

Yhteentoimivuusalusta: Miten saadaan ihmiset ja koneet ymmärtämään toisiaan paremmin?

*Avoin verkkoalusta ihmisen ja koneen ymmärtämien
tietomäärittysten tekemiseen*

Riitta Alkula

20.3.2019

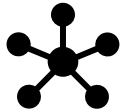


Esityksen sisältö

- Lyhyt oppimäärä
- Yhteentoimivuus ja semanttinen yhteentoimivuus
- Yhteentoimivuusalusta
 - Sanastot-työkalu
 - Koodistot-työkalu
 - Tietomallit-työkalu
- Yhteentoimivuusmenetelmä ja linkitetty data

Lyhyt oppimäärä

- ***Yhteentoimivuus***: eri toimijoiden välillä vaihdetun tiedon merkitys säilytetään ja ymmärretään muuttumattomana tämän tietojenvaihdon aikana
- ***Yhteentoimivuusalusta***: kokonaisuus, joka tarjoaa työkalut tietosisältöjen eli sanastojen, koodistojen ja tietomallien tuottamiseen, ylläpitoon ja jakeluun.
- ***Yhteentoimivuusmenetelmä***: periaatteet, eri osapuolten roolit ja vastuut sekä sovitut käytännöt, joilla tietosisältöjen kuvauksia (metatietoja) tuotetaan ja muiden tuottamia kuvauksia hyödynnetään

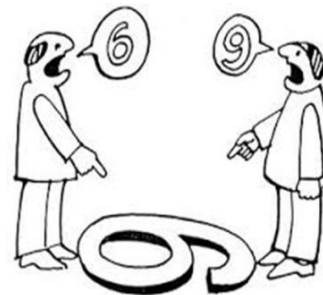


Ongelma: ilmausten monimuotoisuus niin ihmisten kuin koneidenkin välillä

- Vuorovaikutuksessa ei riitä, että siirrämme tietoa paikasta toiseen, jos sen sisältö tarkoittaa vastaanottopäässä jotain muuta kuin lähtöpäässä
- Ihmisetkään eivät välttämättä ymmärrä toisiaan – saati sitten tietokoneet, jos tähän ei rakenneta työkaluja
- Jos meillä ei ole yhteistä kieltä, tulkitsemme tiedon eri tavoin ja ymmärrämme toisemme väärin



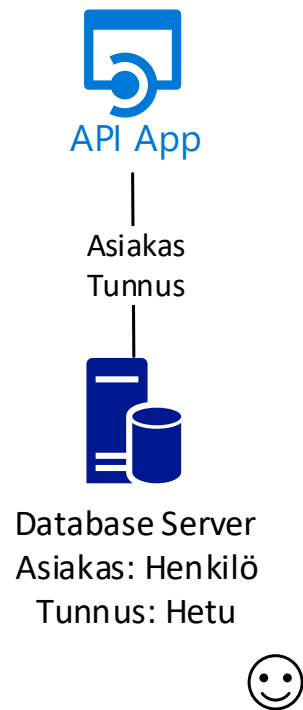
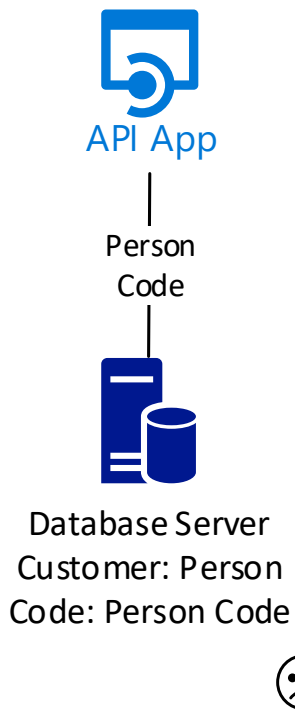
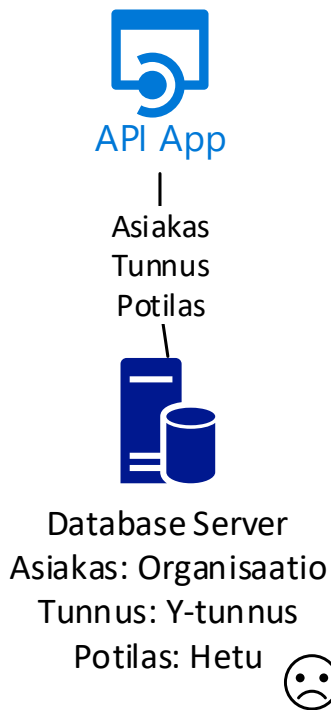
6



