viitoittavat tietä siihen, miten systeeminen muutos voi toteutua.

Ensimmäinen lähtökohta on **ihmiskeskeisyys**, jossa yhteiskunnan toimintaa rakennetaan ihmisten ympärille, vastapainona suuntaukselle, jossa ihminen redusoidaan järjestelmien osaksi ja keskitytään organisaatioiden toimintaedellytyksiin. Yhteiskunnan toiminta perustuu kasvavassa määrin tiedon keräämiseen ja hyödyntämiseen. Kansalaiset eivät ole muutoksen kohde vaan muutoksen tekijöitä. On ratkaiseva ero sillä, suunnitellaanko tiedon keräämisen ja hyödyntämisen mekanismit ihmisten vai organisaatioiden näkökulmasta. Oikeus omaan dataan ja sen hallintaan on rinnastettavissa digitaalisen ajan ihmisoikeudeksi.

Toisaalta My Data on avoimen tiedon kanssa rinnakkainen ideologia, joka korostaa **tiedon hyödynnettävyyttä** ja läpinäkyvyyttä niin, että yksilönsuoja otetaan huomioon. My Data -ajattelussa yksityisyydensuojaan liittyviä ongelmia ratkaistaan vahvistamalla yksilöiden mahdollisuuksia hyödyntää itse omaa dataansa ja hallita, kuinka sitä kerätään, jalostetaan, hyödynnetään ja jaetaan edelleen. Keskeistä on, että henkilötieto on teknisesti helposti käytettävissä.

Kolmas lähtökohta My Datalle on henkilötiedon hyödyntämiseen liittyvien **liiketoimintamallien avautuminen** kehitykselle, kilpailulle ja yhteistoiminnalle. My Data -ajatteluun kuuluu pyrkimys kehittää henkilötiedon ympärille syntyviä sovellus-, palvelu- ja liiketoimintarakenteita suuntaan, jossa yksilöillä on keskeinen päätösvalta hänestä itsestään kerättyyn dataan. Ihmiset voisivat esimerkiksi tarjota heille itselleen tai heidän käyttämiinsä palveluihin kertynyttä dataa myös uusille palveluille. Tästä seuraa datan arvoketjun pilkkoutuminen, mikä tarjoaa yrityksille mahdollisuuksia erikoistumiseen.

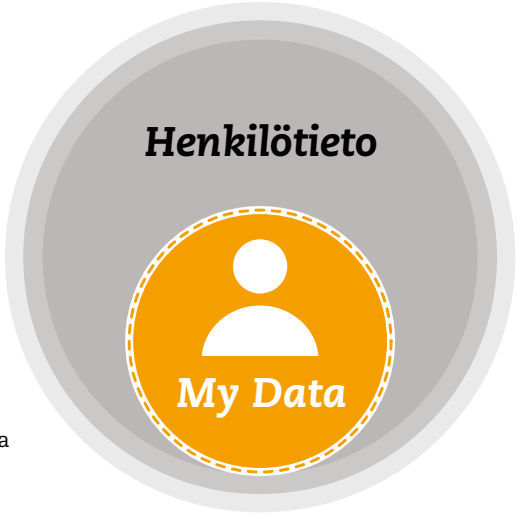


Kuva 0.1: My Data -lähestymisessä pyritään ratkaisemaan henkilötiedon organisointi ihmiskeskeisesti ja näin saavuttamaan henkilötiedon organisoinnin systeeminen yksinkertaistuminen, kun organisaatioiden välistä tiedon luovutusta määrittävät sopimukset korvataan yhdenmukaisilla yksilön itse hallinnoimilla sopimuksilla.

My Data käsitteenä

Kun kysytään, mitä My Data on, niin ensimmäinen helposti lähestyttävä tapa on määritellä se datalähtöisesti: tämä data on My Dataa ja tämä data ei ole My Dataa. Toinen laajempi näkökulma on systeeminen, joka kuvaa haluttua henkilötiedon organisoinnin muutosta laajemmin mukaan lukien periaatteiden, teknologian, liiketoimintamallien ja organisaatioiden toiminnan muutokset. Selvityksessä keskitytään pääasiassa systeemisen muutoksen eri osa-alueiden kuvaamiseen, mutta datalähtöinen määritelmä käsitellään ennen sitä.

My Data
Ihmiskeskeinen malli henkilötiedon organisointiin



Kuva 0.2: Datalähtöisessä määrittelyssä My Data on henkilötiedon osajoukko, joka toteuttaa My Data -rajapinnoille asetetut tiedon saatavuuteen liittyvät ehdot (luku 2). Datalähtöisen määritelmän lisäksi My Data voidaan määritellä systeemisenä muutoksena henkilötiedon organisointiin.

Termiä **henkilötieto** (personal data) käytetään selvityksessä kuvaamaan laajasti kaikkea henkilöön liittyvää tai henkilön toiminnan seurauksena syntyvää dataa. Esimerkiksi videopalvelu Netflix kerää dataa käyttäjien median kulutuksesta: mitä ohjelmia he katsovat, kuinka pitkään, mihin kellon aikaan ja milloin he pitävät taukoja jne. Kauppaketjuilla on ostostietoa, verkkopalvelut keräävät käyttäjädataa, teleoperaattorille jää tietoa puheluista ja puhelimen liikkeistä. Julkishallinnolla on paljon henkilötietoa aina kirjastojen lainaustiedoista ja terveystiedoista rikosrekistereihin. Pankeilla ja luottokorttiyhtiöillä taas on tiedot yksittäisistä maksutapahtumista ja esimerkiksi maksurästeistä. Henkilötiedolla ei tässä siis tarkoiteta vain kaikkein yksilöivimpiä tietoja, kuten nimeä ja osoitetta, eikä välttämättä rajoituta edes henkilöiden tunnistettavuuteen pohjautuvaan lainsäädännön mukaiseen määritelmään (luku 5), joka on laaja sekin. My Datan kannalta ei ole tärkeää määritellä henkilötietoa tarkan rajaavasti vaan keskittyä käytännöllisesti hyödynnettävään tietoon. Henkilötieto-termin käyttäminen laajennetussa ja hieman epätarkasti rajatussa muodossa saattaa häiritä lainsäädäntöä tuntevia lukijoitamme. Vaihtoehtoisesti olisimme voineet käyttää juridisesta termistä eroavaa sanaa henkilödata.

Kaikki henkilötieto voi olla My Dataa, mikäli se on henkilön itsensä saatavilla ja hallittavissa My Data -periaatteiden mukaisesti (luku 1). Mikäli yksilöllä ei ole mahdollisuutta hyödyntää itse ja hallinnoida jonkun muun hänestä keräämää henkilötietoa, niin sitä ei voida vielä kutsua My Dataksi. Näin määriteltynä **My Data on henkilötiedon osajoukko**. Kaikki My Data on henkilötietoa, mutta kaikki henkilötieto ei ole My Dataa. Tavoitteena on, että My Datan osuus henkilötiedosta kasvaisi nopeasti ja että tulevaisuudessa saavutetaan tilanne, jossa organisaatioille, yhteiskunnalle ja yksilölle on käytännöllisin ja taloudellisesti kannattavin vaihtoehto hallinnoida henkilötietoa yhteisesti sovittujen mallien mukaisesti. Täydelliseen kattavuuteen ei kuitenkaan ole realistista tai edes tavoiteltavaa pyrkiä. Esimerkiksi rikolliset mielellään poistaisivat omat tietonsa rikosrekisteristä, mutta sitä ei pidä sallia. Näin rikosrekisteri ei koskaan voi olla täysin My Dataa.

My Data terminä on herättänyt paljon keskustelua, ja on syytä taustoittaa sanavalintaa. Käsitteenä alun perin Britanniasta lähtenyt My Data on vielä vakiintumaton. Käytössä on myös ilmaisuja MyData ja midata. Kotimaisten kielten keskus on ehdottanut suomennokseksi omadata. Tässä selvityksessä käytetään termiä My Data kahdesta syystä. Ensinnäkin ilmiö on kansainvälinen ja toivotaan, että kehitys Suomessa ja maailmalla yhdistyvät suoraviivaisesti toisiinsa. Kun pyritään kansainväliseen yhdessä selkeästi sovittuun ja erityiseen toimintamalliin, on tarkoituksenmukaista käyttää yhteistä nimitystä. Vastaavalla tavalla on toimittu eräiden muidenkin kansainvälisten ja uusien käsitteiden kuten käsit-

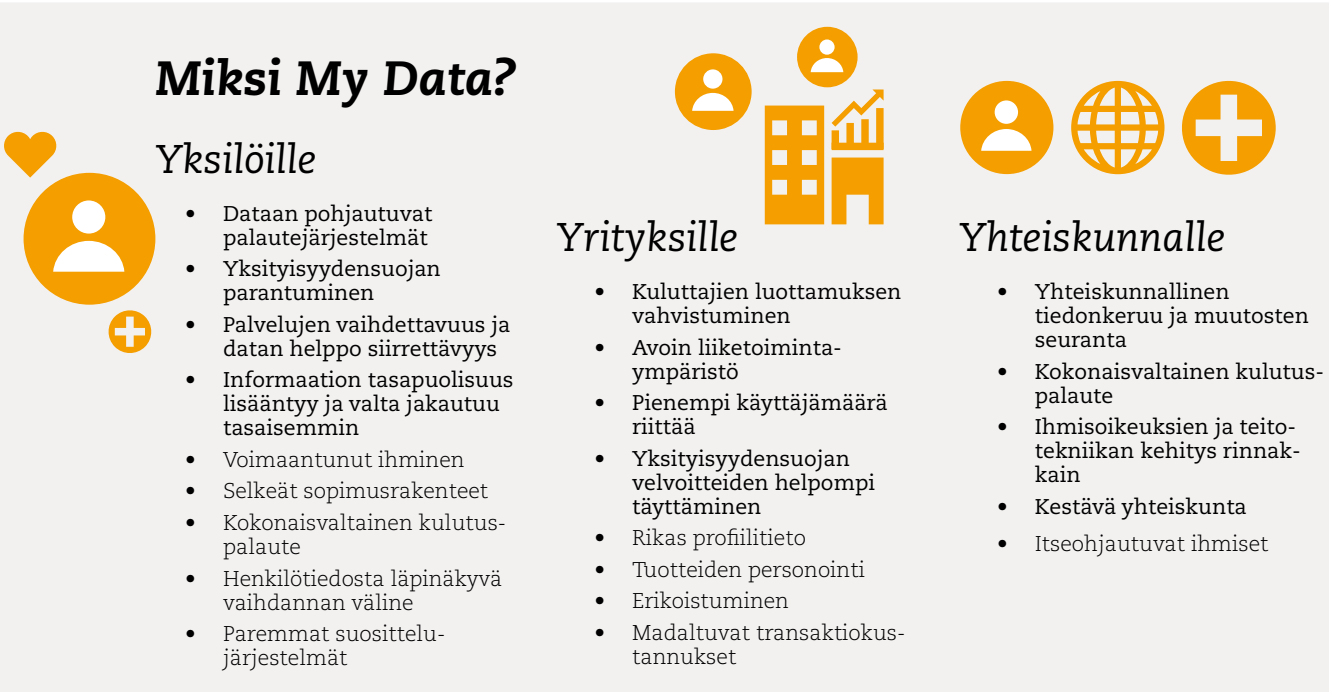
teen big data viestinnässä (esim. LVM 2014 raportti “Big datan hyödyntäminen”[1]). Toiseksi My Data on jo Suomessakin monessa yhteydessä omaksuttu käsite, ja sekaannusten välttämiseksi on paras pitäytyä yleistyneimmässä käsitteessä, kunnes Suomessa mahdollisesti vakiintuu jokin suomenkielinen termi tässä selvityksessä kuvatulle kehitykselle.

My Datan hyötyjä

My Datan kehitys on maailmanlaajuisesti alkutaipaleella. Tällä hetkellä on vasta alustavasti hahmottumassa, millaisia teknologisia ja organisatorisia rakenteita My Datan toteutuminen edellyttää. Yhtenäistä tapaa tai avointa standardia henkilötietojen käsittelyyn ei ole vielä syntynyt, eikä toistaiseksi mikään yksittäinen yritys tai organisaatio ole päässyt henkilötiedon keräämisen, käsittelyn ja yhdistämisen monopoliksi. Monet teknologiaan, organisaatioihin ja periaatteisiin liittyvät yksityiskohdat herättävät kuitenkin jo paljon kiinnostusta ja kiivasta keskustelua. Ennen kuin paneudutaan näihin ja muihin yksityiskohtiin syvemmälle, hahmotellaan tavoitetilaa: Mihin My Datalla pyritään ja mitä hyötyä siitä on?

Hyödyistä puhuttaessa on syytä pitää mielessä, että kaikissa teknologioissa on myös riskejä, joista osaa ei voi edes ennakoida. Mahdollisuuksien ja uhkien toteutuminen ei useimmiten ole riippuvainen teknologiasta itsestään, vaan siitä, miten ja millaisissa yhteyksissä sitä käytetään. Riskien ja hyötyjen vastakkainasettelu ei ole hedelmällistä vaan keskustelu on parempi sitoa käyttötarkoituksiin. Sen sijaan, että puhutaan abstraktisti yksityisyydensuojasta, pitäisikin puhua datan käyttötarkoituksista suhteessa yksityisyyteen. My Dataan liittyviä esteitä ja riskejä käsitellään tarkemmin luvussa 5.

Alla on esimerkkejä esille nousseista My Datan positiivisia vaikutuksia yksilöiden, yritysten ja yhteiskunnan kannalta. Hyötyjen listaus ei ole tyhjentävä, eikä siinä pyritä vertailemaan hyötyjä toisiinsa tai asettamaan niitä tärkeysjärjestykseen.



Kuva 0.3: My Datan hyötyjä yksilön, yritysten ja yhteiskunnan näkökulmasta. My Data on yhteiskunnallinen muutoshanke, jonka vaikutuksia ei voida mitata pelkästään taloudellisin mittarein.

1 http://www.lvm.fi/julkaisu/4417803/big-datan-hyodyntaminen

Ihmisille

- Dataan pohjautuvat palautejärjestelmät, kuten liikunnan, ravinnon, talouden, ajankäytön jne. seurantatyökalut auttavat ihmisiä oppimaan itsestään ja ympäristöstään ja muokkaamaan toimintaansa. My Datan myötä datan saatavuus lisääntyisi. Ihmiset voisivat esimerkiksi seurata ajantasaisesti remonttibudjettiaan kaupan ostojen perusteella tai tarkastella todellista ruokavaliotaan, rakentaa perheen kuukausibudjettia ja asettaa tavoitteita omassa arjessaan.
- Yksityisyydensuoja parantuu, koska My Data -periaatteiden toteutuessa yksilöllä on paremmat mahdollisuudet ymmärtää ja hallita organisaatioiden hänestä keräämää tietoa. Lisäksi organisaatioiden keräämän tiedon läpinäkyvyys paranee, jolloin yksilöt, kansalaisjärjestöt ja muut yhteiskunnalliset toimijat voivat entistä paremmin puuttua yksityisyyttä loukkaaviin väärinkäytöksiin.
- Palveluiden vaihdettavuus ja datan helppo siirrettävyys My Data -periaatteiden mukaan suojaavat yksittäiseen palveluun lukkiutumiselta. Ihmiset voivat tarpeen vaatiessa vaihtaa datan uuteen paikkaan. Tilanne on verrattavissa siihen, että teleoperaattorin vaihtaminen helpottui, kun sai säilyttää vanhan puhelinnumeronsa. Esimerkiksi nykyisin on vaikeaa vaihtaa terveystietoa palveluntarjoajien väillä, koska omia tietojaan ei saa mukaan.
- Informaation tasapuolisuus lisääntyy ja valta jakautuu nykyistä tasaisemmin henkilötietoa keräävien organisaatioiden ja ihmisten välillä. Esimerkiksi verkkopalveluissa ja markkinoinnissa yleisesti käytetyistä suosittelujärjestelmistä tulee läpinäkyvämpiä, kun ihmisillä on pääsy samaan tietoon, johon suositukset pohjautuvat.

Yrityksille

- Kuluttajien luottamus on yhä merkittävämpi arvo yrityksille. Asiakkaat haluavat ymmärtää, mitä heidän datallaan tehdään. My Data lisää yksilön ja organisaation suhteen läpinäkyvyyttä, mahdollistaa luottamuksen rakentamisen ja tuottaa yrityksille goodwill-hyötyä.
- Avoin liiketoimintaympäristö tarjoaa pienille ja uusille toimijoille tasavertaisen mahdollisuuden päästä käsiksi samaan tietoon, mihin nyt pääsevät suuret kansainväliset toimijat. Tasavertaisuus toteutuu, kun ihmiset itse päättävät, ketkä heidän dataansa saavat hyödyntää ja miten. Nykyisin suuret tiedon kerääjät saavat merkittävän edun henkilötietoihin pohjautuvaan liiketoimintaan, koska tiedon kerääminen ja hyödyntäminen kulkevat käsi kädessä.
- Pienempi käyttäjämäärä riittää, koska My Data -palvelut ovat yhteentoimivia ja keskenään vaihdettavia datan helpon siirrettävyyden ansioista. Nykyään verkkopalvelujen menestystä määrittää usein niin sanottu voittaja-saa-kaiken -ilmiö (winner takes all) ja moni hyvä palvelu kuolee, koska ei saavuta kriittistä käyttäjämäärää.
- Yksityisyyden suojan velvoitteiden helpompi täyttäminen edistää palvelujen kehittämistä. My Data -käytäntöjen avulla hyväksynnän pyytäminen henkilötietojen hyödyntämiseen onnistuu suoraan ihmisiltä itseltään erotuksena nykytilaan, jossa kaikkeen henkilötiedon käsittelyyn tarvitaan lakimiesten hyväksyntä.

Yhteiskunnalle laajemmin

- Yhteiskunnallinen tiedonkeruu ja muutosten seuranta parantaa päätöksentekoa. My Data -periaatteiden mukaisesti yksilöillä on mahdollisuus antaa lupa omien tietojensa hyödyntämiseen tutkimuskäytössä. Vaikeiden ongelmien ratkaiseminen helpottuu, kun tutkijat ja päätöksentekijät saavat työkaluja entistä kattavampaan tiedonkeruuseen.

- **Kokonaisvaltainen kulutuspalaute** mahdollistuu, kun dataa voidaan yhdistellä eri lähteistä. Tällä hetkellä yksilön on vaikea tietää kattavasti, mitä hän kuluttaa ja millaisia vaikutuksia sillä on. Palaute auttaa ihmisiä kulutuspäätösten tekemisessä ja ohjaa palvelutuotantoa. Esimerkiksi velvoite autojen kulutuslukemien ilmoittamiseen on vaikuttanut uusien autojen ominaisuuksiin. Teollisuus on muuttanut tuotteitaan sen jälkeen, kun kuluttajille on esitelty selkeitä mittareita, joiden avulla ostopäätöstä voidaan suhteuttaa muuhunkin kuin hintaan.
- **Ihmisoikeuksien ja tietotekniikan kehitys** ovat kulkeneet pitkälti toisistaan erillään. My Data on ihmiselle keino ottaa digitaaliset oikeudet haltuunsa. Ihmisellä on oltava oikeus ja mahdollisuus hallita omaa digitaalista identiteettiään siinä missä hänellä on oikeus ajattelun ja ilmaisun vapauteen kansalaisena.
- **Kestävä yhteiskunta** on taloudellisesti, ekologisesti ja sosiaalisesti kestävällä pohjalla. Tieto on yhteiskunnan keskeinen resurssi, jonka hallinta ja sujuva hyödyntäminen ovat kestävän yhteiskunnan peruspilareita. My Data tarjoaa henkilötiedon organisoinnille pitkäjänteisen ja sektoririippumattoman organisointimallin, joka toimii kestävänä pohjana tietoyhteiskunnalle.

Selvityksen sisältö

My Dataa on hahmoteltu luvussa 1 esitettyjen periaatteiden ja henkilötiedon yksinkertaistetun arvoketjun kautta. Periaatteiden toteutuminen johtaa henkilötiedon arvoketjun pilkkoutumiseen siten, että eri vaiheisiin voi syntyä erikoistuneita toimijoita. Arvoketjun osa-alueita tarkastellaan luvuissa: *My Data -lähteet* (luku 2), *My Data -palveluinfrastruktuuri* (luku 3) ja *My Datan hyödyntäminen* (luku 4).

Kuva 0.4: Luvussa 1. avataan My Datan periaatteet ja My Datan vaikutukset henkilötiedon arvoketjuun. Arvoketjun eri vaiheet (datan tuottaminen, välittäminen ja hyödyntäminen) on tarkemmin käsitelty luvuissa 2–4.

Luku 1. My Data -periaatteet

Oikeus ja mahdollisuus hallita omaa dataa
Ihmisillä on oikeus ja käytännön mahdollisuus hallita omia henkilötietojaan.

Kattava ja käytännöllinen tiedon saatavuus
Henkilötieto on ihmisille itselleen saatavilla koneluettavasti ja riittävän ajantasaisesti rajapintojen kautta.

Hallinnan hajauttaminen ja yhteentoimivuus
My Datan hallinnointi ja säilytys on mahdollista hajauttaa ja palvelut ovat vaihdettavissa, mutta kokonaisuus on yhteentoimiva ja looginen.

Henkilötiedon arvoketjun pilkkoutuminen ihmiskeskeisesti

Luku 2. My Datan lähteet
(rajapinnat ja standardit)
Miten henkilötietoa syntyy?
Miten henkilötieto muuttuu My Dataksi?
Hallittava standardoitu sopiminen
Rajapintoihin liittyvät standardit ja formaatit

Luku 3. Palveluinfrastruktuuri
(välittäminen ja hallinta)
Henkilötiedon organisointitapoja
Hallintapalvelut
Tallennuspalvelut
Autentikaatio- ja luottamuspalvelut
Anonymisointipalvelut

Luku 4. My Datan hyödyntäminen
(sovellukset ja palvelut)
Itsestään ja ympäristöstään oppiminen
Paremmin kohdennetut tuotteet ja markkinointi
My Datan design huomioita

Henkilötiedon luominen, välitys ja hyödyntäminen on mahdollista eriyttää eri paikkoihin ja eri palveluihin.

Luku 1: My data -periaatteet ja henkilötiedon arvoketju
Kansainvälisessä keskustelussa yksityisyyden suojan edistäjät, tekniikan ja juridiikan osaajat, yrittäjät ja muut asiantuntijat, kehittävät yhdessä periaatteita, jotka konkretisoivat My Datan -ajattelun kolmea lähtökohtaa: ihmiskeskeisyyttä, tiedon hyödynnettävyyttä, ja liiketoimintamallien avautumista. Keskeiset periaatteet ovat *1. yksilöiden oikeus ja mahdollisuus hallita omaa dataansa, 2. henkilötiedon kattava ja käytännöllinen saatavuus sekä 3. henkilötiedon hallinnan hajauttaminen ja yhteentoimivuus.* Periaatteet ohjaavat My Data -rajapintojen ja standardien, välittämiseen ja hallintaan liittyvän palveluinfrastruktuurin sekä My Dataa hyödyntävien sovellusten ja palvelujen kehitystä. Luvussa 1 esitellään myös henkilötiedon arvoketju, jonka eri osia käsitellään tarkemmin seuraavissa luvuissa.

Luku 2: My datan lähteet
Potentiaalisia My Datan lähteitä ovat yritysten ja muiden organisaatioiden tietokannat sekä yksilön itse keräämä tieto. Datalähteitä hyödynnetään tulevaisuudessa pääosin rajapintojen kautta. Niiden kehittäminen palveluihin mahdollistaa henkilötiedon käytännöllisen saatavuuden ja hyödynnettävyyden. Henkilötietorajapintoja kehitetään jatkuvasti, mutta niille ei vielä ole yleisiä standardeja. Palvelujen määrän kasvaessa käyttäjän on vaikea muodostaa kokonaiskuvaa siitä, minkä palvelujen välillä hänen tietonsa liikkuu. My Data -periaatteiden mukaiset rajapinnat antavat yksilölle mahdollisuuden hallita omien tietojensa hyödyntämistä ja keräämistä kokonaisuutena tuomalla nykyisiin datarajapintoihin lisäominaisuuksina koneluettavassa muodossa saatavat henkilötiedon käyttöselosteet ja yksityisyysasetusten etähallinnan. Luvussa 2 käsitellään rajapintojen lisäksi myös My Dataan liittyviä standardeja, dataformaatteja ja teknologioita.

Luku 3: My data -palveluinfrastruktuuri
Henkilötiedon arvoketjussa datan lähteiden ja My Data -sovellusten väliin kehittyy palveluinfrastuktuuria, jolla tarkoitetaan muun muassa hallinnointi-, anonymisointi- ja säilytyspalveluita, My Data -tilejä sekä organisaatioita, jotka tarjoavat näitä palveluita. Palveluinfrastruktuurin organisoimiseen on eri vaihtoehtoja, joista selvityksessä parhaaksi nousi ns. operaattorimalli. Se muistuttaa hieman nykyisten pankki- ja telekommunikaatiosektorien toimintaa. Siinä datalähteiden hallintapalvelua tarjoaisivat useat keskenään yhteentoimivat My Data -operaattorit, joista käyttäjät voisivat valita itselleen sopivimman. Operaattorimallin vaihtoehtoja ovat esimerkiksi yksittäinen toimija (aggregaattori) tai hajautetusti ilman ylläpitäviä organisaatioita täysin teknologiapohjaisesti toimiva järjestelmä, jossa kukin ihminen toimii “itseoperaattorina”. Käytännössä nämä eri organisoitumistavat voivat kehittyä ja osin varmasti kehittyvätkin rinnakkain.

Luku 4: My datan hyödyntäminen
Henkilötiedon helppo saatavuus, uudelleenkäytettävyys ja yhdisteleminen yksityisyydensuojan säilyttävällä tavalla mahdollistavat kokonaan uusien sovellusten ja palvelujen kehittämisen. Esimerkiksi liikenteessä, terveys- ja hyvinvointialalla ja työn organisoinnissa on jo käynnissä hankkeita, jotka perustuvat henkilötiedon entistä laajempaan hyödyntämiseen. My Data -periaatteiden mukainen lähestyminen tarjoaa näille hankkeille yhteisen ratkaisun yksityisyyden suojan ja tiedon keruun hyväksyttävyyden ongelmiin. Keskeistä on, että ihminen voi itse toimia oman datansa hyödyntäjänä ja jatkohyödyntäjänä suoraan tai erilaisten palvelujen kautta, mutta yksilön luvalla My Dataa voidaan hyödyntää myös esimerkiksi tutkimuksessa ja julkisella sektorilla. My Data mahdollistaa esimerkiksi profiilitiedon kerryttämisen useista lähteistä ja profiilien siirtämisen palvelusta toiseen.

Luvussa 5 tarkastellaan eri toimijoiden (yksilöt, yritykset, tutkimus, julkishallinto) näkökulmia ja lainsäädännön nykytilannetta suhteessa My Dataan sekä hahmotellaan askelmerkkejä tulevaisuuteen. Viimeisessä luvussa 6. My Data Suomessa tarkastellaan My Dataa kansallisesta näkökulmasta ja linkitetään sitä meneillään olevaan kehitykseen Suomessa.

My Data ja avoin data

Viime vuosina erityisesti julkishallinnon keräämiä tietovarantoja on avattu niiden hyötykäytön lisäämiseksi. Myös ihmisten henkilötiedon alueella selkeillä datan hallinnan periaatteilla, paremmalla yhteentoimivuudella ja datan siirrettävyydellä voidaan saavuttaa mittavia hyötyjä. Tällä hetkellä henkilötieto on rajusti alikäytetty 'raaka-aine' uusille palveluille.

Avoimella datalla ja My Datalla on ilmeisiä liittymäkohtia. Molemmissa tarvitaan yhteisistä periaatteista sopimista, viisasta sääntelyä sekä koneluettavia rajapintoja, standardeja ja palveluita tiedon hallittuun siirtämiseen, varastointiin, käsittelyyn ja analysointiin.

Avoimen tiedon määritelmän mukaan avoin aineisto on teknisesti ja juridisesti kenen tahansa vapaasti käytettävissä, uudelleen käytettävissä ja jaettavissa. Vastaavasti My Data voitaisiin määritellä niin, että se on teknisesti ja juridisesti datan kohteen itsensä vapaasti käytettävissä, uudelleenkäytettävissä ja jaettavissa.

My Datasta voi tulla avointa dataa:

- **muunnosten kautta** – Esimerkiksi suuri osa julkisista tilastoista syntyy kyselytutkimusten tai muiden yksilöiden henkilötietojen pohjalta yhdistelemällä, aggregoimalla ja anonymisoimalla. On tärkeää tietää, mitä haasteita anonymisointiin liittyy.
- **yksilöiden valinnan kautta** – Voi olla ihmisiä, jotka ovat valmiita avaamaan omaa henkilötietoaan hyödyttääkseen muita. Esimerkiksi vaikeasta sairaudesta kärsivä saattaa mielellään jakaa omaa terveysdataansa, jos voi siten edistää lääketieteellistä tutkimusta ja auttaa muita sairastuneita.

My Data -ajattelu korostaa oikeuksien ohella myös käytännön mahdollisuuksia. Esimerkiksi tiedon keräämiseen, hyödyntämiseen ja eteenpäin lähettämiseen liittyvien asetusten ja sopimusehtojen hallinnan pitää olla helppoa.

