



GSFI
PAIKKATIEDOT SUOMEN
TILASTOTUOTANNOSSA

Tiedon laatu yhteistyön ytimessä

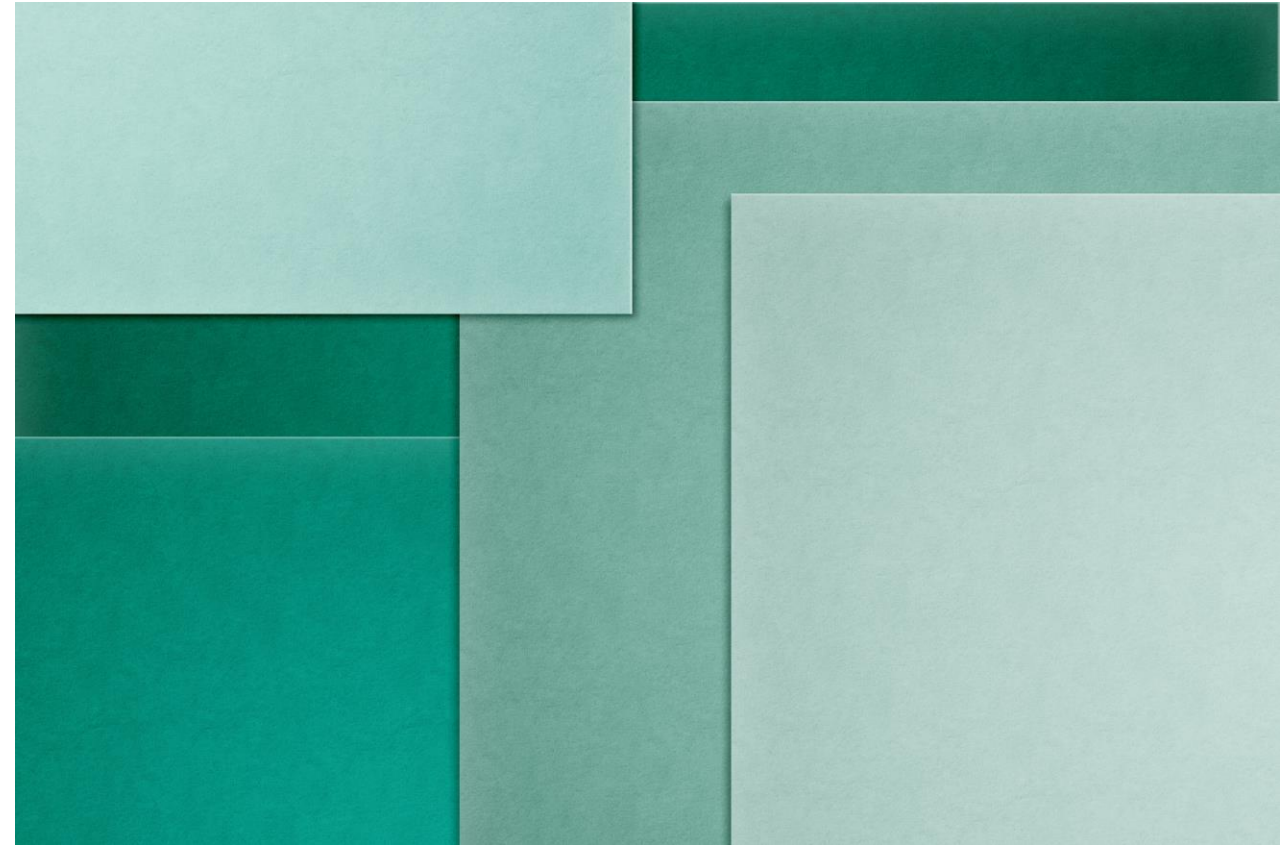
31.7.2024



**Euroopan unionin
rahoittama**

Sisältö

- Johdanto
- Introduction in English
- Tiedon laatu
- Kansallinen tiedon laatukehikko
- Tietoprosessin laatu tiedon laadun tekijänä
- Tiedon laatu yhteistyön liimana: arvoketjutarkastelu
- Tiedon laatukehikko GSGF:n mukaisessa paikkatietojen tuotannossa
- Lisätietoa
- Kiitos
- Lisämateriaalia: CoP, QAF, SIMS, GSBPM ja ISO-laatustandardi



