

Tekninen ohje: Finton tarjoamien ontologioiden ja asiasanastojen käyttöönotto

Ennen teknisen ohjeen lukemista, suosittelemme tutustumaan [Finton ontologioiden käyttöönotto-ohjeistukseen](#).

Tekninen ohje: Finton ja ontologioiden käyttöönotto

Finton rajapintojen käyttö vs. sanaston lataus paikalliseen järjestelmään

Finto tarjoaa avoimen REST-tyyppisen rajapinnan [api.finto.fi](#), joka mahdollistaa sanastotoiminnallisuuksien integroinnin omiin sovelluksiin. Tässä ohjeessa käydään läpi esimerkkien avulla Finton sanastojen käyttöä kuvitteellisessa kuvailujärjestelmässä. Tämän ohjeen ohella kannattaa tutustua rajapinnan [tekniiseen dokumentaatioon](#).

Finton sanastot on mahdollista ottaa käyttöön kahdella tavalla: tekemällä yksittäisiä hakuja rajapinnan kautta tai ottamalla sanastosta paikallinen kopio omaa järjestelmää varten. Nämä vaihtoehdot eivät sulje pois toisiaan vaan ovat osa jatkumoa, jossa yhdessä ääripäässä ontologia (esim. YSO) tuodaan kertaalleen omaan järjestelmään eikä Fintoa tämän jälkeen käytetä ollenkaan, ja toisessa ääripäässä ontologiasta ei tallenneta yhtään mitään vaan kaikki toiminnot käyttävät Finton palveluita verkon yli.

Omaan järjestelmään tuomisessa on etuina ainakin riippumattomuus Finto-palvelun toimivuudesta ja mahdollisesti myös nopeat vasteajat (koska ei tule verkkoviivettä). Haittapuolena on se, että YSO ja muut sanastot kehittyvät koko ajan ja uusien sanastoversioiden hyödyntäminen vaatii raskaita kertapäivityksiä.

Finton palvelujen käytön etuina on se, että saatavilla on aina uusimmat ontologiaversiot sekä valmiiden rajapintojen käytön helppous verrattuna siihen, että vastaavat toiminnot toteutetaan omassa järjestelmässä. Haittana on riippuvuus Finton ja verkkoyhteyksien toiminnasta sekä ääritapauksessa se, että yksinkertaisetkin asiat (esim. tietueiden näyttö, jossa tarvitaan tietoa ontologian termeistä) vaativat rajapintakutsuja Fintoon.

Käytännössä suosittelemme välimuotoa, jossa osaa ontologian tiedoista ylläpidetään omassa järjestelmässä ja osa haetaan Finton rajapintojen kautta tarvittaessa. Esimerkiksi näin:

1. Sisällönkuvailussa käytetään Finton rajapintapalveluja sopivien ontologiakäsitteiden poimintaan
2. Omaan järjestelmään tallennetaan kaikki sellaiset ontologian käsitteet (URI ja termit), joihin on viitattu omassa aineistossa
3. Hakutoiminnoissa käytetään tarvittaessa Finton rajapintapalveluja esim. haun laajennukseen ja hakuvihjeisiin (katso myös...)

Esimerkki Fintoa hyödyntävästä kuvailujärjestelmästä

Esimerkissämme kuvailija käyttää YSOa suomen kielellä kuvaillessaan äänitallennetta. Käyttäjä hakee YSOsta käsitettä hakusanalla *kissa* ja saa nähdäkseen listan käsitteitä joiden termeihin hakuehto on osunut. Hän valitsee sen jälkeen listalta sopivimman käsitteen, joka tässä tapauksessa on *kös saeläimet*. Tässä vaiheessa järjestelmän on syytä ottaa talteen käsitteen päätermi (skos:prefLabel) kaikilla kolmella YSON tukemalla kielellä. Termien lisäksi kuvailutietueeseen on tärkeää tallentaa myös käsitteen URI (yksilöllinen tunniste). Tämän tunnisteiden avulla käsitteen tiedot voidaan myöhemmin tarkistaa mahdollisten muutosten varalta.

Toteutus Finto-rajapinnan kautta:

Käyttäjän syötettyä hakusanan hakulaatikkoon tehdään rajapinnan search-metodiin kutsu:

http://api.finto.fi/rest/v1/ysa/search?query=kissa*&lang=fi&fields=prefLabel

Jonka vastauksista näytetään hänelle lista mahdollisista käsitteistä. Kyselyn lopussa oleva parametri `fields` arvoineen `prefLabel` kertoo rajapinnalle että olemme kiinnostuneet käytettävistä termeistä myös muilla kielillä kuin hakukielellämme (suomi). `Fields`-parametrin avulla voi hakea myös muita skos-ominaisuuksia. Esimerkiksi jos hakutuloksissa halutaan näyttää jokaisen käsitteen yläkäsite, lisätään `fields`-parametriksi myös `broader`:

http://api.finto.fi/rest/v1/ysa/search?query=kissa*&lang=fi&fields=prefLabel+broader

