

TTV Ohjausryhmän muistio

Aika: ke 9.10.2019 klo 12-15

Paikka: nh. Väinämöinen, OKM (Meritullinkatu 10)

Kutsutut:

- Opetusneuvos Jukka Haapamäki (puheenjohtaja), opetus- ja kulttuuriministeriö
- Opetusneuvos Matti Kajaste, opetus- ja kulttuuriministeriö
- Tutkimuksen kehittämispäällikkö, Timo Taskinen Jyväskylän yliopisto
- Tutkimuspäällikkö Seliina Päällysaho, Seinäjoen ammattikorkeakoulu
- Kehittämispäällikkö Minna Liikala, Terveystieteiden tutkimuskeskus
- ~~Ylilääkäri Ari Lindqvist, Helsingin ja Uudenmaan sairaanhoitopiiri ESTE~~
- Tutkimuksen tietoasiantuntija Yrjö Lappalainen, Tampereen teknillinen yliopisto
- Palvelupäällikkö Anne Sunikka, Aalto-yliopisto
- Johtaja Marja Kylämä, Suomen Akatemia
- Ennakointipäällikkö Soile Ollila, Innovaatorahoituskeskus Business Finland
- ~~Tiedejohtaja Kalle Korhonen, Koneen Säätiö ESTE~~
- Tietojärjestelmäasiantuntija Jyrki Ilva, Kansalliskirjasto
- Suunnittelupäällikkö Janne Pölönen, Tieteellisten seurain valtuuskunta (klo 14.15 saakka)
- Kehittämispäällikkö Ari Leppälahti, Tilastokeskus
- Kehityspäällikkö Hanna-Mari Puuska (sihteeri), CSC - Tieteen tietotekniikan keskus

Vierailijat

- Projektipäällikkö Tiina Strengell (sihteeri), CSC - Tieteen tietotekniikan keskus
- Tietoarkkitehti Tommi Suominen, CSC - Tieteen tietotekniikan keskus

Asialista:

1. Kokouksen avaus ja työjärjestyksen hyväksyminen
 - Puheenjohtaja avasi kokouksen klo 12.00
2. Edellisen kokouksen pöytäkirja [Liite 1](#)
 - Edellisen kokouksen pöytäkirja hyväksyttiin.
 - Kerrottiin, että edustajat Tutkijan tiedot -työryhmään on nimetty ja ryhmän ensimmäinen kokous järjestetään lokakuun loppupuolella.
3. Tilannekatsaus (Tiina Strengell) Liite 2

Tilannekatsaus esiteltiin liitteen 2 mukaisesti.

Nimikilpailu

Nimikilpailuun liittyen kysyttiin, onko ajatuksena valita nimi, joka toimii kaikilla kieliversioilla? Asiaa ei olla päätetty, mutta suomen- ja ruotsikielisten nimien olisi hyvä olla toisiaan vastaavat.

Todettiin, että ennen valintaa olisi hyvä selvittää, onko nimiä käytetty muissa yhteyksissä. OKM pyytää Kotukselta (Kotimaisten kielten keskus) konsultointiapua nimen valinnassa. OKM päättää lopullisen nimen marraskuun aikana.

Organisaatioiden tiedot tutkimustietovarantoon

Organisaatioiden tietojen suhteen ajatuksena on, että tarvetta hakea ja päivittää tietoja manuaalisesti olisi mahdollisimman vähän. Kysyttiin, onko Y-tunnus avaintunniste vai millä perusteella tiedot haetaan eri palveluista? Tunnistamiseen pyritään käyttämään Y-tunnusta. Fintossa tai opintopolun koodistoissa ei ole kuitenkaan y-tunnuksia, joten sieltä tiedot täytyy hakea organisaation nimellä. Oltiin sitä mieltä, että nimellä hakeminen voi osoittautua ongelmalliseksi ja erityisesti nimien muuttumisen myötä voi haku olla hankalaa, koska esimerkiksi SA:sta ja HY:sta palveluissa on useita erilaisia muunnoksia.

Organisaatiotietoja haettaessa täytyy arvioida myös datan laatua. Kysyttiin, mihin organisaatioiden luokitus perustuu. Perustuu sektoriluokitukseen. Kysyttiin myös, onko Tilastokeskuksessa oppilaitoskoodi ja y-tunnus yhdistetty? Pitäisi olla yhdistetty ja löytyä organisaatiokoodin kautta. Rahoittajien osalta kerätään tietyt perustiedot.

Tutkimusorganisaatioiden tietojen haun osalta selvitetään tarkemmin erityisesti hakua nimen perusteella.