Johdanto

- Tiedon laatu kuvaa, kuinka hyvin tietoaaineisto vastaa tiedon käyttäjän tarpeisiin.
- Yksi merkittävä haaste eri toimijoiden väliselle ja hallinnonalarajatkin ylittävälle yhteistyölle on saumattoman tiedonkulun varmistaminen. Tieto liikkuu lähteestä kohteisiin erilaisten tietoprosessien mukaisesti. Tietovirtojen eri vaiheissa tieto ei kuitenkaan aina kaikilta osiltaan vastaa seuraavan käyttäjän tarvetta, joten tietoa joudutaan korjaamaan, täydentämään tai muuttamaan. Mitä paremmin eri vaiheiden tiedon käyttäjien tarpeet pystytään huomioimaan jo tiedon alkulähteillä, sitä vähemmän syntyy hukkaa myöhemmissä vaiheissa.
- Tiedon käyttäjien tarpeet eivät välttämättä kuitenkaan ole tiedon tuottajien tiedossa – tai ei ole keinoja kiinnittää niihin huomiota. Tässä työssä ehdotetaan tietovirtojen tarkastelua arvoketjujen avulla ja kansallisen tiedon laatukehikon hyödyntämistä arvonalenemien tunnistamiseksi. Laatukehikon 11 kriteeriä tarjoavat systemaattisen rakenteen tiedon laadun tarkasteluun ketjun taitekohdissa. Tämä tilanneymmärrys voidaan jakaa muiden toimijoiden kanssa ja yhdessä sopia muutoksista, joilla arvonalenemia voidaan vähentää. Tiedon virtaus tehostuu, mikä suoraviivaistaa prosesseja ja säästää aikaa ja resursseja.

Introduction in English

- Data quality describes how well data meets the needs of the user.
- One significant challenge for cooperation between different actors and even crossing the boundaries of administrative branches is ensuring a seamless flow of information. Information moves from the source to the destinations according to various information processes. However, in different stages of information flows, the information does not always meet the needs of the next user. The information has to be corrected, supplemented or changed. The better the needs of information users of different stages can be taken into account already at the initial sources of information, the less waste is created in later stages.
- However, the needs of information users are not necessarily known to information producers - or there are no means to pay attention to them. In this work, it is proposed to examine information flows using value chains and to utilize the national information quality framework to identify impairments. The 11 criteria of the quality framework offer a systematic structure for examining the quality of information at the breaking points of the chain. This understanding of the situation can be shared with other operators and together they can agree on changes that can reduce impairments. The flow of information becomes more efficient, which streamlines processes and saves time and resources.

Tiedon laatu

- Tiedon laadulla kuvataan sitä, kuinka hyvin tietoaaineisto vastaa tiedon käyttäjän tarpeisiin. Tiedon käyttäjä siis arvioi ja päättää, onko tiedon laatu riittävän hyvä.
- Tiedon laadun kuvaaminen ja kehittäminen
 - edistää tiedon ja eri organisaatioiden järjestelmien ja prosessien yhteentoimivuutta
 - tukee tietovarantojen laajaa ja monipuolista käyttöä
 - tuo säästöjä
 - tehostamalla tiedon tuotantoa ja hyödyntämistä
 - vähentämällä päällekkäistä työtä sekä tiedon jälkikäteistä korjaamista ja täydentämistä
 - vähentämällä virheellisten päätösten ja toimenpiteiden riskiä



[Kuvakaappaus ja linkki: Laadukas tieto kaikkien hyödyksi \(YouTube\) - VM ja Tilastokeskus](#)
[Video: Tussitaikurit Oy / Marker Wizards Ltd](#)

