



Lausunto

4.10.2021

VN/16357/2020
VN/16357/2020-MMM-90

Maa- ja metsätalousministeriön lausunto Tiedon laatukehikko -selvityksestä

Laatukriteerikokonaisuus ja sen vaikuttavuus:

Miltä laatukriteerikokonaisuus vaikuttaa?

Julkaisun alussa todetaan, että tämä on ehdotus julkishallinnon tiedon laadun kuvaukseen ja arviointiin tarkoitetusta laatukriteerien ja mittariston muodostamasta kokonaisuudesta. Tavoitteena on, että laatukriteerit ja mittaristo yhdessä tarjoaisivat selkeän ja ymmärrettävän työkalun julkishallinnon tietoaaineistojen, nimenomaan rakenteisen datan, yhteismitallista ja käyttäjälähtöistä laadun kuvaamista ja tarkastelua varten. Tarkasteltava dokumentti on suunnattu erityisesti tiedon laadun kysymysten parissa työskenteleville tiedon tuottajille ja tiedon laadusta kiinnostuneille tiedon käyttäjille. Tavoitteena on tarjota, ensimmäistä kertaa, yhteinen kansallinen kieli ja termistö tiedon laadun käsittelyyn tiedon laadun asiantuntijoille ja siitä kiinnostuneille.

Keskiosässä on siis tiedon laadun kuvaaminen tiedon tuottajille ja tiedon laadusta kiinnostuneille tiedon käyttäjille ja tarjota yhteinen kansallinen kieli ja termistö tiedon laadun käsittelyyn tiedon laadun asiantuntijoille ja siitä kiinnostuneille. Tavoite tarjota yhteinen termistö ja mittaristo tiedon laadun määrittelyyn on kannatettavaa ja kehittäisi osaltaan julkisen hallinnon tiedon hyödyntämistä. Tiedon laadun eri näkökulmien jäsenetty kuvaaminen on hyvä tavoite.

Kokonaisuus vaikuttaa kunnianhimoiselta. Laatukriteeristöä on kuitenkin syytä määritellä, sillä tämä kriteeristö helpottaa tiedon laadun arviointia ja yhteismitallistamista.

Selvityksessä todetaan, että tiedon laatu on monitahoinen käsite. Ilmeisesti tästä syystä tiivistettyä määritelmää käytetyille termille laatu ei ole pystytty kuvaamaan, vaan päädytty joukkoon laatua kuvaavia tiedon ominaisuuksia (laatukriteerejä). Tekstissä todetaan, että nyt kehitetyt laatukriteerit pohjaavat pitkälti ISO 25012 -standardiin, jota käytetään myös useissa muissa maissa tietoaaineistojen laadun kuvauksen taustamallina. Lisäksi laatukriteereihin on tuotu perustietovarannoille tärkeä oikeellisuuden kriteeri sekä asiakasnäkökulmaa.

Standardin kuvauksessa todetaan, että The Quality of a Data Product may be understood as the degree to which data satisfy the requirements defined by the product-owner organization. Specifically, those requirements are the ones that are reflected in the Data Quality model through its characteristics (Accuracy, Completeness, Consistency, Credibility, Currentness, Accessibility...). Tiedon (tietyn aineiston) laatu määrittyy käytännössä aineiston datan ominaisuuksien suhteesta asetettuihin vaatimuksiin. Nämä vaatimukset ovat joko alkuperäiselle tiedon tuottamiselle ja tuotteistamiselle asetettuja (ensikäyttö) tai muissa käyttötapauksessa määriteltyjä (ns. toisiokäyttö).

Postiosoite
Postadress
Postal Address
Maa- ja metsätalousministeriö

Käyntiosoite
Besöksadress
Office

Puhelin
Telefon
Telephone

Faksi
Fax
Fax

s-posti, internet
e-post, internet
e-mail, internet

PL 30
00023 Valtioneuvosto

Hallituskatu 3 A
Helsinki

29 516 001
35 829 516 001

kirjaamo@mmm.fi

Kohdan 2.1 mukaan laatukriteerien ja mittariston muodostaman kokonaisuuden ensisijaisena tavoitteena on tukea tietoaineistojen, erityisesti rakenteisen datan laadun tunnistamista ja kuvaamista tiedon käyttäjille erilaisissa tiedon vaihdon tilanteissa sekä tiedon laadun yhtenäisen vertailun kehittämistä julkishallinnossa. Laatukriteerit yhdessä mittariston kanssa tavoittelevat siis yhdenmukaista kuvausta erilaisten tietoaineistojen laadusta yli organisaatio- ja alarajojen ja auttavat siten käyttäjiä arvioimaan, soveltuuko tietoaineisto aiottuun käyttötarkoitukseen.

