

Tiedon laatukriteerit ja mittaristo

Tiedon laatukehikko, TiHA TP3



Sisällys

1	Johdanto	3
2	Tiedon laadun kuvaaminen	4
2.1	Tavoitteet.....	4
2.2	Tunnistetut hyödyt.....	4
2.3	Rajaukset.....	5
2.4	Ymmärrys tiedon laadusta tietoaineistojen laatukriteerien kautta	5
2.5	Huomio keskeisiin tekijöihin mittariston tuella	7
2.5.1	Mittariston soveltaminen	8
3	Laatukriteeriryhmä: Mitä tieto koskee?	8
3.1	Laatukriteeri: Jäljitettävyys	9
3.1.1	Muutosten hallinta.....	9
3.1.2	Tiedon elinkaari.....	9
3.1.3	Tietolähde.....	9
3.2	Laatukriteeri: Kattavuus	10
3.2.1	Ajallinen tavoitekattavuus	10
3.2.2	Alueellinen tavoitekattavuus.....	10
3.2.3	Puutteelliset kohdeyksiköt	10
3.2.4	Puutteelliset ominaisuustiedot.....	11
3.2.5	Puuttuvat kohdeyksiköt.....	11
3.2.6	Tavoitekohdeyksiköt.....	11
3.2.7	Tavoiteominaisuudet.....	12
3.2.8	Esimerkit: Ylimääräiset kohdeyksiköt	12
4	Laatukriteeriryhmä: Miten tieto kuvaa todellisuutta?	12
4.1	Laatukriteeri: Ajantasaisuus	13
4.1.1	Luontiajankohta.....	13
4.1.2	Muutosajankohta.....	13
4.1.3	Tarkistusajankohta.....	14
4.1.4	Tiedon muokkaaminen päättynyt.....	14
4.1.5	Viiteajankohta.....	14
4.2	Laatukriteeri: Johdon mukaisuus	14
4.2.1	Tieto loogisuustarkastettu.....	14
4.3	Laatukriteeri: Oikeellisuus	15
4.3.1	Menetelmällisesti tuotetut arvot.....	15
4.3.2	Virheelliset ominaisuusarvot.....	15
4.3.3	Väärinluokittelu.....	16
4.4	Laatukriteeri: Tarkkuus	16
4.4.1	Hajonta	16
4.4.2	Poikkeavat havainnot.....	16

5	Laatukriteeriryhmä: Miten hyvin tieto on kuvattu?	17
5.1	Laatukriteeri: Suositustenmukaisuus.....	17
5.1.1	Käytetyt standardit.....	17
5.1.2	Standardienmukaisuus	18
5.2	Laatukriteeri: Ymmärrettävyys.....	18
5.2.1	Aineistokuvaus.....	18
5.2.2	Käsitteiden määritelmät.....	18
5.2.3	Ominaisuuksien tietokuvaukset.....	19
5.2.4	Ymmärrettävyyden asiakaspalaute.....	19
6	Laatukriteeriryhmä: Miten tietoa voi käyttää?	19
6.1	Laatukriteeri: Koneluettavuus.....	20
6.1.1	Kohdeyksikön pysyvä tunniste.....	20
6.1.2	Koneluettavuuden asiakaspalaute	20
6.1.3	Tietoaineiston tietomalli.....	20
6.2	Laatukriteeri: Käyttöoikeudet.....	21
6.2.1	Käyttöoikeus.....	21
6.2.2	Käytön rajoitukset.....	21
6.3	Laatukriteeri: Täsmällisyys.....	21
6.3.1	Määräpäivien noudattaminen	22
6.3.2	Päivitystiheys.....	22
6.3.3	Päivityksessä muuttuneet ominaisuustiedot.....	22
7	Laatukriteerien ja mittariston ylläpito ja kehittäminen	22
7.1	Työkalun hallinnan roolit ja vastuut.....	23
7.2	Laatukriteerien elinkaari.....	24
7.3	Mittareiden elinkaari.....	24
7.4	Muita huomioita.....	25
8	Liitteet.....	25

1 Johdanto

Tämä on ehdotus julkishallinnon tiedon laadun kuvaukseen ja arviointiin tarkoitettusta 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. Tämä 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.

Tavoitteena on, että tästä laatukriteerikokonaisuudesta koostetaan julkisen hallinnon suositus vuoden 2021 lopulla syksyllä toteutettavien pilotointi- ja lausuntokierrosten tulosten huomioimisen jälkeen. Tällöin valitaan myös suosituksen rakenne ja päätetään rajanveto suosituksen ja sitä tukevan ohjeistuksen välillä esimerkiksi mittariston suhteen.

Laatukriteerien ja mittariston kehittäminen on osa Tiedon laatukehikkoa, jota kehitetään Valtiovarainministeriön asettaman Tiedon hyödyntämisen ja avaamisen hankkeen (TiHA, VN/5386/2020) työpaketissa kolme. TiHA-hankkeen tavoitteena on edistää julkisen tiedon entistä laajempaa ja tehokkaampaa hyödyntämistä koko yhteiskunnassa. Hanke laatii ehdotuksen tiedon hyödyntämisen ja avaamisen strategisista tavoitteista julkiselle hallinnolle sekä valmistelee ja toteuttaa tiedon hyödyntämistä ja avaamista edistäviä toimenpiteitä. Hankkeen toimikausi on 30.4.2020–31.12.2022.

Tiedon laatu on monitahoinen käsite. Laatukriteerien määrittelyssä on tavoiteltu kattavaa mutta tiivistä kokoelmaa tiedon laadun eri näkökulmat huomioon ottaen. Tavoitteeseen pääsemisessä keskeisessä roolissa ovat olleet projektiin osallistuneet henkilöt, jotka edustavat laajasti erilaisia tiedon suunnittelun, käsittelyn, luovutuksen, analysoinnin ja laadun tulokulmia. Laatukehikon luomiseen osallistuneet tahot ovat yhdessä keskustellen hakeneet näkemystä siihen, mitä laatu julkishallinnon tietoaaineistojen osalta tarkoittaa ja mitä kaikkia osa-alueita tulee ottaa huomioon tiedon laatua kuvattaessa.

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. Kriteerit ovat myös linjassa Euroopan yhteentoimivuusperiaatteen (EIF), FAIR-periaatteiden ja Euroopan tilastojen käytännesääntöjen kanssa.

Tiedon laatukriteerien soveltamiseen on tuotettu pilotoinnin kautta erillisten mittarien muodostama mittarikokoelma eli mittaristo. Mittariston tuottamisessa on vahvasti nojaututtu pilotointiin osallistuneiden osaamiseen. Pilotoinnin ensimmäinen vaihe tuotti laajan valikoiman mittariehdotuksia ns. mittarisammion. Näistä pilotoinnin toiseen vaiheeseen valikoitui osajoukko. Pilotoinnin toisessa vaiheessa mittaristoa työstettiin edelleen. Tuotetuista laadun mittareista osalla on selkeämpi, osalla löyhempi kytkös yleisiin standardeihin tai muualla käytössä oleviin mittareihin.

Kehittämistyöhön ovat pilotoijina osallistuneet Tilastokeskus, Maanmittauslaitos, Verohallinto, Tulli, Luonnonvarakeskus sekä Valtiokonttori. Hankkeen laajan projektiryhmän kautta työhön ovat osallistuneet edellä lueteltujen lisäksi Digi- ja väestötietovirasto, Kela, Opetushallitus, Patentti- ja rekisterihallitus ja Työterveyslaitos. Lisäksi laatukriteerejä on esitelty muun muassa TiHA-hankkeen

kunnille suunnatussa tilaisuudessa ja ne ovat olleet avoimesti kommentoitavina Tilastokeskuksen sivuilla ennen pilotointia. Tässä dokumentissa on esitelty myös alustava luonnos laatukriteerien ja mittariston ylläpidosta ja kehittämisestä, joka on tuotettu pilotoinnin rinnalla TiHA TP3 -projektiryhmässä, pilottien kokemuksista myös ammentaen.

2 Tiedon laadun kuvaaminen

2.1 Tavoitteet

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. Sisällöltään laadukkaan tiedon lisäksi kriteerien valinnan taustalla ovat vaikuttaneet tavoitteet tietojen helposta löydettävyydestä, yhdisteltävyydestä, yhteentoimivuudesta sekä sujuvasta käyttökokemuksesta.

Laatukriteerien ja mittariston suunnittelussa on laatua pyritty kuvaamaan ymmärrettävästi siten, että aineiston soveltuvuuden tarkastelu onnistuu myös ilman aiempaa kokemusta kyseisestä tietoaineistosta tai tiedon laadun määrittämiseen liittyvää erityisosaamista. Tavoitteena on ollut myös, että kynnys laatukriteerien ja mittariston käyttöönottoon olisi mahdollisimman matala. Tämä tavoite on sidoksissa kehittämistyön lisäksi tavoitteeseen luoda jatkossa selkeä ohjeistus ja käyttöönottosuunnitelma laatukriteerien ja mittariston muodostamalle kokonaisuudelle.

2.2 Tunnistetut hyödyt

Edellisessä kappaleessa kuvatut tavoitteet helpottavat tiedon soveltuvuuden arviointia eri käyttötarkoituksiin ja tukevat siten osaltaan julkishallinnon tietovarantojen nykyistä monipuolisempaa käyttöä. Näin laatukriteerit ja mittaristo yhdessä tarjoavat työkalun, joka osaltaan tukee tiedon avaamista, yhteentoimivuutta ja hyödyntämistä.

Työn aikana on tunnistettu myös muita hyötyjä kuten tiedon laadun seuraaminen ja parantaminen sekä prosessien kehittymisen ohjaaminen, kun mittareilla saadaan tuotettua ajassa vertailtavaa tietoa tietoaineistojen laadusta. Mittariston odotetaan helpottavan tiedon laadun viestintää myös organisaatioiden sisällä.

Yhdessä kehitettävien laatukriteerien ja mittariston myötä myös kansallinen tiedon laadun ymmärrys ja osaaminen kasvaa ja samalla rakennetaan yhteistä kieltä, jolla tiedon laatua kuvataan ja siitä keskustellaan. Tiedon laatu voidaan nyt myös määrittää yhteisesti esimerkiksi kehityshankkeissa ja lainsäädännön valmistelussa. Laatukriteerit on tarkoitettu julkishallinnon työkaluksi, mutta niitä voidaan hyödyntää myös muualla yhteiskunnassamme, kuten julkisissa hankinnoissa.

Pidemmillä aikavälillä kehittyvä laatukriteerien ja mittariston kokonaisuus voisi myös mahdollistaa julkishallinnon tiedon laadun kansallisen ohjauksen ja seurannan, mikäli se katsotaan tarpeelliseksi.

Vastaavaa mallia tiedon laadun kansalliseen määrittämiseen ei ole ollut. Kansainvälisesti verraten olemme aallon harjalla luomassa ensimmäisiä tiedon

laatua määritteleviä kansallisia malleja. Voimme siten osallistua omalla panoksellamme myös kansainväliseen keskusteluun ja kehittämiseen.

2.3 Rajaukset

Laatukriteerien ja mittariston kehittämisessä on tehty seuraavat rajaukset:

- Laatukriteereitä ja mittaristoa kehitetään ainoastaan rakenteiselle datalle.
- Kehitettävän mittariston ensisijainen tavoite on palvella erilaisia tiedonvaihdon tilanteita ja tiedon käyttäjistä puhuttaessa tarkastellaan lähinnä organisaatioiden ulkopuolisia käyttäjiä.
- Laatukriteerit ja mittaristo eivät ota kantaa tavoiteltaviin laatutasoihin.
- Laatutiedon viestintä- ja jakeluratkaisut eivät sisälly tähän kokonaisuuteen.

Tiedon laatuun panostaminen on keskeistä, mutta laatuajattelun täytyy kulkea käsi kädessä myös tietotuote- ja järjestelmäkehityksen sekä asiakaspalvelun kehittämisen kanssa. Tietovarantojen monipuolinen käyttö ja yhdisteltävyys edellyttää myös panostamista rajapintojen ja sähköisten järjestelmien käytettävyyteen ja palveluiden tunnettuuteen ja saatavuuteen.