Kansallinen tiedon laatukehikko

- Kansallisen tiedon laatukehikon 11 laatukriteeriä ja niihin liitetyt mittarit tarjoavat työkalun, jonka avulla voidaan systemaattisesti tarkastella, miten hyvin tuotettu tieto vastaa tiedon käyttäjän tarpeisiin eli kuinka laadukasta tieto on.
- Huomaa, että tiedon laatukehikko tarkastelee myös ajallisia elementtejä eli usein tiedon tuotannossa laadun kilpailijaksi mielletty 'nopeus' on osa tiedon laatua.
- Tiedon tuottajan ja käyttäjän on tärkeä yhdessä tunnistaa kriittisimmät tiedon laadun elementit, jotka mahdollistavat tai estävät tiedon hyödyntämisen.
- Lisätietoa:
 - [Kansallinen tiedon laatukehikko: stat.fi/tiedonlaatu -verkkosivusto](https://stat.fi/tiedonlaatu-verkkosivusto)
 - [Tiedon laatu – laadun merkitys ja kuvaaminen eOppiva-kurssi](#)

Miten tieto kuvaa todellisuutta?

Virheettömyys

Tarkkuus

Kattavuus

Johdonmukaisuus

Ajantasaisuus

Miten tietoa voi käyttää?

Koneluettavuus

Käyttöoikeudet

Oikea-aikaisuus

Miten tieto on kuvattu?

Alkuperäisyys

Metatietojen
ymmärrettävyys

Suosittelunmukaisuus

Tietoprosessin laatu tiedon laadun tekijänä

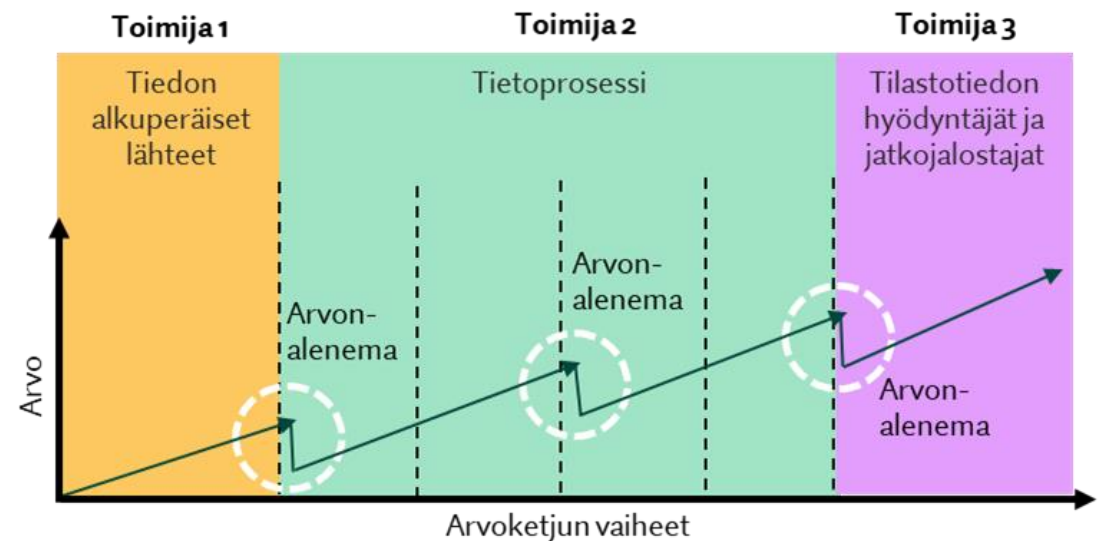
- Tietoprosessi on prosessi, jonka tavoitteena on tiedon tuottaminen, käsittely tai jakaminen (Lähde: [Finto asiasanasto- ja ontologiapalvelu verkossa](#))
- Laadukas prosessi on tehokas, toistettavissa ja vähentää virheiden riskiä. Se on olennainen edellytys laadukkaalle tiedolle. Puutteellinen tai virheellinen tietoprosessi heikentää tiedon laatua.
- Tietoprosessin laatua arvioidaan strategiasta ja tiedon käyttäjien odotuksista johdetuilla mittareilla sekä operatiivisilla mittareilla.
- Tilastoissa käytettyjä viitekehyksiä ja ohjeita:
 - Quality Assurance Framework of the European Statistical System (QAF)
 - European Statistics Code of Practice (CoP)
 - Single Integrated Metadata Structure (SIMS)
 - GSBPM

”Tieto- ja tilastoekosysteemiin vaikuttavat yhä enemmän uudet tietolähteet, tilastollisten liiketoimintaprosessien uudistaminen sekä tietojen ja tilastojen laadunhallinta. Tähän dynamiikkaan liittyy mahdollisuuksia tilastollisen edistyksen saavuttamiseen, mm. parantunut oikea-aikaisuus ja tarkkuus ja siten tehostettu päätöksenteko, mutta myös riskejä liittyen mm. tietosuojaan, tietojen eheyteen, tilastojen laatuun, tietojen saatavuuteen ja tietolähteiden kestävyteen. Näihin haasteisiin on tartuttava.”