1. My Data -periaatteet

My Datan -ajattelun kolmea lähtökohtaa: ihmiskeskeisyyttä, tiedon hyödynnettävyyttä, ja liiketoimintamallien avautumista voidaan konkretisoida periaatteilla *1. yksilöiden oikeus ja mahdollisuus hallita omaa dataansa, 2. henkilötiedon kattava ja käytännöllinen saatavuus sekä 3. henkilötiedon hallinnan hajauttaminen ja yhteentoimivuus.* Nämä periaatteet ohjaavat My Data -rajapintojen ja standardien, välittämiseen ja hallintaan liittyvän palveluinfrastruktuurin sekä My Dataa hyödyntävien sovellusten ja palvelujen kehitystä. Periaatteiden toteutuminen johtaa henkilötiedon arvoketjun pilkkoutumiseen siten, että eri vaiheisiin (keruu – välittäminen – hyödyntäminen) voi syntyä erikoistuneita toimijoita. Tässä kappaleessa kuvataan ensin tarkemmin nämä periaatteet, jonka jälkeen analysoidaan periaatteiden vaikutuksia henkilötiedon arvoketjuun.

1.1. Periaatteet

1.1.1 Oikeus ja mahdollisuus hallita omaa dataa

Periaate: Ihmisillä on oikeus ja käytännön mahdollisuus hallita omia henkilötietojaan.

Henkilötiedon hallinta ei ole joko-tai-kysymys vaan hallintaa voi olla eri tasoista lähtien siitä, tietääkö ihminen edes itseään koskevan henkilötiedon olemassaolosta. Alla on jaoteltu hallinta erilaisiin oikeuksiin. Passiiviset oikeudet liittyvät yksilöiden mahdollisuuksiin tietää, mitä tietoja heistä on kerätty ja miten niitä käytetään. Nykyinen henkilötietolainsäädäntö osin jo antaa näitä oikeuksia yksilöille, mutta käytännön tasolla on paljon parannettavaa. Aktiiviset oikeudet liittyvät yksilön mahdollisuuteen saada data itselleen ja jakaa sitä edelleen sekä poistaa henkilötieto niin halutessaan.

Passiiviset

- **Oikeus tietää** – yksilöllä on oikeus tietää, mitä kaikkea henkilötietoa hänestä on olemassa.
- **Oikeus nähdä** – yksilö pääsee näkemään itseään koskevan henkilötiedon.
- **Oikeus oikaista** – yksilö voi oikaista häntä koskevan väärän henkilötiedon.
- **Oikeus valvoa** – yksilö voi tarkistaa, kuka hänen henkilötietoaan käsittelee ja miksi.

Aktiiviset

- **Oikeus saada** – yksilö saa halutessaan tiedon itselleen uudelleenkäytettävässä muodossa ja saa hyödyntää sitä itse haluamallaan tavalla.
- **Oikeus jakaa** – yksilö voi julkaista tiedon eteenpäin kolmansille osapuolille, ja tähän on tekniset edellytykset.
- **Oikeus poistaa** – yksilö voi poistaa henkilötietonsa sieltä missä ne ovat.

Vaikka kaikkia oikeuksia ei syystä tai toisesta voisi jonkin henkilötiedon osalta soveltaa, niin silti yksilön mahdollisuutta hallita omia henkilötietoja voidaan parantaa tukemalla mahdollisimman monia ylläluetelluista oikeuksista. Esimerkiksi sähköisen lääkemääräyksen ansiosta lääkärit näkevät, mitä voimassaolevia reseptejä potilaalla on. Muiden hyötyjen ohella tämä auttaa välttämään reseptihuijauksia, joissa potilas käy usealla lääkärillä pyytämässä reseptin samaan vai-

vaan ja välittää lääkkeet katukauppaan. Mikäli potilas voisi oman datan hallintaoikeuksien perusteella kieltää reseptitiedon näkymisen eri lääkäreille, tulisivat reseptihuijaukset jälleen mahdolliseksi, mikä ei ole toivottavaa. Ihmisille voitaisiin silti antaa oikeus saada reseptidata koneluettavassa muodossa itselleen muita käyttötarkoituksia varten.

My Data -ajattelu korostaa oikeuksien ohella myös käytännön mahdollisuuksia. Esimerkiksi tiedon keräämiseen, hyödyntämiseen ja eteenpäin lähettämiseen liittyvien asetusten ja sopimusehtojen hallinnan tulee olla helppoa.

Ei puhuta datan omistajuudesta

Arkijärjellä ymmärrettynä My Dataan liittyy ajatus omistajuudesta. Ihmisillä pitäisi olla oikeus omistaa omat tietonsa. Datan omistajuuden käsite on kuitenkin varsin hankala.

Omistaminen on helposti ymmärrettävissä irtaimiston tai kiinteän omaisuuden kohdalla, joiden omistaja voi määrätä niistä muut poissulkevasti. Tuolin omistaja voi yleensä päättää, kuka tuolilla saa istua tai minkä väriseksi tuoli maalataan. Toisen maalle ei saa rakentaa eikä toisen metsästä kaataa puita ilman lupaa.

Tiedon suhteen omistaminen ei ole näin suoraviivaista. Monet ihmiset voivat tietää samoja asioita, emmekä voi estää ketään tietämästä jotain. Taloustieteellisin termein: tieto on kilpailematon hyödyke. Se että yksi ihminen tietää jotain ja käyttää sitä hyödykseen, ei sinällään estä muita samanaikaisesti tietämästä ja hyödyntämästä samaa tietoa. Vastaavasti kun dataa kopioidaan ei yhden kopion käyttö estä muiden kopioiden käyttöä. Datan saatavuutta ja hallintaa voidaan toki rajoittaa niin, että käytännöllisesti vain harvoilla on mahdollisuus sitä hyödyntää.

Pääsääntöisesti tietoon tai dataan ei kohdistu yksinoikeuksia, kukaan ei omista tietoa. Sen sijaan joihinkin tietoihin voi kohdistua esimerkiksi tekijänoikeuksien, liikesalaisuuden tai yksityisyydensuojan takia rajoitetumpia oikeuksia. Tietoon kohdistuvat oikeudet ovat yleensä kielto-oikeuksia: ne antavat oikeudenhaltijalle mahdollisuuden kieltää muita hyödyntämästä tietoa.

Moniin tietoihin voi useilla osapuolilla olla perusteltu intressi. Esimerkiksi kaupalla on asiakassuhteessa hyvä syy saada käyttää keräämiään asiakastietoja, vaikka asiakkailla olisikin samoihin tietoihin oikeuksia, kuten mahdollisuus saada data itselleen tai poistaa data asiakassuhteen päättyessä.

My Datalla tavoitellaan sitä, että ihmisillä on oikeus ja todellinen käytännöllinen mahdollisuus hallita omia tietojaan. Omistajuuden käsitteen sijaan olemme päätyneet siis puhumaan hallinnasta (control). (Pitkänen 2014)

1.1.2 Kattava ja käytännöllinen tiedon saatavuus

Periaate: Henkilötieto on ihmisille itselleen saatavilla koneluettavasti ja riittävän ajantasaisesti rajapintojen kautta.

Henkilörekistereihin liitetyn tarkastusoikeuden nojalla ihmisillä on nykyisinkin mahdollisuus saada itseään koskevat tiedot (luku 5). Nykytasoinen tiedon saatavuus ei ole kuitenkaan riittävää sen jatkohyödyntämiselle tai uudenlaisen ihmiskeskeisen tietoarkkitehtuurin syntymiselle. Lain velvoite tarkastusoikeudesta toteutetaan usein niin, että henkilötieto lähetetään rekisteröidylle postitse, kun rekisteröity tekee tietopyynnön. Esimerkiksi teleoperaattorilta saa oman puhelinliittymän puhelu- ja paikkatiedot maksamalla noin 10 euron toimitusmaksun.

Tietojen tarkastamista laajemmille jatkokäyttötapauksille paperinen ja kallis datatuloste on hyödytön. My Datan minimitoteutus on, että datan saa ladattua koneluettavassa muodossa itselleen. Todellisten innovatiivisten sovellusten näkökulmasta minimitoteutus ei riitä, koska tieto kerätään yksittäisinä paketteina ja sen päivittäminen vaatii yksilöltä useita vierailuja tiedon tarjoajan sivustolla. Tällöin usean tietolähteen ja jatkuvasti päivittyvän tiedon hyödyntäminen ja hallinnointi on epäkäytännöllistä.

Saako yksilö hetkellisen tilanteen mukaisen otoksen henkilötiedoistaan vai saako hän rajapinnan kautta jatkuvan pääsyn ajantasaiseen dataan? Ajantasaisesti ja automaattisesti rajapintojen kautta saatava My Data avaa uusia käyttömahdollisuuksia. Esimerkiksi ostosdata on hyödyllisintä, jos sähköisen ”datakuitin” saa halutessaan automaattisesti heti ostoksen maksettuaan kuten paperikuitin nykyään.

Kattavan ja käytännöllisen tiedon saatavuuden mahdollistavat standardit, toimintatavat ja infrastruktuuri ovat vasta kehittymässä. Vaikka yksittäinen toimija toteuttaisikin mallioppilaan tavoin My Data -rajapinnat palveluunsa, ei se takaisi kokonaisuuden sujuvaa käytettävyyttä, koska yleiskäyttöisiä My Data -hallintapalveluita ei ole vielä tarjolla. Siirtymävaiheen ratkaisuilla voidaan kuitenkin osoittaa ihmiskeskeisen henkilötiedon keräämisen hyötyjä jo ennen kuin organisaatiot alkavat kehittämään tietojärjestelmiensä rajapintoja. Tietoa voidaan kerätä yksinkertaisemmistakin rajapinnoista tai jopa ilman rajapintoja hyödyntäen ohjelmistoja, jotka ovat erikoistuneita keräämään verkkosivuilla julkaistua tietoa.

1.1.3 Hallinnan hajauttaminen ja yhteentoimivuus

Periaate: My Datan hallinnointi ja säilytys on mahdollista hajauttaa ja palvelut voidaan vaihtaa, mutta kokonaisuus on yhteentoimiva ja looginen.

Keskeinen lähtökohta avoimen My Data -infrastruktuurin suunnittelussa on hajauttaminen. Ei haluta, että yksi organisaatio olisi infrastruktuurin tarjoajana monopoliasemassa, kaikki yksilön henkilötieto sijaitsisi yhdessä palvelussa ja hallinnointipalveluita voisi toteuttaa vain yhdellä teknologialla. Hajautus parantaa yksityisyyden suojaa ja mahdollistaa avoimen kilpailun, mikä kiihdyttää teknologioiden ja palvelujen kehitystä. Hajautus on tärkeää myös sosiaalisessa mielessä, koska se mahdollistaa erilaiset identiteettimallit ja autentikaatiotasot. Ihmisellä voi olla esimerkiksi useita My data -tilejä, joista osa on vahvasti tunnistettuja ja virallisia ja osa kevyemmin esimerkiksi vain sähköpostilla tunnistettuja ja osa pseudonyymitilejä, siis yksilöityjä tilejä ilman henkilön omaa nimeä.

Hajautus saattaa lisätä järjestelmien monimutkaisuutta, joskin monissa tapauksissa hajautus on myös keino hallita suuria kokonaisuuksia. Hajautuksissa täytyy pyrkiä siihen, että kokonaisuus on toimiva ja käytettävä. Tekninen ja organisatorinen selkeys on hajautuksen ohella yksi My Data -infrastruktuurin tavoiteltava perusominaisuus. Järjestelmän toiminnan ymmärrettävyys lisää ihmisten luottamusta ja edistää siten järjestelmän hyödyntämistä.

My Data -operointipalvelut, jotka avustavat yksilöä tiedon hallinnoinnissa ja jalostamisessa, pitäisi hajauttaa kuten pankki- tai teleoperaattoripalvelut. Joku voi luottaa talousdatansa pankille ja terveysdatansa terveysalan My Data -operaattorille, kun taas toinen ei luota muuhun kuin omalla koneellaan olevaan avoimen lähdekoodin ohjelmistoon. Hajautuksessa eri operaattoreiden, myös ”itseoperaattoreiden”, pitäisi toimia yhdessä – vastaavasti kuin pankkikortilla maksaminen toimii pankista riippumatta.

Hajautus ja yhteentoimivuus ovat osin toisiaan tukevia ja osin toisiaan hankaloittavia periaatteita. Hajautus on kriittistä avoimen ja kestävän liiketoimintaympäristön kehitykselle, mutta asettaa erityisvaatimuksia yhteentoimivuudelle. Avoimilla standardeilla pyritään siihen, että monet My Datan hallinnointipalvelut ja niitä tarjoavat organisaatiot toimivat hyvin keskenään niin, että palvelujen vaihtaminen ja datan siirtäminen sujuisi vaivattomasti. Avoimia standardeja käsitellään luvussa 2.

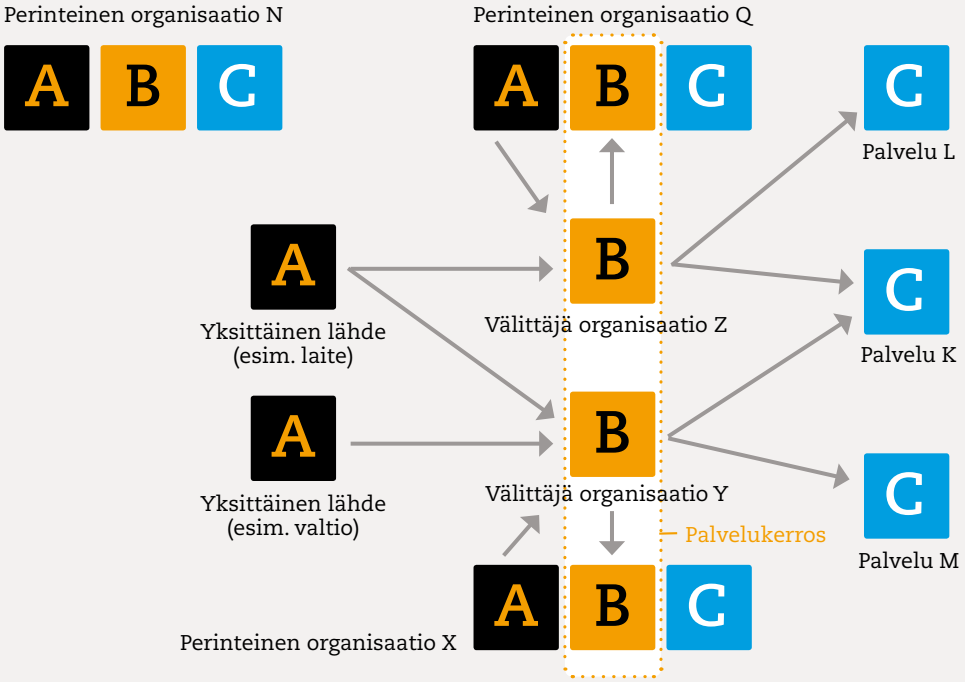
1.2 Henkilötiedon arvoketjun pilkkoutuminen ihmiskeskeisesti

Keskeinen seuraus My Data -periaatteiden toteuttamisesta on henkilötiedon arvoketjujen pilkkoutuminen ja tiedon hallinnan keskittyminen ihmisen ympärille. Tämä avaa mahdollisuuksia uusille toimijoille ja rikkoo perinteisiä sektoreiden ja toimialojen rajoja.

Henkilötiedon jalostuksen arvoketju koostuu **henkilötiedon lähteistä** (luominen, kerääminen), **välittämisestä** (jalostus ja hallinta) ja **hyödyntämisestä**. Dataa voidaan säilyttää kaikissa näissä vaiheissa. Perinteisesti koko arvoketju on yhden organisaation sisällä. Esimerkiksi kun pankin järjestelmiin syntyy tieto henkilön kaikista tilitapahtumista, pankki jalostaa ja välittää tietoa ja tuottaa tiliotteet ja verkkopankkinäkymät asiakkailleen.

My Datan myötä siirrytään yksittäisten organisaatioiden sisään suljetuista arvoketjusta kohti hajautettua henkilötiedon arvoverkostoa, jossa syntyy eri vaiheisiin erikoistuneita ja rinnakkaisia toimijoita. Esimerkiksi suomalainen Balancion[2]- ja yhdysvaltalainen Mint[3]-palvelut hyödyntävät pankkisektorilla henkilötietoa arvoketjun loppupäässä ja tarjoavat käyttäjille tavallista verkkopankkia kokonaisvaltaisemman oman talouden näkymän.

Perinteisessä datan jalostusketjussa yritys tai palvelu, jossa data syntyy, on portinvartija. Esimerkiksi kauppaketjun kanta-asiakaskorttiostosten data syntyy kauppaketjun tietojärjestelmiin. Jos laki ei muuta edellytä, kauppaketju päättää, saako jokin muu taho käyttää dataa. Mikäli palvelulla on My Data -periaatteiden mukainen henkilötietorajapinta, niin portinvartijana toimiikin palvelun käyttäjä; hän päättää itse, mille tahoille hänen tietojaan saa luovuttaa. My Data -ajattelussa tiedon välittämiseen muodostuu yleiset käytännöt, niin että useat eri operaattorit voivat tarjota My Datan hallinnointi-, säilytys- ja välityspalveluja keskenään yhteentoimivasti.



Kuva 1.1: My Data -lähestymisessä henkilötiedon arvoketju hajaantuu eri organisaatioiden yhteistoiminnaksi. Arvoketjun kolme vaihetta ovat henkilötiedon luonti (A), välitys ja jalostus (B) ja hyödyntäminen (C). My Data -mallissa rinnakkaiset välittäjäpalvelut (B) muodostavat hajautetun, mutta yhteentoimivan palvelukerroksen, jonka kautta hyödyntäjät voivat käyttää useasta lähteestä peräisin olevaa tietoa. Palvelukerroksen palvelujen tarjoajia kutsutaan My Data -operaattoreiksi (luku 4).

2 http://www.balancion.com/
3 https://www.mint.com

A Henkilötiedon luonti ja integraatioargumentti

Integraatioargumentti
Henkilötiedolla on monta lähdettä. Yksi lähde sellaisenaan on heikko.

Arvoketjun hajauttaminen mahdollistaa datan yhdistelyn useista lähteistä. Käytännössä, kun pysytään yhden organisaation sisällä, voidaan hyödyntää vain kyseisen organisaation järjestelmissä olevaa dataa, mikä rajoittaa sovellusmahdollisuuksia. Pankkiesimerkissä Balancion- ja Mint-palveluissa voidaan yhdistää tiedot käyttäjän kaikilta pankkitileiltä ja luottokorteista, vaikka ne olisivat eri pankkien tarjoamia. Samaan tapaan ollisi hyödyllistä, jos kulutusseurantapalvelut voisivat yhdistää tietoja ihmisen ostoista kaikissa kaupoissa ja ravintoloissa eikä vain yhdessä kauppaketjussa.

B Henkilötiedon välitys ja yksityisyysargumentti

Yksityisyysargumentti
Henkilötiedon välitys ja jalostus on yksityisyyden kannalta herkkä asia

Välittäjäorganisaatiot kokoaisivat yksilön henkilötietoa useista lähteistä, jalostaisivat ja yksilön suostumuksella tarjoaisivat sitä eteenpäin dataa hyödyntäville palveluille ja sovelluksille. Esimerkiksi yksi syy siihen, miksi lääkärit ja sairaalat eivät ole merkittävästi kiinnostuneita ihmisten itse mittaaman datan hyödyntämisestä on, että datan vastaanottaminen aiheuttaisi niille uuden velvollisuuden huolehtia yksityisyydensuojasta (Sullivan 2014). Tiedon välittämiseen erikoistuneet operaattorit voisivat huolehtia yksityisyydensuojan velvoitteista tiedon hyödyntäjien puolesta. Operaattorit voivat kehittää myös ratkaisuja tiedon laadun parantamiseksi. Operaattoreilla onkin keskeinen rooli My Data -palveluinfrastruktuurin toteuttajina (luku 4).

C Henkilötiedon hyödyntäminen ja innovaatioargumentti

Innovaatioargumentti
Uusien innovaatioiden kehittyminen täytyy olla irrallista vanhoista rakenteista.

Kolmas argumentti hajauttamisen puolesta liittyy uusiin innovaatioihin henkilötiedon hyödyntämisessä. Parhaita käyttötapauksia ei välttämättä keksitä niissä organisaatioissa, joissa henkilötietoa kerätään tai luodaan. Esimerkiksi kattavan henkilökohtaisen talouden analysointipalveluiden kehittäminen ei ole perinteisten pankkien ydinliiketoimintaa, mutta Balancionin ja Mintin kaltaiset toimijat voivat keskittyä juuri siihen.

Vaikka organisaatiot arvoketjun hajautumisen myötä erikoistuvat, niin monet niistä silti toimivat samanaikaisesti sekä datan lähteinä että käyttäjinä. Hajautetussa mallissa nekin organisaatiot, jotka toimivat sekä datan lähteinä että hyödyntäjinä tarjoavat ihmisille kuitenkin mahdollisuuden ohjata data myös erikoistuneille dataoperaattoreille tai muihin palveluihin. (luku 4)

Henkilötiedon arvoketjun pilkkominen on strategisesti merkittävä muutos, joka vaikuttaa joihinkin nykyisin tyypillisiin liiketoimintamalleihin ja toisaalta luo puitteet aivan uudenlaiselle liiketoiminnalle. Esimerkiksi käyttäjistä kerättyyn tietoon ja ilmaisen palvelun avulla käyttäjämäärien kasvattamiseen perustuva liiketoiminta ei toimi ympäristössä, jossa yksilöt halutessaan voivat nopeasti siirtää tiedon toisiin palveluihin. Palveluyritysten pitäisi nähdä arvoketjun pilkkoutuminen mahdollisuutena solmia entistä läpinäkyvämpiä ja syvempään luottamukseen perustuvia asiakkuuksia. Yritykset voisivat myös kehittää liiketoimintamallejaan vahvemman asiakasymmärryksen perusteella ja hyödyntää asiakkaita palveluiden osatuotannossa ja suunnittelussa.

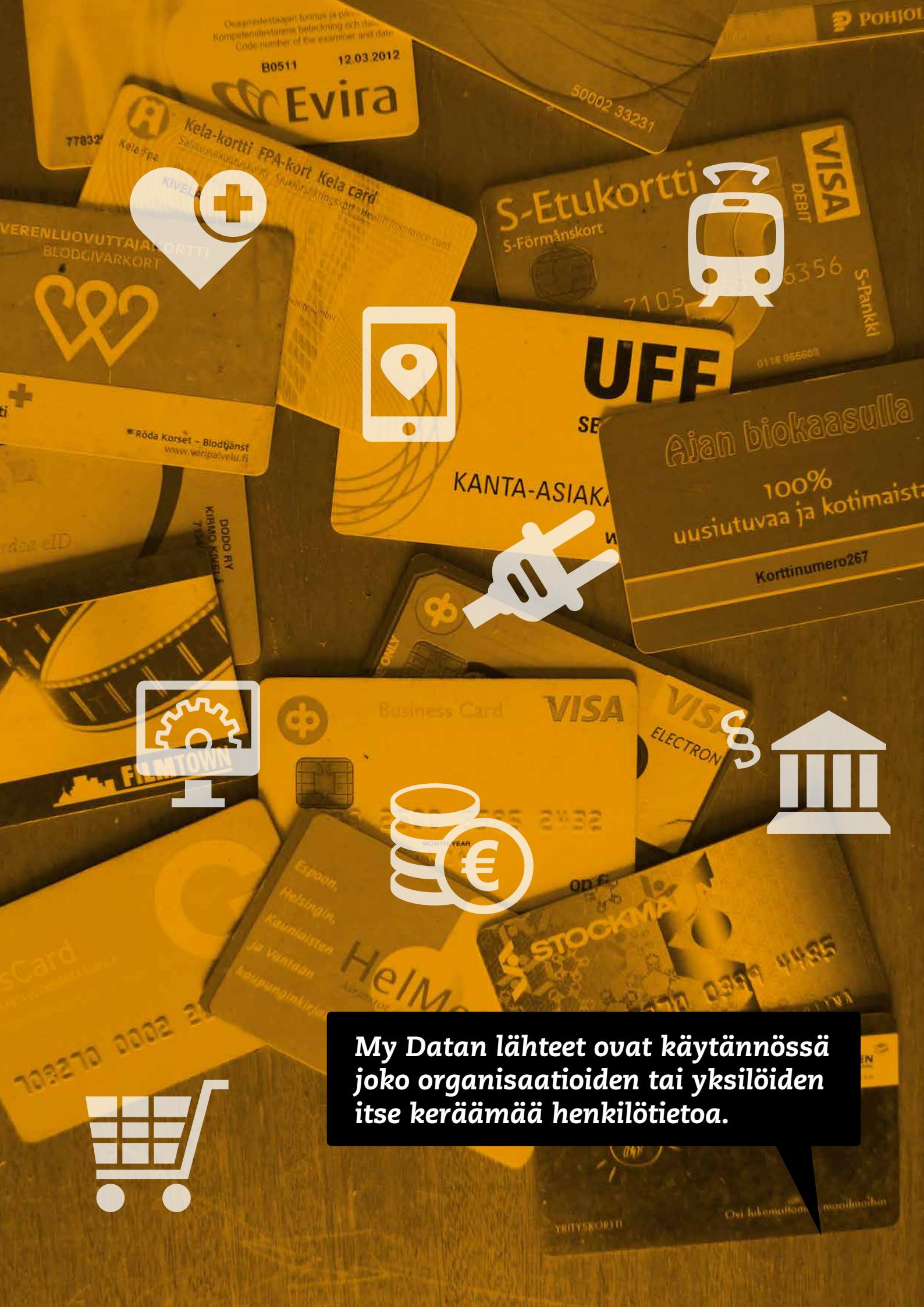
Organisaation My Data -johtosääntö (policy)

My Data -johtosääntöä voidaan nähdä tietoturvaohjeistusten, rekisteriselosteiden ja käyttöehtojen laatimista ohjaavana yleisenä linjauksena toimia My Data -periaatteiden mukaisesti. Johtosäännön ohjaamista rekisteriselosteista ja käyttöehdoista ilmenee, mitä tietoja asiakkaan on mahdollista saada itselleen ja missä muodossa. Vastaavasti johtosäännön ohjaamana laadittu tietoturvaohje toimii yrityksen sisällä selkeänä ohjeistuksena siitä, miten organisaatiossa käytännössä huolehditaan sen hallussa olevan My Data -aineiston tietoturvasta.

Yrityksen tai palveluntarjoajan My Data -johtosääntö voi riippua esimerkiksi toimialasta, koska yrityksiä on monenlaisia ja dataa kerätään eri käyttötarkoituksiin. Esimerkiksi terveysalan yrityksen My Data -johtosääntö saattaa olla yksityiskohtaisempi ja tiukempi kuin ohjelmointiyrityksen, joka tekee sovelluksia ornitologeille 'The Birds I've Seen Lately' -tietokantojen pohjalta.

Osa yrityksistä on myös valmiita jakamaan dataa avoimesti asiakkaalleen. Harvoja esim. terveystietoon liittyviä poikkeuksia lukuunottamatta lainsäädäntö ei estä asiakkaan datan antamista asiakkaalle itselleen. Tulevaisuudessa asiakkaan kannalta hyvästä My Data -johtosäännöstä voi tulla merkittävä kilpailutekijä. Johtosäännön pitää tukea henkilötiedon käsittelyä yrityksessä ja kertoa ulospäin, miten dataa kerätään, käsitellään, jaetaan jne.

Henkilötietolainsäädäntö edellyttää jo nyt henkilörekisteriselosteiden ylläpitämistä. Palveluilla on käyttöehdot, joihin käyttäjät sitoutuvat, usein niitä lukematta. Yritysten tietoturvaohjeistuksista tai niiden liitteenä olevista tiedon luokitteluohjeista yleensä käy ilmi, käsitelläänkö teletunnistetietoa, henkilötietoa tai asiakkaiden luottamuksellista tietoa, saako henkilötunnuksia käsitellä, saako salaisia tietoja käsitellä yrityksen järjestelmissä jne. My Data -periaatteiden mukaiset käytännöt olisivat tarkennuksia tällaisiin asiakkaan tietojen käsittelyohjeisiin.


Evira
Kela-kortti FPA-kort Kela card
S-Etukortti
S-Förmånskort
VISA
DEBIT
S-Pankki
BLODGIVARKORT
Röda Korset – Blodtjänst
UFF
KANTA-ASIAKA
Ajan biokaasulla
100%
uusiutuvaa ja kotimaist
Korttinumero267
Business Card
VISA
ELECTRON
FILMTOWN
My Datan lähteet ovat käytännössä joko organisaatioiden tai yksilöiden itse keräämää henkilötietoa.

2. My Datan lähteet

My Datan lähteet ovat käytännössä joko organisaatioiden tai yksilöiden itse keräämää henkilötietoa. Henkilötietoa syntyy nykyään valtavia määriä erilaisissa prosesseissa. Tässä luvussa esitellään yleiskatsaus siihen, miten henkilötietoa syntyy ja missä muodossa sitä on olemassa nykyisissä tietokannoissa.

