

# Perustietovarantojen laadunvarmistuksen nykykäytännöt

Haastattelujen yhteenveto joulukuu 2020

Tiedon laatukehikko

# Haastattelujen toteutus

Suunnittelussa ja toteutuksessa  
mukana MML, Vero ja TK

# Taustaa

- Nykytilaselvityksessä päädyttiin haastatteluun aiheen moninaisuuden, tietovarantojen ennakoitun kirjavuuden ja erilaisen kielen tähden. TK ohjausryhmä myös kannatti tätä selvitysmuotoa.
- Haastattelussa päätettiin edetä osallistujaorganisaatioiden haastatteluun. Haluttaessa haastatteluja voidaan myöhemmin laajentaa muihin perustietovarantoihin tai tarpeen mukaan selvittää jotain yksityiskohtaisempaa asiaa kyselyillä
- Selvityksen tavoitteena oli luoda kuvaa tietovarantojen laadunvarmistuksen nykytilasta ja tunnistaa mahdollisesti benchmarkattavia käytänteitä ja malleja

# Haastattelun toteutus

- Haastateltavina tässä vaiheessa Tiedon laatukehikon osallistujaorganisaatiot
- Haastateltavat valitsivat itse käsiteltävän tietovarannon / tietovarannot, tietovarannot eivät kata kaikkia tai välttämättä edes olennaisia osia kyseisten organisaatioiden tietovarannoista
- Haastattelu käytiin sivun 6 kehän ympärillä
- Huomioita haastattelujen yhteismitallisuudesta
  - Tämän yhteenvedon tieto perustuu haastatteluun, muuta materiaalia ei pyydetty, jotain materiaalia saatiin haastattelun jälkeen
  - Tarkastellut tietovarannot vaihtelivat suuresti monessakin mielessä (slide 5)
  - Tietovarannon tarkastelu laadun tulokulmasta ei pääpiirteissään ole kovin tuttua; kysytyjä asioita tehdään ”toisella nimellä” ja kaikki olennainen ei välttämättä tullut esille
  - Organisaatioiden käyttämä terminologia eroaa
  - Haastattelijat ja erityisesti kirjaajat vaihtuivat
  - Haastateltavilla ei välttämättä aina ollut itsellään kaikkea tietoa
    - Tavoitteena ei siis ollut tieteellinen tai tilastollinen tutkimus vaan projektin työtä tukeva selvitys
- Haastattelussa ilmenee tietovarannon ylläpitäjän näkökulma, haastattelussa ei siis haastateltu tietovarannon käyttäjiä, joilta olisi voinut nousta erilaisia asioita esille

# Haastateltujen tietovarannot vaihtelevat suuresti mm. seuraavien tekijöiden suhteen

- Lainsäädännön mahdollisuudet ja rajoitteet tietovarannon tiedon käsittelyssä, kansainvälisen lainsäädännön rooli tietojen määrittelyssä
- Tietovarannon koko ja elinkaaren vaihe
- Tiedon käyttötarkoitukset ja tiedon julkisuus (osa puhtaasi ei-julkisia, osa kokonaan julkisia)
- Minkälaisia tietoa keräävät järjestelmät ovat (keskitettyjä – hajautettuja) ja millainen yhteistyömahdollisuus on niiden suuntaan

Käyttäjien tunnistaminen/Asiakastarpeet/Priorisointi

Lainsäädäntö  
Omistaja

Tietojen määrittely

Laatuvaatimukset  
Laatumittarit  
Laatusäännöt  
Yhteentoimivuus –  
vain kerran periaate

Käytettävyyden arviointi  
Kustannus/hyötyanalyysi  
Yhteentoimivuus

Tiedon  
Käytettävyys

(Ensikäyttö organisaation sisällä  
ja erilainen toisiokäyttö)



Tiedon keruu ja käsittely / laadun varmistus

Kyvykkyyksmallit  
Laadun  
perusparantaminen

Laatutulokset/Laatulupaukset  
Laatudashboardit

Metatiedot/  
Laatutiedot

Laadun testaus

Näyteotanta/täydellinen testaus  
Automaattinen laadun eheyden  
testaus

Tietopalvelu/Tuotteistus

(Ensikäyttö organisaation sisällä  
ja erilainen toisiokäyttö)

Automaattinen laadun eheyden  
testaus

# Tarkastellut tietovarannot

- EU-sisäkauppa Intrastat (Tulli)
- Henkilötietovaranto (TK)
- Kansallinen maastotietokanta (MML)
- Kanta tietoaalusta (Kela)
- Kaupparekisteri (PRH)
- Koski (OPH)
- Kuntatalous (Valtiokonttori)
- Tulorekisteri (Vero)
- Väestötietojärjestelmän tietovaranto (DVV)
- Yritystietovaranto / YTY (TK)

