

JHS 179 Kokonaisarkkitehtuurin suunnittelu ja kehittäminen

Liite 7. Semanttisen yhteentoimivuuden menetelmäohje

Versio: 2.0

Julkaistu: 7.2.2017

Voimassaoloaika: toistaiseksi

Sisällys

1Johdanto.....	2
1.1Taustaa.....	2
1.2Termit ja määritelmät.....	3
2Tavoitetila.....	3
3Menetelmän viitekehys.....	4
4Yhteentoimivuusmenetelmä.....	5
4.1Tietokomponentit.....	5
4.2Sanastot.....	6
4.3Koodistot.....	6
4.4Soveltamisprofiilit.....	7
5Yhteentoimivuusmenetelmän soveltaminen.....	8
5.1Työohje.....	8

1 Johdanto

Tässä dokumentissa esitetään menetelmä Suomen julkisen hallinnon tietojen yhteiskäytölle välttämättömän semanttisen yhteentoimivuuden parantamiseksi. Yhteentoimivuusmenetelmä täydentää tietoarkkitehtuurin suunnittelussa, kehittämisessä ja ylläpidossa käytettävää kokonaisarkkitehtuurityön menetelmäjoukkoa. Yhteentoimivuusmenetelmällä on ohjelmistoapuväline, joka ohjaa käyttämään, tuottamaan ja dokumentoimaan uudelleenkäytettäviä tietomääriä. Semanttinen yhteentoimivuus koskee tietojen lisäksi myös prosesseja.

Tämä semanttisen yhteentoimivuuden menetelmäohje on osa julkisen hallinnon tiedonhallinnan ohjausta, ja se liittyy vaiheittain kehitettävään *Julkisen hallinnon metatietopalveluun*, joka on kuvattu *Julkisen hallinnon metatietopalvelun esiselvityksessä*¹.

Noudattamalla yhteentoimivuusmenetelmää julkisen hallinnon organisaatiot edistävät tietojärjestelmien välistä semanttista yhteentoimivuutta tietohallintolain ja sen nojalla annettujen asetusten määräämällä tavalla.

1.1 Taustaa

Tietohallintolain 3 luvun 7 §:ssä todetaan, että julkisen hallinnon viranomaisen on julkisen hallinnon tietojärjestelmien yhteentoimivuuden mahdollistamiseksi ja varmistamiseksi suunniteltava ja kuvattava kokonaisarkkitehtuurinsa, sekä noudatettava laadittua ja ylläpidettyä kokonaisarkkitehtuuria ja sen edellyttämiä yhteentoimivuuden kuvauksia ja määriä sekä toimialakohtaisia tietojärjestelmien yhteentoimivuuden kuvauksia ja määriä.

Tietohallintolain 1 luvun 3 §:ssä määritellään tietojärjestelmien yhteentoimivuus

“tietojärjestelmien tekniseksi ja tietosisällölliseksi yhteentoimivuudeksi muiden julkisen hallinnon viranomaisten tietojärjestelmien kanssa silloin, kun järjestelmät käyttävät samoja tietoja.”

Julkisen hallinnon tietoarkkitehtuurin määrittelydokumentissa (versio 1.0) todetaan, että

”Julkisen hallinnon toiminnan kehittämisessä tähdätään mm. sähköisiin tietointensiivisiin palveluprosesseihin.”

Semanttisen yhteentoimivuuden parantaminen edistää tietointensiivisten prosessien automatisointia. Merkitykseltään yhtenäisten tietojen lisääntyminen ja laajeneva yhteiskäyttö mahdollistavat tietojärjestelmien ja prosessien yhdistämisen uusien tehokkaiden ja käyttäjäystävällisten palvelukokonaisuuksien tuottamiseen. Tästä hyötyvät sekä asiakkaat että julkinen hallinto.

