



Tietojohtamisen pilotti SoteDigi Oy – Kainuun sote loppuraportti

14.11.2019

Jaakko Pentti, Elina Välikangas, Pekka Uotinen, Pirkka-Matti Alanne,
Tero-Matti Kinanen, Kari Natunen, Jari Pekkanen, Elina Sarpola, Marja
Valkonen, Olli Nylander, Marita Pikkarainen ja Pekka Stigell



13.12.2019

Sisällys

1.	Tiivistelmä.....	4
2.	Johdanto	5
2.1	Projektin tausta	5
2.2	Projektin tavoitteet	5
2.3	Projektin tuotokset.....	7
3.	Projektin työvaiheet ja hallinta	9
3.1	Projektin organisointi	9
3.2	Projektin vaiheistus	9
3.3	Projektin aikataulu.....	10
3.4	Projektin resurssit.....	11
3.5	Projektin työskentelytapa.....	12
3.6	Tulosten laadun varmistaminen	12
3.7	Projektin tavoitteiden toteutuminen	13
3.8	Projektin riskit ja niiden toteutumisen arviointi.....	14
3.9	Projektin työmenetelmien arviointi	15
4.	Projektin sisältö ja havainnot	20
4.1	Palveluluokitukset	20
4.2	Käsittemallinnus.....	21
4.3	Tietolähteet ja tiedonsiirto.....	24
4.4	Kustannuslaskenta.....	26
4.4.1	Tavoitteet	26
4.4.2	Prosessi.....	26
4.4.3	Tuloksia.....	31
4.4.4	Havainnot.....	34
4.4.5	Kustannuslaskennan vaihtoehtoisia malleja	36
4.5	Kehittynyt analytiikka ja koneoppiminen	37
4.5.1	Ongelman ja ratkaisun täsmentäminen	37
4.5.2	Koneoppimisen canvas – ajantasainen kokonaiskuva.....	39
4.5.3	Tavoite – päätetapahtuma	40
4.5.4	Tietosisällöt, rajaukset ja havainnot datasta	40
4.5.5	Johdetut muuttujat.....	42
4.5.6	Mallinnus	42

13.12.2019

4.6	Tiedon raportointi ja visualisointi.....	43
4.7	Sidosryhmätyö – yhteistyö eri tahojen kanssa	46
5.	Teknologia	48
5.1	Yleistä	48
5.2	Välineet.....	49
5.3	Tietokannan rakenne.....	50
5.3.1	Tietokannat ja skeemat	50
5.3.2	Data Vault-toteutus	50
5.3.3	Data Martit	51
5.4	Kokemukset eri välineistä.....	53
5.4.1	Mallinnus	53
5.4.2	Tietovarastoautomaatio	55
5.4.3	Tietokanta.....	55
5.4.4	Raportointi ja visualisointi	56
5.4.5	Kustannuslaskenta.....	57
5.4.6	Edistynyt analytiikka, koneoppiminen ja tekoäly	57
5.5	Kokemus Kainuun ympäristöstä	58
5.6	Suorituskykytietoja	58
5.6.1	SQL Server ja Snowflake	59
5.6.2	Analytiikkapalvelin.....	60
6.	Tietoturva ja tietosuoja	61
6.1	Sopimukset	61
6.2	Laki sosiaali- ja terveystietojen toissijaisesta käytöstä eli toisiolaki	61
6.3	Teknologia	65
6.3.1	Pseudonymisointi	66
6.3.2	Anonymisointi.....	66
6.3.3	Jäännösriskin arviointi	67
6.4	Sisällön rajaus	68
7.	Tulokset	69
7.1	Tuloksellisuuden näkökulmat.....	69
7.2	Kehittynyt analytiikka ja koneoppiminen	72
8.	Kokemukset ja suositukset	77
8.1	Yleisiä havaintoja	77

13.12.2019

8.2	Kansallisille toimijoille	78
8.3	Sote-palvelujen järjestäjälle	79
8.4	Kainuun soten kokemukset	79
Lähteet.....		82
Liitteet.....		83
LIITE 1: Projektiin osallistuneet Kainuun soten henkilöt		84
LIITE 2: Käsittemallit.....		85

13.12.2019

1. Tiivistelmä

SoteDigi Oy toteutti Kainuun sosiaali- ja terveydenhuollon kuntayhtymän, jäljempänä Kainuun soten, kanssa yhteistyössä sote-järjestäjän tietojohdamisen pilottiprojektin touko-syyskuussa 2019. Projektin lähtökohtana oli 26.4.2019 hyväksytyn Sosiaali- ja terveystietojen toisiokäytön lain eli toisiolain ja sen tietojohdamisen pykälän soveltaminen ja testaaminen.

Projektin tietosisällöksi rajattiin yli 75-vuotiaiden asukkaiden sosiaali- ja terveystietopalvelut. Tavoitteena oli tuoda näkyväksi näiden palveluiden toiminta ja kustannukset organisaatioriippumattomasti sekä integraation toteutuminen palveluittain, alueellisesti ja kuntatasolla. Toisaalta tavoitteena oli testata tietojohdamisen kehitysprosessia ja välineistöä sekä tuottaa tietosisällöllisesti ja teknisesti mahdollisimman geneerinen malli, joka on toistettavissa kansallisesti. Samalla testattiin toisiolain antamia mahdollisuuksia ja toteutettiin tietoturvan ja -suojaan edellyttämät ratkaisut voimassa olevan ohjeistuksen mukaisesti. Lisäksi toteutettiin koneoppimisen keinoin ennustemalli yli 75-vuotiaiden säännöllisiin palveluihin ohjautumiseen.

Projekti toteutettiin tiiviissä yhteistyössä ja vuorovaikutuksessa Kainuun soten henkilöstön ja SoteDigi Oy:n sekä valittujen alihankkijoiden välillä noudattaen ketterän kehityksen periaatteita. Projektin tavoiteasetanta lähti Kainuun soten strategisista tavoitteista ja projektin aikana pyrittiin jatkuvalla seurannalla varmistamaan, että lopputuloksena syntyy juuri sitä, mihin kysymyksenasettelulla haluttiin vastauksia.

Teknisesti projekti toteutettiin paikallisena on-premise -ratkaisuna hyödyntäen Kainuun soten IT-infrastruktuuria. Toteutettu alustaratkaisu ei edellyttänyt investointeja ja projektissa käytetty välineistö hankittiin koekäyttöön ilman erillisiä lisenssikustannuksia. Lisäksi projektissa testattiin pilvikomponenttina tietokantaa, minne siirrettiin anonymisoitua tietoa testaustarkoituksessa.

Projektin aikana tuotettiin pilotin tietosisältöön rajautuvat käsittemallit. Käsittemallien yhteyteen tallennettiin käsitteiden kuvaus- ja lähdetiedot. Käsitteiden kuvauksissa hyödynnettiin kaikilta osin kansallisia määrittelyjä ja lähteitä, kun niitä oli saatavilla. Käsittemalleista generoitiin loogisia tietomalleja tietovarastototeutuksen käyttöön.

Projektin lopputuloksena tuotettiin kokonaisnäkymä yli 75-vuotiaiden kainuulaisten sote-palvelujen käytöstä ja palvelukäyttöön liittyvistä kustannuksista yli toimiala- ja kuntarajojen. Palvelukäyttö- ja kustannustietoon lisättiin asiakkaiden toimintakyky- ja hoitoisuustietoa. Palveluittain pystyttiin kuvaamaan kustannuksia per asiakas ja vertailemaan eri asiakassegmenttien palvelukäyttöä ja toteutuneita kustannuksia toisiinsa palveluittain. Kuntien välistä kustannusvertailua tehtiin asukaskohtaisella tarkastelulla. Toteutettu koneoppimisen malli tunnisti 86 prosentin tarkkuudella henkilöt, jotka ovat riskissä ohjautua 12 kuukauden kuluttua säännöllisiin palveluihin.

Kaiken kaikkiaan projektissa pystyttiin toteuttamaan lyhyessä ajassa ne tietosisällöt, mitkä tavoitteenasetannassa määriteltiin. Lopputulokset ovat sellaisenaan hyödynnettävissä päätöksenteossa Kainuun soten järjestämistehtävässä. Lisäksi toteutettiin aiotussa laajuudessa tietojohdamisen kehityshankeprojektin metodiikan ja välineistön testausta ja saatiin arvokasta tietoa tarvittavien kyvykkyyksien osalta sekä asiakkaan että toimittajan näkökulmasta. Pilottiprojektin dokumentaatiota tullaan hyödyntämään seuraavien toteutusten suunnittelussa ja toteutuksessa, mikä osaltaan mahdollistaa kehityshankeprosessin tehostamisen ja läpimenoaikojen lyhentämisen jatkossa.

13.12.2019

2. Johdanto

2.1 Projektin tausta

Sote-järjestäjän tietomallityö käynnistyi keväällä 2018 ja valmistui maakuntien yhteistyön tuloksena kesäkuussa 2018. Samanaikaisesti toteutettiin myös kolme muuta sosiaali- ja terveysministeriön (STM) suunnitteluhanketta: valtakunnallisten toimijoiden yhteinen kokonaisarkkitehtuuri, tietojärjestelmien esiselvitysprojekti ja digitaalisten palvelujen arkkitehtuuriprojekti. Lisäksi taustalla on pitkäjänteistä kehitystyötä, mm. STM:n Sote-tieto hyötykäyttöön -strategia sekä useita kansallisesti toteutettuja tietojohdamisen hankkeita kuten:

- Terveyden ja hyvinvoinnin laitoksen (THL) tietopohjahanke
- Sote-KUVA -muutosohjelma
- STM:n Sote-kokonaisarkkitehtuurityö
- Sitran ja maakuntien yhteistyössä kehittämä STM:n omistajuuteen siirtynyt Sote-tietopakettityö

Sote-järjestämisen tietomallityö siirtyi maakuntien valmistelun jälkeen SoteDigi Oy:lle, jossa käynnistettiin jatkotyön hankkeistaminen. Tavoitteena oli tehdä maakuntien kanssa yhteistyössä tietojohdamisen maakunnallista valmistelua tähdäten kansallisesti yhtenäiseen lopputulokseen. Valmistelun alkuvaiheessa, syksyllä 2018 toteutettiin SoteDigin ja Vimaan yhteinen, kaikkiin maakuntiin kohdistunut maakuntakierros. Kierroksella kuultiin maakuntien tilannekuvaa ja tarpeita jatkokehitystyöhön. Taustatyön, tilannekuvan ja tavoitteiden pohjalta laadittiin "Virta"-hankesuunnitelma. Maakunta- ja soteuudistuksen valmistelun kariuduttua päädyttiin toteuttamaan hankesuunnitelman mukainen pilotti.

Pilottiprojekti perustuu sote-palvelujen järjestäjän tietotarpeisiin ja sen yhtenä perustana on 26.4.2019 hyväksytty Laki sosiaali- ja terveystietojen toissijaisesta käytöstä (552/2019) [1]. Lain osalta projektissa sovelletaan tietojohdamisen pykälää §41. Tämän pohjalta pilottiin haettiin sellaista toimijaa, jolla on sekä perusterveydenhuollon, sosiaalitoimen että erikoissairaanhoidon rekisterinpitäjyys samalla toimijalla. Tällaisia ovat maakunnalliset kuntayhtymät, joita ovat Etelä-Karjalan sosiaali- ja terveyspiiri (Eksote), Kainuun sosiaali- ja terveydenhuollon kuntayhtymä (Kainuun sote), Kymenlaakson sosiaali- ja terveyspalvelujen kuntayhtymä (Kymsote), Päijät-Hämeen hyvinvointikuntayhtymä (PHHYKY), Pohjois-Karjalan sosiaali- ja terveyspalvelujen kuntayhtymä (Siunsote) ja Keski-Pohjanmaan sosiaali- ja terveyspalvelukuntayhtymä (Soite). Eri osapuolten kanssa käytyjen keskustelujen pohjalta päädyttiin toteuttamaan pilotti Kainuun maakuntalaajuisen, yhden rekisterinpitäjän alaisuudessa toimivan sote-kuntayhtymän kanssa. Tässä raportissa kuvattu pilotti on siis osa SoteDigin tietojohdamisen kokonaisuutta (Virta-hanke), jossa lähtökohtana järjestäjän tietojohdamisen tarpeet (järjestäjän tietomalli).

2.2 Projektin tavoitteet

Pilottiprojektin tavoitteena oli testata Virta-hankkeen toimintatapoja, menetelmiä ja malleja sekä selvittää tietoteknologisen arkkitehtuurin parhaat käytännöt tämän tyyppisiin hankkeisiin. Hankkeessa testattiin keväällä 2019 voimaantullutta lakia sosiaali- ja terveystietojen toissijaisesta käytöstä. Edelleen pilotin aikana oli tavoitteena valmistella järjestämistehtävässä tarvittavia tietosisältöjä ja testata niihin liittyen palveluluokittelua sekä käsitemallien ja loogisen tietomallin valmistelua siten, että ne täyttävät kansalliset vaatimukset ja toisaalta vaikuttavat tietojohdamisen kautta myös kohdeorganisaation toimintaan ja kehittävät sitä eteenpäin.

13.12.2019

Kysymyksiä, joihin Virta-hankkeen avulla pyrittiin saamaan vastauksia, olivat hankkeen laajuuteen liittyvät asiat kuten: miten pilkkoa ja jäsentää laajan toteutuksen eri osa-alueet ja edetä siivu kerrallaan systemaattisesti, kuinka tekeminen saadaan fokuksittua alueiden kannalta tärkeimpiin asioihin ja kuinka toteutusta saadaan tehtyä tehokkaasti ja nopeasti. Teknologisen arkkitehtuurin kannalta hankkeen valmistelua varten oli oleellista selvittää, millainen teknologinen arkkitehtuuri soveltuu toteutukseen eli kuinka saadaan pääosa tekemisestä kohdistettua tietojen hyödyntämiseen niiden integroinnin ja siirron sijasta.

Valtakunnan laajuisessa kehityksessä on oleellista huomioida, että saadaan kansallinen kehitys ja kaikkien alueiden edustus mukaan toteutukseen, vaikka ollaankin toimimassa kulloinkin yksittäisellä alueella joko asiakkaan tai aihealueen puolesta. Asiakkaan toimintaan vaikuttavan järjestelmän tekeminen vaatii lisäksi, että asiakkaan organisaatio ja eri sidosryhmät saadaan hyvin mukaan tekemiseen, jolloin tulosten käyttäminen mahdollistuu ja saadaan tarvittavat näkökulmat mukaan kehitykseen.

Näistä laajoista hankkeen valmisteluun liittyvistä tavoitteista tiivistettiin pilottiprojektin tavoitteet erittäin suoraviivaiseksi toteutukseksi kuitenkin laajalla ja joustavalla tietosisällöllä.

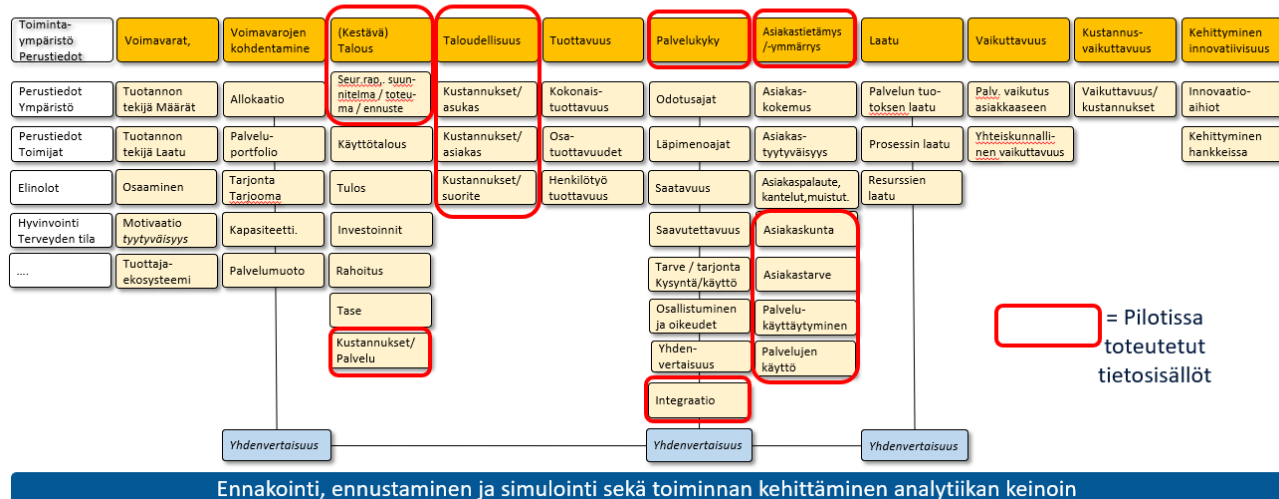
Pilottiprojektin tavoitteet:

- Tuottaa tietosisältö, joka tukee Kainuun soten johtamista tuomalla tietoa päätöksenteon tueksi eri tasoille organisaatiota ja vastaa kansallisia vaatimuksia
- Kehittää toimintatapa SoteDigin tietojohdamisen projektien läpivientiin
- Testata tietojohdamisen arkkitehtuurin teknologisia komponentteja
- Kehittää toisiokäytön lain [1] hyödyntämisen käytäntöjä ja malleja

Taustalla tietosisältöjen valmistelussa oli kansallisessa Sote-järjestäjän tietomallityössä määritetyt tuloksellisuuden näkökulmat ja edelleen maakuntakierroksella saadut alueiden priorisoinnit tietotarpeista (kuva 1). Lisäksi kartoitettiin Kainuun tilannetta ja heidän tarpeitaan. Aihealueeksi valikoitui ikääntyneiden palvelut ja sieltä näkökulmina erityisesti talous ja taloudellisuus sekä asiakkaiden palveluihin liittyvän integraation toteutuminen.

13.12.2019

Palvelujen järjestäjän kokonaisnäkymä ja tuloksellisuuden tietotarpeet – Sote- järjestämisen tietomalli



Kuva 1 Sote-järjestäjän tietomalli

Projektin sisällöllisiä tavoitteita täsmennettiin hypoteesicanvaksen avulla lähtien Kainuun sote kuntayhtymän strategiasta ja sen tavoitteista. Tavoitteena projektilla oli tuottaa tietojohtamisen toteutus, jonka avulla mahdollistetaan organisaation tukeminen tietojohtamisen keinoin niin, että Kainuun soten alueella: "Osallistuva ikäihminen asuu omassa kodissaan lähiverkostonsa sekä tarpeen mukaisten palvelujen turvin".

Työtä ohjaamaan luotiin strategista tavoitetta tukevia hypoteeseja, joista valittiin kolme varsinaisen tietotuotteen kehittämistä ohjaaviksi hypoteeseiksi:

- Yli 75-vuotiaiden muiden kuin säännöllisen kotihoidon asiakkailla sotepalveluiden käyttö on säännöllisen kotihoidon asiakkaiden palvelukäyttöä runsaampaa
- Päivystyksestä ohjaututaan usein perus- tai erikoistason sairaalapalveluihin
- Palvelutarpeiden ennakointi ja riittävä resurssointi varhaisessa vaiheessa hillitsevät kustannuskehitystä

2.3 Projektin tuotokset

Pilottiprojektissa tuotettiin seuraavat lopputuotokset:

- Virta-hankkeeseen kuuluva yhteisesti rajattu tietosisältö, joka hyödyntää Kainuun soten lähdejärjestelmiä
- Tietosisällön määrittely, käsittemallit ja loogiset tietomallit
- Integraatiot lähdejärjestelmiin hyödyntäen olemassa olevia rajapintoja
- Tietojen lataus ja talletus tietoaalustaan
- Tietojen jalostus mittaritiedoksi ja tiedon hyödyntämisen mahdollistaminen

13.12.2019

- Projektin toteutusta ja sen eri vaiheita kuvaava dokumentaatio
- Loppuraportti eli tämä dokumentti

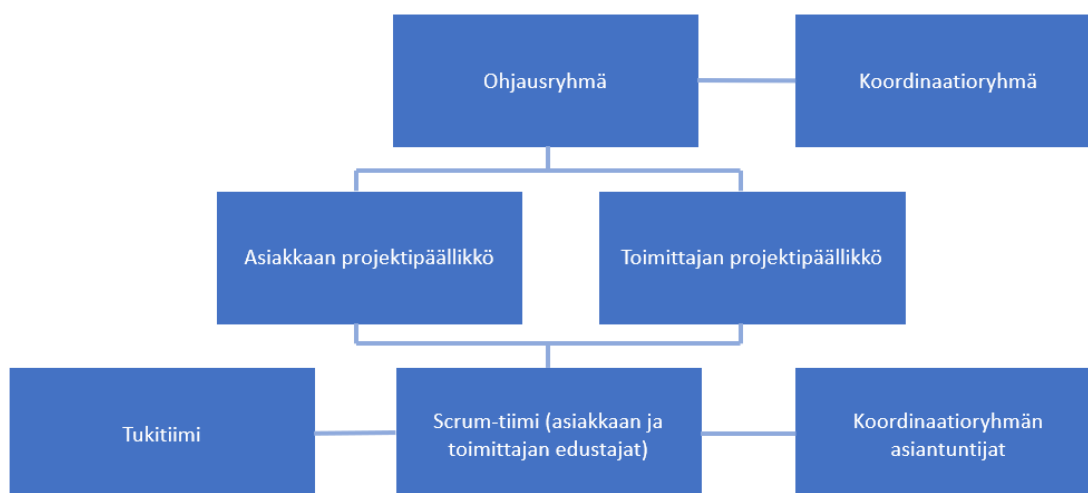
13.12.2019

3. Projektin työvaiheet ja hallinta

Teknisen toteutuksen näkökulmasta pilottiprojekti toteutettiin täysin uutena, puhtaalta pöydältä lähtevänä toteutuksena Kainuun soten tietotekniseen ympäristöön. Tietosisältöjen osalta työtä jatkettiin aiemmin valmistellun sote-järjestämisen tietomallityön pohjalta.

3.1 Projektin organisointi

Projektin osapuolina ovat olleet Kainuun sote ja SoteDigi Oy. SoteDigin alihankkijoina toimivat Vimana Oy, Citrus Solutions Oy ja Prodacapo Finland Oy. Citrus Solutions Oy:n alihankkijana on toiminut Productivity Leap Oy. Kuvassa 2 on esitetty projektin organisaatio.



Kuva 2 Projektin organisaatio

Projektin ohjausryhmässä ovat olleet Maire Ahopelto, Marita Pikkarainen, Eija Tolonen, Ville Turunen ja Kristiina Taskinen Kainuun sotesta ja Jaakko Pentti, Elina Välikangas, Pirkka-Matti Alanne ja Kari Natunen SoteDigi Oy:stä sekä Mikko Huovila STM:stä. Asiakkaan projektipäällikkönä ovat toimineet Kristiina Taskinen ja Niina Komulainen Kainuun sotesta sekä toimittajan projektipäällikkönä Kari Natunen. Asiakkaan puolelta on projektiin osallistunut 34 toiminnan, talouden ja ICT:n asiantuntijaa. Projektin ketterän kehityksen scrum-tiimissä ovat toimineet Kainuun soten, SoteDigin, Vimanan ja kumppaneiden asiantuntijat. Projektin koordinaatioryhmään ovat kuuluneet muut maakunnalliset kuntayhtymät Eksote, Kymsote, PHHYKY, Siun sote ja Soite, sekä STM ja THL. Koordinaatioryhmän jäsenten rooleja on tarkemmin käyty läpi luvussa 4.8 Sidosryhmätyö. Projektin tukitiimissä ovat toimineet ICT-välineistöjen teknologiatoimittajat.

3.2 Projektin vaiheistus

Projekti jaettiin seuraaviin vaiheisiin:

- Pilottiprojektin käynnistys ja tekninen sprintti

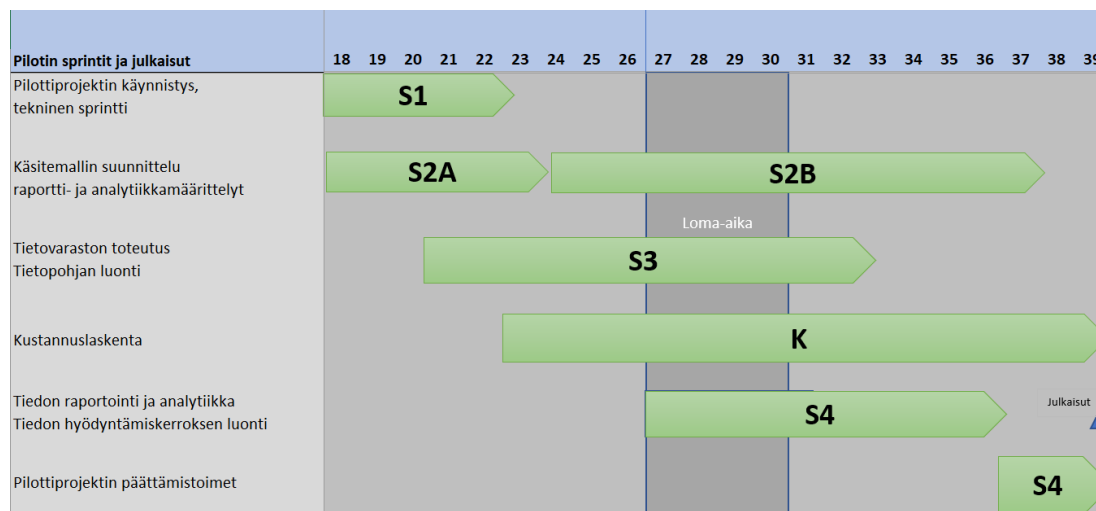
13.12.2019

- Tietotarpeiden ja tietosisältöjen määrittely: sote-järjestäjän tietomallin ja hypoteesicanvaksen kautta rajattiin projektin scope ja toteutukseen liittyvät käsitteet
- Käsittemallinnus ja raportti- ja analytiikkamäärittelyt
- Tietovaraston toteutus ja tietopohjan luonti: automaattisen ETL/ELT-prosessin kautta sekä pseudonymisoitu tietovarasto, että sen pohjalta anonymisoitu raportointikanta
- Tiedon raportointi ja analytiikka ja tiedon hyödyntämiskerroksen toteutus
- Koneoppimisen mallin toteutus
- Tietojen validointi toteutettujen visualisaatioiden avulla
- Pilottiprojektin päättämistoimet

Taustalla tietosisältöjen valmistelussa oli kansallisessa Sote-järjestäjän tietomallityössä määritetyt tuloksellisuuden näkökulmat ja edelleen maakuntakierroksella saadut alueiden priorisoinnit tietotarpeista. Lisäksi kartoitettiin Kainuun tilannetta ja heidän tarpeitaan. Aihealueeksi valikoitui ikääntyneiden palvelut ja sieltä näkökulmina erityisesti talous ja taloudellisuus sekä asiakkaiden palveluihin liittyvän integraation toteutuminen.

3.3 Projektin aikataulu

Projekti käynnistettiin 2.5.2019 ja päätettiin 23.10.2019 Kajaanissa pidettyyn projektin loppuseminaariin. Aktiivinen toteutusvaihe tapahtui aikavälillä 2.5.2019 – 30.9.2019. Projektin Gantt-kaavio esitetty kuvassa 3.



Kuva 3 Projektin Gantt-kaavio.

Karkealla tasolla projekti eteni kuukausittain seuraavasti:

- Toukokuussa asennettiin projektin tietotekninen ympäristö ja kuvattiin sekä hypoteesicanvas että käsittemalli.
- Kesäkuussa toteutettiin tietojen lataaminen tietovarastoon ja määriteltiin raportointi ja kuvattiin kehittyneen analytiikan canvas.
- Heinäkuussa tuotettiin tiedot raportoinnin ja kehittyneen analytiikan toteutuksia varten
- Elokuussa aloitettiin raportoinnin kokonaisuuden toteutus, joka sisältää pääraportoinnin lisäksi kustannuslaskennan, episodianalytiikan ja koneoppimisen mallin tulokset.

13.12.2019

- Syyskuussa laajennettiin projektin toteutusta, lisättiin tietosisältöä ja tehtiin uusia toteutuksia sekä tuotettiin loppuraportti toteutetusta kokonaisuudesta
- Lokakuussa viimeisteltiin visualisaatiot ja projektin loppuraportti

3.4 Projektin resurssit

Toimittajalta projektiin osallistuivat SoteDigi Oy:n, Vimana Oy:n, Citrus Solutions Oy:n ja Productivity Leap Oy:n asiantuntijat.

Projektin toteutustyö jakaantui neljään eri osa-alueeseen:

- Määrittely, suunnittelu, tulosten validointi
- Tietojen lataus tietovarastoon (integrointi- ja ETL/ELT-vaiheet)
- Tiedon hyödyntäminen (visualisaatiot, koneoppimisen mallit)
- Projektin hallinta

Projektin työmäärät jakaantuivat toimittajan osapuolille kuvan 4 mukaisesti.

Osa-alue	SoteDigi FTE	Kumppanit FTE	Yhteensä FTE	Yhteensä %
Määrittely, suunnittelu ja tulosten validointi kaikissa työvaiheissa	2,0	0,8	2,8	42
Tietojen lataus tietovarastoon: integrointi- ja ETL-vaiheet	0,2	0,7	0,9	14
Tiedon hyödyntäminen, tekninen toteutus: visualisaatiot, koneoppimisen mallit	0,0	0,9	0,9	14
Projektin hallinta	1,5	0,5	2,0	30
Yhteensä	3,7	2,9	6,6	100

Kuva 4 Projektin työmäärien jakautuminen tehtäväalueisiin ja osapuolille, FTE=kokopäiväinen resurssi

Kainuun soten ja koordinaatioryhmän työmäärät koostuivat pääosin työpajoista ja niihin osallistumisesta. Kainuun sotea projektissa oli mukana kaikkiaan 34 henkilöä. Lista osallistuneista henkilöistä on esitetty loppuraportin liitteessä 1.

Projektin työpajat olivat avoimia eli kaikki olivat tervetulleita. Projektissa pyrittiin aktiivisesti muodostamaan yli organisaatorajojen koostuvia moniammatillisia ryhmiä valittujen aihealueiden (loppukäyttäjät & johto, toiminta & tietosisältö, ICT-toteutus) ympärille. Tavoitteessa onnistuttiin hyvin. Osallistuminen järjestettyihin kahden tunnin työpajoihin oli laajaa ja aktiivista.

Projektin yhteyksien ja laajemman hyödynnettävyyden kannalta tärkeän koordinaatioryhmän työskentelyä on käyty läpi loppuraportin luvussa 4.8 sidosryhmätyö. Kainuun soten ja koordinaatioryhmän työmäärät on esitetty kuvassa 5.

13.12.2019

Ryhmä	Tunti	Htp
Kainuu	563	75
THL	19	2,5
STM	5	0,6
Siun sote	4	0,5
Eksote	13	1,7
PHHYKY	10	1,3
Soite	6	0,8
Kymsote	4	0,5
Yhteensä	624	83,1

Kuva 5 Kainuun soten ja koordinaatioryhmän työmäärät

3.5 Projektin työskentelytapa

Projektia ohjattiin työryhmätyöskentelyn avulla siten, että projektin alussa muodostettiin kolme ryhmää, jotka osallistuivat projektin määrittelyyn. Pääryhmät olivat:

- Loppukäyttäjä & johto
- Toiminta & tietosisältö
- ICT-toteutus & tietolähteet

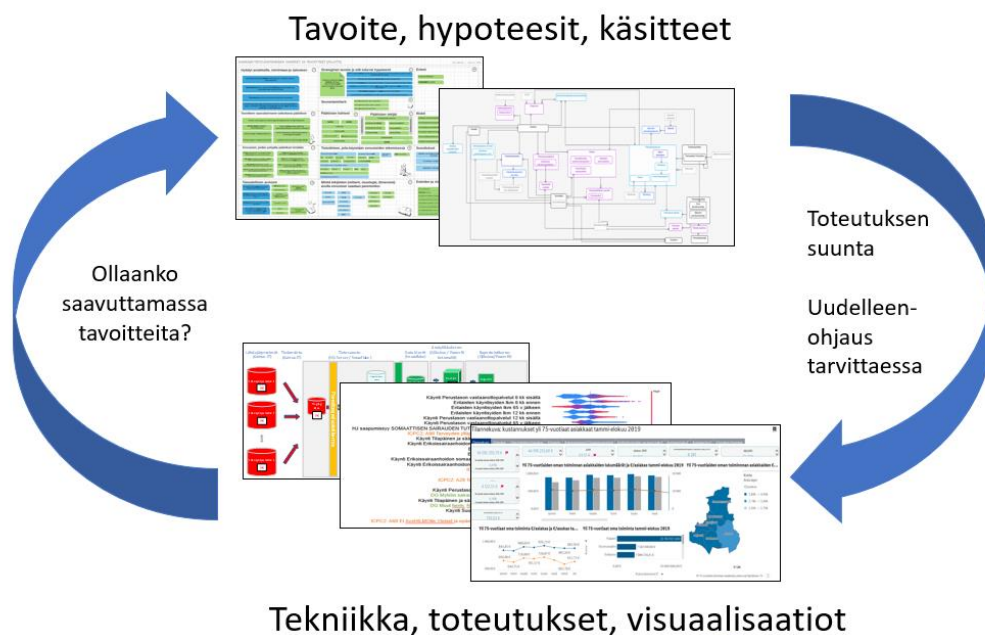
Koko projekti toteutettiin virtuaalisesti Skype:n välityksellä. Työpajoja pidettiin määrittelyvaiheessa viikoittain. Projektin loppuvaiheissa työpajoja harvennettiin ja kestoajkoja lyhennettiin. Työpajoissa käsiteltiin tietotarpeita, hypoteesicanvasta, käsitelmalleja, tietolähteitä ja tietojen saatavuutta sekä visualisaatioiden ja kehittyneen analytiikan määrittelyä ja datan validointia toteutettujen visualisaatioiden avulla. Lisäksi työpajoja järjestettiin seuraavilla aihealueilla: kustannuslaskenta, episodianalytiikka ja ICT-alusta. Kaikista pidetyistä työpajoista laadittiin muistiot.

3.6 Tulosten laadun varmistaminen

Tulosten laatu haluttiin pitää korkealla ja varmistaa, että myös sisällön kannalta saavutetaan tavoitteet. Työskentelyprosessiksi eri aihealueiden etenemisen yhteydessä vakiintui tapa palata takaisin kuvattuun strategiseen tavoitteeseen ja hypoteeseihin. Läpikäynnin tarkoituksena oli varmistaa, että projektin tulosten avulla on mahdollista todentaa hypoteeseja ja niiden avulla auttaa Kainuun sotea pääsemään strategiseen tavoitteeseen.

Kuvassa 6 on esitetty työskentelyn prosessia. Projektissa pyrittiin suoraviivaiseen toteutusmalliin. Lähtöjärjestelmän tiedot ladattiin tietovarastoon ja edelleen käytettäväksi kustannuslaskennan, kehittyneen analytiikan sovelluksen ja visualisaatioiden pohjana. Kuvassa suoraviivainen malli koskee kuvan oikean laidan esittämää etenemistä määrittelyistä projektin lopputuloksiin eli tiedonsiirtoihin ja tiedon hyödyntämiseen.

13.12.2019



Kuva 6 Tulosten laadunvarmistusprosessi.

Kuvan vasemmassa laidassa on kuvattu takaisinkytkentä tavoitteeseen, hypoteeseihin ja käsitteisiin. Tarkoituksena oli säännöllisesti arvioida, ollaanko saavuttamassa tavoitteita. Tarvittaessa toteutusta uudelleen ohjattiin ketterästi kohti tavoitetta, mikä joissakin tapauksissa tarkoitti esim. tietolähteiden lisäämistä.

3.7 Projektin tavoitteiden toteutuminen

Projektissa toteutettiin Virta-hankkeeseen kuuluva yhteisesti rajattu tietosisältö, joka hyödyntää Kainuun soten lähdejärjestelmiä. Projektissa tuotettiin tietovaraston tietosisältöön määrittely käsittemallimuodossa. Käsittemalli koostui kolmesta erillisestä mallista: tuotanto, yleiset käsitteet ja kustannuslaskenta. Mallissa määriteltiin kaikkiaan 110 käsitettä ja niiden määrittelyt pohjautuivat kansallisiin määrittelyihin, niiltä osin kuin kansalliset määrittelyt olivat olemassa.

Lähtötiedot tietovarastoon tuotiin Kainuun soten lähdejärjestelmistä, joita olivat Effica Raportointi, Proconsona, RAIsoft ja TIJO-tietovarasto, josta tuotiin Pegasos HR:n tiedot ja SAP Finance. Lisäksi tuotiin väestötietoja manuaalisesti Tilastokeskuksen nettisivuilta.

Tietosisältöön oleellisena osana kuuluvassa tietovarastossa on kaikkiaan 209 taulua data vault -mallinnuksen mukaan toteutettuna sisältäen 18 hub-, 18 satelliitti-, 18 linkki-, 11 referenssi-, 66 state- ja 67 load-taulua. Tietovarastossa on 68 näkymää ja raportointikannassa 11 taulua. Tietotuotteeseen tehtiin 14 visualisaatiota, jolla loppukäyttäjät voivat tarkastella monipuolisesti tietovarastoon kerättyjä tietoja.

Toteutukseen kuuluu koneoppimisen malli, joka ennustaa etukäteen henkilöiden ohjautumisen säännöllisiin palveluihin. Kehittyneen analytiikan tuloksia on tarkemmin käyty läpi luvussa 4.6. Tietojen välitystä ulkoisesta kohteesta tietovarastoon ja tietojen välitystä tietovarastosta ulkoiseen kohteeseen ei toteutettu rajatusta toteutusajasta johtuen.

13.12.2019

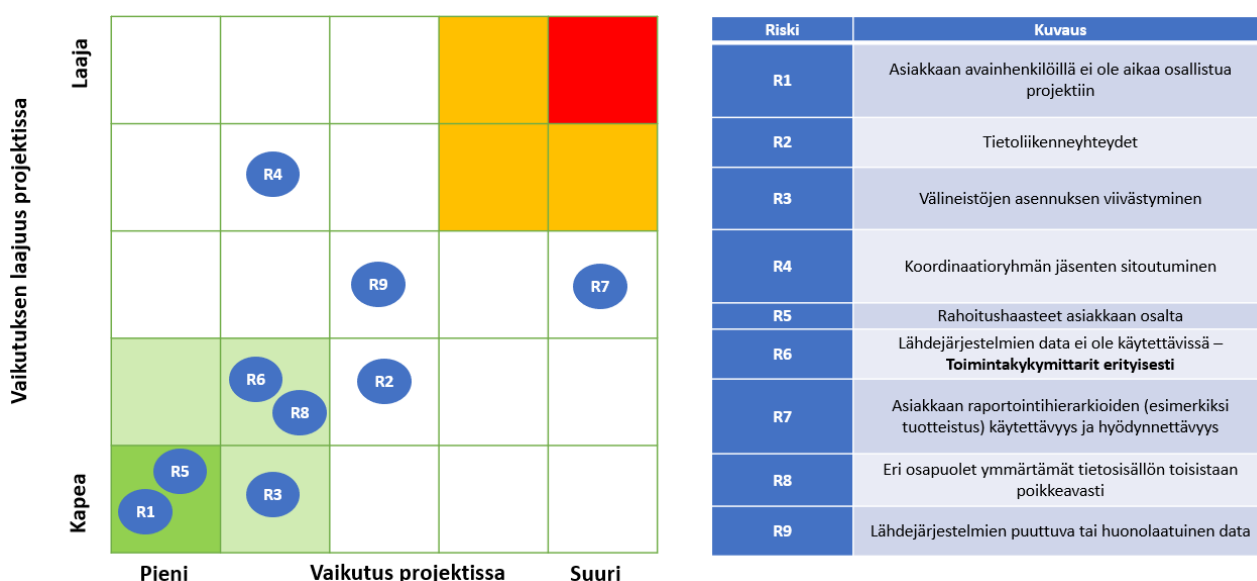
Projektissa testattiin tietöalustan integroitumista Kainuun soten muuhun tietöjohtamisen ympäristöön. Projekti toteutettiin Kainuun soten omassa ICT-ympäristössä. Toteutus on kuvattu tarkemmin luvussa 5. Samassa luvussa kuvataan myös kokemuksia projektissa käytetyistä välineistä ja vertaillaan niitä projektin vaatimuksiin ja tavoitteisiin.

