

# Suositus tietoaineistojen laadun arviointikriteereistä ja mittaristosta

Tiedonhallintalautakunta

Valtiovarainministeriö

© [Copyright-taso tähän](#)

ISBN sid. [VNK täyttää](#)

ISBN pdf [VNK täyttää](#)

ISSN sid. [VNK täyttää](#)

ISSN pdf [VNK täyttää](#)

Taitto Valtioneuvoston hallintoyksikkö, Julkaisutuotanto

Helsinki 2023

Paino PunaMusta Oy, 2023

## Suositus tietoaineistojen laadun arviointikriteereistä ja mittaristosta

VNK täyttää, sarja ja numero	Teema	Lautakunnat
<b>Julkaisija</b>	Valtiovarainministeriö	
<b>Tekijä/t</b>	<a href="#">Napsauta ja kirjoita</a>	
<b>Toimittaja/t</b>	<a href="#">Napsauta ja kirjoita</a>	
<b>Yhteisötekijä</b>	Tiedonhallintalautakunta	
<b>Kieli</b>	suomi	<b>Sivumäärä</b> <a href="#">V NK täyttää</a>
<b>Tiivistelmä</b>	<p>Suosituksessa kuvataan julkisen hallinnon tietoaineistojen laadun arviointikriteerit ja niihin liittyvä mittarikokoelma. Ne tarkastelevat laadun eri osa-alueita ja ulottuvuuksia tietoaineistoihin sisältyvässä datassa sekä aineistojen vaihdon tilanteissa käyttäjän näkökulmasta. Kriteerit auttavat käyttäjiä arvioimaan soveltuuko tietoaineisto aiottuun tarkoitukseen ja mittaristo antaa mahdollisuuden tarkastella yhteismitallisesti aineistojen laatua.</p> <p>Suositus on suunnattu laatuksymysten parissa työskenteleville tiedon tuottajille ja tiedon laadusta kiinnostuneille käyttäjille. Se on kirjoitettu sähköisessä muodossa olevan, julkisen hallinnon tuottaman ja kokoaman rakenteisen tiedon kuvailuun ja arviointiin. Kriteereitä ja mittareita voi soveltuvilta osin käyttää muidenkin tietoaineistojen laadun arviointiin ja seuraamiseen. Suosituksessa ei kuvata kriteerien tai mittariston käyttöönottoa, ylläpitoa ja hallintaa.</p> <p>Suositus tukee tiedonhallintalain veloitteiden täytäntöönpanoa ja osa kriteereistä liittyy suoraan niiden täytäntöönpanoon. Laki julkisen hallinnon tiedonhallinnasta (906/2019) tuli voimaan 1.1.2020. Yleislakina se sisältää säännökset julkisen hallinnon tiedonhallinnan järjestämisestä ja kuvaamisesta. Sen tavoitteisiin on kirjattu tavoite varmistaa viranomaisen tietoaineistojen yhdenmukainen ja laadukas hallinta ja tavoite tietoaineistojen turvallisen ja tehokkaan hyödyntämisen mahdollistamisesta. Tiedonhallintalautakunta hyväksyi suosituksen x.x.2023.</p>	
<b>Klausuuli</b>	<a href="#">V NK täyttää</a>	
<b>Asiasanat</b>	tiedonhallintalautakunta, tiedonhallintalaki, lautakunnat, suositukset, julkinen hallinto, tiedonhallinta, tietovarannot, tietoaineistot, data, metatiedot, kuvailutiedot, laadunarviointi	
<b>ISBN PDF</b>	<a href="#">V NK täyttää</a>	<b>ISSN PDF</b> <a href="#">V NK täyttää</a>
<b>ISBN nid.</b>	<a href="#">V NK täyttää</a>	<b>ISSN painettu</b> <a href="#">V NK täyttää</a>



## Sisältö

Johdanto.....	7
<b>1 Tiedon laatu ja tiedonhallintalaki .....</b>	<b>9</b>
1.1 Laissa säädetyt vaatimukset tiedon laadulle .....	9
1.2 Suosituksen hyödyntäminen .....	10
1.3 Suosituksen suhde muihin tiedonhallintalautakunnan suosituksiin .....	12
<b>2 Tiedon laatukriteerit ja mittaristo .....</b>	<b>14</b>
2.1 Tiedon laatuun liittyvien metatietojen ylläpito .....	15
2.2 Laatukriteerien ja mittariston rajaukset.....	16
2.3 Mittariston soveltaminen.....	16
2.4 Laatukriteeriryhmä: Miten tieto kuvaa todellisuutta? .....	18
2.4.1 Laatukriteeri: Virheettömyys .....	19
2.4.1.1 Menetelmällisesti tuotetut arvot -mittari.....	19
2.4.1.2 Virheelliset ominaisuusarvot -mittari .....	20
2.4.1.3 Väärinluokittelu-mittari.....	20
2.4.2 Laatukriteeri: Tarkkuus .....	21
2.4.2.1 Keskihajonta-mittari.....	21
2.4.2.2 Poikkeavat havainnot -mittari .....	22
2.4.3 Laatukriteeri: Johdonmukaisuus .....	23
2.4.3.1 Tieto loogisuustarkastettu -mittari .....	23
2.4.4 Laatukriteeri: Ajantasaisuus .....	23
2.4.4.1 Viiteajankohta-mittari.....	24
2.4.4.2 Luontiajankohta-mittari.....	24
2.4.4.3 Tarkastusajankohta-mittari.....	25
2.4.4.4 Muutosajankohta-mittari.....	25
2.4.5 Laatukriteeri: Kattavuus.....	26
2.4.5.1 Ajallinen tavoitekattavuus -mittari.....	26
2.4.5.2 Alueellinen tavoitekattavuus -mittari.....	27
2.4.5.3 Tavoitekohdeyksiköt-mittari.....	27
2.4.5.4 Tavoiteominaisuuksissa puutteita -mittari .....	27
2.4.5.5 Puuttuvat kohdeyksiköt -mittari .....	28
2.4.5.6 Ylimääräiset kohdeyksiköt -mittari.....	29
2.4.5.7 Puutteelliset kohdeyksiköt -mittari.....	29

2.4.5.8	Puutteelliset ominaisuustiedot -mittari.....	30
2.5	Laatukriteeriryhmä: Miten tieto on kuvattu? .....	30
2.5.1	Laatukriteeri: Alkuperäisyys.....	31
2.5.1.1	Tietolähde-mittari .....	31
2.5.1.2	Tiedon elinkaari -mittari.....	32
2.5.1.3	Muutosten hallinta -mittari.....	33
2.5.2	Laatukriteeri: Metatietojen ymmärrettävyys .....	33
2.5.2.1	Aineistokuvaus-mittari.....	33
2.5.2.2	Käsitteiden määritelmät -mittari.....	34
2.5.2.3	Ominaisuuksien tietokuvaukset -mittari.....	35
2.5.2.4	Metatietojen ymmärrettävyyden asiakaspalaute -mittari .....	35
2.5.3	Laatukriteeri: Suositustenmukaisuus.....	36
2.5.3.1	Noudatettavat standardit ja suositukset -mittari .....	36
2.6	Laatukriteeriryhmä: Miten tietoa voi käyttää?.....	37
2.6.1	Laatukriteeri: Koneluettavuus.....	37
2.6.1.1	Tietoaineiston tietomalli -mittari.....	38
2.6.1.2	Kohdeyksikön pysyvä tunniste -mittari.....	38
2.6.1.3	Koneluettavuuden asiakaspalaute -mittari .....	39
2.6.2	Laatukriteeri: Tietoaineiston käyttöoikeudet .....	39
2.6.2.1	Käyttöoikeus-mittari.....	40
2.6.2.2	Käytön rajoitukset -mittari.....	40
2.6.3	Laatukriteeri: Oikea-aikaisuus .....	41
2.6.3.1	Määräpäivien noudattaminen -mittari.....	41
2.6.3.2	Päivitystiheys-mittari .....	41
2.6.3.3	Päivityksessä muuttuneet ominaisuustiedot -mittari .....	42
	<b>Liitteet.....</b>	<b>43</b>
	<b>Liite 1 Suosituksessa käytetyt keskeiset käsitteet.....</b>	<b>44</b>
	<b>Liite 2 Laadun arviointikriteerien mittarit.....</b>	<b>48</b>

## Johdanto

Tiedonhallintalautakunnan suosituksessa tietoaaineistojen laatukriteereistä ja mittaris-  
tosta kuvataan tietoaaineistojen laadunhallinnassa hyödynnettäviä arviointikriteerejä ja  
näiden soveltamista tukeva mittarikokoelma. Ne muodostavat työkalun julkisen hallin-  
non tietoaaineistojen laadun kuvailuun ja tarkasteluun organisaatio- ja toimialarajoista  
riippumatta. Suositus ei ota kantaa tavoiteltaviin laatutasoihin eikä kuvaa mitä tiedon  
laadun tulisi viranomaisen eri tehtävissä olla.

Suosituksen tavoitteena on edistää julkisen hallinnon tietovarantojen nykyistä moni-  
puolisempaa käyttöä ja tukea tiedonhallintalaissa säädettyjen vaatimusten toteutta-  
mista. Se auttaa tietoaaineistojen turvallisuuden varmistamisessa, niiden laadun ku-  
vaamisessa ja tunnistamisessa, erilaisissa tiedon vaihdon tilanteissa, tietoa avatta-  
essa sekä tiedon laadun vertailussa ja kehittämisessä.

Julkisen hallinnon tiedonhallinnasta annetun lain (906/2019), myöhemmin tässä suo-  
situksessa tiedonhallintalaki tai TiHL, tavoitteena on varmistaa viranomaisten tietoi-  
aineistojen yhdenmukainen ja laadukas hallinta sekä tietoturallinen käsittely julkisuus-  
periaatteen toteuttamiseksi. Kun viranomaisten tehtävien hoitaminen ja palvelujen  
tuottaminen ovat aina sidoksissa käytettävissä oleviin ajantasaisiin tietoihin, on tie-  
donhallintalaissa säädetty viranomaisen velvollisuudesta varmistaa käyttämiensä tieto-  
aineistojen muuttumattomuus, ajantasaisuus ja virheettömyys (TiHL 15 §), tietoaaneis-  
tojen saatavuudesta (TiHL 15, 19 sekä 24 a ja 24 b §) sekä tietojen hallinnassa käy-  
tettävistä metatiedoista (TiHL 25–27 §). Lisäksi laissa on säädetty tiedonhallintayksik-  
könä toimiville viranomaisille velvollisuus kuvata tiedonhallintaa määrittävä toimin-  
taympäristö tiedon laadun varmistavine tietoturvaluustoimenpiteineen (TiHL 5 §)  
sekä arvioida suunnittelemiensa toimenpiteiden vaikutukset tiedonhallintaan suunni-  
tellessaan muutoksia toimintaansa ja tietojärjestelmiinsä (TiHL 5.3 §). Tietojen laadun  
varmistamiseen liittyy myös tiedonhallintayksikön velvollisuus selvittää olennaiset tie-  
tojenkäsittelyyn kohdistuvat riskit ja mitoittaa tietoturvaluustoimenpiteensä riskiarvi-  
oinnin mukaisesti (TiHL 13.1 §).

Julkisen hallinnon tiedonhallintalautakunnan, myöhemmin tässä tiedonhallintalauta-  
kunta, tehtävänä on edistää tiedonhallintalaissa säädettyjen tiedonhallinnan ja tieto-  
turvallisuuden menettelytapojen ja vaatimusten toteuttamista (TiHL 10 §). Tämän teh-  
tävänsä toteuttamiseksi tiedonhallintalautakunta laatii ja ylläpitää suosituksia, joilla hy-  
viin käytäntöihin pohjautuen ohjataan tiedonhallintayksiköjä ja viranomaisia toteut-  
tamaan tiedonhallintalaissa säädettyt vaatimukset.

Tietoaaineiston laadun kuvaaminen auttaa tiedon käyttäjiä arvioimaan soveltuuko tieto-  
aineisto aiottuun käyttötarkoitukseen. Se myös auttaa arvioimaan tietoaaineistojen

muuttumattomuutta, alkuperäisyyttä, ajantasaisuutta, virheettömyyttä, saatavuuden oikea-aikaisuutta sekä käyttökelpoisuutta.

Laatukriteerien ja mittariston käytön on tarkoitus olla mahdollisimman helppoa. Kriteerit on pyritty kuvaamaan niin, että tietoaineiston metatietojen täydentäminen on mahdollista ilman laadun määrittämisen osaamista tai aiempaa kokemusta. Kuvaukset sisältävät myös kriteerien ja mittariston keskeisimmät soveltamisnäkökulmat.

Arviointikriteerit ja mittaristo on kehitetty erityisesti sähköisessä muodossa olevan, julkisen hallinnon tuottaman ja kokoaman rakenteisen tiedon kuvailuun ja arviointiin. Ne soveltuvat paremmin toiminnan suunnittelussa tai toimintaympäristön tilannekuvan ylläpidossa tai tilastoinnissa käytettävien tietoaineistojen kuin yksittäisten hallinnollisten ratkaisujen tai palveluiden vaatimien tietojen laadun arviointiin. Niitä voidaan soveltuville osin käyttää myös muiden kuin rakenteisten tietojen laadun arviointiin ja seuraamiseen. Suosituksessa ei käsitellä laatukriteerien tai mittariston käyttöönottoa tiedonhallintayksiköissä tai viranomaisissa, ei laatuun liittyvien metatietojen ylläpitoa tai hallintaa eikä laadun erilaisia arviointimenetelmiä.