¹ <https://www.avoindata.fi/data/fi/dataset/julkisen-hallinnon-yhteisen-metatietopalvelun-esiselvitys>

1.2 Termit ja määritelmät

koodisto

fi koodisto
en code list

luettelo luokan ominaisuuden sallituista arvoista

ominaisuus

fi ominaisuus
en attribute

olion tai luokan ominaispiirre

soveltamisprofiili

fi soveltamisprofiili
en application profile

tietomäärittely, joka määrittelee tietojärjestelmän tarvitsemat tai tarjoamat tietosisällöt ja rakenteet siten, että ne ovat semanttisesti yhteentoimivia muiden tietojärjestelmien kanssa, hyödyntämällä olemassa olevia sanastoja ja tietomalleja

tietokomponentti

fi tietokomponentti
en core component

tietomäärittely, joka kuvaa tietoa reaali maailman ilmiöistä ja niiden ominaisuuksista toteutusneutraalilla tavalla ja mahdollistaa uudelleen käytön

tietokomponenttikirjasto

fi tietokomponenttikirjasto
en core vocabulary

järjestetty joukko uudelleen käytettäviä tietokomponentteja

2 Tavoitetila

Semanttisen yhteentoimivuuden menetelmällä on mahdollista toteuttaa tietohallintolaissa tarkoitettu tietosisällöllinen yhteentoimivuus Suomen julkisen hallinnon organisaatioiden välisessä tietojenvaihdossa.

Menetelmän avulla tuotetaan tietohallintolaissa mainitut kokonaisarkkitehtuurin edellyttämät yhteentoimivuuden kuvaukset ja määrittelyt julkisen hallinnon yhteisen tietoarkkitehtuurin osalta. Menetelmä ja sen viitekehys ovat keskeinen osa kansallista palveluarkkitehtuuria ja ohjaavat organisaatioita toimimaan kohti julkisen hallinnon tietoarkkitehtuurin määrittelydokumentin mainitsemaa sähköisten tietointensiivisten palveluprosessien automaatiota.

Noudattamalla menetelmää

- toimialakohtaiset tietokomponentit tehdään näkyviksi ja yhteiskäyttöisiksi.
- harmonisoidut tietokomponentit muodostuvat *tietohallintolain* 4.2 § tarkoitetuiksi yhteentoimivuuden edellyttämien tietoarkkitehtuurimäärittelyjen sisällöiksi.

JUHTA - Julkisen hallinnon tietohallinnon neuvottelukunta

- tiedonvaihdossa tarvittavat tiedot on dokumentoitu yhdenmukaisesti aiheuttamatta välittömiä muutosvaatimuksia olemassa oleviin teknisiin toteutuksiin (esimerkiksi tietokantarakenteisiin).

Julkisen hallinnon organisaatioille suositellaan seuraavaa menettelyä:

