

Erkki Mäkinen (toim.)

**Tietojenkäsittelytieteellisiä tutkielmia
Syksy 2016**



INFORMAATIOTIETEIDEN YKSIKKÖ
TAMPEREEN YLIOPISTO

INFORMAATIOTIETEIDEN YKSIKÖN RAPORTTEJA 51/2016

TAMPERE 2016

TAMPEREEN YLIOPISTO
INFORMAATIOTIETEIDEN YKSIKKÖ
INFORMAATIOTIETEIDEN YKSIKÖN RAPORTTEJA 51/2016
JOULUKUU 2016

Erkki Mäkinen (toim.)

**Tietojenkäsittelytieteellisiä tutkielmia
Syksy 2016**

INFORMAATIOTIETEIDEN YKSIKKÖ
33014 TAMPEREEN YLIOPISTO

ISBN 978-952-03-0348-8 (pdf)

ISSN-L 1799-8158
ISSN 1799-8158

Sisällysluettelo

Ville Hakola	
Big data ja jalkapallon pelianalyysi.....	1
Tuukka Hämynen	
An overview of common server virtualization solutions.....	24
Tuomas Karjanlahti	
Suomen valtion sähköisen hallinnon kehitys 2003–2016.....	42
Evita Laitinen	
Kasvojen symmetrian koneellisen mittauksen mahdollisuudet toispuoleisessa kasvohalvauksessa.....	66
Jarkko Saarinen	
K:n lähimmän naapurin menetelmän käyttökohteet.....	82
Aleksi Sillanmäki	
Big datan visualisoinnin ongelmat ja mahdollisuudet.....	99
Valtteri Ylisalo	
Digitalisaation haasteita sote- ja maakuntauudistuksessa.....	117

Big data ja jalkapallon pelianalyysi

Ville Hakola

Tiivistelmä.

Teknologian kehityksellä on ollut selvä vaikutus jalkapalloon: pelistä on tiedon avulla tullut kilpailullisempi ja erilaisten taktikoiden määrä ja monimuotoisuus on kasvanut. Sensoreiden, kameroiden ja algoritmien avulla kerätään valtava ja epäyhtenäinen määrä tietoa, jota tutkimalla analyttikot ja tutkijat pyrkivät ymmärtämään peliä. Tutkielmassa selvitetään jalkapallon pelianalyysin, tiedonkeräysmenetelmien ja tiedon automatisoidun analyysin nykytilaa ja tulevaisuuden suuntaa. Aineistona käytettiin aiheesta löytyvää kirjallisuutta.

Tutkielman tuloksena voidaan todeta big datan ja pelianalyysin yhdistämisen osoittavan paljon potentiaalia, mutta löydettävissä on myös useita ratkaisemattomia ongelmia. Tutkielmassa ehdotetaan pelianalyysille uutta määrittelmää, jolla pyritään vastaamaan big data -teknologioiden ja modernin jalkapallon asettamiin vaatimuksiin. Lisäksi esitetään pelianalyysiin tarkoitettua big data -järjestelmän yksinkertaistettu malli. Tutkimus pelianalyysista ja big data -teknologioista on tähän asti ollut pääosin tietojenkäsittelytieteilijöiden tekemää ja tulevaisuudessa tutkimukselta kaivataan monitieteellisyyttä ja asiantuntijoiden mukaanottoa.

Avainsanat ja -sanonnat: koneoppiminen, neuroverkot, analyysi, valmennus, puettava teknologia.

1. Johdanto

Nykyään jokaisesta huipputasolla pelattavasta ottelusta kerätään suuret määrät tietoa kentän laidalla olevan teknologian (kamerat ja anturit) ja pelaajilla olevan puettavan teknologian avulla. Yhdestä ottelusta voidaan tallentaa 3 100 000 paikkatietoa, joihin kuuluvat jokaisen pelaajan ja pallon liikkeet ottelun aikana [Memmert *et al.* 2016]. Paikkatietojen lisäksi jokaisen pelaajan fyysistä suorituskykyä on mahdollista mitata läpi ottelun [Rein and Memmert 2016]. Ottelun lisäksi tietoa kerätään harjoituksista, pelaajien vapaa-ajalta ja sosiaalisesta mediasta.

Pelistä kerättävän tiedon lisäksi pelin intensiteetti ja vaatimukset ovat nousseet [Barnes *et al.* 2014; Bush *et al.* 2015; Wallace *et al.* 2014]. Nykyään taktiikan onnistunut toteutus edellyttää pelaajien taktisten, fyysisten ja psyykkisten kykyjen riittävyttä. Joukkueen suorituksen analysointi on noussut yhä tärkeämpään asemaan.

Pelianalyysillä pyritään objektiivisesti ja luotettavasti löytämään tekijöitä, jotka vaikuttavat joukkueen suoritukseen. Pelianalyysia käytetään jalkapallossa oman ja vastustajan joukkueen tekniseen ja taktiseen arviointiin ja taktiikan laadintaan. [Häyrinen 2010]. Pelianalyysia käsitellään tarkemmin luvussa 2.

Valmentajille, analytiikoille ja medialle on nykyään tarjolla monia kaupallisia palveluita, kuten STATS, Statsports, OptaPro, Instat Football ja Squawka, jotka keräävät tietoa huippusarjoista. Luvussa 3 esitetään kaupallisten organisaatioiden, seurojen ja julkisten tutkimusinstituuttien käyttämiä tiedonkeruumenetelmiä.

Vaikka pelianalyysi on helpottunut tiedon määrän ja yksityiskohtaisuuden kasvaessa, on siitä tullut myös ongelma. Tiedon määrä tarkoittaa, ettei sitä ole mahdollista käydä läpi käsin, vaan tarvitaan automatisoitu tapa käsitellä tietoa. Koneoppiminen mahdollistaa suuren tietomäärän automaattisen käsittelyn. Pelianalyysin kannalta mielekkäitä koneoppimisen menetelmiä käsitellään luvussa 4.

Lopuksi luvussa 5 pohditaan big data -järjestelmän käyttöä pelianalyysissa. Luvussa esitetään malli big data -järjestelmälle ja pohditaan järjestelmän vaatimuksia. Lisäksi tarkastellaan yleisesti big datan käyttöön liittyviä ongelmia.

Tässä tutkielmassa keskitytään esittämään big dataa hyödyntäviä menetelmiä, joiden avulla on mahdollista löytää pelianalyysin kannalta relevanttia tietoa suuresta ja epäjohdonmukaisesta tietomäärästä. Lisäksi pyritään pohtimaan pelianalyysin nykytilaa ja tekemään katsaus tulevaisuuteen.

Tutkielmalla pyritään vastaamaan kysymyksiin kuten, miltä modernin jalkapallon pelianalyysi näyttää, mikä on big data -teknologioiden vaikutus jalkapalloon ja kuinka viimeaikaiset kehitykset big data -teknologioissa voivat ratkaista pelianalyysissa ilmeneviä ongelmia.

2. Pelianalyysi

Pelianalyysia on käsitelty hyvin vähän suomeksi, eikä ole täysin selvää, mitkä tekijät kuuluvat pelianalyysin piiriin. Etenkään tieteellistä tutkimusta aiheesta ei ole juurikaan löydettävissä. Tämän vuoksi tässä työssä pelianalyysin määrittämisessä on tukeuduttu pääosin englanninkieliseen kirjallisuuteen.

Vaikka englanninkielistä materiaalia onkin enemmän tarjolla, ei kuitenkaan ole löydettävissä yhtä vakiintunutta termiä pelianalyysia käsiteltäessä. Myös englanninkielinen tutkimus aiheesta on varsin uutta, kun ottaa huomioon jalkapallon historian [Garganta 2009; Rein and Memmert 2016; Sarmiento *et al.* 2014]. Seuraavia englanninkielisiä termejä kuitenkin käytetään jalkapalloon liittyvästä analyysista: "performance analysis" [Mackencie and Cushion 2013],

”tactical analysis”, ”notational analysis” [Rein and Memmert 2016] ja ”match analysis” [Sarmiento *et al.* 2014].

Koska ei ole löydettävissä yksimielisyyttä pelianalyysin määritelmästä, esitetään tässä luvussa pelianalyysille määritelmä, jonka viitekehyksenä toimii teknologian hyödyntäminen valmennuksessa. Määritelmällä pyritään vastaamaan big data -teknologioiden ja modernin jalkapallon asettamiin vaatimuksiin. Määritelmässä korostuu kokonaisvaltainen ja kontekstuaalinen pelianalyysi, jossa huomioidaan pelitapahtumien lisäksi historialliset ja pelin ulkopuoliset tiedot kuin myös pelaajan fysiologia, psykologia ja taktinen tietämys. Lisäksi pohditaan pelianalyysissa esiintyviä ongelmia sekä mahdollisia ratkaisuja.

2.1. Määritelmä

Pelianalyysi voidaan jakaa määrälliseen ja laadulliseen pelianalyysiin. Määrällinen analyysi pohjautuu täysin tilastomerkintöjen pohjalta tehtyyn analyysiin, kun taas laadullisessa analyysissä pyritään ymmärtämään asioiden merkityssuhteita, kuten pelin dynamiikkaa. [Juntunen 2011].

Tässä tutkielmassa läpikäytävät tietojenkäsittelytieteen menetelmät käyttävät molempia pelianalyysin haaroja; suuren tietomäärän avulla pyritään objektiivisesti ja luotettavasti ymmärtämään syy-seuraussuhteita pelin *tapahtumien* taustalla, kuitenkin tekemättä liian suuria yleistyksiä ja unohtamatta kontekstia [Rein and Memmert 2016]. Pelin tapahtumalla tarkoitetaan esimerkiksi pallon riistoa, syötönkatkoa tai maalin tekoa.

Tästä eteenpäin tutkielmassa puhutaan ainoastaan pelianalyysista, sillä sen jako käytetyn metodin mukaan ei ole tutkielman näkökulman kannalta relevanttia. Siitä huolimatta pelianalyysista voidaan valmennuksen kannalta löytää kaksi mielenkiinnon kohdetta: joukkueen taktinen analyysi ja yksittäisen pelaajan analyysi. Molempien tavoitteena on oman tai vastustajan joukkueen vahvuuksien ja heikkouksien tunnistaminen. [Rein and Memmert 2016; Luhtanen 2015.]

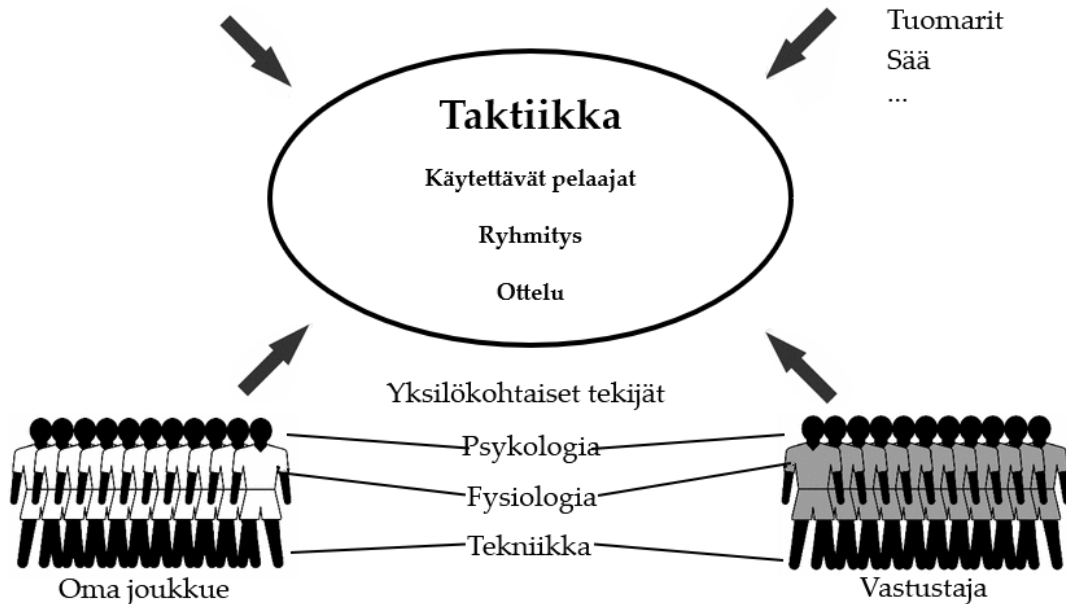
Taktisessa analyysissä (tactical analysis) tutkitaan joukkueen taktiikkaan vaikuttavia tekijöitä, joita on havainnollistettu kuvassa 1. Rein ja Memmert [2016] painottavat kontekstin huomioimista taktisessa analyysissä. Heidän mukaan taktisia päätöksiä tehtäessä tulee huomioida oman ja vastustajan joukkueen pelaajien lisäksi myös peliin vaikuttavat ulkoiset tekijät ja historiallinen tieto.

Historiallinen tieto

Kuntopuntari
Aiemmat ottelut vastustajan kanssa
Aiemmat ottelut muita vastaan

Ulkoiset tekijät

Koti vai vieras
Vastustaja
Sarjasijoitus
Tuomarit
Sää
...



Kuva 1. Taktiikan valintaan vaikuttavia tekijöitä [Rein and Memmert 2016].

Pelaaja-analyysi (player performance analysis) on osa taktista analyysia, sillä joukkueen taktinen onnistuminen on kentällä olevien pelaajien suoritusten summa, mutta sitä voidaan käyttää myös pelaajan henkilökohtaisessa valmennuksessa ja pelaajahankintojen tukena. Myös pelaaja-analyysissa tarkastellaan useita suoritukseen vaikuttavia muuttujia, kuten pelaajan fyysisiä, psyykkisiä, teknisiä ja taktisia ominaisuuksia. [Rein and Memmert 2016; Luhtanen 2015.]

Pelianalyysissa analysoidaan tekijöitä, jotka johtavat joukkueen tai pelaajan suoritukseen. Analysoitavat tekijät liittyvät taktiikkaan, pelaajaan ja ottelutapahtumiin. Kysymyksiä, joihin pelianalyysi etsii ratkaisua, ovat esimerkiksi: mitkä tekijät ovat johtaneet Atlético Madridin menestykseen, mikä ryhmitys antaa joukkueelle parhaimman mahdollisuuden voittaa, kun vastassa on puolustusorientoitunut joukkue tai mitkä ovat seuraavan vastustajan heikkoudet ja kuinka oman joukkueen pelaajat pystyvät niitä hyödyntämään.

Pelianalyysilla pyritään siis ymmärtämään syitä pelin kulkua muuttavien tapahtumien takana ja ennustamaan tulevaa käyttäytymistä pelaajien yksittäisistä suorituksista aina joukkueen kollektiivisiin tekoihin, tutkimalla mikro- ja makrotason muuttujia.

2.2. Pelianalyysin ongelmat

Perinteisesti pelianalyysi on levännyt muutaman *avainindikaattorin* varassa, kuten syöttöprosentin, pallonhallinnan ja maalintekotilanteiden. Perinteisen analyysimenetelmän heikkoutena voidaan pitää kontekstin katoamista: menetelmän on osoitettu olevan riittämätön selitysvoimaltaan, eivätkä valmentajat ole ottaneet sitä hyvin vastaan. [Mackenzie and Cushion 2013; Sarmiento *et al.* 2014.]

Valmentajien ja analyytikkojen käyttämistä menetelmistä ei ole tehty juuriakaan tutkimusta. Mackenzie ja Cushion [2013] toteavat, että useimmiten pelianalyysit perustuvat asiantuntijan (valmentaja tai pelaajatarkkailija) subjektiiviseen näkemykseen ottelusta. Toisaalta 70 % valioliigajoukkueesta käyttää OptaProta [Optasportspro.com]. Myös useassa ei-tieteellisessä artikkelissa viitataan englantilaisten seurojen käyttävän data-analyytikkoja pelianalyysin teossa [Computer-weekly.com; Forbes.com; Theguardian.com]. Maajoukkueista Saksan ja Norjan joukkueet ovat ottaneet big datan osaksi toimintaansa [Cnbc.com; Johansen *et al.* 2013; 2015; Telekom.com]. Lisäksi suomalaisista joukkueista 16 käyttää Instatia [Instatfootball.com].

On siis selvää, että seuroissa ollaan tietoisia menetelmistä ja tietoa käytetään ainakin jollain tasolla. Hämärän peittoon kuitenkin jää, millä tasolla ja kuinka tärkeää roolia tiedon käyttö esittää. Pelianalyysin tutkimuksesta sen sijaan löytyy informaatiota ja seuraavissa kappaleissa puretaan niiden sisältöä.

Niin kompleksisen pelin kuin jalkapallon selittäminen ainoastaan muutamalla muuttujalla johtaa vääjäämättä virheelliseen analyysiin. Siitä huolimatta sekä Mackenzie ja Cushion [2013] että Sarmiento ja muut [2014] raportoivat pelianalyysin tutkimusten jättävän usein huomioimatta kontekstin, jossa peli on pelattu. Esimerkiksi 81 % tutkituista artikkeleista ei ottanut huomioon vastustajaa. Lisäksi 70 % tutkimuksista ei huomionnut, onko peli koti- vai vierasottelu [Mackenzie and Cushion 2013].

Valaistaan muutamaan muuttujaan pohjautuvan analyysin heikkoutta ja kontekstin merkitystä esimerkillä. MM-kisoissa 2014 Saksa voitti Brasilian 7-1. Siitä huolimatta Brasilia oli parempi lähes kaikilla suorituskykyä mittaavilla avainindikaattoreilla. Brasilia voitti muun muassa pallonhallinnan, laukaukset, maalia kohti menevät laukaukset, vaaralliset hyökkäykset, pallonriistot, kulmapotkut, onnistuneet syötöt viimeisellä kolmanneksella ja keskitykset. Saksa oli kuitenkin parempi yhdessä tärkeässä avainindikaattorissa, joka on laukauksien onnistumisaste, jolla kuvataan, kuinka moni laukauksista menee maaliin. [Fifa.com.] Laukauksien onnistumisaste ei kuitenkaan selitä, miksi Saksa oli parempi ottelussa, vaan sen selvittämiseksi tarvitaan kontekstuaalista pelianalyysia.

Saksan voittoa voidaan analysoida *odotetuilla maaleilla* (expected goals). Odotettu maali kuvaa laukauksen todennäköisyyttä johtaa maaliin [Lucey *et al.* 2014]. Todennäköisyys lasketaan pelaajien ja pallon paikkatiedoista ottamalla huomioon konteksti, jossa laukaus tapahtuu: puolustajan etäisyys, ympäröivien pelaajien liikkeet, pelinopeus ja laukovan pelaajan sijainti kentällä vaikuttavat maalinteon todennäköisyyteen.

Ainoastaan ottelutapahtumiin keskittyminen suoritusta analysoitaessa on myös ongelmallista. Pelaajan huonoon suoritukseen tai koko joukkueen suoritustason äkkinäiseen laskuun voi löytyä syy pelikentän ulkopuolelta: esimerkiksi univaje ja jatkuvasti muuttuva unirytmii voivat johtaa pelaajan suorituskyvyn ja vastustuskyvyn heikkenemiseen [Reilly and Edwards 2007]. Pelaajien unirytmii ovat usein muuttuvia ja etenkin peliä edeltävä yö usein nukutaan huonosti [Nédélec *et al.* 2015; Julif *et al.* 2015]. Pelisuoritusta analysoitaessa ei tule keskittyä ainoastaan harjoituksista ja otteluista kerättyihin tietoihin.

Pelin ulkopuoliset tekijät korostuvat varsinkin pienemmissä joukkueissa, joilla ei ole resursseja seurata pelaajien toimintaa harjoitusten ja otteluiden ulkopuolella. Tämä ei koske maailman huippusarjoja, mutta esimerkiksi Veikkausliigassa budjetit ovat usein niin pieniä, etteivät kaikki pelaajat ole ammattilaisia. Tämä tarkoittaa, että pelaajien kuormitus harjoitusten ulkopuolella saattaa vaihdella paljon. Huippusarjoissa taas joukkueet eivät välttämättä ole tietoisia tai kiinnostuneita pelin ulkopuolella tapahtuvista asioista. Esimerkiksi Britanniassa valmentajalla on epätavallisen suuri vaikutusvalta seuran toimintaan ja valmentajien koulutuksessa on usein puutteita [Kelly 2008], mikä voi vaikuttaa siihen, kuinka tärkeänä seura näkee pelianalyysin ja pelaajien vapaa-ajan kuormituksen seurannan.

Pelianalyysin tulisi kuitenkin pohjautua riittävän monipuoliseen ja tarkkaan tietoon, johon lukeutuu niin pelikentältä kuin sen ulkopuoleltakin kerätty tieto [Mackenzie and Cushion 2013; Sarmiento *et al.* 2014.]. Teknologian kehitys mahdollistaa tarkan tiedon keräämisen pelin sisältä ja sen ulkopuolelta ja näin kontekstin säilyttämiseen pelianalyysissa.

3. Tiedon keräys

Tässä luvussa tarkastellaan eri tapoja kerätä tietoa jalkapallo-ottelusta tai harjoituksesta. Kohdissa 3.1 ja 3.2 tarkastellaan pääosin *määrällisen tiedon* keräystä ja kohdassa 3.3 *laadullisen tiedon* keräystä.

3.1. Videoteknologia

Videoteknologialla tarkoitetaan *tietojärjestelmää*, johon kuuluu videokamera ja videokuva prosessoiva ohjelmisto. Järjestelmät sisältävät yleensä useampia

kameroita ja usein järjestelmän käyttö vaatii myös käyttäjän. Videoteknologiaa käytetään pääosin *tilastotietojen* ja *paikkatietojen* (tracking data) keräykseen. Paikkatietoja tallennetaan pelaajista, pallosta ja tuomareista. Videoteknologiaa voidaan käyttää sekä ottelun että harjoituksen aikana. Videoteknologiaa käyttävät järjestelmät ovat yleensä puoliautomaattisia: pelaajien seuraukset toimii automaattisesti, mutta pelaajat täytyy tunnistaa manuaalisesti [Castellano *et al.* 2014].

Numeerisen tiedon lisäksi video tarjoaa luonnollisesti visuaalisen esityksen suorituksesta. Visuaalista tietoa pelistä käytetään usein havainnollistamaan pelin kulkua muuttavia tapahtumia tekemällä *videoyhteenveto*. Videoyhteenveto voi sisältää esimerkiksi kuvaa joukkueen epätasapainoisesta ryhmyksestä, joka edeltää vastustajan tekemää maalia. Videoyhteenvedon tekoa on mahdollista automatisoida, ja sitä on sovellettu niin valmennuksessa kuin televisiolähetyksissä. [Johansen *et al.* 2013; Rehman and Saba 2014.]

Kaikkien pelaajien paikkatietojen keräykseen vaaditaan videokuvaa, jossa näkyy koko kenttä. Vaikka pelaajan seuranta on mahdollista myös tv-kuvasta [Lu *et al.* 2013], menetelmä ei kerää luotettavasti paikkatietoja kaikilta pelaajilta, minkä vuoksi käytetään useampia kameroita keräyksen tarkentamiseksi [Memmert *et al.* 2016; Rehman and Saba 2014].

Videosta voidaan tallentaa useita miljoonia yksittäisiä paikkatietoja. Paikkatietojen määrä riippuu käytettävän videon kuvanopeudesta. Kuvanopeudesta riippuen pelaajan paikkatiedot tallennetaan 10-120 kertaa sekunnissa. Paikkatiedot sisältävät x- ja y-koordinaatit pelaajan kulloisestakin paikasta kentällä. Paikkatiedot itsessään kertovat hyvin vähän pelistä ja niiden analysointi vaatii jatkoprosessointia. Prosessointiin soveltuvia menetelmiä tarkastellaan luvuissa 4 ja 5.

Pelaajan *seuranta-algoritmit* (tracking algorithm) mahdollistavat tarkan, mutta eivät täysin virheettömän, tiedon keräyksen. Nykyiset seuranta-algoritmit eivät pysty erottamaan, onko pelaaja vastustajan etu- vai takapuolella kontaktitilanteessa [Lu *et al.* 2013]. Myös valaistus saattaa synnyttää ongelmia: valaistus voi vaihtua pelin aikana moneen kertaan, katsomot saattavat luoda varjoja ja illalla pelatuissa otteluissa jokaisen stadionin valaistus on toteutettu eri tavalla. Valaistuksen epäjohdonmukaisuus saattaa johtaa pelaajan kadottamiseen tai virheelliseen tunnistukseen. [D’Orazio and Leo 2010; Pettersen *et al.* 2014.]

Pelaajan seurantaan liittyvät ongelmat tarkoittavat, ettei paikkatietojen perusteella voida tehdä pelianalyysia pelin kaikista osa-alueista. Pelianalyysit, jotka koskevat pelaajan tarkkaa sijaintia kentällä joudutaan tekemään vielä manuaalisesti. Tämän päivän teknologialla ja koneoppimisen menetelmillä paikkatiedoilla voidaan kuitenkin tehdä mielekästä pelianalyysia.

3.2. Puettava teknologia

Videoteknologian lisäksi pelaajasta on mahdollista kerätä tietoa puettavan teknologian avulla. Tässä kohdassa tarkastellaan pelianalyysin kannalta mielekkäitä puettavan teknologian sovelluksia. Lisäksi pohditaan teknologioiden toimivuutta tiedonkeruumenetelmänä.

Puettavalla teknologialla (wearable technology) viitataan pelaajan mukana kulkeviin laitteisiin ja sensoreihin, jotka mittaavat fyysikaalista tai kemiallista suureita. Mitattavia suureita ovat muun muassa syke, kiihtyvyys, hapenotto-kyky ja juostu matka. Puettava teknologia mahdollistaa pelaajan fyysisen suorituskyvyn seurannan reaaliaikaisesti ottelun tai harjoituksen aikana [Johansen *et al.* 2013.]. Teknologia mahdollistaa siis paikkatietojen lisäksi fysiologisen tiedon keräyksen urheilijasta.

Puettavan teknologian kaupallisia sovelluksia on nykypäivänä useita: älykellot, aktiivisuusrannekkeet, sykevyöt, hyvinvointisormukset ja monet muut hyvinvointitekniikan sovellukset mahdollistavat tiedon keräyksen urheilijasta sekä suorituksen aikana että vapaa-ajalla.

Suorituksen aikana käytettyjä sensoreita ovat gps, kiihtyvyysanturi, gyrokooppi, digitaalinen kompassi ja sykesensori. Lisäksi tiedon lähettämiseen vaaditaan lähetin, joka voi olla radio- tai bluetooth-lähetin tai jokin muu vastaava. Edellä mainitut sensorit voidaan rakentaa vyön tai liivin sisälle, jolloin tiedonkeruu voidaan suorittaa *huomaamattomasti*. [Johansen *et al.* 2013.] Huomaamattomalla tiedonkeruulla viitataan siihen, ettei käytetty laite vaikuta urheilijan suoritukseen.

Satelliittipaikannukseen perustuvat gps-laitteet ovat yleensä epätarkempia kuin paikalliset, kentän ympärille rakennettavat, seurantajärjestelmät. Toisaalta gps-laitteet ovat halvempia kuin paikallisesti rakennetut järjestelmät. [Johansen *et al.* 2013.] Kumpikin menetelmä kärsii kuitenkin samasta ongelmasta kuin videoteknologia: paikkatietojen tarkkuus ei riitä määrittämään pelaajan tarkkaa sijaintia kontaktitilanteessa.

Mahdollista on myös urheilijan aktiivisuuden seuranta vapaa-ajalla. Unta, askeleita ja palautumista voidaan seurata automaattisesti [Johansen *et al.* 2015]. Palautumisesta kertova tieto voi olla valmentajille hyödyllistä harjoitusten ja kokoonpanojen laatimisessa: Brink ja muut [2010] toteavat, että pelaajan ylikuormitus ja loukkaantumisriksi ovat toisistaan riippuvia.

Zhou ja muut [2016] ovat osoittaneet, että potkutekniikasta on mahdollista kerätä tietoa kenkään istutettavan sensorin avulla. Myös askelluksesta on mahdollista kerätä tietoa puettavan teknologian avulla [Shull *et al.* 2014]. Potkutekniikan ja askelluksen tutkimus on toteutettu vielä täysin laboratorioolosuhteissa, mikä tarkoittaa, että niiden luotettavuutta käytännössä tulee tut-

kia lisää ennen käyttöönottoa. Ensimmäiset tulokset ovat kuitenkin lupaavia. Tulevaisuudessa tiedot askelluksesta ja potkutekniikasta voivat auttaa pelaajan henkilökohtaisessa valmennuksessa ja loukkaantumisten ennaltaehkäisyssä. Hewett ja muut [2013] toteavat, että eturistisideleikkauksen jälkeen urheilijoille usein jää biomekaanisia puutteita, jotka lisäävät riskiä loukkaantua uudelleen. Askelluksesta kerättävä tieto voi auttaa tunnistamaan näitä puutteita.

Myös päähän kohdistuvista iskuista on mahdollista kerätä tietoa kiihtyvyyssantureiden avulla. Antureilla kerättävä tieto on kuitenkin vielä epäluotettavaa, eikä sovellu analyysien tekoon [Nevins *et al.* 2016; Siegmund *et al.* 2016]. Tarvitaan luotettavampia sensoreita ja urheilukohtaisia algoritmeja laskemaan iskun voimakkuus, jotta tietoa voidaan käyttää analyyseissa [Gastin *et al.* 2014].

Puettavan teknologian avulla kerätyt tiedot suorituksesta ja vapaa-ajalta voivat auttaa valmentajia, fysioterapeutteja ja analyytikkoja ottamaan huomioon pelaajan kuormituksen harjoituksia suunniteltaessa. Kuormitustasoon suhteutetut harjoitukset saattavat auttaa loukkaantumisten ja sairauksien ennaltaehkäisyssä, mikä tarkoittaa, että seurat ja pelaajat voivat hyötyä puettavan teknologian käytöstä. [Johansen *et al.* 2015; Ehrmann *et al.* 2016.]

3.3. Laadullisen tiedon keräys

Tähän asti on tarkasteltu lähinnä määrällisen tiedon keräystä, jolla voidaan tuoda objektiivisuutta analyysien tekoon. Myös laadullisella tiedolla, kuten subjektiivisilla arvioilla ja pelaajan itse tuottamalla tiedolla, voidaan tehdä mielekästä analyysia [Johansen *et al.* 2015; Rein and Memmert 2016; Xu *et al.* 2015]. Laadullisia tiedon lähteitä voivat olla asiantuntijan arviot, internet-blogit [Rein and Memmert 2016], sosiaalinen media [Xu *et al.* 2015] ja pelaajien päiväkirjat [Johansen *et al.* 2015].

Tiedonlounhinta, koneoppiminen ja tiedon algoritminen käsittely mahdollistavat relevanttien analyysien tekemisen urheilusuoritukseen epäsuorasti liittyvästä tiedosta. Xu ja muut [2015] löysivät NBA koripalloilijoiden ennen peliä tehtyjen Twitter-julkaisujen kautta välittyvän tunnetilan korreloivan pelaajan suorituksen kanssa. Tunnetila muodostettiin algoritmisesti Twitter-julkaisuista löytyvien sanojen positiivisuuden ja negatiivisuuden perusteella. Sosiaalisen median käyttö julkaisusta välittyvän tunnetilan analysointiin ei onnistu jalkapallossa, koska pelaajat eivät usein hallinnoi omia tilejään. Xu ja muut [2015] osoittavat kuitenkin laadullisen tiedon hyödyllisyyden analyyseissa.

4. Koneoppiminen

Tässä luvussa esitetään muutama pelianalyysin kannalta mielekäs koneoppimisen menetelmä. Tarkoituksena ei ole kattaa kaikkia pelianalyysiin sovellettavia

menetelmiä, eikä selittää niiden toimintaa seikkaperäisesti, vaan antaa kuva, minkälaisia analyyseja ne mahdollistavat.

Suuri osa tutkimuksista, jotka koskevat koneoppimisen menetelmillä toteutettua pelianalyysia, ovat tietojenkäsittelytieteilijöiden tekemiä. Tutkijoiden jalkapallotietämyksen puutteet näkyvät tutkimuksissa usein virheellisinä päätelminä tai rajauksina. Puutteista huolimatta osa tutkimuksista tarjoaa mielekkäitä sovelluksia pelianalyysin tekoon.

Kohdassa 4.1 käsitellään koneoppimista yleisesti ja kohdassa 4.2 esitellään lyhyesti analyyseja, joita koneoppimisen menetelmät mahdollistavat. Lisäksi kohdissa 4.2 ja 4.3 käsitellään tarkemmin neuroverkkojen käyttöä pelianalyysissa. Neuroverkkoihin perehdytään tarkemmin niistä tehtyjen tutkimusten ansiokkuuden ja määrän perusteella.

4.1. Koneoppiminen yleisesti

Koneoppiminen (machine learning) on tietojenkäsittelytieteen osa-alue, jossa keskitytään tiedon avulla tehtävään algoritmiseen oppimiseen, piirteiden tunnistukseen ja ennustusten tekemiseen. Koneoppimisella voidaan rakentaa apriorisia malleja suuresta tietojoukosta. Näin voidaan tunnistaa automaattisesti piirteitä esimerkiksi pelaajien paikkatiedoista. Koneoppimisella pystytään sekä luomaan uutta tietoa että analysoimaan jo olemassa olevaa. [Rein and Memmert 2016.]

Koneoppimisalgoritmit voidaan jakaa oppimistavan mukaan neljään eri kategoriaan: ohjattuun oppimiseen, ohjaamattomaan oppimiseen, puoli-ohjattuun oppimiseen ja vahvistusoppimiseen. Näitä kaikkia oppimistapoja esiintyy pelianalyysissa käytetyissä koneoppimisen menetelmissä.

Ohjatussa oppimisessa (supervised learning) tieto annetaan algoritmillemme syötetulospareina [Sammut and Webb]. Jokainen syöte on siis luokiteltu jo valmiiksi. Algoritmin tehtävä on ainoastaan löytää malli, jolla tietoa luokitellaan. Ohjatun oppimisen menetelmiä käytetään muun muassa sähköpostin roskapostisuodattimissa, kasvontunnistuksessa ja lääketieteessä oireiden diagnosoinnissa [Jordan and Mitchell 2015].

Ohjaamattomassa oppimisessa (unsupervised learning) algoritmi muodostaa mallin tiedon luokitteluun itsenäisesti ilman annettuja tulosluokkia [Sammut and Webb]. Haluttuja tulosluokkia ei siis tiedetä etukäteen kuten ohjatussa oppimisessa. Suuri osa jalkapallon pelianalyysiin käytettävistä menetelmistä kuuluvat ohjaamattoman oppimisen kategoriaan.

Puoli-ohjatussa oppimisessa (semi-supervised learning) algoritmillemme annetaan sekä luokiteltuja että luokittelemattomia syötteitä. Puoli-ohjattua oppimista voidaan käyttää kuvien, puheen ja tekstin luokitteluun. Puoli-ohjattua oppimista

on hyödynnetty myös kognitiivisessa psykologiassa ihmisen oppimisen laskennallisessa mallinnuksessa. [Sammut and Webb.]

Vahvistusoppimisessa (reinforcement learning) keinotekoinen agentti pyrkii käyttäytymään optimaalisesti tuntemattomassa ympäristössä negatiivisen ja positiivisen palautteen avulla. [Sammut and Webb.] Algoritmi voi esimerkiksi oppia pelaamaan tietokonepelejä vahvistusoppimisen avulla. Algoritmi oppii tunnistamaan optimaalisen pelistrategian kussakin pelitilanteessa pelistä saadun palautteen (voitto tai häviö) perusteella. Vahvistusoppimista käytetään peliteoriassa [Tuyls and Now 2005], robotiikassa ja robottijalkapallossa [Kostiadis and Hu 2001].

4.2. Koneoppimisen soveltaminen pelianalyysiin

Tässä kohdassa tarkastellaan muutamaa koneoppimisen menetelmää osittaisen pelianalyysin tekemiseen ja selostetaan menetelmien toimintaa lyhyesti.

Knauf ja muut [2016] osoittavat, kuinka paikkatietoja on mahdollista analysoida ohjaamattomasti *konvoluutioytimillä* (convolution kernels). Menetelmä tunnistaa pelinaloituksia ja maalintekotilanteita paikkatiedoista. Menetelmä mahdollistaa joukkueen pelinrakennuksen analysoinnin. Näin voidaan tutkia, kuinka joukkue pyrkii pääsemään viimeiselle kolmannekselle ja luomaan maalintekotilanteita. Tulevaisuudessa algoritmin kehittäminen niin, että se tunnistaisi onnistuneet ja epäonnistuneet pelinaloitukset, voisi helpottaa taktiikan analysoinnissa.

EM-algoritmia (expectation-maximization algorithm) eli odotusarvon maksimointi -algoritmia on onnistuneesti sovellettu pelianalyysiin. EM-algoritmeilla on mahdollista tunnistaa ohjaamattomasti joukkueen ryhmitys ottelun aikana paikkatiedoista [Bialkowski *et al* 2014]. Ryhmityksen tunnistuksessa on ollut usein ongelmana se, että pelaajan rooli määritellään ainoastaan kerran ottelun alkaessa, vaikka rooli saattaa muuttua ottelun aikana. Bialkowski ja muut [2014] ratkaisivat ongelman päivittämällä pelaajan roolia jatkuvasti ottelun aikana. Näin esimerkiksi laitahyökkääjien paikkojen vaihto ei aiheuta virheellistä ryhmityksen tunnistusta.

EM-algoritmeilla on osoitettu, että joukkueiden ryhmitys on vierasotteluissa puolustusorientoituneempi kuin kotiotteluissa [Bialkowski *et al.* 2014]. Samaa menetelmää käyttäen saattaa olla mahdollista analysoida joukkueen ryhmitystä tietyn tyyppistä vastustajaa vastaan. Esimerkiksi voi olla mahdollista tutkia, kuinka puolustusorientoitunut joukkue ryhmityy toista puolustusorientoitunutta joukkuetta vastaan. Tämä tieto voi auttaa valmentajaa taktisissa päätöksissä. Tulevaisuudessa mielekkäitä tutkittavia asioita voisi olla, mitkä ryhmitykset johtavat useimmin voittoon tai vaikuttaako ryhmityksen onnistumisas- teeseen vastustajan pelityyli.

Montoliu ja muut [2015] käyttivät *Bag-of-Words* algoritmia jalkapallo-ottelun videon koodaukseen, jonka jälkeen he luokittelivat koodattuja pelitilanteita käyttäen *satunnainen metsä* (random forest) -menetelmää. Analyysi pohjautuu täysin videosta saatuun tietoon, eikä siinä käytetä apuna paikkatietoja. Menetelmä tunnistaa ottelusta seuraavat kolme tilannetta: pallonhallinnan, nopean hyökkäyksen ja erikoistilanteen. Menetelmällä voidaan tutkia pelaajien liikkeitä aiemmin mainituissa tilanteissa.

Montoliun ja muiden [2015] lähestymistavasta tekee ainutlaatuisen valmentajien konsultointi. Tutkijat pyysivät apua ammattilaisvalmentajalta jalkapallokentän tärkeimpien alueiden tunnistukseen. Menetelmällä tehdyt analyysit eivät kuitenkaan tarjoa mitään uutta pelianalyysiin, vaan nopeuttavat ainoastaan nykyisiä analyysejä. Lisäksi *Bag-of-Words* algoritmin käyttö vaatii, että videoihin on etukäteen manuaalisesti lisätty selitteitä.

4.3. Neuroverkot

Neuroverkolla (neural network) tarkoitetaan keinotekoisesti luotua verkkoa, joka on abstrakti mallinnus hermoverkosta. Neuroverkoissa esiintyy kaikkia neljää aiemmin esitettyä oppimistapaa. Tässä kohdassa tarkastellut neuroverkot ovat kuitenkin oppimistavaltaan ohjaamattomia. Neuroverkolla voidaan tunnistaa tiedosta relevantteja piirteitä, luoda piirteiden välille merkityksellisiä suhteita ja luokitella piirteet. [Memmert *et al.* 2016.] Tästä syystä neuroverkot soveltuvatkin hyvin monimutkaisen ja rakenteeltaan monimuotoisen tiedon käsittelyyn, kuten jalkapallo-ottelusta kerättyjen paikkatietojen.

Neuroverkoista on löydettävissä kaksi onnistuneesti pelianalyysiin sovellettua mallia: itseorganisoituva kartta [Kohonen 1990] ja dynaamisesti ohjattu verkko [Perl 2002].

Itseorganisoituvassa kartassa (self-organizing map) neuronit luokitellaan niiden välisten suhteiden perusteella klustereiksi, jotka muodostavat kaksiulotteisen kartan [Kohonen 1990]. Itseorganisoituva kartta oppii ohjaamattomasti, kunnes verkossa ei enää tapahdu uusia muutoksia, toisin sanoen verkko on vakiintunut, eikä sitä voida enää muuttaa [Grunz *et al.* 2011]. Itseorganisoituvan kartan heikkous on juuri sen vakiintuminen, joka tekee siitä pitkällä aikavälillä huonosti sovellettavan alati muuttuvan lajin analyysiin. [Memmert *et al.* 2016]

Dynaamisesti ohjattu verkko (dynamically controlled network) pohjautuu itseorganisoituvaan karttaan, ja se mahdollistaa jatkuvan oppimisen. Jatkuvan oppimisen mahdollistaa jokaiseen neuronin sisällytetty ohjausyksikkö, joka kontrolloi neuronissa olevaa muistia. [Perl 2002.] Itseorganisoituvan kartan neuronit ovat siis ainoastaan muistiyksiköitä, jotka järjestetään neuroverkkoa

ohjaavalla algoritmilla, kun taas dynaamisesti ohjatun verkon neuronit pystyvät itsenäiseen ja jatkuvaan oppimiseen.

4.4. Neuroverkon käyttö pelianalyysissa

Neuroverkolle syötettävät paikkatiedot saadaan usein suoraan videosta, mikä tarkoittaa, että jokaisesta pelaajasta saadaan videon kuvanopeudesta riippuen 10-120 x- ja y-koordinaattia sekunnissa [Memmert *et al.* 2016].

Jotta paikkatiedoista saataisiin pelianalyysin kannalta relevanttia tietoa, Grunz ja muut [2011] ehdottavat joukkueen jakamista pelipaikan mukaan hyökkääviin ja puolustaviin pelaajiin. Myös Memmert ja muut [2016] ehdottavat samaa jakoa. Joukkueen jakaminen ainoastaan puolustaviin ja hyökkääviin pelaajiin on kuitenkin hieman ongelmallista. Joukkueen taktiikkaan voi esimerkiksi kuulua, että laitahyökkääjät kaventavat keskelle, mikä jättää laitapuolustajille tilaa nousta hyökkäykseen mukaan. Kyseiselle taktiikalle on ominaista, että laitahyökkääjä pyrkii antamaan maaliin johtavan syötön keskityksenä. Esimerkkitilanteessa ei olekaan enää selvää, onko laitapuolustaja tilanteessa puolustava vai hyökkäävä pelaaja.

Ongelmallisesta jaosta huolimatta Grunz ja muut [2011] löysivät 84 % puolustus- ja hyökkäysryhmyksistä, kun neuroverkon löytämiä muotoja verrattiin asiantuntijan tekemiin. On kuitenkin huomattava, että puolustus- ja hyökkäyspelaajat kattoivat ainoastaan kahdeksan pelaajaa joukkueesta, koska molemmat ryhmät sisälsivät ainoastaan neljä pelaajaa [Grunz *et al.* 2011]. Näin ollen joukkueen maalivahtia ja keskikenttäpelaajia ei huomioitu ollenkaan.

Memmert ja muut [2016] jatkavat Grunzin ja muiden [2011] työtä sisällyttämällä myös keskikenttäpelaajat hyökkääviin ja puolustaviin pelaajiin. Näin saadaan kattavampi kuva joukkueen puolustus- ja hyökkäysryhmyksestä. Joukkueen ryhmitystä analysoimalla voi olla mahdollista määrittää, kuinka hyvin ennen ottelua sovittu taktinen ryhmitys on kussakin tilanteessa toiminut. Memmert ja muut [2016] eivät maininneet otetaanko pelaajan roolin muutos ottelun aikana huomioon vai johtaako tapahtuma virheelliseen ryhmityksen tunnistukseen.

Kun molempien joukkueiden hyökkäys- ja puolustusryhmitykset on määritetty ottelun jokaisessa tilanteessa, voidaan niiden ja pelin tapahtumien avulla selvittää *voittavat ja häviävät ryhmitykset*. Voittavalla ryhmityksellä viitataan ryhmitykseen, joka edeltää joukkueen kannalta suotuisaa tapahtumaa, kuten pallonriistoa tai maalintekotilannetta. Häviävä ryhmitys puolestaan on voittavan vastakohta. [Memmert *et al.* 2016.]

Voittavia ja häviäviä ryhmityksiä tunnistamalla voi olla mahdollista löytää joukkueen taktisia heikkouksia ja vahvuuksia, joita ei välttämättä asiantuntijaakaan huomaa. Toisaalta, vaikka neuroverko tekeekin suurimman osan työstä,

tarvitaan asiantuntijaa tekemään lopullinen arvio analyysista. Asiantuntijoiden mukaanotto neuroverkkojen suunnitteluprosessiin voisi olla hyödyksi. Tarkastelluissa tutkimuksissa ei ollut mainintaa asiantuntijoiden konsultoinnista neuroverkkoja suunniteltaessa. Tulevaisuudessa tutkimuksen tulisikin olla poikkeusteellisempää urheilu- ja tietojenkäsittelytieteiden välillä.

5. Big data

Tässä luvussa käsitellään big dataa ja kuinka pelianalyysiin soveltuva big data -järjestelmä on mahdollista toteuttaa. Big data -järjestelmää käsitellään pelianalyysin näkökulmasta ja pohditaan ongelmia, jotka liittyvät big datan käyttöön.

5.1. Big data yleisesti

Big datalle ei ole löydettävissä yhtä yleistä määritelmää, mutta sillä usein viitataan tietojoukkoon, josta on löydettävissä seuraavat kolme piirrettä: *volyymi*, *heterogeenisyys* ja *nopeus*. Volyymilla kuvataan tietojoukon suuruusluokkaa, heterogeenisyydellä tiedon monimuotoisuutta ja nopeudella tiedon tallennusnopeutta. Jalkapallosta kerättävä tietojoukko sopii tähän määritelmään. [Rein and Memmert 2016.]

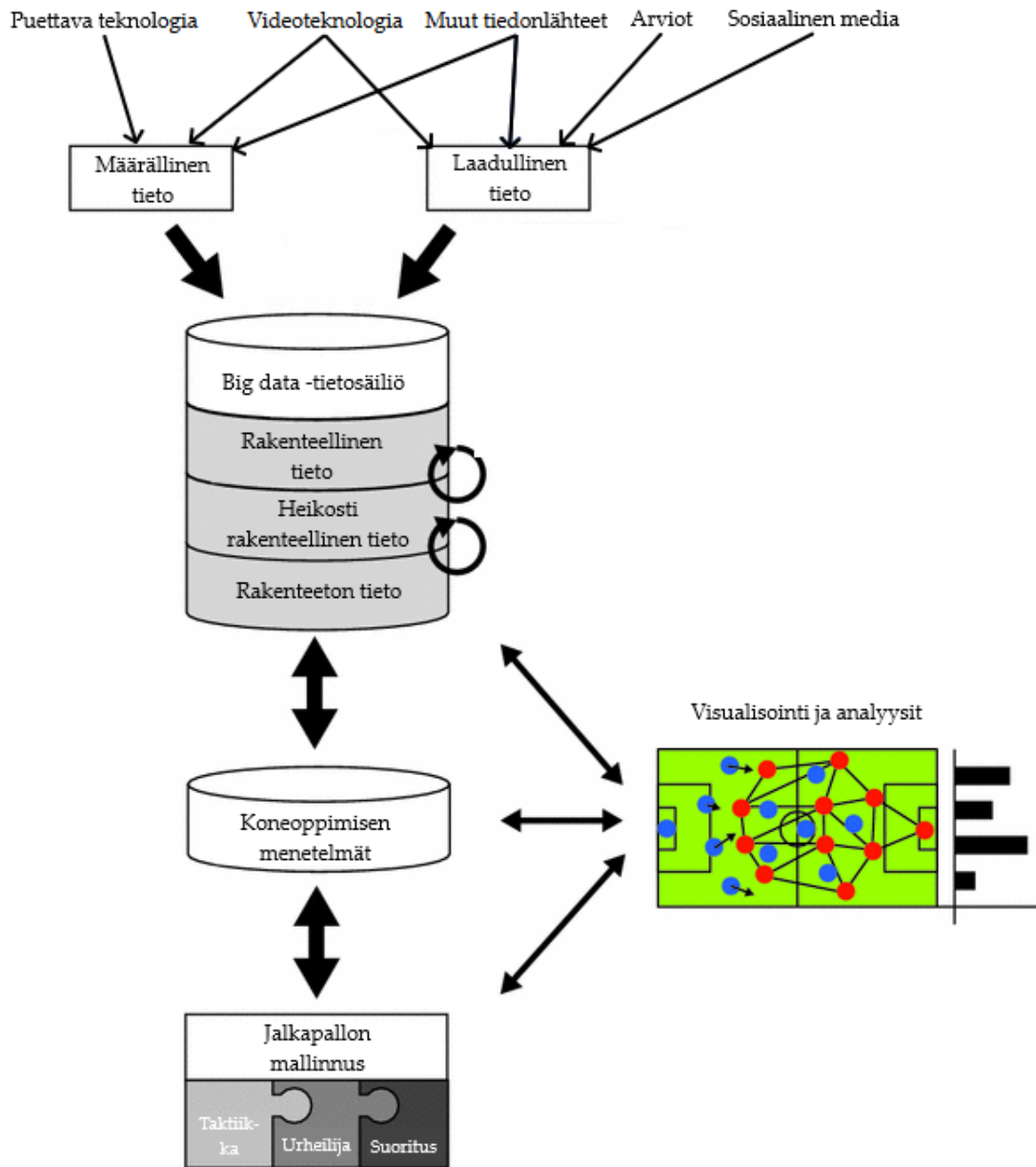
Big datalla viitataan siis kooltaan suureen tietojoukkoon, joka on rakenteeltaan monimuotoista ja jonka tallennusnopeus on suuri. Seuraavaksi esitetään, kuinka jalkapallosta kerättävä tietojoukko sopii tähän kuvaukseen. Otteluista ja harjoituksista kerättävä tietojoukko on suuri: yhdestä Bundesliiga kaudesta otteluista tallennettavat videot, paikka- ja tilastotiedot ovat kooltaan yli 400 gigatavua. Jos seura kerää tietoa myös harjoituksista käyttäen videoteknologiaa, nousee tietojoukon koko suureksi. Kerättävä tieto on myös monimuotoista: jalkapallossa tiedolla voidaan viitata muun muassa paikkatietoihin, videoon, fyysikaaliseen suureeseen tai valmentajan arvioon. Lisäksi tiedon tallennusnopeus vaihtelee nopeista reaaliaikaisista tietovirroista ottelun tai harjoituksen aikana hitaaseen manuaaliseen tiedon tallennukseen, kuten asiantuntijan arvioihin. [Rein and Memmert 2016.]

5.2. Big data -järjestelmä pelianalyysiin

Big data -järjestelmä on tietojärjestelmä, joka voi tallentaa ja prosessoida nopeasti monirakenteista tietoa. Tässä luvussa esitetään yksinkertaistettu big data -järjestelmän malli jalkapallon pelianalyysiin.

Jotta big data -järjestelmä mahdollistaa tässä tutkielmassa määritellyn pelianalyysin teon, tulee sen pystyä keräämään ja säilömään monirakenteista tietoa eri lähteistä ja prosessoimaan tietoa koneoppimisen menetelmillä, jotka pohjautuvat tutkittuun tietoon jalkapallon lainalaisuuksista. Kuvassa 2 esitetään yk-

sinkertaistettu malli pelianalyysiin soveltuvasta big data -järjestelmästä, joka pohjautuu Reinin ja Memmertin [2016] ehdottamaan malliin.



Kuva 2. Yksinkertaistettu malli pelianalyysiin soveltuvasta big data -järjestelmästä, joka pohjautuu Reinin ja Memmertin [2016] ehdottamaan malliin.

Mallissa on kolme tasoa. Ensimmäinen taso käsittää jalkapalloon suorasti ja epäsuorasti liittyvän tiedon keräyksen. Järjestelmään kerätään laadullista ja määrällistä tietoa videoteknologian, puettavan teknologian, sosiaalisen median, asiantuntijoiden ja pelaajien omien arvioiden ja muiden tiedonlähteiden avulla. Toiseen tasoon kuuluu tiedon tallennus big data -tietosäiliöön. Tallennettava tieto voi olla rakenteellista, heikosti rakenteellista tai rakenteetonta [Rein and Memmert 2016].

Viimeisessä tasossa säilöttyyn tietoon sovelletaan koneoppimisen menetelmiä. Koneoppimisen menetelmillä luodaan uutta tietoa ja analysoidaan jo olemassa olevaa. Koneoppimisen menetelmät käyttävät apunaan jalkapallon mallinnusta sekä luovat uusia malleja ja lainalaisuuksia. Jalkapallon mallinnukseen kuuluvat taktiikan, pelaajan ja suorituksen mallinnus, kuten luvussa 3 määriteltiin. Järjestelmän jokaiselta tasolta vaaditaan tiedon visualisointia.

Järjestelmän monitasoisuus tarkoittaa, että järjestelmän luominen vaatii suuren määrän asiantuntemusta. Myös tutkimus- ja kehitystyötä on tehtävä paljon, jotta järjestelmän luominen on mahdollista. [Rein and Memmert 2016.]

5.3. Big data -järjestelmän vaatimukset

Tässä kohdassa tarkastellaan vaatimuksia big data -järjestelmälle, jota voidaan käyttää pelianalyysin tekoon. Tarkoituksena on antaa ainoastaan yleiskuva asioista, joita järjestelmältä vaaditaan. Vaatimuksia tehdessä ei ole otettu huomioon käytännön realiteetteja, vaan pyritty kuvaamaan, mitä ihannetilanteessa järjestelmältä voidaan odottaa.

Vaatimukseen ei ole sisällytetty edellä mainittuja järjestelmän luomiseen liittyviä vaatimuksia, kuten monirakenteisen tiedon käsittelyä tai mahdollisuutta kerätä tietoa usealla eri tavalla. Vaatimuksissa otetaan huomioon kuitenkin ympäristö, jossa järjestelmää tullaan käyttämään. Tässä tapauksessa ympäristöllä viitataan jalkapalloseuraan, koska siellä pelianalyysin teko on yleisintä.

Seuraavat neljä vaatimusta on asetettavissa big data -järjestelmälle:

1. reaaliaikaisuus,
2. helppokäyttöisyys,
3. kontekstin säilytys,
4. tietoturvan vahvuus.

Ensimmäiseksi järjestelmän tulee pystyä tallentamaan ja tarjoamaan tietoa valmentajille ja analyytikkoille reaaliaikaisesti. Jotta valmentaja voi perustaa pelinaikaiset päätöksensä pelistä kerättyyn tietoon, järjestelmän on tarjottava tietoa reaaliajassa. Myös harjoitusten aikana tieto pelaajan suorituskyvystä reaaliajassa voi antaa valmennukselle tärkeää tietoa [Ehrmann *et al.* 2016]. Pelaajan harjoituskuormitusta voidaan muuttaa kesken harjoituksen, jos on selvää, että pelaajan fyysinen kunto ei ole vaadittavalla tasolla.

Toiseksi järjestelmän helppokäyttöisyys on tärkeää, koska järjestelmä tulee ensisijaisesti valmentajien, analyytikkojen ja pelaajien käyttöön. Käyttäjien on päästävä helposti käsiksi analyyseihin. Käyttöliittymien rakennetta ja toimivuutta suunniteltaessa tulee ottaa huomioon, että järjestelmän käyttäjien taustat tulevat olemaan hyvin moninaiset.

Helppokäyttöisyys ei koske ainoastaan järjestelmästä löytyviä käyttöliittymiä, vaan myös tiedonkeruuseen tarkoitettua teknologiaa. Pelaajien ja valmentajien tulee pystyä käyttämään laitteita ilman asiantuntijan paikallaoloa. Tällainen tapaus tulee esimerkiksi kyseeseen Johansenin ja muiden [2015] tutkimuksessa, jossa pelaaja käyttää itsensä mittaukseen tarkoitettua puettavan teknologian sovellusta vapaa-ajalla.

Kolmanneksi järjestelmän tekemisessä analyyseissa on säilytettävä oikea konteksti. Analyyseissa käytettävän tiedon tulee olla mielekästä [Mackenzie and Cushion 2013; Sarmiento *et al.* 2014]. Esimerkiksi Saksan maajoukkueen tehdesä pelianalyysejä tulevasta vastustajastaan, huomioidaan, mitä joukkuetta vastaan tuleva vastustaja on kyseisessä ottelussa pelannut. Jos Saksan tuleva vastustaja on korkeatasoinen, kuten Englanti, ei ole mielekästä tarkastella, kuinka Englanti on pelannut San Marinaa vastaan. San Marino pelaa täysin erilaista jalkapalloa kuin Saksa ja täten ottelusta tehty pelianalyysi voi antaa virheellistä tietoa. Jotta tämän tasoinen kontekstin säilyttäminen saavutetaan, järjestelmän on pystyttävä luokittelemaan useita asioita automaattisesti ja luotettavasti.

Myös järjestelmän tietoturvaan tulee kiinnittää huomiota. Järjestelmästä löytyvä tieto voi olla hyvinkin arkaluonteista, kuten pelaajan terveystiedot. Tämän takia on tärkeää, että järjestelmän luottamuksellisuuteen kiinnitetään huomiota. Käyttäjän oikeuksia tulee rajoittaa siten, ettei käyttäjä näe tietoja, joita hänen ei tarvitse tai kuulu nähdä. Oikeuksia on mahdollista rajoittaa *tunnistus-* tai *todennusjärjestelmällä*.

Lisäksi tiedon *ehyteen* tulee kiinnittää huomiota. Big data -järjestelmän tulee olla tarpeeksi vahva kestämään mahdolliset hyökkäykset [Costa 2014]. Järjestelmässä on monia osia, jotka ovat alttiita hyökkäyksille, kuten tiedonsiirto ja tiedon keräykseen käytetyt laitteet. Tieto ei saa muuttua tiedonsiirron tai käsittelyn aikana.