Jotta organisaatioiden hallussa oleva henkilötieto muuttuisi My Dataksi, eli olisi helposti ja käytännöllisesti saatavilla ihmisille itselleen tulee datan lähteinä toimivissa palveluissa olla **koneluettava ohjelmointirajapinta** (Application Programming Interface API), jonka mukaan eri ohjelmat voivat tehdä pyyntöjä ja vaihtaa tietoja keskenään. Rajapintojen kautta henkilötieto liikkuu vaivattomasti palvelusta toiseen. Parhaimmillaan samaa tietoa voidaan käyttää useissa sovelluksissa lähes reaaliaikaisesti. Henkilötiedon saatavuuteen liittyvistä rajapintojen ominaisuuksista koneluettavuus, standardien mukaisuus ja reaaliaikaisuus ovat tärkeimpiä, ja näistä on jo runsaasti hyviä esimerkkejä eri verkkopalveluissa.

My Data -periaatteista seuraa, että ohjelmointirajapinnalla pitäisi olla myös muita ominaisuuksia kuin henkilötiedon saatavuus rajapinnan kautta. Näistä keskeisimpiä ovat koneluettavat käyttöselosteet ja yksityisyysasetusten etähallinta, jotka yhdessä mahdollistavat **hallittavan ja standardoidun sopimisen.** Näiden ominaisuuksien ympärillä keskustelu on vasta käynnistynyt, ja niiden kehittäminen vaatii vielä runsaasti selkeytystä ja kokeiluja sekä yhteiskehittelyn tuloksena syntyviä uusia standardeja.

Rajapintojen olemassaolo ja ominaisuudet eivät kuitenkaan yksin riitä My Datan -periaatteiden toteutumiseen. Lisäksi tarvitaan henkilötiedon välittämisen ja helpon hallinnan mahdollistava palveluinfrastruktuuri, jota käsitellään tarkemmin seuraavassa luvussa.

2.1 Miten henkilötietoa syntyy?

Henkilötieto voi olla monen tyyppistä. Käytännössä henkilötietoa syntyy esimerkiksi digitaalisissa palveluissa kaikesta vuorovaikutuksesta asiakkaan ja palvelun välillä. Teoriassa tietoa voidaan kerätä loputtomasti, mutta olennaisempaa on tunnistaa, mikä on tarkoituksenmukaista tietoa, minkälaisen tiedon keräämiseen on lupa ja mitä tietoa kannattaa tallentaa ja jalostaa. Yrityksen palveluprosessissa tarvittava, yritykselle hyödyllinen tieto voi olla asiakkaallekin hyödyllistä jo tiedonkeruun läpinäkyvyydenkin vuoksi. Palveluprosesseissa saattaa syntyä myös sellaista tietoa, joka olisi hyödyllistä yksilölle, mutta ei palvelua tarjoavalla organisaatiolle. Tällöin organisaatiota ei voida velvoittaa keräämään kyseistä tietoa. My Datan toteutumisen myötä asiakkaat saattavat kuitenkin kokea heille hyödyllisen henkilötiedon saatavuuden ja tiedon keräämisen heidän puolestaan osaksi palvelun ominaisuuksia.

Tiedon tärkeitä määreitä ovat sen koko, ajallinen kertymistahti ja rakenne. Sijaintitieto ja syketieto ovat esimerkkejä rakenteeltaan yksinkertaisesta aikasarjadatasta; saldotieto päivittyy usein, mutta ei ole luonteeltaan jatkuvaa; osoitetieto on pääasiassa pysyvää, mutta päivittyy satunnaisesti. Datan määrällä mitattuna esimerkiksi geenitieto eroaa kengän numerosta merkittävästi. Henkilötietoa voitaisiin luokitella myös sektoreittain, kuten kuten terveystieto, liikkumistieto, oppimistieto jne., mutta My Data -lähestymisessä pyritään toimintatapoihin, jotka mahdollistavat sektorirajat ylittävän tiedon hyödyntämisen. Käytännössä tämä tarkoittaa, että henkilötieto liikkuu rajapintojen välillä

ja on teknisesti yhteentoimivaa datan määrästä, päivitystahdista ja rakenteesta riippumatta. Tiedon rakenteen pitää olla riittävän kattavasti ja selkeästi kuvattu, jotta esimerkiksi eri lähteistä saatavat aikasarjatiedot voidaan yhdistää toisiinsa.

Mitä on henkilötieto?

Kuva 2.1: Aloja, joissa syntyy paljon henkilötietoa, ovat muiden muassa liikenne, pankki- ja vakuutustoiminta, viestintä ja kommunikaatio (mukaan lukien sosiaalinen media ja muut verkkopalvelut) ja media, kuluttajakauppa ja erityisesti ruokakauppa, terveys- ja hyvinvointiala, energia ja koulutus ja oppiminen. My Data -periaatteiden toteuttaminen mahdollistaa näiden alojen sisäisen henkilötiedon hallinnan organisointia, mutta erityisesti alojen välistä tiedonsiirtoa.

My Data -lähestyminen tuottaa lisäarvoa yhdistämällä eri lähteistä tulevaa henkilötietoa. Kuvassa 2.1 on kuvattu keskeisimmät henkilötiedon osa-alueet. Eri sektoreiden tietoja yhdistämällä palveluita tarjoava organisaatio voi nykyistä kokonaisvaltaisemmin ymmärtää yksilön tarpeita ja toimintaa. Jos palvelua kehitetään vain yhden sektorin sisällä, kerätty tieto antaa vain osittaisen kuvan ihmisen tarpeista ja toiminnasta. Esimerkiksi terveys- ja hyvinvointipalveluissa kliinisen terveystiedon lisäksi kannattaa kurkottaa muiden sektoreiden alueille: talous, liikkuminen ja median käyttö vaikuttavat myös ihmisen terveyteen. Kaikkea kulutus- ja käyttäytymistietoa voidaan hyödyntää ihmisen elämäntapojen mallintamisessa ja nykyistä parempien suositusten ja syvemmän ymmärryksen luonnissa.

2.2 Miten henkilötieto muuttuu My Dataksi?

Johdannossa esitettiin My Datan datalähtöinen määritelmä – My Dataa on se osa henkilötiedoista, mikä on henkilön itsensä saatavilla ja hallittavissa. Yksittäisen datalähteen, kuten vaikkapa pankkipalvelun osalta tämä toteutuu silloin, kun palveluun on olemassa rajapinta, joka täyttää tietyt tiedon saatavuuteen liittyvät ehdot.

Monilla digitaalisessa liiketoiminnassa menestyvillä yrityksillä ja verkkopalveluilla, kuten Facebookilla, Googlella, Amazonilla ja Twitterillä, on jo ohjelmointirajapinta eli API, jonka kautta voi päästä kiinni palveluissa olevaan henkilö-

tietoon. Nämä rajapinnat ovat usein hyvin kuvattuja, ja käyttävät pääasiassa yleisesti tunnettuja avoimia rajapintastandardeja. Nykyisten rajapintojen kautta saatava tietosisältö ei kuitenkaan ole aina kovin kattavaa, koska rajapinnat suunnitellaan ensisijaisesti yritysten ja palvelujen väliseen integraatioon eikä asiakkaiden pääsyyn heitä koskevaan dataan.

Reuben Binns on määritellyt henkilötiedon saavutettavuudelle seuraavanlaiset ehdot, joita hän kutsuu henkilötiedon saatavuuden viiden tähden luokitukseksi (Binns 2013):

1. **Henkilötieto on yksilöille saatavilla maksutta ja digitaalisessa muodossa** – esimerkiksi rajapinnan tai sähköpostin kautta ilmaiseksi ja ilman sitoutumista markkinointiviestien tai vastaavien vastaanottamiseen
2. **Henkilötieto on koneluettavaa** – esimerkiksi CSV-muodossa
3. **Henkilötieto on avoimessa dataformaatissa** – kuten CSV, XML tai JSON mieluummin kuin Excel
4. **Henkilötieto on saatavilla kattavasti** – kaikki henkilötieto on saatavilla samassa paikassa ja kattavasti.
5. **Henkilötieto on saatavilla ajantasaisesti** – joko ajantasaisen ja säännöllisesti päivittyvän rajanpinnan kautta tai sitten reaaliaikaisesti jatkuvana syötteenä

Tiedon saatavuus voidaan määritellä selkeästi ymmärrettäväksi ehtolistaksi. Yllä olevaa listausta ei vielä tunneta kovin laajasti, mutta on oletettavissa, että ennen pitkää vastaavanlainen listaus vakiintuu käyttöön, näin on esimerkiksi käynyt avoimen datan alueella. Jo nyt organisaatio voi luoda henkilötietorajapinnan, joka toteuttaa nämä ehdot, eli My Datan datalähtöisen määritelmän.

My Datan minimitoteutus

Monet yritykset tarjoavat jo yksilölle mahdollisuutta ladata kulutustietonsa verkkosivuilta yleensä joko Excel- tai CSV-muodossa. Esimerkiksi joidenkin suomalaisten energiayhtiöiden ja pankkien verkkopalveluista tiedon lataaminen on jo mahdollista. Lisäksi energiayhtiöiden verkkosivustoilla tarjolla olevilla sovelluksilla voi tehdä erilaisia vertailuja omaan aikaisempaan ja viiteryhmien kulutukseen.

Yhdysvalloissa on kehitetty “green button”- ja “blue button” -konseptit edistämään kuluttajien mahdollisuutta saada energiankulutus- ja terveysdataansa itselleen. Blue button on verkkosivulle sijoitettu symboli, joka osoittaa terveyspalvelun asiakkaille, että he voivat verkon kautta katsella ja ladata itselleen omat terveystietonsa. Alun perin ratkaisu kehitettiin, jotta sotaveteraanit pystyisivät käyttämään erikoislääkäripalveluja eksoottistenkin vammojen hoidossa, ja siirtyä joustavasti erikoistuneelta yksityislääkäriltä toiselle. Green button on vastaava energian kulutustiedon latausnappi.

Tällaista manuaalista latausmahdollisuutta voidaan pitää My Datan minimitoteutuksena, joka on huomattavasti parempi kuin oman datan saaminen pyydettäessä vain paperitulosteena. Nämä ovat kuitenkin vain välivaiheen parannuksia, joissa data pitää edelleen käydä manuaalisesti lataamassa, eikä sitä voi siirtää automatisoidusti rajapintojen kautta sovelluksesta toiseen.

http://energy.gov/data/green-button ja http://www.healthit.gov/bluebutton

2.3 Hallittava ja standardoitu sopiminen

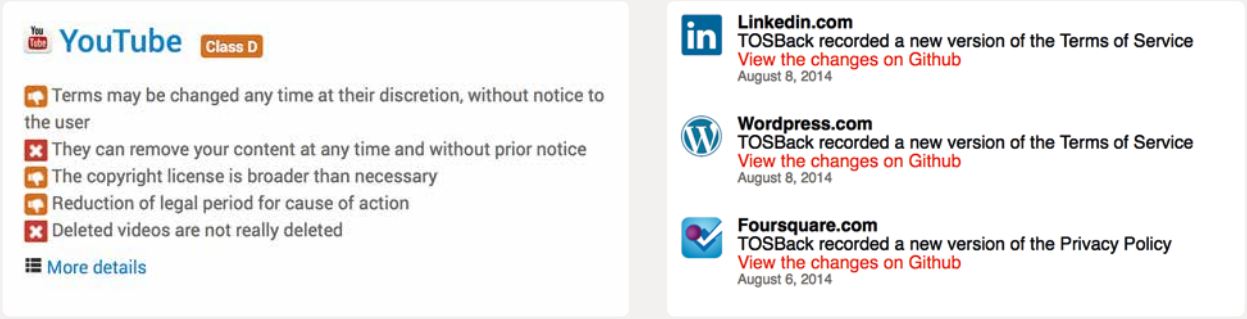
Kun ihminen ottaa käyttöön uuden palvelun, häntä pyydetään hyväksymään palvelun käyttöehdot. Verkkopalvelun hyväksymisnappia painettaessa syntyy palvelusopimus. Jo nykyisin ihmisillä on arkisesti käytössään kymmenittäin erilaisia verkkopalveluja, joiden kaikkien kanssa on tehty palvelusopimus. Esineiden verkon (Internet of Things) edetessä palvelusopimusten ja niihin sisältyvien henkilötiedon käyttölupien määrä kasvaa entisestään. Kun ottaa käyttöön uutta televisiota, saattaa joutua hyväksymään ehdot, joissa lupaa, että oman viihdekulutuksen dataa saa välittää eteenpäin, ja autonavigaattorin käyttäjä saattaa joutua hyväksymään ehdot, joissa oman sijainnin tiedot välittyvät navigaattoriohjelmiston valmistajalle.

Nykyisin eri palveluiden sopimuskäytännöt vaihtelevat huomattavasti, eikä ihmisillä ole käytännöllistä mahdollisuutta hallita esimerkiksi sitä, mitä kaikkia henkilötietoihin liittyviä oikeuksia on palvelusopimusten muodossa antanut eri yrityksille. Sopimisen yhdenmukaistaminen ja hallittavuus lisäisi henkilökohtaisten digitaalisten palvelujen kokonaisuuden ymmärrettävyyttä ja käytettävyyttä ihmisille.

Hallittava ja standardoitu sopiminen edellyttää että palvelujen käyttöehdot ja henkilödatan käyttöselosteet ovat koneluettavia ja toisaalta, että yksityisyysasetuksia on mahdollista muuttaa rajapintojen kautta.

2.3.1 Koneluettavat käyttöehdot ja käyttöseloste

Tutustuminen kaikkien palvelujen käyttöehtoihin ja niiden ymmärtäminen on käytännössä mahdotonta, eikä yksilöllä nykyisin ole muuta neuvotteluvaraa kuin hyväksyä käyttöehdot tai olla käyttämättä palvelua. Käyttöehtojen monimutkaisuudesta on kampanjoitu ja muun muassa ongelmaa esittelevä elokuva *Terms and conditions may apply* on saanut laajaa kansainvälistä huomiota (Hoback 2013). Kehitteillä on myös standardeja ja teknisiä apuvälineitä, kuten Opennotice.org, joka pyrkii ratkaisemaan käyttöehtoihin ja palvelusopimuksiin liittyviä ongelmia rakentamalla yleisesti tunnistetun ja avoimen standardijärjestelmän käyttöehdoille.



Kuva 2.2: Terms of Service; Didn't Read[4] -projektissa on arvioitu ja luokiteltu yleisesti käytettyjen verkkopalveluiden käyttöehtoja ja tehty niistä käyttäjille helpommin ymmärrettäviä koosteita. TOSback[5] -projekti puolestaan tarkistaa automaattisesti palveluiden alati muuttuvia käyttöehtoja ja tallentaa niiden muutoshistorian.

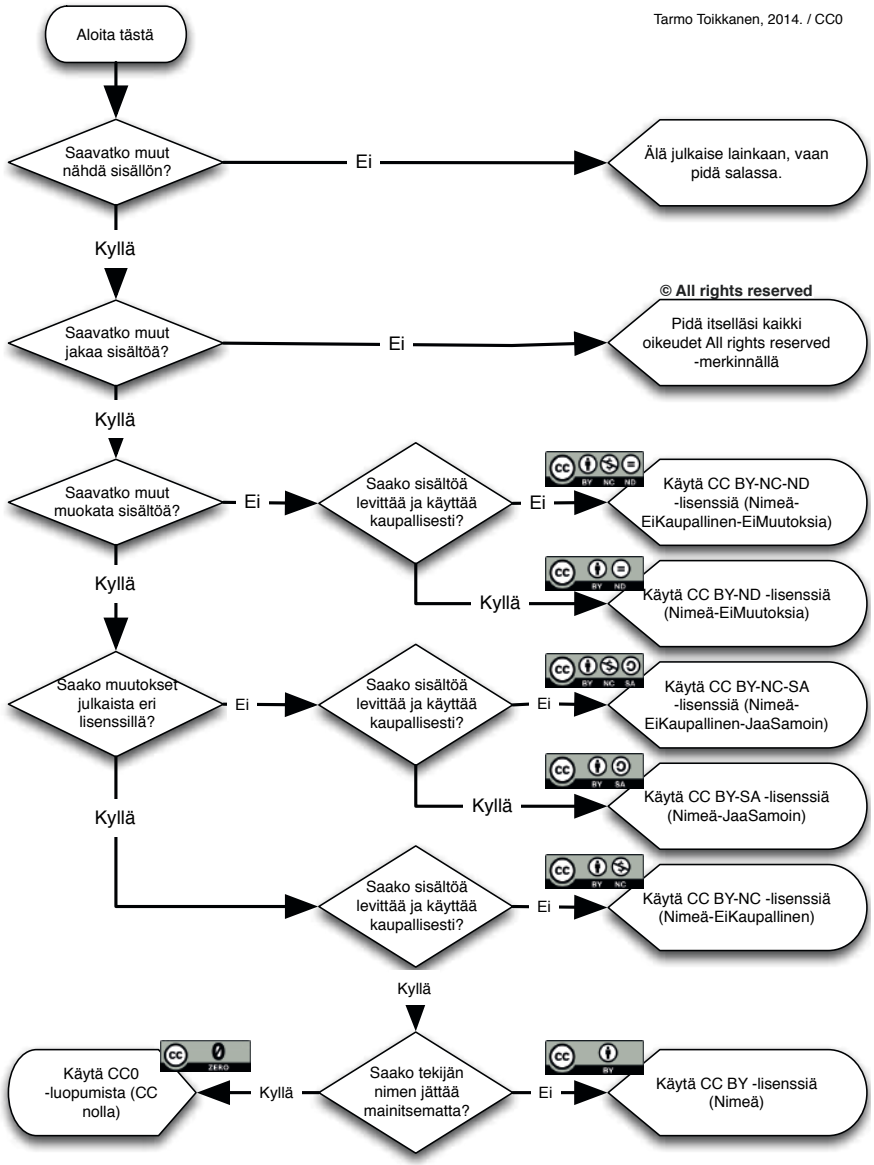
Verkkopalvelujen käyttöehdot jättävät palvelun käyttäjän usein heikkoon asemaan. Esimerkiksi käyttäjien tuottaman sisällön oikeudet saattavat siirtyä palvelun tarjoajalle, ja usein palveluntarjoaja saa yksipuolisesti muuttaa käyttöehtoja. Käyttöehtojen tai kerätyn tiedon käyttötarkoituksen muutokset saattavat

4 https://tosdr.org/
5 https://tosback.org/

liittyä yrityskauppoihin tai laajojen aineistojen myyntiin. Esimerkiksi Moves -älypuhelinsovellus, joka seuraa jatkuvasti käyttäjän liikkeitä, muutti käyttöehtojaan kaksi viikkoa sen jälkeen, kun sovelluksen kehittänyt suomalaisyhtiö myytiin Facebookille. Aiemmin sovellus ei välittänyt käyttäjätietoja eteenpäin, mutta sen jälkeen tiedot menivät Facebookille (Wall Street Journal 2014).

Henkilötiedon rahallisesta arvosta voi saada osviittaa laajojen henkilötietoaineistojen myyntiuutisista. Esimerkiksi amerikkalainen lentoyhtiö Delta sai Skymiles-bonusohjelmansa datasta luottokorttiyhtiö American Expressiltä yli puoli miljardia euroa (Taloussanomat 2014). Yksilö harvoin ymmärtää henkilötietojensa arvoa yrityksille. Äärimmäisenä esimerkkinä ovatkin ne yritykset, joiden arvo määräytyy pääosin asiakkuuksista kerätystä tiedosta. Tällaisten yritysten toiminta tai liiketoimintamallit eivät yleensä näy henkilöille, joista kerättyjä tietoja ne myyvät. Yksityisyydensuojaa digitaalisessa taloudessa käsittelevässä toimintaohjeessa Presidentti Obaman hallinto suosittelee, että asiakkaiden tietoja käsittelevien kolmansien osapuolien (data brokers), jotka eivät ole suorassa kontaktissa asiakkaiden kanssa, tulisi tarjota sitä paremmat työkalut asiakkaille saada tietoa tiedonkäsittelystä, mitä sensitiivisemmästä tiedosta on kyse (White House 2012).

Creative Commons -lisenssin valintaprosessi



Kuva 2.3: Kaavio Creative Common lisenssin toiminnasta esimerkkinä kuinka lisensseistä voidaan tehdä ihmisille ymmärrettäviä ja standardoituja[6]

6 http://creativecommons.fi/lisenssit/valitse-lisenssi/

My Data -mallissa ei ole tarkoitus estää yrityksiä ansaitsemasta henkilötiedolla. Olennaista on tehdä ansaintamallit läpinäkyviksi ja avata yhteiskunnallista keskustelua siitä, millainen henkilötiedon hyödyntäminen on kestävää. Esimerkiksi kohdennettu mainonta, jonka kohdistamisen toimintaperiaatteista tai edes olemassaolosta yksilö ei ole tietoinen on eettisesti arveluttavaa. Usein ihmiset hyötyvät kohdentamisesta, mutta haluaisivat ymmärtää kohdentamisen perusteet. Ihmiset voisivat myös parantaa kohdentamista tuottamalla tai tarjoamalla enemmän tietoa itsestään, mutta tämä vaatii läpinäkyvyyttä ja sen pitää olla vapaaehtoista. My Data -periaatteissa pyritään tällaiseen progressiiviseen suhtautumiseen henkilötiedon hyödyntämisessä.

Jotta käyttäjien olisi helppo ymmärtää käyttöehtoja, niiden pitäisi olla mahdollisimman selkeitä ja rakenteeltaan yhtenäisiä. Nykyisin palvelujen käyttöehdot poikkeavat toisistaan rakenteellisesti, joten niitä on mahdotonta esittää yksinkertaisina valintoina tai visualisointeina. Jatkossa voisimme kehittää yhtenäisiä standardeja rakenteisessa muodossa julkaistaville käyttöehdoille. Niissä pitäisi pyrkiä vastaavanlaiseen yksinkertaisuuteen, johon Creative Commons -lisensseissä on päästy. Valitessaan oikeaa lisenssiä tekijänoikeuksien haltijan tarvitsee vastata vain muutamaan kysymykseen, kuten saavatko muut muokata sisältöä tai käyttää ja levittää sitä kaupallisesti (kuva 2.3).

Käyttöehdoissa voitaisiin kysyä, saako dataa myydä tai luovuttaa eteenpäin, saako dataa luovuttaa viranomaisille, miten kauan dataa säilytetään jne. Vakiomuotoiset käyttöehdot voitaisiin visualisoida vaihtoehtoja kuvaavilla ikoneilla. Esimerkiksi aiemmin Mozilla-säätiölle työskennellyt Aza Razkin on konseptoinut verkkosivuille tarkoitettua ikonisarjaa (kuva 2.4), jolla voi kuvata, miten verkkosivu käyttää tallentamiaan henkilötietoja.

Kuva 2.4: Aza Razkinin tekemä ikonisarja Mozilla-säätiölle kuvaamaan kuinka verkkosivut tallentaa henkilötietoa

Käyttöehdot ovat staattinen dokumentaatio, jota luodessaan organisaatio on pyrkinyt varautumaan erilaisiin henkilötiedon tulevaisuuden hyödyntämismahdollisuuksiin. Tästä syystä niissä ei määritellä kovinkaan rajaavasti, mihin tietoa voidaan käyttää. Palvelusuhteen aikana tapahtuu yleensä paljonkin muutoksia sen suhteen minne tietoa siirretään. Laki rajoittaa tiedon jatkohyödyntämistä alkuperäisestä poikkeaviin tarkoituksiin. Siitä huolimatta yritysten välillä tehdään jatkuvasti kauppaa henkilötiedoilla, yllä esitetty Delta Airlinesin esimerkkitapaus on vain yksi monista vastaavista. Jotta yksilö pysyisi perillä siitä, mitä tiedolla missäkin vaiheessa tehdään, olisi läpinäkyvyyden nimissä tärkeää luoda käytänteitä, joilla yritys voisi ilmoittaa henkilötiedon käytöstä. Tällainen käyttöseloste voisi yksinkertaisemmillaan olla ilmoitus, jossa kerrotaan, että nyt organisaatio on aloittanut tiedon välittämisen toiselle organisaatiolle saatuaan aiemmin asiakkaan suostumuksen tiedon välittämiseen.

2.3.2 Yksityisyysasetusten hallinta rajapintojen kautta

Yksityisyysasetusten hallinta rajapintojen kautta tarkoittaa, että käyttäjällä voi olla yksi paikka, josta hän voi kerralla määritellä yksityisyyteen ja datan käyttöön liittyvät asetukset useammassa käyttämässään palvelussa. Vastaavalla tavalla kuin rajapinnat mahdollistavat nykyisin sen, että esimerkiksi Foursquare-sovelluksesta voi lähettää viestin Twitteriin ja Facebookiin, niin tulevaisuudessa yksilö voisi My data -tilinsä hallintasovelluksesta säätää yksityisyysasetukset kaikkiin niihin palveluihin, jotka tukevat yksityisyysasetusten hallintaa rajapinnassaan.

Nykyiset palvelut eivät vielä mahdollista yksityisyysasetusten hallintaa rajapintojen kautta, vaan asetuksia säädetään palvelun sisällä. Sujuva palvelujen etähallinta edellyttää myös yleisiä käytäntöjä siitä, millaisia yksityisyysasetuksia ja -valintoja eri palveluissa on.

Alla on lueteltu yksityisyysominaisuuksia, joiden asetuksia etähallinnalla voisi säätää:

- Mitä tietoa palvelu saa käyttötilanteessa haltuunsa (esimerkiksi mobiililaitteesta sijainti tai kontaktitietoja)?
- Mitä tietoa palvelun sisällä voidaan näyttää muille käyttäjille?
- Mitä tietoa palvelun tarjoaja voi hyödyntää suoraan liiketoiminnassa kumppaneidensa kanssa (esimerkiksi kohdennettu markkinointi)?
- Mitä tietoa palvelu voi tallentaa käyttäjästä (osa tiedosta, kuten sijainti, voi olla sellaista, jota käyttäjä voi haluta antaa palvelun hyödyntää reaaliaikaisesti, mutta ei tallentaa)?
- Mitä tietoa palvelu lähettää suoraan eteenpäin rajapinnan kautta (reititys toiseen palveluun ilman operaattoria)?
- Mitä tietoa palvelu lähettää käyttäjän omaan tietovarastoon (käyttäjällä voi olla esimerkiksi My data -operaattorin ylläpitämä oma tietovarasto, josta käyttäjä voi edelleen ohjata tietojaan muille tahoille)?

Kuva 2.5: Esimerkki yhdenmukaisista yksityisyysasetuksista on Android-sovellusten "application permissions", mitä jokaisen Android-sovelluksen on kysyttävä käyttäjältä ennen asennusta.

Kuvassa 2.5 esitetty Android-sovellusten yksityisyysasetusten lista on konkreettinen esimerkki siitä, miten henkilötiedon keräämistä voidaan sovelluskohtaisesti luokitella ja muodostaa yhteisesti sovittu malli, jonka mukaan kaikki palveluntarjoajat määrittelevät omaa tiedonkeruutaan. Android-sovellusten yhdenmukainen lista yksityisyysasetuksista on askel oikeaan suuntaan, mutta ihmisillä ei kuitenkaan ole mahdollisuutta vaikuttaa siihen, mitä oikeuksia millekin sovellukselle antaa, ainoa vaihtoehto on olla asentamatta sovellusta, mikäli ei halua antaa kaikkia sovelluksen pyytämiä lupia. Suurin osa ihmisistä ei tarkastele sovellusten käyttöoikeuksia kovinkaan kriittisesti ennen asentamista.

Yhtenäinen malli yksityisyysasetuksiin ja mahdollisuus hallita asetuksia rajapintojen kautta yhdessä mahdollistavat My Data -hallintapalvelujen tekemisen. Tällaisen palvelun avulla yksilö voi saavuttaa keskitetyn hallinnan ja kattavan kuvauksen oman henkilötietonsa virtaamisesta digitaalisessa maailmassa.

Kun rajapinnat toteuttavat My Datan ehdot, organisaatiot tuottavat kattavan käyttöselosteen ja yksityisyysasetuksia voidaan etähallita, olemme lähellä ohjelmallista sopimista; organisaation ja yksilön välillä vallitsee aktiivinen sopimusprosessi, joka ylläpitää tilaa sopimukseen liittyvän tiedon hyödyntämisestä ja mahdollistaa sopimusehtojen muuttamisen hallitusti ja molemminpuoleisesti.

2.4 Rajapintoihin liittyvät standardit ja formaatit

Yleisesti käytössä olevat standardit ovat edellytys helposti toteutettavalle palvelujen väliselle yhteentoimivuudelle. Avoimet standardit ovat puolestaan edellytys avoimille markkinoille. 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.

Henkilötiedon alueella nykytilanne on se, että suurelta osin toimijoiden tietomallit on suunniteltu omien tarpeiden ympärille. Niiden yhteneväisyys esimerkiksi kilpailijoiden tai muiden sektorien toimijoiden kanssa on heikkoa. Kun tiedon yhdistäminen useasta lähteestä yleistyy, kasvaa myös tarve yhtenäisille tietomalleille. Organisaatiot siirtyvät käyttämään yhdessä sovittuja standardeja itse kehittämiensä sijaan. Tämä voi hyödyttää organisaatiota tulevien tietojärjestelmien määrittelyssä ja toteutuksessa.

My Datan kehityksessä tarvitaan avoimia standardeja eri alueilla. Usein puhutaan standardoinnin eri tasoisista, niin että ylemmän tason standardit tukeutuvat alempiin tasoihin. Käytännössä esimerkiksi internetin sovelluksiin ja dataan liittyvät standardit muodostavat toisiaan täydentävän verkoston, eikä niitä voida asettaa selkeään hierarkiaan, vaikka niin sanotussa OSI-mallissa[7] (Open Systems Interconnection model) tätä yritettiinkin.