Tiedätkö, mitkä ovat vaatimukset organisaatiosi käyttämälle tai tuottamalle tiedon laadulle, entä tietoprosessin laadulle?

Tiedon laatu yhteistyön liimana: arvoketjutarkastelu

- Tieto liikkuu lähteestään eri kohteisiin määriteltyjen tietoprosessien mukaisesti ja julkishallinnossa usein myös organisaatorajat ylittäen.
- Kun tieto ei vastaa eri prosessivaiheiden tarvetta eli tiedon laatu ei ole riittävän hyvä, joudutaan tietoa korjaamaan tai muuttamaan – mahdollisesti myös täydentämään erillisillä kyselyillä.
- Arvoketjutarkastelu auttaa ymmärtämään tiedon virtauksen yhteydessä tapahtuvaa tiedon arvon kasvua – ja toisaalta tunnistamaan kohdat, joissa arvo alenee. Painopiste on ketjun eri toimijoiden kokonaiskuvan ja tarpeiden ymmärtämisessä.
- Tiedon laatukehikko tarjoaa yhteisen työkalun tarkastella arvoketjun nivelkohtia ja tunnistaa kustannustehokkaimmat kehittämiskohteet.



Tiedon laatukehikko GSGF:n mukaisessa paikkatietojen tuotannossa

- *GSGF:ssä kiinnitetään huomiota käytännössä kaikkiin tiedon laatukehikon tarkastelemiin laadun näkökulmiin, esimerkiksi*
 - *Johdonmukaisuus ja virheettömyys: kansallisten sijaintiviitetietojen käyttö ja yhdenmukaiset geokoodauksen periaatteet tukevat eri tietoaineistojen keskinäistä johdonmukaisuutta sekä virheettömyyttä.*
 - *Alkuperäisyys: sijaintitiedon alkuperän tuntemiseen ja tietoihin tehtyjen muutosten jäljitettävyyteen panostetaan.*
 - *Metatietojen ymmärrettävyys: läpinäkyvyyteen panostaminen sijainnin kuvaamisessa edellyttää mm. tiedon käyttöä tukevia, kattavia metatietoja.*
 - *Suosittelun mukaisuus: tietoaineistojen ja sen ominaisuustietojen edellytetään myös noudattavan tunnettuja standardeja, käytäntöjä ja säädöksiä ja ne tulee olla kerrottu tietoaineiston yhteydessä.*
- Merkittävin vaikuttamisen tiedon laatuun on prosessin alussa, kun tieto tuotetaan.
- GSGF:n tavoitteena on rakentaa perusinfrastruktuuri, joka varmistaa sijaintia kuvaavan tiedon laadun heti syntyhetkestä lähtien.
- GSGF Suomen tiekartalla tiedon yhtenäinen kuvaaminen alkuperätietoineen on pitkälti laadun kuvaamista ja siitä viestimistä.
- Kansallinen tiedon laatukehikko antaa hyvän nojan tiedon laadunhallinnalle. Jatkossa erityistä huomiota tulee kiinnittää laadun arviointi-, varmistus- ja palautemekanismin kehittämiseen.

Lisätietoa

- EFGS and Eurostat 2024a. Käytännön lähestymistapa tilastojen ja paikkatiedon integroinnin kehittämiseen arvoketjujen ja tiedon laadun avulla. GSFI-projekti, EU 101112903 - 2022-FI-GEOS-GSFI.
- EFGS and Eurostat, 2024b. Paikkatieto Tilastokeskuksen yritystietojärjestelmässä – tapaustutkimus. GSFI-projekti, EU 101112903 - 2022-FI-GEOS-GSFI.
- Rina Tammisto, Godfrey Lowndes, 2024. Practical approach in developing the integration of statistics and geospatial information through value chains and data quality. European Conference on Quality in Official Statistics, Portugal 2024.
- Tiedon laatukehikko, 2022. Tiedon laatukehikko -osahanke (7/2020–12/2022), Tilastokeskus, Tiedon hyödyntämisen ja avaamisen hanke (VN/5386/2020), VM. [stat.fi/tiedonlaatu -verkkosivut](https://stat.fi/tiedonlaatu-verkkosivut)
- [Tiedon laatu – laadun merkitys ja kuvaaminen, eOppiva-kurssi](#)



Kiitos mielenkiinnosta!