2.4 Ymmärrys tiedon laadusta tietoaineistojen laatukriteerien kautta

Tiedon laatu on laaja ja moniulotteinen käsite. Tiedon laatu määritellään yleisesti sen perusteella, kuinka hyvin se soveltuu käyttäjän tarpeisiin. Nyt kehitetyt laatukriteerit ja mittarit tarkastelevat tiedon laadun eri ulottuvuuksia sidottuna ensisijaisesti julkishallinnon tiedonvaihdon eri tilanteisiin ja siihen, mitkä laadun osa-alueet ovat keskeisiä tiedon käyttäjän näkökulmasta.



Kuva 1. Tietoaineistojen laatukriteerit ryhmittäin. Tiedon laatua tarkastellessa katsottiin keskeisiksi näkökulmat, jotka kertovat käyttäjälle siitä, mitä tieto koskee ja miten hyvin tarkasteltava tieto kuvaa todellisuutta. Tiedon käytettävyyden näkökulmasta olennaisista on tietää miten tietoa voi käyttää ja miten hyvin tieto on kuvattu.

Laatukriteerien ja mittariston kehittämistyössä nousi keskeiseksi haasteeksi yhteisen ymmärryksen rakentaminen asiakkaiden tietotarpeiden ja tiedon tuottajien tietolupausten välille. Laatukriteerit auttavat tunnistamaan, mitkä tekijät ovat tiedon laadun näkökulmasta tärkeitä. Siinä käyttäjälle tunnistettiin olevan

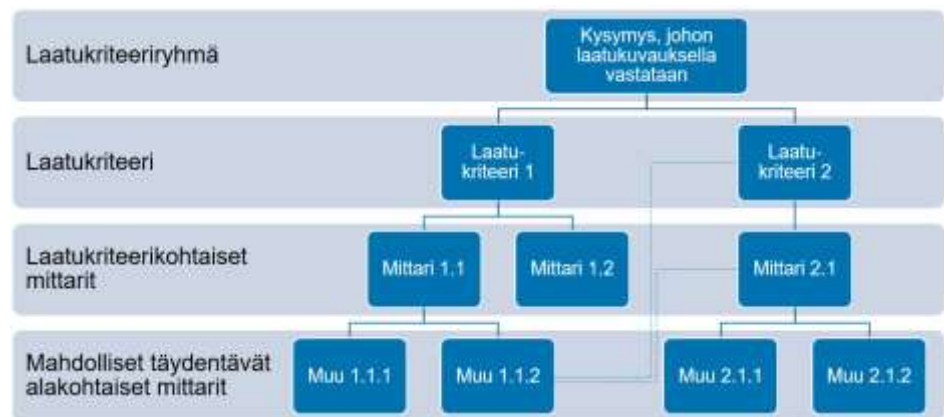
tärkeää saada kattava ja oikea kuva tietoaaineiston sisällöstä, sisällön ja sen kuvauksen laadusta sekä tiedon käyttömahdollisuuksista. Seuraavassa tarkastellaan näitä eri näkökulmia käyttäjälähtöisten kysymysten ja kysymyksiin vastaavien laatukriteerien kautta.

Laatukriteerit kuvattuna ryhmittäin. Ryhmiä on neljä, joissa jokaisessa on kahdesta neljään laatukriteeriä. Mitä tieto koskee? -ryhmässä ovat laatukriteerit jäljitettävyyden ja kattavuuden. Miten tieto kuvaa todellisuutta? -ryhmässä ovat ajantasaisuus, johdonmukaisuus, oikeellisuus ja tarkkuus. Miten hyvin tieto on kuvattu? -ryhmässä ovat suositustenmukaisuus ja ymmärrettävyys. Miten tietoa voi käyttää? -ryhmässä ovat koneluettavuus, käyttöoikeudet ja täsmällisyys.

Metatieto on tietoa, joka kuvailee toista tietojoukkoa. Tiedon laadun kuvaukset voidaan siten nähdä osana tietoaaineistojen (metatieto)kuvauksia. Laatukuvausten ja mittarien tuottamat arvot tarvitsevat tuekseen myös muita tietoaaineistojen kuvausten tietoelementtejä, mm. tavoitekohdejoukon kuvauksen ja sijainnin. Tällainen kattava kuvailutieto on olennaista, jotta laatumittarien antamia arvoja voi tulkita ja ymmärtää. Siksi tähän laatukuvauksen ensimmäiseen versioon on nähty tarpeelliseksi sisällyttää myös sellaisia kuvauslementtejä, jotka laajemmassa metatietomallissa, kuten Rekisteritiedon metatietomallissa tai organisaatioiden omissa metatietomalleissa saattavat sijoittua luontevasti jonkin toisen metatietoelementin kuin laadun alle. Esimerkkinä tällaisesta voisi ajatella joitakin ajantasaisuuden laatukriteeriin liitettyjä mittareita. Tämän lisäksi mittaristossa on mukana useampia mittareita, jotka peräänkuuluttavat keskeisimpien metatietojen määrittelyä ja saatavuutta. Vähintään nämä metatiedot tulisi tarjota tiedon käyttäjälle laatukuvauksen rinnalla.

Laatukuvauksen rakenteeksi valikoitui tiukka hierarkkinen malli (Kuva 3). Sitä pidettiin selkeimpänä ja mittariston tukevan näin parhaiten laatukriteerien ymmärtämistä. Alimmalle tasolle, mahdollisille alakohtaisille täydentäville mittareille haluttiin kuitenkin antaa enemmän vapausasteita. Laatukriteerien ja erityisesti mittariston työstämisessä tunnistettiin paljon organisaatio-, aineistotyyppi- ja alakohtaisia eroja. Mittariston peruskokonaisuudeksi pyrittiin kuitenkin koostamaan joukko sellaisia mittareita, jotka soveltuvat käyttöön mahdollisimman laajasti. Samalla kuitenkin tunnistettiin tarve täydentää laadun kuvaamista toimialakohtaisilla ja muilla mahdollisilla täydentävillä mittareilla.

Hierarkkisesta kuvaamistavasta huolimatta laatu, jota kriteerit ja mittarit kuvaavat, on kokonaisuus, jossa laadun eri elementit vaikuttavat toisiinsa. Laadun parannus jonkin laatukriteerin suhteen voi jopa heikentää toisen laatukriteerin avulla kuvattua tiedon laatua. Esimerkiksi jos tavoitellaan tietoaaineiston täydellistä kattavuutta tai ominaisuustietojen erityisen hyvää tarkkuutta niin tiedon ajantasaisuus yleensä heikkenee.



Kuva 2. Kaaviokuva laatuksivauksun rakenteesta.

Laatuksiteerit ja niihin liittyvät mittarit on esitelty tarkemmin laatuksiteeriryhmittäin kappaleissa 4 – 7. Lisäksi mittarit on kerätty taulukkoon Liitteeseen 2. Liitteeseen 1 kootun tämän dokumentin sanaston rooli on tärkeä ja kehitetään edelleen projektin aikana. Sanaston kehittämällä ja tarkalla sisällön kuvauksella päästään lähemmäs tilannetta, jossa asioista puhutaan samoilla käsitteillä ja ne ymmärretään mahdollisimman samalla tavalla. Yhteisymmärrys on tärkeää laatuksiteerien ja mittarien jatkokehittämisen sekä laadun vertailemisen kannalta.

2.5 Huomio keskeisiin tekijöihin mittariston tuella

Laatuksiteereitä täydentävä mittaristo tukee laatuksiteerien ymmärtämistä ja tiedon laadun tunnistamista luomalla kriteereille konkreettisen sisällön. Mittarien määrä vaihtelee kriteerien välillä, kun osassa kriteereistä on päästy suhteellisen helposti laadun kuvaamisen ytimeen ja toisissa on koettu tarvetta lähestyä kuvaamista useammasta suunnasta. Mittaristo sisältää sekä määrällisiä mittareita että tiedon sisällön ja kuvauksen toteavia mittareita (kyllä/ei). Toteavien mittarien osalta oletetaan, että lisätieto mittarin aiheesta on tiedon käyttäjän löydettävissä muusta tietoaineiston dokumentaatiosta. Tavoitteena on, että mittaristo tarjoaa tiedon käyttäjälle mahdollisuuden tarkastella yhteismitallisesti julkishallinnon tietoaineistojen laatua.

Mittarien ja tiedon laadun kuvauksen elementtien tunnistamisessa on huomioitu erityisesti seuraavat näkökulmat:

- Mittarit ja tiedon laadun elementit kuvaavat yksityiskohtaisemmin laatuksiteerien mukaista laatuksinäkökulmaa käsiteltävästä tietoaineistosta.
- Mittari ja tiedon laadun elementit kuvaavat laatua ja erityisesti laatuksiteeriä selkeästi, yksikäsitteisesti ja ymmärrettävästi.
- Mittarien ja tiedon laadun elementtien tulee olla yhteisiä, kaikille käyttäjille soveltuvia ja hyvin erilaisista tietoaineistoista kohtalaisen helposti tuotettavia.
- Jokaisella laatuksiteerillä on vähintään yksi tiedon mittari tai laadun elementti.
- Kullekin mittarille on sovittu yksi yhteinen suositeltu esittämismuoto tai tunnusluku.

2.5.1 Mittariston soveltaminen

Seuraavissa kappaleissa on kuvattu laatukriteerit sekä kuhunkin kriteeriin liittyvä mittaristo. Tämän dokumentin kuvaukset ovat tiiviit ja esittelevät, mutta kertovat mittareistakin keskeisimmät soveltamisen näkökulmat, joita esimerkit havainnollistavat. Mittareille ja laadun elementeille tuotetaan vielä täsmällisemmät käyttöönotto- ja soveltamisohjeet. Soveltamisohjeilla pyritään tukemaan mittarin arvon tuottamista siten, että saavutetaan keskinäinen vertailukelpoisuus. Näiden lisäksi mittariston soveltamisessa tulee huomioida seuraavat yleiset periaatteet ja neuvot:

- Laatukriteerit ja mittaristo on tarkoitettu joustavaksi työkaluksi; kaikki kriteerit ja siten mittarit eivät välttämättä ole merkityksellisiä kaikissa tilanteissa. Laatuä suositellaan kuitenkin tarkasteltavan tuotetun kokonaisuuden kautta mahdollisesti uusistakin näkökulmista ja tiedon käyttäjän tarpeista lähtien.
- Tiedon käyttäjät eli asiakkaan tunnistaminen: Eri hallinnonaloilla on erilaisia asiakkaita. Pohdi keitä tietoaineiston asiakkaat ovat ja millä perusteilla asiakasnäkökulmaa voidaan tuoda laadun tarkasteluun? Onko tunnistettavissa asiakasryhmiä, joilla on erityisiä laatuun liittyviä tarpeita?
- Mittariston soveltaminen koostettuun tietoon toteutetaan sillä tasolla millä tietoja on kuvattu. Esimerkiksi jos yritystiedot on summattu toimialatasolle, niin puuttuvien kohdeyksiköiden osuus -mittaria sovelletaan puuttuvien toimialojen tasolla.
- Mikäli tarkan kuvauksen, esim. numeroarvon tuottaminen on hankalaa, voi määrittelyssä käyttää myös arviota. Mieluummin siis arvio kuin jätetään kokonaan vastaamatta tarkasteltavan tiedon laadun kannalta merkitykselliseen mittariin.
- Laatukriteerit ja mittaristo ovat tiivis kokonaisuus ja eri elementit linkittyvät toisiinsa siten, että laadun parantuminen yhtäällä voi merkitä heikkenemistä toisaalla. Esimerkiksi jos tavoitellaan tietoaineiston täydellistä kattavuutta tai ominaisuustietojen erityisen hyvää tarkkuutta niin tiedon ajantasaisuus yleensä heikkenee.
- Mittaristoa täydentävien mittarien käyttö on sallittua ja jopa suotavaa.

