

Tiedon laatukehikko -hankkeen Laajan projektiryhmän (=TIHA TP3 hankeryhmän) tapaamismuistio

Aika: Perjantaina 16.10.2020 klo 8.-9.30
Paikka: Teams
Läsnä: Vesa Ylönen (KELA), Antti Jakobsson (MML), Jaana Lehtilä (Vero), Erja Nokkanen (OPH), Kukka-Maaria Berg (OPH), Riitta Alkula (DVV), Sini Slotte (DVV), Olli Ahonen (Valtiokonttori) Susanna Summa (Tulli), Outi Ahti-Miettinen (TK), Jarmo Ranki (TK), Jari Kosonen (TK), Noora Seppä (TK), Jari Nieminen (TK) Essi kaukonen (TK)

1. Tapaamisen avaaminen

Tapaaminen aloitettiin 8.00 tiedon laatuun orientoitumalla ja osallistujaesittelyillä.

2. Tiedon laatu -taustaa

Tilastokeskuksen laatupäällikkö Outi Ahti-Miettinen kertoi tiedon laatukehikon taustoista, tietoaineistojen laatuun kiinteästi liittyvistä asioista sekä tietoaineistojen laadunhallinnan tekijöistä.

3. Organisoituminen ja työtavat

Kokouksen vetäjä Essi Kaukonen kiitti tapaamista edeltäneeseen organisoitumista ja työtapoja koskevaan kyselyyn vastaajia. Vastauksissa saatiin arvokasta tietoa työtapojen kehittämiseksi sekä syksyn tapaamisten sisällöstä. Käytiin läpi seuraavien tapaamisten alustavia aiheita, joita oli jo päivitetty kyselyn tuloksilla.

Tiedon laatukehikon laaja projektiryhmä (=hankeryhmä) kokoontuu työpajanomaisissa vuorovaikutteisissa tapaamisissa työstämään Tiedon laatukehikkoa syyskauden 2020 aikana. Seuraavat tapaamisajat ovat 5.11.2020, 18.11.2020 sekä 3.12.2020. Näiden välillä Laaja projektiryhmä voi osallistua Teams-työtilassa tehtävään työhön, reagoida pari kertaa kuussa ilmestyviin laajan projektiryhmän sisäisiin uutiskirjeisiin ja vastata tapaamisia edeltäviin kyselyihin. Aktiivinen ja proaktiivinen osallistuminen on toivottavaa. Laajan projektiryhmän lisäksi halukkaat voivat edelleen lähteä mukaan Nykytilaselvitystä työstämään ns. teemaryhmään. Jaana Lehtilä (Vero) ja Antti Jakobsson (MML) ovat olleet tässä työssä jo ahkerasti mukana. Vuoden 2021 alussa lähtee näillä näkymin liikkeelle Pilotointiin keskittyvä teemaryhmä.

Projektin aikataulua pidettiin kunnianhimoisena, erityisesti syksyn 2020 osalta.

4. Tiedon laatukehikon kohdentuminen

Työstettiin tiedon laatukehikon kohdentumiseen liittyviä asioita ryhmätyönä ja keskusteltiin rajauksista. Tiedon jalostusketjussa nimenomaan dataan keskittyminen oli kannatettavaa. Kun tarkasteltiin erityyppistä dataa, niin todettiin, että työn helpottamiseksi on ensi vaiheen rajauksessa hyvä rajata tarkastelun ulkopuolelle kuvanteet, biologiset

näytteet (kuten solunäytteet) ja havaittavat fysikaaliset ilmiöt kuten ääni, valo ja lämpö.

Kuvat, video- ja äänitallenteet sekä 3D mallit on hyvä ottaa tarkasteluun, mutta niiden osalta voitaneen odottaa ensi vuodelle. Keskustelussa todettiin, että rakenteellinen aineisto on aluksi hyvä rajaus. Seuraavassa vaiheessa muitakin hyvä huomioida ja pohdittiin myös asiakirjojen roolia projektissa.