5.4. Big datan ongelmat

Järjestelmän rakentamiseen liittyvien ongelmien lisäksi suurimmat ongelmat big datassa liittyvät tiedonkeruuseen, yksityisyyden varjeluun ja tietoturvaan.

Urheilijoiden asenteista teknologiaa ja tiedonkeruuta kohtaan ei ole tehty riittävää tutkimusta. [Baerg 2016; Johansen *et al.* 2015]. On epäselvää mitä tapahtuu esimerkiksi tilanteessa, jossa pelaaja kieltäytyy kokonaan käyttämästä tiedonkeruuseen tarkoitettua laitetta vedoten oman yksityisyytensä suojeluun [Baerg 2016]. Lisäksi pelaaja ei ole välttämättä tietoinen kaikista tiedoista, joita hänestä kerätään tai mihin niitä käytetään [Johansen *et al.* 2015]. Tällä hetkellä ei ole olemassa sääntökirjaa, jonka mukaan toimia tiedon keräyksen, anonymisoinnin tai poistoon liittyvissä asioissa [Baerg 2016].

Nykytilanteessa urheiluseuran voi olla mahdollista yhdistää pelaajan vapaa-ajalta kerätyt tiedot kentällä tapahtuviin suorituksiin ja käyttää tätä hyödyksi sopimusneuvotteluissa [Baerg 2016]. Asia on ongelmallinen, sillä vapaa-ajalta kerättäviä tietoja on juuri tarkoitus käyttää apuna suoritusten analysointiin. Tietojen väärinkäytölle ei kuitenkaan tulisi antaa mahdollisuutta tai ainakin sitä on säädettävä jollain tasolla. Jalkapalloliittojen olisi hyvä pohtia tarkkaan, miten ja kuinka paljon pelaajista kerättävän tiedon käyttöä tulee säädellä.

Tällä hetkellä jalkapallosta kerää ja tallentaa tietoa kaupalliset instituutiot, julkiset tutkimusinstituutit ja jalkapalloseurat. Tiedon kerääjien monimuotoisuus tarkoittaa, että tiedonjako toimijoiden välillä on puutteellista johtuen yksityisyyden, tietoturvan ja kilpailuedun suojelemisesta [Rein and Memmert 2016].

Tiedonjaon puutteellisuus vaikeuttaa big data -järjestelmien tutkimusta. Jos tiedonjako toimijoiden välillä paranee, tutkimus- ja kehitystyö voi nopeutua [Rein and Memmert 2016]. Toisaalta, kuten aiemmin todettiin, urheilijoista kerättävä tieto voi olla hyvinkin henkilökohtaista. Jos tiedonjakoa toimijoiden välillä harkitaan, tulee tiedon anonymisointi tehdä siten, ettei tiedon väärinkäytölle ole mahdollisuutta tai sen mahdollisuus on hyvin pieni.

Jalkapalloseurat saattavat kuitenkin tuntea, että tiedon jakaminen voi heikentää heidän kilpailuetuaan. [Baerg 2016; Rein and Memmert 2016.] Tästä syystä tiedonjako varsinkin kuluva kaudesta voi osoittautua ongelmalliseksi. Ratkaisuna tähän voi olla, ettei seurojen tarvitse jakaa tietoa ennen kuin tiedonkeruusta on kulunut tietty aika. Näin saattaa olla mahdollista minimoida negatiivisia vaikutuksia seuran kilpailullisuuteen, ja toisaalta seurat voivat hyötyä suoritetuista tutkimuksista. Toinen vaihtoehto on tehdä yhteistyötä suoraan seurojen kanssa, kuten Johansen ja muut [2013, 2015] ja Saksan maajoukkue ja Kölnin yliopisto ovat tehneet [Cnbc.com; Telekom.com].

6. Yhteenveto

Pelianalyysin tieteellinen tutkimus ei ole pysynyt teknologian kehityksen vauhdissa. Pelianalyysia ei ole määritelty ja se kulkee usealla eri nimellä kirjallisuudessa. Ei ole olemassa yksimielisyyttä siitä, mitkä asiat vaikuttavat pelin kulkuun. Etenkin suomenkielinen tutkimus aiheesta on olematonta. Tutkielmassa esitetyllä pelianalyysin määritelmällä pyritään vastaamaan näihin ongelmiin.

Myös pelianalyysin automatisoinnin tutkimuksessa on puutteita: tyypillinen piirre tutkimuksissa on, että ne ovat pääasiassa tietojenkäsittelytieteilijöiden käsialaa, mikä johtaa usein virheellisiin analyysiin. Toinen tyypillinen piirre on tutkia asioita, jotka uusi teknologia tai menetelmä mahdollistaa, eikä

kiinnitetä juurikaan huomiota siihen, mitä tulisi tutkia. Sama pätee myös tiedon keräykseen: tietoa ei tule mitata vain sen vuoksi, että se on mahdollista.

Tulevaisuudessa tutkimuksen tulisi olla poikkitieteellisempää, sillä sekä tietojenkäsittely- että urheilutiede hyötyvät yhteistyöstä. Yliopistojen opetussuunnitelmien muutoksilla voi olla mahdollista lisätä yhteistyötä tieteiden välillä, kuten Kölnin yliopistossa on tehty. Myös valmentajien ja muiden asiantuntijoiden mukaanotto voi osoittautua hyödylliseksi.

Big datan ja pelianalyysin integrointia seurojen rakenteeseen saattaa hidastaa jalkapallon vanhoillisuus. Jalkapallossa esimerkiksi maaliviivateknologia otettiin käyttöön arvokisoissa vasta vuonna 2013, vaikka teknologia mahdollisti sen jo paljon kauemmin. Valmentajille big data -järjestelmien ja teknologian käyttö voi tuntua turhalta. Tämä toisaalta tarkoittaa sitä, että big data -järjestelmän käyttö voi antaa joukkueelle kilpailuedun. Yleinen suunta on kuitenkin, että big data ja pelianalyysi tulevat yleistymään.

On nähtävissä, kuinka pieniä osia tutkielmassa esitetystä big data -järjestelmästä on jo toteutettu. Tällä hetkellä tutkielmassa määritellyn pelianalyysin teko vaatii kuitenkin useiden eri järjestelmien käyttöä, mikä aiheuttaa hukkaa seurojen resursseihin. Vaikka järjestelmät voivat auttaa jo nyt pelianalyysissä huomattavasti, vaativat ne vielä paljon kehitys- ja tutkimustyötä. Järjestelmien reaaliaikaisuuteen, käytettävyyteen, analyysien kokonaisvaltaisuuteen ja tietoturvaan tulee kiinnittää huomiota tulevissa tutkimuksissa.

Big datan ja jalkapallon pelianalyysin yhdistäminen osoittaa paljon potentiaalia: tulevaisuudessa kokonaisvaltaisten, objektiivisten ja luotettavien pelianalyysien teko lähes reaaliajassa voi olla mahdollista. Lisäksi big data mahdollistaa jalkapallon mallinnuksen täysin uudella tavalla. Jotta tähän päästään, vaaditaan kuitenkin useiden eri alojen asiantuntijoiden ja tutkijoiden yhteistyötä

Viiteluettelo

- Andrew Baerg. 2016. Big data, sport, and the digital divide theorizing how athletes might respond to big data monitoring. To appear in *Journal of Sport & Social Issues*.
- C. Barnes, D. T. Archer, B. Hogg, M. Bush and P. S. Bradley. 2014. The evolution of physical and technical performance parameters in the English premier league. *Int. J. Sports Med.* 35, 13, 1095-1100.
- Alina Bialkowski, Patrick Lucey, Peter Carr, Yisong Yue and Iain Matthews. "Win at home and draw away": Automatic formation analysis highlighting the differences in home and away team behaviors. In: *Proc. of 8th Annual MIT Sloan Sports Analytics Conference*.

- Michel S. Brink, Chris Visscher, Suzanne Arends, Johannes Zwerver, Wendy J. Post and Koen A. Lemmink. 2010. Monitoring stress and recovery: New insights for the prevention of injuries and illnesses in elite youth soccer players. *Br. J. Sports Med.* 44, 11, 809-815.
- Michael Bush, Chris Barnes, David T. Archer, Bob Hogg and Paul S. Bradley. 2015. Evolution of match performance parameters for various playing positions in the English premier league. *Human movement science* 39, 1-11.
- Julen Castellano, David Alvarez-Pastor and Paul S. Bradley. 2014. Evaluation of research using computerised tracking systems (amisco® and prozone®) to analyse physical performance in elite soccer: A systematic review. *Sports Medicine* 44, 5, 701-712.
- Claude Sammut and Geoffrey I. Webb. 2011. *Encyclopedia of machine learning*. Springer Science & Business Media.
- Computer Weekly. *Foxy Leicester City FC won Premiership with data analytics* 5.12.2016. <http://www.computerweekly.com/news/450296302/Foxy-Leicester-City-FC-won-Premiership-with-data-analytics>.
- Cnbc. Athletes to analysts: How big data gave the German football team a leg up 12.12.2015. <http://www.cnn.com/2016/07/07/euro-2016-sap-and-german-football-team-worked-to-develop-big-data-analytics.html>.
- Fabricio F. Costa. 2014. Big data in biomedicine. *Drug Discov. Today* 19, 4, 433-440.
- Fifa. <http://www.fifa.com/worldcup/matches/round=255955/match=300186474/statistics.html>. Checked 19.12.2016.
- T. D’Orazio and M. Leo. 2010. A review of vision-based systems for soccer video analysis. *Pattern Recognition* 43, 8, 2911-2926.
- Fabian E. Ehrmann, Craig S. Duncan, Doungkamol Sindhusake, William N. Franzsen and David A. Greene. 2016. GPS and injury prevention in professional soccer. *The Journal of Strength & Conditioning Research* 30, 2, 360-367.
- Forbes. *Big Data: The Winning Formula in Sports* 5.12.2016. <http://www.forbes.com/sites/bernardmarr/2015/03/25/big-data-the-winning-formula-in-sports/#3b643a1d26dc>.
- Jlio Garganta. 2009. Trends of tactical performance analysis in team sports: Bridging the gap between research, training and competition. *Revista Portuguesa de Ciências do Desporto* 9, 1, 81-89.
- Paul B. Gastin, Owen C. Mclean, Ray V. Breed and Michael Spittle. 2014. Tackle and impact detection in elite australian football using wearable microsensor technology. *J. Sports Sci.* 32, 10, 947-953

- Andreas Grunz, Daniel Memmert and Jürgen Perl. 2011. Tactical pattern recognition in soccer games by means of special self-organizing maps. *Human Movement Science* 31, 2, 334-343.
- Timothy E. Hewett, Stephanie L. Di Stasi and Gregory D. Myer. 2013. Current concepts for injury prevention in athletes after anterior cruciate ligament reconstruction. *Am. J. Sports Med.* 41, 1, 216-224.
- Mikko Häyrinen. 2010. Pelianalyysi valmennuksen tukena. *Kilpa- ja huippu-urheilun tutkimus keskus KIHU*, http://www.kihu.fi/tuotostiedostot/julkinen/2010_hay_pelianalyysi_sel2_1_34933.pdf
- InStat. <http://instatfootball.com/clients/>. Luettu 5.12.2016.
- Håvard Johansen, Cathal Gurrin and Dag Johansen. 2015. Towards consent-based lifelogging in sport analytic. In: *Proc. of the International Conference on Multimedia Modeling*, 335-344.
- Håvard D. Johansen, Svein A. Pettersen, Pål Halvorsen and Dag Johansen. 2013. Combining video and player telemetry for evidence-based decisions in soccer. In: *Proc. of the icSPORTS*.
- M. I. Jordan and T. M. Mitchell. 2015. Machine learning: Trends, perspectives, and prospects. *Science* 349, 6245, 255-260.
- Laura E. Juliff, Shona L. Halson and Jeremiah J. Peiffer. 2015. Understanding sleep disturbance in athletes prior to important competitions. *Journal of Science and Medicine in Sport* 18, 1, 13-18.
- Jouni Juntunen. 2011. *Jalkapallon lajiansalyysi ja valmennuksen ohjelmointi*. Seminaarityö, liikuntabiologian laitos, Jyväskylän yliopisto.
- Seamus Kelly. 2008. Understanding the role of the football manager in Britain and Ireland: A weberian approach. *European Sport Management Quarterly* 8, 4, 399-419.
- Konstantin Knauf, Daniel Memmert and Ulf Brefeld. 2016. Spatio-temporal convolution kernels. *Mach. Learning* 102, 2, 247-273.
- Teuvo Kohonen. 1990. The self-organizing map. In: *Proc IEEE* 78, 9, 1464-1480.
- Kostas Kostiadis and Huosheng Hu. 2001. KaBaGe-RL: Kanerva-based generalisation and reinforcement learning for possession football. In: *Intelligent Robots and Systems, 2001*. In: *Proc. 2001 IEEE/RSJ International Conference on*, 1, 292-297
- Wei-Lwun Lu, Jo-Anne Ting, James J. Little and Kevin P. Murphy. 2013. Learning to track and identify players from broadcast sports videos. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 7, 1704-1716.
- Patrick Lucey, Alina Bialkowski, Mathew Monfort, Peter Carr and Iain Matthews. 2014. Quality vs quantity: Improved shot prediction in soccer using

- strategic features from spatiotemporal data. In: *Proc. 8th Annual MIT Sloan Sports Analytics Conference*, 1-9.
- Pekka Luhtanen. 2015. Pelianalyysi-pelikäsitys jalkapallossa. Eerikkilän urheilupuisto, <https://www.eerikkila.fi/wp-content/uploads/2015/11/Pelianalyysi-ja-Pelikasitys-Luhtanen.pdf>. Lue-ttu 28.10.2016.
- Rob Mackenzie and Chris Cushion. 2013. Performance analysis in football: A critical review and implications for future research. *J. Sports Sci.* 31, 6, 639-676.
- Daniel Memmert, Lemmink, A. P. M. Koen and Jaime Sampaio. 2016. Current approaches to tactical performance analyses in soccer using position data. To appear in *Sports Medicine*.
- Ral Montoliu, Ral Martn-Flez, Joaquin Torres-Sospedra and Adolfo Martnez-Us. 2015. Team activity recognition in association football using a bag-of-words-based method. *Human Movement Science* 41, 165-178.
- Mathieu Nédélec, Shona Halson, Barthélémy Delecroix, Abd-Elbasset Abaidia, Said Ahmaidi and Gregory Dupont. 2015. Sleep hygiene and recovery strategies in elite soccer players. *Sports Medicine* 45, 11, 1547-1559.
- Derek Nevins, Kasee Hildenbrand, Jeff Kensrud, Anita Vasavada and Lloyd Smith. 2016. Field evaluation of a small form-factor head impact sensor for use in soccer. In: *Procedia Engineering* 147, 186-190.
- OptaPro. <http://www.optasportspro.com>. Checked 5.12.2016.
- Jürgen Perl. 2002. Game analysis and control by means of continuously learning networks. *International Journal of Performance Analysis in Sport* 2, 1, 21-35.
- Svein A. Pettersen, Dag Johansen, Håvard Johansen, Vegard Berg-Johansen, Vamsidhar R. Gaddam, Asgeir Mortensen, Ragnar Langseth, Carsten Griwodz, Håkon K. Stensland and Pål Halvorsen. 2014. Soccer video and player position dataset. In: *Proc. of the 5th ACM Multimedia Systems Conference*, 18-23.
- Amjad Rehman and Tanzila Saba. 2014. Features extraction for soccer video semantic analysis: Current achievements and remaining issues. *Artif. Intell. Rev.* 41, 3, 451-461.
- Thomas Reilly and Ben Edwards. 2007. Altered sleep-wake cycles and physical performance in athletes. *Physiol.Behav.* 90, 2, 274-284.
- Robert Rein and Daniel Memmert. 2016. Big data and tactical analysis in elite soccer: Future challenges and opportunities for sports science. *Springer-Plus* 5, 1, 1-13.

- Hugo Sarmiento, Rui Marcelino, M. T. Anguera, Jorge Campanio, Nuno Matos and Jos C. Leito. 2014. Match analysis in football: A systematic review. *J. Sports Sci.* 32, 20, 1831-1843.
- Pete B. Shull, Wisit Jirattigalachote, Michael A. Hunt, Mark R. Cutkosky and Scott L. Delp. 2014. Quantified self and human movement: A review on the clinical impact of wearable sensing and feedback for gait analysis and intervention. *Gait Posture* 40, 1, 11-19.
- Gunter P. Siegmund, Kevin M. Guskiewicz, Stephen W. Marshall, Alyssa L. DeMarco and Stephanie J. Bonin. 2016. Laboratory validation of two wearable sensor systems for measuring head impact severity in football players. *Ann. Biomed. Eng.* 44, 4, 1257-1274.
- Telekom. *Will data decide the championship?* 12.12.2016
<https://www.telekom.com/en/company/digital-responsibility/details/will-data-decide-the-championship--430380>.
- The Guardian. *How computer analysts took over at Britain's top football clubs* 5.12.2016.
<https://www.theguardian.com/football/2014/mar/09/premier-league-football-clubs-computer-analysts-managers-data-winning>.
- Karl Tuyls and Ann Now. 2005. Evolutionary game theory and multi-agent reinforcement learning. *The Knowledge Engineering Review* 20, 01, 63-90.
- Jarryd L. Wallace and Kevin I. Norton. 2014. Evolution of world cup soccer final games 1966–2010: Game structure, speed and play patterns. *Journal of Science and Medicine in Sport* 17, 2, 223-228.
- Bo Zhou, Harald Koerger, Markus Wirth, Constantin Zwick, Christine Martindale, Heber Cruz, Bjoern Eskofier and Paul Lukowicz. 2016. Smart soccer shoe: Monitoring foot-ball interaction with shoe integrated textile pressure sensor matrix. In: *Proc. of the 2016 ACM International Symposium on Wearable Computers*, 64-71.

An overview of common server virtualization solutions

Tuukka Hämynen

Abstract

As the number of servers in enterprise environments grows, so does the need for server consolidation and resource management. Virtualization is an effective way of maximizing the utilization of available server resources. This thesis provides an overview of six commonly used virtualization solutions, and gives insight to their features and differences. This thesis also discusses conflicting research results regarding the power consumption of various virtualization solutions, and proposes an explanation to the discrepancy. Lastly, a number of likely use cases are outlined for the six virtualization solutions.

Keywords: Server, virtualization, Docker, OpenVZ, Xen, KVM, VMware, ESXi, Hyper-V, virtual machine manager, hypervisor, power consumption.

1. Introduction

With the rise of cloud computing and e-commerce, there is an ever-growing demand for servers and datacenters in enterprise environments [Jiang *et al.* 2016].

Virtualization is often introduced as a way of consolidating multiple physical servers into one physical server, hosting multiple virtual server instances. [Dittner and Rule 2011] This way, the physical resources of the server machine can be better utilized, assuming the servers are not always under full load. For example, a collection of five physical servers that are only 15% utilized, could be consolidated into one physical server running five virtual servers. This way, the physical host server will theoretically be 75% utilized.

Recently, server consolidation has been gaining popularity as a way of improving the energy efficiency of data centers. With this approach, idle servers in data centers can be shut down to reduce energy usage [Jin and Wen 2012].

Virtualization has other benefits as well; since virtualization isolates the host machine from the virtualized instances, a system fault in one instance will not generally affect the others, which increases reliability and security. Adding and decommissioning virtual servers is also simple, since there is no need to install new, or to disassemble old hardware. [Dittner and Rule 2011]

There are many different virtualization solutions on the market, each with their own benefits and limitations. These solutions generally fall into one of three categories: operating-system-level virtualization, paravirtualization, or full vir-

tualization. These can be further divided into two groups: hypervisors and container-based solutions. These categories and groups will be explained in more detail in later chapters.

This thesis is an overview of six commonly used virtualization solutions, and aims to provide insight on their differences in features, use cases, and power consumption.

2. Terminology

The terminology related to virtualization is not always clear. For example, words like *virtual machine manager* and *hypervisor* are sometimes used interchangeably, while at times apparently having distinct differences.

Some definitions are difficult to coherently explain without first introducing the concepts they represent. For this reason, there will be a more detailed explanation of some definitions in Chapter 4. For now, it is enough to know that, in this thesis, hypervisors are considered a subset of virtual machine managers.

In the following chapters, three different categories for virtualization methods will be introduced, followed by two different types of virtualization platforms, as well as the two types of hypervisors. It should be noted that these are not clear-cut categorizations, and some sources will distinguish them differently. However, for the purposes of this thesis, these distinctions are likely to be the easiest to understand.

3. Virtualization methods

There exist many different virtualization solutions, commonly called *virtual machine managers*, or *VMMs* [Dittner and Rule 2011]. These VMMs manage hardware resources, and handle the requests of the guest operating systems or application containers.

Virtual machine managers are generally divided into three categories: *operating-system-level virtualization*, *paravirtualization* and *full virtualization*. These categories will be explained in more detail in the following chapters.

3.1. Operating-system-level virtualization

Operating-system (OS) level virtualization is a method based on a single instance of an operating system, which allows the existence of multiple isolated software containers within that instance.

OS-level virtualization is often compact and efficient, as it only requires a single operating system installation, which is used for managing and updating the system and the software containers, instead of each container needing to run its own OS. This allows the containers to run at native speeds, without any overhead

from having to virtualize hardware, and thus supports all native hardware and OS features that the host OS is configured for. [Dittner and Rule 2011]

One possible drawback of OS-level virtualization is that it does not support hosting mixed OS families on the same server; it is not possible to, for example, run Windows containers on a Linux server [Dittner and Rule 2011].

This thesis covers two OS-level virtualization solutions, Docker and OpenVZ, which will be detailed in Chapter 5.

3.2. Paravirtualization

Paravirtualization provides partial simulation of the underlying hardware in which most, but not all, of the hardware features are simulated.

Operating systems cannot be run in paravirtualized environments without significant modifications to the OS and the kernel, which limits the selection of operating systems available for paravirtualization. Regardless, paravirtualization is easier to implement than full virtualization, due to full virtualization having more specific hardware requirements and limitations. In addition, paravirtualized *virtual machines* (VMs) are often the highest performing for network and disk I/O when no hardware assistance is available. [Dittner and Rule 2011] This thesis covers one paravirtualization solution, Xen.

3.3. Full virtualization

Full virtualization provides complete simulation of the underlying hardware. This results in a system where any software capable of running on the physical hardware can be run in a virtual machine. Thus, full virtualization also supports the widest range of guest operating systems, which can be installed without any modification. It also provides near-native CPU and memory performance. [Dittner and Rule 2011]

Full virtualization completely isolates the virtual machines and the VMM from each other, which increases overall security of the system; a compromised host does not compromise the whole server. It also requires a processor that supports hardware-assisted virtualization (*Intel VT-x* or *AMD-V*). [Dittner and Rule 2011]

This thesis covers three full virtualization solutions, VMware ESXi, Microsoft Hyper-V, and KVM.

4. Virtualization platform types

In this chapter, the major differences between the two different hypervisor types and container-based solutions will be described in brief.

It should be noted that, as mentioned in Chapter 2, there is some ambiguity in the terminology related to virtualization, mainly the terms *hypervisor* and *virtual machine manager*.

Some sources [Dittner and Rule 2011] use the terms *hypervisor* and *virtual machine manager* synonymously, and some list OS-level virtualization methods when discussing hypervisors [Tafa *et al.* 2011].

Other sources [Morabito 2015] draw a clear distinction between hypervisors, which implement either full virtualization or paravirtualization, and *container-based* solutions, which implement OS-level virtualization. This distinction implies that either they do not consider OS-level virtualization methods to be VMMs, or they consider VMMs and hypervisors to be two different things.

For the purposes of this thesis, methods implementing paravirtualization and full virtualization will be considered both hypervisors and VMMs. Methods implementing OS-level virtualization will be considered VMMs, but not hypervisors. As such, hypervisors will be considered a subset of VMMs. This is to highlight the differences between traditional hypervisors and container-based solutions, while still having a collective term for both groups.

4.1. Hypervisors

The x86 architecture uses four *privilege levels*, or *protection rings*, numbered from 0 to 3, as shown in Figure 1.

System memory is divided into segments, and each segment is dedicated to a specific ring. The processor uses the privilege levels to determine what can and cannot be done with code or data within a segment. *Ring-0* is the innermost ring, with total control of the processor. Moving outward, each ring has more restrictions applied to it. *Ring-3*, the outermost ring, is provided only with restricted access. [Dittner and Rule 2011]

The *Supervisor Mode* is the execution mode on an x86 processor that enables the execution of all instructions, including privileged instructions such as I/O and memory management operations. Normally, the operating system would run in Supervisor Mode in Ring-0. [Dittner and Rule 2011]

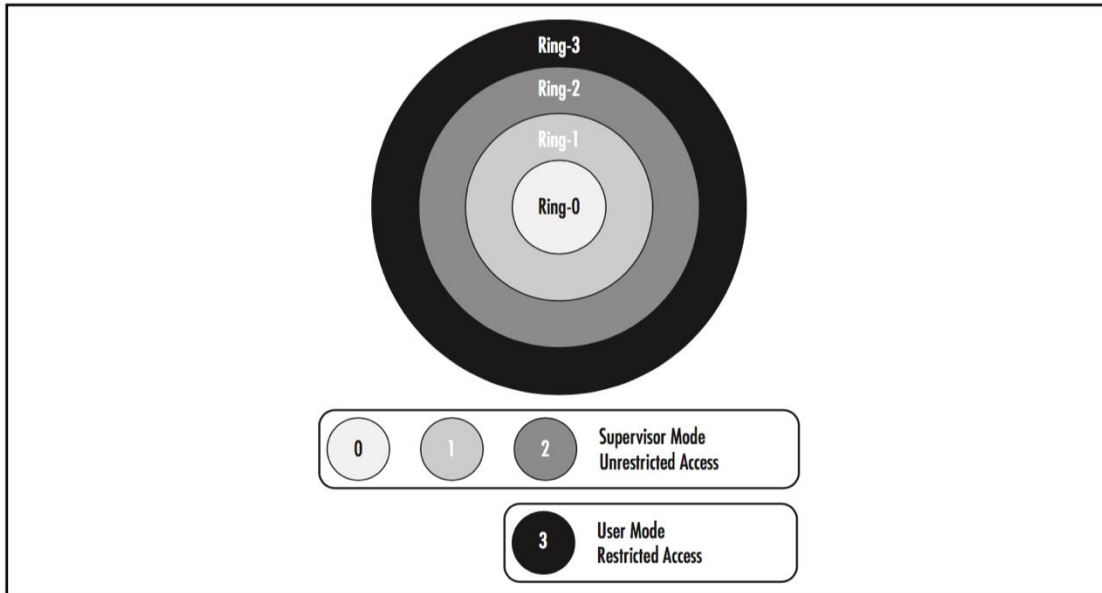


Figure 1. Privilege rings in the x86 architecture [Dittner and Rule 2011].

Since Ring-3 depends on Ring-0, any system compromise or instability directly impacts *User Mode* running in Ring-3. Isolating Ring-0 for each virtualized guest requires moving Ring-0 closer to the guests. Thus, a Ring-0 failure for one guest is unable to impact Ring-0, or consequently Ring-3, of any other guest. The perceived Ring-0 for guest VMs can reside in either Ring-1, Ring-2, or Ring-3. The further away the perceived Ring-0 is from the true Ring-0, the more distant it is from executing direct hardware operations, resulting in reduced performance and independence. [Dittner and Rule 2011]

Virtualization moves Ring-0 up by placing the hypervisor in one of the rings, which in turn presents the Ring-0 implementation to the hosted virtual machines. It is on this presented Ring-0 that guest operating systems run, while the hypervisor handles the actual interaction with the underlying hardware platform. [Dittner and Rule 2011]

Hypervisors are generally divided into two categories, *type-1* and *type-2* hypervisors, based on where on the ring model the hypervisor is placed. These types are demonstrated in Figure 2.

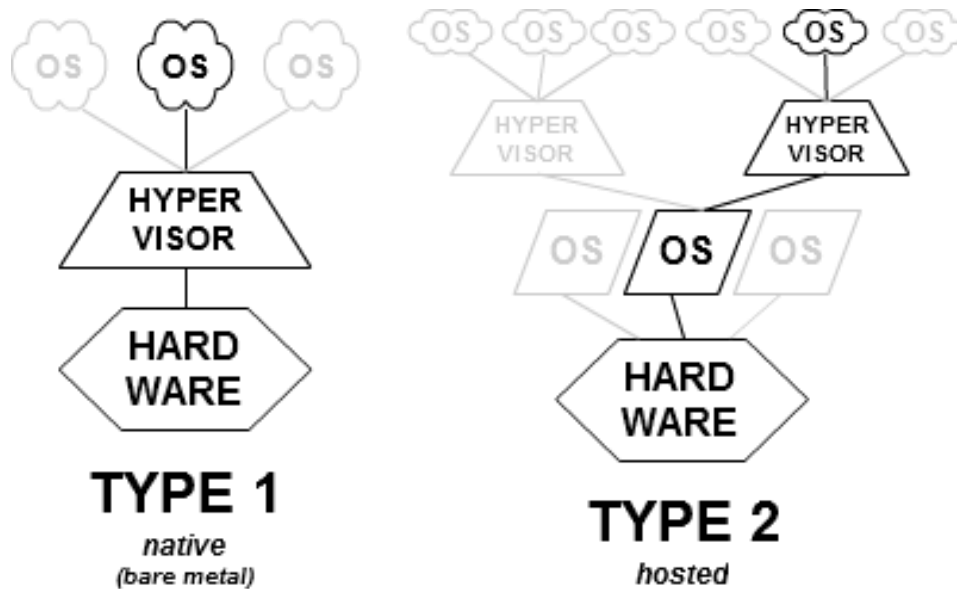


Figure 2. Hypervisor types [Wikipedia 2016b].

A *type-1*, or a *bare metal* hypervisor is software that runs directly on top of a given hardware platform on the true Ring-0. Guest operating systems then run at a higher level above the hardware, allowing for true isolation of each virtual machine. [Dittner and Rule 2011]

A *type-2* hypervisor is software that runs within an operating system, usually in Ring-3. Since there are no additional rings above Ring-3 in the x86 architecture, the presented Ring-0 that the virtual machines run on is as distant from the actual hardware platform as it can be. Although this offers some advantages, it is usually compounded by performance-impeding factors; system calls to the hardware must go through many diverse layers before the operations are returned to the guest operating system. [Dittner and Rule 2011]

4.2. Container-based solutions

Container-based solutions are sometimes regarded as *type-2* hypervisors. However, due to the differences in the core principles between traditional hypervisors and container-based solutions, many sources choose to separate the two. A simplified diagram of the main differences between hypervisor-based solutions and container-based solutions can be seen in Figure 3.

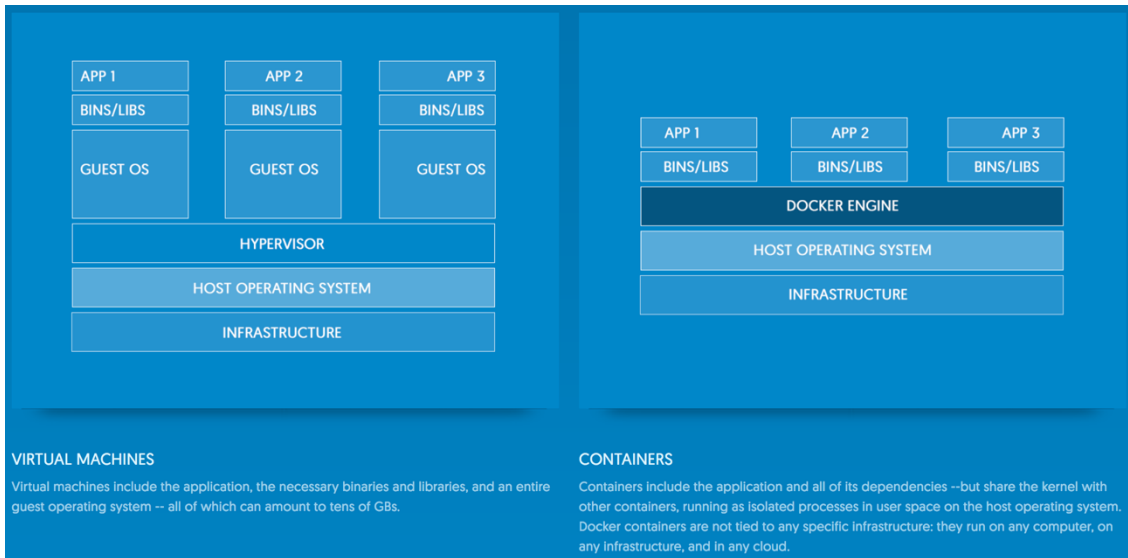


Figure 3. Differences between virtual machines and container-based solutions [Docker 2016b].

Instead of running a full OS on virtual hardware, containers implement virtualization of processes at the operating system level of the underlying host machine, modifying it to provide extra isolation [Felter *et al.* 2015; Morabito 2015]. Thus, the overhead due to virtualized hardware, and virtual device drivers is avoided [Morabito 2015]. Unlike a VM which runs a full OS, a container can be made to run as little as a single process [Felter *et al.* 2015]. This has two advantages: a higher density of virtualized instances, and a smaller disk image [Morabito 2015].

Generally, containers are isolated from each other and the main OS, so that they have no visibility or access to objects outside the container. Processes running inside the container appear to be running on a normal system, although they share the underlying kernel with processes located in other containers. However, if complete isolation is not required, containers can be allowed to share some resources between themselves. [Felter *et al.* 2015]

5. Commonly used virtual machine managers

There is a variety of virtual machine managers available to choose from, depending on the intended use case. In this chapter, the six VMMs chosen for this thesis will be introduced, and their main features briefly explained.

One major feature commonly supported by VMMs is *live migration*, which means that a running virtual machine instance can be moved from one physical host to another with no interruption to service; the VM remains powered on, network connections remain active and applications continue to run while the VM is being moved to another physical host. [Microsoft 2016b; Red Hat 2015]

Out of the six VMMs covered by this thesis, Docker is the only one that does not currently support live migration.

5.1. Docker

Docker is a relatively new container-based virtualization solution, which has been gaining popularity in recent years. It can be considered as a lightweight alternative to hypervisor-based virtualization solutions. [Morabito 2015]

Docker is free, open source, and runs natively on Linux, Mac OS X, and Windows. Due to kernel requirements, Linux-based containers can be run within Linux, OS X or Windows, but Windows-based containers can only be run on Windows. [Docker 2016a]

Docker completely encapsulates software applications, creating a sandboxed virtual environment, without the overhead of a complete virtual machine. This is because Docker does not replicate the full OS, it only loads the libraries and binaries required by the application being virtualized. [Chamberlain and Schommer 2014]

Docker used to rely on the LXC toolset, but recently released its own libraries called *libcontainer*, which it uses to deploy containers with namespaces and control groups. Namespaces in Docker are used to isolate containers, so that each aspect of a container runs in its own namespace and does not have access outside it. Control groups take care of the resource limitation of the hardware, such as vCPU, memory and disk I/O. [van Kessel *et al.* 2016]

5.2. OpenVZ

OpenVZ is an open source container-based virtualization solution, which can create multiple isolated Linux containers on a single server [Che *et al.* 2010; [OpenVZ 2016]. Each container performs like a stand-alone server; they can be rebooted independently, have IP addresses, memory, system libraries, etc. [OpenVZ 2016]

Due to implementing operating-system-level virtualization, OpenVZ requires the kernel of the guest operating system to be the same as the host kernel [Che *et al.* 2010].

OpenVZ runs on a modified Linux kernel that enables virtualization, isolation, and resource management. OpenVZ controls container operations and system resources with parameters that can be changed without the need for rebooting the container. [Che *et al.* 2010]

5.3. Xen

Xen is an open source type-1 hypervisor that supports both paravirtualization and full virtualization, which can be used at the same time on a single hypervisor [Xen Project 2016]. This gives it the flexibility to suit many different needs.

Xen uses a specialized virtual machine (as shown in Figure 4), called *Domain 0*, which has unique privileges [Jin and Wen 2012], such as the capability to access the hardware directly. It handles all access to the I/O functions of the system, and interacts with the other VMs [Xen Project 2016]. Xen hosts most of its unmodified Linux device drivers in Domain 0, which also handles CPU scheduler parameters and resource allocation policies [Che *et al.* 2010].

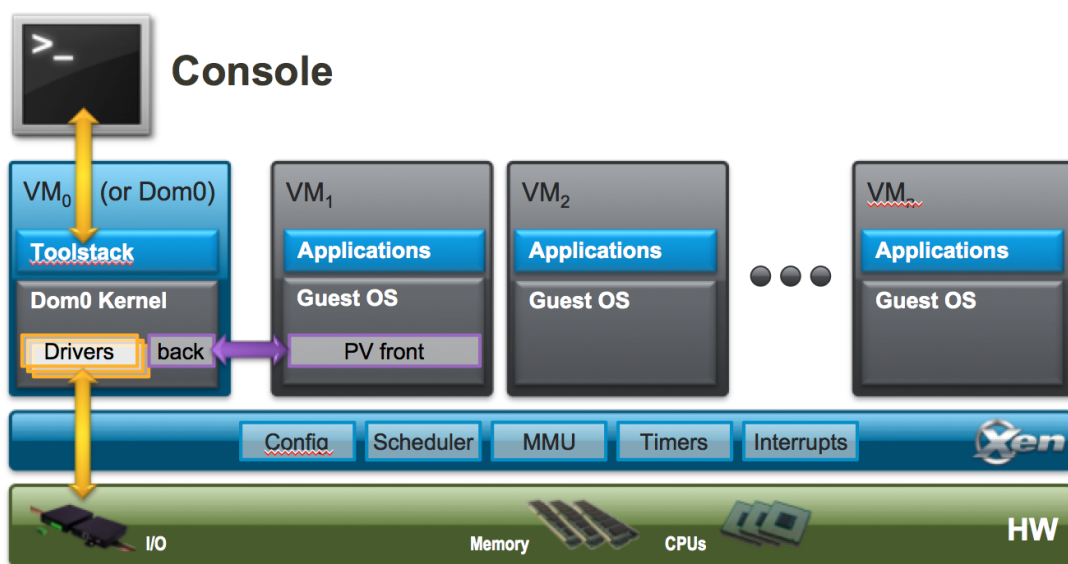


Figure 4. Diagram of the Xen architecture [Xen Project 2016].

Xen provides two communication mechanisms between hypervisor and domain 0; a *synchronous call* using *hypercall* to send messages from domain 0 to Xen, and an *asynchronous event* using a *virtual interrupt* to send notifications from Xen to guest Domain. [Che *et al.* 2010]

The guest operating systems used in Xen must have a modified kernel that supports paravirtualization; such kernels exist for Linux, NetBSD, FreeBSD, and OpenSolaris. It is therefore not possible to virtualize Windows guests in Xen. [Xen Project 2016] The guests are aware of the hypervisor, and can be run without any virtual emulated hardware. Therefore, paravirtualized guests make direct calls without an additional layer, thus having generally less overhead compared to full virtualization. [van Kessel *et al.* 2016; Xen Project 2016]

By running itself in Ring-0 and migrating guest operating systems to Ring-1, Xen separates them from underlying hardware and holds full control of system

resources. When a guest operating system tries to execute a sensitive privileged instruction, the physical processor will trap it into Xen. [Che *et al.* 2010]

5.4. KVM

KVM, or *Kernel-based Virtual Machine*, is an open source full virtualization solution, which requires *Intel VT-x* or *AMD-V* enabled CPUs. KVM is implemented as a loadable kernel module, which converts the Linux kernel into a bare metal hypervisor [Red Hat 2015]. A simplified diagram of the KVM architecture is shown in Figure 5.

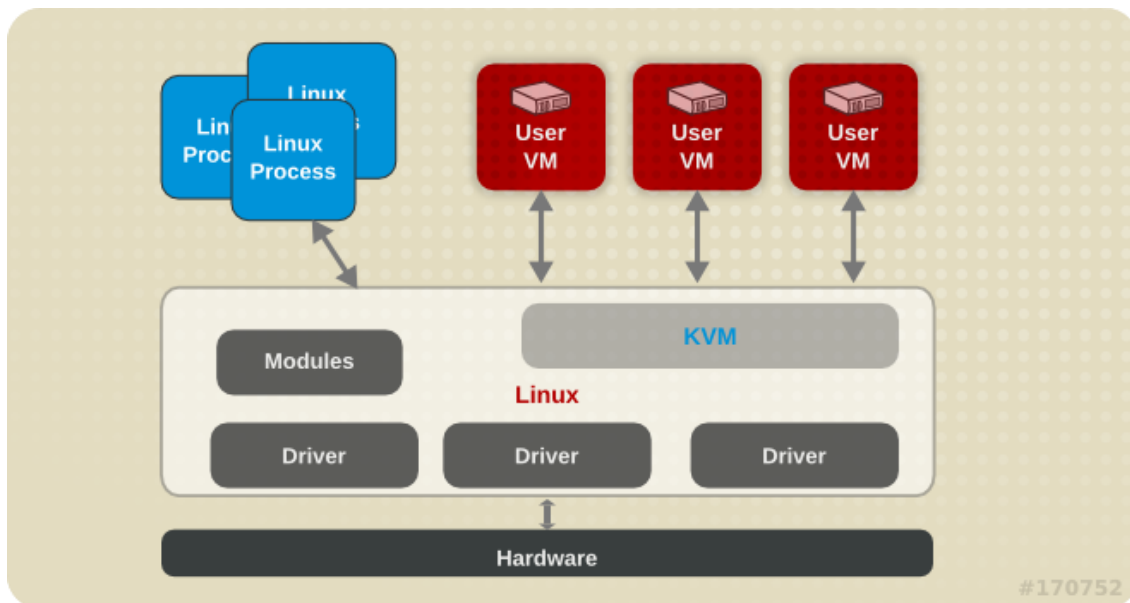


Figure 5. Diagram of the KVM architecture [Red Hat 2016]

KVM augments the traditional *kernel* and *user* modes of Linux with a new mode called *guest*, which has its own *kernel* and *user* modes and oversees the code execution of the guest operating systems [Che *et al.* 2010]. The guest VMs are implemented as regular Linux processes, and as such are subject to the standard Linux scheduler. Thus, KVM enjoys all advantages of the standard Linux kernel. [Che *et al.* 2010; Red Hat 2015]

I/O virtualization is handled by employing a modified QEMU to simulate I/O behavior, or sometimes triggering the real I/O if necessary. [Che *et al.* 2010; Red Hat 2015]

KVM typically uses full virtualization, but can be configured for paravirtualization as well [Jin and Wen 2012]. This makes the classification of KVM difficult, because it is technically designed to be type-2 hypervisor, but then modifies the Linux kernel to function much like a type-1 hypervisor.

KVM supports a wide variety of guest operating systems: Linux, Windows, NetBSD, FreeBSD, OpenSolaris, Solaris x86 and MS DOS [Red Hat 2015].

5.5. VMware ESXi

VMware ESXi is a type-1 hypervisor that uses full virtualization. It is designed to behave more like firmware rather than traditional software, and does not have a conventional operating system or a service console. VMware has instead created APIs (which can be seen in Figure 6) through which monitoring and management tasks can be performed. Scripting environments such as vCLI and PowerCLI allow the remote execution of scripts and commands. [VMware 2016b]

Since ESXi does not have a console OS, it is very compact, making it easy to install and configure, which leaves more resources for the virtualized guests. Because of its minimal attack-surface, it also does not require constant patching, reducing security risks. [Nishikiori 2011]

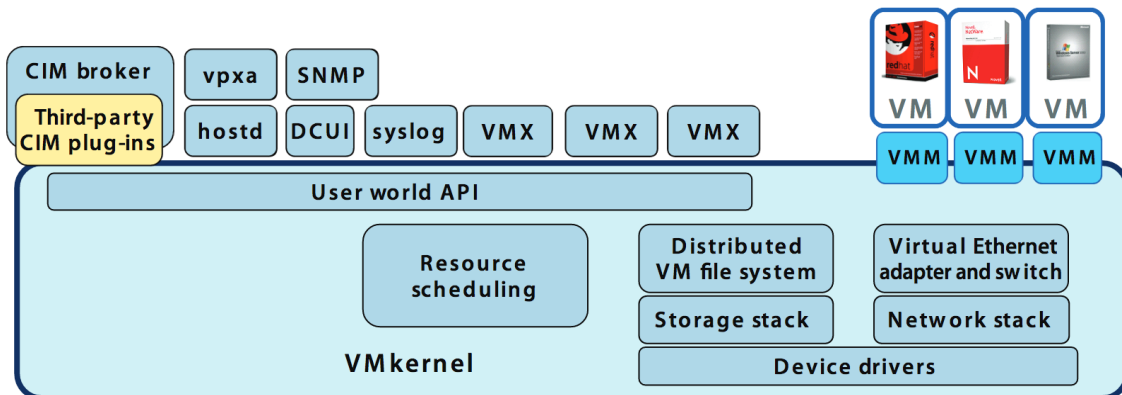


Figure 6. Diagram of the ESXi architecture [VMware 2007].

ESXi handles all the hardware access for each VM, hosting drivers for all the hardware (storage, network, and input devices) that the VMs need to access. Because all device drivers reside within the hypervisor, it is necessary for VMware to be very strict about which systems will support their hypervisors and which will not. Therefore, ESXi will only run on a selective number of systems. [Fayyad-Kazan *et al.* 2013]

ESXi supports a wide variety of guest operating systems, including Linux, Windows, Mac OS, FreeBSD, Solaris and MS DOS [VMware 2016a].

VMware's virtualization solutions are the most widely used in enterprise environments, with an estimated market share for server virtualization of approximately 70% [Murugesan and Bojanova 2016].

5.6. Hyper-V

Microsoft Hyper-V is a type-1 hypervisor for 64-bit versions of Windows Server, which uses full virtualization. It exists in two variants: as a stand-alone product called Hyper-V Server and as an installable component in Windows Server. [Microsoft 2016b]

Much like KVM, Hyper-V also requires a processor with hardware-assisted virtualization support [Fayyad-Kazan *et al.* 2013].

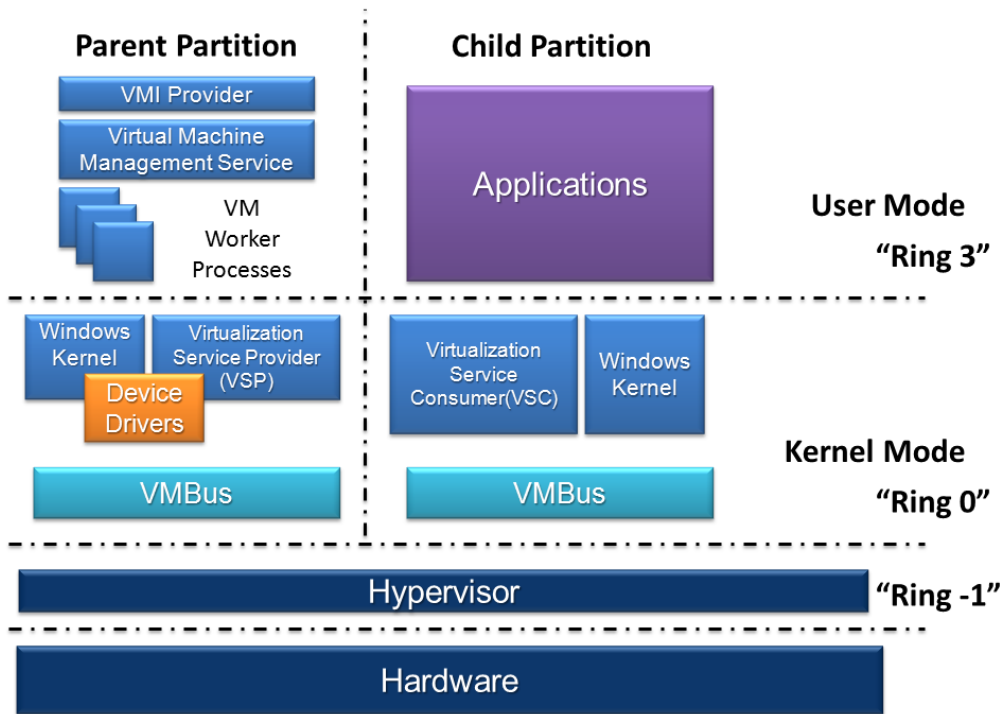


Figure 7. Diagram of the Hyper-V architecture [Wikipedia 2016a].

Hyper-V implements virtualization in terms of a *partition* (see Figure 7). A partition is a logical unit of isolation, supported by the hypervisor, in which each guest operating system executes. The hypervisor requires at least one *root partition*, running a supported version of Windows Server 64-bit Edition. The root partition in turn creates the *child partitions*, which host the guest operating systems, using the hypercall API. [Microsoft 2016a]

Partitions have a virtual view of the processor and run in virtual memory private to each guest partition. The hypervisor handles interrupts to the processor, and redirects them to the relevant partition. [Microsoft 2016a]

6. Performance and power consumption

Servers consume a significant amount of electricity, even when they are idle. Thus, server consolidation via virtualization has been suggested as a way for data centers to reduce their power consumption [Jin and Wen 2012].

It was estimated that datacenters in the USA consumed 61 million MWh electricity in 2006, which was double the consumption in 2000, and was expected to double again by the end of 2011 [Jin and Wen 2012]. In a report issued by the National Renewable Energy Laboratory it was estimated that U.S. data center

electricity usage was approximately 91 million MWh in 2013 [Morabito 2015]. The slowed-down increase in power consumption can be speculated to be partly due to advances in producing energy-efficient hardware, as well as more widespread adoption of server virtualization. It has been demonstrated that by optimizing data center operations via virtualization, energy consumption can be cut by up to 20% [Jin and Wen 2012].

Jin and Wen [2012] found that virtualized servers generally consume more electricity than physical ones, for both computing and networking intensive tasks. According to their research, the energy overhead of virtualized servers increases as the utilization of physical resources increases.

However, they also found that a physical server could consume more electricity than a virtualized server does. For example, in their research concerning Xen and KVM, Jin and Wen [2012] found that when the number of VMs is one more than the number of CPU cores, the energy overhead is negative for Xen-based servers.

Jin and Wen [2012] observed that, from a set of 5 test instances on a physical quad-core server, the first four finished first and the last instance is completed much later. This was further verified by the observation that, in the physical server, the CPU usage maintained at nearly 100% until the first four instances were completed, after which it dropped to approximately 25%. By contrast, in the virtualized servers, all the instances were completed at almost the same time while consuming less electricity. This suggests that there is an optimal number of virtual instances for each server, which will reduce their overall power consumption.

The power efficiency of virtualized servers has been studied extensively. The results of the studies appear mixed in some areas. This can be partly due to the studies being spread across many years, which, because virtualization is a relatively new and rapidly developing technology, can mean that there have been various upgrades and optimizations for different virtualization methods between the studies.

6.1. Power consumption at idle

Multiple studies have found that the power consumption of different servers varies when the server is powered on, but the virtual machines are idle.

Docker and OpenVZ have generally been found to have the lowest power consumption at idle, very close to the power consumption of an identical non-virtualized server [Morabito 2015; Shea *et al.* 2014].

Xen has received mixed results. Jin and Wen [2012] measured only a 0.32% increase with two VMs, and a 0.47% increase with three VMs, compared to the

physical server. Morabito [2015] further supports these findings by stating that, although he measured Xen to consume the most power compared to KVM and Docker, he found the differences between the three to be almost negligible. However, Shea *et al.* [2014] measured a 41% increase in power consumption for Xen compared to the physical server. They propose that this is due to Xen's Domain 0 being a virtual machine, which does not allow the Linux kernel to properly manage the power of the system. Unlike KVM and OpenVZ, which take advantage of the standard Linux power saving system.

KVM has received mixed results as well. In some cases, KVM based servers have been measured to incur a much higher overhead than other solutions. For example, Jin and Wen [2012] measured an increase of 11.6% with three VMs, and 9.55% with two VMs. Contrarily, others [Morabito 2015; Shea *et al.* 2014] have observed KVM to have a similar idle power consumption compared to the physical server.

ESXi has been observed to have 50% smaller idle power consumption while running two VMs, compared to two physical servers. From this it can be extrapolated that ESXi has a similar idle power consumption to a single identical physical server. [Liu *et al.* 2012]

Data comparing the idle power consumption of Hyper-V to an identical non-virtualized server could not be found. This suggests that, if such research data exists, it is rare. This could be partly due to Hyper-V being able to run only on Windows servers, which requires a paid license.

6.2. Power consumption in general

Research covering the power consumption of various virtualization methods has been inconclusive, and drawing accurate conclusions from conflicting results can be difficult [Che *et al.* 2010; Fayyad-Kazan *et al.* 2013; Felter *et al.* 2015; Jin and Wen 2012; Morabito 2015; Tafa *et al.* 2011].

Jiang *et al.* [2016] suggest a possible explanation for this problem; based on their research, it appears that the efficiency of current virtualization solutions depends significantly on the server hardware used as the host machine.

They found that one hypervisor may have the highest power consumption for some, or all, workload levels on one set of hardware, but have the lowest power consumption for some, or all, workload levels on another set of hardware. For example, Hyper-V was observed to have the highest power consumption on an HP DL380 G6 server at all workload levels, while it had the lowest power consumption for fair and very heavy computing workloads on an Intel S2600GZ server. [Jiang *et al.* 2016]

Since most of the studies done on the power consumption of VMMs have been done using a single type of server, the explanation is likely correct.

7. Use case comparison

Due to differences in architecture, some VMMs will likely be better suited for some tasks than other VMMs.

Container-based solutions, such as Docker and OpenVZ, are ideal for running multiple isolated copies of the same application. Since they do not virtualize a full OS, but rather only the specific components required by the application, there is less wasted space on the server. Containers are also useful for developing and testing software, because they are easier to set up, simple to share with other developers if needed, and can be expected to run identically regardless of the environment. [Docker 2016b]

Paravirtualization and full virtualization solutions generally have the widest range of supported guest operating systems [VMware 2016a; Red Hat 2015]. For this reason, they are usually best suited for large-scale hosting, or for environments which require running multiple different applications, possibly on multiple different operating systems, perhaps even on different versions of the same operating system.

Because many hypervisors require hardware-assisted virtualization support from the CPU, they might not be a feasible option for some small-scale operations, which could be better served by a container-based solution.

There are also some specific requirements to keep in mind for some VMMs. Xen, for example, requires the guest VMs to have a paravirtualization-enabled kernel. It is therefore not possible to run a Windows guest on a Xen hypervisor.

Hyper-V, on the other hand, specifically requires the host and guest servers to run Windows, which significantly limits which applications can be virtualized. If, for example, some Linux-exclusive applications are required, Hyper-V is not a realistic option.

KVM presents users with a convenient setup of full virtualization, as it does not require any modifications to the guest operating systems, and, since it is a kernel module, can be installed on an already running Linux server without the need to completely reformat the system. KVM is probably more suitable for environments where full virtualization and ease-of-use are wanted.

8. Conclusion

There are various virtualization solutions, each with their benefits and restrictions. In this thesis, I have briefly covered the differences between the three

types of virtualization methods; described the main differences between hypervisors and container-based solutions; given an overview of the main features of six commonly used virtual machine managers, and discussed the research surrounding their performance and power consumption.

Based on the features, restrictions and requirements of the virtualization methods and the VMMs themselves, I have suggested some likely use cases for each type of VMM, as well as noted some specific cases for Xen, Hyper-V and KVM.

Researchers investigating the power consumption of virtualization solutions have come to conflicting conclusions. This shows that there is still much to discover about the intricacies of virtualization and its relation to the hardware it is implemented on.

The results indicate that, currently, modern virtualization solutions perform differently based on the hardware they are run on, as well as the tasks they are used to perform. When compared to other VMMs, one VMM may have the highest power consumption for some or all workload levels on one platform, but have the lowest power consumption at some of all workload levels on another platform.

Finding a virtualization solution that performs the intended tasks as efficiently as possible on the intended hardware appears currently to be mostly trial-and-error. Therefore, companies looking to consolidate their servers via virtualization should likely first focus on the possible hardware requirements and desired features, when deciding which virtualization solution is right for them, and only later focus on finding the optimal configurations in terms of the number of virtual machines and servers.

References

- Ryan Chamberlain and Jennifer Schommer. 2014. Using Docker to support reproducible research. DOI: [http://dx.doi.org/10.6084/m9.figshare, 7, 14](http://dx.doi.org/10.6084/m9.figshare.7.14).
- Jianhua Che, Congcong Shi, Yong Yu, and Weimin Lin. 2010. A synthetical performance evaluation of OpenVZ, Xen and KVM. In: Proc. of the 2010 IEEE Asia-Pacific Services Computing Conference, 587-594.
- Rogier Dittner and David Jr. Rule. 2011. *The Best Damn Server Virtualization Book Period: Including Vmware, Xen, and Microsoft Virtual Server*. Syngress.
- Docker. 2016a. *Docker Engine frequently asked questions (FAQ)*. Accessed December 12, 2016. <https://docs.docker.com/engine/faq/>.
- Docker. 2016b. *What is Docker?* Accessed December 17, 2016. <https://www.docker.com/what-docker>.

- Hasan Fayyad-Kazan, Luc Perneel, and Martin Timmerman. 2013. Benchmarking the performance of Microsoft Hyper-V server, VMware ESXi and Xen hypervisors. *Journal of Emerging Trends in Computing and Information Sciences*, 922-933.
- Wes Felter, Alexandre Ferreira, Ram Rajamony, and Juan Rubio. 2015. An updated performance comparison of virtual machines and Linux containers. In: *Proc. of the 2015 IEEE International Symposium On Performance Analysis of Systems and Software*, 1-15.
- Congfeng Jiang, Dongyang Ou, Yumei Wang, Xindong You, Jilin Zhang, Jian Wan, Bing Luo, and Weisong Shi. 2016. Energy efficiency comparison of hypervisors. In: *Proc. of the 7th International Green & Sustainable Computing Conference*. IEEE.
- Yichao Jin and Yonggang Wen. 2012. Energy efficiency and server virtualization in data centers: an empirical investigation. In: *Proc. of the 2012 IEEE Conference on Computer Communications Workshops*, 133-138.
- Lu Liu, Osama Masfary, and Nick Antonopoulos. 2012. Energy performance assessment of virtualization technologies using small environmental monitoring sensors. *Sensors* 12, 5, 6610-6628.
- Microsoft. 2016a. *Hyper-V Architecture*. Accessed December 9, 2016. [https://msdn.microsoft.com/en-us/enus/library/cc768520\(v=bts.10\).aspx](https://msdn.microsoft.com/en-us/enus/library/cc768520(v=bts.10).aspx).
- Microsoft. 2016b. *Hyper-V Technology Overview*. Accessed December 17, 2016. <https://technet.microsoft.com/windows-server-docs/compute/hyper-v/hyper-v-technology-overview>.
- Roberto Morabito. 2015. Power consumption of virtualization technologies: an empirical investigation. In: *Proc. of the 2015 IEEE/ACM 8th International Conference on Utility and Cloud Computing*, 522-527.
- San Murugesan and Irena Bojanova. 2016. *Encyclopedia of Cloud Computing*. John Wiley & Sons.
- Masaaki Nishikiori. 2011. Server virtualization with VMware vSphere 4. *Fujitsu Scientific and Technical Journal*, 356-361.
- OpenVZ. 2016. *OpenVZ Virtuozzo containers Wiki*. Accessed December 17, 2016. https://openvz.org/Main_Page.
- Red Hat 2016. 1.3. *KVM Overview*. Accessed December 19, 2016. https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Virtualization_Tuning_and_Optimization_Guide/sect-Virtualization_Tuning_Optimization_Guide-Introduction-KVM_Overview.html.

