

Tämä dokumentti on tarkoitettu uudistettavan JHS179-suosituksen tietoarkkitehtuuriosion liitteeksi.

LUONNOS

JHS179 liite 7 (alustava numerointi)

Julkisen hallinnon semanttisen yhteentoimivuuden viitekehys

Yleistä

Tässä dokumentissa kuvataan Suomen julkisen hallinnon semanttisen yhteentoimivuuden viitekehys, joka muodostuu uudelleenkäytettävien tietomääritysten laadinnassa käytettävästä menetelmästä (yhteentoimivuusmenetelmä) ja menetelmän tulosten dokumentaatiotavasta tähän soveltuvalla välineistöllä.

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

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 määrittämiä sekä toimialakohtaisia tietojärjestelmien yhteentoimivuuden kuvauksia ja määrittämiä.

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:

Julkisen hallinnon toiminnan kehittämisessä tähdätään mm. sähköisiin tietointensiivisiin palveluprosesseihin. Tietointensiivisyyden etuna on mahdollisuus monien prosessien tai osaprosessien jopa täydelliseen automatisointiin. Automatisoinnin tarkoituksena on siirtää koneiden hoidettavaksi monia sellaisia rutiiniluonteisia tehtäviä, joiden hoitamiseen tällä hetkellä tarvitaan virkailija tai asiakas. Tavoitteena on siten lisätä hallinnon tehokkuutta ja vähentää asiakkaan vaivaamista sellaisissa tehtävissä, jotka voidaan siirtää automatisoitujen prosessien hoidettavaksi.

Sipilän hallitusohjelmassa mainitaan hallituskauden yhtenä tavoitteena käyttäjälähtöiset tuottavuutta ja tuloksellisuutta nostavat, yhden luukun digitaaliset palvelut. Hallitusohjelmaan kirjattu tavoitteeksi digitaalisen liiketoiminnan kasvuympäristön rakentaminen. Tässä kuvattu menettely tukee edellä mainittuja tavoitteita. edistää Sipilän hallitusohjelman mukaisesti uuden teknologian, digitalisaation ja uusien liiketoimintakonseptien käyttöönottoa.

Tavoitetila

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

Viitekehyksen avulla tuotetaan tietohallintolaissa mainitut kokonaisarkkitehtuurin edellyttämät yhteentoimivuuden kuvaukset ja määrittelyt julkisen hallinnon yhteisen tietoarkkitehtuurin osalta. Viitekehys on myös keskeinen tekijä osana kansallista palveluarkkitehtuuria ja julkisen hallinnon tietoarkkitehtuurin määrittelydokumentin mainitsemassa sähköisten tietointensiivisten palveluprosessien automaatioissa.

Noudattamalla tähän viitekehykseen kuuluvaa menetelmää

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

Tavoitetilassa julkisen hallinnon organisaatiot noudattavat seuraavaa menettelyä:

1. Tunnista ja kuvaat tilanne, jossa toimijan tietotarpeet vaativat paitsi toimijan omien tietovarantojen myös muiden toimijoiden tietovarannon tietojen hyödyntämistä.
2. Laadi tilanteesta tarvittavat toiminta-arkkitehtuurikuvaukset ja johda näistä toiminnan tietotarpeet.
3. Laadi toiminnan tietotarpeista Semanttisen yhteentoimivuuden viitekehyksen mukainen sovellusprofiili/soveltamisprofiili.
4. Hyödynnä sovellusprofiilin perusteella tuotettuja teknisiä tietorakenteita (siirtoformaatti) edellä kuvatussa prosessissa (tietojenvälityksessä eri tietovarantojen välillä).

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

Organisaatioiden uudistaessa elinkaarensa lopussa olevia tietojärjestelmiä voidaan menettelyn avulla laatia jatkossa uusittavien järjestelmien ja tietovarantojen tietomääritykset semanttisesti yhteentoimivalla tavalla.