Myös My Data -standardikehitystä on syytä lähestyä käytännöllisesti ilman kerrosrakennetta. Standardit kyllä tukeutuvat toisiinsa ja rakentavat kokonaisuutta yhdessä, mutta eivät hierarkkisessa järjestyksessä. Ei siis ole välttämättä mitään “pohjimmaisia” My Data -standardeja vaan eri standardeilla on erilaisia funktioita, joita yhdistelemällä saavutetaan systeemin tasolla hyvä yhteentoimivuus. Alla on listattu muutamia My Dataan liittyviä standardeja.

- Rajapintastandardit (REST)
- Dataformaatit, henkilötiedon tietomallit ja semantiikat (XDI, RDF ja W3C Linked data)
- Tiedon salaus ja kryptausmenetelmät (Bitcoin)
- Todentamiseen ja valtuuttamiseen liittyvät standardit (OAuth)
- Profiilitietoon liittyvät standardit (Orcid)
- Tiedon välittämiseen ja delegointiin liittyvä hallinnointistandardi (Respect Network)
- Tiedon lisensointiin ja käyttöehtoihin liittyvät standardit (Creative Commons, Open Notice)

Näiden standardien kehitystyö on jatkuvaa ja useimmiten avointa. Alkuvaiheessa on hyvä ymmärtää standardien kehityksen tila ja tehdä olemassa olevia standardeja hyödyntäviä kokeiluja, joiden avulla selviävät eri standardien mahdollisuudet ja yhteensopivuus. Uusien standardien kehityksessä on pyrittävä laajennettavuuteen, jotta vältytään tilanteelta, jossa lukittu, vanhanaikainen ja jäykkä standardi estää innovaation.

7 http://en.wikipedia.org/wiki/OSI_model

Standardien välillä vallitsee jatkuva kilpailu. Esimerkiksi yritykset pyrkivät edistämään omia tai itselleen edullisia standardeja. My Datan suunnittelussa on keskeistä, että toiminnan ytimessä olevat standardit ovat selkeitä ja yksinkertaisia, ja järjestelmä muuten kykenee hyödyntämään useita erilaisia kilpailevia standardeja.

Datakuitti – yksinkertainen tapa toteuttaa rajapinta

Datakuitti voisi olla joillekin yrityksille mahdollisimman helppo ensiaskel kohti My Data rajapinnan kehittämistä. Kun käyn kaupassa, saan kuitin lompakkooni. Vastaavalla tavalla voisin saada datakuitin datalompakkooni automaattisesti jokaisesta ostostapahtumasta tai koosteena vaikkapa kerran kuukaudessa. Datakuitti sisältäisi kattavasti tiedot, jotka yritys on kerännyt minusta asiakkaana ostotapahtuman yhteydessä. Datakuittiin voitaisiin liittää muutakin saatavilla olevaa tietoa kuten ostossisältöön liittyviä tuotekoodeja, ravintoselosteita, takuutietoja jne.

Yrityksen taustajärjestelmiin pitäisi avata mahdollisuus datakuitin automaattiseen lähettämiseen. Lisäksi yrityksen tulisi kuvata prosessit, joilla asiakas ottaa datakuitin käyttöönsä. Kevyimmillään datakuitin välittäminen voidaan toteuttaa vaikkapa sähköpostilla, jolloin datalompakkona toimisi tavallinen sähköpostitili. Käytännössä asiakas saa kuitin sähköpostiinsa, kun kirjoittaa kassapäätteeseen sähköpostiosoitteensa. Apple alkoi tarjoata tällaisia sähköpostitse lähetettäviä datakuitteja jo vuonna 2005. Sen jälkeen monet muutkin kauppaketjut ja palveluntarjoajat, kuten Urban Outfitters, Nordstrom, Macy's, Dick's Sporting Goods, Dillard's and Avis ovat seuranneet perässä (Charski 2013). Edistyneet syöteformaatit avaavat lisää mahdollisuuksia, mutta edellyttävät käyttäjiltä erikoistuneita työkaluja ja palveluita datakuittien vastaanottamiseen.

Ero paperikuitteja pursuavan nahkalompakon ja datalompakon välillä on se, että jälkimmäiseen voidaan asentaa hyödyllisiä ohjelmia, jotka käsittelevät ja havainnollistavat tietoa. Datalompakossa voi toimia vaikkapa reaaliaikainen talousseurantaohjelma. Luonnollisesti käyttäjä voi itse valita, mitä ohjelmia datalompakkoonsa asentaa, mutta kauppias voi myös suositella ohjelmia, jotka erityisesti ottavat huomioon hänen lähettämänsä datan. Datakuitin voi toteuttaa monella tapaa, ja datakuittia toteutettaessa on tärkeää, että data on hyvin määriteltyä, koneen luettavassa ja avoimessa muodossa eikä esimerkiksi pelkästään kuvaksi printtatuna pdf-tiedostona.

Olennaista on, että kontrolli omaan dataan on yksilöillä itsellään, ja että infrastruktuuripalveluiden tarjoajia on useita ja että palvelut ovat yhteentoimivia ja vaihdettavissa.

3. My Data -palveluinfrastruktuuri

Tässä luvussa esitellään erilaisia organisoitumistapoja, jotka mahdollistavat henkilötiedon yhdistämisen eri lähteistä. Organisoitumistapa vaikuttaa siihen, kuinka helposti hyödynnettävää henkilötieto on, kuinka läpinäkyvää tiedon käyttö on, kuinka hyvin rakenteet tukevat avointa kilpailua ja yhteiskehittämistä sekä kuinka ihmiskeskeistä ja yksilöä tukevaa henkilötiedon hyödyntäminen tulevaisuudessa on.

Hahmoteltu avoin My Data -infrastruktuuri on näkemys siitä, mitä teknisiä ja organisatorisia perusratkaisuja tarvitaan, jotta My Datan periaatteet voisivat toteutua. My Data -infrastruktuuri on palvelu- ja tietoinfrastruktuuria, mutta yksinkertaisuuden vuoksi nimitämme sitä myös pelkästään infrastruktuuriksi.

Tavoitteena on luoda luotettava ja pelkistetty infrastruktuuri, joka on avoin uusille toimijoille ja uusille innovaatioille. Ihmisille tämä tarkoittaa, että on olemassa helppokäyttöisiä infrastruktuuripalveluita kuten henkilötietoasetusten etähallinnointipalveluita, My Datan säilytyspalveluita ja omien profiilien ylläpitopalveluita. Nämä palvelut ovat paloittain laajennettavissa, ja palveluja voi helposti vaihtaa, koska data liikkuu niiden välillä My Data -rajapintojen ja operaattorien välisen yhteentoimivuuden ansiosta. Infrastruktuurin ominaisuuksien on taattava ihmisille muun muassa mahdollisuus suojata ja poistaa omaa tietoaan.

Kuka tai ketkä tarjoavat hallinnointipalveluita, ja pitäisikö näitä keskeisten infrastruktuuripalveluiden tuottajia valvoa tai säännellä kuten esim. pankkeja ja teleoperaattoreita säännellään nykyään? Muutaman miljoonan kuluttajan rikkaan profiilin hallinnointi on liiketoiminnallinen mahdollisuus, joka houkuttelee monia. Kukapa ei haluaisi olla koko kehittyvän toimialan keskiössä? Sääntelyä voidaan kehittää ajan myötä, mutta nyt alkuvaiheessa on syytä keskittyä luomaan infrastruktuurin pohja niin, ettei sinne synny perustavanlaatuisia valuvikoja, jotka myöhemmin haittaisivat tai estäisivät My Data -periaatteiden toteutumisen. Olennaista on, että kontrolli omaan dataan on yksilöillä itsellään, infrastruktuuripalveluiden tarjoajia on useita ja palvelut ovat yhteentoimivia ja vaihdettavissa.

3.1 Henkilötiedon organisointitapoja

Ennen kuin lähdetään tutkimaan, kuvaamaan ja kehittämään My Data -infrastruktuuria, on syytä kysyä, tarvitaanko infrastruktuuria ylipäänsä. Edellisissä luvuissa on kuvailtu My Data -rajapintoja ja sovelluksia. Eikö olisi helpompaa, jos nämä rajapinnat ja sovellukset keskustelisivat suoraan keskenään ilman välissä olevaa infraa? Nykykehitys onkin jo monin paikoin kehittymässä tällaisen orgaanisesti laajentuvan infrastruktuurittoman niin sanotun API-ekosysteemin suuntaan. Toisaalla taas syntyy yksittäisiä sektorikohtaisia aggregaattoreita, jotka pyrkivät keskittämään ja harmonisoimaan henkilötietoa omaan palveluunsa.

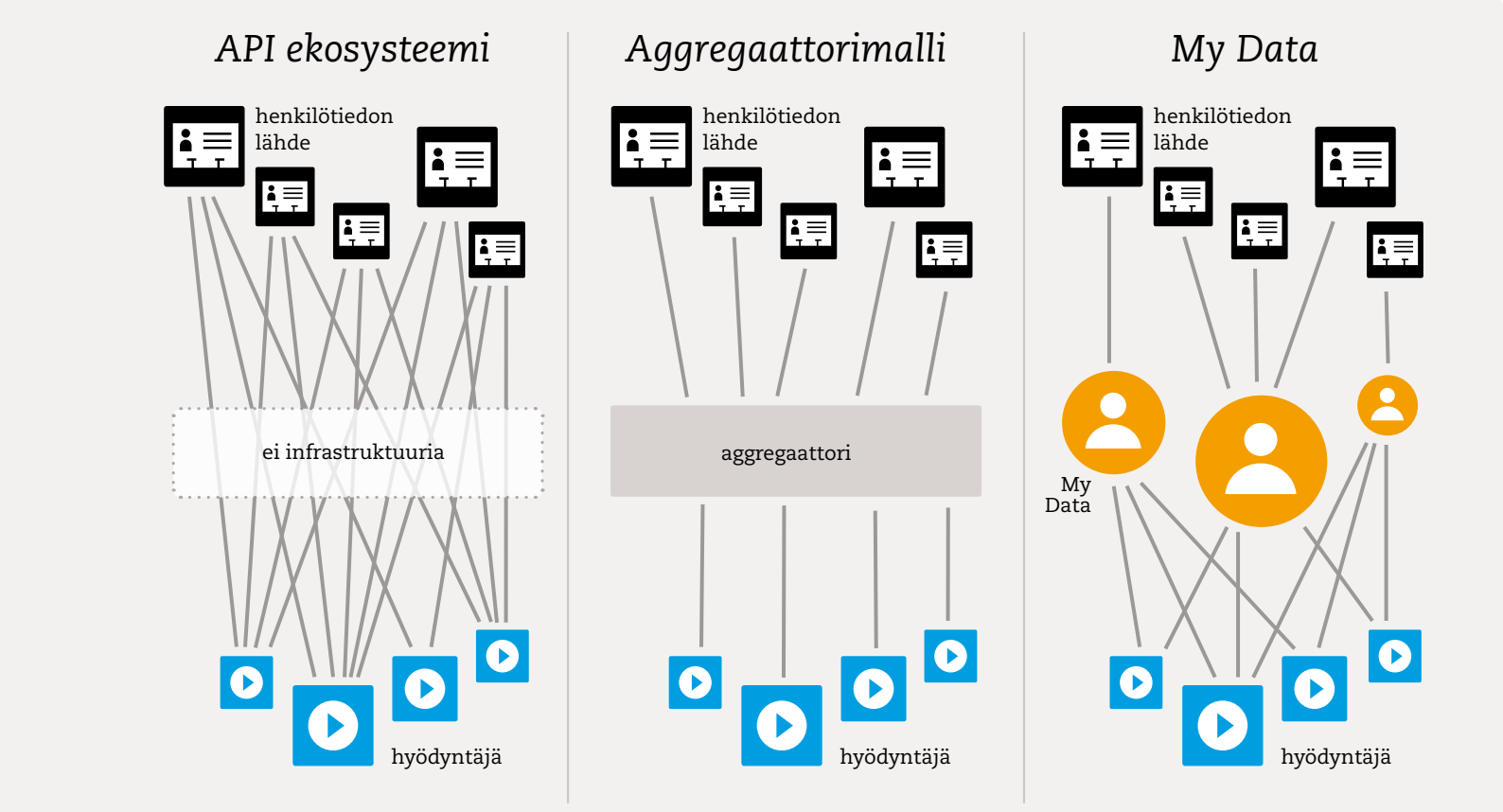
Vaihtoehtoisia henkilötiedon organisointitapoja on:

- **Infrastruktuuriton API-ekosysteemi,** jossa yksilö hallinnoi palveluiden yhdistymistä jokaisen palvelun sisällä erikseen (esimerkkinä nykyisten verkkopalveluiden APIt)
- **Organisaatiokeskeinen aggregaattorimalli**, jossa yksi organisaatio ottaa keskeisen roolin tiedon integraattorina (esimerkkinä Googlen ja Applen suljetut ekosysteemit tai aggregaattorit Taltioni[8] tai Mydex[9])

8 http://taltioni.fi/
9 https://mydex.org/

- **Avoin My Data -palveluinfrastruktuuri**, jossa henkilötiedon hallinta voidaan hajauttaa useamman keskenään yhteentoimivan operaattorin palveluun
- **Kansalliset järjestelmät**, jossa rakennetaan yksi alusta ja jonka operoinnissa julkishallinto ottaa keskeisen roolin henkilötiedon kerääjänä, jakelijana ja hyödyntämisen mahdollistajana (esim. kansalaistili, palveluväylä tai Viron X-road)

Nykykehityksessä näkyy viitteitä useista eri organisointitavoista, jotka tulevaisuudessa eivät välttämättä sulje pois toisiaan. Esimerkiksi on olemassa useita toimijoita, jotka pyrkivät toteuttamaan normaalia aggregaattorimallia avoimempaa mallia (kuten The Gooddata[10], Respect Network[11]), mutta My Data -infrastruktuurin puuttuessa pyrkivät asettumaan itse keskeiseen asemaan, eivätkä näin toteuta vaihdettavuuden periaatetta. Kuva 3.1 visualisoi kolmen keskeisen mallin eroja. Seuraavissa kappaleissa analysoidaan näitä malleja laajemmin.



Kuva 3.1: Erilaisia henkilötiedon yhdistämisen mahdollistavia organisointitapoja: vasemmalla infrastruktuuriton API-ekosysteemi, jossa kaikki datalähteiden ja sovellusten suhteet määritellään erikseen datalähteiden yksityisyysasetuksista, keskellä aggregaattorimalli, jossa yksittäinen toimija kerää ja harmonisoi dataa useasta lähteestä ja jakelee eteenpäin, oikealla avoin My Data -infrastruktuuri, jossa välittäjäorganisaatioita voi olla useita ja ne ovat kaikki yksilön palveluksessa. Huom. kuvassa viivat voivat kuvata datan liikkumista tai luottamussuhdetta. Osa datasta on käytännöllistä kerätä yhteen (datapankki), mutta osa tallennetaan syntypaikassa ja välittäjäorganisaatio ainoastaan huolehtii käyttöluvista (dataoperaattori). Tarpeettomien datakopioiden tekemistä pyritään välttämään.

10 https://thegooddata.org/
11 https://www.respectnetwork.com/

3.1.1 Infrastruktuuriton API-ekosysteemi

Rajapintojen laajasti yleistyessä voidaan puhua API-ekosysteemistä, missä uusia yhteyksiä eri palveluiden välille voidaan luoda nopeasti ja ketterästi ilman erillisiä taustalla olevien organisaatioiden kahdenvälisiä sopimuksia. Esimerkiksi REST[12] (RESTful API) on tapa tehdä rajapintoja, joka on yleistynyt nopeasti, ja sen toteuttamiseen ja hyödyntämiseen on paljon toteutustekniikoita ja oppeja, vaikka lähestymisen takana ei olekaan yksittäistä merkittävää edistäjäorganisaatiota. RESTin yleistyminen on esimerkki API-ekosysteemin itseohjautuvuudesta ja ketteryydestä. Yleisesti voidaan sanoa, että API-ekosysteemi edistää tiedon virtaamista, luo uutta liiketoimintaa ja on hyödyllinen edistysaskel digitaalisten palveluiden kehityksessä.

Nykyisellään rajapintojen ominaisuudet vaihtelevat merkittävästi. Kunkin palvelun toteuttaja voi vapaasti määrätä API:n protokollan ja ominaisuudet. Yleensä API:n kehittämisen motiivi on tehdä omasta palvelusta mahdollisimman keskeinen osa laajempaa palvelukokonaisuutta ja nähdä API:a hyödyntävät muut palvelut oman palvelun laajennuksina. Kun yksittäiset API:t on toteutettu palveluntarjoajan omien motiivien mukaisesti, usean palveluntarjoajan API:en yhdistämisestä muodostuu haastava kokonaisuus. Rajapintojen tekninen yhdisteltävyys varmasti paranee ajan myötä, kun syntyy yhteisiä rajapintastandardeja.

Henkilötiedon hallittavuuden ja ihmiskeskeisyyden kannalta infrastruktuurittoman API-ekosysteemin ongelma on palveluiden ja niiden välisten yhteyksien suuri määrä. Kun palvelujen välisten yhteyksien määrä kasvaa, niiden ylläpidettävyys monimutkaistuu. Yksilölle ei enää muodostu kokonaiskuvaa oman tietonsa sijainnista ja tiedon liikkumisesta palvelujen välillä, ja yleensäkin käyttäjien mahdollisuus hallita ja hyödyntää tietoa on heikko.

Infrastruktuurittomassa systeemissä ainoa tapa hallita, mitä tietoa mistäkin palvelusta menee muihin, on kirjautua jokaiseen palveluun erikseen, ja etsiä se paikka, jossa näytetään, mille kaikille muille palveluille on myönnetty lupa rajapinnan kautta lukea käyttäjän dataa. Nykyisin verkkoaktiivisella kuluttajalla on käytössään ehkä parikymmentä palvelua, jotka voivat rajapintojen kautta kytkeytyä muihin palveluihin. Näitä ovat esimerkiksi sosiaalisen median palvelut, monet mobiilisovellukset ja internetiin kytketyt elektroniset laitteet kuten digivaa'at ja aktiivisuusmittarit. Muutaman palvelun erillinen hallinta on vielä mahdollista, mutta erilaisen henkilökohtaisen ja kotiin liittyvän anturiteknologian yleistymisen myötä hallittavia datalähteitä voi tulevaisuudessa olla kymmenien sijaan satoja tai tuhansia, ja silloin tiedonhallintaa helpottavan infrastuktuurin tarve on jo ilmeinen.

3.1.2 Organisaatiokeskeinen aggregaattorimalli

Tällä hetkellä valtaosa henkilötiedosta sijaitsee organisaatioiden hallinnoimissa tietokannoissa ja tiedon käyttö tapahtuu yksilöille läpinäkymättömällä tavalla. Tästä hyvänä esimerkkinä ovat vaikka kansainväliset verkkomainonnan yritykset ja niiden väliset verkostot. Monelle organisaatiolle henkilötieto, asiakkuudet ja profiilit ovat kauppatavaraa. Myös julkisella sektorilla henkilötietoja yhdistetään valtaosin organisaatioiden välisiin sopimuksiin perustuvina yksittäisinä järjestelmäintegraatioina.

Kun meillä ei vielä ole avoimiin standardeihin perustuvaa henkilötiedon yhteentoimivuutta, niin yksittäiset globaalisti toimivat yritykset laajentavat kukin omaa henkilötiedon ekosysteemiään ja pyrkivät suuren käyttäjävolyymin voimalla de facto -standardin asemaan. Liian pitkälle vietynä tässä kehityskulussa on riski, että muiden pelureiden tulo markkinoille estyy tai ne joutuvat alihankkijan asemaan ilman valinnan tai vaikuttamisen mahdollisuuksia.
Viimeisien vuosina eri sektoreilla on alkanut tapahtua kehitystä, joka mahdol-

12 http://en.wikipedia.org/wiki/Representational_state_transfer

listaa useiden toimijoiden keräämän tiedon yhteiskäyttöä. Terveyssektorilla tästä useita esimerkkejä eri maissa, kuten sveitsiläinen Healthbank[13], brittiläinen Patients Know Best[14] ja suomalainen Taltioni. Tällaisessa rakenteessa yritykset perustavat yhteisen palvelun, josta tulee tiedon välittämisen keskipiste. Keskittäminen edistää tiedon yhdistämistä ja uusien käyttötapojen kehittämistä, mutta samalla järjestelmä tulee riippuvaiseksi yksittäisestä toimijasta, joka päättää toiminnan tavoitteista ja tekemisen tavoista.

3.1.3 Avoin My Data -palveluinfrastruktuuri

My Data -lähestyminen eroaa rajapintaekosysteemistä siinä, että ihmisillä on suora hallinta tiedon välitykseen yleiskäyttöisten infrastruktuuripalvelujen avulla. Tieto virtaa infrastruktuurin välittämänä keskitetymmin (tämä on eri asia kuin keskitetysti, mikä ei ole tavoiteltavaa), mikä helpottaa ja tehostaa hallintaa ja luo paremmat edellytykset sovellusten kehittämiselle.

Toisin kuin aggregaattorimallissa My Data -mallissa on kuitenkin useita kilpailevia infrastruktuuripalveluita, jotka toimivat yhteen avoimuuden ja standardoinnin ansiosta. Ihminen voi valita itselleen sopivimmat palvelut muun muassa tiedon tallennukseen ja useiden datalähteiden keskitettyyn hallinnointiin. Hallintapalvelut voivat tulla usealta toimittajalta tarpeiden mukaan. Malli muistuttaa tapaa, jolla rahavirtoja hallitaan nykypäivänä. Kaikki käyttävät samaa rahaa, eri toimijoiden välillä on yhteistoiminnan periaatteita, ja pankin vaihtaminen onnistuu verrattain helposti. Seuraavaksi käsitellään keskeiset My Data -infrastruktuuripalvelut ja niin sanottu operaattorimalli.

3.2 Infrastruktuuripalveluita

Infrastruktuuripalveluilla tarkoitetaan sellaisia peruspalveluita, jotka mahdollistavat toisaalta yksilön kontrollin omaan dataansa ja toisaalta helpottavat yksittäisten My data -sovellusten toteuttamista, kun jokaisen sovelluksen kehittäjän ei tarvitse erikseen toteuttaa samoja ominaisuuksia omaan palveluunsa. Infrastruktuuripalveluita ovat muun muassa hallintapalvelut, tallennuspalvelut ja autentikaatiopalvelut.

3.2.1 Hallintapalvelut

Hallintapalvelu on keskeinen osa My Data -infrastruktuuria. Hallintapalvelu on paikka, missä My Data -rajapinnat, datan varastointi, sovellusten jatkohyödyntäminen, yksilöiden oma datan hyödyntäminen ja datan anonymisointi kohtaavat.

Edistyneissä hallintapalvelussa tulisi olla seuraavanlaisia ominaisuuksia:

- Rajapintoihin liittyminen ja tarvittava autentikaatio
- Yksityisyysasetusten etähallinta
- Henkilötiedon tallennus yksilön niin halutessa (ks. tallennuspalvelut)
- Tiedolle tehtävät operaatiot ja mahdollisuus operoida paikallisia sovelluksia
- Tiedon jakelu anonymisointia ja sitä seuraavaa julkistamista varten

13 http://healthbank.ch/
14 http://www.patientsknowbest.com/

Hallintapalvelun hyvä toiminnallisuus vaatii ylläpitäjältä suhteellisen laajaa palvelupakettia. On epärealistista, että jokainen henkilötietoa hyödyntävä taho voisi toteuttaa kaikki listatut ominaisuudet. Hallintapalvelun toiminnan analogiana voisi käyttää sähköpostijärjestelmän toimintaa. Saman sähköpostitilin sähköposteja voidaan vastaanottaa ja lähettää usealla eri ohjelmalla, ja tiettyjen ohjelmien avulla voidaan hallinnoida useita sähköpostitilejä.

3.2.2 Tallennuspalvelut

My Data -lähestymisessä keskeistä on henkilötiedon hallinnan organisoiminen. Yksilön luvalla tieto voi virrata suoraan palvelusta toiseen ilman, että sitä tallennetaan välillä. Tällaisia palvelujen välisiä kytköksiä hallinnoidaan edellä esitellyllä hallintapalvelulla. On kuitenkin useita syitä, miksi yksilö voi haluta tallettaa ja arkistoida tietoa itselleen, tätä varten tarvitaan tallennuspalveluja.

My Data -infrastruktuurissa voi olla erilaisia tallennuspalveluita. Näitä voidaan kutsua tileiksi lainaten metaforaa, joka on tuttu pankkitileistä, sähköpostitileistä ja asiakastileistä. My Data -tilien ominaisuuksia on esitelty alla olevassa listassa.

- My Data -tileille voi kerätä ja tallettaa omat tiedot erilaisista henkilötietorajapinnoista.
- My Data -tilejä voi olla useita, osa tileistä voi olla usean henkilön yhteisiä ja osa voi olla pseudonyymitilejä.
- Tileillä on erilaisia autentikaatiotasoja.
- Tilin hallintaohjelma pitää automaattisesti rekisteriä siitä, mitä tietoa on luovutettu eteenpäin, ja osaa hallita tiedon eteenpäin luovuttamista.
- Tilin yhteydessä voi ajaa erilaisia sovelluksia datan muokkaamiseen, kuvaamiseen ja analysointiin.
- Tilin yhteydessä voi olla dataa jalostavia toimintoja, jotka voivat tuottaa esimerkiksi profilointeja datasta ilman raakadatan eteenpäin luovutusta.
- Tilejä voi linkittää toisiinsa ja niitä voi hallinnoida suoraan tai erillisen My Data -hallintapalvelun avulla. Osa tileistä voi sijaita kotikoneella, osa palveluntarjoajan huomassa.
- Tili voi olla kryptattu datavarasto, jolloin tallennuspalvelun tuottajalla ei ole pääsyä informaatiosisältöön.

Paikalliset ja siirrettävät sovellukset

Paikallisilla sovelluksilla tarkoitetaan ohjelmia, jotka eivät lähetä käyttäjän dataa eteenpäin toiselle verkkopalvelimelle vaan toimivat siten, että sovellus tai dataa analysoiva koodi ladataan verkosta datan luo.

Paikalliset sovellukset voivat toimia käyttäjän päätelaitteella tai käyttäjän palvelimella tai, mikäli data on säilössä, jossain pilvipalvelussa. Silloin paikallisuudella tarkoitetaan, että sovelluksen ajoympäristö on datatilin yhteydessä. Olennaista on, ettei käyttäjän dataa tarvitse siirtää sieltä, missä se oli alunperin.

Paikallisten sovellusten toimintalogiikka on siis käänteinen verrattuna sovelluksiin, joiden luo data lähetetään. Useimmiten paikallista sovellusta käytetään selaimella kuten mitä tahansa verkkosovellusta, eikä käyttökokemus eroa perinteisestä verkkopalvelimilla toimivista palveluista ja sovelluksista.

Kun datan määrä on suuri, niin on käytännöllisempää tuoda ohjelmakoodi datan luokse. Päätelaitteiden ja selaimien ominaisuudet ovat kehittyneet niin paljon, että moni toiminnallisuus, jonka toteuttaminen aiemmin oli mahdollista vain palvelimella, voidaan nykyisin tehdä päätelaitteessa.

Unhosted https://unhosted.org/
OwnCloud https://owncloud.org/

3.2.3 Autentikaatio- ja luottamuspalvelut

My Data -infrastruktuurissa on monenlaisia autentikaatiotarpeita. Ihmisten ja heidän My Data -tilien autentikoimisen lisäksi myös data ja palvelut pitää pystyä autentikoimaan. Kun ihminen operaattorin kautta välittää dataa eteenpäin, niin vastaanottavan palvelun pitää tietää, onko data todella sitä, mitä sen väitetään olevan, vai onko sitä muuteltu matkalla. Esimerkiksi tulevaisuuden rekrytointipalvelu ottaa vastaan datana opiskeluhistorian ja haluaa vahvistuksen tiedon oikeellisuudesta. Vastaavasti ihmisten oikeusturvan kannalta on merkityksellistä, että teknisesti voidaan varmistaa, etteivät erilaiset palvelut vääristele ihmisten dataa heidän tietämättään.

Teknologian kehitys saattaa tuoda ratkaisun siihenkin, miten voidaan todistaa toisaalla syntyneen datan autenttisuus myöhemmin ja eri käyttökontekstissa. Proof of existence[15] -verkkopalvelu on konseptiehdotus siitä, miten Bitcoin-kryptovaluutan pohjalla olevaa teknologiaa voidaan hyödyntää aikaleimaamaan dokumentti niin, että myöhemmin voidaan tarkistaa bitin tarkkuudella, onko dokumentti sama vai onko sitä peukaloitu matkalla. Dokumentista lasketaan kryptografinen sormenjälki, joka tallennetaan hajautetusti verkkoon. Jos dokumentti on bitilleen sama kuin alkuperäinen, niin myöhemmin uudelleen laskettuna sormenjälki täsmää, mutta jos pilkkuakin on muutettu, niin muutos paljastuu, koska sormenjälki ei enää täsmää.