- Red Hat. 2015. KVM – Kernel-based Virtual Machine. Accessed December 17, 2016. <https://www.redhat.com/en/resources/kvm-%E2%80%93-kernel-based-virtual-machine>.
- Ryan Shea, Haiyang Wang, and Jiangchuan Liu. 2014. Power consumption of virtual machines with network transactions: measurement improvements. In: *Proc. of the IEEE INFOCOM 2014-IEEE Conference on Computer Communications*, 1051-1059.
- Igli Tafa, Ermal Beqiri, Hakik Paci, Elinda Kajo, and Aleksander Xhuvani. 2011. The evaluation of transfer time, cpu consumption and memory utilization in xen-pv, xen-hvm, openvz, kvm-fv and kvm-pv hypervisors using ftp and http approaches. In: *Proc. of the 2011 Third International Conference on Intelligent Networking and Collaborative Systems*, 502-507.
- Jeroen van Kessel, Arie Taal, and Paola Grosso. 2016. Power efficiency of hypervisor-based virtualization versus container-based virtualization. University of Amsterdam.
- Wikipedia. 2016a. *Hyper-V – Wikipedia*. Accessed December 17, 2016. <https://en.wikipedia.org/wiki/Hyper-V>.
- Wikipedia. 2016b. *Hypervisors – Wikipedia*. Accessed December 17, 2016. <https://en.wikipedia.org/wiki/Hypervisor>.
- VMware. 2007. The Architecture of VMware ESXi. *VMware.com*. Accessed December 20, 2010. https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/techpaper/ESXi_architecture.pdf.
- VMware. 2016a. *VMware Compatibility Guide*. Accessed December 17, 2016. <http://www.vmware.com/resources/compatibility/search.php?action=base&deviceCategory=software>.
- VMware. 2016b. *VMware ESX and ESXi 4.1 Comparison (1023990)*. Accessed December 7, 2016. https://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalID=1023990.
- Xen Project. 2016. *Xen Project Software Overview*. Accessed December 17, 2016. https://wiki.xenproject.org/wiki/Xen_Overview.

Suomen valtion sähköisen hallinnon kehitys 2003–2016

Tuomas Karjanlahti

Tiivistelmä

Sähköinen hallinto on jatkuvasti kehittyvä ilmiö, joka jaetaan neljään osaan: katalogisointi, transaktiot, vertikaalinen integraatio ja horisontaalinen integraatio. Kehitystä voidaan lisäksi tarkastella Web 2.0 näkökulmasta.

Suomen sähköisen hallinnon kehittämiseksi on vuosina 2003-2016 tehty kolme hallitusvetoista hanketta. Vanhasen ensimmäinen hallitus julkisti syksyllä 2003 hallituksen strategia-asiakirjassa tietoyhteiskuntaohjelman, jonka tarkoituksena oli lisätä kilpailukykyä ja tuottavuutta. Keväällä 2009 asetettiin sähköisen asioinnin ja demokratian vauhdittamisohjelma eli SADe-hanke. Sen tarkoituksena oli parantaa Suomen tietoyhteiskunta- ja asiointipalvelujen kehitystä.

Kolmas hanke on yhä käynnissä oleva vuonna 2014 aloitettu Sipilän hallituksen kärkihanke julkisten palveluiden digitalisoinnista käyttäjäläheiseksi. Digitalisaatiolla on tavoitteena saada aikaan tuottavuusloikka, jolla tarkoitetaan tuottavuuden merkittävää parantamista. Sähköistä hallintoa pyritään kehittämään luomalla yhden luokun palvelumalli, jossa tärkeänä osana on Kansallinen Palveluarkkitehtuuri -hanke.

Tässä tutkielmassa esitellään sähköisen hallinnon kehitysprojekteja, niiden aikaansaannoksia ja niissä kohdattuja ongelmia.

Avainsanat ja -sanonnat: Sähköinen hallinto Suomessa, Sähköisen hallinnon kehittäminen, Web 2.0, Kansallinen palveluarkkitehtuuri

1. Johdanto

Suomen sähköisen hallinnon kehittämiseksi on vuosina 2003–2016 tehty kolme hallitusvetoista hanketta. Vanhasen ensimmäinen hallitus julkisti syksyllä 2003 hallituksen strategia-asiakirjassa tietoyhteiskuntaohjelman, jonka tarkoituksena oli lisätä kilpailukykyä ja tuottavuutta. Tämän lisäksi sen tavoitteena oli lisätä sosiaalista ja alueellista tasa-arvoa, kansalaisten hyvinvointia ja elämänlaatua, sekä säilyttää Suomen asema yhtenä tieto- ja viestintäteknologian johtavista tuottajista ja hyödyntäjistä. [Hallituksen strategia-asiakirja]

Valtiovarainministeriö asetti keväällä 2009 sähköisen asioinnin ja demokratian kehittämishanke -nimisen ehdotuksen pohjalta samannimisen vauhdittamisohjelman (SADe-hanke). Sen tarkoituksena oli parantaa Suomen tietoyhteiskunta -palvelujen ja asiointipalvelujen kehitystä kolmella toimintakokonaisuudella. Ensimmäinen oli *"jatkaa arjen tietoyhteiskuntaohjelmaa"*, toinen *"käynnistää sähköisen asioinnin ja demokratian ohjelma"* sekä kolmas *"uudelleen organisoida julkisten tietojärjestelmien määrittely, hankinta ja ylläpito"*. SADe-ohjelman tavoitteena

oli mahdollistaa kansalaisten ja yritysten sähköinen asiointi vuoteen 2013 mennessä kattaa kaikki keskeiset palvelut. [SADe toteuttamissuunnitelma 2009]

Kolmas hanke on yhä käynnissä oleva vuonna 2014 aloitettu Sipilän hallituksen kärkihanke julkisten palveluiden digitalisoinnista käyttäjäläheiseksi. Digitalisaatiolla on tavoitteena saada aikaan tuottavuusloikka, jolla tarkoitetaan tuottavuuden merkittävää parantamista [Sipilä ja Vehviläinen 2015]. Sähköistä hallintoa pyritään kehittämään ”Digitalisoidaan julkiset palvelut” -hankkeella, jonka tarkoitus on tehdä sähköiset palvelut käyttäjälähtöiseksi ja ensisijaisesti digitaaliseksi toimintatapoja uudistamalla. Hankkeen ensimmäiset vaiheet on tarkoitus ottaa käyttöön vuoden 2017 aikana ja projektin tulisi olla kokonaan valmis vuoden loppuun mennessä.

Kärkihankkeella on kolme toimenpidettä, joista ensimmäisessä luodaan kaikkia julkisia palveluja koskevat digitalisoinnin periaatteet. Toisessa hallinnonalat ja kuntien hallinnot sitoutetaan sisäisten prosessien uudistamiseen. Kolmannessa luodaan yhden luukun palvelumalli ja uudistetaan tiedonhallintalainsäädäntö. Kohtaan kolme sisältyy KaPA, eli Kansallinen Palveluarkkitehtuuri. Kansallisessa palveluarkkitehtuurissa pyritään luomaan yhteen toimiva digitaalisten palvelujen infrastruktuuri, jonka avulla tiedon siirto organisaatioiden ja palvelujen välillä on helppoa. [Palveluarkkitehtuuri]

Tutkimuksen toisessa kappaleessa tarkastellaan, kuinka tietokirjallisuus määrittelee sähköistä hallintoa ja sen kehittämistä. Kolmas kappale määrittelee, millä periaatteilla hallintoa tulisi luoda ja kehittää tulevaisuudessa. Neljännessä kappaleessa perehdytään tarkemmin jo alussa mainittuihin Suomen sähköisen hallinnon kehityshankkeisiin ja tarkastellaan niissä tehtyjä ratkaisuja tutkimustiedon teoriataustaan peilaten. Viidennessä kappaleessa pohditaan tähän asti tehtyjen ratkaisujen onnistumista ja sähköisen hallinnon tulevaisuuden kehittämistä Suomessa. Kokonaisuutena tutkimuksessa tarkastellaan Suomen sähköisen hallinnon kehittymisen lisäksi sitä, millä periaatteilla sitä tulisi yleisesti kehittää. Tutkimus pyrkii myös luomaan katsauksen tulevaisuuden sähköisen hallinnon kehittämiseen.

2. Sähköisen hallinnon luominen ja kehittäminen

Layne ja Lee [2001] ovat luoneet neliportaisen mallin sähköisen hallinnon (eGovernment tai e-government) luomiseksi ja kehittämiseksi. Sähköinen hallinto on jatkuvasti kehittyvä ilmiö, joka jaetaan neljään osaan: ensimmäinen on katalogisointi, toinen transaktiot tai vuorovaikutukset, kolmas vertikaalinen integraatio ja neljäs horisontaalinen integraatio. Sun ja muut [2014] vievät horisontaalisen ja vertikaalisen integraation vaiheita vielä syvemmälle ja kehittävät sähköistä hallintoa Web 2.0 -näkökulmasta.

Ensimmäisessä vaiheessa tiedot hallinnon tarjoamista palveluista ja prosesseista pyritään viemään Internetiin ja esimerkiksi kaavakkeet tai informaatio on saatavilla verkon kautta. Se ikään kuin kerää ”katalogiin” hallinnon tiedot ja esittää ne verkossa. Toisessa vaiheessa palvelu tukee vuorovaikutusta Internetin välityksellä, kuten rahaliikennettä tai verkossa täytettäviä lomakkeita. Pyrkimyksenä on minimoida ihmisten tekemä työ niin, että maksut tai tallennetut tiedot kohdistetaan suoraan niille tarkoitettuihin järjestelmiin ja tietokantoihin automaattisesti, minimaalisella tai olemattomalla manuaalisella työllä. [Layne and Lee 2001]

Laynen ja Leen [2001] mukaan kansalaisten vaatimukset ja yhteiskunnan muutokset tarkoittavat, että hallitusten on mentävä yhä pidemmälle sähköisen hallinnon suunnittelussa. Kolmannessa vaiheessa, eli vertikaalisessa integraatiossa paikalliset järjestelmät on linkitetty korkeamman tason järjestelmiin, jotka toimivat keskenään samankaltaisilla toiminnoilla. Neljännessä vaiheessa, eli horisontaalisessa integraatiossa erilaiset järjestelmät toimivat yhteen, eli kaikki tuotteet ovat saatavilla yhden luukun takaa, jolloin voidaan puhua yhden luukun palvelumallista (one stop shopping). Sähköisen hallinnon toteutuksessa ratkaisevat hyödyt saadaan hallinnon eri tasojen taustaprosessien integraatiosta.

Vertikaalisessa ja horisontaalisessa integraatiossa voidaan puhua sähköisen hallinnon lisäksi hallinnon digitalisoimisesta. Digitalisaation periaatteiden mukaan ei ole riittävää, että olemassa olevia toimintoja siirretään sellaisenaan verkkoon. Prosesseja tulisi uudelleenorganisoida pysyvästi, sekä hallinnon toimintamalleja tarkastella uudelleen. Tällä tavalla voidaan saavuttaa pitkällä tähtäimellä sähköisen hallinnon täydet hyödyt, kun organisaatiolliset muutokset yhdistyvät teknologisten kanssa. [Layne and Lee 2001]

Useimmat vuorovaikutusta vaativat tietokannat ovat hajallaan eri paikoissa, eli ne ovat lokalisoituja ja fragmentoituneita. Luonnollinen kehitys tulee Laynen ja Leen [2001] mukaan olemaan hajautettujen eritasoisten (vertikaalisten) hallinnon järjestelmien erilaisten toimintojen (horisontaalisten) integraatio. He toteavat myös, että oletettavasti vertikaalinen integraatio tapahtuu ensin, koska kiulu hallinnon palveluiden välillä on suurempi kuin hallinnon eri tasojen välillä. Useat virastot kommunikoivat enemmän kansallisten ja paikallisten eri tasojen osapuolien kanssa kuin muut virastot samalla hallintotasolla. Esimerkiksi poliisilaitos on yhteydessä paikallisen syyttäjänviranomaisen tai vankeinhoitolaitoksen kanssa enemmän kuin eri kaupungissa toimivan poliisilaitoksen.

Laynen ja Leen [2001] mukaan vertikaalisen integraation tarkoitus on saumattomasti yhdistää valtion järjestelmät muiden toimijoiden, kuten kuntien kanssa. Verkkopalvelujen saavutettavuuden kannalta kansalaisille on tarjottava hyvät verkkoyhteydet. Lisäksi muita teknisiä haasteita asettavat muun muassa

autentikointi, elektronisen datanvaihdon soveltuvat tiedostomuodot ja järjestelmän haavoittuvuus ulkopuolisille. Lisäksi he huomauttavat, että on tärkeää ottaa huomioon vertikaalisen integraation aiheuttama hallinnon rajaviivojen häilyminen. Tämä voi aiheuttaa sen, että kansalliset ja kunnalliset toiminnot saattavat olla kansalaisen näkökulmasta vaikeammin erotettavissa. On myös huomioitava, että digitalisaation edetessä valtion työntekijöiden rooli muuttuu prosessien suorittajasta enemmänkin niiden valvojaksi. Heidän toimintonsa rajat voivat ylittyä osastolta toiselle entisen selkeän lokeroinnin sijaan.

Kolmannen tason kehittyminen vaatii hallinnon usealta asteelta joustavuutta ja tietokantojen kehittämistä myös tasojen yhteisiin tarpeisiin. Virastojen tulee myös olla avoimempia jakamaan tietoaan. Layne ja Lee [2001] tähdentävät kuitenkin, että vaikka parantuneella tietojen jakamisella saavutettaisiin parantuneet hallinnon olosuhteet, on yksityisyys- sekä luottamuksellisuusongelmat otettava huomioon ennen palvelujen kehittämistä.

Layne ja Lee [2001] toteavat, että kansalaisen näkökulmasta informaatioteknologian täysi hyöty voidaan saavuttaa vain horisontaalisesti integroimalla hallinnolliset palvelut yli erilaisten toiminnallisten muurien. Esimerkiksi Suomessa työtön henkilö hakiessaan töitä käyttää työvoimatoimiston palveluita ja hakiesaan toimeentulotukea KELA:n palveluita. He toteavat myös, että kun eri toimialojen tietokannat kommunikoivat keskenään ja yhden laitoksen tiedot ovat saatavilla toisessa, parantaa horisontaalinen integraatio palvelujen käyttäjän toimintaa huomattavasti. Sen mahdollistama järjestelmien integraatio pystyy automatisoimaan prosesseja, joissa vuorovaikutus yhden viraston kanssa saattaa johdattaa automaattisiin tietojen tarkistuksiin tai muutoksiin toisessa virastossa.

3. Sähköisen hallinnon kehitys tulevaisuudessa

Sun ja muut [2014] lähestyvät sähköisen hallinnon tulevaisuutta Web 2.0 -näkökulmasta määrittelemällä viitekehyksen sähköiselle hallinnolle 2.0 (E-Government 2.0, jatkossa tekstissä SH 2.0). He ehdottavat, että SH 2.0:n kehitys on yhteydessä sosiaalisuuden, toiminnollisuuden ja datamäärän kasvuun ja niistä luotaviin innovaatioihin. Kuten tähän asti, myös tulevaisuudessa tavoitteena on parantaa hallinnon toimintojen tehokkuutta ja vaikuttavuutta koko yhteiskunnassa.

3.1. Web 2.0

Sunin ja muiden [2014] mukaan ICT:n tärkein tehtävä on luoda innovaatioita uuden informaatioteknologian avulla ja ratkaista ongelmia, joita hallinto ei pystyisi ratkaisemaan yksin. He ehdottavat, että tähän sopiva ratkaisu on Web 2.0, koska

se on sosiaalinen kehitysaste, eikä pelkästään teknologinen. Web 2.0 mahdollistaa ja kannustaa osallistumiseen sovellusten ja palveluiden kautta, joiden sisältöä voi käyttää uusissa konteksteissa. Se muovaa verkkoa välineestä, jossa on informaatiota, ja sitä vain kulutetaan alustaksi, jossa sisältöä luodaan, jaetaan, muokataan, käytetään uudelleen ja jaetaan eteenpäin.

Web 2.0:lle ei ole olemassa tarkkaa määritelmää, mutta sitä voidaan kuvata eri sovellusten kautta. Siinä toimivat sovellukset hyödyntävät käyttäjien tuottamaa dataa, mahdollistavat heidän muokata sovellusta mieleisekseen sekä sovellukset toimivat usealla eri laitteella ja tarjoavat mielekkäitä käyttäjäkokemuksia [O'Reilly 2007]. Näitä palveluita nykyään ovat lähes kaikki sosiaalisen median palvelut kuten Facebook ja Twitter sekä esimerkiksi blogit.

Parhaimmillaan Web 2.0 -teknologioiden avulla hallintoa voidaan kehittää palvelemaan kansalaisia itse hallinnon sijaan jakamalla informaatiota ja palveluita paremmin esimerkiksi eri virastojen välillä. Tämän lisäksi voidaan luoda uusia palveluita, jotka vastaavat entistä paremmin kansalaisten tarpeita sekä helpottavat heidän vuorovaikutustaan hallinnon kanssa. Uudet teknologiat toimivat myös katalyyttinä uudistaa hallintoa ja sen toimintatapoja korvaamalla perinteisiä työskentelytapoja uusilla, tehokkaammilla ja vaikuttavammilla prosesseilla, rakenteilla ja kommunikaatitavoilla. [Sun *et al.* 2015]

Tarjotakseen parempia palveluita hallintojen tulisi tarkastella, kuinka muovata itsestään joustavampia organisaatioita. Näin ne pystyisivät vastaamaan vaihteleviin olosuhteisiin ja löytämään uusia tapoja toteuttaa tehtäviään. Web 2.0 antaa mahdollisuuden luoda SH 2.0, jonka avulla kansalaisille voidaan tuoda hallinnolliset toiminnot entistä saavutettavammaksi. Se poistaa virastojen rajoja ja tukee avoimuutta, läpinäkyvyyttä sekä kansalaisten osallistamista. Hallinnon ja kansalaisten kuilun kaventaminen mahdollistaa, että kansalaiset voivat osallistua päätöksentekoon entistä paremmin. He voivat ehdottaa uusia palveluita, jättää uusia ehdotuksia ja päästä hyödyntämään hallinnollista dataa. Se myös kannustaa avoimuuden kulttuuriin ja läpinäkyvyyteen, jossa hallinto on valmis ottamaan kansalaiset mukaan ja kuuntelemaan heitä. Hallinnon avoimuutta lisääisi myös julkisen sektorin informaation, joka ei ole arkaluontoista, vapauttaminen julkiseen käyttöön. [Sun *et al.* 2015]

Käyttäjien osallistaminen onkin tärkeä osa hallinnon tulevaisuutta. Sähköisen hallinnon kautta käyttäjillä on suora kanava raportoida tyytyväisyydestään hallinnon palveluihin, osallistua käsiteltävistä asioista käytävään keskusteluun ja lisäksi vaikuttaa hallinnon toimintoihin. Osallistamisen ohella sähköinen hallinto parantaa yhteistyötä eri hallintojen, kansalaisten ja yritysten välillä. [Sun *et al.* 2015]

3.2. Sähköisen hallinnon jatkokehitys

Sun ja muut [2015] ehdottavat viitekehystä sähköisen hallinnon jatkokehittämiseen, joka koostuu neljästä osasta: prosessien integraatiosta (process integration), resurssien yhtenäistämisestä (resource integration), taustaprosessien integraatiosta (back-office integration) sekä käyttäjälle näkyvien prosessien integraatiosta (front-office integration).

3.2.1. Prosessien integraatio

Luodessa SH 2.0 on ymmärrettävä, että kyse on digitalisaatiosta, ei jo olemassa olevien prosessien sähköistämisestä. Digitalisaation myötä jo olemassa olevia prosesseja tulee tarkastella kriittisesti ja pohtia, kuinka ne voidaan toteuttaa sähköisillä välineillä tehokkaasti. Prosessien integraatiossa ensimmäinen askel onkin löytää käyttäjille tarpeelliset palvelut ja päättää, kuinka ne toteutetaan tehokkaasti uusilla menetelmillä. Käyttäjille tärkeitä prosesseja ja palveluita voidaan kutsua yleisiksi liiketoimintaprosesseiksi (Common Business Processes, CBPs). [Sun *et al.* 2015]

Ennen liiketoimintaprosessien suunnittelua ensimmäinen askel on luoda visio ja strategia sähköisten palveluiden tuottamisesta, jonka tulee olla kansallinen valtion sähköistä hallintoa luotaessa. Toiseksi on kuvattava korkeatasoinen suunnitelma, jossa kuvataan tällä hetkellä tarjottavat palvelut sekä projektien toteutumisen seuraamiseen ja arviointiin luotavat työkalut. [Sun *et al.* 2015]

Tämän jälkeen määritellään luodun vision ja tarpeen perusteella toteutettavat komponentit. Hallinnossa olevat laajat toimintamallit on pilkkottava pienemmiksi, hallittaviksi osiksi ja niille on tehtävä jako ensi- ja toissijaisten prosessien, sekä taustalla ja näkyvillä olevien prosessien välillä. Lopulta yleiset liiketoimintaprosessit käsittävät kokonaan sähköisen hallinnon, mutta on tärkeää ottaa huomioon siihen liittyneiden organisaatioiden määrät ja toimintatavat. Ne prosessit, jotka ovat eri organisaatioissa, mutta joilla on yhteinen päämäärä, on uudelleenorganisoidava. [Sun *et al.* 2015]

3.2.2. Resurssien integraatio

Resurssien integraatio on tärkeä osa hallinnon avoimuuden kehitystä. Sen kehittäminen on mahdollista laajentamalla kansalaisten oikeuksia hallinnolliseen tietoon, esimerkiksi sen vapaaseen uudelleenkäyttöön, uudelleenjulkaisuun, uudelleenmäärittämiseen sekä tiedon jalostamiseen. Tätä voidaan kutsua Informaatiotietohallinnan reformiksi (Information Resource Management Reform eli IRMR). Sen päätavoitteena on järjestää ja standardisoida kansalliset prosessit keräämällä, jakamalla ja lajittelemalla tietoa esimerkiksi muutettaessa yksittäistä

resurssienhallintoprotokollaa koko hallinnon kattavaksi ja uudistettaessa osastollisia toimintamalleja tai standardeja sekä väliohjelmistoja (middleware). Informaationhallintakäytäntöjen tulisi mahdollistaa ”luo kerran, käytä aina” -periaatteen. [Sun *et al.* 2015]

Resurssien yhdistämisessä suositellaan käytettävän palveluorientoituneita arkkitehtuuriratkaisuja (SOA) ohjaamaan datan ja ICT:n käyttöä läpi vaihtelevien kanavien. SOA koostuu palveluista, jotka kommunikoivat keskenään. Kommunikaatio sisältää joko yksinkertaista tiedonvälitystä tai monimutkaista koordinaointia, minkä vuoksi palveluiden yhdistäminen on tärkeää. SOA:n sopivaksi tekemiä tärkeitä ominaisuuksia ovat sen avoin rajapinta, jakautuvuus, prosessikeskeinen ja väljä kytkentäinen integroidakseen kaikki palvelut yhteen. Avoin rajapinta on erittäin tärkeä, sillä sen avulla on mahdollista liittää eri sovelluksia sekä palveluita yhteen saumattomasti. [Sun *et al.* 2015]

SH 2.0:ssa tulisi pyrkiä kehittämään ketterälle alustalle (agile platform), koska se on vaivattomampi kuin muut tarjolla olevat mallit. SOA on tarkoitettu toimimaan laajojen jakautuneiden järjestelmien kanssa. Skaalautuvuus, virheensietokyky sekä muutosten ja virheiden vaikutusten minimointi järjestelmässä ovat tärkeitä ominaisuuksia, jotka ylläpidettävän järjestelmän tulee sisältää. SOA:n yksi pääominaisuuksista on väljä kytkentä (loose coupling), eli yksittäisten elementtien vähäinen tieto toisista yksittäisistä elementeistä. Väljän kytkennän tarkoitus on vähentää riippuvuutta, jolloin muutokset tai virheet koskevat vain yhtä osaa järjestelmässä ja vaikutukset koko järjestelmään ovat vähäiset. [Sun *et al.* 2015]

SOA mahdollistaa eri virastojen itsenäisten palveluiden ja ohjelmien ketterät yhteydet välittämällä, reitittämällä ja hoitamalla kommunikointia niiden välillä. Hallinto ja sen useat virastot tarvitsevat myös yhtenäisen kehitys-, ylläpito-, valvonta- sekä hallintarajapinnan komponenttien välille. Sunin ja muiden [2015] mukaan vain SOA voi tarjota yhtenäisen yksittäisen rajapinnan kaikkiin nykyisiin ja tuleviin integraatiovaatimuksiin. Se vähentää riskiä virheisiin ja reagointiaikaa liiketoiminnan muuttuviin vaatimuksiin, nopeuttaa markkinointia ja alentaa omistajuuden kokonaishintaa sujuvoittamalla yritysten, kansalaisten ja hallinnon välistä byrokratiaa.

Web 2.0 auttaa luomaan hallinnollisesta informaatiosta saavutettavaa ja käytettävää. Se luo hallinnosta informatiivisen, osallistavan ja läpinäkyvän esimerkiksi verkosta löytyvien oppaiden tai sähköisten äänestyspalvelujen avulla. Lisäksi se auttaa muun muassa tuomaan esiin hallinnon sisällä innovatiivista kulttuuria, tukemaan yhteistyötä virastojen välillä ja luomaan arvostusta verkko- ja informaatiotekniikkaan liittyville aloitteille. [Sun *et al.* 2015]

Kokonaisuutena sähköisen hallinnon kehittäminen ja prosessien uudelleenjärjestely tarvitsevat täydellisen uudelleenajattelun siitä mitä informaatiotoiminnot hallinnossa tarkoittavat ja miten informaatioteknologian keinoilla ne voidaan toteuttaa. SH 2.0 luo uusia tiedonhallinnan tehtäviä, kuten auditoinnin ja hallintomekanismien uudelleenjärjestelyn sekä reaaliaikaisen tietovirran pankkeihin. Kehityksessä tulee ottaa huomioon kulujen ja hyötyjen suhde sekä erityisesti käyttäjien tarpeet. [Sun *et al.* 2015]

3.2.3. Taustaprosessien integraatio

Sähköisessä hallinnossa tulisi erottaa taustaprosessit sekä etualalla olevat näkyvät prosessit. Taustaprosessit ovat pyyntöjen käsittelyjä tai toimintojen rekisteröintiä, näkyvät prosessit taas ovat vuorovaikutusta asiakkaan kesken. Taustaprosessit yleensä tarvitsevat eri virastojen välistä tiedonvaihtoa. Lisäksi nykyiset SH 2.0 -projektit voivat sisältää sopimuksia, joissa hallitus ja yksityinen toimija tarjoavat eri tasoja projektin toiminnoista. Esimerkiksi joissain projekteissa hallitukset voivat olla vastuussa projektin taustalla olevasta verkkorakenteesta ja työvälineiden tarjoamisesta, kun taas tätä verkkoa käyttävä palvelu on yksityisen kumppanin vastuulla. Toisissa tapauksissa taas yksityinen kumppani voi olla vastuussa esimerkiksi verkkoinfrastruktuurin suunnittelusta, rahoituksesta, asennuksesta, rakentamisesta tai jopa hallinnoinnista yhdessä sähköisen hallinnon palvelujen kanssa. [Sun *et al.* 2015]

Taustaprosessien integraatiossa palvelujenhallinta-alusta vastaa vertikaalisten, horisontaalisten, ulkoisten ja sisäisten palveluiden yhdistämisestä. Palvelunhallinta-alustan rinnalla toimii sovelluksenhallinta-alusta, joka tukee, ylläpitää ja tehostaa virastojen tietojärjestelmien sovellusten toimintaa. Yhdistetty palvelunhallinta-alusta luo saumattoman tiedonsiirtokyvyn, alentaa hallinnointikustannuksia, parantaa palvelun laatua ja tehostaa toimintoja. [Sun *et al.* 2015]

Tehokkaan, avoimen ja luotettavan hallinnon luomiseksi olisi tärkeää poistaa eristäytyneitä rakenteita virastojen väliltä käyttämällä uusia yhteiskäyttöisiä työkaluja ja lähestymistapoja, joita Web 2.0 tarjoaa. Yhteiskäyttöisillä työkaluilla tarkoitetaan esimerkiksi wikejä eli tietoa sisältäviä sivustoja joiden rakennetta ja sisältöä voidaan muokata suoraan selaimesta, blogeja sekä reaaliajassa toimivia viestintäsovelluksia. [Sun *et al.* 2015]

3.2.4. Etualan prosessien integraatio

Kansalaiselle näkyvien toimintojen integraatio tarkoittaa verkkopalvelujen tuottamista ja yhtenäistämistä kansalaisille sekä yritykselle Internetin tai muiden digitaalisten keinojen kautta. Verkkopalveluita luodessa on varmistettava niiden esteettömyys ja helppokäyttöisyys. Lisäksi on huolehdittava, että palveluissa

oleva tieto on oikeaa ja päivitettyä. Helppokäyttöisyys lisää asiakastyytyvyyttä ja palvelun uudelleenkäytön mahdollisuutta. Lisäksi esimerkiksi hallinnollisten palvelujen oppaat, lomakkeet sekä verkossa tehtävät kyselyt ovat tärkeitä kansalaiselle verkossa olevia palveluita. Samalla hallinnon osastojen tulisi kerätä palautetta ja ehdotuksia sekä vastata niihin mahdollisimman nopeasti parantaakseen palveluitaan entisestään. Kun kansalaisella on käytössään laajat verkkopalvelut, määrittää niiden integraatio- sekä palvelutaso niiden onnistumisasteen. [Sun *et al.* 2015]

Yhteentoimivuus on myös tärkeä ominaisuus, joka mahdollistaa sovellusten toiminnan niiden käyttökanavasta riippumatta. Monikanavainen malli mahdollistaa pääsyn palveluihin usean reitin kautta, ja näin voidaan tarjota uusia sovelluksia kehittämättä täysin uusia kanavia. Jos kanavat ja taustaprosessit ovat integroituja, eri kanavat voivat tukea toisiaan parantaen palvelujen laatua ja toimintaa samaan aikaan. Tästä johtuen täydellinen tiedon, järjestelmien ja prosessien integraation omaava ympäristö on vaatimuksena tehokkaalle sähköiselle hallinnolle. [Sun *et al.* 2015]

Luotaessa sähköisen hallinnon järjestelmää tulee ymmärtää mitä on tarpeen tehdä ja kuinka. Tämä kuitenkin riippuu poliittisista, kulttuurisista, hallinnollisista, teknologisista ja taloudellisista ympäristöistä. SH 2.0 vaikuttaa koko hallinnon laajuisesti tarjoamalla integraation ja lisäksi parannuksia sen informaatioteknisille ratkaisuille. Yhteiset rakenteet mahdollistavat virastojen omien järjestelmien kommunikoinnin keskenään, mikä estää vakavien teknisten ongelmien ilmentymisen erilaisten ratkaisujen johdosta. Se takaa yhden luukun tarjota palvelut sidosryhmille eli yrityksille, järjestöille kansalaisille ja muille hallinnollisia palveluja käyttäville toimijoille. Järjestelmien standardisointi yli virastojen vähentää resurssihukkaa ja luo saumattoman sähköisen hallinnon 2.0 -arkkitehtuurin. [Sun *et al.* 2015]

4. Sähköisen hallinnon kehitysohjelmat Suomessa 2003-2016

Suomen sähköisen hallinnon kehittämiseksi on vuosina 2003-2016 tehty kolme hallitusvetoista hanketta: Arjen tietoyhteiskuntaohjelma 2003-2007, Sähköisen asioinnin ja demokratian vauhdittamisohjelma SADe-hanke 2008-2015 ja Digitalisoidaan julkiset palvelut 2015-2018. Kaikki hankkeet sisälsivät useita erillisiä projekteja joten pyrin tarkastelemaan hankkeita pääosin yleisellä tasolla. Viimeisimmässä Digitalisoidaan julkiset palvelut -hankkeessa keskityn tarkemmin yhden luukun palvelumallin luomiseen, jossa tärkeänä osana on Kansallinen Palveluarkkitehtuuri -hanke [Digitalisoidaan julkiset palvelut].

4.1. Arjen tietoyhteiskuntaohjelma 2003-2007

Vanhasen ensimmäinen hallitus julkisti syksyllä 2003 Hallituksen strategia-asiakirjassa (jatkossa HSA tai strategia-asiakirja) tietoyhteiskuntaohjelman, jonka tarkoituksena oli lisätä kilpailukykyä ja tuottavuutta. Tämän lisäksi sen tavoitteena oli lisätä sosiaalista ja alueellista tasa-arvoa, kansalaisten hyvinvointia ja elämänlaatua sekä säilyttää Suomen asema yhtenä tieto- ja viestintäteknologian johtavista tuottajista ja hyödyntäjistä. [Hallituksen strategia-asiakirja 2003]

Ohjelman avulla pyrittiin luomaan kansalaisille pääsy nopeiden tietoliikenneyhteyksien piiriin. Tämän lisäksi pyrittiin kehittämään kansalaisten tietoyhteiskuntavalmiuksia luomalla uusia toimintatapoja ja sähköistämällä valtion omia palveluja sekä hallintoa. Ohjelma pyrki myös edistämään vastaavaa kehitystä kuntasektorilla ja elinkeinoelämässä koulutuspoliittisilla toimenpiteillä, tutkimukseen ja tuotekehitykseen panostamalla, sekä tukemalla muita tietoyhteiskuntasektorin toimijoita. Kehitystä pyrittiin edistämään myös osallistumalla tietoyhteiskuntakehitystä edistäviin keskeisiin hankkeisiin. [Hallituksen strategia-asiakirja 2003]

Jaeger ja Thompson [2003] pitävät kriittisenä sähköisen hallinnon kehittämisen kannalta sitä, että kansalaiset osaavat käyttää teknologioita ja että heillä on mahdollisuus siihen. Tämä on taattava esimerkiksi laitteiden käyttömahdollisuuksilla julkisissa rakennuksissa, kuten kirjastoissa, sekä tarjoamalla koulutusta uusien teknologioiden käyttöön. Hallinnollisiin palveluihin ja tietoihin pääsyä ei voida evätä siksi, ettei kansalainen pystyisi käyttämään hallinnon sähköisiä palveluita.

Tietoyhteiskuntaohjelman vaikuttavuustavoitteessa sivuttiinkin näitä teesejä. Tavoitteisiin kuului muun muassa tietoyhteiskunnan palveluiden hyödyntäminen asuinpaikasta tai sosiaalisesta asemasta riippumatta, tietotekniikan perustaitojen sekä medialukutaidon hankkimisen mahdollistaminen ja sähköisen asioinnin tietoturvallisuus. Lisäksi suomalaisen tietoyhteiskunnan olisi oltava kansainvälisissä vertailuissa huippuluokkaa. [Hallituksen strategia-asiakirja, 2003]

Strategia-asiakirjan [2003] toteutussuunnitelmassa tietoliikenneyhteyksiä lähdettiin kehittämään niin, että vuonna 2005 nopeat tietoliikenneyhteydet olisivat olleet kaikkien kansalaisten saatavilla. Kansalaisten tietoyhteiskuntavalmiuden ja palveluiden käytön tehostamista hallitus halusi parantaa muun muassa vahvistamalla mediakasvatusta koulutuksen eri sektoreilla, luomalla tietotekniset perusvalmiudet sekä kehittämällä viestintäkampanjoita. Strategia-asiakirjan mukaan haluttiin tukea myös kouluihin ja kirjastoihin hankittavia asiakaspäätteitä, joissa olisi tarjolla Internetyhteydet. Jos kansalaiset eivät tiedä palveluiden saatavuudesta verkosta, eivät he myöskään niitä todennäköisesti käytä [Jaeger

and Thompson 2005]. Lisäksi heidän mukaansa sähköisen palvelun, sekä sen käyttämisen mahdollistamisen on oltava yleispalvelu (universal service), jotta kaikilla kansalaisilla on oikeus sähköisen hallinnon palveluihin.

Tutkimuksen ja tuotekehityksen osa-alueella toteutus suunnitelmassa haluttiin taata riittävä tarvittavan osaamisen saatavuus varaamalla rahoitusta tieto- ja viestintäteknologian koulutukseen. Organisaatioiden toimintatapojen muutosta edistettiin työelämän kehittämistoimilla. Lisäksi suunnitelmassa haluttiin kehittää esimerkiksi virtuaalista yliopistoa sekä turvallisuusviranomaisten koulutusta verkkotuettuun muotoon. [Hallituksen strategia-asiakirja 2003]

Julkishallinnossa tapahtuvaa sähköistä asiointia haluttiin kehittää kiinnittämällä huomiota kansalaisille tarjottaviin sähköisiin palveluihin ja niiden käyttäjälähtöisyyteen, helppokäyttöisyyteen sekä tietoturvaan. Strategiassa määriteltiin selvitystyön aloittaminen Internetin kautta tarjottavista valtionhallinnon ja kuntasektorin palveluista toteutusajankohtineen sekä arkkitehtuuriratkaisuihin. Tämän lisäksi tavoitteena oli jo sovittujen verkkopalveluiden määrittely ja käyttöönotto julkishallinnossa, Suomi.fi-portaalin jatkokehitys ja tunnettavaksi tekeminen sekä YritysSuomi.fi-portaalin kehittäminen interaktiivisten palvelujen suuntaan.

Sähköistä asiointia haluttiin kehittää myös laajentamalla Lomake.fi-palvelun käyttöä sekä jatkokehittämällä ja markkinoimalla TYVI-palvelua, eli tietovirrat yritysten ja viranomaisten välillä -palvelua. Tämän lisäksi kaikkien kirjastotyyppien yhteisen kansallisen tietoportaalien asteittainen kehittäminen, asevelvollisten palveluiden sähköistäminen, Suomen viisumijärjestelmän uusiminen ja kytkeminen osaksi Euroopan Schengen -viisumijärjestelmää sekä esimerkiksi XML-strategian toteuttaminen monen muun prosessin ohella, olivat strategiassa lueteltuja konkreettisia toimenpiteitä sähköisen asioinnin kehittämiseksi. [Hallituksen strategia-asiakirja 2003]

Sähköiseen liiketoimintaan, sisältöihin ja palveluihin liittyvistä hankkeista tärkeimpiä oli selvitys kansallisen sähköisen toiminnan nykytilasta ja tulevista haasteista. Selvityksen pohjalta laadittaisiin mahdollinen strategia ja toimeenpano uuden teknologian, osaamisen ja uusien toimintatapojen hyödyntämiseksi elinkeinoelämässä. Paikallisten, alueellisten ja kansallisten kehityskohteiden aloitteiden koordinointi on myös yksi huomio, jonka Jaeger ja Thompson [2003] nostavat esiin. Tehokkaassa sähköisessä hallinnossa eri tasojen kehitysprosesseja tulee miettiä koko prosessin kannalta, ei pelkästään yhden viraston ja sen tarpeiden pohjalta.

Hallituksen strategia-asiakirjassa [2003] haluttiin myös edistää sähköistä las- kutusta ja siihen liittyvää standardointityötä, tieto- ja viestintäteknologian sekä uusien liiketoimintamallien ja uusien työn organisointitapojen käyttöönoton

edistämistä pk-yrityksissä. Myös sähköisen palvelukeskustoiminnan edistäminen, selvitys julkisen sektorin tiedon uudelleenkäytöstä ja sen kaupallisesta hyödyntämisestä olivat sähköiseen liiketoimintaan liittyviä hankkeita.

Valtion omaa tietohallintoa, taloushallintoa sekä henkilöstöhallintoa pyrittiin kehittämään valtionhallinnon asianhallinnan sähköistämällä ja tietojärjestelmien yhtenäistämällä. Lisäksi valtion hankintatoimi haluttiin sähköistää, tietotekniikan hankintaohjeet täsmentää, hallinnollisten asioiden tiedonsiirto toteuttaa kaupallisten yritysten palveluilla ja turvallisuusviranomaisten sisäinen tiedonsiirto toteuttaa valtion omistaman tiedonsiirtoyhtiön palveluilla. Lisäksi valtion tietohallintoon liittyviä hankkeita olivat esimerkiksi perusrekisterin käytettävyyden ja yhteiskäytön kehittäminen sekä poikkeusoloihin valmistautumisen ohjeistaminen. [Hallituksen strategia-asiakirja, 2003]

Layne ja Lee [2001] toteavat, että teknisesti heterogeenisten tietokantojen integraatio eri virastojen välillä vaatii mahdollisesti teknisiltä toteutuksiltaan ristiriitaisten tietojärjestelmien yhteen liittämiseen liittyvien ongelmien ratkaisua. Tämän toteuttaminen on monelle valtiolle kompastuskivi, eikä kyse ole pelkästään tekniikasta, vaan myös yhdistymisprosessien johtamisesta. Horisontaalinen integraatio vaatii myös muutosta hallinnon johtajien ajattelutavassa. Heidän tulisi päästä pois ajatuksesta, että heidän oman virastonsa tietojen vaihtotarpeet ovat tärkeimmät, eikä muilla toimijoilla ole niin suurta merkitystä. Internetin myötä yksittäin erikoistuneet hallinnon prosessit eivät ole välttämättä tehokkaita, vaikuttavia tai kansalaisystävällisiä. Laynen ja Leen [2001] mukaan toiminnallinen erikoistuminen ei ole sopiva hallinnointirakenne sähköisessä hallinnoinnissa.

Suomalaista tietoyhteiskuntaa rakentamassa -loppuraportin [2007] mukaan hanke onnistui lisäämään kansalaisten Internetin käyttöastetta sekä käyttötaitoja luomalla kotitalouksille mahdollisuuden päästä verkkoon tai lisäämällä julkisten asiakaspäätteiden määrää. Lisäksi hanke onnistui lujittamaan suomalaisten luottamusta sähköisen asioinnin palveluihin. Raportin mukaan Suomi oli OECD-vertailun kärkipäässä vertailtaessa yritysten Internetin hyödyntämistä. Lisäksi työntekijöiden tietoyhteiskuntaosaaminen oli raportin mukaan korkealla. Suomi oli myös pärjännyt hyvin vuoden 2006 kesäkuussa julkaistussa kansainvälistä sähköistä hallintoa koskevassa vertailussa, kuten myös kansainvälisissä kilpailukyky- ja tietoyhteiskuntavertailuissa. Yhteensä toimenpiteitä ja hankkeita ohjelmassa oli noin 150.

Valtiontalouden tarkastusvirasto suoritti vuonna 2008 toiminnankertomuksen alueellisten tietoyhteiskuntahankkeiden toteutuksesta. Siinä todetaan, että etenkin julkisen hallinnon sähköisen asioinnin edistämisen eli JUPA-hankkeen

toteutusvaiheelle asetetut tavoitteet olivat epärealistisia suhteessa asetettuun aikatauluun, käytettävissä oleviin henkilöstöresursseihin, toteutettuun hankeohjausmalliin ja käytössä olleeseen rahoitukseen nähden. Lisäksi toiminnankertomus toteaa, että ohjauksesta vastanneet toimijat eivät olleet sitoutuneet riittävästi hankkeen tavoitteisiin, mikä hankkeen toteutusvaiheessa johti alueellisten projektien tavoitteiden karsimiseen. Tästä johtuen projektissa saavutetut tulokset jäivät keskeneräisiksi ja vaatimattomiksi alkuperäisiin tavoitteisiin nähden. Tarkastuksen kokonaisarvio on se, että JUPA-hanke alueellisten sähköisten palveluiden kehittämistä edistävänä hankkeena jäi tuloksiltaan vaatimattomaksi. [Alueellisten tietoyhteiskuntahankkeiden toteutus]

4.2. SADe – Sähköisen asioinnin ja demokratian vauhdittamisohjelma 2008-2015

Valtiovarainministeriö asetti keväällä 2009 sähköisen asioinnin ja demokratian kehittämishanke -nimisen ehdotuksen pohjalta samannimisen vauhdittamisohjelman (SADe-hanke). Sen tarkoituksena oli parantaa Suomen tietoyhteiskunta- ja asiointipalvelujen kehitystä kolmella toimintakokonaisuudella. Ensimmäinen oli *”jatkaa arjen tietoyhteiskuntaohjelmaa”*, toinen *”käynnistää sähköisen asioinnin ja demokratian ohjelma”*, sekä kolmas *”uudelleen organisoida julkisten tietojärjestelmien määrittely, hankinta ja ylläpito”*. [SADe toteuttamissuunnitelma 2009]

SADe-hankkeelle asetettiin tiukat vaatimustavoitteet. Pää tavoite oli saada vuoteen 2013 mennessä kansalaisten ja yritysten sähköinen asiointi kattamaan kaikki keskeiset palvelut. Sähköisen asioinnin todentaminen tehtiin seuraamalla erilaisia mittareita, joihin kuuluivat esimerkiksi kehitettävien ja käyttöönotettujen palvelukokonaisuuksien määrä, hallinnonala- tai virastokohtaisten sähköisten palveluiden suunnitelmat ja niiden määrät, tai vuosittaiset kansalaisen asiointitilin käyttömäärät. SADe-ohjelma kesti yhteensä kuusi vuotta, jonka aikana tehtiin seitsemän hanketta ja luotiin 42 palvelua. Ohjelman kustannukset olivat 70 miljoonaa euroa. [SADe toteuttamissuunnitelma 2009]

Toinen tavoite oli kehittää julkisen hallinnon tietojärjestelmien yhteentoimivuutta luomalla yhtenäiset tietojärjestelmäarkkitehtuurit ja valtakunnalliset tietojärjestelmät. Jaeger ja Thompson [2003] toteavat, että paikallisten, alueellisten ja kansallisten kehityskohteiden aloitteiden koordinointi on tärkeää tehokkaassa sähköisessä hallinnossa. Heidän mukaan eri tasojen kehitysprosesseja tulee miettiä koko prosessin kannalta, ei pelkästään yhden viraston ja sen tarpeiden pohjalta SADe-ohjelmalla tuli myös aikaansaada kustannussäästöjä sekä resurssien vapauttamista prosessien ja toimintojen kehittämiseen niiden suorittamisen sijasta [SADe toteuttamissuunnitelma 2009]. SADe-ohjelman loppuraportin [2015] mukaan poikkihallinnollisten palvelukokonaisuuksien toteuttaminen epäonnistui, sillä kehitettävät palvelut valittiin virasto- tai hallinnonrajoja noudattamalla.

Raportti toteaa kuitenkin, että SADe loi mahdollisuuden jatkokehitykselle tulevaisuudessa.

Jaegerin ja Thompsonin [2003] mukaan olisi myös tärkeää kehittää mittareita, joilla sähköisen hallinnon suorituskykyä pystytään tarkastelemaan. Niiden tulisi ottaa huomioon kansalaisen tarpeet löytää, ymmärtää ja käyttää heille tarpeellista tietoa. Arvioiden tuottamien johtopäätösten perusteella tulisi löytää myös positiivisia ominaisuuksia jotka lisäävät sähköisten palvelujen käyttöä.

SADe-toteuttamissuunnitelman [2009] mukaan projektissa mittareilla seurattavia asioita olivat olemassa olevien palveluiden kattavuus vuoden 2010 loppulla ja tukipalveluita käyttävien käyttäjäorganisaatioiden määrä, eli kuinka moni organisaatio käyttää esimerkiksi Vetumaa ja sähköisiä kilpailuttamisen välineitä. Lisäksi kriteerinä mainittiin vielä ”JHS-suositusten määrä ja vastaavuus tarpeisiin, sekä niiden hyödyntämisen laajuus”.

Vetuma on kansalaisen tunnistus ja maksamispalvelu [Vetuma] ja JHS-suosituksilla tarkoitetaan valtion- ja kunnallishallinnon tietohallintoa koskevia menettelytapoja, määrittelyitä tai ohjeita. JHS-suositukset hyväksyy JUHTA eli julkisen hallinnon tietohallinnon neuvottelukunta ja suositukset laatii sen alainen asiantuntijajasto [JHS-suositukset]. SADe-loppuraportin [2015] mukaan ohjelman suunnittelu ja raportointi olivat ylimitoitettua eikä niillä ollut hankkeiden tuloksellisen etenemisen kannalta parantavaa vaikutusta. Lisäksi tavoitteiden ja raportoitavien asioiden välillä oleva yhteys jäi heikoksi.

Kahden päätavoitteen lisäksi toteuttamissuunnitelman [2009] lisätavoitteena oli, että koko valtionhallinto ottaa kaikki laskut vuoden 2009 loppuun mennessä vastaan sähköisenä. Lisäksi tavoitteena oli edistää videoneuvotteluyhteyksiä ja etäpalvelujen käyttöä koko yhteiskunnassa, minkä avulla julkisen hallinnon kotimaan matkoista oli tarkoitus karsia 20 prosenttia vuoteen 2013 mennessä.

Valtiovarainministeriön SADe-tuloskoosteen [2015] mukaan ohjelman rahallinen hyöty syntyi pääosin prosessien sujuvoittamisesta sekä täytettyjen tietojen virheettömyydestä. Parhaiten toteutuneita palveluita olivat esimerkiksi Opintopolku.fi- sekä Kansalaisaloite.fi-palvelut. Riippumattoman SADe-arvioinnin [2015] mukaan ohjelma vahvisti sähköisten palveluiden merkitystä julkishallinnossa sekä loi uudenlaista kulttuuria, jossa sähköisillä palveluilla on entistä suurempi painoarvo koko hallinnon kehittämisessä. Tästä huolimatta ohjelmassa oli loppuraportin mukaan osallisena useita palveluita, jotka vastasivat heikosti ohjelman tavoitteita ja näillä kustannussäästöjä vain vähäisesti tai ei lainkaan.

4.3. Kansallinen palveluarkkitehtuuri 2015-2018

KaPA:n kolme pääosaa ovat palveluväylä, palvelunäkymät sekä kansallinen sähköinen tunnistautuminen. Muita osia ovat tietovarantojen lisääminen osaksi arkkitehtuurirakennetta sekä niiden yhteen toimivuuden varmistaminen. Tämän lisäksi toteutukseen kuuluvat yritysten, muiden organisaatioiden ja henkilöiden roolien ja valtuutusten hallintaan liittyvät toiminnot. Kokonaisuuden täydentää tarvittavien lainsäädäntötoimien valmistelu sekä hallintamallien luomisesta huolehtiminen palveluarkkitehtuurin ylläpitämiseksi ja valvomiseksi. Kansallisella palveluarkkitehtuurilla on tiukat vaikuttavuustavoitteet, kuten aiemmillaikin kehittämisohjelmilla. [Palveluarkkitehtuuri]

Sähköisen hallinnon järjestelmien tulee olla keskenään toimivia, eli niissä käytettävät järjestelmät, ohjelmistot tai tekniikat eivät saa olla vanhentuneita tai keskenään sopimattomia. Esimerkiksi keskenään erilaiset tiedostoformaatit tai toimintatavat voivat aiheuttaa ongelmia, jos tiedostojen muokkaus toisella sovelluksella ei onnistu tai sitten käyttäjät joudutaan erikseen kouluttaa käyttämään erilaisia järjestelmiä. Järjestelmien yhteensopimattomuudesta aiheutuu ylimääräistä työtä niin viranomaisille kuin palvelujen käyttäjillekin. Lisäksi hallinnon on pystyttävä tarjoamaan kattava infrastruktuuri niin verkkoyhteyksien kuin laitteidenkin muodossa, jotta uutta teknologiaa ja viestintävälineitä pystytään käyttämään. [Sun *et al.* 2015]

4.3.1. Kansalaisten, yritysten ja yhteisöjen asiointi viranomaisten kanssa

Kansallisen palveluväylän tarkoituksena on helpottaa kansalaisten, yritysten ja yhteisöjen asiointia viranomaisen kanssa sekä muuttaa sitä tietoturvallisemmaksi [Perustiedot]. Suomessa kansalaisten mahdollisuudet käyttää sähköisen hallinnon palveluita ovatkin erittäin hyvät ja Euroopan huipputasoa. Negreiro [2015] siteeraa Eurostatin tilastoja vuodelta 2014, joiden mukaan Suomessa on mahdollisuus Internetyhteyteen 89 %:lla väestöstä. Vakituksia Internetin käyttäjiä Suomessa on noin 90 % väestöstä; sellaisia, jotka eivät koskaan käytä Internetiä on noin 5 % ja satunnaisia Internetin käyttäjiä loput. Onkin tärkeää, että digitaalinen kuilu valtion sisällä on mahdollisimman pieni.

Digitaalinen kuilu (digital divide) aiheutuu siitä, että kaikilla kansalaisilla ei ole osaamista tai mahdollisuutta opiskella toimintaa IT-ympäristöissä. Tilannetta, jossa on kansalaisia joilla on tietotekniikka- ja etenkin Internet-taidot sekä niitä, joilla niitä ei ole, kutsutaan digitaaliseksi jakautumiseksi. [Sun *et al.* 2015]

Sun ja muut [2015] toteavat myös, että kehitettäessä sähköistä hallintoa on sillä joko mahdollisuus tasoittaa digitaalisen jakautumisen eroja ja mahdollistaa kansalaisten parempi osallistuvuus tai sitten kasvattaa niitä ja lisätä esteitä. Huonon informaatioteknologian omaavissa olosuhteissa olevat kansalaiset eivät

pysty oppimaan tärkeitä tietokonetaitoja eivätkä pääse osaksi sähköisen hallinnon hyötyjä. Niinpä hallinnon onkin tärkeää luoda sellaiset olosuhteet, että kaikilla kansalaisilla on mahdollisuus hyötyä sen tarjoamista palveluista.

Van Dijk ja Hacker [2011] havainnollistavat neljä toimintamallia, jotka lisäävät digitaalista kuilua. Ensimmäinen on teknologiakielteisyys, josta aiheutuu tietoteknisen kokemuksen puute. Toinen on laitteiden puute, kolmas taitojen puute huonon koulutuksen tai olosuhteiden vuoksi ja neljäs käyttömahdollisuuksien puute. Van Dijk ja Hacker [2011] huomauttavat, että digitaalinen kuilu ei ole yksiselitteinen eikä staattinen ilmiö, vaan koostuu useasta, olosuhteista riippuvista tekijöistä ja muuttuu koko ajan, esimerkiksi uusien laitteiden johdosta.

Lisäksi Jaeger ja Thompson [2003] huomauttavat, että kun sähköisen hallinnon palvelut on luotu, on kansalaisia tiedotettava niistä. Elleivät kansalaiset tiedä palveluiden saatavuudesta verkosta, eivät he myöskään niitä todennäköisesti käytä. Lisäksi verkossa olevan materiaalin tulee olla ajantasaista ja hyödyllistä. Sähköisen palvelun on oltava yleispalvelu (universal service), jotta kaikilla kansalaisilla on oikeus sähköisen hallinnon palveluihin.