3 Laatukriteeriryhmä: Mitä tieto koskee?

Tiedon käytettävyyttä ja laatua arvioidessa tarvitaan tietoa siitä, mitä ilmiötä tiedolla on tarkoitus kuvata eli mitä tieto koskee. Tietoaineiston aineistokuvaus sisältää yleensä tämän tiedon, mutta tiedon laatuksessa on haluttu korostaa keskeisimmät tarkasteltavat näkökulmat.

Tiedon kattavuuden kautta kuvataan sitä kohdejoukkoa, jota tietoaineiston on tarkoitus kuvata ja sitä, miten hyvin tavoiteltu kohdejoukko on tietoaineistossa mukana. Toisaalta tiedon kattavuuden voidaan ajatella kertovan myös siitä, miten hyvin tavoiteltu aineiston kattavuus on saavutettu sekä kohdeyksiköiden että ominaisuuksien näkökulmasta. Jäljitettävyys taas kertoo siitä, onko tiedon alkuperä ja muutokset tiedossa ja jälkikäteen tarkasteltavissa.

Mitä tieto koskee?



Kuva 3. Laatuksiteerien mittarit kysymyksen ”Mitä tieto koskee?” alla.

3.1 Laatuksiteeri: Jäljitettävyys

Synonyymit: *kiistämättömyys*

Kuvaus: Jäljitettävyys kertoo siitä, että tietoaineistoon ja sen tietoihin tehdyt muutokset voidaan jäljittää. Tiedon alkuperä tunnetaan.

Esimerkit: Tiedon alkuperä ja muutostapahtumat on kuvattu, muutosten aikaleimat ovat saatavilla. Tieto voidaan todistaa kiistämättömäksi, tietoaineiston tiedot voidaan todentaa uudelleen.

3.1.1 Muutosten hallinta

Muutosten hallinta -mittari kertoo, että tietoaineistossa tai ominaisuustiedoissa tapahtuvia rakenteellisia tai lähdetietojen muutoksia seurataan.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: kyllä/ei

Mittarin arvo kertoo siitä, tallennetaanko tieto tietoaineiston tai ominaisuustietojen muutoksista.

Lisäksi tulee huomioida, että silloin kun muutoksia tallennetaan, on hyvä tuoda esille, miten muutoshistoriaa on mahdollista päästä käsittelemään.

Esimerkit: Tietoaineiston muutoksista pidetään lokia tai tieto on löydettävissä metatietojen kautta.

3.1.2 Tiedon elinkaari

Tiedon elinkaari -mittari kertoo, onko tietoaineiston elinkaari määritelty ja kuvaus siitä saatavilla.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: kyllä/ei

Mittarin arvo kertoo, onko tiedon elinkaari määritelty ja kuvattu. Tiedon elinkaari kuvaa tiedon muutokset luomisesta tiedon poistamiseen ja se sisältää tiedon lähteiden tai tuottamisen muutokset, esimerkiksi laskentatavan muutokset.

3.1.3 Tietolähde

Tietolähde-mittari kertoo kuinka suurella osalla kohdeyksiköitä tai ominaisuustietoja lähdetiedot ovat saatavilla.

- Arviointitaso: ominaisuus ja tietoaineisto

- Arvon esitysmuoto: prosenttia

Mittaria soveltaessa tulee huomioida, että mittarin arvo ilmoitetaan saatavilla olevien lähdetietojen suhteena kaikkiin tietoihin. Tarkastelussa keskitytään tarkastelemaan tietolähdettä, josta tiedot on suoraan saatu laatua arvioivalle taholle. Mahdolliset aiemmat tiedon vaiheet on tällöin mahdollista selvittää tiedon siirtymisen ketjuna.

Lisäksi tulee huomioida, että tiedon käyttäjää kiinnostaa tietää tietolähteet ja miltä osin lähdetietoja puuttuu. Tietolähteen osalta on olennaisinta kertoa edellinen taho, joka tiedon on toimittanut, jolloin haluttaessa laatutarkastelussa voidaan ottaa huomioon ketjumaisesti tiedon laatuun liittyviä näkökulmia eri toimijoiden osalta. Jos tietolähteen vaillinaisuus koskee esimerkiksi tiettyä kohdeyksikköjoukkoa, olisi tietoaineiston kuvaustiedoissa hyvä kuvata puutteita.

3.2 Laatuksiteeri: Kattavuus

Synonyymit: *Täydellisyys*

Kuvaus: Kattavuus kuvaa tietoaineiston tavoitellun ajallisen ja alueellisen kattavuuden sekä tavoitellut kohdeyksiköt ja ominaisuustiedot. Toisaalta kattavuus kertoo miltä osin tietoaineisto sisältää tavoiteltuja tietoja.

Esimerkit: Tietoaineisto kattaa määritellyn tarkastelualueen yksiköt, esimerkiksi kaikki Suomen yritykset. Alueellinen kattavuus kertoo, ovatko kaikki käsiteltävät alueet mukana, esimerkiksi kaikki Suomen kunnat tai kattaako tietoaineisto Ahvenanmaan. Ylipeitto kertoo, että tietoaineistossa on siihen kuulumattomia yksiköitä. Alipeitto kertoo, että tietoaineistosta puuttuu siihen kuuluvia yksiköitä. Vastuskato on alipeittoa. Toisaalta kattavuus kertoo myös sisältääkö tietoaineisto kaikki määritellyt ominaisuustiedot tietoaineistossa oleville kohdeyksiköille, esimerkiksi tietoaineistossa olevien Suomen kuntien väkiluku- ja pinta-aliatiedot tai onko kaikilla yrityksillä tiedossa esimerkiksi osoite- tai liikevaihtotieto.

3.2.1 Ajallinen tavoitekattavuus

Ajallinen tavoitekattavuus -mittari kertoo, että tietoaineiston tavoiteltava ajallinen kattavuus ja tiheys on kuvattu.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kyllä/ei

Mittarin arvo kertoo siitä, onko tietoaineiston ajallinen tavoitekattavuus kuvattu esimerkiksi aineistokuvauksessa.

3.2.2 Alueellinen tavoitekattavuus

Alueellinen tavoitekattavuus -mittari kertoo, että tietoaineiston tavoiteltava alueellinen kattavuus ja tiheys on kuvattu.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kyllä/ei

Mittarin arvo kertoo siitä, onko tietoaineiston alueellinen tavoitekattavuus kuvattu esimerkiksi aineistokuvauksessa.

3.2.3 Puutteelliset kohdeyksiköt

Puutteelliset kohdeyksiköt -mittari kuvaa tietoaineiston puuttuvia ominaisuustietoja omaavien kohdeyksiköiden osuuden kaikista kohdeyksiköistä.

- Arviointitaso: ominaisuus ja tietoaineisto

- Arvon esitysmuoto: prosenttia

Mittaria soveltaessa tulee huomioida, että mittarin arvo ilmoitetaan yhdenkin puuttuvan ominaisuustiedon omaavien kohdeyksiköiden lukumäärän suhteena kaikkien kohdeyksiköiden lukumäärään. Rakenteellista puuttuvuutta ei oteta tämän mittarin kohdalla huomioon, vaan tarkastellaan puuttuvuutta silloin, kun jokin arvo on kyseiselle kohdeyksikölle relevantti.

Lisäksi tulee huomioida, että rakenteellista puuttuvuutta on tärkeä kuvata tarkemmin tiedon käyttäjälle tiedon muussa kuvauksessa.

3.2.4 Puutteelliset ominaisuustiedot

Puutteelliset ominaisuustiedot -mittari kuvaa kuinka suuri osuus kohdeyksiköistä saa puuttuvan arvon ominaisuuden suhteen tarkasteltuna. Mittari kuvaa sitä, kuinka kattavasti tietoaaineistossa on arvoja tietyille ominaisuudelle.

- Arviointitaso: ominaisuus ja tietoaaineisto
- Arvon esitysmuoto: prosenttia

Mittaria soveltaessa tulee huomioida, että mittarin arvo ilmaistaa puuttuvan ominaisuustiedon omaavien kohdeyksiköiden lukumäärän suhteena kaikkien kohdeyksiköiden lukumäärään ominaisuuksittain tarkasteltuna. Rakenteellista puuttuvuutta ei oteta tämän mittarin kohdalla huomioon, vaan tarkastellaan puuttuvuutta silloin, kun jokin arvo on kyseiselle kohdeyksikölle relevantti.

Lisäksi tulee huomioida, että rakenteellista puuttuvuutta on tärkeä kuvata tarkemmin tiedon käyttäjälle tiedon muussa kuvauksessa.

3.2.5 Puuttuvat kohdeyksiköt

Puuttuvat kohdeyksiköt -mittari kuvaa tietoaaineiston alipeittoa eli tietoaaineiston tavoitekohdejoukosta puuttuvien kohdeyksiköiden prosentuaalista osuutta.

- Arviointitaso: tietoaaineisto
- Arvon esitysmuoto: prosenttia
- Mittarin lähde: ISO 19157, Norja

Mittaria soveltaessa tulee huomioida, että arvo ilmoitetaan tietoaaineistosta (tavoiteperusjoukosta) puuttuvien kohdeyksiköiden lukumäärän suhteena tietoaaineistoon kuuluvien kohdeyksiköiden lukumäärään.

Lisäksi tulee huomioida, että valmiiksi koostetun tiedon, esim. tilasto, osalta puuttuvuus on yleensä korjattu tilastollisin menetelmin. Mikäli täsmällisen arvon laskeminen on haastavaa, voidaan antaa kokemukseen perustuva arvio prosenttiarvoksi. Mikäli tietoaaineiston puuttuvista kohdeyksiköistä on käsitys, niin puuttuvuutta olisi tärkeää kuvauksen tietoaaineiston kuvaukseen.

3.2.6 Tavoitekohdeyksiköt

Tavoitekohdeyksiköt-mittari kertoo siitä, että tietoaaineiston tavoitekohdeyksiköt kuvattu kertoen kohdejoukon rajauksista.

- Arviointitaso: tietoaaineisto
- Arvon esitysmuoto: kyllä/ei
- Mittarin lähde: ISO 19157, Norja

Mittaria soveltaessa tulee huomioida, että tavoitekohdeyksiköiden kuvauksessa on otettu huomioon olennaiset muutkin rajaukset kuin ajallisen ja alueellisen kattavuuden näkökulma.

Esimerkit: Tietoaineisto kattaa tiettyjen toimialojen yritykset.

3.2.7 Tavoiteominaisuudet

Tavoiteominaisuudet-mittari kertoo, onko tietoaineiston ilmiön kuvaamiseen tarvittavat ominaisuudet kuvattu.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kyllä/ei

Mittaria soveltaessa tulee huomioida, että tarkoituksena on kuvata ne ominaisuudet, jotka ovat keskeisiä tarkasteltavan ilmiön kannalta ja joiden osalta tietoa on pyritty keräämään tietoaineistoon mahdollisimman kattavasti.

3.2.8 Esimerkit: Ylimääräiset kohdeyksiköt

Ylimääräiset kohdeyksiköt -mittari kuvaa tietoaineiston ylipeittoa eli tietoaineistossa mukana olevien, mutta tietoaineiston tavoiteperusjoukkoon kuulumattomien ja siten ylimääräisten kohdeyksiköiden prosentuaalista osuutta.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: prosenttia
- Mittarin lähde: ISO 19157, Norja

Mittaria soveltaessa tulee huomioida, että arvo ilmoitetaan ylimääräisten kohdeyksiköiden lukumäärän suhteena tietoaineiston tavoiteltavien kohdeyksiköiden lukumäärään.

Lisäksi tulee huomioida, että valmiiksi koostetun, esim. tilaston, tiedon osalta ylipeitto on yleensä pyritty jo korjaamaan. Mikäli täsmällisen arvon laskeminen on haastavaa, voidaan antaa kokemukseen perustuva arvio prosenttiarvoksi. Mikäli tietoaineisto sisältää ylimääräisiä kohdeyksiköitä ja näitä on mahdollista kuvata, olisi tämä kuvaus tärkeää sisällyttää tietoaineiston kuvaukseen.