Sähköinen luotettava identiteetti on tärkeä mahdollistaja kansalaiselle ja organisaatioille. Mikäli luottamus toisen osapuolen sähköisen agentin autenttisuuteen puuttuu, mitään tietoa ei uskalleta antaa My Data -rajapintojen kautta. Autentikaatio ja luottamuspalvelut ovat keskeinen osa My Data -infrastruktuuria. Osin My Data -tilien ja oikeuksienhallintaprotokollan pitää ottaa huomioon tietoturva ja autentikaatiokysymykset jo standardin tasolla, mutta tärkeää olisi, että järjestelmä mahdollistaisi luottamuspalvelujen rakentamisen perustan päälle. On mahdollista, että toiset haluavat enemmän joustavuutta, vaikka se tarkoittaisi tietoturvan tasosta tinkimistä, ja toisille taas on ehdottoman tärkeää saada maksimaalinen tietoturva.

3.2.4 Anonymisointipalvelut

My Datan hyödyntäminen perustuu usein mahdollisuuteen yhdistellä eri lähteistä peräisin olevaa henkilötietoa. Tämä vaatii, että tieto yhdistyy siihen liittyvään ihmiseen. Monien käytännön sovellusten kannalta olisi kuitenkin tärkeää, jos yksilö voisi käyttää omaa autenttista dataansa eri yhteyksissä ilmiantamatta omaa identiteettiään. Anonymisoinnin tarve tulee yleisimmin vastaan, kun tehdään populaatiotason tutkimusta. Tutkijat ja data-analyytikot haluaisivat hyödyntää ja yhdistellä laajoja datamassoja, mutta heitä ei käytännössä kiinnosta yksittäinen yksilö. Jotta mahdollistetaan laajat big data -tutkimukset ja tutkimusaineistojen yhdistäminen yksilön tietosuojaa heikentämättä, My Data -infrastruktuurissa on oltava luotettavia anonymisointipalveluita, jotka mahdollistavat data yhdistämisen ja anonymisoinnin tutkimustarpeisiin.

Erilaisten de-anonymisointitekniikoiden kehitys ja lisääntyneet mahdollisuudet yhdistää dataa useista lähteistä ovat siirtämässä ja hämärtämässä rajoja yksilöivän henkilötiedon ja anonyymin tiedon välillä. Tämä on johtanut kiivaaseen keskusteluun, jossa toiset hyvin perustein väittävät, ettei aukoton anonymisointi ylipäätään ole mahdollista, ja että kaikkea alunperin yksilöihin liittyvää tietoa pitää käsitellä henkilötietona lain säätämällä tavalla. Toiset taas väittävät yhtä lailla hyvin perustein, että näin tiukka tulkinta rajoittaisi monia nimettömien tietojen käyttötapoja, joissa hyödyt ovat selvästi suurempia kuin yksityisyyden suojan menetyksestä aiheutuvat haitat. Anonymisointipalvelujen toteuttaminen on teknisesti ja sosiaalisesti haaste. Todelliset hyödyt syntyvät, kun palvelulla on

15 http://www.proofofexistence.com/

käytössä kymmenien tai satojen tuhansien ihmisten dataa. Kun tähän kokoluokkaan päästään, voidaan alkaa tehdä tiedettä ja älykkääseen hallintoon liittyvää tutkimusta ennen näkemättömällä tasolla.

My Data ja Big data

Big datalla viitataan äärimmäiseen suureen ja nopeasti karttuvaa tiedon määrään, jonka kerääminen, tallennus ja analyysi vaativat uusia käsittelymenetelmiä. Toisaalta big datan voi käsittää myös tiedon paradigman muutoksena. Sen myötä yrityksissä ja hallinnossa voidaan yhä useammin tehdä päätöksiä, jotka perustuvat suoraan kerättyyn ja mitattuun tietoon. Tutkimuksessa on mahdollista muodostaa teoriaa uusilla tavoilla, kun datamassojen analysointi ja yhdistely on entistä helpompaa. Laajojen datamassojen sovellusmahdollisuudet ovat lähes rajattomat ja tiedon hyödyntämisestä on tullut yhä vahvemmin kilpailukyvyn edellytys alasta riippumatta. (LVM 2014)

Suuria datamääriä syntyy mm. internettiin kytketyistä laitteista, anturijärjestelmistä, sosiaalisesta mediasta, verkon yli tehtävistä transaktioista, yritysten liiketoimintaan liittyvistä ohjaus- ja raportointijärjestelmistä jne. Keskeinen osa big datasta on ihmisten käyttäytymisdataa, joka perustuu asiakkaan tunnistamiseen. Big data -keskustelussa korostetaan henkilötietojen analysoinnin ja hyödyntämisen mahdollisuuksia organisaatioiden näkökulmasta. Ihmisten näkökulma on supistettu usein vain vaatimukseen siitä, että yksityisyydensuoja säilytetään. Asiakkaan kiinnostusta saati oikeutta omiin tietoihinsa ei big data -keskustelussa ole juurikaan tuotu esille.

Henkilöihin liittyvässä tiedossa My Data ja big data ovat kaksi toisiaan täydentävää näkökulmaa, ”ihmisnäkökulma” ja ”yritysnäkökulma”. My Data tuo läpinäkyvyyttä ja sitä kautta hyväksyttävyyttä henkilöihin liittyvien datamassojen käsittelyyn ja antaa konkreettisia keinoja yksityisyydensuojan toteuttamiseen. Ilman tätä ihmisnäkökulmaa monet big datan hyödyntämismahdollisuudet katoavat, koska ne eivät ole yksilöiden suojan kannalta hyväksyttäviä.

3.3 My Data -operaattori

On lähes välttämätöntä, että infrastruktuurilla on jonkinlaisia välittäjäorganisaatioiden roolissa olevia toimijoita, jotka muun muassa ylläpitävät ja kehittävät edellä esitettyjä keskeisiä peruspalveluita. Tällaisista organisaatioista käytetään yleisnimitystä My data -operaattori. Muissa yhteyksissä käytetään nimitystä datapankki, databroker tai data-aggregaattori kuvaamaan samanlaista tai osin samanlaista toimijaa. Tässä raportissa tarkoitamme My Data -operaattorilla yksilön infrastruktuuripalveluita ylläpitävää organisaatiota. Operaattori ylläpitää ’My Data -tilejä’, joilla henkilötietojen käyttöön liittyviä oikeuksia hallitaan. Tilin hoitoon voi kuulua tiedon varastointia tai välittämistä.

Operaattorimalli perustuu luottamukseen ihmisten ja operaattoriorganisaatioiden välillä. Koska henkilötiedon kontrolli halutaan säilyttää yksilöillä itsellään, tarkoittaa se, että My Data -operaattorien tulisi olla ihmisten eikä yritysten palveluksessa. Jokaisella voisi olla luotettuja My Data -operaattoreita eri luonteisille tiedoilleen: terveystiedot, omaisuustiedot, kuluttajaprofiilitiedot, liikkumisprofiilitiedot jne. Yksilöt antavat datan välittämiseen, jalostamiseen ja säilyttämiseen liittyviä tehtäviä näille operaraattoreille, koska eivät halua ylläpitää omia datavarastoja kotikoneellaan, vaikka se teknisesti olisi mahdollista. Halutessaan ihminen voisi hallita kaikkia tietojaan myös vain yhden operaattorin kautta.

Operaattorimalli toimii myös sovellustason kilpailun mahdollistajana. Sovelluskehittäjille operaattorit tarjoavat riittävän suuren asiakaspotentiaalin ja yhtenevän sovellusintegraatiopisteen, mutta yksilöitä palvelevilla operaattoreilla ei kuitenkaan ole intressiä sulkea markkinoita esimerkiksi tukemalla vain tiettyjä yksittäisiä sovelluksia.

Kypsän My Data -infrastruktuurin toteuttajina ja hyödyntäjinä voi olla lukuisia osin keskenään kilpailevia ja osin toistensa toimintaa täydentäviä operaattoreita. Operaattorin organisoitumismalli vaikuttaa siihen, minkälaisella arvoketjulla se rakentaa omaa (liike)toimintaansa ja minkälainen investointikyky ja motivaatio sillä on. Seuraavassa esimerkkejä operaattorin ansaintamalleista:

- Yksilö maksaa palvelusta (tilinhoitomaksu).
- Operaattori saa tuloja toimiessaan sovellusten jakelukanavana (jakelija).
- Operaattori saa tuloja välittäessän yksilön puolesta henkilötietoa (aggregaattori-operaattori).
- Operaattori tarjoaa lisäarvopalveluita (ensisijainen palveluntarjoaja).

Kun infrastruktuuri on vasta hahmottumassa, on tärkeää kiinnittää huomiota siihen, millaiset operaattorit saavat jalansijaa alkuvaiheessa, koska niiden toiminta vaikuttaa keskeisesti infrastruktuurin kehityssuuntaan. Seuraavaksi on listattu erilaisia operaattoreiden organisoitumiseen vaikuttavia ominaisuuksia:

- Valtiollisesti lisensoitu operaattori – Nykyisiin teleoperaattoreihin tai pankkeihin vertautuva malli, jossa valtiollinen taho myöntää operaattorille luvan tarjota yksilöille henkilötiedon hallintapalveluita, autentikaatiota ja mahdollisesti myös yrityksille ja organisaatioille datan varmentamista. Lisensointi ei ole rajoittunut operaattorin organisoitumismalliin, eli osuuskunta voi olla yhtälailla lisensin haltija kuin kaupallinen yrityskin.
- Verkosto – Vertaisperiaatteilla ilman operaattoria toimiva infrastruktuuri (p2p-pohjainen itseoperaattoreiden verkosto), joka nojautuu täysin teknologiaan. Liiketoimintarakenteiden kehittyessä myös vertaismalliin perustuville infrastruktuureille ilmaantuu usein operaattoreita, jotka tekevät vertaisverkon käytöstä yksinkertaisempaa ja helpompaa asiakkailleen. Näille järjestelmille on kuitenkin ominaista, että operaattoria ei ole pakko käyttää, vaan yksittäinen henkilö (vertainen) voi toimia myös itse omana operaattorinaan (rahan käyttö ilman pankkia vielä onnistunee, mutta matkapuheluita ei voi soittaa ilman operaattoria).
- Yritysten yhteenliittymä – Ryhmä yrityksiä (todennäköisesti henkilötiedon lähteitä ja hyödyntäjiä) perustaa yhteenliittymän ja käynnistää operaattoritoiminnan. Tästä hyvänä esimerkkinä on Taltioni.
- Ihmisten yhteenliittymä – Osuuskunnat, joissa ihmiset olisivat omistajia, voisivat olla myös My Datan hallinnointiin joissain tapauksissa sopiva malli. Yhteisten resurssien hoitoon perustettiin jo 1800-luvulla esim. vesi ja puhelinosuuskuntia. Esimerkiksi henkilötietoa käsittelevä sveitsiläinen healthbank.ch on organisoinut toimintansa yksilöiden omistamaksi osuuskunnaksi.
- Valtiollinen operaattori – Esimerkiksi valtion hallinnoimat kansalaistili, KanTa, Ruotsin terveystili yms. ovat esimerkkejä valtiollisten operaattorien (yleensä jokin virasto) hallinnoimista järjestelmistä. Monet teleoperaattorit ovat alun perin käynnistyneet kansallisina toimijoina,mutta ne on yksityistetty liiketoimintakentän kilpailun ja kansainvälistymisen vuoksi.
- Itsenäinen kaupallinen osakeyhtiö – My Data -operaattori voi olla täysin kaupallisilla periaatteilla toimiva organisaatio. Tällaisen organisaation kasvun ja kansainvälistymisen edellytykset ovat tarjolla olevien rahoitusmallien takia mahdollisesti parhaat. Alkuvaiheessa täysin kaupallisilla interesseillä itsenäisesti toimiva taho voi olla hankalassa tilanteessa, koska osa avoimien standardien ominaisuuksista saattaa olla kompromissi pysyvän kilpailuedun tai liiketoiminnan suojaamisen kannalta. Osakeyhtiömalli toimiikin mahdollisesti paremmin, kun operaattorimalli

on yleistetty ja markkinoiden muoto, yleisimmät liiketoimintamallit ja koko alkaa hahmottua.

Eikö olisi yksinkertaisempaa, jos olisi vain yksi iso tietokanta?

My Data -infrastruktuurin keskeisiin tavoitteisiin kuuluu henkilötiedon hallinnointi- ja hyödyntämispalvelujen organisointi yksinkertaisesti ja ihmiskeskeisesti. Operaattorimalli monimutkaistaa asioita. Ei olekaan enää yhtä pistettä, jonka kautta kaikki tieto yhdistyy, vaan on useita rinnakkain toimivia operaattoreita.

Tällaisen avoimen, tietojen ja palvelujen yhteentoimivaa vaihdettavuutta tukevan operaattorikentän synnyttämistä pidetään kovana urakkana. Miksi ei vain tyydyttäisi kansainvälisten, keskenään kilpailevien, mobiiliviestintää tai sosiaalisen median palveluita tarjoavien yhtiöiden walled-garden tyyppisiin ekosysteemeihin My Data -tiedon säilytys ja kehityspaikkoina - tai tavoiteltaisi yhtä, vaikkapa kansallista, toimijaa keskitetyksi operaattoriksi?

Keskeinen yhden organisaation malliin liittyvä ongelma on koko järjestelmän riippuvuus tästä yhdestä toimijasta (single point of failure). Kun on yksi taho, niin ongelmien syntyessä ne koskevat kaikkia, ja seuraukset voivat olla katastrofaalisia.

Useampi operaattori mahdollistaisi myös ketterän ja monipuolisen palvelukehityksen ja vaihtoehtoisten kilpailevien infrastruktuuripalveluiden rinnakkaisen kehittymisen. Toiset kaipaavat enemmän suojaa, kun taas toiset arvostavat järjestelmän keveyttä ja vapautta tehdä asioita itsenäisesti. Yhden organisaation malli saattaa helposti muuttua jäykäksi ja hitaaksi, eikä sovellu kevyiden käyttötapausten ketterään toteuttamiseen.

Kun olemme keskustelleet erilaisista henkilötiedon organisointitavoista kansainvälisissä kontekstissa, on tullut selväksi, että monessa muussa maassa kansalaisilla on huomattavasti vähemmän luottamusta hallintoon kuin Suomessa. Näissä maissa kansallisesti organisoitu tietojärjestelmä ei soveltuisi arkielämää ja yksilönvapauttaa korostavien sovellusten keskiöksi. On jopa poikkeuksellista, että Suomessa osa uskoo keskitetyn mallin mahdollisuuksiin. My Datan kannalta on keskeistä, että toimintamalleilla on mahdollisuus kansainvälisesti laajamittaiseen vaikuttavuuteen. Operaattorimalli on näin eri näkökantojen valossa arvioitu kestävimmäksi ja kansainvälisen yhteisen toimintatavan kannalta parhaaksi lähestymistavaksi.

Itsestään ja ympäristöstään oppiminen on tärkeä My datan hyödyntämisalue.

4. My Datan hyödyntäminen

Aiemmissa luvuissa olemme esitelleet My Datan hyötyjä menemättä konkreettisesti niihin käyttötapoihin, joita My Datalle olisi. Oikeuksien vahvistuminen ja datan parempi hallinta tuskin innostavat suuria joukkoja, vaan ilmiötä edistävät voimat tulevat ensisijaisesti siihen liittyvien arkea helpottavien käyttötapausten kautta.

Mahdollisten käyttötapausten kenttä on laaja ja sitä voidaan jaotella esimerkiksi sektoreittain: omaan terveyteen, omaan talouteen tai omaan liikenteeseen jne. liittyvät sovellukset. Jaotteluna voisi yhtä hyvin olla käyttäjäryhmät, kuten yksittäiset ihmiset, tutkijat, julkishallinto jne. My Dataan pohjautuen voidaan tehdä vakavamielisiä hyötysovelluksia, mutta myös viihdettä ja pelejä. Esimerkiksi Lontoon julkisen liikenteen matkakortilla voi pelata omaan matkadataan pohjautuvaa sosiaalista Chromaroma[16] -peliä verkossa.

Yksittäisten esimerkkien listaamisen ja luokittelun sijaan tässä luvussa käydään läpi kaksi laajempaa hyödyntämisaluetta, joita tarkastellaan syvemmin erityisesti siitä näkökulmasta, miten ne muuttuisivat My Datan myötä.

Itsestään ja ympäristöstään oppiminen on tärkeä My Datan hyödyntämisalue. Siinä omaan toimintaan liittyvää dataa kerätään ja hyödynnetään toiminnan muuttamiseksi. Maailmanlaajuisesti leviävä itse mittaamisen (Quantified Self) trendi on yksi esimerkki tästä alueesta. My Data mahdollistaisi ihmisten toiminnan tuloksena automaattisesti syntyvän laajan niin sanotun digitaalisen jalanjälkidatan hyödyntämisen oman oppimisen tukena.

Yritysten näkökulmasta keskeinen My Datan hyödyntämisalue on tuotteiden parempi kohdentaminen ja markkinointi. Nykyään puhutaan paljon profiloinnista ja big datasta, joilla yritykset pyrkivät paremmin ymmärtämään asiakkaitaan. My Data tuo tähän keskusteluun ihmisnäkökulman. Ihmisten itse hallinnoimat profiilit mahdollistaisivat tarkempia suosittelujärjestelmiä ja läpinäkyvyys lisäisi luottamusta yrityksiä kohtaan.

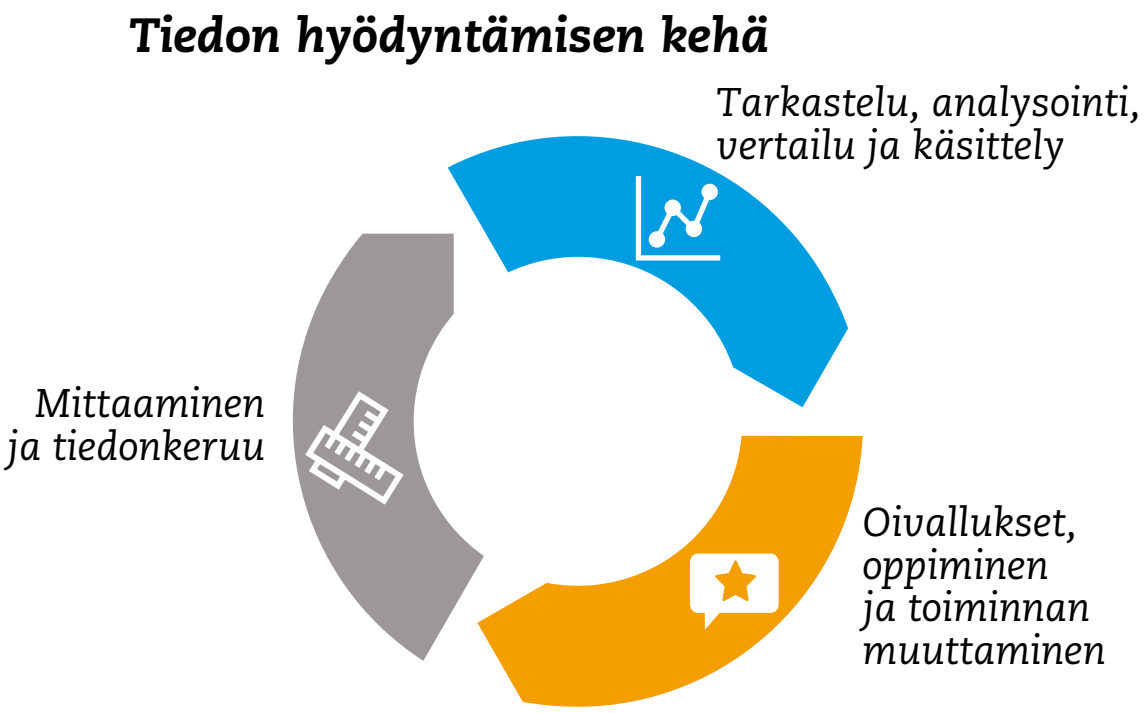
On syytä korostaa, että valitut kaksi hyödyntämisaluetta eivät pyri esittämään kattavasti kaikkea sitä, mitä My Datalla voidaan tehdä. Tulevaisuuden merkittävien käyttötapojen ennakointi on hankalaa, mutta nykyhetkestä katsottuna vaikuttaa siltä, että ainakin näillä kahdella alueella tapahtuu paljon. Luvun loppuun on koottu muutamia design -huomioita My Data -sovelluksista.

4.1 Itsestään ja ympäristöstään oppiminen

Toiminnasta saatava palaute on oppimisen ja itsensä ymmärtämisen edellytys. Tyypillisesti palaute on aistinvaraista tai muilta ihmisiltä saatavaa sosiaalista palautetta, mutta monessa tapauksessa perinteiset palautteen muodot ovat vajavaisia. Itsensä mittaamisen työkalut ja sovellukset mahdollistavat palautteen saamisen asioista, joita olisi muuten hankala havaita, kuten vaikkapa mihin hurahtavat päivän aikana tietokoneen ääressä käytetyt tunnit. Omaan toimintaan liittyvää tietoa kertyy automaattisesti moniin paikkoihin kuten pankin ja ruokakaupan tietojärjestelmiin. Tällainen niin sanottu digitaalinen jalanjälki, mikäli se olisi saatavilla My Datana, voitaisiin jalostaa oppimista tukevaksi palautteeksi samoin kuin aktiivisesti itse mitattu datakin.

16 http://www.chromaroma.com/about

Riippumatta siitä, onko oppimista tukeva tieto kerätty mittaamalla vai saatu suoraan palveluntarjoajien rajapinnoista, sen hyödyntäminen etenee samalla tavalla. Alla olevassa kuvassa on esitetty sykli, jossa tiedon keräämistä seuraa sen tarkastelu ja analysointi ja viimeisenä vaiheena tiedon reflektointi eli tietoon liittyvien ajatusten ja tunteiden käsittelyyn. Tässä syklissä syntyy oivalluksia ja oppimista, jotka muuttavat toimintaa.



Kuva 4.1: Itsestään oppimisen sykli. Kuinka henkilötieto jalostuu yksilön omassa käytössä opeiksi ja oivalluksiksi.

Ihminen voi seurata omaa käyttäytymistään ja kulutustaan ja vertailla sitä omaan aikaisempaan dataan tai viiteryhmien vastaavaan dataan. Lisäksi tietoa on mahdollista rikastaa kirjaamalla ylös esimerkiksi mittaushetkeen liittyviä muita tapahtumia, kokemuksia ja tuntemuksia.

Perimmäinen tavoite itsensä ja oman toimintansa kuten kulutuksen mittaamisessa on yleensä oppiminen ja toiminnan ohjaaminen tavoiteltavaan suuntaan. Migreeniherkkä voi pyrkiä välttämään tilanteita, joissa on aiemmin saanut kohtauksen, maratoonari haluaa optimoida harjoittelun ja levon suhdetta, perheellinen saattaa seurata rahan menoa säästääkseen ja niin edelleen.

Kerätty data ja sen analysointi ovat apuvälineitä, vasta reflektio muuttaa toimintaa. Reflektio on yleiskäsite ajatus- ja tunneprosesseille, joilla ihminen selvittää kokemuksiaan. Siihen kuuluu oman käyttäytymisen tiedostaminen ja muuttaminen sekä vaihtoehtoisten tulkintojen ja lähestymistapojen luominen.

Oman datan seuraaminen, tulkinta ja vertailu herättävät ihmisessä ajatuksia ja tunteita. Ihminen saattaa pohtia datasta esille nousseita kysymyksiä läheistensä tai oman tukiverkostonsa kanssa. Reflektointi yksin ja yhdessä muuttaa käyttäytymistä. Esimerkiksi päivittäinen painon seuranta voi auttaa laihduttamaan, vaikka muita tietoisia laihduttamiseen tähtääviä toimia ei tehtäisi.

Seuraavaksi käsitellään itsensä mittaamista ja automaattisesti kertyvää niin sanottua digitaalista jalanjälkeä tapoina saada oppimista tukevaa ihmisen toimintaa kuvaavaa dataa. Itsensä mittaamisella kerätty data on hyvä lähtökohta toimintatapamuutoksille, mutta vielä suuremmat mahdollisuudet aukeavat, mikäli My Data -periaatteiden noudattamisen myötä ihmiset saavat laajemmin

käyttöönsä erilaisiin palveluihin ja rekistereihin automaattisesti kertyvää dataa. Digitaalinen jalanjälkidata, joka pohjautuisi vaikka ostos-, liikkumis-, energiankulutus- ja terveystietoihin, mahdollistaisi kokonaisvaltaiset jalanjälkimittarit.

4.1.1 Itsensä mittaaminen

Quantified Self on maailmanlaajuisesti nopeasti kasvava trendi, jossa ihmiset uusien laitteiden ja teknologian avulla keräävät, analysoivat ja hyödyntävät kaikkea mahdollista omaan kehoon, ympäristöön ja toimintaan liittyvää dataa. Älypuhelinsovellusten, puettavien sensorien, verkkopalveluiden ja muiden teknisten apuneuvojen kanssa itsensä mittaajat keräävät dataa, joka liittyy esimerkiksi ruokailuun, mielialoihin, fyysiseen ja henkiseen suorituskykyyn, ajan- ja rahan käyttöön, sosiaaliseen vuorovaikutukseen, liikkumiseen, nukkumiseen jne. Keskeisenä ajatuksena on yksilöllisen mittaustiedon hyötykäyttö terveyden, hyvinvoinnin ja elämänlaadun parantamisessa.

Kyse ei ole vain terveys- ja teknologiaintoilijoiden liikkeestä. Urheilijat ovat jo pitkään mitanneet muun muassa sykettä, tehoa ja omia suorituksiaan. Fysiologisessa ja psykologisessa valmennuksessa ja kuntoutuksessa käytetään mittareita. Tietotyöläiset saattavat mitata ajankäyttöä, stressitasoa ja erilaisen viestinnän ja kommunikaation volyymeja. Askelmittareiden, verkkoyhteydellä varustettujen henkilövaakojen, aktiivisuusrannekkeiden ja muiden vastaavien kulutuselektroniikkatuotteiden sekä niihin liittyvien verkkopalveluiden kehittyminen on tuonut itsensä mittaamisen monille osaksi arkipäivää. Alla olevaan taulukkoon on koottu otos asioista, joita nykyisillä työkaluilla ja sovelluksilla on mahdollista mitata.

Taulukko 4.2: Esimerkkejä datasta, jota itsensä mittaamisen työkaluilla ja sovelluksilla voi kerätä.

Kulutus • Kalorinkulutus ja ravinto • Ravintoaineet • Alkoholinkulutus • Tupakointi • Kahvi • Vesi • Lääkkeet	**Kehon toiminta** • Kehon PH • Kuukautiskierto • Raskaus • Ulostaminen	**Fyysinen aktiivisuus** • Liikuntasuoritukset • Nukkuminen, unen laatu • Matkustaminen • Seksuaalinen aktiivisuus • Hampaiden pesu • Asento • Askelmäärä ja aktiivisuus
Oireet • Päänsäryt • Kivut • Astmakohtaukset • Allergiat	**Tila ja aika** • Sijainti • Korkeus • Aika • Mitä näet	**Fysiologiset suureet** • Syke • Verensokeri • Kehon lämpö • Verenpaine • Paino ja kehon koostumus • Hengitys
Mielen hyvinvointi • Mieliala • Stressi • Vireystila	**Ajankäyttö** • Median kulutus • Tietokoneohjelmien käyttö • Työaika • Sähköpostiviestien määrä • Puhelut ja muu viestintä	**Muuta** • Rahan käyttö • Ostokset • Sosiaaliset verkostot • Oppimistulokset • Liikenne • Energiankulutus

4.1.2 Digitaalinen jalanjälki ja kulutuslaskurit

Kun itsensä mittaamisessa ihminen aktiivisesti kerää tietoa antureiden, sovellusten ja muiden teknisten apuvälineiden avulla, niin digitaalinen jalanjälki syntyy automaattisesti. Jokainen verkkohaku, ostostapahtuma, matkakortin tai kellokortin leimaus tallentuu. My Datan myötä ihminen itse voisi hyödyntää digitaalista jalanjälkidataansa. Esimerkiksi kirjastosta voisi saada oman lainaushistorian datana ja syöttää sen virtuaaliseen lukupiirisovellukseen.

Eräs ihmisille hyödyllinen ja yhteiskunnallisesti merkittävä digitaalisen jalanjälkidatan sovellusalue on kokonaisvaltainen kulutuksen seuranta. Kuluttaminen jakautuu lähes aina useisiin palveluihin. Jos kerätään tietoa vain yhdestä lähteestä, ei saada kuvaa kokonaisuudesta. Esimerkiksi ravitsemuksen terveysvaikutuksia analysoiva sovellus antaa virheellisen kuvan, jos siinä on mukana vain elintarvikekaupan ostodata, mutta ei tietoa työpaikkaruokailusta, ravintolakäynneistä tai ostoksista kulman karkkikaupassa. Tällä hetkellä ravinnon kulutusprofiilin kerääminen onnistuu itse ylläpidettävillä laskureilla, joita saa esimerkiksi älypuhelinsovelluksina. Niiden ongelmana on kirjaamisen vaivalloisuus ja puutteelliset tiedot eri tuotteiden tarkoista ravintosisällöistä.