5. Tiedon laatukehikko ja tietovarannot

Projektipäällikkö Essi Kaukonen kertoi etenemisajatuksista, jonka mukaan aluksi eli vuoden 2020 aikana keskityttäisiin perustietorekistereihin ja perustietovarantoihin. 2021 tarkastelu laajennettaisiin muihin julkisiin tietovarantoihin ja 2022 ylipäätään tietovarantoihin. Projektipäällikkö kertoi, että tämän jaottelun taustalla on käytetty Perustietovarantojen viitearkkitehtuuria ja julkisen hallinnon kokonaisarkkitehtuurityötä.

Etenemisjärjestystä pidettiin hyvänä, joskin siinä nähtiin riskinä mahdollinen vääristymä alkuvaiheen tiedon laatukehikossa. Keskustelussa todettiin, että muut hankkeet, kuten ns. high value data settien avaaminen ja tiedonhallintamalli noudattavat samaa järjestystä. Keskustelussa todettiin myös, että on hyvä lähteä ensin perustietovarannoista, koska monet muut tietovarannot rakentavat tämän päälle ja että perustietovarannoissa on eniten yhteneviä osia, jotka oletettavasti toistuvat muuallakin. Perustietovarannoissa tiedon laatu on tärkeää, jossain muussa tietovarannossa esimerkiksi nopeus voi painottua tiedon (muun) laadun kustannuksella. Todettiin toisaalta, että samaa laatukriteeristöä tulisi pystyä käyttämään erityyppisissä aineistoissa, toisaalta ajateltiin, että muiden tietovarantojen osalta laatukriteereihin voi sitten laajentua erityispiirteitä.

6. Tiedon laadun merkitys avoimessa datassa, viranomaisten välisessä tiedon vaihdossa, käyttöluvan tai sopimukseen perustuvassa tiedonvaihdossa sekä organisaation sisäisessä laadunvarmistuksessa

Tämän aiheen ympärillä käydyssä keskustelussa todettiin muun muassa seuraavaa:

- Avoimessa datassa käyttäjiltä ei saada palautetta tiedon laadusta, sopimuksen varaisessa tietoluovutuksessa tämä nousee keskusteluissa luontevasti esille. Keskusteluissa olisi hyvä tarkastella sekä tiedon käytön että tuotannon puolelta nousevaa palautetta.
- Jokaisessa näissä tiedonvaihdon tilanteissa on tärkeää tietää, mitä tieto sisältää. Pitäisikö tiedon kuvaaminen olla yhtenäisempää?
- Tietosuojaan nämä suhtautuvat eri tavoin, tärkeä erottava tekijä.
- Myös viranomaisten välisessä tiedon vaihdossa tehdään määrämuotoiset sopimukset ja aineistot usein asiakkaalle ”valmiiksi”. Tämä kommentti vertautui sopimuksen ja / tai käyttöluvan varaiseen tiedonvaihtoon.