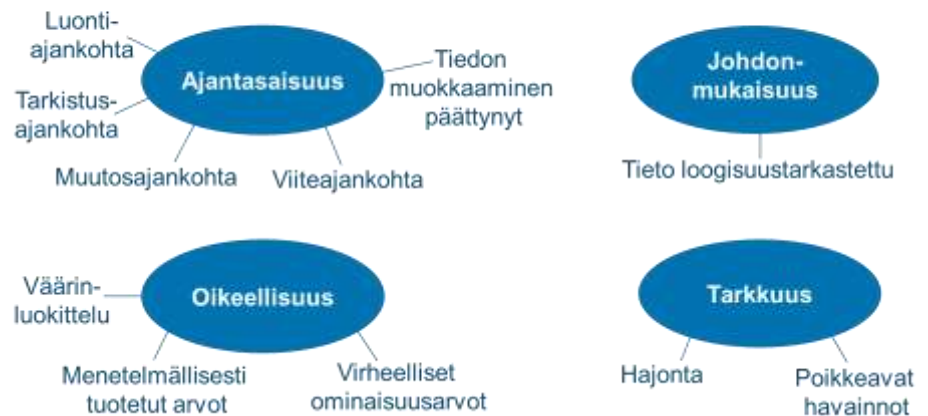
4 Laatuksiteeriryhmä: Miten tieto kuvaa todellisuutta?

Laadukkaan tiedon tulisi kuvata tavoiteltua ilmiötä mahdollisimman oikeellisesti, tarkasti ja ajantasaisesti. Ilmiö kuvataan yleensä tietomäärittelyjen avulla ja tämä laatuksiteeriryhmä kertoo miten hyvin se vastaa määriteltyä todellisuutta.

Ajantasaisuuden laatuksiteeri sisältää huomattavan määrän eri mittareita siksi, että tiedon ajantasaisuutta voidaan tarkastella monesta eri näkökulmasta. Tiedon tulisi olla mahdollisimman tuoretta suhteessa viiteajankohtaan, mutta toisaalta päivittämätönkään tieto ei välttämättä ole huonolaatuista, jos ominaisuusarvoissa ei ole tapahtunut muutoksia.

Tiedon tulisi kuvata todellisuutta mahdollisimman tarkasti ja oikeellisesti eli systemaattiset vääristymät tai muut virhelähteet tulisi tunnistaa ja niiden vaikutus korjata. Johdonmukainen tieto ei sisällä sisäisiä ristiriitoja.

Miten tieto kuvaa todellisuutta?



Kuva 4. Laatuksien mittarit kysymyksen ”Miten tieto kuvaa todellisuutta?” alla.

4.1 Laatuksien mittarit: Ajantasaisuus

Kuvaus: Ajantasaisuus kuvaa tietoaineiston tietojen täsmällisyyttä ajan suhteen. Ajantasaisuus on sitä parempi mitä lähempänä tiedon viiteajankohta on nykyhetkeä. Viiteajankohta on se ajankohta, jota tieto koskee.

Esimerkit: Tietoaineiston tietoon liittyvä viiteajankohta on kerrottu tiedon yhteydessä. Sen avulla pystytään päättämään tiedon tuoreus. Viiteajankohta voi olla esimerkiksi vuoden alun ja lopun välinen aika tai yhden tietyn päivän tilanne. Tiedon tuotannossa tarkastellaan erilaisia tiedon tarkastus- ja muutosajankohtia.

4.1.1 Luontiajankohta

Luontiajankohta-mittari kertoo päivämäärän, joka on kohdeyksikön tai ominaisuuden luontiajankohta.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: päivämäärä

Mittarin arvo kertoo ominaisuuden tai tietoaineiston luontiajankohdan. Mittari vastaa myös kysymykseen mistä päivämäärästä lähtien ominaisuus on ollut tietoaineistossa tai mistä päivämäärästä alkaen tietoaineistoa on koottu. Kohdeyksikön luontiajankohta ei välttämättä ole sama kuin tietoaineiston luontiajankohta.

Lisäksi kannattaa huomata, että luontiajankohta on tieto, joka on yleensä osa metatietoja.

4.1.2 Muutosajankohta

Muutosajankohta-mittari kertoo päivämäärän, joka on kohdeyksikön tai ominaisuuden muutosajankohta.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: päivämäärä; ei relevantti

Mittarin arvo kertoo ominaisuuden tai tietoaineiston muutosajankohdan. Mittari vastaa myös kysymykseen, milloin tietoja on viimeksi päivitetty.

Lisäksi kannattaa huomata, että muutosajankohta on tieto, joka on yleensä osa metatietoja.

4.1.3 Tarkistusajankohta

Tarkistusajankohta-mittari kertoo päivämäärän, joka on kohdeyksikön tai ominaisuuden tarkistusajankohta.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: päivämäärä; ei relevantti

Mittarin arvo kertoo ominaisuusarvojen tai tietoaineiston tietojen viimeisimmän tarkistusajankohdan. Mittari vastaa myös kysymykseen, milloin tietoja on viimeksi tarkistettu.

Lisäksi kannattaa huomata, että tarkistusajankohta on tieto, joka on yleensä osa metatietoja.

4.1.4 Tiedon muokkaaminen päättynyt

Tiedon muokkaaminen päättynyt -mittari kertoo päivämäärän, jonka jälkeen tietoja ei ole enää korjattu tai muuten muutettu.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: päivämäärä; ei relevantti

Mittarin arvo kertoo ajankohdan, jolloin tietoaineiston tietoja voidaan pitää lopullisina tietoina eikä niitä enää muuteta. Kerrottuun päivämäärään asti tietoja voidaan korjata virheiden tullessa ilmi.

Esimerkit: Tilastojen julkaisemisen lopullisen tilaston päivämäärä.

4.1.5 Viiteajankohta

Viiteajankohtaa -mittari kertoo ajankohdan, jota tietoaineisto kuvaa.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: päivämäärä; ei relevantti
- Mittarin lähde: Norja

Mittarin arvo kertoo, kuinka paljon aikaisempaan ajankohtaan liittyviin tapahtumiin tietoaineisto perustuu. Mittarin arvosta on mahdollista laskea tietoaineiston käsittelyn aiheuttaman viive tiedon julkaisuun tai käyttöön saamiseen.

Esimerkit: Tilaston viiteajankohta, joka voi olla useammankin kuukauden takainen, koska tietojenkeruu ja -käsittely vaatii aikansa ennen julkaisemista.

4.2 Laatuksiteeri: Johdonmukaisuus

Synonyymit: *säännönmukaisuus, tiedon looginen eheys*

Kuvaus: Johdonmukaisuus kertoo siitä, että tietoaineisto on yhtenäinen ja ristiriidaton. Johdonmukaisuudella voidaan kuvata myös eri tietoaineistojen keskinäistä johdonmukaisuutta.

Esimerkit: Esimerkiksi seuraavat tilanteet eivät ole johdonmukaisia: asuinrakennuksella ei ole yhtään asuntoa tai henkilön avioliiton solmimispäivä on aikaisempi kuin syntymäpäivä. Johdonmukaisuutta voidaan tarkistaa tarkistussäännöillä/laatusäännöillä.

4.2.1 Tieto loogisuustarkastettu

Tieto loogisuustarkastettu -mittari kuvaa sitä, onko aineistoa tarkasteltu loogisuusehdoin tietoaineiston koostamisessa tai käsittelyssä.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: kyllä/ei

Mittaria soveltaessa tulee kertoa, onko tietoja tarkistettu loogisuusehtoja käyttäen.

Lisäksi tulee huomioida, että erityisesti tietoaineistotasolla olisi hyvä kuvata tarkemmin miltä osin loogisuusehtoja on käytetty. Tiedon käyttäjää kiinnostaa myös yksityiskohdat käytetyistä loogisuusehdoista ja ne olisi hyvä tuoda tietoaineiston kuvauksessa esille.

Esimerkit: Ominaisuuden ”toimipaikan osoite” osalta on käytetty loogisuusehtoa, että vastaako toimipaikan sijaintiosoitteen postinumero toimipaikan sijaintikuntaa.

4.3 Laatuksiteeri: Oikeellisuus

Synonyymit: *virheettömyys*

Kuvaus: Oikeellisuus kuvaa sitä, miten tietoaineiston tiedot vastaavat todellisuutta. Tiedon oikeellisuutta tarkastelemalla voidaan saada kiinni myös systemaattisia vääristymiä tietoaineistossa.

Esimerkit: Se tieto, mikä operatiiviseen päätökseen päätyy, on käytännössä paras käsitys siitä, mikä tieto on oikea. Tieto on oikeellista esimerkiksi tilanteessa, jossa verotuksessa kerrottu palkka on sama kuin todellisuudessa maksettu palkka.

4.3.1 Menetelmällisesti tuotetut arvot

Menetelmällisesti tuotetut arvot -mittari kuvaa tietoaineiston menetelmällisesti tai korvaavalla tiedolla tuotettujen ominaisuusarvojen osuuden kaikista arvoista.

- Arviointitaso: ominaisuus
- Arvon esitysmuoto: prosenttia; ei relevantti

Mittarin arvo lasketaan menetelmällisesti eli laskennallisesti tai korvaavaa tietoa käyttäen tuotettujen ominaisuusarvojen lukumäärän suhteena ominaisuuden kaikkien arvojen lukumäärään. Tässä tarkoitetaan sellaisella korvaavalla tiedolla tai laskennallisella menetelmällä tuotettuja arvoja, jotka eivät ole täsmällisesti kohdeyksikön saamia oikeita arvoja. Tiedon korjaaminen tiedonantajalta selvittämällä saadulla oikealla arvolla ei kuulu tämän mittarin piiriin.

Esimerkit: Tulotiedon täydentäminen luovuttajaimputoinnilla käyttäen sukupuolta, ikää, koulutusta ja ammattia arvon luovuttajan valinnan kriteerinä.

4.3.2 Virheelliset ominaisuusarvot

Virheelliset ominaisuusarvot -mittari kertoo tietoaineiston virheellisiä ominaisuusarvoja saavien kohdeyksiköiden osuuden kaikista kohdeyksiköistä.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: prosenttia; ei relevantti
- Mittarin lähde: ISO 19157 (id 63 JHS 160)

Mittaria soveltaessa tulee huomioida, että arvo ilmoitetaan virheellisen ominaisuusarvon saavien kohdeyksiköiden lukumäärän suhteena kaikkien kohdeyksiköiden lukumäärään.

Lisäksi tulee huomioida, että valmiiksi koostetun tiedon osalta tunnistetut virheelliset arvot on yleensä jo korjattu. Mikäli mittarin täsmällisen arvon laskeminen on haastavaa, voidaan antaa kokemukseen perustuva arvio prosenttiarvoksi.

4.3.3 Väärinluokittelu

Väärinluokittelu-mittari kertoo tietoaineiston virheellisesti luokiteltujen kohdeyksiköiden osuuden kaikista kohdeyksiköistä.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: prosenttia; ei relevantti
- Mittarin lähde: ISO 19157 (id 63 JHS 160)

Mittaria soveltaessa tulee huomioida, että arvo ilmoitetaan väärinluokiteltujen kohdeyksiköiden lukumäärän suhteena kaikkien kohdeyksiköiden lukumäärään. Väärinluokitelluksi voidaan määritellä myös ne kohdeyksiköt, joille ei ole saatu korjattua puuttuvaa tietoa ja joiden puuttuvuus ei ole rakenteellista puuttuvuutta eli ominaisuus on kohdeyksikölle relevantti.

Lisäksi tulee huomioida, että valmiiksi koostetun tiedon osalta väärinluokittelu on yleensä jo valmiiksi korjattu. Mikäli täsmällisen arvon laskeminen on haastavaa, voidaan antaa kokemukseen perustuva arvio prosenttiarvoksi.

Esimerkit: Rakennuksen käyttötarkoitus on määritelty väärin n% kaikista rakennuksista.

4.4 Laatuksiteeri: Tarkkuus

Synonyymit: *harhattomuus*

Kuvaus: Tarkkuus kuvaa sitä, miten hyvin tietoaineiston tiedot vastaavat sitä mitä tavoitellaan. Tarkkuus kuvaa sitä, kuinka hyvin tieto osuu oikeaan.