Tapahtumia

Tutkimustietovaranto on esillä syksyn ja talven aikana muutamassa tapahtumassa. RDA-tapaamista varten on tehty tutkimustietovarannosta posterit, jota voidaan hyödyntää muuallakin. Tutkimustietovaranto on esillä myös Digisfäärissä sekä euroCRIS-tapaamisessa marraskuussa. Tammikuussa järjestetään PIDapalooza, jonne on ehdotettu esitystä PID –mikropalvelusta. Kysyttiin, onko tullut esiin PIDia, jota käytetään yleisesti infroille. Ei ole tullut vastaan. Nyt on ajatuksena antaa infroille URN-tunnukset, joita on annettu myös aiemmin tutkimusinfrastruktuurien tietopankissa oleville infroille.

4. Tutkimusinfrastruktuurien tietopankin tietovirrat (Tiina Strengell) [Liite 3](#)

Tutkimusinfrastruktuurien tietopankkia esiteltiin tietojen keräämiseen ja aiheesta toteutettuun kyselyyn liittyen liitteen 3 mukaisesti.

Kerrottiin, että Jyväskylän yliopistolla on otettu käyttöön infrojen varausjärjestelmä. Kommentoitiin, että tietojen päivitys on hankalaa ja pitäisi miettiä käytettävyyttä huomioiden, miten infraa pystyisi päivittämään.

Ongelmallinen kysymys on myös, kuinka tutkimusinfrastruktuuri määritellään. AMK:illa ei ole samanlaiset kriteerit kuin yliopistoilla ja niiden osilta käyttötarkoitus olisi hyvä ottaa huomioon.

Keskusteltiin tutkimusinfrastruktuurien master datan keräämisestä ja säilyttämisestä tutkimustietovarannossa. Käytännössä infrastruktuurien tietoja tällöin kerättäisiin ja ylläpidettäisiin tutkimustietovarannossa. Ehdotettiin, että tämän rinnalle voisi luoda Justusta vastaavan palvelun. Kustannukset palvelun suhteen olisi kuitenkin mietittävä tarkemmin. Kustannukset voivat olla esteenä pienempien organisaatioiden liittymiselle.

Päätös: CSC valmistelee seuraavaan kokoukseen toteutussuunnitelman tutkimusinfrastruktuurien syöttöpalvelusta kustannusarvioineen sisältäen alustavan ehdotuksen palvelun sisältämistä toiminnallisuuksista, hallinnointimallista ja kustannusten jakomallista.

5. Tutkimustietovarannon kehitys pitkällä aikavälillä (Hanna-Mari Puuska) [Liite 4](#)

Esiteltiin alustavia suunnitelmia tutkimustietovarannon jatkokehitykseen liitteen 4 mukaisesti. Liitteessä esiteltyjä kohtia ei ehditä tehdä hankkeen aikana, mutta niitä olisi hyvä edistää jatkossa. Hanke kestää ensi vuoden loppuun asti ja portaalin lanseeraus on ajateltu ajoitettavan ensi keväälle.

Uudet ja laajennetut tietokokonaisuudet

Tutkimusorganisaatioiden hanketietoihin liittyen kommentoitiin, että jos halutaan muutkin, kuin rahoituspäätöstiedot, täytyy saada tiedot organisaatioista. Tiedot yliopistojen ja organisaatioiden välillä eivät ole yhteneviä.

Kommentoitiin, että tutkimustietovarannon tietoja voisi rikastaa esimerkiksi Julkaisufoorumin julkaisukanavatiedoilla. Tietoa voisi kerätä myös APC –maksuista. Patenttietoja voisi löytyä Euroopan patenttitoimiston tietokannasta (EPO – PATSTAT).

Kysyttiin, kerääkö yliopistot patenteja CRIS-järjestelmiin. Tutkijan täytyy olla itse aktiivinen, mikäli haluaa patenttinsa näkyviin. Aalto-yliopistolla ja Turun yliopistolla on olemassa tietokanta patenteille, mutta näitä ei ole hyödynnetty.

Erillislainsäädännön laatiminen on edennyt siten, että sitä on alettu kirjoittamaan. Tähän liittyen on hyvä ennakoida, mitä sen osalta on otettava huomioon. Esimerkiksi kielteisten päätösten välittäminen ja tiedon rikastaminen on hyvä ottaa huomioon erillislakia laadittaessa.

Yhteydet kansainvälisiin palveluihin

OpenAIREen liittyen kysyttiin, onko tarkoitus jatkossa viedä aineistot VIRTAan vai toimitetaanko tutkimusaineistot METAXin kautta. Ajatuksena on, että julkaisutiedot menevät edelleen VIRRAn kautta ja jatkossa aineistojen metatiedot METAXin kautta, ja kummatkin olisivat hyödynnettävissä tutkimustietovarannon kautta.