# Havainnot haastattelusta

Havainnot nostettu haastattelukehän  
mukaisissa osissa



# Asiakastarpeet/Lainsäädäntö /Tietovarantovastuu

- Perustietovarantojen asiakastarpeet/käyttötapaukset perustuvat pääsääntöisesti lainsäädäntöön (kansainväliseen ja/tai kansalliseen)
  - Asiakastarpeiden kartoituksia tehdään ilmeisesti harvakseltaan, usein systemaattisesti uuden järjestelmän kehittämisen yhteydessä. Toisiokäytön tarpeiden miettiminen tuli esille vain joissain tapauksissa. Jatkokehitystarpeita ei välttämättä systemaattisesti kerätä (muuttaisi lainsäädäntöä)
- Osassa tietovarannoissa julkinen luotettavuusolettama (esim. Kaupparekisteri, Väestötiedot). Oletetaan, että tieto on julkisesti luotettavaa ja virheistä voidaan joutua jopa maksamaan korvauksia -> Pitäisikö tästä tehdä laatuvaatimus?
- Osassa tietovarannoissa ei sallita virheen korjausta rekisterinpitäjän toimesta vaan virheen voi ainoastaan korjata tiedon tuottaja, osassa hyvinkin vahva kiistämättömyysvaatimus esim. Vero, eli tiedon tulee olla alkuperäinen tieto, tai muutostiedot tallella. Lainsäädäntö kieltää rekisterinpitäjää korjaamasta suoraan tietoa..
- Osassa tietovarannossa selkeä tiedon vastuu selkeyttää myös tiedon laatuvaatimusta esim. Tulli, MML, osassa ei ole vastuuta tiedon sisällöstä esim. VTJ/Rakennukset ja KELA/Potilastiedot
- Lainsäädäntö jäykistää laadun ja tietosisällön kehittämistä asiakastarpeiden mukaiseksi ja joskus jopa estää sen. Joissain tilanteissa (uusi) lainsäädäntö on myös mahdollistaja, esimerkiksi kuntatalouden taksonomat.
  - Lainsäädäntöä usein tarvitaan velvoittamaan tiedon lähteitä (kunnat, yritykset, henkilöt jne) luovuttamaan tietoja, mutta toisaalta lainsäädännön muuttaminen uusien tarpeiden ilmetessä ei ole helppoa
- Perusrekisterien lainsäädäntö ei velvoita perusrekisterien tietojen käyttöön master datana ja lisäksi toisaalta muu lainsäädäntö saattaa jopa kieltää perusrekisterien käytön tiedon laadun parantajana

# Tiedonmäärittely

- Haastatelluissa organisaatioissa ei ole tunnistettu/sanoitettu tiedon määrittelyssä tiedon määrittelyn ja laadun roolia
- Yhteentoimivuus monesti huomioitu sanastotyössä
- Selkeiden tietomäärittelyjen tekeminen parantaa tiedon laatua, kuten yhteisesti sovitut taksonomiat/koodistot/luokitus
- Pääsääntöisesti ei ole määritelty laatuvaatimuksia (vrt. edellä julkinen luotettavuuskohta)
  - Laatuvaatimuksia saattaa kuitenkin olla ministeriöiden tulossopimuksessa (esim. MML ja PRH), yleensä ne koskevat täydellisyyttä
  - MML:ssä määritelty laatuvaatimukset myös tarkkuudelle (sijainti), täydellisyydelle kohteittain ja ominaisuuksille
- Laatusääntöjä (tarkistussääntöjä) määritelty tiedontuotantojärjestelmiin