Esimerkit: Esimerkkejä tarkkuudesta ovat mitattujen arvojen hajonta ja tietoaineiston poikkeavien havaintojen osuus sekä lisäksi luokittelun osuvuus ja mittaustaso esim. desimaalit, ajan tai koordinaattien mittaustaso.

4.4.1 Hajonta

Hajonta-mittari kuvaa kuinka laajalle alueelle ominaisuusarvot ovat jakautuneet.

- Arviointitaso: ominaisuus
- Arvon esitysmuoto: vaihteluväli; keskihajonta; muu hajontaa kuvaava tunnusluku

Mittaria soveltaessa on tarkoitus kuvata mahdollisimman selkeästi mille arvoalueelle ominaisuusarvot vaihtelevat.

Lisäksi tulee huomioida, että eri aloilla, esimerkiksi paikkatiedon osalta, tämän mittarin arvoille on yksityiskohtaisempia laskukaavoja.

4.4.2 Poikkeavat havainnot

Poikkeavat havainnot -mittari kuvaa tietoaineiston poikkeavien havaintojen osuutta kaikista kohdeyksiköistä.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: prosenttia
- Mittarin lähde: ISO 25024

Mittaria soveltaessa tulee huomioida, että mittarin arvo ilmoitetaan poikkeavien havaintojen lukumäärän suhteena kaikkien kohdeyksiköiden lukumäärään. Poikkeavalla havainnolla tarkoitetaan kohdeyksikköä, joka saa suuresti havaittujen arvojen enemmistöstä poikkeavan arvon. Poikkeava havainto voi olla aito oikea arvo tai virheellinen arvo.

Esimerkit: Tarkastellessa henkilöiden tulotietoja, hyvin suurituloiset ovat aineistossa poikkeavia havaintoja. Mikäli suurituloisimmat otetaan mukaan tulojen keskiarvon laskemiseen, voi saatu arvo olla huomattavasti suurempi kuin keskimääräisen aineistoon kuuluvan henkilön tulot ovat.

5 Laatuksiteeriryhmä: Miten hyvin tieto on kuvattu?

Laadukkainkin tietoaineisto ilman tiedon kuvausta voi olla pahimmillaan jopa käyttökeltvotonta. Tiedon käytettävyys vaatii aineistokuvauksen ja ominaisuuksien kuvaukset, jotta tietoaineistolle ja ominaisuuksien arvoille on mielekäs ja ymmärrettävä tulkinta. Tietojen yhteensovittamisen kannalta erilaisten suositusten: standardien, käytäntöjen ja säädösten noudattaminen on olennaista.

Miten hyvin tieto on kuvattu?



Kuva 5. Laatuksiteerien mittarit kysymyksen "Miten hyvin tieto on kuvattu?" alla.

5.1 Laatuksiteeri: Suositustenmukaisuus

Synonyymit: *yhteensovivuus, semanttinen yhdenmukaisuus, yhdenmukaisuus*

Kuvaus: Suositustenmukaisuus kertoo siitä, että tietoaineisto ja sen ominaisuustiedot noudattavat tunnettuja standardeja, käytäntöjä ja säädöksiä ja ne on kerrottu tietoaineiston yhteydessä.

Esimerkit: Kansallista yhdenmukaisuutta tukee muun muassa yhtenäisten kansallisten sanastojen ja koodistojen käyttö tietoaineistojen tietoja suunniteltaessa. Kansainvälistä yhdenmukaisuutta tukevat esimerkiksi EU:n vahvistamat standardiluokitukset ja ISO-kielikoodisto.

5.1.1 Käytetyt standardit

Käytetyt standardit -mittari kertoo ovatko tietoaineiston noudattamat standardit, käytännöt ja säädökset listattuna tietoaineiston kuvauksissa.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: kyllä/ei
- Mittarin lähde: INSPIRE / FAIR periaatteet

Mittaria soveltaessa tulee huomioida, että standardeja on monen tasoisia: yleisiä ja toimialakohtaisia sekä kansallisia ja kansainvälisiä. Myös ominaisuuksien osalta yleisten koodistojen käyttö tulee ottaa huomioon. On erittäin tärkeää, että noudatettavien suositusten lista on tiedon käyttäjän saatavilla. Myös tieto standardeista, joita ei noudateta, voi olla tiedon käyttäjälle hyödyllinen tieto.

Esimerkit: Ominaisuus noudattaa kansainvälistä koodistostandardia, mutta kyseistä koodistoa on täydennetty organisaatiokohtaisesti lisäluokilla. Lisäluokat

on määritelty koodistostandardista erilliseksi siten, että ne täsmentävät standardia mutteivat ole ristiriidassa sen kanssa.

5.1.2 Standardienmukaisuus

Standardienmukaisuus -mittari kertoo, noudattavatko tietoaineiston ja ominaisuuksien sisällöt ja koodistot yleisiä ja toimialakohtaisia standardeja, käytäntöjä ja säädöksiä.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: kyllä/ei

Mittaria soveltaessa tulee huomioida, että tavoitteena on antaa mahdollisimman realistinen arvio standardien mukaisuudesta suhteessa tietoaineiston kannalta olennaisiin standardeihin, käytäntöihin ja säädöksiin. Tarkastelussa tulee muistaa tiedon käyttäjän näkökulma.

5.2 Laatuksiteeri: Ymmärrettävyys

Synonyymit: *tulkittavuus, käsitettävyys*

Kuvaus: Ymmärrettävyys kuvaa sitä, miten kattavasti tietoaineistolla on sellaisia metatietoja, jotka auttavat tiedon ymmärtämisessä sitä käytettäessä.

Esimerkit: Tietoaineisto ja sen ominaisuustiedot on kuvattu metatietokuvauksessa riittävällä tasolla siten, että tietosisällön ja merkityksen ymmärtäminen helpottuu. Ominaisuustiedoissa käytetyt koodistot on kirjattu ja ne ovat datan kanssa yhteneväiset. Koodistojen kuvaukset ovat saatavilla esimerkiksi linkkien kautta. Olennaiset käsitteet on kuvattu ja linkit tarpeellisiin sanastoihin on liitetty metatietokuvauksiin.

5.2.1 Aineistokuvaus

Aineistokuvaus-mittari kertoo, onko tietoaineiston aineistokuvaus saatavilla ja mitä kieliversioita siitä on saatavilla.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kieliversiot

Mittaria soveltaessa tulee huomioida, että arvo ilmaistaan kertomalla ne kielet, joilla aineistokuvaus on saatavilla.

Lisäksi tulee huomioida, että kieliversion voi mainita vaikkei kaikki kieliversiot olisikaan yhtä kattavia. Aineistokuvauksen osalta myös määritelmät ja yhtenäinen termistö ovat tärkeitä – siten käyttäjä voi olla varma, että tietoaineistot ovat vertailukelpoisia.

5.2.2 Käsitteiden määritelmät

Käsitteiden määritelmät -mittari kertoo, onko tietoaineiston käsitteet määritelty selkeästi ja saatavilla sekä mitä kieliversioita niistä on saatavilla.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: kieliversiot

Mittaria soveltaessa tulee huomioida, että arvo ilmaistaan kertomalla ne kielet, joilla käsitteiden kuvaukset ovat saatavilla. Käsitteiden määrittelemisen selkeästi on tärkeä osa tiedon kuvausta, jotta voidaan varmistua siitä, että samoilla sanoilla ymmärretään sama asia.

Lisäksi tulee huomioida, että kieliversioiden voi mainita vaikei kaikki kieliversiot olisikaan yhtä kattavia.

5.2.3 Ominaisuuksien tietokuvaukset

Ominaisuuksien tietokuvaukset -mittari kertoo, onko tietoaineiston ominaisuuksien kuvaukset ja koodistot saatavilla ja mitä kieliversioita niistä on saatavilla.

- Arviointitaso: ominaisuus
- Arvon esitysmuoto: kieliversiot

Mittaria soveltaessa tulee huomioida, että arvo ilmaistaan kertomalla ne kielet, joilla kuvaus on saatavilla.

Lisäksi tulee huomioida, että kieliversioiden voi mainita vaikei kaikki kieliversiot olisikaan yhtä kattavia. Kuvauksien osalta myös määritelmät ja yhtenäinen termistö ovat tärkeitä – siten käyttäjä voi olla varma, että tietoaineistot ovat vertailukelpoisia.

5.2.4 Ymmärrettävyyden asiakaspalaute

Ymmärrettävyyden asiakaspalaute -mittari kertoo, että tietoaineiston ymmärrettävyydestä on mahdollisuus antaa palautetta olemassa olevan palauteväylän kautta tai kohdennetulla asiakaskyselyllä.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kyllä/ei

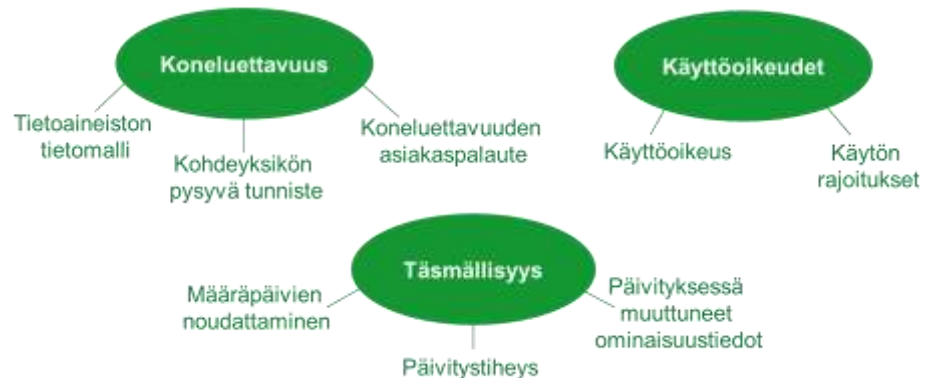
Mittaria soveltaessa tulee huomioida, että mittarin on tarkoitus kuvata sitä, onko palautetta mahdollista antaa ja että palautetta voidaan tarvittaessa hyödyntää tietoaineiston kehittämisessä.

Lisäksi tulee huomioida, että palautteen yhteenveto tai palautteen aiheuttamat muutokset ovat hyödyllistä lisätietoa tiedon käyttäjille.

6 Laatuksiteriryhmä: Miten tietoa voi käyttää?

Tietoaineistojen käytölle on rajoituksia sen suhteen mikä taho tietoja voi tarkastella, mihin käyttötarkoitukseen niiden käyttö on sallittu ja missä muodossa tiedot ovat saatavilla. Rajoitteita aiheuttavat esimerkiksi tiedonkeruussa tehty tiedonantajan informointi ja tietosuojat. Olennaista on myös varmistua, että tieto on saatavilla silloin, kun se on luvattu. Tiedon käytettävyyttä tarkastellaan myös erityisesti koneluettavuuden näkökulmasta.

Miten tietoa voi käyttää?



Kuva 6. Laatuksien mittarit kysymyksen ”Miten tietoa voi käyttää?” alla.

6.1 Laatuksien mittarit: Koneluettavuus

Kuvaus: Koneluettavuus kuvaa, onko tietoa rakenteistettu siten, että sitä voidaan käsitellä koneellisesti ja käsittely on mahdollista eri tietojärjestelmissä.

Esimerkit: Tietoaaineisto on rakenteisessa formaatissa, esim. csv, json, xml. Tietoaaineiston rakenne on kuvattu, esim. skeema.

6.1.1 Kohdeyksikön pysyvä tunniste

Kohdeyksikön pysyvä tunniste -mittari kertoo, että tietoaaineiston kohdeyksiköillä on vähintään tietoaaineistokohtainen pysyvä tunniste, jonka perusteella kohdeyksiköt voidaan erottaa toisistaan.

- Arviointitaso: tietoaaineisto
- Arvon esitysmuoto: kyllä/ei

Mittaria soveltaessa tulee huomioida, että tarkastelussa riittää, että pysyvä tunniste on vähintään tietoaaineistokohtainen. Luonnollisesti yhtenäinen ja pysyvä tunniste kansallisesti tai jopa kansainvälisesti olisi laajemman hyödynnettävyyden kannalta eduksi.