Kysyttiin, miten saataisiin yliopiston omasta Pure-järjestelmästä siirrettyä aineistot METAXiin. Viisi Pure-järjestelmää käyttävää yliopistoa on miettinyt asiaa yhdessä. Käytännössä aineistojen siirto ei onnistu muuten kuin siten, että METAX haravoisi tiedot tai rakentaisi rajapinnan, jonka kautta tiedot voitaisiin toimittaa. ***Sovittiin, että***

Anne Sunikka laittaa OKM:lle pyynnön haravoinnin tarpeesta. METAXin ja VIRRA:n osalta toimintamallit ovat hiukan erilaiset, eikä kannata tehdä toista aineistotietovarantoa VIRTAAan, koska METAX on jo olemassa.

Kansallisiin ja kansainvälisiin tutkimuspalveluihin liittyen kysyttiin, mikä kuvioon liittyen toimii jo tällä hetkellä. Tällä hetkellä mm. Tietoarkisto ja Kielipankki haravoidaan Käytännössä täytyisi selvittää tarkemmin, minkä lähteiden haravointi kannattaisi toteuttaa.

Esittelystä puuttuivat APC-maksut, mutta muuten listausta pidettiin hyvänä koontina. Ajatuksena on se, että TTV toimii välittäjänä kansallisiin ja kansainvälisiin palveluihin.

ORCIDin hyödyntämisen osalta toivottiin, ettei organisaation tarvitsisi olla konsortion jäsen. Tämä on hyvä ottaa huomioon neuvotellessa ORCIDin kanssa, jonka toiminta perustuu jäsenmaksuihin, joten tämä on otettava huomioon kumminkin päin. AMK:eille voi tämä voi olla kynnyksikysymys.

6. Julkisen hallinnon yhteentoimivuusvälineistön hyödyntäminen (Tommi Suominen, CSC)

Tommi Suominen esitteli julkishallinnon yhteentoimivuusvälineistöä ja sen käyttöä tutkimustietovarannossa liitteen 5 mukaisesti.

Tunnistettujen haasteiden osalta kysyttiin, mikä yhteentoimivuusvälineistön tämän suhde on Fintoon ja sen ontologioihin. Finto olisi hyvä viedä lähteeksi yhteentoimivuusvälineistöön. Yhteentoimivuusvälineistö on koko julkishallinnon yhteinen, mutta tällä hetkellä vain osa julkishallinnosta käyttää sitä.

Jyväskylän yliopistolla kehitetään sanastoja, tietomalleja ja koodistoja erillään, jolloin ne helposti erkaantuvat. Ehkä jatkossa niiden yhteinen käyttö yhdistyy ja vakiintuu. Toivottiin, ettei sanasto turhaan laajentuisi duplikaateilla, vaan yhdistettäisiin samojen termien eri versioita.

Kieliversioihin liittyen kysyttiin, miten ne toimivat yhteen. Käsitteet on määritelty kieliriippumattomiksi, mutta näillä on eri kieliversioita. Nyt on määritelty kielistä suomi, ruotsi, englanti, mutta käsitteistöä voidaan laajentaa myös muille kielille. Kieliversioita voi lisätä itse.

7. Yritysten saaman rahoituksen näyttäminen portaalissa (Hanna-Mari Puuska) [Liite 6](#)

Edellisen kokouksen keskustelua jatkettiin tutkimusorganisaatioiden tietojen siirron osalta. Tämän yhteydessä kerrottiin Business Finlandin näkemyksiä.

Mitä tietoja yrityksistä kerätään BF:n rahoituspäätöksissä?

Business Finlandilla yrityksiltä kerätään yritystä koskevaa ja hakemukseen liittyvää tietoja. Tämän lisäksi kerätään lisäksi vientilukuja ja verotustietoja. Kaikki tiedot eivät kuitenkaan ole julkisia. Tietokantatasolla tutkimustietovarantoon vietävissä rahoituspäätöksissä on yrityksiä osapuolina, mutta tätä tietoa ei nyt näy, vaikka tieto on olemassa. Business Finlandilta löytyy esimerkiksi hankkeita (mm. CO -innovation) ja

hakemuksia, joissa on mukana yrityksiä ja yrityksille tehdään omia rahoituspäätöksiä. Hankkeessa voi olla kolme osapuolta. Tälläkin hetkellä yritysten tutkimusrahoitus julkaistaan, mikäli se kuuluu ohjelmaan. Julkisia tietoja ovat yrityksen nimi ja rahoituksen määrä, mutta hankkeiden nimien tai tarkempien sisältöjen osalta on tehty linjaus, ettei niitä julkaista. Tämän osalta kaivattaisiin tietoa siitä, mistä olisi suurin lisäarvo tutkimustietovarannolle. Tällä hetkellä tutkimustietovarantoon tulevissa tiedoissa ei ole yritysten nimiä.