Toinen esimerkki kokonaisvaltaisesta kulutuslaskurista on henkilökohtaisen liikkumisen kustannuksia seuraava sovellus. Liikkumiskustannukset voivat koostua mm. julkisen liikenteen matkalipuista, auton käyttökustannuksista, auton vuokrasta, taksimatkoista, parkkimaksuista, matkoihin käytetyistä ajasta, kotiinkuljetuksen maksuista jne. Ympäristön kannalta olisi edullisinta, jos ihmiset käyttäisivät vähemmän yksityisautoa, mikä usein voisi olla edullisinta myös ihmisille itselleen. Nykyisin liikkumiseen käytetyistä kustannuksista saatika ympäristövaikutuksista ei helposti saa kokonaiskuvaa tai välitöntä palautetta; ihmisiltä puuttuu tietoa, jonka pohjalta he voisivat muuttaa omaa toimintaansa. Ruotsissa on rohkaisevin tuloksin pilotoitu urbaaneille talouksille suunnattua UbiGo[17] -palvelua, joka näyttää liikkumisen kustannukset sekä mahdollistaa samasta mobiilisovelluksesta helposti mm. auton vuokraamisen, matkalippujen ostamisen ja muiden liikennevaihtoehtojen varaamisen “liikenne palveluna” -periaatteen mukaisesti. Tällaisten sovellusten tekeminen ja skaalaaminen tulisi monin verroin helpommaksi, jos ihmisten liikkumisen ja rahan käytön jalanjälkidataa olisi My Datana.

Palaute on edellytys oppimiselle ja oppien jalostamiseksi tavoiksi. Kalorinkulutuksen ja liikkumisen kustannusten lisäksi ihmiset voivat olla kiinnostuneita seuraamaan yleisesti oman kulutuskäyttäytymisensä taloudellisuutta, terveellisyyttä, ekologisuutta tai eettisyyttä. Kokonaisvaltaisen kulutusmittarin avulla voi löytää ne osa-alueet, joita muuttamalla voi helpoimmin saavuttaa halutun lopputuloksen. Esimerkiksi mistä kannattaa säästää, jos talous on tiukalla tai millä kulutusmuutoksilla on todellisuudessa suurin merkitys omaan hiilijalanjälkeeni? Markkinoilla on nykyisin lukuisia mobiilisovelluksia, jotka auttavat tekemään tietoisia kulutuspäätöksiä tarjoamalla tuotteista lisätietoa viivakoodin skannauksella. My Datan ja laajan jalanjälkidatan hyödyntäminen tekisi tällaisten sovellusten käytöstä helppoa, kun voitaisiin automaattisesti analysoida jo tehdyt ostokset.

17 http://web.viktoria.se/ubigo/las-mer/about-english/

4.2 Paremmin kohdennetut tuotteet ja markkinointi

Asiakkaiden profilointi ja paremmin kohdennettujen tuotteiden ja markkinoinnin kehittäminen motivoivat yrityksiä. Verkkopalveluiden suosittelujärjestelmät ovat tehokkaita markkinoinnin välineitä, ja verkkomainonta perustuukin merkittävältä osin ihmisten profilointiin. Yleisiä tunnettuja esimerkkejä ovat esimerkiksi Amazonin kirjasuositukset, Googlen ja Facebookin mainokset, Netflixin elokuvasuositukset sekä Grouponin henkilökohtaiset tarjoukset. Suosittelussa on kaksi keskeistä ongelmaa. Ensinnäkin käyttäjä voi kokea epämiellyttävänä ja epäluottamusta herättävänä sen, ettei hän tiedä, mihin suositus pohjautuu. Toisaalta monesti suositukset ovat aika heikkoja, koska suosittelujärjestelmällä on rajallisesti tietoa ihmisen toiminnasta; yleensä tieto on kerätty vain kyseisen palvelun sisältä eikä ihminen itse ole määrittänyt preferenssejään.

Profiloinnin ja suosittelun käytännöt voivat muuttua merkittävästi My Datan myötä. Ihmiset voisivat itse hallita omaa profiilitietoaan ja kerryttää sitä useasta lähteestä. Yritykset, jotka tarjoaisivat asiakkailleen heidän My Datansa, voivat rakentaa vahvaa luottamussuhdetta. Vaihtokaupassa yritys voisi saada asiakkaalta rikasta profiilitietoa, jota tämä on kerryttänyt muiden yritysten ja palveluiden käyttäjänä ja itse määrittämällä. Kun käyttäjä itse hallitsee tietoa, jonka pohjalta suosituksia tehdään, tulee suosittelusta vuorovaikutteista ja läpinäkyvää nykyisen tungettelevan verkkomainonnan sijaan.

Markkinoiden avoimuuden kannalta olennainen muutos olisi, että pienetkin toimijat, joiden ei muutoin olisi käytännössä mahdollista kerätä profilointitietoja asiakkaistaan, voisivat tarjota ihmisten itse hallinnoimien profiilien pohjalta kehittynyttä datapohjaista suosittelua. Parhaimmillaan kysyntä ja tarjonta voisivat kohdata pienilläkin markkinoilla.

4.2.1 Profilointi ja suosittelu nykyisin

Nykyisin yritykset pääsevät vaivattomasti käsiksi vain tietoon oman yrityksensä transaktioista, jolloin ne eivät voi tuottaa kovinkaan yksityiskohtaista profiilia asiakkaastaan. Joillakin kauppaketjuilla on suhteellisen laaja palveluvalikoima, mutta silti niiden kontakti ihmiseen on rajallinen suhteessa ihmisen kokonaistoimintaan.

Profilointia voi täydentää ostamalla lisädataa esim. verkkoseurantaa tarjoavilta yrityksiltä. Tällä hetkellä yritykset käyttävätkin rahaa, vaivaa ja joskus myös huonosti päivänvaloa kestäviä keinoja potentiaalisten asiakkaiden profilointiin. Vaikka dataa olisi paljonkin, niin profilointi on pinnallista, koska se perustuu ihmisestä hänen tietämättään kerättyyn tietoon; jos ostan vaippoja, niin ennustetaan, että haluan ostaa myös vauvanruokaa. Tälläistä profilointia kutsutaan implisiittiseksi profiloinniksi eksplisiittisen, käyttäjän itse määrittämän profiloinnin sijaan. Panostus asiakkaiden implisiittiseen profilointiin on viime aikoina ollut suurta ja se on keskeinen big datan kehitystä eteenpäin vievä voima.

Joissain tapauksissa implisiittinen profilointi toimii. Esimerkiksi Yhdysvalloissa kauppaketju on ostos- ja osoitetietoa yhdistelemällä päätellyt teinitytön olevan raskaana ennen kuin hänen vanhempansa tiesivät siitä (Hill 2012). Onko kuitenkaan yhteiskunnallisesti tai yksilön kannalta hyvä, että yritykset tuottavat profilointeja meistä tietämättämme ja sen perusteella vaikuttavat kulutuskäyttäytymiseemme? Jos haluan muuttaa kulutustottumuksiani, niin kuinka viestin tästä organisaatioille muuttaakseni historiallisen profiilini tuottaman lokeroinnin?

Profiloinnilla on muitakin käyttötarkoituksia kuin verkkomainonta. Esimerkiksi Yhdysvalloissa jotkin tietokantayritykset keräävät verkosta ja erilaisista tietolähteistä ihmisten henkilötietoa ja jalostavat siitä automaattisesti potentiaalisten työntekijöiden tietokantoja yritysten rekrytoinnin tueksi. Mikäli tällaiseen tietokantaan päätyy yksilöstä väärää tietoa ja päätelmiä, ne saattavat haitata hänen työnsaantimahdollisuuksiaan. Tätä mekanismia kutsutaan algoritmivankilaksi (Davidov 2014). Kun entistä enemmän taloudellisia ja ihmisten elämään vaikuttavia päätöksiä tehdään automaattisesti datan ja algoritmien avulla, riskinä on, että syrjivät käytännöt yleistyvät esimerkiksi työmarkkinoilla, lainojen ja vakuutusten antamisessa tai vuokra-asuntomarkkinoilla (White House 2014).

4.2.2 Ihmisten itse hallitsema profiilitieto

Profiloinnilla yritykset pyrkivät tavoittamaan oikeat kohderyhmät ja tarjoamaan niille kiinnostavia tuotteita ja palveluita. Pääsy laajaan käyttäjätietoon mahdollistaa merkittäviä palveluinnovaatioita. Samaan päästään myös, jos ihmiset itse jakavat haluamansa tiedot itsestään ja kertovat, minkälaisista tuotteista tai palveluista he ovat kiinnostuneita ja millä ehdoilla. Etuna yritysten tekemään profilointiin nähden olisi, että ihmiset voisivat kerryttää omaan profiiliinsa huomattavasti rikkaampaa dataa kaikista My Datan lähteistään ja heillä on motiivi "hoivata" itse hallittuja ja monikäyttöisiä profiilejaan.

Kuva 4.3: Käyttäjän suostumuksella yritykset voivat saada käyttäjästä rikasta profiilitietoa, jonka vastineeksi yritys pystyy tuottamaan käyttäjälle parempaa palvelua ja palveluun liittyvää viestintää.

Henkilökohtainen, rikas profiilitieto voi olla vaihdannan väline siten, että kuluttaja antamalla palveluntarjoajalle pääsyn osaan henkilötiedostaan saa vastineeksi paremmin tarpeitaan vastaavaa palvelua. Joissain tapauksissa profiilitietojen luovuttaja voisi saada suoraan rahaa tililleen tai alennusta tuotteen hinnasta. Esimerkiksi: *"Saat autovakuutuksen hinnasta pois 10 %, jos annat meille dataa, joka kuvaa ajokäyttäytymistäsi ja osoittaa sen turvalliseksi."* Vastaavasti luovuttamalla yksilölle häntä koskevan datan yritys voisi saada vaihtokaupassa pääsyn asiakkaan laajempaan profiilitietoon.

My Datan myötä ihmisen oma kontrolli ja datan käytön läpinäkyvyys lisään-

tyisivät ja ihmisille alkaisi kehittyä selkeä kuva henkilötietojensa arvosta. Riskinä on, että mikäli markkinat eivät ole tasapainoiset, voi syntyä tilanteita, joissa yritykset vaativat yksilöiltä dataa. Esimerkki autovakuutuksesta saattaakin kääntyä muotoon: *"Kuluttajat, jotka eivät luovuta ajokäyttäytymisdataa luokitellaan riskikuljettajiksi ja he joutuvat maksamaan vakuutuksesta 20 % korkeamman hinnan."*

Monikäyttöisten ja siirrettävien profiilien helppo luominen ja hyödyntäminen voisi olla osin jopa My Data infrastruktuurin osa-alue. Mikäli profiileille luotaisiin yhtenevät toimintatavat voisi profiileihin sisältyä myös vaihdettavuuteen ja anonymiteettiin liittyviä ominaisuuksia. Yrityksille kynnyskysymys on, miten asiakkaiden itsensä hallinnoimat profiilit vaikuttavat myyntiin ja asiakashankinnan kustannuksiin. Jos asiakkaan tarkoitushakuisesti määrittämän profiilin perusteella syntyy kauppa, niin yritykselle se ei ole ongelma; hyvän kaupankäynnin lopputuloksena sekä myyjä että asiakas ovat tyytyväisiä. Tällöin profiilia voi pitää yksilön tapana kommunikoida omia tarpeitaan kaupalle. Voidaan olettaa, että asiakkaiden eri yrityksille suuntaamat profiilit loppujen lopuksi tukevat kaupankäyntiä ja pitkäaikaista asiakassuhdetta. Jos asiakas antaa väärää tai huonoa tietoa itsestään, hän saa myös huonosti itselleen sopivia tarjouksia, mikä ei palvele kumpaakaan osapuolta.

4.2.3 Luottamus ja läpinäkyvät suositukset

Datan käsittelyn läpinäkyvyys lisää ihmisten luottamusta dataa keräävään yritykseen. Dataa ei tarvitse kerätä vaivihkaa, vaan ihmiset luovuttavat tietoa, jos tietävät, mitä yritys datalla tekee, ja näkevät sen hyödylliseksi itselleen tai muuten arvokkaaksi. Käyttäjällä pitäisi kuitenkin olla aito oikeus myös kieltäytyä käyttäjätiedon luovuttamisesta. Hyvä esimerkki on TomTom-autonavigaattori. Sen käyttäjät voivat itse päättää, lähettävätkö tietoa TomTomille siitä, missä ja millä nopeudella liikkuvat. Tiedon perusteella navigaattori pystyy päättelemään ruuhkakohdat ja suositella parempia reittejä. Jos ihmiset luottavat TomTomiin, he osallistuvat liikennetiedon kartuttamiseen ja lisäävät järjestelmän arvoa kaikille käyttäjille.

Luottamuksen saavuttaminen vaikuttaa siihen, miten yritykset toteuttavat liiketoimintaansa. Asiakkuutta ei kannatakaan ajatella yksittäisenä ostotapahtumana vaan elinkaarena. Esimerkiksi verkkokauppaan palaava asiakas on yleensä arvokkaampi kuin ensimmäistä kertaa asioiva. Luottamuksen rakentamisessa My Datan tarjoaminen asiakkaille on suora tie hyvään dialogiin.

My Data -periaatteiden mukaisesti toimiva verkkokauppa saattaisi kysellä asiakkaalta hänen mieltymyksiään sekä kerätä automaattisesti tietoa hänen klikkailuistaan, kuten nykyisetkin verkkokaupat. Olennainen ero olisi, että verkkokauppa palauttaisi kertyneen informaation ja siitä tehdyn tulkinnan läpinäkyvästi asiakkaalle. Salakuuntelusta tulisikin asiakkaan kuuntelua ja datan keräämisestä hyväksyttävää tai jopa toivottavaa – ei räätälissäkään kukaan häiriinny siitä, että mittatilauspukua varten tehdään tarkkoja mittauksia.

Toimittajatiedon hallinta (Vendor Relationship Management VRM)

Asiakastiedon hallinta on keskeinen osa yritystoimintaa. Yritysasiakkaiden tiedot ovat yleensä jollain tavoin hallittavissa ja asiakastietoa voi ostaa muun muassa hakemistoyrityksistä. Mitä pienemmistä ja monilukuisemmista asiakkaista asiakaskunta koostuu, sitä haastavampaa on asiakastiedon hallinta. Kuluttajakaupassa ajantasaisen asiakasrekisterin ylläpito lähestyy jo mahdottomuutta. Monet yritykset haluaisivat ehkä päästä eroon oman asiakasrekisterin jatkuvasta ylläpidosta. Niille riittäisi, että asiakaskontaktin syntyessä ne saisivat ajantasaiset tiedot suoraan sähköisesti asiakkaalta omaan järjestelmäänsä tai jopa niin, ettei asiakastietoa edes tallennettaisi, sitä vain käytettäisiin sillä hetkellä, kun sitä tarvitaan. My Data mahdollistaisi tämän, koska asiakkailla voisi olla oma My Data -profiili.

My Dataan läheisesti liittyvä ajatus on niin sanottu toimittajatiedon hallinta (Vendor Relationship Management VRM). Sen lisäksi, että kuluttajasta on tietoja useiden yritysten asiakastietojärjestelmissä (Customer relationship management CRM), niin kuluttajalla itsellään voisi olla järjestelmä, johon hän tallentaa tiedot yrityksistä, joiden asiakas hän on – aina autokorjaamoista parturi-kampaamoihin. Tällainen henkilökohtainen toimittajarekisteri voisi sisältää sopimuksia, takuukuitteja ja historian yhteydenpidosta kunkin yrityksen kanssa. Sen avulla, kun ihminen esimerkiksi muuttaa asunnosta toiseen, hän voisi jakaa automaattisesti uudet yhteystiedot kerralla kaikille niille yrityksille, joiden kanssa haluaa jatkaa yhteydenpitoa. Asiakkaiden hallitsema VRM ja yrityksen CRM -järjestelmä voisivat täydentää toisiaan ja vaihtaa tietoa tarpeen mukaan.

Omassa hallinnassa olevan asiakasprofiilin ja toimittajienhallintajärjestelmän avulla kuluttaja-asiakkaatkin voivat nykyistä helpommin kilpailuttaa yrityksiä. Sen sijaan, että asiakas käyttää aikaa parhaiden tarjousten metsästämiseen, hän voikin tehdä julkiseen profiiliinsa liitetyn ostotarjouksen. Nykyisin asiakkaan ostotarjouksiin pohjautuvia järjestelmiä on joillain yksittäisillä toimialoilla. Esimerkiksi Tilausajot[18] -palvelussa asiakas täyttää tiedot tarvitsemansa kyydin ajankohdasta ja reitistä, ja saa sen jälkeen eri kuljetusyrityksiltä tarjouksia sähköpostiinsa.

18 http://tilausajot.net

4.3 My Datan design-huomioita

Yllä esitetyt kaksi My Datan hyödyntämisaluetta – itsestään oppiminen ja palveluiden ja viestinnän parempi kohdentaminen, ovat vain kaksi nostoa siitä miten My Dataa voidaan hyödyntää. My Dataa voidaan hyödyntää laajamittaisesti esimerkiksi tutkimuksessa, resurssien käytön ja tuotannonohjauksen optimoinnissa ja tietotyössä. Seuraavaksi esitellään yksilöiden monimuotoisuuteen ja sosiaalisuuteen liittyviä design-huomioita, jotka koskettavat useita My Datan hyödyntämisalueita.

4.3.1 Palvelumuotoilu ja itse tekeminen

Yksi My Datan onnistumisen edellytys on, että sovellukset ja infrastruktuuripalvelut, kuten dataoperaattorit ja datapankit, ovat ihmisille ymmärrettäviä ja helppokäyttöisiä. Suurinta osaa ihmisiä ei kiinnosta data itsessään, vaikka sitä olisi saatavilla, eivätkä he jaksa nähdä vaivaa oman datansa hallinnointiin. Todennäköisesti vain harvat jaksaisivat koota yhteen tietoja useasta lähteestä, vaikka palveluissa olisi avoimet My Data -rajapinnat. Tarvitaan loppuun asti tuotteistettuja, sujuvia ja helppokäyttöisiä palveluja yksinkertaistamaan monimutkaista maailmaa helpommin hahmotettavaksi ja hallittavaksi.

Käytettävyyden ja palvelumuotoilun merkitystä ei voi aliarvioida. Samalla on kuitenkin vältettävä aliarvioimasta ihmisiä itseään ja heidän uteliaisuuttaan ja kyvykkyyttään. Jos kaikki on valmiiksi pureskeltu, niin ihmisten oma ymmärrys ja voimaantuminen saattaa jäädä puolitiehen. Esimerkiksi Quantified Self -liike on selvä signaali siitä, että on olemassa paljon ihmisiä, jotka ovat valmiita näkemään vaivaa ja haluavat itse tehdä asioita omalla datallaan.

Hyvä palvelumuotoilu jättää tilaa itse tekemiselle. My Data -palveluja suunniteltaessa on hyvä pyrkiä selkeyteen peruskäyttäjille, mutta samalla jättää mahdollisuus itse tekemiseen, kuten omien analyysien ja metriikoiden luomiseen niille, joita se kiinnostaa. Vaikka kokeilijoiden ja tekijöiden marginaaliryhmä ei ole suuri, se on kuitenkin tärkeä uuden synnyttäjänä. Vaikka tänään kaupan hyllyltä löytyy hyvin tuotteistettuna se, minkä tekeminen vielä vähän aikaa sitten vaati vaivannäköä, niin tänään kokeilijat ovat jo tekemässä sitä, mikä ehkä tulevaisuudessa on valtavirtaa. Kokeiluista voi nousta suurta joukkoa palvelevia tuotteita ja erityisryhmille palveluja, joita muut eivät tee.

Jotta voidaan mahdollistaa helppokäyttöisyys ja toisaalta syvällinen käyttö on tärkeää ymmärtää, että tieto, metatieto, ja sopimuksien tieto olisi pidettävä erillisenä palvelulogiikasta ja käyttöliittymistä. Tämä mahdollistaa vaihtoehtoiset käyttöliittymät. Tällainen suunnitteluparadigma on yleinen ja hyvin tunnettu tietojärjestelmien puolella, mutta se ei monesti toteudu käytännössä yksilön kannalta palveluissa.

4.3.2 My Datan sosiaaliset näkökulmat

Kun puhutaan palveluista, joista ihmiset aidosti voivat valita, haluavatko niitä vai eivät, suunnittelun on lähdettävä ihmisistä ja heidän tarpeistaan. Yhteisöllisyyden ja yhdessä tekemisen merkitys ihmisille on perustarve, joka saattaa helposti unohtua, kun puhutaan vain yksilöistä ja heidän My Datastaan. Henkilötiedon jakaminen voi olla keskeistä erilaisissa vertaisuuteen pohjautuvissa tukipalveluissa, kuten valmennusryhmien toiminnassa ja yleisesti tavassa miten ystäväpiirit ja yhteisöt viestivät toisilleen.

Nykyisin markkinoilla on paljon yhteisöllisiä sovelluksia, joihin ihmiset osallistuvat omalla datallaan. Esimerkiksi suosittu Waze-mobiilipalvelu[19] mahdollis-

19 https://www.waze.com/

taa sen, että oman sijaintitiedon voi jakaa reaaliaikaisesti tutuille, ja koordinoida vaikkapa kaikkien saapumisajankohdat määränpäähän. Myös monille muille urheiluun, terveyteen ja itsensä mittaamiseen liittyville palveluille (kuten esimerkiksi Patienslikeme[20] ja Runkeeper[21]) keskeistä on tiedon jakaminen ja vertailu.

Nykysovelluksissa data on kuitenkin yleensä lukittuna eikä sitä voida viedä mukanaan uuteen yhteisöön. Jakamisen oikeudet palvelun sisällä sosiaalisissa ryhmissä on osa yksityisyysasetuksia, joita pitäisi pystyä My data -periaatteiden mukaisesti hallitsemaan keskitetysti ja yhtenevien käytäntöjen mukaisesti.

My Dataan liittyy myös huoltajuuteen ja asioiden hoitamiseen liittyviä sosiaalisia näkökulmia esimerkiksi toiminnassaan tukea kaipaavien vanhusten tai lasten osalta. Henkilötieto voi olla tarpeen, jotta huoltaja tai asioiden hoitaja voi auttaa tai suojella holhottavaa, mutta toisaalta holhottavalla on oltava mahdollisuus myös yksityisyyteen. Näihin tapauksiin on tärkeä selvittää ja luoda hyviä käytäntöjä.

Meidän data vai yritysten data?

Yksinkertaisimmin My Datan käsitteen ymmärretään koskevan henkilöistä kerättävää dataa. Myös lainsäädännöllisestä näkökulmasta My Data -ajattelu rajautuu helpoimmin luonnollisiin henkilöihin. Henkilötietolainsäädännössä käytetään termiä rekisteröity (englanniksi data subject), joka tarkoittaa nimenomaan luonnollista henkilöä, josta tietoa on kerätty, vaikka sen arkikielisesti voisi ymmärtää tarkoittavan myös muunlaisia datasubjekteja, kuten yrityksiä.

Yksilön data on kuitenkin usein hankalasti määriteltävää. Esimerkiksi kodin sähkönkulutusmittarin tai autoon asennetun anturin data koskee kaikkia kodin asukkaita tai auton käyttäjiä. Ostoksia tehdään usein koko perheelle, vaikka maksaja onkin yksi ihminen. Myöskään puhelinliittymän laskujen maksaja ei aina ole sama henkilö, joka puhelinta todellisuudessa käyttää. Pitäisikö osin puhua myös ‘Our Datasta’?

My Datan mukaisia hyötynäkökulmia ajatellen myös organisaatiot, kuten yritykset, voisivat hyötyä, mikäli saisivat niitä koskevan datan helposti käytettävässä muodossa. Esimerkiksi taloyhtiöiden on nykyisin hankalaa vaihtaa isännöitsijää, koska isännöitsijäyrityksen järjestelmissä on kaikki taloyhtiön data (Ympäristöministeriö 2014). Myös julkishallinnon kanssa asioidessaan, kuten veroilmoitusta tehdessään, yritykset kerryttävät yritysten “omaa” dataa, jolla voisi olla paljonkin hyödyllisiä jatkokäyttömahdollisuuksia.

Selkeyden vuoksi tässä selvityksessä termillä My Data viitataan ensisijaisesti luonnollisten henkilöiden dataan. Samoja My Data -periaatteita voidaan kuitenkin soveltaa myös useamman ihmisen yhteiseen jaettuun dataan ja organisaatioiden dataan.

20 http://www.patientslikeme.com/
21 http://runkeeper.com/

5. Toimintaympäristö ja askelmerkit

Edellisissä luvuissa kuvatut tulevaisuuden visiot My Data -periaatteista, -rajapinnoista, -sovelluksista ja -palveluinfrastruktuurista saattavat tuntua kaukaisilta ja saada kuulijan lausumaan *"kyllä, mutta…"*. Visioon pääseminen edellyttää uudenlaisia toimintatapoja ja asenteita sekä jatkuvasti uusiutuvaa, mahdollistavaa ja kokeilevaa politiikka. Onko tällaiseen muutokseen halukkuutta yrityksissä ja muissa organisaatioissa, entä ihmisillä itsellään – mikä motivoi ja mikä estää?

Tässä luvussa kuvataan My Data -kehityksen lähtötilannetta lainsäädännössä, sääntelyssä ja toimijakentässä. Lainsäädännöllä ja muulla sääntelyllä voi olla kiihdyttävä tai hidastava vaikutus, mutta lainsäädäntö yksin ei saa aikaan muutosta. Sääntelyn haasteena on erityisesti liiketoiminnan kansainvälistyminen. Vaikka sääntelyä olisi, se ei aina sovellu globaaleihin palveluihin. Kehitystä on tehtävä monella eri alueella. Tarvitaan yhteiskunnallisesti tuettua kehitystä, monien toimijoiden kykyä nähdän yhteistoiminnan etuja, teollisesti merkittäviä ja investointikyvykkäitä toimijoita, jotka näkevät My Datan kehittämisen mielekkäänä sekä ennen kaikkea yksilöitä, jotka ovat kiinnostuneita ja valmiita toimimaan aktiivisesti sekä oman tiedon hyödyntäjinä että omien oikeuksien puolustajina. My Data joko kehittyy tai on kehittymättä sen mukaan, miten eri toimijat tarttuvat ajatukseen.

Luonnollisesti suhtautuminen My Dataan ei ole, eikä sen pidäkään olla, varauksettoman positiivista. *Esteitä My Datan edessä* -alalukuun on koottu muutamia yleisesti esille nousseita huolia ja My Dataa vastustavia näkökulmia. Kaikkea ei tarvitse ratkaista kerralla eikä jokaiseen kysymykseen voi vastata tyhjentävästi etukäteen. Varmasti matkan varrella nousee uusia kysymyksiä ja näkemys My Datan toteutumisesta muuttuu.

5.1 Lainsäädäntö ja sääntely

Selvityksessä on käytetty henkilötieto-sanaa hyvin laajassa ja väljässä merkityksessä, kuten johdannossa esiteltiin. Tietosuojaan ja yksityisyydensuojaan liittyvän lainsäädännön kulmakivenä on kuitenkin tarkempi määrittely. Esimerkiksi EU:n tietosuojaryhmä on antanut 26-sivuisen lausunnon *henkilötiedon* käsitteestä (EU 2007). Henkilötietolainsäädäntö koskee vain juridisen määritelmän mukaisten *henkilötietojen* käsittelyä – ei muiden tietojen. Kun tässä luvussa tarkoitetaan nimenomaisesti lainsäädännöllisen määrittelyn mukaista *henkilötietoa*, se on kirjoitettu kursiivilla. Suuri osa kaikesta henkilöitä koskevasta tiedosta, joka voi olla muutettavissa My Dataksi, mahtuu myös lainsäädännön määritelmän sisälle, mutta ei välttämättä kaikki.

Henkilötietoon liittyvän sääntelyn ja lainsäädännön kehittämisen haasteena on tasapainottaa datan hyödyntämiseen ja tietosuojaan ja yksityisyyden suojaan liittyvät toimet. Kaikki henkilötiedon kerääminen ja hyödyntäminen saattaa heikentää yksityisyydensuojaa, mikäli lakeja ja hyviä käytäntöjä ei noudateta. Toisaalta henkilötiedon keräämisellä ja käytöllä voidaan myös parantaa yksilöiden elämänlaatua ja jopa pelastaa ihmishenkiä esimerkiksi ensiaputyössä.

Nykyinen lainsäädäntö antaa ihmisille oikeuden tarkistaa omat tietonsa henkilörekisterin pitäjältä ja pyytää korjaamaan tai tietyissä tilanteissa myös poistamaan tiedot. EU:n perusoikeuskirjaan on kirjattu, että *"jokaisella on oikeus tutustua niihin tietoihin, joita hänestä on kerätty, ja saada ne oikaistuksi"*. Käytännössä rekisterinpitäjä pyydettäessä usein lähettää paperitulosteena tai PDF-tiedostona otteen, josta tiedot käyvät ilmi.

Lainsäädäntö lähtee siitä, että tiedot on annettava ”ymmärrättävässä muodossa”, mutta toistaiseksi lainsäädäntö ei anna taustatukea, mikäli haluaisi saada datan itselleen uudelleenkäytettävässä muodossa. Kansalaisten vaatimus omien tietojen saamiseksi nykyistä monikäyttöisemmässä muodossa on voimistunut. Esimerkiksi sosiaalisen median palveluissa oleva tieto halutaan siirtää palvelusta toiseen tai vain omaan käyttöön. My Datan toteutumisen kannalta olisi toivottavaa, että yrityksiä ja organisaatioita kannustetaan ja ohjataan avaamaan henkilötietorajapintoja yksilöille itselleen.

