



Tutkimusinfrastruktuurit tutkimustietovarannossa -työryhmä

Kokous 17.9.2021

Aika: klo 9.30–11.00

Läsnä:

- Beatrice Bilcu, Aalto-yliopisto
- Juha Herrala, Tampereen yliopisto
- Timo Hellgren, VTT
- Niku Kivekäs, Ilmatieteen laitos
- Jaakko Kuha, Turun yliopisto
- Anttoni Lehto, Turun ammattikorkeakoulu
- Marko Peura, Helsingin yliopisto
- Seliina Päällysaho, Seinäjoen ammattikorkeakoulu
- Merja Särkioja, Suomen Akatemia
- Timo Taskinen, Jyväskylän yliopisto
- Hanna Tolonen, Terveiden ja hyvinvoinnin laitos
- Walter Rydman, CSC, sihteeri + valmistelija

Päätökset

- (1) Käsitellään tulevissa kokouksissa omana aiheenaan tutkimusinfrastruktuurien avoimuuden seuranta ja vaikuttavuutta. Valmistellaan asia yhdessä TSV:n koordinaation kanssa.
- (2) Valmistellaan keskusteltavaksi linjaus tutkimusinfrastruktuurien määrittelemiseksi tutkimustietovarannossa. Otetaan huomioon myös yhteydet kansainvälisiin tutkimusinfrastruktuureihin. Millainen tietokokonaisuus tutkimustietovarannossa ymmärretään tutkimusinfrastruktuuriksi?
- (3) Päivitetään tutkimusinfrastruktuuritiedon käyttötapaukset, -tarkoitukset ja käyttäjäryhmät.
- (4) Valmistellaan keskusteltavaksi ehdotus tutkimusinfrastruktuurien luokittelutiedoista tutkimustietovarannossa.
- (5) Käsitellään tietojen toimituksen prosesseja tulevissa kokouksissa.
- (6) Selvitetään ajatusta modulaarisemmasta mallista.

Muistio

1. Lyhyen esittäytymiskierroksen jälkeen todettiin kaikkien ryhmän jäsenten olevan paikalla.
2. Käytiin lävitse tutkimusinfrastruktuurien nykytilanne tutkimustietovarannossa ja tiivis historia.
 - Tutkimustietovarannossa yhdistyy toisaalta Finn-ARMAn tutkimusinfrastruktuuriryhmän aiemmin tekemä työ yhteisen infrojen metatietomallin eteen ja toisaalta ATT-hankkeen yhteydessä toteutettu tiekarttainfrastruktuurien tietopankki, jonka tiedot soveltuvin osin vietiin tutkimustietovarantoon.
 - Nykyinen tietomalli muodostuu ajatukselle tutkimusinfrastruktuurista, jolla on yksi tai useampi palvelu, joita voidaan tarjota yhdestä tai useammasta paikasta. Infrastruktuurit voivat myös liittyä toisiinsa.
 - Tutkimusinfrastruktuurien tiedot toimitetaan yksinkertaisella syöttölomakkeella. Suunnitteluvaiheessa sen rinnalle kaavailtiin toisena tapana suoraa tiedonsiirtoa toimijoiden tietojärjestelmistä. Ratkaisu jätettiin kuitenkin suunnittelematta, koska tutkimusinfrastruktuurien tietoja ei useimmiten ole saatavissa riittävän yhdenmukaisina.
 - Nykyisessä mallissa infrastruktuurien luokittelu on mahdollista muutamilla koodistoilla:
 - Tieteenalat
 - Laitteisto/aineisto/palvelu
 - ESFRI-domain
 - MERIL-luokittelu. *(tämä on ilmeisesti jäänyt vain dokumentoiduksi, mutta ei nykyään juurikaan käytetä)*
 - Pysyviä tunnisteita infrastruktuureilla ei vielä tutkimustietovarannossa ole. Poikkeuksena aiemman tietopankin infrastruktuurit, joilla on URN-tunniste.
3. Nykytilan kuvauksen jälkeen keskusteltiin laajemmin tutkimusinfrastruktuuritiedon ylläpitämisen tavoitteista tutkimustietovarannossa. Keskustelussa nousi esille seuraavia asioita:
 - Tällä hetkellä varannossa olevat tutkimusinfrastruktuurit ovat hyvin heterogeenisia eikä kokonaiskuva oikein hahmotu. Tavoitteet ovat kuitenkin oikean suuntaiset. Tutkimusinfrastruktuurien tietojen yhdistämisestä yhteen paikkaan saataville on pelkästään hyötyjä.