Palvelujen laadun takaaminen ja hallinnon avoimuuden kehittäminen on toinen tärkeä vaikuttavuustavoite. Suurella osalla kansalaisista ei ole tarkkaa ymmärrystä siitä, kuinka hallinnollisia päätöksiä tehdään tai hallinto toimii. Tämä vaikuttaa hallinnon läpinäkyvyyteen ja saavutettavuuteen. Sähköinen hallinto pystyy kuitenkin onnistuessaan parantamaan näitä demokratian kannalta tärkeitä asioita. Lisäksi hallinnon on otettava huomioon kaikki kansalaiset, riippumatta heidän fyysisestä kyvykkyydestään, eli verkkopalvelujen on oltava esteettömiä. [Sun *et al.* 2015]

Valtion monikielisyys on myös huomioon otettava tekijä. Kaikki sähköiset palvelut pitää tarjota valtion virallisilla kielillä. Ylipäätään viestintä ja kommunikointi asiakkaiden kanssa tulisi olla sujuvaa ja aktiivista, sillä sähköisessä viestinnässä asiakas on helpompi jättää huomiotta. Tähän liittyen on myös varmistettava, että hallinnon työntekijät eivät käytä teknologiaa välttelemään vastuuta tekemistään. Esimerkiksi huonoja työtuloksia tai aikataulun venymisiä ei voida perustella teknologisilla haasteilla, jos ne ovat perusteettomia. [Jaeger and Thompson 2003]

Palveluiden esteettömyys on taattava eli esimerkiksi kansalaiset, joilla on jokin vamma, joka vaikeuttaa oppimista tai ylipäätään tietotekniikan käyttöä, on myös otettava huomioon. Heille on mahdollistettava tasavertaiset hallinnolliset palvelut. [Jaeger and Thompson 2003]

4.3.2. Sähköisten palvelujen kustannustehokkuuden parantaminen

Sähköisten palvelujen kustannustehokkuuden ylläpitäminen koko niiden elinkaareksi on myös Valtiovarainministeriön hankkeelle asettama tavoite. Koko julkista hallintoa koskien tavoitteena on tietojen ja tietojärjestelmien yhteen toimivuuden parantaminen. Lisäksi tavoitteena on edistää yritysten mahdollisuuksia julkisiin tietovarantoihin ja palveluihin sekä tukea kansantaloutta julkisen hallinnon tehostamisella ja uusien liiketoimintamahdollisuuksien luomisella yksityiselle sektorille. [Perustiedot]

Sun ja muut [2015] toteavat, että SH 2.0:ssa mallin kestävyys ja sen kulu-rakenteen suunnittelu ja budjetointi on haastavaa, koska toimintaympäristö muuttuu jatkuvasti. Kestävän ratkaisun saavuttaminen on mahdollista vain kehittämällä bisnesmalli tai tekemällä kustannus-hyötyanalyysi, jossa hallinto varmistaa, että investoitava ohjelma on kestävä ja tuottaa varmasti säästöjä. [Sun *et al.* 2015]

He nostavat esille myös, sisällönhallintajärjestelmän (CM) tai viitekehyksen. Heidän mukaan se helpottaa viranomaisten tehokasta hallinnointia ja auttaa heitä tunnistamaan hallinnoinnissa sillä hetkellä olevia esteitä. CM toimii viranomaisten tukena auttaen heitä tekemään hyödyllisiä analyyskejä ja ymmärtämään eri sosiaalisten tai taloudellisten kehitystrendien vaikutuksia koko hallintoon. Smith ja McKeen [2003] toteavat, että sisällönhallintajärjestelmillä (ECM, Electronic Content Management) ei ole tarkasti vakiintunutta määritelmää, mutta niitä voidaan kuvata esimerkiksi seuraavasti: *"Ne ovat strategioita, työkaluja, prosesseja ja taitoja jotka organisaatio tarvitsee huolehtiakseen kaikista sen informaatiotyypeistä koko niiden elinajan."*

4.3.3. Palveluväylä

Palveluväylä toteutetaan teknisesti Virossa käytössä olevan X-Roadin pohjalta ja samoilla periaatteilla [Palveluväylä]. X-Road on luotu Virossa jo vuonna 2001 [X-Road] ja siitä on nyt menossa kuudes kehitysversio, joka tulee myös käyttöön Suomessa. [Kalja *et al.* 2015]. Se on eri hallinnonalojen tietokannat ja eri organisaatiot tietoturvalisestisesti yhdistävä tekninen ympäristö. Se mahdollistaa turvallisen tietojenvälityksen niin yritysten ja hallinnon kuin yksittäisten henkilöiden ja hallinnon välillä. [Kalja *et al.* 2015]. Palveluväylä koostuu tiedonvälityskerroksesta, joka määrittää, miten tietoja ja palveluja välitetään eri tietojärjestelmien välillä, sekä tiedonvälityspalvelusta. Tiedonvälityspalvelu mahdollistaa muiden siihen liittyneiden organisaatioiden palvelujen ja tietovarantojen hyödyntämisen liityntäpalvelimien sekä palvelun keskuspalvelimien välillä. [Palvelukuvaus]

Sun ja muut [2015] huomauttavat, että hallitsevat organisaatiot ovat vastuussa valtavana määrästä käyttäjien henkilökohtaisia tietoja. Ne keräävät sidosryhmiltään valtavan määrän dataa päivittäisissä toiminnoissa, jolloin niillä on vastuu käyttää ja suojella käyttäjien yksityisiä tietoja vastuullisesti. Lisäksi tietoturva on otettava huomioon jo suunnitteluvaiheessa sen kustannuksista huolimatta, sillä ilman sitä kansalaiset eivät ota palveluja käyttöön. Projektien tulee luoda luottamusta virastojen, heidän hallintojen sekä sidosryhmien välillä, eikä päinvastoin tietoturvaongelmien johdosta.

Julkinen palveluväylä sisältää eri vyöhykkeitä, jotka ovat rajattuja, mutta tietyllä kriteerillä yhdistettyjä kokonaisuuksia. Vyöhykkeet rajaavat palveluväylää eri verkkoihin, kuten esimerkiksi VY-, KY- tai TUVE-verkkoihin [Palvelukuvaus]. VY-verkko, eli valtion yhteiset tietoliikennepalvelut tarjoaa valtionhallinnon organisaatioille luotettavat ja turvalliset yhteydet muihin valtionhallinnon palveluntarjoajiin ja organisaatioihin. Tietoturvapalvelut sisältävät tietoturvaohjeiden havainnointi- ja estojärjestelmän, palomuuripalvelut, haittaohjelmien ja roskapostin suodatuksen sekä palvelunestohyökkäyksiä torjunnan. [VY-verkko]

KY-verkolla taas tarkoitetaan kuntien keskitettyä verkkoratkaisua, joka yhdistää eri kuntien tietoliikenneverkot toisiinsa. Se mahdollistaa kuntien kytkeytymisen valtion virastoihin, laitoksiin, tietovarantoihin ja tietopalveluihin. Lisäksi se mahdollistaa kuntien kytkeytymisen toisiinsa Internetin välityksellä, muodostaa tietyille ryhmälle alueverkkoja tai muita yhteenliittymiä, tarjota toisilleen palveluita sen kautta, sekä muodostaa toisistaan riippumattomia verkkoja. [KY-verkko]

TUVE-verkolla tarkoitetaan julkisen hallinnon turvallisuusverkkoa, joka on käytössä valtion ylimmällä johdolla ja yhteiskunnan turvallisuuden kannalta tärkeillä viranomaisilla. Viranomaisverkko täyttää korkean turvallisuuden vaatimukset ja siihen kuuluvat viestintäverkko, laitetilat ja laitteet sekä yhteiset tieto- ja viestintätekniset palvelut [Turvallisuusverkkotoiminta]. Verkkoa käyttävät esimerkiksi tasavallan presidentin kanslia, tietyt ministeriöt, puolustusvoimat, poliisi ja rajavartiolaitos [TUVE-laki].

Palveluväylän toiminta perustuu siihen liittyneiden järjestelmien tarjoamien palveluiden julkaisemiseen SOAP-rajapintoina. Järjestelmien väliset palvelupyynnöt ja vastaussanomien kulkevat salattuina SOAP-sanomina liityntäpalvelinten välillä [Sovitinpalvelu]. SOAP eli Simple Object Access Protocol on XML-kieleen pohjautuva tietoliikenneprotokolla, jota X-Road-ratkaisun liityntäpalvelimen rajapintatoteutuksessa käytetään. SOAP toimii useiden eri protokollien yli, mutta X-Roadin tapauksessa sitä käytetään vain http-protokollan yli. [Sanasto]

SOAP-pohjaisten järjestelmien liittäminen palveluväylään tarkoittaa yksinkertaisimmillaan liityntäpalvelimen pystyttämistä, sekä sovitinpalvelun toteuttamista tarvittavien SOAP-otsikkotietojen lisäämiseksi. Sovitinpalvelun on myös tarjottava liitettävien palveluiden rajapintakuvaukset eli WSDL:t. REST-mallisten järjestelmien ja palveluiden tulee myös luoda SOAP-sanomamuotoinen yhteys käytettävän viestirajapinnan kautta, jotta liittyminen on mahdollista. Koska palvelu ei ole suoraan REST-yhteensopiva, jokaista palveluväylään liitettävää yksittäistä REST-palvelua varten on toteutettava vastaava SOAP-sovitin, jolloin muunnoksen tekeminen on sovitinpalvelun vastuulla. [Sovitinpalvelu]

REST on järjestelmäarkkitehtuuri, joka pyrkii minimoimaan viiveen ja verkon kommunikaation samalla, kun se säilyttää tehokkaasti viestien itsenäisyyden ja skaalautuvuuden. Määritelmä saadaan aikaan luomalla rajoitteita kiinnikkeille. REST mahdollistaa viestien tallentamisen välimuistiin ja niiden uudelleenkäytön, dynaamisen komponenttien vaihdettavuuden sekä tapahtumien prosessoinnin kolmansien osapuolien avulla. [Fielding 2000]

5. Sähköisen hallinnon kehittämisen onnistuminen ja sen tulevaisuus

Kun arjen tietoyhteiskuntaohjelma aloitettiin 2003, oli sähköinen hallinto silloin katalogisointi- ja transaktiomallin välimaastossa, jossa suurin osa tiedoista löytyi verkosta, mutta vuorovaikutusta hallinnon kanssa verkon kautta oli vain vähän. Ohjelman pyrkimyksenä oli lisätä tätä vuorovaikutusta sekä luoda lisäksi vertikaalista integraatiota eri virastojen välille. Vertikaalisen integraation kehitystä jatkettiin ja horisontaalista pyrittiin aloittamaan SADe-ohjelman yhteydessä. KaPA keskittyy lähes pääsääntöisesti horisontaaliseen integraatioon ja pyrkii tuomaan kaikki sähköisen hallinnon palvelut yhden luukun taakse.

Sähköisen hallinnon kehittämiseksi tehtyjen projektien pyrkimykset ovat olleet tutkitun tiedon mukaisia ja edenneet oikeassa järjestyksessä. Toisaalta, voidaan pohtia olisiko ollut järkevää luoda kansallinen palveluväylä jo alkuvaiheessa ikään kuin selkärangaksi, jonka päälle palvelut olisi rakennettu, kuten Virossa toimittiin. Arjen tietoyhteiskuntaohjelmassa ja SADe-hankkeessa luotiin ensin palvelut, ja sitten nyt vasta KaPA:ssa ne yhdistävä tekninen ratkaisu. Jos palveluväylä olisi luotu aiemmin, olisi nykyisiä palveluita pystytty käyttämään tehokkaammin jo useita vuosia sitten sekä olisi voitu välttää kahden erilaisen tunnistuspalvelun luominen. Ennen sähköisen hallinnon palveluihin tunnistautumiseen käytössä ollut palvelu Vetuma poistuu käytöstä uuden Suomi.fi-tunnistautumisen tieltä.

Sähköisen hallinnon kehittämisessä riittää haasteita, joista suurimpana on kehittämisprojektien johtaminen. Hankkeet eivät aina ole saavuttaneet vaikuttavuustavoitteitaan, koska niillä on ollut päällekkäisiä tavoitteita eli projektit on valittu heikosti tai ne on koordinoitu valtakunnallisella tasolla huonosti. Tämä on toisaalta tutkimustiedon mukaista, sillä usein eri virastojen välisen horisontaalisen integraation toteuttaminen on haastavaa johtuen teknisistä, mutta etenkin johtamiseen liittyvistä ongelmista [Layne and Lee 2001]. Lisäksi esimerkiksi Sunin ja muiden [2015] mukaan sähköisen hallinnon kehitysprosessissa on tärkeää arvioida jatkuvasti kriittisesti, että ovatko tavoitteet saavutettavissa sen hetkisen projektiaikataulun puitteissa. Lisäksi tietotekniikan kehityksen vauhti eli koko ajan muuttuva toimintaympäristö asettaa rajoitteita valtiovetoiselle kehitystoiminnalle ja saattaa hidastaa sitä.

Palveluväylässä käytettävät tietotekniset ratkaisut ovat standardien mukaisia, joten ne kestävät varmasti tulevaisuuden kehityksen. X-Roadissa käytettävät teknologiat ovat yleisesti suositeltuja ja muissakin vastaavissa järjestelmissä käy-

tössä olevia teknologioita. Näin ollen tämän ei pitäisi muodostua ongelmaksi tulevaisuudessa. Lisäksi standardisoidut ratkaisut tarkoittavat sitä, että ne ovat tietoturvallisia. Tietoturvallisuuskäytännöistä tarkasteltuna tietojen tallentaminen hajautettuihin tietokantoihin estää esimerkiksi tilanteen, jossa ulkopuolinen kolmas taho voisi päästä käsiksi yhdellä kertaa kaikkiin tietokannassa olevan toimijan tietoihin.

Sunin ja muiden [2015] mukaan uuden teknologian käyttöönotossa voidaan myös törmätä lainsäädännöllisiin ongelmiin, joten kehitysprosessin aikana tai sitä ennen tulisi lainsäädäntöä tarkastaa niin, että se mahdollistaa sähköiseen hallintoon liittyvät prosessit, kuten sähköisen vuorovaikutuksen, sekä sähköiset dokumentit. Suomessa tehtyjen kehitysprojektien lähtökohtana on ollut yleensä jokin projektia ennen säädetty laki tai ellei näin ole ollut, on projektin alkuvaiheessa tai sen aikana pyritty selvittämään lainsäädännölliset esteet ja korjaamaan ne.

Layne ja Lee [2001] toteavat, että horisontaalinen integraatio antaa valtiolle ja mahdollisesti yrityksille suhteessa enemmän tietoa kuin kansalaiselle. Siksi on tärkeää, että kansalainen ei koe oloaan tarkkailluksi. Hänelle ei myöskään saisi tulla kokemusta, että hänestä pienissä osissa kerättävää sähköistä tietoa käytettäisiin tähän tarkoitukseen. Kansalaisen tulisi säilyä kontrollissa hänestä ulkopuolisille näkyvistä tiedoista. Palveluväylässä esimerkiksi huoltajalla voi olla oikeudet lapsensa käyttöoikeuksiin, tai yhdistyksen nimenkirjoittajalla yhdistykseen, mutta muuten tietoturva, yksityisyys sekä käyttöoikeudet on määritelty tarkasti KaPA-laissa.

Tutkimusta tehdessä oli mielenkiintoista huomata, että verrattaessa tutkimustuloksia vuosituhanen alusta, eli lähinnä Laynen ja Leen [2001] aivan viimeaikaiseen tutkimukseen Sun ja muut [2015] oli siellä paljon yhteneväisyyksiä.

Tulevaisuudessa sähköisen hallinnon kehittämistä tulee viedä yhä enemmän Web 2.0 -palvelujen tarjoamien mahdollisuuksien suuntaan. Etenkin hallinnon läpinäkyvyys, kansalaisten osallistaminen sekä digitaalisen kuilun kaventaminen ovat tärkeitä huomioita, jotka tulee pitää mielessä jatkokehityksessä. Suomessa digitaalinen kuilu on suhteellisen pieni, mutta koska uusia teknologioita kehittyi koko ajan, on siihen kiinnitettävä jatkuvasti huomiota koulutuksella ja ajantasaisella laitteistolla. Palvelujen helppokäyttöisyys ja esteettömyys on taattava myös tulevaisuuden projekteissa. Organisaatioiden joustavuuden ja niiden välisen kommunikoinnin kehittäminen parantaa hallinnon taustaprosessien toimivuutta. On myös huomioitava, että digitalisaation myötä on kokonaisia toimintaketjuja tarkasteltava uudelleen alusta loppuun, ei pelkästään siirtää jo olemassa olevia toimintamalleja sähköisiksi.

Lisäksi tulevaisuudessa on verkkoyhteyksien nopeus, laatu ja hinta pyrittävä pitämään sopivana sekä tasapuolisena ympäri maan, jotta kaikilla kansalaisilla olisi pääsy sähköisen hallinnon piiriin. Palveluorientoituneiden ratkaisujen sekä ketterien alustojen käyttäminen on ensisijaisen tärkeää. Sähköisen hallinnon palvelujen käyttäminen eri laitteilla, kuten mobiililla, tulee myös olla mahdollista.

Palveluväylä on ollut toiminnassa Virossa jo yli vuosikymmenen. Siitä odotetaan digitalisaation vauhdittajaa myös Suomessa, koska se tarjoaa tietokannat useille eri toimijoille yhden palvelun takaa ja mahdollistaa näin tarvittavan tiedon hakemisen eri lähteistä. Se mahdollistaa asiakkaalle yhdellä tunnistautumisella pääsyn tietoihin, joihin hänellä ei välttämättä ole ennen ollut mahdollisuutta.

X-Road palveluväylä vaikuttaa lupaavalta ja on jo Virossa usean vuoden aikana toimivaksi todettu. Suomen ottaessa sen käyttöön, ovat Suomen ja Viron tietoliikenneyhteydet helpompi liittää toisiinsa. Tulevaisuudessa olisikin Suomen kannalta järkevää saada X-Road palveluväylä koko Euroopan Unionin laajuiseksi ratkaisuksi. Vielä pidemmällä aikavälillä pyritään varmasti löytämään maailmanlaajuinen ratkaisu yhdistää valtioiden sähköiset hallinnot toisiinsa.

Viiteluettelo

- Alueellisten tietoyhteiskuntahankkeiden toteutus, Valtiontalouden Tarkastusvirasto, 2008. Tarkistettu 2.12.2016. http://www.vtv.fi/julkaisut/tulokselli-suustarkastuskertomukset/2008/alueellisten_tietoyhteiskuntahankkeiden_toteutus.4044.xhtml.
- Digitalisoidaan julkiset palvelut, Valtiovarainministeriö. Tarkistettu 22.10.2016. <http://vm.fi/digitalisoidaan-julkiset-palvelut>.
- Dijk, J. v., & Hacker, K. (2003). The digital divide as a complex and dynamic phenomenon. *The Information Society*, 19(4), 315-326.
- Roy Thomas Fielding. 2000. *Architectural Styles and the Design of Network-based Software Architectures*. Ph. D. Dissertation, University of California, Irvine.
- Fielding, R., & Reschke, J. Hypertext transfer protocol (HTTP/1.1): Message syntax and routing. Retrieved 14.11.2016 from <https://tools.ietf.org/html/rfc7230>.
- Hallituksen strategia-asiakirja 2003, Tietoyhteiskuntaohjelma, Valtioneuvoston Kanslia, Hallituksen poikkihallinnolliset politiikkaohjelmat ja politiikat, Valtioneuvoston kanslian julkaisusarja 14/2003, ss. 29-40.
- Jaeger, P. T., & Thompson, K. M. (2003). E-government around the world: Lessons, challenges, and future directions. *Government Information Quarterly*, 20(4), 389-394.
- JHS-suositukset. Tarkistettu 22.11.2016. <http://www.jhs-suositukset.fi/>.

- Kalja, A., Robal, T., & Vallner, U. (August 2015). New generations of Estonian eGovernment components. pp. 625-631.
- KY-Verkko, Kuntien tietoliikenneverkot yhdistävä sopimus Soneralle. Tarkistettu 15.11.2016. <http://www.kunnat.net/fi/Kuntaliitto/media/tiedotteet/2014/09/Sivut/2014-09-01-ky-verkko-soneralle.aspx>.
- Layne, K., & Lee, J. (2001). Developing fully functional E-government: A four stage model. *Government Information Quarterly*, 18(2), 122-136.
- Negreiro M. 2015. *Bridging the digital divide in the EU* retrieved 11.12.2016 from [http://www.europarl.europa.eu/RegData/etudes/BRIE/2015/573884/EPRS_BRI\(2015\)573884_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/BRIE/2015/573884/EPRS_BRI(2015)573884_EN.pdf).
- TUVE-laki, Suomen säädöskokoelma, Oikeusministeriö, Laki julkisen hallinnon turvallisuusverkkotoiminnasta. Tarkistettu 15.11.2016. <http://vm.fi/documents/10623/360832/TUVE-laki/102a6ef8-9c32-4cd9-af31-0c1e31ce8a93>.
- O'Reilly, T. (2007). What is web 2.0: Design patterns and business models for the next generation of software. Retrieved from <https://papers-ssrn-com.hellos.uta.fi/abstract=1008839>.
- Palveluarkkitehtuuri, Valtiovarainministeriö. Tarkistettu 22.10.2016. <http://vm.fi/palveluarkkitehtuuri>.
- Palvelukuvaus, Palveluväylä. Tarkistettu 15.11.2016. <https://esuomi.fi/palveluntarjoajille/palveluvayla/palvelukuvaus/>.
- Palveluväylä, Valtiovarainministeriö. Tarkistettu 22.10.2016. <http://vm.fi/palveluvayla>.
- Perustiedot, Palveluarkkitehtuuri, Valtiovarainministeriö. Tarkistettu 8.11.2016. <http://vm.fi/palveluarkkitehtuuri/perustiedot>.
- SADe-tuloskooste 2009-2015, SADe-ohjelma digitalisaation vauhdittajana, Valtiovarainministeriö. Tarkistettu 27.11.2016. <http://vm.fi/documents/10623/1997070/SADe+tuloskooste+2009-2015/81933d82-0fdb-44eb-af64a4059d29f2ca>.
- SADe-toteuttamissuunnitelma, Valtiovarainministeriön muistio, 2009. Tarkistettu 11.12.2016. <http://vm.fi/documents/10623/1464506/Liite%201,%20SADe%20toteuttamissuunnitelma%202009-2014/696dde1e-69de-4f97-b513-802dc4564da7>.
- SADe-ohjelman loppuraportti, Owalgroupp, 2015. Tarkistettu 11.12.2016. <http://vm.fi/sade-ohjelman-taustamateriaaleja>.
- Sanasto, Valtiovarainministeriö, Kansallinen palveluväylä. Tarkistettu 2.10.2016. <https://confluence.csc.fi/display/Palveluvayla/Sanasto>.

- Sipilä Juha ja Vehviläinen Anu, Digitalisaatiolla tuottavuusloikka, 2015. VM/1184/00.01.02.02/2015.
- Smith, H. A., & McKeen, J. D. (2003). Developments in practice VIII: Enterprise content management. *The Communications of the Association for Information Systems, 11*(1), 41.
- Sovitinpalvelu, Palveluväylä. Tarkistettu 20.11.2016. <https://esuomi.fi/palvelun tarjoajille/palveluvayla/tekninen-aineisto/sovitinpalvelu/>.
- Sun, P., Ku, C., & Shih, D. (2015). An implementation framework for E-government 2.0. *Telematics and Informatics, 32*(3), 504-520.
- Suomalaista tietoyhteiskuntaa rakentamassa-loppuraportti, Valtioneuvoston kanslia. Tarkistettu 2.12.2016. http://vnk.fi/documents/10616/622950./2007_Suomalaista+tieto+yhteiskuntaa+rakentamassa.pdf/ecff3693-a556-4223-9a25-c54c0668d3c9?version=1.0.
- Turvallisuusverkkotoiminta. Tarkistettu 15.11.2016. <http://vm.fi/turvallisuusverkkotoiminta>.
- Vetuma. Valtori. Tarkistettu 22.11.2016. <http://www.valtori.fi/palvelut/vetuma>
- VY-verkko, Valtion yhteiset tietoliikennepalvelut. Tarkistettu 15.11.2016. http://www.valtori.fi/fi-FI/Palvelut/Tyoskentely-ympariston_palvelut/Valtion_yhteiset_tietoliikennepalvelut_VY-verkko.
- X-Road. Retrieved 22.10.2016 from <https://e-estonia.com/component/x-road/>.
- X-Road tiedonsiirtoprotokolla. Tarkistettu 21.11.2016. <https://esuomi.fi/palvelun tarjoajille/palveluvayla/tekninen-aineisto/rajapintakuvaukset/x-road-tiedonsiirtoprotokolla/>.

Kasvojen symmetrian koneellisen mittauksen mahdollisuudet toispuoleisessa kasvohalvauksessa

Evita Laitinen

Tiivistelmä

Toispuoleinen kasvohalvaus on vamma, jonka voi saada lähes kuka vain yllättäen koska vain, eikä sen varmaa aiheuttajaa tiedetä. Kasvojen symmetriaa arvioidaan vertaamalla kasvojen halvaantunutta puolta kasvojen terveeseen puoleen levossa ja liikkeessä. Nykyään tämän tekee silmänmääräisesti lääkäri erilaisten lomakepohjaisten pisteytysjärjestelmien avulla, vaikka saman arvioinnin voisi suorittaa tietokoneohjelmalla objektiivisesti. Kasvojen symmetrian mittaamiseen kliinisessä ympäristössä ei kuitenkaan ole tällä hetkellä käytössä objektiivista ja yleispätevää mittausmenetelmää.

Tutkielmassa luodaan katsaus kasvojen symmetrian mittaamiseen ja esitellään erilaisia olemassa olevia koneellisia mittaustapoja. Tarkoitus on avata näiden koneellisten mittausmenetelmien mahdollista soveltuvuutta toispuoleisen kasvohalvauksen vaikeusasteen mittaamiselle kliinisissä ympäristöissä, eli käytännössä yleensä lääkärin seurantakäynneillä. Tutkielmassa esitellään myös nykyään käytössä olevia lomakepohjaisia subjektiivisia mittausmenetelmiä ja esitetään perusteluja sille, miksi nämä perinteiset mitta-asteikot eivät ole luotettavia.

Avainsanat ja -sanonnat: kasvohalvaus, kasvojen symmetrian mittaaminen, pisteytysjärjestelmä, konenäkö, elektromyografia, kapasitiivinen menetelmä.

1. Johdanto

Kasvojen symmetrian mittaamisesta saattaa tulla ensimmäisenä mieleen tutkimukset, joissa todetaan henkilön kasvojen symmetrisyyden vaikuttavan siihen, miten miellyttäväksi muut kokevat hänen ulkonäkönsä. Kun kasvojen symmetriaa lisätään kuvamanipulaation avulla, kasvot arvioidaan viehättävämmiksi ja epäsymmetriaa lisäämällä vähemmän viehättäviksi [Rhodes *et al.* 1998]. Symmetrian mittaaminen on kuitenkin erittäin tärkeää silloin, kun kasvojen symmetrisyys muuttuu äkkiä jonkin sairauden tai vaikkapa tapaturman seurauksena. Toispuoleinen kasvohalvaus on eräs tällainen yhtäkkisen epäsymmetrian aiheuttaja. Kliinisten käyttötapojen lisäksi kasvojen symmetrian mittausta ollaan käytetty esimerkiksi yksilön tai ilmeiden tunnistamisessa [Singh and Nandi 2012; Mitra *et al.* 2007].

Koska kasvot ovat ihmiselle tärkeä ja näkyvä osa jokapäiväistä vuorovaikutusta ympäristön kanssa, niiden symmetria pyritään palauttamaan erilaisin keinoin kirurgisista toimenpiteistä aina lääkehoitoon ja ajan kanssa itsestään tapahtuvaan parantumiseen [Kanerva ja Pitkäranta 2006]. Tällöin se, että potilaan kasvojen tilan edistymistä pystytään seuraamaan luotettavasti ja tarkasti, on tärkeää. Nykyään toispuoleisen kasvohalvauksen vaikeusasteen mittaamiseen käytetään yleisesti subjektiivisia mitta-asteikkoja, joiden avulla lääkäri arvioi halvaantuneen puolen toimintaa verrattuna terveeseen. Subjektiivisia mittaamenetelmiä käytettäessä tulokset voivat vaihdella eri havainnoitsijoiden välillä ja havainnoitsijoiden välinen yksimielisyyden aste voi olla merkittävän alhainen. Lisäksi monet lomakepohjaiset mittarit eivät ota tarpeeksi hyvin huomioon kasvohalvauksen toissijaisia oireita tai kasvojen ajallista eli temporaalista epäsymmetriaa.

Tutkielmassa kerrotaan aluksi luvuissa 2 ja 3 yleisesti toispuoleisesta kasvohalvauksesta sekä epäsymmetrian mittaamisesta. Osioissa pyritään selvittämään, miksi kasvohalvauksen arvioiminen vaatii arviointimenetelmältään niin paljon. Luvussa 4 paneudutaan subjektiivisiin mittareihin sekä niiden puutteisiin, eli siihen, miksi ne eivät vastaa kasvohalvauksen arvioimisen tarpeisiin. Lopuksi luvussa 5 esitellään erilaisia olemassa olevia koneellisia kasvojenmittausmenetelmiä, joita voitaisiin soveltaa kasvojen symmetrian objektiiviseen arvioimiseen kasvohalvauspotilailla, sekä tarkastellaan näidenkin menetelmien heikkouksia ja vahvuuksia.

2. Kasvohalvaus

Kasvohalvaus on yleisimmin toispuoleinen, eli vain toinen puoli kasvoista halvaantuu. Tällainen halvaus aiheuttaa kasvoihin todennäköisesti huomattavaa epäsymmetriaa. Kasvohalvaus ilmaantuu yleensä yhtäkkiä ja se saattaa heikentää henkilön itsetuntoa ja vaikeuttaa syömisen, juomisen ja ilmehtimisen sujuvuutta [Robinson *et al.* 2012]. Lisäksi se voi aiheuttaa ongelmia kyynel- ja syljen erityksessä [Reizen *et al.* 2009]. Toispuoleista kasvohalvausta on olemassa kahta eri tyyppiä: perifeerinen ja sentraalinen kasvohalvaus. Sentraalisessa kasvohalvauksessa erityisesti kasvojen jommankumman puolen alaosan lihasten toiminta on heikentynyt. Otsan ja silmien toiminta taas on molemmilla puolilla normaali. Tämän tyyppisessä halvauksessa neurologiset selvittelyt ovat tarpeen, sillä sen syyt ovat yleensä aivoissa, ja ne voivat johtua esimerkiksi aivohalvauksesta. [Kanerva ja Pitkäranta 2006.]

Perifeerinen kasvohalvaus tarkoittaa sitä, että kasvojen toisen puolen kaikki lihakset ovat kokonaan tai osittain halvaantuneet. Vika ei ole tällöin keskushermostossa, vaan käskyt lähtevät aivoista normaalisti, mutteivät pääse kasvoihin

asti. Tämä johtuu siitä, että kasvohermossa, joka hermottaa kasvojen lihaksia, on jokin sen toimintaan vaikuttava häiriö. Kasvojen oikealla ja vasemmalla puolella on oma kasvohermonsensa, jolloin esimerkiksi häiriö vasemmassa kasvohermossa vaikuttaa koko kasvojen vasemman puolen toimintaan oikean puolen toimiessa normaalisti. Kasvohermo on herkkä, joten sen toiminta voi häiriintyä monestakin syystä, esimerkiksi tapaturman tai infektion seurauksena. Usein syy kasvohermon häiriöön jää kuitenkin selvittämättä, sillä halvauksen aiheuttajaa ei aina pystytä tunnistamaan. Tällöin puhutaan idiopaattisesta kasvohalvauksesta. [Kanerva ja Pitkäranta 2006.]

2.1. Bellin pareesi

Bellin pareesi tai Bellin halvaus on yleisin perifeerinen kasvohalvaus. Yleisesti mitä vain perifeeristä kasvohalvausta, jonka tarkkoja syitä ei tunneta, kutsutaan Bellin pareesiksi. Pitkän aikaa Bellin pareesin etiologia oli tuntematon, mutta nykyään sen aiheuttajiksi epäillään melko yleisesti eräitä herpesviruksia. Halvaus on myös yhdistetty muun muassa flunssaan, krooniseen keskikorvan tulehdukseen ja pään vammoihin [NINDS 2003].

Pareesi on melko yleinen, sillä noin joka 60:s suomalainen sairastuu siihen. Vuosittain uusia tapauksia ilmaantuu 20–30 jokaista 100 000 suomalaista kohti. Useimmat Bellin pareesiin sairastuneista paranevat täydellisesti tai ainakin osittain. Niistä, joilla kasvojen toisen puolen toiminta ei ole täysin kadonnut, 95 % paranee täysin. Täydellisen halvauksen saaneista noin 50 % paranee täysin ja lopuistakin suurin osa paranee hyvin. Noin 30 %:lle sairastuneista jää halvauksesta oireita, ja noin 4 %:lla sairastuneista pareesi ei parane ollenkaan. [Kanerva ja Pitkäranta 2006.]

2.2. Synkinesia

Halvauksen jälkitiloihin voi liittyä toissijaisia oireita kuten synkinesiaa, joka tarkoittaa jonkin muun kasvojen lihaksen aktivoitumista tahattomasti samaan aikaan, kun potilas yrittää liikuttaa toista lihasta. Synkinesian esiintymisen yleisyys on vaihdellut eri tutkimuksissa 9–55 %:n välillä. [Kanerva ja Pitkäranta 2006.] Synkinesian ajatellaan yleisesti johtuvan lihasten väärinhermottumisesta, kun hermotoiminta alkaa korjautua kasvojen halvaantuneella puolella.

Kasvohalvauksen kuntoutukseen käytetään joskus kasvojen lihasten jumppaamista. Oikeanlaisella jumppaamisella on saatu hyviä tuloksia sekä yleisesti koko halvauksen että synkinesian eli myötäliikkeiden suhteen [Robinson *et al.* 2012], mutta jumppaamista ei kuitenkaan suositella aloitettavaksi heti pareesin ilmaantumisen jälkeen, sillä se voi jopa lisätä synkinesian syntymistä [Kanerva & Pitkäranta 2006]. Kasvohalvauksen arvioimiseen käytetyn mittausmenetel-

män pitäisi olla niin tarkka, että sillä pystytään seuraamaan erilaisten interventioiden vaikutusta monipuolisesti: sekä halvauksen yleisen tilan, että toissijaisten oireiden osalta, oli intervention vaikutus sitten parempaan tai huonompaan päin.

3. Epäsymmetrian arvioiminen yleisesti

Ihmiskasvot ovat symmetriset, kun ne jaetaan kahtia pystysuoralla linjalla. Harvan ihmisen kasvot ovat kuitenkin täysin symmetriset. Lievä epäsymmetria on ihmiskasvoissa normaalia, eikä se ole yleensä silmäänpistävää tai haittaa ihmisten elämää. Kasvohalvauksessa kasvojen epäsymmetria on kuitenkin usein niin merkittävää, että se vaikuttaa huomattavasti ihmisen elämään.

Kasvojen epäsymmetriaa arvioidaan vertaamalla kasvojen halvaantuneen puolen toimintaa kasvojen terveeseen puoleen, sillä normaalisti kasvot näyttävät levossa eli staattisina ja liikkeessä eli dynaamisina lähes tai täysin symmetrisiltä. Koneellisesti kasvojen puoliskojen välistä symmetriaa voidaan verrata esimerkiksi peilaamalla toinen puoli kasvoista ja vertaamalla sitä toiseen, ei-peilattuun, puoliskoon [Harguess and Aggawal 2011]. Perinteisillä subjektiivisilla mittareilla tämä arviointi tehdään silmänmääräisesti ja kasvojen symmetriaa arvioidaan yleensä erikseen levossa ja liikkeessä.

Kasvohalvauksessa kasvojen liikkeen symmetria voidaan jakaa kahteen eri tyyppiin: liikkeen suuruuden sekä liikkeen ajoituksen eli temporaalisuuden epäsymmetria. Kasvohalvauksessa kasvojen toisen puolen lihasten toiminta on heikentynyt tai kokonaan poissa, jolloin tämän puolen lihakset todennäköisesti tuottavat vähemmän liikettä kuin kasvojen terveen puolen lihakset. Liikettä siis saattaa olla havaittavissa samaan aikaan molemmilla puolilla kasvoja, mutta toisella puolella liike on heikentynyt. Temporaalinen epäsymmetria taas ilmenee liikkeiden eriaikaisuutena: kasvojen toisen puolen lihakset liikkuvat eri aikaan kuin terveen puolen lihakset ja liike tulee esimerkiksi jäljessä.

Symmetrisyyteen vaikuttavat myös myötäliikkeet eli synkinesia, mikä tarkoittaa tahdonalaisen liikkeen aikana esiintyvää tahatonta liikettä. Yleisin tällainen myötäliike kasvohalvauksessa on kasvojen halvaantuneen puolen suupielen nousu samaan aikaan, kun potilas yrittää sulkea silmänsä. [Kanerva ja Pitkäranta 2006.]

Kasvohalvauksen vakavuusasteen arvioiminen on tärkeää, koska sen avulla päätetään, millaisiin interventioihin ryhdytään, jos halvauksen aiheuttajaa ei tiedetä: lievissä tapauksissa ei välttämättä tehdä mitään, mutta pitkään jatkuneissa ja vakavissa tapauksissa voi olla syytä ryhtyä kirurgisiin toimenpiteisiin kasvojen liikkeen korjaamiseksi. Vakavuusaste pitää pystyä arvioimaan tarkasti ennen

interventiota ja sen jälkeen, jotta intervention vaikutuksia pystytään seuraamaan. Fattah ja muut [2015] esittävät eri tutkimusten pohjalta kootun listan täydellisen mittajärjestelmän vaatimuksista:

1. Mahdollistaa kasvojen mittauksen alueittain.
2. Pystyy mittaamaan staattisia ja liikkuvia kasvoja.
3. Huomioi kasvohalvauksen toissijaiset oireet (esim. synknesia).
4. Tuottaa toistettavia tuloksia, joilla on korkea havainnoitsijoiden sisäinen ja välinen yksimielisyys.
5. Herkkä muutosten ja interventioiden seuraamiselle.
6. Sopii kliiniseen käyttöön.

4. Subjektiiiset mittarit

Kasvohalvausten tapauksessa lääkäri arvioi kasvojen symmetriaa erilaisin lomakkein ja arvoasteikoin. Nämä lomakkeet ovat subjektiivisia, eli käytännössä arvioija eli lääkäri arvioi liikkeen suuruutta tai symmetriaa oman havaintonsa perusteella. Eri havainnoitsijat voivat inhimillisistä syistä saada merkittävästikin erilaisen tuloksen samalla mittausmenetelmällä ja -kohteella. Suositut lomakepohjaiset mitta-asteikot ovat nopeita ja helppoja käyttää, mutta niiden avulla ei välttämättä saada luotettavaa tulosta. Seuraavaksi esitellään yleisimmin käytettyjen mittareiden lisäksi muitakin subjektiivisia mittareita ja kerrotaan tarkemmin niiden heikkouksista.

4.1. House-Brackmann ja Sunnybrook

Suomessa kasvohalvausten seurantakäynneillä käytetään yleisimpiä subjektiivisia lomakepohjaisia mittareita, kuten House-Brackmannia, joka kehitettiin 1900-luvun lopulla. House-Brackmann -menetelmän avulla lääkäri arvioi halvauksen vaikeutta kuudella eri tasolla (I–VI) normaalista symmetrisestä liikkeestä (taso I) aina täydelliseen halvaantumiseen (taso VI). Esimerkiksi tasoon II: *Lievä vajaatoiminta* kuuluvat seuraavat ominaisuudet:

- Havaittavissa lievää heikkoutta lihaksissa; voi esiintyä myös hieman synknesiaa.
- Levossa kasvot näyttävät symmetrisiltä.
- Liikkeessä otsa liikkuu lähes täydellisesti, molemmat silmät sulkeutuvat täysin ja suun alueella on havaittavissa hieman epäsymmetrisyyttä.

Arvio siis tehdään näiden ominaisuuksien perusteella kerralla koko kasvoille. House-Brackmann on yleisimmin käytetty pisteasteikko kasvohermon toiminnan mittaamiseen, koska sillä on pitkät perinteet, se on helppokäyttöinen, eikä niin laaja kuin monet muut menetelmät [Niziol *et al.* 2014].

House-Brackmannin lisäksi toinen yleinen mittari on lomakepohjainen asteikko nimeltään Sunnybrook Facial Grading Scale (ks. kuva 1). Sunnybrook-lomakkeella arvioidaan kasvojen halvaantuneen puolen lihasten liikettä terveeseen verrattuna viiden eri ilmeen aikana, synknesian esiintyminen joka ilmeen kohdalla erikseen sekä näiden lisäksi arvioidaan kasvojen symmetria levossa. Kasvojen tahdonalaista liikettä arvioidaan asteikolla yhdestä viiteen (1–5) ja synknesian vakavuutta nolasta kolmeen (0–3). Näiden pisteiden avulla lasketaan kokonaistulos, joka kertoo lääkärille halvauksen vakavuuden. Sekä Sunnybrook-että House-Brackmann -mittareiden luotettavuutta on tutkittu ja niitä vastaan on esitetty kritiikkiä. Kritiikkiin palataan hieman myöhemmin tässä tutkielmassa.

Sunnybrook Facial Grading System									
Resting Symmetry		Symmetry of Voluntary Movement					Synkinesis		
Compared to normal side		Degree of muscle EXCURSION compared to normal side					Rate the degree of INVOLUNTARY MUSCLE CONTRACTION associated with each expression		
Eye (choose one only) normal 0 narrow 1 wide 1 eyelid surgery 1		Standard Expressions Forehead Wrinkle (FRO) 1 2 3 4 5 <input type="checkbox"/> Gentle eye closure (OCS) 1 2 3 4 5 <input type="checkbox"/> Open mouth smile (ZYG/RIS) 1 2 3 4 5 <input type="checkbox"/> Snarl (LLA/LLS) 1 2 3 4 5 <input type="checkbox"/> Lip Pucker (OOS/OOI) 1 2 3 4 5 <input type="checkbox"/> Total <input type="checkbox"/>					NONE: No synkinesis or mass movement MILD: Slight synkinesis MODERATE: Obvious but not disfiguring synkinesis SEVERE: Disfiguring synkinesis/Gross mass movement of several muscles		
Cheek (naso-labial fold) normal 0 absent 2 less pronounced 1 more pronounced 1									
Mouth normal 0 corner dropped 1 corner pulled up/out 1		Gross Asymmetry Severe Asymmetry Moderate Asymmetry Mild Asymmetry Normal Symmetry Total <input type="checkbox"/>					Total <input type="checkbox"/>		
Total <input type="checkbox"/> Resting symmetry score Total X 5 <input type="checkbox"/>		Voluntary movement score: Total X 4 <input type="checkbox"/>					Synkinesis score: Total <input type="checkbox"/>		
Patient's name _____ Dx _____ Date _____		Vol mov't score <input type="checkbox"/> - Resting symmetry score <input type="checkbox"/> - Synk score <input type="checkbox"/> = Composite score <input type="checkbox"/>							

Ross, Fradet, Nedzelski 1992

Kuva 1. Sunnybrook Facial Grading System englanniksi

4.2. Muita subjektiivisia mittareita

Suosituimpien mittareiden – House-Brackmannin ja Sunnybrookin – lisäksi on kehitetty lukuisia muitakin subjektiivisia, lomakepohjaisia mittausten menetelmiä. Seuraavaksi esitellään näistä keskeisimmät, sillä erilaisia mittareita on olemassa niin monta, ettei kaikista niistä ole mahdollista kertoa tässä tutkielmassa. Muita mittareita ovat esitelleet esimerkiksi Niziol ja muut [2014] sekä Fattah ja muut [2014].

Sydney on Sunnybrookin kaltainen kasvohalvauksen pisteytysmittari, jolla arvioidaan erikseen kasvojen lihasten liike viiden eri ilmeen aikana ja synkinesian esiintyminen joka ilmeen kohdalla erikseen, lisäksi arvioidaan kasvojen symmetria levossa. Sydneyn ja Sunnybrookin välinen suurin ero on siinä, että Sydney-menetelmässä käytetään asteikkona arvoja nolasta kolmeen (0–3). Nolla tarkoittaa täydellistä halvaantuneisuutta ja kolme normaalia liikkuvuutta.

Vrabec ja muut [2009] päivittivät House-Brackmann-asteikon Facial Nerve Grading Scale 2.0:een (FNGS 2.0). Mittarin responsiivisuutta kasvojen liikkeisiin lisättiin laajentamalla arviointi koskemaan kasvojen eri osia erikseen: uudella versiolla lasketaan arvot erikseen kulmakarvojen alueelle, silmien alueelle, ”hymyrypyille” nenän molemmin puolin sekä suun alueelle. Näitä kasvojen osia arvioidaan halvauksen vakavuuden suhteen asteikolla yhdestä kuuteen (1–6). Myös synkinesia arvioidaan erikseen asteikolla nolasta kolmeen (0–3) sen perusteella, miten sitä yleisesti esiintyy kasvoilla. Lopulta arvot lasketaan yhteen ja saadaan jokin summa väliltä 4–24, joka määrää, mihin House-Brackmannin kuudesta tasosta kasvohalvaus vakavuudessaan kuuluu. Facial Nerve Grading Scale 2.0:a ei ole vielä vertailtu Sunnybrookin kanssa, ja se vaatii vielä lisää tutkimusta havainnoitsijoiden välisessä luotettavuudessa [Fattah *et al.* 2014].

Erilaisia mittareita – niin subjektiivisia kuin koneellisiakin – on kehitetty paljon, eivätkä läheskään kaikki niistä ole päätyneet kliiniseen käyttöön. On esitetty, että House-Brackmannin perinteet ovat niin pitkällä ja elävät niin vahvoina, että uuden menetelmän pitäisi olla sen kanssa vertailukelpoinen, jotta näitä kahta menetelmää pystyttäisiin käyttämään toistensa tukena [Fattah *et al.* 2014]. Uuteen menetelmään voi olla hankalaa siirtyä ilman tällaista siirtymävaihetta.

4.3. Subjektiivisten mittareiden ongelmat

Kasvohalvausta arvioivissa mittareissa tärkeää on toistettavuus ja eri arvioijien välinen ja sisäinen yksimielisyys [Niziol *et al.* 2015; Kanerva *et al.* 2006; Fattah *et al.* 2014]. Tämä tarkoittaa käytännössä, että eri lääkärit arvioivat saman potilaan samalla tavalla (arvioijien välinen yksimielisyys; interobserver reliability/ agreement) ja että myös sama lääkäri arvioi saman kuvamateriaalin samasta potilaasta myöhemmin samalla tavalla kuin aikaisemmin (arvioijien sisäinen yksimielisyys; intraobserver reliability/ agreement).

Kanerva ja muut [2006] saivat tutkimuksessaan havainnoitsijoiden eli koulutettujen lääkäreiden väliseksi yksimielisyydeksi House-Brackmann-asteikkoa käyttäessä ensimmäisellä kierroksella 48 % ja toisella kierroksella 50 %, kun taas Coulson ja muut [2005] saivat samaksi arvoksi 44 %. Tämä kertoo siitä, että sama potilas voi hyvinkin saada erilaisen arvion kasvojensa symmetrisyydestä riippuen siitä, kuka häntä arvioi House-Brackmann-mittarin avulla.

Niziol ja muut [2015] pyrkivät katselmuksessaan selvittämään parhaan mahdollisen olemassa olevan mittaamenetelmän arvioimaan kasvohalvauspotilaan kasvoja ennen ja jälkeen kasvojen liikkuvuuden palauttavan kirurgisen leikkauksen. Katselmuksen perusteella House-Brackmannin soveltuvuus tähän tarkoitukseen oli keskinkertainen. He nostivat esille muun muassa havainnoitsijoiden välisen erimielisyyden, eikä menetelmä ollut myöskään tarpeeksi herkkä mittaamaan potilaan kasvoissa tapahtuvia muutoksia kuntoutumista seurattaessa.

House-Brackmannin ongelma on, ettei se pysty täysin mittaamaan kasvojen liikkuvuutta, eikä se ole tarkka, sillä sen avulla arvioidaan kasvoja yleisesti kokonaisuutena. Edustavampi tulos sen avulla saadaan, jos asteikkoa sovelletaan erikseen eri puolille kasvoja, kuten otsaan, silmiin, nenään ja suuhun. Reizen ja muut [2009] vertasivat toisiinsa House-Brackmannin avulla tehtyjä arvioita potilaiden kasvoista kokonaisuutena ja osissa. Tutkimuksessa korkeinta yksimielisyyttä (58,8 %) oli koko kasvojen ja kasvojen keskiosan arvioiden välillä. Tämä johtui todennäköisesti siitä, että koko kasvoja arvioidessa kiinnitettiin ensimmäisenä huomiota kasvojen keskiosaan ja se vaikutti eniten kokonaisarvion tekemiseen. Koko kasvojen ja pelkän otsan välillä arviot kohtasivat 35,1 %:ssa tapauksista, eli luku oli matalin kaikista. Havainnoista käy ilmi, ettei yleisesti koko kasvoille tehty arvio anna luotettavaa kuvaa kasvojen eri osien tilasta.

Vrabecin ja muut [2009] mainitsevat yhdeksi House-Brackmannin ongelmista sen, että ero tasojen III ja IV välillä on häilyvä ja epäselvä, eikä arviota kasvoista kokonaisuutena ole välttämättä helppoa antaa, sillä toiminta saattaa olla normaalimpaa kasvojen ylä- kuin alaosassa. Facial Nerve Grading Scale 2.0 korjaa suurimman osan aiemmin esitetyistä House-Brackmannin ongelmista, mutta sekään ei silti ole objektiivinen. Sen avulla arvioutuna havainnoitsijoiden välinen yksimielisyys oli 67 %.

Kanerva ja muut [2006] pitivät Sunnybrook-mittaria hieman luotettavampana kuin yleisemmin käytettyä House-Brackmannia. Samanlaisiin tuloksiin tulivat myös Fattah ja muut [2014] katsauksessaan, jossa he kokosivat muiden tutkimustuloksia eri mittareista. Sunnybrookin on myös osoitettu olevan herkkä eli soveltuvan paremmin interventioiden vaikutusten seuraamiseen kuin House-Brackmannin [Ross *et al.* 1996]. Sunnybrook on House-Brackmannia luotettavampi mittari todennäköisesti siksi, että se on tarkoitettu arvioimaan kasvoja monipuolisemmin: erikseen levossa ja liikkeessä, lisäksi se huomioi synknesian paremmin. Muun muassa Coulson ja muut [2005] kuitenkin huomasivat keskenään samankaltaisten Sunnybrook- ja Sydney-mittareiden sisäisen ja välisen luotettavuuden olevan alhainen, eli eri arvioijat saivat niillä erilaisia tuloksia samoista kohteista. Luotettavuus oli melko hyvä tahdonalaisia kasvonilmeitä arvi-

oitaessa, mutta erityisen alhainen luotettavuus mittareiden välillä oli juuri synkinesiaa arvioitaessa. Eli sillä, miten ja millä perusteella synkinesiaa arvioidaan, voi olla suuri vaikutus saatuihin tuloksiin.

Coulson ja muut [2005] ehdottavat, että synkinesian arvioimiseksi tulisi antaa tarkemmat sanalliset ohjeet, joissa kuvaillaan, miten synkinesia mahdollisesti ilmenee ja miten sitä voidaan havainnoida yksilöllisesti. Niziol ja muut [2014] ehdottavat, että mittausmenetelmän, joka on tarkoitettu kasvojen symmetrian mittaamiseen ennen ja jälkeen kasvoleikkauksen, tulisi ottaa huomioon potilaan oma näkemys ja tyytyväisyys kasvojensa symmetriaan. Lisäksi he painottavat, että menetelmän on tärkeää olla nopea ja yksinkertainen.

5. Koneellinen mittaaminen

Kasvojen symmetrian arvioiminen koneellisesti on ihmisen silmänmääräiseen arvioimiseen verrattuna objektiivinen tapa, sillä inhimilliset arviointivirheet tai mielipiteet eivät vaikuta tietokoneen tekemään arvioon. Objektiivinen mittari saa samasta kohteesta aina saman tuloksen, eli havainnoitsijoiden sisäiset ja väliset arviointierot tuloksissa eivät tuota ongelmia. Subjektivisten mittareiden kohdalla mainitaan usein esimerkiksi havainnoitsijoiden saama koulutuksen määrä ja sen vaikutus saatuihin tuloksiin, joissa vähemmän kokeneet tietyn mittarin käyttäjät saavat epäluotettavampia tuloksia kuin kokeneemmat käyttäjät [Fattah *et al.* 2014; Reizen *et al.* 2009]. Objektiiviseksi mittariksi voidaan myös laskea mittojen ottaminen kasvoista käsin, kuten Manktelowin viivoittimella mittaaminen [Niziol *et al.* 2014]. Seuraavaksi keskitytään kuitenkin tietokonesovellusten avulla toimiviin objektiivisiin mittareihin, joista esitellään muutama toisistaan eroava tapa mitata kasvojen symmetriaa. Lisäksi esitetään näiden koneellisten menetelmien ongelmia.

5.1. Konenäkö

Konenäkö (machine vision) mahdollistaa sen, että tietokoneohjelma pystyy automaattisesti analysoimaan sille välitetyn kuvan tai videokuvan sisällön. Kuva oikeasta maailmasta välittyy tietokoneelle yleensä jonkinlaisen kameran välityksellä, jonka jälkeen tietokoneohjelma analysoi kuvan ja antaa käyttäjälle palautteen analyysistä. Myös etukäteen kuvattu materiaali voidaan useimmiten välittää tietokoneohjelmalle. Tähän tekniikkaan perustuvia sovelluksia kasvojen symmetrian tai niiden yleiseen arvioimiseen on olemassa lukuisia erilaisia ja ne ovat olleet suosiossa jo kauan (ks. [O'Mara and Owens 1996]). Seuraavaksi esitellään näistä menetelmistä muutama ajankohtainen.

Facial Action Coding System (FACS) on vanha järjestelmä kasvon ilmeiden tunnistamiseen ja mittaamiseen, jossa kasvojen ilmeet tunnistetaan liiketapahtumien eli toimintayksiköiden (action unit, AU) avulla [Ekman and Friesen 1976]. Aiemmin järjestelmä toimi silmänmääräisesti arvioituna ja oli siksi riippuvainen havainnoitsijan saamasta koulutuksesta, mutta nykyään siihen pohjautuvia menetelmiä käytetään konenäöllä tietokonepohjaisesti mm. animaatiossa. Koska toimintayksikköjä on molemmin puolin kasvoja, FACSin avulla voi myös mitata kasvojen symmetriaa, mutta vaikka sen käyttöä kliinisissä ympäristöissä on tutkittu (ks. [Haase *et al.* 2015; Hamm *et al.* 2011]), se ei ole vielä yleistynyt käytössä.

Glasgow Facial Palsy Scale (GFPS) on ilmainen tietokoneohjelma, joka mittaa House-Brackmann-asteikon arvot videokuvasta konenäön avulla. Se on siis objektiivinen versio yleisimmin käytetystä subjektiivisesta mittarista, ja tämän takia vartenotettava mittaustapa. Menetelmällä on saatu hyviä tuloksia, koska tietokoneen objektiivisuudella on pystytty korjaamaan havainnoitsijoiden välinen erimielisyys, josta House-Brackmannia on eniten kritisoitu [Niziol *et al.* 2015].

Cao ja muut [2015] ovat kehittäneet menetelmän kasvojen reaaliaikaiseen korkean tarkkuuden 3D-mallintamiseen. Menetelmä käyttää kasvojen kolmiulotteiseen mallintamiseen optista virtausta (optical flow), jota monet muutkin samankaltaiset kasvojen mittaamisen menetelmät käyttävät (ks. [Chen *et al.* 2014; Wang *et al.* 2004]). Menetelmä on tarkoitettu viihdekäyttöön ja muun muassa animaation avuksi, sillä ihmiskasvojen ja etenkin autenttisten ilmeiden animointi on vaikeaa. Menetelmä luo kaksiulotteisesta web-kameran kuvasta reaaliajassa liikkuvan kolmiulotteisen kasvomallin, jossa aikaisemmista samanlaisista menetelmistä poiketen näkyvät selkeinä myös kasvojen rypyt. Toispuoleisessa kasvohalvauksessa halvaantunut puoli kasvoista ei liiku, eikä sille puolelle muodostu ryppejä. Tämän takia menetelmä voisi sopia myös kliiniseen käyttöön, jos siihen liitettäisiin ominaisuus, jonka avulla voitaisiin vertailla kasvojen puoliskoja keskenään.

Microsoft X-Box-pelikonsolin Kinect-sensoriin pohjautuvat kasvonmittausmenetelmät näkevät videokuvan sijasta infrapunavalolla. Infrapunavalolla mitattaessa valaistuksella ei ole väliä, vaan menetelmä toimii pimeässäkin huoneessa. Laite heijastaa eteensä tiheän, pienistä pisteistä muodostuvan verkon, jonka avulla se tunnistaa pinnanmuodot ja liikkeen. Menetelmällä pystytään tunnistamaan kasvonilmeitä matalalaatuisella 3D-sensorilla [Li *et al.* 2013], joten se soveltuisi sopivan oheisohjelman kanssa myös kasvojen symmetrian arvioimiseen keskinkertaisissa valaistusolosuhteissa.

5.2. Elektromyografia

Elektromyografia (EMG) on yleisimmin käytetty objektiivinen mittausmenetelmä kasvojen toiminnan mittaamiseen [Mannarelli *et al.* 2012]. EMG mittaa lihasten lähettämiä sähköimpulsseja suoraan iholle kiinnitettävillä elektrodeilla, joten se tunnistaa pienimmätkin liikkeet myös ihon alta. EMG:n avulla voidaan mitata tarkasti sekä tahdonalaista että itsestään tapahtuvaa liikettä kasvoissa.

Kasvojen alueelta mitattava EMG eroaa hieman muualle kehoon suoritettavasta mittauksesta. Muualle kehoon käytetään yleensä neulamyografiaa, jossa suoraan lihaksiin pistellään ohuita neuloja, joiden sisällä elektrodit ovat. Tätä voidaan kutsua myös lihaksensisäiseksi elektromyografiaksi (intramuscular EMG). Kasvoihin voidaan käyttää neulaa tai mukavuussyistä myös toista tapaa, jossa elektrodit on sijoitettu huovasta tai muusta pehmeästä materiaalista valmistettuihin pieniin lappuihin, jotka asetetaan kasvojen lihaksien kohdalle ihon pinnalle. Lihaksensisäinen EMG on jonkin verran herkempi ja tarkempi kuin ihon pinnalta tapahtuva EMG-mittaus.

Elektroneuromyografia (ENMG) eli hermo- ja lihassähkötutkimus on tutkimus, jossa on yhdistetty EMG ja elektroneurografia (ENOG tai ENG). Molempia voidaan käyttää kasvohalvausta arvioitaessa, sillä ENG mittaa kasvohermon toimintaa ja EMG lihasten toimintaa [Mannarelli *et al.* 2012]. Elektroneurografia ei kuitenkaan kuulu rutiinitutkimuksiin, ja sen toimimisesta toipumisen ennustajana ollaan saatu vaihtelevia tuloksia [Kanerva ja Pitkäranta 2006].

Elektromyografia mittaa lihasten liikkeet ajallisesti tarkasti, eli sen avulla voidaan havainnoida temporaalista epäsymmetriaa. EMG:tä käytetään yleensä tapauksissa, joissa halutaan tarkkaa tietoa esimerkiksi jonkin lihaksen toiminnan heikentymisestä. EMG:n avulla on esimerkiksi pystytty mittaamaan vegetatiivisessa tilassa olevien potilaiden erittäin hienovaraisia ja muuten huomaamattomia kasvonliikkeitä, kun heille on kerrottu eri tunteita herättäviä tarinoita [Fiacconi and Owen 2016].

5.3. Kapasitiivinen menetelmä

Kapasitiivinen menetelmä eroaa edellisistä menetelmistä siinä, ettei se perustu konenäköön, eikä se vaadi toimiakseen kasvoihin kiinnitettäviä elektrodeja. Kapasitiivinen menetelmä mittaa kasvonliikkeitä nimensä mukaisesti kapasitiivisella menetelmällä, joka tunnistaa korkeuden vaihtelut jollakin tietyllä pinnalla. Kapasitiivinen menetelmä on myös tuttu älypuhelinien ja tablettilaitteiden kosketusnäytöistä. Rantasen ja muiden [2013] kehittämässä menetelmässä henkilön päähän asetetaan kuulokkeiden kaltainen mittaustaite, josta haaroittuu molemmille puolille kasvojen eteen kolme sormenmallista ulokeparia: ylimmät ulokeet mittaavat kulmakarvojen aluetta, keskimmäiset poskien aluetta ja alin pari

mittaa leukaa ja suun aluetta. Ulokkeiden sisältämät anturit mittaavat kasvojen liikkeet sähköisen kentän avulla, joka tunnistaa etäisyysvaihtelut kasvojen pinnan ja mittaasanturin välillä. Laite lähettää keräämänsä datan tietokoneeseen Bluetoothin avulla.