# Tiedon keruu ja käsittely ja laadunvarmistus: osa 1/2

- Tiedon laatusääntöjen avulla pyritään varmistamaan rakenteisen tiedon eheys,
  - Tiedon keruussa vanhat ja uudet lähdejärjestelmät rinnakkain, ja tämä aiheuttaa haasteita
  - osalla ongelmana PDF-tyyppiset tiedot (esim PRH tilinpäätöstiedot ja kuolemansyytodistukset osittain), joiden laatua ei voida varmistaa
- Virheen korjaus ei aina ole mahdollista rekisterinpitäjän toimesta
  - Saattaa seurata pyyntö originaalitiedon toimittajalle korjata tieto, pyyntä saattaa lähteä varannon ylläpitäjältä tai tiedon hyödyntäjiltä, esim Jiralla tiketointijärjestelmällä voidaan seurata yhteisesti tietyn tietovarannon korjauspyyntäjä (koordinointi järjestelmän kautta)
  - On erityyppisiä rekisterejä, operatiivinen rekisteri voi laittaa korjauspyynnön, kokoavilla tietovarannoilla tämä ei ole yhtä suoraviivastai tai mahdollistakaan
  - Joskus tiedon toimittamatta jättämisestä seuraa (yritykselle) rangaistus/sanktio (tullaustiedot), mutta aina sekään ei riitä motivoimaan toimittamaan tietoa, sakan suuruus vaivaan nähden ei ehkä riittävän iso
  - Tiedon vastaanotossa automaattitarkastukset TK
- Laadunvarmistuksen automatisointia erityisesti tiedon keruussa
  - Laatusäännöt/tarkistussäännöt määrittävät eteenpäinmenon
  - Pääsääntöisesti ei ilmeisesti seurata / raportoida ongelmien laajuutta
- Järjestelmähaasteet tiedon keruun päässä
  - järjestelmiä paljon ja myö essm kuntia paljon, myös järjestelmätoimittajia on paljon aiheuttaa haasteita tiedon laadulle
  - Mikä järjestelmä on originaali/ masterdatajärjestelmä
  - Esimerkiksi sijaintitietojen ja potilastietojen osalla haasteita
- Reaaliaikainen tiedon keruu,
  - Tarve siirtyä tiedon keruun helpottumisen takia
  - mahdollistuu automatisoinnin ja rajapintojen kautta
  - helpompaa myös tiedonluovuttajalle

# Tiedon keruu ja käsittely ja laadunvarmistus: osa 2/2

- Kattavuus ja puuttuvan tiedon korvaaminen
  - Muutama viimeinen prosentti aiheuttaa suurimman työn tiedon keruussa. Tavoitetason asettaminen (tietty %) voisi olla hyvä. Järkevä tavoitetaso riippuu tietoaineistosta.
  - Estimointiin tilastolliset laatu-arvot, laadun kuvaus näiden tietojen osalta (tilastojen puolella näitä on käytössä)
  - perusrekisterien puolella puuttuvan tiedon korjaamista ei pääsääntöisesti esiinny, kattavuuteen pyritään pureutumaan esimerkiksi tiedontuotantoprosessin parannuksella (esim. nopeampi tiedon keruu)
  - Kattavuutta seurattu myös ns. laatudashboardilla
- Käyttöönottojen yhteydessä tehtävät koulutukset parantavat tiedon laatua (esimerkiksi Varda, Koski, Kuntatalous), klinikat (esim. Koski), chat-tuki tietoja täyttäessä (esim. Tulorekisteri)
- Tekoälyn käyttö: chatbot yhteydenoton alkuvaiheessa
- Laadunparannusohjelmia joissain käynnissä (esim. MML, DVV), joissain toiminta satunnaisempaa
- Valvonta/Laatuprosesseja joissain organisaatioissa (esim. Tulli)
- Masterdatan käyttö
  - Pitäisikö pystyä hyödyntämään? Masterdataa ei ole määritelty tai sen käyttö ei ole mahdollista lainsäädännön rajoitusten takia

# Laadun testaus

- Yleisesti ei ole tunnustettu, että voitaisiin tehdä erikseen laadun arviointia
  - Pääsääntöisesti ei tehdä erillistä aika ajoin tehtävää laadun testausta
  - Esimerkiksi vertailu toiseen aineistoon, on tehty esimerkiksi Veron ja DVV:n tietojen vertailua TK tietoihin
  - Mistä impulssi laadun testaukseen?
  - Mitä laadun testauksesta voi seurata?
  - Kallista, tehokkaampaa tehdä näyteotannalla
  - Vaikea ehkä keksiä menetelmää, kun tiedonlähde ilmoitus, mihin voi verrata
  - Tilastollisia menetelmiä voisi hyödyntää enemmän operatiivisissa kannoissa (poikkeavat havainnot, loogisuus suhteessa toiseen muuttujaan)
- Joitain esimerkkejä löytyy
  - MML tekee säännöllisesti sisäisesti ja ulkoisesti KMTK:n osalta -> tuloksena tietoa laatuvaatimusten saavuttamisesta
  - Esim. laatukontrollitiedustelu kohdennettuna pieniin yriyksiin TK YritystietovarantoYTY