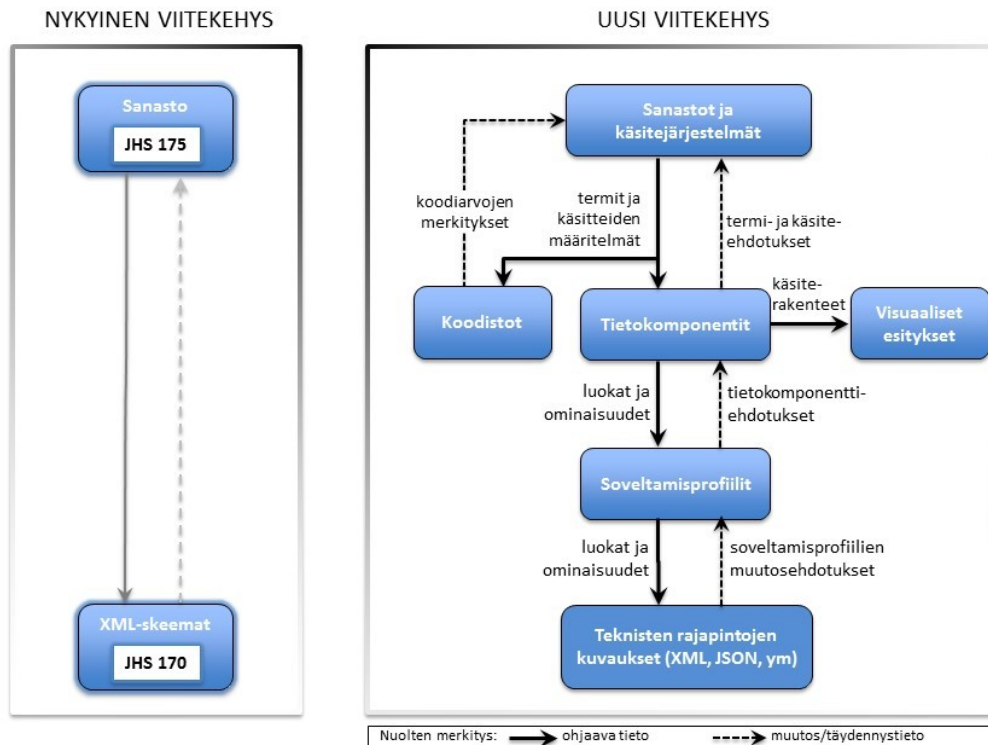
- Tunnista ja kuvaava tilanne, jossa käyttäjän tai toimijan tietotarpeet vaativat paitsi toimijan omien tietovarantojen myös muiden toimijoiden tietovarannon tietojen hyödyntämistä.
- Laadi tilanteesta tarvittavat toiminta-arkkitehtuurikuvaukset ja johda näistä toiminnan tietotarpeet.
- Laadi toiminnan tietotarpeista yhtentöimivuusmenetelmän mukainen soveltamisprofiili. Soveltamisprofiili on tietyn tietotarpeen toteutustavasta riippumaton tietosisällön kuvaus.
- Toteuta soveltamisprofiiliin perustuen tarvittavat tekniset tietorakenteet (esimerkiksi tietojenvälitykseen eri tietovarantojen välillä). Ks. myös *liite 9 Integraation ja rajapintojen kuvaus*.

Noudattamalla tätä menettelyä julkisen hallinnon organisaatio varmistaa, että sen omien tietovarantojen tietosisältöjen merkitys säilyy muuttumattomana myös tilanteissa, joissa tietosisältöä hyödynnetään organisaation ulkopuolella.

Organisaatioiden uudistaessa elinkaarensa lopussa olevia tietojärjestelmiä voidaan menettelyn avulla varmistaa uusittavien järjestelmien ja tietovarantojen tietomäärittysten semanttinen yhteentoimivuus.

3 Menetelmän viitekehys

Yhteentoimivuusmenetelmä ja siihen kehitetty ohjelmistoapuväline perustuvat seuraavassa esitettävään viitekehykseen, jonka keskeisiä osia ovat sanastot, tietokomponentit, soveltamisprofiilit sekä koodistot. Viitekehys korostaa sanastotyön roolia toimintatapojen, tietojen ja tietojärjestelmien yhteentoimivuuden parantamisessa.