Projektissa selvitettiin ja kehitettiin Sosiaali- ja terveystietöjen toissijaisen käytön lain (552/2019) [1] hyödyntämiseen liittyviä käytäntöjä, sopimusmalleja ja sen mahdollistamia tiedon hyödyntämistapoja. Tätä osuutta on analysoitu luvussa 6.2. Projektin tietosisältöihin liittyvät tavoitteet sekä niiden toteutuminen ja tulokset on vedetty yhteen luvussa 7.

3.8 Projektin riskit ja niiden toteutumisen arviointi

Projektin alussa tunnistettiin projektin riskit ja kategorisoitiin ne riskien vaikutuksen ja todennäköisyyden avulla luokkiin asteikolla 1–5. Kolme tärkeintä tunnistettua riskiä olivat: välineistön asennuksen viivästyminen, asiakkaan avainhenkilöillä ei ole aikaa osallistua projektiin ja tietoliikenneyhteyksien avaaminen.

Projektin päättyessä todettiin, että korkeimman prioriteetin riskit eivät toteutuneet (kuva 7). Välineistöt saatiin asennettua sopivaan aikaan projektin työskentelyn kannalta. Kainuun soten asiantuntijat osallistuivat hyvin työpajoihin ja kokouksiin. Tietoliikenneyhteyksien avaamiset onnistuivat oikea-aikaisesti ja nopeasti. Projekti saatiin kaikkiaan hyvin liikkeelle.



Kuva 7 Riskien toteutumisen vaikutukset projektiin

Suurimmat vaikutukset projektiin olivat asiakkaan valmiiden raportointihierarkioiden käytettävyydellä. Valmiiden, järjestäjän toimintojen seurantaan ja vertailuun käytettävien hierarkioiden puuttuminen (R9) vaikutti koko projektiin jossain määrin. Asiakkaan raportointihierarkiat ja siihen liittyvät luokitukset eivät vastanneet järjestämistehtävän mukaista tarvetta ja ne jouduttiin rakentamaan projektissa.

Lähdejärjestelmien puuttuva tai huonolaatuinen data (R9) vaikutti projektiin tiettyjen tietosisältöjen osalta. Lähdejärjestelmien tietosisältöjen selvittely hidasti jossain määrin projektin etenemistä. Jatkossa

13.12.2019

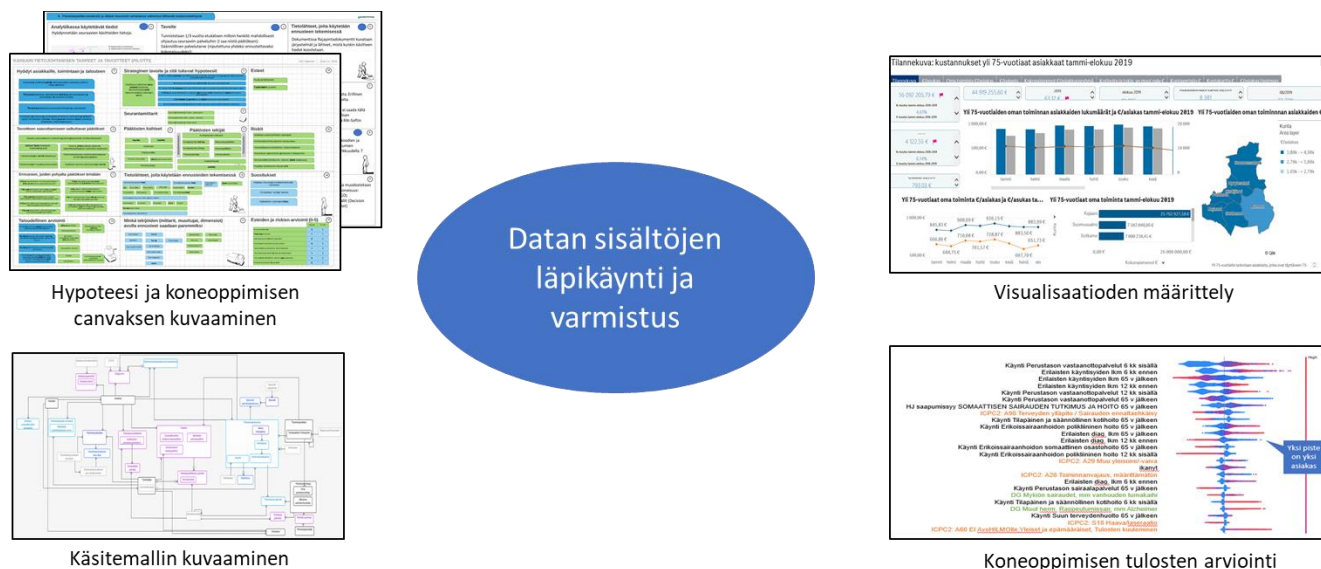
lähdejärjestelmien sisältöä kannattaa tarkastella siihen soveltuvilla data preparation -välineillä heti projektin alusta lähtien osana käsitelmäärittelyä.

Koordinaatioryhmä sitoutui projektiin osittain kiitettävästi. Koordinaatioryhmän osallistuminen ei kuitenkaan toteutunut aiotussa laajuudessa. Eksote ja PHHYKY olivat kuntayhtymistä aktiivisimmat, osittain myös sen vuoksi, että näillä alueilla on vastaavia hankkeita menossa. Osallistumisen aktiivisuuteen vaikutti kommenttien perusteella myös se, että useilla alueilla on haasteita yhtymän talouden osalta ja myös seuraavien vuosien talousarvioiden laadinta osui samaan aikaan projektin kanssa. Vaikutus projektiin erityisesti laajemman maakuntakohtaisen näkyvyyden ja vertailtavuuden kannalta oli keskimääräistä suurempi, mutta ei estänyt projektin edistämistä.

Tietoliikenneyhteydet olivat joissakin käyttötapauksissa hitaita, samoin kuin käytettävien VPN-yhteyksien määrä osittain riittämätön. Muilta osin riskit olivat sekä toteutumiseltaan että vaikutuksiltaan vähäisiä.

3.9 Projektin työmenetelmien arviointi

Projektin työmenetelmiä on arvioitu projektin sisällöllisten ja projektin teknisen toteutuksen kannalta. Projektin sisällöllisessä määrittelyssä käytettiin sote-järjestämisen tietomallia, hypoteesicanvasta, käsitemallia, datan sisältöjä todellisen järjestelmistä siirrettävän ja tietovarastoon siirretyn tiedon osalta, datan pohjalta toteutettuja visualisaatioita ja koneoppimisen mallin tuloksia. Projektin sisällölliset työmenetelmät on esitetty karkealla tasolla kuvassa 8.



Kuva 8 Projektin sisällölliset työmenetelmät

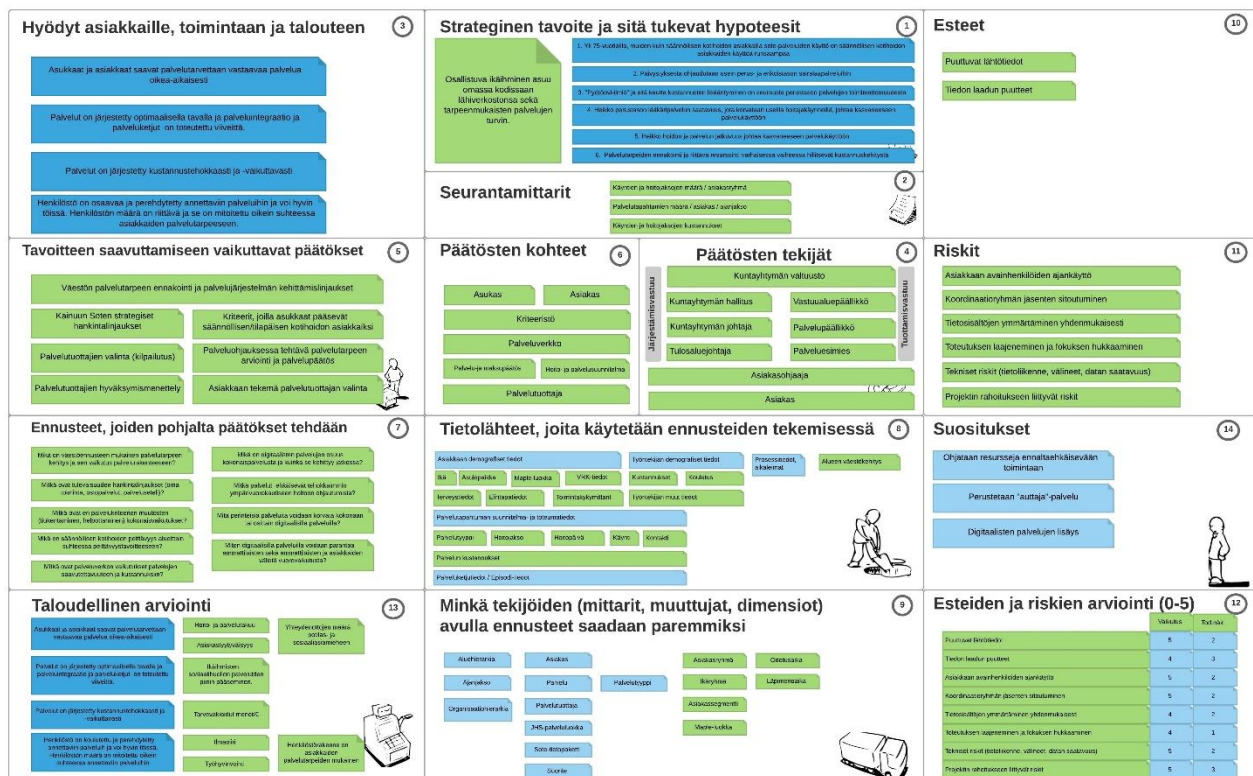
Projektin tavoitteiden sekä kehitysprosessiin liittyvien muuttujien hahmottamisen välineenä käytettiin hypoteesicanvas-menetelmää. Keskeistä tässä menetelmässä on tiivistää projektiin ja sen tavoitteiden saavuttamiseen liittyvät asiat yhdelle sivulle selkokielisenä, mikä osaltaan helpottaa kommunikointia kehitystiimin, asiakkaan ja eri toimijoiden välillä.

Hypoteesicanvas tuotettiin Kainuun sotien strategiakartan pohjalta. Kainuun sotien strategiakartassa vuosille 2017–2020 on kuvattu erityisesti mittaristoja, joita otettiin mukaan pilotissa tehtyyn määrittelyyn soveltuvin

13.12.2019

osin. Määrittely käynnistyi ikääntyneiden kotona asumisen tukemista kuvaavasta tavoitteesta ja sen avulla työpajoissa asiakkaan kanssa tuotettiin projektin määrittelyn tarvitsemat kuvaukset:

- Strategiset tavoitteet sekä oletukset, hypoteesit, joiden tietosisältö tuottaa informaatiota päätöksenteon tueksi
- Mittarit, joilla onnistumista mitataan (strategiasta)
- Hyödyt, joita saavutetaan tavoitteeseen pääsyn ansiosta: asiakkaisiin, henkilöstöön, toimintaan talouteen liittyen (strategiasta johdettavaa tietoa)
- Päätöksentekijät, jotka vaikuttavat tavoitteeseen pääsyyn (myös muita kuin ylin johto)
- Päätökset, jotka vaikuttavat tavoitteeseen pääsyyn
- Päätösten kohteet
- Päätöksentekoa tukevat ennusteet
- Tietolähteet raportoinnin ja ennakoivan analytiikan toteuttamiseksi
- Tekijät (mittarit, muuttujat ja dimensiot) joiden avulla ennusteita voidaan parantaa
- Esteet, riskit ja niiden arviointi
- Taloudellinen arviointi
- Suositukset projektin tulosten pohjalta



Kuva 9 Hypoteesicanvas

Hypoteesicanvas tuki tavoiteasetantaan liittyvää keskustelua. Tavoitteen pohjalta kuvattiin hypoteesit eli oletukset, joilla ajatellaan olevan vaikutusta tavoitteeseen pääsyyn. Pilottiprojektin rajoitukset tehtiin vasta sen jälkeen, kun sekä hypoteesicanvas ja myöhemmin tässä dokumentissa kuvattu käsitelmä oli muodostettu. Tällöin otettiin huomioon pilotin rajallinen toteutusaika.

13.12.2019

Strategisena tavoitteena oli:

”Osallistuva ikäihminen asuu omassa kodissaan lähiverkostonsa sekä tarpeenmukaisten palvelujen turvin”

Hypoteesien muodostamista varten Loppukäyttäjä ja Johto-työpajassa analysoitiin laajasti tavoitteeseen pääsyyn vaikuttavia tekijöitä. Hypoteeseiksi tunnistettiin seuraavat oletukset:

1. Yli 75-vuotiailla, muiden kuin säännöllisen kotihoidon asiakkailla kotihoidon asiakkailla, sote-palveluiden käyttö on säännöllisen kotihoidon asiakkaiden palvelukäyttöä runsaampaa.
2. Päivystyksestä ohjaututaan usein perus- ja erikoistason sairaalapalveluihin.
3. ”Pyöröovi-ilmiö” ja sitä kautta kustannusten lisääntyminen on seurausta perustason palvelujen toimimattomuudesta.
4. Heikko perustason lääkäripalvelujen saatavuus, jota korvataan useilla hoitajakäynneillä, johtaa kasvaneeseen palvelukäyttöön.
5. Heikko hoidon ja palvelun jatkuvuus johtaa kasvaneeseen palvelukäyttöön.
6. Palvelutarpeiden ennakointi ja riittävä resurssointi varhaisessa vaiheessa hillitsevät kustannuskehitystä.

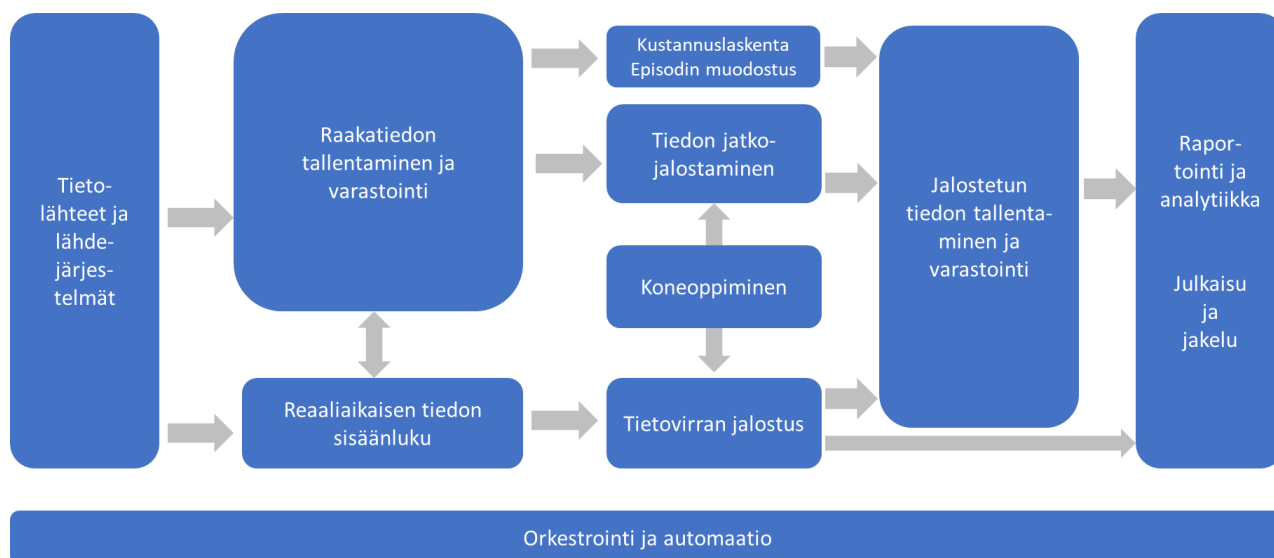
Pilotin rajatussa toteutuksessa päädyttiin tarkemmin käymään läpi kolmea hypoteesia ja kehittämään järjestelmä niiden pohjalta. Kolmeksi toteuttavaksi hypoteesiksi valittiin hypoteesit 1, 2 ja 6. Näistä kahta ensimmäistä päädyttiin toteuttamaan ja analysoimaan tehtävien visualisaatioiden kautta ja viimeistä palvelutarpeiden ennakointiin liittyvää hypoteesia kehittyneen analytiikan ja koneoppimisen avulla.

Sisällölliset työmenetelmät toimivat hyvin ja takasivat projektin laadun varmistamisen. Tätä on tarkemmin käsitelty kappaleessa 3.6. Kun työmenetelmiä arvioidaan, kehittämiskohteena jatkon kannalta on erityisesti datan tietosisällön selvitys aikaisemmassa vaiheessa projektia.

Kun tarvittavat lähdejärjestelmät on saatu projektissa hyödynnettäväksi, olisi projektin laadun varmistamiseksi tuotettava aikaisessa vaiheessa tietoa tietovaraston tietosisällöistä arvioitavaksi sisältöasiantuntijoille. Tämä voidaan toteuttaa esimerkiksi niin, että valitulla raportointivälineellä tuotetaan tietoja asiantuntijoiden käyttöön nopeasti ja ketterästi datan jalostuksen eri vaiheista. Tiedon perusteella voidaan aikatauluttaa toteutusta, tilata tarvittavia uusia rajapintoja toteutusta varten tai vaikuttaa pidemmän aikavälin tavoitteisiin kuten kirjauskäytäntöjen kehittämiseen.

Teknisiä työmenetelmiä tarkastellessa lähtökohdaksi voidaan ottaa projektissa käytetty karkean tason arkkitehtuuri (kuva 10).

13.12.2019

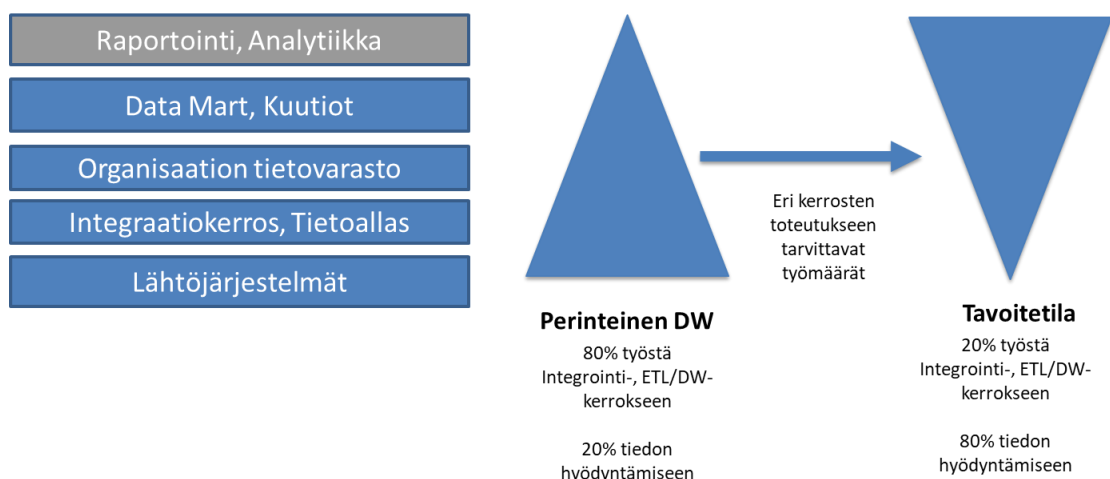

Kuva 10 Projektin tekniset työmenetelmät

Käsitemallinnus toimi hyvin linkkinä projektin sisällöllisen määrittelyn ja teknisen toteutuksen välillä. Sisällöllinen määrittely tehtiin toiminnan näkökulmasta ottamatta huomioon teknisen toteutuksen mahdollisesti tuomia rajoitteita.

Käsitemallin pohjalta kuvattiin lataukset raakatiedon tallettamiseen ja varastointiin tietolähteistä ja lähdejärjestelmistä. Kaikki tietojen siirrot toteutettiin automaattisen ETL/ELT:n mahdollistamalla välineistöllä. Välineen avulla siirrettiin data myös raportointikantaan (kuvassa 10: jalostetun tiedon tallentaminen ja varastointi). Raportointikantaan siirrettäessä jouduttiin tekemään enemmän käsittelysääntöjä ja muunnoksia. Datan siirron automatisointi mahdollistaa sen, että käytännössä projektin työpanosta voidaan paremmin kohdistaa tiedon hyödyntämiseen.

Kuvassa 11 on kuvattu muutos perinteisestä tiedonjalostusprosessista suhteessa automatisoituun prosessiin. Karkeasti arvioituna perinteisessä mallissa 80% projektien työmäärästä menee datan siirtoon lähdejärjestelmistä hyödyntämistä varten ja 20% pystytään käyttämään tiedon hyödyntämisen. Tämän suhteen muuttuminen on erittäin tärkeää sisällöllisen laadun varmistamisessa. Kun projektin sisältöasiantuntijat päätyvät siihen, että projektissa tarvitaan uutta tietoa, tiedon saaminen käytettäväksi on tuntien tai päivien toteutus, ei viikkojen tai kuukausien.

13.12.2019


Kuva 11 Painopiste työskentelyssä tiedon hyödyntämiseen

Projektissa saavutettiin 50/50 jakautuminen pääosin siitä syystä, että tietojen latausvaiheeseen laskettiin myös tietojen tuottaminen kustannuslaskennan rajapintaan normaalista poikkeavana prosessina. Ilman rajapinnan tuottamista tietovaraston tiedoista kustannuslaskennan järjestelmään olisi saavutettu tavoitteena ollut 20/80-jakautuminen eli selvästi suurempi kohdistuminen tiedon hyödyntämiskerrokseen.

Pilottiprojektissa päädyttiin aikataulusyistä siihen, että tietoallasta ei toteutettu vaan data ladattiin suoraan lähdejärjestelmistä käytettäväksi tietovarastoinnissa. Jos käytettävissä olisi tietoallas, olisi siihen ennestään ladattu kaikkien tarvittavien tietojärjestelmien sisällöt. Näin data olisi välittömästi käytettävissä riippumatta siitä, mistä järjestelmistä data tulee. Vastaava pätee myös koneoppimisen ja kustannuslaskennan malleihin. Jos näissä tarvitaan lisää tietoa tai tiedossa on muutostarpeita, on uuden tietosisällön lataaminen helppoa ja tehokasta ilman aikaa vieviä integrointeja.

Teknisiin työmenetelmiin ja välineisiin voisi tietoa lisäksi lisätä datan valmisteluun liittyvät data preparation -välineet. Näiden avulla datan sisältöihin päästään aikaisessa vaiheessa kiinni jo ennen kuin niitä on siirretty tietovarastoon. Tämä nopeuttaa aiemmassa kappaleessa kuvastusti sisältöasiantuntijoiden työtä ja teknistä toteutusta, koska välineiden avulla voidaan etukäteen hahmottaa paremmin tietolähteiden sisältöä.

13.12.2019

4. Projektin sisältö ja havainnot

4.1 Palveluluokitukset

Kansallisesti on käytössä useita erilaisia sosiaali- ja terveydenhuollon palveluluokituksia, joita on laadittu eri käyttötarkoituksia varten ja eri lainsäädäntöjen ohjaamana. Luokituksia ovat mm. JHS-palveluluokitus, THL:n eri palveluluokitukset ja Sote-tietopakettiluokitus, yhteensä 17 kansallista luokitusta. Lisäksi alueittain on laadittu omia luokituksia, joiden avulla johdetaan ja ohjataan toimintaa. Tietojohdamisen näkökulmasta tulisi päästä kansallisesti yhteneväiseen luokitukseen, joka mahdollistaa tiedon vertailun. Eri luokitusten määrä tulisi minimoida eri raportointeja varten toteutettavan tietotuotannon helpottamiseksi.

Järjestämistehtävässä käytettävällä palveluluokittelulla tulisi olla mahdollista vastata sekä alueellisiin että kansallisiin tietotarpeisiin. Pidemmän aikavälin toiminnan ja talouden tietojen vertailtavuuden vuoksi luokituksen tulee olla organisaatoriippumaton ja keskeisiltä osiltaan vakaa, niin että luokittelutekijät eivät jatkuvasti muutu. Lisäksi luokittelun tulee olla yhteensopiva kustannuslaskennan tarpeiden kanssa. Näistä syistä johtuen esim. tämän hetkinen JHS 200, Kuntien ja kuntayhtymien palveluluokitus, ei välttämättä sovellu käytettäväksi muuttuvassa toimintaympäristössä.

Tässä pilotissa hyödynsimme mahdollisuuksien mukaan Sote-tietopakettiluokitusta ja sen määrittelyjä, koska kyseisessä työssä on ollut laajasti mukana eri alueiden edustusta ja luokituksen valmistelussa on pyritty noudattamaan organisaatoriippumatonta rakennetta. Palveluiden sisällön tarkastelussa hyödynnettiin tietopakettien toiminto- ja palveluryhmätason hierarkiaa.

Pilotissa palvelujen sisällöt muodostettiin tarkastelemalla ja yhdistämällä toimipaikkoja toiminnan ja taloushallinnon näkökulmasta, kiinnittäen edelleen toimipaikat palveluihin. Järjestämistehtävän tietotarpeiden mukaisesti palvelukäytön ja palvelujen sisältöjen vertailun pohjaksi syntyi 18 eri palvelua, joissa yli 75-vuotias henkilö voi olla asiakkaana. Osittain sisältöjä tarvittaessa avattiin myös tarkemmalle tasolle. Koko väestötason tarkastelussa käytetty luokittelu ei ole kattava.

- Erikoissairaanhoidon polikliininen hoito
- Erikoissairaanhoidon somaattinen osastohoito
- Kuntoutuspalvelut
- Mielenterveyspalvelut asumispalvelut
- Mielenterveyspalvelut avohoito
- Muut kotona asumista tukevat palvelut
- Omaishoito
- Ostopalvelut
- Perustason sairaalapalvelut
- Perustason vastaanottopalvelut
- Psykiatrinen osastohoito
- Päihdepalvelut avohoito
- Päivystyspalvelut
- Suun terveydenhuolto
- Tehostettu palveluasuminen
- Tilapäinen ja säännöllinen kotihoito
- Vammaispalvelut
- Ympäri vuorokautinen laitoshoido

13.12.2019

4.2 Käsittemallinnus

Sosiaali- ja terveydenhuollon käsitteitä ja niiden määrittelyjä hyödynnetään useisiin eri tarkoituksiin. Kansallisesti virallisten sosiaali- ja terveydenhuollon käsitteiden määrittelystä vastaa THL. Lisäksi sote-toimialaa koskevaa käsittemäärittelyä tehdään mm. Kelassa, Sosiaali- ja terveysministeriössä ja Valvirassa sekä alueellisissa kehittämishankkeissa. Pilotin käsittemäärittelyssä hyödynnettiin seuraavia kansallisia lähteitä:

- Sote-sanastot, sisältäen Kanta-sanaston, Sosiaalialan tiedonhallinnan sanaston, Lääkityksen sanaston ja Käytönhallinnan sanaston
- Hoitoilmoitus-oppaiden käsitteet
- Toimia-tietokannan käsitteet
- Koodistopalvelu
- Lait ja asetukset, esimerkiksi kuntalaki
- Muut kansalliset julkaisut, esimerkiksi Kuntaliiton ICPC2-luokituksen opas ja THL:n ICD10-tautiluokitus

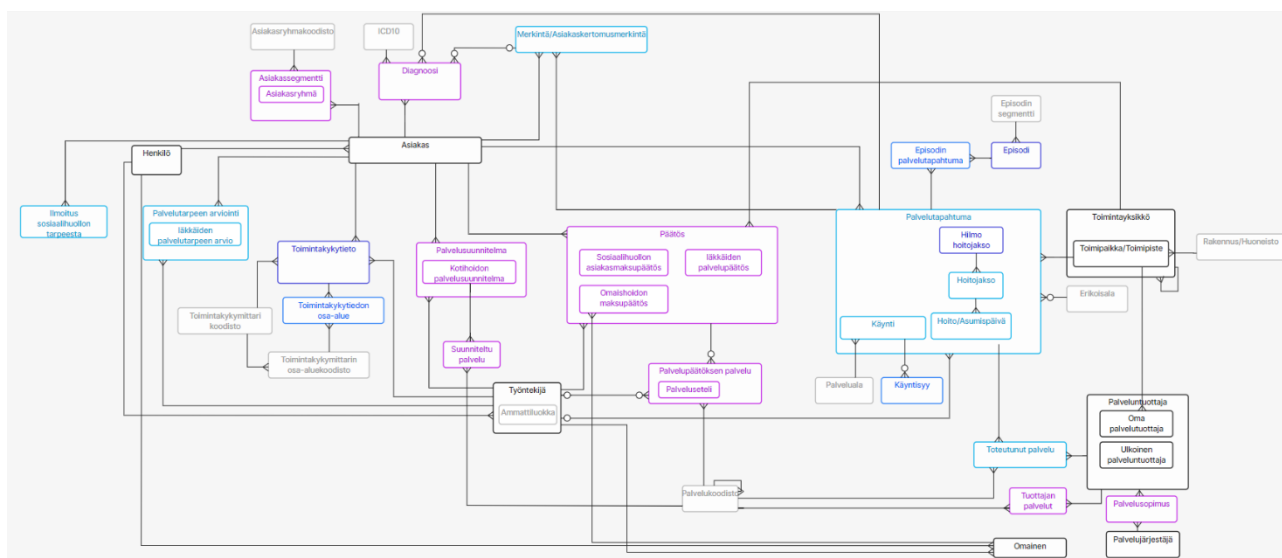
Käsittemallinnuksella tarkoitetaan menetelmää, jolla tietyn kohdealueen käsitteitä ja tietoja mallinnetaan liiketoiminnan näkökulmasta, sovellus- ja tekniikkariippumattomalla tavalla. Tuloksena syntyy käsitteet ja niiden keskinäiset riippuvuudet kuvaava graafinen käsittemalli sekä käsitteiden määritykset [2].

Käsittemallinnuksen perustan muodostivat Kainuun soten ikäihmisten palveluiden strateginen tavoite ja sen edellyttämät tietotarpeet. Mallinnuksen sisältö käsitteiden osalta rajautui pilotin toteutuksen mukaisesti ikäihmisten koko palvelukokonaisuuden avainkohtiin, kattaen niin sosiaalihuollon, perusterveydenhuollon kuin erikoissairaanhoidon palvelut. Kotihoidon suunnitteluun, toteuttamiseen ja arviointiin liittyvät käsitteet ovat käsittemallin keskiössä.

Pilotissa käsittemallinnusta tehtiin työpajatyöskentelynä toimittajan ja Kainuun soten asiantuntijoiden kesken. Mallinnuksessa käytettiin kansallisia käsitteitä ja käsitteiden lähdetiedot tallennettiin käsitteiden yhteyteen. Työn aikana hyödynnettiin mm. THL:n, STM:n sekä pilotin koordinaatioryhmän jäsenten asiantuntemusta käsittemäärittelyssä. Käsitemalleja tarkennettiin työn edetessä saatujen täsmennysten perusteella.

Mallien laajuuden vuoksi ne jaettiin kolmeksi erilliseksi malliksi: tuotanto, yleiset käsitteet ja kustannuslaskenta. Tuotanto-malliin kuvattiin toiminnan kannalta keskeiset käsitteet ja niiden väliset suhteet, joita ovat mm. asiakkaan hoidon ja palvelun suunnitteluun, toteutukseen ja arviointiin sekä palveluiden järjestämiseen ja tuottamiseen liittyvät käsitteet (kuva 12).

13.12.2019



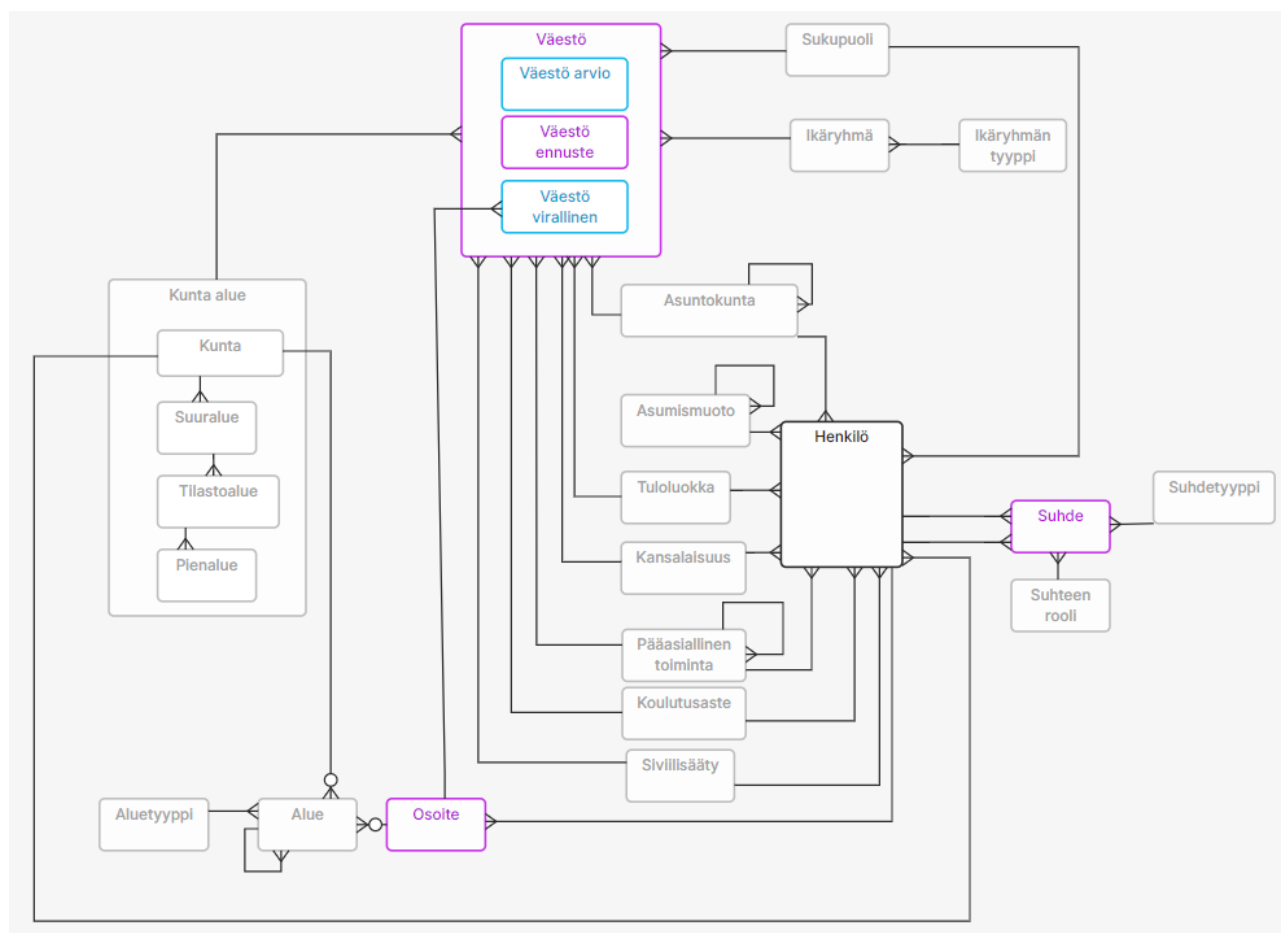
Kuva 12 Tuotanto-käsitemalli

Kustannuslaskentamallissa on kuvattu ikäihmisten palvelukokonaisuuteen liittyvät kustannuslaskennan käsitteet, joita on käytetty pilotin kustannuslaskennan toteutuksessa. Keskeisiä käsitteitä tässä mallissa ovat mm. palvelutapahtuma, palvelun toteutunut kustannus ja kirjanpidon toteuma. (kuva 13)

[illegible]

Yleiset käsitteet -malli sisältää yleisiä käsitteitä, jotka ovat riippumattomia käytettävästä toimintaympäristöstä, kuten väestöön, maantieteellisiin alueisiin ja henkilöön liittyvät käsitteet (kuva 14).

13.12.2019



Kuva 14 Yleiset käsitteet -käsittemalli

Käsittemalleista generoitiin tietovaraston looginen tietomalli Wherescape-työkalun avulla. Pilotissa muodostuneet käsittemallit on viety yhteentoimiva.suomi.fi tietomallit-työkaluun tietojohdaminen-kategoriaan. Yhteentoimivuusalustasta on kerrottu tarkemmin kappaleessa 5.5.1.

4.3 Tietolähteet ja tiedonsiirto

Projektissa käytettiin tietolähteinä Kainuun sotien käytössä olevia tietojärjestelmiä ja lisäksi tehtiin erilaisia luokitteluja ja tietojen yhdistelyjä varten manuaalisesti ylläpidettäviä tietolähteitä.

- SAP ja TIJO SAP
- SAPissa on kirjanpito ja taloustiedot
- TIJO on Kainuussa aikaisemmin toteutettu tietojohdamisen järjestelmä, johon on myös viety taloustietoja
- Lifecare-potilastietojärjestelmän tiedot Effica Raportoinnin kautta
- Proconsona: sosiaalipalveluiden järjestelmä
- Pegasos HR: Pegasos HR:ssä on henkilöstöön liittyviä tietoja
- RAI-toimintakyky mittarit: RAI-järjestelmän toimintakyky tiedot
- Manuaaliset tietolähteet
- Palveluiden luokittelu

13.12.2019

- Kustannuspaikkojen ja yksiköiden yhdistäminen eri vuosien organisaatorakenteella
- Poistettavat yksiköt
- Prodacapon konversiotaulut
- Väestötiedot tilastokeskuksesta

Tiedonsiirto lähdejärjestelmistä toteutettiin kahdella eri tavalla. SAPista tuli päivittäin dataa Ensemblen avulla csv-tiedostona, nämä luettiin levyltä tietokantaa BULK INSERTillä ajastetusti. Muutoin tietojen haku tehtiin suorina tietokantakyselyinä avattuihin rajapintoihin. Yhteydet perustettiin linked server -objekteina ja niissä käytettiin käyttäjätunnukseksi teknistä tunnusta, jolle oli annettu riittävät oikeudet lähdejärjestelmiin. Manuaaliset tietolähteet ja RAI-toimintakykymittarit luettiin tietokantatauluihin pääasiassa kertasiirtoina, jotka toistettiin tarvittaessa, jos tiedot muuttuivat.

SAPin tiedot siirretään CSV-tiedostoina ajastetusti joka yö. Siirto hoidetaan Kainuun Ensemble-ohjelmiston avulla. Kun tiedot on saatu tiedostopalvelimelle (kehityspalvelimen D:\Data_in\SAP -kansio), ne ladataan tietokantatauluun käyttäen BULK LOAD -komentoa tietovarastopalvelimella. Ladatut taulut on kuvattu liitteenä olevassa arkkitehtuuridokumentissa.

Projektin aikana huomattiin, että Ensemblen kautta tulee vain kuluvaan kauden SAP-tiedot. Koska kyseessä oli aikanaan TIJO-projektia varten rakennettu tiedonsiirto, niin historiatiedot olivat saatavissa **TIJO:n kannasta**. Projektissa päätettiin, että historiatiedot on helpompi hakea TIJOsta kuin tehdä muutos SAP-poimintaan. Siksi TIJOa varten perustettua liittymää käytettiin myös SAP-historiatietojen osalta. Ladatut taulut on kuvattu liitteenä olevassa arkkitehtuuridokumentissa.