Suositus on suunnattu tietojen laadunvarmistamisesta vastaaville viranomaisille sekä tiedonhallintalain soveltamisalaan kuuluville julkista hallintotehtävää hoitaville yksityisille toimijoille. Suositusta voivat hyödyntää kaikki tietoaineistojen ja niiden laatukysymysten parissa työskentelevät tiedon tuottajat sekä tiedon laadusta kiinnostuneet tiedon käyttäjät.

Suosituksen laatukriteerit pohjaavat pitkälti ISO 25012 -standardiin ”Software engineering – Software product Quality Requirements and Evaluation (SQuaRE) – Data quality model”, jota käytetään useissa maissa tietoaineistojen laadun kuvauksen taustamallina. Käsitteistö on tarkistettu suhteessa tietosuoja-asetukseen, saavutettavuutta koskevaan sääntelyyn sekä tiedonhallintalakiin. Lisäksi laatukriteereihin on otettu mukaan tiedon virheettömyyden kriteeri. Laatukriteerien laadinnassa on huomioitu Eurooppalaiset yhteentoimivuusperiaatteet (European Interoperability Framework, EIF, COM(2017)134)<sup>1</sup>, FAIR-periaatteet (FAIR Guiding Principles for scientific data management and stewardship, 2016)<sup>2</sup> ja Euroopan tilastojen käytännesäännöt (European Statistics Code of Practice, CoP)<sup>3</sup>.

<sup>1</sup> <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2017%3A134%3AFIN>.

<sup>2</sup> <https://www.nature.com/articles/sdata201618>.

<sup>3</sup> <https://ec.europa.eu/eurostat/web/products-catalogues/-/ks-02-18-142>.



# 1 Tiedon laatu ja tiedonhallintalaki

Julkisen hallinnon tiedonhallinnasta annetun lain, (906/2018) tiedonhallintalaki, tarkoituksena on varmistaa viranomaisten tietoaineistojen yhdenmukainen ja laadukas hallinta sekä tietoturvallinen käsittely julkisuusperiaatteen toteuttamiseksi. Tiedonhallintalaissa tietoaineistojen laadukkaalla hallinnalla tarkoitetaan erityisesti viranomaisen käyttämien tietoaineistojen ajantasaisuuden, virheettömyyden ja käyttökelpoisuuden valmistamista.

Tietoaineistolla tarkoitetaan tässä suosituksessa tiedonhallintalain mukaisesti asiakirjoista ja muista vastaavista tiedoista muodostuvaa tiettyyn viranomaisen tehtävään tai palveluun liittyvää tietokokonaisuutta.

## 1.1 Laissa säädetyt vaatimukset tiedon laadulle

Viranomaisen velvollisuudesta varmistaa tietoaineistojensa alkuperäisyys, ajantasaisuus ja virheettömyys tarpeellisin tietoturvaluustoimenpitein on säädetty tiedonhallintalain 15.1 §:n 3 kohdassa. Asiasta huolehtiminen on tärkeää viranomaisen toiminnan asianmukaisuuden varmistamiseksi sekä hallinnossa työskentelevien ja hallinnon asiakkaiden oikeusturvan toteuttamiseksi ja varmistamiseksi.

Tiedonhallintalaissa ei säännellä tarkemmin mitä nämä virheettömyys- ja ajantasaisuusvaatimukset ovat. Laillisuusvalvonnassa on todettu itsestään selvänä lähtökohdaksi, että viranomaisen laatimien asiakirjojen ja ylläpitämien tiedostojen on oltava oikeita, virheettömiä ja ajantasaisia. Tämän on katsottu tarkoittavan myös sitä, että viranomaisten ylläpitämien tietojärjestelmien on oltava sellaisia, että niistä on tulostettavissa oikeaa tietoa ja tietojen keräämiseen tarkoitettun ohjelman tulee käsitellä järjestelmän sisältämiä tietoja siten, että tietojen keräämisen tuloksena syntyneet tiedot ovat oikeita.<sup>4</sup>

Tietosuoja-asetuksen 5 (1) artiklan d kohdan ”täsmällisyys” mukaan henkilötietojen käsittelyn periaatteena on vaatimus henkilötietojen täsmällisyydestä ja ajantasaisuudesta tarpeen mukaan. Tietojen käsittelyssä on siis toteutettava kaikki mahdolliset kohtuulliset toimenpiteet sen varmistamiseksi, että käsittelyn tarkoituksiin nähden epätarkat ja virheelliset henkilötiedot poistetaan tai oikaistaan viipymättä. Tiedon laatuvaatimuksista on kyse myös tietosuoja-asetuksen 5 (1) artiklan f kohdassa ”eheys ja

---

<sup>4</sup> Ks. esim. AOKS OKV/1242/1/2013, 28.4.2014 sekä EOAK/5372/2019, 22.6.2020.

luottamuksellisuus”, jonka mukaan henkilötietoja on käsiteltävä, käyttäen asianmukaisia teknisiä tai organisatorisia toimia, tavalla, jolla varmistetaan henkilötietojen asianmukainen turvallisuus, mukaan lukien suojaaminen luvattomalta ja lainvastaiselta käsittelyltä sekä vahingossa tapahtuvalta häviämiseltä, tuhoutumiselta tai vahingoittumiselta.

Tiedonhallintalain ja tietosuojasetuksen tietojen laatua koskevat vaatimukset liittyvät jokaisen viranomaisen velvollisuuteen huolehtia asioiden asianmukaisesta selvittämisestä päätöksenteon yhteydessä. Hallintolain 31 §:ssä säädetään viranomaisen velvollisuudesta huolehtia asian asianmukaisesta ja riittävästä selvittämisestä hankkimalla asian ratkaisemiseksi tarpeelliset tiedot sekä selvitykset. Hallintolain esitöiden mukaan asian riittävällä selvittämisellä tarkoitetaan sitä, että viranomainen hankkii sellaiset tiedot ja selvitykset, joilla se arvioi olevan merkitystä asian ratkaisemiselle. Viranomaisen on tapauskohtaisesti arvioitava näiden selvitysten laajuus ja tarve.<sup>5</sup>

Tiedonhallinnan yleislainsäädännössä ei säädetä tarkemmin, mitä tiedon laatua koskevat vaatimukset ovat tai millä menettelyillä ja toimenpiteillä viranomaisen olisi mahdollista varmistaa tietojensa laatu. Siksi tässä suosituksessa on esitetty kansainvälisiin standardeihin pohjautuva menettelytapa tiedon laadun kuvaamiseen.

## 1.2 Suosituksen hyödyntäminen

Suosituksen sisältämät laatuksiteerit ja mittarit tarkastelevat tiedon laadun eri ulottuvuuksia julkisen hallinnon rakenteisessa datassa, siitä muodostettujen tietoaineistojen vaihdon tilanteissa sekä sitä, mitkä laadun osa-alueet ovat näissä aineistoissa ja tilanteissa keskeisiä tiedon käyttäjän näkökulmasta. Liitteeseen 1 on kuvattu keskeiset suosituksessa käytetyt käsitteet.

Laatukriteerit tarjoavat viranomaiselle apuvälineen varmistaa, että sen toisille viranomaisille luovuttamat tiedot vastaavat tiedon vastaanottajan laatuvaatimuksia. Yhteinen kriteeristö mahdollistaa tietoja luovuttavan ja vastaanottavan viranomaisen kuvata laatuvaatimuksiaan yhteismitallisella tavalla. Yhteismitallisuus puolestaan parantaa viranomaisten yhteistä ymmärrystä tiedon laadun vaatimuksista sekä laadun turvaavista toimenpiteistä.

Laatukriteerit auttavat tunnistamaan, mitkä tekijät ovat tiedon laadun näkökulmasta tärkeitä yhteisen ymmärryksen rakentamisessa. Tiedon käyttäjälle on tärkeää saada

---

<sup>5</sup> Tarkemmin HE 72/2002 vp, s.92–93.

kattava ja oikea kuva tietoaineiston sisällöstä, sisällön ja sen kuvauksen laadusta ja osuvuudesta sekä tiedon käyttömahdollisuuksista.

Viranomaisen on kuitenkin ensin tunnistettava tehtävässään tarvitsemansa tiedon olemassa olevat laatuvaatimukset, jotka tulee huomioida ja ottaa mukaan kuvailtaessa tietoaineiston laatua suosituksen mukaisilla laatukriteereillä. Tällaisia olemassa olevia laatuvaatimuksia ovat esimerkiksi tehtävään liittyvästä erityislaista tai hallinnon ja tiedonhallinnan yleislaeista muodostuvat vaatimukset. Näiden vaatimusten tunnistamisen jälkeen viranomainen voi sekä sovittaa nämä vaatimukset suosituksessa esitettyyn viitekehykseen että hyödyntää suositukseen sisältyviä muita laatukriteereitä tietoaineistonsa laadun kuvailussa.

Vaatimukset tiedon ajantasaisuudesta tai tarkkuudesta muodostuvat aina viranomaisen tehtävän, sitä koskevan sääntelyn sekä tietojen käyttötarkoituksen kautta. Suosituksen kriteerien ja mittaristojen soveltaminen on helpompaa kun tehtävä, sääntely ja arvioitu käyttötarkoitus ovat tiedossa.

Tiedon laatu on kokonaisuus, jossa laadun eri elementit vaikuttavat toisiinsa. Laadun parannus jonkin laatukriteerin kohdalla voi heikentää toisen laatukriteerin avulla kuvattua tiedon laatua. Jos esimerkiksi tavoitellaan tietoaineiston täydellistä kattavuutta tai ominaisuustietojen erityisen hyvää tarkkuutta, niin tiedon ajantasaisuus yleensä heikenee.

Suosituksessa kuvattu menetelmä ei kerro tai kuvaa mitä tiedon laadun tulisi viranomaisen eri tehtävissä olla, vaan se tarjoaa apuvälineen soveltaa hyviä käytäntöjä tietoaineistojen laadun kuvaamiseen ja varmistamiseen. Suosituksen ISO 25012 -standardiin pohjautuva malli on pyritty muodostamaan joustavaksi, jotta sitä voidaan hyödyntää niin yksittäisiin tietoaineistoihin kuin laajemman toiminnan suunnittelussa ja arvioinnissa hyödynnettävissä tietoaineistoissa tai tilastoinnin tietojenkäsittelyn tarpeisiin.

Suosituksen valmistelussa tehdyistä sisällöllisistä rajauksista johtuen se soveltuu paremmin toiminnan suunnittelussa, toimintaympäristön tilannekuvan ylläpidossa tai tilastoinnissa käytettäviin tietoaineistoihin kuin yksittäisten hallinnollisten ratkaisujen tiedonhallintaan

Tiedon laadun kuvaukset voidaan liittää osaksi tietoaineiston metatietoja. Metatieto, tiedonhallintalaissa kuvailutieto, on tietoa, joka kuvailee tietojoukkoa. Laatukriteerien ja mittarien tiedot tarvitsevat tuekseen myös muita metatietojen tietoelementtejä, jotta tiedon käyttäjä ymmärtää ja tulkitsee laatutiedot ja mittarien arvot oikein.

## 1.3 Suosituksen suhde muihin tiedonhallintalautakunnan suosituksiin

Tietoaineistojen laatukriteerejä ja niihin liittyviä mittareita koskeva suositus tarjoaa viranomaiselle apuvälineitä, paitsi kuvata tietoaineistojensa laatuun liittyviä ominaisuuksia niiden kuvailutiedoissa, myös tietoaineistojensa laadun turvaavien toimenpiteidensä arviointiin. Viranomaisen tulee kuvata tiedonhallintamallissaan (TiHL 5 §) näitä tietojensa käsittelyn ja tietovarantojen ylläpidon vaatimia tietoaineistojen laadun turvaavia toimenpiteitä. Tiedonhallintamallin laatimista ja ylläpitoa tukee tiedonhallintalautakunnan suositus tiedonhallintamallista (Valtiovarainministeriön julkaisu 2020:29). Tiedonhallintamallissa kuvatut tiedot tiedonhallintayksikön tietovarannoista ja niiden sisältämistä tietoaineistoista tietoryhmittäin sekä tietoaineistojen käsittelyyn liittyvistä tietoturvaluustoimenpiteistä muodostavat luontevan soveltamiskohteen tässä suosituksessa kuvatulle laatukriteeristölle.

Laatukriteereitä ja -mittareita voidaan hyödyntää osana tiedonhallinnan muutosvaikutusten arviointia, muutoksen liittyvien tai sen kohteena olevien tietoaineistojen turvallisuuden varmistamista (TiHL 15 §), osana tietoaineistojen saatavuutta (TiHL 19 sekä 24 a ja 24 b §) tai tietojen hyödyntämistä koskevien vaatimusten vaikutusarviointia (TiHL 5.3 §). Tiedonhallintalautakunnan suositus tiedonhallinnan muutosvaikutusten arvioinnista (Valtiovarainministeriön julkaisu 2020:53) kuvaa hyviä käytäntöjä arvioida muutosten vaikutuksia tiedonhallintayksikön tiedonhallintaan. Viranomaisen tehtävän ja siihen liittyvän tietotarpeen muuttuessa voidaan laatukriteereitä käyttää apuna esimerkiksi verrattaessa muutoksen vaikutusta lähtötilanteeseen tai arvioitaessa käsiteltävien tietojen laatuvaatimuksia muutoksen jälkeisessä tilanteessa. Tietoaineistojen laadun varmistaminen ja laatua seuraavat kontrollit ovat osa tiedonhallintayksikön toiminnan ja toimintaympäristön aktiivista seuranta- ja arviointia, jonka perusteella tiedonhallintayksikkö valitsee muun muassa tiedonhallinnassa tarvittavat tietoturvaluustoimenpiteet (TiHL 13.1 §). Tietojen laadullisten ominaisuuksien tunnistaminen, tietojen laatukriteerit sekä laatua kuvaavat tiedot osana tietoaineistojen meta-tietoja ovat apuvälineitä laissa mainittujen laatuvaatimusten toteuttamiseksi.