Lähde: Tuntematon tekijä

"Asiakas"

Kysely: Asiakas 200319+059G



Ratkaisuna semanttinen yhteentoimivuus



- Tarvitsemme semanttista yhteentoimivuutta hallitsemaan luonnollisen kielen moniselitteisyyttä
- Määriteltyä tietoa tarvitaan yhteisen ymmärryksen muodostamiseen – myös eri rajapinnoissa
- Kyse on paljon muustakin kuin tietojärjestelmien teknisestä rakentamisesta

Kuinka semanttista yhteentoimivuutta voidaan toteuttaa

Sovitetaan yhteisesti käytetyt:

- käsitteet ja sanastot
- koodistot ja luokitukset
- tietokomponentit (joiden pohjalta tuotetaan omat soveltamisprofiilit)

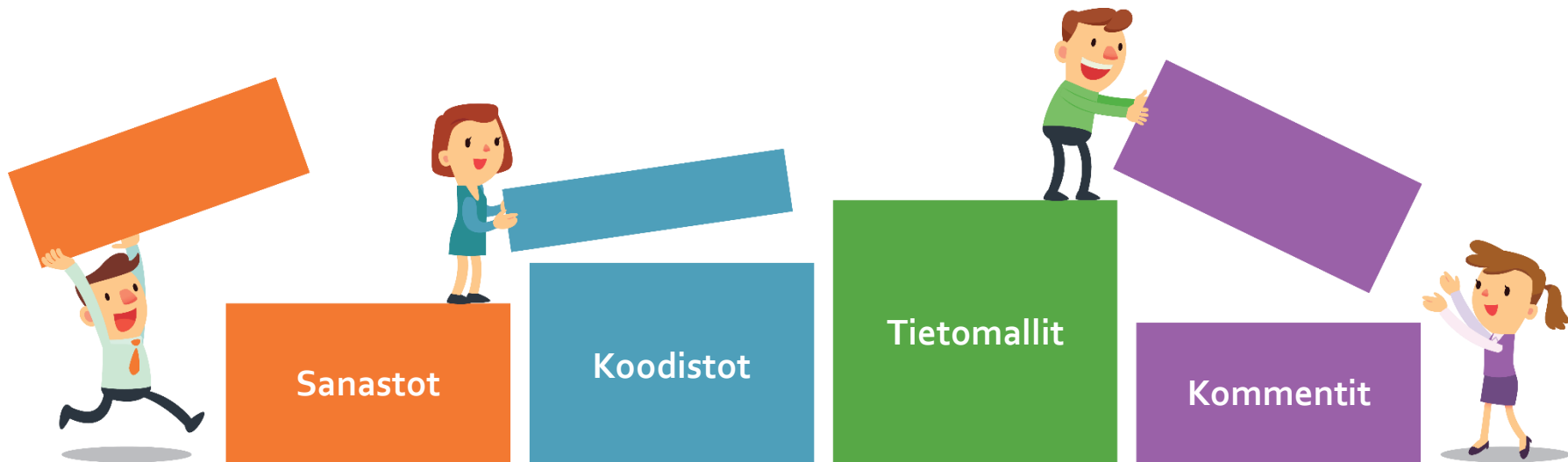
- Linkitetään mahdollisuuksien mukaan kansainvälisiin määrittäksiin ja standardeihin



Tieto ≠ metatieto

- **Yhteentoimivuusalusta on metatietoympäristö.**
Yhteentoimivuusalustalle ei tallenneta varsinaista dataa vaan tietoa tiedosta eli metadataa.
- **Metatietoa** on esimerkiksi vaikkapa kirjan (nimi, tekijä, vuosi, kustantaja...)
- **Metatietoja** ei luoda ad hoc itse, vaan ensisijaisesti käytetään yhteisesti sovittuja metatietoresursseja (kuten käsitteitä, koodeja ja tietomallien luokkia ja ominaisuuksia).