Lifecaren tiedot haetaan suoraan Effican Raportointi -rajapinnassa olevasta SQL-palvelimen tietokannasta. Palvelin on perustettu linked server -objektina sekä kehitys- että testipalvelimelle, nimenä on käytetty palvelimen IP-osoitetta, joka on dokumentoitu arkkitehtuuridokumentissa. Rajapinnasta ladatut taulut on kuvattu erikseen liitteenä olevassa arkkitehtuuridokumentissa.

Proconsonan tiedot haetaan suoraan rajapinnassa olevasta SQL-palvelimen tietokannasta. Palvelin on perustettu linked server -objektina kehityspalvelimelle nimellä PROCONSONA. Linked serveriä varten kehityspalvelimelle on myös asennettu Oraclen client ja ODBC-ajurit, jotta siirto onnistuu. Nämä asensi CGI. Ladatut taulut on kuvattu liitteenä olevassa arkkitehtuuridokumentissa.

Pegasos HR:n tiedot otettiin TIJO-järjestelmästä, johon ne oli siirretty aikaisemmin. Tätä varten ei siis luotu uutta liittymää. Pegasoksen tiedot haetaan suoraan rajapinnassa olevasta SQL-palvelimen tietokannasta. Palvelin on perustettu linked server -objektina sekä kehitys- että testipalvelimelle, nimenä on käytetty palvelimen IP-osoitetta, joka on dokumentoitu arkkitehtuuridokumentissa. Ladatut taulut on kuvattu liitteenä olevassa arkkitehtuuridokumentissa.

RAI-toimintakykytietoja ei saatu rajapinnasta ladattua, joten niiden kohdalla päädyttiin siihen, että tiedot otettiin CSV-tiedostoina kertasiirtona ja niistä muodostettiin tietokantataulut load-alueelle.

Manuaalisesti koottuja tietolähteitä oli useita. Palveluiden luokittelu tehtiin Excel-tilin avulla. Luokittelu perustui suorituspaikkana toimineeseen yksikköön. Eri järjestelmissä olevien organisaatietietojen yhdistäminen tehtiin erillisellä Kainuun ylläpitämällä Excel-tilillä. Taulukossa oli SAPin, Pegasoksen ja Lifecaren organisaatietietojen keskinäinen vastaavuus eri vuosien organisaatorakenteen mukaisesti. Poistettavat yksiköt listattiin omaan tauluunsa. Prodacapoa varten tehtiin omia konversiotauluja, nämä ovat myös manuaalisesti ylläpidettäviä. Väestötiedot haettiin Tilastokeskuksesta. Tätä käyttötarkoitusta varten ei rakennettu suoraa liittymää tilastokeskuksen tietoihin vaan tiedot tallennettiin manuaalisesti ylläpidettävään Excel-tiliin ja luettiin siitä sisään.

13.12.2019

Tiedot ladattiin lähdejärjestelmästä pääosin raakana, sellaisenaan ilman rajoituksia. Tästä poikettiin rekisterinpitäjän Puolangan kunnan tietojen kohdalla, ne rajattiin pois jo tiedonsiirtovaiheessa. Tällä tavoin saatiin aikaiseksi tilanne, jossa Puolangan rekisterinpitäjyyden alaiset tiedot eivät missään vaiheessa ole tietovarastossa. Poistaminen tehtiin teknisesti yksiköiden kautta eli määriteltiin yksiköt, joiden tietoja ei oteta mukaan. Katso myös kappale 5.4.

4.4 Kustannuslaskenta

Pilotissa toteutettiin kustannuslaskentaa kahdella eri tavalla. Kotihoidon kustannuslaskennan toteutus on kuvattu kappaleessa 4.4.2 ja muun palvelukäytön kustannuslaskenta sekä vaihtoehtoisia kustannuslaskennan malleja on kuvattu kappaleessa 4.4.5. Kustannuslaskentaa voidaan toteuttaa monen toimittajan valmistuotteilla. Pilotissa valittiin implementoitavaksi yhden valmisohjelmatoimittajan ohjelmisto. Saatua kokemusta voidaan soveltaa myös muihin ohjelmistoihin.

4.4.1 Tavoitteet

Kustannuslaskenta, joka sisälsi sekä suoritelaskennan että episodianalytiikan, suunniteltiin toteutettavaksi erillisenä työkokonaisuutena, mutta kuitenkin kiinteänä osana Kainuun pilottia. Suoritelaskennan tavoitteena oli tehdä näkyväksi Kainuun soten organisaation vanhuspalvelujen ja siellä erityisesti kotihoidon asiakaskohtaisesti tarjottavien palvelujen nykytilaa ja kulurakennetta vuoden 2018 toiminta- ja taloustietojen pohjalta. Suoritelaskennan menetelmäksi valikoitui melko työläs, mutta samalla potentiaalisesti tarkka toimintolaskennan viitekehys. Tähän kuuluvalla KPP-laskennalla, eli kustannus per potilas tai asiakas - laskennalla keskityttiin erityisesti työvoimakustannuksiin, mutta myös merkittäviin yleiskustannuseriin kuten tiloihin, laitteisiin, tarvikkeisiin ja materiaaleihin ml. lääkeannosjakelu ja ateriapalvelu.

Episodianalytiikan eli hoitoketjuanalytiikan tavoitteena oli ryhmitellä Kainuun soten asiakkaiden saamat palvelut loogisesti ja ajallisesti episodeiksi ja nämä episodit edelleen episodisegmentteihin ison kuvan täsmentämiseksi asiakasvirtojen ja palvelujen integraation nykytilan osalta. Episodianalytiikalla pyrittiin tekemään näkyväksi palvelujen järjestäjän asiakas- ja palveluohjauksen toimintaa ja palvelujen ketjuuntumista tarkastelujaksolla yli palvelu- ja vastuualuearajojen. Tätä kautta haluttiin konkretisoida Kainuun sotea leimaavia palvelujen järjestämiseen liittyviä ilmiöitä ja tuottaa uutta tietoa palvelujen järjestämisen johtamisen tueksi. Episodianalytiikan tarkastelu oli laajempi kuin suoritelaskennan, keskittyen 75 vuotta täyttäneiden palvelukäyttöön rajoittumatta kotihoidon suoritteisiin.

Kustannuslaskennan kautta jalostetun ja rikastetun Kainuun soten toiminta- ja taloustiedon lisäksi pyrittiin saamaan yleisesti oppia sosiaali- ja terveystieteiden järjestämistehtävässä kustannuslaskennan mallinnuksesta, työkaluista ja teknisistä ratkaisuista sekä tämän tyyppiseen tietojohdamisen kehittämiseen soveltuvasta projektinhallintamallista.

4.4.2 Prosessi

Suoritelaskenta toteutti Kainuun soten kotihoidon 75 vuotta täyttäneiden asiakkaiden osalta perussuoritteiden muodostamisen ja toimintolaskennan kustannusten aiheuttamisperusteisesti. Perussuoritteissa yhdistettiin talous- ja toimintatiedot erikseen parametroidun kustannuslaskentamallin mukaisesti. Mallin läpi syötettiin Kainuun soten toiminnasta kerätty rajattu lähtötietoaineisto käsitelmallinnuksen ja tietolähteiden kartoituksen pohjalta, keskittyen olemassa oleviin rajapintoihin.

13.12.2019

Kustannuslaskentamalli otti huomioon Kainuun soten organisaation ja kustannuspaikat, tilikartan ja tilit, resurssit, toiminnot, välisuoritteiden kustannusajurit, välisuoritteet ja perussuoritteet. Tulokset tallennettiin Prodacapo Costing DataMart -tietovarastoon raportointia ja jatkohyödyntämistä varten.

Episodianalytiikka eli hoitoketjuanalytiikka otti vastaan hinnoitellut ja asiakkaisiin kiinnitetyt suoritelaskennan tuottamat perussuoritteet (käynnit ja hoitojaksot), linkitti ne toisiinsa hoitoketjuiksi asiakkaittain vuoden 2018 palvelutapahtumien osalta, ja ryhmitteli näin saadut hoitoketjut episodiprofiileittain segmenteiksi. Episodiprofiileja parametroidiin viisi 75 vuotta täyttäneistä: kaikki asiakkaat, kaikki kotihoidon asiakkaat, muistisairaat, päivystystä käyttäneet ja paljon palveluja tarvitsevat. Tulokset tallennettiin Prodacapo Region DataMart -tietovarastoon raportointia ja jatkohyödyntämistä varten.

Episodianalytiikassa raportoitiin tulokset rajoittumatta suoritelaskennan hinnoittelemiin kotihoidon suoritteisiin. 75 vuotta täyttäneiden asiakkaiden kotihoidon ulkopuolisen palvelukäytön perussuoritteet hinnoiteltiin keskiarvoistaen ilman kustannuslaskentamallin apua. Kustannuslaskennan tulokset tallennettiin projektin lopussa myös koko pilotin tietojohdamisen tietovarastoon.

Kustannuslaskennan suunnitteluvaihe tehtiin toukokuussa 2019. Toteutusvaihe kesti kesäkuusta syyskuun loppuun. Pilotin päättämistoimet suoritettiin lokakuussa. Projektin tehtävät vaiheistettiin alkamaan kustannuslaskentamallin toteutuksen kannalta optimaalisella tavalla (ks. kuva 15) samalla noudattaen SoteDigi Oy:n tuotekehitysprosessia. Kustannuslaskentamallia oli mahdollista kehittää iteratiivisesti tarkentaen. Jo aloitettuihin tehtäviin palattiin toteutuksen myöhemmissä vaiheissa useaan otteeseen. Näin ollen mallin eri osa-alueet ja kustannuslaskenta valmistuivat projektin tuotoksina käytännössä vasta syyskuun lopussa. Mallia ei lukittu miltään osin ennen projektin päättymistä. Tämä kaikki antoi mahdollisuuksia mallin ketterään kehittämiseen eri työryhmissä. Toisaalta mallin valmiusasteen arvioiminen oli työlästä lyhytkestoisessa pilottitoteutuksessa työn vielä ollessa kesken. Useamman toteutuserän sisältävä pidempi tuotantototeutus sallisi väliversioiden visualisoinnin, demot ja sitä kautta palautteen keruun kehitysprosessiin seuraavan iteraation pohjaksi. Nyt toteutettiin vain yksi täysi iteraatiokierros.

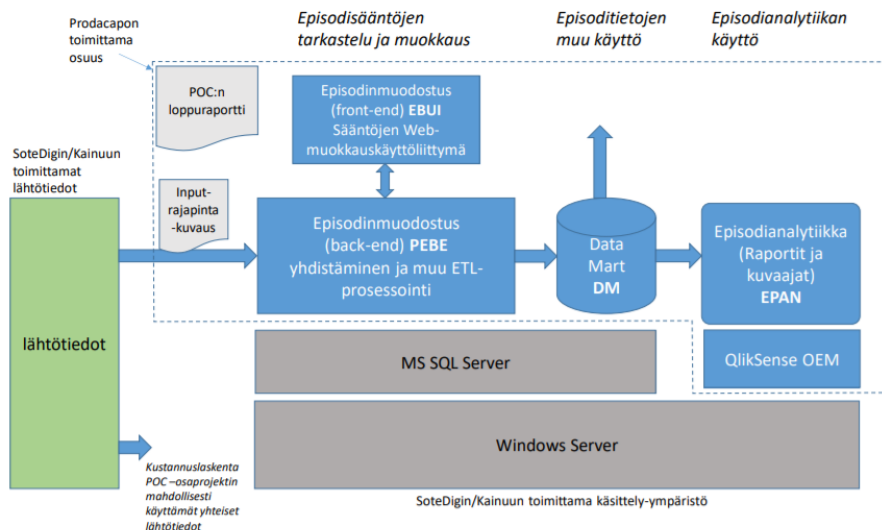
13.12.2019

Päätösportti / Tehtävän nimi	Pvm	2019											
		1	2	3	4	5	6	7	8	9	10	11	12
P0: Hyötyjen analysointi	1.2.2019	▲											
P1: Esiselvitys	2.5.2019				▲								
Hoitoketjuanalytiikka - Aloitus: Sopimus, alustava suunnitelma, tuote-esittelyt	2.5.2019 - 17.5.2019				■								
P2: Suunnittelu	20.5.2019				▲								
Hoitoketjuanalytiikka - Määrittely: Projektisuunnitelma, tietosisältökuvaus	20.5.2019 - 24.5.2019				■								
Hoitoketjuanalytiikka - Toteutus - Järjestelmän asennus ja konfigurointi: DM, PEBE, EBU, EPAN	20.5.2019 - 14.6.2019				■								
Suoritelaskenta - Toteutus - Suoritteiden määrittely	20.5.2019 - 28.6.2019				■								
Suoritelaskenta - Toteutus - Tietosisältökuvausten läpikäynti	20.5.2019 - 28.6.2019				■								
P3: Toteutus	27.5.2019				▲								
Hoitoketjuanalytiikka - Toteutus: Tietojen poiminta, poiminnan validointi	27.5.2019 - 7.6.2019				■								
Suoritelaskenta - Toteutus - Organisaatorakenne	27.5.2019 - 28.6.2019				■								
Suoritelaskenta - Toteutus - Tilit ja tiliryhmät	27.5.2019 - 28.6.2019				■								
Suoritelaskenta - Toteutus - Kustannukset	27.5.2019 - 28.6.2019				■								
Suoritelaskenta - Toteutus - Kustannuspaikkojen linkitykset	27.5.2019 - 28.6.2019				■								
Suoritelaskenta - Toteutus - Resurssien määrittely	3.6.2019 - 28.6.2019				■								
Suoritelaskenta - Toteutus - Toimintojen määrittely	3.6.2019 - 28.6.2019				■								
Suoritelaskenta - Toteutus - Yleiskustannukset / sisäiset palvelut: jakosäännöt	3.6.2019 - 28.6.2019				■								
Suoritelaskenta - Toteutus - Kustannusajureiden määrittely	3.6.2019 - 28.6.2019				■								
Hoitoketjuanalytiikka - Toteutus - Episodin muodostus	17.6.2019 - 16.8.2019					■							
Suoritelaskenta - Toteutus - Asennukset: DM, backend, frontend	1.7.2019 - 31.7.2019						■						
Suoritelaskenta - Toteutus - Tuotantodatan lataaminen	1.8.2019 - 30.8.2019							■					
Suoritelaskenta - Toteutus - Tiliryhmältä resursseille Costingiin	1.8.2019 - 30.8.2019							■					
Suoritelaskenta - Toteutus - Tuotantodatan yhdistely	5.8.2019 - 30.8.2019							■					
Suoritelaskenta - Toteutus - HR-resurssien yhdistäminen	5.8.2019 - 30.8.2019							■					
Suoritelaskenta - Toteutus - Resurssilta toiminnolle Costingiin	5.8.2019 - 30.8.2019							■					
Suoritelaskenta - Toteutus - Välisuoritteiden luominen	12.8.2019 - 30.8.2019							■					
Suoritelaskenta - Toteutus - Yleiskustannukset / sisäiset palvelut: jakosääntöjen lisääminen laskentasääntöihin	12.8.2019 - 30.8.2019							■					
Hoitoketjuanalytiikka - Koulutus - Raportoinnin koulutus, episodimäärittysten esittely	19.8.2019 - 23.8.2019								■				
Hoitoketjuanalytiikka - Testaus - Koekäyttö	26.8.2019 - 20.9.2019									■			
Suoritelaskenta - Toteutus - Määrävalidointi	2.9.2019 - 30.9.2019									■			
Suoritelaskenta - Toteutus - Hintavalidointi	9.9.2019 - 30.9.2019									■			
Suoritelaskenta - Toteutus - Mahdolliset uudelleenlataukset	9.9.2019 - 30.9.2019									■			
Suoritelaskenta - Toteutus - Mahdolliset laskentasääntöjen tarkennukset	9.9.2019 - 30.9.2019									■			
Hoitoketjuanalytiikka - Lopetus - Tulokset, loppuraportti, ympäristön purku	23.9.2019 - 30.9.2019										■		
P4: Lopetus	30.9.2019										▲		
P5: Projekti loppuu	30.9.2019										▲		
Jälkiarviointi	30.11.2019												▲

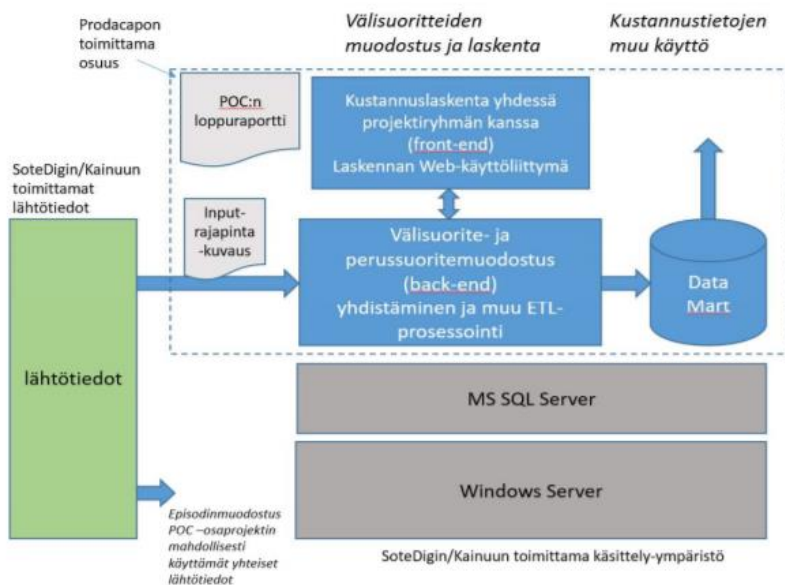
Kuva 15 Kustannuslaskennan toteutuksen vaiheistus (SoteDigi Oy, Thinking Portfolio -työkalu)

Kustannuslaskenta toteutettiin Kainuun pilotin kehitysympäristöön asennetuilla kustannuslaskennan erityisillä sovelluksilla ja työkaluilla. Suoritelaskennalla ja hoitoketjuanalytiikalla oli kummallakin oma arkkitehtuuriratkaisunsa (ks. kuva 16 ja kuva 17).

13.12.2019



Kuva 16 Hoitoketjuanalytiikan toteutuksen ratkaisuar kitehtuuri (Prodacapo Finland Oy)



Kuva 17 Suoritelaskennan toteutuksen ratkaisuar kitehtuuri (Prodacapo Finland Oy)

Projektin esiselvitys- ja suunnitteluvaiheissa käytettiin paljon aikaa pilotin käsitteiden työstämiseen. Toiminnan ja yleisten käsitteiden lisäksi tehtiin erillinen kustannuslaskennan käsitteellinen malli (ks. kappale 4.2, kuva 13), jotta saatiin ymmärrys tämän osa-alueen käsitteistä ja niiden välisistä suhteista, attribuuteista sekä niiden virallisista tietolähteistä.

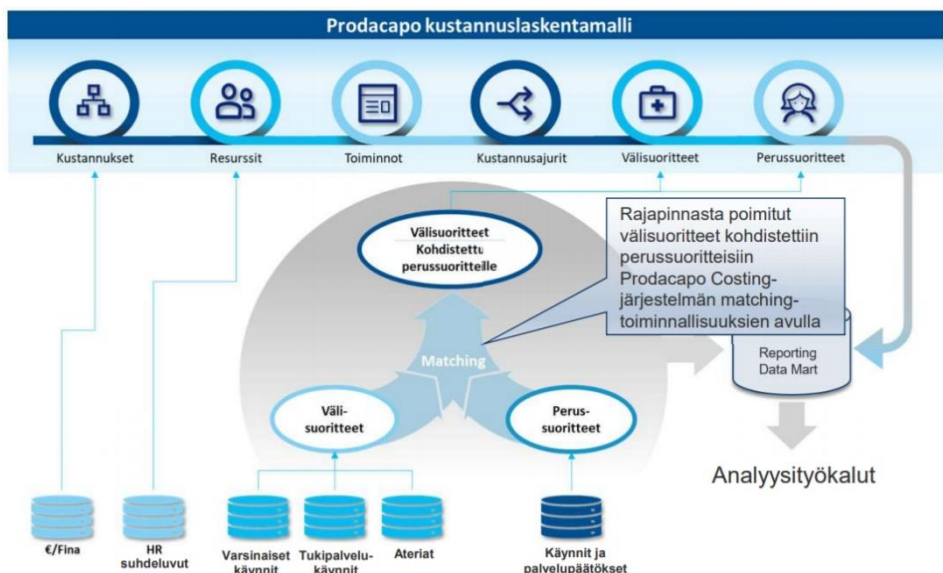
Kustannuslaskentamallin rakentamista ja toiminnan suoritteiden kustannuslaskennan toteuttamista varten toimitettiin suuri joukko lähtötietoja Kainuun soten taloudesta ja toiminnasta kustannuslaskennan määrittämään lähtötietojen rajapintaan. Tämä vaati ennakoituakin enemmän tietojen konversioita. Lähtötiedot toimitettiin pääsääntöisesti tärkeysjärjestyksessä rajauksen mukaisesti. Lähtötiedot toimitettiin mm. seuraavilta osa-alueilta:

- Kainuun soten organisaatio ja kustannuspaikat organisaatiotyyppittäin

13.12.2019

- Kainuun soten kirjanpito, tilikartta ja tiliryhmät
- Kustannusten jakosuhteet kustannuspaikoittain ja tileittäin tiliryhmiltä resursseille, toiminnoille ja välisuoritteille
- Toimintatiedot: päätökset, kontaktit, hoitopäivät ja -jaksot aikaleimoiheen

Iteratiivinen kustannuslaskentamalli ja -prosessi on esitetty kuvassa 18. Malli yhdistää toiminnan ja talouden tiedot ja hinnoittelee malliin valitut välisuoritteet (ks. kuva 19) minuutti- ja kappalemääräperusteisesti. Välisuoritteet siirtävät sosiaali- ja terveydenhuollon organisaation toiminnasta aiheutuneet kulut oikeassa suhteessa mallin läpi ajetuille toiminnan perussuoritteille eli asiakkaan saamille palveluille. Mallia tulee päivittää säännöllisesti organisaation toiminnan kehittyessä. Mallin läpi on mahdollista ajaa hinnoiteltavaksi uusi jakso asiakkaille tarjottuja toiminnan perussuoritteita. Perussuoritteet puolestaan toimivat lähtötietona episodianalytiikalle, joka järjestää nämä perussuoritteet asiakkaittain episodeiksi ja edelleen episodit segmentteihin asiakasryhmittäin. Jos yksityiskohtaista kustannuslaskentaa ei ole suoritettu läpi organisaation kaikille palveluille, tulee keskiarvohinnat palveluittain tuottaa jollain muulla laskentatavalla ja tuoda episodianalytiikan lähtötiedoksi.



Kuva 18 Tuotteistettu kustannuslaskennan prosessimalli (Prodacapo Finland Oy)

Todellisiin minuutteihin perustuvat välisuoritteet:	Kappalemääriin perustuvat välisuoritteet:
1. Fysioterapeutin käynti	1. Lounasateria 1
2. Lähihoitajakäynti	2. Lounasateria 2
3. Sairaanhoidajakäynti	3. Lounas palvelukeskus
4. Kotihoidon asiointipalvelu–Lähihoitaja	4. Peruskustannus per käynti
5. Kotihoidon asiointipalvelu–Sairaanhoidaja	
6. Lähihoitajan saattajapalvelu	
7. Sairaanhoidajan saattajapalvelu	
8. Turvpuhelin-asennus–Lähihoitaja	
9. Turvpuhelin-asennus–Sairaanhoidaja	
10. Turvpuhelin hälytyskäynti–Lähihoitaja	
11. Turvpuhelin hälytyskäynti–Sairaanhoidaja	
12. Turvpuhelin hälytyskäynti viikonloppu–Lähihoitaja	
13. Turvpuhelin hälytyskäynti viikonloppu–Sairaanhoidaja	
14. Turvpuhelin hälytyskäynti maksuton–Lähihoitaja	
15. Turvpuhelin hälytyskäynti maksuton–Sairaanhoidaja	

Kuva 19 Kustannuslaskennassa määritetyt välisuoritteet (Prodacapo Finland Oy)

13.12.2019

Pilotissa koodattiin viisi pääsegmenttiä. Lisäksi segmentti ”kaikki kotihoidon asiakkaat” jaettiin neljään alasegmenttiin (ks. kuva 20).

Projektissa keskityttiin tarkastelemaan Kainuun maakunnan yli 75-vuotiaiden kotihoidon asiakkaiden prosesseja yhdestä perussuoritelähteestä viiden episodisegmentin kautta:

- 01-Kaikki asiakkaat
 - Poimitaan kaikkien yli 75-vuotiaiden asiakkaiden suoritteet
- 02-Kaikki KH-asiakkaat
 - Poimittavat asiakasryhmät:
 1. Säännöllisen kotihoidon asiakkaat
 2. Tilapäisen kotihoidon asiakkaat
 3. Tukipalveluasiakkaat perustuen palveluvaimiin
 4. Kaikki muut asiakkaat (yli 75-vuotiaat palveluja saaneet, jotka eivät kuulu mihinkään em. ryhmistä)
- 03-Muistisairaat
 - Poimitaan kaikki yli 75-vuotiaat asiakkaat, joilla esiintyy muistisairautta osoittava ICD-10-diagnoosikoodi. Episodinmuodostuksessa hyödynnetään THL:n määrittelemiä muistisairauksien pää- ja aladiagnoseja.
- 04-Päivystystä käyttäneet
 - Poimitaan kaikki yli 75-vuotiaat asiakkaat joilla esiintyy vähintään yksi päivystyskäynti
 - ESH-vastaanottolaji: Päivystyskäynti tai tilastoitumaton päivystyskäynti
 - PTH-vastaanottolaji: Päivystysvastaanotto
- 05-Paljon palveluja käyttäneet
 - Poimitaan yli 75-vuotiaat asiakkaat joilla vähintään 20 tapahtumaa (suoritetta) / vuosi

Kuva 20 Pilotissa määritetyt episodisegmentit (Prodacapo Finland Oy)

4.4.3 Tuloksia

Kuvassa 21 on Costing-sovelluksen näkymä yksikköhinnoiteltuihin Kainuun soten kotihoidon välisuoritteisiin organisaatioyksiköittäin ja niiden perussuoritteille kohdennettavissa oleviin yksikkömääriin. Valittuna on yhden organisaatioyksikön kotihoidon lähihoitajakäynti, jota on jaettavissa perussuoritteille noin 6,24 miljoonaa minuuttia ja vastaavasti 5,27 miljoonaa euroa. Välisuoritteen lähihoitajakäynti yksikkökustannus (kustannus per minuutti) on moninkertainen päivätoiminnan osalta saatuihin yksikkökustannuksiin verrattuna. Tätä selittää päivätoiminnan tarjoaminen pääsääntöisesti ryhmille.

Välisuoriteen kuvaus

Organisaatiotyyppi

Organisaatioyksikön tunnistus

Välisuoriteen tunnistus

Yksikkö

Määrä

Kokonaiskustannus

Yksikkökustannus

ATERIA

FYSIOTERAPEUTTI

LAHIHOITAJA

PERUS

SAIRAANHOITAJA

Välisuoriteen kuvaus: Fysioterapeutti käynti

Välisuoriteen kuvaus: Hammashoitaja- tai hammashoitaja käynti

Välisuoriteen kuvaus: Kotihoidon asiantuntijapalvelu

Välisuoriteen kuvaus: Lounas palvelukeskus

Välisuoriteen kuvaus: Lounasateria 1

Välisuoriteen kuvaus: Lounasateria 2

Välisuoriteen kuvaus: Lähihoitajakäynti

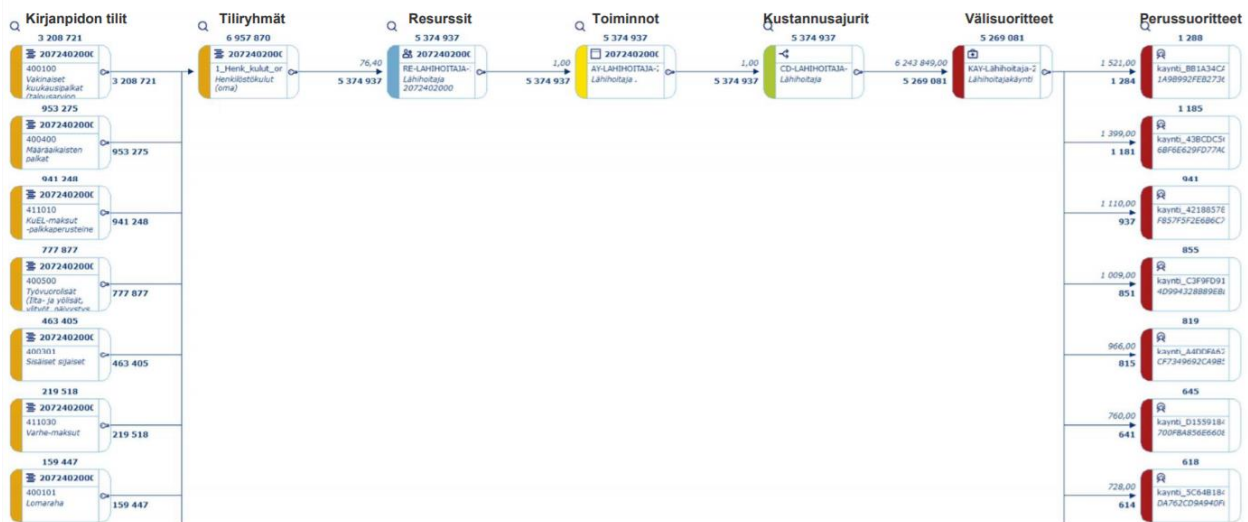
KOTIHOITO	2072402000	KAY-Lähihoitaja-2072402000	Minuutit	6 243 849,00	5 269 081,04	0,84			0,84	
KOTIHOITO	2073203000	KAY-Lähihoitaja-2073203000	Minuutit	2 756 009,00	2 182 916,33	0,79			0,79	
KOTIHOITO	2073207000	KAY-Lähihoitaja-2073207000	Minuutit	2 610 718,00	2 359 689,07	0,90			0,90	
KOTIHOITO	2074208000	KAY-Lähihoitaja-2074208000	Minuutit	2 246 209,00	2 037 401,36	0,91			0,91	
KOTIHOITO	2074204000	KAY-Lähihoitaja-2074204000	Minuutit	1 204 480,00	1 118 980,64	0,93			0,93	
KOTIHOITO	2074201000	KAY-Lähihoitaja-2074201000	Minuutit	920 943,00	744 855,27	0,81			0,81	
KOTIHOITO	2074202024	KAY-Lähihoitaja-2074202024	Minuutit	709 308,00	635 089,10	0,90			0,90	
KOTIHOITO	2074206000	KAY-Lähihoitaja-2074206000	Minuutit	525 312,00	447 800,19	0,85			0,85	
PAIVATOIMINTA	2076604000	KAY-Lähihoitaja-2076604000	Minuutit	431 010,00	35 879,20	0,08			0,08	
PAIVATOIMINTA	2076608000	KAY-Lähihoitaja-2076608000	Minuutit	285 690,00	32 827,90	0,11			0,11	
PAIVATOIMINTA	2076603000	KAY-Lähihoitaja-2076603000	Minuutit	255 450,00	37 013,02	0,14			0,14	
PAIVATOIMINTA	2076602000	KAY-Lähihoitaja-2076602000	Minuutit	240 610,00	34 304,49	0,14			0,14	
PAIVATOIMINTA	2076607000	KAY-Lähihoitaja-2076607000	Minuutit	234 360,00	33 565,98	0,14			0,14	
PAIVATOIMINTA	2076601000	KAY-Lähihoitaja-2076601000	Minuutit	46 610,00	7 152,12	0,15			0,15	
PAIVATOIMINTA	2076606000	KAY-Lähihoitaja-2076606000	Minuutit	42 025,00	7 152,15	0,17			0,17	
KOTIHOITO	2072402030	KAY-Lähihoitaja-2072402030	Minuutit	140,00	0,00	0,00			0,00	
					Summa=14 983 707,84					

Kuva 21 Välisuoritteille kohdistetut kokonais- ja yksikkökustannukset (Prodacapo Finland Oy, Costing-sovellus)

13.12.2019

Kuvissa 22, 23 ja 24 on käyty esimerkein läpi kustannusten kohdentumista ja jäljittämistä kustannuslaskentaprosessin läpi. Esimerkeissä on käsitelty jo kuvassa 21 esiin nousutta yksittäistä välisuoritetta lähihoitajakäynti.

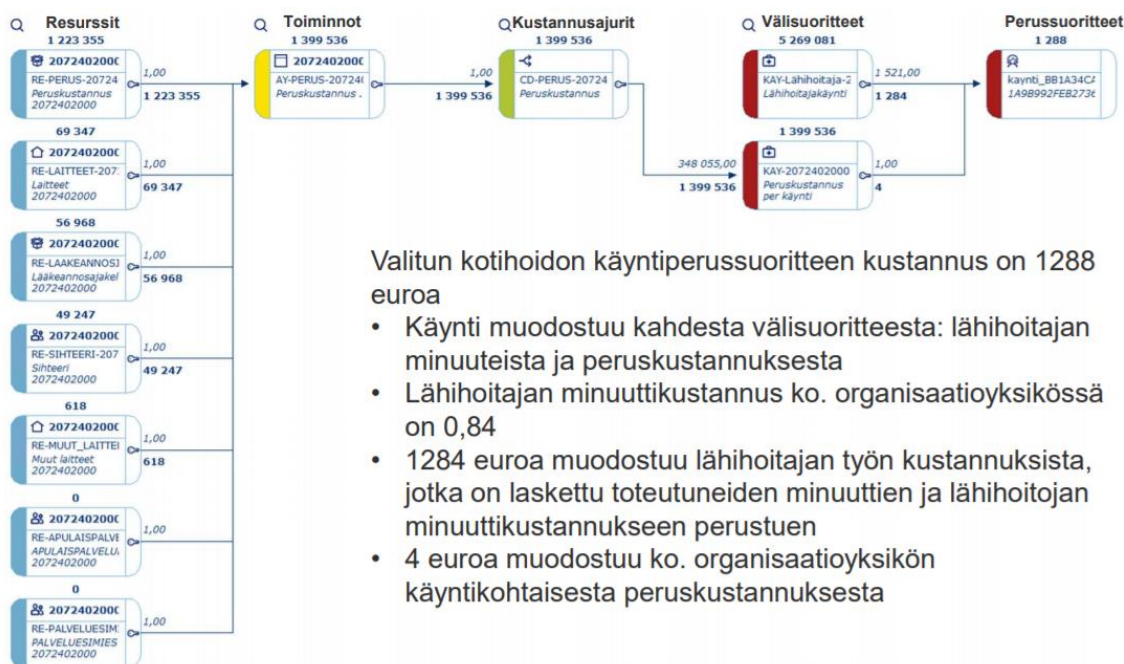
Kuvan 22 esimerkissä nähdään miltä kirjanpidon tileiltä tiliryhmien, resurssien ja toimintojen kautta kyseisen välisuoritteiden aiheuttamat kustannukset tulevat. Toisaalta taas nähdään mille asiakkaille tarjotuille perussuoritteille (palveluille) tätä kyseistä välisuoritetta (lähihoitajan käynti) on ohjattu perussuoritteiden kustannusvaikutuksen toteen näyttämiseksi. Perussuoritteista on kaaviossa näkyvissä vain murto-osa, koska välisuoritetta on niille jaettavana yhteensä noin 5,27 miljoonaa euroa.



Kuva 22 Suoritelaskennan laskennan jäljitys (esimerkki 1). Lähihoitajan minuuttikustannuksen muodostuminen ja kohdistuminen perussuoritteille (Prodacapo Finland Oy, Costing-sovellus)

Kuvan 23 esimerkissä nähdään edelleen samaan tapaukseen liittyen, miten asiakkaalle tarjottuun perussuoritteeseen lisätään lähihoitajakäynnin lisäksi myös käyntikohtaista peruskustannusta neljä euroa välisuoritteena. Tämä kohdennettavissa oleva välisuorite kokonaisarvoltaan noin 1,40 miljoonaa euroa on johdettavissa kuvan vasemmassa laidassa listatuilta varsinaista toimintaa tukevilta resursseilta, joita ovat tilat, laitteet, materiaalit ja tietyt yhteiset henkilöstöresurssit kuten palveluesimies.

13.12.2019



Kuva 23 Suoritelaskennan laskennan jäljitys (esimerkki 2). Kotihoidon käyntiperussuoritteen kustannuksen muodostuminen (Prodacapo Finland Oy, Costing-sovellus)

Kuvan 24 esimerkissä nähdään edelleen samaan asiaan liittyen miten ko. organisaatioyksikössä lähihoitaja-resurssia arvoltaan yhteensä noin 5,37 miljoonaa euroa käytetään muuhunkin kuin vain välisuoritteelle lähihoitajakäynti, jonka kohdistettavissa oleva potti on 5,27 miljoonaa euroa. Kuvan oikeassa reunassa on listattu kaikki välisuoritteet, joille tässä yhteydessä on mallinnettu ja kohdistuu käyttöä. Toki suurin osa ko. resurssista kuluu varsinaisilla lähihoitajakäynneillä, mutta osa työstä menee asiointi-, asennus- ja hälytyskäynneille.



Kuva 24 Suoritelaskennan laskennan jäljitys (esimerkki 3). Lähihoitajan työn kohdentuminen eri välisuoritteille (Prodacapo Finland Oy, Costing-sovellus)

Kuvassa 25 on siirretty hoitoketjuanalytiikan puolelle ja listattu vuoden 2018 tuotantoaineiston pohjalta saadut tunnusluvut kaikille määritetyille episodisegmenteille. Pääsegmenttejä 75 vuotta täyttäneille on viisi:

13.12.2019

- 01 – kaikki asiakkaat
- 02 – kaikki kotihoidon (kh) asiakkaat
- 03 – muistisairaat
- 04 – päivystystä käyttäneet
- 05 – paljon palveluja tarvitsevat

Lisäksi segmentit ”02 – kaikki kotihoidon (kh) asiakkaat” ja ”04 – päivystystä käyttäneet” on jaettu lisäsegmentteihin analyysiä varten. Episodisegmenttien tunnusluvuista löytyvät segmenttikohtaiset asiakasmäärät, episodimäärät, suoritelmäärät, kokonaiskustannukset, kustannukset per episodi, episodin keskimääräinen läpimenoaika ja hoitopäivät. Episodien keskimääräinen kesto on taulukon mukaisesti vaihdellut episodisegmenteittäin vajaan puolesta vuodesta lähes puoleentoista vuoteen. Yksittäisen asiakkaan yksittäinen hoitoketju eli episodi voi kuulua useampaan episodisegmenttiin. Määritellyt episodisäännöt saattavat nostaa yksittäisen episodin useaan episodisegmenttiin; esimerkiksi paljon palveluja tarvitseva asiakas on voinut käyttää myös päivystystä tarkastelujaksolla ja episodi voi kuulua näin ollen episodisegmentteihin ”paljon palveluja tarvitsevat” ja ”päivystystä käyttäneet”.