# Palvelut ja tuotteet

- Palvelujen ja tuotteiden laatu pääsääntöisesti johdettavassa tietovarannon laadusta
- Testausprosesseja on käytössä tietojärjestelmien ja rajapintojen toimivuudesta, testiaineistoja myös käytössä
- Palvelujen laatumittareita ei yleisesti käytössä esim. Palvelun saatavuus
- Automaattitestausta, esim. Veron tietopalvelussa erityisesti eheyttä, vertaillaan aiempaan
- Tiketointijärjestelmä käytössä asiakaspalautteen hallinnassa auttaa hallitsemaan tiedon laadun korjauksia, esim. Kuntatalous
- Muutosten vienti ja hallinta voi olla haasteena. Asiakasjärjestelmissä tieto voi olla osin vanhentunutta verrattuna taustatietovarantoon. (laatuvirheet, ns turhat kyselyt, prosessiongelmat, käyttäjien haasteet)
- Masterdata-ajattelu tässäkin voisi auttaa
- Palvelujen maksullisuus voi olla haaste käyttäjän päässä esimerkiksi tiedon päivittymisen suhteen
- Pääsääntöisesti ei 24/7 palvelua tarjolla (kopiot)

# Metatieto ja tiedon laatu

- Laatutietoja ei yleisesti saatavilla eikä jaeta käyttäjille metatietojen yhteydessä
- Metatietoja voi olla organisaation sisällä, ei keskitetysti julkaistu eikä niihin pääsääntöisesti aseteta vaatimuksia
  - Avoin data.fi julkaisussa oletettavasti oletusarvoisesti metatiedot, avoin data.fi liittyen ei ole keskitettyä valtion metatietojärjestelmää tai vaateita niiden sisällölle?
  - MML julkaisee paikkatietojen osalta metatiedot, mutta ei laatutietoja (perustuu INSPIRE direktiivin vaatimukseen)
- Laatulupauksia ei yleensä ole tehty tietovarantotasolla, esim. SVT tuottajilla laatulupaukset tietotuotteen tasolla?
- Laadun visualisointeja (Laatu dashboardeja tai tuotekortteja) ei yleensä ole, Kuntataloudella laadun dashboard

# Käytettävyys

- Useimmat organisaatiot jotka ovat vastuussa tietovarannosta keskustelevat sidosryhmien kanssa ja tekevät asiakastytyväisyystutkimuksia
  - Paikoin sidosryhmien vaikutus tietotarpeisiin ja tietojärjestelmän kehityksen ohjaukseen, esim. Tulorekisteri ja Kuntatalous
  - Tiedon laatu ja vaikutus käytettävyyden paranemiseen ei pääsääntöisesti selkeästi esillä asiakastytyväisyystutkimuksissa tai niiden jatkotoimissa
  - Asiakas löytää virheen – konkreettinen palaute tiedon korjaukseen, mahdollisesti isompia toimenpiteitä, vastuu virheen korjaamisesta voi olla asiakkaalla (oman tietonsa osalta)
- Lainsäädäntö ja resurssit estävät joskus reagoimasta asiakastarpeisiin
- Käyttö ja tiedon avaaminen ovat lisänneet tiedon laatua
- Avoin data ei tuota tuloja: hyvä palvelu ja laadunparannus ei ehkä siksi ole tiedontuottajan prioriteeteissa. Toisaalta tiedon avaaminen voi parantaa laatua palautteen kautta.
- Osassa tietovarantoja ei ole avointa dataa tai se on hyvin vähäistä
  - Tietovarantojen tietosisältö vaikuttaa myös tähän
  - Tilastotietoa tietoineiston käytön asteesta voisi tarjota avoimen datan palvelusta