Tässä työssä pyritään siis luoman enemmänkin työkaluja toisiokäyttöön liittyvälle tiedon ominaisuuksien arvioinnille. Tämä asetelma olisi hyvä tuoda selkeämmin esille johdannossa ja muualla tekstissä. **Kysymyksessä on enemmänkin tiedon ominaisuuksien kuvaamisesta kuin varsinaisen laadun kuvaamisesta, koska laatu määrittyy näiden ominaisuuksien kautta tapauskohtaisesti.** Tähän tiedon (tietyn aineiston) ominaisuuksien arviointiin ja vertailuun laadittu kehikko näyttää tarjoavan hyvät lähtökohdat. Tästä syystä **MMM pitäisi valitussa terminologiassa 'laatukriteeri' termin sijasta parempana terminä esimerkiksi 'laatutekijä' termiä tai muuta tiedon ominaisuuksiin viittaavaa termiä.** Laatukriteeri-termistä syntyy se mielikuva, että se on tiedon laadulle asetettava vaatimus, joka ohjaa tiedon tuottamista. Joidenkin toisiokäyttöön tarkoitettujen tuotteistusten osalta näin on varmasti ajateltukin, mutta alkuperäisen tiedontuotannon laadulliset vaatimukset syntynevät edelleen toiminnan vaatimuksista. Tätä kuvastaa myös kohta 2.3, jossa todetaan, että laatukriteerit ja mittaristo eivät ota kantaa tavoiteltaviin laatutasoihin.

Termiä 'laatu' käytetään melko väljästi tiedon ominaisuuksien synonyyminä esim. kappaleessa 2.1: *sekä tiedon laadun yhtenäisen vertailun kehittämistä julkishallinnossa.* Voitaneen arvioida kuinka laadukkaasti tiedon ominaisuudet on kuvattu, mutta ei sitä, miten tieto soveltuu hyvinkin laajaan kirjoon erilaisia (toisio)käyttötapauksia. Laatu-termin käyttöä olisi edelleen tarkistettava ja pohdittava milloin tarkoitetaan laatua ja milloin jotakin tiedon ominaisuutta.

Julkaisussa todetaan, että pidemmällä aikavälillä kehittyvä laatukriteerien ja mittariston kokonaisuus voisi myös mahdollistaa julkishallinnon tiedon laadun kansallisen ohjauksen ja seurannan, mikäli se katsotaan tarpeelliseksi. **MMM katsoo, että tämä on mahdollista niissä tapauksissa, joissa tuetaan tietovarantojen toisiokäyttöä ja/tai tietoa tuotetaan varta vasten yleiseen käyttöön. Erilaisten julkishallinnon prosessien tiedon laadun vaatimusten tulee kuitenkin jatkossakin määrittää pääosin toiminnan vaatimuksista ei vaihtelevan toisiokäytön vaatimuksista.**

Kuinka hyvin laatukriteerit ja mittarit tukevat yhteismitallista tiedon laadun tunnistamista ja siitä keskustelemista julkishallinnossa?

Vrt. edellinen asiakohda. Laatukriteerit ja mittarit tukevat tiedon käytettävyyden arviointia ja laatua erilaisissa käyttötapauksissa, mutta ei sovellu esimerkiksi jonkin julkisen hallinnon prosessin vaatimusten määrittelyyn kuin yksittäisissä tapauksissa (niissä, joissa jotakin perustietoa nimenomaisesti tuotetaan yleiseen käyttöön).