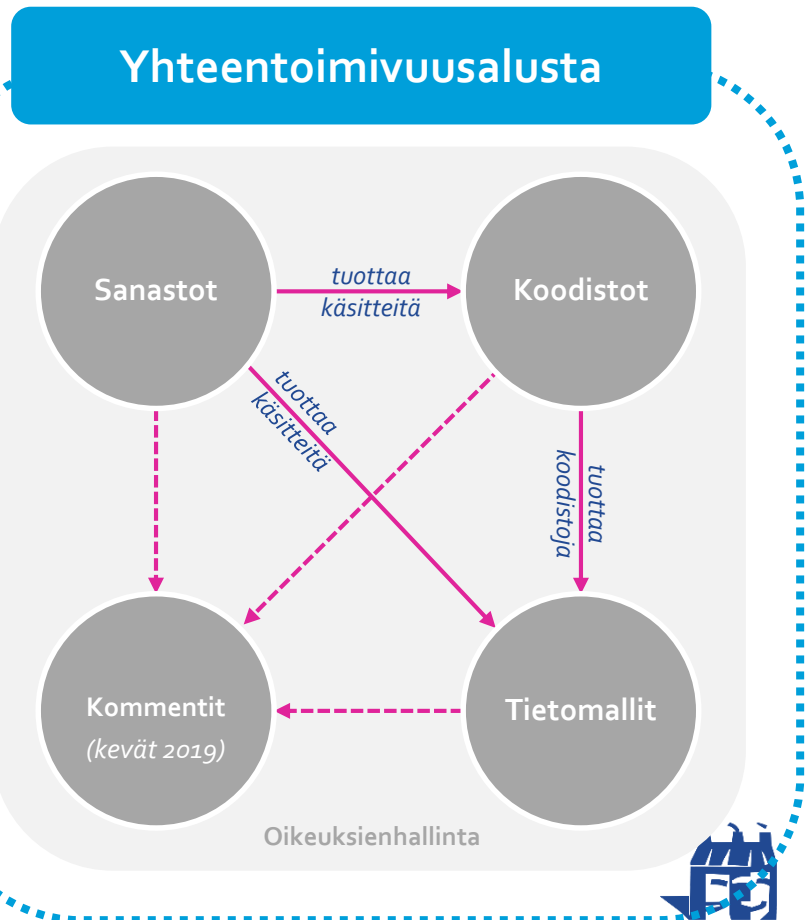
Yhteentoimivuusalusta ja sen työkalut



Yhteentoimivuusalusta

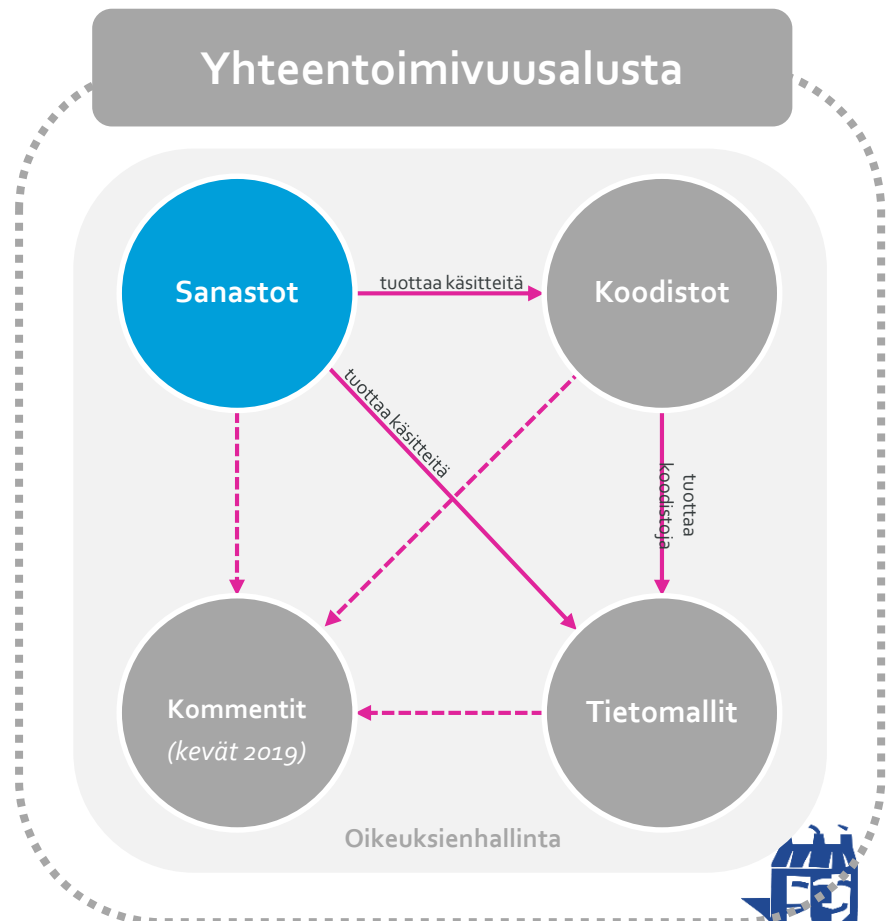
<http://yhteentoimiva.suomi.fi>

- Yksi näkymä julkishallinnon ja joiltain osin liittyvien toimijoiden tietomäärityksiin ja metatietoihin
- Poistaa tarpeen rakentaa vastaavia digitaalisia metatiedon jakelukanavia itse
- Työkalut on integroitu toisiinsa
- Hyödynnetään pysyviä URI-tunnisteita
- Kaikella tietosisällöllä on vastuorganisaatio



Sanastot-työkalu

- Luodaan, ylläpidetään ja julkaistaan käsitteitä, käsitteiden kuvauksia sekä käsitteiden metatietoja
- Luodaan, ylläpidetään ja julkaistaan terminologisia sanastoja
- Kuvataan ja visualisoidaan käsitteiden välisiä suhteita
- Voidaan tuoda käsitteitä CSV- ja XML-tiedostojen avulla
- Selataan ja tarkistetaan olemassa olevia käsitteitä ja niiden määrittelyjä ja muita tietoja
 - Esimerkiksi synonyymit, kuten virallinen *varhaiskasvatus* vs. arkikielinen *päivähoito*
