

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

Aika: Torstaina 5.11.2020 klo 14.30-16.15
Paikka: Teams
Läsnä: Vesa Ylönen (KELA), Antti Jakobsson (MML), Marja Rantala (MML), Jaana Lehtilä (Vero), Alpo Lahtinen (Vero), Erja Nokkanen (OPH), Kukka-Maaria Berg (OPH), Riitta Alkula (DVV), Sini Slotte (DVV), Olli-Pekka Penttilä (Tulli), Sari Kuisma-Kari (PRH), Outi Ahti-Miettinen (TK), Jarmo Ranki (TK), Jari Kosonen (TK), Noora Seppä (TK), Jari Nieminen (TK), Jennika Leino (TK), Mikko Saloila (TK), Essi kaukonen (TK)

1. Tapaamisen avaaminen

Tapaaminen aloitettiin 14.30 työpajan ohjelman esittelyllä.

2. KV benchmarkkaus julkishallinnon tiedon laatukehikoiden suhteen

Mikko Saloila esitteli kansainvälistä vertailua julkishallinnon tiedon laatukehikoiden suhteen, sekä siinä mukana olevia maita ja niiden tilannetta. Benchmarkkauksessa on hyvä huomioida, että julkisen tiedon avaruudet eroavat merkittävästi: paikallishallinnoilla on erilaisia vastuita eri maissa ja lainsäädäntö eroaa eri maissa merkittävästi. Tiedon laatukehikot ja niiden sisältö käsitetään eri tavoin. Tiedon hallinta ja mahdollinen tietovastaavan rooli yhdistetään eri tavoin tiedon laadun hallintaan eri maissa. Toisaalta tiedon laadun taustamalleina käytetään ilmeisesti samoja lähteitä.

Saloila kertoi, että kiinnostavia maita on rajattu eri lähteistä saatujen tietojen perusteella. Näistä pisimmälle edenneiden maiden kanssa aikataulutetaan verkkokokouksia. Virtuaalitapaaminen on ollut hyvä tapa saada lisätietoa. Tiedon laatukehikot ovat varsin alkuvaiheessa ja suoraan verkosta niistä ei ole paljoakaan tietoa saatavilla. Kanadan kanssa palaveri on jo pidetty, Uusi-Seelanti ja Norja ovat vielä tulossa. Australia on tunnistettu, mutta palaveriin sen kanssa ei ole vielä tässä vaiheessa pyritty. Australiassa julkisen hallinnon laatukehikkona on osittain käytetty Australian tilastoviraston laatukäsikirjaa. Palaveritarpeita arvioidaan vielä tarkemmin uudelleen Uuden-Seelannin ja Norjan kanssa pidettävien kokousten jälkeen. Verkkokokousten pohjalta katsotaan mitä tietoa on saatu, ja onko tarkoituksenmukaista kartoittaa vielä uusia maita.

Keskustelussa pohdittiin myös Data Stewardshipin ja Data Custodianshipin eroja ja tätä on hyvä jatkaa. Näissä molemmissa on kyse tiedon hallinnasta ja tiedosta huolehtimisesta. Projektipäällikkö totesi, että Data Stewardship on käytössä tilastoalalla. Tiedon laatukehikkoa kehitettäessä on tärkeää saavuttaa yhteinen ymmärrys erityisesti valituista suomenkielisistä käsitteistä, kuten tietovastaava ja tiedon omistaja.

3. Taustamallit

Jennika Leino esitteli Tiedon laatukehikon ja erityisesti laatukriteerien taustalla vaikuttavia potentiaalisia taustamalleja. Tässä vaiheessa on tunnistettu ja tarkasteltu seuraavia: ISO/IEC 25012 Data

quality model (SQuaRE), ISO 8000 (Global standard for Data Quality and Enterprise Master Data), ISO19157 Geographic information — Data quality, OGD principles of open government data (14 Principles of Open Government Data, Open Government Data Principles), DCAT, DQV vocabulary (development in progress, extension to the DCAT vocabulary). Laatuksiteerejä on tarkasteltu myös useissa artikkeleissa.

Projektin tähänastisessa työssä malleja on tunnistettu, käyty läpi ja vertailtu keskenään. Mallien laatuksiteerejä on arvioitu ja käännetty alustavasti suomeksi työstämistä varten sekä tarjottu ensivaiheessa relevanteilta vaikuttavilla laatuksiteerejä osallistujien pohdittavaksi kyselyyn ja sen kautta tähän työpajaan.

Lopuksi Essi Kaukonen esitteli vielä hyvin karkean taulukkomuotoisen varmentamattoman näkemyksen siitä, miten Norjan, Viron ja Kanadan tiedon laatuksiteerit määppäytyvät keskenään ja ISO/IEC 25012 Data quality model (SQuaRE) -malliin.

Työpajassa ei noussut esille uusia tiedon laadun taustamalleja tarkasteluun, mutta mietittiin palvelun laadun ja tuotelaadun mallien hyödyntämistä. Pohdittiin, että palvelu/tuoteajattelu voisi toimia hyvin ainakin, jos kyse on päivittyvästä datasta tai jotakin tarkoitusta varten laaditusta ja kuvaillusta tiedosta. Keskustelussa todettiin, että tietopalvelun laatumalleja on ja että tietotuotelaatua löytyy esimerkiksi tilastojen laatumalleista.