Business Finlandin tietoihin liittyen kiinnostavana tietona pidettiin erityisesti sitä, kuinka paljon yritykselle myönnetään rahaa. Toinen kiinnostava tieto on kuvaus projektista ja se, mikä yritys on toimijana. Nämä ovat tärkeitä tietoja, jos halutaan esitellä mitä kukakin osapuoli tutkii.

Kiinnostavaa olisi myös, minkä yritysten kanssa yliopistot tai eri yksiköt tekevät yhteistyötä. Tämä on tärkeä tieto myös Business Finlandin kannalta. Yrityksillä on intressi löytää yhteistyöhankkeita ja tutkijoita. Yleisesti ottaen verkostot olisi hyvä saada näkyviin.

Tilastokeskuksella on Business Finlandia pienemmän luvut eivätkä tilastot ole yksi yhteen. Tilanne on erilainen kuin muilla organisaatioilla ja tarpeelliset tiedot riippuvat siitä mitä halutaan tutkimustietovarannossa kuvata. Tutkimusorganisaatiot saavat itse kertoa osallistumisestaan, mutta BusinessFinlandilla ei ole aina valtuutta paljastaa nimiä.

EU-hanketietojen osalta on mahdollista saada tietoon myös summat. Muiltakin osin EU-data on hyvä esimerkki, millaista tietoa yrityksistä voitaisiin näyttää.

Edelläkävijäyrityksiä, jotka haluavat julkaista tutkimustaan, oletettavasti löytyy. Tutkimukset eivät ole aina ole salaisia ja suuret yritykset julkaisevat usein tietoja itsekin.

Business Finlandin palautteen mukaan tutkimustietovaranto on tärkeä projekti, jotta pystytään näyttämään suomalaista tutkimusosaamista ja voidaan houkuttaa osaamista tänne.

Tulisiko yrityksiltä kerätä myös muuta tietoa?

Kysyttiin, tekevätkö yritykset julkaisuja ilman yliopistoja? Oppilaitokset ovat yleensä mukana, mutta joitakin julkaisuja tehdään myös pelkästään yrityksissä. Sellaiset yritykset haluaisivat oletettavasti tutkimustaan esille. Ulkoiset organisaatiot eivät näy myöskään VIRTAn tiedonkeruissa.

Yhteisjulkaisuissa haasteena se, että ei tiedetä missä toimipaikassa tutkimusta on tehty. Yhtä lailla näissä voi olla ulkomaisia yliopistoja mukana. Käyttäjryhmien (yritysten) keskeinen tarve on kilpailijoiden seuranta. Kysymyksenä on myös se, millä tiedot löydetään. Julkaisutiedonkeruuseen pystyy liittymään yksittäinen organisaatio, mutta voi kuitenkin olla, ettei yritykset tee niin paljon julkaisuja, että tähän lähtisi. Merkittävydeltään rahoituspäätösten keruu olisi varmasti suurinta.

Yritysten kiinnostus julkaista tutkimustaan voi riippua tapauksista. Mikäli tutkimus on kriittistä omalle tutkimusalueelle, voi olla, että eivät halua julkaista, mutta joissain tapauksissa voi olla myös tarve tuoda tutkimustaan esille. On myös yrityksiä mukana

ilman julkista rahoitusta, esimerkiksi opiskelijaprojektit. Myös nämä olisi hyvä huomioida. Yritysten näkemyksiä voisi kartoittaa esimerkiksi niille suunnatulla kyselyllä.

Tulevaisuudessa tutkimustietovarannon portaalissa voisi tehdä nostoja yrityksen projekteihin ja yhteistyöhön liittyen.

Päätökset: Jatketaan hanketietovarannon ja Business Finlandin sinne toimittamien tietojen osalta nyt tehdyillä määrittelyillä ja julkaistaan tiedot, jotka ovat nytkin julkisia. Pyritään tuomaan esille yritysten mahdollisuuksia yrityksissä tehdyn tutkimuksen julkaisemiseen tutkimustietovarannon portaalissa. Arvioidaan myöhemmin, onko tarvetta näyttää portaalissa myös muita yritykseen liittyviä tietoja.

Jatketaan keskustelua Business Finlandin kanssa EU–datasta tehtävien esimerkkien avulla. Sen jälkeen käydään keskustelua tietojen toimittajien kanssa.

Pyritään luomaan keskusteluyhteyttä suoraan yrityksiin ja kysytään muutamilta eniten rahoitusta saavilta yrityksiltä näkemyksiä tietojen keräämiseen ja esittämiseen liittyen.

8. Muut asiat
9. Seuraava kokous 4.12.2019
 - Sovittiin, että myös ensi vuonna on kuusi kokousta. Laitetaan näistä kutsut lähiaikoina.
10. Kokouksen päättäminen
 - Puheenjohtaja päätti kokouksen klo 14.30