

Asia: VN/16357/2020

Lausuntopyyntö julkisen hallinnon tiedon laatukriteereistä

Laatukriteerikokonaisuus ja sen vaikuttavuus

Miltä laatukriteerikokonaisuus vaikuttaa?

Laatukriteeristö nojaa suurelta osaltaan kansainväliseen ISO 25012 -standardiin (Data Quality), joten se on jo siksi näkökulmien ja käsitteiden puolesta vakuuttava, hyvin jäsenelty ja pitkälle työstetty. Kriteerit ovat myös linjassa Euroopan yhteentoimivuusperiaatteen (EIF), FAIR-periaatteiden ja Euroopan tilastojen käytännesääntöjen kanssa.

Kriteeristö on tilastokäsitteistön ja tilastotieteen pohjalta tehty, mikä näkyy käytetyissä käsitteissä? Rajoittaako tämä käsitteistön soveltuvuutta muun tyyppiseen tietoon?

Laadullisiin sisältöihin, joita ei voi numeerisesti arvioida eivät sovellu samat laatukriteerit kuin tilastodataan. Esimerkiksi asianhallintajärjestelmiin laatukriteeristö ei välttämättä sovellu kaikilta osin.

Tilastotieteellisen painotuksen ongelmana on se, että yhteistä kansallista kieltä ei vielä tällä versiolla laatukriteereistä saavuteta.

Asianhallintajärjestelmien kannalta kriteeristöä tulisi tarkastella myös sanallisen (rakenteiset asiakirjat) ja prosessuaalisen aineiston kannalta. Osaa kriteereistä, kuten laskentaa vaativia osuuksia, ei tällaisenaan voi käyttää asianhallintajärjestelmissä ja rakenteisissa asiakirjoissa, joissa on pääosin sanallista, ei numeerista tietoa. Olisi ollut toivottavaa, että esimerkiksi Kansallisarkisto olisi ollut mukana laatukriteerien laadinnassa.

Lähdetietokannoista avoimeksi dataksi siirrettäessä datan laatutekijät vähenevät. Lähdetietokannassa (tuotantokannassa) on yleensä esim. jäljitettävyyys olemassa, mutta tiedon

julkaisun ja jatkojalostamisen jälkeen jäljitettävyyden metatietoja ei enää ole mukana (esim. edk-avoin data).

Suosituksen mukaisuus on lähtökohtaisesti oleellinen tieto yhteentoimivuuden ja organisaatioiden välisten integraatioiden edistämiseksi. Kansallisten sanastojen ja koodistojen yhtenäinen käyttö on tässä keskeinen asia.

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

Eduskunnan kanslia tukee lähtökohtaisesti tiedon avoimuutta, avoimen datan tuottamista ja avointen rajapintojen rakentamista sekä tiedon yhteiskäyttöisyyttä ja tietojärjestelmien yhteen toimivuutta.

Laatukriteeristö on johdannon mukaan laadittu rajattuihin käyttötarkoituksiin: tukemaan rakenteellisen ja koneluettavan tiedon laadun arviointia sekä avoimen tiedon tuottamista, joka on lähtökohtaisesti julkista dataa, ja lisäksi fokus on erityisesti organisaation ulkopuolisiin tiedon tuottamiseen tai tiedon vaihdon näkökulma. Avoimen datan tuottamisen kannalta onkin tärkeää, että lähtödatan laatu on hyväksyttävällä tasolla ja ettei laatu oleellisesti vähene, jos se siirretään avoimeksi dataksi.

On erinomaista, että laatukriteeristöä saadaan tiedon yhteiskäyttöisyyttä tukemaan myös konkreettisempia tarkastelutapoja. Seuraava kysymys kuuluu, miten ja missä yhteyksissä laatukriteereitä voidaan käytännössä hyödyntää? Ilmeisesti käyttötarkoituksina ovat uusien tai jo käytössä olevien tietojärjestelmien ja avoimen datan suunnittelu ja kehittäminen sekä auditointi?