4. Tiedon laatuksiteerien työstö

Projektipäällikkö kertoi pajaosuuden aluksi, että tiedon laatuksiteerien käsittelyn kanssa päätettiin lähteä liikkeelle jo nyt, jotta niiden työstämiselle jää riittävästi aikaa. Huoli tiukasta aikataulusta oli tullut esille jo edellisessä pajassa. Samaan aikaan projektissa ovat edelleen menossa nykytilaselvitys ja kv benchmarkkaus, mikä tuo oman haasteensa. Alustavat tiedon laatuksiteerit on tarkoitus muodostaa tämän vuoden puolella, ja näitä voidaan sitten iteroida ja täydentää 2021 pilottien ja muun kokemuksen ja yhteistyön perusteella.

Projektipäällikkö esitteli kaikille projektin osallistujille lähetetyn kyselyn tuloksia. Vastauksissa nostettiin esille kansallisten määritysten suhde (toimialakohtaisiin) kansainvälisiin vaateisiin. Kyselyn vastauksista ilmeni myös, että vastaajilta on kysely tiedon laadusta kaikkea mahdollista, liittyen esimerkiksi yhteentoimivuuteen, ymmärrettävyyteen, tulkintaan, käsittelyyn, kattavuuteen, luotettavuuteen ja virheisiin.

Kyselyssä esitetyistä mahdollisista laatuksiteereistä erityisesti kattavuus oli kirvoittanut kommentteja. Tietoturvaluisuus nousi esille ainoana uutena ehdokkaana mahdolliseksi tiedon laatuksiteeriksi, se nostettiin mukaan päivän pajan työstöön ajankohtaisena ehdotuksena. Muut kolme tässä työpajassa työstettyä kriteeriä olivat Ymmärrettävyys, Saatavuus ja Tekninen käytettävyys.

Työpajan alussa käytiin läpi joitakin pajassa käytettäviä käsitteitä. Muuttuja on näistä termistöltään kirjavin. Muuttuja-termin lisäksi

osallistujaorganisaatioissa on käytössä seuraavia synonyymejä: ominaisuus, attribuutti, kohde.

Työstö tehtiin Flinga-työkalussa vapaassa kierrossa neljässä rinnakkaispalaverissa, siten, että kullakin palaverilla oli oma vetäjä Flinga-työkalun ääressä. Väline koettiin toimivaksi. Todettiin, että virtuaalipaja toimii yhtä hyvin kuin entisajan työpajat, joissa moderaattori kirjaa kommentteja. Toivottiin, että kierto rinnakkaispalaverista toiseen tapahtuisi tietynä ajanhetkenä, jolloin palaverinvaihtoon voisi liittää vetäjän esittelemän tiivistetyn esityksen siihenastisesta keskustelusta.

Työstön ja sen palautekeskustelun pohjalta vaikuttaa siltä, että Tietoturvaan liittyvät laatukriteerit (Saatavuus, Luottamuksellisuus ja Eheys) on hyvä siirtää hetkeksi sivuun, samoin Tietosuoja / Tietosuojaus. (Näiden osalta on hyvä seurata myös TiHA TP1 työpaketin tietoturvaa ja tietosuojaa koskevia ohjeita ja linjauksia sekä TiHA TP4 työpaketin API-linjauksia ja ohjeistuksia.)

Tekninen saatavuus sai työpajakäsittelyssä sisällökseen Koneluettavuus / Siirrettävyys -elementtejä ja menee jatkokäsittelyyn tuolla sisällöllä. Ymmärrettävyys oli saanut kyselyssä eniten ”ääniä” selkeänä laatukriteerinä ja tämäkin kriteeri siirtyy jatkokäsittelyyn. Jatkokäsittely tehdään näillä näkymin jollakin toisella välineellä.

Työpajan keskustelussa toivottiin selkeämpää liitosta taustamalleihin, ja jatkossa laatukriteerin taustalla vaikuttava taustamalli ja sen englanninkielinen nimi ja kuvauskin olisi hyvä pitää esillä jossain muodossa. Pohdittiin yhteentoimivuutta tiedon laatukriteerinä sekä sitä, että miten tiedonvaihto seuraavien kesken vaikuttaa esimerkiksi ymmärrettävyyteen: julkinen-julkinen tai julkinen-yksityinen, julkinen-kansalainen, kansalainen-yksityinen. Käytettävyys koettiin vaikeaksi sen subjektiivisen luonteen takia. Keskustelussa todettiin myös, että Nykytilaselvitys valmistuttuaan voi antaa osviittaa Tiedon laatukehikon ja laatukriteerien sisällölle.

Keskustelussa siirryttiin myös tiedon laatukriteerejä laajemmallekin tasolle ja tuotiin Eurooppalaisen yhteentoimivuuden viitekehyksen EIF:n lisäksi esille EIRA, jonka yhteentoimivuuden rakennuspalikoista saisi lihaa EIF:in neljän näkökulman ympärille, esim. erilaisia katalogeja. Kaivattiin erilaisia näkökulmia, esimerkiksi jaottelulla julkishallinto-perustietovarannot-tuotantoprosessit ja henkilöt - palvelut/tuotteet - asiakkaat/kansalaiset. Tiedon laatukriteerien yksityiskohtainen työstäminen ketterästi palastellen toi arvokasta yhteistä käsitystä käsitellyistä aiheista, mutta asiaan siirtyminen ilman laajaa kuvaa aiheutti haasteita kokonaisuuden näkymättömyyden takia. Tähän on hyvä panostaa yksityiskohtaisen työstämisen rinnalla.

5. Seuraava tapaaminen

Seuraava tapaaminen on 19.11.2020

Jakelu:

Tiedon laatukehikko TiHA TP3 Teams-työtila sekä Tiedon laatukehikko verkkosivusto