

Tiedon laatukehikko -osahankkeen ohjausryhmän kokouksen pöytäkirja

Aika: Maanantai 1.3.2021 klo 9:00-10:50
Paikka: Skype-kokous

Osallistujat: Sami Kivivasara, lainsäädäntöneuvos, yksikön päällikkö, VM puheenjohtaja
Olli-Pekka Rissanen, johtava asiantuntija, VM varapuheenjohtaja

Jäsenet:

Heli Koski, tutkimusjohtaja, Elinkeinoelämän tutkimuslaitos
Juuso Rahkola, Analytiikkayksikön päällikkö, Kansaneläkelaitos
Antti Markkanen, kouluneuvos, Opetushallitus (varajäsen)
Tapio Matinmikko, johtaja, Oulun kaupunki
Helena Laaksonen, Tietoarkiston johtaja, Tampereen yliopisto
Maarit Leinonen, tutkimuspäällikkö, Terveystieteiden ja hyvinvoinnin laitos
Kirsti Pohjanpää, ryhmäpäällikkö, Tilastokeskus
Outi Ahti-Miettinen, kehittämisspäällikkö, Tilastokeskus (varajäsen)
Jani Martikainen, ylitarkastaja, Verohallinto

Pysyvät asiantuntijat:

Essi Kaukonen, Tiedon laatukehikko -projektipäällikkö, Tilastokeskus sihteeri

1. Kokouksen avaus ja esityslistan hyväksyminen

Puheenjohtaja Kivivasara avasi kokouksen 9.00. Asialista hyväksyttiin sellaisenaan.

2. Alustavat tietoaineistojen laatukriteerit

Projektipäällikkö Essi Kaukonen esitteli alustavat tietoaineistojen laatukriteerit, jotka ovat ensimmäinen ja merkittävin osa tiedon laatukehikkoa. Laatukriteerit kohdistuvat rakenteiseen dataan. Ne pohjaavat pitkälti ISO 25012 -standardiin ja ovat seuraavat: oikeellisuus, tarkkuus, ajantasaisuus, johdonmukaisuus, kattavuus, täydellisyys, ymmärrettävyys, suositustenmukaisuus, jäljitettävyys, koneluettavuus ja oikea-aikaisuus.

Kaukonen kertoi, että laatukriteerit olivat avoimesti kommentoitavana helmikuun ajan. Kommenttien määrä oli odotetusti vaisuhko. Yleisesti yhteydenotoissa sekä laatukriteerejä että kommentointitapaa pidettiin hyvinä. Kriteereiden pilotoinnin ja mittareiden kehittämisen jälkeen voisi harkita toista avointa kommentointikierrosta kesän aikana. Näillä näkymin virallinen lausuntokierros on tähtäimessä syksyllä 2021 ja tavoitteena on laatukriteerien julkaisu yhtä aikaa API-linjausten kanssa loppuvuodesta 2021 kansallisena suosituksena. Tiedon laatukehikon kokonaisuus, ml. hallintamalli, kehitty rinnalla tarpeen ja TiHA-tavoitteiden mukaisesti.

Ohjausryhmän keskustelussa nostettiin esille yksittäisiä laatukriteerejä sekä niiden välisiä yhteyksiä koskevia huomioita. Kattavuuden ja täydellisyyden ero mietitytti ja laajemmin erot ISO 25012 -standardiin nähden. Nähtiin, että standardia tarkempaa jaottelua ei kannata lähteä tekemään. Kaukonen kertoi, että pilotoinnissa arvioidaan ja jatkokehitetään myös itse laatukritee-

reja. Yhdisteltävyyttä tiedusteltiin ja projektipäällikkö totesi, että sitä edesauttavia tekijöitä löytyy esimerkiksi koneluettavuuden ja ymmärrettävyyden takaa, vaikka suoraa yhdisteltävyyden kriteeriä sinänsä ei ole kriteerien joukossa.

Todettiin, että tiedon reaaliaikaisuus ja koneluettavuudesta huolehtiminen ajassa on tärkeää. Toisaalta todettiin, että reaaliaikaisuus ja kattavuus ovat toistensa kilpakumppaneita, sillä äärimmäinen ajantasaisuus heikentää muiden kriteerien kautta kuvattua tiedon laatua. Reaaliaikaisuuden tavoittelun suhteen tietoaineistot eroavat toisistaan, rekisterien osalta tapahtumaa kirjatessa on jo itse tapahtumasta saattanut kulua aikaa ja esimerkiksi tilastotuotannossa ei tavoitella reaaliaikaisuutta. Tiedusteltiin myös, miten tietoaineistojen laatukriteerit suhteutuvat tietoturvallisuuden kriteereihin. Projektipäällikkö kertoi, että tietoturva ja tietosuoja olivat kriteerien alkumäärittelyissä mukana, mutta eivät edenneet kriteereiden sisältöön, lähinnä siksi, että ovat asioita, joista on joka tapauksessa huolehdittava hyvin joka tilanteessa.