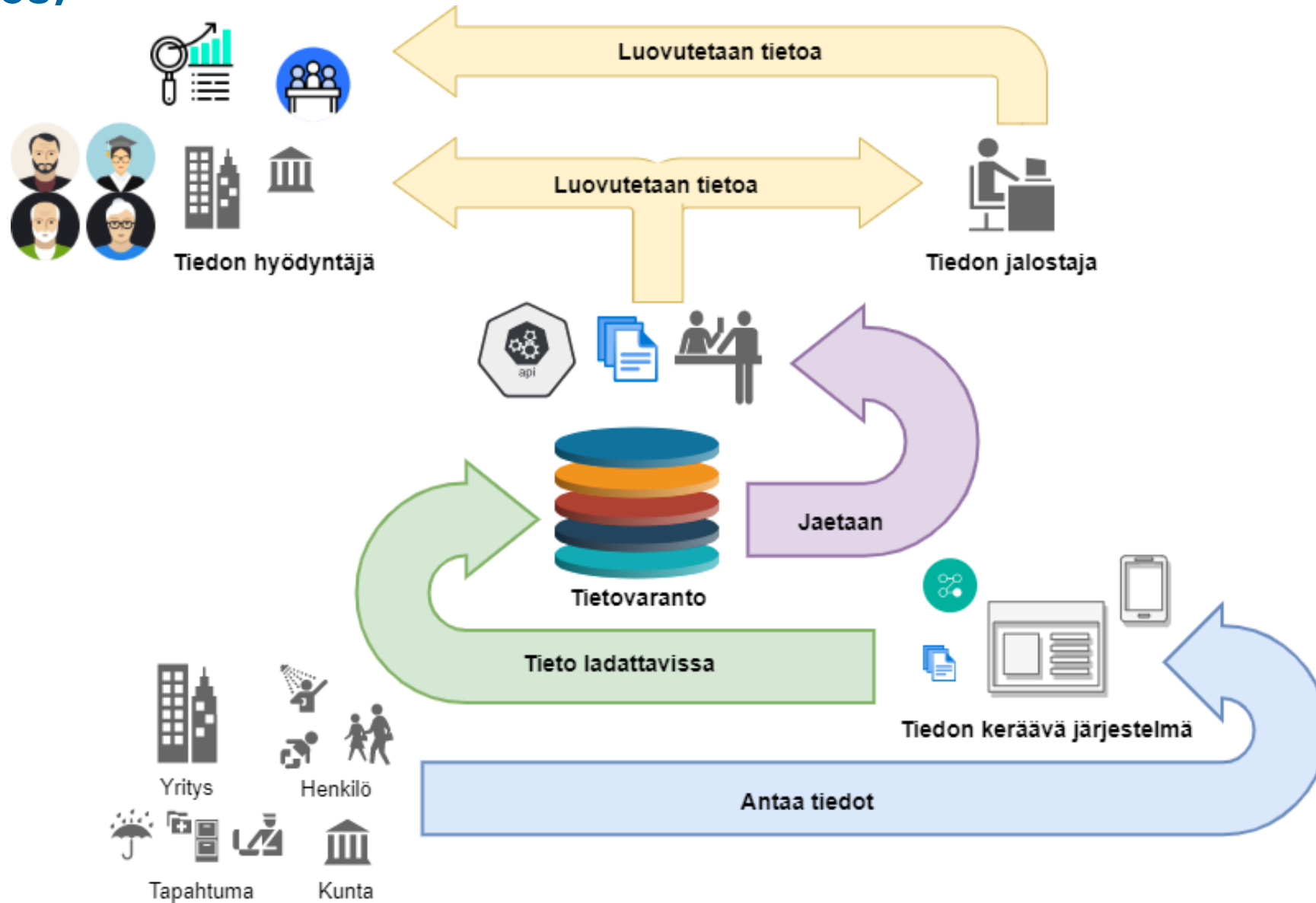
# Muita havaintoja

- Tiedon keräämisen vaiheessa ei ole velvoitetta ottaa muuta kuin primaarikäyttöä huomioon, se saatetaan jopa kieltää lailla, eli tietoa saa kerätä vain tiettyyn tarkoitukseen. Toisiokäyttöä ei pääpiirteissään huomioida laajasti.
- My datan hyödyntäminen tiedon laadun parantajana:
  - Esim. Koskidatassa opiskelija pystyy jakamaan oman tietonsa eteenpäin, linkkinä tai tieto heille joille valitsee
  - Kannassa voi rajoittaa oman tiedon käyttöä