Esitys on tuotettu osana EU-rahoitteista GSFI-projektia.

Title: Tiedon laatu yhteistyön ytimessä

Project: The GSGF in Finland – Integration of geospatial and statistical information in Finland (GSFI)

Grant agreement number: 101112903 - 2022-FI-GEOS-GSFI

It is permitted to copy and reproduce the content in this report. When quoting, please state the source. © GSFI and Eurostat 2024



**Euroopan unionin
rahoittama**



Lisämateriaalia

- European Statistics Code of Practice (CoP)
- Quality Assurance Framework of the European Statistical System (QAF)
- Single Integrated Metadata Structure (SIMS)
- GSBPM
- Kansainväliset ISO-standardit

European Statistics Code of Practice (CoP)

- Viitekehys eurooppalaisten tilastojen kehittämiseksi, tuottamiseksi ja levittämiseksi.
- Julkaistu 2005, viimeisin päivitys 2017.
- Tavoite: Varmistaa, että Euroopan tilastot ovat luotettavia, ajantasaisia ja tarkkoja.
- Keskeiset periaatteet:
 - Ammatillinen riippumattomuus: Tilastojen tuottaminen ilman ulkopuolista vaikutusta.
 - Puolueettomuus ja objektiivisuus: Tilastojen tulee olla puolueettomia ja objektiivisiä.
 - Luotettavuus: Tilastojen tulee perustua luotettaviin menetelmiin ja lähteisiin.
 - Relevanssi ja ajantasaisuus: Tilastojen tulee vastata käyttäjien tarpeisiin ja olla ajantasaisia.
 - Läpinäkyvyys: Tilastojen tuotantoprosessien tulee olla avoimia ja dokumentoituja.
- 16 keskeistä periaatetta ja 84 indikaattoria parhaiden käytäntöjen ja standardien arvioimiseksi.
- Lisätietoa: [Euroopan tilastojen käytäntösäännöt: European Statistics Code of Practice – Revised version 2017](#)

Quality Assurance Framework of the European Statistical System (QAF)

- Asiakirja, joka tukee CoP:n toteuttamista tarjoamalla ohjeita, menetelmiä ja työkaluja, jotka auttavat tilastoviranomaisia varmistamaan tilastojen korkean laadun.
- Julkaistu 2011, viimeisin päivitys 2019.
- Tarjoaa kokoelman menetelmiä, työkaluja ja hyviä käytäntöjä, joita ehdotetaan käytettäväksi tai jotka ovat jo käytössä joissakin ESS:n tilastoviranomaisissa.
- Sisältää myös tiedon laatuindikaattoreita, mutta tilastollisille tuloksille ei aineistoille:
 - Relevanssi: Tilastot vastaavat käyttäjien tarpeisiin.
 - Tarkkuus ja luotettavuus: Tilastot ovat tarkkoja ja luotettavia.
 - Ajantasaisuus ja täsmällisyys: Tilastot julkaistaan ajallaan ja täsmällisesti.
 - Yhtenäisyys ja vertailukelpoisuus: Tilastot ovat yhtenäisiä ja vertailukelpoisia eri aikojen ja alueiden välillä.
 - Saatavuus ja selkeys: Tilastot ovat helposti saatavilla ja ymmärrettäviä¹²³.
- Lisätietoa: [Laadunvarmistuskehikko: Quality Assurance Framework – Eurostatin verkkosivusto](#)

Single Integrated Metadata Structure (SIMS)