- Hyödynnetään käsitteiden määritelmiä tietomallien ja koodistojen kuvauksissa













ETUSIVU

Sanastot-työkalulla kerätään ja ylläpidetään julkishallinnon yhteisiä terminologisia sanastoja ja käsitteitä. Esimerkiksi sanastotyöntekijät ja tietomallintajat voivat määritellä Sanastoissa käsitteitä ja käsitteiden suhteita toisiinsa, minkä lisäksi käsitteiden väliset suhteet saa myös visualisoitua kuvaksi. Kansalainen puolestaan voi selata ja tarkistaa käsitteitä ja niiden määrityksiä.

Hae sanastoja

Rajaa tietoaueella

-  Eläkkeet (1)
-  Koulutus (2)
-  Luonnonvarat, eläimet ja kasvit (1)
-  Oikeusturva (1)
-  Perheiden palvelut (1)
-  Rakennettu ympäristö (4)
-  Sosiaalipalvelut (1)
-  Turvallisuus (1)
-  Työ ja työttömyys (2)
-  Yleiset tieto- ja hallintopalvelut (8)

Rajaa hakutuloksia

16 hakutulosta

Ammatti ja työ

 Terminologinen sanasto

Sanasto, joka kuvaa ammatin ja työn sekä liitännäisten käsitteiden välisiä suhteita ensisijaisesti kansallisen ammattinimikoodiston määrittelytarpeiden pohjalta.