Ajantasaiset tiedot tietoaineistojen laadusta auttavat tiedonhallintayksikköä tietovarantoihin liittyvien riskien tunnistamisessa ja hallinnassa (TiHL 13.1 §). Laatukriteerien soveltaminen tiedon elinkaaren eri vaiheissa ja laatuvaatimusten mukaiset tiedon laatuun liittyvät tietojen käsittelytoimet auttavat viranomaista määrittämään tietoaineiston elinkaareen kohdistuvia riskejä ja mitoittamaan tietoturvaluustoimenpiteet tehdyn riskiarvioinnin mukaisesti. Tietoaineistojen turvallisuuden varmistamisesta sekä riskienhallintaan liittyvistä menettelyistä saa lisätietoa ja ohjeistusta tiedonhallintalautakunnan suosituskokoelmasta tietyjen tietoturvaluussäännösten soveltamisesta

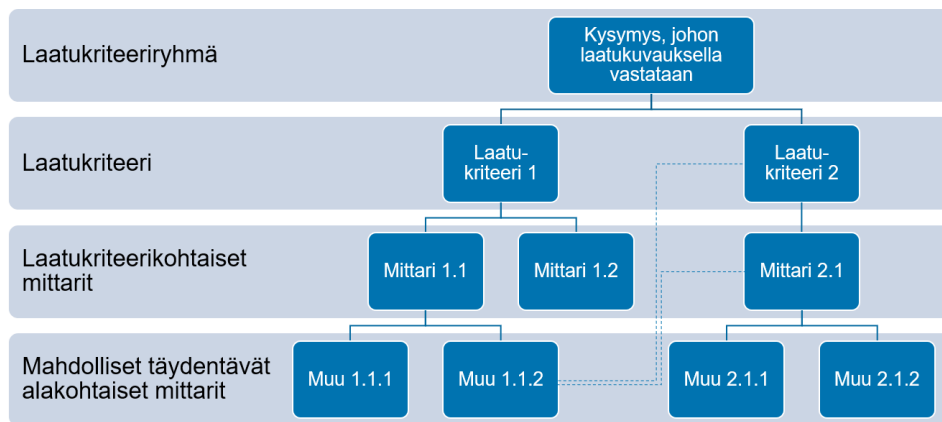
(Valtiovarainministeriön julkaisu 2021:65). Tietoaineiston saavutettavuuteen sisältyvä oikea-aikaisuus on perusedellytys myös tietojen luovuttamisessa teknisten rajapintojen ja katseluyhteyksien välillä. Teknisistä rajapinnoista ja katseluyhteisistä on julkaistu oma suosituksensa (Valtiovarainministeriön julkaisu 2021:21).

## 2 Tiedon laatukriteerit ja mittaristo

Laatukriteerit on tässä suosituksessa jaettu kolmeen ryhmään sen mukaan miten arvioinnin kohteena oleva tieto kuvaa todellisuutta, miten tieto on kuvattu ja miten tietoa voi käyttää. Jokaiseen kriteeriin liittyy kuvattua laadun ominaisuutta tarkentavat mittarit, joiden määrä vaihtelee kriteereittäin.

Suositukseseen on valittu joukko mittareita, jotka soveltuvat laadun kuvaamiseen mahdollisimman laajasti. Mittarien kuvaukset on pidetty tiiviinä ja esittelevinä. Ne sisältävät myös keskeisimmät soveltamisnäkökulmat ja -esimerkit. Kriteerit ja mittarit voidaan kuvata hierarkkisella mallilla, joka on esitetty kuviossa 1. Laatukriteerit mittareineen on esitelty tarkemmin laatukriteeriryhmittäin luvuissa 2.4.–2.6. Lisäksi mittarit on esitetty koosteena liitteessä 2.

**Kuvio 1.** Tiedon laatukriteerien hierarkkinen malli.



Hierarkkista mallia voi tarvittaessa syventää ja laajentaa suosituksessa kuvattua tarkemmilla toimialakohtaisilla ja muilla mahdollisilla mittareilla organisaation tarpeiden mukaan.

Suositus sisältää mittareita, joiden käyttö ja tietosisällön oikea ymmärtäminen edellyttävät tietoaineiston keskeisimpien metatietojen kuvausta ja saatavilla oloa laatua kuvailevan metatiedon rinnalla. Tällaisia ovat esimerkiksi tietoaineistoihin sisältyvien kohdeyksiköiden rajaukset tai tiedossa olevat puutteet tai ylimääräiset kohdeyksiköt sekä tietoaineistoon liittyvät tietolähteet, ks. esimerkiksi 2.4.5.3 Tavoitekohdeyksiköt,

2.4.5.5 Puuttuvat kohdeyksiköt, 2.4.5.6 Ylimääräiset kohdeyksiköt sekä 2.5.1.1 Tietolähde. Näiden tietojen tarjoaminen yhdessä tietoaineiston muiden metatietojen kanssa auttaa tiedon käyttäjää tai käyttöä harkitsevaa laatukuvauksen hyödyntämisessä.

Suositukseseen sisältyy kuvauselementtejä, jotka laajemmissa metatietomalleissa tai tiedonhallintayksiköiden omissa metatietomalleissa saattavat sijoittua luontevasti muun kuin laatu-metatietoelementin yhteyteen tai hierarkkisesti sen alle. Tällaisia voivat olla esimerkiksi ajantasaisuuden laatukriteeriin liitetyt mittarit luontiajankohta, tarkastusajankohta ja muutosajankohta.

## 2.1 Tiedon laatuun liittyvien metatietojen ylläpito

Tiedonhallintayksikön ja viranomaisen kannattaa määritellä laatukriteerien kuvaamiseen ja niiden ylläpitoon liittyvät vastuut ja roolit. Tietoaineistojen laatuun liittyviä metatietoja ylläpidetään ja päivitetään tiedonhallintayksikössä säännöllisesti, jotta tiedot pysyvät ajan tasalla.

Tässä suosituksessa ei kuvata tai ohjeisteta millä menettelyillä tai missä muodossa laatuun liittyvät metatiedot tuotetaan, ylläpidetään tai julkaistaan. Niissä kannattaa hyödyntää tiedonhallintayksikössä aiemmin tehtyjä kuvauksia, tietoaineistojen olemassa olevia metatietoja ja niiden hallinta- ja tallennustyökaluja sekä tiedonhallintamallia, joka on kuvaus tiedonhallintayksikön toimintaympäristön tiedonhallinnasta.

Tiedonhallintayksikkö voi rikastaa tässä suosituksessa esitettyä tietosisältöä tarpeen mukaan, jotta se palvelee yhä paremmin tiedon käyttäjien tarpeita. Jatkokehittämiseen tai kuvauksen laadun arviointiin voidaan ottaa mukaan myös sidosryhmien edustajia.

Tietoaineistojen laatuun liittyvissä metatiedoissa on kiinnitettävä erityistä huomiota niiden selkeyteen ja ymmärrettävyyteen tiedon hyödyntäjän näkökulmasta. Metatietojen kuvausten tulee täyttää myös saavutettavuuden vaatimukset. Lisätietoja saavutettavuusvaatimuksista löytyy osoitteesta <https://www.saavutettavuusvaatimukset.fi>.

Laatukuvaukset liittyvät tiedonhallintayksikön riskien hallintaan kertoen tietoaineistojen sisällöistä ja niihin liittyvistä riskeistä. Ajantasaiset tietoaineistojen laadun kuvaukset auttavat tiedonhallintayksikköä tietoaineistoihin liittyvien riskien tunnistamisessa ja hallinnassa.

## 2.2 Laatuksiteerien ja mittariston rajaukset

Suosituksen sisältämien laatuksiteerien ja niiden mittariston osalta on tehty seuraavat rajaukset:

- Laatuksiteerit ja mittaristot koskevat vain rakenteista dataa.
- Mittariston tarkoitus on palvella tiedonhallintayksiköiden erilaisia tiedonvaihdon tilanteita ja tiedon käyttäjistä tarkastellaan lähinnä tiedonhallintayksikön ulkopuolisia käyttäjiä.
- Laatuksiteerit ja mittaristo eivät ota kantaa tavoiteltaviin laatuksitasoihin eiväkä ne kerro, mitä tiedon laadun tulisi olla viranomaisten eri tehtävissä.

## 2.3 Mittariston soveltaminen

Suosituksen sisältyvä tietoaisteistojen laatuksiteerien mittaristo tukee tiedon laadun tunnistamista ja kriteerien ymmärtämistä yhdistämällä kriteereihin konkreettista sisältöä. Mittarit on kuvattu kunkin laatuksiteerin kohdalla ja niiden määrä vaihtelee kriteereittäin. Mittarit on esitelty tarkemmin laatuksiteerikohtaisesti luvuissa 2.4–2.6. Lisäksi kooste kaikista mittareista on liitteessä 2.

Osassa laatuksiteereitä on sekä kriteerissä että sen mittareissa päästy suhteellisen helposti laadun kuvaamisen ytimeen ja toisissa on lähestytty kuvaamista useammasta suunnasta.

Mittaristo sisältää sekä määrällisiä että tiedon sisällön ja kuvauksen toteavia (kyllä/ei) mittareita. Toteavien mittarien osalta oletetaan, että lisätieto mittarin aiheesta on tiedon käyttäjän löydettävissä muusta tietoaisteiston dokumentaatiosta, esimerkiksi saatavilla oleista metatiedoista.

Tavoitteena on, että mittaristo tarjoaa tiedon käyttäjälle mahdollisuuden tarkastella yhteismitallisesti julkisen hallinnon tietoaisteistojen laatuksia.



### Mittariston soveltamisen yleiset periaatteet ja ohjeet

- **Suosituksen laatukriteerit ja mittaristo on tarkoitettu joustavaksi työkaluksi:** Kaikki kriteerit ja mittarit eivät välttämättä ole merkityksellisiä kaikissa aineistoissa ja tilanteissa. Laatua suositellaan tarkasteltavan uusistakin näkökulmista ja käyttäjän tarpeista lähtien.
- **Tunnista tiedon käyttäjät ja hyödyntäjät:** Tiedon käyttäjiä ovat esimerkiksi viranomaiset, jotka tarvitsevat tiettyjä, määriteltyjä tietoja tehtäviensä toteuttamiseen, hallinnon asiakkaat sekä muut viranomaiset ja toimijat, jotka tarvitsevat tietoja tehtävänsä toteuttamiseksi. Tiedon hyödyntäjillä on toisistaan poikkeavia vaatimuksia tiedon laadun osalta. Nämä ryhmät erilaisine tarpeineen on hyvä tunnistaa mittaristoa käytettäessä.
- **Aloita mittariston soveltaminen tietoaineistoa koskevan sääntelyn sisältämällä laatuvaatimuksilla:** Viranomaisten eri tehtävissä ja käyttötarkoituksissa käsiteltävien tietojen laatuvaatimukset perustuvat aina tehtävää ja siinä käsiteltäviä tietoja koskevaan sääntelyyn. Tehtävän luonne ja siinä tarvittavien tietojen käyttötarkoitus muodostavat pohjan erilaisille tehtäväkohtaisille laatuvaatimuksille, jotka on hyvä kuvata tietoaineiston kuvailutiedoissa. Esimerkiksi tiedon ajantasaisuutta ja virheettömyyttä koskevat laatuvaatimukset ovat erilaiset hallinnon asiakkaan oikeuksiin suoraan vaikuttavaa hallintopäätöstä tehtäessä kuin suunniteltaessa viranomaisen omaa toimintaa. Tiedon laatua kuvatessa on siis otettava huomioon tiedonhallinnan yleislainsäädännöstä, tietosuojasääntelystä sekä erityislainsäädännöstä muodostuvat vaatimukset.
- **Mittariston soveltaminen toteutetaan aina sillä tasolla millä tietoja on arvioitavassa tietoaineistossa kuvattu:** Esimerkiksi jos yritystiedot on summattu toimialatasolle, niin puuttuvien kohdeyksiköiden osuus -mittaria sovelletaan puuttuvien toimialojen tasolla.
- **Mikäli mittarin tarkan kuvauksen, esimerkiksi numeroarvon, tuottaminen on hankalaa, voi määrittelyssä käyttää myös arviota.** Mieluummin siis arvio kuin tiedon laadun kannalta merkityksellisen mittarin kokonaan käyttämättä jättäminen. Kun mittarin arvon kohdalla käytetään arviota, tulee se mainita ja perustella.
- **Laatukriteerit ja mittaristo ovat tiivis kokonaisuus ja niiden eri elementit linkittyvät toisiinsa** tavalla, jossa laadun parantuminen yhtäällä voi merkitä sen heikkenemistä toisaalla. Esimerkiksi jos tavoitellaan tietoaineiston täydellistä kattavuutta tai ominaisuustietojen erityisen hyvää tarkkuutta niin tiedon ajantasaisuus yleensä heikkenee.

Kaikkia mittareita yhdistävät seuraavat asiat:

- Mittari kuvaa yksityiskohtaisemmin laatukriteerin mukaista laatu näkökulmaa kohteenaan olevasta tietoaaineistosta.
- Mittari kuvaa laatua ja erityisesti laatukriteeriään selkeästi, yksikäsitteisesti ja ymmärrettävästi.
- Mittarien tulee olla yhteisiä, kaikille käyttäjille soveltuvia ja hyvin erilaisista tietoaaineistoista kohtalaisen helposti tuotettavia.
- Jokaisella laatukriteerillä on vähintään yksi mittari.
- Kullekin mittarille on sovittu yksi yhteinen suositeltu esittämismuoto tai tunnusluku.