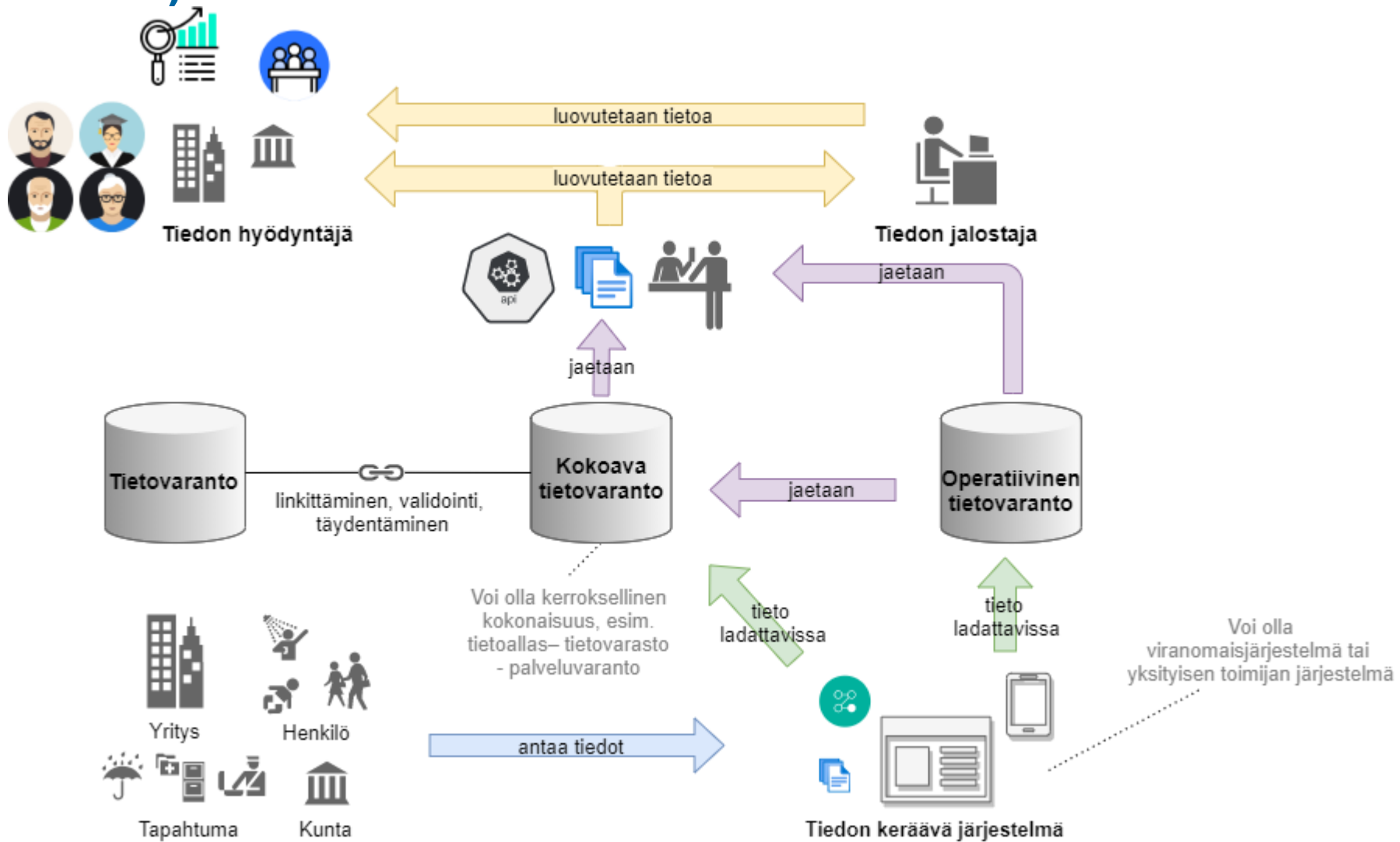
# Julkisen hallinnon tietovirrat tiedon laadun tulokulmasta: nykytilahaastatteluissa tunnistettuja kipupisteitä

Nämä ovat pääosin asioita joihin tiedon  
laatukehikko ei voi tai ei voi ainakaan  
yksin vastata...

# Julkisen hallinnon tietovirrat tiedon laadun näkökulmasta, yleiskuva (luonnos)



# Julkisen hallinnon tietovirrat tiedon laadun näkökulmasta, tarkennettu kuva (luonnos)



# Tieto on pääpiirteissään laadukkaampaa tilanteissa joissa:

vrt. edellinen kuva

- Tietovarannolla on vahva omistajuus / vastuutaho tiedon ja tietoarkkitehtuurin tulokulmasta
- Operatiivinen järjestelmä on kokoavan Tietovarannon vaikutuspiirissä
  - Esimerkiksi reseptikeskus on Kannan vaikutuspiirissä toisin kuin potilastietojärjestelmät
  - Esimerkiksi kuntien rakennustiedot eivät ole VTJ:n vaikutuspiirissä
- Tietyn tiedon keräämisen järjestelmäkirjo on suppea
  - Esimerkiksi reseptikeskuksessa kerätään kaikki reseptitietoja, tilinpitojärjestelmiä on paljon ja eri IT-toimittajien
- Tietovaranto voi hakea master-tiedot toisesta Tietovarannosta
  - Esimerkiksi osoitetieto ja muut sijaintitiedot voisi noutaa toisesta tietovarannosta laadukkaampina – lainsäädäntö ei tätä kaikissa tilanteissa salli
- Tiedon kohde / tiedonantaja (esim. yritys/kunta/muu organisaatio/henkilö) on myös tiedon hyödyntäjä
  - Esimerkiksi oppilaitokset ja opiskelijat Koski-tietokannassa, esimerkiksi HSL voi käyttää suoraan Koskien tietoja opiskelijan luvalla, jolloin opiskelija saa opiskelualennukset
  - Esimerkiksi henkilö täydentää omia Kantatietojaan
- Tietovarannon tietoja käytetään toisessa Tietovarannossa tai toisessa palvelussa
  - Esimerkiksi ajoneuvorekisterin tietojen täyttäjät saa PRH:n tietoja näkyville ja haluaa tiedot ajoneuvorekisterissä nähdessään korjata alkuperäisen tietovarannon tietoja
- Tietovarannolla on ylipäätään useita ja erilaisia Tiedon hyödyntäjiä ja Tiedon jalostajia
  - Esimerkiksi Tilastokeskuksen Tutkijapalvelu, Kuntatalouden tietopalvelu, Tulorekisteri
- Jos pystytään resursoimaan ja rahoittamaan palvelun ja tiedon laadun kehittymistä
  - Esimerkiksi laadunkehityshankkeet / perusparannukset; ei ratko kuitenkaan jatkuvaa parantamista, jonka tulisi olla sisäänrakennettu