- Tärkein ratkaistava kysymys on päättää mistä “asiasta” tietoja kerätään ja luoda kaikille yhteinen linja. Mikä on tutkimusinfrastruktuuri, jonka halutaan näkyvän tutkimustietovarannossa? Eri toimijoilla voi olla erilaiset intressit näkyvyydelle.
 - Tutkimusinfrastruktuurien erilaisuus ja organisaatioiden erilaiset tulkinnat nähtiin merkittäväksi vaikeudeksi. Kun on täysi vapaus päättää, toimitetaan tietoja merkittävästi erilaisilla karkeustasoilla. Se johtaa nykyisenkaltaiseen vaikeasti hahmotettavaan kokonaisuuteen.
 - Esimerkkinä mittausasemien verkosto, joista jokaisen voi hahmottaa yhtenä infrastruktuurina. Toisaalta yhdessä ne kuuluvat laajempaan infrastruktuuriin ja voivat vielä tämän ohella kuulua myös johonkin toiseen infrastruktuuriin.
 - Miten tällaisen infrastruktuurin tiedot ilmoitetaan? Mitä tietoja toimitetaan? Kuka toimittaa? Kuinka tiedot linkittyvät kokonaisuudeksi?
- Kun tutkimusinfrastruktuurien tietojen keräämisen fokus on päätetty, on suunniteltava prosessit yhdenmukaisesti.
 - Miten varmistetaan tietojen pysyminen ajantasaisena? Jossain vaiheessa prosessia manuaalinen työ on välttämätöntä. Miten asia ratkaistaan vaivattomasti?
- Edelliseen liittyen on selvennettävä kohderyhmiä. Kenelle tutkimusinfrastruktuurien tiedot on tarkoitettu?
 - Todettiin, että päivitetään ja kuvataan käyttötapaukset, käyttäjäryhmät ja käyttötarpeet nykyistä tarkemmin.
 - Tutkimustietovaranto pystyy oikein suunniteltuna vastaamaan myös raportoinnin vaatimuksiin (esimerkiksi kuinka palveluja on käytetty, mitä niillä on toteutettu, jne.
 - Käyttäjäryhmiä:
 - Oma organisaatio ja muut tutkimustoimijat.
 - Tutkijat
 - Yritykset (tutkimusinfrastruktuurien ostopalvelut). Kuinka välitetään viesti tätä tarkoitusta varten informatiivisesti?
 - Rahoittajat
 - Viestinnälliset tarpeet - tutkimustietovaranto viestinnän välineenä.
- Tutkimusinfrastruktuurien avoimuuden seuranta ja vaikuttavuus on asiana ajankohtainen. Sitä edistää TSV:n avoimen tieteen koordinaatio.
 - Avoimuuden osalta tarvitaan yhteisiä linjauksia ottaen huomioon myöhemmin esille tulevat hierakiatasot.
 - Aiemmin tutkimustietovarantoa suunniteltiin kenties idealistisesta lähtökohdasta, jossa kaikki siellä näkyvät tutkimusinfrastruktuurit ovat jollain

tavalla avoimia käytön suhteen (infrastruktuurin ulkopuolinen pääsee käyttämään vähintään tietyt ehdot täytettyään).