Kapasitiivisella menetelmällä mitataan samanaikaisesti erikseen kasvojen vasenta ja oikeaa puolta, ja se havaitsee objektiivisesti epäsymmetrian kasvoliikkeiden suuruudessa, eli sen voidaan katsoa soveltuvan kasvohalvauksen vaikeuden mittaamiseen ja epäsymmetrian arvioimiseen. Laitteella on myös mahdollista mitata ajoitukseltaan epäsymmetrisiä liikkeitä, eli se havaitsee temporaalisen epäsymmetrian. Laite soveltuu myös synknesian mittaamiseen, sillä eri puolilla kasvoja sijaitsevat elektrodit tuottavat havainnoistaan dataa erikseen, ja tätä dataa voidaan vertailla muiden elektrodien datan kanssa, jolloin nähdään helposti, jos jokin lihas liikkuu silloin, kun sen ei pitäisi.

5.4. Objektiivisten mittareiden ongelmat

Koneellisilla mittareillakin on kuitenkin omat ongelmansa: mittaustavat voivat olla kalliita ja joskus aikaavieviä, eikä niitä tämän takia pystytä käyttämään joka paikassa. Kalliiden oheislaitteiden hankkiminen ei välttämättä tunnu kannattavalta, kun ilmaisiakin paperilomakkeita on olemassa. Esimerkiksi 3D-kasvomallinnus voi olla kallista.

Kaikkia tässä tutkielmassa esiteltyjä koneellisia mittaajärjestelmiä ei olla vielä yritetty soveltaa kliiniseen ympäristöön ja vastaamaan toispuoleisen kasvohalvauksen erikoistarpeisiin. Kaikki niistä eivät sellaisenaan sovellu kasvohalvauksen ja sen toissijaisten – mutta aivan yhtä huomionarvoisten – oireiden, kuten synknesian automaattiseen arvioimiseen. Esimerkiksi House-Brackmanniin perustuva Glasgow Facial Palsy Scale ei ota toissijaisia oireita huomioon yhtään sen paremmin kuin alkuperäisenkään versio [Niziol *et al.* 2014]. Mittaamiseen sopivaa teknologiaa on olemassa siis jo nyt, mutta koska useita niistä ei ole tarkoitettu kasvojen epäsymmetrian mittaamiseen, niistä puuttuu oheisohjelma, joka mittauksen perusteella laskisi automaattisesti kvantitatiivisia arvoja, joihin potilaan myöhempiä arvoja tai toisten potilaiden arvoja pystyttäisiin vertaamaan. Arvioinnin automatisointi säästäisi paljon aikaa, joka muutenkin laskeaan koneellisten mittareiden ongelmaksi [Fattah *et al.* 2014].

Jotkin konenäköön pohjautuvista tietokonesovelluksista, jotka on varta vasten kehitetty kasvohalvausten mittaamiseen, ovat ilmaisia ja vaativat vain tietokoneen sekä jonkin videointilaitteen, jolloin niiden käyttöönoton uskoisi olevan melko helppoa. Tällainen ohjelma on esimerkiksi aikaisemminkin mainittu Glasgow Facial Palsy Scale. Videopohjaiset mittaajärjestelmät vaativat kuitenkin kunnolla toimiakseen tietynlaisen valaistuksen ja paikallaan pysyvän kuvauskohteen, jotta ohjelma pystyy tunnistamaan mittaamisen kannalta tärkeät pisteet

potilaan kasvoilta. Konenäköpohjaiset menetelmät vaativat usein lisäksi monta työvaihetta kuvan tai videokuvan käsittelemiseen. Näihin työvaiheisiin kuuluu esikäsitteily, kasvojen ja kasvonpiirteiden tunnistus, kuvan suoristaminen sekä edellä kuvatun FACS-menetelmän toimintayksiköiden tunnistaminen [Rantanen *et al.* 2013]. Mitä enemmän menetelmä vaatii kasvojen käsittelyyn erilaisia työskentelyvaiheita, sitä herkempi se todennäköisesti on virheille ja sitä enemmän ohjelman käyttö vaatii aikaa ja tietokoneresursseja.

Kapasitiivisen menetelmän huonot puolet liittyvät lähinnä mittauslaitteeseen, jota pidetään päässä. Laitetta ei ole helppo olla huomaamatta, joten se voi häiritä potilasta, kun tämä yrittää liikuttaa kasvojaan. Laite saattaa myös liikkua hieman potilaan päässä, jolloin se voi mitata sellaistaakin liikettä, joka ei johdu potilaan kasvojenliikkeistä. Tällaiset liikkeet tulisi pystyä tunnistamaan ja suodattamaan pois, jotta laite mittaisi täysin luotettavasti. Kohina (noise) saattaa myös vaikuttaa hieman mittaustuloksiin. Muita menetelmän tarkempia haittapuolia ei olla vielä raportoitu. [Rantanen *et al.* 2012.]

6. Yhteenveto

Olemassa olevat subjektiiviset mittausmenetelmät eivät sovi toispuoleisen kasvohalvauksen luotettavaan arvioimiseen, minkä takia niiden tilalle suositellaan koneellista objektiivista mittaustapaa. Täydellisen mittarin piirteet Fattahin ja muiden [2015] mukaan ovat seuraavat: halpa, nopea, huomaamaton, herkkä, tarkka, objektiivinen ja kvantitatiivinen. Tällaisen mittausmenetelmän pitäisi siis:

1. Pystyä maksamaan hintansa takaisin.
2. Olla nopea käyttää, jotta aikaa säästyy.
3. Olla sen verran huomaamaton, ettei se häiritse potilasta hänen ilmehtiessään.
4. Olla herkkä erilaisille ilmeille ja muutoksille kasvoissa, ja pystyä huomioimaan myötäliikkeet eli synkinesia.
5. Mitata tarkasti ja aluekohtaisesti, jotta kasvojen eri osien parantumista voidaan seurata.
6. Olla objektiivinen, jotta tulokset ovat luotettavia ja keskenään vertailtavissa.
7. Antaa kvantitatiivisia eli määrällisiä ja siten kuvaajien avulla tulkittavissa olevia tuloksia.

Vaikka koneelliset menetelmät vaativat toimiakseen aikaa ja osaamista käyttäjältään, edut subjektiivisiin menetelmiin verrattuna ovat myös suuret. Tietoko-

nepohjaisten mittausmenetelmien epäsuosio kliinisessä käytössä johtuu todennäköisesti siitä, ettei tarpeeksi helppokäyttöistä ja yleispätevää koneellista mittausmenetelmää olla vielä kehitetty. Fattahin ja muiden [2015] mukaan nykyteknologialla ei pystytä saavuttamaan täydellistä kasvohalvauksen mittausmenetelmää. Väitteen kanssa on vaikeaa olla eri mieltä, sillä täydellisyys on paljon vaadittu. Jo nyt on kuitenkin olemassa paljon lupaavaa teknologiaa. Erittäin hyvä kasvojen epäsymmetrian mittausmenetelmä, joka sopii juuri toispuoleisen kasvohalvauksen tarpeisiin, on varmasti mahdollista toteuttaa.

Viiteluettelo

- Marjolaine Baude, Emilie Hutin and Jean-Michel Gracies. 2015. A bidimensional system of facial movement analysis and reliability in adults. *BioMed Research International* 6/16, 2015, 1–8.
- Jiansheng Chen, Chang Yang, Yu Deng, Gang Zhang and Guangda Su. 2014. Exploring facial asymmetry using optical flow. *IEEE Signal Processing Letters* 21, 7, 792–795.
- Susan E. Coulson, Glen R. Croxson, Roger D. Adams and Nicholas J. O'Dwyer. 2005. Reliability of the "Sydney," "Sunnybrook," and "House-Brackmann" facial grading systems to assess voluntary movement and synkinesis after facial nerve paralysis. *Operative Techniques in Otolaryngology-Head and Neck Surgery* 135, 6, 865–871.
- Paul Ekman and Wallace V. Friesen. 1976. Measuring facial movement. *Environmental Psychology and Nonverbal Behavior* 1, 1, 56–75.
- Adel Y. Fattah, Anthony D. R. Gurusinghe, Javier Gavilan, Tessa A. Hadlock, Jeff R. Marcus, Henri Marres, Charles C. Nduka, William H. Slattery and Alison K. Snyder-Warwick. 2015. Facial nerve grading instruments: Systematic review of the literature and suggestion for uniformity. *Plastic & Reconstructive Surgery* 135, 2, 569–579.
- Chris M. Fiacconi and Adrian M. Owen. 2016. Using facial electromyography to detect preserved emotional processing in disorders of consciousness: A proof-of-principle study. *Clinical Neurophysiology* 127, 9, 3000–3006.
- D. Haase, L. Minnigorde, G.F. Volk, J. Denzler and O. Gunthias-Lichius. 2015. Automated and objective action coding of facial expressions in patients with acute facial palsy. *European Archives of Oto-Rhino-Laryngology* 272, 5, 1259–1267.
- Jihun Hamm, Christian G. Kohler, Ruben C. Gur and Ragini Verma. 2011. Automated facial action coding system for dynamic analysis of facial expressions in neuropsychiatric disorders. *Journal of Neuroscience Methods* 200, 2, 237–256.

- Josh Harguess and J. K. Aggarwal. 2011. Is there a connection between face symmetry and face recognition? In: *Proc. of the 2011 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 66–73.
- Mervi Kanerva ja Anne Pitkäranta. 2006. Perifeerinen kasvohalvaus. *Duodecim* 122, 2267–2274.
- Mervi Kanerva, Tuija Poussa and Anne Pitkäranta. 2006. Sunnybrook and House-Brackmann facial grading systems: Intrater repeatability and inter-rater agreement. *Operative Techniques in Otolaryngology - Head and Neck Surgery* 135, 6, 865–871.
- Billy Y. L. Li, Ajmal S. Mian, Wanquan Liu, Aneesh Krishna. 2013. Using Kinect for face recognition under varying poses, expressions, illumination and disguise. In: *IEEE Workshop on Applications of Computer Vision*, 186–192.
- Greg Mannarelli, Garrett R. Riffin, Paul Kileny and Bruce Edwards. 2012. Electrophysiological measures in facial paresis and paralysis. *Operative Techniques in Otolaryngology - Head and Neck Surgery* 24, 4, 236–247.
- Sinjini Mitra, Nicole A. Lazar and Yanxi Liu. 2007. Understanding the role of facial asymmetry in human face identification. *Statistics and Computing* 17, 1, 57–70.
- National Institute of Neurological Disorders and Stroke. 2003. Bell’s palsy fact sheet. http://www.ninds.nih.gov/disorders/bells/detail_bells.htm. Checked 14.12.2016.
- Rafal Niziol, Francis P. Henry, Jonathan I. Leckenby and Adriaan O. Grobbelaar. 2015. Is there an ideal outcome scoring system for facial reanimation surgery? A review of current methods and suggestions for future publications. *Journal of Plastic, Reconstructive & Aesthetic Surgery* 68, 4, 447–456.
- David O’Mara and Robyn Owens. 1996. Measuring bilateral symmetry in digital images. In: *Proc. of the Digital Signal Processing Applications*, 1, 151–156.
- Ville Rantanen, Pekka Kumpulainen, Hanna Venesvirta, Jarmo Verho, Oleg Špakov, Jani Lylykangas, Akos Vetek, Veikko Surakka and Jukka Leikkala. 2012. Capacitive facial activity measurement. In: *Proc. of XX IMEKO World Congress*, 1–6.
- Ville Rantanen, Hanna Venesvirta, Oleg Spakov, Jarmo Verho, Akos Vetek, Veikko Surakka and Jukka Leikkala. 2013. Capacitive measurement of facial activity intensity. *IEEE Sensors Journal* 13, 11, 4329–4338.
- Shari D. Reizen, James S. Babb and Anil K. Lalwani. 2009. Significance and reliability of the House-Brackmann grading system for regional facial nerve function. *Operative Techniques in Otolaryngology - Head and Neck Surgery* 140, 2, 154–158.

- Mara Wernick Robinson, Jennifer Baiungo, Mark Hohman and Tessa Hadlock. 2012. Facial rehabilitation. *Operative Techniques in Otolaryngology-Head and Neck Surgery* 23, 4, 288–296.
- Brenda G. Ross, Gaeton Fradet and Julian M. Nedzelski. 1996. Development of a sensitive clinical facial grading system. *Operative Techniques in Otolaryngology - Head and Neck Surgery* 114, 3, 380–386.
- W. S. W Samsudin and K. Sundaraj. 2012. Image processing on facial paralysis for facial rehabilitation system: A review. In: *Proc. of the 2012 IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, 259–263.
- Avinash Kumar Singh and G. C. Nandi. 2012. Face recognition using facial symmetry. In: *Proc. of the Second International Conference on Computational Science, Engineering and Information Technology (CCSEIT)*, 550–554.
- S. Wang, H. Li, F. Qi and Y. Zhao. 2004. Objective facial paralysis grading based on Pface and eigenflow. *Medical and Biological Engineering and Computing* 42, 5, 598–603.
- Jeffrey T. Vrabec et al. 2009. Facial nerve grading system 2.0. *Operative Techniques in Otolaryngology - Head and Neck Surgery* 140, 4, 445-450.

K:n lähimmän naapurin menetelmän käyttökohteet

Jarkko Saarinen

Tiivistelmä

Lähimmän naapurin menetelmä on koneoppimisessa käytettävä luokittelumenetelmä. Menetelmä perustuu datan esiintymien etäisyyksiin vektoriavaruuksessa, jonka eri ulottuvuudet kuvaavat datan muuttujia. Menetelmä on yksinkertaisuutensa ja erilaisiin käyttötapauksiin mukautuvuutensa vuoksi hyvin laajassa käytössä. Tässä tutkielmassa perehdytään tarkemmin menetelmän toimintaan ja tarkastellaan sen käyttökelpoisuutta muutamassa erilaisessa käyttötapauksessa. Tarkastelujen pohjalta havaitaan, että menetelmä soveltuu parhaiten yksinkertaisiin, datan sisältöä luokitteleviin käyttötapauksiin. Monimutkaisemat ja datan semantiikkaa käsittelevät luokittelutapaukset tuovat menetelmän yksinkertaisuuden tuomat heikkoudet esille.

Avainsanat ja -sanonnat: Koneoppiminen, luokittelu, k -NN

1. Johdanto

Lähimmän naapurin menetelmä (Nearest Neighbor Rule) on 1950-luvun alussa esitelty tilastollinen luokittelumenetelmä [Fix and Hodges 1951], jota käytetään nykyisin laajasti erilaisissa koneoppimisjärjestelmissä. Luokittelulla tarkoitetaan tuntemattomien datan esiintymien (instanssien) sijoittamista johonkin luokkaan [Duda *et al.* 2000]. Esimerkiksi roskapostifiltteri voi luokitella sähköposteja luokkiin ”roskaposti” ja ”ei-roskaposti”.

Lähimmän naapurin menetelmän toiminta perustuu siihen, että ennalta tunnetut datan instanssit ja tuntematon, luokiteltava datan instanssi kuvataan n -ulotteiseen avaruuteen. Avaruuden ulottuvuuksien lukumäärä n määräytyy yksittäisen datainstanssin luokittelussa olennaisten ominaisuuksien, eli piirteiden, määrän mukaan. Tuntematonta instanssia lähimmäksi sijoittuvat tunnetut instanssit, eli naapurit, kertovat luokittelun tuloksen. Yksinkertaisimmillaan luokan valintaan käytetään vain lähintä naapuria, mutta usein etsitään k kappaletta lähimpiä naapureita, jotka äänestävät luokittelun tuloksen. Kun luokittelussa käytetään useampaa kuin yhtä lähintä naapuria, puhutaan usein $k:n$ lähimmän naapurin menetelmästä, englanniksi k -Nearest Neighbor Rule (lyhennetään k -NN).

Tutkielman luvussa 2 esittelen tarkemmin $k:n$ lähimmän naapurin menetelmän toimintaperiaatetta ja sen käytön eri vaiheita. Luvussa 3 käydään läpi joitain esimerkkejä koneoppimisjärjestelmistä, joissa menetelmää ollaan käytetty. Nel-

jännessä luvussa vertaillaan esiteltyjen käyttötapauksien tapoja hyödyntää lähimmän naapurin menetelmää ja viidennessä luvussa tehdään vertailussa tehtyjen havaintojen perusteella yhteenveto menetelmän hyvistä ja huonoista puolista.

2. K :n lähimmän naapurin menetelmä

K :n lähimmän naapurin menetelmä on parametriton, ohjattua oppimista käytävä, instanssiperusteinen luokittelumenetelmä. Puretaan edellinen lause osiin.

Parametrittomuus tarkoittaa tilastollisissa menetelmissä sitä, että tilastoinnissa ei tehdä oletuksia muuttujien todennäköisyysjakaumasta [Duda *et al.* 2000]. Luokittimen tapauksessa tämä tarkoittaa käytännössä sitä, että tunnetussa datassa esiintyvää luokkajakaumaa ei käytetä perusteena luokittelussa. Kun k -NN-algoritmi etsii piirreavaruudesta luokiteltavan instanssin lähimpiä naapureita, koko opetusaineiston luokkajakaumaan ei kiinnitetä huomiota.

Ohjatusti oppiva koneoppimismenetelmä saa opetusdatana tunnettujen datan instanssien piirrevektorien lisäksi instanssien luokat [Duda *et al.* 2000]. Näin järjestelmän ei tarvitse itse tehdä luokkapäätelmiä opetusdatalle, vaan sen tehtäväksi jää ainoastaan luoda malli, jonka perusteella se voi luokitella tuntemattomia instansseja.

Instanssiperusteinen oppiminen on yksi laiskan oppimisen (lazy learning) muoto. Koneoppimisjärjestelmä ei tee saamastaan opetusdatasta eksplisiittistä yleistämismallia, vaan vertailee luokiteltavaa instanssia suoraan opetusdatan instanssien kanssa [Alpaydin 2009]. Tämä tekee opetusvaiheesta nopean, sillä järjestelmä ei käsittele opetusdataa mitenkään, se vain säilötään muistiin. Tämä tekee mukautuvasta oppimisesta (online learning) helppoa, sillä algoritmillemme tarvitsee vain syöttää lisää opetusdataa. Toisaalta muistinkäyttö on luonnollisesti varsin suurta. Myöhemmissä luvuissa mainittavat kaksi muuta luokittelumenetelmää, neuroverkot ja tukivektorikone, eroavat k -NN-luokittimesta olennaisesti nimenomaan käyttämällä opetusvaiheessa luomaansa eksplisiittistä luokittelumallia raa'an opetusdatan sijaan.

Lähimmän naapurin menetelmän toiminta perustuu siihen perusoletukseen, että samaa luokkaa edustavat instanssit omaavat vastaavanlaisia piirteitä. Kun instanssien piirteet sijoitetaan vektoreihin ja nämä vektorit kuvataan pisteiksi avaruuteen, niin samaa luokkaa edustavien instanssien pisteet sijoittuvat lähelle toisiaan. Kun luokiteltavaksi annetaan tuntematon instanssi, kuvataan se samaan avaruuteen opetusdatan kanssa. Luokiteltavaa instanssia lähimpänä olevia opetusdatan instanssien luokkia voidaan sitten käyttää instanssin luokan todennäköisyysjakauman muodostamiseen [Fix and Hodges 1951], ja luokittelun tulokseksi valitaan jakauman mukaan todennäköisin luokka.

Seuraavissa kohdissa perehdytään tarkemmin lähimmän naapurin menetelmän toiminnan ja käytön eri vaiheisiin.

2.1. Datan valmistelu ja esiprosessointi

Ennen k -NN-luokittimen käyttöönottoa päätetään luokittelussa käytettävät piirteet. Hyviä piirteitä ovat sellaiset muuttujat, jotka saavat tiettyä luokkaa edustavissa instansseissa samankaltaisia arvoja. Tämä on edellytys luokittimen olettamalle saman luokan instanssien läheisyydelle piirreavaruudessa. Paljon erilaisia muuttujia sisältävästä datasta valitaan usein käyttöön vain luokittelun kannalta olennaisimmat muuttujat luokittelussa käytettäväksi piirteiksi (feature selection) [Alpaydin 2009]. Huonosti valitut piirteet voivat aiheuttaa luokitteluun kohinaa, eli siirtää piirrevektorin sijoitusta avaruudessa ilman hyvää syytä. Lisäksi hyvin moniulotteisissa avaruuksissa kohdataan usein ulottuvuuksien kirousta (curse of dimensionality), jolloin vektorit kuvautuvat helposti kauas toisistaan avaruudessa, hämärtäen läheisyyden käsitettä [Alpaydin 2009]. Näiltä ongelmilta vältytään helpoiten valitsemalla luokitteluun mukaan mahdollisimman pieni, mutta silti kattava määrä relevantteja piirteitä.

Monimutkaista dataa luokiteltaessa usein on käytännöllistä tehdä varsinaisesta datasta yksinkertaisempi malli luokittelijalle. Tämä voidaan tehdä esim. piirreirroituksella (feature extraction), jossa datan muuttujista johdetaan uusia piirteitä [Alpaydin 2009]. Esimerkiksi kuvadataa luokiteltaessa piirteitä voidaan johtaa laskemalla pikselidatasta kuvan esittämää kohdetta kuvaavia muuttujia, ja käyttämällä niitä luokittelussa pikselidatan sijaan [Babu *et al.* 2014].

2.2. Opetusvaihe

Varsinaisen k -NN-luokittimen käyttö alkaa opetusdatan syöttämisellä. Opetusdatan instanssien piirteet sijoitetaan piirrevektoreihin, jotka syötetään luokittelijalle. Opetusdatan piirrevektorien lisäksi luokittimelle annetaan instanssien luokat, kuten ohjattu oppiminen edellyttää.

Ohjatun oppimisen ja instanssiperusteisuuden ansiosta k -NN-luokittimen opettaminen on nopea operaatio. Yksinkertaisimmillaan luokitin tallentaa tämän opetusdatan sellaisenaan muistiinsa ja on heti valmis käyttöönottoa varten. Usein kuitenkin opetusdata sijoitetaan erilliseen tietorakenteeseen, kuten kd-puuhun [Grother *et al.* 1997], jonka on tarkoitus pienentää luokitteluvaiheessa läpikäytävän opetusdatan määrää opetusvaiheen nopeuden kustannuksella. Tietorakenteen valintaan vaikuttaa myös tarve mukautuvalle oppimiselle, jolloin opetusdataa lisätään tai poistetaan luokittimen käyttöönoton jälkeen. Jos mukautuvuudelle ei ole tarvetta, ei opetusdataan ole tarvetta koskea opetusvaiheen jälkeen.

2.3. Luokittelu

Tuntemattoman instanssin luokittelu voidaan jakaa karkeasti kahteen vaiheeseen, lähimmäisten naapurien etsintään ja äänestämiseen. Naapurien etsintävaiheessa opetusdata ja luokiteltava instanssi kuvataan piirreavaruuteen ja avaruudesta etsitään k (k on positiivinen kokonaisluku) opetusdatan instanssia, joiden pisteet ovat lähimpänä luokiteltavan instanssin pistettä. Löydetyt opetusdatan instanssit äänestävät luokittelun tuloksen. Yksinkertaisimmassa tapauksessa, kun k :n arvo on yksi, luokittelun tulokseksi valitaan suoraan lähimmän naapurin luokka.

Tutustutaan tarkemmin eri tapoihin käytännössä toteuttaa etsintävaihe ja äänestysvaihe.

2.3.1. Lähimpien naapurien etsintä

Naapurien etsintävaiheessa haetaan piirreavaruudessa luokiteltavaa instanssia "lähinnä" olevat opetusdatan instanssit. Lähimmäisyyden määrittelyyn on kuitenkin monia tapoja, ja käytetty määritelmä on kriittinen osa parametrittomien luokittelumenetelmien toimintaa [Alpaydin 2009]. Erityyppisten piirteiden kanssa käytettäviksi sopivat erilaiset määritelmät.

Jatkuvista muuttujista koostuvien piirrevektorien etäisyyksien laskemiseen voidaan käyttää perinteisiä, yleisesti etäisyyden laskemiseen käytettyjä funktioita. Yksi tällainen funktio on euklidinen etäisyys, joka on koulun geometrian tunneilta tuttu vektorin pituuden kaava yleistettynä n -ulotteiseen avaruuteen. Euklidisen etäisyyden kaava n -ulotteisessa avaruudessa pisteiden p ja q välillä on

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2}.$$

Diskreettejä, varsinkin luokitteluasteikollisia piirteitä käytettäessä perinteiset etäisyysfunktioit eivät ole aina mielekkäitä. Jos esimerkiksi yksi piirre olisi henkilön lempiväri, niin miten eri värien "etäisyyksiä" toisiinsa voidaan mielekkäästi verrata? Entä miten "värien etäisyydet" vaikuttavat muuten mielekkäästi järjestettävistä muuttujista koostuvien piirrevektorien sijoittumiseen piirreavaruudessa? Tämänkaltaisten ongelmien vuoksi kategoristen muuttujien "etäisyytenä" käytetään erilaisia datainstanssien samankaltaisuutta vertailevia metriikkoja, kuten Hammingin etäisyyttä [Alpaydin 2009].

Etäisyyksien laskentaa voidaan abstraktoida antamalla luokittimen tutkia ja oppia käytettävä etäisyysfunktio. Tähän tarkoitukseen on kehitetty mm. neighborhood components analysis -menetelmä [Goldberger et al. 2004]. Myös vaihtoehtoisia tapoja määritellä etäisyys, kuten neuroverkkoja [Vajda and Szocs 2014], on tutkittu.

Opetusdatan sisältäessä usein tuhansia instansseja on luokittelun nopeuden kannalta hyvä pyrkiä vähentämään vertailuun mukaan otettavia instansseja. Tämä voidaan tehdä esim. opetusdatan säilömiseen käytetyn tietorakenteen avulla [Grother *et al.* 1997] tai jakamalla opetusdata osiin ja tutkimalla vain jollain menetelmällä olennaiseksi päätettyä osaa [Yamamoto *et al.* 2014a].

2.3.2. Luokittelutuloksen äänestys

Kun k lähintä naapuria on löydetty, voidaan äänestää luokittelun tulos. Äänestysvaiheeseen mukaan otetut naapurit äänestävät omaa luokkaansa luokittelun tulokseksi, ja eniten ääniä saanut luokka valitaan tulokseksi. Kun k :n arvo on yksi, eli äänestysvaiheeseen otetaan mukaan vain yksi opetusdatan instanssi, äänestyksen tulokseksi saadaan triviaalisti kyseisen instanssin luokka.

Kun k on suurempi kuin yksi, tarvitsee naapurien ääniä oikeasti laskea. Yksinkertaisimmillaan jokainen äänestykseen otettu naapuri antaa yhden tasaver-
taisen äänen [Alpaydin 2009]. Tällaisissa äänestyksissä voi kuitenkin esiintyä tasapelitilanteita, jotka joudutaan ratkomaan esimerkiksi arpomalla. Arvontaan perustuva luokittelu ei kuitenkaan kuulosta erityisen mielekkäältä, joten tasapelitilanteita pyritään välttämään. Yleinen keino välttää tasapelejä on valita pariton k :n arvo, sillä sekaannusta on usein kahden "naapuriluokan" välillä [Alpaydin 2009]. Tällaisissa tapauksissa siis useammankin luokan välinen luokittelu muistuttaa kahden luokan välistä luokittelutapausta, jolloin pariton äänimäärä useimmiten estää tasatilanteet.

Naapurien ääniä voidaan halutessa painottaa etäisyyden perusteella siten, että lähimpien naapurien äänet ovat merkitsevempiä. Painottamalla lähimpien naapurien ääniä pyritään vähentämään kaukaisten naapurien tuomaa kohinaa äänestyksessä [Alpaydin 2009]. Otetaan esimerkiksi tilanne, jossa luokitellaan harvinaisen luokan instanssia, joka osuukin avaruudessa lähelle saman luokan opetusinstansseja. Jos k on tarpeeksi suuri ja oikean luokan instansseja ei ole kuin muutama, niin kauempana olevat, yleisemmän luokan instanssit silti voittavat painottamattoman äänestyksen ja antavat siten väärän luokittelutuloksen. Painotukseen voidaan käyttää esimerkiksi etäisyyden käänteislukua [Cunningham and Delany 2007].

Edellisen kappaleen esimerkistä nousee kysymys siitä, mikä on hyvä k :n arvo. k :n arvon valintaan vaikuttaa monet tekijät. Suurempi k yleisesti ottaen tasoi-
taa rajoja luokkien välillä [Alpaydin 2009], jolloin yksittäisten (ja mahdollisesti poikkeavien) instanssien vaikutus äänestyksessä vähenee. Liian suuri k toisaalta tuo mukaan kohinaa valitsemalla äänestykseen liian kaukaisia naapureita. Käyttökohteeseen sopivinta k :n arvoa voidaan etsiä esimerkiksi ristiinvalidoimalla testattuja arvoja [Alpaydin 2009].

2.4. Suorituskyky

Luokittimien tarkkuus riippuu monista tekijöistä, kuten edellisissä luvuissa kuvatuista toteutusyksityiskohdista ja opetusdatan laadusta. Jos oletetaan, että käytössä on kaikilta osin toimiva luokittimen toteutus, voidaan sen tarkkuutta tutkia tilastollisesti. Luokittimien tarkkuutta verrataan usein bayesilaiseen virhemäärään (bayes error rate), eli pienimpään mahdolliseen teoreettiseen luokittimen virhemäärään. Useamman luokan välillä luokittelevan k -NN-luokittimen virhemäärä asettuu välille

$$R^* \leq R_{k-NN} \leq R^* \left(2 - \frac{MR^*}{(M-1)} \right),$$

jossa R^* on bayesilainen virhemäärä, R_{k-NN} luokittimen todellinen virhemäärä ja M on mahdollisten luokkien määrä [Cover and Hart 1967]. Tilanteesta, jossa $M = 2$ ja R^* lähestyy nollaa, saadaan kahden luokan k -NN-luokittimen maksimivirhemääräksi kaksi kertaa bayesilainen virhemäärä.

Tulosten tarkkuuden lisäksi luokittimien suorituskykyä voidaan mitata mm. muistinkäytössä ja nopeudessa. Yleisellä tasolla voidaan todeta, että laiskan oppimisen johdosta k -NN-luokittimet käyttävät paljon muistia, sillä ne joutuvat säilyttämään varsinaista opetusdataa sellaisenaan muistissa. Naiivisti kaikki opetusdatan instanssit läpikäymällä k lähintä naapuria löydetään ajassa $O(dn^2)$ [Duda *et al.* 2000], jossa d on piirteiden määrä ja n opetusdatan instanssien määrä. Useissa reaalimaailman toteutuksissa on kuitenkin erilaisia optimointeja, kuten aiemmin mainittu opetusdatan säilöminen kd -puuhun, joilla pyritään vähentämään läpikäytävää opetusdataa ja näin nopeuttamaan naapurien etsimistä.

2.5. Mahdollisia ongelmakohtia

Opetusdataa kerätessä ja valitessa tulee olla tarkkana, ettei niiden pohjalta tehty yleistämismalli kärsi ylisovittamisesta (overfitting) [Duda *et al.* 2000]. Ylisovittamisesta kärsivä yleistämismalli ei kuvaa niinkään koko populaatiota, vaan pelkästään opetusdataa. Näin voi käydä esimerkiksi tilanteissa, joissa reaalimaailmassa esiintyviä limittäisyyksiä luokkien välillä ei esiinny opetusdatassa, tai kun reaalimaailmassa harvinaisia instansseja esiintyy suhteessa enemmän opetusdatassa.

Piirteiden saamien arvojen vaihteluväli riippuu piirteen tyyppistä. Esimerkiksi ihmisen ikä voi saada keskimäärin arvoja 0-100 vuoden välillä ja pituus 40-230 cm välillä. Kun vaihtelevilla skaaloilla olevia piirteitä käytetään sellaisenaan etäisyyksien laskemisessa, saavat suuremman vaihteluvälin omaavat piirteet helposti suuremman painotuksen etäisyyslaskuissa. Tämä ongelma voidaan ratkaista esimerkiksi normalisoimalla piirteiden arvot siten, että niiden varianssit ovat samoja [Alpaydin 2009].

3. Käyttökohteet

Koneoppimisjärjestelmillä ja niiden luokittimilla on aina jokin tietty tehtävä, joiden täyttämiseen suunnitellaan. Seuraavissa kohdissa esitellään joitain reaali-maailman koneoppimisjärjestelmiä, joiden toteutuksen osana on käytetty k -NN-luokitinta. Eri käyttökohteista annetaan yleistason kuvaus, niiden luokitteluun liittyvät ongelmat ja ratkaisut sekä arviota k -NN-perustaisen luokittelun soveltuvuudesta niihin. Pyrin myös tuomaan esille mielenkiintoisia luokittimien toteutusyksityiskohtia niiltä osin kuin niitä on saatavilla.

3.1. Videopelit

Videopelien tekoälyt hyödyntävät harvoin koneoppimismenetelmiä päätöksenteossaan, vaan toimivat pääasiassa erilaisten sääntöpohjaisten algoritmien pohjalta [Rhalibi *et al.* 2009]. Usein tästä ei ole haittaa ja joissain peleissä nimenomaan halutaan, että vastustajat toimivat aina kaavamaisesti samalla tavalla. Koneoppimista hyödyntävien vastustajien käyttö soveltuu kuitenkin sellaisiin peleihin, joissa ihmisvastustajan korvikkeeksi luodun tietokonevastustajan toiminnan liiallinen ennustettavuus voi heikentää pelikokemusta. Yksi esimerkki tällaisesta pelityypistä ovat taistelupelit, joissa kaksi pelaajaa yrittävät hyökkäyksillään tiputtaa toisen pelaajan elämäpisteet nolnaan. Eri pelitilanteisiin aina samalla tavalla reagoivan tietokonevastustajan heikkouksia on helppo hyväksikäyttää, sillä se ei opi ja yritä uusia asioita samalla tavalla kuin ihmisvastustaja [Yamamoto *et al.* 2014a]. Japanilainen tutkimusryhmä on järjestänyt vuodesta 2013 taistelupelitekoälykilpailua [Lu *et al.* 2013] ja julkaissut artikkeleita koneoppimista ja k -NN-luokitinta päätöksenteossaan hyödyntävästä tekoälystä, MizunoAI:sta [Yamamoto *et al.* 2014a].

MizunoAI:n toiminta perustuu vastustajan toimintojen ennustamiseen ja pelin etenemisen simulointiin ennustuksen pohjalta [Yamamoto *et al.* 2014a]. Vastustajan toiminnon ennustaminen tapahtuu kahdessa osassa. Ensin ennustetaan, aikooko vastustaja hyökätä ja sitten ennustetaan todennäköisimmät hyökkäykset. Jos vastustajan ennustetaan hyökkäävän, simulaattori käy läpi kaikki Mizunon hahmon käytettävissä olevat toiminnot, ja suorittaa sen toiminnon, joka simulaattorin mukaan aiheuttaa Mizunon kannalta parhaan muutoksen hahmon elämäpisteissä, eli se käytännössä pyrkii minimoimaan omat pistemenetyksensä ja maksimoimaan vastustajan pistemenetykset.

Mizuno käyttää k -NN-luokitinta vastustajan todennäköisimmän hyökkäyksen ennustamiseen [Yamamoto *et al.* 2014a]. Luokitinta opetetaan mukautuvasti aina vastustajan hyökätessä. Käytettävät piirrektorit sisältävät pelaajahahmon etäisyyden toisiinsa kaksiulotteisella koordinaatistolla ja luokkina käytetään

vastustajan käyttämiä hyökkäyksiä. Vastustajan todennäköisimmän hyökkäyksen ennustaminen perustuukin puhtaasti siihen, mitä hyökkäyksiä vastustaja on pelin nykytilaa vastaavalta etäisyydeltä aiemmin käyttänyt.

MizunoAI käyttää opetusdatan tallentamiseen yhden tietorakenteen sijaan neljää tietorakennetta [Yamamoto *et al.* 2014a]. Pelaajahahmot voivat olla joko maassa tai ilmassa, ja kukin tietorakenne vastaa eri tilojen yhdistelmää (maa – maa, maa – ilma, ilma – maa ja ilma – ilma). Syytä tälle ratkaisulle ei artikkelissa kerrota, mutta sillä on todennäköisesti tekemistä eri tilanteiden toistensa poisulkevuudessa pelilogiikan ja päätöksenteon kannalta. Vastustajan ollessa ilmassa tämä ei pysty tekemään maahyökkäyksiä, ja esimerkiksi maasta tehdyt heittohyökkäykset osuvat ainoastaan maassa olevaan vastustajaan. Vastustajan toimintojen voidaan siis olettaa muuttuvan vastustajan ja Mizunon tilojen yhdistelmien mukaan. Kun luokittelussa käytetään vain nykyisen pelitilanteen kannalta olennaista opetusdataa, nopeutetaan algoritmin toimintaa ja vähennetään epäolennaisen datan tuomaa kohinaa luokittelussa.

Nykytilannetta vastaavia opetusdatan tilanteita etsiessä Mizuno käy läpi kaikki nykyistä pelaajien tilaa vastaavat opetusdatan instanssit [Yamamoto *et al.* 2014a], ja etäisyysfunktiona käytetään euklidista etäisyyttä [Yamamoto *et al.* 2014b]. Mizunon luokittimen äänestyksessä ei käytetä painotuksia, vaan kaikkien k -lähimmän naapurin äännet ovat samanarvoisia [Yamamoto *et al.* 2014b]. Äänestyksessä tasatilanteet ovat sallittua, ja kaikki äänestyksessä eniten ääniä saaneet hyökkäykset lähetetään simulaattorille [Yamamoto *et al.* 2014a]. Tutkimuksessa k :n arvoja testattiin välillä 1-11, ja tulokset olivat vastustajasta kerätyn datan määrästä ja vastustajan taipumuksista riippuvaisia [Yamamoto *et al.* 2014a]. Testeissä Mizuno voitti vuoden 2013 taistelupelitekoälykilpailun kolme parasta osallistujaa.

3.2. Konenäkö

Konenäkö on tietokoneen tekemää kuvan sisällön analysointia, ja luokittimet ovat olennainen osa konenäköjärjestelmiä. Konenäön käyttökohteita on mm. viivakoodinlukijat, laaduntarkkailu, esineiden laskenta ja käänteinen kuvahaku.

Kuvien luokittelu voi perustua kuvan pikselidataan [LeCun *et al.* 1995] tai esiprosessointivaiheessa kuvasta laskettuihin piirteisiin [Babu *et al.* 2014]. LeCunin ja muut [1995] toteuttivat osana erityyppisten käsinkirjoitettujen numeroiden luokittimien vertailua euklidista etäisyyttä käyttävän k -NN-luokittimen, joka käytti piirrevektoreina 20×20 pikselin kokoisten harmaasävykuvien pikselidataa, jossa jokainen pikseli vei 4 tavua tilaa. Piirreavaruudessa oli siis 400 ulottuvuutta, ja 60 000 kuvan testidataa vei 12 megatavua tilaa. Muistivaatimukset ovat moninkertaisia testissä mukana olleisiin neuroverkkopohjaisiin luokittimiin, joiden luokittelumallit veivät alle puoli megatavua tilaa [LeCun *et al.* 1995].

k -NN-luokitin hävisi 97,6% luokittelutarkkuudellaan suurimmalle osalle testatuista neuroverkoista, mutta toisaalta yhtä lukuun ottamatta niiden opetusajaksi kesti vähintään kaksi viikkoa. Artikkelissa todetaankin, että realistinen k -NN-luokitin käyttäisi piirrevektoreita pikselidatan sijaan, jolla siis viitataan piirreirroitukseen.

Babu ja muut [2014] lähestyvät numeroiden tunnistamisen ongelmaa laskeamalla kuvista 10 piirrettä esiprosessointivaiheessa, joita ovat mm. silmukoiden ja eri suuntien "vesisäiliöiden" määrä. Vesisäiliöllä tarkoitetaan koveraa aluetta, johon kuvan yhdestä laidasta vastakkaiselle laidalle valuva neste jäisi jumiin. Esimerkiksi numeromerkin 5 alaosassa on usein vasemmanpuolinen vesisäiliö ja yläosassa oikeanpuolinen säiliö. Tämän tyylinen piirteiden etsintä vaatii paljon opetusdataa, sillä esimerkiksi huonosti suljetut silmukat voivat muodostaa vesisäiliöitä [Babu *et al.* 2014]. Viidenkymmentuhannen kuvan opetusaineistolla euklidista etäisyyttä käyttänyt k -NN-luokitin pääsi 96,94% tarkkuuteen, kun $k = 1$ [Babu *et al.* 2014]. Luokittelun tarkkuus on siis vertailukelpoinen LeCunin ja muiden [1995] k -NN-luokittimen saavuttamaan 97,6% tarkkuuteen, mutta huomattavasti pienemmän piirremäärän ansiosta luokittelun pitäisi olla nopeampaa ja muistivaatimusten pienemmät.

Edelliset esimerkit numeromerkkien tunnistamisesta vaativat toimiakseen tarkkaan valittuja kuvia, joissa on vain yksi numeromerkki. Tarkastellaan monimutkaisempaa kuvadataa käyttävänä esimerkkinä Szummerin ja Picardin [1998] luomaa luokitinta, joka päättää, onko luokiteltava valokuva otettu sisätiloissa vai ulkona. Heidän luokittimensa perustuu kuvan jakamiseen 4×4 matriisiin, ja kaikille matriisin osille tehdään erikseen osion väreihin ja tekstuureihin perustuvat k -NN luokittelut. Erikseen tehtävissä väri- ja tekstuuriluokittelussa on etuina mahdollisuus käyttää kyseisille piirteille sopivia etäisyysfunktioita ja piirreavaruuksien pienemmät ulottuvuudet [Szummer and Picard 1998]. Etäisyysfunktioina käytettiin väreille histogrammien leikkausta ja tekstuureille erityistä tekstuurien luokitteluun käytettävää funktiota. Koko kuvan luokittelun tulos saadaan yhdistämällä osatulokset erillisessä luokitteluvaiheessa, joka toteutettiin yksinkertaisella enemmistöäänestyksellä. Lopullinen luokitin ylsi 90,3% tarkkuuteen, ja väärin luokitelluissa kuvissa oli usein toiselle luokalle ominaista sisältöä, kuten sisälle tuotu joulukuusi [Szummer and Picard 1998]. Pelkkää väridataa hyödyntämällä saatiin 75,4% luokitteluista oikein, ja pelkillä tekstuureilla saatiin 86,2%, eli osatulosten yhdistäminen selvästi paransi luokittelutarkkuutta.

Esimerkkejä vertaamalla huomaa luokittimien toteutuksista helposti sen, että valokuvien luokittelussa on jouduttu pilkkomaan pienempiin osiin, mutta pelkkiä numeroita luokiteltaessa riittää yksi luokittelukierros. Toisaalta ongelmatkin ovat erilaisia, numeroita luokiteltaessa tutkitaan pelkästään kuvan

sisältöä ja valokuvan ottopaikkaa luokitellessa kuvan sisällön merkitystä. Szummerin ja Picardin [1998] mukaan vielä 1990-luvun lopulla konenäkö tutkimuksessa ei juurikaan keskitytty kuvien semantiikkaan. Pääsyyksi tähän he arvelivat ongelman vaikeutta.

3.3. Tekstin luokittelu

Viimeisenä esimerkkinä tarkastelemme tekstien luokittelua. Johdannossa mainitun roskapostifiltterin lisäksi tekstin koneellista luokittelua voidaan käyttää esimerkiksi automaattiseen artikkelien ja Internet-sivujen kategoriointiin. Tekstien luokittelussa nousee myös esille datan sisällön lisäksi datan semantiikan ymmärtäminen, sillä esimerkiksi homonyymit tai sarkasmi muuttavat yksittäisten sanojen merkitystä kontekstin mukaan. Tässä luvussa esiteltävät k -NN-luokittimet toimivat kuitenkin puhtaasti teksteistä löytyvien sanojen pohjalta, ja minkäänlaista eksplisiittistä tekstin semanttisen merkityksen tutkintaa ei tehdä missään vaiheessa.

K -NN pohjaisia sähköpostifilttereitä ovat tutkineet mm. Tretyakov [2004] ja Lai [2007], joiden lähestymistavat tutkimusongelmaan ja tulokset ovat samankaltaisia. Heidän toteutustensa piirrevektorit sisältävät piirteen kaikille lähdeaineistossa esiintyville sanoille, ja niiden arvot sisältävät binäärisenä tiedon siitä, esiintyykö kyseinen sana sähköpostissa vai ei. Tämä ratkaisu johtaa todella suuriin ulottuvuuksiin, Tretyakovin [2004] tapauksessa ulottuvuuksia on huimat 21 700. Kumpikaan ei suoraan nimeä käyttämäänsä etäisyysfunktia, mutta Tretyakov [2004] kertoo käyttävänsä tekstien erilaisuutta mittaavien metriikkojen sijaan "vektorien etäisyyttä" ja Lai [2007] sanoo käytännössä kaikkien k -NN-luokittimien käyttävän euklidista etäisyyttä, antaen ymmärtää niin olevan myös hänen luokittimensa tapauksessa.

Molemmat tutkijat päätyivät samaan lopputulokseen: k -NN-pohjaiset sähköpostifiltterit eivät pärjää vertailussa tukivektorikonetta käyttäville ratkaisuille [Tretyakov 2004; Lai 2007]. Tretyakovin [2004] testissä k -NN-luokittimen 90,8% luokittelutarkkuus hävisi tukivektorikoneelle 7,3 prosentilla, ja Lain [2007] lukuisissa testeissä k -NN-luokittimen noin 70-80% tarkkuus oli useimmiten noin 10% tukivektorikonetta jäljessä.

Tretyakov [2004] nostaa esille myös virheiden eriarvoisuuden: ajoittain filteristä läpi pääsevä roskaposti on pienempi ongelma kuin väärin perustein roskapostiksi merkitty asiallinen sähköposti. Nämä "väärät positiiviset" olivatkin hänen testeissään k -NN-toteutuksen suurin ongelma. Muokattuaan roskapostiksi lajittelun vaatimaan äänestysvaiheessa määränemmistön (51 naapurista 35 piti olla roskaposteja) hän pääsi eroon "vääristä positiivisista" mutta lajittelun tarkkuus tippui 69,3 prosenttiin [Tretyakov 2004]. Tukivektorikoneen luvut ennen ja jälkeen "väärin positiivisten" pois kitkemistä olivat 98,1% ja 90,8%. Lai

[2007] puolestaan teki useita testejä eri sähköpostin osien (otsakkeet, otsikko, leipäteksti, kaikkien yhdistelmä) ja piirreirroitustekniikoiden yhdistelmillä. Kaikissa testeissä, joissa ei oltu poistettu kielen yleisimpiä sanoja (stop words) k -NN-suodattimen tulos heikkeni 70-82% tarkkuudesta 32-40% tarkkuuteen, kun tukivektorikoneella poisjätettyjen sanojen vaikutus oli vain parin prosentin luokkaa ja luokittelutarkkuus yleisestikin yli 90%.

Tretyakovin ja Lain k -NN-sähköpostifiltterit olivat toteutuksiltaan varsin naiveja, eivätkä yrittäneet optimoida tai muokata menetelmää tekstien luokitteluun paremmin soveltuvaksi. Lain [2007] tuloksista havaitaan etenkin ulottuvuuksien vähentämisen tärkeys, ja Yang ja Pedersen [1997] ovatkin tutkineet "stop wordien" poistamisen lisäksi muita keinoja vähentää luokittelussa huomioitavien sanojen määrää. Tilastomenetelmiin perustuvien khiin neliö -testin ja informaatiohyödyn (information gain) perusteella epäolennaiseksi todettujen sanojen poisjättäminen todettiin tehokkaaksi, kuten myös liian harvassa tekstissä esiintyvien sanojen poisjättäminen. Nämä kolme menetelmää pystyivät tiputtamaan luokittelussa käytettyjen sanojen määrää kahdeksasosaan alkuperäisestä ennen kuin luokittelutulokset kärsivät [Yang and Pedersen 1997]. Toinen mahdollinen lähestymistapa tuloksien parantamiseen on painottaa sanojen vaikutusta etäisyysfunktiossa [Han *et al.* 2001]. Hanin ja muiden testeissä opetusdatan pohjalta painotusten luominen paransi tuloksia painottamattomaan etäisyysfunktioon nähden käytetystä testiaineistosta riippuen noin 5-15%. Optimoinneista huolimatta k -NN-luokittimet kuitenkin yltävät vain yksittäisissä testitapauksissa yli 90% tuloksiin [Yang and Petersen 1997; Han *et al.* 2001].

4. Käyttökohteiden vertailua

Tarkastellaan seuraavaksi edellisessä luvussa esiteltyjen luokittimien käyttökohteiden ja toteutusten samankaltaisuuksia ja eroavaisuuksia. Ensimmäisenä eroavaisuutena huomataan käyttökohteiden olevan ensisilmäyksellä luonteeltaan varsin erilaisia, esimerkiksi videopelivastustajan seuraavan liikkeen ennustaminen ja sähköpostien filttäminen ovat täysin erilaisia tehtäviä ihmisen kannalta. Vektorien etäisyyksiin perustuvassa luokittelussa ei kuitenkaan oteta juurikaan kantaa itse kohdealueeseen, joten nämä ihmisen kannalta varsin erilaiset tapaukset muuttuvat koneellisen luokittimen näkökulmasta varsin samanlaisiksi tapauksiksi.

4.1. Toteutusten vertailu

Luokittimien toteutuksissa eroavaisuuksia löytyy piirteiden valinnassa, etäisyysfunktioissa, k :n valinnassa ja erilaisissa käyttökohteellisissa optimoinneissa ja laajennuksissa.

4.1.1 Piirteiden valinta

Luokittelussa käytettävät piirteet riippuivat pitkälti käyttökohteesta ja saatavilla olevasta datasta. Piirteiden valinta voi olla yksinkertainen prosessi, kuten MizunoAI:n luokittimella [Yamamoto *et al.* 2014a], jonka luokittelutehtävää voidaan kuvailla kysymyksellä ”mitä hyökkäystä vastustaja todennäköisesti käyttää tältä etäisyydeltä”. Vastaus kysymykseen saadaan yksinkertaisesti pelihahmojen etäisyydellä kaksiulotteissa pelimaailmassa. Toisaalta sähköpostifiltterien [Tretyakov 2004; Lai 2007] piirrevalintatehtävä on todella vaikea, sillä yksittäisiin sähköposteihin liittyy todella paljon dataa, josta luokittelun kannalta olennaisimpien piirteiden valitseminen tai johtaminen ei ole triviaali tehtävä. Molemmat esiteltyt roskapostifiltterit tyytyivät hieman brute force -tekniikkaa muistuttavaan tapaan ottaa piirteiksi lähes kaikki aineistossa esiintyvät sanat.

LeCunin ja muiden [1995] ja Babun ja muiden [2014] erilaiset lähestymistavat samaan käsikirjoitettujen numeromerkkien tunnistamisongelmaan näkyvät selvimmin juuri luokittelussa käytetyissä piirteissä. LeCunin ja muiden [1995] raa’asta pikselidatasta muodostetut piirrevektorit sisälsivät 400 piirrettä, kun taas Babun ja muiden [2014] vektorit sisälsivät 10 pikselidatasta piirreirroituksella johdettua piirrettä. Hyvin erilaisista lähestymistavoista huolimatta molemmat luokittimet olivat tarkkoja ja tulosten tarkkuus prosenttien sisällä toisistaan (Babu ja muut 96,9% ja LeCun ja muut 97,6%). Käyttökohde ei siis ilmeisesti kärsi suurten ulottuvuuksien tuomista ongelmista. Toisaalta pitää muistaa, että Babun ja muiden pienempien piirrevektorien välisten etäisyyksien laskemisen pitäisi olla selvästi nopeampaa ja muistivaatimusten vähäisempiä.

4.1.2 Etäisyysfunktiot

Euklidinen etäisyys oli yleisin käytetty etäisyysfunktio. Se mainittiin muutama artikkelissa nimeltä [LeCun *et al.* 1995; Babu *et al.* 1995] ja on löydettävissä MizunoAI:n lähdekoodista [Yamamoto *et al.* 2014b]. Kyseessä on varsin yleispätevä metriikka, jolla saadaan vähintäänkin kohtuullisia tuloksia monenlaisissa käyttökohteissa. Lain [2007] kommentti siitä, että ”käytännössä kaikki” k -NN-luokittimet käyttävät euklidista etäisyyttä on kuitenkin hieman liioiteltu.

Esitellyistä käyttötapauksista ainoastaan Szummer ja Picard [1998] mainitsivat suoraan käyttävänsä jotain muuta etäisyysfunktioita kuin euklidista etäisyyttä. Heidän kahdesta osaluokittimesta koostuva luokitin käytti toisessa osassa värihistogrammien leikkausta ja toisessa erityistä tekstuurien luokittelumetriikkaa. Mielenkiintoisesti hekin kuitenkin testasivat euklidista etäisyyttä värihistogrammien etäisyysmetriikkana ensimmäisissä testeissään. Tulokset olivat pääosin prosenttien sisällä toisistaan, ja parempi metriikka riippui k :n arvosta ja ja

käytetystä väriavaruudesta [Szummer and Picard 1998]. Tässä tapauksessa etäisyysfunktion vaihtamisesta ei siis saatu merkittävää hyötyä.

4.1.3 K :n arvon valinta

Käytettyjen k :n arvojen esittely ja varsinkin niiden perusteluvat jäivät monissa tutkimusartikkeleissa varsin pienelle huomiolle. Pahimmillaan sitä ei joko mainittu ollenkaan [Lai 2004] tai mainitaan vain ohimennen [LeCun *et al.* 1995; Tretyakov 2004]. Monet esittelivät taulukossa joitain testattuja arvoja [Szummer and Picard 1998; Yamamoto *et al.* 2014; Babu *et al.* 2014], mutta esiteltäväksi valittuja arvoja ei siltikään perusteltu.

Luokittimien käyttämät k :n arvot olivat suurimmilta osin varsin pieniä, väliltä 1-13 [LeCun *et al.* 1995; Szummer and Picard 1998; Yamamoto *et al.* 2014, Babu *et al.* 2014]. Opetusaineistojen koko vaihteli pientä k :n arvoa käyttäneissä luokittimissa noin tuhannesta [Szummer and Picard 1998] kuuteenkymmeneen tuhanteen [LeCun *et al.* 1995]. Suurinta k :n arvoa ($k = 51$) käytti Tretyakov [2004], jonka opetusaineisto oli myös noin tuhannen instanssin kokoinen. Sopivan k :n arvon ja opetusaineiston koon välistä yhteyttä ei siis pysty tällä otannalla juurikaan päättelemään.

4.1.4 Käyttökohdekohtaiset optimoinnit ja muunnokset

Muutamassa tutkimuksessa luokitinta muokattiin erilaisista syistä toimimaan hieman erilaisesti kuin miten luvussa 2 kuvailtu k -NN-luokittimen perusmalli toimii. Pienempiä muutoksia tehtiin muokkaamalla yksittäisiä menetelmän osia, kuten opetusdatan tallennustapaa [Yamamoto *et al.* 2014a] ja vaihtamalla äänestysmenetelmää [Tretyakov 2004]. MizunoAI:n tulee toimia reaaliajassa 60 kertaa sekunnissa päivittyvän pelin tahdissa [Yamamoto *et al.* 2014], joka on varsin vaativa tavoite ja johon tutkittavan opetusdatan minimoiminen auttaa varmasti. Tretyakov [2004] puolestaan halusi tuoda parametrillisille luokittimille tyypillisemmän tavan hienosäätää tuloksia k -NN-luokittimeen muuttamalla äänestysmenetelmän vaatimaan määräänemmistön. Näillä pienillä muokkauksilla pyrittiin muovaamaan luokittimesta paremmin omaan käyttökohteeseen sopiva.

Isoin eroavaisuus perinteiseen k -NN-luokittimeen nähtiin Szummerin ja Picardin [1998] luokittimessa, jossa kaksi k -NN-luokitinta olivat vain osia isommassa luokittelujärjestelmässä.

4.2. Tulosten vertailu

Luokittimien toteutusyksityiskohtien lisäksi on mielenkiintoista vertailla niillä aikaansaatuja tuloksia. Näiden vertailujen pohjalta voidaan tehdä suuntaa antavia päätelmiä siitä, että millaisiin luokittelutehtäviin k -NN-luokitin soveltuu hyvin ja milloin on syytä harkita muita menetelmiä.

Yleisesti ottaen esiteltyistä k -NN-luokittimista kaikkia muita paitsi sähköpostifilttereitä pidettiin ainakin jollain tasolla onnistuneina. Etenkin kirjoitettujen numeroiden luokittelijat pääsivät todella hyvin tuloksiin, molemmilla luokittelun tarkkuus oli yli 96% [LeCun *et al.* 1995; Babu *et al.* 2014]. Näiden kahden luokittimen menestyminen ei yllätä, sillä ne ovat esiteltyistä luokittelutilanteista yksinkertaisimpia, niiden kun ei tarvinnut ottaa kantaa datan semantiikkaan ja ne saivat useita kymmeniä tuhansia opetuskuvia. Datan luokittelu sen sisällön merkityksen pohjalta sai yleisesti huonompia tuloksia, vain binäärinen valokuva-luokitin [Szummer and Picard 1998] ja yksittäiset testitapaukset roskapostifilteröinnissä [Lai 2007] pääsivät yli 90% tarkkuuksiin tulosten pääasiassa ollessa 70-80% luokkaa [Tretyakov 2004; Lai 2007].