- Euroopan tilastojärjestelmän (ESS) kehittämä ja käyttämä, metatietojen vaihtoa ja levittämistä standardoiva metatietorakenne.
- Julkaistu 2014.
- Tavoitteena on auttaa tilastojen käyttäjiä tulkitsemaan ja käyttämään tilastoja oikein.
- SIMS on suunniteltu standardoimaan metatietojen vaihto ja levitys, mikä parantaa tilastotietojen laatua ja yhtenäisyyttä.
- SIMS-rakenne perustuu 19 korkean tason käsitteeseen, joiden avulla voidaan tuottaa yhdenmukaistettuja, tilastotulosten sisältöä ja laatua kuvaavia vertailumetatietoja. Määritellyt käsitteet ovat yhteensopivia SDMX-standardin yleisen terminologian kanssa.
- SIMS-rakenteen avulla Eurostat ja kansalliset tilastovirastot voivat luoda yhtenäisiä ja kattavia metatietoraportteja, jotka tukevat tilastotietojen laatua ja läpinäkyvyyttä.
- Lisätietoja: [Yhtenäinen integroitu metatietorakenne: Reference metadata reporting standards – Eurostatin verkkosivusto](#)

Generic Statistical Business Process Model (GSBPM)

- YK:n Euroopan talouskomission UNECE:n ja eurooppalaisten tilastovirastojen ylläpitämä ja virallisten tilastojen tuottajien tilastotuotannossaan noudattama yleinen tilastojen tuotantoprosessin kuvaus.
- Julkaistu 2009, viimeisin päivitys 2019
- Tavoitteena on kuvata ja määritellä liiketoimintaprosessit, joita tarvitaan virallisten tilastojen tuottamiseen. GSBPM tarjoaa standardoidun viitekehyksen ja yhdenmukaistetun terminologian, joka auttaa tilastollisia organisaatioita modernisoimaan tilastotuotantoprosessejaan sekä jakamaan menetelmiä ja komponentteja.
- Malliin sisältyy useita laatuksiteereitä, jotka auttavat varmistamaan tilastotuotannon laadun eri tuotantoprosessin vaiheissa.
- Lisätietoja:
 - [Generic Statistical Business Process Model \(GSBPM, v5.1\) – UNECE:n verkkosivusto](#)
 - [Quality Indicators for the Generic Statistical Business Process Model \(GSBPM\) – SCRIBD -digitaalisten dokumenttien kirjasto](#)

Kansainväliset ISO-standardit

- ISO 9000 –standardisarja
 - on johtava kansainvälinen laadunhallinnan standardisto, johon myös alan peruskäsitteiden termit ja määritelmät sisältyvät.
 - keskittyy laatujohtamiseen ja laadunvarmistukseen: auttaa organisaatioita dokumentoimaan ja ylläpitämään tehokasta laatujärjestelmää.
 - sisältää useita standardeja, joista tunnetuin on ISO 9001, joka määrittelee vaatimukset laadunhallintajärjestelmille.
 - ISO 9001 -standardin keskeiset periaatteet koskevat asiakaskeskeisyyttä, johtajuutta, ihmisten osallistumista, prosessilähtöisyyttä, jatkuvaa parantamista, päätöksentekoa faktojen perusteella sekä suhteiden hallintaa.
- ISO25012
 - määrittelee yleisen tietojen laatumallin tietojärjestelmissä säilytettävälle tiedoille.
 - kuuluu SQuaRE (Software product Quality Requirements and Evaluation) -sarjaan
 - jakaa tietojen laatuominaisuudet kahteen pääluokkaan:
 - Sisäsyntyinen tietojen laatu: viittaa tietojen laatuun itsessään, kuten tarkkuuteen, täydellisyyteen ja johdonmukaisuuteen.
 - Järjestelmäriippuvainen tietojen laatu: viittaa siihen, miten tietojen laatu saavutetaan ja säilytetään tietojärjestelmässä, esimerkiksi laitteiston ja ohjelmiston avulla.
 - määrittelee 15 tietojen laatuominaisuutta, jotka ovat tärkeitä eri sidosryhmille. Näitä ominaisuuksia voidaan käyttää tietojen laatuvaatimusten määrittämiseen, laadun mittaamiseen ja arviointiin.
- Kansallisen laatukehikon laatuksiteerit pohjaavat pitkälti ISO 25012 –standardiin.