Laatukriteeristön yhteydessä ei myöskään tuoda esille laatutyön vaatimia kustannuksia tai miten ja kenelle mahdolliset hyödyt yhteishankkeissa jakaantuvat. Johtamisen kannalta kustannus-hyöty -suhteen arviointi on organisaatiolle merkittävä kysymys, koska toiminnassaan ja sen kustannusvaikutuksien arvioinnissa on lähtökohtaisesti priorisoitava organisaation ydintoimintaa.

Laatukriteeristön fokus avoimen datan tuottamiseen ja koneluettavaan tietoon herättää myös kysymyksen siitä, onko myös laatutyöstä tulossa jatkossa velvoittavaa, koska avoimeen dataan liittyvät velvoitteet ovat jatkuvasti laajentuneet

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

Mahdollisia käyttäjiä olisi esimerkiksi tiedon jatkojalostamisen yhteydessä ja tai yhdistettäessä eri viranomaislähteiden tietoa erilaisiin portaaleihin

(esim. <https://seco.cs.aalto.fi/projects/semparl/>)

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

Laatukriteerit voivat antaa hyvän pohjan tiedon laadun tarkasteluun organisaation sisäisten tietojärjestelmien suunnittelussa edellyttäen, että ne soveltuvat kyseiseen aineistoon.

Laatukriteeristön kehyksissä voisi evaluoida nykyiset tietojärjestelmät ja tietojärjestelmistä tuotetut jatkojalostuksen muodot (esim. avoin data ja portaalit)

Yhteishankkeissa laadun arviointia voisi käyttää eri organisaatioiden tietojärjestelmien väliseen vertailuun.

Voisiko tiedon laatukriteeristö toimia tulevaisuudessa myös valmissovellusten yhtenä vertailukriteerinä? Jos sovellus tukisi laatukriteereitä, se olisi tuotteelle hyvä kilpailuetu etenkin julkishallinnon asiakaskunnan piirissä.

Kriteeristöstä voisi yleisellä tasolla saada tukea myös tietojärjestelmien suunnittelussa eräänlaisena tarkistuslistana näkökulmista, jotka tulisi tietosisältöjä tarkastellessa huomioida.

Miten kriteereitä olisi tarkoitus käyttää niin, että niiden pohjalta voitaisiin tehdä yhteismitallista laadun tunnistamista ja siitä keskustelua? Tiedon laadun arvioinnin toimintamalli ei vielä tule ilmi annetuista kriteereistä. Jotta näistä kriteereistä olisi hyötyä tiedon yhteismitallisuuden tunnistamiseen, niitä olisi myös käytettävä yhtenäiseen ja vertailukelpoiseen tapaan. Mikä se toimintamalli olisi?

Tiedon käyttäjän tulokulma laatukriteeristöön ja mittaristoon

Miltä laatukriteerien ja mittarien kokonaisuus vaikuttaa tiedon käyttäjän näkökulmasta?

Yksittäinen loppukäyttäjä (esim. verkkopalvelun kansalaiskäyttäjä) ei saa kovin paljon irti tiedon laatukriteereistä, jos heillä ei ole kuvauksia siitä, miten näitä on hyödynnetty.

Laatukriteerit vaikuttavat ensisijaisesti tietojärjestelmien suunnittelijoiden ja tietoa arvioivien asiantuntijoiden työvälineeltä.

Loppukäyttäjän kannalta oleellisia tietoja on vain vähän. Viimeinen päivityshetki lienee ainoita laatutietoja, jotka ovat satunnaisten kansalaisen käytössä.

Tuottaako se riittävän kuvan yksittäisen tietoaineiston laadusta sen käyttäjän tueksi?

- Käyttäjä, jota laatukriteeristö hyödyttää, on pääasiassa tietoa hyödyntävä asiantuntija tai tietojärjestelmien suunnittelija, joka voi tarkistaa, millä tasolla tiedot ovat ja ymmärtää metatiedot.