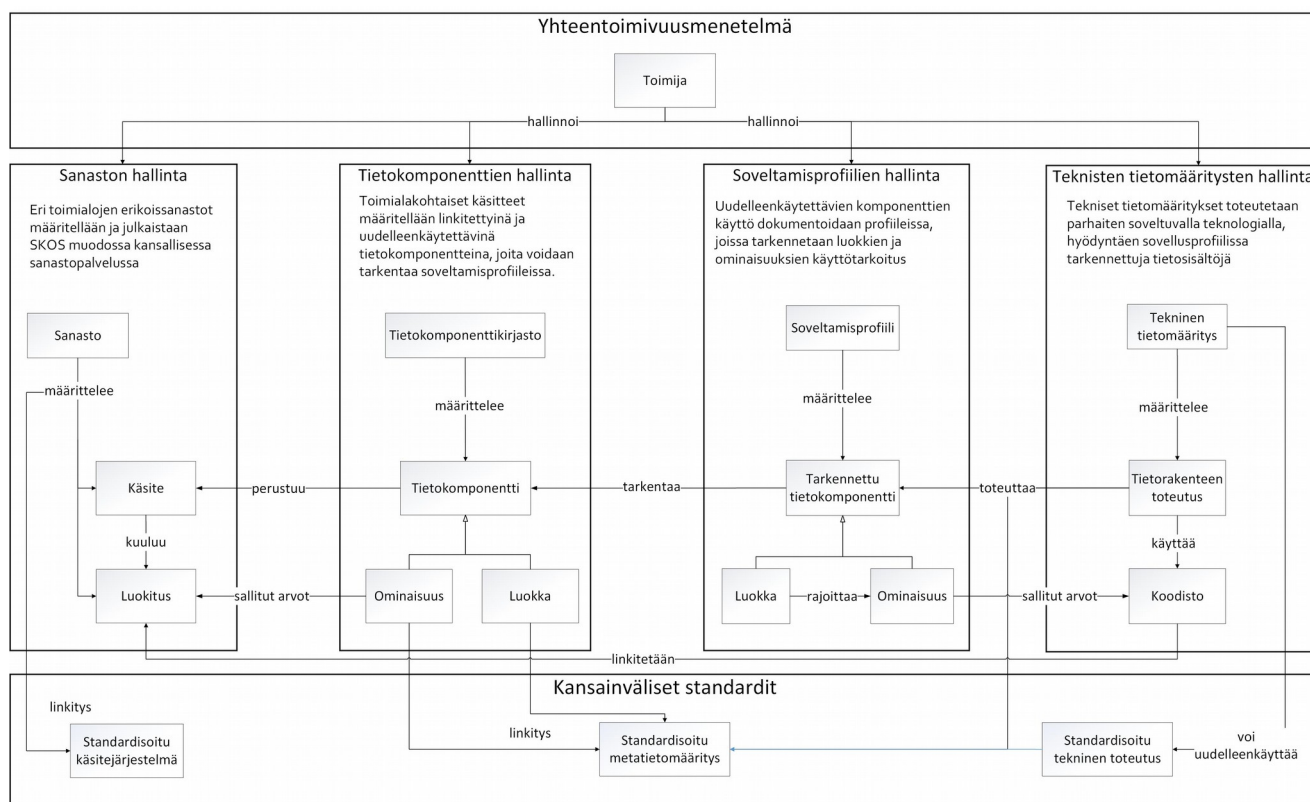


Kuva 1. Yhteentoimivuusmenetelmän viitekehys.

4 Yhteentoimivuusmenetelmä

Yhteentoimivuusmenetelmän avulla julkisen hallinnon organisaatiot voivat kuvata tietovarantojaan siten, että niistä on siirrettävissä tai niihin on vastaanotettavissa tietoa, jonka tietosisältö on merkitykseltään yksiselitteistä eli kaikkien osapuolten yhtenäisesti ymmärrettävissä. Yhteentoimivuusmenetelmällä määriteltyjen tietojen siirtämisessä säilyy tiedon muodostajan tarkoittama sisällöllinen merkitys. Seuraavassa kuvataan menetelmä tällaisten kuvausten muodostamiseksi.

Ennen itse menetelmän kuvaamista esitetään yhteentoimivuusmenetelmän osatekijät ja tuotokset, joita menetelmän avulla luodaan, sekä näiden tuotosten väliset riippuvuudet.



Kuva 2. Yhteentoimivuusmenetelmä.

Yhteentoimivuusmenetelmän kontekstissa ”käsite” viittaa terminologisen sanastotyön menetelmän mukaisesti mielikuvaan, jonka ihminen muodostaa tarkkaillessaan ympäröivää todellisuuttaan, vertaillen sen yksittäisiä ilmiöitä toisiinsa, tunnistaessaan ilmiöiden yhtäläisyyksiä ja eroja ja luokitellensa niitä samankaltaisten ilmiöiden ryhmiin. Tästä ihmisen ajattelutoiminnalle luontaisesta prosessista käytetään nimitystä käsitteellistäminen. Sanastoissa määritellään käsitteen sisältö (määritelmä), käsitteen nimi (termi) ja suhteet toisiin käsitteisiin. Suhteiden kautta käsitteet muodostavat käsitejärjestelmiä. Näistä käytetään joskus virheellisesti nimitystä käsitelmä.