	Asiakkaat	Episodit	Suoritteet	Kustannukset yhteensä	Kustannukset / episodi	Läpimenoaika keskimäärin, vrk	Päiviä yhteensä
01 - Kaikki asiakkaat	10 683	10 683	1 074 560	67 720 069	6 339	445	42 958
02 - 01 Säännöllisen kotihoidon asiakkaat	2 052	2 052	866 104	32 857 157	16 012	503	17 830
02 - 02 Tilapäisen kotihoidon asiakkaat	2 896	3 788	610 353	32 939 474	8 696	329	34 964
02 - 03 Tukipalvelusasiakkaat	1 847	2 127	746 007	29 264 755	13 759	467	10 329
02 - 04 Kaikki muut asiakkaat	3 873	5 035	918 333	41 802 510	8 302	325	39 408
02 - 05 Muut kuin säännöllisen KH-asiakkaat	10 668	10 668	1 074 590	67 718 395	6 348	445	42 958
02 - Kaikki kh-asiakkaat	3 700	3 700	932 696	44 551 533	12 041	504	39 960
03 - Muistisairaat	913	913	242 092	10 321 422	11 305	510	12 003
04 - 01 Säännöllisen kotihoidon asiakkaat päivystys indx tapahtuma ei indx jälk	1 537	5 226	606 672	23 781 062	4 551	124	13 769
04 - Päivystystä käyttäneet	6 605	6 605	886 147	54 909 505	8 313	478	35 177
05 - Paljon palveluja tarvitsevat	4 452	4 452	210 804	35 764 600	8 033	316	32 000

Kuva 25 Episodisegmenttien tunnusluvut (Prodacapo Finland Oy, Region-sovellus, Qlik Sense -työkalu)

4.4.4 Havaintoja

Kustannuslaskennan pilotti toi esiin asioita, jotka on hyvä pitää mielessä kun sote-palvelujen järjestämistehtävää varten ryhdytään alueilla rakentamaan lisää kustannuslaskennan toiminnallisuutta eri palvelukokonaisuuksille. Nämä havainnot liittyvät etupäässä kustannuslaskennan kehittämisen prosessiin. Kustannuslaskennan toteuttamisen sisällöllisten tulosten hinta- ja määrävalidointi jäi pilotin kireästä aikataulusta johtuen suurelta osin keskeneräiseksi. Näin ollen kustannuslaskennan sisällöllisten tulosten pohjalta on pilotissa tehty vain hyvin rajallisesti johtopäätöksiä.

Mallinnuksen laajuus. Suoritelaskenta on suositeltavaa tehdä samassa laajuudessa episodianalytiikan kanssa. Ideaalitalanteessa suoritelaskenta on tehty alueen koko sote-toiminnan laajuisena, mutta työmäärällisesti tämä vaatii todella isoa resursointia. Täytyy muistaa, että suoritelaskennan yhdistämä talouden ja toiminnan tieto eli asiakaskohtaisesti hinnoitellut perussuoritteet (asiakkaan saama palvelu) muodostavat perustan episodianalytiikalle, jonka syötetietona toimivat juuri perussuoritteet. Jos kustannuslaskentamalli ja sen pohjalta suoritelaskenta tehdään suppeammin kuin episodianalytiikka, tulee episodianalytiikassa väistämättä eteen puuttuvan suoritelaskennan ja sitä kautta kustannustiedon korvaaminen muilla, kevyemmällä kustannusten laskenta- tai arviointimenetelmillä. Kainuun pilotissa vanhuspalvelujen kotihoidon ulkopuoliset suoritteet piti hinnoitella kustannuslaskentamallin ulkopuolella.

Mallinnuksen tarkkuus. Mitä tarkempi malli, sitä enemmän sen kehittäminen ja ylläpitäminen vie aikaa ja rahaa. Kustannuslaskentamallin tarkkuuden tulee olla tarkoituksenmukainen ja sidoksissa sote-palvelujen

13.12.2019

järjestämistehtävän tietojohdamisen tarpeisiin (strateginen johtaminen). Mallinnusta voidaan jatkaa eri yhteydessä ja eri resurssein tarkemmalle tasolle palvelemaan palvelujen tuottamistehtävän tarpeita (operatiivinen johtaminen). Toki mallin tulee olla rakenteeltaan sellainen, että se on yhteensopiva ylöspäin kansalliselle tasolle (ohjaus) ja alaspäin tuottamistasolle.

Mallinnuksen taloudellisuus. Pareto-periaatteen (20/80-sääntö) mukaisesti tulisi keskittyä ensin niihin alueisiin, joilla mallinnus tuottaa suurimman tuoton panostukseen nähden. Helpot voitot kehitystyön alkumetreillä myös motivoivat tiimiä jatkamaan kohti haastavampia osa-alueita. Esimerkiksi, työvoimavaltaisella sote-sektorilla palkkasummat näyttelevät suurta roolia palvelujen kokonaiskustannuksissa ja ne voi olla hyvä mallintaa ensimmäisenä, ja nämä tulisi saada ladattua automaattisesti henkilöstöhallinnon järjestelmästä.

Kustannuslaskenta mukaan canvas-ajatteluun. Pilotissa hyödynnettiin laajasti liiketoiminnan hypoteesicanvasta ja edistyneen analytiikan koneoppimisen canvasta. Kustannuslaskenta kannattaa kytkeä osaksi koko toiminnan ja talouden raportoinnin valmistelua. Canvas-ajattelu jäsentää monimutkaisia kokonaisuuksia ja niiden välisiä syy-seuraus-ketjuja sekä auttaa tavoitteiden viestimisessä ja kirkastamisessa pitkin matkaa.

Referenssien hyödyntäminen. Paikallisia ja alueellisia toteutuksia on tehty paljon eri välineillä myös sote-sektorilla. Toki niiden hyödyntämisessä päivittäisessä johtamistyössä on suuria eroja. Kokemuksia kannattaa vaihtaa eri toimijoiden kesken, jotta saadaan aikaan eväät rakentaa aikaa kestäviä ja mahdollisimman monia tahoja palvelevia yhtenäisiä kustannuslaskennan ratkaisuja ilman suuria oppirahoja.

Organisaation muutospisteiden mallinnus. Organisaatiot uudistuvat, kehittyvät ja muuttuvat joka vuosi. Kun organisaation rakenne, kustannuspaikat tai esimerkiksi tilikartta on muuttunut kustannuslaskennassa tarkasteltavalla ajanjaksolla, kannattaa kiinnittää huomiota mallinnuksen jatkuvuuteen näiden muutospisteiden yli. Tässä auttavat konversiotaulut ja muu dokumentaatio.

Ostopalvelujen suorit tiedot näkyviksi. Palvelujen järjestäjä saattaa ostaa hyvin suuren osan, jopa yli puolet tarjoamistaan palveluista yksityisiltä palveluntuottajilta suoraan ostopalveluina tai sitten välillisesti asiakkaille tarjottavien palvelusetelien muodossa. Näin ollen ostopalvelut ovat merkittävä osa kustannuskertymää ja kustannuslaskentaa. Jos laskutuksen yhteydessä ei saada toiminnan suorit tiedot, hukataan mahdollisuus kohdistaa ja jyvittää tämä merkittävä ostopalvelujen kustannuserä aiheuttamisperusteisesti toimintolaskennan keinoin tuotetuille palveluille. Silloin kustannuslaskenta jää auttamatta torsoksi. Ostopalvelujen laskutuksen yhteydessä tuleekin ohjeistaa yksityisiä palveluntuottajia toimittamaan laskun ohessa myös riittävä erittely yksittäisille suoritteille saakka. Ostopalvelutuotannon rakenteen tulisi avautua samalla tavoin näkyväksi kuin oman julkisen palvelutuotannon.

Suoritteiden kirjaamiskäytännöt. Suoritteiden kirjaamiskäytännöt vaihtelevat organisaatioissa ja tämä vaikuttaa suoraan kustannuslaskennan mahdollisuuksiin. Kustannuslaskentaa kehitettäessä tuleekin olla kattava näkemys organisaation vallitsevista suoritteiden kirjaamiskäytännöistä.

Organisaation tarjoama kehitysapu. Kustannuslaskentaa ei voi mallintaa ja toteuttaa puhtaasti ulkoapäin. Organisaation tulee sitoutua tähän tekemiseen ja varata siihen riittävät omat resurssit. Kustannuslaskennan työkalupakki, työtapo ja työnohjaus voi tulla ulkoa, mutta tietämys kohdeorganisaation toiminnasta on riittävän syvällistä vain talon omalla väellä. Samalla kun organisaatio sitoutuu tähän ponnistukseen, se myös nostaa omaa kyvykkyyttään kustannuslaskennassa ja rakentaa omistajuuden kustannuslaskennan itsenäiseen tuottamiseen, hyödyntämiseen ja jatkokehittämiseen.

13.12.2019

Integroitu projektin työnohjaus. Kaikki toimijat kannattaa ohjata käyttämään yhteistä projektin työnohjauksen alustaa. Tämä alusta liimaa asiantuntijoiden tekemiset toisiinsa ajallisesti ja sisällöllisesti, auttaa työjonon priorisoinnissa, työnjaossa ja viestinnässä. Yhteinen työnohjauksen alusta mahdollistaa tehokkaan työnteon myös hajautetussa kehitysorganisaatiossa.

Integroidut kustannuslaskennan osaprojektit. Jos kustannuslaskenta koostuu sopimuksellisista tai organisatorisista syistä johtuen useasta eri osaprojektista, kannattaa kuitenkin linkittää ne riittävällä määrällä koordinaatiota ja tarkistuspisteitä yhdeksi tavoitteelliseksi kokonaisuudeksi. Tämä lisää työnteon taloudellisuutta ja läpinäkyvyyttä.

4.4.5 Kustannuslaskennan vaihtoehtoisia malleja

Seuraavassa on esitelty vaihtoehtoisia malleja, joiden avulla kustannuksia voidaan suhteuttaa joko asiakkaiden tai asukkaiden määrään. Tämä mahdollistaa palveluiden järjestäjälle seurata eri palveluryhmien kustannuskehitystä pidemmällä ajanjaksolla.

Kustannusten kohdistaminen asiakkaille

Projektissa laskettiin kustannukset karkealla tasolla hyödyntäen kappaleessa 4.1 kuvattua palveluluokitusta. Kustannukset saatiin SAP-järjestelmästä tilitasolla. Ensimmäisessä vaiheessa laskettiin kunkin SAP-kustannuspaikan osalta kustannusten summa. Yhteen SAP-kustannuspaikkaan oli kytketty pääsääntöisesti yksi Lifecare järjestelmän kustannuspaikka. Vuoden 2018 tietojen osalta poikkeuksen muodosti kolme SAP-kustannuspaikkaa, joihin oli kytketty kaksi Lifecare-järjestelmän kustannuspaikkaa. Lifecare järjestelmässä kustannuspaikat jakaantuivat suorituspaikkoihin, joihin oli kirjattu käynnit ja hoitojaksot. Hoitojaksot purettiin hoitopäiviksi.

Jokaisen kustannuspaikan osalta laskettiin kaikkien Lifecare kustannuspaikkaan kytkettyjen suorituspaikkojen käyntien ja hoitopäivien määrät yhteen muodostaen kyseisen kustannuspaikan suoritemäärän. Tämän jälkeen jaettiin SAP-kustannuspaikan kustannukset siihen kytkettyjen Lifecare-järjestelmän kustannuspaikkojen suoritteiden lukumäärällä. Saatu tulos muodosti yksikkökustannuksen kyseisen Lifecare-järjestelmän kustannuspaikan osalta. Tämän jälkeen laskettiin kullekin henkilölle käyntien ja hoitojaksojen kustannukset hyödyntäen edellä laskettua yksikkökustannusta. Kun kustannukset oli kohdistettu henkilölle, niin tämän pohjalta pystyttiin poimimaan ja visualisoimaan yli 75-vuotiaiden aiheuttamat kustannukset kappaleen 4.6 mukaisesti asiakaskohtaisina kustannuksina.

Koska ostopalvelujen ja omaishoidon osalta projektissa tehtyjen rajausten pohjalta ei ollut käytössä tietoa kehen kustannukset kohdistuvat, esitettiin kustannukset näiden kustannuspaikkojen osalta visualisoinneissa kokonaiskustannuksina eli niissä ei ollut eroteltu yli 75 -vuotiaiden osuutta. Projektissa ei myöskään käytetty eri suorituspaikkojen osalta suoritteiden suhteellista painokerrointa, jolla olisi voinut saada mahdollisesti tarkempia tuloksia.

Kustannukset per asiakas on mahdollista laskea myös toisella tavalla. Kirjanpidon kustannuspaikat ja kaikki niiden kustannukset sijoitetaan edellä mainitun luokittelun mukaisille palveluille, ja jos yksi kustannuspaikka tuottaa useamman palveluluokituksen palveluita, niin prosentuaalinen arvio kunkin palvelun osuudesta kyseisen kustannuspaikan osalta. Näin saadaan kunkin palvelun kokonaiskustannukset, jotka jaetaan kyseistä palvelua käyttäneiden asiakkaiden lukumäärällä ja saadaan keskimääräinen kustannus per asiakas

13.12.2019

palveluittain. Tässä laskentatavassa ei pystytä kohdistamaan euroja oikeaan asiakkaaseen, mutta saadaan palveluittain keskimääräinen kustannus, jonka muutosta voidaan seurata ajanjaksoilla.

Kustannusten kohdistaminen asukkaille

Jos kustannuksia ei voida kohdistaa asukkaille, niin vaihtoehtoisesti kirjanpidosta voidaan koota kustannukset kustannuspaikoille. Tämän jälkeen voidaan määritellä erillisen taulukon avulla mitä palveluja kukin kustannuspaikka tuottaa, kappaleessa 4.1 kuvatun palveluluokituksen mukaisesti, ja jos yksi kustannuspaikka tuottaa useamman palveluluokituksen palveluita, niin prosentuaalinen arvio kunkin palvelun osuudesta kyseisen kustannuspaikan osalta. Tämän pohjalta saadaan laskettua palveluluokkien kokonaiskustannukset. Kun käytössä on väestörekisterikeskukselta saatu todellinen tai tilastokeskukselta saatu virallinen asukaslukumäärä, niin tällöin voidaan laskea kustannukset asukaskohtaisesti. Tämä mahdollistaa kunkin palvelun osalta kustannuskehityksen seuraamisen järjestäjän näkökulmasta. Jos kustannuspaikat voidaan edelleen kohdistaa tietylle alueelle, niin tällöin päästään vertailemaan kustannuksia alueittain.

4.5 Kehittynyt analytiikka ja koneoppiminen

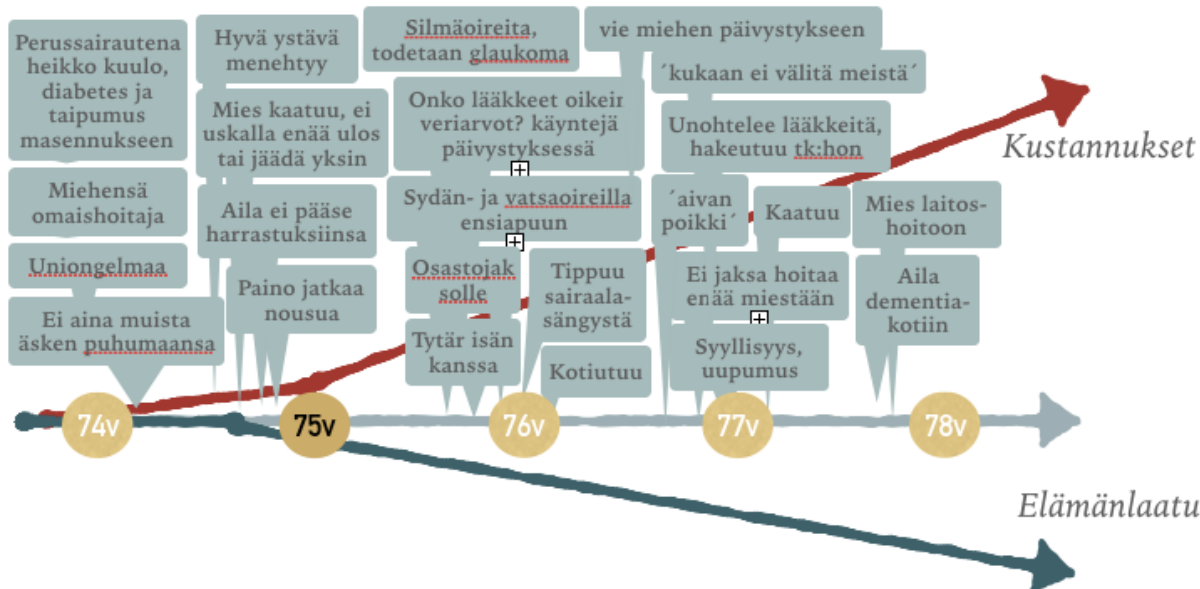
Kehittyneen analytiikan soveltamisen kohteeksi valittiin hypoteesicanvaksista hypoteesi numero 6: Palvelutarpeiden ennakointi ja riittävä resurssointi varhaisessa vaiheessa hillitsevät kustannuskehitystä.

4.5.1 Ongelman ja ratkaisun täsmentäminen

Mallinnusratkaisun kehityksessä keskeistä on perehtyä ratkaistavaa ongelmaa ympäröivään ilmiöön. Täsmennetään tavoite, tässä ei-toivottu lopputulema, jota halutaan ennakoida ja tunnistetaan datalähteet, jotka voisivat kertoa ilmiöstä. Kuvassa 26 kuvattu kiihtyvä palvelutarve kuvitteellisen 75-vuotiaan Ailan elämäntapahtumien kautta.

13.12.2019

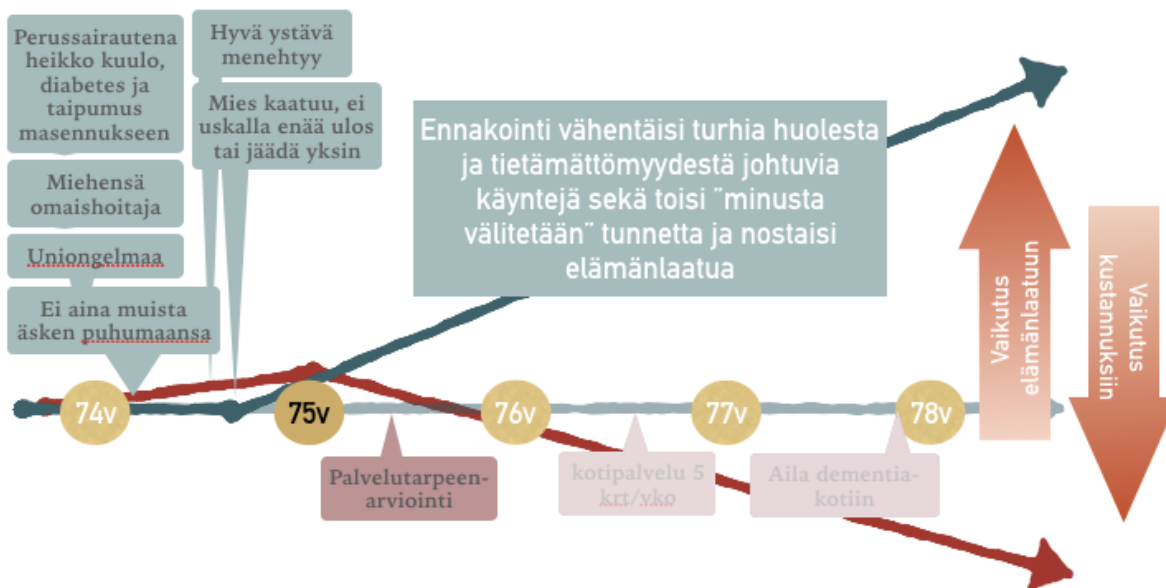
AILA 75V - KIIHTYVÄ PALVELUTARVE



Kuva 26 Kiihtyvä palvelutarve

Heti alussa on tärkeää tunnistaa myös arjen käytänteet ja miettiä miten tuloksia tullaan hyödyntämään (kuva 27), mitä tulosten hyödyntämisellä voitaisiin saavuttaa, miten tulosten hyödyntäminen vaikuttaisi olemassa oleviin prosesseihin ja miten tuotantokäyttö toteutettaisiin.

AILA 75V - ENNAKOINNILLA VAIKUTTAVUUTTA



Kuva 27 Ennakoinnilla vaikuttavuutta

13.12.2019

Tavoite on tärkeä rajata täsmällisesti ja kuvata selkeästi niin ajallisesti kuin sisällöllisesti, esim. ”kuka on ohjautumassa 12 kuukauden sisällä säännöllisiin palveluihin” tai ”kuka vaarassa siirtyä 12 kuukauden sisällä Maple 5 -luokkaan”.

Yleensä on tunnistettavissa optimaalinen kohta tai sweet spot (kuva 28), jossa ei-toivottuun tapahtumaan on optimaalista vielä vaikuttaa – eli ei olla liian peruuttamattomassa tilanteessa, mutta dataa on jo alkanut kertyä, jotta tapahtuman varhaiset kehitysvaiheet ovat datasta jo tunnistettavissa.

ENNAKOINNIN ”SWEET SPOT”



Kuva 28 Ennakoinnin sweet spot

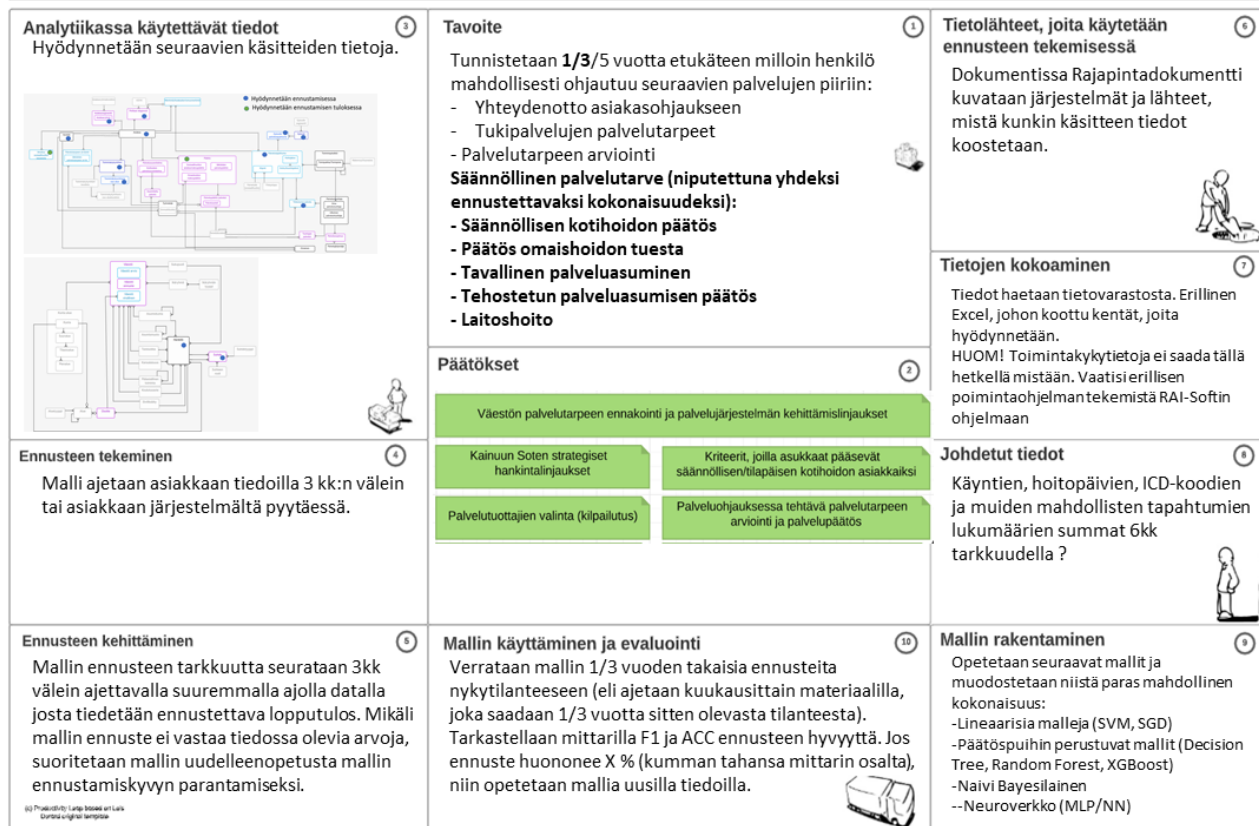
4.5.2 Koneoppimisen canvas – ajantasainen kokonaiskuva

Tässä projektissa edistyneen analytiikan työn tavoitetta analysoitiin kehittyneen analytiikan canvaksen avulla. Ennustemallin tavoitteeksi asetettiin tunnistaa etukäteen ja ajoissa henkilön ohjautuminen tiettyihin säännöllisiin palveluihin. Tuotettavan tiedon pohjalta voidaan esimerkiksi lisätä ymmärrystä ja tehdä väestötason linjauksia tulevasta palvelutarpeesta, suunnitella kohderyhmäkohtaisia palveluja ja ennaltaehkäiseviä toimia sekä pyrkiä siirtämään säännöllisiin palveluihin ohjautumisen ajankohtaa. Tämän tyyppisillä päätöksillä on merkittävä vaikutus kustannuskehitykseen, mutta mahdollisesti myös asukas- ja asiakastytytyvyyteen. Jatkossa ennuste voidaan päivittää ja ennustevertailulla on mahdollista päästä kiinni mm. interventioiden kustannusvaikuttavuuteen.

13.12.2019

6. Palvelutarpeiden ennakoiti ja riittävä resursointi varhaisessa vaiheessa hillitsevät kustannuskehitystä

productivity leap



Kuva 29 Kehittyneen analytiikan canvas

4.5.3 Tavoite – päätetapahtuma

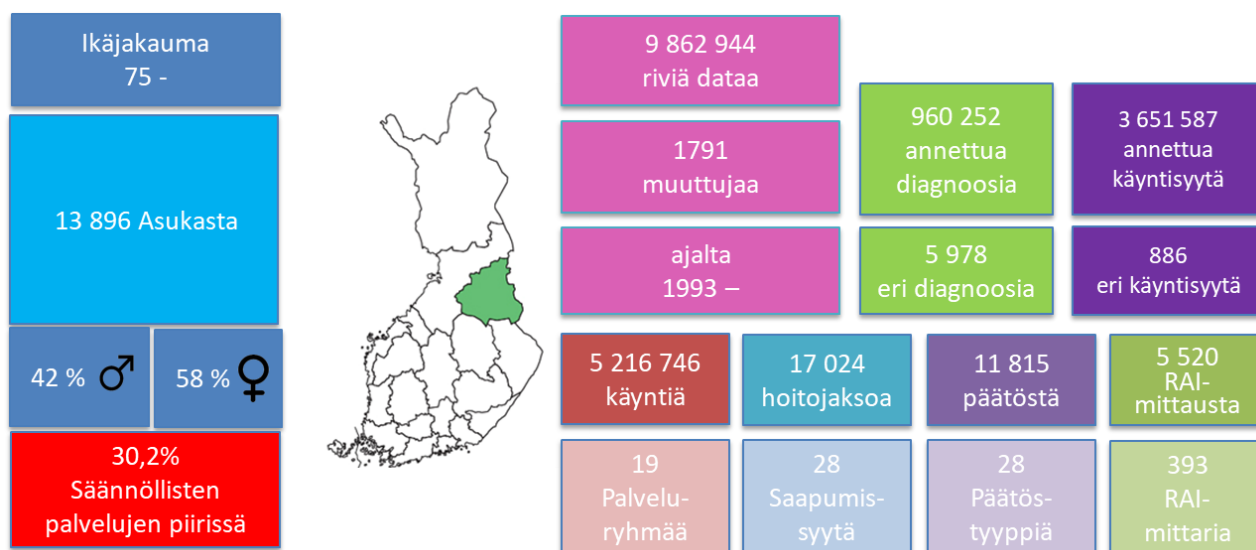
Analysoinnin pohjalta tavoitteeksi (kuva 29) täsmentyi tuottaa malli, jolla tunnistetaan 12 ja 36 kuukautta etukäteen milloin henkilö mahdollisesti ohjautuu seuraaviin palveluihin eli saa jonkin seuraavista päätöksistä:

- Säännöllinen kotihoito
- Omaishoidon tuki
- Tavallinen palveluasuminen
- Tehostetun palveluasuminen
- Laitoshoido

4.5.4 Tietosisällöt, rajaukset ja havaintoja datasta

Opetuksessa käytettävissä ollut tietosisältö on esitetty infograafi-muodossa kuvassa 30.

13.12.2019



Kuva 30 Koneoppimisen mallin opetuksessa ja testauksessa käytetty aineisto.

Data-aineistoon poimittiin Kainuun soten alueelta kaikki 75-vuotiaat elossa olevat henkilöt sekä kaikki 75-vuotiaat, jotka ovat päätyneet säännöllisiin palveluihin tarkastelujaksolla, vaikka olisivat olleet jo tarkasteluajankohdan hetkellä kuolleita. Mallinnuksessa on hyödynnetty kaikkien poimittujen asukkaiden osalta data sen jälkeen, kun he ovat täyttäneet 65 vuotta. Henkilöille, joilla oli säännöllisen palvelun päätös, poimittiin tiedot vuotta tai kolmea vuotta ennen päättymistä palveluun (riippuen ennustettavasta tavoitteesta). Niiden henkilöiden osalta, joilla ei ollut säännöllisen palvelun päätöstä, tiedot poimittiin ajankohtaan kuusi kuukautta ennen lataushetkeä. Säännöllisissä palveluissa oli näillä rajauksilla 4 383 henkilöä.

Ennusteen pohjana käytettiin tietovarastoon käsitemallin pohjalta kerättyä tietoa. Kainuun sote pitkän historian ansiosta yhtenäistä tietoa on hyvin käytettävissä. Kuvassa 31 on esitetty kunkin päätietoryhmän osalta se mistä vuodesta eteenpäin tietoa saatiin koneoppimisen mallin aineistoon.

Tietoryhmä	Tietoja ajalta
Diagnoosit (ICD-10)	1993 -
Käynti	2003 -
Käyntisyy (ICPC-2)	2011 -
Hoitojakso	2000 -
Toimintakykyluokitus	2012 -
Päätökset	2003 -
Perustiedot	Voimassa oleva tilanne

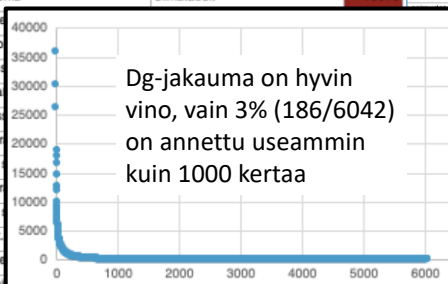
Kuva 31 Koneoppimisen data-aineiston laajuus.

Edistyneessä analytiikassa dataa hyödynnetään uudella tavalla ilmiöiden ymmärtämisessä. Toiminnassa kertynyttä dataa voidaan hyödyntää paitsi yksilötason ennusteissa ja ryhmätason ymmärryksen lisäämisessä, myös interventiodien vaikuttavuustarkastelussa. Samalla tämä asettaa uusia odotuksia kirjauskäytännöille ja niiden harmonisoinnille.

13.12.2019

Diagnoosit ICD-10

ICD10kodi	diagnoosilyhenne	diagnoosisuomeksi	tautiryhmälyhenne	tautikylälyhenne	ikm	sairaudet
I10	Essentiaalinen verenpainetauti	Essentiaalinen (primaarinen) verenpainetauti	Verenpainsairaudet	Verenkiertoelinten sairaudet	35956	149
I48	Eteisvärinä tai eteislepat	Eteisvärinä tai eteislepat	Muut sydänsairaudet	Verenkiertoelinten sairaudet	30331	418
H25.1	Vanhuden tumakaihi	Vanhuden tumakaihi	Myöskin sairaudet	Silmätaudit	26192	1
I25.1	Ateroskleroottinen sydänsairaus	Ateroskleroottinen sydänsairaus	Iskeemiset sydänsairaudet	Verenkiertoelinten sairaudet	18814	1229
Z97.4	Aikais. asenn. ulk. kuulolaite	Aikaisemmin asennettu ulkoinen kuulolaite	Anamneesiin liitt. terveysvaar	Muita tekijöitä	17975	2185
H40.10	Kaps.glaukooma,eksfolaat-sdr	Kapsulaarinen glaukooma ja ekxfolaatio-oireyhtymä	Glaukooma	Silmätaudit	16615	63
R06.0	Hengenahdistus	Hengenahdistus	Verenki			16
H40.13	Pitkäaik. avokulmaglaukooma	Pitkäaikainen avokulmaglaukooma	Glaukooma			113
E11.9	Aik. diabetes ei komplikaatioit	Aikuistyyppin diabetes ilman komplikaatioita	Diabetes			218
I70.2	Raajojen valtim. ateroskleroosi	Raajojen valtimoiden ateroskleroosi	Perif. va			32
J18.9	Määrittämätön keuhkokuume	Määrittämätön keuhkokuume	Influenssa			3
M17.1	Muu prim. polven nivelrikko	Muu primaarinen polven nivelrikko	Nivelsair			129
N40	Eturauhasen benign. liikakasvu	Eturauhasen hyvänlaatuisen liikakasvu	Miehen			3
M05.8	Muu seroposit. nivelreuma	Muu seropositivinen nivelreuma	Nivelsair			2
C61&	Eturauhassyöpä	Eturauhassyöpä	Miehen			4
J45.9	Määrittämätön astma	Määrittämätön astma	Pitkäaik.			1
R05	Yskä	Yskä	Verenki			32
M48.0	Selkärangan ahtauma	Selkärangan ahtauma	Selkäsai			35
M17.0	Prim. mol. polven nivelrikko	Primaarinen molempien polvien nivelrikko	Nivelsairaudet	Tuki&liik. elin ja sidek. sair	7276	



Kuva 32 Diagnoosien jakautuminen

Datan tutkimisvaiheessa havaittiin, että suuressa osassa tietoryhmiä eri kategorioiden jakauma oli hyvin vino eli osaa kategorioista/koodista on käytetty merkittävästi enemmän kuin toisia. Kuvassa 32 tilannetta on kuvattu diagnoosikoodien osalta. Merkittävää oli myös, että ne tarkemman tason koodit, joita oli käytetty, nousivat herkästi mallinnuksessa merkitsevien muuttujien joukkoon. Diagnoosi-, hoitojakso-, käyntisyy- yms. koodien valinnalla, kirjauskäytännöillä, on siten keskeinen merkitys mallinnustuloksiin.

4.5.5 Johdetut muuttujat

Tietovarastosta löytyvän raakadatan lisäksi tuotettiin asukaskohtaisia johdettuja muuttujia. Kaikkiaan kustakin opetusaineistoon mukaan otetusta yli 75-vuotiaasta henkilöstä tuotettiin 1 791 datapistettä/attribuuttia/johdettua muuttujaa. Raakadataa olivat esimerkiksi ikätiedot, toimintakykytiedot tai yksittäisten diagnoosien tiedot. Johdettuja muuttujia taas tekijät, joita muodostettiin useampaa eri tietoa yhdistämällä tai niistä johtamalla. Esimerkkinä johdetuista tiedoista on eri käyntisyyt ja niiden määrät viimeisen 6 tai 12 kuukauden aikana sekä toisaalta koko ajalta asiakkaan täytettyä 65 vuotta. Lisäksi kullekin tietoriville tuotettiin päätetapahtumatieto siitä, oliko henkilö ohjautunut säännöllisiin palveluihin 12 tai 36 kuukautta kyseisen ajankohdan jälkeen.

4.5.6 Mallinnus

Data-aineisto jaettiin kahteen osaan, josta 75 prosenttia käytettiin opetusaineistona ja 25 prosenttia testiaineistona. Opetusaineistolla koneoppimisen mallia opetettiin tunnistamaan henkilöt, jotka saivat säännöllisen palvelun päätöksen määrätyn ajanjakson sisällä. Testiaineistolla testattiin miten hyvin malli osasi ennustaa sitä, päätyikö asiakas säännöllisiin palveluihin vai ei, vertaamalla ennustetta tapahtuneeseen.

Mallinnusta tehtiin sekä toimintakykytietoa sisältävällä aineistolla että aineistolla, josta toimintakykytiedot puuttuvat. Vertailun tuloksena todettiin, että aineisto jossa toimintakykytiedot eivät ole mukana tuottaa parempia ennustetuloksia kuin toimintakykytiedot sisältävä aineisto.

13.12.2019

4.6 Tiedon raportointi ja visualisointi

Raportoinnin tarkoituksena on kertoa organisaation johdolle sekä toiminnan että talouden toteumaa, tilannekuvaa ja ennustetta ja lisäksi ennakoida tulevaa palvelutarvetta. Ylimmän johdon raportoinnissa organisaation tilaa tarkastellaan laajana kokonaisuutena yksittäisten tapahtumien sijaan. Hyvä johdon raportointijärjestelmä mahdollistaa tietoon porautumisen tarvittaessa, koska se auttaa hahmottamaan syitä lukujen taustalla. Eri näkökulmia ja tiedon tasoja hyödyntäen voidaan samaa toteutusta käyttää sekä palvelujen järjestäjän että palvelutuotannon tarpeisiin eli strategisessa ja operatiivisessa johtamisessa. Toisaalta hyvässä johdon raportoinnissa tulee pystyä tarkastelemaan eri muuttujia suhteessa toisiinsa. Samoin raportoinnin on mahdollistettava vertailu oman organisaation sisällä, mutta myös suhteessa muihin samankaltaisiin organisaatioihin ja toimintoihin.

Tiedon saanti tulee olla dynaamista, jatkuvasti päivittyvää, poimintoina koottavan tiedon sijaan, mikä puolestaan edellyttää osittain myös tiedonsaantiin liittyvien prosessienkin nopeuttamista, esim. laskutus. Edelleen raportointi ei voi olla pelkästään toteutuneiden tietojen tutkimista, vaan siihen tulee kyetä liittämään myös tulevaisuuden ennustaminen sekä ennustemallien ja erilaisten skenaarioiden laatiminen. Kaikessa tarkastelussa tulee olla dimensioina sekä oma toiminta että ostopalvelut, jolloin kokonaisuuteen saadaan näkyväksi koko toteutunut toiminta ja kaikki käytetyt eurot.

Järjestämistehtävän tasoisessa tiedossa tarkastellaan palvelukäyttöä ja niihin liittyviä kustannuksia pääsääntöisesti väestötasolla asiakassegmenteittäin. Palvelujen järjestäjä tarvitsee tietoa myös organisaatiohierarkian mukaisesti eri vastuualueiden näkökulmasta ja siten, että eri vastuualueet ja niiden eurot ja toiminta näkyvät erillisinä koko organisaation kokonaisuudessa. Myös tämä tieto koostuu sekä omaa toimintaa että ostopalveluita koskevasta tiedosta. Tieto on välttämätöntä paitsi vastuualueiden johtamisen, myös koko sosiaali- ja terveyspalveluiden strategisen johtamisen ja koko toiminnan yhteen sovittamisen kannalta.

Pilotin visualisoinneista yleisesti

Qlik Sense -raportit ja visualisoinnit ovat lähtökohtaisesti dynaamisia eli raporttien käyttäjällä on mahdollisuus tehdä itse valintoja mittareiden ja dimensioiden avulla. Oletusarvoisesti tehdyt valinnat vaikuttavat kaikkiin raporteilla näkyviin graafeihin ja taulukoihin. Tämä mahdollistaa hyvin moniulotteisen tietojen tarkastelun.