## 2.4 Laatukriteeriryhmä: Miten tieto kuvaa todellisuutta?

Tiedon todellisuutta kuvaava laatukriteeriryhmä muodostuu tiedon virheettömyydestä, tarkkuudesta, johdonmukaisuudesta, ajantasaisuudesta ja kattavuudesta.

Todellisuutta kuvaavan tiedon tarkastelu lähtee liikkeelle ilmiöstä, jota koskevaa tietoa halutaan tarkastella tai käyttää. Tietoaaineiston hyödynnettävyyden arviointiin liittyvät tiedon sisällölliset ominaisuudet: mitä näkökulmia tarkasteltavasta ilmiöstä tiedon halutaan ja tarvitaan kuvaavan.

Tiedon laatua käyttötilanteessa ja käyttäjälleen kuvaa se, miten hyvin tieto vastaa tavoiteltua sisältökokonaisuutta. Laadukkaan tiedon tulisi kuvata tavoiteltua ilmiötä mahdollisimman virheettömästi ja tarkasti niin, että tietoaaineiston systemaattiset vääristymät tai muut virhelähteet on tunnistettu ja niiden vaikutus on korjattu. Johdonmukainen tieto ei sisällä sisäisiä ristiriitoja.

Ajantasaisuus-laatukriteerissä on mukana useita näkökulmia tiedon tuoreuden tarkasteluun. Tiedon tulisi olla mahdollisimman tuoretta suhteessa viiteajankohtaan eli ajankohtaan, jota tieto kuvaa. Toisaalta päivittämätönkään tieto ei välttämättä ole huonolaatuista, jos ominaisuuden arvoissa ei ole tapahtunut muutoksia. Tiedon ajantasaisuus on merkittävässä roolissa tietoaaineistojen luovutuksissa teknisten rajapintojen tai katseluyhteyksien avulla.

Tiedon kattavuuden avulla kuvataan sekä kohdejoukkoa, jonka ominaisuuksia tietoa-  
ineiston on tarkoitus kuvata, että sitä, miten hyvin tavoiteltu kohdejoukko on tietoi-  
neistossa mukana. Yleensä tietoa-ineiston aineistokuvaus sisältää laajasti tietoa siitä,  
mikä tietoa-ineiston tavoiteltu sisältö on. Kattavuuden laatukriteeriin sisältyy myös nä-  
kökulma siitä, miten laajasti näitä tavoiteltuja ominaisuustietoja kuvattuun tietoa-ineis-  
toon sisältyy.

## 2.4.1 Laatukriteeri: Virheettömyys

Synonyymi: oikeellisuus

Kuvaus: Virheettömyys kuvaa sitä, miten tietoa-ineiston tiedot vastaavat todellisuutta.  
Tiedon virheettömyyttä tarkastelemalla voidaan saada kiinni myös tietoa-ineiston sys-  
temaattisia vääristymiä. Tiedonhallintalain 15.1 §:ssä säädetysti viranomaisen tulee  
varmistaa tietoa-ineistojensa virheettömyys, joten virheettömyyteen liittyvät tiedot tulee  
lähtökohtaisesti kuvata tietoa-ineiston metatietoihin.

Esimerkki: Tieto on virheetöntä tilanteessa, jossa verotuksessa kerrottu palkka on  
sama kuin todellisuudessa maksettu palkka.

### 2.4.1.1 Menetelmällisesti tuotetut arvot -mittari

Menetelmällisesti tuotetut arvot -mittari kertoo tietoa-ineiston menetelmällisesti tai kor-  
vaavaa tietoa käyttäen tuotettujen ominaisuuden arvojen osuuden ominaisuuden kai-  
kista arvoista.

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

Mittarin arvo tuotetaan menetelmällisesti, siis laskennallisesti, tai korvaavaa tietoa  
käyttäen tuotettujen ominaisuuden arvojen lukumäärän suhtena ominaisuuden kaik-  
kien arvojen lukumäärään. Menetelmällisesti tuotettujen arvojen käyttämisestä on esi-  
merkiksi edellisen vuoden tietojen käyttäminen tietoa-ineistojen puuttuvien tietojen täy-  
dentämisessä.

Menetelmällisesti saatu tai korvaava tieto tarkoittavat mittarin kohdalla sellaisia arvoja, jotka eivät ole täsmällisesti kohdeyksikön tietoaineistossa saamia oikeita arvoja. Tiedon korjaaminen suoraan tiedonantajalta selvittämällä saadulla oikealla arvolla ei kuulu tämän mittarin tarkasteluun.

Esimerkit: Pienten yritysten toimintaa kuvaavia tietoja tarkastellaan harvemmin kuin kerran vuodessa ja arvoina käytetään aiemmassa tiedustelussa saatuja tietoja. Tilastoaineiston käsittelyssä tulotiedon täydentäminen luovuttajaimputoinnilla käyttäen sukupuoliä, ikää, koulutusta ja ammattia arvon luovuttajan valinnan kriteerinä.

#### 2.4.1.2 Virheelliset ominaisuusarvot -mittari

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

- Arviointitaso: ominaisuus
- Arvon esitysmuoto: prosenttia/ei relevantti
- Mittarin tausta: ISO 19157 Geographic information - Data quality

Mittarin arvo ilmoitetaan virheellisen ominaisuusarvon saavien kohdeyksiköiden lukumäärän suhteena kaikkien ominaisuuden arvon saavien kohdeyksiköiden lukumäärään.

Arvon tuottamisessa voi hyödyntää vertailuun soveltuvaa tietoaineistoa tai tehdä laadun tarkastusta otosperusteisesti. Mikäli mittarin täsmällisen arvon laskeminen on haastavaa, voidaan prosenttiarvoksi antaa kokemukseen perustuva arvio.

Esimerkki: Tietoaineistossa tiedetään olevan päivittämättömiä tietoja ja kokemuksen perusteella osataan arvioida päivitysten kohdistuvan 5 % ominaisuuden arvoista.

#### 2.4.1.3 Väärinluokittelu-mittari

Väärinluokittelu-mittari kertoo tietoaineiston virheellisesti luokiteltuja ominaisuuden arvoja saavien kohdeyksiköiden osuuden kaikista kohdeyksiköistä.

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

- Mittarin tausta: ISO 19157 Geographic information - Data quality

Mittarin arvo ilmoitetaan väärinluokiteltuja ominaisuuden arvoja saavien 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.

Arvon tuottamisessa voi hyödyntää vertailuun soveltuvaa tietoaineistoa tai tehdä laadun tarkastusta otosperusteisesti. Mikäli mittarin täsmällisen arvon laskeminen on haastavaa, voidaan antaa kokemukseen perustuva arvio prosenttiarvoksi.

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

## 2.4.2 Laatuksiteeri: Tarkkuus

Synonyymi: harhattomuus

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

Esimerkit: Mitattujen arvojen hajonta tai tietoaineiston poikkeavien havaintojen osuus sekä lisäksi luokittelun osuvuus ja mittaustaso, esimerkiksi desimaalit, ajan tai koordinaattien mittaustaso. Väärässä mittakaavassa käytettynä tietoaineiston, esimerkiksi kartta- tai kuvamuotoisen aineiston, tarkkuus voi muuttua ja vaikuttaa aineiston käytettävyyteen.

### 2.4.2.1 Keskihajonta-mittari

Keskihajonta-mittari kuvaa ominaisuuksien arvojen vaihtelua keskiarvon ympärillä. Mittari antaa tiedon käyttäjälle käsitystä ominaisuuden arvojen hajonnasta, siis siitä, miten laajaa ominaisuuden arvojen vaihtelu on.

- Arviointitaso: ominaisuus
- Arvon esitysmuoto: keskihajonta

Mittaria soveltaessa on hyvä ottaa huomioon, että keskihajonnan arvon rinnalla tulisi tiedon käyttäjän saatavilla olla myös tieto ominaisuuksien arvojen keskiarvosta, jotta arvojen vaihtelua voisi paremmin hahmottaa. Ilman tietoa keskiarvosta vaihtelu jää tarkasteltavasta asteikosta erilliseksi. Keskiarvo ei ole suosituksessa mukana erillisenä mittarina, koska se ei sinällään kuvaa tarkkuutta. Keskiarvon tuottaminen on olennaista myös siksi, että sekä keskiarvoa että keskihajonnan arvoa tarvitaan tietoaineiston poikkeavien havaintojen määrittämisessä.

Esimerkit: Keskihajonnan ollessa pientä, ominaisuuden arvot ovat keskittyneet lähelle keskiarvoa ja vastaavasti suuren keskihajonnan tapauksessa arvot ovat enemmän hajallaan. Tämä voi liittyä epätarkkuuteen tai olla ominaisuudelle tyypillistä vaihtelua.

#### 2.4.2.2 Poikkeavat havainnot -mittari

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

- Arviointitaso: ominaisuus
- Arvon esitysmuoto: prosenttia
- Mittarin tausta: ISO 25024 Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - Measurement of data quality

Mittarin arvo ilmoitetaan poikkeavien havaintojen lukumäärän suhteena kaikkien kohdeyksiköiden lukumäärään.

Poikkeavalla havainnolla tarkoitetaan kohdeyksikköä, joka saa havaittujen arvojen enemmistöstä suuresti poikkeavan arvon. Poikkeava havainto voi olla oikea tai virheellinen arvo.

Tyypillisenä raja-arvona poikkeavalle havainnolle pidetään 2,5 kertaa keskihajonnan verran keskiarvosta mitattuja arvoja ja tätä raja-arvoa käytetään myös tämän mittarin yhteydessä. Poikkeavia havaintoja ovat siis arvot, jotka ovat pienempiä kuin arvo, joka saadaan, kun keskiarvosta vähennetään 2,5 kertaa keskihajonta tai suurempia kuin arvo joka saadaan, kun keskiarvoon lisätään 2,5 kertaa keskihajonta. Monissa tilastollisissa ohjelmistoissa poikkeavat havainnot on helppo määritellä koneellisesti.

Esimerkki: Tarkastellessa henkilöiden tulotietoja, huomattavan suurituloiset ovat henkilöaineistossa poikkeavia havaintoja. Huomattavan suurituloisten tulotiedot voivat aiheuttaa haastetta analyyseissä.

## 2.4.3 Laatuksriteeri: Johdonmukaisuus

Synonyymit: säännönmukaisuus, tiedon looginen eheys

Kuvaus: Johdonmukaisuus kertoo tietoaineiston yhtenäisyydestä ja ristiriidattomuudesta. Sitä voidaan tarkistaa tarkistus- tai laatusäännöillä. Johdonmukaisuudella voidaan kuvata myös eri tietoaineistojen keskinäistä johdonmukaisuutta.

Esimerkit: Seuraavat tilanteet eivät ole johdonmukaisia: asuinrakennuksella ei ole yhtään asuntoa tai henkilön avioliiton solmimispäivä on aikaisempi kuin syntymäpäivä.

### 2.4.3.1 Tieto loogisuustarkastettu -mittari

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 tai laatusääntöjä käyttäen. Tietoaineistotasolla on hyvä kuvata tarkemmin miltä osin loogisuusehtoja on käytetty. Tiedon käyttäjää kiinnostavat myös yksityiskohdat käytetyistä loogisuusehdoista ja ne olisi hyvä tuoda kuvauksessa esille.

Esimerkki: Ominaisuuden ”toimipaikan osoite” osalta on käytetty loogisuusehtoa vastaako toimipaikan sijaintiosoitteen postinumero toimipaikan sijaintikuntaa.

## 2.4.4 Laatuksriteeri: Ajantasaisuus

Synonyymit: -

Kuvaus: Ajantasaisuus kuvaa tietoaineiston tietojen täsmällisyyttä ajan suhteen. Ajantasaisuus on sitä parempi mitä lähempänä tiedon viiteajankohta on nykyhetkeä. Viranomaisen tulee tiedonhallintalain 15.1 §:ssä säädetysti varmistaa tietoaineistojensa ajantasaisuus, joten tietoaineiston metatietojen tulee lähtökohtaisesti sisältää tietoa tietoaineiston ajantasaisuudesta.

Esimerkit: Tiedon tuotannossa tarkastellaan erilaisia tiedon tarkastus- ja muutosajankohtia.

#### 2.4.4.1 Viiteajankohta-mittari

Viiteajankohta-mittari kertoo ajankohdan, jolloin tietoaineiston sisältämä tieto on kerätty eli ajankohdan, jota tieto kuvaa. Se on kerrottu tiedon yhteydessä. Tietoaineiston käsittely aiheuttaa viivettä ja tietoa on voitu kerätä aiemmin kuin tietoaineisto on valmis tiedon käyttäjän käsiteltäväksi. Viiteajankohdan avulla pystytään päättelemään tiedon tuoreus.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: ajankohta/ei relevantti

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

Esimerkki: Tilaston viiteajankohta voi olla useammankin kuukauden takainen, koska tietojenkeruu ja -käsittely vaatii aikansa ennen julkaisemista. Viiteajankohta voi olla esimerkiksi vuoden alun ja lopun välinen aika tai yhden tietyn päivän tilanne.

#### 2.4.4.2 Luontiajankohta-mittari

Luontiajankohta-mittari kertoo ajankohdan joka on kohdeyksikön tai ominaisuuden luontiajankohta. Mittari vastaa myös kysymykseen mistä ajankohdasta, esimerkiksi päivämäärästä, lähtien ominaisuus on ollut tietoaineistossa tai mistä alkaen tietoaineistoa on koottu. Mittarin tarkoituksena on lisäksi kertoa siitä, miten pitkältä ajanjaksoilta tietoja on saatavilla.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: ajankohta



Mittarin soveltamisessa tulee ottaa huomioon, että luontiajankohta on tieto, joka on yleensä osa tietoaineiston metatietoja.