Luonnos

Työ ja työttömyys

Julkisen hallinnon ydinsanastoryhmä

Asemakaavoituksen käsitteitä

 Terminologinen sanasto


Sanasto keskittyy kaavoituksen peruskäsitteisiin, joille tarvittiin yhteisesti sovitut määritelmät asemakaavoituksessa käytettäviä tietojärjestelmiä ja järjestelmien keskinäistä yhteentoimivuutta varten.

Luonnos

Rakennettu ympäristö

Sanastokeskus TSK

Digikaavoituksen käsitteitä

 Terminologinen sanasto


Sanasto liittyy käynnissä olevaan digikaavoituksen ja kaavan kansallisen tietomallin kehittämiseen. Kehittämistyötä tehdään Maankäyttöpäätökset-hankkeen Kuntapilotti-projektissa. Hanke liittyy maankäyttö- ja rakennuslain uudistamiseen. Hanksivut: <http://maankaytto.paikkatietoalusta.fi/kuntapilotti>. Sanasto on luonnos. Lausuntokierros toteutetaan keväällä 2019. Arvioitu valmistumisajankohta 06/2019. Lisätietoja: Satu Taskinen, etunimi.sukunimi@ym.fi.

Luonnos

Rakennettu ympäristö

Ympäristöministeriö

Identiteetin hallinta

 Terminologinen sanasto

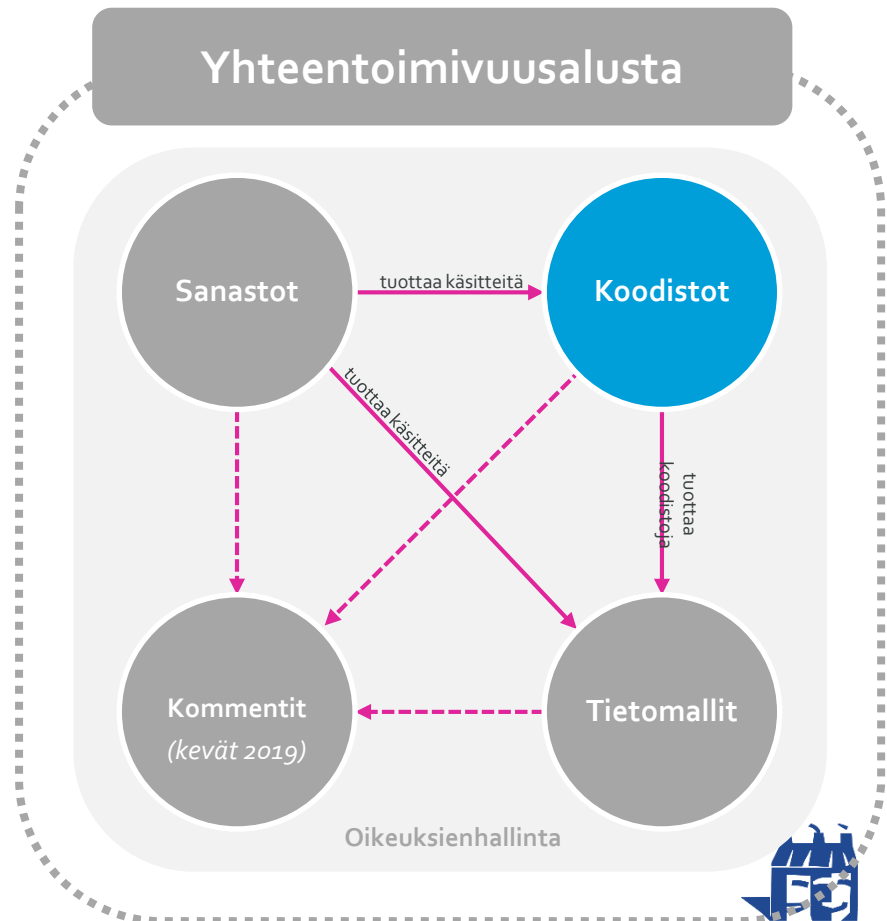
Yleiset tieto- ja hallintopalvelut

Luonnos

Julkisen hallinnon ydinsanastoryhmä

Koodistot-työkalu

- Luodaan, ylläpidetään ja julkaistaan koodistoja (myös hierarkkisia), koodeja sekä näiden metatietoja
- Yhdistetään samaan aiheeseen liittyvät koodistot yhteen saman otsikon (rekisterin) alle
- Luodaan koodistoista laajennuksia, versioita ja variantteja
- Voidaan tuoda tai viedä koodistoja CSV- tai Excel-tiedostoina
- Koodistoja voidaan hyödyntää REST -rajapinnan kautta