Nykyisen lainsäädännön lähtökohtana on yksityisyyden suojaaminen ja rekisterin tietosisällön oikeellisuus, ei niinkään *henkilötiedon* potentiaalisten hyötykäyttöjen realisointi. My Data -ajattelu pyrkii tuomaan hyödyntämisnäkökulman tasaveroisena yksityisyyden suojan rinnalle. Tästä näkökulmasta nykylainsäädäntö on riittämätöntä.

5.1.1 Henkilötietolaki

Lainsäädännössä henkilötiedon määritelmä kattaa kaiken sellaisen tiedon, mikä voidaan suoraan tai eri tietoja yhdistelemällä liittää yksilöön, ja näin ollen tietoa voidaan käyttää kuvaamaan yksilöä: *”kaikenlaisia luonnollista henkilöä taikka hänen ominaisuuksiaan tai elinolosuhteitaan kuvaavia merkintöjä, jotka voidaan tunnistaa häntä tai hänen perhettään tai hänen kanssaan yhteisessä taloudessa eläviä koskeviksi”* (HLT 1999 3 §:n 1 kohta). Olennaista on, voidaanko henkilö tunnistaa arvioitavana olevan tiedon perusteella vai ei. EU:n henkilötietodirektiivin (EU 1995) mukaan *”sen määrittämiseksi, onko henkilö tunnistettavissa, olisi otettava huomioon kaikki kohtuullisesti toteutettavissa olevat keinot, joita joko rekisterinpitäjä tai joku muu voi kyseisen henkilön tunnistamiseksi käyttää”*.

Tätä määritelmää on pidetty lukittuna ja sääntely on tapahtunut joko - tai -lähestymistavalla: henkilön tunnistamisen mahdollistavien *henkilötietojen* käyttöä ja käsittelyä on rajoitettu henkilötietolainsäädännöllä ja muiden kuin henkilötietojen käyttö on vapaampaa. Henkilötietolaissa (HLT 1999 8 §) luetellaan yhdeksässä kohdassa ne edellytykset, jolloin *henkilötietoja* saa käsitellä (käsittely on sallittu ainoastaan, jos jokin näistä edellytyksistä toteutuu). Yleisesti käytettyjä ovat asiakassuhteeseen perustuva käsittely ja rekisteröidyn yksiselitteisesti antama suostumus, mutta mukana listalla ovat myös mahdollisuudet henkilötietojen käsittelyyn tietosuojalautakunnan luvalla sekä käsittely yksittäistapauksissa rekisteröidyn elintärkeän edun suojaamiseksi.

Henkilötietoa saa lain mukaan käyttää lähtökohtaisesti vain siihen tarkoitukseen, mihin se on kerätty. Yksityisyyden suojan kannalta on perusteltua, etteivät organisaatiot voi noin vain, varsinkaan ihmisten tietämättä, alkaa käyttää henkilötietoa uusiin tarkoituksiin. Toisaalta taas datan järkevät hyödyntämistarkoitukset saattavat nousta esiin vasta jälkikäteen. Suostumuksen pyytäminen jälkikäteen on usein käytännössä vaikeaa, vaikka uusi käyttötarkoitus olisikin rekisteröityjen mielestä hyväksyttävä. Lain hengen ja tarkoituksen mukaisesti henkilötietojen käyttötarkoituksen tulisi aina olla mahdollisimman tarkasti määritelty (EU 2013). Todellisuudessa kuitenkin käyttötarkoituksen kuvaava rekisteriseloste monesti kirjoitetaan mahdollisimman sallivaksi. Laveaan muotoon kirjoitetuista rekisteriselosteista rekisteröidyn on vaikea saada selkoa siitä, mitä kaikkea käytännössä tiedoille voi tehdä.

Pyrkimyksenä lainsäädännössä on löytää tasapaino jatkokäytön ja alkuperäisen käytön välillä. Esimerkkejä sallituista jatkokäyttötarkoituksista ovat tieteellinen tutkimus ja tilastointi, missä tieto on alunperin saatettu kerätä muuta tarkoitusta varten.

My Data -lähestymistavassa tasapainoa haetaan sitä kautta, että ihminen toimii ikään kuin *”kontrollipisteenä”* oman datansa käytön suhteen. Silloin kaikkia mahdollisia jatkokäyttötarpeita ei tarvitse etukäteen luetella, sillä ihminen

itse saa käsitellä omaa dataansa vapaasti. Nykyisen luettelon jatkoksi riittää yksi kohta: ihminen saa oman datansa itselleen, jolloin hän voi antaa sen eteenpäin niihin tarkoituksiin, joihin itse haluaa. Tässä lähestymistavassa luotetaan yksilöön eikä pyritä suojelemaan häntä etukäteen.

5.1.2 EU:n uusi yleinen tietosuoja-asetus

EU:ssa on parhaillaan valmisteilla uusi tietosuoja-asetus, josta tulee kaikissa jäsenvaltioissa suoraan sovellettavaa lainsäädäntöä. On mahdollista, että asetus antaisi kansalaisille nykyistä vahvemman oikeuden saada omat tiedot käyttöönsä koneluettavassa muodossa ja siirtää ne niin halutessaan helposti palvelusta toiseen. Alkuperäisessä Euroopan komission ehdotuksessa henkilötiedon siirrettävyyttä käsiteltiin artiklassa 18 Oikeus datan siirrettävyyteen.

Alkuperäinen Euroopan komission ehdotus

18 artikla – Oikeus siirtää tiedot järjestelmästä toiseen

1. Kun henkilötietoja käsitellään sähköisin keinoin, jäsennellyssä ja yleisesti käytetyssä muodossa, rekisteröidyllä on oikeus saada rekisterinpitäjältä jäljennös käsiteltävistä tiedoista yleisesti käytetyssä sähköisessä ja jäsennellyssä muodossa, joka antaa rekisteröidylle mahdollisuuden käyttää tietoja edelleen.

2. Kun rekisteröity on itse antanut henkilötiedot ja niiden käsittely perustuu suostumukseen tai sopimukseen, rekisteröidyllä on oikeus siirtää kyseiset henkilötiedot ja kaikki muut rekisteröidyn itsensä antamat, automatisoituun käsittelyjärjestelmään tallennetut tiedot toiseen järjestelmään, yleisesti käytetyssä sähköisessä muodossa, sen rekisterinpitäjän estämättä, jonka hallusta henkilötiedot poistetaan.

My Datan toteutumisen kannalta alkuperäinen ehdotus olisi ollut perusteltu. Sellaisenaan se ei kuitenkaan tule toteutumaan, ja se onkin jo parlamentin käsittelyssä muuttanut muotoaan (ks. infoboksi nykytilanteesta). Elinkeinoelämä lobbaa asetuksen tiukkuutta vastaan vetoamalla muun muassa yritysten kasvavaan hallinnolliseen taakkaan. Lähes kaikki yritykset ja yhteisöt ovat nykyisin tavalla tai toisella henkilörekisterin pitäjiä, ja tiukkenevat säännökset edellyttäisivät investointeja tietojärjestelmiin.

Asetusluonnoksia käsittelevät sekä jäsenmaiden hallitusten edustajat ministerineuvostossa että poliitikot avustajineen Euroopan parlamentissa. Tätä kirjoitettaessa ei ole varmuutta siitä, miten hyvin tuleva asetus tukee My Datan toteutumista. Alla oleva on selvitystä tehtäessä muodostunut näkemys siitä, mikä olisi toivottava lopputulema EU:n tietosuoja-asetuksen muotoilulle siten, että My Datan kannalta oleellinen oikeus koneluettavaan dataan olisi mukana, mutta kaikkia pienikokoisia rekisterinpitäjiä ei kuritettaisi uusilla velvollisuuksilla.

Järkevästi rajaamalla rekisterinpitäjän velvollisuuksia My Datan hyödyt olisi toteutettavissa ilman pelättyjä haittavaikutuksia. Rajaus voisi tapahtua seuraavilla tavoilla:

- **aikarajaus** – ei koske ennen asetuksen voimaantuloa kerättyä aineistoa
- **kokorajaus** – koskisi vain tiettyä henkilömäärää suurempia rekistereitä ja pienet rekisterit jäisivät vain tarkastusoikeuden piiriin
- **tarkkuusrajaus** – koskisi vain tiettyä sopivaa tarkkuutta esim. sijaintitiedoissa datamäärien hallinnan järkevöittämiseksi
- **digitaalisuusrajaus** – mitään paperista ei tarvitse digitoida ja mitään, mikä ei ole rakenteellisessa muodossa rekisterissä, ei erikseen tarvitse muuttaa rakenteelliseen muotoon

Rekistereiden yhteentoimivuutta (vaatimus siitä, että kaikki data olisi jossakin tietyssä muodossa) ei kannata lainsäädännön tasolla pyrkiä ohjaamaan, koska tämä olisi varma keino lisätä lakivelvoitteiden noudattamisen kustannuksia, eikä hyödyistä voisi mennä takuuseen. Yhtenevät dataformaatit ja sisällöllinen tai semanttinen yhteentoimivuus kehittynee markkinoilla ensin toimialojen sisällä ja sitten niiden välillä. Olennaista olisi, että ihmisillä on oikeus saada heitä koskeva data rakenteellisessa koneluettavassa muodossa. Soveltaen voitaisiin käyttää esimerkiksi julkishallinnon tietojen uudelleenkäyttöä käsittelevän direktiivin (Public Sector Information Directive PSI) kuvailua koneluettavuudesta.

PSI direktiivin johdanto – kappale 21 (EU 2013 B)

Asiakirjan olisi katsottava olevan koneellisesti luettavassa esitysmuodossa, jos se on sellaisessa tiedostomuodossa, jonka rakenne mahdollistaa sen, että ohjelmistot pystyvät helposti yksilöimään, tunnistamaan ja poimimaan siitä tiettyjä tietoja. Koneellisesti luettavassa muodossa oleviin tiedostoihin koodatut tiedot ovat koneellisesti luettavia tietoja. Koneellisesti luettavissa olevat esitysmuodot voivat olla avoimia tai yksityisiä; ne voivat olla virallisia standardeja, mutta se ei ole välttämätöntä. Asiakirjojen, jotka on koodattu sellaiseen tiedostomuotoon, joka rajoittaa automaattista käsittelyä siksi, että tietoja ei saada poimittua niistä lainkaan tai ei saada poimittua helposti, ei olisi katsottava olevan koneellisesti luettavassa esitysmuodossa. Jäsenvaltioiden olisi tarvittaessa kannustettava avointen koneellisesti luettavien esitysmuotojen käyttöön.

5.2 Toimijakenttä

Tässä on tarkasteltu yksilöiden, julkisten organisaatioiden, palveluita tuottavien yritysten, tutkimuslaitosten ja muiden yhteisöjen kuten kansalaisjärjestöjen, standardisointiorganisaatioiden ja kansainvälisten verkostojen motiiveja ja rooleja osallistua My Datan edistämiseen. Tarkastelu ei ole tyhjentävä, sillä kaikkien eri toimijaryhmien edustajia ei ole vielä tavoitettu aiheen äärelle. Olennaista on nähdä toimijakentän laajuus; My Data ei ole vain teknologiayritysten pelikenttä, julkishallinnon agenda tai yksilöiden oikeus. Kansallisella tasolla ja kansainvälisesti My Datan edistäjien on tunnistettava, kuinka eri tahot voidaan saattaa rikastavaan toimintaan niin, että suunta on yhteinen, vaikka sitä ei keskitetysti kukaan johdakaan.

5.2.1 Yksilöiden rooli

Alusta alkaen on tärkeää löytää My Data -sovelluksia jotka palvelevat konkreettisesti yksilöitä olivat he sitten urheilijoita, tietotyöläisiä, terveydestään tai yleisesti itsensä mittaamisesta ja itsestään oppimisesta kiinnostuneita. Ihmisoikeuksiin tai tulevaisuuden kestävään kehitykseen liittyvät hyödyt saattavat kiinnostaa yksittäisiä ihmisiä, mutta jokapäiväiset hyödyt saavat todennäköisesti suuremmat joukot liikkeelle.

Ihmisten yhteistoiminta on My Data -kehityksen alkuvaiheessa erittäin keskeinen muutosvoima. Yksi tapa organisoida yhteistoimintaa voisi olla osuuskunta, joka voisi pilotoida My Data -operaattorin toimintaa. Osuuskunnat ovat toimineet muun muassa pankkitoiminnassa, veden saannissa, kaupan alalla ja telekommunikaatiossa alkuvaiheen toiminnan käynnistäjänä. Yksilöiden rooli korostuu myös erilaisten My Datan käyttötapausten esikaupallisessa kehittämisessä, jota yritysmuotoinen palvelukehitys seuraa.

5.2.2 Julkisten organisaatioiden rooli

Julkisen sektorin rooli My Datan edistäjänä liittyy ensisijaisesti My Datan huomioimiseen politiikan eri osa-alueilla aina oikeuspolitiikasta liikennepolitiikkaan ja elinkeinopolitiikasta koulutuspolitiikkaan. Esimerkiksi koulutuspoliittisena tavoitteena tulisi pitää kansalaisten datalukutaidon, datavalveutuneisuuden ja digitaalisen yksityisyyden suojaamisen taitojen kehittämistä. Jokainen sektori voi toki tehdä oman My Data -strategian, mutta kuten aiemmin on esitetty, niin My Data -lähestyminen on pohjimmiltaan sektorirajat ylittävää. Politiikan tulisi tukea sitä, että erilaiset My Data -palvelujen muodot enenevässä määrin korvaavat nykyisen sektoripohjaisen rakenteen.

Julkishallinnon organisaatiot siinä kuin yrityksetkin keräävät ja käsittelevät paljon henkilötietoa, joka olisi mahdollista tarjota ihmisille itselleen My Data -periaatteiden mukaisesti. Julkisella sektorilla henkilötiedon käsittely pohjautuu suurelta osin lainsäädännön velvoitteisiin eikä esimerkiksi ihmisten itsensä antamaan suostumukseen, mutta yhtä kaikki, ihmisille hyödyllinen data tulisi antaa heille, ja tässä julkistoimijat voivat olla jopa tien avaajia. My Data -lähestyminen voisi myös helpottaa julkisten ja yksityisten palveluiden yhtyeentoimivuutta, kun ihminen saisi itse siirtää datan palvelusta toiseen.

Julkinen sektori kannustaa ja valvoo yrityksiä. Kannustimena voi toimia esimerkiksi kansallinen hanke kilpailukyvyn parantamiseksi tai muutokseen sopeutumiseksi. Myös lakimuutoksilla ja ohjeistuksilla ja sääntelyllä voidaan luoda puitteita, jotka kannustavat My Data -liiketoiminnan kehittämiseen.

5.2.3 Yritysten rooli

Erilaisilla yrityksillä toiminnan laajuudesta (kansallinen, globaali, kasvuyritys) ja toimialasta riippuen on keskenään erilaiset edellytykset ja tavoitteet, ja siten rooli My Datan edistämisessä ja hyödyntämisessä.

Kaikki henkilötietoa käsittelevät yritykset edistävät My Dataa avaamalla henkilötietorajapintoja. Osa yrityksistä on keskeisessä asemassa uusien henkilötiedon keruujärjestelmien ja My Datan hallintajärjestelmien kehityksessä. Suuria toimijoita kiinnostaa yleisesti, miten toimintamallit ja standardit kehittyvät. Oma lukunsa ovat asiakkuudenhallintajärjestelmiä (CRM) toimittavat yritykset, joilla on mahdollisuus toteuttaa omiin tuotteisiinsa My Data -rajapinnat ja myydä niitä edelleen asiakasyrityksilleen. Startup-yritysten pääasiallinen rooli on demonstroida, miten saataville tulevalla My Datalla voi tehdä erilaisia hyödyllisiä sovelluksia.

My Data yhdistettynä muuhun palveluinfrastruktuurin kehittämiseen mahdollistaa long tail -ilmiön, joka tarkoittaa yhä pienempien asiakastarpeiden tehokasta palvelua. My Datan vaikutus verkkokaupan ja asioinnin logistiikkaan ja markkinointiin on merkittävä ja koskettaa yrityksiä eri toimialoilla.

Motiiveina henkilötietorajapintojen avaamiseen sekä infrastruktuuripalveluiden ja -sovellusten tuotekehitykseen voivat toimia:

- Markkinointinäkökulma ja brändihyöty (eettinen ja asiakasystävällinen reilu toimija, medianäkyvyyttä)
- Edelläkävijyys suhteessa kilpailijoihin (uutuusarvo asiakkaille)
- Uudet tavat kiinnittää asiakkuuksia ja lisätä asiakkaiden luottamusta (kanta-asiakasohjelmien tehostaminen – vahva profiili)
- Uudenlainen globaali kasvumahdollisuus, jos kotimarkkina on edelläkävijänä kansainvälisessä kehityksessä
- Perinteisen rakenteen haastaminen ja horisontaalinen asema markkinoilla
- Sensitiivisen tiedon ongelmilta välttyminen, kun toiminta perustuu yksilön antamaan valtuutukseen

Yllä lueteltujen yritysten omien motiivien lisäksi myös ulkoiset vaikuttimet kuten sääntely, julkinen rahoitus ja yleinen mielipideilmasto vaikuttavat yritysten innokkuuteen lähteä mukaan muutokseen. Liiketoiminnallinen perustelu löytyy usein siitä, että My Data -rajapintojen ja muiden mahdollisesti investointeja vaativien muutosten toteuttaminen katsotaan osaksi palveluiden digitalisaatiota tai asiakastietojen parempaa hallintaa, jotka ovat molemmat itsessään merkittäviä yritysten kilpailukykytekijöitä, joihin kannattaa sijoittaa.

5.2.4 Tutkimuslaitosten rooli

Yliopistojen tehtäviin kuuluu koulutus ja uuden tiedon luominen sekä toimiminen yhteiskunnallisen kehityksen kommentoijana ja uusien innovaatioiden synnyttäjänä. My Data on merkittävä kaikkien näiden tehtävien kannalta. Tutkimus- ja kehitystoimintaa tapahtuu myös yliopistojen tutkimusyksiköiden ulkopuolella muissa tutkimuslaitoksissa ja yrityksissä. Tutkimuksen rooli My Datan edistäjänä on kuitenkin pitkälti sama riippumatta siitä, missä se tapahtuu. Tutkimuslaitokset ovat erittäin kiinnostuneita uusista tavoista kerätä ja hyödyntää henkilötietoa.

- **Uuden tiedon luominen** – My Data helpottaa yksilöiden itse keräämän henkilötiedon hyödyntämistä tutkimuksessa. Käytännössä edellytyksenä on My Data -infrastruktuurin tarjoama helppo anonymisointikäytäntö, jonka avulla datan tutkimuksellinen arvo voidaan säilyttää tutkittavien yksityisyyden suojaa vaarantamatta.
- **Tutkimus opastamassa yhteiskunnan kehitystä** – Henkilötiedon kerääminen ja hyödyntäminen muuttavat yhteiskunnan rakenteita, siksi tarvitaan muiden muassa humanistista, psykologista, oikeustieteellistä ja yhteiskuntatieteellistä tutkimusta My Datan hyödyistä ja haitoista.
- **Akateemisia My Data innovaatioita** – Mikäli henkilötiedon jalostamisen ja välittämisen infrastruktuureista tehdään My Data -mallin mukaisesti avoimia, niin silloin tiedeyhteisöllä voi olla merkittävä rooli erilaisten teknisten komponenttien kehityksessä sekä standardien luomisessa yhdessä yritysten ja muiden organisaatioiden kanssa. Mikäli henkilötietoinfrastruktuuri on suljettu, tiedeyhteisön mahdollisuus vaikuttaa kehitykseen on huomattavasti rajallisempi.

5.2.5 Muiden yhteisöjen rooli

Monet My data -sovelluksia kehittävät yritykset ovat riippuvaisia siitä, että muut organisaatiot kuten pankit, kauppaketjut jne. avaavat rajapintoja. Sovelluksia kehittävät yritykset ovat kuitenkin hankalassa välikädessä pyytämässä rajapintojen avaamista. Vaatimuksen pitäisi tulla mieluiten puolueettomalta taholta, ihmisiltä itseltään, kuten esimerkiksi pankkien ja kauppaketjujen asiakkailta. Sovellusten kehittäjät tarvitsevat siis promoottoreita ja aktiivisia puolestapuhujia, jotka puolueettomasta roolistaan käsin kaatavat raja-aitoja ja innovaatioiden onnistumisen esteitä. Yhtenä järjestöjen ja vapaasti organisoitujen kansanliikkeiden roolina on koota yhteen ihmisiä tekemään joukkoistettuja datapyyntöjä ja mielenilmaisuja. Tällaisia hankkeita on jo käynnissä, kuten esimerkiksi Terveystieto.me[22], joka pyrkii kokoamaan 10 000 suomalaista yhdessä pyytämään terveystietonsa My Datana. Vastaavalla tavalla esimerkiksi ympäristöjärjestöt ja eettiset liikkeet voisivat kerätä toimijoita yhteen ja aktivoida kansalaisia, jotta kattavaan kulutusprofiiliin perustuvien jalanjälkilaskureiden toteuttaminen olisi mahdollista.

22 http://terveystieto.me

My Datan kehittymisen kannalta on ensiarvoisen tärkeää edistää rajapintojen ehtojen määrittelyä, muodostaa lisenssimalleja ja rakenteita, joilla yksityisyysasetusten ja tiedon hallinnan voi tehdä keskitetysti monen palvelun osalta sekä formaatteja ja määrityksiä, jotka auttavat ymmärtämään, milloin tieto on laadultaan riittävää. Henkilötietorajapintojen avaamiseen tähtäävien kampanjoiden ohella myös standardien valmisteluun vaikuttaminen on keskeinen rooli kansalaisjärjestöille ja muille yhteisöille. Tekninen kehitys on globaalia ja yhteisöjen pitää vaikuttaa myös kansainvälisissä standardointiorganisaatioissa kuten W3C:ssa.

5.3 Esteitä My Datan edessä

Yleisin My Data -ajattelua kohtaan esitetty kritiikki on, että se saattaisi entisestään kiihdyttää suuntausta, jossa ihmisiltä vaaditaan jatkuvasti enemmän henkilötietoa - voiko olla esimerkiksi niin, että tulevaisuudessa kohtuuhintaista vakuutusta on mahdotonta saada antamatta ensin kattava ja rikas profiilidata vakuutusyhtiölle? My Datan toteutuminen ei poista tarvetta alakohtaiselle sääntelylle. Edelleen voi olla tarpeellista säännellä esimerkiksi sitä, millaista tietoa vakuutuksenottajilta saa pyytää. Sääntelyä ja yhteisiä sopimuksia toimeenpanevien ja valvovien viranomaisten ja myös kansalaisjärjestöjen tehtävä on pitää huolta, etteivät organisaatiot väärinkäytä yksilöiden suostumusta henkilötiedon käyttöön, ja tarvittaessa tiedottaa ja reagoida väärinkäytöksiin. My Datan myötä tiedon siirtäminen palveluntarjoajalta toiselle olisi kuitenkin henkilön itsensä käsissä toisin kuin tällä hetkellä.

Toinen yleisesti esille noussut kysymys on, pitäisikö joissain tilanteissa ainakin pohtia, pitääkö ihmistä suojella hänestä kerätyltä tiedolta. Ihmiselle saattaa paljastua omasta datastaan asioita, joiden käsittelyyn hän ei ole valmistautunut. Tästä esimerkkinä on geenitieto, jonka pohjalta voidaan arvioida esimerkiksi kohonnutta alttiutta sairauksille, joista osaan ei ole parannuskeinoa. Kansalaisen pitäisi itse pystyä päättämään, haluaako hän esimerkiksi tietää, onko hänellä geeneistä johtuva kohonnut riski sairastua johonkin vakavaa tautiin kuten Alzheimeriin. Holhoavan suojelemisen sijaan tulisi varmistaa, että ihmisille on tarvittaessa tarjolla tukea tiedon käsittelyssä, esim. läheisiltä, valmentajalta, opettajalta, ystäviltä, vertaisryhmältä. Mikäli tällaista tukea ei ole saatavilla ja ihminen jää yksin datansa kanssa, se voi aiheuttaa negatiivisia seurauksia. Tämän takia esimerkiksi lääkäreille opetetaan ihmissuhdetaitoja: miten kertoa potilaan terveydestä huonoja uutisia potilaalle. Lääkäri on siinä hetkessä tukihenkilö.

Yllä olevien lisäksi muutamat muut uhkakuvat nousevat usein esille My Dataa sivuavissa keskusteluissa. Niitä on lyhyesti lueteltu alla.

- Ihmiset eivät osaa huolehtia yksityisyydestään ja jakavat vahingossa dataansa laajemmin kuin ymmärtävätkään tai jakavat ymmärtäen, mitä tekevät, mutta laajemmin kuin heille itselleen on hyväksi.
- Markkinoille tulee sovelluksia, jotka ovat virheellisiä ja antavat ihmisten dataan pohjautuen valheellisia, vääriä tai jopa haitallisia tulkintoja.
- Yksilöt määrittävät tai seuraavat liikaa itseään datan kautta, jolloin aito elämyksellinen ja aistinvarainen tieto ja itsetuntemus jäävät entistä pienempään rooliin.
- Ihmiset alkavat itse tulkita omaa terveysdataansa eivätkä hae ammattilaisen apua.
- Siirrettäessä tietoja yksilön kautta sisältö saattaa vääristyä, joko tahallisesti tai tahattomasti.
- Tiedon konteksti katoaa, mikä johtaa vääriin tulkintoihin. Karrikoituna esimerkkinä tutkija saattaa ”tykätä” Facebookissa natsisivuista tutkimuksen tekemisen tarkoituksessa, mutta kontekstista tietämättä tehty profilointi saattaa johtaa väärään päätelmään.

- Väärennetyillä sertifikaateilla toisina esiintyvät palvelut saattavat huijata ihmisiä luovuttamaan henkilötietojaan ja käyttää niitä yksilön edun vastaisesti.

Eri toimijoiden kanssa käydyissä keskusteluissa uhkakuvat ovat nousseet esille, mutta yleisesti ottaen niitä ei pidetä My Datan toteutumisen esteinä, vaan pikemminkin seikkoina, jotka on hyvä ottaa huomioon, kun tulevaisuuden tietoinfrastruktuuria suunnitellaan ja toteutetaan.

Uhkakuvia suurempia esteitä tai hidasteita ovat muutamat My Data -ajattelun kanssa ristiriitaiset näkökulmat, jotka ovat toistuneet eri yhteyksissä. Näitä näkökulmia käsitellään seuraavaksi.

5.3.1 Data on keino pitää asiakas

Yritys tai palvelu, joka on onnistunut pitämään asiakkaan pitkään, on voinut myös kerryttää ison datahistorian. Esimerkiksi vuosikausia kestäneen pankkiasiakkuuden myötä pankin järjestelmiin on kertynyt henkilön tilitapahtumien historia koko tältä ajalta. Monessa tapauksessa data ja historia myös sitovat henkilöä kyseisen yrityksen asiakkaaksi tai palvelun käyttäjäksi. Esimerkiksi Facebookista irtautuminen olisi kivuliasta vaikka markkinoille tulisi parempi palvelu, koska siellä on tuttavaverkostoja, viestintähistoriaa, valokuvia yms. jo pitkältä ajalta.

Asiakkaiden kannalta helppo mahdollisuus vaihtaa palvelua tuntuu luonnolliselta ja oikeutetulta, mutta yrityksen näkökulmasta liiketoimintariski voi kasvaa suureksi, jos kuka tahansa uusi tulokas markkinoilla voi kaapata koko asiakaskunnan datoineen hetkessä.

Tulevaisuuden mahdollinen markkinatilanne, jossa suurin osa yrityksistä toimii My Data -periaatteiden mukaisesti, ohjaisi todennäköisesti loputkin yritykset ja uudet tulokkaat toimimaan samalla periaatteella, koska asiakkaat eivät valitsisi yritystä, joka haluaa lukita asiakkaan ja hänen datansa. Nykytilanteessa edelläkävijäyritysten ongelmana on (first mover problem), että muut yritykset olisivat kyllä halukkaita hyödyntämään asiakkaan dataa, mutta eivät luovuttamaan sitä asiakkaalle vapaasti käytettäväksi.

5.3.2 Pyrkimys keskipisteeksi

Ajatus rajapintojen kautta liikkuvasta datasta, joka mahdollistaa ketterän organisaation ulkopuolisen sovelluskehityksen, toivotetaan usein tervetulleeksi uudistuksena, joka mahdollistaa paremman palvelutarjonnan yrityksen asiakkaille, kun muut toimijat voivat tarjota täydentäviä palveluita.

Luontainen ajatus on nähdä oma organisaatio tällaisen ekosysteemin keskiössä. Lukuisilla powerpoint-kalvoilla yritys on keskellä ja muita kolmansia osapuolia reunoilla hyödyntämässä dataa ja rikastamassa palveluita. Useinkaan ei nähdä kirkkaasti, että toimivasta ekosysteemistä on merkittävää arvoa, vaikka ei itse olisikaan keskiössä. Erikoistumalla kukin voi tehdä sitä, mitä parhaiten osaa ja tukeutua ulkopuolisiin toimijoihin niissä asioissa, joita ei ole rahkeita toteuttaa riittävällä laadulla itse. Jos joku on My Data -systeemin keskiössä, niin sen tulisi olla yksilö, joka dataansa kontrolloi.