Mittarin arvo kertoo ominaisuuden tai tietoaineiston luontiajankohdan. Kohdeyksikön luontiajankohta ei välttämättä ole sama kuin tietoaineiston luontiajankohta. Luontiajankohdan lisäksi aineiston kuvauksissa on tärkeää kertoa vertailukelpoisten tietojen ajanjaksot.

Esimerkki: Luontiajankohta on päivämäärä, jolloin viranomaisen asiakirja tai tietoaineisto on luotu. Esimerkiksi kaupparekisteriotteen luontiajankohta.

#### 2.4.4.3 Tarkastusajankohta-mittari

Tarkastusajankohta-mittari kertoo ajankohdan, esimerkiksi päivämäärän, joka on kohdeyksikön tai ominaisuuden tarkistusajankohta. Mittari vastaa myös kysymykseen, milloin tietoja on viimeksi tarkistettu.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: ajankohta/ei relevantti

Mittarin soveltamisessa tulee ottaa huomioon, että tarkastusajankohta on tieto, joka on yleensä osa tietoaineiston metatietoja.

Mittarin arvo kertoo ominaisuuden arvojen tai tietoaineiston tietojen viimeisimmän tarkistusajankohdan.

Esimerkki: -

#### 2.4.4.4 Muutosajankohta-mittari

Muutosajankohta-mittari kertoo ajankohdan, esimerkiksi päivämäärän, joka on kohdeyksikön tai ominaisuuden muutosajankohta. Mittari vastaa myös kysymykseen, milloin tietoja on viimeksi päivitetty. Jatkuvasti tai usein päivittyvien tietoaineistojen kohdalla mittari ei ole mielekäs, mutta tietojen muutosajankohdat on syytä olla saatavilla tietojen elinkaaren todentamiseksi.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: ajankohta/ei relevantti

Mittarin soveltamisessa tulee huomioida, että muutosajankohta on tieto, joka on yleensä osa tietoaaineiston metatietoja.

Mittarin arvo kertoo ominaisuuden arvojen tai tietoaaineiston muutosajankohdan.

Esimerkki: -

## 2.4.5 Laatuksiteeri: Kattavuus

Synonyymi: Täydellisyys

Kuvaus: Kattavuus kuvaa tietoaaineiston tavoitellun ajallisen ja alueellisen kattavuuden sekä tavoitellut kohdeyksiköt ja ominaisuustiedot. Toisaalta kattavuus kertoo miltä osin tietoaaineisto sisältää tavoiteltuja tietoja tai sisältääkö se kaikki määritellyt ominaisuustiedot tietoaaineistossa oleville kohdeyksiköille.

Esimerkit: Tietoaaineisto kattaa määritellyn tarkastelualueen yksiköt, esimerkiksi kaikki Suomen kunnat tai yritykset. Tietoaaineiston määriteltyjen ominaisuustietojen osalta esimerkiksi tietoaaineistossa olevien Suomen kuntien väkiluku- ja pinta-ala tiedot tai onko kaikilla yrityksillä esimerkiksi osoite- tai liikevaihtotieto. Alueellinen kattavuus kertoo, ovatko kaikki käsiteltävät alueet mukana, esimerkiksi kaikki Suomen kunnat tai kattaako tietoaaineisto Ahvenanmaan.

### 2.4.5.1 Ajallinen tavoitekattavuus -mittari

Ajallinen tavoitekattavuus -mittari kertoo, että tietoaaineiston tavoiteltava ajallinen kattavuus ja tiheys on kuvattu. Ajallisella tavoitekattavuudella tarkoitetaan aikaväliä, jota tietoaaineiston on tarkoitus kuvata, sekä tiheyttä, jolla ominaisuuksien arvoja on mitattu.

- Arviointitaso: tietoaaineisto
- Arvon esitysmuoto: kyllä/ei; lisätieto/tarkennus voidaan kuvata

Mittarin arvo kertoo onko tietoaaineiston ajallinen tavoitekattavuus kuvattu esimerkiksi aineistokuvauksessa.

Esimerkki: FinTerveys 2017 -tutkimuksen perusjoukko on Manner-Suomessa asuva vuonna 2017 18 vuotta täyttänyt väestö.

### 2.4.5.2 Alueellinen tavoitekattavuus -mittari

Alueellinen tavoitekattavuus -mittari kertoo, että tietoaineiston tavoiteltava alueellinen kattavuus ja tiheys on kuvattu. Alueellisella tavoitekattavuudella tarkoitetaan sitä maantieteellistä aluetta, jonka tietoa tietoaineiston on tarkoitettu kuvaavan. Tiheydellä viitataan siihen, miten tarkasti aluetaso on tietoaineistossa kuvattu.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kyllä/ei; lisätieto/tarkennus voidaan kuvata

Mittarin arvo kertoo onko tietoaineiston alueellinen tavoitekattavuus kuvattu esimerkiksi aineistokuvauksessa.

Esimerkki: FinTerveys 2017 alueellinen kattavuus on Manner-Suomi.

### 2.4.5.3 Tavoitekohdeyksiköt-mittari

Tavoitekohdeyksiköt-mittari kertoo siitä, että tietoaineiston aineistokuvauksessa on selkeästi määritelty ajallisen ja alueellisen tavoitekattavuuden lisäksi muut tietoaineiston rajaukset.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kyllä/ei; lisätieto/tarkennus voidaan kuvata
- Mittarin tausta: ISO 19157 Geographic information - Data quality

Mittaria soveltaessa tulee ottaa huomioon tavoitekohdeyksiköiden kuvauksessa kirjatut olennaiset kohdeyksiköiden rajaukset. Ajallisen ja alueellisen kattavuuden mittarit ovat mukana omina mittareinaan. Katso tarkemmin alaluvut 2.4.5.1 Ajallinen tavoitekattavuus -mittari ja 2.4.5.2 Alueellinen tavoitekattavuus -mittari.

Esimerkit: Tietoaineisto kattaa tiettyjen toimialojen yritykset. Tietoaineisto sisältää tiedot rakennuksista, joiden koko on yli 10 m<sup>2</sup>.

### 2.4.5.4 Tavoiteominaisuuksissa puutteita -mittari

Tavoiteominaisuuksissa puutteita -mittari kertoo tietoaineiston kuvaaman ilmiön kannalta keskeisten ominaisuuksien puutteista arvioitavassa tietoaineistossa. Puutteellisuutta on tärkeä kuvata tarkemmin esimerkiksi tietoaineistokuvauksessa.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kyllä/ei

Mittaria soveltaessa tulee ottaa huomioon, että tietoaineiston laadun arvioinnin kannalta on tärkeää tunnistaa tietoaineiston kuvaamat ilmiöt ja näiden ilmiöiden kuvaamiseen tai mittaamiseen tarvittavat ominaisuudet. Mikäli tietoaineistosta puuttuu joku tietty tarkastelun näkökulma eli joitakin ominaisuuksia, on tämä kuvattava selkeästi tiedon käyttäjälle.

Esimerkki: -

#### 2.4.5.5 Puuttuvat kohdeyksiköt -mittari

Puuttuvat kohdeyksiköt -mittari kuvaa tietoaineiston alipeittoa eli tietoaineiston tavoitekohdejoukosta puuttuvien kohdeyksiköiden prosentuaalista osuutta. Mikäli tietoaineisto sisältää puuttuvia kohdeyksiköitä ja tätä puuttuvien kohdeyksiköiden joukkoa on mahdollista kuvata, on tärkeää sisällyttää kuvaus tietoaineiston kuvaukseen.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: prosenttia
- Mittarin tausta: ISO 19157 Geographic information - Data quality

Mittarin arvo ilmoitetaan tietoaineistosta puuttuvien kohdeyksiköiden lukumäärän suhteena tietoaineistoon kuuluvien kohdeyksiköiden lukumäärään. Mikäli täsmällisen arvon laskeminen on haastavaa, voidaan antaa kokemukseen perustuva arvio prosenttiarvona.

Mittaria soveltaessa tulee ottaa huomioon, että valmiiksi koostetun tiedon, esimerkiksi tilaston, osalta puuttuvuus on yleensä korjattu tilastollisin menetelmin. Tällöin voidaan olettaa, ettei tarkasteltavassa tietoaineistossa ole puuttuvia kohdeyksiköitä. Puuttuvia kohdeyksiköitä voi kuitenkin halutessaan tarkastella raportointitasolla esimerkiksi toimiala-tasolla yritystason sijaan.

Esimerkki: -

#### 2.4.5.6 Ylimääräiset kohdeyksiköt -mittari

Ylimääräiset kohdeyksiköt -mittari kuvaa tietoaineiston ylipittoa eli tietoaineistossa mukana olevien, mutta tietoaineistoon kuulumattomien ja siten ylimääräisten kohdeyksiköiden prosentuaalista osuutta. Mikäli tietoaineisto sisältää tietyn tyyppisiä ylimääräisiä kohdeyksiköitä ja tätä ylimääräisten kohdeyksiköiden joukkoa on mahdollista kuvata, on tärkeää sisällyttää kuvaus niistä tietoaineiston kuvaukseen.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: prosenttia
- Mittarin tausta: ISO 19157 Geographic information - Data quality

Mittarin arvo ilmoitetaan ylimääräisten kohdeyksiköiden lukumäärän suhteena tietoaineiston tavoiteltavien kohdeyksiköiden lukumäärään. Mikäli täsmällisen arvon laskeminen on haastavaa, voidaan antaa kokemukseen perustuva arvio prosenttiarvoksi.

Mittaria soveltaessa tulee ottaa huomioon, että valmiiksi koostetun tiedon, esimerkiksi tilaston, osalta ylipittoa on yleensä pyritty jo korjaamaan.

Esimerkit: Tietoaineistossa on samoja kohdeyksiköitä useamman kerran, koska tietoja on tuotu tietoaineistoon useasta lähteestä ja kaikilla ei ole sama kohdeyksikön tunnistus käytössä. Otostutkimuksen kohdeyksiköiden otannassa on mukana maasta muuttaneita henkilöitä, koska osoitetiedot eivät ole päivittyneet otoksen pohjana olevaan tietokantaan.

#### 2.4.5.7 Puutteelliset kohdeyksiköt -mittari

Puutteelliset kohdeyksiköt -mittari kuvaa tietoaineiston puuttuvia ominaisuustietoja omaavien kohdeyksiköiden osuuden kaikista kohdeyksiköistä. Tarkastelussa keskitytään kyseiselle kohdeyksikölle relevantteihin puuttuviin arvoihin. Esimerkiksi lapsille ei kuulukaan olla ammattiluokituksen arvoa.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: prosenttia

Mittarin arvo ilmoitetaan yhdenkin puuttuvan ominaisuustiedon omaavien kohdeyksiköiden lukumäärän suhteena kaikkien kohdeyksiköiden lukumäärään.

Rakenteellista puuttuvuutta, eli tilannetta, jossa jotain ominaisuuden arvoa ei kuulu-kaan tulla tietoaaineiston tietylle kohdeyksikölle tai kohdeyksikkö ei sellaista voisi mis-sään tapauksessa saada, ei oteta tämän mittarin kohdalla huomioon. Esimerkiksi tu-loeriä sellaisen tuon osalta, jota tarkasteltava henkilö ei saa.

Mittaria soveltaessa tulee ottaa huomioon, että mittari ei välttämättä tuota mielekkäitä arvoja koostettua tietoa tietoaaineistotasolla arvioitaessa. Vastaavasti laajojen yksikkö-tason tietoaaineistojen tai pitkien aikasarjojen tapauksessa puutteellisten kohdeyksiköi-den osuus voi kasvaa erittäin suureksi, koska yksittäisiä tietoja saattaa puuttua mel-kein jokaiselta kohdeyksiköltä. Ominaisuuksittaisen tarkastelun kohdalla on hyvä ku-vata tietoaaineistoon sisältyvää rakenteellista puuttuvuutta, jos se kohdistuu tiettyyn kohdejoukkoon.

Esimerkki: -

#### 2.4.5.8 Puutteelliset ominaisuustiedot -mittari

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

- Arviointitaso: tietoaaineisto
- Arvon esitysmuoto: prosenttia

Mittarin arvo ilmaistaan puuttuvan ominaisuustiedon omaavien kohdeyksiköiden luku-määrän suhteena kaikkien kohdeyksiköiden lukumäärään ominaisuuksittain tarkastel-tuna.

Rakenteellista puuttuvuutta, ei oteta tämän mittarin kohdalla huomioon, vaan tarkas-tellaan puuttuvuutta silloin, kun jokin arvo on relevantti kyseiselle kohdeyksikölle.

Esimerkki: -

## 2.5 Laatuksiteeriryhmä: Miten tieto on kuvattu?

Tiedon kuvaamista käsittelevä laatuksiteeriryhmä muodostuu tiedon alkuperäisyy-destä, metatietojen ymmärrettävyydestä ja suositustenmukaisuudesta.

Laadukkainkin tietoaineisto ilman sen sisältämän tiedon kuvausta voi pahimmillaan olla käyttökeltovotonta. Tiedon alkuperäisyys kertoo ovatko tiedon alkuperä ja muutokset tiedossa ja jälkikäteen tarkasteltavissa. Tietoaineiston metatietojen ymmärrettävyys vaativat sekä aineistokuvauksen että ominaisuuksien kuvaukset, jotta tietoaineistolle ja sen sisältämien ominaisuuksien arvoille on muodostettavissa mielekäs ja ymmärrettävä tulkinta. Tietoaineistojen ja niiden sisältämien tietojen toisiin tietoaineistoihin ja tietoihin yhteensovittamisen kannalta erilaisten standardien, suositusten ja hyvien käytäntöjen noudattaminen on olennaista.