Kriteeristö helpottaa tiedon laadun tunnistamista. Tiedon laadun yhteismitallistaminen edellyttää kuitenkin laatukriteeristön lisäksi tiedon semanttisen yhteentoimivuuden kehittämistä. Tällä hetkellä datalla ja tiedolla käsitetään eri toimialueilla jossain määrin eri asioita. Semanttisen yhteentoimivuuden puute muodostaa haasteen laatukriteeristön käytölle. Mutta laatukriteeristö antaa toisaalta työkalun, jota voidaan hyödyntää myös tiedon semanttisen yhteentoimivuuden kehittämisessä. Tämä liittyy suoraan laatukriteeriryhmään 5 ja kriteeriin 5.2.2.

Laatumittarit antavat kehikon tunnistaa tiedon laatua yhteismitallisesti. Sen lisäksi tarvitaan kuitenkin koulutusta niiden yhteismitallisen käytön varmistamiseksi. Keskustelua tarvitaan myös organisaation käytössä olevien laatumäärittelyjen kytkemisestä yhteiseen laatukehikkoon täydentävinä alakohtaisina mittareina.

Laatukriteerit ja mittarit, erityisesti oikeellisuutta ja tarkkuutta kuvaavat mittarit, luovat hyvän pohjan tiedon laadun tunnistamiseen ja siitä keskusteluun. Kaikista aineistoista näitä ei kuitenkaan ole mahdollista tuottaa.

Missä yhteyksissä laatukriteeristön mukainen yhteinen tiedon laadun tunnistaminen ja määrittely tukisi julkishallinnon tiedon hyödyntämistä ja avaamista?

Vrt. edelliset asiakohdat. Kuvattu laatutekijöiden kuvattu kehikko tukee tiedon hyödyntäjiä, niiden arvioidessa tiedon (tietyn aineiston) soveltuvuutta johonkin käyttötapaukseen. Tällä tavalla kehikko auttaa välillisesti myös tiedon avaamista. Laatukriteeristö on huono termi, koska se luo mielikuvan tiedolle asetettavista vaatimuksista eikä tiedon kuvaamisen osa-alueesta (laatutekijästä tms.).

Kun suunnitellaan uuden toisen viranomaisen tuottaman aineiston hyödyntämistä viranomaispäätöksessä tai tutkimuksessa, yhtenäiset laatukriteerit helottaisivat työtä.

Yhteisesti määritelty laatukriteeristö parantaa tiedon yhteentoimivuutta sekä osaltaan mahdollistaa ja helpottaa rajapintojen käyttöä.

Miten hyvin laatukriteeristö ja mittaristo toimisi tiedon laadun parannuksen tukena pidemmällä tähtäimellä?

Vrt. edelliset asiakohdat. Laatukriteeristö ja mittaristo voisivat ohjata toisiokäyttöön tarkoitettua tuotteistusta (esim. rajapinnoissa), jos toisiokäyttö on selkeästi rajautuvaa ja laajaa. Se voi ohjata alkuperäistä tiedon tuottamista niissä tapauksissa, joissa tietoa tuotetaan alun perin yleiseen ja laajaan käyttöön. Aina joudutaan kuitenkin punnitsemaan mitä käyttötapauksia painotetaan.

Tiedon laadun parantamiseen vaikuttaa pitkälti käytettävissä olevien resurssien määrä. Laatua on hyvä seurata, mutta se ei välttämättä paranna laatua, jos resursseja ei tiedon tuotantoon ja/tai ylläpitoon pystytä lisäämään.

Laatukriteeristö helpottaa osaltaan tiedon laadun parantamista, mutta kriteeristön kehittämisen pitää kytkeytyä muuhun yhteentoimivuuden kehittämiseen tähtäävään työhön.

Tiedon käyttäjän tulokulma laatukriteeristöön ja mittaristoon:

Miltä laatukriteerien ja mittarien kokonaisuus vaikuttaa tiedon käyttäjän näkökulmasta?

Vrt. edelle. Julkaisussa todetaan kohdassa 2.4, että laatukriteerit auttavat tunnistamaan, mitkä tekijät ovat tiedon laadun näkökulmasta tärkeitä. Siinä käyttäjälle tunnistettiin olevan tärkeää saada kattava ja oikea kuva tietoaineiston sisällöstä, sisällön ja sen kuvauksen laadusta sekä tiedon käyttömahdollisuuksista.