6.1.2 Koneluettavuuden asiakaspalaute

Koneluettavuuden asiakaspalaute -mittari kertoo, että tietoaaineiston koneluettavuudesta on mahdollisuus antaa palautetta olemassa olevan palauteväylän kautta tai kohdennetulla asiakaskyselyllä.

- Arviointitaso: tietoaaineisto
- Arvon esitysmuoto: kyllä/ei

Mittaria soveltaessa tulee huomioida, että mittarin on tarkoitus kuvata sitä, onko palautetta mahdollista antaa ja että palautetta voidaan tarvittaessa hyödyntää tietoaaineiston kehittämisessä.

Lisäksi tulee huomioida, että palautteen yhteenveto tai palautteen aiheuttamat muutokset ovat hyödyllistä lisätietoa tiedon käyttäjille.

6.1.3 Tietoaaineiston tietomalli

Tietoaaineiston tietomalli -mittari kertoo, onko tietoaaineisto kuvattu rakenteisesti.

- Arviointitaso: tietoaaineisto

- Arvon esitysmuoto: kyllä/ei

Mittarin arvo kertoo, onko aineiston rakenne kuvattu tietomallin/skeeman tai vastaavan standardin mukaisesti. Mikäli tietoaineisto on kuvattu tietomallin mukaisesti, se on koneluettava.

Lisäksi tietoaineiston tietomallin osalta olisi hyvä kertoa minkä tietomallin tai standardin mukaan tietoaineisto on kuvattu.

Esimerkit: Kansallisen maastotietokannan rakennustietojen rakenne on kuvattu JHS 210-standardissa. Tietoaineiston rakenne on kuvattu skeemalla (esim. xml, json).

6.2 Laatuksiteeri: Käyttöoikeudet

Kuvaus: Käyttöoikeudet kuvaa sitä, miten tietoaineiston käyttöoikeus on määritelty ja mitä aineistolla voi tehdä eli mihin käyttötarkoituksiin tietoaineistoa voi hyödyntää.

Esimerkit: Esimerkiksi tietoaineisto on tietyin rajoituksin käytettävissä tieteelliseen tutkimukseen. Avoin data on lisensoitu.

6.2.1 Käyttöoikeus

Käyttöoikeus-mittari kertoo, miten tietoaineiston käyttö on rajoitettu eli kenen käytettävissä tietoaineisto on.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: käyttöoikeuden rajaus

Mittarin arvo kertoo, kenen on mahdollista saada tietoaineisto käyttöönsä. Esimerkiksi avoindata, julkinen tieto, käyttöluvalla, sopimusperusteinen tai virkakäyttö. Käyttöä voi rajoittaa tietosuojanäkökulmat tai tiedonkeruussa tiedonantajalle annettu informaatio.

Esimerkit: Tilastotieto on julkista tietoa. Tilastokeskuksen yksikköaineistot ovat käyttöluvalla käytettävissä esimerkiksi tutkimustarkoituksessa.

6.2.2 Käytön rajoitukset

Käytön rajoitukset -mittari kertoo, mihin kaikkiin käyttötarkoituksiin tietoaineiston tietoa on luvallista käyttää.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: käyttötarkoituksen rajaus

Mittarin arvo kertoo, millaiseen käyttötarkoitukseen tietoaineistoa saa käyttää. Esimerkiksi virkakäyttö päätöksentekoon, virkakäytössä referenssinä, käyttöluvan tai sopimusehtojen rajoissa tai avoimen datan lisenssi.

Lisäksi olisi hyvä lyhyesti kuvata mahdollisia rajoituksia tiedon käytön osalta.

Esimerkit: Tilastokeskuksen yksikköaineistoista saa tuottaa anonymoituja, summattuja tutkimustuloksia. Julkaistua tietoa saa käyttää kunhan lähde mainitaan.

6.3 Laatuksiteeri: Täsmällisyys

Synonyymit: *Oikea-aikaisuus*

Kuvaus: Täsmällisyys tarkoittaa sitä, että tietoaineisto on käytettävissä ilmoitettuna ajankohtana ja riittävän tiheästi tietoaineistossa tapahtuviin muutoksiin nähden.

Esimerkit: Tietojen julkistamisaika ja julkistamistiheys on ilmoitettu. Julkistamisaikataulujen muutoksista kerrotaan etukäteen.

6.3.1 Määräpäivien noudattaminen

Määräpäivien noudattaminen -mittari kuvaa tietoaineiston suunnitellun toimitusaikataulun seuranta toteutuneeseen toimitusaikatauluun.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: toimitusviive; ei relevantti
- Mittarin lähde: Euroopan tilastojen käytännesäännöt

Mittaria soveltaessa ilmoitetaan toimitetun aineiston viive suhteessa sovittuun määräpäivään eli ”toteutunut toimituksen päivämäärä – sovittu päivämäärä”.

Lisäksi viivetilanteessa tiedon käyttäjälle voidaan lisätietona kuvata viiveen syy.

6.3.2 Päivitystiheys

Päivitystiheys-mittari kuvaa tietoaineiston säännöllistä päivitystiheyttä.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: sanallinen ilmaus
- Mittarin lähde: ISO 19139 MDMaintenanceFrequencyCode ja maintenanceNote

Mittarin arvon tulee kuvata säännöllisen päivityksen tiheyttä. Annettavat ilmaukset esim. reaaliaikainen, jatkuva, viikoittain, kuukausittain, kerran vuodessa tai muu päivitysrytmi.

Esimerkit: Tiestöä ja nimistöä päivitetään **jatkuvasti**. Hallintorajoja ja rakennuksia päivitetään **vuosittain**. Muut kohteet päivitetään karttalehdittäin määräaikaisen ajantasaistusprosessin yhteydessä **5-10 vuoden välein**.

6.3.3 Päivityksessä muuttuneet ominaisuustiedot

Päivityksessä muuttuneet ominaisuustiedot -mittari kuvaa muuttuneiden ominaisuusarvojen osuutta kaikista tietoaineiston ominaisuusarvoista verrattaessa päivitettyä tietoaineistoa edelliseen versioon.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: prosenttia

Mittarin on tarkoitus kertoa siitä, kuinka suuri vaikutus päivityksen odottamisella on mahdollisesti jo käytössä olevan tietoaineiston sisältöön.

7 Laatuksiteerien ja mittariston ylläpito ja kehittäminen

Tässä on kuvattu alustavat ajatukset siitä, miten laatuksiteerien ja niihin liittyvän mittariston ylläpito ja hallinta voitaisiin toteuttaa. Tavoitteena on, että laatuksiteerien ja mittariston muodostama kokonaisuus kehittyy ja muokkautuu kokemuksen ja tarpeiden mukaisesti. Tämä vaatii selkeät vastuut työkalusta huolehtimiselle. Koska laatuksiteerit, ja erityisesti mittaristo, on ajassa kehittyvä työkalu, tulee myös niiden elinkaari suunnitella. Ylläpito vaatii myös fyysisen järjestelmän, josta käyttäjät voivat selailun lisäksi hakea laatuksiteerit ja mittarit omaan käyttöönsä rajapinnan kautta.

Tässä esitetty kuvaus koskee vain ja ainoastaan laatuksiteerit ja mittaristo - työkalun ylläpitoa ja hallintaa. Nyt on jo tunnistettu, että voisi olla hyödyllistä tarkastella kokonaisuutena myös julkishallinnon tiedon laadun edistämisen

toimijoita ja toimintamalleja. Vahva verkostotyö voisi toimia tiedon laadun edistämisen ja osaamisen jakamisen vauhdittajana.

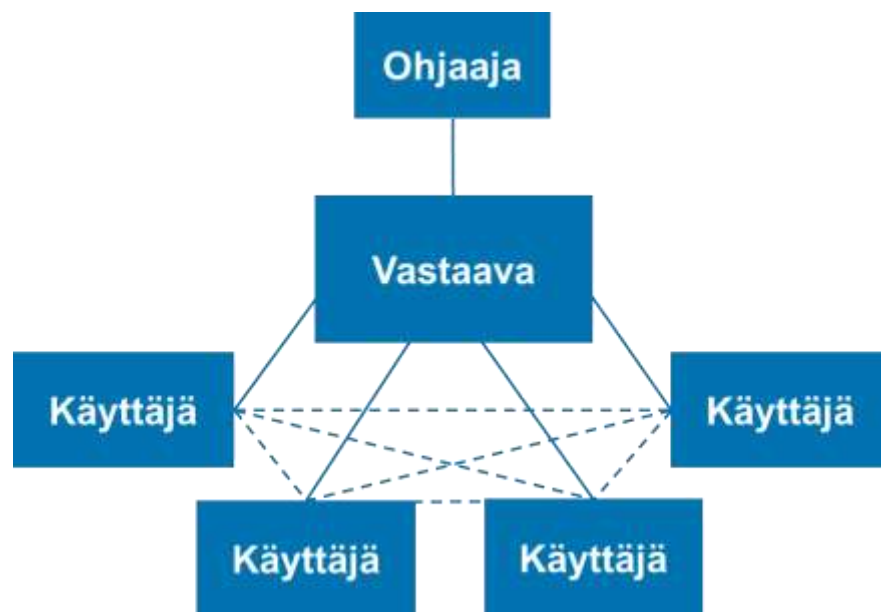
Tässä yhteydessä ei oteta kantaa yksittäisen organisaation tiedon laadun kuvauksen tai hallinnan järjestämiseen. Tiedon laadun tavoitteiden asettaminen ja sen yhteydet laadunparannuksen toimenpiteisiin rajautuu työkalun hallinnan kuvauksen ulkopuolelle.

Alla esitetty kuvaus on suppea alustava hahmotelma. Vastuutoimijoiden nimeämisen jälkeen sitä on tarpeen täydentää. Laatuksiteereitä ja mittareita voidaan hyödyntää ilman tietojärjestelmää, jossa niitä ylläpidetään, mutta täysimääräinen hyödyntäminen ja jakokehityksen turvaaminen edellyttäne jatkossa ylläpidon järjestelmän rakentamista.

7.1 Työkalun hallinnan roolit ja vastuut

Laatukriteerit ja mittaristo -työkalun hallinnassa on tunnistettu kolme erilaista roolia. Toimijoita, jotka huolehtivat näistä rooleista, ei ole vielä nimetty.

Laatukriteerit ja mittaristo on koko julkishallinnon yhteinen työkalu. Suositus hyödyntää tätä työkalua tiedon laadun kuvauksessa ja arvioinnissa koskee aikanaan julkishallinnon toimijoita. Siten Käyttäjät tarkoittavat tässä julkishallinnon organisaatioita ja muita julkishallinnon toimijoita. Myös yksityisen sektori voi hyödyntää laatukriteereitä ja mittareita; lähtökohtaisesti he eivät kuitenkaan ole alla mainittuja Käyttäjiä, jotka osallistuvat työkalun kehittämiseen.



Kuva 7. Laatukriteerit ja mittaristo -työkalun hallinnassa tunnistetut roolit.

Laatukriteerein ja mittariston hallinnan rooleja ovat:

- Ohjaaja, joka vastaa laatukriteerikokonaisuuden ylläpidon ja kehittämisen arvioinnista ja ohjauksesta.
- Vastaava, joka vastaa laatukriteeri- ja mittaristo -työkalun ylläpidosta, saatavilla olosta sekä soveltamisen tuesta ja seurannasta. Vastaava huolehtii myös mittariehdotuksia sisältävän mittarisammion hallinnoinnista, ja ylläpidosta ja kehittämisestä. Vastaava huolehtii yhteistyössä verkoston kanssa työkalun linjauksista ja kehittämisestä.

Vastaava viestii Käyttäjille laatukriteerien ja mittariston päivityksistä ja raportoi päivitysprosessin etenemisestä ja päivityksistä Ohjaajalle.

- Käyttäjä, joka vastaa laatukriteerien ja mittareiden käytöstä omassa organisaatiossaan sekä palautteen ja käyttökokemuksensa jakamisesta. Käyttäjä osallistuu päivityksiin oman kriteeristön ja mittariston käyttökokemuksensa pohjalta.

