

2021-05-03 Kokous

Kello 12–13

Osallistujat (16)	Organisaatio
Hanna Koivula	CSC
Maisa Borg	CSC
Jussi Pajari	Jyväskylän yliopisto
Tiina Sipola	Oulun yliopisto
Juha Hakala	Kansalliskirjasto
Mikko Heikkinen	LUOMUS
Tarja Mäkinen	Kansalliskirjasto
Pekka Nygren	Suomen metsätieteellinen seura
Turkka Näppilä	Tampereen yliopisto
Heidi Troberg	Vaasan yliopisto
Tuula Pääkkönen	Kansalliskirjasto
Tanja Lindholm	Helsingin yliopisto
Tuomas Alaterä	Tietoarkisto
Ari Asmi	Helsingin yliopisto
Jessica Parland-von Essen	CSC
Marjut Vuorinen	THL

Edellisestä kokouksesta: sovittu, että jatketaan ohjeiden muotoilua tässä 3.5.2021 kokouksessa sekä 5.5.2021 Avoimen tieteen kevätpäivien työpajassa. Viime kerralla käsitelty kohdan I1 ((Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.) ohjeita palveluntarjoajalle, data-asiantuntijalle ja tutkijalle.

Muistio

- **Avoimen tieteen kevätpäivät: työryhmälle varattu työskentelyaika 5.5. klo 13–14**
- Jatketaan tänään ohjeiden tekemistä eri rooleille kohdasta I2 (Meta)data use vocabularies that follow FAIR principles
- I2 Palveluntarjoajalle
 - Sanastojen oltava koneellisesti noudettavissa
 - Sanaston itsensä pitäisi olla avoimen rajapinnan kautta käytettävissä (esim. Finton sanastot)
 - Alakohtaisia sanastoja liian kapea ilmaisu entä paikan, henkilöiden ja organisaatioiden nimet lisätään myös "yleisiä sanastoja"
 - Ohjeisiin voisi muotoilla asiaa PIDEistä, sanastoilla oltava pysyvä tunniste
 - Sanastolla on pysyvä tunniste mutta pitääkö jokaisella sanaston termillä olla pysyvä tunniste?
 - PIDin määrittely? Voitaiko kiertää tässä dokumentissa? Esim. joidenkin mielestä Cool URIt lukeutuisivat PIDEiksi.
 - Monet ammatilliset sanastot voivat olla käytännössä PDF-dokumentteja, jolloin parhaimmassa tapauksessa viitattavana on DOI ja pahimmassa URL
 - Ohjeissa on hyvä asettaa rima tavoitetasolle
 - Monia käyttötapauksia ja filosofioita PIDEistä: esim. lajipuolella käsitteen muuttuessa annetaan uusi PID, toisaalta välttämättä ei
 - Aina kun mahdollista, pitäisi käyttää olemassa olevia sanastoja
 - Omat sanastot tulisi perustaa olemassa oleviin sanastoihin sikäli kun mahdollista
 - I2-kriteerin suosituksessa tulkitaan usein niin, että linkitetään FAIR Data Pointiin. Se ei kuitenkaan ole ainoa vaihtoehto ja on myös yksittäinen palvelu, joka voi kadota.
 - Järjestelmissä pitää olla mahdollisuus, että esitetään uusia käsitteitä ja tutkijoiden pitäisi olla mukana päivittämässä sanastoja
- I2 Data-asiantuntijalle/kouluttajalle
 - Kouluttajan tulee tuntea yleiset periaatteet ja keskeisimmät sanastot
 - Ei ole tarpeen tai mahdollista tuntea erikoisalojen sanastoa
 - Tärkeimmät järjestelmät oltava tiedossa
 - Usein kirjastoissa luetteloijilla paras käsitys sanastoista
 - Tuntee keskeisimmät yleiset sanastot, semanttiset hakupalvelut ja osaa ohjata kunkin alan sanastojen piiriin
 - "semanttisen yhteentoimivuuden yleiset periaatteet" tätä käsitettä olisi hyvä avata
 - Miten dataa ilmaistaan niin, että se ymmärretään samalla tavalla kuin muut samankaltaiset datat
- **Mihin jäätin: I2 ohjeet tutkijoille**
- Tapahtumavinkki: Kulttuuriperintölaitosten aineistojen hankinta, käyttö ja jatkokäyttö -työpaja pe 21.5.2021 9:00 - 12:15 <https://www.kansalliskirjasto.fi/fi/tapahtumat/kulttuuriperintölaitosten-aineistojen-hankinta-kaytto-ja-jatkokaytto>