Laatutavoitteiden asettamisesta käytiin runsasta keskustelua. Ohjausryhmässä todettiin, että kaikkiin tilanteisiin sopivia yleisiä laatuvaatimuksia ei voida asettaa, koska aineistot ovat erilaisia ja myös käyttötarkoitukset ja niiden vaatimukset tiedon laadulle eroavat toisistaan.

Keskusteltiin myös rekisterien, tiedon laadun ja päätöksenteon välisistä yhteyksistä. Keskustelussa todettiin, että rekisterinpitäminen on kallista, mutta kansalliset rekisterit ovat tärkeitä ja niiden laatu on keskeistä, myös siksi, että niiden pojalta tehdään esim. terveydenhuollossa kattavia poliittisia ja terveydellisiä johtopäätöksiä. Todettiin, että automaattinen päätöksenteko edellyttää laadukasta tietoa. Kelassa on pohdittu, että tarkkuuden pitäisi olla lähes 100 %, mikäli automaattista päätöksentekoa otetaan käyttöön heidän tilanteissaan.

Ohjausryhmä näki tärkeäksi kiteyttää kriteerien ryhmittelyä tai mahdollisesti tiivistää itse kriteereitä, eli olisi helpompi lähestyä neljää-viittä kriteeriä tai selkeästi nimettyä kriteerijoukkoa/ulottuvuutta. Laatukriteerien helppo lähestyttävyyys on tärkeää käyttöönottoa tukevana ominaisuutena. Kaukonen vie evästyksen eteenpäin laatukriteerien iteroivaan kehittämiseen.

3. Tietoaineistojen laatukriteerien pilotointi

Projektipäällikkö esitteli tietoaineistojen laatukriteerien pilotoinnin suunnitelman: tavoitteet, karkean aikataulun sekä työstötavat ja tuotokset. Pilotoinnin tavoitteena on saada ymmärrystä alustavien laatukriteerien käytöstä ja soveltuvuudesta eri aloilla ja erilaisissa tilanteissa ja näin vahvistaa laatukriteerien käyttöönottoa ja mahdollisuuksia tulla osaksi organisaatioiden jatkuvaa toimintaa. Pilotoinnin aikana on tarkoitus tunnistaa mahdolliset laatukriteerien kehittämistarpeet sekä kehittää ehdotus laatukriteereihin liitettäväksi mittaristoksi, joka konkretisoi laatukriteereitä, tukee tiedon laadun arviointia ja kehittämistä sekä helpottaa tiedon laadun viestintää niin käyttäjien kuin organisaation johdonkin suuntaan.

Pilotoinnin toteutus jakaantuu kahteen vaiheeseen. Pilotoinnin 1. vaihe tuottaa mittariaihiot tiiviissä 18.3.2021 päättyvässä jaksossa. Tämän jälkeen käynnistyy pilotoinnin 2. vaihe, jossa testataan tuotettuja mittariaihioita ja tarvittaessa tuotetaan myös uusia. Pilotoinnin toteutus päättyy toukokuussa, jolloin siirrytään yhteenvetovaiheeseen, jolloin arvioidaan myös mahdollinen jatkopilotoinnin tarve. Pilotoinnin 1. vaiheessa ovat mukana Maanmittauslaitos, Vero ja Tilasto-

keskus omilla aineistoillaan. 2. vaiheeseen lähtevät lisäksi mukaan ainakin Tulli, Luonnonvara-keskus ja Digi- ja väestötietovirasto. Keskustelut mahdollisten osallistujien kanssa ovat vielä käynnissä.

Pilotointisuunnitelmaa ja erityisesti pilottien keskinäistä yhteistyömallia pidettiin hyvänä. Pilottoijien joukkoon kaivattiin ohjausryhmässä vielä uusia toimijoita, kuten Kelaa ja THL:ää. Juuso Rahkola kertoi, että Kelassa ollaan projektin työssä mukana ja seurataan kehitystyötä, vaikka nyt ei ole mahdollisuutta osallistua itse pilotointiin. Projektipäällikkö kertoi, että THL:n VALTAVA-hankkeen kanssa keskustelut mahdollisesta pilotoinnista ovat vielä kesken.

