

Finton ontologioiden käyttöönotto-ohjeistus järjestelmäkehitykseen

Finton tarjoamat ontologiat ovat merkittävä väline kotimaisten tietojärjestelmien ja tietosisältöjen välisen yhteentoimivuuden edistämiseksi. Ontologiat on suunniteltu vastaamaan nykyhetken verkkoympäristön ja linkittyvän datan vaatimuksiin ja ne sisältävät huomattavia etuja asiasanastoihin ja vapaaseen asiasanoitukseen nähden. Kaikki edut eivät realisoitu automaattisesti, kun Finton ontologioita otetaan järjestelmässä käyttöön. Ne saadaan kattavasti ja käyttäjää parhaalla tavalla palvelevasti käyttöön vain, jos niihin kiinnitetään erityistä huomiota tietojärjestelmän teknisessä toteutuksessa. Järjestelmien ja aineistojen keskinäinen yhteentoimivuus ei myöskään toteudu, ellei ontologioita hyödynnetä yhdenmukaisesti. Tässä dokumentissa esitellään yhteentoimivuuden kannalta ensisijaiset ontologioihin liittyvät järjestelmän ominaisuudet sekä näitä täydentävät lisäominaisuudet.

Suosittelavat järjestelmän ominaisuudet

- **Pysyvien URI-tunnisteiden käyttö**

Finton ontologioissa, jokaisella käsitteellä on oma pysyvä URI-tunnisteensa. Monet ontologian edut tulevat mahdolliseksi juuri URI-tunnisteiden kautta. On suositeltavaa, että järjestelmä tallentaa käsitteiden URIt. Tämä tekee esimerkiksi termien päivittämisen helpoksi. Kielen rikkaus ja elävyys tuovat kuvailuun ja tiedonhakuun omia haasteitaan, mutta ontologian URIt kiinnittävät samaan käsitteeseen sen eri ilmaisutavat, eri kieliset vastineet ja muut käsitteeseen liittyvät tiedot.

- **Tuki ohjaustermeille**

Ontologiassa jokaisella käsitteellä on yksi käytettävä termi ontologiaan sisältyvillä kielillä. Käsitteellä voi lisäksi olla käytettävään termiin ohjaavia termejä (esim. synonyymeja), joita ei käytetä kuvailussa. Näin synonymia ei hankaloita käyttäjän työtä vaan toimii hänen tukenaan. Esimerkiksi YSO-ontologiassa ammeet-käsitteellä on ohjaustermeinä kylpyammeet ja amme. Kuvailujärjestelmissä kuvailija on ohjattava käyttämään ohjaustermin sijaan käytettävää termiä. Hakujärjestelmässä tiedonhakijalle on tarjottava samat hakutulokset riippumatta siitä tekikö tiedonhakija haun käytettävällä termillä vai siihen viittaavalla ohjaustermillä.

- **Tuki monikielisyydelle**

Useat Fintossa julkaistut ontologiat ovat monikielisiä. Tämä tarkoittaa, että käsitteillä on termivastine niillä kielillä, jotka kyseisessä ontologiassa ovat käytössä. Järjestelmissä voidaan käyttää joko kaikkia ontologian sisältämiä kieliä tai vain tiettyjä valittuja kieliä. Kuvailujärjestelmä on toteuttava siten, että kuvailu tarvitsee tehdä vain yhdellä kielellä ja hakujärjestelmällä haun voi tämän jälkeen tehdä millä tahansa tuetulla kielellä. Kuvailijalle on tarjottava valittavaksi kieli, jolla kuvailu tehdään. Hakujärjestelmässä tiedonhakijalle on tuotettava hakutuloksiksi samat aineistot huolimatta siitä, millä kielellä hän haun on tehnyt. Tiedonhakijalle on tarjottava valittavaksi kieli, jolla haku tehdään.