Kerrallaan näytettävä raportointikokonaisuuden eli yhden visualisointisivun sisältö ja rakenne tulee suunnitella etukäteen. Yhdelle sivulle mahtuu rajallinen määrä tiedon näyttämiseen liittyviä objekteja kuten graafeja, taulukoita ja KPI-lukuja. Tyypillisesti sivun rakenne on sellainen, että siinä on esimerkiksi yläreunassa tärkeitä mittareita KPI-lukuina ja pari, kolme graafia tai taulukkoa, joiden sisältöä voidaan muuttaa mittareita ja dimensioita muuttamalla. Aloitussivuna käytetään yleensä erillistä tilannekuvasivua, joka voi olla kokonaan tai osittain staattinen. Tilannekuvasivulla (kuva 33) näytetään tyypillisesti tietyn hetken tilanteen tärkeimmät luvut ja kuvaajat eli se pyritään toteuttamaan niin, että käyttäjä näkee siitä tilannekatsauksen yhdellä silmäyksellä. Tilannesivun lisäksi sovellus koostuu useammasta muusta visualisointisivusta, joilla voi tarkastella tarkemmin tilannesivulla esitettyjen yksittäisten lukujen tietoja ja joilla käyttäjä voi tehdä lisärajauksia eli tarkastella osajoukkojen tietoja.

13.12.2019

Tilannekuva: kustannukset yli 75-vuotiaat asiakkaat tammi-elokuu 2019

Kainuun sote



Kuva 33 Tilannekuva

Tiedot jaettiin pilotissa tarkastelunäkökulmien perusteella kokonaisuuksittain neljälletoista eri sivulle, joita olivat mm:

- Tilannekuva
- € per asukas
- Oma toiminta € per asiakas
- € per palvelu
- Kokonaismenot per asiakkuusryhmä
- Kotihoito ja tukipalvelut vs. muut asiakkaat
- Päivystyksestä ohjautuminen
- Ennustava analytiikka
- Hoitoisuus (RUG-22)
- Toimintakyky (MAPLe5)

Pilotissa oli käytettävissä vuoden 2018 tiedot kokonaisuudessaan ja vuodelta 2019 tammi-elokuu. Syyskuu 2019 rajattiin pois, koska syyskuun kustannukset eivät olleet vielä kattavasti käytettävissä, vaikka lähtökohtaisesti kolmen kvartaalin tarkastelu olisi ollut mielekkäämpää. Vertailu edelliseen vuoteen toteutettiin kuukausittaisella vertailulla tammi-elokuu 2018 vs. 2019, mutta käyttäjä pystyi myös vaihtamaan vertailtavan ajankohdan esimerkiksi yksittäiseksi kuukaudeksi. Raportointia varten raportointikantaan ja Qlik Senseen ladattiin pelkästään yli 75-vuotiaiden asiakkaiden tiedot.

Raportoinnin ytimessä olivat hypoteesikanvaksen rajausten mukaisesti yli 75-vuotiaat ja erityisesti kotihoidon asiakkaat (säännöllinen, tilapäinen ja tukipalvelu). Raportoinnissa päästiin pisimmällä oman toiminnan kuvaamisessa.

Yli 75-vuotiaiden asiakkaiden palvelujen käyttö

Projektissa lähdettiin liikkeelle kuukausitasolle summatuista yli 75-vuotiaiden asiakkaiden kustannus-, asiakas- ja suoritemääristä. Summatason tietojen käytössä tuli vastaan rajoitteena asiakaslukumäärien saatavuus. Koska tiedot oli summattu kuukausittain, ei todellisia asiakaslukumääriä ollut mahdollista laskea eri käyttötapauksissa oikein raporteilla esimerkiksi vuositasolla. Raportointivälineet, kuten pilotissa käytetty

13.12.2019

Qlik Sense, pystyvät parhaiten hyödyntämään dataa silloin, kun tarvittava data ladataan järjestelmään alimmalla mahdollisella tasolla ja annetaan välineen tehdä summaukset ja aggregoinnit automaattisesti käyttäjän tekemien rajausten mukaisesti. Tämän vuoksi varsinaisissa toteutetuissa pilotin visualisoinneissa käytettiin kahteen kertaan pseudonymisoitua asiakastasoista dataa, mistä oli poistettu yksilöivät tekijät, kuten ikä ja sukupuoli. Vaikka suurin osa datan käsittelystä tehtiin jo aiemmin ETL/ELT-latauksissa, niin Qlik Senseen latauksen yhteydessä datalle tehtiin vielä joitain raportoinnin edellyttämiä muunnoksia, kuten kaikkien muiden kuin kotikunnaltaan kainuulaisten asiakkaiden leimaaminen ulkokuntalaisiksi.

Keskeisimmät mittarit olivat toiminnan kokonaismenot euroina, oman toiminnan menot euroa/asiakas ja oman toiminnan asiakaslukumäärät. Kokonaismenoissa näytettiin kaikki kyseisen väestöryhmän palvelukäyttöön liittyneet kustannukset. Euroa/asiakas tietoon liittyen pystyttiin näyttämään vain oman toiminnan asiakaskohtaiset keskimääräiset kustannukset, koska omaishoidon ja ostopalveluiden asiakaslukumääriä ei saatu pilotissa raportoinnin käyttöön suoritettujen puuttumisen vuoksi. Muita mittareita olivat oman toiminnan euroa/suorite ja suoritteiden lukumäärä. Mittareissa olennaiseksi tarkastelunäkökulmaksi nousi kustannukset ja niihin liittyvä muutos edellisen vuoden vastaavaan ajankohtaan verrattuna. Pelkkä absoluuttisten kustannusten vertailu ilman suhteuttamista asiakas- tai suoritelmääriin tai niiden muutokseen ei kuitenkaan ole riittävää, koska isommilla kunnilla myös kokonaismenot ovat aina luonnollisesti pieniä kuntia suuremmat.

Keskeisimmät dimensiot olivat kunta, palvelu ja asiakkuus. Kuntatietona käytettiin asiakkaan kotikuntaa. Erillisen ryhmän muodostivat ulkokuntalaiset, jotka olivat käyttäneet Kainuun soten palveluja. Palveluiden tietoja raportoitin palvelukokonaisuuksittain organisaatioriippumattomalla rakenteella, jossa hyödynnettiin Sote-tietopakettien luokittelua, ks. kappale 4.1. Lisäksi pilottia varten muodostettiin kotihoidon palvelujen lähempää tarkastelua varten oma neljäloukkaainen asiakkuusryhmittely: säännöllisen kotihoidon asiakkaat, tilapäisen kotihoidon asiakkaat, tukipalveluasiakkaat ja muut asiakkaat, jotka eivät kuuluneet mihinkään kolmeen ensimmäiseen asiakasryhmään. Kotihoidon palvelujen tarkemmassa tarkastelussa jokainen asiakas luokiteltiin johonkin neljästä asiakasryhmästä kuukausitasolla. Asiakkuuksien palvelukäyttöä oli mahdollista tarkastella myös asiakasryhmien yhdistelmänä, esim. Säännöllisen ja tilapäisen kotihoidon sekä tukipalvelun asiakkaat vs. muut asiakkaat. Kansallista määrittelyä edellä kuvattuun asiakasluokitteluun ei tällä hetkellä ole.

Hoitoisuus (RUG-22) ja toimintakyky (MAPLe5)

Hoitoisuus (RUG-22) ja toimintakyky (MAPLe5) tiedot koostettiin manuaalisesti raporttina ja ladattiin kertalatauksena raportointikantaan, jossa kumpikin tieto liitettiin asiakasfaktaan. Tietoja raportoitin omilla välilehdillä mm. vertailemalla eri kuntia ja suhteessa palvelukäyttöön. Integraatio RAlsoft:iin tulisi toteuttaa, koska hoitoisuus ja toimintakykytiedot tuovat merkittävää lisäarvoa.

Päivystyksestä ohjautuminen

Pilotissa toteutettiin päivystyksen käyttöön liittyvä tarkastelu. Päivystyksestä ohjautuminen muodosti oman erillisen kokonaisuuden, koska ohjautumiseen liittyviä tietoja ei pilotissa liitetty asiakkuusryhmittelyyn. Visualisaation keskeisimpiä mittareita olivat päivystyskäyntien ja ohjautumisten lukumäärät ja prosenttiosuudet sekä hoitojaksojen mediaanikestot. Ohjautumista ja hoitojaksojen mediaanikestoja vertailtiin yksikkö, palvelu ja kuntatasolla. Päivystyskäynniksi tulkittiin käynti keskussairaalan yhteispäivystyksessä tai päivystyksellisesti terveysasemien vastaanotoilla. Yksikkötietona käytettiin kotiuttavaa yksikköä, koska mm. nopean diagnostiikan yksikön palvelukäyttö olisi sekoittanut laajemman tarkastelun. Olennaista oli tarkastella ohjautumisen jakautumista ja mediaanikestoja perusterveydenhuollon ja erikoissairaanhoidon sairaalapalvelujen välillä.

13.12.2019

Väestöennuste

Väestöennusteita ei tuotu tietovarastoon, joten väestötason mittareita ei pystytty toteuttamaan kattavasti. Pilotissa käytettiin Tilastokeskuksen väestön ennakkotietoa vuodelle 2019. Tiedot koostettiin Tilastokeskuksen sivuilta manuaalisesti ja ladattiin suoraan Qlik Senseen omana tietolähteenä. Väestöennuste yhdistettiin asiakasfaktaan, ajankohtaan ja kuntaan, jolloin tietoja voitiin näyttää pilotissa ajankohta- ja kuntadimensioiden suhteen. Väestöennusteen tärkein mittari oli euroa/asukas, jossa kustannukset olivat yli 75-vuotiaiden asiakkaiden kokonaismenot euroina.

Mikäli haluttaisiin joustavasti käyttää asukastason ja eri väestösegmenttien tietoja palvelutarpeiden ennakkoinnissa ja tietojen vertailussa, tulisi olla integraatio Väestörekisterikeskuksen järjestelmään tai asukasmäärät tulisi olla ladattavissa rajoituksetta Väestörekisterikeskuksen järjestelmästä. Asukasmäärät olisi hyvä saada vähintäänkin kuukausitasolla kunnittain sekä kuluvalle että edellisille vuosille.

Ennustava analytiikka

Ennustavan analytiikan tuloksista ladattiin raportointikantaan 12 ja 36 kuukauden ennusteet, mitkä liitettiin asiakasfaktaan. Segmenttitietoja ei pilotissa ehditty siirtää raportoinnin käyttöön. Analytiikan tulosten raportointi muiden tulostietojen raportoinnin yhteydessä tuottaa merkittävää lisäarvoa, joten analytiikan tulokset tulisi aina kytkeä raportointiin.

4.7 Sidosryhmätyö – yhteistyö eri tahojen kanssa

Projektin koordinaatioryhmän tehtävä oli katselmoida ja kommentoida pilottiprojektissa tehtyjä suunnitelmia ja lopputuloksia ja arvioida ja tuottaa näkemyksiä mahdollisia pilotointiprojektin jälkeisiä tulevia laajennuksia varten. Koordinaatioryhmään kuuluivat kaikki maakuntalaajuiset kuntayhtymät eli Etelä-Karjalan sosiaali- ja terveyspiiri (Eksote), Kymenlaakson sosiaali- ja terveyspalvelujen kuntayhtymä (Kymsote), Päijät-Hämeen hyvinvointiyhtymä (PHHYKY), Pohjois-Karjalan sosiaali- ja terveyspalvelujen kuntayhtymä (Siun Sote) ja Keski-Pohjanmaan sosiaali- ja terveyspalvelukuntayhtymä (Soite) sekä Sosiaali- ja terveysministeriö (STM) ja Terveiden ja hyvinvoinnin laitos (THL).

Jokaisella koordinaatioryhmän jäsenorganisaatiolla oli oma vastuualueensa, jonka kautta kyseinen toimija toi työhön asiantuntemustaan. STM:n edustaja toimi ohjausryhmässä ja STM:stä saatiin asiantuntemusta mm. toisiokäytön lain testaamiseen. Samoin Sote-tietopakettien käyttöä varten saatiin ohjausta STM:stä, jossa on kyseisen mallin omistajuus. THL:n rooli oli ohjata käsitteiden ja tietosisältöjen määrittelytyössä, jotta päästään mahdollisimman pitkälle kansallisesti yhteneväiseen malliin. Samoin kustannuslaskennan mallin suunnittelussa hyödynnettiin THL:n asiantuntemusta.

Kainuun pilotin aikana yhteistyö THL:n kanssa toteutettiin pääsääntöisesti sähköpostien ja puheluiden välityksellä sekä muutaman Skype-palaverin avulla. Tulevaisuuden osalta on jo käyty keskusteluja kestävämpien yhteistyömallien toteuttamiseksi. Erityisesti Toivo-ohjelman organisoinnissa on syytä suunnitella sekä resurssien että toimintatapojen näkökulmasta yhteistyön toteutusta ja huomioida osana Valtava- ja Virta-hankkeita.

Eksoten roolina oli kommentoida analytiikkaa ikääntyneiden palvelujen kehittämisessä ja tuoda asiantuntemusta tiedon hyödyntämiseen ja raportointiin. Kymsote juuri toimintansa käynnistäneenä arvioi kokonaisuuden toteutuskelpoisuutta äskettäin käynnistyneiden tai käynnistyvien organisaatioiden näkökulmasta. PHHYKY:ssä on toteutettu laaja hanke palveluohjauksen kehittämiseksi ja siksi heidän rooliin muotoutui tuoda asiantuntemusta asiakas- ja palveluohjauksen prosessien kehittämiseen.

13.12.2019

ikääntyneiden palveluissa. Siunsoten asiantuntemusta pyydettiin kustannuslaskentamallien kehittämiseen ja kommentointiin. Soitessa puolestaan on tehty työtä toimintälähtöisen kokonaisarkkitehtuurin kuvaamisessa ja kehitetty palveluprosesseja integraatio huomioiden. Näin ollen heidän roolikseen tuli tietojohdamisen kytkentä kuntayhtymän kokonaisarkkitehtuuriin sekä asiantuntijuus prosessien kehittämisestä ja integraatiosta.

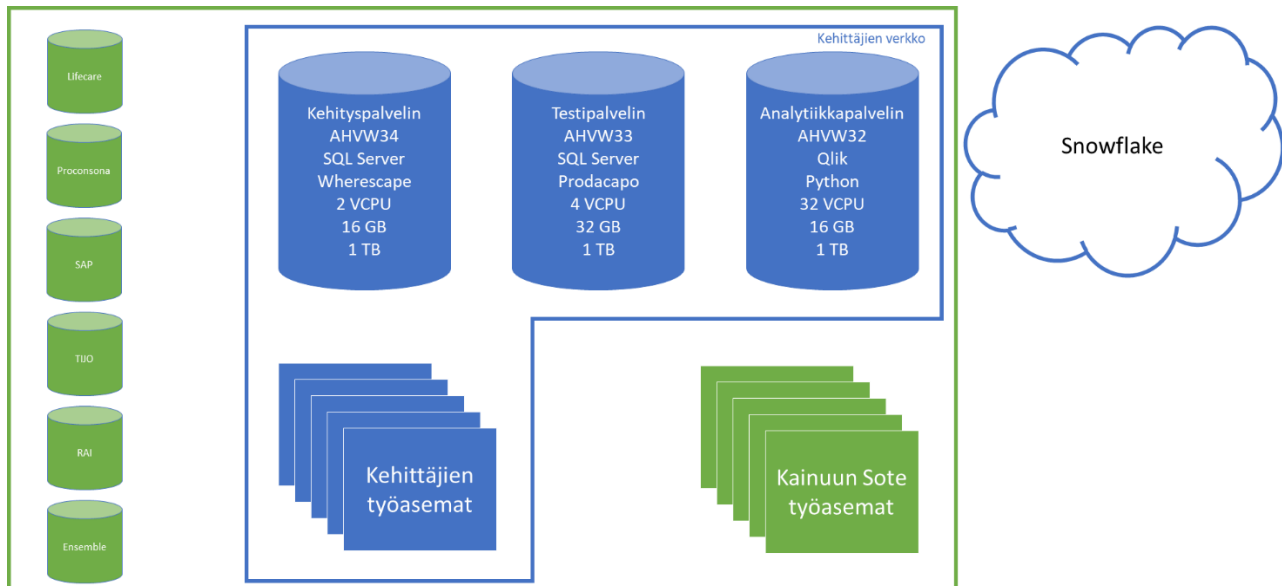
Edellä kuvatun kaltainen asiantuntijuuden jakaminen ja laajemman osallistujajoukon kuuleminen valmistelussa on merkittävää kansallisessa toteutuksessa. Koordinaatioryhmän jäseniltä saatiin arvokasta tietoa ja osaamista projektiin. Aikataulujen sovittelussa oli ongelmia, mutta jatkossakin tulee tavalla tai toisella varmistaa laaja osallistujajoukko mukaan projektien toteutukseen.

13.12.2019

5. Teknologia

5.1 Yleistä

Projektissa käytettiin Kainuun omaa ympäristöä ja lisenssejä soveltuvin osin. Sitä täydennettiin testikäyttöön hankituilla työkaluilla ja pilviympäristöllä. Ensisijainen kehitysympäristö ja -työkalut valittiin tunnetuista ja käyttöön soveltuvista välineistä, lisäksi mahdollisuuksien mukaan suoritettiin vertailua ja testausta muilla välineillä. Ajan ja resurssien rajallisuudesta johtuen kaikkea ei päästy vertailemaan projektin aikana.

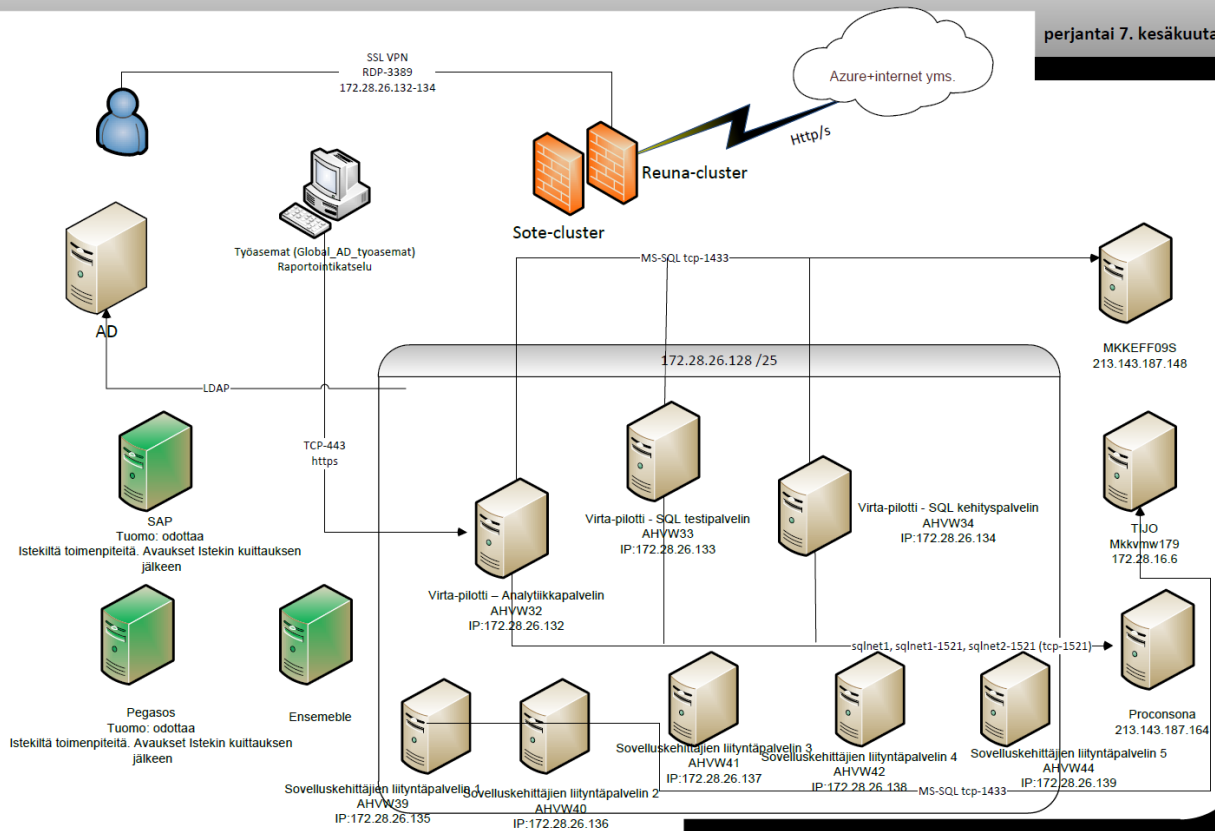


Kuva 34 Yleiskuva ympäristöstä

13.12.2019

Virta-pilotti, SoteDigi

perjantai 7. kesäkuuta 2019



Kuva 35 Yleiskuva verkosta

5.2 Välineet

Testaukseen valituista työkalut pyydettiin toimittajilta testikäyttöön veloituksetta tarjouspyynnön kautta. Ensisijaiset työkalut olivat:

- SQL Server: tietovaraston ja datamartien on-premise-ympäristössä oleva tietokanta
- Snowflake: datamartien pilviympäristössä oleva tietokanta
- Ellie: käsitemallinnus
- Wherescape: tietovaraston automaatio
- Qlik Sense: raportointi
- Prodacapo: kustannuslaskenta
- Python: edistynyt analytiikka, koneoppiminen, tekoäly
- Lisäksi muita testattuja välineitä olivat:
 - QPR ja Yhteentoimivuusalusta: mallinnus
 - Power BI: raportointi

13.12.2019

5.3 Tietokannan rakenne

5.3.1 Tietokannat ja skeemat

Ympäristöjä oli kaksi, kehitys ja testi, joissa kummassakin kaksi kantaa, load, stage ja raw data vaultin sisältävä VIRTADWDEV ja informartin alueet load, stage ja dm sisältävä VIRTADMDEV. Vastaavasti testiympäristössä kannat VIRTADWTEST ja VIRTADMTEST. Skeemoina kuten edellä mainittu DW-kannassa load, stage ja dv, DM-kannassa load, stage ja dm.

5.3.2 Data Vault-toteutus

Data Vault -toteutuksessa käytettiin WhereScape-tuotteen ominaisuuksia hyväksi mahdollisimman pitkälle parhaita käytäntöjä noudattaen. WhereScape-tuote sisältää kaksi erillistä välinettä, 3D:n, jolla tehdään tietokannan suunnittelu ja RED, jolla generoidaan 3D:llä tuotettujen kantataulujen latausskriptit ja muokataan tarvittavat ETL/ELT-osuudet ja orkestroidaan dataputkia.

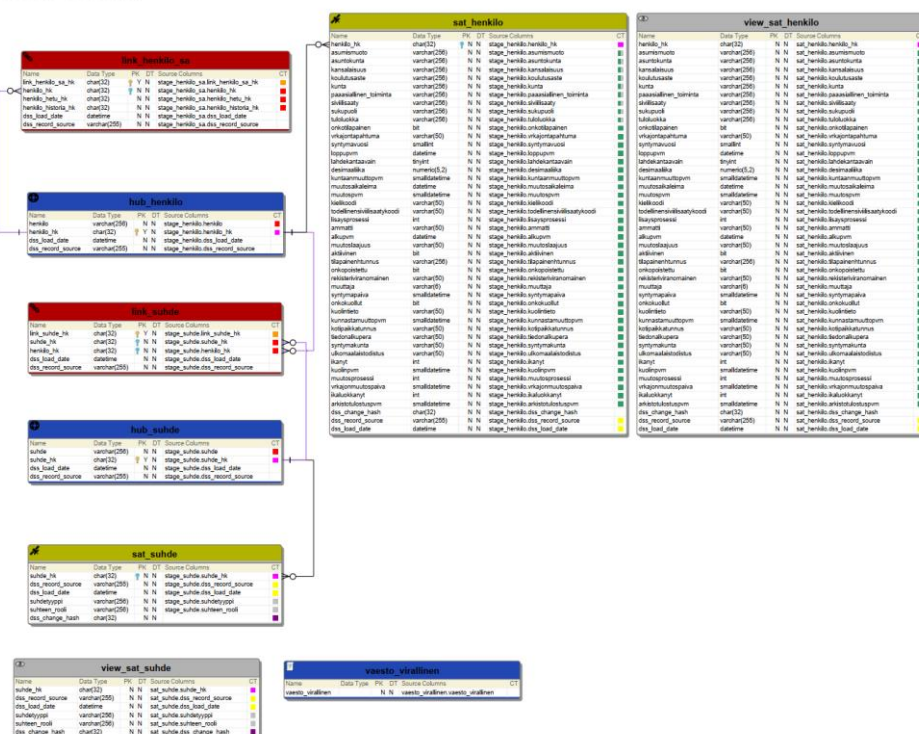
Ensimmäisenä työnä valmisteltiin tietovaraston lähtötiedot. Toteutus tapahtui WhereScape RED-välineellä. Kun lähtötiedot oli luotu load-alueelle tietovarastokantaan, niille tehtiin yksisuuntainen pseudonymisointi. Load-alueella oli siis sekä henkilötietoa, että pseudonymisoitua tietoa. Tietovaraston staging-alueelle ja siitä eteenpäin vietiin aina vähintään pseudonymisoitua tietoa.

Data Vaultin toteutus lähti liikkeelle Elliessä tuotetuista käsitelmalleista. Ellie-välineestä saatiin ladattua ulos mallit xml-tiedostoiksi, jotka vietiin WhereScape 3D-välineeseen. Käsitelmallin käsitteen tyyppi (master, sopimus, tapahtuma, referenssi) on tärkeä ominaisuus, jonka perusteella WhereScape osaa generoida Data Vault -kantaan oikeat taulutyypit.

Kun käsitelmalli oli tuotu WhereScape 3D:hen, siitä konvertoitiin looginen malli perustuen etukäteen laadittuihin konversiosääntöihin. Sääntöjen perusteella loogisessa mallissa jokaisesta käsitteestä muodostui oma taulu, joka sisälsi pääavaimet eli liiketoiminnan avaimet, sekä vierasavaimet, mikäli käsitteestä oli viittaus toiseen käsitteeseen.

Looginen malli toimi mallinnuksen master-tasona, jossa lähtöjärjestelmien tiedot yhdistettiin käsitelmallin tietoihin. Kun lisättiin uusia malleja tai tehtiin jokin muutos olemassa olevaan malliin, muutokset tehtiin aina loogiseen malliin. Kun loogiseen malliin oli tehty lähtötietojen määppäys, eli jokainen kenttä sisälsi tiedon siitä, mistä lähteestä se tulee, voitiin konvertoida Data Vaultin staging -aluemalli. Kun staging-aluemalli oli valmis, siitä voitiin konvertoida raw data vault -malli. Sekä staging, että raw data vault -mallit vietiin WhereScape RED -välineeseen export-toiminnolla, jolloin 3D -väline kirjoitti REDin metatietokantaan tarvittavat tiedot staging ja raw data vault -alueen tauluista. Tällä menetelmällä saatiin tietovarasto toteutettua inkrementaalisesti pala kerrallaan ja automatisoitua eri latausvaiheet mahdollisimman pitkälle. Seuraavassa kuvassa on esimerkki tehdystä data vault-rakenteesta.

Data vault: Master - DV kanta - Yleiset



5.3.3 Data Marting

Kun tieto oli saatu haluttuun muotoon DW-kannan stage -alueelle, se vietiin 1:1 DM-kannan load -alueelle. Tämän jälkeen käytettävissä oli kahta faktaa vastaavat load-alueen taulut valmiina ladattavaksi faktatauluihin. Ennen faktatauluihin lataamista oli kuitenkin yhdistettävä dimensioid faktaan tulevaan tietoon, joka tapahtui RED-välineen ominaisuuksia hyödyntäen ja tapahtui DM-kannan stage-alueella. Tällöin esim. SCD2 -muotoiset dimensioid otettiin huomioon faktan latauksien generoinnissa, eli väline kysyi mitä faktan päivämääräkenttää käytetään dimension rivin voimassaolon hakuun. Kun dimensioid oli yhdistetty faktan lähtötietoihin, eli käytännössä haettu surrogaatit dimensiotauluista faktan lähtötiedoista löytyville luonnollisille avaimille, voitiin faktataulu ladata dm-alueella. Data marteja toteutettiin kaksi, joissa faktatauluja fact asiakas ja fact paiv ohjaus (kuvat 37 ja 38). Dimensiotauluja tehtiin yhteensä 7 kpl.

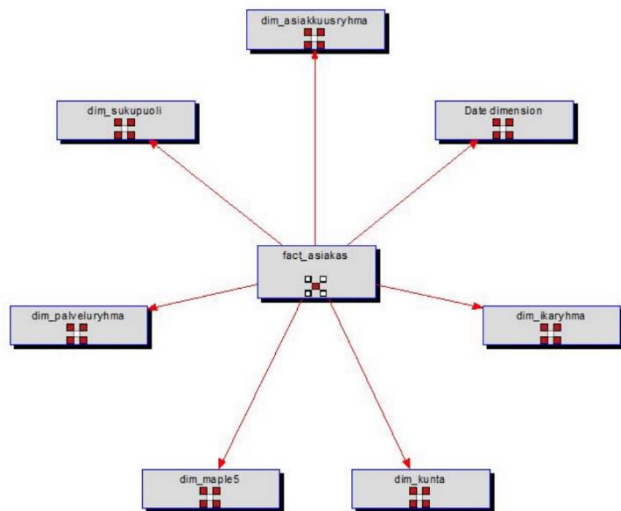
13.12.2019

Table Definition

Display Name	Column Name	Measure	Attribute	Comments
ika	ika	Y		
alkupvm	ajankohta		Y	
palvelukm	palvelukm	Y		
henkilo	henkilo		Y	
kustannus	kustannus	Y		
rug22nursing	rug22nursing	Y		
ai_tulos_12kk	ai_tulos_12kk	Y		
ai_tulos_36kk	ai_tulos_36kk	Y		

Column Documentation

No documentation found

Schema

Dimensions

The Dimensions associated with the table fact_asiakas are :

- *dim_asiakkuusryhma*
- *Date dimension*
- *dim_ikaryhma*
- *dim_kunta*
- *dim_maple5*
- *dim_palveluryhma*
- *dim_sukupuoli*

Kuva 37 Fact_asiakas ja sen käyttämät dimensiot.

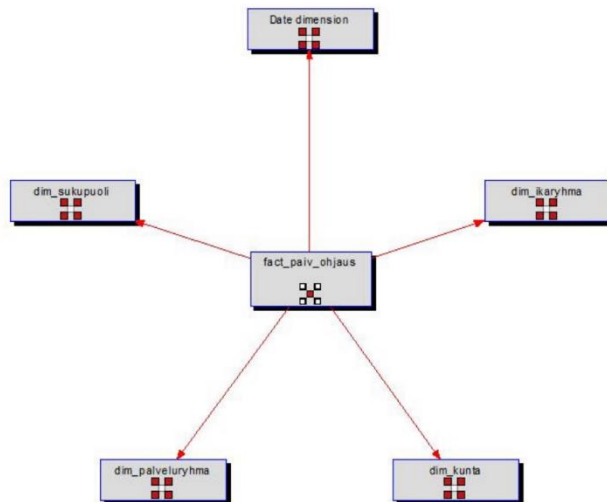
13.12.2019

Table Definition

Display Name	Column Name	Measure	Attribute	Comments
aikavali	aikavali	Y		
osastojaksoalkuhetki	osastojaksoalkuhetki		Y	
osastojaksoloppuhetki	osastojaksoloppuhetki		Y	
kayntiaikaleimahetki	kayntiaikaleimahetki		Y	
yksikkolyhenne	yksikkolyhenne		Y	
yksikkonimi	yksikkonimi		Y	
henkilo	henkilo		Y	
kaynti_hk	kaynti_hk		Y	
esh_pth	esh_pth		Y	

 Column Documentation
 No documentation found

Schema



Dimensions

The Dimensions associated with the table fact_paiv_ohjaus are :

- Date dimension
- dim_ikaryhma
- dim_kunta
- dim_palveluryhma
- dim_sukupuoli

Kuva 38 Fact_paiv_ohjaus ja sen käyttämät dimensiot

5.4 Kokemukset eri välineistä

Yleisesti ottaen valitut välineet sopivat hyvin projektin käyttöön. Seuraavissa kappaleissa on arvioitu tarkemmalla tasolla välineiden käyttöä ja niistä saatuja kokemuksia.

5.4.1 Mallinnus

Ellie

Väline oli yksinkertainen opetella, koska symboleja ja erilaisia tekemisen vaihtoehtoja on rajallinen määrä. Oli hyvä, että mallinnus oli vastuutettu yhdelle henkilölle, koska tällöin mallin kokonaisuus pysyi hallinnassa. Toisaalta mallin tekeminen ja muuttaminen oli ryhmätöitä. Mallista saadaan esille keskeiset käsitteet, joiden varaan jatkotyö pitää ensisijaisesti ankkuroida. Malleista voidaan purkaa ulos vielä mallikohtaisia havaintoja: asiakas, toteutunut palvelu, päätöstiedot, toimintakykytiedot, aluejako, väestöennuste, tiedot kirjanpidosta, laskennallinen työaika, suoritelaskenta. Mallinnuksen aikana havaittiin, että malli voi tulla liian monimutkaiseksi ja vaikeaksi hallita. Haaste ratkaistiin jakamalla mallinnus kolmeksi erilliseksi tietomallitauluksi. Projektin alussa rakennettiin oma sovellusohje mallinnusta varten, joka sisälsi keskeiset

13.12.2019

työtavat, ja nimeämisstandardit. Ellien mallinnus noudattaa pääasiassa ER-mallinnusta ja sen symboleja, mutta erikoisuus on käsitteiden tyypitys (Hovi Data Framework). Tämä tukee tietovarastojen tietomallinnusta, mutta tässä olisi voinut valita yleisesti käytössä olevan merkintätavan eikä luoda välinetoimittajan omaa uutta standardia.

Oleellisin ominaisuus tämän projektin kannalta Elliessä oli se, että käsittemallien avulla saadaan automaattisesti generoitua tietovaraston rakenteet ja tietosisällöt. Ellieen kuvataan käsitteiden tyypitys, jonka pohjalta saadaan tehtyä looginen tietomalli automaattisen ELT-välineen puolella ja edelleen linkattu lähdejärjestelmät tietosisältöön ja tehtyä lataukset. Ellie- ja Wherescape-integrointi varmisti tehokkaan ja nopean toteutuksen projektissa.

Välineellä tehdyt käsittemallit ovat automaattisesti hyödynnettävissä tietovaraston rakentamisessa mallin export-toiminnan kautta. Samoin sanastot ovat vietävissä ulos ja siirrettävissä toiseen välineeseen tarvittaessa.

Muita huomiota mallinnuksesta:

- Uudet projektiin osallistujat pääsevät hyvin mukaan kokonaisuuteen käsittemallin kautta.
- Eri käsittemallien liittyminen toisiinsa saattaa joissakin tilanteissa jäädä epäselväksi – olisi hyvä kuvata mallien väliset suhteet (mallin kuvaukseen kannattaa kirjoittaa kuvaus mistä on kyse).
- Tekemisen kannalta helpompi hahmottaa kokonaisuutta ja helpottaa keskusteluja eri osapuolten välillä.

Havaittuja puutteita:

- Väline oli vielä vahvasti kehitteillä: pitkin matkaa toimitettiin ehdotuksia välineen ominaisuuksien parantamiseksi
- Versiohistoriaa ei ole
- Attribuutteja ei saa näkyville itse mallissa
- Ei mallin käyttöoikeuksien hallintaa

Yhteentoimivuusalusta (YTI)

Yhteentoimivuusalusta on valtiovarainministeriön (VM) Yhteinen tiedon hallinta -hankkeessa (YTI) kehitetty ja VRK:n ylläpitämä alusta, joka on käytettävissä osoitteessa <https://yhteentoimiva.suomi.fi/fi/>. Yhteentoimivuusmenetelmä on yhteinen tapa tuottaa, hallita ja ylläpitää digitaalisten palveluiden ja tietovirtojen taustalla tarvittavia tietomäärittelyksiä ja metatietoja. Menetelmän avulla luodaan ja ylläpidetään tiedon semanttista yhteentoimivuutta eli tiedon käsittelyä siten, että sen merkitys säilyy tietoa vaihdettaessa. Menetelmää tukeva Yhteentoimivuusalusta mahdollistaa yhtenäisen tietosisältöjen määrittelyn sekä toimijoiden tehokkaan ja läpinäkyvän yhteistoiminnan tiedonhallinnassa. Yhteentoimivuusalusta koostuu sanastoista, koodistoista ja tietovirroista ja muussa tiedonhallinnassa tarvittavista tietomalleista. Kullekin työkalulle on tehty omat prosessikuvaukset yhteentoimivuusmenetelmän mukaiselle käytölle.

Yhteentoimivuusalusta ei varsinaisesti ole mallin piirtotyökalu, vaan sanastojen, koodistojen ja käsitteiden (luokkien) ja niiden suhteiden ylläpitotyökalu, joka pystyy visualisoimaan käsittemallin. Yhteentoimivuusalusta on erinomainen käsitteiden uudelleenkäytävyydessä ja periytymisessä. Se on myös hyvä väline tietojen jakamiseen eri organisaatioiden kesken, jokainen näkee mitä muualla on tehty. Prosessi mallien hallinnassa ei kuitenkaan ole täysin selvä, esimerkiksi hyväksymiskäytäntöjä olisi syytä tarkentaa.

13.12.2019

Väline on suhteellisen tuore ja sellaisenaan kirjoitushetken versiota on jonkin verran liian hankala käyttää, jotta sillä voisi tehdä mallintamista workshopissa. Välineen kanssa työskentely on paljon toistoja vaativaa, koska käytettävyyttä ei ole vielä hiottu mallintajan kannalta parhaaksi mahdolliseksi.

YTI:n sisäinen tietomalli on erittäin monipuolinen ja mahdollistaa lähes minkä tahansa tietomallin kuvaamisen. Toisaalta tämä on tuonut tiettyihin paikkoihin kompleksisuutta, mainittakoon näistä esimerkkeinä relaatioiden nimeäminen ja niiden välisten suhteiden hallinta. Kehitysvaiheestaan johtuen väline ei ole paras mahdollinen sovelluskohtaiseen mallintamiseen, vaan mallit on helpompi tehdä ensin muulla välineellä.

Muita havaittuja puutteita olivat:

- Käsitteitä ei tyypitetä vaan kaikki ovat samanlaisia keskenään (ei arkkityyppejä tai stereotyypppejä)
- Sisältää versiot, mutta versionhallinta ei tule esiin selkeästi
- Ei mallin käyttöoikeuksien hallintaa

QPR

QPR on parhaimmillaan kokonaisarkkitehtuurin ja prosessien mallintamisessa, mutta soveltuu myös tieto- ja käsittemallinnukseen. Väline on monipuolinen ja räätälöitävissä moneen eri tarkoitukseen. Se sisältää myös versionhallinnan ja käyttöoikeuksien hallinnan. Se on kuitenkin käyttöliittymältään hieman hankala ja hidas, ja sellaisenaan soveltuu heikommin esimerkiksi työskentelyyn työpajassa. Valtionhallinnon pohjassa ei ollut valmiita malleja tietovaraston tietomallinnukseen, nämä toiminnallisuudet vaativat räätälöintiä. Käsittemallit eivät myöskään olleet automaattisesti muunnettavissa/vietävissä tietovaraston rakentamiseen tarkoitettujen välineiden käyttöön.

5.4.2 Tietovarastoautomaatio

Wherescape

Tietovaraston automatisaatioväline Wherescape nopeuttaa erittäin paljon ELT-vaiheen toteutusta:

- Manuaalisen tietokantasuunnittelun tarve poistuu, koska väline on integroitu Ellie-käsittemallinnukseen ja tietokanta muodostetaan automatisoidun prosessin avulla
- Tietokannan hallinnointi automatisoituu
- Erillisen testauksen tarve vähenee, esim. kohdekenttä vs. lähdekenttä -tietotyyppi ja pituusvertailuja ei tarvitse erikseen koodata manuaalisesti

Wherescape yhdenmukaistaa työtavat eri kehittäjien välillä, jolloin jokainen tietää, miten työt ja työvaiheet tehdään sekä virheet vähenevät. Myös Data Vault -satelliittien voimassaolevan rivin näkymän muodostus voidaan automatisoida.