Ohjausryhmä nosti esille sen, että sidosryhmiltä olisi tärkeää saada näkemystä, siitä, miten he näkevät tiedon laadun. Ettei laatu ole vain tietoaineiston omistajan asia. Keskustelussa todettiin myös, että käyttäjällä tulee olla mahdollisuus ilmoittaa laatueroista. Projektipäällikkö kertoi, että pilotoinnissa on tavoitteena tarkastella asiaa myös sidosryhmien näkökulmasta yksittäisissä piloteissa ja eri tiedon vaihdon tilanteissa.

Keskustelussa todettiin, että tiedon laatu on haasteellinen asia rekisterinpitäjälle, sillä kirjautustavat ovat usein hyvin erilaisia tiedon alkulähteillä. Tarvitaankin vuoropuhelua rekisterinpitäjien, tiedon jalostajien ja käyttäjien kesken siitä, mikä on riittävää tiedon laatua ja mitä asioita tiedon laadussa halutaan painottaa. Tiedon laadun parannuksessa on tarkastelussa usein myös vastuun kohdentuminen ja kustannukset.

Ohjausryhmässä nähtiin, että tiedon laatukehikko kertoo tarinaa siitä, millainen laatu on kulloinkin kyseessä, sillä tiedon luotettavuus ja laatu ei ole aina sitä mitä luullaan. Keskusteltiin myös tavoiteltavasta laadun tasosta, sillä laatu on aina kallista. Eri tilanteissa tavoiteltava laatu vaihtelee. Keskustelussa pohdittiin laajemminkin tiedon laadun tavoitetasoa määrittelyä. Riittävän laadun tason määrittää loppujen lopuksi käyttäjä. Tiedolla on lähestulkoon aina monta erilaista käyttäjää, ja osalla on erilaisia tarpeita tiedon laadun suhteen. Tiedon käyttötarkoitus vaikuttaa tavoiteltavaan tiedon laatuun. Käyttäjä tekee lisäksi itse arvovalinnan eri laatukriteerien, esimerkiksi ajantasaisuuden ja kattavuuden, välillä. Todettiin, että tämän takia käyttäjien tarpeet on lähes mahdotonta huomioida tyhjentävästi ja siksi ratkaisuna tiedon laatukehikossa toteutetaan tiedon laadun yhtenäinen kuvaus, jonka pohjalta käyttäjä voi arvioida onko laatu hänelle riittävää.

Ohjausryhmä tiedusteli myös metadatan suhdetta laatukriteereihin. Projektipäällikkö kertoi, että metadata sisältyy ymmärrettävyyden kriteeriin, mutta suhde metadataan on moninainen. Laatukriteerit itsessään tulee myös kuvata tietoaineistojen metatiedoissa. Ohjausryhmä näki, että kaikki laatutyön keskeiset tulokset tulisi rutiininomaisesti viedä osaksi aineiston kuvailua. Projektipäällikkö totesi, että yhteistyö myös tällä saralla TiHA-työpakettien kesken on tärkeää.

Ohjausryhmäkeskustelun pohjalta sidosryhmätarkastelu ja käyttäjätulokulma on hyvä pitää pilotoinnissa esillä ja käytännössä nostaa myös pilotoinnin loppuraportissa tarkasteltavaksi asiaksi. Projektipäällikkö vie nämä evästyksensä pilotoinnissa huomioitaviksi.

4. Muut asiat (10 min)

Projektipäällikkö kertoi, että maaliskuun alussa julkaistaan tiedon laatuun keskittyvä jakso [Tiedon äärellä - tietopolitiikasta tekoihin](#) -podcast -sarjassa. Keskiviikkona 14.4. järjestetään kuntien keskustelutilaisuus, jossa on tarkoitus keskustella erityisesti tiedon laadusta. Torstaina 22.4. on tiedossa sidosryhmätilaisuus, jossa tarkastelussa tiedon laatu muiden TiHA-työpakettien ohella.

Olli-Pekka Rissanen muistutti lisäksi 9.3.2021 pidettävästä tiedon hyödyntämisen ja avaamisen strategisten linjausten tilaisuudesta, joka pohjustaa linjausten lausuntokierroksen. Osallistuminen on ollut aktiivista, edelliseen työpaketin 1 tilaisuuteen osallistui noin 300 henkilöä. Työpakettien 4 API-linjaukset tulevat puolestaan syksyllä lausuttavaksi.

5. Seuraavat kokoukset (5 min)

Seuraava kokous on torstaina 17.6.2021 klo 14-16. Ohjausryhmä kokoontuu neljännesvuosittain. Ohjausryhmäkokoukset ovat etäkokouksia ainakin kevätkaudella.

6. Kokouksen päättäminen

Puheenjohtaja kiitti hyvää ja asiantuntevasta keskustelusta ja evästyksistä ja päätti kokouksen 10.50.