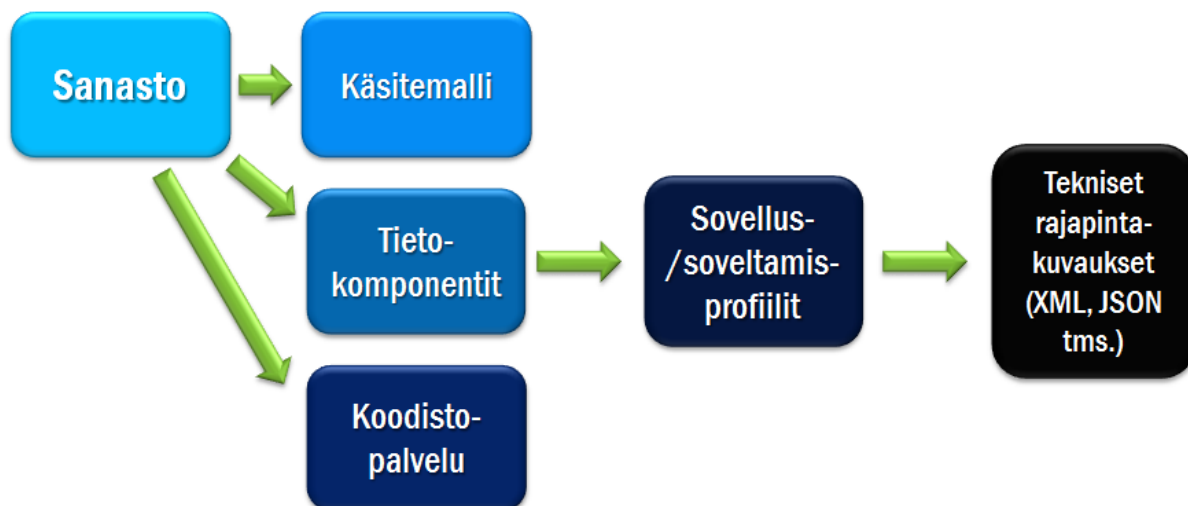
Viitekehysten osat

Semanttisen yhteentoimivuuden viitekehys koostuu tietojärjestelmien semanttista yhteentoimivuutta toteuttavasta menetelmästä (yhteentoimivuusmenetelmä) ja menetelmän tulosten dokumentaatiotavasta tähän soveltuvalla välineistöllä.

Nykyisen JHKA Tietoarkkitehtuurin lähestymistapa



Muutos: "Semanttisen yhteentoimivuuden viitekehys"



Yhteentoimivuusmenetelmä

Yhteentoimivuusmenetelmä on menetelmä, jonka 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. Tiedon siirtäminen yhteentoimivuusmenetelmän mukaisesti säilyttää myös sen muodostajan tarkoittaman sisällöllisen merkityksen. Seuraavassa kuvataan menetelmä tällaisten kuvausten muodostamiseksi.

Ennen itse menetelmän kuvaamista kuvataan kuitenkin yhteentoimivuusmenetelmän osatekijät eli ne tuotokset, joita menetelmän avulla synnytetään sekä näiden tuotosten väliset riippuvuudet.

Yhteentoimivuusmenetelmän käsitteistö

core component - tietokomponentti

(UN/CEFACTS CCTS / ADMS) *context-free semantic building block*

tietomäärittäminen, joka kuvaa tietoa reaali maailman ilmiöistä ja niiden ominaisuuksista toteutusneutraalilla tavalla ja uudelleenkäytön mahdollistaen

core vocabulary - tietokomponenttikirjasto

(ISA) *simplified, re-usable and extensible data models that capture the fundamental characteristics of an entity in a context-neutral fashion.*

joukko tietokomponentteja

application profile - sovellus-/soveltamisprofiili

(Dublin core) *metadata record which meets specific application needs while providing semantic interoperability with other applications on the basis of globally defined vocabularies and models.*

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

attribute - ominaisuus

(ISO 19788) *characteristic of an object or entity*

olion tai tarkoitteen ominaispiirre

code list - koodisto

(ADMS) *complete set of data element values of a coded simple data element*

luettelo luokan ominaisuuden sallituista arvoista

Tietokomponentit

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

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.

Sanastot

Tietokomponenttikirjaston luokat ja ominaisuudet ovat lähtökohtaisesti kontekstisidonnaisia, joten yhteiskäyttöisyyden turvaamiseksi niitä kuvaavat käsitteet on määriteltävä erikseen mahdollisimman neutraalilla tavalla. Tätä varten laaditaan erillisiä sanastoja, joissa tietokomponenttien kuvaamiseen tarvittavat käsitteet on määritelty mahdollisimman kontekstivapaasti.

Käytettävä sanasto tulee olla julkaistuna sovitussa formaalissa 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 JHS175 -suosituksen mukaisessa sanastotyöprosessissa terminologisen sanastotyön periaatteiden mukaisesti.

Julkisen hallinnon yhteiseen sanastoon määriteltävät käsitteet eli ns. ydinkäsitteet on harmonisoitava YSR:n eli julkisen hallinnon ydinsanastoryhmän 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 tietoarkkitehtuuriryhmän?] valtuuttama taho, esim. jokin JHS175: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

Koodistot

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

Esimerkki:

Tietokomponentti Henkilön (luokka: henkilö) ominaisuus Siviilisäätö voi saada määritellyt arvot "Avioliitto", "Leski", "Rekisteriöity parisuhde", "Naimaton" jne. Nämä arvot tallennetaan koodistoon "Siviilisäätö", joka julkaistaan julkisen hallinnon koodistopalvelussa.

Sovellusprofiilit/Soveltamisprofiilit

Sovellus-/soveltamisprofiili on tietomääritys, jota käytetään:

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

Sovellus-/soveltamisprofiilin muodostamisessa tulee noudattaa kansainvälisesti sovittuja käytäntöjä (*Guidelines for Dublin Core Application Profiles*). Keskeisin periaate on se, että yhteisiä tietokomponentteja voidaan tarkentaa sovellus-/soveltamisprofiileissa eri käyttötarkoituksiin muuttamatta tietokomponenttien semanttista merkitystä. Sovellus-/soveltamisprofiileissa tehdyt rajoitukset ja tarkennukset eivät saa olla ristiriidassa tietokomponenteille annettujen määritysten kanssa.