Tähän tarkoitukseen kehikko vaikuttaa toimivalta. MMM katsoo, että tässä käytetty termi 'tekijä' (laatutekijä) otettaisiin käyttöön kriteeri-termin sijaan.

Laatukriteeristö on laaja. Kaikki mittarit eivät välttämättä ole kaikissa käyttötapauksissa relevantteja. Esimerkiksi määräpäivien noudattaminen. Tämän vuoksi **olisi perusteltua mahdollistaa joustavuutta mittareiden käytössä.**

Tuottaako se riittävän kuvan yksittäisen tietoaineiston laadusta sen käyttäjän tueksi?

Monissa käyttötapauksissa tarvitaan yksityiskohtaisempaa tietoa tiedon tai yksittäisen tietokohteen syntyvästä ja tiedon elinkaaresta. Yleiseen tiedon hakemiseen ja käytön arviointiin kehikko soveltuu hyvin. Varsinaiseen käyttöönottoon tarvittaneen usein yksityiskohtaisempaa tietoa. Tämä nostaa esille kysymyksen mikä on tietoaineistoa kuvaavan metatiedon ja tämän tiedon laadullisia ominaisuuksien kuvaavan tiedon suhde?

Laatukriteerit kattavat hyvin eri tiedon laadun näkökulmia, minkä johdosta mittaristo antaa käyttäjälle eväitä tutkia tietoaineiston käytettävyyttä omassa käyttötarkoituksessa.

Kaipaisimme kuitenkin jonkinlaista luokittelua tiedon laadun luotettavuudesta. Esimerkiksi Lajitietokeskuksen lajit.fi -palvelun lajihavainnoissa on eri tasoisia havainnoita. Osa havainnoista on ko. lajin spesialistin tekemiä havainnoita, osa havainnoista on kansalaishavainnoita. Havaintojen luotettavuudessa voi tästä syystä olla hyvinkin suurta eroa, millä on suurta vaikutusta tiedon hyödynnettävyyteen esimerkiksi viranomaisten päätöksenteossa.

Tiedon tuottajan tulokulma laatukriteeristöön ja mittaristoon:

Miltä laatukriteerien ja mittarien käyttöönotto tiedon avaamisen yhteydessä vaikuttaisi tiedon tuottajan näkökulmasta?

Vrt. edelliset asiakohdat. Tietoa tuotetaan erilaisissa prosesseissa ja erilaisiin käyttötarpeisiin. Lähtökohtaisesti julkisen hallinnon toimijat keskittyvät hoitamaan tehtävänsä ja määritellyt tietojärjestelmät tukevat näitä prosesseja. Tietoon kohdistuva vaatimukset määrittyvät toiminnan vaatimusten kautta. Osa prosesseista toki tähtää yleiseen käyttöön tarkoitettun tiedon tuottamiseen. Uusi kehikko toisi tiedon tuottajille yhtenäisen tavan tietoaineistojen laadullisten ominaisuuksien kuvaamiseen. Yleisellä tasolla kehikko vaikuttaa toimivalta, mutta tiedon tuottajien on usein ylläpidettävä yksityiskohtaisempaa tietoa tiedon tai yksittäisen tietokohteen ominaisuuksista, syntyvästä ja tiedon elinkaaresta. Tämä nostaa esille kysymyksen mikä on tietoaineistoa kuvaavan laajemman metatiedon ja tämän tiedon laadullisia ominaisuuksien kuvaavan tiedon suhde? Olisi myös hyvä pohtia mikä suhde tällä kehikolla on tiedonhallintalain mukaisiin vaatimuksiin tiedonhallinnallisen ja –kartan ylläpitoon.

Tavoitteena kokonaisuus on tiedon tuottajan näkökulmasta hyvä. Mittareita on melko paljon, joten niiden kaikkien käyttöönotto vaatii aikaa. Lisäksi osa mittareista voi olla sellaisia, ettei niiden osalta pystytä tuottamaan tietoa tai se vaatisi mittavia panostuksia.

