

Tietojen tuottaminen

(in English: <https://wiki.eduuni.fi/x/M5X0Gw>)

Integraatio organisaation omasta aineistojen metatietovarannosta aloitetaan aina suunnittelu- ja käyttöönottoprojektilla yhdessä CSC:n Fairdata-tiimin kanssa (fairdata (at) csc.fi).

- Projektissa on mukana CSC:n puolelta tekninen yhteyshenkilö, joka toimii yhteistyössä organisaation omien teknisten yhteyshenkilöiden kanssa, sekä etenemistä varmistamassa Metaxin tuoteomistaja

Projektin aikana integraatio toteutetaan Metaxin demoympäristöön, ja sovittuna ajanhetkenä otetaan demoympäristössä testattu toteutus käyttöön tuotantoympäristössä.

- Organisaation aineistojen kuvailutiedot tuodaan Metaxiin omaan aineistokatalogiinsa, jonka CSC luo organisaatiota varten.
 - Tarkempi kuvaus Metaxin tietokentistä ja pakollisista tiedoista löytyy täältä: <https://wiki.eduuni.fi/x/7w7PCg>
- CSC luo organisaatiolle service-tunnuksen (käyttäjätunnus ja salasana) aineistojen kuvailutietojen tuomiseksi yo. aineistokatalogiin.
- Organisaatio luo tarvittavat API-kutsut (sanomat), joilla aineistojen kuvailutiedot tuodaan.

Projektin eteneminen, esimerkki

Aloituskokouksessa sovitaan yhdessä organisaation kanssa tarkemmat projektivaiheet ja aikataulu.


Alla on esitetty esimerkki, millaisin vaihein projekti yleensä etenee.


1. Suunnittelu- ja testausvaihe



2. Tuotantoonmenovaihe



 = CSC:n vastuulla

 = Organisaation vastuulla

Esimerkkiprojekti tehtävittäin

Tehtävä	Kuvaus	Ympäristö	Vastuu
Organisaation oman aineistokatalogin luominen, demo	CSC luo organisaatiolle oman aineistokatalogin, jonne aineistojen kuvailutiedot tuodaan. Sovitaan yhdessä katalogin nimi ja muut perustiedot.	Demo	CSC
Organisaation Service-tunnuksen luominen, demo	CSC luo organisaatiolle käyttäjätunnuksen ja salasanan, joilla on oikeus tuoda ja päivittää organisaation aineistokatalogiin aineistojen kuvailutietoja	Demo	CSC
Sanomamallit, API-rajapintojen dokumentaatio, reunaehdot integraatiolle	CSC toimittaa sanomamallit ja API-rajapintojen dokumentaation, sekä sopii yhdessä organisaation kanssa integraation reunaehdoista (PID-käytännöt, ym.)	-	CSC
Sanomamalleihin ja API-rajapintoihin tutustuminen ja sanomien suunnittelu	Organisaatio tutustuu Metaxin API-rajapintoihin ja sanomamalleihin, ja kerää avoimet kysymykset yhdessä läpikäytäväksi	-	Organisaatio
Tarkempi sanomasisältöjen ja API-rajapintojen läpikäynti yhdessä ja määppästarpeiden tunnistaminen	Käydään yhdessä läpi organisaation tunnistamat API-rajapintojen käyttöön ja sanomien muodostamiseen liittyvät avoimet kysymykset. Tunnistetaan tässä vaiheessa myös määppästarpeet (käytettyjen koodistojen ja tunnisteiden muuntaminen organisaation käyttämistä arvojoukoista Metaxin tunnistamiin arvojoukkoihin sekä muutostarpeet metadatumallien välillä).	-	Organisaatio (CSC vahvasti tukena)