# Nykytilaselvityksen pohjalta hyödynnettäviä ja kehitettäviä ajatuksia

# Nostoja tiedon laatua tukevista hyvistä käytänteistä

## ❖ Lainsäädäntö tiedon laadun tukena

- Julkinen luotettavuus perusrekistereissä käytössä oleva oletama, virhevastuu perusrekisterin ylläpitäjällä (Perusrekisterit)
- Taksonomioista / koodistoista lailla säättäminen: Kuntatalouden tietopalvelu
- Sanktiot määritelty laissa (EU-Intrastat tiedonkeruu)

## ❖ Laatudashboard tukemassa läpinäkyvyyttä

- Kartta kertoo, mitkä kunnat ovat jo antaneet tiedot (Kuntatalouden tietopalvelu)

## ❖ Tiketointijärjestelmän hyödyntäminen palautteen hallinnassa

- Tiedonantajille annettavan palautteen käsittelyn organisointi ja läpinäkyvyys (Jira / Kuntatalouden tietopalvelu)

## ❖ Tuki tiedonantajille / tietoa keräävän järjestelmän käyttäjille

- Koulutus ja tuki (Kuntatalouden tietopalvelu)
- Säännölliset käyttäjäinfot (Koski)

## ❖ Ohjeistuksen generointi automaattisesti koodista (Koski)

- Aina kun ohjelmistoon tulee muutoksia, generoituvat muutokset automaattisesti ohjeisiin (Koski)

## ❖ Masterdatan käyttö

- Masterdatan käytön lisääntyminen myös parantaa masterdatan laatua, esimerkiksi Kaupparekisterin yritysten osoitetietojen käyttö Ajoneuvorekisterissä

## ❖ Tietojen määrittely: sanastot ja koodistot/luokitukset

- Määrittely tarpeen mukaan yhteentoimivasti seuraavilla tasoilla tietovarantotaso, kansallinen taso, kansainvälinen toimialataso (Kanta, Kuntatalous, Tilastot, EU Sisäkauppa)

## ❖ Laatumetatietojen julkaiseminen

- Sijaintitarkkuus tallennetaan kiinteistörajapyykille talletettava laatumetatieto (Kiinteistörekisteri)

## ❖ Automatisoidut tarkistukset

- Laatu/ tarkistussäännöt käytössä tiedon vastaanotossa (Tulorekisteri, Kuntatalous, TK aineistojen vastaanotto, EU-Intrastat Tulli)
- Formalisoitujen laatusääntöjen käyttö (KMTK)
- Skeemojen hyödyntäminen: validointi ja mäppäys toiseen skeemaan (edellä mainitut)

# Havainnot ja tiedon laatukriteereihin liittyen

- Perustietovarantojen käytössä olevat tiedon laatukriteerit vaihtelevat
  - Tiedon laatukriteerejä on perustietovarannoissa jonkin verran käytössä
  - Terminologia ja käytettyjen laatukriteerien sisältö eroaa ainakin osin
  - Laatukriteerien kypsyys ja mahdollinen kansainvälinen tausta vaihtelee
- Oikeellisuus, Kattavuus ja Ajantasaisuus joissain tietovarannoissa käytössä
  - Onko sisältö kaikissa varannoissa sama?
  - Oikeellisuus korostuu perusrekistereissä
  - Ajantasaisuudessa monta eri puolta, useita eri mittareita
- Tarkkuus
  - Esimerkiksi sijaintitarkkuus
- Tiedon tekninen validointi pitkällä uusissa tietovarannoissa
  - benchmarkattavaa esimerkiksi Koneluettavuuden / Siirrettävyyden mittaamisessa?
    - Käytetään myös termiä looginen eheys, toisaalta tietojen välinen eheys on myös looginen eheys)
- Yhteentoimivat sanastot ja koodistot/taksonomiat useassa tietovarannossa käytössä tai työn alla
  - Ymmärrettävyys / analysoitavuus ja Vertailukelpoisuus (osa semanttista yhteentoimivuutta)
- Sisällön saatavuus, voi olla palvelun laatukriteeri?
- Miten tieto kuvaa kohdettaan, onko tämä esillä muutoin kuin oikeellisuuden kautta?