ETUSIVU

Koodistot työkalu on tekninen alusta julkishallinnon yhteisten koodistojen ja luokituksien hyödyntämiseen ja ylläpitämiseen. Koodistot kieliversioineen ovat ajantasaisia, luotettavia sekä yli toimialarajojen kaikkien selattavissa ja omaan käyttöön ladattavissa. Yhteisten koodistojen käyttö mahdollistaa yhteentoimivuuden järjestelmien välillä sekä tilastoinnissa.

Hae koodistoa

Kohdista haku myös koodeihin laajennuksiin

Rajaa tietoaalueella

-  Asuminen (1)
-  Koulutus (32)
-  Luonnonvarat, eläimet ja kasvit (3)
-  Rakennettu ympäristö (32)
-  Työ ja työttömyys (4)
-  Työnantajan palvelut (1)
-  Verotus ja julkinen talous (43)
-  Yksityinen talous ja rahoitus (1)
-  Yleiset tieto- ja hallintopalvelut (16)

Rajaa hakutuloksia

Kaikki organisaatiot

Kaikki rekisterit

Kaikki tilat

Kaikki koodistotyytit

130 hakutulosta

Aineiston elinkaaren tapahtumat

 Tutkimusaineistojen koodistot

Koulutus

Luonnos

CSC - Tieteen tietotekniikan keskus

Aineiston tila Tulorekisteri


 Tulorekisteri

Verotus ja julkinen talous

Luonnos

Verohallinto

Aineiston tilaajan tyyppi Tulorekisteri

 Tulorekisteri

Verotus ja julkinen talous

Luonnos

Verohallinto

Aineiston tyyppi, tulorekisteriin toimitettava Tulorekisteri

 Tulorekisteri

Verotus ja julkinen talous

Luonnos

Verohallinto

Aineiston tyyppi, tulorekisteristä jaettava Tulorekisteri

 Tulorekisteri

Verotus ja julkinen talous

Luonnos

Verohallinto

Alueen käyttötarkoituksen laji

 JHS

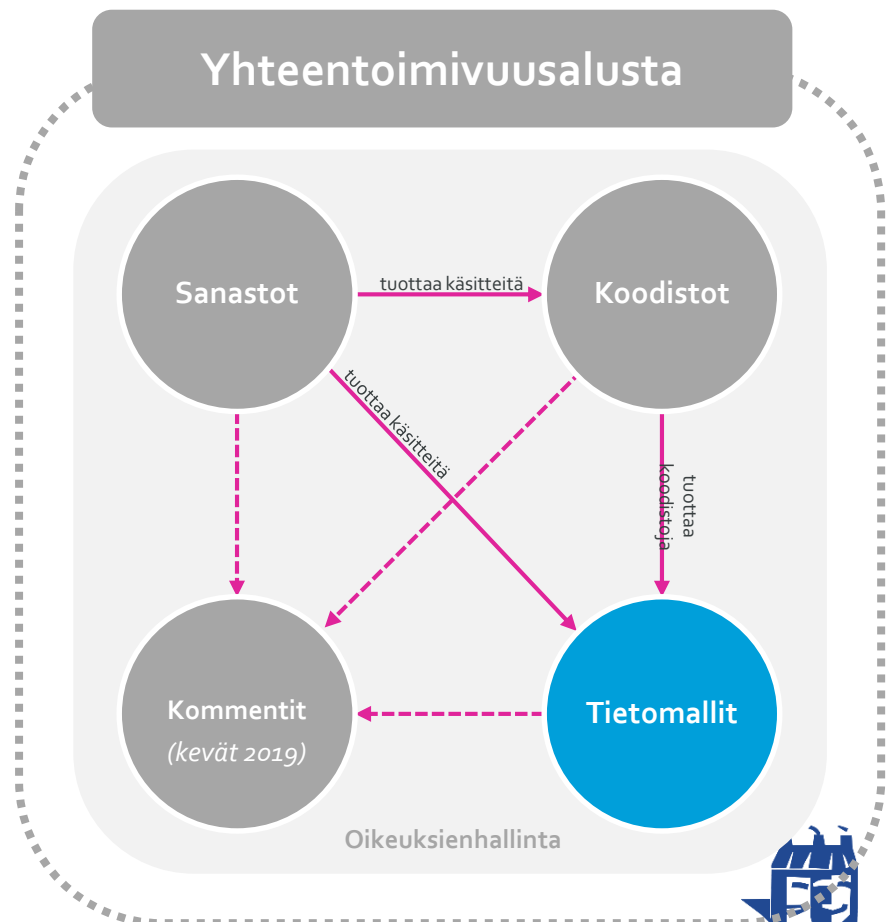
Asemakaavassa osoitettujen maankäyttö- ja rakennuslain 55 §:n mukaisten alueiden lajit.