Käytännössä tekniset ja strategiset ratkaisut, joita tehdään esimerkiksi henkilötietorajapintojen toteuttamiseksi, ovat pitkälti samoja riippumatta siitä, nähdäänkö yrityksen olevan keskellä vai reunalla. Mielen mallina kuitenkin organisaatiokeskeinen ajattelu on ristiriidassa My Datan ihmiskeskeisen näkökulman kanssa. Se saattaa estää näkemästä yrityksen kannalta merkityksellisiä My Datan mahdollisuuksia, jotka aukeavat nimenomaan asiakkaiden ollessa

keskiössä. Vaikka oman yrityksen näkeminen kaiken keskellä ei ole varsinaisesti vaarallista, niin suuressa osassa tapauksia se on epärealistista. Vain muutamat yritykset kerrallaan voivat olla globaalisti toimivien ekosysteemien napoina. Tasapainoinen avoin ekosysteemi antaa huomattavasti suuremmalle määrälle yrityksiä tilaa olla mukana.

5.3.3 Meidän datasta ei ole muille iloa

Vaikka My Datan perusperiaatteisiin suhtauduttaisiinkin myönteisesti eikä organisaatioilla olisi liiketaloudellisia intressejä henkilötietoihin, niin usein juuri oman organisaation hallinnoimiin henkilötietoihin saatetaan silti suhtautua protektionistisesti. Datalle ei nähdä mitään hyötykäyttöä alkuperäisen käyttötarkoituksen ulkopuolella.

Esimerkiksi hoitotyössä on ehdottoman tärkeää, että potilastieto on oikeaa ja historiallisesti kattavaa. Diagnoosia tehdessään lääkäri luottaa laboratoriotestien ja mittausten tuloksiin, jos hän tietää, että ne on tehty oikein ja oikeissa olosuhteissa. Tiedon oikeellisuuden voi taata esimerkiksi virallinen sähköinen potilastietokanta, mutta siitä otettu My Data -kopio olisi altis tahalliselle tai tahattomalle manipuloinnille. Lääketieteellisen käytön ja hoitotyön kannalta siis virallinen järjestelmä on luotettavin ja My Datalle ei välttämättä nähdä tarvetta.

Potilastietojärjestelmän osalta voisi ajatella, että vaikka lääketieteellisessä käytössä olisikin vain virallinen järjestelmä, niin datan saaminen ulos My Datana mahdollistaisi, että ulkopuoliset tahot voisivat kehittää sovelluksia ja käyttöliittymiä eri päätelaitteille ja pienille kohderyhmille kuten näkövammaisille, kielitaidottomille, diabeetikoille jne. Potilas voisi esimerkiksi ajaa oman lääkityksensä validaattorisovellukseen ja saada tietoa, onko päällekkäisyyksiä tai ongelmia, ja konsultoida sen jälkeen omaa lääkäriään mahdollisista jatkotoimista.

Ulkomaalainen terveydenhuollon organisaatio voisi saada potilastiedon käyttöönsä potilaan välittämänä, vaikka virallista yhteyttä kotimaan potilastietojärjestelmään ei olisi. Tässä tapauksessa datan autenttisuutta ei voitaisi välttämättä varmistaa, mutta tilanne olisi parempi kuin ei dataa ensinkään.

Potilas voisi itse tarkastella eri verikokeissa mitattua hemoglobiiniarvoa ja verrata sen kehittymistä ruokavalionsa muutoksiin. Hemoglobiiniarvot hän saisi kätevästi rakenteellisessa muodossa potilastietojärjestelmän My Data -ominaisuuden ansiosta.

Vaikka esimerkkinä yllä käytettiin potilastietoa, niin vastaava alkuperäiseen käyttötarkoitukseen lukkiutuminen on yleinen vaiva kaikilla sektoreilla. Vaikka ei itse heti keksikään mitään mielekästä ulkopuolista käyttöä oman organisaation hallinnoimalle henkilötiedolle, olisi hyvä asennoitua avoimen uteliaasti siihen, mitä ihmiset itse ja heidän valtuuttamaan muut oman alan ulkopuoliset toimijat voisivat keksiä.

5.3.4 Suojellaan tietoa ihmiseltä itseltään

Henkilötietojen käsittelyyn liittyvät lainsäädännölliset velvoitteet ovat yrityksille tuttuja. On yllättävää, kuinka usein tietosuojaan liittyvät vastuut nostetaan mahdollisten My Dataa estävien asioiden listalle, vaikka kysymys on siitä, että yksilölle itselleen vain annetaan pääsy omiin tietoihinsa koneluettavassa muodossa.

Esimerkiksi pankit eivät ole avanneet henkilöasiakkaille pääsyä omiin pankkitietoihin rajapintojen kautta, vaikka yritysasiakkaat voivat kytkeä taloushallinto-ohjelmistonsa pankin järjestelmiin. Toimintatapaa on perusteltu henkilöasiakkaiden tietosuojalla. Verkkopankissa henkilöasiakkaatkin voivat manuaalisesti ladata tilitapahtumat omalle koneelleen. Tietosuojamielessä omalle ko-

neelle lataaminen ei eroa siitä, että pankkitunnuksilla tunnistautunut henkilö valtuuttaisi tiedonsiirron rajapinnan kautta vaikkapa omaan henkilökohtaiseen taloushallinto-ohjelmistoonsa.

Toisaalta tietosuojaan liittyviä velvoitteita myös halutaan välttää. My Data -periaatteilla voitaisiin kääntää tilanne sekä teknisesti että juridisesti niin, ettei palvelun tarjoajan tarvitsisi pitää dataa, vaan palvelulla olisi yksilön luvalla pääsy tarvittaessa käyttämään dataa. Nykyisen tietosuojalainsäädännön terminologiaa käyttäen ihminen toimisi itse rekisterinpitäjänä ja olisi vastuussa omasta datastaan. Tällainen yksinkertainen malli voisi avata innovaatiolukkoja monilla alueilla.

EU:n tietosuoja-asetuksen tilanne

Viimeisimmässä Euroopan parlamentin käsittelyssä olleessa tietosuoja-asetuksen versiossa My Dataa läheisimmin koskeva säännös löytyy artiklasta 15 (European Parliament 2014), mutta sitä on merkittävästi heikennetty komission ehdotukseen nähden rajaamalla yksilöiden oikeus koskemaan vain sellaisia tietoja, jotka henkilö on itse syöttänyt.

> *Article 15 / Right to access and to obtain data for the data subject*
>
> *2a. **Where the data subject has provided** the personal data where the personal data are processed by electronic means, the data subject shall have the right to obtain from the controller a copy of the provided personal data in an electronic and interoperable format which is commonly used and allows for further use by the data subject without hindrance from the controller from whom the personal data are withdrawn. Where technically feasible and available, the data shall be transferred directly from controller to controller at the request of the data subject.*

Yllä oleva Euroopan Parlamentin hyväksymä muotoilu ei kuitenkaan vielä anna lopullista kuvaa tulevasta lainsäädännöstä. Artiklasta keskustellaan edelleen neuvostossa ja neuvoston päästyä artiklasta yksimielisyyteen, neuvottelut jatkuvat vielä instituutioiden välillä.

My Datan kehittymiseen tarvitaan avoimuutta, yhteistoimintaa ja luottamusta yksilöiden ja organisaatioiden välille.

6. My Data Suomessa

Aiemmissa luvuissa on yleisellä tasolla kuvattu My Datan mahdollisuuksia ja toteutumisen edellytyksiä ottamatta kantaa erityisesti kansallisiin kysymyksiin. Yksi selvityksen lähtökohta on kuitenkin hahmottaa Suomen roolia My Datan systeemisessä kehittämisessä.

My Datan kehitys on kansainvälistä. Tällä hetkellä näkyvimmät My Data -projektit lähtevät Yhdysvalloista tai Britanniasta kuten esimerkiksi Midata-ohjelma[23]. EU:ssa valmistellaan parhaillaan uutta tietosuoja-asetusta, joka vaikuttanee henkilötiedon käsittelyyn, mutta sen voimaantuloa ei kannata odotella toimettomana. Henkilötiedon helppoa saatavuutta ja parempaa yhteentoimivuutta eri palvelujen ja järjestelmien välillä kannattaa tavoitella asetuksista riippumatta.

Nyt on aika toimia, jotta suomalaiset ovat mukana kehittämässä uusia henkilötiedon hyödyntämisen käytäntöjä. Suomella on mahdollisuus rakentaa kansainvälisesti merkittävä tukijalka My Data -pohjaiselle osaamiselle ja teknologioille. Pienessä, korkeasti koulutetussa maassa yhteisistä käytännöistä on helpompi sopia kuin suuremmissa maissa. Onnistuneet ratkaisut voidaan skaalata maailmalle.

Globalisaatio muuttaa suomalaista yhteiskuntaa ja palveluja myös perinteisillä aloilla, jotka ovat yleensä olleet kotimaisten toimijoiden hallussa. Pelko isojen monikansallisten yritysten liiketaloudellisen vallan kasvamisesta on ajankohtaista ja aiheellista. Tietoinfrastruktuurin organisoiminen avoimuuden periaatteilla voi merkittävästi vaikuttaa kotimaisten toimijoiden kykyyn vastata kansainvälisten yritysten haasteeseen, joka perustuu pääasiassa suljettuihin järjestelmiin ja suuren koon tuomaan etuun.

My Datan kehittymiseen tarvitaan avoimuutta, yhteistoimintaa ja luottamusta yksilöiden ja organisaatioiden välille. Suomalaisesta yhteiskunnasta löytyy näitä ominaisuuksia. Suomessa on vahvat perinteet avoimessa tietoteknisessä kehityksessä kuten avoimen datan ja avoimen lähdekoodin alueilla. Toisaalta myös yritysten ja valtion yhteistyöllä on saatu tuloksia. GSM-standardin nopea omaksuminen ja sitä seurannut mobiiliteollisuuden kehittyminen ovat hyviä esimerkkejä siitä, miten Suomesta voidaan panna alulle kansainvälisesti ja kaupallisesti merkittäviä informaatioteknologian standardeja.

Henkilötiedon alueella ratkaistavat kysymykset ovat erilaisia, eikä vanhoihin GSM- ja teleoperaattorin maailman esimerkkeihin ja menestystarinoihin pidä tuudittautua liian syvälle. Toiminnan edistämiseksi on syytä luoda rohkea mutta uskottava visio siitä, miten Suomi voi kehittää uusia toimintamalleja.

Pyrkikäämme siihen, että suomalaiset yritykset ovat varhain liikkeellä ja suomalainen sääntely ohjaa kehitystä oikeaan suuntaan. *'Voi kysyäkin, voisiko Suomi tehdä tietosuojasta ja yksilön oikeuksien kunnioituksesta myös vientituotteen'*, päätti Heinäluoma valtiopäivien avajaispuheensa (Eduskunta 2014). Ja vastaus on "kyllä", jos niin haluamme. Tämä edellyttää, että yritykset lähtevät varhaisessa vaiheessa mukaan kehittämään uutta liiketoimintaa - soveltuvin osin yhdessä julkisen sektorin kanssa. Tässä jos missä *Public Private People Partnership* -ajattelua tarvitaan.

Selvityksen lisäksi tarvitaan muitakin toimia: rakenteellisia ja lainsäädännöllisiä muutoksia, tutkimusta ja tuotekehitystä, toimintatapojen ja asenteiden uudistamista sekä ennen kaikkea tiivistä vuorovaikutusta eri tahojen ja sektoreiden välillä. Tarvitaan myös innokkaita ihmisiä ja organisaatioita, jotka ovat valmiita kokeilemaan ja oppimaan kokeiluista.

---

23 https://www.gov.uk/government/news/the-midata-vision-of-consumer-empowerment

6.1 My Dataan liittyviä kansallisia kehityskulkuja

Henkilötiedon uudelleen organisointi ihmiskeskeisesti on keskeinen teema, kun tiedon ja henkilötiedon määrä ja merkitys yhteiskunnassa kasvaa. Alle on koottu muutamia selvitystyön aikana esille nousseita kansallisia kehityskulkuja, joilla on selkeästi liittymiä My Data lähestymiseen.

Palveluväylä

Valtiovarainministeriön johdolla tehtävä palveluväylä on keskeinen kansallinen tietoarkkitehtuurin kehittämishanke. Toteutuessaan palveluväylä mahdollistaa henkilötietolähteiden käytön eri järjestelmissä ja yhdistämisen nykyistä yksinkertaisemmin. Palveluväylä mahdollistaa rajapintojen ja järjestelmien avaamista ja siten tukee My Datan kehitystä. Yksistään palveluväylä ei kuitenkaan takaa yksilölle pääsyä tietoon tai anna valtaa vaikuttaa oman tiedon välittämiseen ja käsittelyyn. Tätä varten palveluväylään tulisi tehdä ihmiskeskeisesti toteutettu liittymä, jota kautta My datan periaatteita voitaisiin paremmin toteuttaa suhteessa palveluväylän kautta siirrettävään henkilötietoon.

Big dataan liittyvä tutkimus ja kehitys

Viime aikoina big data ja erilaisten datan käsittelyalustojen (platformien) kehitys ovat olleet suosittuja tutkimus- ja tuotekehitysteemoja ja saaneet merkittävästi rahoitusta ja huomiota niin kansallisesti (esimerkiksi Digilen D2I ohjelma[24]) kuin Euroopan tasolla (esimerkiksi Horizon 2020, Big Data and Open Data innovation take-up, ICT-15-2014[25]). Big data ei ole eksakti tutkimuskäsite, vaan suurien tietomassojen analytiikkaa kuvaava markkinointikäsite. Suomessa big dataan liittyvää rahoitusta myönnetään niin Tekesin, Akatemian kuin valtioneuvoston strategisten linjausten ja ministeriöiden toimesta. Koska My Datalla ja big datalla on yhteys (ks. tietolaatikko johdanto-luvussa) esimerkiksi anonymisointirakenteiden kautta, voidaan osa big data tutkimukseen ja kehitykseen kanavoidusta rahoituksesta ohjata My Dataa -infrastruktuurin kehittämiseen. Käytännössä big datan tutkimuksen kannalta My Data voidaan nähdä ratkaisuna yksityisyys- ja suostumuskysymyksiin, jotka ovat useassa yhteydessä nousseet big datan osalta keskeiseksi haasteeksi.

Turvallinen infrastruktuuri

Suomi voisi olla tietoturvallinen innovatiivinen datan turvasatama, jossa on etuina merikaapelit, konesalit ja turvallinen infrastruktuuri. Yllä on jo viitattu eduskunnan puhemies Eero Heinäluoman kysymykseen voisiko tietosuojasta ja yksilön oikeuksien vaalimisesta tulla Suomelle vientituote. Turvallinen infrastruktuuri ei tarkoita pelkästään fyysistä infrastruktuuria, vaan myös tapaa millä tieto- ja viestintätekniikkaan liittyvää infrastruktuuria organisoidaan ja hallitaan. My Data on ehdotus kuinka laskentaan ja tiedon käsittelyyn voidaan kehittää infrastruktuuri, jossa yksilö on keskiössä, ja johon liittyvät teknologi-

24 http://www.datatointelligence.fi/
25 http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/87-ict-15-2014.html

at voivat muodostaa kansainvälisen markkinan. Suomi voisi olla houkutelevan alusta uusien My Data -palveluiden pilotoinnille ja kehitykselle kuten Suomi oli vuosituhannen alussa pilottimarkkina monille mobiilialan innovaatioille.

Kansallinen terveystili

Laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä (Asiakastietolaki 2007) velvoittaa terveydenhuollon organisaatiot tallentamaan potilastiedot valtakunnallisesti keskitettyyn arkistoon. Kansallinen Terveysarkisto KanTa[26] on lain pohjalta valmisteltu järjestelmä, joka mahdollistaa muun muassa sähköisen reseptin ja tarjoaa kansalaisille käyttöliittymän omien potilastietojen tarkasteluun. Laki ja järjestelmä eivät kuitenkaan anna potilaille mahdollisuutta tallentaa omia potilastietojaan tai siirtää niitä toiseen sovellukseen. Lakia ja KanTa -järjestelmän toteutusta tulisi tältä osin pikaisesti muuttaa niin, että se toimisi My Data -periaatteiden mukaisesti. Eduskunnassa on ollut käsiteltävänä myös kirjallinen kysymys kansallisesta terveystilistä, johon vastuuministeri on vastannut, että omien terveystietojen hallintaan tarkoitetun tilin rakentaminen kuuluu suunnitelmaan[27]. Ministerin vastauksessa kerrotaan myös, että ratkaisu sisältää rajapinnan kaupallisillekin sovelluksille. Terveydenhuollon My Dataan liittyvät myös käynnissä olevat potilastietojärjestelmähankkeet, kuten Apotti[28]. Olisi erittäin toivottavaa, että My Data -periaatteet vaikuttaisivat potilastietojärjestelmien kehitykseen ja määrittelyyn.

26 http://www.kanta.fi/
27 http://www.eduskunta.fi/faktatmp/utatmp/akxtmp/kk_587_2014_p.shtml#VASTAUS
28 http://www.hel.fi/hki/apotti/fi/Etusivu

6.2 Suomalaistoimijoita

Suomessa on useita My Dataan liittyviä toimijoita, hankkeita ja aloitteita. Open Knowledge Finland ylläpitää listaa suomalaisista My Data -toimijoista osoitteessa fi.okfn.org/wg/my-data/suomessa/

Listaus sellaisena kuin se oli selvityksen julkaisun aikaan syyskuussa 2014:

Nimi ja url	Kuvaus
Arkkeo http://www.arkkeo.com yritys	Arkkeo on henkilökohtainen arkistointipalvelu, joka tallentaa kauppojen, vakuutusyhtiöiden, matkatoimistojen yms. toimijoiden lähettämät kuitit, takuutodistukset ja matkaliput. Arkkeo rakentaa parhaillaan palveluverkostoa matkailualan, vähittäiskaupan, terveydenhuollon, rahoitus- ja vakuutusalan yritysten sekä palveluorganisaatioiden kanssa, jotta voisit automaattisesti saada kaikki tärkeät henkilökohtaiset dokumenttisi suoraan Arkkeo-tilillesi.
Balancion http://www.balancion.com yritys	Balancion on verkossa toimiva palvelu henkilökohtaisen talouden hallintaan, suunnitteluun ja seurantaan. Palvelun avulla käyttäjät voivat noutaa tilitapahtumatietonsa eri verkkopankeista Balancion-palveluun, omaan käyttöönsä. Palvelu esittää käyttäjän tulot ja menot jaoteltuna erilaisiin tulo- ja menoluokkiin.
Digital Health Revolution http://cht.oulu.fi/news/digital-health-revolution-was-launched-in-oulu hanke	Monialaisen Digital Health Revolution -tutkimushankkeen visiona on mahdollistaa yksilöstä saatavan tiedon hyötykäyttö osana personoituja, ennaltaehkäiseviä palveluja sekä parantaa kansalaisen mahdollisuuksia omaehtoiseen hyvinvoinnin ylläpitoon. Hankkeen lähtökohtana on, että yksilö voi hyödyntää omia tietojaan itse valitsemissaan käyttökohteissa sekä kontrolloida tietojensa käyttöä.
Foodie http://foodie.fm yritys	Foodie.fm on personoitu sosiaalinen osto- ja ruokasuosituspalvelu. Foodien suositusteknologia oppii käyttäjien valinnoista, joihin pohjautuen palvelu suosittelee kunkin käyttäjän yksilöllisiin tottumuksiin parhaiten sopivia reseptejä ja tuotteita.
Glome http://glome.me yritys	Glome on anonyymi datalompakko, eräänlainen dataa keräävä asiakkuuskortti, jonka sisältämästä tiedosta kuluttaja päättää itse ja saa etuja valitsemansa tiedon jakamisesta digitaalisessa maailmassa. Glomen avulla yritykset saavat paikkasidonnaista tietoa kuluttajan digitaalisesta jalanjäljestä eli tarkkaa tietoa siitä, mistä kuluttaja on kiinnostunut.
Huge http://huge.fi yritys	Huge on tapa vastaanottaa etuja ja kerätä ostohyvityksiä liikkeistä, ravintoloista tai kahviloista. Mukana olevat yritykset lähettävät käyttäjälle henkilökohtaisia etuja, joita käyttämällä käyttäjä kerryttää ostohyvitystä.
Liikennelabra http://liikennelabra.fi/ hanke	Liikennelabra on liikenne- ja viestintäministeriön käynnistämä kokeiluhanke, jonka tavoitteena on edistää älyliikenteen palvelumarkkinan syntymistä.

Nimi ja url	Kuvaus
Open Knowledge Finland http://okf.fi/my-data yhteisö	Open Knowledge Finlandin My Data -työryhmä on avoin kaikille My Datasta kiinnostuneilla. Ryhmässä joukko yritysten ja julkishallinnon edustajia, tutkijoita sekä muita aiheesta kiinnostuneita on kasvattanut ymmärrystä ihmiskeskeisemmän henkilödatan käsittelyn mahdollisuuksista. Aihetta on kartoitettu niin teknisesti, liiketoimintanäkökulmasta kuin ihmisten oikeuksienkin kannalta.
Personal web http://personalweb.me yritys	Personal Web (henkilökohtainen verkko) on lähestymistapa omien tietojen hallintaan. Se on kuten World Wide Web (maailmanlaajuinen verkko), mutta kunkin oma, henkilökohtainen ja yksityinen.
Pivo http://www.pivolompakko.fi yritys	Pivo on mobiililompakko. Se tallentaa tärkeät pankkitiedot, maksukorttit sekä suosikkikauppojen tiedot. (OP-ryhmä asiakkaille)
Taltioni http://taltioni.fi yritys	Taltioni-terveystili on kuin henkilökohtainen työkalupakki, jossa säilyvät terveyteen ja hyvinvointiin liittyvät käyttäjän tiedot.
Younited http://www.f-secure.com/fi/web/home_fi/younited yritys	Younited on tiedostojen hallintapalvelu. Esim. musiikki, valokuvat ja videot ovat käytettävissä puhelimessa, tietokoneella ja tabletissa.

Digital Health Revolution -ohjelma

Yksilöllisen terveys- ja hyvinvointitiedon hyödyntäminen kansalaisten terveyden ylläpitämiseksi ja sairauksien ennaltaehkäisemiseksi on maailmanlaajuinen suuntaus, jonka eturintamassa myös Suomi on vahvasti mukana. Monialaisen Digital Health Revolution -tutkimushankkeen visiona on mahdollistaa yksilöstä saatavan tiedon hyötykäyttö osana personoituja, ennaltaehkäiseviä palveluja sekä parantaa kansalaisen mahdollisuuksia omaehtoiseen hyvinvoinnin ylläpitoon.

Hankkeen lähtökohtana on, että yksilö voi hyödyntää omia tietojaan itse valitsemissaan käyttökohteissa sekä kontrolloida tietojensa käyttöä. Hankkeen tavoitteena on yhdistää systemaattisesti eri lähteistä saatavaa tietoa, mukaan lukien yksilön genomitieto, terveys- ja hyvinvointiseurantatieto sekä arkikäyttäytymiseen liittyvä tieto eli ihmisen digitaalinen jalanjälki. Tavoitteena on älykkäästi jalostaa yksilöllinen terveystieto yhteiskunnan ja yksilön hyödyksi uusina palveluratkaisuina.

Hankkeessa tullaan rakentamaan henkilötietoon perustuvia käyttäjälähtöisiä, ennaltaehkäiseviä palveluratkaisuja palvelupilottien avulla. Hankkeella pyritään perinpohjaiseen asenneilmapiirin muutokseen liittyen digitaalisen tiedon hyödyntämisen tapoihin. Siksi tekemisen ytimessä on mydata.fi-yhteisön rakentaminen, mikä samalla mahdollistaa uuden liiketoiminnan ja uusien palveluratkaisujen kehittämisen eettisiä ja lainmukaisia säännöksiä noudattaen.

Pitkällä aikajänteellä hankkeen tulokset mahdollistavat henkilötietoon perustuvien ennaltaehkäisevien palveluratkaisujen integroitumisen myös osaksi julkista terveydenhoitoa ja tulevaisuuden palvelurakennetta.

http://www.tekes.fi/nyt/uutiset-2014/tekes-rahoittaa-kolme-uutta-strategista-tutkimusavausta/

http://cht.oulu.fi/news/digital-health-revolution-was-launched-in-oulu

Lähteet

Binns, R. (2013). 5 Stars of Personal Data Access. Retrieved September 9, 2014, from http://www.reubenbinns.com/blog/5-stars-of-personal-data-access

Charski, M. (2013). More Retailers View E-Receipts as Customer Relationship Channel. Retrieved September 6, 2014, from http://data-informed.com/more-retailers-analyze-e-receipts-as-customer-relationship-channel

Davidov, B. (2014, February 20). Welcome to Algorithmic Prison. The Atlantic. Retrieved from http://www.theatlantic.com/technology/archive/2014/02/welcome-to-algorithmic-prison/283985

Eduskunta. (2014, April 2). Valtiopäivien avajaiset 4.2.2014 täysistunnon pöytäkirja. Retrieved February 20, 2014, from http://www.eduskunta.fi/faktatmp/utatmp/akxtmp/ptk_1b_2014_p.shtml

EU. (1995a, October 24). Eeuroopan parlamentin ja neuvoston direktiivi 95/46/EY - johdantokappaleen kohta 26. Retrieved from http://eur-lex.europa.eu/legal-content/FI/TXT/?uri=CELEX:31995L0046

EU. (2007). Lausunto 4/2007 henkilötietojen käsitteestä. Retrieved from http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_fi.pdf

EU. (2013a). Article 29 Data Protection Working Party - opinion on “Purpose limitation.” Retrieved September 7, 2014, from http://idpc.gov.mt/dbfile.aspx/Opinion3_2013.pdf

EU. (2013b, B). Euroopan parlamentin ja neuvoston direktiivi 2013/37/EU, julkisen sektorin hallussa olevien tietojen uudelleenkäytöstä annetun direktiivin 2003/98/EY muuttamisesta. Retrieved September 10, 2014, from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:175:0001:0008:FI:PDF

European Parliament. (2014, December 3). European Parliament legislative resolution of 12 March 2014 on the General Data Protection Regulation. Retrieved from http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT%20TA%20P7-TA-2014-0212%200%20DOC%20XML%20V0//en

FTC, 2014. (n.d.). FTC Recommends Congress Require the Data Broker Industry to be More Transparent and Give Consumers Greater Control Over Their Personal Information. Retrieved September 6, 2014, from http://www.ftc.gov/news-events/press-releases/2014/05/ftc-recommends-congress-require-data-broker-industry-be-more

Hill, K. (2012, February 16). How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did. Retrieved September 4, 2014, from http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/

HLT. (1999). Henkilötietolaki (1999/523). Retrieved from http://www.finlex.fi/fi/laki/ajantasa/1999/19990523#L2P8

Hoback, C. (2013). Terms and Conditions May Apply (2013) - IMDb. Retrieved from http://www.imdb.com/title/tt2084953/

LVM. (2014). Liikenne- ja viestintäministeriö - Big datan hyödyntäminen. Retrieved September 4, 2014, from http://www.lvm.fi/julkaisu/4417803/big-datan-hyodyntaminen

Newman, N. (2013, September 24). Taking on Google's Monopoly Means Regulating Its Control of User Data. Retrieved September 5, 2014, from http://www.huffingtonpost.com/nathan-newman/taking-on-googles-monopol_b_3980799.html

Pitkänen, O. (2014). Sinun tietosi eivät ole sinun: rekisteröidyn oikeus hyödyntää omia henkilötietojaan. Oikeus, (2/2014), 202–214.

Poikola, A., Hintikka, K. A., & Kola, P. (2010). Julkinen data - johdatus tietovarantojen avaamiseen. Retrieved from http://www.lvm.fi/julkaisu/1155483/julkinen-data-johdatus-tietovarantojen-avaamiseen

quantifiedself.fi. (2014). Quantified Self & Biohacking Finland | Itsetietoisuuteen numeroiden avulla. Retrieved September 4, 2014, from http://quantifiedself.fi/

Razkin, A. (2010). Privacy Icons: Alpha Release. Retrieved from http://www.azarask.in/blog/post/privacy-icons/

Sullivan, M. (2014, August 15). Guess what? Doctors don't care about your Fitbit data. Retrieved September 6, 2014, from http://venturebeat.com/2014/08/15/guess-what-doctors-dont-care-about-your-fitbit-data/

Taloussanomat. (2014, November 8). Ilmailuala hakee toivoa kanta-asiakasohjelmista. Retrieved September 4, 2014, from http://www.taloussanomat.fi/uutiset/2014/08/11/lentoyhtio-myi-tietoja-asiakkaistaan-teki-valtavan-tilin/201411102/12

Wall Street Journal. (2014, June 5). After Facebook Deal, Moves App Changes Privacy Policy. Retrieved September 4, 2014, from http://blogs.wsj.com/digits/2014/05/05/after-facebook-deal-moves-app-changes-privacy-policy/

White House. (2012). Consumer data privacy in a networkrd world. Retrieved September 9, 2014, from http://www.whitehouse.gov/sites/default/files/privacy-final.pdf

White house. (2014). Big Data: Seizing opportunities, preserving values. Retrieved September 4, 2014, from http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf

World Economic Forum. (2013). Unlocking the Value of Personal Data. Retrieved August 28, 2014, from http://www.weforum.org/reports/unlocking-value-personal-data-collection-usage

Ympäristöministeriö. (2014). Sähköisen asunto-osakerekisterin toimintamalli. Retrieved from http://www.ym.fi/download/noname/%7B76F44244-DE60-4B65-95F2-634D6A857096%7D/99026

6. Avoimen datan infrastruktuuri