Vaikka luokittelutulosten oikeellisuus on usein hyvä ja helposti vertailtava luku, ei siihen voida aina täysin sokeasti luottaa. MizunoAI:n ennustuksista vain 21% toteutui välittömästi, ja 49% seuraavan puolen sekunnin sisällä [Nakagawa *et al.* 2015]. Mizunoa pidettiin kuitenkin onnistuneena ja tehokkaana kokeiluna, sillä saavutettu tarkkuus riitti voittamaan muita kilpailussa menestyneitä tekoälyjä [Yamamoto *et al.* 2014a]. Toisaalta varsin kunnioitettava oikeellisuusprosentti voi myös olla harhaanjohtava, kuten Tretyakovin [2004] ensimmäisissä testeissä, joissa käytettiin normaalia enemmistöäänestystä. Tällöin luokittelun tarkkuus oli 90,8%, mutta virheistä kaksi kolmasosaa oli normaalien sähköpostien luokittelua roskaposteiksi [Tretyakov 2004], tilanne jota Tretyakovin mielestä ei ikinä saanut tapahtua. Kun luokitin muokattiin toimimaan Tretyakovin toivomalla tavalla, tippuikin luokittelun tarkkuus 69,8 prosenttiin.

Kolmessa tutkimuksessa vertailtiin k -NN-luokittimia muihin luokittelumenetelmiin, kuten tukivektorikoneeseen ja neuroverkkoihin. Kaikki vertailut päättyivät samaan lopputulokseen: muut menetelmät suoriutuvat samoista tehtävistä k -NN-luokitinta paremmin [Tretyakov 2004; Lai 2007; LeCun *et al.* 1995]. Etenkin roskapostifiltterien kohdalla vertailu luokittelutarkkuudesta näyttää karuja lukuja: eroa on 10-20% tukivektorikoneen hyväksi [Tretyakov 2004; Lai 2007]. LeCunin ja muiden [1995] vertailuissa k -NN-luokittimen tarkkuus on puolestaan vain reilun prosentin huonompi kuin neuroverkoilla, mutta vaati kymmeniä kertoja enemmän muistia ja luokittelu vei sekunnin kahdeskymmenesosan sijaan kokonaisen sekunnin. k -NN voitti vertailuissa ainoastaan opetusajassa, jota laiskasti oppivalla menetelmällä ei käytännössä ole, kun taas neuroverkoilla opetusajat vaihtelivat kolmesta päivästä kahteen viikkoon.

5. Yhteenveto

Tutkielmassa esiteltiin k :n lähimmän naapurin menetelmän toimintaperiaatetta ja käyttöä erilaisissa luokittelutehtävissä. Esiteltyt luokittelutehtävät olivat selvästi toisistaan poikkeavia, ja monissa tapauksissa menetelmän toimintaa jouduttiin hienosäätämään paremmin käyttökohdetta palvelevaksi, mutta perusperiaatteeltaan menetelmää pystyttiin käyttämään hyvin erilaisissa ympäristöissä. Tutkielmaa kirjoittaessa olikin mielenkiintoista havaita, miten ihmisen kannalta hyvin erilaiset ongelmat pystytään pienellä vaivalla muuttamaan numeroiksi ja ratkaisemaan samalla kaavalla.

Osassa esitellyistä tutkimuksista k -NN-menetelmällä saatiin hyviä tuloksia [Babu *et al.* 2014], ja joissain se puolestaan todettiin testatulle kohdealueelle käytökelvottomaksi sellaisenaan [Tretyakov 2004]. Minkä tyyppisiin luokitteluongelmiin k -NN sitten yleisellä tasolla soveltuu? Empiirisesti asiaa voisi lähestyä pohtimalla, että voisiko ihminen mielekkäästi käyttää samoja datan piirteitä instanssien luokitteluun. Esimerkiksi numeromerkin muotoa kuvailevista piirteistä numeron päättely [Babu *et al.* 2014] kuulostaa hyvin järkeenkäyvältä, kun taas yksittäisten sanojen esiintymisten suhteet sähköpostissa [Tretyakov 2004; Lai 2007] kuulostaa hyvin raskaalta tavalta tunnistaa roskaposteja. Tällä ajattelutavalla voidaan päätellä, että menetelmä soveltuu kohteisiin, joissa on maltillinen määrä instansseja selkeästi kuvaavia piirteitä ja joissa ei pyritä etsimään dataan piilotettuja suurempia totuuksia, kuten semanttisia merkityksiä. Päätelmä on yllättävän hyvin linjassa esiteltyjen luokittimien tulosten kanssa.

Vaikka k :n lähimmän naapurin menetelmän ikä ja heikkoudet modernimpiin menetelmiin nähden tulivat tutkielmassa esille, on sillä kuitenkin käyttöä vielä 2010-luvulla. Menetelmä on paperilla yksinkertainen ja helposti ymmärrettävissä, joten siihen tutustumalla pääsee helposti perille erilaisista koneoppimiseen liittyvistä konsepteista. Oikeassa tutkimustyössä puolestaan nopea opetus-aika tekee erilaisten testaamisesta nopeampaa ja useissa luokitteluongelmissa tulokset ovat niin hyviä, ettei monimutkaisempien menetelmien käyttöönoton viemä aika ole perusteltua.

Viiteluettelo

- [Alpaydin 2009] Ethem Alpaydin 2009. *Introduction to Machine Learning, second edition*. MIT Press.
- [Babu *et al.* 2014] U. Ravi Babu, Y. Venkateswarlu and Aneel Kumar Chinthu 2014. Handwritten digit recognition using k -nearest neighbour classifier. In: *Proc. of the 2014 World Congress on Computing and Communication Technologies*, 60-65.

- [Cover and Hart 1967] Thomas Cover and Peter Hart 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* 13, 1, 21-27.
- [Cunningham and Delany 2007] Padraig Cunningham and Sarah Delany 2007. *k-Nearest neighbour classifiers*. Technical Report UCD-CSI-2007-4, University College Dublin.
- [Duda *et al.* 2000] Richard O. Duda, Peter E. Hart and David G. Stork 2000. *Pattern Classification, 2nd edition*. Wiley.
- [Fix and Hodges 1951] Evelyn Fix and Joseph L. Hodges Jr. 1951. Discriminatory analysis – nonparametric discrimination: consistency properties. Technical Report 4. USAF School of Aviation Medicine.
- [Goldberger *et al.* 2004] Jacob Goldberger, Sam Roweis, Geoffrey Hinton and Ruslan Salkhutinov 2004. Neighbourhood component analysis. In: *Advances in Neural Information Processing System 17*, 513-520.
- [Grother *et al.* 1997] Patrick Grother, Gerald Candela and James Blue 1997. Fast implementations of nearest neighbor classifiers. *Pattern Recognition* 30, 3, 459-465.
- [Han *et al.* 2001] Eui-Hong Han, George Karypis and Vipin Kumar 2001. Text categorization using weight adjusted k-nearest neighbor classification. In: *Proc. of the 5th Pacific-Asia Conference on Knowledge (PAKDD'01)*, 53-65.
- [Lai 2007] Chih-Chin Lai 2007. An empirical study of three machine learning methods for spam filtering. *Knowledge-Bases Systems* 20, 3, 249-254.
- [LeCun *et al.* 1995] Yann LeCun, Larry Jackel, Léon Bottou, Corinna Cortes, John Decker, Harris Drucker, Isabelle Guyon, Urs Müller, Eduard Säcker, Patrice Simard and Vladimir Vapnik 1995. Learning algorithms for classification: a comparison on handwritten digit recognition. In: *Neural Networks: the Statistical Mechanics Perspective*, World Scientific, 261-276.
- [Lu *et al.* 2013] Feiyu Lu, Kaito Yamamoto, Luis Nomura, Syunsuke Mizuno, YoungMin Lee and Ruck Thawonmas 2013. Fighting game artificial intelligence competition platform. In: *Proc. of the 2nd Global Conference on Consumer Electronics*, 320-323.
- [Nakagawa *et al.* 2015] Yuto Nakagawa, Kaito Yamamoto, Chu Chun Yin, Tomohiro Harada and Ruck Thawonmas 2015. Predicting the opponent's action using the k-nearest neighbor algorithm and a substring tree structure. In: *Proc. of the 4th Global Conference on Consumer Electronics*, 533-534.
- [Rhalibi *et al.* 2009] Abdennour El Rhalibi, Kok Wai Wong and Marc Price 2009. Artificial intelligence for computer games. *International Journal of Computer Games Technology*, Article ID 251652, 3 pages.

- [Szummer and Picard 1998] Martin Szummer and Rosalind Picard 1998. Indoor-outdoor image classification. In: *Proc. of the 1998 IEEE International Workshop on Content-Based Access of Image and Video Database*, 42-51.
- [Tretyakov 2004] Konstantin Tretyakov 2004. Machine learning techniques in spam filtering. In: *Proc. of Data Mining Problem-oriented Seminar, MTAT 3*, 177, 60-79.
- [Vajda and Szocs 2014] Szilard Vajda and Barna Szocs 2014. A neural network based distance function for the k-nearest neighbor classifier. In: *Proc. of the 14th International Conference on Frontiers in Handwriting Recognition*, 429-433.
- [Yamamoto *et al.* 2014a] Kaito Yamamoto, Syunsuke Mizuno, Chun Yin Chu and Ruck Thawonmas 2014. Deduction of fighting-game countermeasures using the k-nearest neighbor algorithm and a game simulator. In: *Proc. of the 2014 Conference on Computational Intelligence and Games*, 437-441.
- [Yamamoto *et al.* 2014b] Kaito Yamamoto, Syunsuke Mizuno, Chun Yin Chu and Ruck Thawonmas 2014. MizunoAI source code. <http://www.ice.ci.ritsumei.ac.jp/~ftgaic/index-2h.html>. Checked 6.11.2016.
- [Yang and Pedersen 1997] Yiming Yang and Jan Pedersen 1997. A Comparative study on feature selection in text categorization. In: *Proc. of the 14th International Conference on Machine Learning (ICML'97)*, 412-420.

Big datan visualisoinnin ongelmat ja mahdollisuudet

Aleksi Sillanmäki

Tiivistelmä.

Tämän tutkielman pääaihe on informaation visualisointi. Tutkielmassa keskitytään visualisoinnin muistettavuuteen ja big datan visualisointiin. Tutkielma tuo esiin visualisoinnin ongelmia yleisesti sekä tarkemmin big dataa käsitellessä. Tutkielmassa annetaan myös ratkaisuja havaittuihin ongelmiin.

Avainsanat ja -sanonnat: Informaation visualisointi, big data, muistaminen, data

1. Johdanto

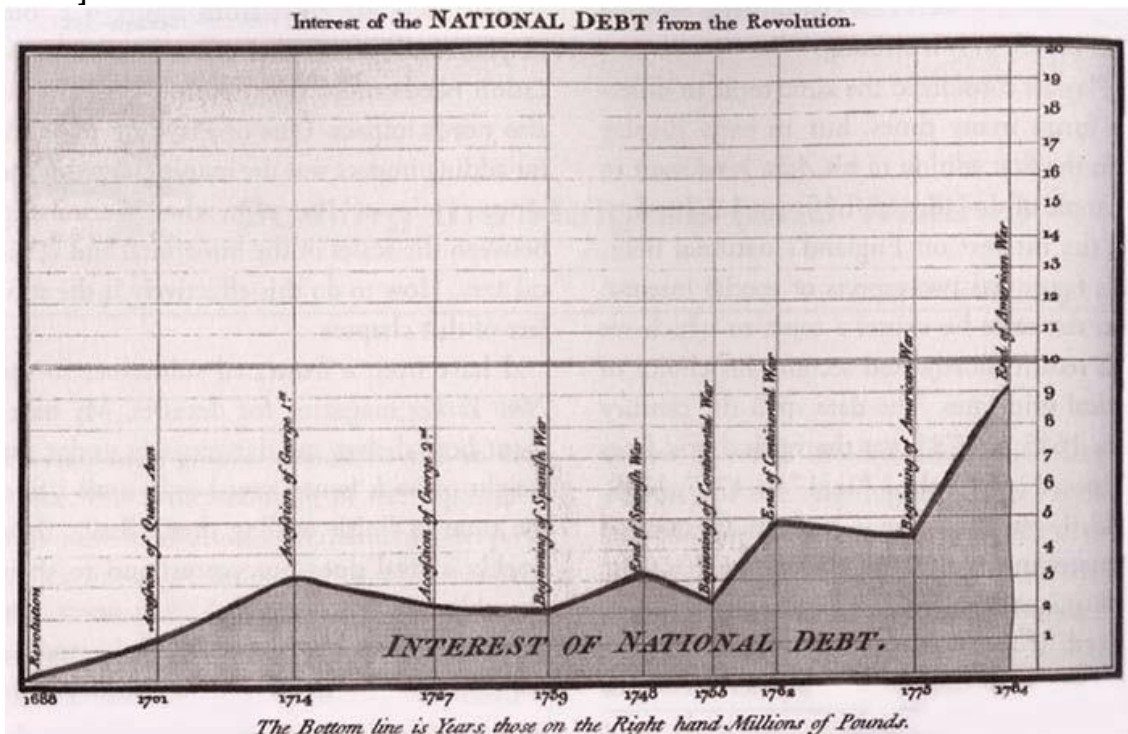
Tutkielmassa perehdytään olemassa olevan kirjallisuuden avulla hyvän visualisoinnin ominaisuuksiin ja mahdollisiin ongelmiin sekä niiden ratkaisuihin. Tämän tutkielman kolmannessa luvussa määritellään tiedon eri tasot ja tasojen termit ennen datan visualisointiin siirtymistä. Tutkielmassa käytetään vierasperäisiä ja vakiintuneita termejä data sekä big data. Data-sanaa käytetään, jotta tiedon eri tasojen hahmottaminen ja big data -termiin yhdistäminen olisi helpompaa. Big datasta on olemassa myös suomenkielisiä termejä kuten massadata ja iso data, mutta big data on vakiintunut ja yleisesti käytössä oleva termi.

Datan visualisointia tutkielmassa käsitellään tarkemmin kuvioroinan ja tyyllillisten määritelmien kautta kuin hyvien visualisointikaavioiden kautta. Eri kaavioita ovat esimerkiksi viiva- ja pylväsdiagrammi. Datan visualisointi on ongelmallinen aihe sen monimuotoisuuden vuoksi. Onkin tärkeää määritellä hyvän visualisoinnin ohjeita sekä ratkaista visualisoinnin ongelmia. Alussa määritellään hieman hyvän visualisoinnin ominaisuuksia ja perehdytään tarkemmin visualisoinnin muistettavuuteen. Tutkielman loppupuolella käsitellään visualisointia big datan apuna, sen ongelmia big dataa käytettäessä ja ratkaisuja ongelmiin.

Kuvioroina on hyvin kiistelty aihe datan visualisoinnissa. Kuvioroinasta on usein hyötyä visualisoinnissa, mutta sen käyttö ei ole helppoa. On suuri mahdollisuus epäonnistua kuvioroinan käytössä ja onkin tärkeää käyttää harkintaa kuvioroinan lisäämisessä. Big datan visualisoinnissa suurimpia ongelmia tuo big datan suuruus. Suurten datamäärien kanssa ongelmaksi muodostuu myös esitettävän datan suuri tilantarve näytöllä. Ratkaisuna voi olla vuorovaikutteinen visualisointi, mutta nykyisten tietokoneiden laskentateho ei ole riittävän nopeaa visualisoinnin reaaliaikaiselle vuorovaikutukselle. Laskentatehoa pystytään nopeuttamaan erilaisilla keinoilla, mutta lopullisia ratkaisuja ei ole vielä löydetty.

2. Lyhyesti visualisoinnin historiasta

Datasta on tehty taulukoita jo ennen ajanlaskumme alkua, mutta kvantitatiivisen datan visualisoinnin ajatus syntyi vasta 1600-luvulla Rene Descartesin kehittäessä kaksisuuntaisen koordinaatiston. William Playfairia voidaan pitää varhaisen pylväs- sekä piirakkadiagrammien keksijänä. Monet nykyaikanakin käytetyt erilaiset diagrammit ja graafit perustuvat Playfairin 1700-luvulla kehittämiin kaaviokuviin, joista yksi esimerkki on kuvassa 1 oleva viivadiagrammi. [Few 2013.]



Kuva 1. William Playfairin visualisointi vuodelta 1786 [Few 2013].

Tufte [2007] pitää datan visualisointia yllättävän uutena tieteen tutkimuskohdeena siihen nähden, että sitä käytetään nykyään laajasti niin sanomalehdissä kuin internetissä ja esimerkiksi opetusmateriaaleissa. Vastoin Tuften [2007] ajatuksia on datan visualisointi nykymuodossaan kuitenkin yllättävän vanha, noin 200–300 vuotta. Lascaux'in luolamaalauksia voidaan kuitenkin pitää eräänlaisina visualisointeina ja ne ovat noin 17 000 vuotta vanhoja. Vanhahtavan vaikutuksen tuo datan visualisoinnin muodostamisen helppous nykyaikaisilla tietotekniikan välineillä.

Hyvin suunniteltu datan visualisointi on usein paras ja tehokkain tapa selittää suuriakin tietokokonaisuuksia. Hyvä visualisointi helpottaa myös lukijan kykyä ymmärtää ja oppia haluttu tieto esitetystä datasta. [Tufte 2007.]

3. Mitä tieto on?

Ennen kuin paneudutaan syvemmin datan visualisointiin, on hyvä määritellä itse data ja muut tiedon tasot. Data on lajittelematonta informaatiota, jota ei pystytä ymmärtämään. Informaation hierarkia (DIKW) määrittää tiedon neljä tasoa: *data* (data), *informaatio* (information), *tietämys* (knowledge) ja *ymmärrys* (wisdom). Näiden alin taso data on merkitykseltään ja arvoltaan pienin ja ylin taso ymmärrys suurin. Informaatio on dataa, joka on käsitelty ihmisen ymmärtämään muotoon, ja se on merkityksellistä tiedon vastaanottajalle. Tietämys on usein monimutkaisempi ja moniselitteisempi kuin data tai informaatio. Tietämys voidaan kuitenkin nähdä eräänlaisena datan ja informaation yhdistelmänä, jota voidaan käyttää päätösten teon apuna. Tietämykseen tarvitaan informaation lisäksi kokemus ja taidot. Tietämys vastaa miten-kysymykseen. Hierarkian huipulla oleva ymmärrys on tiedon tasoista abstraktein ja vastaa miksi-kysymykseen. [Rowley 2007.]

Informaation visualisoinnin avulla voidaan nousta hierarkiassa ylöspäin. Datasta voidaan luoda informaatiota visualisoimalla se. Abstraktista datasta muodostetaan ihmisen ymmärtämä näkymä eli *visualisointi*.

4. Visualisoinnin peruskäsitteitä

Edward Tuftea voidaan pitää yhtenä nykyajan merkkihenkilönä informaation visualisoinnin alalla. Hänen kirjaansa *The Visual Display of Quantitative Information* on viitattu lähes 10 000 kertaa joulukuussa 2016 Google Scholarilla tehdyn haun mukaan. Kirja sisältää paljon esimerkkejä erilaisista eri informaatioista luoduista graafisista esityksistä. Tuften [2007] mukaan hyvä graafinen esitys ei häiritse varsinaisen datan lukemista tai ymmärtämistä. Grafiikan on hyvä tukea visualisoitavaa dataa esimerkiksi auttamalla vertailemaan kahta eri tietokokonaisuutta ja paljastamaan sekä pieniä yksityiskohtia että antamaan hyvän kokonaiskuvan esitettävästä datasta. Huono grafiikka taas saattaa johtaa lukijaa harhaan ja väärin tulkintoihin.

Datan visualisointi onkin tehty parhaiten silloin, kun sen avulla saadaan luotua raa'asta datasta informaatiota ja visualisoinnilla saadaan vastaanottajalle parhaimmassa tapauksessa jopa eräänlainen ymmärrys visualisoidusta datasta.

Suurimman osan ihmisen havainnointikyvystä vie näköaisti, ja sen vuoksi visualisointi on erittäin hyvä keino välittää lukijalle tietoa. Tämä ei kuitenkaan tarkoita, että kaikki graafiset esitykset olisivat keskenään yhtä hyviä [Simon 2014].

Informaation visualisoinnin tärkeys tulee hyvin ilmi siitä, kuinka paljon ihmiset lukevat verkkosivuilta. Lukemisen prosentuaalinen määrä annetusta tekstistä vähenee jyrkästi tekstimäärän kasvaessa, ja puolet tekstistä luetaan vasta,

kun sitä on 111 sanaa tai vähemmän [Nielsen 2008]. Informaation muuttaminen tekstimuotoisesta visuaalisempaan muotoon on siis kannattavaa tiedon välittämisen ja ymmärtämisen kannalta.

4.1. Infografiikat visualisoinnin muotona

Infografiikka (infographics) eli visualisoitu data tyylliteltynä on tullut yhdeksi suosituimmaksi tavaksi esittää dataa suurille yleisöille. Infografiikkaa antaa yleiskuvan esiteltyyn aiheeseen datan avulla. Infografiikan tarinallisen kerronnan avulla voidaan kiinnittää yleisön huomio. [Chang *et al.* 2015.]

Infografiikat eroavat siis datan visualisoinnista siten, että infografiikoita luodessa otetaan huomioon grafiikan ulkoasu. Infografiikoita tutkimalla voidaan kuitenkin selvittää myös yleisesti visualisointien laadullisia ominaisuuksia.

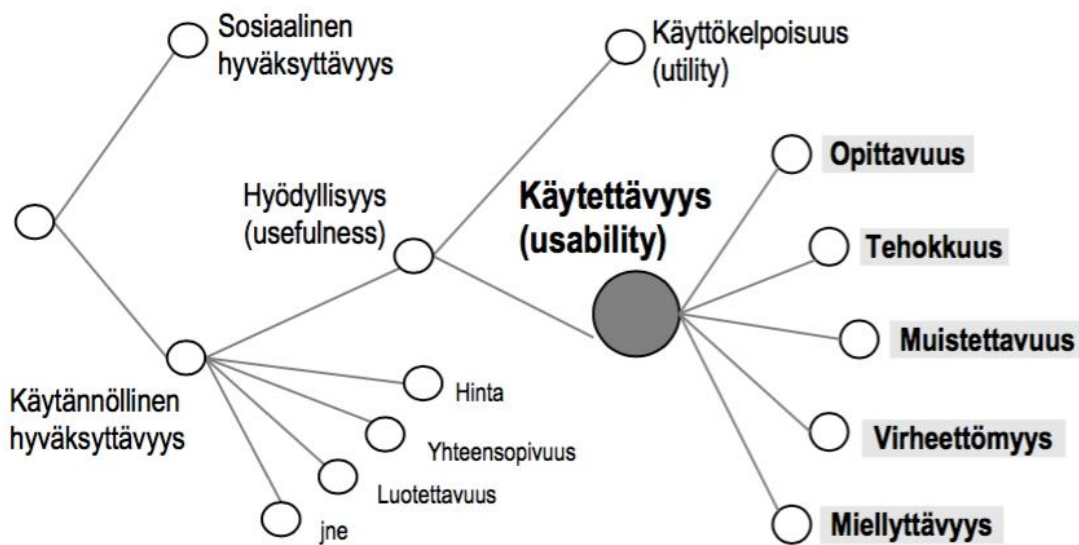
Yhtenä yllättävänä tutkimustuloksena on sukupuolen, iän ja koulutustason vaikutus infografiikoiden mielekkyyden kokemiseen. Ihmiset muodostavat infografiikoiden mielekkyydestä mielipiteensä jo ensimmäisen 500 millisekunnin aikana grafiikan näkemisestä. Yleisesti ihmiset suosivat infografiikoissa enemmän värikkäämpiä ja yksinkertaisempia kaavioita. Ihmisen muisti onkin rajoittunut muistamaan vain noin 3–5 erilaista objektiota ja kuviota [Ware 2004]. Miehet suosivat värittömämpiä ja monimutkaisempia kaavioita naisia enemmän. Miehiin ei myöskään vaikuta suuri tekstimäärä niin paljon kuin naisiin. Naiset kokevat kuvien katselun mielekkäämmäksi suurien tekstimäärien sijaan. Suunnittelijoiden kannattaa keskittyä hyvään ensivaikutelmaan esimerkiksi lisäämällä kontrastia ja värikylläisyyttä eri grafiikan värien välillä. Yleisesti yksinkertaisemmat eli vähemmän turhia kuvia ja tekstiä sisältävät kaaviot ovat ihmisille miellyttävämpiä. [Chang *et al.* 2015.]

Changin ja muiden [2015] esiin tuomat ihmisten mieltymykset ja niistä luodut ohjenuorat infografiikoiden osalta näen hyvin käytettäväksi kaikessa datan visualisoinnissa, ei pelkästään infografiikoissa. Visualisointeja luodessa on otettava huomioon monet eri käyttäjät. Onnistuneet visualisoinnit ottavat huomioon niin miehet ja naiset kuin myös värisokeat ja muut esteelliset käyttäjät.

4.2. Visualisoinnin laadullisia määritelmiä

Mielestäni yksi hyvä ohje laadukkaisiin visuaalisiin tuloksiin on Nielsenin [1993] määrittelemä käytettävyyden osatekijöiden malli, joka on kuvassa 2. Malli ei ole tarkoitettu suoraan visualisoinnin määrittelyyn, mutta siinä olevat osatekijät on hyvä ottaa huomioon visualisointeja käsitellessä. Käytettävyys määritellään lopujen lopuksi käyttäjän kokemana kokemuksena [Aula *et al.* 2005]. Kokemuksen yksilöllisyys tuo haasteita laadukkaiden visualisointien tekemiseen. Visuali-

sointi ei saa väärentää datan tietoja [Tuftte 2007]. Jokaiselle käyttäjälle ei samanlainen visualisointi toimi yhtä hyvin. Huomioon täytyy ottaa, kenelle kyseinen datan visualisointi on suunnattu. Kokenut käyttäjä pystyy lukemaan vaikeasti tulkittavia visualisointeja kokemattomia helpommin. Vaikeammin tulkittavat visualisoinnit voivat usein sisältää yksinkertaisempia enemmän dataa. Hyvä käytettävyys on nykyisin perusvaatimus oikeastaan joka osa-alueella. [Simon 2014.]



Kuva 2. Malli käytettävyyden osatekijöistä [Nielsen 1993] suomennos [Aula *et al.* 2005].

Datamuste (data-ink) on grafiikan osa, jota ei voida poistaa menettämättä sen sisältöä. Muu grafiikan muste on se grafiikan osa, joka ei välitä tietoa käytetystä datasta, esimerkiksi ylimääräiset koordinaatiston viivat. Näiden välistä suhdetta kutsutaan *datasuhteeksi*. [Tuftte 2007].

Tuften [2007] mukaan *datasuhde* (data-ink ratio) kannattaa pyrkiä maksimoimaan ja jokaisella mustepisaralla tai pikselillä pitäisi olla syy olemassaoloon. Jokaisen mustepisaran pitäisi kertoa uutta tietoa datasta. Toinen vaihtoehto parempaan visualisointiin on pyrkiä vähentämään ylimääräistä datamustetta sekä poistamaan sitä grafiikan mustetta, jolla ei ole tekemistä varsinaisen datan kanssa.

Vastoin Tuften [2007] ohjeita Bateman ja muut [2010] eivät löytäneet mitään eroa ihmisten tulkinnassa vertaillen turhalla datamusteella koristeltuja ja minimalistisia kaavioita keskenään. Ihmiset jopa muistivat kaavioiden sisällön pitkällä aikavälillä vahvan kuvituksen avulla paremmin. Tutkimuksessa kävi ilmi, että ihmiset käyttivät enemmän aikaa koristeellisen kuvituksen katseluun, mikä saattaa hyvinkin olla syynä kuvioiden sisällön parempaan muistamiseen. Kuvat johtavatkin parempaan tulokseen sekä tunnistamisessa että mieleen palauttamisessa.

Kuvien ja visuaalisten elementtien avulla voidaan kiinnittää paremmin lukijan huomio visualisointiin. Käytettäessä kuvia ja visuaalisia elementtejä visualisoinneissa lukijan muistaminen saattaa parantua ja lukijan huomio voidaan ohjata tärkeimpiin osa-alueisiin.

Tietokoneiden yleistyessä 1900-luvulla *kuvioroina* (chartjunk) lisääntyi tietokoneohjelmien tuodessa valmiita työkaluja kaavioiden tekemiseen. Kuvioroinalla tarkoitetaan juuri sitä visualisoinnin osaa, jolla ei ole varsinaista tekemistä datan kanssa. Kuvioroina on siis datan kannalta turhaa datamustetta. Tietokonegrafiikkaa käytettäessä on kuitenkin suunnittelijan vastuulla, selkeyttääkö kyseisten grafiikoiden käyttö kaaviota vai ei. Esimerkiksi kaavion taustalla oleva koordinaatisto ei saisi häiritä itse datan lukemista vaan pysytellä taustalla, esimerkiksi mustavalkoisessa kuvassa himmeämpänä harmaana datan ollessa mustalla. [Tuft 2007.]

Batemanin ja muiden [2010] lisäksi Inbar ja muut [2007] selvittivät, pitävätkö ihmiset enemmän minimalistisista vai koristeellisuutta sisältävistä visualisoinneista. Heidän tuloksissaan ihmiset pitivät selvästi enemmän koristeellisesta kaaviosta kuin Tuften [2007] ohjeiden mukaisesta kaaviosta.

Voidaan siis päätellä, että äärimmäinen kuvioroinan poistaminenkaan ei ole hyväksi, vaan tarvitaan lukijan huomion herättämistä ja muistettavuuden tukeamista erilaisilla apukeinoilla kuten väreillä ja kuvioilla. Kaavioissa kuitenkin koristeellisuuden täytyy tukea datan luettavuutta ja pyrkiä lisäämään ymmärrystä.

4.3. Visualisoinnin tyyli

Aikaisemminkin mainitut Bateman ja muut [2010] selvittivät ihmisten mieltymyksiä visualisoinneista. Ihmiset pitivät miellyttävimpinä koristeellisia kaavioita. Erittäin minimaaliset visualisoinnit eivät olleet ihmisten mielestä mielenkiintoisia. Visuaalisesti koristeelliset kaaviot olivat myös helpoimmat ja nopeimmat muistaa sekä kuvailta. Lyhyellä aikavälillä tai kaavioiden lukuhetkellä ei ollut suuria eroja pelkistettyjen ja koristeellisten kaavioiden välisissä tulkinnoissa. Myöhemmin, noin 2–3 viikon ajan päästä, koristeellisten kaavioiden sisältö muistettiin huomattavasti pelkistettyjä kaavioita paremmin. Tulokset vahvistavat omaa näkemystäni visualisoinnin tyylin merkityksestä. Tyylillä voidaan vahvistaa lukijan oppimista sekä esitetyn datan muistamista. Visualisoinneissa täytyy näkyä sen eri variaatiot eikä niinkään tyylin eri variaatiot [Tuft 2007].

Miksi Batemanin ja muiden [2010] tutkimus on osittain Tuften [2007] määrittämiä vastaan ja kuvioroina saattaakin olla hyväksi? Yksi syy saattaa olla, että ihmiset katsovat kaavioiden dataa yhtä kauan huolimatta kuvioroinan määrästä. On kuitenkin tärkeää huomata, etteivät Bateman ja muut [2010] rajoittaneet kaavioiden katseluaikaa. Saattaakin olla, että ihmiset olisivat käyttäneet vähemmän

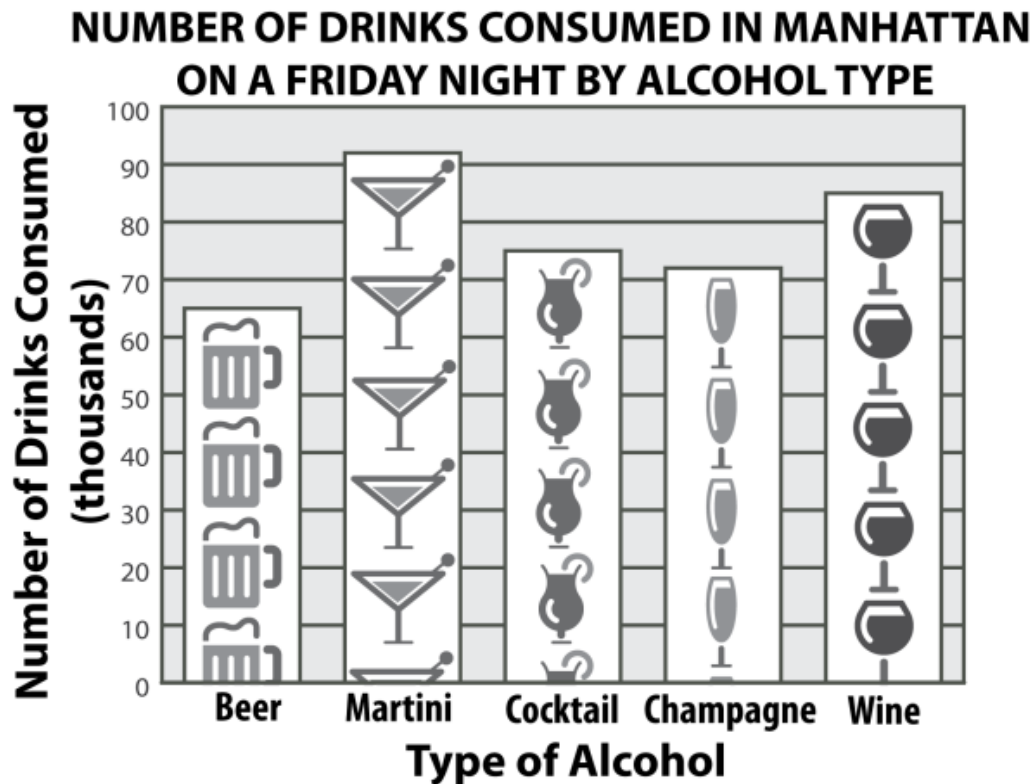
aikaa datan katsomiseen koristeellisissa visualisoinneissa, jos aika olisi ollut rajattu. Bateman ja muut [2010] myös käyttivät kaavioita, joissa kuvioroina liittyi tiukasti kuviossa esitettyyn dataan. Heidän käyttämässään kaavioissa ei siis ollut lukijaa häiritsevää grafiikkaa, vaikka se ei suoraan dataa esittänytkään. Tärkeäksi kriteeriksi nouseekin grafiikan ja datan yhteys. Grafiikan täytyy tukea käyttäjää ja olla yhteydessä dataan, jotta se ei häiritse lukijaa vaan on hyödyksi.

Kuvat itsessään aktivoivat muistin eri osa-alueita tekstin sijaan eivätkä täten vähennä luetun tekstin muistettavuutta vaan tuovat lisääpua muistamiseen. Koristeellisuuden avulla tuotettu kaavioiden erilaisuus saattaa herättää ihmisten tunteita. Koristeellisuuden ei kuitenkaan pidä aiheuttaa liian suuria tunteita. Voimakkaat koristeellisuudet voivat aiheuttaa järkytystä tai vastenmielisyyttä kuviota kohtaan [Bresciani and Eppler 2008]. Muistaminen ja mieleen palauttaminen vain vahvistuvat eri ärsykkeistä syntyneistä muistoista. Mikäli koristeellisuus ja ylimääräiset visuaaliset elementit kaaviossa pystytään liittämään esitettyyn dataan, parantuu ihmisten muistaminen huomattavasti. [Bateman *et al.* 2010.]

Pidän yhtenä tärkeänä kaavion tyyliohjeena sitä, että esitetty kuvioroina on tiukasti sidoksissa kaavion dataan. Kuvioroinalla voidaan myös ohjata lukijaa tiettyyn ajattelutapaan tai mielipiteeseen annetusta datasta. Kuvioden suunnittelijoilla onkin vastuu siitä, miten data koetaan. Kuvioroinan herättämät tunteet ja visuaaliset esitystavat auttavat palauttamaan kuvion sisältöä mieleen.

Kuvilla voidaan herättää erilaisia mielikuvia kyseisestä datasta muistamisen avuksi. Kuvassa 3 on esimerkki kuvioroinasta, joka ei vaikuta kuvion tulkintaan eikä herätä erityistä mielikuvaa tai ohjaa lukijaa. [Bateman *et al.* 2010.]

Yksinkertaisena esimerkkinä mielikuvan herättämisestä voisi olla elefantin käyttö suurta asiaa kuvatessa, se auttaa muistamaan asian olleen suurta. Vastakohtaisesti esimerkiksi hiiri koetaan pieneksi, ja se saattaa herättää mielikuvia pienestä asiasta.



Kuva 3. Kuvioroinaa, joka ei herätä mielikuvia, mutta auttaa esimerkiksi muistamisessa [Bateman *et al.* 2010].

Datalle hyvän ja merkityksellisen tyylin löytäminen ei kuitenkaan ole helppo tehtävä edes kokeneelta suunnittelijalta. Turvallisuutta vaativissa tehtävissä kuten lentokoneissa suosituksena on minimaalinen tyyli siellä esitetyissä kaavioissa, jotta vältetään kriittiset tulkintavirheet. Täytyy kuitenkin muistaa, että myös Tuften [2007] suosimilla minimalistisilla kaavioilla pystytään tuomaan omaa mielipidettä tai näkökantaa esiin datan ja otsikoiden valinnalla, korostamalla datan tiettyä osaa tai järjestää osat tiettyyn järjestykseen. Minimaaliset kuviot eivät siis välttämättä ole sen objektiivisempia kuin paljon kuvioroinaa sisältävät. Koristeelliset kuviot saattavat kuitenkin antaa enemmän arvoa lukijalle. [Bateman *et al.* 2010.]

Few [2011] kirjoitti kattavan kritiikin Batemanin ja muiden [2010] metodologiasta. Hän kritisoi esimerkiksi tutkimuksessa käytettyjä kaavioita, jotka olivat äärimmäisiä esimerkkejä kuvioroinasta. Toinen viidestä metodologisesta ongelmasta oli heidän käyttämänsä 20 yliopisto-opiskelijaa, jotka Few [2011] näki ongelmallisiksi luotettavien tulosten saamiseksi. Hän kritisoi myös kaavioiden yksinkertaisuutta ja näki ongelmaksi, että koristeelliset kaaviot oli tehnyt taitava graafinen taiteilija. Tämän onkin yksi mielenkiintoinen haaste hyvän kuvioroinan tekemisessä. Hyvä kuvioroina vaatii taitavan tekijän. Few [2011] antaa esimerkin erittäin huonosta visualisoinnista ja vertaa sitä minimalistiseen visuali-

sointiin samasta datasta. En kuitenkaan näe tämän olevan verrattavissa Batemanin ja muiden [2010] tuloksiin. Heidän käyttämänsä kaaviot ovat äärimmäisiä esimerkkejä kuvioroinan käytöstä, mutta ne eivät ole huonosti tehtyjä. Huonosti tehty kuvioroina saattaa pilata minimalistisen kaavion. On siis ehkä suurempi riski käyttää kuvioroinaa kuin jättää se käyttämättä. Kuvioroinalla voidaan Batemanin ja muiden [2010] tutkimuksenkin mukaan parantaa esimerkiksi muistettavuutta.

4.4. Visualisoinnin muistettavuus

Few [2011] määrittelee kuvioroinan hieman tarkemmin kuin Tufte [2007]. Kuvioroina ei kuitenkaan ole sellaista grafiikka, joka antaa tietoa datasta ja olisi näin hyödyksi lukijalle. Graafiset elementit voivat kuitenkin olla hyödyksi kolmella eri tavalla. Grafiikalla voidaan saada ihmiset lukemaan sisältö, korostaa tiettyjä kaavion kohtia ja lisätä muistettavuutta. Koristeellisuus kuitenkin vain lisää kaavion tehokkuutta eikä tuo lukijalle uutta tietoa.

Borkin ja muut [2013] tutkivat visualisointien muistettavuutta 2070 erilaisella visualisoinnilla 261 osallistujan avulla. Visualisoinnit, jotka sisälsivät kuvia tai muita kuvioita, auttoivat muistamaan kuvion sisällön paremmin. Testihenkilöt muistivat parhaiten visualisoinnit, joissa oli ihmiselle tuttuja ja tunnistettavissa olevia kuvia. Kuvioiden värien määrällä oli myös merkitystä muistettavuuteen. Mitä enemmän värejä visualisoinneissa oli, sitä paremmin ne paransivat muistettavuutta. Mustavalkoisissa visualisoinneissa muistettavuus on heikointa.

Kun vastoin Tuften [2007] ja Fewn [2013] suosituksia datasuhde pidettiin suurena ja suosittiin minimalistisia visualisointeja, niin testitulokset antoivat ainakin muistettavuuden osalta viitteitä, että värikkäät ja kuvioroinaa sisältävät visualisoinnit ovat parempia. Borkin ja muut [2013] kuitenkin tutkivat visualisointeja kuvina, eivätkä datan visualisointeina. He eivät siis tutkineet, kuinka hyvin osallistajat ymmärsivät visualisoinnin tai kuinka sitoutuneita he olivat.

Vastoin Borkinin ja muiden [2013] hypoteeseja uniikit ja harvemmin käytetyt visualisoinnit, kuten puudiagrammit ja koordinaatistot, olivat paremmin muistettavia kuin perinteisemmät pylväs- ja viivadiagrammit. He toteavat, etteivät pylväs- ja viivadiagrammit ole niin luonnollisia kuin esimerkiksi tiheyskartta ja puudiagrammi. Nämä harvinaisemmin käytetyt diagrammit ovat lähempänä ihmisen käsittämää luonnollisuutta datassa ja näin saattavat olla syynä niiden parempaan muistamiseen. Mielenkiintoinen yksittäinen havainto oli myös se, että esimerkiksi pyöreitä reunoja pidetään luonnollisempina ja pyöreitä kuvioita sisältävät visualisoinnit saivat korkeammat muistamisen arviot.

Infografiikoihin luokitellut visualisoinnit olivat kaikkein muistettavimpia. Grafiikoiden luokittelu oli tehty infografiikoihin, tieteellisiin julkaisuihin, uutismediaan sekä joidenkin julkisten tahojen laatimiin visualisointeihin. Grafiikoiden muistettavuus oli tässä järjestyksessä, infografiikoiden ollessa muistettavimpia. Muistettavuus on kuitenkin vasta yksi osa hyvää visualisointia. Tärkeää olisi saada ihmisille jäämään muistiin visualisoinnin sisältö eikä pelkästään visualisointi itsessään. [Borkin *et al.* 2013.]

Uniikkien visualisointien paremmuus on yksi syy hyvää muistettavuutta tukevien visualisointien ohjeiden määrittelyn vaikeuteen. Hyvän tyylin määritelmät unohtuvat toisinaan suunnittelijoita heidän käyttäessään uusia teknologioita [Tufta 2007]. Uudet teknologiat kyllä mahdollistavat erilaisia lähestymistapoja datan esittämiseen, mutta suunnittelijoiden täytyy pitää mielessä myös visualisointien lukijat. Esimerkiksi liiallinen 3D-grafiikan käyttö usein häiritsee lukijaa. Visualisointi tukee luettavan materiaalin muistamista, sillä visuaalinen muisti on erillään kielellisestä muistista [Ware 2004]. Muistettavuuden näenkin yhtenä tärkeänä kriteerinä visualisointeja käytettäessä, varsinkin kun niitä käytetään opetustarkoituksessa.

Kuvioroina ja epäoleellinen datamuste ovat haitallisempia suuria datamääriä sisältävissä visualisoinneissa kuin vähän informaatiota sisältävissä [Tufta 2007]. Erityistä varovaisuutta on syytä noudattaa suuria informaatiomääriä sisältävien visualisointien teossa. Suuret datamäärät tuovat uusia mahdollisuuksia visualisoinnin parissa.

5. Big data

Tietoa kerätään nykypäivänä valtavia määriä älypuhelimista, sosiaalisesta mediasta ja uusista laitteista, eikä yrityksillä ole välttämättä tarvittavia resursseja tietomäärän käsittelyyn [Simon 2014; EMC Education Services 2015]. Suurta määrää monimutkaista, moniulotteista ja järjestämätöntä dataa kutsutaan *big dataksi* [Brauner 2015]. Big datan monimutkaisuus ja moniulotteisuus tulee monista eri lähteistä saadusta datasta, esimerkiksi eri laitteista tai verkkosivuista [EMC Education Services 2015]. Big datalle ei kuitenkaan ole olemassa yhtä ja tiettyä määritelmää. Big data usein määritellään vähintään kolmen V:n avulla: volyyymi, vauhti ja vaihtelevuus (volume, velocity, variety) [EMC Education Services 2015].

Perinteiset datankäsittelymenetelmät eivät riitä big datan luomien ongelmien ratkaisuun, ja yritysten tarvitseekin luoda uusia datankäsittelymenetelmiä. Big data antaa yrityksille uutta dataa lähes kaikesta. Uuden datan avulla saadaan uusia mahdollisuuksia yrityksille lisäarvon luomiseen ja kilpailukykyyn paranta-

miseen. Esimerkiksi Facebookissa vuonna 2012 julkaistiin 700 tilapäivitystä sekunnissa. Tilapäivityksistä Facebook saa tietoa käyttäjän tekemisistä ja pystyy kohdentamaan asiakkaalle mainontaa. Esimerkiksi naisen vaihtaessa suhdetilansa sinkusta kihlatuksi, aktivoituu Facebookin yhteistyökumppaneiden mainokset häämekoista tai hääsuunnittelusta. [EMC Education Services 2015.]

Big dataa ei kuitenkaan saada hyötykäyttöön niin hyvin kuin voisi olla mahdollista. Tiedonhallinnon ammattilaisista vain 12 prosenttia käsitteli big dataa vuonna 2012. Tiedon visualisointi on big datan vuoksi entistäkin tärkeämpää suuren tietomäärän vuoksi. Vuoteen 2015 mennessä internetiin luotu tietomäärä oli lähes 8 tsettabittiä, joka vastaa noin 800 miljardin kannettavan tietokoneen kovalevytilaa. Tästä tietomäärästä suuri osa on järjestämätöntä ja käsittelemätöntä dataa. Datan määrä kasvaa kiihtyvää tahtia. Vuonna 2013 datasta 90 prosenttia oli luotu edellisten kahden vuoden aikana. Vuonna 2012 Google vastaanotti yli 2 miljoonaa hakua, lähes 600 uutta nettisivua luotiin joka minuutti ja jokaisesta näistä syntyy uutta dataa. [Simon 2014.]

Esineiden internet (internet of things) luo lisää suuria määriä dataa ja vaatii ihmisten vuorovaikutusta koneiden kanssa. Data monimutkaisista teollisuuslaitteista tai kotikäyttöön tarkoitetuista laitteista saattaa olla hyvin monimutkaista [Brauner 2015].

Data on kasvavassa määrin yhteydessä toiseen internetissä olevaan dataan, vaikei läheskään kaikki data ole vieläkään yhteydessä toisiinsa. Esineiden internet tuo lisää dataa saataville, mutta kuinka sitä voidaan hyödyntää? Laitteiden välinen datan vaihto vaikeuttaa saadun datan ymmärrettävyyttä. Datan visualisointi on yksi keino analysoida ja hyödyntää kerättyä dataa. [Simon 2014.]

Hyvään vuorovaikutukseen koneen ja ihmisen välillä tarvitaan hyvää visualisointia.

5.1. Big datan visualisoinnin ongelmia

Tutkielman alkupuoliskolla mainitut Tufte [2007] sekä Simon [2014] tuovat ilmi visualisoinnin olevan yksinkertainen ja luonnollinen tapa ihmiselle käsitellä tietoa. Tämän vahvistavat myös Gorodov ja Gubarev [2013]. Kaikista perinteisemmät visualisointitavat big dataa käsitellessä eivät ole kaikkein tehokkaimpia tai sopivimpia tapoja konkreettisiin tehtäviin [Gorodov and Gubarev 2013]. Myöskään nykyiset visualisointityökalut eivät mahdollista laajaa datan käsittelyä ja reaaliaikaista vuorovaikutusta [Heer *et al.* 2012].

Myöskään Choon ja Parkin [2013] mukaan visualisointiin tarkoitettut uusimmatkaan työkalut eivät välttämättä toimi big datan kanssa halutusti. Ihmisen hyvä kyky hahmottaa asioita ja tehdä päätöksiä visuaalisen analyysin kautta on

lisännyt big datan visualisointiin kohdistunutta mielenkiintoa. Viimeisimmät visualisointityökalut tai tekniikat eivät valitettavasti toimi big datan kanssa niin hyvin kuin voisi. Syynä tähän Choon ja Parkin [2013] mukaan on kaksi pääongelmaa.

Ensimmäiseksi ihmisen havainnointikyky ei ole parhaimmillaan suurten tietomäärien kanssa ja oleellisin tieto datasta saattaa jäädä huomaamatta. Big datan suuren määrän vuoksi ihmiset eivät välttämättä saa poimittua sitä tietoa, jota datan esittäjä tai sen lukija haluaisi löytää. [Choo and Park 2013.]

Toinen ongelma on näytöt, joilla dataa pitäisi esittää. Näyttöjen koko on usein liian pieni valtaviin tietomassojen esittämiseksi, ja datan visualisoinnista saattaa tulla sotkuista ja epäselvää. Big datan ahtaminen verrattain pieniin näyttöihin ei ole helppo tehtävä. Ihmisten hahmotuskyky ei ole parhaimmillaan, kun suuresta datasta pienellä näytöllä pitäisi löytää oleellinen. [Choo and Park 2013.]

Heer ja muut [2013] mainitsevat myös kaksi pääongelmaa, joista ensimmäinen on hahmottamiskyky. Hahmottamiskyvyllä he tarkoittavat samaa kuin edellä mainittu Choon ja Parkin [2013] toinen pääongelma näyttöjen koosta. He mainitsevat ongelmaksi myös datan kanssa käydyn vuorovaikutuksen. Choon ja Parkin [2013] määrittelemä suuren tehon vaatimus vastaa tätä ongelmaa. Ongelmana on reaaliaikainen vuorovaikutus dataan, kun tietoa käsittelevät koneet eivät pysty reagoimaan käyttäjän nopeudella.

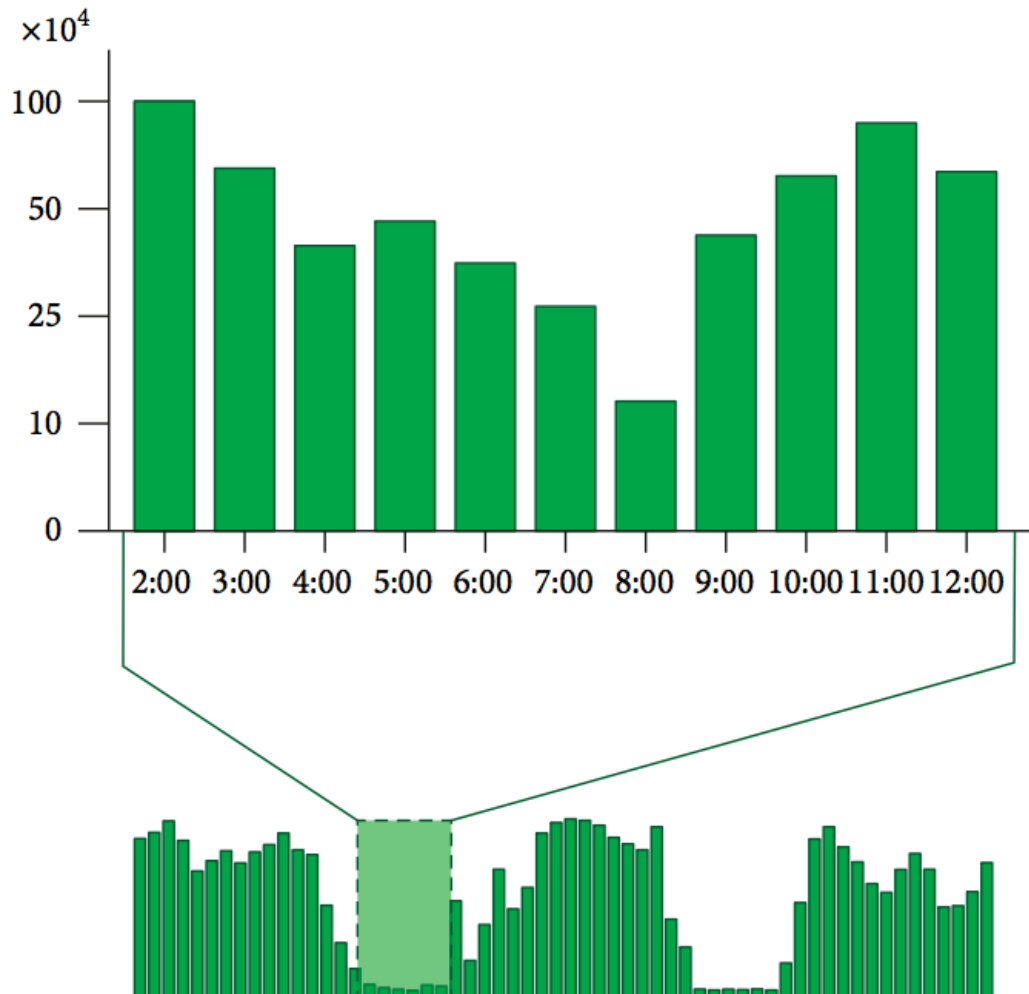
Gorodov ja Gubarev [2013] taasen mainitsevat viisi big datan visualisointiin liittyvää ongelmaa, joista ensimmäinen on *visuaalinen melu* (visual noise). Visuaalisella melulla tarkoitetaan sekaista esitystä datasta, joka johtuu aineiston sisältämien datan lähekkäisistä suhteista. Lukija ei välttämättä pysty erottamaan eri data-alkioita toisistaan. Toisena Gorodov ja Gubarev [2013] mainitsevat ongelman, jota he kutsuvat nimellä *suuren kuvan havainnointikyky* (large image perception). Tällä voitaisiin ratkaista heidän ensimmäiseksi mainitsema lähekkäisen datan ongelma, käyttämällä suurempaa ja erotuskyvyltään parempaa näyttöä. Ongelmaksi tulee Choon ja Parkin [2013] mainitsema ihmisen havainnointikyky. Lukija kohtaisi vaikeuksia hyödyllisen tiedon löytämisessä. Yllä olevia ongelmia voitaisiin ratkaista käyttämällä vain osaa datasta ja piilottamalla loput. Ongelmaksi syntyisi kuitenkin *informaation katoaminen* (information loss). Visualisointia lukevaa analyttikkoa jonkin datan puuttuminen saattaa harhauttaa, eikä analyttikko välttämättä löytäisi tarpeellista tietoa aineistosta. Suuret datamäärät vaativat tehokkaat tietokoneet sitä käsittelemään. Gorodov ja Gubarev [2013] nimittävät neljättä heidän viidestä ongelmastaan *suuren tehon vaatimukseksi* (high performance requirements). Viides ongelma on *nopea kuvan vaih-*

tuvuus (high rate of image change). Monitorointitehtävissä oleva henkilö ei välttämättä pysty reagoimaan datan määrän nopeaan vaihteluun tai sen tiiviyyteen näytöllä.

5.2. Ratkaisuja big datan visualisoinnin ongelmiin

Visualisoinnin metodit eivät yleensä synny tyhjästä. Metodeja kehitetään usein vanhojen metodien pohjalta uusiin käyttötarkoituksiin. Nykyiset metodit eivät toimi kovin hyvin big datan visualisoinnissa.

Useampi kuin yksi visuaalinen esitys datasta ei monesti ole riittävä. Analytiikat saattavat vertailla samasta datasta saatavia eri visualisointeja. Visualisointien olisi hyvä olla myös yhteyksissä toisiinsa. Esimerkiksi analytiikon valitessa yhden visualisoinnin alueen korostuu myös muiden visualisointien vastaava alue. Tämä mahdollistaa suodattimien käyttämättä jättämisen ja nopeuttaa analytiikon työtä. Suodattimet täytyy kuitenkin olla käytettävissä, jotta analytiikat voivat nähdä sekä koko visualisoidun datan että haluamansa osan siitä. Esimerkki ratkaisumallista on kuvassa 4. [Gorodov and Gubarev 2013.]



Kuva 4. Tarkennettu näkymä datan yleiskuvasta [Gorodov and Gubarev 2013].

Isoin syy suodattamisen tarpeellisuuteen on aikaisemminkin mainittu ihmisen havainnointikyvyn rajallisuus suurten tietomäärien tapauksessa. Suuria tietomääriä ei havainnoida parhaalla mahdollisella tavalla ja datan esittäminen pienissä määrissä helpottaakin tiedon omaksumista. Tällaisiin isommasta datasta tehtäviin lähentämisiin ja rajaamisiin tarvitaan datan uudelleen laskemista, jotta datasta saadaan tarvittavia tarkempia tietoja.

Big datan laskenta-aika on erittäin suurta big datan koon vuoksi. Laskennallisia metodeja parantamalla voidaan datasta antaa kompakti ja merkityksellinen tieto lukijalle. Ryhmittely ja erilaiset koneoppimisen sekä tiedon louhinnan metodit ovat luonnollisia big datan visualisoinnissa. [Choo and Park 2013.]

Ryhmittely on tärkeää, jotta analyttikkoa ei johdateta väärään suuntaan. Ryhmittelyllä voidaan myös ehkäistä datasta syntyviä väärinkäsityksiä. [Gorodov and Gubarev 2013.]

Reaaliaikaiseen visualisointiin eivät ulottuvuuksien vähentäminen, ryhmittely ja erilaiset koneoppimisen sekä tiedon louhinnan metodit kuitenkaan yksinään ole riittävän tehokkaita. Suuri datamäärä vaatii interaktiivisemmän visualisoinnin kuin pieni datamäärä [Simon 2014]. Tärkeintä laskennallisissa metodeissa olisi saada ne riittävän nopeiksi suurten datamäärien reaaliaikaisiin ja interaktiivisiin visualisointeihin. [Choo and Park 2013.]

Mikäli liiallista laskenta-aikaa voidaan nopeuttaa, saadaan lukijalle mahdollisuus reaaliaikaiseen vuorovaikutukseen datan kanssa ja laskennalliset metodit huomattavasti nopeammiksi. Ihmiset eivät useinkaan välitä liian tarkoista numeraalisista arvoista. Esimerkiksi monelle riittää tieto piin desimaaliarvosta kahden desimaalin tarkkuudella 3,14. Nykyiset kotitietokoneetkin mahdollistavat paljon useampien desimaalien laskemisen. Desimaaleja vähentämällä voidaan laskenta-aikaa nopeuttaa. Näytön resoluutiot eivät myöskään ole niin suuria, että dataa kannattaisi laskea monen desimaalin tarkkuudella. Laskettuja desimaaleja ei pystytä näytöllä esittämään, ellei visualisointia lähennetä. [Choo and Park 2013.]