Sovellus-/soveltamisprofiileja muodostetaan esimerkiksi seuraavissa tapauksissa:

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

Yhteentoimivuusmenetelmän soveltaminen

Seuraavassa kuvataan tarkemmalla tasolla miten yhteentoimivuusmenetelmää sovelletaan käytännössä.

Työohje

1. Tunnista ja kuvaa tilanne, jossa toimijan tietotarpeet vaativat paitsi toimijan omien tietovarantojen myös muiden toimijoiden tietovarannon tietojen hyödyntämistä.
2. Kuvaa toimijan tietotarpeeseen sopiva tietomääritys eli **sovellus-/soveltamisprofiili**
3. Sovellus-/soveltamisprofiili kuvataan hyödyntämällä julkisen hallinnon yhteistä tietokomponenttikirjastoa. Tietokomponenttikirjasto sisältää luokat ja ominaisuudet, joita tulee käyttää määrityksissä. Yhteisillä tietokomponenteilla tarkoitetaan sekä KMR-ryhmässä harmonisoituja luokkia ja ominaisuuksia että toimialakohtaisia luokkia ja ominaisuuksia. Ominaisuuksien arvoalueet on mahdollisuuksien mukaan määritelty julkisen hallinnon yhteisessä koodistopalvelussa.
4. Tarkastele tietokomponentteja tietokomponenttikirjastossa.
 - a. Jos sopivanoloinen tietokomponentti löytyy, tarkasta, että tietokomponentille annettu tietosisältö (luokka, ominaisuudet ja niiden merkitykset) vastaa käyttötarvetta.
 - b. Jos tietokomponentille annettu tietosisältö vastaa tarvetta, niin hyödynnä tietokomponentti kokonaisuutena tai sen soveltuvat ominaisuudet sovellus-/soveltamisprofiilissa.
 - c. Jos tietokomponentille annettu tietosisältö ei vastaa tarvetta, laajenna olemassa olevaa tietokomponenttia.
 - i. Laajentaminen tapahtuu lisäämällä tietokomponenttikirjastoon uusi tarvittava ominaisuus olemassa olevaan luokkaan.
 - ii. Tarkenna tarvittaessa ominaisuuden arvoaluetta määrittelemällä uusi koodisto julkisen hallinnon yhteiseen koodistopalveluun.
 - iii. Käytä laajennettua komponenttia tai uutta ominaisuutta osana omaa sovellus-/soveltamisprofiilia.
 - iv. Ilmoita lisätty ominaisuus KMR-ryhmälle käsittelyä varten.
 - v. Laajennokset tulevat ehdokkaiksi yhteiseen tietokomponenttikirjastoon ja ne
 1. joko hyväksytään yhteisen tietokomponenttikirjaston osaksi ja lisätään yhteiseen kirjastoon uusi **yläominaisuus**
 2. todetaan tapauskohtaiseksi määritykseksi, vain tietyssä sovellus-/soveltamisprofiilissa käytettäväksi.

- d. Jos tietokomponenttia ei löydy, luo tarvittava tietokomponentti osaksi tietokomponenttikirjastoa.
 - i. Luo komponentti määrittelemällä tarvittava luokka hyödyntämällä sanastoa ja siellä määriteltyjä käsitteitä.
 - ii. Lisää luokkaan tarvittavat ominaisuudet hyödyntämällä sanastoa ja siellä määriteltyjä käsitteitä.
 - iii. Tarkenna tarvittaessa ominaisuuden arvoaluetta määrittelemällä koodisto.
 - iv. Käytä uutta komponenttia osana omaa sovellus-/soveltamisprofiilia.
 - v. Ilmoita uusi komponentti KMR-ryhmälle käsittelyä varten.
 - vi. Uudet komponentit tulevat ehdokkaiksi yhteiseen tietokomponenttikirjastoon ja ne joko
 1. hyväksytään yhteisen tietokomponenttikirjaston osaksi ja lisätään yhteiseen kirjastoon uusi luokka ja sen ominaisuudet.
 2. todetaan tapauskohtaiseksi määriykseksi, vain tietyssä sovellus-/soveltamisprofiilissa käytettäväksi.
 - e. Muodosta soveltuvan työkalun avulla valituista tietokomponenteista sovellus-/soveltamisprofiili eli toteutuskohtainen tietomäärittely.
5. Muodosta sovellus-/soveltamisprofiilin tietomäärittelyksen pohjalta tekniset tietorakenteet.

~~Yhteinen semanttinen määppäys sanastojen avulla mahdollistaa tarvekohtaisten määrittelysten hallitun kytkemisen yhteisiin määrittelyisiin.~~