4.1 Tietokomponentit

Tietokomponentit koostuvat luokista ja niiden ominaisuuksista. Luokkien ja ominaisuuksien merkitys määritellään sanastossa, jonka esittämiin käsitteisiin komponentista viitataan luokka- ja ominaisuuskohtaisesti.

JUHTA - Julkisen hallinnon tietohallinnon neuvottelukunta

- Tietokomponentit ovat itsessään aina luokkia, mutta ne voivat muodostaa keskinäisen hierarkian ylä- ja alaluokkasuhteilla.
- Ominaisuuksille on määritelty tiedon esitysmuoto, joka voi olla joko yleinen tietotyyppi tai viittaus toiseen luokkaan.
- Tietokomponenttiin lisätään myös kuvaukset luokan ja sen ominaisuuksien merkityksestä komponenttikohtaisessa kontekstissa eli selitys sille, mitä tietoa komponentilla on tarkoitus välittää tietovarannosta toiseen.
- Tietokomponenttien määrittelyssä käytetään kansainvälisiä, kansallisia ja organisaation omia tietokomponenttien standardeja.
- Tarkennettu tietokomponentti (ks. kuva 2) on tietokomponentti, jota on rajoitettu tai laajennettu sovellusaloittain ominaisuuksilla käyttökontekstiin soveltuvaksi.

4.2 Sanastot

Tietokomponenttien luokat ja ominaisuudet ovat lähtökohtaisesti kontekstisidonnaisia, joten yhteiskäyttöisyyden turvaamiseksi niitä kuvaavat käsitteet on määriteltävä erikseen mahdollisimman neutraalilla tavalla. Tätä varten laaditaan erillisiä sanastoja (esimerkiksi julkisen hallinnon yhteiset sanastot ja toimialakohtaiset sanastot), joissa tietokomponenttien kuvaamiseen tarvittavat käsitteet on määritelty mahdollisimman kontekstivapaasti.

Käytettävä sanasto tulee olla julkaistuna sovitussa muodossa ja siinä esiintyvään käsitteeseen tulee pystyä viittaamaan käsitteen yksilöivällä tunnisteella.

Mikäli komponenttien tarvitsemia käsitteitä ei ole julkaistu julkisen hallinnon yhteisessä tai toimialakohtaisissa sanastoissa, pitää käsite määritellä sanastoon. Tietokomponenttien luokkia ja ominaisuuksia kuvaavien käsitteiden määrittely tehdään *JHS 175 -suosituksen* mukaisessa sanastotyöprosessissa terminologisen sanastotyön periaatteiden mukaisesti (www.jhs-suositukset.fi/suomi/jhs175).

Julkisen hallinnon yhteiseen sanastoon määriteltävät käsitteet eli ns. ydinkäsitteet on harmonisoitava *julkisen hallinnon ydinsanastoryhmän (YSR)* toimesta, sillä nämä käsitteet on usein määritelty eri tavalla eri kohdealueilla tai toimialoilla.

Esimerkki:

- *henkilötunnus* ja *ammatti* ovat ydinkäsitteitä, niiden määrittelystä vastaa YSR.
- *palkka* ja *etuus* ovat yhteisiä käsitteitä, niiden määrittelystä vastaa *julkisen hallinnon tietohallinnon neuvottelukunnan (JUHTA)* nimeämä ja valtuuttama taho, esim. jokin *JHS 175:n* mukainen ns. intressiyhteisö.
- *opiskeluoikeus* on kohdealuekohtainen eli jonkin intressiyhteisön (opetustoimi) hallinnoima käsite, joka julkaistaan myös julkisen hallinnon yhteisessä sanastossa, mutta sen määrittely on käsitettä hallinnoivan tahon vastuulla.

4.3 Koodistot

Koodistoja ja koodiarvoja hyödynnetään tietotyyppinä määrittelemään ominaisuuksille yhteisesti sovitut ja hyväksytyt arvoalueet. Koodiarvon merkitysisältö määritellään sanastossa, jonka sisältämiin käsitteisiin koodistosta viitataan koodiarvotasolla.