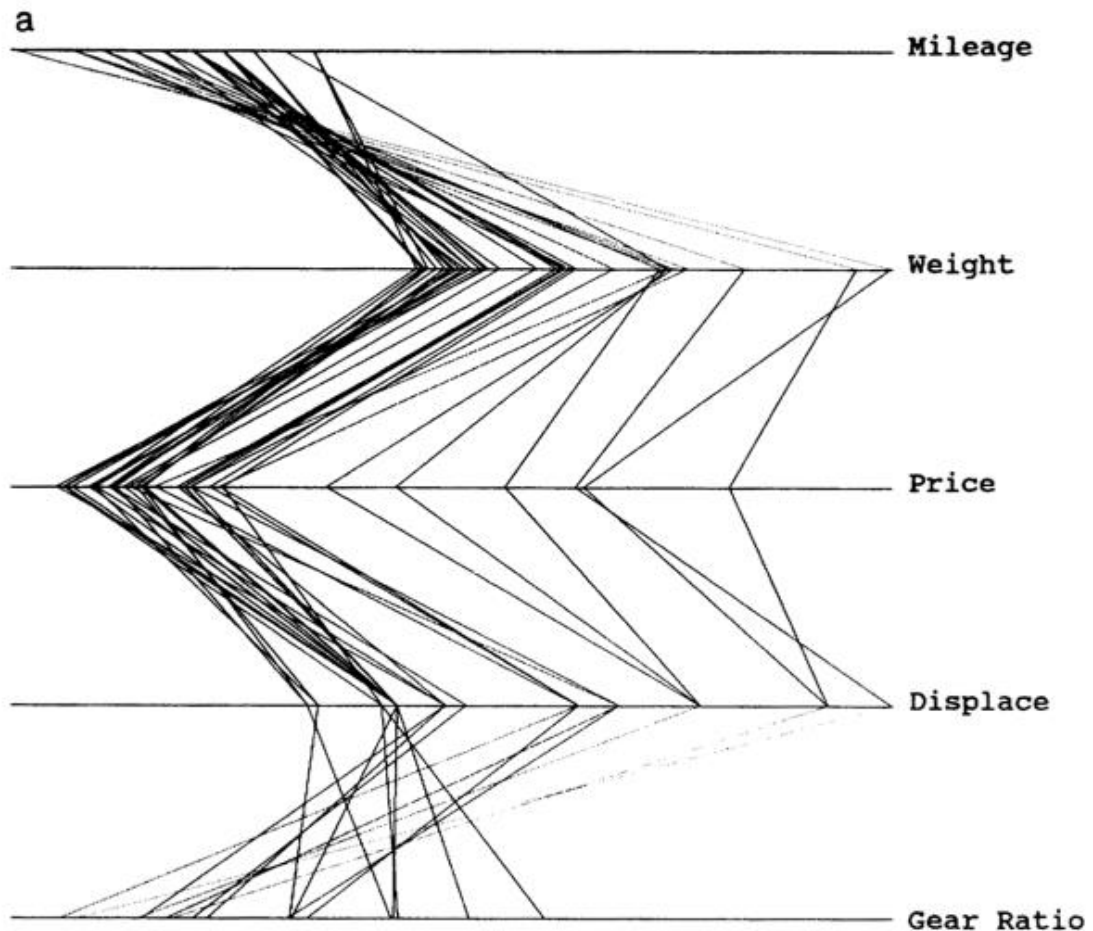
Choo ja Park [2013] vähensivät datan laskemisen tarkennuskertoja kahdesta yhteen. He eivät löytäneet merkittäviä eroja datasta muodostetusta hajontakuviosta. Mahdollista olisi esimerkiksi määrittää tietylle resoluutiolle vaadittu minimitarkkuus, jotta visualisointi olisi riittävän tarkka halutulla resoluutiolla. Iterointikertojen vähentäminen nopeuttaa merkittävästi pienitehoisten prosessorien visualisointia esimerkiksi älypuhelimissa.

Toki laskentaa täytyisi tehdä reaaliajassa käyttäjän ollessa vuorovaikutuksessa visualisointiin. Visualisoinnin lähentämisessä tarvitaan tarkempaa tietoa datasta ja sitä voidaan laskea iteroimalla data uudelleen lähennettäessä [Choo

and Park 2013]. Visualisointia voidaan tehdä reaaliajassa, kun käyttäjä lisää parametreja ilman, että koko visualisointi tehtäisiin uudelleen [Choo and Park 2013].

Gorodov ja Gubarev [2013] selvittivät parhaimpia visualisointimetoodeja big datalle. Esimerkiksi *ruutupu*-visualisointi (treemap) ei sovi hierarkkisen luonteensa takia kaiken datan esittämiseen. Ruutupu on hierarkkisesti järjestetty ja suhteessa toisiin pieniin neliöihin visualisoitu graafinen esitys. Ruutupu-visualisointia näkee usein esimerkiksi kovalevyn vapaan tilan visualisoinnissa. Kovalevyllä oleva data on kooltaan hierarkkisessa järjestyksessä, joten ruutupu on luonnollinen visualisointi tapa sen esittämiseen.

Parhaimmaksi tavaksi visualisoida big dataa osoittautui *rinnakkaiskoordinaatisto* (parallel coordinates). Rinnakkaiskoordinaatisto muodostuu objekteja kuvaavista, rinnakkaisista ja pystysuorista akseleista. Pystysuorat akselit ovat yhteydessä toisiinsa poikittaisilla viivoilla, jotka kuvaavat niiden sijaintia suhteessa toisiin objekteihin. Kuvassa 5 on esimerkki rinnakkaiskoordinaatistosta. Rinnakkaiskoordinaatisto tuki kaikkia kolmea big datan ominaisuutta: volyymia, vauhtia ja vaihtelevuutta. Gorodov ja Gubarev [2013] käyttivät vaihtelevuudesta nimitystä *data dynamics*.



Kuva 5. Amerikkalaisten autojen ominaisuudet merkitty rinnakkaiskoordinaatistoon [Wegman 1990].

Big data tuo yrityksille sekä tutkijoille paljon uusia mahdollisuuksia. Mahdollisuuksia tutkia, löytää ja hyödyntää aivan uusia tuloksia, joita ei olisi voinut havaita pienellä datamäärällä. Visualisointi on merkittävässä roolissa suurten datamäärien hahmottamisessa ja mahdollistaa eri asioiden välillä ja piilossa olevien kytkösten havaitsemisen. Big data tuo visualisoinnin saralle uusia jo osittain löydettyjä ongelmia sekä on verrattain uusi aihe tieteen parissa eikä yhtenäisiä ohjeita ei big datan visualisointiin ole vielä määritelty. Havaittuihin ongelmiin on ratkaisuja jo löydetty, mutta nopeaa vauhtia kasvava datamäärä melko varmasti tuo lisää ongelmia. Ongelmiin olisikin hyvä keksiä kestäviä ratkaisuja.

6. Yhteenveto

Kuvioroina pystytään tietotekniikan avulla luomaan helpommin kuin koskaan aikaisemmin. Aikaisempien tutkimusten mukaan kuvioroina, datamuste ja datasuhde ovat vielä hieman kiistellyjä aiheita. Edellä on annettu hyvän visualisoinnin määritelmiä erityisesti kuvioroinaa sisältävien visualisointien parissa.

Kuvioroinan hyödyllisyyttä muistettavuuden ja opittavuuden kannalta olisi hyvä tutkia tulevaisuudessa. Kuvioroinan käytössä piilee mahdollisuus parantaa visualisoinnin muistettavuutta ja opittavuutta.

Big data tuo lisää haasteita suurella määrällään ja reaaliaikaisuudellaan. Kuvioroinaa täytyy suurten datamäärien kanssa käyttää harkiten. Nykyisissä tutkimuksissa ei ole määritelty big datan visualisointeja kuvioroinan tai muun tyyllittelyn näkökulmasta. Lukemani perusteella eriäviä mielipiteitä ilmaantuu, ja yhtenäisten ohjeiden puuttuminen hankaloittaa laadukasta ja yhtenäistä big datan visualisointia. Suurten datamäärien muistettavuus on mielenkiintoinen tulevaisuuden tutkimuskohde. Kuinka saadaan suurien datamäärien tärkeimmät osat muistettaviksi? Tähän kysymykseen tulevaisuuden tutkimukset toivottavasti löytävät ratkaisuja. Nykyisin big datalle hyväksi luokitellut visualisoinnit ovat vaikealukuisia, eivätkä tue kovin hyvin muistettavuutta. Ratkaisut muistettavuuden ongelmiin voidaan hyvinkin saada kuvioroinan avulla.

Viiteluettelo

- Aula Anne, Majaranta Päivi ja Ovaska Saila. 2005. *Käytettävyyystutkimuksen menetelmät*. Tietojenkäsittelytieteiden laitos, Tampereen yliopisto.
- Bateman Scott, Brooks Christopher, Genest Aaron, Gutwin Carl, Mandryk Regan L. and Mcdine David. 2010. Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts. In: *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, April 10–15, 2573–2582.
- Brauner Philipp, Holzinger Andreas, Schaar Anne K., Valdez André C. and Ziefle Martina. 2015. Reducing Complexity with simplicity - Usability Methods for Industry 4.0. In: *Proc. of 19th Triennial Congress of the IEA*, 1-8.
- Brescina Sabrina and Eppler Martin J. 2008. *The Risk of Visualization: a Classification of Disadvantages Associated with Graphic Representations of Information*. Institute of Corporate, University of Lugano.
- Chang Remco, Harrison Lane and Reinecke Katharina. 2015. Infographic Aesthetics. In: *Proc. of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, 1187–1190.
- Choo Jaegul and Park Haesun. 2013. Big-Data Visualization Customizing Computational Methods for Visual Analytics with Big Data. *IEEE Computer Graphics & Applications* 33, 4, 22-28.
- EMC Education Services. 2015. *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. John Wiley & Sons.
- Few Stephen. 2011. The Chartjunk Debate – A Close Examination of Recent Findings. *Visual Business Intelligence Newsletter*, 1-11.

- Few Stephen. 2013. Data Visualization for Human Perception. In: Interaction design foundation, *The Encyclopedia of Human-Computer Interaction*, 35.
- Heer Jeffrey, Hellerstein Joseph M., Kandel Sean and Paepcke Andreas. 2012. Enterprise data analysis and visualization: An interview study. *IEEE Transactions Visualization and Computer Graphics* 18, 12, 2917–2926.
- Heer Jeffrey, Jiang Biye and Liu Zhicheng. 2013. ImMens: Real-time visual querying of big data. *Computer Graphics Forum* 32, 3, 421–430.
- Inbar Ohad, Meyer Joachim and Tractinsky Noam. 2007. Minimalism in Information Visualization – Attitudes towards Maximizing the Data-Ink Ratio. In: *Proc. of the ECCE 2007 Conference*, 28-31.
- Nielsen Jakob. 1993. *Usability Engineering*. Academic Press Inc.
- Nielsen Jakob. 2008. How Little Do Users Read? Nielsen Norman Group, 06 May 2008. <http://www.nngroup.com/articles/how-little-do-users-read/>. Checked 28.9.2016.
- Simon Phil. 2014. *The Visual Organization: Data Visualization, Big Data, and the Quest for Better Decisions*. Wiley.
- Rowley Jennifer. 2007. The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science* 33, 2, 163–180.
- Tufte Edward R. 2001. *The Visual Display of Quantitative Information*, Graphics Press.
- Ware Colin. 2004. *Information Visualization: Perception for Design*, Morgan Kaufmann.
- Wegman Edward J. 1990. Hyperdimensional Data Analysis Using Parallel Coordinates. *Journal of the American Statistical Association* 85, 411, 664–675.

Digitalisaation haasteita sote- ja maakuntauudistuksessa

Valtteri Ylisalo

Tiivistelmä

Nykyajan tietoyhteiskunnassa digitalisaatiosta eli digitaalisten teknologioiden yleistymisestä arkielämässä on tullut monen alan toiminnan mullistaja ja selkäranka. Myös terveydenhuollon ala elää murroksessa digitalisaation edetessä, ja digitalisaatio onkin eräs keskeisimpiä elementtejä käynnissä olevassa sosiaali- ja terveydenhuollon palvelurakennemuutoksessa, jonka on tarkoitus astua voimaan vuonna 2019. Uudistukselle on asetettu useita digitalisaatioon liittyviä tavoitteita, jotka ovat tiiviisti sidoksissa sen onnistuneeseen läpivientiin.

Tutkielmassa otetaan kirjallisuuskatsauksen avulla selvää, mitä nämä tavoitteet ovat, mitä ongelmia uudistuksen ja yleisestikin terveydenhuollon digitalisaatiossa saattaa ilmetä ja millaisia ratkaisuja näihin ongelmiin voidaan löytää. Tutkimustulosten perusteella uudistuksen ja terveydenhuollon alan digitalisaatioon liittyy hyvin moniulotteisia ongelmia niin kokonaisuuden, tiedonkulun, käytettävyyden kuin tietoturvan suhteen. Tuloksista voidaan päätellä, että vaikka hyviä ja perusteellisia ratkaisuja näihin onkin olemassa, ongelmien ratkaiseminen on silti vaikeaa, koska ratkaisujen toteuttaminen käytännössä on haastavaa.

Avainsanat ja -sanonnat: digitalisaatio, tietojärjestelmä, sote- ja maakuntauudistus, terveydenhuolto.

1. Johdanto

Sote- ja maakuntauudistus, eli koko nimeltään sosiaali- ja terveydenhuollon palvelurakennemuutos ja maakuntauudistus, on massiivinen julkisen sektorin meneillään oleva hallinnollinen hanke. Tällä hetkellä uudistuksesta vastaa Juha Sipilän hallitus, joka hallitusohjelmansa mukaisesti määrittelee, että siinä on kyse sosiaali- ja terveyspalvelujen saatavuuden ja laadun yhdenvertaisuuden parantamisesta, palvelujen järjestämisen ja ostokustannusten hillitsemisestä sekä eri vastuiden selkeyttämisestä terveydenhuollon eri toimijoiden välillä. Hallinnollisten muutosten seurauksena iso osa kuntien palveluiden vastuusta siirtyy maakunnille ja samalla järjestetään myös uudet maakuntavaalit, jotka tulevat selkeyttämään tällä hetkellä osittain sekavaa maakuntahallintoa ja tekemään siitä demokraattisempaa. Tämä helpottaa myös kuntien vastuita esimerkiksi sähköisten palveluiden järjestämisessä. [Alueuudistus.fi]

Uudistus on Suomen suurimpia sekä yhteiskunnallisesti merkittävimpiä hallintouudistuksia itsenäistymisen jälkeen ja sen taloudellinen raha-arvo liikkuu

miljardeissa. Siitä on kiistelty paljon viime vuosina ja poliittiset erimielisyydet aiheesta olivat lähellä kaataa maan hallituksen marraskuussa 2015. Uudistukseen liittyy hyvin monenlaisia ulottuvuuksia ja tavoitteita, jotka voivat tuntua tavallisille kansalaisille vaikeilta hahmottaa: syynä ei ole ollut vain uudistuksen suuruus ja monimutkaisuus, vaan myös sitä koskenut puutteellinen tiedottaminen. Myös jatkuvasti uutisoidut ongelmat uudistuksesta heikentävät nekin kansalaisten luottamusta uudistukseen ja uskottavuutta sen tavoitteiden saavuttamisessa.

Uudistuksen yksi kivijaloista on digitalisaatio, ja siten siinä korostuu erilaisien digitaalisten palvelujen ja toimintamallien tarkastelu, uudistaminen sekä yhteensovittaminen. Kyseessä on erittäin kallis hanke, joten on tärkeää suunnitella esimerkiksi sitä koskevien tietojärjestelmien kokonaisuus huolella. Jos tässä ei onnistuta, vaarana on, että uudistuksella ei kyetä saavuttamaan sille asetettuja tavoitteita. Tästä syystä uudistuksen digitalisaation onnistuminen on erittäin tärkeää.

Tutkielman tarkoituksena on kirjallisuuskatsauksen avulla selvittää, mitä uudistukselle asetetut digitaaliset tavoitteet tarkoittavat, mitä ongelmia niiden saavuttamisessa saattaa ilmetä ja millaisia ratkaisuja niihin on olemassa. Samalla käsittelen yleisiä terveydenhuollon alan digitalisaatioon liittyviä ongelmia ja ratkaisuja, sillä ne auttavat ymmärtämään paremmin uudistuksen digitalisaatiota. Tarkoituksena on luoda yleiskuva käsittelemistäni asioista syventymättä mihinkään osa-alueeseen kovin yksityiskohtaisella tasolla.

Luvussa 2 esittelen ensin uudistuksen digitalisaation keskeisiä tavoitteita ja niitä strategioita, joita hallitus sekä keskeisessä asemassa olevat asiantuntijat ovat esittäneet näiden tavoitteiden saavuttamiseksi. Luvussa 3 puolestaan tarkastelen niitä digitalisaatioon liittyviä ongelmia, joita uudistuksessa saatetaan kohdata sekä yleisesti tämänhetkisiä ongelmia terveydenhuollon digitalisaatiossa. Luvussa 4 tuon esille mahdollisia ratkaisuja niihin ongelmiin, joita luvussa 3 on käsitelty. Lopuksi päätän tutkielman yhteenvetoon, jossa kertaan työn pääkohdat ja pohdin vielä hieman laajemmin koko uudistusta ja sen tulevaisuuden näkymiä.

2. Digitalisaation tavoitteet uudistuksessa

Sote- ja maakuntauudistuksessa ajankohtaiset yhteiskunnalliset haasteet ovat pitkälti sidoksissa sen digitalisaation tavoitteisiin ja samoin koko terveydenhuollon alan digitalisaatioon. Jotta uudistuksen ja koko terveydenhuollon alan digitalisaation ongelmia voidaan laajasti tarkastella, on syytä ensin määritellä, mitä digitalisaatio ylipäänsä tarkoittaa ja millaisia siihen liittyviä tavoitteita uudistuksessa on asetettu.

Digitalisaatiosta (digitalization) ei ole olemassa yksiselitteistä määritelmää tai käsitystä. Tässä tutkielmassa tukeudun Visalan [2011] esittämään määritelmään

digitalisaatiosta, jossa sen katsotaan olevan erilaisten digitaalisten teknologioiden yleistymisen ihmisten arkielämässä. Se mielletään osaksi todellisuutta, jossa teknologia on sulautunut ympäristöön keskeiseksi osaksi arkipäivää ja se toimii erilaisten toimintojen instrumenttina. Tällaista määritelmää täydentäisiin vielä sisällyttämällä siihen digitaalisen teknologian laajan sosiaalisen hyväksymisen yhteiskunnassa lähes kaikkien ihmisryhmien keskuudessa.

Digitalisaatiossa eräitä keskeisimpiä elementtejä ovat tietojärjestelmät eli tietotekniset systeemit, jotka mahdollistavat jonkin toiminnan tietojenkäsittelyn. Ne ovat periaatteessa tietotekninen apuväline datan keräämiseen, tallentamiseen, jakamiseen ja hyödyntämiseen, jolla mahdollistetaan digitaaliset toimintatavat erilaisten organisaatioiden ja muiden toimijoiden keskuudessa.

Digitalisaatio on nimetty yhdeksi Sipilän hallituksen viidestä kärkihankkeesta, ja hallituksen strategisen ohjelman mukaisesti se yhdistyy vankasti sote- ja maakuntauudistuksen sisältöön. Uudistuksen tavoitellusta noin kolmen miljardin säästöistä noin viidenneksen on tarkoitus syntyä digitalisaation oikeanlaisesta hyödyntämisestä. Tämä noin 500–600 miljoonan säästö pyritään saamaan aikaan saavuttamalla uudistuksen digitalisaation kuusi tavoitetta, jotka on esitetty kuvassa 1. [Valtioneuvosto.fi]



Kuva 1. Digitalisaation tavoitteet uudistuksessa [Alueuudistus.fi].

Käsittelen seuraavaksi jokaista digitalisaation tavoitetta omissa kohdissaan, lukuun ottamatta tavoitetta asiakastietojen sujuvasta siirtymisestä eri tuottajien

välillä. Tätä poikkeusta perustelen sillä, että tämä tavoite liittyy kiinteästi ja oleellisesti tavoitteeseen asiakastietojen käyttämisestä organisaatio- ja aluerajoista riippumatta: sitä ei ole siten mielekäästä käsitellä erillisenä kokonaisuutena.

2.1. Kustannustehokas palvelurakenne digitalisaation avulla

Hallituksessa on katsottu, että digitalisaation avulla kansalaisten tarvitsemia palveluita voidaan tuottaa ja tehostaa uusilla innovatiivisilla, teknologian mahdollistamilla toimintatavoilla, jotka tukevat paremmin kansalaisten yhdenvertaisuutta palveluiden suhteen [Valtioneuvosto.fi]. Korhosen [2016] mukaan sote- ja maakuntauudistuksen tavoiteltu kustannustehokas ja joustava palvelurakenne pyritäänkin toteuttamaan hyödyntämällä juuri uusimpia toimintatapoja sekä trendejä, joita on syntynyt teknologian nopean edistymisen myötä. Niiden avulla koordinoidaan ja tuetaan koko sosiaali- ja terveydenhuollon palveluiden järjestämisen ja tarjoamisen palvelurakennetta niin, että se palvelee paremmin kansalaisten etuja.

Yksi keskeinen kustannustehokkaan palvelurakenteen saavuttamiseen käytettävä instrumentti on syksyllä 2015 käynnistynyt Una-hanke, jonka tavoitteena on tuottaa kansallisessa yhteistyössä sellaisia yhteentoimivia tietojärjestelmäratkaisuja, jotka parhaiten tukevat uusia digitaalisia toimintatapoja. Hankkeella pyritään siirtymään nykyisestä jäykästä palvelurakenteen kehittämismallista uuteen, *moduulitekniikkaa* (modular design) tukevaan malliin. Se tarkoittaa mallia, jossa eri tietojärjestelmien osia suunnitellaan ja kehitetään keskenään samankaltaisilla standardoiduilla tekniikoilla. [Andersson 2016]

Anderssonin [2016] mukaan tulevaisuuden terveydenhuollon palvelurakenne ja tietojärjestelmäkokonaisuus ja sen vaatimusmäärittely tullaan näillä näkymin toteuttamaan Una-hankkeen arkkitehtuurilinjauksiin. Uudistuksessa järjestelmiä ja niiden osia määritellään yhteistyössä erityisesti markkinoiden kanssa. Jatkuvasti muuttuvien toimintamallien tarpeisiin pyritäänkin vastaamaan hyödyntämällä markkinoilla saatavilla olevaa terveydenhuollon ja teknologian osaamista.

2.2. Sujuva asiakastietojen käyttö organisaatio- ja aluerajojen välillä

Sote- ja maakuntauudistuksen digitalisaatio edellyttää sujuvaa kansalaisten eli asiakkaiden tiedonkulkua eri tietojärjestelmien välillä [Alueuudistus.fi]. Korhonen ja Virtanen [2015] toteavat, että asiakkaiden omia potilastietoja asetetaan uudistuksessa nykyistä tehokkaampaan hyötykäyttöön kehittämällä olemassa olevien järjestelmien kommunikaatiöväyliä eri palveluntarjoajien välillä. Järjestelmistä tullaan myös poistamaan päällekkäisyyksiä ja niistä integroidaan toimivampi kokonaisuus, joka sisältää nykytilaan verrattuna vähemmän järjestelmiä.

Järjestelmien sisältävät laajat tietovarannot tullaan myös päivittämään siten, että ne tukevat paremmin tiedonkulun sujuvuutta.

Nissisen [2010] mukaan uudistuksessa pyritään kehittämään uusia mahdollisuuksia asiakastietojen käyttöön yli organisaatio- ja aluerajojen. Tämän pitäisi tulla onnistumaan mahdollisimman hyvin kaikissa sosiaali- ja terveyshuollon organisaatioissa ja niiden käyttämissä tietojärjestelmissä, jotta palvelukokonaisuus palvelisi paremmin kansalaisten etuja. Organisaatio- ja aluerajoja pyrittäneen madaltamaan niin, että niissä asetetaan asiakastietojen käsittelyn sujuvuus etusijalle. On kuitenkin huomioitava, ettei tästä madaltamisesta saisi koitua tarpeetonta haittaa järjestelmiään päivittäville sosiaali- ja terveyshuollon organisaatioille.

Aiemmin esitellyn Una-hankkeen eräs keskeisiä rooleja on pyrkiä luomaan sellainen tietojärjestelmien kokonaisuus, jossa mahdollistetaan asiakkaille helppo pääsy heitä koskevaan ajantasaiseen tietoon riippumatta tiedon sijainnista ja siitä, missä sitä on viimeksi käsitelty. Tämä saadaan aikaan tekemällä järjestelmien rajapinnoista avoimempia. Järjestelmien kokonaisuutta suunniteltaessa Una-hankkeessa pyritään jatkamaan sosiaali- ja terveydenhuollon organisaatioissa jo käynnistettyä yhteistoimintaa sekä yhteensovittamaan asiakaskeksisiä toimintamalleja ja digitaalisia palveluita. Näiden lisäksi Una-hankkeessa pyritään korostamaan organisaatio- ja aluerajojen näkymättömyyttä kehitetyissä digitaalisissa palveluissa, jotta tiedot eivät jämahdä kiinni organisaatioiden omiin rakenteisiin. [Andersson 2016]

2.3. Digitaalisten palvelujen oma käyttö

Sote- ja maakuntauudistuksessa hyvin toteutettu digitalisaatio edellyttää sellaisen digitaalisten palvelujen kehittämistä, jotka ovat mahdollisimman helppokäyttöisiä ja turvallisia kansalaisen omasta taustasta riippumatta. Tähän on jo kiinnitetty paljon huomiota sekä yksityisen että julkisen sektorin digitaalisissa palveluissa siten, että kansalaisten antamaa palautetta palveluista on kuunneltu aktiivisesti esimerkiksi sosiaalisen median kautta.

Digitaaliset palvelut ovat kansalaisille helppo, nopea ja tehokas tapa hoitaa viranomaisasioita paikasta ja ajasta riippumatta. Samalla kun digitaalisten palvelujen suosio ja käyttö eri muodoissa lisääntyy, myös julkisen palvelutuotannon toiminta tehostuu ja tämän myötä julkinen talous säilyy kestävämmällä pohjalla. Digitaalisten palveluiden merkitys on todellakin tunnustettu viranomaistojen keskuudessa ja niihin on suunnattu paljon valtion resursseja, sillä niiden toimivuus tukee hyvin viranomaistojen omien tavoitteiden saavuttamista. [Vm.fi]

Lähtökohtana on, että julkishallinnon asiointiin liittyvien digitaalisten palveluiden tulisi aina olla mahdollisimman toimivia, helppokäyttöisiä ja turvallisia

kaikille kansalaisille. Julkishallinnon asiakkuusstrategian mukaisesti viranomaisten tuleekin huolehtia siitä, että digitaalinen palvelukanava olisi asiakkaille aina paras mahdollinen vaihtoehto hoitaa asioitaan heidän omista lähtökohdistaan riippumatta. Tämän päämäärän saavuttamiseksi digitaalisen palvelukanavan kehittämisessä keskeistä on ennen kaikkea käyttäjäkeskeinen suunnittelu, palveluprosessien uudistaminen, palvelujen yhteentoimivuus sekä tietoturva ja -suoja. [Vm.fi]

Eräs tärkeimmistä kansalaisten käytössä olevista digitaalisista palveluista on Kansallisen Terveysarkiston ylläpitämä valtakunnallinen verkkopalvelu Oma-kanta, josta voi helposti hakea omat terveydenhuollon kirjaamat tiedot, esimerkiksi tuoreimman sähköisen reseptin yksityiskohtaiset tiedot [Kanta.fi]. Omakannasta on tärkeää nostaa esille, että kaikkien terveydenhuollon organisaatioiden henkilöstön pitäisi käyttää aktiivisesti Omakantaa siten, että asiakkaiden tiedot myös todella siirtyvät sinne. Uudistuksen myötä Omakantaa kehitetään niin, että myös etänä annetut terveydenhuollon palvelut pyritään saamaan entistä lähestyttävimmiksi verrattuna perinteisiin vastaanottoihin, vaikka etäpalvelut ovatkin jo pitkään olleet arkitoimintaa sosiaali- ja terveydenhuollon alalla [Korhonen 2016].

Hallituksen suunnitelmien mukaan uudistuksella edistetään riskitöntä omien hoito- ja lääketietojen tarkastelumahdollisuutta ajasta ja paikasta riippumatta. Digitaalisia palveluita tarjoavalla sektorilla ei saisi olla tähän merkitystä, vaan jokaiselle kansalaisille tulee tarjota hänen käyttämässä palvelussa sellainen suojattu näkymä omista tiedoista, jonka turvallisuuteen he voivat luottaa. Tämä saavutetaan rakentamalla semanttinen, yhteensopiva ja tietosuojatasoltaan virheetön tietojärjestelmäkokonaisuus, joka tukee suojattua dataa potilaan tiedoista kaikissa digitaalisissa palveluissa. [Alueuudistus.fi]

Kansalaisilla on tällä hetkellä hyvät perusvalmiudet sekä myönteinen asenne käyttää digitaalisia palveluita hyödyksi esimerkiksi hakemalla verkosta tietoa oman sairautensa hoitoon [Jauhiainen *et al.* 2014]. Digitaalisten palvelujen oman käytön yhteydessä uudistuksessa pyritään huolehtimaan siitä, että palvelujen käyttäminen olisi mahdollisimman helppoa eri ihmisryhmissä. Laiho [2011] toteaa, että toistaiseksi siinä on vielä parantamisen varaa, sillä tällä hetkellä esimerkiksi nuorten ja vanhempien ikäluokkien digitaaliset taidot poikkeavat paljon toisistaan. Nuoret omaksuvat erilaisten digitaalisten palvelujen käytön huomattavasti helpommin ja nopeammin kuin vanhemmat ihmiset, jotka puolestaan saattavat kokea tällaisen toiminnan vaikeaksi ja turhauttavaksi.

On tärkeää nostaa esille, että edellä mainituilla ikäihmisillä on motivaatiota kehittää tietoteknisiä taitojaan, koska he saattavat muutoin kokea jäävänsä ulkopuoliseksi yhteiskunnasta: monien kohdalla tällainen syrjäytyminen on jo tapahtunut [Laiho 2011]. Osallistumisen halukkuutta voitaisiin edistää järjestämällä heille sellaista tietoteknisten taitojen opetusta, johon on matala kynnys osallistua

ja jossa apua tarjotaan kädestä pitäen. Tällaisessa opetuksessa kannattaisi antaa digitaalisten palvelujen käyttöavun lisäksi myös yleisluontoisia tietoturvaohjeita siitä, mitä tietoja ikäihmisten kannattaa syöttää ja tallentaa erilaisiin palveluihin ja mitä ei.

2.4. Vahva valtakunnallinen ohjaus ICT-ratkaisuihin

Hallitus on esittänyt, että sote- ja maakuntauudistuksessa maakuntien oman tietohallinnon ja ICT-toiminnan ohjausta varten muodostetaan valtakunnallinen tietohallinnon ohjaustoiminto, joka ohjaa digitaalisten palvelujen toimivaa hankintaa, suunnittelua, kehittämistä, järjestämistä ja tarjontaa. Tämä ohjaustoiminto tulee tärkeäksi osaksi muuta maakuntien valtakunnallista ohjausta. [Alueuudistus.fi]

Oikarisen ja Larsion [2016] mukaan ICT-ratkaisujen kehittämisessä keskeistä on olla tietoinen niistä ohjaavien toimijoiden tavoitteista. Tätä pyritään edistämään aktiivisella vuoropuhelulla eri toimijoiden välillä. Ohjausta tullaan johtamaan yhtenä kokonaisuutena, jossa varmistetaan ratkaisujen yhteentoimivuus ja kustannustehokkuus. Siinä tullaan vaatimaan tiivistä viestintää ministeriöiden ja maakuntien välillä niistä digitaalisiin palveluihin liittyvistä ratkaisuista, joita uudistuksella tavoitellaan. Mielestäni tällainen ohjaus näyttäytyy vahvasti sellaisena, että ratkaisut tulevat olemaan toimialakohtaisia, joissa on tärkeää tunnistaa erilaisten organisaatioiden strategiset tavoitteet ja visiot.

Uudistuksen myötä kaikissa ICT-ratkaisuissa pyritään käyttämään sellaisia ohjausvälineitä, jotka pystyvät vaikuttamaan tehokkaasti myös terveydenhuollon organisaatioiden tietohallinnon omaa toimintaan ja tarkoituksena on löytää keinoja siihen, miten jokaisen organisaation tietohallinnon toimintaa voidaan parhaiten edistää [Oikarinen ja Larsio 2016]. Vaikka jokaisen keinon pitäisi olla ensisijaisesti tietohallinnon omia prosesseja tukeva, samalla on myös huomiotava niihin kohdistetut valtakunnallisen ohjaustason asettamat määräykset koskien esimerkiksi prosessien budjetointia, mittaamista ja raportointia.

Korhosen [2016] mukaan uudistuksessa edistetään terveyskeskusten ICT-ratkaisuja kokonaisuudessaan niin, että eri toimijoiden prosessit kytketään enemmän yhteen ja niitä kehitetään mahdollisimman paljon rinnakkain. Tätä pyritään varmistamaan kehittämällä lainsäädäntöä paremmaksi niin, että ICT-ratkaisujen kehittäminen olisi joustavampaa. Yhteisiin ICT-ratkaisuihin pyritään, jotta digitaaliset palvelut olisivat kansalaisten näkökulmasta valtakunnallisesti mahdollisimman yhdenvertaisia. Tämä vaikuttaa palvelevan paremmin kansalaisten etua, kun maakuntien väliset alueelliset erot digitaalisten palveluiden suhteen

eivät tule olemaan suuret: jokaisessa maakunnassa pystytään tarjoamaan kansalaisille mahdollisimman laadukkaita digitaalisia palveluita.

2.5. Yhteinen ICT-palvelukeskus ja yhteiset hankinnat

Hallituksen suunnitelmien mukaan jokainen maakunta tulee sote- ja maakuntauudistuksen myötä vastaamaan tulevaisuudessa sekä oman tietohallinnon että digitaalisten palveluidensa järjestämisestä. Maakunnissa tulee edistää uusimman hoitoteknologian käyttöönottoa ja digitalisaatiota niiden järjestämissä yksiköissä, jotka tulevat tarjoamaan terveydenhuollon palveluita kansalaisille. [Alueuudistus.fi]

Maakuntien digitaalisten palveluiden tuottaminen pyritään keskittämään pääosin yhteiseen valtakunnalliseen ICT-palvelukeskukseen. Tämä tehdään, jotta digitaalisten palveluiden laatu, yhtenäinen toimintalogiikka ja kokonaisvaltainen yhdenvertaisuus voidaan taata mahdollisimman hyvin maakunnan taloudellisesta tilasta ja sen henkilöstön ICT-ammattitaidosta riippumatta. ICT-palvelukeskuksella tulee olemaan tärkeä rooli yhdenvertaisten digitaalisten palveluiden järjestämisessä kaikissa maakunnissa. [Korhonen 2016]

Joitakin digitaalisia palveluita voidaan tulla mahdollisesti kilpailuttamaan valtakunnallisen yhteishankintayksikön kautta, mikäli se nähdään tarpeelliseksi maakunnassa [Alueuudistus.fi]. Lisäksi samalla tavoin kuin kuntien välillä on tehty yhteistyötä digitaalisten palveluiden järjestämisessä ja tarjoamisessa, myös maakuntatasolla tai maakuntien ja kuntien välillä sosiaali- ja terveystaloudellisia palveluita voidaan hoitaa samalla tavalla [Stm.fi]. Ilmeisesti maakunnille tullaan antamaan melko paljon vapauksia sen suhteen, millaisella toiminnalla ne muodostavat oman palveluverkkonsa ja minkälaisia eri toimijoita tässä palveluverkossa saa olla mukana.

Toisaalta vaikka digitaalisten palvelujen järjestämisvastuu olisikin jollakin toisella maakunnalla tai palvelun tuottaisi jokin toinen maakunta, niin lähtökohteisesti maakuntien kannattaisi rahoittaa ne itse. Tätä perustelen sillä, että omakustanteinen rahoitus kalliiden digitaalisten palveluiden osalta edesauttaa maakunnan kykyä pitää oma taloutensa vakaana etenkin, kun näillä näkymin niille ei ole tulossa omaa verotusoikeutta, jolla paikata budjettialijäämää taloudellisesti huonoina aikoina.

Hallitus on esittänyt, että uudistuksessa pyritään kehittämään mahdollisimman yhtenäistä kokonaisuutta digitaalisten palveluiden tietovarannoille, jotta esimerkiksi yhteisiin ICT-hankkeisiin tarvittavaa tietoa on helpompi varastoida. Samalla edistetään sidosryhmien osalta toimivaa yhteistyöverkostoa, jotta digitalisaatiota voidaan edistää koko valtion tasolla ja maakunnissa parhaalla mahdollisella tavalla. Tämä vaatii maakuntien ja valtion yhteisen ICT-palvelukeskuksen tehtävien ja palveluiden selkeää määrittelyä ja sujuvaa viestintää niistä.

Toimivan kokonaisuuden valmistelun onnistumiseksi uudistuksessa tullaan tarvitsemaan myös uudistuksen tavoitteista johdettuja tarkentavia linjauksia, jotka liittyvät ICT-palvelukeskuksen ohjelmaan ja sen toimintaan. [Kuusisto ja Kantola 2016]

3. Ongelmia uudistuksen ja terveydenhuollon digitalisaatiossa

Tässä luvussa tarkastelen sote- ja maakuntauudistuksessa ilmeneviä ja yleisesti terveydenhuollon alalla havaittuja digitalisaation ongelmia. Tutkielman kannalta ongelmien rajaaminen on hieman tulkinnanvaraista, joten keskityn vain niihin, joihin hallituksen asettamat digitalisaation tavoitteet liittyvät. Ongelmat on jaettu neljään eri kohtaan: kokonaiskuvan puuttuminen, tiedonkulun ongelmat, käytettävyysongelmat sekä tietoturvariskit. Näillä kaikilla osa-alueilla on merkittävä vaikutus uudistuksen digitalisaatiossa.

Ongelmista tarkastelen ensiksi kokonaiskuvan puuttumista, sillä sitä kautta on luontevaa syventyä niihin matalamman tason ongelmiin, joita käsitellen seuraavissa kohdissa. Jokaisesta ongelmasta pyrin antamaan riittävän perusteellisen ja oikean kuva, jotta vastaavia ratkaisuja näihin ongelmiin voidaan esittää luvussa 4.

3.1. Kokonaiskuvan puuttuminen

Terveydenhuollon tietojärjestelmien, ja ehkäpä laajemmin koko terveydenhuollon, kokonaisuudesta ei näytä koskaan olleen kenelläkään kovin selkeää kuvaa, vaan tämä kokonaisuus on vuosikymmenten saatossa muotoutunut nykyiseen tilaansa ilman sen suurempaa organisoitua. Mäki-Lohiluoma [2016] arvioi, poliittiset päättäjät eivät ole saaneet pysyvää otetta terveydenhuollon kokonaisuudesta, koska neljän vuoden kestoiset hallituskaudet ovat olleet liian lyhyitä, jotta siitä voitaisiin saada aikaan perusteellisia ja pitkäjänteisiä ratkaisuja. Ongelmana lienee se, että seuraavat hallituskaudet ovat harvoin sitoutuneita edellisen hallituksen päätöksiin ja saattavat pahimmassa tapauksessa romuttaa aikaansaadut päätökset kokonaan.

Terveydenhuollon kokonaisuuskuvan puuttumisen yksi syy näyttääkin siis löytyvän syvältä poliittisen järjestelmän rakenteista. Ongelmallisuutta lisää se, että näiden rakenteiden muuttaminen on lähes mahdotonta riippumatta siitä, ketkä ovat hallitusvastuussa ja muuttamista tuskin muutenkaan tehtäisiin terveydenhuollon kokonaisuuden edistämisen nimissä. Lisäksi monien eri toimijoiden erilaiset, keskenään ristiriitaiset poliittiset tavoitteet näyttävät tekevän tästä kaikista vielä vaikeampaa.

Kokonaiskuvan puuttumisen lisäksi terveydenhuollon alalla suuren mittaluokan tietojärjestelmiä näyttää olevan suhteellisen hankalaa hallinnoida, mikä vaikeuttaa toimivan digitalisaation mahdollistamista. Järjestelmiä on erittäin paljon ja ne ovat muodostaneet sirpaloituneen kokonaisuuden, joka on kaikkea

muuta kuin selväpiirteinen. Moni järjestelmä on suunniteltu ja kehitetty lyhytnäköisesti, jolloin mahdollisiin ongelmiin tulevaisuudessa ei olla varauduttu. Pidemmällä aikavälilläkään niitä ei ole osattu ennakoida, koska teknologian kehittyminen on ollut nopeampaa kuin on osattu kuvitella. [Koivu 2016]

Seppälä [2014] esittää potilastiedon hallinnointiin liittyen, että potilastietojärjestelmien kehittäminen on ollut vaikeaa, koska potilastiedot ovat monimuotoisia ja ne ovat harvoin täysin yhteensopivia järjestelmien rakenteiden kanssa. Tätä vaikeuttaa entisestään se, että potilastietoja käytetään hyvin monenlaisissa käyttö- ja toimintaympäristöissä.

Kokonaisarkkitehtuuri (enterprise architecture) tarkoittaa käsitteellistä mallia ja kokonaisnäkemyä yhden tai useamman organisaation tilasta ja rakenteesta. Sote- ja maakuntauudistuksessa kokonaisarkkitehtuurista vastaavilla henkilöillä olisi erinomainen tilaisuus yrittää paikata terveydenhuollon tietojärjestelmien kokonaisuuden puutteen epäkohtaa. Smedberg [2014] toteaa, että nykyisessä kokonaisarkkitehtuurin suunnittelutyössä erityisesti asiakaslähtöisyys on jäänyt liian vähälle huomiolle ja sen merkitystä ei ole vielääkään osattu ymmärtää riittävän hyvin. Se on elänyt pitkälti vain erilaisissa toiveissa ja visioissa eikä sen eteen ole tehty riittävästi konkreettista työtä.

Kuusisto ja Kantola [2016] arvioivat, että uudistuksen edellyttämän ICT-muutoksen massiivisuus luo suuria vaatimuksia niin resurssien, tietotaidon kuin riskien hallinnalle. Jos näiden muodostamasta kokonaisuudesta ei osata luoda oikeaa kuvaa, vaarana on, että niitä kohdistetaan väärin osa-alueisiin. Hukkaan heitetyt resurssit, arvokkaan tietotaidon hyödyntämättä jättäminen, kustannusten alimitoitus ja haitallisten riskien realisoituminen eivät ole kenenkään etu. Virtanen ja Stenvall [2010] katsovat, että tämänkaltaisia erehdyksiä syntyy yleensä siksi, että ei ole osattu tunnistaa riittävän hyvin niiden merkitystä, todennäköisyyttä eikä heijastusvaikutuksia.

Mäki-Lohiluoma [2016] korostaa, että edellä mainittuihin ongelmiin ei ole varaa, koska terveydenhuollon kustannukset kasvavat kiihtyvällä tahdilla: epäonnistuminen uudistuksen ICT-muutoksessa ajaisi kansantaloutta yhä huonompaan tilaan. Tästä syystä uudistuksen avainhenkilöillä näyttää olevan suuri vastuu kokonaisuuden onnistumisesta.

Uudistuksen koko tietojärjestelmäkokonaisuus voidaan nähdä valtakunnallisella tasolla suurena sosioteknisenä järjestelmänä, joka koostuu ihmisistä, laitteista, ohjelmistoista ja organisaatioista. Sommerville [2007] toteaa, että tällaisessa lähestymistavassa keskeistä on tunnistaa ne ongelmat, joita ilmenee näiden neljän asian yhteensovittamisessa. Vastaavasti Nykänen ja muut [2016] esittävät, että terveydenhuollon tietojärjestelmäkokonaisuus voidaan nähdä ekosysteeminä, joka koostuu organisaatioista, toimijoista, vuorovaikutuskanavista ja järjestelmistä. Valitettavasti näyttää siltä, että päättäjät eivät ole ottaneet riittävästi

hyötykäyttöön tällaisia lähestymistapoja digitalisaatioon sosiaali- ja terveysalalla: ongelmana on voinut olla sellaisen osaamisen löytäminen, jolla vastattaisiin jatkuvasti muuttuvan järjestelmäkokonaisuuden hallintaan sekä tunnistettaisiin kokonaisuuden eri osien merkitykset ja suhteet muihin osiin.

Kaiken kaikkiaan systemaattisesti rakennettu ja huolellisesti integroitu tietojärjestelmien kokonaisuus auttaisi palvelemaan paremmin koko terveydenhuollon alaa. Tällaiseen toimintaan olisi voitu tietysti pyrkiä jo vuosia ennen uudistusta, mutta ainakin nyt uudistuksen olevan käynnissä on noussut selvemmin esille, että kokonaiskuva on ollut todella puutteellinen.

3.2. Tiedonkulun ja valinnanvapausmallin ongelmat

Terveydenhuollon alalla tiedonkulun suurimpia ongelmia on se, että tietoa ei saada siirrettyä tietojärjestelmästä toiseen, jolloin päädytään ongelmalliseen tilanteeseen, jossa tieto ei ole saatavilla silloin kun sitä tarvittaisiin. Tiedon puuttumisen syynä saattaa olla esimerkiksi se, että se on kirjattuna eri organisaation järjestelmään eikä ole saatavilla, koska järjestelmälle ei ole kehitetty mahdollisuutta kommunikoida jonkin toisen järjestelmän kanssa, jossa tietoa tarvittaisiin. Tämä saattaa aiheuttaa jopa uhkaavia ongelmia potilaan turvallisuudessa ja hoidossa tietyissä tilanteissa. Esimerkiksi lääkityksen historian tiedon puute voi aiheuttaa potilaalle pahimmassa tapauksessa täysin väärän diagnoosin. [Helenius 2011]

Tiedonkulku eri organisaatioiden välillä voi aiheuttaa potilaalle uhkaavia tilanteita myös silloin, kun se on kirjattu vääränlaisena. Syynä saattaa olla datan tyypistäminen niin, että sen alkuperäinen merkitys on muuttunut tai jos se on kerätty vääränlaisella menetelmällä. Helenius [2011] korostaakin tiedon merkityksen vaillinaisuuden ja huonon laadun voivan aiheuttaa ikäviä tilanteita, joissa tieto on kulkenut vääränlaisena tietojärjestelmästä toiseen: tätä voidaan luonnehtia rikkinäisen puhelimen metaforalla.

Terveydenhuollon toimialalla tiedonkulun ja toimivan digitalisaation yksi este on jäykät organisaatio- ja aluerajat. Koivu [2016] toteaa, että potilastiedot saattavat sijaita hajautettuina eri potilastietojärjestelmissä ja niiden siirtäminen organisaatioista toiseen ei ole aina kovin helppoa. Näyttää siltä, että potilastietoja on ollut vaikea saada kerättyä yhteen paikkaan niin, että se ei olisi massiivisen työmäärän takana. Tämä johtunee siitä, että terveydenhuollon henkilöstöllä ei välttämättä ole riittävää pitkäjänteisyyttä ymmärtää organisaatio- ja aluerajojen alentamisen tärkeyttä potilastietojen siirtämisen sujuvuudessa, jotta se palvelisi paremmin potilaiden etuja.

Nykänen [2016] esittää, että vaikka tieto kulkeekin potilastietojärjestelmästä toiseen oikeassa muodossa, joissakin järjestelmissä se dokumentoidaan tai varastoidaan eri tietorakenteisiin hyvin monimutkaisesti, jolloin sen hakeminen voi

olla työlästä. Tätä vaikeuttaa vielä entisestään se, jos terveydenhuollon organisaatio ei ole määritellyt tiedonkäytön ja -haun oikeuksia tavalla, joka palvelisi parhaiten kutakin henkilöryhmää. Esimerkiksi hoitajilla saattaa olla lääkäreihin verrattuna rajatummat käyttöoikeudet tietojen hakuun tilanteessa, jossa hoitaja tarvitsisi samoja tietoja kuin lääkärikin.

Sen lisäksi, että potilastietojärjestelmät ovat tiedonkulun osalta puutteellisia, moni pitkään käytössä ollut järjestelmä näyttää olevan elinkaarensa loppuvaiheessa. Ajan myötä useista tietojärjestelmistä tulee *perinnejärjestelmiä* (legacy systems) eli sellaisia järjestelmiä, jotka ovat liian vanhanaikaista nykytarpeisiin nähden. Sote- ja maakuntauudistuksen yhteydessä tällaiset järjestelmät joutuvatkin uudelleentarkasteluun ja niistä on luovuttava, jos niiden teknologia ei pysty vastaamaan modernin toimintaympäristön vaatimuksiin. Toisaalta tällainen toiminta on ongelmallista, jos perinnejärjestelmän tarjoamat toiminnot ovat tärkeitä organisaation toiminnalle eikä järjestelmästä haluta luopua. [Sommerville 2007]

Narratiivinen potilaskertomus (patient narrative) tarkoittaa sellaista vapaassa muodossa kirjattua tekstiä potilaan tilaan liittyen, joka kirjataan potilastietojärjestelmiin. Ukkonen [2016] korostaa narratiivisen tekstin roolin tärkeyttä potilastietojärjestelmissä ja arvioi, että sitä ei ole vielä samalla tavoin hyödynnetty kuin tavallista rakenteista kirjaamismuotoa. Potilaskertomukset eivät esimerkiksi aina kulje sujuvasti nykyisten käytössä olevien järjestelmien välillä, mikä on ongelmalliselta, sillä vapaamuotoiset tekstit ovat tehokas keinoja kuvata potilaan terveydentilaa tilanteissa, joissa siitä olisi muuten hankalaa informoida.

Tarkastelen tässä kohdassa vielä tiedonkulun ongelmiin liittyen hallituksen kaavailemaa valinnanvapausmallia, sillä uudistuksen digitalisaatiolla on merkittävä rooli tämän valinnanvapausmallin toteuttamisessa. Suunnitteilla oleva valinnanvapausmalli tarkoittaa sitä, että asiakas saa itse vapaasti rekisteröityä hänelle mieluisimpaan terveydenhuollon palveluja tarjoavan yksikköön, joka voi toimia joko julkisella, yksityisellä tai kolmannella sektorilla. Toistaiseksi tämä tulee tapahtumaan vain perusterveydenhuollon osalta eikä siten esimerkiksi tule koskemaan erikoissairaanhoidon osalta.

Käytännössä edellä mainittu valinnanvapauteen perustuva toimintamalli saadaan toimimaan vain, jos potilaiden tietoja voidaan käsitellä vaivattomasti yli organisaatio- ja aluerajojen eri terveydenhuollon palveluja tarjoavien yksiköiden välillä. Valinnanvapausmallin hyödyt siis vesittyvät, jos tiedonkulku tietojärjestelmien välillä ei onnistu sujuvasti. Ongelmana on, että näillä näkymin koko malli tullaan toteuttamaan valtakunnallisella tasolla eikä esimerkiksi paikallisella kokeilulla, mikä tarkoittaa sitä, että sen astuttua voimaan tiedonkulku eri

järjestelmien välillä pitäisi olla sujuvaa kaikkien valinnanvapautta koskevien terveydenhuollon palveluja tarjoavien yksiköiden kohdalla.

Valinnanvapausmallin eräs ongelma on myös sen vaatimus terveydenhuollon toimivasta laaturekisteristä, joka tarkoittaa käytännössä laajaa julkisesti näkyvillä olevaa tietokantaa kaikista terveydenhuollon palveluja tarjoavista yksiköistä. Laaturekisterin hyöty olisi kiistaton, sillä sen avulla asiakkaat voisivat vertailla eri terveydenhuollon palveluja tarjoavien yksiköiden hintatasoa, laatua, saatavuutta sekä muita tärkeitä ominaisuuksia keskenään. Tämä vertailu on hankalaa, jos laaturekisteri on tiedoiltaan puutteellinen tai se ei ole vaivattomasti kaikkien saatavilla [Mäki-Lohiluoma 2016]. Tämä epäkohta korostuu erityisesti ikäihmisten kohdalla, joiden tietotekniset taidot ovat puutteellisia ja jotka yleensä vieläpä tarvitsevat eniten hoitoa [Laiho 2011].

Käytännössä laaturekisterin puuttuminen saattaa aiheuttaa asiakkaiden näkökulmasta hankalan tilanteen, jossa vallitsee *epäsymmetrinen informaatio* (information asymmetry). Se tarkoittaa tilannetta, jossa toisella osapuolella on enemmän informaatiota jostakin asiasta kuin toisella. Jos laaturekisteriä ei siis ole saatavilla verkossa siinä vaiheessa, kun terveydenhuolto avataan käytännössä vapaille markkinoille valinnanvapausmallin nimissä, niin ollaan asiakkaiden kannalta ongelmallisessa tilanteessa: he eivät välttämättä saa riittävästi informaatiota eri terveydenhuollon palveluja tarjoavia yksiköistä.

3.3. Käytettävyysongelmat

Sote- ja maakuntauudistuksessa on keskityttävä myös tietojärjestelmien käytettävyyteen, jotta esimerkiksi potilastietoja tarkasteleva terveydenhuollon henkilöstö ja asiakkaat pystyvät käyttämään järjestelmiä mahdollisimman sujuvasti. Järjestelmien käytettävyyden ongelmat ovat luonteeltaan usein moniulotteisia ja puutteellisten käyttöliittymien ongelmia ei ilmeisesti ole aina helppo tunnistaa. Kuitenkin keskeisimmät käytettävyyden puutteet näyttävät johtuvan siitä, että oikeita henkilöitä ei oteta mukaan käyttöliittymien suunnitteluun, kehittämiseen ja testaukseen riittävän tiiviisti eikä ajoissa.

Viitanen ja Nieminen [2009] toteavat, että terveydenhuollon tietojärjestelmien käytettävyyden kehittäminen kokonaisuutena on vaikeaa, jos ei osata huomioida riittävästi järjestelmän eri käyttötilanteita. Ilman oikeanlaista käytettävyyden suunnittelua siitä tulee enemmän epämääräistä arvailua, jolloin toimivia ratkaisuja havaitaan usein vasta silloin, kun järjestelmien käyttäjät antavat negatiivista palautetta niiden käytettävyydestä. Vaikuttaa siltä, että tällaiset ongelmat aiheutuvat usein siitä, että käyttäjiä ei vaivauduta ottamaan mukaan käytettävyyden aktiiviseen kehitykseen, mikä kuitenkin johtaisi kaikkien kannalta parempaan lopputulokseen.

Terveydenhuollon alalla käytettävyys- ja loppukäyttäjän näkökulman huomiointi tietojärjestelmähankinnoissa jää yleensä vähäiseksi. Käytännössä oikeita

käytettävyystarpeita ei osata tunnistaa järjestelmien *hankintaprosessin* (acquisition process) aikana. Hankintaprosessi tarkoittaa siis hankinnan menetelmällistä suunnittelua ja toteuttamista niin, että hankinta kilpailutetaan ja paras tarjous valitaan. Käytettävyys- ja loppukäyttäjän näkökulman yhdistämistä järjestelmien hankintaprosessissa näyttää olevan vaikeaa, kun järjestelmien loppukäyttäjien käyttötarkoitukset ovat olleet epäselvät. [Kaipio *et al.* 2015]. Tämä tarkoittanee, että järjestelmän käyttötilanteita ei ole osattu kartoittaa riittävästi.

Nissinen [2010] katsoo, että terveydenhuollon aluetietojärjestelmien, eli usean kunnan yhteisten tietojärjestelmien käytettävyydessä on paljon parantamisen varaa. Ongelmat johtuvat käytön toiminnan ja tiedonsaannin hitaudesta sekä puutteellisesta käyttöliittymästä. Toisaalta moni lääkäri näkee aluetietojärjestelmän silti toimivaksi apuvälineeksi, ja he kokevat saavansa sen käytöstä hyötyä työssään. Toistaiseksi koulutuksen lisääminen järjestelmien käytöstä sen mielekkyyden lisäämiseksi ei ole tuottanut toivottua tulosta lääkäreiden keskuudessa, vaikka he näkevätkin tällaiset keinot pääosin positiivisina yrityksinä parantaa järjestelmien käytön mielekkyyttä.

Tällä hetkellä terveydenhuollon tietojärjestelmien käyttö aiheuttaa lääkäreillä useita ongelmia. Viitasen ja muiden [2011] mukaan näyttää siltä, että lääkärit kokevat järjestelmien käytön tuskastuttavan rasittavaksi ja hitaaksi. Järjestelmät vastaavat lähetettyihin pyyntöihin melko hitaasti ja niiden satunnaiset kaatumiset aiheuttavat menetetyt ajan lisäksi myös tallentamattomien tietojen katoamista. Ei siis ihme, jos jo valmiiksi kiireiset lääkärit kokevat järjestelmät turhauttaviksi, jos niiden hitaan käytön lisäksi ne vielä aiheuttavat tilanteita, joissa lyhyellä aikavälillä tehty työ valuu kokonaan hukkaan.

Lääkäreiden kokemuksia ja näkemyksiä terveydenhuollon tietojärjestelmistä on esitetty Viitasen ja muiden [2011] kokoamassa taulukossa 1. Siitä voi havaita, että lääkärit eivät koe heidän näkökulmansa huomioimista riittäväksi järjestelmien käyttöliittymää suunniteltaessa. He kokevat puutteellisten järjestelmien käytön hidastavan potilastyötä, johon toivottaisiin jäävän enemmän aikaa. Järjestelmiä haukutaan yksinkertaisesti huonoiksi työvälineiksi, ja pieni osa lääkä-

reistä jopa esittää toiveita palata paperiaikaan, jolloin niitä ei ollut käytössä lainkaan. Monia lääkäreitä vaivaa pitkälti se, että heidän antamansa palaute järjestelmistä ei yleensä johda toivottuihin parannuksiin.