Laatukriteerien ja mittariston käyttökokemus on tärkein laatukriteerien kehittämiseen vaikuttava tekijä ja tässä Käyttäjien rooli on kehityksessä siksi merkittävä. Vastaava ja Käyttäjä toimivat tiiviissä verkostoyhteistyössä laatukriteerien kehittämisen ja soveltamisen onnistumisen varmistamiseksi.

Edellä mainitut roolit on tunnistettu ylätasolla. Osallistuvissa organisaatioissa voi olla tarpeen tunnistaa erilaisia rooleja myös organisaation sisällä; edellä esitetty kolmijako voi sopia myös tähän. Myös verkostoja ja ryhmiä voi olla eri tasoisia ja ne voivat toteuttaa eri tehtäviä.

7.2 Laatukriteerien elinkaari

Laatukriteerien ja mittariston selkeä elinkaari on tärkeää, jotta laatukriteereitä ja mittareita voidaan käyttää ja kehittää yhteisesti ja yhtenäisesti.

Laatukriteerien elinkaareissa on tässä vaiheessa tunnistettu kolme vaihetta:

- ehdotettu kriteeri
- hyväksytty kriteeri
- passivoitu kriteeri

Ehdotuksia voivat luoda sekä Vastaava että Käyttäjät. Niitä voi nousta käyttökokemuksen lisäksi myös toimintaympäristön seurannasta tai taustastandardeista.

Kriteerien hyväksynnän tekee Vastaava yhdessä verkoston kanssa myöhemmin määriteltävin toimintatavoin. Hyväksytyn laatukriteerin merkitys ei muutu, mutta sen kuvaustietoihin voidaan tehdä pieniä muutoksia, kuten lisätä esimerkkejä. Hyväksytyn kriteerin pysyvyys on tärkeää, sillä se on jo käytössä.

Kriteeri on periaatteessa pysyvä, mutta mikäli se koetaan käyttökelvottomaksi voidaan se passivoida, jolloin siihen liittyvät mittarit passivoituvat samalla. Passivointi tulee olla perusteltu. Päätöksen tekee Vastaava yhdessä verkoston kanssa myöhemmin määriteltävin toimintatavoin. Laatukriteeriä ei kuitenkaan poisteta kriteeristöstä, vaan passivoinnista huolehditaan esimerkiksi voimassaoloajoin. Tämä siksi, että kriteeri on jo ollut käytössä ja voi edelleen olla käytössä osalla Käyttäjistä.

7.3 Mittareiden elinkaari

Mittareiden elinkaareissa on tässä vaiheessa tunnistettu kolme vaihetta:

- ehdotettu mittari
- hyväksytty mittari
- passivoitu mittari

Vastaava ja Käyttäjät voivat ehdottaa uusia mittareita käyttökokemukseen ja toimintaympäristönmuutoksiin ja standardeihin pohjautuen.

Mittarin hyväksyy Vastaava yhdessä verkoston kanssa myöhemmin määriteltävin toimintatavoin. Hyväksytty mittari voidaan ottaa käyttöön ja siksi sen

pysyvyyteen tulisi tähdätä. Hyväksytyin mittarin kuvaustietoja voi päivittää. Mikäli mittarin olennainen sisältö muuttuu, tulee rinnalle luoda uusi mittari tai uusi mittariversio.

Mittari voidaan passiivoida, mikäli se koetaan toimimattomaksi. Passivointi tulee olla perusteltu. Päätöksen tekee Vastaava yhdessä verkoston kanssa myöhemmin määriteltävin toimintatavoin. Mittaria ei kuitenkaan poisteta mittaristosta, vaan passivoinnista huolehditaan esimerkiksi voimassaoloajoin. Tämä siksi, että mittari voi jo olla käytössä osalla Käyttäjistä.

7.4 Muita huomioita

Mittarisammio on laatukriteerien ja mittareiden pilotoinnissa hyväksi havaittu työkalu, johon koostettiin mittariehdotuksia. Mittarisammion kaltaista työkalua voitaisiin hyödyntää mm. mittariehdotusten keräämisessä ja hallinnoinnissa jatkossa. Mittarisammion ylläpito ja kehittäminen olisi tällöin Vastaavan vastuulla.

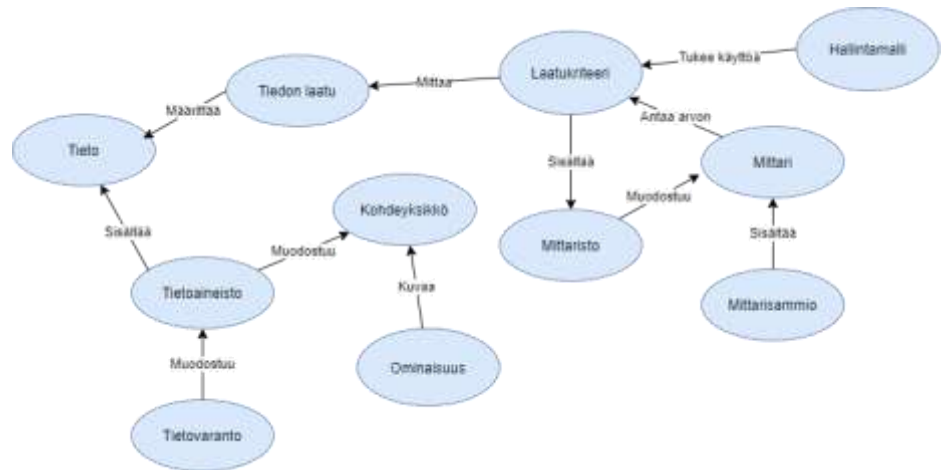
Myös Käyttäjät voisivat hyödyntää mittarisammion mittareita tiedon laadun kuvauksessa varsinaisen mittariston täydentäjänä. Mittarisammion mittarit voivat olla käyttökelpoisia tiettyjen tilanteiden mittareita, vaikka niitä ei mittaristoon ole liitetty. Sammio tukee myös jatkuvaa laatuajattelua: esimerkiksi jokin mittari voidaan todeta tässä hetkessä toimimattomaksi, mutta samalla myös haastaa pohtimaan ja olla myöhemmin ajankohtainen. Mittaristossa ja siihen kohdistuvissa tarpeissa voi olla ajallista vaihtelua ja se, että käytössä olevia mittareita voidaan peilata ajan myötä sammioon kertyneisiin mittareihin, tukee osaltaan kehittämistarpeiden tunnistamista.

Laatukriteerien ja erityisesti mittariston ja mittarisammion täysimääräisen hyödyntämisen ja kehittämisen tueksi tarvittaneen ylläpitojärjestelmä, josta laatukriteerit ja mittarit voidaan noutaa omiin järjestelmiin. Haku- ja selailutoiminnot lienevät tarpeelliset, varsinkin jos mittarisammiota ylläpidetään samassa järjestelmässä. Pysyvät tunnisteet helpottavat erityisesti mittareiden käyttöä. Hyvä huomioda, että tiedon laadun kuvauksen järjestelmät ovat tästä ylläpidon järjestelmästä erillinen asia.

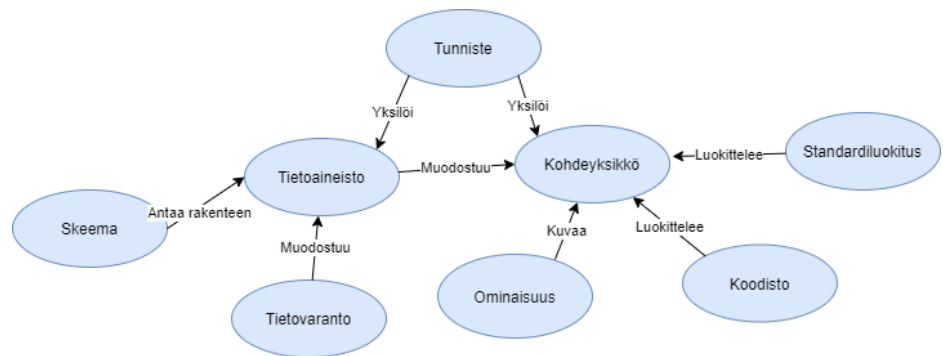
8 Liitteet

LIITE 1 – Sanasto

Dokumentissa esiintyvien termien kuvaukset



Kuva 8. Kuvitus tämän dokumentin tiedon laadun kuvaukseen liittyvistä termeistä.



Kuva 9. Kuvitus tämän dokumentin rakenteiseen dataan liittyvistä termeistä.

Taulukko 1. Dokumentissa esiintyvien termien kuvaukset.

Termi (synonyymi)	Kuvaus
kohdeyksikkö (luokka, tilastoyksikkö, yksikkö, kohde)	Tietoa ineistossa tarkasteltava yksikkö.
koodisto	Kokoelma ryhmiä, jotka eroavat toisistaan tietyiltä piirteiltään.
laatuksiteeri	Tekijä, joka kuvaa laatuksiteeriuksen laatua tietystä näkökulmasta.
mittari	Laatuksiteeriuksen näkökulmaa konkretisoiva ja tietoa ineiston ominaisuutta mittaava mittari. Kukin mittari viittaa tiettyyn laatuksiteeriin. Yhdelle laatuksiteerille voidaan soveltaa useita mittareita.
mittarisammio	Mittariyhdistelmien, mittariyhdistelmien muodostama kokonaisuus, josta hankkeen aikana poimittiin sopivin kokonaisuus laatuksiteeriuksen mittariksi. Jatkossa mittarisammio toimii paitsi mittariyhdistelmien myös mittariksi täydentävien mittareiden kokoelmana.
mittaristo	Tiedon laatuksiteeriuksen soveltamiseen tuotettu erillisten mittarien muodostama mittarikoelma.
ominaisuus (ominaisuus, attribuutti, kohde, muuttuja)	Kohdeyksikköä kuvaava tieto.
perustietovaranto	Perustietovaranto on alemman jalostusasteen tietovaranto tai jalostetun tiedon taustalla oleva tietovaranto.
rakenteinen data	Konelluttava tieto, johon on liitetty metatietojen avulla määritetty rakenne, jonka avulla tietoa voidaan jäsentää.
skeema	Tietoa ineiston rakenteen määräämuotoinen esitys.

standardiluokitus (<i>luokitusstandardi</i>)	Luokitus suositus, joka perustuu esimerkiksi EU-direktiiveillä vahvistettuihin kansainvälisiin standardeihin.
tiedon käyttäjä	Tietoa käyttävä henkilö. Tiedon käsittelyprosessin näkökulmasta prosessin viimeinen, tietoa todella käyttävä henkilö. Tiedon käyttäjä voi myös edelleen jatkojalostaa tietoa, mutta tämä on uusi käsittelyprosessi verrattuna tiedon laadun kuvaukseen.
tiedon laatu	Se, missä määrin jonkin tietoaineiston ominaisuudet täyttävät vaatimuksia tai ta voitteita. Tiedon soveltuvuus siihen käyttötarkoitukseen, johon käyttäjä aikoo sitä käyttää ja tiedon tuottaja sitä tarjoaa.
tieto	Tiedolla voi viitata monenlaisiin asioihin: merkkijono, viesti, tosiasia, havainto, tulkinta tai käsitys. Tässä tiedolla tarkoitetaan ilmiöistä koottua alimman jalostusasteen tietoa tai siitä jalostamalla tuotettua informaatiota.
tietoaineisto	Aineisto, joka sisältää tietovälineelle tallennettuja tietoja.
tietovaranto	Tiettyä tarkoitusta varten muodostettu, loogisesti tai fyysisesti yhteenkuuluvien tietojen muodostama tietoaineisto tai tietoaineistojen kokoelma.
tunniste	Yksilöimiseen käytettävä merkkijono.

Taulukko 2. Dokumentissa esiintyvien tilastollisten termien kuvaukset.