Mäppäykset	Tehdään mäppäystaulukot kuhunkin muunnosta vaativaan koodiarvojoukkoon liittyen. (Organisaation käyttämä arvo <-> Metaxin tunnistama arvo) <ul style="list-style-type: none"> Organisaatio toimittaa käyttämänsä arvot CSC:lle, ja CSC muodostaa ehdotuksen mäppäyksestä. Organisaatio hyväksyy lopullisen mäppäyksen. Metadatatallien välinen mäppäys voidaan tehdä joko yhteistyössä organisaation ja CSC:n välillä, tai organisaatio voi tehdä sen itsenäisesti.	-	CSC / Organisaatio
Sanomien muodostus	Organisaatio toteuttaa sanomat mäppäysmuunnoksineen	-	Organisaatio
Testaus	Organisaatio testaa sanomamuodostuksen ja API-rajapintojen käytön Metaxin demo-ympäristöön. CSC auttaa tarvittaessa tuotujen aineistojen kuvailutietojen katselmoinnissa. Huom! Aineistojen kuvailutietojen tuonin lisäksi testataan myös tietojen päivitys ja poistaminen.	Demo	Organisaatio
Testaushavaintoihin reagoiminen	Organisaatio kommunikoi CSC:lle, jos testauksessa havaitaan sellaisia haasteita tai muutostarpeita, joissa CSC:n tukea tarvitaan. Esim. sanomasisällöissä tai rajapintojen kutsuissa,	Demo	Organisaatio (CSC vahvasti tukena)
Tuotantoonmenosta sopiminen	Testauksen loppuvaiheessa sovitaan yhdessä tuotantoonmenon ajankohta.	Tuotanto	Organisaatio
Organisaation oman aineistokatalogin luominen, tuotanto	CSC luo organisaatiolle oman aineistokatalogin, jonne aineistojen kuvailutiedot tuodaan.	Tuotanto	CSC
Organisaation Service-tunnuksen luominen, tuotanto	CSC luo organisaatiolle käyttäjätunnuksen ja salasanan, joilla on oikeus tuoda ja päivittää organisaation aineistokatalogiin aineistoja	Tuotanto	CSC
Aineistojen kuvailutietojen tuonti tuotantoon, vaihe 1	Organisaatio tuo ensimmäiset aineistojen kuvailutiedot tuotantoympäristöön (1 tai korkeintaan muutama).	Tuotanto	Organisaatio
Aineistojen kuvailutietojen katselmointi tuotannossa	Varmistetaan, että 1-vaiheessa tuodut aineistojen kuvailutiedot ovat siirtyneet tuotantoon oikein ja hyvälaatuisina. Huom! Tässä vaiheessa varmistetaan myös, miltä aineistojen kuvailutiedot näyttävät Tiedejatutkimus.fi-portaalissa.	Tuotanto	Organisaatio (CSC vahvasti tukena)
Aineistojen kuvailutietojen tuonti tuotantoon, vaihe 2	Organisaatio käynnistää integraation täysimittaisesti (aineistojen kuvailutietojen tuonti, päivitys, poistaminen)	Tuotanto	Organisaatio

Ohjeita sanoman muodostamiseen:

- Käyttäjätunnus ja salasana välitetään sanomassa Base64 encoodattuna autentikointi-stringinä
 - Key = 'Authorization'
 - Value = 'Basic '+ base64 enkoodattu *login:password*
- Monet Metaxin kentät toimivat ns. "referenssidata" -tyyppisesti, eli arvona on tunniste (identifier) taustalla toimiviin erilaisiin koodistoihin
 - Tällaisia kenttiä ovat mm. lisenssi, pääsyoikeus, asiasanat, tieteenala, aineiston kieli, toimijoiden organisaatiotiedot, infrastruktuuri ja rahoittajaorganisaatio
 - Tällaisia tietoja tuotaessa on tärkeää, että tunnisteina käytetään Metaxin tunnistamia arvoja
 - Tarvittaessa CSC auttaa organisaatiota käytössä olevien tunnisteiden muuntamisessa Metaxin tunnisteisiin (ns. "arvojoukkojen mäppäys")
- Sivulla https://metax.demo.fairdata.fi/swagger/v2/#/dataset_api on lista API-operaatioista, joita voidaan käyttää. Samalta sivulta pääsee myös koekäyttämään operaatioita:
 - Avaa haluamasi operaatio klikkaamalla
 - Klikkaa "Try it out" -painiketta
 - Täytä haluamasi parametrit, ja aseta "dryrun" -parametrin arvoksi true (tällöin operaatioita ei oikeasti suoriteta Metaxiin saakka)
 - Klikkaa "Execute"

Joihinkin operaatioihin vaaditaan käyttäjätunnus, jota ei sellaisenaan voi sivulle antaa. Ilman käyttäjätunnustakin pääsee kuitenkin näkemään oikean syntaksin kutsulle.

- Muita hyödyllisiä operaatioita:
 - Lista kaikista aineistoista tietyssä katalogissa: https://metax.demo.fairdata.fi/rest/v2/datasets?data_catalog=<tähän katalogisi tunniste>&limit=1000
 - Esimerkki aineistosta (pakolliset kentät täytettynä): <https://metax.demo.fairdata.fi/docs/v1/datasets.html#retrieve-minimal-valid-dataset-template>
- Esimerkkisäältä kutsusta, jolla aineiston kuvailutiedot tuodaan Metaxiin: [Example_post.py](#)
 - Tarkempi kuvaus Metaxin tietokentistä (esimerkein) ja pakollisista tiedoista löytyy täältä: <https://wiki.eduuni.fi/x/7w7PCg>

Muita linkkejä

Referenssidatit: https://metax.fairdata.fi/docs/v2/reference_data.html
Yleinen Metaxin dokumentaatio: <https://metax.fairdata.fi/docs/>