5.4.3 Tietokanta

SQL Server

Microsoft SQL Serveriä käytettiin projektin aikana pääasiallisena tietokantana, jossa kehitystyö suurimmaksi osaksi tehtiin. On-premise-mallisessa toteutuksessa monimutkaisemmat kyselyt saattavat osoittautua hitaiksi, jolloin kehittäminen hidastuu. Kyselyitä pitää usein optimoida ja jopa rakentaa erikseen indeksejä, jotta ne toimivat tehokkaasti. Tällaisessa ympäristössä tarvitaan erillinen admin tekemään muutoksia ympäristöön eli tarvittaessa lisäämään tehoa ja levytilaa manuaalisesti.

13.12.2019

Snowflake

Snowflake on pilvessä toimiva tietokanta, joka oli projektin aikana käytössä anonymisoidussa data martissa ja sen kehitystyössä. Snowflaken pilviympäristö skaalautuu lähes rajattomasti sekä alas- että ylöspäin. Sen avulla voidaan helposti jakaa tietoja eri tarvitsijoille ilman erillisten liittymien rakentamista. Yksi ympäristö riittää eli ei tarvita erikseen kehitys- ja tuotantoympäristöjä. Tämä johtuu siitä, että tuotantoympäristö on helppo kopioida (kloonata) muutamassa sekunnissa kehitystyötä varten. Tallennuskapasiteetti on käytännössä rajaton eikä kenenkään tarvitse sitä erikseen hallita. Myös tehonmuutokset voidaan asettaa kehittäjien käytettäväksi, jolloin tätäkään varten ei tarvita erillistä adminia. Snowflake toimii suoraan tietolähteenä raportointiin ja visualisointiin ilman muutoksia.

Snowflaken avulla dataa pystyy jakamaan eri toimijoiden kanssa joustavasti ja tehokkaasti. Tässä projektissa ominaisuutta ei hyödynnetty, mutta laajemmin eri sidosryhmien välisissä tietojen toimituksissa ominaisuus olisi hyödyllinen. Tällöin päästään tietojen erillisestä toimituksesta siihen, että tiedot ovat koko ajan käytettävissä yhdessä sovituilta osilta. Näin ollen esimerkiksi THL pystyisi hakemaan tiedot kuntayhtymän tietovarastokokonaisuudesta tarvittaessa. Kuntayhtymä voisi samalla tavalla kuin nykyään kuitenkin varmistua jaettavissa olevasta tietosisällöstä. Vastaavasti kuntayhtymä voisi toimia myös muiden tarvittavien laajojen tiedonvaihtojen kanssa.

5.4.4 Raportointi ja visualisointi

Qlik Sense

Qlik Sense on raportointi- ja visualisointiväline, jolla visualisaatioista saadaan monipuolisia. Jos malli on yksinkertainen, välineellä on nopea ja helppo toteuttaa visualisaatioita.

Qlik Sense -ympäristö toteutettiin on-premise eli se asennettiin analytiikkapalvelimelle Kainuun IT:n tätä projektia varten pystyttämään ympäristöön. Palvelin mitoitettiin ensin liian pieneksi ja projektin alussa oli ongelmia ja hitautta johtuen ympäristön riittämättömyydestä. Tämä parani myöhemmin kun ympäristön resursseja parannettiin.

Muita havaittuja ongelmia olivat:

- Perusominaisuuksilla ei saada toteutettua kaikkia tarvittavia piirteitä, vaan tarvitaan perusversion lisäksi extensioita.
- Kustomointia ei pysty joustavasti tekemään (esim. valintalistat) – ominaisuuksia puuttuu, vaatii ylimääräistä työtä.
- Datan lataus ei toimi automaattisesti.
- Kuntayhteydet haastavia: faktataulujen yhdistämisestä voi aiheutua lisätyötä.

PowerBI

PowerBI on Microsoftin raportointi ja visualisointiväline. PowerBI:tä testattiin lähinnä yhteydellä Snowflakeen ja sillä tehtiin yksi testiraportti. Muutoin sitä ei testattu laajasti projektissa.

Yhteys Power BI Desktopin ja Snowflaken välille onnistui helposti, kun Snowflaken ODBC-ajuri oli ensin asennettu kehittäjän koneelle. Sen sijaan, jos raportin julkaisi pilveen, niin pilvessä raportin päivittäminen ei vielä onnistu ilman gatewayn asentamista. Tämä ei johdu PowerBI:stä vaan Snowflakesta. Tulevissa projekteissa olisi hyvä evaluoida myös PowerBI tarkemmin.

13.12.2019

5.4.5 Kustannuslaskenta

Prodacapo Costing ja Prodacapo Region

Prodacapon Costing-sovelluksen käyttöliittymä toimi suoritelaskennan laskentamallin parametrien määrittely- ja säätötyökaluna sekä apuna mallin määrä- ja hintavalidoinnissa. Alkuvaiheessa mallin parametreja määritettiin ja tätä määrittelyä kommunikoiitiin toimittajan ja asiakkaan välillä toimittajan standardoimalla parametrien lomakepohjalla (MS Excel).

Sovellus mahdollisti usean kehittäjän samanaikaisen sovelluskäytön, sujuvan mallin päivittymisen ja versioinnin - ja siten paransi työn tuottavuutta hajautetussa kehitystiimissä. Sovelluksen logiikka ohjasi myös itsessään työnkulkua ja mallin evoluutiota – sovellukseen on rakennettu oikeellisuustarkistuksia mallin latausten ja päivitysten yhteydessä ja se ohjaa virheenkorjauksessa.

Prodacapon Region -sovelluksen käyttöliittymä toimi episodianalytiikan episodisegmenttien/episodiprofiilien määrittely- ja säätötyökaluna. Toimittaja koodasi määritellyt episodiprofiilit asiakkaan puolesta käyttäen tätä työkalua. Käyttökokemus jäi siis tämän työkalun osalta asiakkaalle (SoteDigi Oy ja Kainuun sote) syntymättä.

Mallin prosessoitavaksi syötettävän asiakasorganisaation toiminnassa syntyneen aineiston muotovaatimukset oli eritelty toimittajan input-tietosisältökuvauksella (MS Excel). Suoritelaskennalla ja episodianalytiikalla oli kummallakin oma kuvauksensa. Kuvauksia jäsennettiin lisää Kainuun soten tapaukseen liittyen projektin edetessä useita kertoja mikä teetti merkittävästi lisätyötä projektissa. Nämä kustannuslaskentatoimittajan rajapinnat tulisi kuvata jatkossa tarkemmin ja standardoida myös asiakkaan palvelutoiminnan substanssin näkökulmasta. Olisi myös hyvä visualisoida aikaisessa vaiheessa, miten rajapintaan tuotettavat tiedot vaikuttavat kustannuslaskennan lopputulokseen. Nyt aikataulu ei sallinut mallin ja mallin ajotulosten väliversioiden esittelyä visualisointeihin asti.

Väline itsessään ei ratkaise tietojohdamisen ongelmatiikkaa vaan toimiva tietojohdamisen ratkaisu edellyttää koko organisaation kattavan tietovaraston olemassaoloa. Prodacapon kustannuslaskenta ja episodianalytiikka sisältää rajatun tietokokonaisuuden ja sitä on mahdoton todellisuudessa laajentaa koko organisaation tiedot sisältäväksi laajaksi, nykyaikaiseksi tietovarastoksi.

Tässä pilottiprojektissa toteutettu rakenne, jossa tiedot kustannuslaskentamoduulia varten tuotetaan tietovarastosta, on toimiva. Tämä luo joustavuutta myös kustannuslaskentamoduulin hallintaan. Tällöin potilastietojärjestelmän ja sen rajapintojen muutoksia ei tarvitse tehdä kustannuslaskennan moduuliin, vaan riittää pelkästään tietotalusta tehdyn rajapintamuutokset. Tietovarasto voi tuottaa edelleen samassa muodossa tietoja kustannuslaskentaa varten.

5.4.6 Edistynyt analytiikka, koneoppiminen ja tekoäly

Python, Scikit-learn, XGBoost, Pandas, R

Valituilla välineillä saatiin hyvin kaikki tavoitteena olevat algoritmit testattua. Koneoppimisen canvaksessa todettiin käytettävät algoritmit ja sovittiin niiden arvioinnissa käytettävät indikaattorit F1 ja Acc. Kaikki käytettävät välineet olivat open source -välineitä ja siten laajasti käytettävissä ja hyödynnettävissä eri toimijoiden toimesta jatkotyössä. Python-kieli on tällä hetkellä eniten käytetty ohjelmointikieli kehittyneen analytiikan alueella ja projektista saatujen kokemusten perusteella soveltuu erinomaisesti tähän alueeseen.

13.12.2019

Koneoppiminen on olemassa olevien kirjastojen hyödyntämistä ja tämän vuoksi kirjastot, joista löytyy laajasti eri algoritmeja, ovat oikeastaan ainoa vaihtoehto toteuttaa tehokkaasti malleja. Projektissa käytettiin seitsemää eri algoritmia, joista XGBoost-pohjainen valittiin jatkotyöhön ja esitetyt tulokset pohjautuvat siihen. R:ää käytettiin asiakassegmentointiin.

Teknisen toteutuksen kannalta kannattaa on-premise-ympäristön sijasta käyttää pilviympäristöjä. Pilviympäristöt skaalautuvat paremmin ja pystytään hetkellisesti opetusajojen aikana käyttämään enemmän resursseja ajojen tekemiseen. On-premise-ympäristöissä ajot saattavat kestää päiviä, koska tilapäistä resurssien lisäys on haastavaa tehdä. Projektissa datan käsittelyyn haastetta toi se, ettei kanta osannut käsitellä koko opetusaineistoa yhtenä isona tauluna, vaan piti jakaa useammaksi kapeammaksi tauluksi ja yhdistää vasta python-ohjelmassa.

Jatkon kannalta on keskeistä miettiä miten ja missä ympäristössä sekä millaisia vaatimuksia analytiikkamallien versionhallinta, säännöllinen ajo, tulosten vertailu, mallin kehitys ja muut tuotannollistamiseen liittyvät aiheet työkaluille tuovat. Esimerkiksi jos halutaan seurata ennusteiden lisäksi segmenttien asiakasmääriä, tulee luoda prosessi, joka asettaa asiakkaat alussa määriteltäviin segmentteihin. Lisäksi on määriteltävä, kuinka usein segmenttien määrittäminen ajetaan sekä missä ja miten tulokset esitetään, sekä kenellä on oikeus tietojen tarkasteluun.

5.5 Kokemus Kainuun ympäristöstä

Kainuun oma on-premise-ympäristö toimi hyvin koko projektin ajan. Haasteita todettiin olevan koneoppimisen mallin ajamisessa, mutta ympäristöä ei toisaalta alun perin suunniteltukaan sitä varten. Resurssien lisääminen analytiikkapalvelimelle paransi tilannetta merkittävästi.

Yhtäaikaisten etäyhteyksien (RDP) määrä rajoitti tekemistä jonkin verran. Jatkossa kannattaa kiinnittää huomiota siihen, että on käytettävissä riittävästi RDP-yhteyksiä palvelinta kohti. Tämä estää tekemistä, jos toteuttajien määrä ylittää käytettävissä olevien yhteyksien määrän.

ICT-alustaryhmä kokoontui viikoittain käsittelemään avoimia asioita ja jos jotain kehitettävää oli, niin muutokset tehtiin viipymättä. Projektin alussa määriteltiin ympäristön minimitaso, jota matkan varrella säädettiin tarpeen mukaan virtualisoidussa ympäristössä. Käytännössä muutokset kohdistuivat analytiikka- ja testipalvelimien muistin ja prosessoreiden määrän kasvattamiseen.

Projektin aikana Kainuun soten IT:n oma valvonta huolehti järjestelmän valvonnasta ja tarvittaessa nosti asioita esiin esimerkiksi palvelinten suorituskykyyn liittyen. Samoin Kainuun soten IT vastasi aina nopeasti, mikäli projektissa havaittiin ongelmia ympäristön suorituskyvyssä. Kainuu sote tarjosi ympäristön käyttöön, työvälineiden asennus tehtiin projektin toimesta.

5.6 Suorituskykytietoja

Projektin aikana testattiin SQL Serverin suorituskykyä, Snowflaken suorituskykyä ja analytiikkapalvelimen suorituskykyä koneoppimisen mallin opettamisessa. Testit eivät ole kattavat eikä niissä ole edes yritetty tehdä mittavaa suorituskykyvertailua kahden eri teknologian välillä. Pääasiallinen tarkoitus testaamiselle oli saada käsitys siitä, miten yksinkertainen käyttötapaus skaalautuu Snowflaken pilviympäristössä ja toisaalta saada jonkinlainen ymmärrys, miten vastaava käyttötapaus toimii on-premise-ympäristössä. Jos testauksesta

13.12.2019

olisi haluttu tehdä kattava, olisi pitänyt tehdä se monella erilaisella käyttötapauksella ja usealla yhtäaikaaisella kyselyllä. Tällainen testaus ei kuulunut projektin laajuuteen.

5.6.1 SQL Server ja Snowflake

SQL Serverin ja Snowflaken suorituskyvyn vertailuun valittiin data martin muodostaminen yhdellä faktalla ja kolmella dimensiolla. Faktataulussa oli 133 miljoonaa riviä.

SQL serverin kokoonpano oli seuraava: 2 VCPU, muistia 16 GB. Snowflake-testit ajettiin virtuaalisilla tietovarastoilla XS, S, M, L, XL, 2XL, 3XL ja 4XL. Snowflaken automaattista klusterointia ei testattu sopivan testitapauksen puuttuessa. Snowflaken sisäinen cache poistettiin käytöstä testin ajaksi, jotta saataisiin vertailukelpoisia tuloksia. Testit toistettiin kahdesti, jotta varmistuttiin, että saadut ajat olivat samat molemmilla kerroilla.

Tulokset olivat seuraavat:

Kokoonpano	Käytetty aika (sekuntia)
Snowflake X-Small: 1 credit/hour	18
Snowflake Small: 2 credit/hour	12
Snowflake Medium: 4 credit/hour	9
Snowflake Large: 8 credit/hour	6
Snowflake X-Large: 16 credit/hour	6
Snowflake 2X-Large: 32 credit/hour	6
Snowflake 3X-Large: 64 credit /hour	7
Snowflake 4X-Large: 128 credit/hour	8
SQL Server: 2 VCPU 16 GB	280 - 440
SQL Server: 4 VCPU 32 GB	270 - 390

Havainnoista nähdään, että Snowflaken käyttämä aika tämän testiajon suorittamiseen vähenee, kun laskentatehoa kasvatetaan. Ajan pieneneminen ei kuitenkaan ole lineaarista eikä suoraan suhteessa laskentatehon kasvattamiseen. Lisäksi aika ei lopulta kaikkein tehokkaimmissa vaihtoehtoisissa enää vähene lainkaan vaan hieman jopa kasvaa, kun mennään kaikkein tehokkaimpaan mahdollisuuteen. Tämän voisi ajatella johtuvan testiaineiston koosta ja yksinkertaisesta rakenteesta. Yksi yhden käyttäjän suorittama 133 miljoonan rivin kysely on yksinkertaisesti liian pieni aineisto mittaamaan suorituskyyä laajasti. Isoon laskentatehoon liittyy prosessoinnin jakaminen usealle eri nodelle ja näiden hallintaan ilmeisesti menee osa tehoista. Toisaalta myös itse levyille kirjoittaminen on työvaihe, johon ei prosessoritehon lisääminen vaikuta.

SQL Serverin osalta ei saatu vertailukelpoisia tuloksia kahden eri kokoonpanon välille. Suoritusajat vaihtelivat liikaa johtuen mahdollisesti palvelimen muusta samanaikaisesta käytöstä. SQL Serverin kokoonpanoa ei tässä pilotissa millään tavoin optimoitu, emme voi olla varmoja, miten hyvin se sai resursseja virtuaalikoneympäristössä, jota ei välttämättä oltu suunniteltu tietokantapalvelimien käyttöön. SQL serverin toimintaa testattiin myös lisäämällä indeksit tauluihin, mutta tällä ei ollut merkittävää vaikutusta suoritusaikoihin. Käytännössä kyselyn suuri rivimäärä tarkoittaa sitä, että suurin osa ajasta kuluu levyille kirjoittamiseen, johon prosessoritehon lisääminen ei vaikuta.

Testi ei siis ollut millään tavalla kattava eikä nyt saatuja tuloksia kannata tulkita muuten kuin tämän yhden yksinkertaisen käyttötapauksen kautta. Tästä huolimatta testitapaus on hyvin tyypillinen tietovarastokäytössä, joten siinä mielessä testi on kuitenkin aiheellinen ja todellista tilannetta kuvaava. On myös hyvin tyypillistä, että SQL-palvelin asennetaan virtuaaliympäristöön. On-premise-ympäristön optimointia pitäisi kuitenkin tehdä enemmän, tällä olisi varmasti vaikutusta testin lopputulokseen. Toki

13.12.2019

esimerkiksi modernien pilvessä toimivien tietokantojen osalta pitää harkita, onko tällainen käytätapa parasta ratkaista tällä tavoin. Pilviympäristössä ei välttämättä ole syytä suorituskvyn kannalta kopioida dataa uuteen tauluun, pelkkä näkymä saattaa riittää hoitamaan saman asian.

5.6.2 Analytiikkapalvelin

Analytiikkapalvelimen testit ajettiin koneoppimisen aineistolla kokoonpanoilla:

- Kehittäjän oma työasema 3,9GHz (16 Threads)
- Virtuaalipalvelin 2,2GHz (32 Threads)
- Virtuaalipalvelin 2,2GHz (2 Threads)

Tulokset olivat seuraavat:

	Pure Windows on Ryzen 2700X @3.9GHz 16 Threads	VM Windows on Xeon Gold 5120 @2,20GHz 32 Threads	VM Windows on Xeon Gold 5120 @2,20GHz 2 Threads
Kokonaisaika	2 h 20 m	2 h 48 m	1-2 vrk
Datan esikäsittely	1 m 27 s	2 min 53 s	N/A
Mallin piirteiden haku (opetus)	2 h 18 m	2 h 44 m	N/A
Lopullisen mallin opetus	1 s	1 s	N/A
Graafien piirtäminen	9 s	10 s	N/A
Kertaa pöytäkoneen aika	1	1,2	8-16

Voidaan siis todeta, että eniten koneoppimisen mallin opettamisen nopeuteen vaikuttaa laskentateho, erityisesti prosessien määrällä on merkitystä.

13.12.2019

6. Tietoturva ja tietosuoja

Projekti toteutettiin Kainuun Soten tietoturvapoliittikan mukaisesti, jossa määriteltiin tekniset ja organisatoriset suojaamistoimet. Projektissa käytetyt työkalut, tuotteet, sovellukset tai palvelut takasivat oletusarvoisesti, että käsittely rajoittuu vain käsittelyn tarkoituksen kannalta tarpeellisiin henkilötietoihin. Käsittelyn rajaamisessa huomioitiin tiedon määrä, käsittelyn laajuus, tietojen säilytysaika ja tietojen käyttöön oikeutettujen henkilöiden lukumäärä. Lisäksi projektissa kiinnitettiin huomiota tietojen minimointiin eli vain käsittelyn kannalta tarpeelliset tiedot kerättiin. Lisäksi tiedon minimointi huomioitiin tiedon anonymisoinnissa ja pseudonymisoinnissa, näistä lisää kappaleessa 6.3. Koska projektissa kerättiin ja käsiteltiin henkilötietoja, solmittiin kaikkien tämän tiedon käsittelijöiden kanssa salassapito- ja tietoturvasitoumus. Kaikki tiedot pilotin jälkeen jäivät rekisterinpitäjälle ja toisaalta rekisterinpitäjyys tietojen osalta pysyi Kainuun sotessa.

Henkilötietojen käsittelyn käyttötarkoituksena oli pilotoida keväällä 2019 voimaan tullutta lakia sosiaali- ja terveystietojen toissijaisesta käytöstä (tietojohdamisen pykälä) ja tehdä pilotin aikana toteutus, joka vaikuttaa tietojohdamisen kautta myös kohdeorganisaation toimintaan ja kehittää sitä eteenpäin.

Projektissa perustettiin uusi rekisteri minkä osalta tehtiin projektin alkuvaiheessa tietosuojaseloste. Tämän tietosuojaselosteen kautta täyttyi velvollisuus informoida rekisteröidyille kyseinen tietojenkäsittelytapa. Uuden rekisterin kautta syntyi myös tarve arvioida tästä rekisteröidyille aiheutuvat riskit. Tämä työ on käynnissä osana Kainuun soten normaalia riskienarviointiprosessia (tietosuoja koskeva riskienarviointi), millä arvioidaan suunnitellun käsittelytoiminnan vaikutuksia henkilötietojen suojaan.

6.1 Sopimukset

Projektin sopimuskumppaneina olivat Kainuun sosiaali- ja terveydenhuollon kuntayhtymä (Kainuun sote, Asiakas) ja SoteDigi Oy (Toimittaja). Sopimuksessa määriteltiin projektin aikataulu, osapuolten yleiset vastuut, projektissa tuotettavat lopputulokset, projektin kustannuksiin liittyvät vastuut, tuotosten omistus- ja immateriaalioikeudet sekä luottamuksellisuuteen ja salassapitoon liittyvät kysymykset. Sopimuksen liitteenä olivat projektisuunnitelma sekä salassapito- ja tietoturvasitoumukset, jotka Toimittajan ja alihankkijoiden puolelta allekirjoittivat kaikki luottamuksellisten henkilötietojen kanssa tekemisissä olleet henkilöt. Tämän tueksi Kainuun sote toimitti erillisen lupapäätöksen ”Lupa henkilötietoja sisältävän aineiston poimintaan ja käsittelyyn” käsittäen Toimittajan sekä projektissa mukana olleet alihankkijat.

SoteDigi Oy:n alihankkijoina projektissa toimivat Citrus Solutions Oy ja Prodacapo Finland Oy. Molempien osalta tehtiin erilliset sopimukset, lisäksi Citrus Solutions Oy:llä oli projektissa omina alihankkijoinaan Productivity Leap Oy sekä Ari Hovi Oy.

6.2 Laki sosiaali- ja terveystietojen toissijaisesta käytöstä eli toisiolaki

Laki sosiaali- ja terveystietojen toissijaisesta käytöstä 552/2019 [1] (myöhemmin toisiolaki) tuli voimaan presidentin esittelystä 26.4.2019. SoteDigi Oy:n tietojohdamisen pilotti on ensimmäinen toteutus, jossa lakia ja sen mahdollistamaa tietojen yhdistämistä ja tuottamista tietojohdamisen käyttöön käytetään laajasti hyödyksi ja tämän vuoksi asian käsittelyyn on myös projektissa käytetty paljon aikaa.

13.12.2019

SoteDigi Oy:n ja Kainuun sote -kuntayhtymän pilotissa toisiolaista hyödynnettiin laissa olevaa tietojohdamisen pykälää 41 (kuvasssa 39). Tietojohdamisen pykälän mukainen toiminta ei vaadi erillistä toisiokäytön lain määrittelemää tietolupaa tietojen yhdistämistä varten. Pykälän mukainen perustelu pohjautuu palvelunantajan oman tiedon hyödyntämiseen sen oman toiminnan johdamisen tarkoituksiin.

Tietojen käsittely lain nojalla ilman tietolupaa**41 §****Tietojohdaminen**

Sosiaali- tai terveydenhuollon palvelunantajalla on oikeus salassapitovelvoitteiden estämättä sekä tietosuoja-asetuksen 9 artiklan 2 kohdan h alakohdan nojalla käsitellä ja yhdistellä tunnisteellisesti asiakastietoja, jotka ovat syntyneet sen omassa toiminnassa tai ovat sen omiin rekistereihin tallennettuja, jos se on välttämätöntä palvelunantajan vastuulla toteutettavan palvelutoiminnan tuottamista, seuranta, arviointia, suunnittelua, kehittämistä, johtamista ja valvontaa varten.

Jos palvelunantajalle on tarpeen vastuullaan toteutettavan palvelutoiminnan taikka palveluketjujen arviointia, suunnittelua tai kehittämistä varten verrata omaa toimintaansa muiden palvelunantajien toimintaan, Tietolupaviranomainen voi tuottaa tarvittavan vertailuaineiston aggregoituna tilastotietona 45 §:n mukaisesti 3 §:n 9 kohdassa tarkoitetun tietopyynnön perusteella.

Sen lisäksi, mitä 1 ja 2 momentissa säädetään, kunnalla tai kuntayhtymällä on tietojohdamisen tarkoituksessa oikeus käsitellä ja yhdistellä tunnisteellisesti myös asiakastietoja, jotka on tallennettu terveydenhuoltolain (1326/2010) 9 §:n 1 momentissa tarkoitettuun yhteisrekisteriin.

Kuva 39 Tietojohdaminen, 41§ [1]

Toisiokäytön laki helpottaa käsittelyä toimijoiden kannalta, mutta sen mahdollistamia hyötyjä ei saada tehokkaasti käyttöön, jos toimijat joutuvat erikseen pohtimaan jokaisen yksityiskohdan, tarvittavat lomakkeet, sopimukset ja käytännöt. Tämän helpottamiseksi tässä projektissa yhtenä tavoitteena oli testata toisiolain mahdollisuuksia käytännössä.

Toisiokäytön laki parantaa merkittävästi sote-tietojen hyödyntämisen mahdollisuuksia. Ilmiselvät hyödyt tulevat tutkimustyön tehostamisessa lain mahdollistaessa yhden luukun periaatteen tietolupien käsittelyssä ja eri rekisterinpitäjän tietojen koostamisessa tutkijoille. Sen lisäksi merkittävänä hyötynä - oikein ja tehokkaasti käytettynä - tulee olemaan lain mahdollistama tietojohdamisen tehostaminen. Tehokas tietojohdaminen mahdollistaa sote-alueen toimijoiden paremman ja laadukkaamman toiminnan, koska tietoa eri alueilta voidaan yhdistää, ymmärtää ja käyttää johtamisessa, resurssien allokoinnissa ja toiminnan ohjaamisessa.

Toisiokäytön lain perusteluissa on kuvattu hyvin lainsäätäjän tavoitteet. Lain tavoitteena on ollut parantaa palvelunantajan tosiasiallisia edellytyksiä hoitaa tehtävänsä ja samalla mahdollistaa asukkaiden paremmat palvelut ja rekisteröidyn perustuslain mukainen oikeus sosiaali- ja terveystietojen saamiseen. Toisiolaki mahdollistaa tietojen tehokkaan hyötykäytön toimintaprosessien kehittämisessä ja potilaan tai asiakkaan hoidon ja palvelujen yhteensovittamisessa. Lisäksi toisiolaki edesauttaa tietojen koostamisessa asiakassegmentointia varten sekä muissa vastaavissa tuottavuutta ja vaikuttavuutta parantavissa käyttötarkoituksissa. Perustelut on koostettu kuvaan 40.

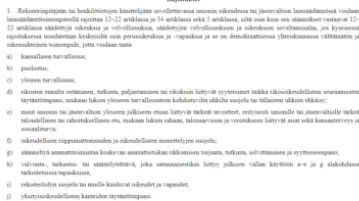
13.12.2019

41§ tavoite ja perustelut

2016/679 (Tietosuoja-asetus)
artikla 6 kohta 1 alakohta e



2016/679 (Tietosuoja-asetus)
artikla 23 kohta 1 alakohta e



Tavoitteena parantaa
palveluntajan tosiasiallisia
edellytyksiä hoitaa tehtävänsä

552/2019
(Toisiolaki) 41§



Oltava välttämätön ja
oikeasuhtainen
toimenpide, perusteltavissa kun
tavoitteena taata yleiseen julkiseen
etuu liittyvät tärkeät tavoitteet,
esimerkiksi kansanterveys ja
sosiaaliturva.

552/2019 41§ Lain perustelut:

Näin toimintaa, tuotantoa ja taloutta kyetään ohjaamaan paremmin. Tuotantoprosessien kehittämiseen, potilaan hoidon kokonaisuuden yhteensovittamiseen, asiakassegmentointiin ja muihin vastaaviin tuottavuutta ja vaikuttavuutta parantaviin käyttötarkoituksiin tarvittavien tunnistettujen tietojen tulee olla hyödynnettävissä johtamistiedon pohjana reaaliaikaisesti, jatkuvaluontoisesti ja kumuloituen, osana organisaatioiden päivittäistä toimintaa ja ilman erillisten tutkimushankkeiden käynnistämistä.

Palveluntajalla tulee olla käytettävissään tietoa, jonka pohjalta se voi vastata terveydenhuoltolaissa (1326/2010), sosiaalihuoltolaissa (1301/2014) sekä alan erityislainsäädännössä säädettyjen taikka sopimukseen pohjautuvien tehtäviensä ja velvollisuuksiensa asianmukaisesti hoitamisesta.

Näin ollen tietojohdamisen mahdollistaminen on välttämätöntä sekä kansanterveyden ja sosiaaliturvan turvaamiseksi että yksittäisten rekisteröityjen perusoikeuden eli perustuslain 19 §:ssä säädetyn sosiaaliturvan turvaamiseksi.

Nämä kaksi tietosuoja-asetuksen 23 artiklan mukaista perustetta kulkevat käsi kädessä: tehokas ja vaikuttava tietojohdaminen johtaa siihen, että asiakkaat saavat parempaa palvelua ja näin turvataan rekisteröidyn oikeutta terveyteen ja sosiaaliseen hyvinvointiin sekä potilasturvallisuuteen.

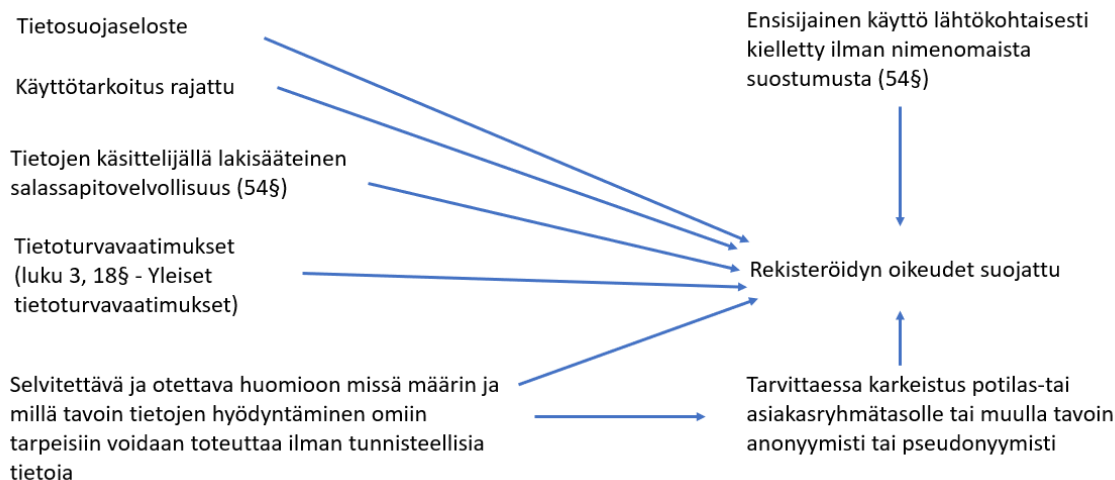
Kuva 40 Tietojohdamisen 41§ tavoite ja perustelut

Palveluntajilla on siis lain säätämät mahdollisuudet hyödyntää tietoja tietojohdamisessa tehokkaasti. Sote-tieto on erittäin sensitiivistä, ja sitä tulee käsitellä äärimmäisen huolellisesti ja mahdollisia riskejä välttämällä. Keskusteluissa eri sidosryhmien kanssa on yhteisesti todettu, että laki on uusi ja sen vuoksi lain tulkintaa ja siihen liittyviä oikeuskäytäntöjä ei juuri ole olemassa. Tilanteessa, jossa lain tulkintaa ei käytännössä ole vielä laajassa mitassa tehty, nousevat lain perusteluissa kuvatut periaatteet ja tavoitteet sekä niiden konkreettinen koeponnistaminen tärkeään asemaan. Näin voidaan käydä läpi sitä mihin tilanteisiin ja millä perusteilla lainsäätäjät on lain laatinut.

Lain tulkintojen lisäämiseksi pilotointi- ja kokeiluprojektit ovat tärkeitä. Ne nostavat keskusteluun konkreettisia, todelliseen tietoon pohjautuvia esimerkkitapauksia. Konkreettisten tapausten avulla voidaan arvioida riskit, mutta havaitaan myös mahdollisuudet toiminnan ennakkoinnin, tehostamisen, tuottavuuden ja vaikuttavuuden kannalta. Keskustelujen ja kokeilujen pohjalta lain soveltaminen ja siihen liittyvät käytännöt vakiintuvat.

Pilottiprojektissa esiinnousseita kysymyksenasetteluja on käyty läpi Kainuun soten, STM:n ja tietosuojavaikuttetun toimiston kanssa ja lisäksi lukuisissa keskusteluissa eri foorumeilla. Toisiolain hyödyntämisen osalta loppuraporttiin on nostettu muutama konkreettinen kysymys. Seuraavassa nämä yksittäiset kysymykset on käyty läpi ja kuvattu niihin liittyvää, pilottiprojektissa eri tahojen kanssa käytyä pohdintaa. Rekisteröidyn oikeuksien turvaamisen varmistamiseen liittyvät tekijät toisiokäytön lain kannalta on esitetty kuvassa 41.

13.12.2019


Kuva 41 Rekisteröidyn oikeuksien suojaaminen

Voidaanko asiakassegmentointia tehdä?

Käytyjen keskustelujen ja selvittelyjen perusteella on selvää, että asiakassegmentointia voidaan tietoturva vaatimusten mukaisessa tietoteknisessä ympäristössä toisilain perusteella tehdä. Toisilain perusteluissa [8] todetaan tietojohdamisen tavoitteena olevan parantaa palvelunantajan tosiasiallisia mahdollisuuksia hoitaa tehtäväänsä. Tehokas ja vaikuttava tietojohdaminen johtaa siihen, että asiakkaat saavat parempaa palvelua ja näin turvataan rekisteröidyn perusoikeus terveyteen ja sosiaaliseen hyvinvointiin ja potilasturvallisuuteen. Jotta toimintaa ja taloutta kyetään ohjaamaan paremmin, on tietojen oltava käytettävissä reaaliaikaisesti, jatkuvaluonteisesti ja kumuloituen osana organisaation päivittäistä toimintaa ja ilman erillisten tutkimushankkeiden käynnistämistä.

Voidaanko asiakassegmentoinnin suorittamista varten tehdä palvelunantajan omaan käyttöön listaa, joka sisältää tunnistelliset asiakastiedot?

Edelleen lain ja sen perustelujen pohjalta näyttää siltä, että palvelunantajan omaan käyttöön voidaan tehdä lista, joka sisältää tunnistelliset asiakastiedot. Asiakassegmenttien todentamisen ja varmistamisen kannalta on välttämätöntä, että tunnistelliset asiakastiedot tietoturva vaatimusten mukaisessa tietoteknisessä ympäristössä ovat käytettävissä palvelunantajalla. Palvelunantaja voi niiden avulla tutkia tietoja, kehittää palveluprosesseja ja yhteensovittaa asiakassegmenttiin kuuluvien henkilöiden hoidon ja palvelun kokonaisuutta. Tietojen käsittelijällä on joka tapauksessa lakisääteinen salassapitovelvollisuus ja tämän ansiosta tietoja käytetään vain käyttötarkoituksen mukaisesti.

Voidaanko tämän listan perusteella tehdä ryhmätason interventioita?

Laista ei selvitysten mukaan suoraan löydy tähän sopivia perusteluja. Jos tunnistetaan ennakolta riskiryhmä, jonka jäsenet hyötyisivät esim. ennaltaehkäisevästä interventiosta, voidaanko näihin henkilöihin ottaa yhteyttä? Millainen interventio olisi ryhmätasolla asiakassegmentin perusteella mahdollinen?

Voidaanko tämän listan perusteella suunnitella ja tehdä interventioita henkilötasolla?

Laista ei myöskään löydy tähän käyttötapaukseen suoraan sopivia perusteluja. Kun rekisterinpitäjä hyödyntää omassa toiminnassaan syntynyttä tietoa toissijaisessa tarkoituksessa, voidaanko lain nojalla edes poikkeustapauksissa käyttää saatua tietoa ensisijaisessa tarkoituksessa yksittäisten henkilöiden suhteen? Pykälässä 55 on kuvattu menettelytavat, kun Tietolupaviranomaisen myöntämän tietoluvan kautta saadun

13.12.2019

aineiston perusteella identifioidaan henkilöllä tai henkilöillä kliinisesti merkittävä löydös, mutta tässä ei ole kuvausta menettelytavoista kun palvelunantaja tekee vastaavan löydöksen toisiokäytön lain käyttötarkoituksessa ilman Tietolupaviranomaisen myöntämää tietolupaa.

Lääkinnällinen laite

Jos toisiokäytön lain tarkoituksessa kerätyn datan hyödyntämiseksi tehdään sovellus, joka antaa hyvinvointiin ja terveydentilaan liittyviä ennusteita tai suosituksia, lähestytään lääkinnällisen laitteen määritelmää. Kysymys kuuluu, missä vaiheessa tällainen sovellus muuttuu lääkinnälliseksi laitteeksi?

Palvelun järjestäjän mahdollisuus hyödyntää terveydenhuollon yhteisrekisterissä olevaa tietoa

Lähtökohtaisesti laki sallii sosiaali- ja terveydenhuollon palvelunantajan yhdistämään tunnisteellisesti asiakastietoja, jotka ovat syntyneet sen omassa toiminnassa tai ovat sen omiin rekistereihin tallennettuja. Toisiokäytön lain pykälässä 41 todetaan että kunnalla tai kuntayhtymällä on tietojohdamisen tarkoituksessa oikeus käsitellä ja yhdistellä tunnisteellisesti myös asiakastietoja, jotka on tallennettu terveydenhuoltolain (1326/2010) 9 §:n 1 momentissa tarkoitettuun yhteisrekisteriin. Voiko siten palvelunjärjestäjä yhdistää tunnisteellisesti erikoissairaanhoidon tietoja omiin perustason sosiaali- ja terveydenhuollon tietoihinsa?

Palvelun järjestäjän mahdollisuus hyödyntää ostopalveluissa syntynyttä tietoa

Ostopalvelutilanteessa terveydenhuollon palvelunantaja, jolta palvelu on ostettu, toimii ostopalvelun tuottajana ja toteuttaa järjestäjän tilaaman palvelun. Ostopalveluna tuotetut potilasasiakirjat kuuluvat kuitenkin palvelun järjestäjän rekisteriin. Tämä ei kuitenkaan tarkoita, että toiminnasta syntyneet tiedot olisivat automaattisesti käytettävissä järjestäjän tietojohdamisen järjestelmissä.