Esimerkki:

Tietokomponentti Henkilön (luokka: henkilö) ominaisuus Siviilisääty voi saada määritellyt arvot “Avioliitto”, “Leski”, “Rekisteröity parisuhde”, “Naimaton” jne. Nämä arvot tallennetaan koodistoon “Siviilisääty”. Koodistot tallennetaan ja ylläpidetään pääsääntöisesti eri toimijoiden koodistopalveluissa, joista niiden tulisi olla myös koneluettavasti saatavilla.

4.4 Soveltamisprofiilit

Soveltamisprofiili on tietomäärittäminen, jota käytetään:

- Yleisesti jonkin olemassa olevan tietomallin tarkennuksena tai laajennuksena toiseen käyttötarkoitukseen.
- Yhteentoimivuusmenetelmässä myös uutena tietojärjestelmä- tai käyttötapauskohtaisena tietomallina yhteisiä tietokomponentteja hyödyntäen.

Soveltamisprofiilin muodostamisessa tulee noudattaa kansainvälisesti sovittuja käytäntöjä (*Guidelines for Dublin Core Application Profiles*). Keskeinen periaate on se, että yhteisiä tietokomponentteja voidaan tarkentaa soveltamisprofiileissa eri käyttötarkoituksiin muuttamatta tietokomponenttien merkityssisältöä.

Tietokomponentin kuvauksen tulee vastata käsitteen merkityssisältöä ja edelleen tarkennetun tietokomponentin kuvauksen tulee noudattaa tietokomponentin määrittäystä.

Esimerkki semantiikan säilyttävistä määrittäyksistä ja määrittäjäristiriidasta:

Jäätelötehtaan päätuote on jäätelö, *maitopohjainen jäädytetty jälkiruoka, jonka pääraaka-aine on kerma*. Tehdas on kuitenkin päättänyt laajentaa tuotevalikoimaansa sorbeteilla ja jogurttijäätelöillä. Sorbetti on *jäädytetty jälkiruoka, jonka pääraaka-aine on hedelmä- tai marjamehu*. Jogurttijäätelö eli jääjogurtti on *jäädytetty jälkiruoka, joka pääraaka-aine on maitopohjainen jogurtti*.

