

Kevättyöpaja 2021-03-18

Avoimen tieteen kevättyöpaja 18.3. kello 13-14.30
<https://avointiede.fi/fi/tapahtumat/avoimen-tieteen-kevattyopaja-2021>

Aihe: Semanttinen yhteentoimivuus

Ohjelma

13–13.15 Johdatus yhteentoimivuuteen FAIR-periaatteena (Hanna Koivula, CSC)

13.15–14 Breakout-groupit työskentelevät Flingassa (<https://edu.flinga.fi/s/EQT7U76>), tehtävä:

- 1) Mitä periaate tarkoittaa *käytännössä*
 - palvelun tarjoajalle (esim repositorio)
 - data managerille ja stewardille
 - tutkijalle

Mieti asiaa FAIR-periaatteiden I- (Interoperable) kohdan kannalta:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles.
- I3. (meta)data include qualified references to other (meta)data.

- 2) Muotoilkaa muutamia ehdotuksia hyväksi käytännöiksi

14–14.10 tauko

14.10–14.30 purku

Tavoitteet

- Työpajan myötä raakasetti ajatuksia semanttisen implementoinnin tukemiseksi
- Käsitteiden selkiyttäminen itsellemme: esim. mikä on dataa, mikä metadataa (ja metadatan eri tasot), mitä vaaditaan toistettavuuteen, käytännön ohjeistusten muotoilua
- Avoimen tieteen kevättyöpajan jälkeen, 22. päivä työryhmän kokouksessa voidaan käsitellä työpajan antia ja katsoa seuraavia askelmerkkejä (esim. mahdolliset tuotokset)

Ryhmäjako yhteentoimivuuden alakohtien mukaan

- Interoperable: The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.
- **Breakout group 1: I1.** (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- **Breakout group 2: I2.** (Meta)data use vocabularies that follow FAIR principles
- **Breakout group 3: I3.** (Meta)data include qualified references to other (meta)data

Muistiinpanot kevättyöpajan sessiosta

(Pohjana [yhteinen Google docs -dokumentti](#))

1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

1. **Palvelun tarjoajalle**
 - a. Pelkkä koneluettavuus ei riitä - myös ihmisten on ymmärrettävä
 - b. Eri alojen sanastojen käyttöönotto joustavasti palveluissa
 - c. Palvelun/palveluntarjoajan/taustajärjestelmän pitää pystyä generoimaan tutkijan kuvailutietojen perusteella semanttisesti yhteensopivaa metadataa
 - d. Palveluntarjoaja käyttää jotain linkitetyn datan skeemaa (esim. RDF tai JSON LD)
 - e. Tieteenalakohtaiset sanastot tulee ontologisoida ja ylläpitää kytkettynä ontologiapalveluun
 - i. Tulee olla tietoinen, millä aloilla on tärkeää käyttää suomalaisia vs. kansainvälisiä sanastoja
2. **Data-asiantuntijalle / kouluttajalle**
3. **Tutkijalle**
 - a. Harvemmin tutkijan käsissä, usein kiinni tarjolla olevista palveluista

2. (Meta)data use vocabularies that follow FAIR principles.

- **Palvelun tarjoajalle**

- Rajapintoja eri systeemeihin helposti linkitettävissä
- Palvelun tarjoajan määritelmä? Tietokannan vai sanaston ylläpitäjä?
- Omia sanastoja ja tietomalleja rakennettaessa tulee kiinnittää huomiota että ne julkaistaan (ja kuvataan) FAIR periaatteita noudattaen
- Ontologioiden ja rajapintojen ylläpidon tarve
- Miten huolehditaan viitattavuudesta, pitäisikö sanastosta ottaa vuosikerta (ettei viitata alati muuttuvaan asiaan)
- Sanaston on itsessään oltava FAIR
- **Data-asiantuntijalle / kouluttajalle**
 - Kouluttajan rooli: ymmärtäisi yleiset periaatteet ja merkityksen, esim. Kirjastohenkilökunnan ei ole realistista tietää kaikkia sanastoja
 - Lähituessa tieteenalasanastojen tuntijat
 - Sanastorekistereitä, esim. BARTOC (<https://bartoc.org/>) ja LOV (<https://lov.linkeddata.es/dataset/lov/>)
 - Esimerkkilista ontologioista: https://www.w3.org/wiki/Lists_of_ontologies
 - <https://bioportal.bioontology.org/ontologies>
 - <https://www.ebi.ac.uk/ols/ontologies/uo>
- **Tutkijalle**
 - Mitkä ovat valintakriteerit tutkijan näkökulmasta sanaston valinnalle? Onko organisaatiolla resursseja/kapasiteettia antaa neuvontaa tässä tapauksessa?
 - Tutkijan olisi hyvä antaa palautetta, jos sanasto ei ole riittävä tai kaipaa muutoksia.
 - Mitä rikkaampi kuvailu niin sen parempi/kuvailun rikastaminen pitkin matkaa
 - (Meta)tietojen säilyttämispaikka voi vaikuttaa sanastojen käyttämisen mahdollisuuteen
 - Oman alan sanaston hallinta osa tutkijan ammattiosaamista
 - Kerrottava, mihin sanasto perustuu

3. (Meta)data include qualified references to other (meta)data.

- **Palvelun tarjoajalle**
 - Palvelun tarjoajan olisi huolehdittava, että dataan viittaaminen on mahdollista
 - Kuvaus mahdollistaa viittaamisen julkaisuihin ja muihin datoihin
 - Tarjoaa luokitteluja, sanaston, PIDit
 - Lisenssit
- **Data-asiantuntijalle / kouluttajalle**
 - Kouluttajan on tunnettava eri alat, ymmärtää itse ja osaa selkeästi selittää vaihtoehdot
 - Yhdessä tutkijan kanssa asioiden selvittäminen: voisiko tämä sopia teille, että tiedosta tulisi yhteentoimivampaa
 - Ontologioissa on implisiittistä oletusta tieteenaloista, mikä on haastavaa
- **Tutkijalle**
 - Läpinäkyvyys, helppokäyttöisyys haaste myös palvelukehittäjille
 - Tietää mistä saa apua
 - On tietoinen/kiinnostunut asiasta
 - Osaavat vaatia käyttäjäystävällistä ja ymmärrettävää palveluja

Keskustelua seuraavista askelista:

- Tehdään lista ontologioista/niitä kokoavista palveluista
- Muuta keskusteltavaa
 - Eri aloilla on erilaisia odotuksia - puhuttava myös riskeistä ja hyvä sisällyttää ne ohjeistuksiin
 - Minimivaatimuksen täyttäminen tärkeintä ns. nyrkkisääntöjen luominen
 - Tarvitaan ohjeistusta, mitä kannattaa tehdä ja mitä ei
 - Suomenkielinen sana semantic artefactille?
 - Koneluettava malli, jolla mallinnetaan tietoa
 - Ei sama asia kuin tietomalli
 - Google translate: semanttinen esine
 - Tuotos, asia, työkalu, työväline...
 - Palataan käsitteen suomentamiseen myöhemmin