- **Ontologioiden säännöllinen päivittäminen**

Ontologian sisältämät termit muuttuvat ajan myötä, uusia käsitteitä otetaan ontologiaan ja vanhojen käsitteiden sisällössä saattaa tapahtua muutoksia. Kuvailu- ja hakujärjestelmiin on järjestettävä ontologioiden säännöllinen päivitys, jotta kuvailijalla ja tiedonhakijalla olisi käytössään ajantasainen ja ajankohtainen termistö. Lisäksi on varmistettava, että terminmuutosten päivittäminen onnistuu myös takautuvasti jo kuvailtuun aineistoon.

- **Muuhun kuin kuvailukäyttöön tarkoitettujen käsitteiden pois rajaus**

Ontologiat sisältävät myös hierarkisoivia, ryhmitteleviä ja koostavia käsitteitä, joita ei ole tarkoitettu kuvailukäyttöön. Lisäksi ontologia voi sisältää vanhentuneita käsitteitä, jotka sisältävät uudelleenohjauksen käytössä olevaan käsitteeseen. Nämä toimivat ontologian ylläpidon, selailun ja konepäättelyn tukena ja ne on merkitty omilla käsitetyypeillään. Järjestelmän on rajattava nämä käsitteet pois siten, että niitä ei tarjota sisällönkuvailijan tai tiedonhakijan käyttöön. Niitä voidaan kuitenkin hyödyntää välillisesti tiedonhaussa (ks. Lisäominaisuuksia).

- **Ontologian ja vapaiden asiasanojen ristiriidattomuus**

Ontologia mahdollistaa, että kaikki kuvailijat käyttävät tietystä aiheesta samaa termiä eivätkä esimerkiksi tämän synonyymeja. Mikäli tietojärjestelmässä käytetään ontologian ohella vapaita asiasanoja, on varmistettava että nämä eivät muodostu päällekkäisiksi tai ristiriitaisiksi: ei saa syntyä tilannetta, jossa sekä ontologiassa että vapaana asiasanana esiintyy identtinen termi tai kaksi samaan käsitteeseen viittaavaa eri termiä ja näillä tulee eri hakutulokset.

Lisäominaisuuksia

- **Käsitesuhteiden hyödyntäminen haussa.**

Hakijalle voidaan tuottaa hakutuloksiin myös aineistoa, joka on kuvailtu haussa käytetyn käsitteen alakäsitteellä (tai joissain tapauksissa myös muilla käsitesuhteilla yläkäsitteellä). Vaihtoehtoisesti hakijalle voidaan tarjota valittavaksi myös ylä- tai alakäsitteitä esim. fasettitoiminnolla. Näistä on erityisesti hyötyä silloin, kun alkuperäinen haku tuottaa tuloksia vain vähän tai ei lainkaan. Hierarkiaa ja muita suhteita voidaan hyödyntää myös hakutulosten ryhmittelyssä tai esim. learning to rank -menetelmien avulla.

- **Käsitesuhteiden hyödyntäminen suosittelussa.**

Sen lisäksi, että järjestelmä voi suositella lukijalle dokumentteja, jotka on kuvailtu samoilla käsitteillä, voidaan hyödyntää myös älykkäämmiin käsitteiden hierarkkisia tai muita suhteita suositusten rikastamiseksi, tarkentamiseksi tai järjestämiseksi.

- **Käsitesuhteiden hyödyntäminen kuvailujärjestelmässä.**

Kuvailujärjestelmä voi tarjota kuvailutyön tueksi tarkempaa tietoa valittavasta käsitteestä, esim. käsitteen hierarkian, määritelmän, linkkejä muihin käsitteisiin. Tämä auttaa kuvailijaa arvioimaan käsitteen soveltuvuutta asiayhteyteen ja tarjoaa pääsyn läheisiin käsitteisiin.

- **Ohjeet kuvailijalle**

Kuvailujärjestelmän aihekuvaluun ja ontologioiden käyttöön liittyvät perustiedot on syytä selventää ohjeistuksessa. Ohjedokumentaation lisäksi järjestelmään voidaan lisätä kuvailutyötä tukevia selitteitä.

Muuta luettavaa

- Tietomalli
- YSO-pohjaisten ontologioiden kehityslinjaukset
- Ontologioiden käyttöönotto-ohje