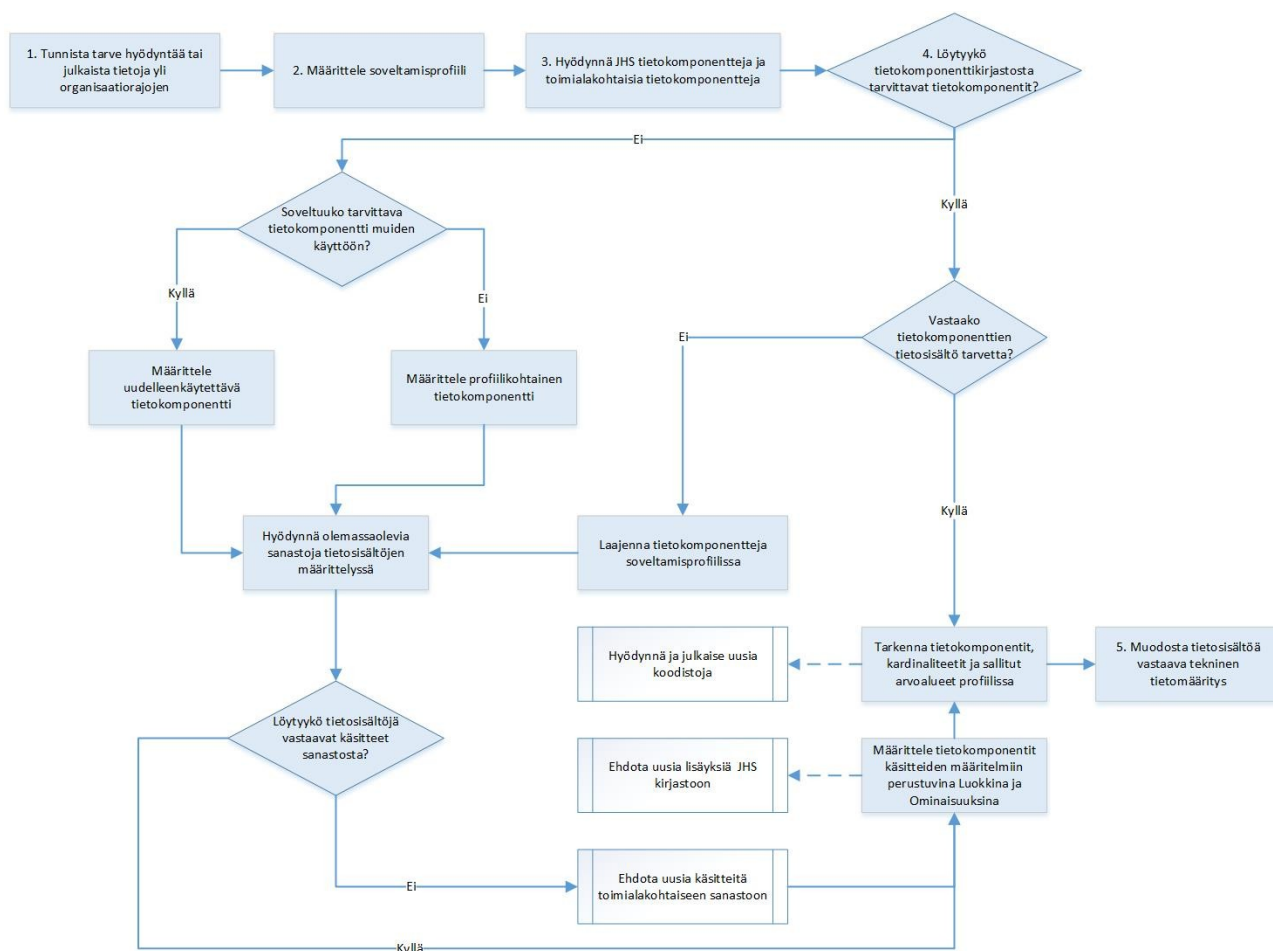
Näitä tuotteita tehdas voi kuluttajamainonnassaan totuudenmukaisesti yleisesti kutsua termillä ”jälkiruoka”, *aterian viimeinen ruokalaji länsimaisessa ruokaperinteessä*. Sen sijaan se ei voi jatkaa mainontaa pelkän alkuperäisen tuotenimikkeen *jäätelö* alla, sillä sorbetin ja jääjogurtin valmistuksessa ei käytetä kermaa ja siten ne eivät määrittelymukaisesti ole jäätelöitä.

Soveltamisprofiileja muodostetaan esimerkiksi seuraavissa tapauksissa:

- kun kyse on tietojenvaihdosta sellaisten toimijoiden välillä, joiden tietovarannot eivät ole yhteismitallisia.
- halutaan yhtenäistää eri tietomalleissa käytettyjä tietomäärittäyksiä.
- halutaan tarkentaa yhteisiä tietokomponentteja toimialakohtaisesti.
- halutaan määrittellä, mikä tieto on toimialakohtaista.

5 Yhteentoimivuusmenetelmän soveltaminen

Seuraavassa kuvassa yhteentoimivuusmenetelmän prosessi esitetään työnkulkukaaviona.



Kuva 3. Työnkulkukaavio.