## 2.5.1 Laatuksriteeri: Alkuperäisyys

Synonyymi: jäljitettävyys, kiistämättömyys

Kuvaus: Alkuperäisyys kertoo, että tietoaineistoon ja sen tietoihin tehdyt muutokset voidaan jäljittää. Tiedon alkuperä tunnetaan. Viranomaisen tulee tiedonhallintalain 15.1 §:ssä säädetysti varmistaa tietoaineistojensa alkuperäisyys, joten tietoaineiston metatietojen tulee lähtökohtaisesti sisältää tietoa tietoaineiston alkuperäisyydestä.

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

### 2.5.1.1 Tietolähde-mittari

Tietolähde-mittari kertoo kuinka suurella osalla kohdeyksiköitä tai ominaisuustietoja lähdetiedot ovat saatavilla. Tietolähteen lisäksi mittarin yhteydessä voidaan kuvata tietoihin tehtyjä muokkauksia. Kuvauksien avulla tiedonkäyttäjä saa lisätietoa tiedon prosessointihistoriasta. Jos tietolähde on vaillinainen ja vaillinaisuus koskee esimerkiksi tiettyä kohdeyksikköjoukkoa, on tietoaineiston kuvaustiedoissa hyvä kuvata näitä puutteita.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: prosenttia

Mittarin arvo ilmoitetaan saatavilla olevien lähdetietojen suhteena kaikkiin tietoihin.

Mittaria soveltaessa on hyvä ottaa huomioon, että siinä keskitytään tarkastelemaan tietolähdettä, josta tiedot on suoraan saatu. Mahdolliset aiemmat tiedon vaiheet on tällöin mahdollista selvittää tiedon aiemman raportoinnin laaturaporttien kautta.

Lisäksi on hyvä pitää mielessä, että tietolähteet ja tiedot puuttuvista lähdetiedoista kiinnostavat tietoaineiston käyttäjää. Tietolähteistä on olennaista kertoa tiedon toimitanut taho, jolloin laatutarkastelussa voidaan haluttaessa ottaa huomioon ketjumaisesti tiedon laatuun liittyviä näkökulmia myös aiempien toimijoiden osalta.

Esimerkki: -

### 2.5.1.2 Tiedon elinkaari -mittari

Tiedon elinkaari -mittari kertoo, onko tietoaineiston elinkaari määritelty ja kuvaus siitä saatavilla. Mittari kuvaa tiedon muutokset sen luomisesta tiedon poistamiseen tai arkistointiin ja se sisältää tiedon lähteiden tai tuottamisen muutokset, esimerkiksi laskentatavan muutokset.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kyllä/ei

Mittarin arvo kertoo, onko tiedon elinkaari määritelty ja kuvattu.

Mittaria soveltaessa tulee ottaa huomioon, että tiedon syntyyn ja tuottajaan liittyvät asiat ovat olennaisia tietoja tiedon käytettävyyden tarkastelussa.

Tiedon elinkaari alkaa tiedon tuottamis- tai vastaanottovaiheessa ja jatkuu tiedon pysyvään säilyttämiseen päättyen sen arkistointiin tai tuhoamiseen. Elinkaari kattaa siis kaikki tiedon käsittelyn vaiheet: tiedon tuottaminen tai vastaanotto, säilytys, käyttö, jatkaminen, siirto sekä arkistointi tai tuhoaminen. Elinkaariajattelun lähtökohtana on tiedon suunnitelmallinen ja riskilähtöinen käsittely sekä hallinta osana tiedonhallintayksikön toimintaa.

Tiedon elinkaaren osalta on hyvä muistaa, että tietoaineistoja käsitellään useassa eri sijainnissa ja tietojärjestelmässä tai laitteistossa, joista jokaisessa tiedolla voi olla oma elinkaarensa, ja että tietoaineistoon sisältyvän tiedon elinkaari on yleensä pidempi kuin yksittäisen tietoaineiston tai tietojärjestelmän elinkaari.

Esimerkki: -



### 2.5.1.3 Muutosten hallinta -mittari

Muutosten hallinta -mittari kertoo, että ominaisuustiedoissa tapahtuvia rakenteellisia tai lähdetietojen muutoksia seurataan. Muutoksia on hyvä kuvata yleisellä tasolla myös tietoaaineistojen metatiedoissa.

- Arviointitaso: ominaisuus
- Arvon esitysmuoto: kyllä/ei

Mittarin arvo kertoo siitä, tallennetaanko tieto ominaisuustietojen muutoksista muutoksia tehtäessä siten, että ominaisuus- tai lähdetietojen muutoksia on mahdollista tarkastella myöhemminkin.

Mittaria soveltaessa on hyvä tuoda esille, miten muutoshistoriaa on mahdollista päästä tarkastelemaan.

Esimerkit: Tietoaaineiston ominaisuuksien muutoksista pidetään lokia tai tieto on löydettävissä tietoaaineiston metatiedoista.

## 2.5.2 Laatuksiteeri: Metatietojen ymmärrettävyys

Synonyymit: tulkittavuus, käsitettävyyys

Kuvaus: Metatietojen ymmärrettävyys kuvaa tässä sitä, miten kattavasti tietoaaineistolla on metatietoja, jotka auttavat tiedon sisällön ja merkityksen oikeassa tulkinnassa sen käyttöä suunniteltaessa tai sitä käytettäessä.

Esimerkit: Tietoaaineisto ja sen ominaisuustiedot on esitetty metatiedoissa riittävällä tasolla ja tarkkuudella 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 tietoaaineiston metatietoihin.

### 2.5.2.1 Aineistokuvaus-mittari

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

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kieliversiot

Mittarin arvo ilmoitetaan kertomalla ne kielet, joilla aineistokuvaus on saatavilla.

Mittaria soveltaessa on hyvä pitää mielessä, että kieliversio voi mainita vaikka kaikki kieliversiot eivät olisikaan yhtä kattavia. Aineistokuvauksen osalta myös määritelmät ja yhtenäinen termistö ovat tärkeitä. Niiden avulla tiedon käyttäjä voi olla varma, että tietoaineistot ovat vertailukelpoisia.

Esimerkki: -

### 2.5.2.2 Käsitteiden määritelmät -mittari

Käsitteiden määritelmät -mittari kertoo, onko tietoaineiston käsitteet määritelty selkeästi ja saatavilla sekä mitä kieliversioita on saatavilla. Käsitteiden määrittelemine on tärkeä osa tiedon kuvausta, jotta voidaan varmistua siitä, että samoilla sanoilla ymmärretään sama asia. Käsitteillä tarkoitetaan niitä tietoaineistoon liittyviä asioita, joita ei voida välttämättä suoraan mitata.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kieliversiot

Mittarin arvo ilmaistaan kertomalla ne kielet, joilla käsitteiden kuvaukset ovat saatavilla.

Mittaria soveltaessa on hyvä pitää mielessä, että kieliversio voi mainita vaikka kaikki kieliversiot eivät olisikaan yhtä kattavia.

Esimerkki: Hyvinvointi-käsitteen osalta kuvataan tietoaineiston määritelmä hyvinvoinnille. Eli kuvataan ne näkökulmat, joiden kautta hyvinvointia mitataan käsiteltävässä tietoaineistossa.

### 2.5.2.3 Ominaisuuksien tietokuvaukset -mittari

Ominaisuuksien tietokuvaukset -mittari kertoo, ovatko tietoaineiston ominaisuuksien kuvaukset ja koodistot saatavilla ja mitä kieliversioita niistä on saatavilla. Ominaisuuksilla tarkoitetaan tässä yhteydessä tietoaineiston mitattavia ja arvoja saavia ominaisuuksia.

- Arviointitaso: ominaisuus
- Arvon esitysmuoto: kieliversiot

Mittarin arvo ilmaistaan kertomalla ne kielet, joilla kuvaus on saatavilla.

Mittaria soveltaessa on hyvä pitää mielessä, että kieliversio voi mainita vaikka kaikki kieliversiot eivät olisikaan yhtä kattavia. Kuvauksien osalta myös määritelmät ja yhtenäinen termistö ovat tärkeitä. Niiden avulla käyttäjä voi olla varma, että tietoaineistot ovat vertailukelpoisia.

Esimerkki: Kunnat aluejakona -paikkatietoaineisto, jossa ominaisuustiedot on kuvattu sarakeittain: Kuntanro/ID(integer), Kuntanimi (char) ja saatavilla olevat kieliversiot, pinta-ala jne.

### 2.5.2.4 Metatietojen ymmärrettävyyden asiakaspalaute -mittari

Metatietojen ymmärrettävyyden asiakaspalaute -mittari kertoo, että tietoaineiston ja sen metatietojen ymmärrettävyydestä on mahdollisuus antaa palautetta olemassa olevan palauteväylän kautta tai kohdennetulla asiakaskyselyllä.

Saatua palautetta seurataan ja palautteeseen reagoidaan. Palautteen yhteenveto tai palautteen aiheuttamat muutokset tietoaineistossa ovat hyödyllistä lisätietoa tiedon käyttäjille. Palautetta voidaan myös hyödyntää tietoaineiston kehittämisessä.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kyllä/ei

Mittaria soveltaessa tulee muistaa, että mittarin on tarkoitus kuvata sitä, onko tiedon käyttäjän mahdollista antaa palautetta ja saada vastauksia antamaansa palautteeseen.

Esimerkki: -

## 2.5.3 Laatuksiteeri: Suositustenmukaisuus

Synonyymit: yhteensopivuus, yhdenmukaisuus, yhteentoimivuus

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

Esimerkit: Kansallista suositustenmukaisuutta tukee muun muassa yhtenäisten kansallisten sanastojen ja koodistojen käyttö tietoaineistojen tietoja suunniteltaessa. Kansainvälistä suositustenmukaisuutta tukevat esimerkiksi EU:n vahvistamat standardiluokitukset ja ISO-kielikoodisto (ISO 639).

### 2.5.3.1 Noudatettavat standardit ja suositukset -mittari

Noudatettavat standardit ja suositukset -mittari kertoo ovatko tietoaineiston noudattamat standardit, suositukset ja julkaistut hyvät käytännöt listattuna tietoaineiston kuvauksissa. Tilanteessa, jossa yleisestä normistosta poiketaan, on tärkeä tuoda poikkeama esiin tietoaineiston metatiedoissa.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: kyllä/osittain/ei
- Mittarin tausta: INSPIRE-paikkatietodirektiivin soveltamissäädökset / FAIR-periaatteet

Mittaria soveltaessa tulee ottaa huomioon, että standardeja ja suosituksia on monenlaisia: yleisiä ja toimialakohtaisia, kansallisia ja kansainvälisiä. Näistä vähintään keskeisimmät tulisi luetella esimerkiksi aineistokuvauksen yhteydessä. Ominaisuuksien osalta myös yleisten koodistojen käyttö tulee ottaa huomioon tietoaineistojen suunniteltaessa ja kerta tietoaineiston yhteydessä.

Esimerkki: Ominaisuus noudattaa kansainvälistä koodistostandardia, mutta kyseistä koodistoa on täydennetty organisaatiokohtaisesti lisäluokilla. Lisäluokat on määritelty koodistostandardista erilliseksi siten, että ne täsmentävät standardia, mutta eivät ole ristiriidassa sen kanssa.

## 2.6 Laatuksiteeriryhmä: Miten tietoa voi käyttää?

Tiedon kuvaamista käsittelevä laatuksiteeriryhmä muodostuu tietoaineiston koneluettavuudesta, tietoaineiston käyttöoikeuksista ja käytettävissä oloon liittyvästä oikea-aikaisuudesta.

Koneluettava tietoaineisto on rakenteistettu siten, että sitä voidaan käsitellä koneellisesti ja käsittely on mahdollista eri tietojärjestelmissä. Koneluettavien tietoaineistojen kuvailutietojen tulee niidenkin olla koneluettavassa muodossa. Tietoaineiston tulee olla rakenteistettu sen metatiedoissa mainitun tietomallin, skeeman tai vastaavan standardin mukaisesti, jotta koneluettavuus on mahdollista ja toteutuu.

Tietoaineistojen käytölle on rajoituksia sen suhteen mikä taho voi tietoja tarkastella, mihin käyttötarkoitukseen niiden käyttö on sallittu ja missä muodossa tiedot ovat saatavilla. Rajoitteita aiheuttavat esimerkiksi tiedonkeruussa tehty tiedonantajan informointi ja tietosuoja. Tiedonantajille suorien tiedonkeruiden yhteydessä kerrotut ja kuvatut tietojen käyttötarkoitukset rajaavat kerättyjen tietojen käyttöä ja yhdistämistä mahdollisiin muihin tietoaineistoihin koko tiedon elinkaaren ajan.

Tiedon oikea-aikaisuus tarkoittaa tietoaineiston käytettävissä olemista ilmoitettuna ajankohtana ja riittävän tiheästi tietoaineistossa tapahtuviin muutoksiin nähden.

### 2.6.1 Laatuksiteeri: Koneluettavuus

Synonyymi: -

Kuvaus: Koneluettavuus kuvaa, onko tietoaineisto rakenteistettu siten, että sitä voidaan käsitellä koneellisesti ja sen käsittely on mahdollista eri tietojärjestelmissä. Viranomaisen on tiedonhallintalain 19.2 §:ssä säädetysti huolehdittava, että tietoaineisto on saatavilla ja hyödynnettävissä koneluettavassa muodossa kuvailutietoineen, jos se on alkujaan tallennettu koneluettavaan muotoon tai voidaan muuttaa alkuperäisestä muodostaan suoraan koneluettavaan muotoon. Tällaisten tietoaineistojen metatietojen tulee lähtökohtaisesti sisältää tietoa tietoaineiston koneluettavasta muodosta.

Esimerkit: Tietoaineisto on rakenteisessa formaatissa. Tällaisia ovat esimerkiksi csv-, json- ja xml-tiedostomuodot. Tietoaineiston rakenne on kuvattu esimerkiksi tietomallina tai skeemana.