Tilastollinen termi (synonyymi)	Kuvaus
Alipeitto, tilastomenetelmäkäsité	Alipeitto liittyy tutkimuksessa käytettyyn perusjoukkoon. Perusjoukosta täytyy olla käytettävissä kehikko eli luettelo kohteista, joita koskevia tietoja halutaan otantatutkimuksella kerätä. Alipeitolla tarkoitetaan sitä, että käytettävissä olevasta kehikosta puuttuu osa perusjoukon eli tutkimuksen kohdejoukon alkiosta. Esim. henkilöt, joilla ei ole puhelinta, puuttuvat puhelinhaastattelun kehikosta.
Imputointi (<i>paikkaus</i>), tilastomenetelmä	(<i>paikkaus</i>) tarkoittaa havaintoaineistossa olevan puutteen tai poikkeavan havainnon korvaamista jolla imputointimenetelmällä. Imputointimenetelmät ovat mm. looginen imputointi (korjataan loogisesti mahdolliset virheet, esim. lapsi ei voi olla vanhempi kuin vanhempansa), hot deck -imputointi, jossa puutteelle haetaan arvo toiselta vastaajalta, cold deck -imputointi, jossa havainto haetaan saman vastaajan edellisestä vastauksesta, regressio- ja muut malliperusteiset menetelmät, jossa käytetään tilastollista mallia ennustamaan arvo puutteelle.
Perusjoukko (<i>populaatio</i>), tilastomenetelmäkäsité	Perusjoukko on tutkimuksen kohteena oleva ryhmä, jota koskevia tietoja halutaan kerätä, esim. äänestysikäiset kansalaiset. Siitä käytetään myös tarkempaa nimitystä kohdeperusjoukko. Siitä erillään on kehikkoperusjoukko, jolla tarkoitetaan tutkimuksen käytössä olevan rekisterin tai muun luettelon peittämää kohdejoukkoa, joka ei kuitenkaan aina vastaa yksi yhteen kohdeperusjoukkoa (ks. ylipeitto, alipeitto).
Poikkeava havainto (<i>outlier</i>), tilastomenetelmäkäsité	poikkeava havainto (<i>outlier</i>) on sellainen muuttujan arvo, joka poikkeaa suuresti havaintojen arvojen valtaenemmistöstä ja se voi olla joko oikea tai virheellinen. Poikkeavilla arvoilla voi olla merkittävästi vääristävä vaikutus käytettyihin tilastollisiin tunnuslukuun, kuten keskiarvoon, hajontaan, regressiosuoraan
Ylipeitto, tilastomenetelmäkäsité	ylipeitto tarkoittaa otoskehikkoon sisältyviä kohdeyksiköitä, jotka eivät enää kuulu kohdeperusjoukkoon, esim. laitoihin siirtyneitä, kuolleita, maasta muuttaneita. Tällaisia tapauksia esiintyy aina jonkin verran otosyksiköiden joukossa, koska rekisterit, joista otokset poimitaan, eivät ole täysin ajantasaisia.

Taulukko 3. Dokumentissa esiintyvien lyhenteiden selitteet.

Lyhenne	Selite
EIF	Euroopan yhteentoimivuusperiaate
Euroopan tilastojen käytännönsäännöt	The European Statistics Code of Practise
FAIR	FAIR-periaatteet: Tiedot ovat Löydettäviä (Findable), Saavutettavia (Accessible), Yhteentoimivia (Interoperable) ja Uudelleenkäytettäviä (Re-usable)
ISO 19139	Geographic information - Metadata - XML schema implementation
ISO 19157	Geographic information - Data quality
ISO 25012	Software engineering - Software product Quality Requirements and Evaluation (SQuaRE) - Data quality model
ISO 25024	Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - Measurement of data quality
ISO-kielikoodisto	ISO 639 määrittelemä kielikoodis

LIITE 2 - Mittaristo

Mittarit laatukriteeriryhmittäin

Taulukko 4. Mitä tieto koskee? -kysymykseen vastaaviin laatukriteereihin liitetyt mittarit.

Mittarin nimi	Kuvaus	Esitysmuoto	Arviointitaso	Laatukriteeri
Muutosten hallinta	Tietoa ineistossa tai ominaisuustiedoissa tapahtuvia rakenteellisia tai lähdetietojen muutoksia seurataan.	kyllä/ei	Tietoa ineisto ja ominaisuus	Jäljitettävyys
Tiedon elinkaari	Tiedosta on kuvattu mistä lähtien tietoa on kerätty, mitä muutoksia tietoa ineistoon on tehty ja mihin asti tietoa ineisto on käytettävissä.	kyllä/ei	Tietoa ineisto ja ominaisuus	Jäljitettävyys
Tietolähde	Tietoa ineiston, kohdeyksikön tai ominaisuuden lähdetieto tallennetaan.	prosenttia	Tietoa ineisto ja ominaisuus	Jäljitettävyys
Ajallinen ta voitekattavuus	Tietoa ineiston tavoiteltava ajallinen kattavuus ja tiheys on kuvattu.	kyllä/ei	Tietoa ineisto	Kattavuus
Alueellinen ta voitekattavuus	Tietoa ineiston tavoiteltava alueellinen kattavuus ja tiheys on kuvattu.	kyllä/ei	Tietoa ineisto	Kattavuus
Puutteelliset kohdeyksiköt	Yhdenkin puuttuvan ominaisuustiedon sisältävien kohdeyksiköiden osuus tietoa ineiston kaikista kohdeyksiköistä.	prosenttia	Tietoa ineisto ja ominaisuus	Kattavuus
Puutteelliset ominaisuustiedot	Puuttuvan ominaisuustiedon omaavien kohdeyksiköiden osuus kaikista kohdeyksiköistä ominaisuuksittain tarkasteltuna	prosenttia	Tietoa ineisto ja ominaisuus	Kattavuus
Puuttuvat kohdeyksiköt	Tietoa ineistosta puuttuvien kohdeyksiköiden prosentuaalinen osuus	prosenttia	Tietoa ineisto	Kattavuus
Tavoitekohdeyksiköt	Tietoa ineiston tavoitekohdeyksiköt on kuvattu aineistokuvauksessa	kyllä/ei	Tietoa ineisto	Kattavuus
Tavoiteominaisuudet	Tietoa ineiston tavoiteominaisuuksien joukko on kuvattu aineistokuvauksessa	kyllä/ei	Tietoa ineisto	Kattavuus
Ylimääräiset kohdeyksiköt	Tietoa ineistossa olevien ylimääräisten kohdeyksiköiden prosentuaalinen osuus	prosenttia	Tietoa ineisto	Kattavuus

Taulukko 5. Miten tieto kuvaa todellisuutta? -kysymykseen vastaaviin laatuksiteereihin liitetyt mittarit.

Mittarin nimi	Kuvaus	Esitysmuoto	Arviointitaso	Laatukriteeri
Luontiajankohta	Kohdeyksikön tai ominaisuuden luontiajankohta	päivämäärä	Ominaisuus	Ajantasaisuus
Muutosajankohta	Kohdeyksikön tai ominaisuuden muutosajankohta	päivämäärä	Ominaisuus	Ajantasaisuus
Tarkistusajankohta	Kohdeyksikön tai ominaisuuden tarkistusajankohta	päivämäärä	Ominaisuus	Ajantasaisuus
Tiedon muokkaaminen päättynyt	Kohdeyksikön tai ominaisuuden korjaamisen päättymisajankohta	päivämäärä	Tietoaineisto	Ajantasaisuus
Viiteajankohta	Ajankohta, jota tieto kuvaa	päivämäärä	Tietoaineisto ja ominaisuus	Ajantasaisuus
Tieto loogisuustarkastettu	Tiedon keräämisessä tai käsittelyssä on käytetty loogisuustarkastelua	kyllä/ei	Tietoaineisto ja ominaisuus	Johdonmukaisuus
Menetelmällisesti tuotetut arvot	Menetelmällisesti tai korvaavaa tietoa käyttäen tuotettujen arvojen osuus ominaisuuden kaikista arvoista	prosenttia	Ominaisuus	Oikeellisuus
Virheelliset ominaisuusarvot	Virheellisten ominaisuusarvojen suhde ominaisuusarvojen kokonaismäärään	prosenttia	Tietoaineisto ja ominaisuus	Oikeellisuus
Väärinluokittelu	Virheellisesti luokiteltujen kohdeyksiköiden suhde kaikkiin kohdeyksiköihin	prosenttia	Tietoaineisto ja ominaisuus	Oikeellisuus
Hajonta	Miten laajalle alueelle ominaisuuden arvot ovat jakautuneet	Vaihteluväli, keskihajonta, muu hajontaluku	Ominaisuus	Tarkkuus
Poikkeavat havainnot	Poikkeavien havaintojen lukumäärän suhde kaikkien kohdeyksiköiden lukumäärään	prosenttia	Tietoaineisto ja ominaisuus	Tarkkuus

Taulukko 6. Miten hyvin tieto on kuvattu? -kysymykseen vastaaviin laatuksiteereihin liitetyt mittarit.

Mittarin nimi	Kuvaus	Esitysmuoto	Arviointitaso	Laatukriteeri
Käytetyt standardit	Käytetyt standardit löytyvät metatiedoista	kyllä/ei	Tietoaineisto ja ominaisuus	Suosituksenmukaisuus
Standardienmukaisuus	Noudatetaanko tietoaineiston osalta standardeja	kyllä/ei	Tietoaineisto ja ominaisuus	Suosituksenmukaisuus
Aineistokuvaus	Onko tietoaineiston aineistokuvaus saatavilla? Aineistokuvausten kieliversiot	kieliversiot	Tietoaineisto	Ymmärrettävyys
Käsitteiden määritelmät	Ovatko keskeiset käsitteet määritellyt ja kuvattu saatavilla? Näiden kieliversiot	kieliversiot	Tietoaineisto ja ominaisuus	Ymmärrettävyys
Ominaisuuksien tietokuvaukset saatavilla	Ovatko ominaisuuksien kuvaukset ja koodistot saatavilla? Näiden kieliversiot	kieliversiot	Ominaisuus	Ymmärrettävyys
Ymmärrettävyyden asiakaspalaute	Ymmärrettävyydestä on mahdollista antaa asiakaspalautetta, joka otetaan huomioon	kyllä/ei	Tietoaineisto	Ymmärrettävyys

Taulukko 7. Miten tietoa voi käyttää? -kysymykseen vastaaviin laatukriteereihin liitetyt mittarit.

Mittarin nimi	Kuvaus	Esitysmuoto	Arviointitaso	Laatukriteeri
Kohdeyksikön pysyvä tunniste	Tietoa ineiston kohdeyksiköillä on vähintään tietoa ineistokohtainen pysyvä tunniste	kyllä/ei	Tietoa ineisto	Koneluettavuus
Koneluettavuuden asiakaspalaute	Koneluettavuudesta on mahdollista antaa asiakaspalautetta, joka otetaan huomioon	kyllä/ei	Tietoa ineisto	Koneluettavuus
Tietoa ineiston tietomalli	Tietoa ineisto on rakenteisesti kuvattu	kyllä/ei	Tietoa ineisto	Koneluettavuus
Käyttöoikeus	Tietoa ineiston käyttöoikeuden rajoitukset: esim. käsittelijä, käsittelyympäristö	käyttöoikeuden rajoitus	Tietoa ineisto	Käyttöoikeudet
Käytön rajoitukset	Tietoa ineiston käyttötarkoitusten rajoitukset. Mahdolliset avoimen datan lisenssit tai käyttöehdot.	käyttötarkoituksen rajoitus	Tietoa ineisto	Käyttöoikeudet
Määräpäivien noudattaminen	Tietoa ineiston toimitusaikataulun seuranta toteutuneeseen toimitusaikatauluun.	toimitusviive	Tietoa ineisto	Täsmällisyys
Päivitystiheys	Tietoa ineiston säännöllinen päivitystiheys	sanallinen ilmaus	Tietoa ineisto	Täsmällisyys
Päivityksessä muuttuneet ominaisuustiedot	Päivityksessä muuttuneiden ominaisuusarvojen osuus kaikista tietoa ineiston ominaisuusarvoista	prosenttia	Tietoa ineisto	Täsmällisyys