kissa

kissaeläimet

kissamaki

kissanhuuto-oireyhtymä

kissankello

kissanäyttelyt

kissanäyttelyt

kissa

uri poista

Kun käyttäjä valitsee listasta käsitteen *kissaeläimet* tallennetaan käsitteen tiedoista järjestelmään vähintään URI sekä prefLabelit eli käytettävät termit eri kielillä. Nämä tiedot voidaan tallettaa myös kuvailtavaan tietueeseen siten, että kuvailun kohde on mahdollista löytää riippumatta siitä minkä kieliseen termiin tiedonhakijan termihaku kohdistuu. Tämän lisäksi voidaan tehdä haluttaessa kaikki käsitteen tiedot noutava kutsu rajapinnan data-metodille. Tämä on hyödyllinen jos käsite halutaan tallettaa järjestelmään kaikkine tietoineen tai tarvitaan ominaisuuksia, jotka eivät ole SKOSin mukaisia, joten niitä ei voi kysyä suoraan `fields`-parametrin kautta.

<http://api.finto.fi/rest/v1/yso/data?uri=http://www.yso.fi/onto/yso/p864>

Tämän kutsun vastauksesta poimitaan järjestelmästä tarpeista riippuen sopivat kentät.

Voit rajoittaa rajapinnan palauttavat käsitteet esimerkiksi `skos:Concept`ihin (tavalliset käsitteet), jolloin saat vain kuvailussa sallitut käsitteet etkä esimerkiksi käsiteryhmiä.

http://api.finto.fi/rest/v1/yso/search?query=*filosofia&lang=fi&type=skos:Concept

Toteutus tuomalla koko sanasto ohjelmistoon:

Sanaston voi ladata omaan järjestelmään REST-rajapinnan data-metodin kautta. Sanastodata on Fintossa tarjolla sekä RDF/XML- että Turtle-formaateissa. Finton sanastojen (esim. YSO) etusivujen alalaidassa on latauslinkit joista koko sanaston voi ladata joko RDF/XML tai Turtle -serialisoituna. Kumpikin sisältää saman informaation eli samat RDF-kolmikot, jotka noudattavat SKOS-tietomallia, mutta sanastosta riippuen käytössä voi olla myös tietomallin laajennuksia kuten ISO 25964 / iso-thes, jota käytetään esim. YSOssa käsiteryhmien mallintamiseen.

Järjestelmän tulee lukea termit suoraan RDF-datasta. Vähintään talteen on otettava URIt, käsitteen käytettävät termit (prefLabelit) eri kielillä ja mahdollisesti myös ohjaustermit (altLabelit). Nämä on talletettava järjestelmässä siten, että mahdolliset muutokset termeihin tai muihin käsitteen ominaisuuksiin voidaan päivittää helposti. Huomioitavaa on myös, että mahdolliset muutokset käsitteiden termeihin on vietävä myös järjestelmän kuvailutietoihin.

Näiden käsittelyyn kannattaa käyttää jotain valmista RDF-kirjastoa. Esimerkkeinä mainittakoon käyttämämme kirjastot: [EasyRdf](#) PHPlle, [RDFLib](#) Pythonille ja [Apache Jena](#) Javalle. Emme suosittele RDF/XML-muotoisen datan käsittelyä pelkällä XML-parserilla, koska RDF/XML:llä on mahdollista ilmaista sama tieto eri tavoin emmekä voi taata, että Finton kautta julkaistu RDF/XML-data noudattaa ikuisesti nykyistä ilmaisutapaa.

Toteutuksessa huomioitavat erityistapaukset

Käytöstä poistetut / korvatut käsitteet

Joissakin tilanteissa käsite voidaan korvata toisella käsitteellä. Korvatut käsitteet eivät näy hakutuloksissa, mutta kuvailutietueisiin niitä on saattanut jäädä ajalta ennen deprekointia. Näin on käynyt esimerkiksi YSON (Yleinen suomalainen ontologia) käsitteelle [tunteminen](#) joka on korvattu käsitteellä [tunteet](#). Jotkin poistetut käsitteet ovat ns. koostekäsitteitä, jotka kuvaavat YSAsa peräisin olevia monimerkityksisiä termejä jotka on YSOssa jaettu erillisiksi käsitteiksi.

Rajapinnassa tämä voidaan huomata deprekointiominaisuudesta (`owl:deprecated true`), käsitteen tyyppistä (`rdf:type skosex:DeprecatedConcept`) ja suhteesta (`yso:p307 dc:isReplacedBy p3485`). Kun järjestelmä huomaa käytetyn käsitteen deprekoidun, tulisi sen vaihtaa kuvailutietuiden tunteminen-käsitteen viittaukset tunteet-käsitteeseen. Joissakin tapauksissa yhdellä deprekoidulla käsitteellä voi olla useita korvaavia käsitteitä. Esimerkki käytöstä poistetusta käsitteestä:

<http://api.finto.fi/rest/v1/yso/data?uri=http://www.yso.fi/onto/yso/p307>