Pseudonymisointi ja anonymisointi

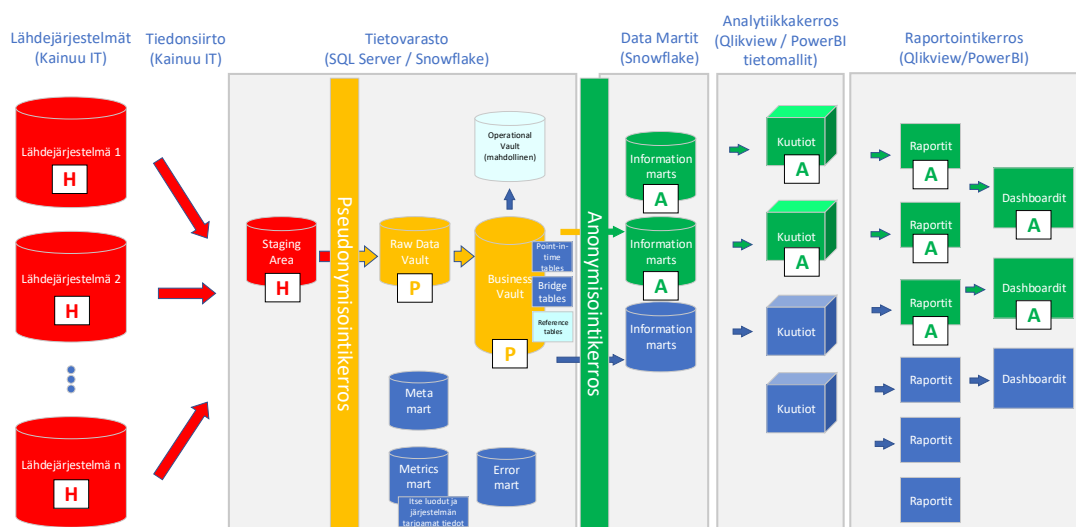
Keskusteluissa on lisäksi tullut esille kysymys siitä millaisia pseudonymisoinnin ja anonymisoinnin käytäntöjä on sovellettava tapauksessa, jossa lain pykälän 41 mukaisesti asiakastietoja voidaan käsitellä tunnisteellisesti tietojohdamisen tukemisessa. Lain perustelujen mukaan on selvitettävä ja otettava huomioon missä määrin ja millä tavoin tietojen hyödyntäminen omiin tarpeisiin voidaan toteuttaa ilman tunnisteellisia tietoja. Tällöin tarvittaessa tietoja karkeistetaan ja käsitellään muulla tavoin anonymisti tai pseudonymisti. Kaikissa tapauksissa tiedon arvosta menetetään jotain. Kysymys kuuluu, estääkö se tiedon hyötykäytön tietojohdamisen tueksi palvelunantajan palvelutehtävän mahdollistamiseksi? Missä tapauksissa pseudonymisointi, anonymisointi tai muu tietojen karkeistus on tarpeen ottaen huomioon, että toimitaan tietoturva vaatimusten mukaisessa ympäristössä niiden henkilöiden toimesta, joilla on lakisääteinen salassapitovelvollisuus?

Pilotissa valittiin varmuuden vuoksi erittäin tarkka ja tietosuojaan huomioiva käsittelytapa, jota on kuvattu luvussa 5.3. Edellä listatut kysymykset on tarkoitettu jatkokeskustelun pohjaksi.

6.3 Teknologia

Tietosuojaan toteutuminen on esitetty kuvassa 42 seuraavasti: punaisella ja H-kirjaimella merkityt kohteet sisältävät yksilöityä henkilötunnusellista tietoa. Keltaisella ja P-kirjaimella merkityt kohteet sisältävät yksilöityä pseudonymisoitua tietoa. Vihreällä ja A-kirjaimella merkityt kohteet sisältävät anonymisoitua tietoa. Sinisellä merkityt kohteet sisältävät sellaista tietoa, joka ei ole henkilötietoa.

13.12.2019



Kuva 42 Tietosuojan toteutuminen.

6.3.1 Pseudonymisointi

Pseudonymisointi toteutettiin yksisuuntaisena, ts. sitä ei voi palauttaa takaisin. Toinen vaihtoehto olisi ollut tallettaa alkuperäinen avain pseudoavaimen kanssa omaan suojattuun tauluun, mutta tässä projektissa siihen ei nähty tarvetta. Pseudonymisointiin käytettiin SQL Serverin SHA2_512-hash-funktiota suolattuna avaimella. Hashauksen lisäksi datasta poistettiin kaikki suorat tunnisteet eli mahdolliset nimet, osoitteet, puhelinnumerot ja sähköpostiosoitteet. Pseudonymisoinnin jälkeen tietovaraston stage- ja dv-alueella on vain pseudonymisoitua tietoa.

6.3.2 Anonymisointi

Anonymisointi tehtiin dv- ja data mart-kerrosten välissä. Anonymisointi tehtiin Tampereen yliopiston Aineistonhallinnan käsikirjan periaatteiden mukaisesti <https://www.fsd.uta.fi/aineistonhallinta/fi/>. Osa data martissa olevista faktoista oli tämän jälkeen summattua tietoa ja henkilötason faktat eivät sisältäneet tunnistamisen mahdollistavia tietoja.

Anonymisointia tarvittiin henkilötasoisessa faktataulussa fact_asiakas. Siihen haluttiin raportointia varten seuraavat tiedot, joita arvioitiin anonymisoinnin näkökulmasta seuraavasti:

Sarake	Kuvaus	Anonymisointitarve
Ikä	Henkilön ikä tapahtumahetkellä	Luokitellaan ikäryhmiin
Ajankohta	Tapahtuman päivämäärä	Säilytetään päivätason seuranta varten
Palvelulkm	Annettujen palveluiden lukumäärä ajankohdassa	Mittaritietoa, summataan valittujen dimensioiden mukaan
Henkilö	Henkilön yksilöllinen tunniste	Estetään yhdistäminen muihin tietoihin muuttamalla avain
Kustannus	Kustannukset ajankohdassa	Mittaritietoa, summataan valittujen dimensioiden mukaan
Rug22nursing	Mittaritietoa, asiakasrakenneluokitus	Mittaritietoa, summataan valittujen dimensioiden mukaan
AI_tulos_12kk	Edistyneen analytiikan tulos vuoden ennusteessa	Henkilön tietoa
AI_tulos_36kk	Edistyneen analytiikan tulos kolmen vuoden ennusteessa	Henkilön tietoa
Asiakkuusryhmä	Asiakkuuden ryhmä	Luokittelu riittävä alustavan arvion mukaan
Palveluryhmä	Annetun palvelun	Luokittelu riittävä alustavan arvion mukaan
Maple5-luokka	Maple5-luokka	Henkilön tietoa, luokittelu riittävä alustavan arvion mukaan
Sukupuoli	Sukupuoli	Henkilön tietoa, mahdollisesti poistettava

13.12.2019

Ikäryhmä	Ikäryhmä	Henkilön ikäryhmä tapahtumahetkellä, luokittelu riittävä alustavan arvion mukaan (vanhuspalveluiden kyseessä ollen vain kaksi luokkaa, 75-84 v ja 85v-)
Kunta	Asiakkaan kunta	Henkilön tietoa, riittävä jaottelu arvion mukaan

6.3.3 Jäännösriskin arviointi

Tämän perusteella muodostettiin faktataulu ja sen sisältämää tietoa arvioitiin erikseen eri dimensioiden ja niiden yhdistelmien kautta. Eri asiakkaita faktataulussa on 33 312. Tällä tavalla ryhmiteltynä saadaan yhteensä 5 950 erilaista ryhmää. Kymmenen suurimman ryhmän sisältö näyttää seuraavalta:

Ikäryhmä	Kunta	Maple5	Palveluryhmä	Sukupuoli	Asiakkuusryhmä	Määrä
75-84v	KAJAANI	Unknown	Perustason vastaanottopalvelut	Nainen	Kaikki muut asiakkaat	56 455
75-84v	KAJAANI	Unknown	Perustason vastaanottopalvelut	Mies	Kaikki muut asiakkaat	41 803
75-84v	KUUMO	Unknown	Perustason vastaanottopalvelut	Nainen	Kaikki muut asiakkaat	24 615
75-84v	KAJAANI	Unknown	ESH poliklininen hoito	Nainen	Kaikki muut asiakkaat	19 739
75-84v	KUUMO	Unknown	Perustason vastaanottopalvelut	Mies	Kaikki muut asiakkaat	19 667
75-84v	SOTKAMO	Unknown	Perustason vastaanottopalvelut	Nainen	Kaikki muut asiakkaat	19 167
75-84v	SUOMUSSALMI	Unknown	Perustason vastaanottopalvelut	Nainen	Kaikki muut asiakkaat	18 856
85v-	KAJAANI	Unknown	Perustason vastaanottopalvelut	Nainen	Kaikki muut asiakkaat	18 133
75-84v	KAJAANI	Unknown	ESH poliklininen hoito	Mies	Kaikki muut asiakkaat	16 841
Tuntematon	KAJAANI	Unknown	ESH poliklininen hoito	Nainen	Kaikki muut asiakkaat	16 145

Havaittiin, että ryhmissä on lopussa 3 004 erilaista ryhmää, joissa on henkilöitä 5 tai vähemmän. Myös yksittäisiä henkilöitä sisältäviä ryhmiä on 1 580. Näistä havainnoista johtuen päätettiin kokeilla ryhmittelyä siten, että poistetaan sukupuoli. Tällöin saatiin aikaiseksi 4 017 erilaista ryhmää, joista 1 969 sisältää alle 5 henkilöä ja 1 029 yksittäisiä henkilöitä. Tähän on otettu esimerkki yksittäisiä henkilöitä sisältävistä ryhmistä:

Ikäryhmä	Kunta	Maple5	Palveluryhmä	Asiakkuusryhmä	Määrä
85v-	JÄMSÄ	Unknown	ESH poliklininen hoito	Tilapäinen kotihoito	1
85v-	JÄMSÄ	Unknown	Päivystyspalvelut	Kaikki muut asiakkaat	1
85v-	JOENSUU	Unknown	Päivystyspalvelut	Kaikki muut asiakkaat	1
85v-	KAJAANI	Kohtalainen avuntarve	ESH somaattinen osastohoito	Kaikki muut asiakkaat	1
85v-	KAJAANI	Kohtalainen avuntarve	Muut kotona asumista tukevat palvelut	Tilapäinen kotihoito	1
85v-	KAJAANI	Kohtalainen avuntarve	Suun terveydenhuolto	Tilapäinen kotihoito	1
85v-	KAJAANI	Kohtalainen avuntarve	Tehostettu palveluasuminen	Tilapäinen kotihoito	1
85v-	KAJAANI	Kohtalainen avuntarve	Tehostettu palveluasuminen	Tukipalveluasiakkaat	1
85v-	KAJAANI	Kohtalainen avuntarve	Vammaispalvelut	Säännöllisen kotihoidon asiakas	1
85v-	KAJAANI	Lievä palvelutarve	Muut kotona asumista tukevat palvelut	Kaikki muut asiakkaat	1
85v-	KAJAANI	Lievä palvelutarve	Kuntoutuspalvelut	Tilapäinen kotihoito	1
85v-	KAJAANI	Lievä palvelutarve	Perustason sairaalapalvelut	Tilapäinen kotihoito	1
85v-	KAJAANI	Kohtalainen avuntarve	Ympäri vuorokautinen laitshoito	Tilapäinen kotihoito	1
85v-	HYRYSALMI	Suuri palvelutarve	Perustason vastaanottopalvelut	Tukipalveluasiakkaat	1
85v-	HYRYSALMI	Suuri palvelutarve	Päivystyspalvelut	Kaikki muut asiakkaat	1

Arvioitaessa summatason faktan anonymisointia todettiin, että näistä tiedoista ei saa selville yksittäistä henkilöä. Henkilötason fakta poikkeaa tästä siten, että jokainen summataso avautuu vielä tapahtumapäivän tasolle. Projektissa todettiin, että tapahtumapäivämäärien lisääminen ei merkittävästi paranna mahdollisuuksia tunnistaa henkilöä yllämainituilla tiedoilla. Faktataulussa on käytetty yksilöivänä tunnisteena eri avainta kuin tietovarastossa, jolloin sitä ei voida yhdistää mihinkään muuhun henkilöä koskeviin tietoihin, vaikka tällaisiin olisi pääsy.

Lisäksi arvioitiin, että mikäli noiden tietojen perusteella toimintaympäristöä ja henkilöitä tarkasti tunteva henkilö pystyisikin tunnistamaan tiedoista yksittäisen henkilön, niin kuinka merkittävää tietoa hän siitä saisi.

13.12.2019

Todettiin, että tiedon merkittävyys on vähäinen, koska se rajoittuu pääasiassa tietoon, jonka hän tietäisi työtehtäviensä vuoksi muutenkin.

Jäännösriskiä arvioitiin myös edistyneen analytiikan toteutuksen kohdalla. Itse ennustemallinnuksessa jäännösriskiä ei synny, koska esim. 12 kuukauden mallissa riskissä olevia oli 375 henkilön ryhmä. Ennustemallinnusta seuranneessa asiakassegmentoinnissa sen sijaan syntyi pieniä jopa yhden henkilön ryhmiä. Segmentoinnin tarkoitus oli tässä kyetä kuvaamaan ryhmiä siten, että ryhmille voisi suunnitella kohdennettuja toimia säännöllisiin palveluihin päätyminen ehkäisemiseksi tai lykkäämiseksi. Kun pyritään löytämään ryhmäkohtaisia, yleistettävissä olevia ratkaisuja, ei hyvin pienet ryhmät ole edes arjen toteutuksen kannalta relevantteja, joten pienimmät ryhmät jätettiin tässä tulosten ja siten käsittelyn ulkopuolelle. Joissain tarkasteluissa juuri tällaiset pienet ”ääriryhmät”, jotka sisältävät hyvinkin tunnistettavia henkilöitä, voivat olla relevantteja ja aihe on siksi tärkeä käsitellä ja ratkaista huolella tapaus- ja mallinnustavoitekohtaisesti.

6.4 Sisällön rajaus

Puolangan kunta ei kuulu Kainuun soteen sosiaalihuollon ja perusterveydenhuollon osalta, joten niihin liittyviä tietoja ei voitu pilotin aikana käyttää. Kunta kuitenkin käyttää samaa perusjärjestelmää Kainuun Soten kanssa, joten Puolangan tiedot piti poistaa siirrettävästä aineistosta. Tämä poistaminen tehtiin pilotin aikana siten, että rajaus tehtiin jo tietojen poimintavaiheessa. Näin edes tietovaraston load-alueelle ei tullut Puolangan tietoja. Poistaminen tehtiin teknisesti yksiköiden kautta.

Koska rajaamisen peruste on tuotetun palvelun toimipaikka eikä asiakkaan oma kotikunta, niin aineistossa on rajausten jälkeen edelleen sellaisia henkilöitä, joiden kotikunta on muualla kuin Kainuun soten alueella. Rajaus ei poista myöskään Puolangalla asuvia henkilöitä sikäli kuin he ovat käyttäneet Kainuun soten palveluita.

Rajaamisen periaatteet olivat seuraavat:

- kontakti (poistetaan määriteltyjen yksiköiden kontaktit)
- kh_kontakti (poistetaan kh_kontaktit, joiden kontakti on määritelty poistettavaksi, kts. edellinen)
- osastojakso (poistetaan määriteltyjen yksiköiden osastojaksot)
- osastohilmo (poistetaan osastohilmot, joiden osastojaksolla oleva yksikkö on määritelty poistettavaksi)
- potilasdiagnoosi (poistetaan määriteltyjen yksiköiden potilasdiagnoosit)
- kayntisyys (poistetaan käyntisyys, joiden kontakti on määritelty poistettavaksi)
- kontakti_palvelu (poistetaan kontakti_palvelut, joiden kontakti on määritelty poistettavaksi)
- kh_paatos (poistetaan määriteltyjen alueiden päätökset)
- kh_palvelupaatoshistoria (poistetaan määriteltyjen alueiden päätökset)
- kh_palvelupaatos (poistetaan määriteltyjen alueiden päätökset)

13.12.2019

7. Tulokset

7.1 Tuloksellisuuden näkökulmat

Sote-palvelujen järjestäjän tietotarpeisiin vastattiin yhteisesti sovittujen lähtökohtaisten rajausten mukaisesti. Pilotissa tavoiteltiin tietosisältöjä maakuntien tietomallityössä määrittelemistä ja myöhemmin priorisoiduista tuloksellisuuden näkökulmista. Tietoa tuotettiin kestävästä taloudesta, taloudellisuudesta, palvelukyvyistä sekä asiakastietämyksestä ja -ymmärryksestä ja osittain myös voimavarojen kohdentamisesta (vrt. kuva 1, sivu 6). Kaikessa tietosisällössä tarkasteltiin yli 75-vuotiaiden väestöryhmää ja heidän koko palvelukäyttöään sekä niihin liittyviä kustannuksia. Perinteisen organisaatiolähtöisen tarkastelun sijaan tarkasteltiin Kainuun soten toimintaa ja kustannuksia yli vastuualuerajojen asiakassegmentin mukaisesti. Tavoitteena oli saada kokonaiskuva kyseisen väestöryhmän tilanteesta niin toiminnan kuin talouden näkökulmasta ja palvelutarpeiden ennakkoinnista. Tietoja vertailtiin ajanjaksojen 1-8/2018 ja 1-8/2019 välillä.

Tilannekuvassa tarkasteltiin kokonaiskustannuksia liittyen yli 75-vuotiaiden väestöryhmään. Kokonaiskustannukset pitivät sisällään sekä oman toiminnan, ostopalveluiden, palvelusetelillä tuotettujen palvelujen, että omaishoidon kustannukset. Vertailussa kuvattiin myös oman toiminnan kokonaiskustannuksia ja asiakaskohtaisten kustannusten muutosta tarkasteluvuosien välillä kunnittain ja palveluittain. Oman toiminnan osalta kustannusten prosentuaalinen kasvu on ollut kuluvaan vuoden aikana viime vuoden vastaavaan ajankohtaan verrattuna 6,14%, ostopalveluiden ja omaishoidon kustannuskasvu on ollut 1,9%. Kokonaismenot ovat kasvaneet 4,61%. Oman toiminnan osalta asiakasmäärässä on vuosien välillä pientä vähennystä (0,84%) eli kysynnän kasvu ei selitä kustannusten kasvua. Ostopalvelujen asiakasmääriä ei ole käytettävissä, joten siltä osin ei pystytty kuvaamaan kustannuksia €/asiakas.

Toisessa näkymässä (kuva 43) tarkasteltiin kokonaiskustannusten jakautumista kaikille yli 75-vuotiaille asukkaille. Kyseinen tarkastelu oli mahdollista toteuttaa vain meneillään olevan vuoden osalta, koska pilotissa ei toteutettu integraatiota väestörekisterikeskuksen järjestelmään ja ennakkotieto väestöstä oli saatavissa manuaalisesti lisäten vain vuodesta 2019. Vuosien välinen vertailu olisi tuonut järjestämistehtävään lisäinformaatiota.

Sekä toimintaa että kustannuksia tarkasteltiin koko sote-kuntayhtymän tilanteen lisäksi vertaillen niitä palveluittain ja kuntayhtymän kunnittain toisiinsa. Myös kuntien välistä kustannusvertailua vaikeutti mahdollisuus liittää asiakastietoja väestörekisterikeskuksen tietoihin. Näin ollen kustannustietojen suhteuttaminen väestöön ei ollut mahdollista ja absoluuttisista luvuista on vaikea tulkita kuntien todellista tilannetta kunnittaisen väestöpohjan suuren kokoeron vuoksi. Meneillään olevan vuoden osalta asukaskohtaisissa keskiarvokustannuksissa kuntien väliset erot olivat suurimmillaan jopa 30 prosenttia. Raportoinnissa tarkasteltiin kunnittain ja palveluittain myös asiakkaiden toimintakykyä ja hoitoisuutta. Sen sijaan eri kuntien alueella olevien toimintayksiköiden toimintamallien sisällöstä ei tietoa ollut käytettävissä tarkemmalla tasolla.

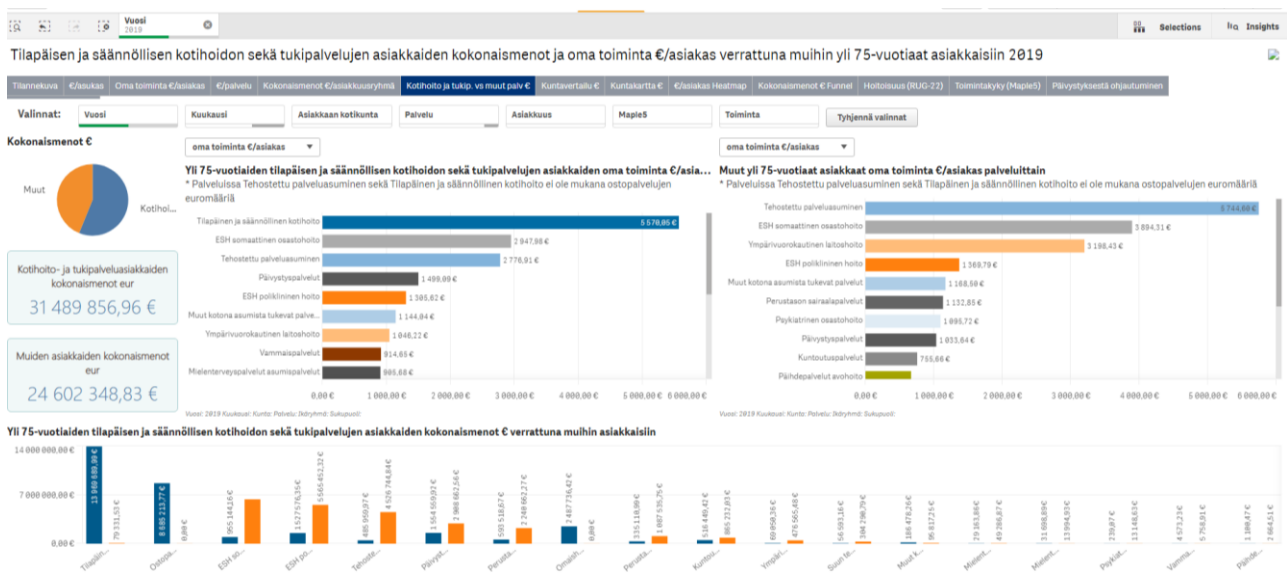
13.12.2019

€/asukas kunnittain yli 75-vuotiaat 2019



Kuva 43 Yli 75-vuotiaiden kustannukset €/asukas kunnittain tammi-elokuussa 2019.

Yhtenä tarkastelunäkökulmana oli eri palveluiden käyttö ja niistä aiheutuneet kustannukset sekä palveluiden integraatio väestötasolla palveluiden välillä. Kyseisellä tarkastelulla (kuva 44) pystyttiin tekemään havaintoja säännöllisissä palveluissa olevien asiakkaiden muusta palvelukäytöstä verrattuna ei-säännöllisissä palveluissa oleviin asukkaisiin. Tarkastelua tehtiin erityisesti kotihoidon asiakkuuksille. Näyttää siltä, että säännöllisen kotihoidon palvelu hillitsee muuta palvelukäyttöä.



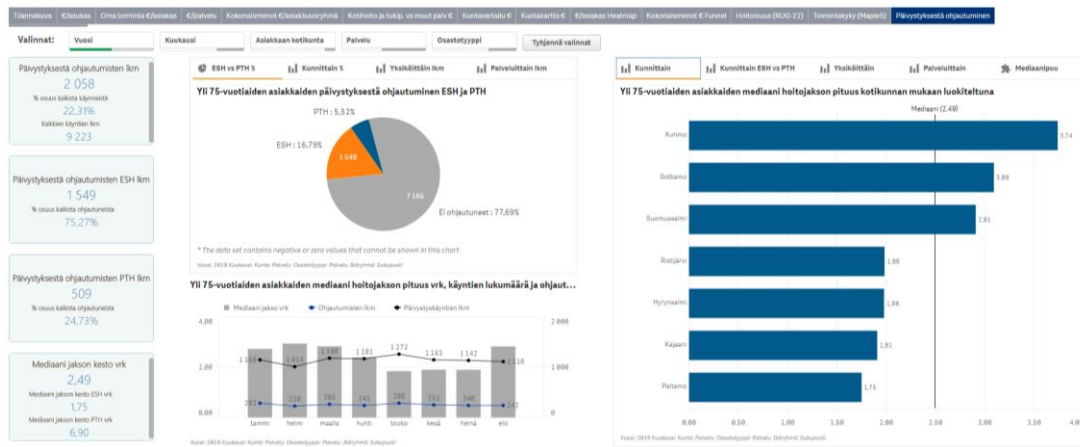
Kuva 44 Tilapaisen ja säännöllisen kotihoidon asiakkaiden palvelukäytön vertailu muiden yli 75-vuotiaiden palvelukäyttöön.

Kainuun sote -kuntayhtymässä haluttiin tarkastella yli 75-vuotiaiden päivystyspalveluiden käyttöä ja asiakkaiden ja potilaiden jatko-ohjautumista päivystyksestä muihin palveluihin, erityisesti perustason sairaalahoitoon ja erikoissairaanhoidon osastohoitoon (kuva 45). Samoin tarkasteltiin osastolla viipymisen mediaaniaikaa asiakkaiden ohjautuessa edellä kuvattuihin palveluihin. Kyseisen väestöryhmän päivystyspalvelua käyttäneistä henkilöistä perustason sairaalapalveluihin ohjautui 5,52% ja

13.12.2019

erikoissairaanhoidon osastoille 16,79%. Mediaaniajat osastohoidossa vaihtelivat välillä 1,61-8,67 päivää ollen erikoissairaanhoidossa kautta linjan selkeästi lyhyempiä. Asukkaiden kotikuntien välillä samoin oli selkeää vaihtelua hoitojaksojen mediaanipituuksissa.

Päivystyksestä ohjautuminen ja hoitojakson mediaani yli 75-vuotiailla asiakkailla



Kuva 45 Yli 75-vuotiaiden kainuulaisten päivystyspalvelun käyttö, päivystyksestä ohjautuminen ja siitä seuranneiden osastohoitojaksojen mediaanipituuksien kunnittain.

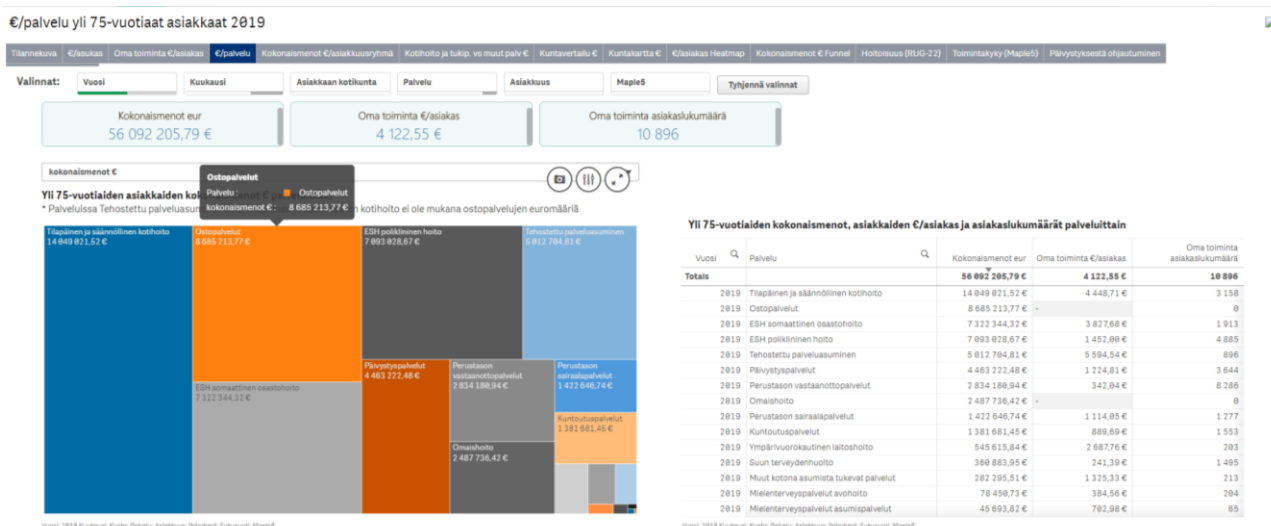
Pilotissa toteutettiin manuaalinen lataus RAIsoft-järjestelmän tiedoista ja näin oli mahdollista tarkastella toimintakykytietoa suhteessa kustannuksiin MAPLe5-luokituksen mukaisesti (kuva 46) sekä tarkastella palveluissa olevien asiakkaiden toimintakyvyn eroja kunnittain. Samoin asiakkuuksien hoitoisuutta tarkasteltiin RUG22-hoitoisuusluokituksen avulla. Tulokset osoittivat kunnittain eroja asiakkaiden toimintakyvyssä ja antavat pohjaa vielä varmistaa, että asiakkaat saavat oikeanlaista palvelua suhteessa toimintakykyyn. Kaikilta osin hoidosta tai palvelusta aiheutuneet kustannukset eivät kohdanneet suoranaisesti asiakkaan toimintakyvyn kanssa. Samoin hoitoisuuksissa oli kunnittaista eroavuutta. Osittain syntyi kuvaa, että jossakin kunnissa resurssit ja palvelutarve kohtasivat paremmin ja toisissa taas eivät kohdanneet kaikilta osin yhtä linjakkaasti.



Kuva 46 Yli 75-vuotiaiden toimintakykytiedot kunnittain MAPLe 5 -luokituksen mukaan kunnittain ja palveluittain.

13.12.2019

Kainuun sotessa on käytössä palveluseteli- ja ostopalvelujärjestelmä. Tällä hetkellä järjestelmä pitää sisällään pääsääntöisesti palvelusetelillä tuotetut palvelut. Ostopalveluista on tällä hetkellä saatavissa vain kokonaiskustannukset eikä niitä ole mahdollista liittää asiakkuuteen. Näin ollen kokonaiskustannuksia ei kyetä osoittamaan palveluittain niistä palveluista, joissa toteutuu myös ostopalvelua tai palvelusetelillä tuotettua palvelua. Se vaikeuttaa järjestämistehtävässä tarvittavan kokonaiskuvan muodostamista (kuva 47).



Kuva 47 Kokonaiskustannusten jakautuminen palveluittain. Ostopalvelut kuvattuna omana kokonaisuutenaan, koska tietoa ei pystytty kytkemään asiakkuuksiin eikä siten sisällyttämään palveluihin.

7.2 Kehittynyt analytiikka ja koneoppiminen

Projektissa tuotetut kehittyneen analytiikan mallit ennakoivat jo erittäin hyvin säännöllistä palvelutarvetta. Mallit tunnistavat ne asiakkaat, joihin kannattaa 12 ja 36 kuukautta etukäteen kohdentaa toimia säännöllisiin palveluihin ohjautumisen ehkäisemiseksi tai lykkäämiseksi. Lisäksi tuotettiin kuvaukset datapohjaisesti muodostetuista asiakasryhmistä interventoiden ja kohdennettujen toimenpiteiden mahdollistamiseksi.

12 kuukauden ennustemalli

Opetuksen jälkeen muodostettua mallia testattiin erillisellä testiaineistolla, jota koneoppimisen malli ei ollut aiemmin nähnyt. Malli osuu oikeisiin lopputuloksiin 88,1 prosentin tarkkuudella kaikista ennustettavista henkilöistä eli ennustaa oikein ne, jotka ohjautuvat säännöllisiin palveluihin ja toisaalta ne, jotka eivät tarvitse säännöllisiä palveluja 12 kuukauden sisällä (tarkkuus). Mallin ennustamista, säännöllisen palveluihin ohjautuvista henkilöistä 85,6 prosenttia myös todellisuudessa saa 12 kuukauden sisällä säännöllisen palvelun päätöksen (täsmällisyys). Malli tunnistaa vuoden päästä säännöllisen palvelun päätöksen saavista henkilöistä 75,6 prosenttia (tunnistuskky).

Työryhmässä arvioitiin, miksi joukossa on myös niitä, jotka eivät tule saamaan päätöstä, vaikka datan pohjalta siltä vaikuttaa. Yhtenä mahdollisena syynä tuli esille se, että osaa tämän tyyppisistä henkilöistä hoidetaan kotona ilman, että hoitaja saa omaishoidon tukea. Se että kaikkia päätöksen saavia ei tunnisteta, saattaa johtua siitä, ettei kaikista henkilöistä ole tarvittavaa tietoa käytettävissä. Tai sitten syy palvelutarpeelle on äkillinen, esimerkiksi tapaturmasta tai akuutista sairastumisesta johtuva. Tätä malli ei etukäteen pysty

13.12.2019

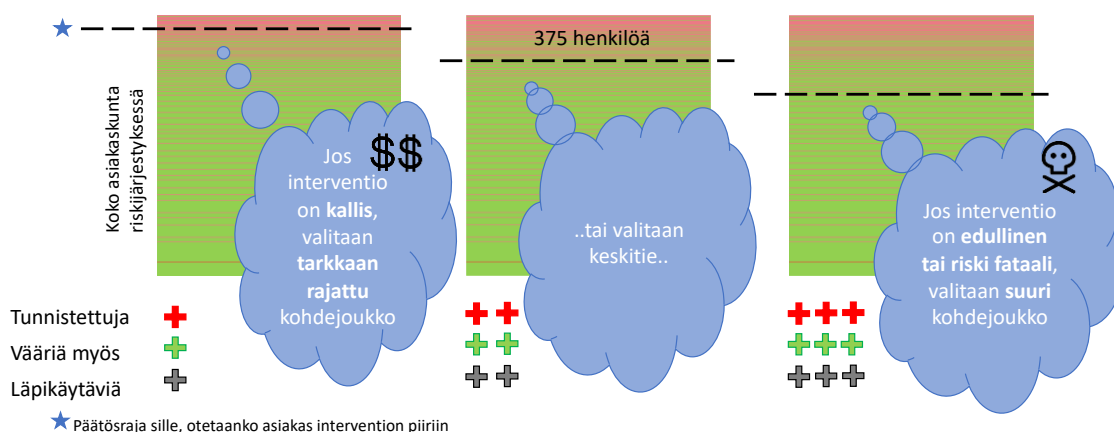
ennustamaan. Tulosten tarkempi arviointi vaatisi lisäselvitystä ja esimerkkitapausten tarkempaa läpikäyntiä asiakastasolla.

36 kuukauden ennustemalli

Vastaavasti toteutettiin myös 36 kuukauden malli opettamalla ensin opetusaineiston avulla malli ja sen jälkeen testaamalla opetettua mallia erillisellä testiaineistolla. Tämä malli on osittain parempi kuin 12 kuukauden malli. Malli osuu oikeisiin lopputuloksiin 89,5 prosentin tarkkuudella kaikista ennustettavista henkilöistä eli ennustaa oikein ne, jotka ohjautuvat säännöllisiin palveluihin ja jotka eivät tarvitse säännöllisiä palveluja 36 kuukauden sisällä (tarkkuus). Mallin ennustamista, säännöllisen palveluihin ohjautuvista henkilöistä 87,9 prosenttia myös todellisuudessa ohjautuu säännöllisiin palveluihin 36 kuukauden sisällä (täsmällisyys). Malli tunnistaa kolmen vuoden päästä säännöllisen palvelun tarpeessa olevat henkilöt 75,6 prosentin todennäköisyydellä (tunnistuskky).

Koska säännöllisiin palveluihin ohjautumisen jakauma on hyvin vino, eli riskissä olevia on merkittävästi vähemmän kuin ei-riskissä olevia, on keskeistä tarkastella sekä mallin tunnistuskkyä, että täsmällisyyttä eri päätösrajan arvoilla (kuva 48).

Tiedon käyttötarkoitus ratkaisee mallivalinnan



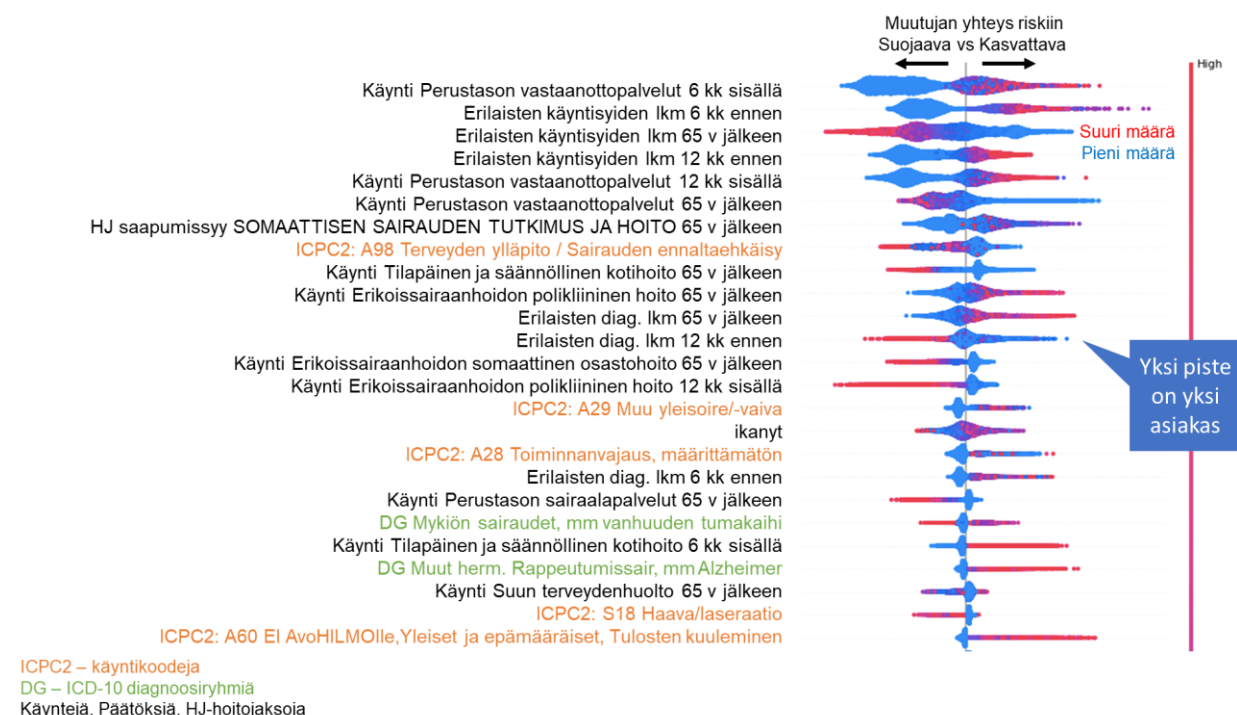
Kuva 48 Mallin valinta.

Päätösrajaa säätämällä voidaan vaikuttaa siihen, miten moni riskissä oleva tunnistetaan ja miten monta asiakasta on kaikkiaan otettava interventioiden piiriin näiden riskissä olevien tunnistamiseksi. Mitä korkeampi päätösraja, sitä parempi täsmällisyys eli joukossa on vähemmän niitä, jotka eivät olleetkaan riskissä. Samaan aikaan korkeampi päätösraja laskee tunnistuskkyä eli useampi riskissä oleva jää päätösrajan ulkopuolelle eli jää tunnistamatta. Tarkoitukseen soveltuva päätösraja saadaan eri toimijoiden keskustelussa. Käytettävä päätösraja voi myös vaihdella eri interventioiden suunnittelussa, eri tarkoituksiin voidaan valita eri joukkoja samojen ennusteiden pohjalta.

13.12.2019

Ennustemallin ajaminen syyskuun 2019 datalla

Opetettuun ennustemallin ajettiin seuraavaksi uusin mahdollinen 2019 data-aineisto. Tulosten mukaan uusia 12 kuukauden sisällä säännöllisten palveluihin ohjautumisen riskissä olevia tunnistettiin aineistosta 375 henkilöä. Mallin ennustetulosta tarkasteltiin mm. SHAP-kaavion (kuva 49) avulla. SHAP-kaavio kertoo mallin merkitsevimmät muuttujat, sekä muuttujien vaikutuksen suunnan ja suuruuden kunkin ilmentymän kohdalla. Kaaviosta nähtiin, että suuret kuin toisaalta pienet määrät käyntejä, diagnooseja, hoitokasjoja tai tiettyjä ei-säännöllisiä päätöstyypppejä ennakoit tapauksesta riippuen säännöllistä palvelutarvetta. Merkitsevien muuttujien osittainen ristiriitaisuus kertoo siitä, että taustalla on hyvin erilaisia asiakasryhmiä ja asiakaspolkuja.



