

Metadatasuositus julkaisuarkistojen tekstiaineistolle

Tämä on metadatasuositus suomalaisiin julkaisuarkistoihin tallennettavalle tekstiaineistolle. Suositus perustuu Dublin Core -standardiin, joka on tunnettu ja hyvin tuettu kansainvälinen elektronisten resurssien kuvailuun kehitetty formaatti. Dublin Core -kenttien monimerkityksistä semantiikkaa on suosituksessa kuitenkin tarkennettu tunnettujen metadatan aggregoijien vaatimusten ja standardista puuttuvien, tekstiaineiston kuvailussa välttämättömien kenttätarkenteiden osalta. Alunperin suosituksessa oli 62 kenttää, vuonna 2018 tehdyn päivityksen jälkeen kenttiä on yhteensä 77. Näiden kenttien toivotaan kattavan kaikki eri tekstityypeille asetetut vaatimukset. Kyseessä on yleinen metadatasuositus, jossa esitetyt kentät eivät ole ristiriidassa aiemman, edelleen käyttökelpoisen elektronisia opinnäytteitä koskevan metadataformaatin (Saijos ym. 2006) kanssa. Kenttien sisältöjä on kuitenkin päivitetty, muun muassa opinnäytetyyppien osalta. Kentät on listattu aakkosjärjestyksessä, koska yleisiä vähimmäisvaatimuksia kaikelle julkaisuarkistoon tallennettavalle tekstiaineistolle on mahdoton määritellä. Suosituksen sisältyy myös lista tärkeistä metadatakentistä sekä esimerkkejä kenttien käytöstä erilaisten aineistojen kuvailussa.

Ota yhteyttä

Dublin Core -metadata: dc-posti@helsinki.fi

Suosituksen tavoitteet

Metadatasuosituksen tavoitteena on yhtenäistää julkaisuarkistoihin tallennettavan tekstiaineiston kuvailua. Yhtenäisyys tehostaa tiedonhakua sekä kuvailutietojen kotimaista ja kansainvälistä vaihtoa esimerkiksi osana OAI-pohjaisia palveluita. Suosituksessa on pyritty käyttämään mahdollisimman paljon jo olemassa olevia ratkaisuja sen sijaan, että oltaisiin kehitetty kokonaan uusia käytäntöjä. Joissain kentissä on lisäksi mainittu Suomessa yleisimmin käytettyyn DSpace-julkaisuarkistojärjestelmään liittyviä huomioita. Mukaan on otettu vain sellaista tietoa, jota on pidetty tarpeellisenä mahdollisimman monessa julkaisuarkistossa. Tarvittaessa julkaisuarkistojen ylläpitäjät voivat laajentaa tätä yhteisesti sovittua joukkoa omien tarpeidensa mukaan, sekä ehdottaa lisäyksiä formaattiin.

Suosituksen käyttötarkoitus

Julkaisuarkistojen metadatasuositus on tarkoitettu ensisijaisesti elektronisen tekstiaineiston kuvailuun. Jos elektronisen teoksen rinnalla on olemassa painettu versio, voidaan sen perustiedot tallentaa rinnakkaisen ilmiäsen kenttiin. Suositusta voidaan soveltaa myös sellaisille teoksille, joista ei ole olemassa elektronista versiota. Jos pelkän painetun version pohjalta kuvaillusta teoksesta ilmestyy myöhemmin myös elektroninen versio, on kuvailutietoja muutettava siten, että elektroninen versio on ensisijainen kuvailun kohde. Painettuun versioon liittyvät tiedot siirretään tällöin rinnakkaista ilmiästä koskeviin kenttiin.

Muita huomioita

- Dublin Core Metadata Initiative:n suosittelemia DCterms-kenttiä ei otettu tässä suosituksessa käyttöön, sillä kentät eivät olisi yksinään riittäneet tarpeeksi rikkaan ja yksiselitteisen metadatan luomiseksi kaikelle tekstiaineistolle.
- Kenttien suuresta määrästä huolimatta tutkijoiden ja tutkimuslaitosten yhdistämiseen, eli niin sanottujen affiliaatioiden merkitsemiseen, ei löydetty tyydyttävää ratkaisua.
- Metadatasuosituksen yhteyteen lisätään myöhemmin ohjeita rajapintojen muokkaukseen keskeisten julkaisuarkistosta ulospäin välitettävien metadataformaattien tuottamiseksi.
- Suositusta päivitetään tarpeen vaatiessa, kenttien muutokset merkitään erilliselle alisivulle [9 Muutokset metadatasuositukseen](#).

Lähteitä/Lisätietoa:

Projektiryhmän yhteystiedot:

Työdokumentit yhteistyöverkoston kiwi-alueella: <https://www.kiwi.fi/x/hYOFAw>

Jyrki Ilva (etunimi.sukunimi[at]helsinki.fi)

[OpenAIRE Guidelines for Literature Repositories](#)

[OKM:n tiedonkeruun käsikirjat](#)

[Dublin Core - metatietostandardi](#) (Kansalliskirjaston ohjesivusto)

[Kansallinen metadataformaatti elektronisille opinnäytteille](#) (2006)

[Dublin Core metadataformaatin käyttöohje suomeksi](#) (PDF, 2022)

[Dublin Core Metadata Initiative](#)

[DC Scholarly Works Application Profile](#)

Suositus on koottu Kansalliskirjaston Tieteen avoin julkaiseminen (TAJUA) -hankkeen ja siinä perustetun asiantuntijaverkoston kanssa yhteistyössä. Verkostoon kuuluu julkaisuarkistoasiantuntijoita kaikista suomalaisista yliopistoista, AMKIT-konsortioista ja muista tutkimusorganisaatioista. Suosituksen kokoamiseen osallistuivat erityisesti: Jyrki Ilva (Kansalliskirjasto), Esa-Pekka Keskitalo (Kansalliskirjasto), Lassi Lager (Kansalliskirjasto), Antti Laurila (Itä-Suomen yliopisto), Minna Marjamaa (AMKIT-konsortio), Mikko Pennanen (Turun yliopisto), Päivi Rosenström (Kansalliskirjasto), Tiina Tolonen (AMKIT-konsortio), Tanja Vienonen (Kansalliskirjasto), Samu Viita (Kansalliskirjasto) ja Irene Ylönen (Jyväskylän yliopisto). Asiantuntija-apuna projektin eri vaiheissa olivat: Juha Hakala, Ulla Ikäheimo ja Jukka Lehmus Kansalliskirjastosta.