### 2.6.1.1 Tietoaineiston tietomalli -mittari

Tietoaineiston tietomalli -mittari kertoo, onko tietoaineisto kuvattu rakenteisesti. Tietomallin osalta on hyvä kertoa minkä tietomallin tai standardin mukaan tietoaineisto on kuvattu.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kyllä/ei, lisätieto/tarkennus käytetty tietomalli tai standardi

Mittarin arvo kertoo, onko aineiston rakenne kuvattu tietomallin tai skeeman tai vastaavan standardin mukaisesti.

Esimerkit: Tietoaineiston rakenne on kuvattu skeemalla, esimerkiksi xml tai json. Tietoaineistossa on käytetty Planned Land Use (PLU) -tietomallia.

### 2.6.1.2 Kohdeyksikön pysyvä tunniste -mittari

Kohdeyksikön pysyvä tunniste -mittari kertoo, että tietoaineiston kohdeyksiköillä, esimerkiksi viranomaisen asiakirjalla, on vähintään tietoaineistokohtainen pysyvä tunniste, jonka perusteella kohdeyksiköt voidaan erottaa toisistaan.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kyllä/ei

Mittaria soveltaessa tulee huomioida, että tarkastelussa riittää tiedonhallintalain edellyttämä tietoaineistokohtainen pysyvä tunniste. Julkisessa hallinnossa on käytössä laaja joukko erilaisia yksilöiviä tunnuksia, joista osa voi olla organisaatio- tai palvelukohtaisia. Kansallisesti yhtenäisen pysyvän tunnisteiden tai kansainvälisen tunnisteiden käyttö on tietoaineiston ja sen erilaisten kohdeyksikköjen laajemman hyödynnettävyyden kannalta eduksi.

Tunnisteiden pysyvyyden lisäksi myös tunnisteiden ainutlaatuisuuteen tulee kiinnittää huomiota ja pyrkiä siihen, ettei tietoaineistossa olisi samaa kohdeyksikköä useampan kertaan.

Esimerkit: Kiinteistötunnus on neliosainen numerosarja, jonka näkee kiinteistöä koskevista otteista kuten lainhuutotodistuksesta. Euroopan unionin paikkatietoinfrastruktuuri -direktiivin (INSPIRE) soveltamissäädökset edellyttävät paikkatietokohteen tunnuksen julkaisemista HTTP URI-muodossa. Muita kansainvälisiä pysyviä tunnisteita

ovat esimerkiksi DOI, URN sekä ORCID-tunnus eli tutkijatunniste, joka on tutkijan yksilöivä digitaalinen tunniste.

Viranomaisen asiakirjojen metatiedoista ja palveluissa syntyvien asiakirjojen metatiedoissa käytettävistä pysyvistä tunneista saa lisätietoja ja tarkempia esimerkkejä tiedonhallintalautakunnan suosituksista asiankäsittelyn metatiedoista (Valtioneuvoston julkaisuja 2021:33) ja viranomaisten asiakirjojen metatiedoista palveluita tuottaessa (Valtioneuvoston julkaisuja 2022:42).

### 2.6.1.3 Koneluettavuuden asiakaspalaute -mittari

Koneluettavuuden asiakaspalaute -mittari kertoo, että tietoaineiston koneluettavuudesta on mahdollisuus antaa palautetta olemassa olevan palauteväylän kautta tai kohdennetulla asiakaskyselyllä. Mittari kuvaa onko tiedon käyttäjän mahdollista antaa palautetta ja saada vastauksia antamaansa palautteeseen. Saatua palautetta myös seurataan ja palautteeseen reagoidaan. Palautteen yhteenveto tai palautteen tuottamat muutokset ovat hyödyllistä lisätietoa tiedon käyttäjille. Palautetta voidaan hyödyntää tietoaineiston koneluettavuuden parantamisessa.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: kyllä/ei

Esimerkki: -

## 2.6.2 Laatuksiteeri: Tietoaineiston käyttöoikeudet

Synonyymi: -

Kuvaus: Tietoaineiston käyttöoikeudet kuvaa miten tietoaineiston käyttöoikeus on määritelty ja mitä aineistolla voi tehdä eli mihin käyttötarkoituksiin tietoaineistoa voi hyödyntää.

Tiedon käsittelyn perusteet ja tarkoitus tunnistetaan ja kuvataan viranomaisessa tiedon tuottamisen tai vastaanoton yhteydessä. Samassa yhteydessä tunnistetaan käsittelyä koskevat erityisvaatimukset, jotka voivat tulla lainsäädännöstä tai toisen organisaation tiedon käsittelylle asettamista vaatimuksista. Erityisvaatimuksia asettaa esimerkiksi henkilötietoja koskeva tietosuojalaki (1050/2018) ja EU:n yleinen tietosuoja-asetus ((EU) 2016/679).

Esimerkit: Tietoaineisto on tietyin rajoituksin käytettävissä tieteelliseen tutkimukseen. Avoimena datana saatavilla oleva tietoaaineisto on lisensoitu.

### 2.6.2.1 Käyttöoikeus-mittari

Käyttöoikeus-mittari kertoo kenelle tietoaaineiston käyttö on mahdollista eli kenen tai millaisessa käsittely-ympäristössä se on käytettävissä.

- Arviointitaso: tietoaaineisto
- Arvon esitysmuoto: käyttöoikeuden rajaus

Mittarin arvo kertoo, kenen on mahdollista saada tietoaaineisto käyttöönsä. Tietoaaineisto voi olla esimerkiksi avointa dataa, julkista tietoa tai määräaikaikaisella käyttöluvalla, lisenssillä, sopimusperusteisesti tai virkakäytön perusteella käytettävissä.

Mittarin arvo löytyy tiedonhallintayksikön olemassa olevista tiedoista, useimmiten tietoaaineistoa koskevista metatiedoista. Tiedonhallintayksikössä tiedon käsittelyn perusteet ja tarkoitus tunnistetaan ja kuvataan tiedon tuottamisen tai vastaanoton yhteydessä.

Esimerkit: Tilastokeskuksen yksikköaineistot ovat käyttöluvalla käytettävissä esimerkiksi tutkimustarkoituksessa. Suomalaiset väylätiedot ja niiden metatiedot ovat saatavilla avoimina tietoaaineistoina.

### 2.6.2.2 Käytön rajoitukset -mittari

Käytön rajoitukset -mittari kertoo, mihin kaikkiin käyttötarkoituksiin tietoaaineistoa on luvallista käyttää. Tietoaaineiston metatietoihin on hyvä kuvata lyhyesti mahdollisia rajoituksia. Tietoaaineistojen käyttö tapahtuu siihen hyväksytyissä ja asetettujen vaatimusten mukaisissa tietojärjestelmissä, laitteissa ja käsittely-ympäristöissä.

- Arviointitaso: tietoaaineisto
- Arvon esitysmuoto: käyttötarkoituksen rajaus

Mittarin arvo kertoo, millaiseen käyttötarkoitukseen tietoaaineistoa saa käyttää. Mittarin arvo löytyy tiedonhallintayksikön olemassa olevista tiedoista, useimmiten tietoaaineistoa koskevista metatiedoista.



Esimerkki: Tilastokeskuksen käyttöluvallisista yksikköaineistoista saa tuottaa anonyymejä, summattuja tutkimustuloksia. Julkaistua tietoa saa käyttää vapaasti kun lähde mainitaan.

## 2.6.3 Laatuksiteeri: Oikea-aikaisuus

Synonyymi: Täsmällisyys

Kuvaus: Oikea-aikaisuus tarkoittaa sitä, että tietoaineisto on käytettävissä ilmoitettuna ajankohtana ja riittävän tiheästi tietoaineistossa tapahtuviin muutoksiin nähden. Oikea-aikaisuus sisältyy tiedonhallintalain 15.1 §:n sääntelyyn tietoaineistojen saatavuudesta, joten tietoaineiston metatietojen tulee lähtökohtaisesti sisältää tietoa tietoaineiston oikea-aikaisuudesta.

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

### 2.6.3.1 Määräpäivien noudattaminen -mittari

Määräpäivien noudattaminen -mittari kuvaa tietoaineiston suunnitellun toimitusaikataulun toteutumista.

- Arviointitaso: tietoaineisto
- Arvon esitysmuoto: toimitusviive/ei relevantti
- Mittarin tausta: Euroopan tilastojen käytäntesäännöt

Mittaria soveltaessa ilmoitetaan toimitetun aineiston viive suhteessa sovittuun määräpäivään eli ”toteutunut toimituksen päivämäärä – sovittu päivämäärä”. Lisäksi viivitelanteessa voidaan lisätietona kuvata viiveen syy.

Esimerkki: -

### 2.6.3.2 Päivitystiheys-mittari

Päivitystiheys-mittari kuvaa tietoaineiston säännöllistä päivitystiheyttä. Mittari liittyy tiedon elinkaareen, katso tarkemmin 2.5.1.2 Tiedon elinkaari –mittari, ja tietoaineiston mahdollisesti sisältämien aikasarjamuotoisten tietojen elinkaareen.

- Arviointitaso: ominaisuus ja tietoaineisto
- Arvon esitysmuoto: sanallinen ilmaus
- Mittarin tausta: ISO 19139 Geographic information - Metadata - XML schema implementation: MDMaintenanceFrequencyCode ja maintenanceNote

Mittarin arvo kuvaa säännöllisen päivityksen tiheyttä. Käytettävät ilmaukset voivat olla esimerkiksi 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.

### 2.6.3.3 Päivityksessä muuttuneet ominaisuustiedot -mittari

Päivityksessä muuttuneet ominaisuustiedot -mittari kuvaa tietoaineiston muuttuneiden ominaisuuden arvojen osuutta kaikista tietoaineiston ominaisuuden arvoista verrattaessa päivitettyä tietoaineistoa edelliseen versioon.

Mittarin tavoitteena on kuvata päivitysten aiheuttaman muutoksen suuruutta tietoaineistossa, siis kuinka suuri vaikutus esimerkiksi päivityksen odottamisella on mahdollisesti jo käytössä olevan tietoaineiston sisältöön. Myös jatkuvasti päivittyvien tietojen kohdalla muutosten suuruuden kuvaaminen on hyödyllinen tieto tiedon käyttäjälle.

- Arviointitaso: ominaisuus
- Arvon esitysmuoto: prosenttia

Mittarin arvoa tuotettaessa on ominaisuustasolla mahdollista määritellä muuttuneiden arvojen osuus ominaisuuksittain tarkasteltuna. Päivityksessä muuttuneiden ominaisuustietojen osuutta voi tarkastella esimerkiksi päivityksen kannalta mielekkään ajanjakson ajalta tarkasteltuna. Tietoaineistotasolla muutosta on usein helpompi kuvata sanallisesti.

Esimerkki: Yritysrekisterin tietopalvelussa tietoaineistotoimituksen jälkeen uuteen aineistotoimitukseen on lisätty uudet aloittaneet yksiköt ja niiden ominaisuustiedot sekä vanhojen yksiköiden osalta niiden ominaisuustiedoissa tapahtuneet muutokset sekä mahdolliset lopetustiedot. Kuntaliitoksen perusteella tehdyt muutokset tietoaineistoon.

## Liitteet

## Liite 1 Suosituksessa käytetyt keskeiset käsitteet

Käsite (synonyymi)	Selite
Alipeitto, tilastomenetelmäkäsite	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.
EIF	Ks. Euroopan yhteentoimivuusperiaatteet.
Euroopan tilastojen käytännesäännöt (The European Statistics Code of Practise)	Euroopan tilastoja koskevat käytännesäännöt muodostavat Euroopan tilastojärjestelmän yhteisen laatujärjestelmän perustan. Ne perustuvat 16 periaatteeseen, jotka koskevat institutionaalista ympäristöä, tilastoprosesseja ja tilastotuotteita. Kullekin periaatteelle on määritelty hyvän käytännön indikaattoreita ja standardeja, joiden sisältämien ohjeiden ja viitteiden avulla voidaan tarkastella käytännesääntöjen täytäntöönpanoa ja samalla lisätä Euroopan tilastojärjestelmän avoimuutta.
Euroopan yhteentoimivuusperiaatteet (The European Interoperability Framework, EIF)	Euroopan komission tiedonantona (COM(2017)134), annettu joukko suosituksia, joissa määritetään miten julkinen hallinto, yritykset ja kansalaiset vaihtavat tietoa keskenään Euroopan unionin alueella ja jäsenmaiden rajat ylittäen.
FAIR-periaatteet	FAIR-lyhenne muodostuu englannin kielen sanoista findable (löydettävä), accesible (saavutettava), interoperable (yhteentoimiva) sekä re-usable (uudelleenkäytettävä) FAIR-periaatteet tarkoittavat, että tutkimuksen lähdeaineistojen, menetelmien ja tietotuotteiden tulee olla löydettävissä, saatavilla, siirrettävissä tai yhdistettävissä ja uudelleen käytettävissä. Itse periaatteet ovat valtaosin teknistä infrastruktuuria ja digitaalisia palveluita

	koskevia teknisiä eritelmiä. Tarkemmin: <a href="https://www.nature.com/articles/sdata201618">https://www.nature.com/articles/sdata201618</a>
Imputointi (paikkaus), tilastomenetelmä	Imputointi eli paikkaus tarkoittaa havaintoaineistossa olevan puutteen tai poikkeavan havainnon korvaamista jollain imputointimenetelmällä. Menetelmiä ovat mm. looginen imputointi, jossa korjataan loogisesti mahdolliset virheet, esimerkiksi 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 sekä regressio- ja muut malliperusteiset menetelmät, jossa käytetään tilastollista mallia ennustamaan arvo puutteelle.
INSPIRE	INSPIRE on vuonna 2007 voimaan tulleen Euroopan yhteisön paikkatietoinfrastruktuurin perustamisesta koskevan direktiivin (Infrastructure for Spatial Information in Europe, 2007/2/EY) lyhenne. Direktiivin sääntelyn tarkoitus on mahdollistaa paikkatietojen yhteentoimivuus yli organisaatioiden ja valtioiden rajojen. Täytäntöönpantu Suomessa mm. lailla paikkatietoinfrastruktuurista (421/2009) sekä asetuksella paikkatietoinfrastruktuurista (725/2009).
Kohdeyksikkö (luokka, tilastoyksikkö, yksikkö, kohde)	Tietoaineistossa tarkasteltava yksikkö.
Koodisto	Kokoelma ryhmiä, jotka eroavat toisistaan tietyiltä piirteiltään.
Laatukriteeri	Tekijä, joka kuvaa laatukuvauksen laatua tietyistä näkökulmasta.
Luovuttajaimputointi	Ks. Imputointi (paikkaus).
Mittari	Laatukriteerin näkökulmaa konkretisoiva ja tietoaineiston ominaisuutta mittaava mittari. Kukin mittari viittaa tiettyyn laatukriteeriin. Yhdelle laatukriteerille voidaan soveltaa useita mittareita.
Mittaristo	Tiedon laatukriteerien soveltamiseen tuotettu erillisten mittarien muodostama mittarikokoelma.