Kansalaiskäyttöä varten laatukriteerien mukainen arvio olisi tehtävä tiiviisti ja ymmärrettävässä muodossa, jotta siitä olisi hyötyä.

Laatukriteereiden pohjalta tiedon laadusta saadaan riittävän laaja-alainen käsitys, jos vain näitä laatukriteereitä pystytään hyödyntämään. Nyt osa kriteereistä on sellaisia, joita emme saa omasta aineistostamme tuotettua.

Laatukriteeritietojen tuottaminen ja datan validoinnin pitäisi olla pääosin automatisoitua. Tämä edellyttää mittavia investointeja ja lisäresursseja. Lisäksi olisi pohdittava, miten nämä siirretään avoimeen dataan.

Tiedon tuottajan tulokulma laatukriteeristöön ja mittaristoon

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

- Voisi olla hyödyllinen tiedon loppukäyttäjille ja tiedon hyödyntäjille, jos kuvaukset voitaisiin tehdä avoimesta datasta.

Vaikeaa myös arvioida etukäteen, mitä hyötyjä ja kuinka paljon jatkossa laatukriteerien mukaisista tiedoista olisi eri käyttäjäryhmille.

Tulisiko laadun arvioinnista kuvaukset verkkosivuille samaan tapaan kuin tietojärjestelmäkuvauksista ja tietosuojaselosteista? Vai miten ne tulisivat käyttäjille esille?

Mihin laatutiedot tallennetaan? Kun tietoja muunnetaan ja siirretään tallennusjärjestelmästä toiseen tiedon avaamisen yhteydessä, niin tekeekö jokainen toimija omat arvionsa?

Mitä laatukriteerien ja mittarien käyttöönotto tiedon tuottajan näkökulmasta edellyttäisi?

Vaikeaa arvioida, miten paljon resursseja tiedon laadun kuvaukset vaatisivat, ketkä ne tekisivät ja ketkä niitä ylläpitäisivät. Voitaisiinko osa työstä automatisoida, vai olisiko se manuaalinen prosessi.

Jos olisi tarkoitus ylläpitää näitä kuvauksia, pitäisi niiden tuottaminen ohjeistaa.

Vanhaan aineistoon laatukriteereitä saisi vain osittain

Uuteen aineistoon sovitettuna laatukriteeristön hyödyntäminen vaatisi useita ohjelmamuutoksia

Kaikki kriteerit eivät olisi käytettävissä asianhallinta- ja rakenteisten asiakirjojen aineistossa

Laatukriteeristön ja mittariston hallinta

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

Kriteeristön tulee pysyä johdonmukaisena ja eheänä, eikä niitä kannattaisi laajentaa liikaa. Olisi hyödyllisempää priorisoida laatukriteerejä, koska vaarana voi olla, että yksityiskohtaisuus johtaa siihen, ettei kriteerejä käytetä

Laatukriteerien käytön edut tulisi konkretisoida

Yhteisen kansallisen kielen tarkentaminen olisi tarpeen siten, että käsitteistö soveltuisi myös rakenteisiin asiakirjoihin ja muuhun kuin tilastoaineistoon.

Minkälaista tukea ja koordinaatiota odostat tältä taholta?

Mittariston sopivuus erilaisiin aineistotyyppeihin tulee varmistaa (esim. teksti, ääni, video, kuva).

Millainen verkostotoiminta edistäisi laatukriteeristön ja mittariston käyttämistä ja kehittämistä parhaiten?

Ohjeistusta ja konkreettisia esimerkkejä laatukriteeristön ja mittariston käyttämisestä eri tapauksissa tarvittaisiin.

Koulutustilaisuudet tai työpajat voisivat tukea tätä.

Karhula Päivikki
Eduskunta - Lausunnon kokoajana Päivikki Karhula, johtava
tietoasiantuntija, Eduskunnan kirjasto