16.10.2020

- Tutkijoilla tiedon jatkojalostus on tärkeää, uudelleen luokitus on ongelma, samoin riittävät aikasarjat. Metatiedot tärkeitä.
- Tiedon tulisi syntyä kuranttina. Esimerkiksi kuntien talousraportoinnissa tärkeää saada tasalaatuista tietoa.
- Organisaatioiden sisällä liitos prosesseihin ja prosessien laatuun on tärkeää. Se, miten varmistetaan tiedon laatu, on osa toimintaa. Automatiikkaa voidaan rakentaa laadunvarmistukseen osaksi tiedontuotantoprosessia siihen, mihin soveltuu.
- Tärkeää tietää käyttötarkoitukset mahdollisimman hyvin, jotta laadunvarmistuksessa tehdään oikeat priorisoinnit.
- Pohdittiin toisiokäyttö-käsitteen määritelmää.
- Kun tiedon tuottaja ja käyttäjä vain yhdessä organisaatiossa, tiedon laadunhallinta on helpompaa ja motivoivampaa: säästö tulee suoraan itselle.
- Kaikissa tilanteissa tiedon elinkaari tärkeää ja tiedosta huolehtiminen. Huomioi sen, että ei hävitä tietoa, joka on toisella tarpeellista.
- My data on aivan oma asiansa ja osa organisaatioiden tiedonhallintaa. (My data on rajattu TiHA hankkeen ulkopuolelle.)
- Avoimessa datassa käytettävyys on tärkeää; tieto tarvitsee tarjota eri muodoissa. Metatiedot tärkeitä, jotta käyttäjät ymmärtävät voivatko käyttää kyseistä tietoa. Toisaalta pohdittiin, että avoimessa datassa riittää, että datat ovat rajapinnalla ja metatiedot kuvattu.
- Avoimen tiedon pitää olla julkista. Eli jaetaan erilaista tietoa kuin mitä muissa tiedonvaihdossa käsitellään. Kaikki julkinen tieto ei ole julkaistua tietoa.
- Ajantasaisuus voi olla avoimen datan jakelussa riski, jos ei ole kiinteästi yhteyttä tiedon tuotantoon vaan erillinen jakeluvaihe.
- Todettiin, että lainsäädäntö rajoittaa tiedon toimitusta.
- Laeissa käsitteet saattavat olla epämääräisiä, vaatii käytännössä paljon selvittämistä. Onko lainsäädännössä määritelty tiedon laatua? Joissain on määritelty tiedon laatukriteerejä, ei kuitenkaan laajasti.

7. Sidosryhmät

Sidosryhmistä ehdittiin käydä vain lyhyttä keskustelua. Todettiin, että emme voi puhua vain valtion sisäisestä tiedon vaihdosta, sillä silloin kunnat ja sairaalat rajautuvat pois. Viranomaisten välinen tiedon vaihto on parempi termi. Todettiin, että yksityiset toimijat ovat esimerkiksi Vardan yhteydessä huomioitava sidosryhmä. Sopimukseen ja/tai käyttölupaan perustuvassa tiedon vaihdossa Finndata on uusi toimija.

8. Tiedon laatukehikkoon sisältyvät asiat

Projektipäällikkö kertoi, että minimissään hankesuunnitelman mukaan Tiedon laatukehikkoon kuuluu sen hallintamalli, laatukriteerit sekä mittarit. Keskusteltiin tiiviisti tapaamisen lopulla muusta mahdollisesta sisällöstä. Kyvykkyydet oli jo aiemmin tapaamisen chatissa nostettu esille. Todettiin että nykytilaselvityksen näkökulmat ja anti on hyvä huomioida tiedon laatukehikossa. Ehdotettiin yhteisiä tarkistuslistapohjia/malleja. Mietittiin myös työvälinettä, jolla laatu voidaan tarkistaa, sellaisiin tilanteisiin, mihin automatiikka soveltuu. Todettiin, että RPA robotiikka Poc voisi kiinnostaa useita. Nähtiin myös, että Tiedon laatukehikko projektin aikana tehtävät selvitykset ovat myös tärkeää projektin antia, eivät pelkästään projektin lopputuotokset.

Lisäksi todettiin, että tulevien laatukriteerien täsmentäminen ja määrittäminen siten, että kaikki ymmärtävät ne samalla tavalla, on hyvin tärkeää.

9. Nykytilaselvitys ja sen haastattelut

Antti Jakobsson kertoi nykytilaselvityksestä tietovarantojen laadunvarmistuskäytänteisiin liittyen sekä tämän selvityksen etenemisestä. Nykytilaselvityksen haastattelut kohdistuvat tässä vaiheessa perustietovarannoista nimenomaan Laajan projektiryhmän osallistujaorganisaatioihin. Tavoitteena on käydä nämä haastattelut 10.11.2020 mennessä. Näistä ollaan yhteyksissä puolin ja toisin lähiaikoina/ lähipäivinä.

10. Seuraava tapaaminen

Seuraava tapaaminen on 5.11.2020

Jakelu:

Tiedon laatukehikko TiHA TP3 Teams-työtila sekä Tiedon laatukehikko verkkosivusto