5.1 Työohje

1. Tunnista ja kuvaava tilanne, jossa toimijan tietotarve vaatii paitsi toimijan oman tietovarannon myös muiden toimijoiden tietovarantojen hyödyntämistä. Tietojen hyödyntämisen helpottamiseksi niitä selittävät tietomäärittelyt tulee kuvata yhteisellä tavalla ja julkaista avoimesti.
2. Kuvaava soveltamisprofiilina toimijan tietotarve tai toimijan omaa tietovarantoa kuvaava tietomäärittely.
3. Soveltamisprofiili kuvataan hyödyntämällä julkisen hallinnon yhteistä tietokomponenttikirjastoa ja toimialakohtaista kirjastoa. Tietokomponenttikirjasto sisältää luokat ja ominaisuudet, joita tulee käyttää määrittelyissä. Yhteisillä tietokomponenteilla tarkoitetaan sekä *KMR-ryhmässä* harmonisoituja luokkia ja ominaisuuksia, että toimialakohtaisia luokkia ja ominaisuuksia (*KMR Käsittemalliryhmä*²). Ominaisuuksien arvoalueet on mahdollisuuksien mukaan määritelty koodistopalveluissa.
4. Tarkastele tietokomponentteja tietokomponenttikirjastossa.

² <https://wiki.julkict.fi/julkict/juhta/juhta-n-jaostot/tietoarkkitehtuuriryhma/kmr-kasitemalliryhma>

JUHTA - Julkisen hallinnon tietohallinnon neuvottelukunta

- Jos sopivanoloinen tietokomponentti löytyy, tarkasta, että tietokomponentille annettu tietosisältö (luokka, ominaisuudet ja niiden merkitykset) vastaa käyttötarvetta.
 - Jos tietokomponentille annettu tietosisältö vastaa tarvetta, hyödynnä tietokomponentti kokonaisuutena tai sen soveltuvat ominaisuudet soveltamisprofiilissa.
 - Jos tietokomponentille annettu tietosisältö ei vastaa tarvetta, laajenna olemassa olevaa tietokomponenttia.
 - Laajentamisessa lisätään tietokomponenttikirjastoon uusi tarvittava ominaisuus ehdokkaana olemassa olevaan luokkaan.
 - Tarkenna tarvittaessa ominaisuuden arvoaluetta määrittelemällä uusi koodisto sopivaan koodistopalveluun.
 - Käytä laajennettua komponenttia tai uutta ominaisuutta osana omaa soveltamisprofiilia.
 - Ilmoita lisätty ominaisuus KMR-ryhmälle käsittelyä varten.
 - Laajennokset tulevat ehdokkaiksi yhteiseen tietokomponenttikirjastoon ja ne joko
 - o hyväksytään yhteisen tietokomponenttikirjaston osaksi ja lisätään yhteiseen kirjastoon uusi yläominaisuus tai
 - o todetaan tapauskohtaiseksi määriytykseksi, vain tietyssä soveltamisprofiilissa käytettäväksi.
 - Jos tietokomponenttia ei löydy, luo tarvittava tietokomponentti osaksi tietokomponenttikirjastoa.
 - Luo komponentti määrittelemällä tarvittava luokka hyödyntämällä sanastoa ja siellä määriteltyjä käsitteitä.
 - Lisää luokkaan tarvittavat ominaisuudet hyödyntämällä sanastoa ja siellä määriteltyjä käsitteitä.
 - Tarkenna tarvittaessa ominaisuuden arvoaluetta määrittelemällä koodisto.
 - Käytä uutta komponenttia osana omaa soveltamisprofiilia.
 - Ilmoita uusi komponentti KMR-ryhmälle käsittelyä varten.
 - Uudet komponentit tulevat ehdokkaiksi yhteiseen tietokomponenttikirjastoon ja ne joko
 - o hyväksytään yhteisen tietokomponenttikirjaston osaksi ja lisätään yhteiseen kirjastoon uusi luokka ja sen ominaisuudet tai
 - o todetaan tapauskohtaisiksi määriytyksiksi, vain tietyssä soveltamisprofiilissa käytettäväksi.
 - Muodosta soveltuvan työkalun avulla valituista tietokomponenteista soveltamisprofiili eli toteutuskohtainen tietomäärittäminen.
5. Muodosta soveltamisprofiilin tietomäärittäksen pohjalta tekniset tietorakenteet.

Käyttämällä yhteisiä sanastoja tarvekohtaiset määrittäykset voidaan kytkeä merkitykseltään yhteisiin määrittäyksiin.