# Esiin nousevat ajatukset laajemmin Tiedon laatukehikkoa koskien

- Laatukehikko voisi toimia organisaation laatutoiminnan kirittäjänä esim. Kyvykyysmallin / Kyvykkyystestin kauttakautta
- Tiedon laadun (ja tietopalvelun laadun) viitearkkitehtuuri esim. EIRAA hyödyntäen
- Tiedon johtamisen ja tiedon laadun yhdistäminen?
- Laadunvarmistuksen prosessit ylätasolla, sisältäen ylätason tarkistusksymyksiä?
- Kaikki edellä olevat yhdistäen yhteen kehikkoon?
- Perustietovarantojen kokonaisuuden kuvaus, esimerkiksi dataekosysteemin laatu-ulottuvuudet, ml. lainsäädäntö

# Tiedon laatuun merkittävästi vaikuttavat tekijät, joihin vaikutetaan tiedon laatukehikon ulkopuolella

- Paikoin vanhentunut ja moninainen lainsäädäntö heikentämässä tiedon laatua
  - Kysytään vain kerran ja my data ajattelun tuki ei ole lainsäädännössä riittävää
  - Erityislainsäädännön yksityiskohdat haaste, voisiko olla yleisempää lainsäädäntöä tiedon hyödyntämisestä?
  - Tietoturvan osalta jako: julkinen ja salainen tieto ← tarvittaisiinko uusia kategorioita, esimerkiksi ”vain viranomaiskäyttöön”
  - Yhteinen sanastotyö lainsäädännön taustalle
- Masterdatan hyödyntäminen perustietovarantojen kesken alkuvaiheessa
  - Lainsäädäntö estää paikoin masterdatan hyödyntämisen
  - Perusrekisterien käyttövelvoite?
  - Tieto kerätään vain kerran ja ylläpidetään vain kerran ajatusten tukeminen, korjaustieto voisi olla kaikkien käytettävissä
  - Vaikuttaisi tiedon käsittelyyn ja käyttöön tiedon laatua parantavasti
- Julkisen hallinnon kokonaisuuden tasolla hyödyn arviointi puuttuu
  - Tiedon laaja käyttö (myös muu kuin primaarikäyttö) olisi hyvä huomioida jo tiedon määrittelyvaiheessa
  - Miten arvioidaan hyötyjen konkretisoituminen myös muille, ei vain tietyille tietovarannolle tai kehittämiskohteelle
  - Missä vaiheessa laadun parannus kannattaa kustannusnäkökulmasta tehdä tiedon virrassa?
  - Mikä on intensiivi parantaa tiedon laatua ekosysteemitulokulmasta ja mihin asti laatua parannetaan (viimeiset parannukset maksavat eniten)
- Koko tietoekosysteemi tiedon alkulähteillä tulisi huomioida tietovarantoja ja tiedon laatua kehitettäessä
  - Kunta – valtio välillä mm tietojärjestelmien ja lainsäädännön kirjavuus aiheuttaa haasteita tiedon laadulle
  - Ylipäättään vaikea vaikuttaa esimerkiksi lukuisiin tietoa kerääviin järjestelmiin ja it-toimittajiin
- My data -asian pohdintaa myös tiedon laadun tulokulmasta
  - Voi lisätä myös luottamusta viranomaiseen toimintaan oman vaikutuksen kautta, vaikutus myös tiedon laatuun
  - Suostumuksen antaminen oman tiedon käyttöön, vaikutukset tietojen hyödyntämiseen laajemmin
- Laadun parannus edellyttää resurssien kohdentamista, voi vapauttaa resursseja jostain muualta

# Johtopäätökset selvityksen osalta

- Selvityksestä oli hyötyä sekä projektin kehityskohteisiin että laajemman kuvan ymmärtämiseen
- Tietovarantojen tyypittelystä (esim. kokoava/operatiivinen) ja laadun tarkastelusta erikseen näiden kohdalla voisi olla hyötyä
- Hyviä käytänteitä voisi jalostaa laajemmin hyödynnettäväksi / jaettavaksi. Muutoinkin selvityksestä on irrotettavissa useita hyödynnettäviä asioita jatkojalostettavaksi, kuten tiedon laadun ”iskulauseet”.
- Selvityksen antia voisi mahdollisesti esitellä jossakin alan seminaarissa, esim. ICA tai SEMIC?