Ominaisuus (ominaisuus, attribuutti, kohde, muuttuja)	Kohdeyksikköä kuvaava tieto.
Poikkeava havainto (outlier), tilastomenetelmäkäsite	Poikkeava havainto (outlier) on sellainen muuttujan arvo, joka poikkeaa suuresti havaittujen arvojen valtaenemmistöstä ja se voi olla joko oikea tai virheellinen. Poikkeavilla arvoilla voi olla merkittävästi vääristävä vaikutus käytettyihin tilastollisiin tunnuslukuihin, kuten keskiarvoon tai hajontaan.
Perusjoukko (populaatio), tilastomenetelmäkäsite	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. alipeitto, ylipeitto).
Rakenteellinen puuttuvuus	Rakenteellisella puuttuvuudella tarkoitetaan tilannetta, jossa kohdeyksikkö ei missään tapauksessa voisi saada arvoa jollekin ominaisuudelle kuten esimerkiksi tuloeriä sellaisen tulon osalta, jota tarkasteltava henkilö ei saa.
Rakenteinen data	Koneluettava tieto, johon on liitetty metatietojen avulla määritetty rakenne, jonka avulla tietoa voidaan jäsentää.
Skeema	Tietoaineiston rakenteen määrämuotoinen esitys.
Standardiluokitus (luokitusstandardi)	Luokitussuositus, 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 prosessiin, jossa käytetyn tiedon laatu on alkujaan kuvattu.
Tiedon laatu	Se, missä määrin jonkin tietoaineiston ominaisuudet täyttävät vaatimuksia tai

	tavoitteita. Tiedon soveltuvuus siihen käyttötarkoitukseen, johon käyttäjä aikoo sitä käyttää ja tiedon tuottaja sitä tarjoaa.
Tiedonhallintamalli	Tiedonhallintamalli kuvaa tiedonhallinnyksikön ja sen toimintaympäristön tiedonhallinnan. Sitä ylläpidetään palvelujen, asiantuntijien ja tietoaineistojen hallinnan suunnittelemiseksi ja toteuttamiseksi, tiedonsaantia koskevien oikeuksien ja rajoitusten toteuttamiseksi, moninkertaisen tietojen keruun vähentämiseksi, tietojärjestelmien ja tietovarantojen yhteentoimivuuden toteuttamiseksi sekä tietoturvallisuuden ylläpitämiseksi. Tiedonhallintalaki 906/2019 5 §.
Tiedonhallintayksikkö	Viranomaisen, jonka tehtävänä on järjestää tiedonhallintansa tiedonhallintalain vaatimusten mukaisesti. Tiedonhallintalaki 906/2019 2 §.
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	Asiakirjoista ja muista vastaavista tiedoista muodostuva tiettyyn viranomaisen tehtävään tai palveluun liittyvä tietokokonaisuus.
Tietovaranto	Viranomaisen tehtävien hoidossa tai muussa toiminnassa käytettäviä tietoaineistoja sisältävä kokonaisuus. Tiettyä tarkoitusta varten muodostettu, viranomaisen tehtävien hoidossa tai muussa toiminnassa käytettävien, loogisesti tai fyysisesti yhteenkuuluvien tietojen muodostama tietoaineisto tai tietoaineistojen kokoelma.
Tunniste	Yksilöimiseen käytettävä merkkijono.
Ylipeitto, tilastomenetelmäkäsité	Ylipeitto tarkoittaa otoskehikkoon sisältyviä kohdeyksiköitä, jotka eivät enää kuulu kohdeperusjoukkoon, esim. laitoksiin siirtyneitä, kuolleita, maasta muuttaneita. Tällaisia tapauksia esiintyy aina jonkin verran otosyksiköiden joukossa, koska rekisterit, joista otokset poimitaan, eivät ole täysin ajantasaisia.

## Liite 2 Laadun arviointikriteerien mittarit

Nimi	Kuvaus	Esitysmuoto	Arviointitaso	Laatukriteeri
<b>Miten tieto kuvaa todellisuutta? -laatukriteeriryhmä: Virheettömyys</b>				
Menetelmällisesti tuotetut arvot	Menetelmällisesti tai korvaavaa tietoa käyttäen tuotettujen arvojen osuus ominaisuuden kaikista arvoista	prosenttia	Ominaisuus	Virheettömyys
Virheelliset ominaisuusarvot	Virheellisiä ominaisuusarvoja saavien kohdeyksiköiden osuus kaikistakohdeyksiköistä	prosenttia	Ominaisuus	Virheettömyys
Väärinluokittelu	Virheellisesti luokiteltuja ominaisuuden arvoja saavien kohdeyksiköiden osuus kaikista kohdeyksiköistä	prosenttia	Ominaisuus	Virheettömyys
<b>Miten tieto kuvaa todellisuutta? -laatukriteeriryhmä: Tarkkuus</b>				
Keskiahjonta	Ominaisuuksien arvojen vaihtelu keskiarvon ympärillä	keskiahjonta	Ominaisuus	Tarkkuus
Poikkeavat havainnot	Poikkeavien havaintojen lukumäärän osuus kaikista kohdeyksiköistä	prosenttia	Ominaisuus	Tarkkuus
<b>Miten tieto kuvaa todellisuutta? -laatukriteeriryhmä: Johdonmukaisuus</b>				
Tieto loogisuus-tarkastettu	Tiedon keräämisessä, koostamisessa tai käsittelyssä on käytetty loogisuusehtoja.	kyllä/ei	Ominaisuus ja tietoaineisto	Johdonmukaisuus
<b>Miten tieto kuvaa todellisuutta? -laatukriteeriryhmä: Ajantasaisuus</b>				
Viiteajankohta	Ajankohta, jota tietoaineisto kuvaa	ajankohta	Ominaisuus ja tietoaineisto	Ajantasaisuus



Luontiajankohta	Kohdeyksikön tai ominaisuuden luontiajankohta	ajankohta	Ominaisuus ja tietoaineisto	Ajantasaisuus
Tarkistusajan-kohta	Kohdeyksikön tai ominaisuuden tarkistusajankohta	ajankohta	Ominaisuus ja tietoaineisto	Ajantasaisuus
Muutosajankohta	Kohdeyksikön tai ominaisuuden muutosajankohta	ajankohta	Ominaisuus ja tietoaineisto	Ajantasaisuus
<b>Miten tieto kuvaa todellisuutta? -laatukriteeriryhmä: Kattavuus</b>				
Ajallinen tavoite-kattavuus	Tietoaineiston tavoiteltava ajallinen kattavuus ja tiheys on kuvattu.	kyllä/ei	Tietoaineisto	Kattavuus
Alueellinen tavoite-kattavuus	Tietoaineiston tavoiteltava alueellinen kattavuus ja tiheys on kuvattu.	kyllä/ei	Tietoaineisto	Kattavuus
Tavoitekohdeyksiköt	Tietoaineiston aineistokuvauksessa on selkeästi määritelty ajallisen ja alueellisen tavoite-kattavuuden lisäksi muut tietoaineiston rajaukset.	kyllä/ei	Tietoaineisto	Kattavuus
Tavoiteominaisuuksissa puutteita	Tietoaineistosta puuttuu sen kuvaaman ilmiön kannalta keskeisiä ominaisuuksia	kyllä/ei	Tietoaineisto	Kattavuus
Puuttuvat kohdeyksiköt	Tietoaineiston alipeitto eli tietoaineiston tavoitekohdejoukosta puuttuvien kohdeyksiköiden prosentuaalinen osuus	prosenttia	Tietoaineisto	Kattavuus
Ylimääräiset kohdeyksiköt	Tietoaineiston ylipeitto eli tietoaineistossa mukana olevien, mutta tietoaineiston tavoiteperusjoukkoon kuulumattomien kohdeyksiköiden prosenttiosuus	prosenttia	Tietoaineisto	Kattavuus
Puutteelliset kohdeyksiköt	Puuttuvan ominaisuustiedon sisältävien kohdeyksiköiden osuus kaikista kohdeyksiköistä	prosenttia	Ominaisuus ja tietoaineisto	Kattavuus

Puutteelliset ominaisuustiedot	Puuttuvan ominaisuustiedon omaavien kohdeyksiköiden osuus kaikista kohdeyksiköistä ominaisuuksittain	Prosenttia	Tietoaineisto	Kattavuus
<b>Miten tieto on kuvattu? -laatukriteeriryhmä: Alkuperäisyys</b>				
Tietolähde	Tietoaineiston, kohdeyksikön tai ominaisuuden lähdetieto on saatavilla	prosenttia	Ominaisuus ja tietoaineisto	Alkuperäisyys
Tiedon elinkaari	Tietoaineiston elinkaari määriteltä ja kuvaus siitä on saatavilla	kyllä/ei	Tietoaineisto	Alkuperäisyys
Muutosten hallinta	Ominaisuustietojen rakenteellisia tai lähdetietojen muutoksia seurataan	kyllä/ei	Ominaisuus	Alkuperäisyys
<b>Miten tieto on kuvattu? -laatukriteeriryhmä: Metatietojen ymmärrettävyys</b>				
Aineistokuvaus	Kieliversiot, joilla tietoaineiston aineistokuvaus on saatavilla	kieliversiot	Tietoaineisto	Metatietojen ymmärrettävyys
Käsitteiden määrittelmät	Kieliversiot, joilla tietoaineiston keskeisten käsitteiden määrittelmät ovat saatavilla	kieliversiot	Tietoaineisto	Metatietojen ymmärrettävyys
Ominaisuuksien tietokuvaukset	Kieliversiot, joilla ominaisuuksien kuvaukset ja koodistot ovat saatavilla	kieliversiot	Ominaisuus	Metatietojen ymmärrettävyys
Metatietojen ymmärrettävyyden asiakaspalaute	Metatietojen ymmärrettävyydestä on mahdollista antaa asiakaspalautetta, saatua palautetta seurataan ja siihen reagoidaan	kyllä/ei	Tietoaineisto	Metatietojen ymmärrettävyys
<b>Miten tieto on kuvattu? -laatukriteeriryhmä: Suositustenmukaisuus</b>				
Noudatettavat standardit ja suositukset	Tietoaineiston noudattamat standardit, suositukset ja hyvät käytännöt ovat listattuna tietoaineiston kuvauksissa	kyllä/osittain/ei	Tietoaineisto	Suosituksenmukaisuus
<b>Miten tietoa voi käyttää? -laatukriteeriryhmä: Koneluettavuus</b>				

Tietoaineiston tietomalli	Tietoaineisto on rakenteisesti kuvattu ja tieto käytetystä mallista.	kyllä/ei	Tietoaineisto	Koneluettavuus
Kohdeyksikön pysyvä tunniste	Tietoaineiston kohdeyksiköillä on vähintään tietoaineistokohtainen pysyvä tunniste	kyllä/ei	Tietoaineisto	Koneluettavuus
Koneluettavuuden asiakaspalautte	Koneluettavuudesta on mahdollista antaa asiakaspalautetta, saatua palautetta seurataan ja siihen reagoidaan	kyllä/ei	Tietoaineisto	Koneluettavuus
<b>Miten tietoa voi käyttää? -laatukriteeriryhmä: Tietoaineiston käyttöoikeudet</b>				
Käyttöoikeus	Tietoaineiston käyttöoikeuden rajaukset, esimerkiksi käsittelijä, käsittely-ympäristö on kuvattu	käyttöoikeuden rajaukset	Tietoaineisto	Käyttöoikeudet
Käytön rajoitukset	Tietoaineiston käyttötarkoitusten rajoitukset: esim. avoimen datan lisenssit tai käyttöehdot on kuvattu.	käyttö-tarkoituksen rajaus	Tietoaineisto	Käyttöoikeudet
<b>Miten tietoa voi käyttää? -laatukriteeriryhmä: Oikea-aikaisuus</b>				
Määräpäivien noudattaminen	Tietoaineiston suunnitellun toimitusaikataulun toteutuminen	toimitusviive	Tietoaineisto	Oikea-aikaisuus
Päivitystiheys	Tietoaineiston säännöllinen päivitystiheys	sanallinen ilmaus	Ominaisuus ja tietoaineisto	Oikea-aikaisuus
Päivityksessä muuttuneen ominaisuustiedot	Päivityksessä muuttuneiden ominaisuusarvojen osuus kaikista tietoaineiston ominaisuusarvoista	prosenttia	Ominaisuus	Oikea-aikaisuus