Kategoriat ja alakategoriat	Kommenttien lukumäärä
Kategoria 1: Näkemyksiä kehittämisen nykytila	(yht. 124)
• Käyttäjien näkökulmaa ei huomioida riittävästi	74
• Palaute ja huomautukset eivät johda muutostoiimiin	28
Kategoria 2: Kokemukset kehittämistyöhön osallistumisesta	35
Kategoria 3: Kokemukset tietojärjestelmien käytöstä työvälineinä	(yht. 1191)
• ”Nykyiset tietojärjestelmät ovat huonoja”	379
• Kokemuksia tietojärjestelmien sopivuudesta kliiniseen työhön	216
• Esimerkkejä käyttöön liittyvistä ongelmista	207
• Positiivisista käyttökokemuksista kertovat kommentit	158
• ”Tietojärjestelmät hidastavat potilastyötä”	126
• ”Liian monia erillisiä järjestelmiä käytössä”	57
• Aluetietojärjestelmiin liittyvät kommentit	37
• Tietojärjestelmien hankintapäätöksiin liittyvät kommentit	10
Kategoria 4: Tulevaisuuden visiot ja kehittämissuhteet	(yht. 2854)
• Kuvaus potilastietojärjestelmään toivotuista toiminnallisista ominaisuuksista	930
• Kuvaus hyvän potilastietojärjestelmän ominaispiirteistä (adjektiiveja)	656
• Toiveena yhtenäiset ja integroidut järjestelmät	453
• Toiveena vain yksi tietojärjestelmä Suomeen	292
• Yleisiä toiveita tulevaisuuden järjestelmiin liittyen	142
• Toive: tietojärjestelmien tulisi tukea organisaatioiden välistä yhteistyötä	135
• Vastaukset, joissa potilaan rooli tietojärjestelmien käyttäjänä nostettu esiin	86
• ”Käyttäjät mukaan kehitystyöhön!”	78
• ”Keskeneräisiä tietojärjestelmiä ei tulisi ottaa käyttöön”	58
• Toiveena paluu paperiaikaan	15
• Toiveena siirtyminen täysin paperittomaan terveydenhuoltoon”	7
• Ehdotus: Jokaisella terveydenhuollon organisaatiolla tulisi olla nimetty kehittämisvastuulääkäri	2

Taulukko 1. Lääkärien näkemyksiä tietojärjestelmistään [Viitanen *et al.* 2011].

Terveydenhuollon henkilöstön osalta lääkärit eivät kuitenkaan ole ainoa ryhmä, joka on ilmaissut negatiivisia ajatuksia työssään käyttämistään tietojärjestelmistä, sillä esimerkiksi Nykänen [2016] arvioi, että lääkäreiden ohella myös sairaanhoitajien ja lähihoitajien käyttämissä järjestelmissä on paljon parantamisen varaa erityisesti niiden vuorovaikutus- ja käyttöliittymäsuunnittelun osilta. Kokemukset ja näkemykset järjestelmistä ovat olleet monilla hoitajilla pääosin negatiivisia ja he ovat toivoneet voivansa osallistua enemmän heidän käyttämiensä järjestelmien suunnittelu-, kehitys- ja testausvaiheisiin.

3.4. Tietoturvariskit

Potilastietojärjestelmissä tietoturvan taso ei näytä koskaan olleen kovin korkea, ja tämä ongelma on vain korostunut digitalisaation myötä. Viime vuosina monet uutiset puutteellisesta tietoturvan tasosta ovat vähintäänkin heikentäneet luottamusta järjestelmien tietoturvaa kohtaan. Tiedotusvälineissä on uutisoitu paljon tapauksista, joissa järjestelmissä on voitu käyttää sellaisia suojaamattomia lo-kinäkymiä, joiden käytöstä ei jää jälkeä historiatietoihin. Tällaiset tapaukset ovat

saaneet monet kansalaiset huolestuneiksi arkaluontoisten tietojensa yksityisyydestä, etenkin kun tapauksia näyttää ilmenevän lähes säännöllisesti.

Lain mukaan potilaalle tulee taata oikeus tietää, kuka tai ketkä hänen tietojaan katselevat [Finlex 9.2.2007/159]. Joissakin tapauksissa lainsäädäntö potilastietojen tietoturvesta ja -suojusta lähtee jo EU-tasolta asti pakottavien direktiivien muodossa [EUR-Lex 95/46/EC]. Tätä lainsäädäntöä ei pystytä aina noudattamaan, koska käytössä olevia järjestelmiä ei valitettavasti ole kehitetty sellaisiksi, että ne tukisivat vahvaa tietosuojaa potilastietojen osalta. Monissa potilastietojärjestelmissä arkaluontoisia tietoja voikin helposti urkkia ilman, että toiminnasta jää käytännössä mitään digitaalista jalanjälkeä historiatietoihin.

Tietoturvan osalta lainsäädäntö itsessään on siis kunnossa, ongelma on sen noudattamisessa. Rousku [2014] toteaa, että lainsäädännön ohjeita yritetään kyllä noudattaa henkilöstön puolelta, mutta siinä ei onnistuta täydellisesti. Esimerkiksi terveydenhuollon organisaation toimintatapojen muutoksissa ei toisinaan muisteta huomioida nykyistä lainsäädäntöä, ja ikävimmässä tapauksissa siitä ei edes välitetä, koska se koetaan hankalaksi toteuttaa organisaation toiminnan osalta.

Terveydenhuollon tietoturvan kannalta ongelmallista on, että toiminta tapahtuu nykyään enemmän *jokapaikan tietotekniikan* (ubiquitous computing) eli ympäristöön sulautuneen tietotekniikan avulla. Siinä piilee erityisen suuria riskejä suojaamattoman tiedon käytöstä väärin tarkoituksiin. Sen yleistyessä tulevaisuudessa potilastietoja saatetaan käsitellä hyvin erilaisissa laitteissa, joten sote- ja maakuntauudistuksessa on kiinnitettävä erityistä huomioita laitteiden luotettavaan tietoturvaan siltä osin, että potilastiedot olisivat aina suojattuna kaikissa käyttöympäristöissä. Arkaluontoisen luonteensa takia potilastietoja pitäisi pystyä käsittelemään aina luottamuksella eri käyttöympäristöissä niin, että potilaiden hyvinvointi ei vaarannu [Ruotsalainen *et al.* 2014].

Terveydenhuollon datan määrän jatkuva kasvaminen ja sen liikkuminen eri tietojärjestelmien välillä on huolestuttavaa, jos datan siirtäminen ei noudata lainsäädäntöä: tällaisissa tapauksissa Valvira eli Sosiaali- ja terveystieteiden valvontavirasto voi ryhtyä toimiin uhkasakoilla, jotka voivat olla rahamäärällisesti hyvinkin merkittäviä. Tämä johtuu siitä, että vuodesta 2007 lähtien potilastietojärjestelmien on pitänyt täyttää tietyt laissa määritellyt toiminnot ja ominaisuudet, joilla potilastiedot tulee kerätä ja niitä käsitellä [Finlex 9.2.2007/159].

Tietoturvallisuuden lisäksi lienee syytä tarkastella hieman myös kyberturvallisuutta ja siihen liittyviä ongelmia terveydenhuollossa. Erityisen tärkeää on nostaa esille *esineiden internet* (Internet of Things), joka on toimintamalli, jossa yksittäiset laitteet yhdistetään toisiinsa internetin välityksellä. Nämä laitteet eivät rajoitu vain tietokoneisiin, vaan niitä voivat olla nykyään jo mitkä tahansa elektroniset laitteet, jotka voidaan kytkeä tietoverkkoon. Kuten muillakin aloilla, esineiden internetin toimintamahdollisuuksissa piilee suuria tietoturvariskejä

myös terveydenhuollon alalla. Ulkopuolisen pääsy tällaisten laitteiden tietoverkkoon voi lamaannuttaa koko systeemin, jos sen tietoturva on suunniteltu huonosti.

Esineiden internetiin liittyvien mahdollisuuksien merkitys tulee yhä vain korostumaan tulevaisuudessa ja niistä on käyty keskustelua uudistuksen yhteydessä. Uudistuksen digitalisaatiossa on nostettu esille esineiden internetin tuomia mahdollisuuksia ja sovelluskohteita, ja nämä ovat osa myös muita hallituksen kärkihankkeita [Valtioneuvosto.fi]. Ajankohtaista poliittista keskustelua seurattessa näyttää siltä, että ainakin osa poliittisista päättäjistä on osannut tunnistaa esineiden internetiin liittyvät tietoturvakysymykset ja niiden tärkeyden.

Terveydenhuollon alalla kokonaisuudessaan on erittäin tärkeä pohtia esineiden internetin mahdollisuuksien lisäksi myös sen riskejä. Jos oman terveytemme toiminnan ylläpitoa edistäviä laitteita hakkeroidaan, uhkana on terveyden kannalta merkittävien riskien realisoituminen. Näissä asioissa tarkastelun keskiöön tulisikin asettaa aina yksityisyys- ja tietoturvakysymykset. Ongelmallista on ennen kaikkea se, että useat laitteet, joita voitaisiin kytkeä esineiden internetiin, eivät yksinkertaisesti pysty tukemaan tarvittavaa tiedonsalausta riittävän hyvin. [Lake *et al.* 2014]

4. Ratkaisuja uudistuksen ja terveydenhuollon digitalisaatioon

Tässä luvussa tarkastelen joitakin sote- ja maakuntauudistuksen sekä terveydenhuollon digitalisaation mahdollisia ratkaisuja, joilla voisi niihin ongelmiin, joita edellisessä luvussa käsitelin. Osa näistä ratkaisuista on hallituksen strategisen ohjelman mukaisia, osa asiantuntijoiden tutkimustietoa ja osan omaa pohdintaani. Tarkoituksena ei ole etsiä täydellisiä ratkaisuja kaikkiin uudistuksen ongelmiin, vaan olen nostanut esille niistä vain joitakin keskeisiä, joita katson mielekkääksi käsitellä.

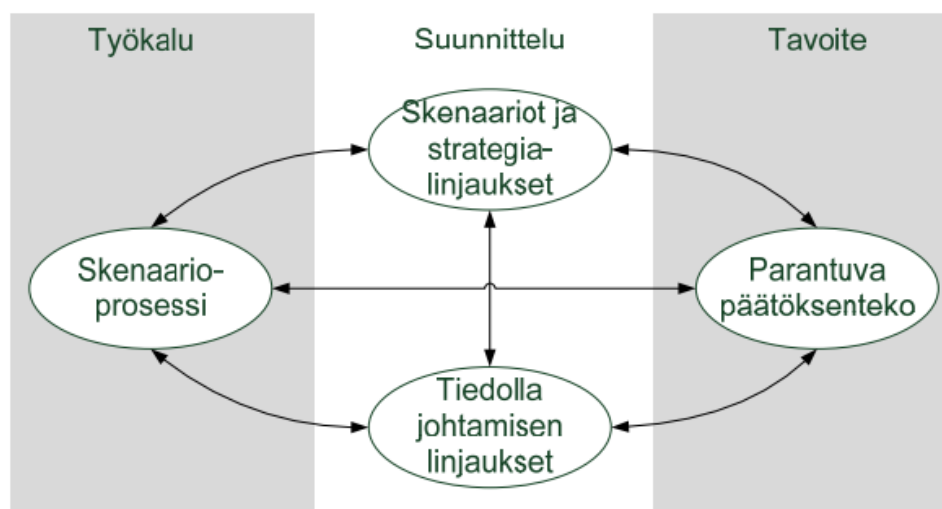
Ratkaisujen käsittelyjärjestys on sama kuin edellisessä luvussa käsiteltyjen ongelmien, eli ratkaisuja esitetään uudistuksen digitalisaatioon liittyvään kokonaisuuteen, tiedonkulkuun, käytettävyyteen sekä tietoturvaan. Jokainen ratkaisu on pyritty kuvaamaan riittävän oikeellisesti ja tarkasti, jotta se antaa uskottavan kuvan edellisessä luvussa esitettyjen ongelmien vastaamiseen.

4.1. Tietojärjestelmien kokonaisuuden hallinnan edistäminen

Sote- ja maakuntauudistuksen tietojärjestelmien kokonaisuudesta tarvitaan riittävän oikeellinen ja tarkka kuva, jotta uudistuksen digitalisaation tavoitteet voidaan saavuttaa. Vesiluoma [2012] toteaa, että yleisesti ottaen kokonaisuuden suunnittelua ei kannata kuitenkaan rakentaa yhden tulevaisuudentilan varaan,

vaan olisi tärkeää hahmottaa erilaisia skenaarioita sekä tunnistaa niitä polkuja, joita pitkin kulkemalla näihin tulevaisuudentiloihin voitaisiin päästä.

Eräs tulevaisuudentilan tunnistamista varten kehitetty skenaariomenetelmä on päätöksentekoa edistävä skenaarioprosessin malli, jossa aloitetaan ensin määrittämällä mahdolliset skenaariot saatavilla olevan tutkimustiedon perusteella. Sen jälkeen laaditaan johtamisen yksityiskohtaiset strategiset linjaukset, joiden perusteella tehdään sellaisia päätöksiä, jotka tukevat haluttujen skenaarioiden realisoitumista. Malli (ks. kuva 2) on iteratiivinen eli eri elementit tukevat syklimäisesti ja toistuvasti toisiaan.



Kuva 2. Päätöksentekoa edistävä skenaarioprosessin malli [Vesiluoma 2012].

Kun selkeät skenaariot erilaisista tulevaisuudentiloista ymmärretään, voidaan paremmin tehdä päätöksiä liittyen varsinaisiin tietojärjestelmiin. Koivun [2016] mukaan uudistuksen järjestelmien kokonaisuutta ajateltaessa on ensisijaisen tärkeää hahmottaa, mitkä käytössä olevista järjestelmistä säilytetään ja mitkä otetaan pois käytöstä. Tärkein huomio lienee se, että eri järjestelmien määrän minimointi koko valtakunnan tasolla helpottaa niiden muodostaman kokonaisuuden hahmottamista eri toimijoiden keskuudessa. On kuitenkin ymmärrettävää, milloin liian vähäisestä järjestelmien kokonaismäärästä ei koidu enää oleellista hyötyä pitkälläkään aikavälillä.

Uudistuksessa on tehtävä päätös siitä, mitkä terveydenhuollon tietojärjestelmät tulevat säilymään valtakunnallisina ja mitkä paikallisina. Järjestelmien tulisi olla ensisijaisesti valtakunnallisia, sillä tällaiset järjestelmät ovat kustannustehokkaampia kuin hajautetut, paikalliset järjestelmät. Paikallisten järjestelmien käyttäjillä tulisi aina olla käytössään valtakunnallisen tason tieto kaikissa tilanteissa. [Koivu 2016]

Terveydenhuollon tietojärjestelmien hajauttaminen tai vaihtoehtoisesti niiden keskittäminen on tärkeä osa kokonaisuutta. Hajautettujen järjestelmien hyvä

puoli on se, että mahdollisten järjestelmien kaatumiskustannusten riskit minimoidaan, kun koko kokonaisuus ei ole yhden järjestelmän varassa. Toisaalta tämä vaatii, että eri järjestelmien on kommunikoitava sujuvasti keskenään. Tästä syystä voi olla perusteltua, että käytetään vain yhtä valtakunnallista järjestelmää, sillä tällöin vältetään kalliilta järjestelmäintegraation kustannuksilta ja potilaan tietoja on helpompi seurata, kävi potilas sitten terveydenhuollon palveluja tarjoavassa yksikössä missä päin Suomea tahansa [Koivu 2016]. Toisaalta haitta- puolena on, että mahdollisen järjestelmähyökkäyksen tilanteessa kaikkien suomalaisten potilastiedot ovat uhattuina.

Terveydenhuollon tietojärjestelmissä toimiva integraatio on elintärkein ta- voite, mutta valitettavasti tämän integraation määrittäminen on harvoin yksi- selitteistä. Usein integraatiolla käsitteenä voidaan tarkoittaa eri yhteyksissä eri asioitakin ja siksi olisi tehtävä selväksi, mitä sillä kussakin tilanteessa tarkoitetaan [Michelsen *et al.* 2015]. Uudistuksessa näytetään puhuvan usein integraatiosta vain palveluiden kokonaisuuden yhteensovittamisena, eikä niinkään tieto- järjestelmätason kokonaisuuden yhteensovittamisena. Toisaalta voi olla, että pal- veluiden integraatiosta puhuttaessa oletetaan sen sisältävän myös tietoteknisen puolen yhteensovittamisen, jolloin siihen ei oteta kantaa, ehkäpä koska siitä ei ole kovin syvällistä ymmärrystä.

Nykänen ja muut [2016] arvioivat, että uudistus synnyttää merkittäviä muu- tostarpeita sekä paikallisiin että alueellisiin tietojärjestelmiin. Näiden muutostar- peet voidaan määrittää terveydenhuollon organisaation strategisiksi tavoitteiksi, jotka eivät kuitenkaan saa olla irrallaan organisaation muista tavoitteista. Näiden strategisten tavoitteiden kannalta lienee tärkeää, että ne on konkreettisesti mää- ritelty ja ne osataan ilmaista ja selittää selkeästi niille toimijoille, jotka jatkossa vastaavat järjestelmiin kohdistuvista muutoksista uudistuksen täytäntöönpanon jälkeen. Toimijoilta tulee edellyttää vahvaa ohjausta, johtamista ja yhteisvas- tuuta.

Strategisten tavoitteiden saavuttaminen on uudistuksen digitalisaation kan- nalta tärkeää, sillä tällöin tietojärjestelmistä pystytään muodostamaan helpom- min sellainen toimiva kokonaisuus, jossa mahdolliset muutostarpeet voidaan valtakunnan tasolla tunnistaa. Ovaskainen [2011] toteaa, että terveydenhuol- lon alalla on tärkeää pystyä luomaan sellainen inhimillinen ilmapiiri, että muu- tostarpeista uskalletaan aktiivisesti viestiä eri organisaatioiden välillä, jotta voi- daan yhteistyössä luoda sellaisia järjestelmiä, jotka palvelevat paremmin kansa- laisten terveyden etua.

Seppälä [2014] katsoo terveydenhuollon muuttuvan perinteisestä organisaa- tiokeskeisestä mallista lähemmäksi asiakaskeskeistä mallia, jossa hoitoa pitäisi osata suunnitella entistä enemmän asiakkaan omien tarpeiden mukaan. Kansa- laisten kasvava tietoisuus ja kiinnostus hyvinvoinnistaan on johtanut siihen, että erilaisten digitaalisten toimijoiden määrä terveydenhuollon alalla on kasvanut.

Kaikille kansalaisille, erityisesti vanhemmille ikäryhmille, pitäisi siis osata tarjota entistä laadukkaampia väyliä sellaiseen toimintaan, joissa he voivat aktiivisemmin toimia osallisena digitaalista terveydenhuollon toimintaa.

Sosiaali- ja terveysministeriössä on jo keväällä 2016 kehitetty valmiiksi kokonaisarkkitehtuuria kuvaava malli uudistuksen kokonaisvaltaisista tietojärjestelmäratkaisuista. Nämä ratkaisut on esitetty kuvassa 3. Tätä tarkasteltaessa huomataan, että maakuntien vastuulle jää asiakas- ja potilastietojärjestelmät sekä niihin liittyvät yhteiset digitaaliset palvelut ja niiden kehittämistä tukevat sidoshankkeet, kuten aiemmin käsitelty Una-hanke.

Kuvasta 3 havaitaan myös, että uudistuksessa kaikki digitaaliset palvelut tulevat osaksi *kansallista palveluväylää* (national service channel), jossa pyritään mahdollisimman avoimiin rajapintoihin ja integraatoratkaisuihin. Se tarkoittaa käytännössä tiedonvälityskerrosta, joka määrittää sen, miten tietoja ja palveluja välitetään eri tietojärjestelmien välillä: sen avulla julkishallinto ja yritykset voivat käyttää hyödyksi muita siihen liittyneitä tietopalveluita ja -varantoja [Vm.fi]. Oikarinen [2013] toteaa, että tällaisissa palveluväylissä yleisesti on tärkeää kiinnittää huomioita erityisesti niiden ylläpitoon, tarjottuihin teknisiin tukipalveluihin sekä selkeään tiedonvälitykseen.

Sosiaali- ja terveydenhuollon tietojärjestelmäratkaisut



Kuva 3. Sosiaali- ja terveydenhuollon järjestelmäratkaisut [Alueuudistus.fi].

Smedberg [2015] esittää, että terveydenhuollon tietojärjestelmien oikeellinen suunnittelu ja kehittäminen vaativat ensisijaisesti asiakaslähtöisyyttä siten, että järjestelmistä vastuussa olevat organisaatiot kiinnittävät huomiota asiakkaan kokemuksiin järjestelmän hallinnasta, tehokkuudesta ja yleisestä tyytyväisyydestä.

Näkisin, että samalla tavalla edellä esitetyn tietojärjestelmien kokonaisarkkitehtuurin hyötyjä voidaan edistää myös asiakaslähtöisyydellä. Erityisesti kaikki valinnanvapautta tukevat palvelut hyötyvät merkittävästi asiakaslähtöisyydestä niiden kehittämisessä. Hallituksen kaavaileman valinnanvapausmallinkin perusidea lähtee siitä, että asiakas on se, joka rekisteröityy hänelle mieluisimman terveydenhuollon palveluja tarjoavan yksikön piiriin. Tämän päätöksen tekee siis asiakas itse, ei esimerkiksi kunta tai maakunta.

4.2. Sujuvan tiedonkulun ja valinnanvapausmallin mahdollistaminen

Sote- ja maakuntauudistuksessa tiedonkulun sujuvuutta voidaan edistää monin tavoin, mutta tärkeintä olisi saada terveydenhuollon eri tietojärjestelmät kommunikoimaan keskenään mahdollisimman sujuvasti ja niin, että siirretty tieto siirtyy sellaisenaan paikasta toiseen juuri sellaisena kuin se on tarkoitettu. Tämä on erittäin keskeistä myös hallituksen kaavailemassa valinnanvapausmallin toimivuudessa, jota käsittelemme tarkemmin tuonnempana.

Virtasen ja Stenvallin [2010] mukaan järjestelmistä on keskeistä tunnistaa se, miten paljon tietoa eri järjestelmissä on saatavilla ja kuinka helposti sitä voi pyytää hallinnollisesti korkeamman tason järjestelmistä. Eri organisaatioissa pitäisi olla tietoisia siitä, millaisilla toimintatavoilla tietoa voidaan välittää ja ottaa vastaan sekä eri järjestelmien että eri hallintotasojen, kuten paikallisten ja valtakunnallisen tasojen, välillä. Tämä on tiedonkulun onnistumisen kannalta erittäin tärkeää.

Rönkkö ja muut [2016] arvioivat, että tulevaisuuden terveydenhuollossa tietojärjestelmien osalta toimivassa tiedonkulussa tarvitaan ensisijaisesti datan oikeanlaista hyötykäyttöä ja tiedonlouhinnan osaamista. Organisaatioissa pitäisi tehdä selkeitä päätöksiä siitä, millaista dataa asiakkaista halutaan kerätä ja millä tavoin. Tämän jälkeen tulisi tarkistaa sen edellyttämät kirjaamiskäytännöt ja varmistaa, että oma tietohallinto kykenee suoriutumaan näistä käytännöistä.

Liaw ja muut [2014] pitävät kerätyn datan laatua potilaista erityisen tärkeänä, sillä mitä enemmän kerätty data vastaa todellisuutta, sitä oikeellisempia ratkaisuja voidaan potilaita koskevissa päätöksissä tehdä. Heidän mukaansa kerätyn datan laadunvarmistamiseksi siitä tulisi esittää ainakin seuraavia kysymyksiä: miksi ja mistä sitä on kerätty, kuka sitä on kerännyt, näyttääkö se järkevältä ja miten sitä on aiemmin mahdollisesti käsitelty. Ovaskainen [2011] toteaa, että terveystietojen oikeellisen keräys- ja analysointimenetelmän olevan tärkeää juuri potilaan kannalta, koska tehdyt päätökset kohdistuvat ennen kaikkea heihin. Hän myös korostaa, että potilaille tulisi aina taata mahdollisuus seurata heitä koskevia tietoja reaaliaikaisesti.

Korhonen ja Virtanen [2015] toteavat, että koska maakunnan omistama ICT-palvelukeskus tulee tuottamaan tarvittavat digitaaliset palvelut, se tulee tarvit-

semaan selkeän kokonaiskuvan siitä, miten tieto eri tietojärjestelmien välillä kulkee ja missä muodossa. Siispä uudistuksessa tarvitaan perusteellista kokonaiskuvan hahmottamista tiedonkulun osalta, jotta se saadaan erilaisten digitaalisten palvelujen välillä mahdollisimman sujuvaksi.

Aiemmin esitelty uudistukseen liittyvä valinnanvapausmalli kuulostaa hyvältä, mutta toimiakseen kunnolla se vaatii oikeanlaisia ratkaisuja tietojenkäsittelyn osalta. Kansalaisille eli asiakkaille on tarjottava helppo ja yhdenvertainen mahdollisuus verrata keskenään eri terveydenhuollon palveluja tarjoavien yksiköiden tietoja, esimerkiksi niiden tarjoamia hoitomahdollisuuksia tai hintatasoja. Asiakkaille on tarjottava täysin läpinäkyvä tieto jokaisen terveystieteiden tiedoista internetissä. Edellisessä luvussa esitelty laaturekisteri olisi saatava verkkoon viimeistään siinä vaiheessa, kun valinnanvapausmallin ensimmäinen vaihe astuu voimaan näillä näkymin vuonna 2019.

Mäki-Lohiluoma [2016] arvioi, että valinnanvapausmallin tarvitsema laaturekisteri auttaa lisäämään läpinäkyvyyttä eri toiminnan tasoille, kun terveydenhuollon palveluja tarjoavien yksiköiden tarvitsemat tiedot ovat kaikkien saatavilla. Erityisen tärkeää on saada laaturekisteri julkisesti näkyville verkkoon ilman vaikeita tunnistautumiskaavoja, jotta kuka tahansa pystyy vertailemaan terveydenhuollon palveluja tarjoavan yksikön hinta- ja laatu-tietoja keskenään. Tämä on tärkeää, jotta asiakas voi itse päättää, mikä on hänelle sopivin terveydenhuollon palveluja tarjoava yksikkö missäkin tilanteessa.

Suunnitteilla oleva valinnanvapausmalli ja sen vaikutukset ovat vielä monilta osin auki, mutta muita pohjoismaita vertailemalla voi saada tarkempaa ymmärrystä siitä, miten se tulisi vaikuttamaan asiakkaisiin. Ainakin Ruotsissa ja Tanskassa vastaavanlaisissa valinnanvapausmallissa asiakkaiden vaikutusmahdollisuudet ovat lisääntyneet sekä palvelujen saatavuus, jatkuvuus, laatu ja monimuotoisuus ovat keskimäärin parantuneet, vaikka alueellisia eroja näiden suhteen onkin havaittu. Toisaalta joidenkin asiakkaiden on ollut vaikea valita heille sopivaa palveluntuottajaa. Kuitenkin tätä valitsemisen vaikeutta on helpottanut laaturekisteri, joka ollut verkosta helposti löydettävissä ja auttanut selvittämään toiminta- ja laatueroja terveydenhuollon palveluja tarjoavien yksiköiden välillä. [Sosiaaliala.fi]

4.3. Tietojärjestelmien käytettävyyden parantaminen

Sote- ja maakuntauudistuksessa digitaalisten tavoitteiden saavuttaminen onnistuu paremmin, jos digitaalisista palveluista tehdään entistä helppokäyttöisempiä. Niitä käyttää hyvin suuri joukko erilaisia ihmisryhmiä, joten ne tulee suunnitella mahdollisimman hyvin kaikkien näiden ryhmien käyttötarpeet huomioiden. Tämä koskee sekä terveydenhuollon asiakkaita että henkilöstöä.

Viitasen ja Niemisen [2009] mukaan käytettävyyssuositusten tulisi olla lähtökohtaisesti vaihteleviin käyttötilanteisiin soveltuvia. Jo suunnitteluvaiheessa

pitäisi tunnistaa ja huomioida mahdollisimman monenlaiset käyttötilanteet tietojärjestelmien käytössä. Tätä voidaan tukea laatimalla *käyttötapauskaavioita* (use case diagrams), joiden tarkoituksena on havainnollistaa järjestelmien erilaisia käyttötilanteita kaavioilla, joissa kuvataan jonkin toimijan ja järjestelmän välistä vuorovaikutusta. Käyttötapauskaavioilla pyritään tekemään selväksi tämän vuorovaikutuksen eri vaiheet ja myös se, millaisena käyttäjä itse näkee käyttötilanteen. Niitä on helppo laatia jo suunnitteluvaiheessa, sillä ne eivät vaadi varsinaista ohjelmointityötä.

Pitkänen ja Pitkäranta [2014] ovat pohtineet käytettävyyden arviointia organisaatioiden toiminnalle parhaiten sopivien tietojärjestelmien valinnassa. He esittävät, että organisaation järjestelmähankinnan onnistuminen edellyttää tarjolla olevien järjestelmien riittävää ja oikeanlaista vertailua. Järjestelmien vertailussa niiden käytettävyyden tarkastelu tulisi tehdä näkökulmasta, jota organisaatio pitää tärkeimpänä. Organisaatiossa olisikin hyvä keskustella, mitä asioita järjestelmissä ja niiden käytettävyydessä pidetään tärkeimpinä.

Käyttäjäkokemusten parempaa lopputulosta voidaan arvioida paremmin keräämällä tutkimustietoa käytettävyydestä tutkimuksista. Näissä tutkimuksissa tulisi kiinnittää erityistä huomiota siihen, millaisilla mittareilla ja ajanjaksoilla käytettävyydestä on suoritettu sekä millaiset henkilöt tutkimuksen ovat tehneet. Tietojärjestelmien käytettävyyttä pystytään kaiken kaikkiaan kehittämään paremmin käyttäjäpalautteen avulla, jos osataan kiinnittää huomioita oikeisiin asioihin jo varhaisessa vaiheessa. [Pitkänen and Pitkäranta 2016]

Ekholm ja Kinnunen [2016] esittävät, että tietojärjestelmien käytettävyyden vaatimusmäärittelyä voitaisiin tarkastella teoreettisilla malleilla, joita voi hyödyntää mahdollisten ongelmien ehkäisemiseksi järjestelmien käyttöönottovaiheen jälkeen. Erään heidän esittämänsä mallin mukaan järjestelmien käytön vaikeuksiin voidaan vastata keskittymällä yksilötasolla järjestelmien helppokäyttöisyyteen. Se voisi tapahtua kartoittamalla tarkasti yksilöiden, mm. järjestelmiä käyttävien lääkärin, aiempia kokemuksia erilaisista järjestelmistä.

Terveydenhuollon tietojärjestelmien käytettävyyden kohdalla pätee ennen kaikkea se, että käytettävyyttä arvioitaessa tulisi tunnistaa kaikkien järjestelmää käyttävien osapuolten tarpeet ja tavoitteet. Coplien ja Bjørnvig [2010] korostavatkin erityisesti varhaisen vaiheen merkitystä siten, että jo silloin olisi hyvä ottaa kaikki järjestelmiä käyttävät sidosryhmät mukaan kehittämään yhdessä järjestelmiä: isompi joukko havaitsee virheitä helpommin ja alkuvaiheessa havaitut virheet ovat yleensä helpompi korjata kuin myöhemmässä vaiheessa havaitut. Tällä voisi ratkaista esimerkiksi Viitasen ja muiden [2011] esittämän ongelman, että

terveydenhuollon henkilöstö kokee olevansa ulkopuolella järjestelmien käytettävyyden suunnittelusta, kehittämisestä ja testauksesta.

4.4. Tietoturvariskien minimointi

Sote- ja maakuntauudistuksessa on tärkeää tarjota kansalaisille käyttöön sellaisia digitaalisia palveluita, jotka ovat tietoturvan osalta mahdollisimman luotettavia. Tämä koskee luonnollisesti sekä asiakkaita että terveydenhuollon henkilöstöä. Potilastietojen salassapito ja tietojen käyttöoikeudet tulee pitää keskiössä kaikissa tietoturvaratkaisuissa, ja erityiset sellaiset tiedot on pidettävä vahvasti salassa, joista yksittäinen henkilö voitaisiin tunnistaa [Kanta.fi].

Potilaiden henkilötietoja pitää pystyä käsittelemään turvallisesti siten, että henkilöiden oikeusturva tai yksityisyyden suoja ei vaarannu missään tilanteessa eikä ole riskiä siitä, että arkaluontoisia tietoja voisi päätyä kolmansille osapuolille. Organisaatiossa on varmistettava, että sen henkilöstö noudattaa organisaation omia tietojenkäsittelyn tietoturvaohjeita täsmällisesti ja seuraa aktiivisesti niissä tapahtuvia muutoksia. [Rousku 2014]

Kouri ja muut [2009] pitävät tärkeänä, että organisaatioissa erityisesti henkilöllisyyden todentaminen pitäisi olla mahdollisimman vahvasti suojattua. Tätä voidaan edistää erityisesti kaksivaiheisella tunnistautumisella, jossa käyttäjätili on lukittuna salasanan lisäksi jonkin muun suojauksen takana: yleisimmin käytettyjä lienevät PIN-koodit, ja suuremmissa organisaatioissa myös fyysiseen ominaisuuteen liittyvä tunnistautuminen, esimerkiksi henkilön sormenjälki tai ääni.

Uudistuksessa tietoturva- ja tietosuojaratkaisut tulevat pitkälti nojautumaan jo olemassa olevaan lainsäädäntöön, vaikka niitä uudistuksen myötä päivitetäänkin. Eräs keskeinen laki on sosiaali- ja terveydenhuollon asiakastietojen sähköistä käsittelyä koskeva laki, jolla pyritään ohjaamaan terveydenhuollon organisaatioita käyttämään vain yhtä potilastietojen arkistointijärjestelmää. Tällöin kaikkien tietojen ollessa samassa paikassa niille on helpompi luoda yhtenäinen tietoturvaratkaisu. Toinen tavoite mainitulla lailla on myös tiedonsaannin edistämisen mahdollisuudet, sillä tiedonhaku on sujuvampaa, kun ne on varastoitu yhteen arkistointijärjestelmään. [Finlex 9.2.2007/159]

Toinen keskeinen laki on laki julkisen hallinnon tietohallinnon ohjauksesta. Sen tarkoituksena on turvata julkisten tietohallintotehtävien hoitaminen käyttäen sellaisia tietoteknisiä menetelmiä ja keinoja, jotka ovat mahdollisimman turvallisia ja noudattavat tietosuojalle määriteltyjä standardeja. Julkishallinnon tietojärjestelmien yhteentoimivuuden mahdollistamiseksi ja varmistamiseksi tietohallintoa koskeva kokonaisarkkitehtuuri on suunniteltava ja kuvattava siten, että

siihen sisältyvät tietoturvaratkaisut täyttävät myös muun tietoturvalainsäädännön vaatimukset. [Finlex 10.6.2011/634]. Tämä koskee luonnollisesti myös terveydenhuollon alaa, sillä pääosa sen yksiköistä on julkisia.

Anttila ja Liimatta [2011] arvioivat, että sosiaali- ja terveysalan henkilöstön tietoturvaan ja -suojaan liittyvää taitoa ja osaamista voidaan parhaiten kehittää järjestämällä aktiivisesti verkkokoulutusta henkilöstön käyttämien tietojärjestelmien käytöstä. Erilaiset verkkokurssit ovatkin tehokas tapa järjestää koulutusta tietoturvasta ja -suojasta suurille henkilöstömäärille, sillä verkkokurssien suoritustapa mahdollistaa joustavan, ajasta ja paikasta riippumattoman, mahdollisuuden päivittää omaa osaamista.

Tietoturvataitoa päivittäviin verkkokursseihin osallistumista tulisi valvoa ja tehdä siitä osittain pakollista, jotta niistä saadaan mahdollisimman paljon hyötyä. Myös Ovaskainen [2011] toteaa, että täydennyskoulutuksen järjestäminen koko terveydenhuollon henkilöstölle on tehokas keino edistää tietoturvaosaamisen lisäksi myös laajempaa tietämystä siitä, miten tällaista osaamista voisi omaloitteisesti kehittää oman yksikön toiminnan ulkopuolella.

Rousku [2014] korostaa, että mahdollisia tietojärjestelmiin kohdistuvia tietoturmuksia varten organisaatioissa tulee laatia toimintasuunnitelma, joka takaa järjestelmien käytön vakauden myös häiriötilanteiden aikana. Mitä kriittisempi järjestelmä on, sitä perusteellisempi toimintasuunnitelman tulisi olla. Tällaisen suunnitelman tulisi aina sisältää järjestelmän toipumista edistävä prosessi, jonka avulla teknisiä häiriöitä voidaan systemaattisesti korjata ja näin palauttaa järjestelmä sen normaalitilaan. Tämä pitänee paikkansa myös potilastietojärjestelmien kohdalla: toimintasuunnitelmaa laatiessa tulisi varmistaa, että mahdollisissa häiriötilanteissa lain edellyttämät tiedot potilaista eivät vaarannu.

Uudistuksessa on huomioitava esineiden internetin tuomat mahdollisuudet niin, että tunnistetaan myös esille nousevat riskit, koska uudistuksen vaikutukset tulevat olemaan kauaskantoiset: siten on hyvä tiedostaa, millaiselta internet voi näyttää vuosien päästä. Esimerkiksi Lake ja muut [2014] arvioivat, että tulevaisuudessa esineiden internetin yleistymisen myötä tietojärjestelmien potilastietoja on osattava suojata yhä monimuotoisemmilta hyökkäyksiltä, kun terveydenhuollon alalla potilastietoja käsittelevät laitteiden määrä tulee merkittävästi kasvamaan. Erityisesti monitasoinen suojaus laitteiden välillä on tärkeää, kun kyse on arkaluonteisesta tiedosta. Epätietoisuuden vähentämiseksi potilaiden olisi myös hyvä saada tietää mahdollisimman tarkasti, millaisissa laitteissa ja mitä tarkoitusta varten heidän tietojensa käsitellään.

Näyttää siltä, että terveydenhuollon tietojärjestelmissä tapahtuva potilastietojen käsittely noudattaa melko hyvin nykyistä lainsäädäntöä. Toisaalta joissakin järjestelmissä on tietoturvan ja -suojan osalta vielä melkoisesti parantamisen varaa ja tämän merkitys tulee esineiden internetin yleistymisen myötä korostumaan. Uudistusta toteutettaessa on siis valvottava ja tarvittaessa kontrolloitava,

että nämä puutteellisetkin käytössä olevat järjestelmät saadaan noudattamaan kaikilta osin myös sitä uudistunutta tietosuojaa koskevaa lainsäädäntöä, jonka on tarkoitus astua voimaan 1.1.2019. On tietysti myös pidettävä huolta siitä, että terveydenhuollon henkilöstö noudattaa jatkossakin yleisiä tietoturvaan liittyviä ohjeita, esimerkiksi vahvojen salasanojen käyttöä ja epäilyttävien sähköpostien avaamiseen välttämistä, mahdollisimman hyvin.

5. Yhteenveto

Tässä tutkielmassa olen tarkastellut sote- ja maakuntauudistuksen sekä yleisesti terveydenhuollon alan digitalisaatiota ja sen haasteita useista eri näkökulmista. Vaihtelevat näkökulmat ovat olleet keskeinen osa uudistuksen digitalisaation todellisuuden ymmärtämistä, sillä ne sisältävät erilaisia oletuksia ilmiöistä, jotka liittyvät sekä uudistukseen että terveydenhuollon digitalisaatioon. Tarkoituksena on ollut myös avata tärkeimmät käsitteet selväpiirteisesti, jotta näitä monimutkaisia ilmiöitä on helpompi ymmärtää. Erilaisten näkökulmien löytämistä on edistetty etsimällä tietoa kattavasta kirjosta eri lähteitä. Lähdekokonaisuus on pyritty kokoamaan mahdollisimman erityyppisistä lähteistä, jotta on saatu muodostettua synteesi.

Tutkielman lähestymistapa on ollut luonteeltaan jokseenkin monitieteinen, sillä siinä on korostunut tietojenkäsittelytieteen lisäksi piirteitä myös yhteiskuntatieteiden aloista. Halusin viestiä tällä sitä, että digitalisaation haasteiden tunnistaminen ja niihin vastaaminen näin laaja-alaisen uudistuksen kontekstissa vaatii monitieteistä lähestymistapaa, sillä tietojenkäsittelytiede itsessään ei tarjoa riittävää kuvaa haasteiden kokonaisvaltaisesta tarkastelusta.

Uudistuksen digitalisaatiota on käsitelty hyvin kokonaisvaltaisesti ja yleisluontoisesti eikä ajatuksena ollut syventyä mihinkään tiettyyn osa-alueeseen kovin yksityiskohtaisesti, vaan halusin mieluummin luoda yleiskuvaa aiheesta. Jokaisesta osa-alueesta on kuitenkin pyritty luomaan oikea ja riittävä kuva, jotta niiden merkityksiä on mielekästä vertailla keskenään. Mahdollisissa jatkotutkimuksissa aiheesta voisin asettaa tutkimuksen rajat johonkin yksittäiseen osa-alueeseen ja tutkia sitä tarkemmalla ja yksityiskohtaisemmalla tasolla. Voisin esimerkiksi perehtyä terveydenhuollon alan digitalisaation strategiseen johtamiseen ja siihen liittyviin kysymyksiin.

Tekemäni kirjallisuuskatsauksen tulosten perusteella voi selvästi todeta, että uudistuksen digitaaliset tavoitteet ovat hyvin merkittäviä koko uudistuksen kannalta. Tavoitteiden ja koko terveydenhuollon digitalisaation haasteisiin on osattava vastata riittävällä osaamisella sekä asioiden välisten merkityssuhteiden laajemmalla ymmärryksellä. Tietojärjestelmien kokonaisuuden, tiedonkulun, käytettävyyden ja tietoturvan ongelmista tulee osata tunnistaa ja priorisoida

kaikkein keskeisimmät, sillä niiden ratkaisemiseksi on käytössä niukka määrä resursseja ja rajallinen määrä aikaa.

Osa esille nostamiini ongelmiin esitetyistä ratkaisuista oli asiantuntijoiden esittämiä, kun taas osa oli omaa pohdintaani, jossa perustelin ratkaisuja oman tietämykseni ja näkemysten pohjalta. Ratkaisut ovat kaiken kaikkiaan hyviä, mutta haasteena on niiden toteuttaminen käytännössä: tiedetään kyllä hyvin, mitä asioita pitäisi tehdä, mutta ei tiedetä riittävän hyvin sitä, miten näitä asioita tehdään. Vaikeutta lisää se, että ei ole olemassa kuin yksi maailmantila, jolloin ei voida tarkkaan tietää, miten paljon ongelmien ratkaisemiseen lopulta vaikutti ehdotettujen ratkaisujen toimeenpano ja miten paljon taas muut asiat. On myös tärkeää vielä mainita, että sote- ja maakuntauudistuksessahan koko Suomi asetetaan uudistuksen myötä koelaboratorioksi, jolloin voi olla vaikeaa arvioida ratkaisujen toimivuutta, kun maakuntien väliset ja jopa sisäiset erot ovat merkittäviä.

Näkisin, että nykyterveydenhuollon keskeinen haaste on pitkälti se, että se asettaa yhä suurempia odotuksia ihmisille sen suhteen, miten laadukasta terveydenhuoltoa he saavat niukkojen resurssien puitteissa. Koska nykyresursseilla on vaikea vastata kasvaviin kustannuksiin, on haettava uusia toimintamalleja hoitaa terveydenhuolto koko maassa, ja tähän koko digitalisaatio onkin erinomainen instrumentti: sitä täytyy vain osata hyödyntää oikealla tavalla. Se ei tietysti ratkaise kaikkia ongelmia, mutta näkisin sen olevan loistava keino tehostaa terveydenhuollon palveluiden järjestämistä tulevaisuudessa. Erityisesti esineiden internetissä on paljon potentiaalia alan haasteisiin vastaamisessa, ja onneksi valtion tasolla sekä kunnissakin on jo suhtauduttu myönteisesti ja lähes intohimoisesti tällaisten keinojen tarjoamiin mahdollisuuksiin.

Uudistus tulee olemaan yhteiskunnallisesti merkittävä ja sillä tulee olemaan kauaskantoisia vaikutuksia, jotka tulevat koskemaan käytännössä kaikkia kansalaisia ja lukemattomia eri organisaatioita ja muita yhteisöjä. Sitä ei ole varaa jättää tekemättä tai aloittaa kokonaan alusta, koska muutoin kansalaisten luottamus poliittiseen järjestelmään, koko kansantalous ja Suomen kyky uudistaa hallintoa kärsivät. Sen onnistunut läpivienti riippuu pitkälti avainhenkiöiden kyvyistä tehdä rohkeita ja ennakkoluulottomia päätöksiä, joiden päämääränä on uudistaa Suomea ja vastata nopeasti muuttuvan maan vaatimuksiin.

Kaiken kaikkiaan on toivottavaa, että koko uudistus suunnitellaan erittäin huolellisesti ja yksityiskohtaisesti ja tehdyistä päätöksistä todella pidetään kiinni, jotta se saadaan vietyä onnistuneesti läpi ei ainoastaan digitalisaation osalta,

vaan myös kokonaisuudessaan. Koko uudistuksen tarkoituksena on kuitenkin viime kädessä edistää ihmisten elämänlaatua.

Viiteluettelo

- Alueuudistus. <http://alueuudistus.fi/soteuudistus/digitalisaatio>. Luettu 2.12.2016.
- Johanna Andersson. 2016. *UNA-hanke – Sotetietojärjestelmäkokonaisuus palvelumuutoksen tukena*. Esitys FellmanCampuksella Lahdessa 23.5.2016. STTY.
- Lotta Anttila ja Sirpa Liimatta. 2011. Tietoturva- ja tietosuojasaaminen koko henkilöstön perustaidoksi. *FinJeHew* 3, 1, 29-32.
- James Coplien and Gertrud Bjørnvig. 2010. *Lean Architecture for Agile Software Development*. Wiley.
- Sami Ekholm ja Ulla-Mari Kinnunen 2016. Tietojärjestelmän käyttöönottoa tukevat teoreettiset mallit terveydenhuollossa. *FinJeHew* 8, 2-3, 63-73.
- EUR-Lex. <http://eur-lex.europa.eu/homepage.html>. *Law of protection of personal data 95/46/EC*.
- Finlex. <http://www.finlex.fi/fi/>. *Laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä 9.2.2007/159*.
- Finlex. <http://www.finlex.fi/fi/>. *Laki julkisen hallinnon tietohallinnon ohjauksesta 10.6.2011/634*.
- Jaakko Helenius. 2011. *Terveydenhuollon tietojärjestelmien aiheuttavat uhat potilaille ja henkilökunnalle*. Pro gradu -tutkielma. Informaatitieteiden yksikkö, Tampereen yliopisto.
- Annikki Jauhiainen, Päivi Sihvo, Helena Ikonen ja Pirjo Rytkönen. 2014. Kansalaisilla hyvät valmiudet sähköisiin terveystietoihin. *FinJeHew* 6, 2-3, 70-78.
- Johanna Kaipio, Tinja Lääveri ja Mari Tyllinen. 2015. Menettelyprosessi käytettävyyden ja loppukäyttäjänäkökulman integroimiseksi tietojärjestelmähankintaan: Tapaus Apotti. *FinJeHew* 7, 2-3, 104-121.
- Kansallinen terveysarkisto. <http://www.kanta.fi/omakanta>. Luettu 2.12.2016.
- Annariina Koivu. 2016. *Tietotekniikka ja tietojärjestelmät terveydenhuollossa – Esimerkkejä kehittyneistä ja kehittyvistä maista*. Esitelmä Tampereen yliopistolla 6.4.2016.
- Maritta Korhonen. 2016. *Sote-uudistus ja digitalisaatio – mikä muuttuu tietohallinnossa, tietojärjestelmissä ja ICT-palveluiden tuottamisessa?* Esitelmä Sosiaali- ja terveysministeriön tilaisuudessa Helsingissä 27.1.2016.
- Maritta Korhonen ja Teemupekka Virtanen. 2015. Digitaalisuus ja asiakaslähtöisyys sosiaali- ja terveydenhuollossa – kansalaisen omat tiedot hyötykäyttöön. *FinJeHew* 7, 4, 237-239.
- Pirkko Kouri, Kristiina Junntila, Kaija Saranto and Anneli Ensio. 2009. Personal health information management around the world - Finland. In: *Proc. of the*

Post-Congress Workshop of the 10th International Nursing Informatics Congress, 152-157.

- Tuija Kuusisto ja Pekka Kantola. 2016. *ICT-palvelukeskusselvitys sosiaali- ja terveyshuollon uudistuksen näkökulmasta: Selvityshenkilöiden loppuraportti 2016-54.* Valtioneuvosto, Sosiaali- ja terveysministeriö.
- Kari-Pekka Mäki-Lohiluoma. 2016. *Valinnanvapaus terveydenhuollossa – Mihin suuntaan Suomen-mallia kehitetään?* Esitelmä Kelan auditoriossa Kansaneläkelaitoksen päätoimitalolla 17.5.2016. KELA.
- Maija Laiho. 2011. *Tietokone ikäihmisen arjessa.* Lisensiaatintutkimus. Valtiotieteellinen tiedekunta, Helsingin yliopisto.
- David Lake, Rodolfo Milito, Monique Morrow and Rajesh Vargheese. 2014. Internet of Things: architectural framework for eHealth security. *JICT* 3, 4, 301-328.
- Siaw-Teng Liaw, Christopher Pearce, Harshana Liyanage, Gladys SS Cheah-Liaw and Simon de Lusignan. 2014. An integrated organisation-wide data quality management and information governance framework: theoretical underpinnings. *BCS Health* 21, 4, 199-206.
- Kai Michelsen, Helmut Brand, Peter Achterberg and John Wilkinson. 2015. *Promoting better integration of health information systems: best practices and challenges.* Synthesis Report-2015. WHO Regional Office for Europe, HEN.
- Sari Nissinen. 2010. Terveydenhuoltohenkilöstön kokemuksia aluetietojärjestelmän käytettävyydestä perusterveydenhuollossa. *FinJeHew* 2, 3, 101-106.
- Pirkko Nykänen. 2016. Health care documentation - could we integrate medical doctors and nurses documentation? *International Journal on Biomedicine and Healthcare* 4, 1, 50-51.
- Pirkko Nykänen, Mari Tyllinen, Tinja Lääveri, Antto Seppälä, Johanna Kaipio ja Marko Nieminen. 2016. *Ekosysteemi ja menetelmällinen ohjeisto terveydenhuollon tietojärjestelmäpalvelun hankintaan.* Raportti 45/2016. Tampereen yliopisto, Informaatiotieteiden yksikkö.
- Tommi Oikarinen. 2013. *Julkisen hallinnon kokonaisarkkitehtuuri – Tilannekatsaus.* Esitelmä Valtionvarainministeriön tilaisuudessa Helsingissä 18.4.2013.
- Tommi Oikarinen ja Antti Larsio. 2016. *Maakunta- ja sote-uudistuksen tietohallinnon ohjauksen selvityshanke - ohjausmalliehdotus.* Raportti 35/2016. Valtioneuvosto, Sosiaali- ja terveysministeriö.
- Päivi Ovaskainen. 2011. Tieto-ohjaus terveydenhuollossa. Teoksessa: Mika Laakkonen, Suvi Lamminpää ja Jarno Malaprade (toim.), *Informaatioteknologian filosofia*, 309-327. Lapin yliopistokustannus.
- Janne Pitkänen ja Matti Pitkäranta. 2014. Käytettävyyden arviointi ja käytettävyydestä soveltaminen terveydenhuollon tietojärjestelmien valinnassa. *FinJeHew* 6, 2-3, 103-110.

- Janne Pitkänen and Matti Pitkäranta. 2016. Improving meaningful use and user experience of healthcare information systems towards better clinical outcomes. *FinJeHew* 8, 2-3, 98-106.
- Kimmo Rousku. 2014. *Kyberturvaopas – Tietoturva kotona ja työpaikalla*. Talentum.
- Pekka Ruotsalainen, Pirkko Nykänen, Antto Seppälä and Bernd Blobel. 2014. Truth-based information system architecture for personal wellness. In: *Proc. of MIE2014.: e-Health – For Continuity of Care*.
- Iлона Rönkkö, Ulla-Mari Kinnunen, Sami Kiviharju ja Risto Mäkinen. 2016. Potilastiedot hyötykäyttöön perusterveydenhuollossa – tarvitaan kysymisen taitoa, dataa ja tiedonlouhinnan osaamista. *FinJeHew* 8, 1, 14-29.
- Antto Seppälä. 2014. *Context-aware and trust-based personal wellness information framework for pervasive health*. Ph. D. Dissertation, School of Information Sciences, University of Tampere.
- Jari Smedberg. 2014. *Asiakslähtöisyyden merkitys tietojärjestelmien kehittämisessä – tarkastelu sosiaalihuollon kokonaisarkkitehtuurin kontekstissa*. Pro gradu - tutkielma. Informaatitieteiden yksikkö, Tampereen yliopisto.
- quality management and information governance framework: theoretical underpinnings. *BCS Health* 21, 4, 199-206.
- Kai Michelsen, Helmut Brand, Peter Achterberg and John Wilkinson. 2015. *Promoting better integration of health information systems: best practices and challenges*. Synthesis Report-2015. WHO Regional Office for Europe, HEN.
- Sari Nissinen. 2010. Terveystietojärjestelmien käytettävyydestä perusterveydenhuollossa. *FinJeHew* 2, 3, 101-106.
- Pirkko Nykänen. 2016. Health care documentation - could we integrate medical doctors and nurses documentation? *International Journal on Biomedicine and Healthcare* 4, 1, 50-51.
- Pirkko Nykänen, Mari Tyllinen, Tinja Lääveri, Antto Seppälä, Johanna Kaipio ja Marko Nieminen. 2016. *Ekosysteemi ja menetelmällinen ohjeisto terveydenhuollon tietojärjestelmäpalvelun hankintaan*. Raportti 45/2016. Tampereen yliopisto, Informaatitieteiden yksikkö.
- Tommi Oikarinen. 2013. *Julkisen hallinnon kokonaisarkkitehtuuri – Tilannekatsaus*. Esitelmä Valtionvarainministeriön tilaisuudessa Helsingissä 18.4.2013.
- Tommi Oikarinen ja Antti Larsio. 2016. *Maakunta- ja sote-uudistuksen tietohallinnon ohjauksen selvityshanke - ohjausmalliehdotus*. Raportti 35/2016. Valtioneuvosto, Sosiaali- ja terveysministeriö.
- Päivi Ovaskainen. 2011. Tieto-ohjaus terveydenhuollossa. Teoksessa: Mika Laakonen, Suvi Lamminpää ja Jarno Malaprade (toim.), *Informaatioteknologian filosofia*, 309-327. Lapin yliopistokustannus.
- Janne Pitkänen ja Matti Pitkäranta. 2014. Käytettävyyden arviointi ja käytettävyydestä soveltaminen terveydenhuollon tietojärjestelmien valinnassa. *FinJeHew* 6, 2-3, 103-110.

- Janne Pitkänen and Matti Pitkäranta. 2016. Improving meaningful use and user experience of healthcare information systems towards better clinical outcomes. *FinJeHew* 8, 2-3, 98-106.
- Kimmo Rousku. 2014. *Kyberturvaopas – Tietoturva kotona ja työpaikalla*. Talentum.
- Pekka Ruotsalainen, Pirkko Nykänen, Antto Seppälä and Bernd Blobel. 2014. Truth-based information system architecture for personal wellness. In: *Proc. of MIE2014.: e-Health – For Continuity of Care*.
- Iлона Rönkkö, Ulla-Mari Kinnunen, Sami Kiviharju ja Risto Mäkinen. 2016. Potilastiedot hyötykäyttöön perusterveydenhuollossa – tarvitaan kysymisen taitoa, dataa ja tiedonlouhinnan osaamista. *FinJeHew* 8, 1, 14-29.
- Antto Seppälä. 2014. *Context-aware and trust-based personal wellness information framework for pervasive health*. Ph. D. Dissertation, School of Information Sciences, University of Tampere.
- Jari Smedberg. 2014. *Asiakaslähtöisyyden merkitys tietojärjestelmien kehittämisessä – tarkastelu sosiaalihuollon kokonaisarkkitehtuurin kontekstissa*. Pro gradu -tutkielma. Informaatiotieteiden yksikkö, Tampereen yliopisto.
- Jari Smedberg. 2015. Käsitteellinen tarkastelu asiakaslähtöisyydestä tietojärjestelmien kehittämisessä. *FinJeHew* 7, 2-3, 149-157.
- Ian Sommerville. 2007. *Software Engineering*. Personal Education M.U.A.
- Sosiaaliala. <http://www.sosiaaliala.fi/>. *Kauppakamarin selvitys valinnanvapausjärjestelmistä Ruotissa ja Tanskassa*.
- Sosiaali- ja terveysministeriö. <http://stm.fi/etusivu>. *Alustavat ehdotukset sote- ja maakuntauudistuksen lainsäädännöksi - Tiivistelmät lakiluonnosten keskeisistä asioista*.
- Anna-Elina Ukkonen. 2016. Sana on vapaa - narratiivisen tekstin rooli sähköisissä potilastietojärjestelmissä. *FinJeHew* 8, 2-3, 120-126.
- Valtioneuvosto. <http://valtioneuvosto.fi/etusivu>. *Ratkaisujen Suomi: Pääministeri Juha Sipilän hallituksen strateginen hallitusohjelma 29.5.2015*.
- Valtiovarainministeriö. <http://vm.fi/julkisen-hallinnon-ict>. Luettu 2.12.2016.
- Sari Vesiluoma. 2012. Strategiaskenaarioita tiedolla johtamisen linjauksiin. *FinJeHew* 4, 2, 116-123.
- Johanna Viitanen, Susanna Martikainen, Mikko Korpela ja Tinja Lääveri. 2011. *Lääkärien kokemuksia ja näkemyksiä terveydenhuollon tietojärjestelmien kehittämisestä*. Raportti 13/2011. Aalto- yliopiston teknillinen korkeakoulu, Strategisen käytettävyyden tutkimusryhmä.
- Johanna Viitanen ja Marko Nieminen. 2009. Terveydenhuollon tietojärjestelmien käytettävyys. *FinJeHew* 1, 3, 130-136.
- Petri Virtanen ja Jari Stenvall. 2010. *Julkinen johtaminen*. Tietosanoma.

Seppo Visala. 2011. Miten informaatioteknologia muuttaa olemustamme? Teoksessa: Mika Laakkonen, Suvi Lamminpää ja Jarno Malaprade (toim.), *Informaatioteknologian filosofia*, 141-160. Lapin yliopistokustannus.