Kuva 49 Esimerkki datapisteiden/muuttujien vaikuttavuusjärjestyksessä (SHAP-kaavio) – listan ensimmäiset kuvassa.

Tulosten hyödynnettävyys paranee merkittävästi entisestään asiakasryhmien ja -polkujen lisätarkasteluilla. Interventoiden suunnittelemisen kannalta on hyödyllistä ymmärtää kokonaisuutta: millaisia yksilöitä, ja tai millaisten ilmentymien ympärille kunkin ryhmän asiakkaat kuuluvat, eli kuvata ns. arkkityypit.

Kustannusten kannalta on myös tärkeää tunnistaa henkilöt, joilla ei ole ennusteiden mukaan riskiä ohjautua säännöllisiin palveluihin 12 tai 36 kuukauden sisällä. Ennustemallien avulla voidaan tunnistaa myös mahdolliset ylipalvellut asiakassegmentit ja tätä kautta mahdollistaa toimenpiteiden kohdistaminen rajallisten resurssien tilanteessa oikeisiin kohteisiin.

Asiakassegmentointi

Tulokset ovat jo nykyisellään hyödynnettävissä eli malli voi kertoa keihin 375 henkilöön kannattaa panostaa interventioita säännöllisiin palveluihin päättymisen lykkäämiseksi tai ehkäisemiseksi. Tulosten hyödynnettävyys paranee merkittävästi entisestään asiakasryhmien ja -polkujen lisätarkasteluilla.

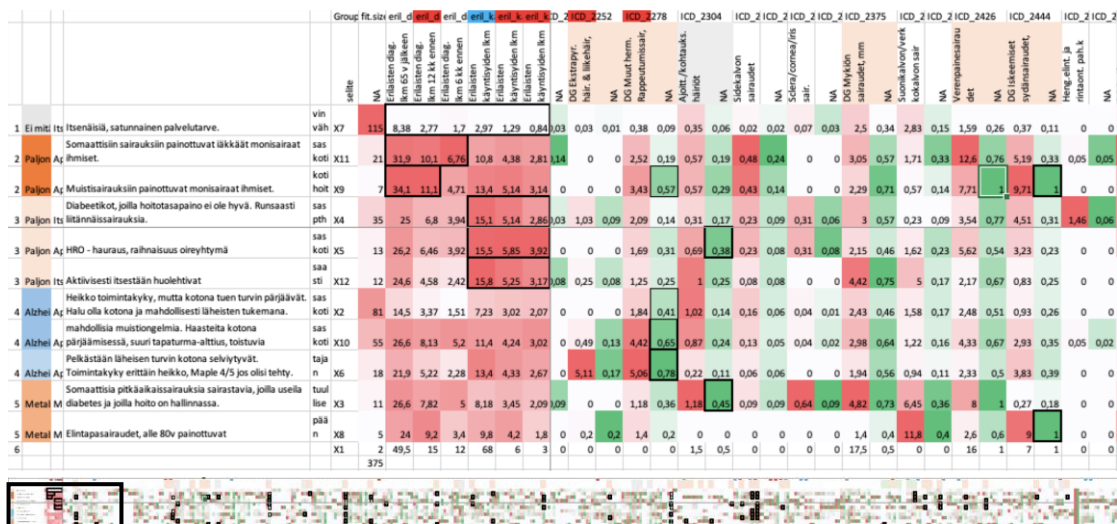
13.12.2019

Interventioiden kohdentamisen helpottamiseksi 375 riskissä olevaa asiakasta päätettiin ryhmitellä eli segmentoida K-means -menetelmällä (kuva 50).

K-means jakaa asiakkaat keskenään mahdollisimman samanlaisiin, mutta toisista ryhmistä erottuviin ryhmiin. K-means on hyvin herkkä lukuarvojen suuruuksille ja skaalauksesta huolimatta herkästi laittaa samoihin ryhmiin asiakkaita, joilla on samankaltaisia määriä, esimerkiksi hyvin suuria tai pieniä määriä, mitä tahansa käyntejä tai diagnooseja.

Interventioiden suunnittelemisen kannalta on kuitenkin tärkeä ymmärtää pelkkien käyntimäärien lisäksi kokonaisuutta, eli millaisia yksilöitä, ja tai millaisten ilmentymien ympärille kuhunkin ryhmään kuuluvat asiakkaat kuuluvat, eli kuvata ns. arkkityypit. Arkkityyppien nimeämiseksi tehtiin useita segmentointeja eri muuttujapainotuksin, ennen mielekkäiden ja arjessa tunnistettavien ryhmien hahmottumista.

Asiakkaiden segmentointi



Kuva 50 Asiakkaiden segmentointi.

Arkkityypit datan pohjalta jaoteltiin seuraaviin asiakassegmentteihin:

Asiakassegmentti	Osuus (prosenttia)
Itsenäisiä, satunnainen palvelutarve	27
Heikko toimintakyky, mutta kotona tuen turvin pärjäävät. Halu olla kotona ja mahdollisesti läheisten tukemana.	19
Fyysinen toimintakyky heikko, ehkä masennusta ja mahdollisia muistiongelmia. Haasteita kotona pärjäämisessä, suuri tapaturma-alttius, toistuvia sairaalahoitojaksoja.	13
Diabeetikot, joilla hoitotasapaino ei ole hyvä. Runsaasti liitännäissairauksia.	8
Somaattisiin sairauksiin painottuvat iäkkäät monisairaajat ihmiset	5
Pelkästään läheisen turvin kotona selviytyvät. Toimintakyky erittäin heikko, Maple 4/5 jos olisi tehty.	4
HRO - hauraus, raihnaisuus oireyhtymä	3

13.12.2019

Aktiivisesti itsestään huolehtivat	3
Somaattisia pitkäaikaissairauksia sairastavia, joilla useilla diabetes ja joilla hoito on hallinnassa	3
Muistisairauksiin painottuvat monisairaajat ihmiset	2
Elintapasairauksia, alle 80v painottuvat	1

Ryhmät ovat keskenään hyvinkin erilaisia. Esimerkiksi suurin ryhmä, johon kuuluu 27 prosenttia riskissä olevista, keskeisin nimittävä tekijä on, että heillä on muihin verrattuna hyvin vähän eli vain satunnaista palvelutarvetta. Ennustemallin merkitsevissä muuttujissa pieni palvelukäyttö koko ajalla 65 ikävuoden jälkeen oli yksi merkittävimpiä riskitekijöitä ja kuvaa siis erityisesti tämän ryhmän riskiä. Toisaalta tästä ryhmästä selkeästi poiketen löytyi muutamia ryhmiä, joissa asiakas vaikuttaa olevan jo hyvinkin haastavassa tilanteessa liitännäissairauksineen. He tuntuvat kuitenkin pärjäävän kotona joko tuen ja tai omaisen, jolla ei ainakaan vielä ole omaishoitajapäätöstä, turvin tai sinnittelevät yksin jatkuvan tapaturma-alttiuden varjossa.

Tärkeä on huomata myös, että suurimman ryhmän, satunnaisten palvelutarpeen ryhmässä, voi olla myös lievempiä versioita muihin ryhmiin kuuluvista. Tämän varmistaminen vaatisi jatkoselvityksiä ja esimerkiksi potilaspolkutarkasteluja sekä yksittäisten tapausten tarkempaa analysointia.

Asiakasryhmien kuvaaminen tuo sote-asiantuntijoille parhaimmillaan kuvaukset kohderyhmistä ja näin mahdollisuuden suunnitella täsmälliset toimenpiteet kunkin asiakasryhmän tarpeita vastaavasti. Ennuste voidaan ajaa yhä uudestaan ja kohdennetut toimet voidaan toteuttaa aina uusille tarpeeseen nousseille asiakasryhmille, joissain tapauksissa jopa automatisoidusti ja digitaalisesti. Kun toimenpiteet on toteutettu, voidaan toteumaa verrata ennusteeseen ja voidaan alkaa seurata, mitata ja evaluoida toimenpiteiden vaikuttavuutta. Näin voidaan siirtyä ennakoivan ja ohjaavan analytiikan käyttöön sekä analytiikkaratkaisun tuotannollistamiseen.

Hypoteesin jatko-osa, eli toimet ja kustannuskehityksen hillitseminen olisi seuraava vaihe. Viemällä työn tulokset käytäntöön ja luomalla prosessit jatkuvalla riskiherätteiden seurannalle, toimenpiteiden toteutukselle ja vaikuttavuuden mittaamiselle, päästään vaiheeseen, jossa analytiikan tuloksin voidaan ohjata toimintaa.

13.12.2019

8. Kokemukset ja suositukset

Pilottiprojektissa tehtyjen havaintojen perusteella muodostettiin suosituksia palvelemaan kansallisen tason tietojohdamisen jatkokehittämistä. Näiden suositusten tavoitteena on tukea eri toimijoita tietojohdamiseen liittyvissä suunnittelu- ja kehittämistehtävissä. Suositukset palvelevat sote-palveluiden järjestäjiä järjestämistehtävän toteuttamisessa ja toisaalta osaltaan varmistavat kansallisesti yhtenäistä kokonaisuutta sekä arkkitehtuurin, että tietosisältöjen näkökulmasta. Järjestämistehtävää ei tule tarkastella pelkästään organisaattiorakenteen mukaisesti, vaan myös asiakassegmenteittäin yli toimintorajojen.

8.1 Yleisiä havaintoja

Usean toimijan yhteistyötä vaativassa projektissa tarvitaan suunnitteluun ja toteutukseen väline, jonka avulla työtä ohjataan koko prosessin ajan. Pilotissa käytettiin hypoteesicanvas- ja koneoppimisen canvas-malleja jäsentämään koko tietojohdamisen kuva lähtien strategiasta ja päätyen tiedon hyödyntämiseen saakka. Kokemusten perusteella mallin hyödyntäminen auttoi säilyttämään ”punaisen langan” läpi projektin kaikilla työhön osallistuneilla.

Tietojohdamisen kehitystyössä tarvitaan edustus ainakin palvelutoiminnasta, taloushallinnosta ja tietohallinnosta. Järjestämistehtävän tietosisältöjen valmistelussa tarvitaan toiminnan osaamista yli toimintorajojen. Eri toimijoiden yhteisillä työpajoilla on oma roolinsa kehitystyössä, etenkin sen alkuvaiheessa. Projektin ketteryuden, taloudellisuuden ja tehokkuuden toteutumiseksi tulee osallistujien roolit ja tarvittavat osaamiset määritellä huolellisesti projektin alussa. Projektinhallinnan näkökulmasta on hyvä käyttää kaikille toimijoille yhteisiä työnohjauksen työkaluja. Tuotetun tiedon ja analyysien tarkastelu ja hyödyntäminen tulee toteuttaa yli organisaatio- ja toimintorajojen monialaisesti ja moniammatillisesti.

Projektissa haettiin tehokkaita ja ketteriä toteutustapoja, jotka mahdollistavat laajojen toteutusten tekemisen siten, että niitä voidaan edistää siivu kerrallaan. Projektin kokemusten pohjalta tällainen on mahdollista käyttämällä oikein valittuja työvälineitä. Muutosjoustavuus eli siivuittain ja osa-alueittain tekeminen mahdollistuu käytetyn tietomallinnustavan ansiosta. Data vault mahdollistaa tietovaraston sisällön laajentamisen siten, että uusi osuus voidaan lisätä entiseen kokonaisuuteen ilman, että aiemmin rakennettua tarvitaan uudelleen organisoida. Projekti suosittelee siis data vault -mallinnusmenetelmän valitsemista kehitystyön pohjaksi tietojohdamisen hankkeissa.

Data vault -mallinnus tuottaa tietovarastoon paljon tauluja, tässä projektissa tuotettiin 209 taulua jakautuen hub-, satelliitti-, linkki-, referenssi-, state- ja load-tiluihin. Tietojen lataus perinteisellä ETL-prosessilla data vault-mallin mukaiseen tietovarastoon on paljon resursseja vaativaa ja virhealtista. Toteutussiivujen tekemisen tulee suurissa kokonaisuuksissa olla nopeaa ja tehokasta, ja sen lisäksi tulee pyrkiä minimoimaan käsityönä tehtäviä virheille alttiita työvaiheita. Projektin kokemusten pohjalta suositellaan automaattisen ELT-ohjelmiston käyttämistä tietojen lautaukseen.

Laajan kokonaisuuden hallinnassa tulee saada kaikki organisaation sidosryhmät ymmärtämään toteutettavaa kokonaisuutta. Tietojohdamisen toteutuksessa on tärkeää ymmärtää järjestelmässä olevat ja sinne ladatut tietosisällöt. Projektissa käytettiin käsittemallinnusta tietosisältöjen ymmärtämisessä. Käsittemallinnuksen kautta sekä johto- ja loppukäyttäjät että ICT pystyvät hahmottamaan helposti mitä tietosisältöjä kokonaisuudessa on ja mitä sinne on tulossa. Käsittemallinnus lisää eri osapuolten ymmärtämistä ja kommunikointia ja sen käyttöä suositellaan.

13.12.2019

Perinteisesti tietojohdamisen projekteissa käsittemallinnus jää irralliseksi osaksi ja saatetaan tehdä jopa jälkikäteen. Lisäksi käsittemallinnusta saatetaan tehdä tietovaraston näkökulmasta kuvaten käytännössä tietovaraston taulurakennetta. Projekti suosittelee käsittemallinnuksen tekemistä toimintalähtöisesti, ammattilaisten ymmärtämin käsittein, ja näin muodostettujen käsittemallien automaattista integrointia data vault -kannan luontiin. Näin ketju käsittemallista tietovaraston rakenteisiin ja tietosisältöön saadaan saumattomaksi ja joustavaksi.

Käyttäjiltä saatava palaute projektin aikaisessa vaiheessa on tärkeää. Käyttäjille tulee aikaisessa vaiheessa muodostaa raportteja ja visualisaatioita jopa keskeneräisestä tietosisällöstä. Projektin kokemusten pohjalta tämä kannattaa tehdä, jo siinä vaiheessa kun tietosisältöjä ollaan lataamassa tietovarastoon. Näin käyttäjät ja tietosisältöjen asiantuntijat pystyvät ottamaan kantaa mahdollisiin ongelmiin ja visualisaatioita voidaan käyttää keskustelujen pohjana. Visualisaatiot luovat myös kehittäjille yhteisen näkymän ja alustan keskustelun käymiseen ja ladattavan datan tietosisältöihin.

Projektin aikana organisaatiota kannattaa valmistaa kehittyneen tietojohdamisen järjestelmän käyttöönottoon ja hyödyntämiseen. Projektin kokemusten pohjalta suositellaan tietojohdamisen nykytilan selvitystä ja arviointia, ja sen pohjalta tehtävän kehittämissuunnitelman laatimista. Kehittämissuunnitelmassa tulisi huomioida tietojohdamisen strategia ja muodostaa hallintamalli, jolla tietojohdamisen kokonaisuutta hallitaan ja kehitetään.

Lähdejärjestelmien puuttuva tai huonolaatuinen data vaikutti projektiin tiettyjen tietosisältöjen osalta. Lähdetietojen selvittely hidasti projektin etenemistä. Jatkossa lähdejärjestelmien tietosisältöä kannattaa tarkastella siihen soveltuvilla data preparation -välineillä heti projektin alusta lähtien osana käsitemäärittelyä. Myös kirjaamiskäytännöt vaihtelevat organisaatioissa ja tämä voi vaikuttaa datan laatuun. Puutteellisten tai virheellisten kirjausten osalta voidaan arvioida, onko niitä mahdollista täydentää esim. robotiikan tai edistyneen analytiikan keinoin. Tämä tulisi ottaa huomioon jo projektin suunnitteluvaiheessa.

Virtuaalisesti toteutettuna projektissa säästettiin matkakustannuksissa, mutta tämä aiheutti toisaalta jossain määrin tehottomuutta osallistujien ajankäytön ja ongelmanratkaisun näkökulmasta. Tulevissa toteutuksissa kannattaa arvioida onko joitakin työtapoja syytä järjestää yhteisissä tiloissa.

8.2 Kansallisille toimijoille

Kansallisesti tulisi päästä tilanteeseen, jossa useita päällekkäisiä tiedonkeruita eri tiedon tarvitsijoille ei tarvittaisi. Tämä edellyttää kansallisesti yhteensopivia tietosisältöjä. Tiedonkeruiden yhtenäistämistarve on tunnistettu myös Toivo-ohjelman valmistelussa.

Yksi keskeinen maakuntien raportointivelvoitteita keventävä ja tietosisältöjen yhdenmukaistumista palveleva toimenpide on useiden eri käyttötarkoituksia varten laadittujen luokitusten yhdenmukaistaminen mahdollisimman pitkälle. Palvelujen järjestämistoiminnossa tarvittavien tietosisältöjen samantasoinen toimittaminen valtionhallintoon palvelee myös kansallisen tason vertailua. Pilotissa hyödynnettiin Sote-tietopakettien palveluryhmä- ja toimintotason luokituksia, jotka palvelevat järjestämistehtävässä tarvittavaa tietosisältöä ja toimivat myös kansallisella tasolla.

Sote-toimialaan liittyviä käsitteitä on jo kattavasti koottu sotesanastot.fi -verkkosivustolle. Käsitteiden ja niiden kuvausten keskittäminen edelleen yhdelle verkkosivustolle parantaisi kansallista yhdenmukaisuutta ja helpottaisi niin alueellisesti kuin kansallisesti toimivia tahoja. Erityistä huomiota tulisi kiinnittää käsitemäärittelyyn, jossa toiminnan kannalta limittyy sekä terveydenhuoltolain että sosiaalihuoltolain mukainen toiminta. Tämän pilotin aikana esimerkiksi kaikille kotihoidon toimintaympäristöön liittyville

13.12.2019

käsitteille ei ollut saatavilla selkeää määritelmää terveydenhuollon eikä sosiaalihuollon keskeisistä sanastoista. Jatkossa olisikin tärkeää saada kansalliset määrittelyt ketterämmin käyttöön tietojohdantamisen toteutusprojektien aikana.

Kustannuslaskennan kansallisen mallin työstäminen edellyttää alueiden ja kansallisen tason toimijoiden yhteistyötä. Jo pilotin alkuvaiheessa tiedettiin, ettei kansallista mallia kustannuslaskennan toteuttamiselle ole olemassa kattavasti. Pilotin aikana käytiin keskusteluja THL:n kanssa yhtenäisen kustannuslaskennan mallin tarpeellisuudesta. Jotta kansallisesti päästään luotettavaan tiedon vertailtavuuteen, tulee palveluiden kustannusten laskeminen toteuttaa alueilla samoilla periaatteilla.

8.3 Sote-palvelujen järjestäjälle

Sote-palvelujen järjestäjän johtaminen perustuu organisaation strategisiin tavoitteisiin. Tietojohdantamisen tuella tavoitellaan organisaation strategisten tavoitteiden saavuttamista ja sen tulisi linkittyä tavoitteiden seurantaan, hallintaan, ohjaukseen ja toiminnan kehittämiseen.

Tällä hetkellä ostopalvelujen osalta toimintatietojen kiinnittäminen kustannuksiin on useissa organisaatioissa hankalaa, koska ostopalvelujen toimintatietoja ei ole saatavissa riittävän tarkalla tasolla suoraan järjestelmistä. Sote-palvelujen järjestäjän tulee saada reaaliaikaista tietoa toiminnan ja talouden suunnitteluun, seurantaan, ohjaukseen, valvontaan ja kehittämiseen. Kokonaiskuvan muodostamista varten tarvitaan tietoa sekä omasta toiminnasta että ostopalveluista ja palvelusetelillä hankituista palveluista eli käytännössä kaikesta järjestämistehtävän sisällöstä.

Sisäisenäkin työnä toteuttava tietojohdantamisen kehitystyö kannattaa asettaa projektiksi tehokkaan etenemisen varmistamiseksi. Työssä on hyvä kuvata tavoitteet, osallistujien roolit, työn kokonaisaikataulu ja eteneminen sekä resursointi. Projektina etenevä kehitystyö luo toteutukseen läpinäkyvyyttä koko organisaation näkökulmasta ja toisaalta varmistaa projektin loppuun saattamisen aikataulussa. Lisäksi projektimuotoinen toteutus mahdollistaa paremman riskienhallinnan ja varmistaa päämäärän saavuttamisen.

Analytiikan ja koneoppimisen käyttöönottoihin ja hyödyntämiseen kannattaa panostaa. Tähän liittyy keskeisesti moderni ja mukautuva tekninen kokonaistoteutus, organisaation toimintatavan syvälinen muutos, datan kerääminen interventioista, vaikuttavuusmittaus ja mallin sekä prosessien jatkokehitys. Tietosisältöihin kannattaa soveltaa edistyneen analytiikan tarjoamia uusia ja alati kehittyviä mahdollisuuksia, kuten ennakkointia, aihetunnistusta, vaillinaisen tai virheellisen tiedon tarkistusta ja täydentämistä, ryhmien tunnistusta, vaikuttavuusmittausta ja mittareita. Joka tapauksessa myös edistyneen analytiikan soveltaminen edellyttää hyvää tietojohdantamisen perusratkaisua ja sen avulla tuotettua kokonaisymmärrystä palveluista, palvelukäytöstä ja kustannuksista koko organisaatiossa.

8.4 Kainuun soten kokemukset

Kainuun soten näkökulmasta pilotti toteutui hyvin. Yhteistyö oli antoisaa ja pilotin toteuttamisen toimintamalli osoittautui tehokkaaksi. SoteDigi pystyi organisoimaan toimivan yhteistyön kokoamalla laajan asiantuntijajoukon Kainuun sotealaisten lisäksi.

Pilotin toteutusaika oli lyhyt ja se haastoi jähmeän julkisen organisaation joustavuuden. Kainuun sote osoitti ketteryyttään mukautumalla pilotin työryhmien viikoittaisiin tapaamisiin. Arvioidut työpanokset ylittivät soten osalta. Tämä kertoo erittäin innokkaasta osallistumisesta niin johdon kuin operatiivisen toiminnan

13.12.2019

osalta. Pilottiin osallistui kolmekymmentä sotelaista Kainuusta. Pilotin substanssi koettiin tärkeäksi ja erittäin ajankohtaiseksi.

Pilotti sujui esimerkillisesti nostamalla esille monia muutostarpeita strategian laadinnasta sen toteutukseen. Pilotissa lähdettiin liikkeelle vanhuspalvelujen strategisesta tavoitteesta ”Osallistuva ikäihminen asuu omassa kodissaan lähiverkostonsa sekä tarpeenmukaisten palvelujen turvin”. Suunnitelman ja toteutuksen edetessä huomattiin, että tavoitteen toteuttaminen vaatii muidenkin kuin vaan vanhuspalvelujen tulosalueen mukana oloa. Strategisen tavoitteen tulee olla koko soten tavoite, eikä niin, että jokainen tulosalue johtaa oman toisistaan hieman poikkeavan tavoitteen organisaation päätavoitteesta.

Pilotissa käytetty hypoteesicanvas todettiin loistavaksi työvälineeksi kompleksisen asian kuvaamiseen. Prosessin aikana jouduttiin aika ajoin palaamaan lähtöruutuun, jotta strateginen tavoite ei unohtuisi. Canvaksen avulla on helppo hahmottaa kokonaisuus, se auttaa pitämään huomion koko pilotoinnin ajan oikeassa asiassa sekä takaa hypoteesien vaatiman tiedon löytämisen.

Prosessin aikana nousi esille myös kirjaamisen tärkeys. Todettiin, että siinä on meillä Kainuussa edelleen kehitettävää mm. diagnoosien kirjaamisen osalta. Lisäksi ostopalvelujen kohdentaminen asiakkaisiin vaati toimintojen uudistamista. Tämä on tärkeää mm. siksi, että oman toiminnan ja ostopalvelujen tiedot saadaan vertailukelpoisiksi ja toisaalta palvelujen järjestämistehtävää varten kokonaiskuva tilanteesta.

Suurin osa tapaamisista järjestettiin etänä Skypeä käyttäen. Näin hiilijalanjälki jäi pitkistä etäisyyksistä huolimatta pieneksi. Yhteydet toimivat pääsääntöisesti hyvin. Työryhmien tapaamiset olivat hyvin suunniteltuja ja aikataulutettuja – toiminta oli tehokasta. Tapaamiset johdettiin napakasti ja kokouskäytännöt olivat hallussa. Etäkokouksiin oli helppo osallistua ja päästä mukaan prosessiin.

Pilotista muodostui tietojohtamisen oppikokemus koko organisaatiolle. Lisäksi saatiin käytännössä kokemusta koneoppimisesta ja sen hyödyntämisestä. Sosiaali- ja terveystietojen toissijaisen käytön lain tietojohtamisen pykälä todettiin tärkeäksi, mutta ko. lain soveltamiseen kaivataan edelleen tukea. Kaiken kaikkiaan Kainuun sote voi hyödyntää pilotissa saatuja tuloksia oman toiminnan tehostamisessa, ennakkoinnissa, palvelujen suunnittelussa ja strategian laadinnassa. Pilotissa toteutettu tiedolla johtamisen malli on laajennettavissa koskemaan muitakin asiakkaita kuin yli 75-vuotiaita. Kainuun soten yhtymähallitus teki jo lokakuun lopulla päätöksen, että pilotissa käytetty tietojohtamisen ympäristö pidetään edelleen pystyssä ja kehitystyötä jatketaan sekä laajennetaan tietosisältöä.

On hienoa, että pieni yhteinen kehittämisaskel on mahdollista vielä ottaa SoteDigin kanssa pilotin jälkeen. Toteutamme yhteistyössä tiedolla johtamisen kyvykkyyden kartoituksen testaamalla tietojohtamisen arviointimallin Kainuun sotessa marras-joulukuussa. Tämä luo hyvän pohjan osaamisen ja koko tietojohtamisen edelleen kehittämiseen. Seuraavassa taulukossa on kuvattu Kainuun soten kokemukset pilotista.

13.12.2019

Plussat	Miinukset
• Erilainen, tehokas ja antoisa tapa tehdä yhteistyötä eri toimijoiden kanssa • Mielenkiintoinen ja innostava • Yhteistyötapa on mahdollistanut asiantuntemuksen jakamisen eri tasoilta ja eri toimijoilta • Iso joukko (n. 30) Kainuun soten johtoa ja työntekijöitä osallisena • Tietojohtamisen vaikuttavuuden parantaminen • Vahvistaa yhdessä toimimisen malleja myös soten eri toimijoiden kesken • Oppimisprosessi • Haastaa organisaation oppimaan • Pilotti on antanut oppia uuden lain täytäntöönpanosta • On saatu uusi työväline – hypoteesicanvas – kehittämistyön välineeksi • Yhteistyö on ollut helppoa ja sujuvaa SoteDigin ja kumppaneiden kanssa • Hiilijalanjälki on jäänyt pieneksi – lähes koko valmistelu toteutettu etänä • Kustannukset tulivat näkyväksi suoritteille ja asiakkuuksille, kustannuslaskenta • Lopputuloksen myötä monta asiaa on tullut näkyväksi • Ketteryys, nopeasti kokeillaan seuraavaa vaihtoehtoa tarvittaessa • Toteutuksen monimutkaisuus on lujittanut uskoa toteutukseen ja tekemiseen, benchmarking	• Pilotin toteutusaika on lyhyt, nopea tekeminen vaatii nopeaa reagointia ja ”hyppäämistä asiasta toiseen” • Operatiivisessa työssä olevan henkilön ajan irrottaminen pilotointiin on ollut haasteellista • Paljon menossa asioita yhtä aikaa • Pilotin toteutus myös kesällä – sijaiset • Työskentelytapa, tietyt työpajat olisi hyvä saada samaan fyysiseen tilaan, sujuvuus • Alusta saakka olisi pitänyt olla laajempi ja moniammatillinen (yli tulosalueiden) • Huono ajankohta • Tietosuojan ja tietoturvan ottaminen huomioon heti prosessin alussa • Johdon osallistuminen, punaisen langan pysyttävä mukana

13.12.2019

Lähteet

- [1]. Laki sosiaali- ja terveystietojen toissijaisesta käytöstä 552/2019: <https://www.finlex.fi/fi/laki/alkup/2019/20190552>
- [2]. Datapedia – Data-alan termit avattuna. 2019. <https://www.arihovi.com/materiaalit/datapedia-data-alan-termit-avattuna/>
- [3]. Aineistohallinnan käsikirja, Tampereen yliopisto <https://www.fsd.uta.fi/aineistohallinta/fi/>
- [4]. Sote-uudistuksen vaikuttavuus- ja kustannustieto-ryhmän työryhmämuistio: Ehdotukset sote-ohjauksen mittareista ja tietopohjan varmentamisesta, Sosiaali- ja Terveysministeriön raportteja ja muistioita, 2019:33, http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161493/STM_2019_33_Ehdotukset_sote_ohjauksen_mittareista.pdf?sequence=1&isAllowed=y
- [5]. Pekkanen Teemu (2019): julkaisematon diplomityö, Informaatioteknologian ja viestinnän tiedekunta, Tampereen yliopisto
- [6]. Tietosuoja laki 1050/2018: <https://www.finlex.fi/fi/laki/alkup/2018/20181050>
- [7]. Tietosuoja-asetus 2016/679: <https://eur-lex.europa.eu/legal-content/FI/TXT/?uri=CELEX%3A32016R0679>
- [8]. Hallituksen esitys eduskunnalle laiksi sosiaali- ja terveystietojen toissijaisesta käytöstä sekä eräiksi siihen liittyviksi laeiksi HE 159/2017 vp: https://www.eduskunta.fi/FI/vaski/HallituksenEsitys/Documents/HE_159+2017.pdf
- [9]. Tietosuojavaltuutetun toimisto, tietosuojaperiaatteet: <https://tietosuoja.fi/tietosuojaperiaatteet>

Liitteet

Liite 1: Lista Kainuun osallistujista

Liite 2: Kuvat käsitteellisestä tuotannosta, Yleiset käsitteet, Kustannuslaskenta

LIITE 1: Projektiin osallistuneet Kainuun soten henkilöt**Osallistuja**

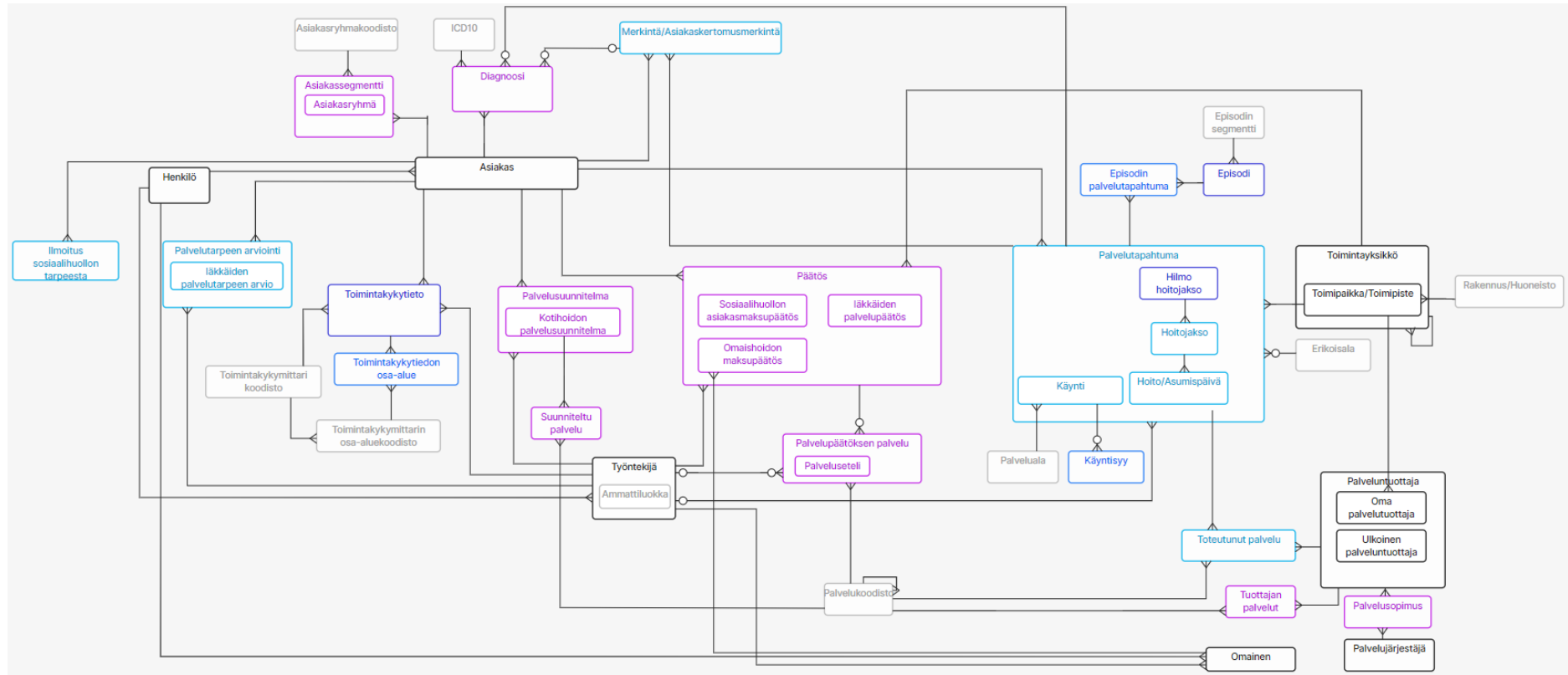
1. Maire Ahopelto
2. Aija-Sisko Hakkarainen
3. Markku Hankio
4. Kaisu Heikkinen
5. Marja Heikkinen
6. Leila Heikkinen
7. Ritva Heinistö
8. Margit Honkila
9. Marjut Hovinen
10. Jaana Hyvärinen
11. Pirkko Hyvönen
12. Tuomo Härkönen
13. Hilkka Karivuo
14. Esa Karjalainen
15. Sari Kela
16. Rauno Ketomäki
17. Niina Komulainen
18. Marja-Leena Lukkari
19. Sirkku Malinen
20. Merja Moisanen
21. Jaana Mäklin
22. Mari Niskanen
23. Marita Pikkarainen
24. Taneli Pikkarainen
25. Sanna Pyykkö
26. Eija Rämä
27. Teemu Röpelin
28. Seija Salomaa
29. Kimmo Syvävirta
30. Kristiina Taskinen
31. Eija Tolonen
32. Ville Turunen
33. Jari Väisänen
34. Johanna Heikkinen

Projektiryhmä / Osa-alue

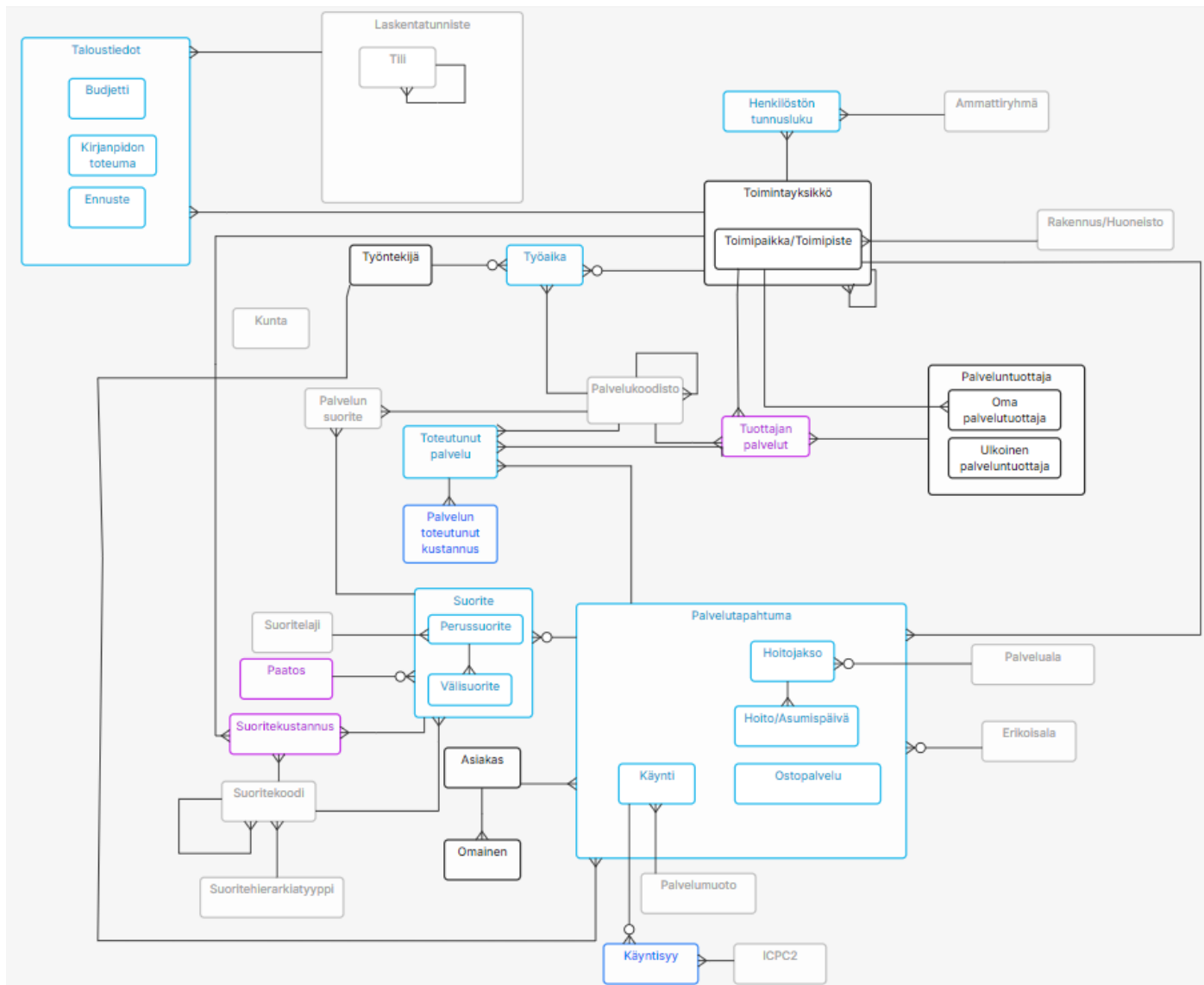
Loppukäyttäjä ja johto
ICT-toteutus ja tietolähteet
ICT-alustaryhmä
ICT-toteutus ja tietolähteet, kustannuslaskenta
Toiminta ja tietosisältö, kustannuslaskenta
ICT-toteutus ja tietolähteet
Toiminta ja tietosisältö, kustannuslaskenta
ICT-toteutus ja tietolähteet
Toiminta ja tietosisältö
Kustannuslaskenta
Toiminta ja tietosisältö, kustannuslaskenta
ICT-alustaryhmä
Projektinhallinta, tietosuoja
ICT-alustaryhmä
Kustannuslaskenta
ICT-toteutus ja tietolähteet
Projektinhallinta, kustannuslaskenta
ICT-toteutus ja tietolähteet
ICT-toteutus ja tietolähteet
Toiminta ja tietosisältö, kustannuslaskenta
Loppukäyttäjä ja johto
ICT-toteutus ja tietolähteet
Loppukäyttäjä ja johto
Toiminta ja tietosisältö, kustannuslaskenta
Kustannuslaskenta
Toiminta ja tietosisältö, kustannuslaskenta
ICT-alustaryhmä
Toiminta ja tietosisältö
ICT-alustaryhmä
Projektinhallinta, kustannuslaskenta
Loppukäyttäjä ja johto
ICT-alustaryhmä, ICT-toteutus ja tietolähteet
ICT-alustaryhmä
Kustannuslaskenta

LIITE 2: Käsitteellit

Tuotanto -käsitteellit



Kustannuslaskenta -käsitelmä



Yleiset käsitteet -käsittemalli