Luonnos

Rakennettu ympäristö

Tietomallit-työkalu

- Kuvataan organisaation ulospäin jaettavien tietojen sisältöjä ja rakenteita eli ylläpidetään ja julkaistaan soveltamisprofieileja
- Kuvataan kansallisesti yhteisiä rekistereitä, lomakkeita tms. tietovirroissa kulkevaa tietoa
- Laaditaan soveltamisprofiili hyödyntämällä uudelleen käytettäviä tietokomponentteja
- Yleistetään soveltamisprofieileissa kuvattuja sopivia luokkia, attribuutteja ja assosiaatioita tietokomponenteiksi (tietoalueelle tai kansalliselle tasolle)
- Hyödynnetään toisilla työkaluilla määriteltyjä käsitteitä ja koodistoja
- Tuotetaan tietomalleista automaattisesti tekniset kuvaukset eri tiedostomuodoissa (JSON, XML...)



ETUSIVU

Tietomallit-työkalu on julkishallinnon organisaatioiden tietomallintajien ja tietoarkkitehtien työkalu, jolla kuvataan tietojärjestelmien ja rajapintojen tietosisäilytys- ja -rakennetta. Tietomallit-työkalu tekee julkishallinnon julkisten tietojen loogisesta tietomallityöstä avointa ja julkista, minkä vuoksi tietomallien hyödyntäminen organisaatioarajojen yli helpottuu.

Hae tietomallia ?

Hakusana (tietomallin nimi)



Tarkennettu haku

Rajaa tietoaueella ?

- Asuminen (3)
- Demokratia (1)
- Elinkeinot (1)
- Koulutus (11)
- Liikenne (1)
- Luonnonvarat, eläimet ja kasvit (1)
- Oikeusturva (1)
- Rakennettu ympäristö (6)
- Turvallisuus (1)
- Yksityinen talous ja rahoitus (1)
- Yleiset tieto- ja hallintopalvelut (29)

Suodata hakutuloksia ?

Kaikki organisaatiot

Kaikki tyypit

Kaikki tilat

45 hakutulosta

Aineistoeditorin tietomalli

Soveltamisprofili

Luonnos

Aineistoeditori on kansallinen aineistojen kuvailujärjestelmä. Valitut aineistot julkaistaan avoimessa Aineistokatalogi-palvelussa. Molemmat palvelut ovat beta-testausvaiheessa. ks. <https://aineistoeditori.fi/>

Koulutus Yleiset tieto- ja hallintopalvelut

Terveysten ja hyvinvoinnin laitos

Application profile for the Interoperability Workbench (en)

Soveltamisprofili

Luonnos

Data model used in the Interoperability Workbench for Core vocabularies and Application profiles. Notice that in this phase of development this model is VERY UNSTABLE and changes often! (en)