- Päätettiin käsitellä asiaa tulevilla kokouksilla ja kutsutaan TSV:n avoimen tieteen koordinaatioon mukaan.
- Tutkimusinfrastruktuurien hierarkiasta syntyy polveileva ja yksityiskohtainen keskustelu.
 - Korkeakoulun sisäiset hierarkiatasot tutkimusryhmästä laajoihin yhteistyöinfrastruktuureihin. Hierarkiatasot vastaavat erilaisiin tiedontarpeisiin organisaation sisällä.
 - Tutkijalle tietyn analyysimenetelmän tai laitteen saatavuus voi olla tärkeä, mutta ei tarvetta tietää mihin infrastruktuuriin ne kuuluvat.
 - Korkeakoulua kiinnostaa useimmiten keskitaso ja sitä ylempät hierarkiatasot esimerkiksi laajempien ja paremmin resursoitujen infrakokonaisuuksien muodostamiseksi. Toisaalta tutkimuksen tukemiseksi tarvitaan myös alemman tason tietoa.
 - Halutaanko tutkimustietovarannolla palvella näitä kaikkia tarpeita vai keskittyä vain joihinkin?
 - Kansallisen tason yläpuolella on vielä kansainvälinen ulottuvuus. Monella tutkimusinfrastruktuurilla on kansainvälisiä yhteyksiä ja toimintoja - suomalaiset tutkimusinfrastruktuurit ovat osa kansainvälistä kokonaisuutta joilla on omat portaalinsa.
 - Kuinka kansainvälinen taso otetaan tutkimustietovarannossa huomioon?
 - Viestinnällisestä näkökulmasta laboratoriotasolta on syytä nousta korkeammalle hierarkiatasolle. Näytetään laitteistot yhteen nivovat infrastruktuurit ja näiden verkostot.
 - Onko syytä valita tässä vaiheessa jokin hierarkiataso, jolla lähdetään liikkeelle?
- Infrastruktuurien muodostamien verkostojen ja kytkösten esiin tuominen on tärkeää valitusta hierarkiatasosta riippumatta. Näiden kautta saadaan näkyväksi suuremmat kokonaisuudet. Liian matalalla tasolla hukataan seurattavuus.
 - Muina huomioina mainittiin myös Suomen kansainvälisten tutkimusinfrastruktuurien jäsenyydet sekä yhteydet EOSCiin.
- Kysymys pysyvistä tunnisteista tutkimusinfrastruktuureille nähtiin asiaksi, joka on tärkeä ratkaista pian. Erityisesti, jos infrastruktuurit halutaan nähdä verkostona.
 - Toimivilla tutkimusinfrastruktuureilla voi myös olla omia tunnisteitaan. Esimerkiksi ACTRIS-infrastruktuurilla on DOI-etuliite sillä tehtyihin aineistoihin.

- Tunnisteet eivät silti riitä. Huomiota on kiinnitettävä myös nimeämisen selkeyteen.
- Tutkimusinfrastruktuurit tarvitsevat myös niiden löytämistä ja hahmottamista helpottavia luokittelutietoja.
 - Luokittelutiedot ja muut tutkimusinfrastruktuureihin liittyvät koodistot otetaan käsiteltäväksi tulevilla kokouksilla.
- Tutkimusinfrastruktuuriolosuhteiden voisi toteuttaa nykyistä modulaarisempaa pysyviä tunnisteita hyödyntäen. Tutkimusinfrastruktuurin tietojen toimittaminen voisi muodostua helpommaksi, jos tiukasta infra/palvelu-rakenteesta luovutaan. Tutkimusinfrastruktuuri voisi olla yhtäaikaista itsenäinen tutkimusinfrastruktuuri, jonkin toisen infrastruktuurin palvelu sekä liittyä laajempaan tutkimusinfrastruktuuriin.

Tietoja toimitettaessa annetaan tarvittavat kuvailutiedot, luokittelutiedot ja luodaan kytkökset muihin infroihiin (liittyvä infra tai hierarkkinen "suunnattu" linkki, jolloin muodostuu palvelu). Järjestelmä antaa pysyvän tunnisteiden toimitetulle kokonaisuudelle. "Modulien" väliset linkit muodostavat verkoston. Samalla helpottuu hankalasti määriteltävä kysymys: Mikä on tutkimusinfrastruktuuri?