Esimerkiksi luonnonvaratiedon osalta tiedon tuottajan kannalta metatietojen kuvaaminen esitetyllä tavalla on joiltakin osin haasteellista koska esimerkiksi kohdejoukko ei ole hyvin tunnettu. Kriteeristö tulee edellyttämään paljon työtä, ei niinkään yhden tuottajan osalta vaan yhteisesti yhteentoimivuuden takaamiseksi.

Mitä laatukriteerien ja mittarien käyttöönotto tiedon tuottajan näkökulmasta edellyttäisi?

Edellyttää nykyisten metatietokuvausmallien yhteensovittamista tämän uuden tiedon laadullisia ominaisuuksien kuvaavan kehikon kanssa. Joissakin tapauksissa tämä voi vaatia lisätyötä. **Olisi hyvä myös selkeyttää miten tämän kehikon käyttö kohdennetaan tietoon. Kuvataanko tietovarannon ominaisuuksia yleisesti vai kohdennettuna johonkin tietoaineistoon, tietokantaan, tietotuotteeseen tms.**

Kehikossa näyttää olevan sisäänrakennettu odotus, että tietoa hallitaan eikä vain kuvata näiden tekijöiden mukaisesti. **Mikäli kaikista keskeisistä tietovarannoista (tai tarkennettuna aineistoihin, tietokantoihin ym.) kootaan täydelliset tiedot, vaaditaan tiedon tuottajilta tähän huomattava työ- ja laskentapanos.**

Lisäksi tarvitaan tukea kuvausten tekemiseen mittariston yhtenäisen käytön varmistamiseksi.

Laatukriteeristön ja mittariston hallinta:

Millaiset tehtävät ovat mielestäsi laatukriteeristöä ja mittaristoa ylläpitävän ja kehittävän tahon tärkeimmät tehtävät?

Kehittävän tahon tulee osata suhteuttaa kehikon käyttö toisaalta tiedon käyttäjien odotuksiin ja tiedontuottajiin kohdistettaviin laatuvaatimuksiin. Selkeät todelliset hyödyt on pystyttävä osoittamaan ja kanavoitava osa syntyvästä hyödystä tiedontuottajien lisätyön kattamiseen.

Oleennaista on nähdä ero tiedon alkuperäisen käytön ja toisiokäytön välillä, jotta viranomaisten alkuperäinen toiminta ei häiriinny tai siihen kohdistu hallitsemattomia vaatimuksia. Tiedon jalostamisen jälkimarkkinoiden rooli (=datatalous?) pitää tunnustaa osana kokonaisuutta.

Laatukriteeristön dokumentaation ylläpito on tärkeää. Dokumentaatioissa on hyvä olla useita erilaisia (tilasto, viranomaisrekisterit, paikkatieto) esimerkkejä havainnollistamassa kuinka mittareita tulisi käyttää. Lisäksi tarvitaan tukea mittariston käyttöön ottaville tahoille.

Tämä osio ohuimmin kuvattu. Yksi tärkeimmistä tehtävistä on varmistaa laatukriteeriryhmän 5 toteutuminen. Tämä edellyttää laajaa yhteistyötä koko julkishallinnolta, etteivät tiedot jää

ymmärrettäviksi ja suositusten mukaisiksi ainoastaan siilojen sisälle tai tiedon tuottajan näkökulmasta.

Minkälaista tukea ja koordinaatiota odotat tältä taholta?

Vrt. edelliset asiakohdat. **Keskittymistä tiedon hyödyntämistä (toisiokäyttöä) tukevaan toimintaan ja tätä varten tehtävään tiedon kuvaamisen edistämiseen.** Tähän voitaisiin mahdollisesti kehittää myös automaattisia tietoteknisiä välineitä?

Ks. edellinen dokumentaatiota koskeva kommentti.

Millainen verkostotoiminta edistäisi laatukriteeristön ja mittariston käyttämistä ja kehittämistä parhaiten?

Mukana tulisi olla sekä tiedon jatkokäyttäjiä että tiedon tuottajia, jotta odotukset ja tarjonta saadaan tasapainoon. Olisiko tätä verkostoyhteistyötä tehtävä myös tietovarantokohtaisesti?

Lausunto toimitettu myös Lausuntopalvelu kautta.

Tietohallintojohtaja

Antti Vertanen
allekirjoitettu sähköisesti

Liitteet

Jakelu

Tiedoksi