Yleiset tieto- ja hallintopalvelut

Yhteentoimivuusalustan ylläpito

Asunto-osakerekisteri ASREK

Soveltamisprofili

Luonnos

Asunto-osakerekisteri on rekisteri- ja palvelukokonaisuus, jonka avulla voidaan rekisteröidä osakehuoneistoja koskevat tiedot ja omistajamerkinnot sähköisesti sekä tukea uusien digitaalisten palvelujen kehittymistä. ASREK-järjestelmään liittyy osakeomistuksen sähköistämisen lisäksi tietoja useista jo olemassa olevista tietovarannoista. Asunto-osakerekisteri (ASREK) muodostaa perusrekisteriluontaisen loogisen kokonaisuuden asuntojen ja asunto- osakeyhtiöiden teknisistä ja taloudellisista tiedoista.

Asuminen Rakennettu ympäristö

Ympäristöministeriö



*Yhteentoimivuusalustassa tietomalli
ei siis ole pelkkä kuva.*

*Se on linkitettyä dataa:
kokoelma yksilöityjä, eri tahojen
määrittelemiä ja yhteisessä käytössä
olevia resursseja.*

Yhteentoimivuusmenetelmä

- Yhteisissä palveluissa ja prosesseissa käytettävät käsitteet on määriteltävä asianmukaisesti ja koneluettavassa muodossa
- Yhteisten tietomäärittelyjen tulee perustua sanastossa määriteltäviin käsitteisiin
- Tietomäärittelyt ovat uudelleenkäytettäviä tietokomponentteja (luokka, ominaisuus ja assosiaatio)
- Tietojärjestelmien väliset rajapinnat kuvataan soveltamisprofiilissa uudelleenkäyttämällä tietomäärittelyksiä
- Soveltamisprofiili muodostuu tiettyä asiayhteyttä tai käyttötarkoitusta varten kuvatuista uudelleenkäytettävistä tietomäärittelyksistä, eli luokista, ominaisuuksista ja niille asetetuista rajoituksista

Esimerkkejä tilanteista, joissa voit hyödyntää yhteentoimivuusalustaa

- Julkaiset sanastoja, koodistoja ja tietomalleja sekä rajapintakuvauksia myös oman organisaatiosi ulkopuolelle
- Haluat eroon sekalaisista ja hankalasti ylläpidettävistä sanasto- ja koodisto-Exceleistä
- Käynnistät hankkeen, jossa kerätään tietoja ja rakennetaan rekistereitä kansalliseen käyttöön (perustietovarannot ja -rekisterit)
- Haluat harmonisoida (meta)tietojen kuvauksia muiden kanssa → voi olla myös manuaalinen "lomake", jonka tietosisältö harmonisoidaan

Yhteentoimivuusmenetelmä ja –alusta tulevaisuuden tekijöinä

- Digitaalinen maailma rakentuu yhä vahvemmin ekosysteemeille ja kitkattomasti liikkuvalla, pysyvästi määritellylle tiedolle.
- Linkitetty tieto rakentaa linkitettyä, digitaalisesti rajatonta maailmaa, jossa tekoälyn merkitys tulee olemaan aivan kriittinen.
- Tekoälyä, käytännössä neuroverkkoja, pitää silti opettaa. Tämä vaatii määriteltyjä käsitteitä, malleja ja niiden määriteltyjä suhteita.
- Tekoälyn kannalta juuri **tiedon merkityksen eheys** on kriittistä; tekoälylle on voitava kertoa, että erilaisista ilmiöistä, kuten sanoista, huolimatta asioilla voi olla sama merkitys.
 - Eri tietoaueilla (knowledge domains) samaa ilmiötä tai oliota kutsua eri termillä, mutta merkityksen ilmaiseva käsitelmäärittely ja koneen tulkittavissa olevat käsitesuhteet rakentavat siltoja eri tietoaueiden väliin ja mahdollistavat tiedon mielekkään käsittelyn myös neuroverkoissa.

Kiitos!



Ota yhteyttä!

<http://yhteentoimiva.suomi.fi>
yhteentoimivuus@vrk.fi
[Yhteentoimivuus.slack.com](https://yhteentoimivuus.slack.com)