

Työryhmän kokous 6.9.2023

Aika: Ke 6.9.2023 klo 12:00-13:30

Etäyhteys: <https://cscfi.zoom.us/j/67872931510?pwd=Q1dtR3Bk3hCa01WTJvXVxVWx2b3BMZz09>

Kutsutut:

- Jukka Haapamäki, opetusneuvos, opetus- ja kulttuuriministeriö (puheenjohtaja)
- Viivi Aumanen, erityisasiantuntija, Jyväskylän yliopisto
- Laura Himanen, projektipäällikkö, CSC (sihteeri)
- Riikka Heikkinen, tietoasiantuntija, Tampereen yliopisto
- Timo Hellgren, asiantuntija, VTT
- Laura Mure, tietoasiantuntija, Aalto-yliopisto
- ~~Laura Niemi, kehittämisasiantuntija, Turun yliopisto~~
- Susanna Näreaho, erityisasiantuntija, Metropolia
- ~~Janne Pölönen, suunnittelupäällikkö, Tieteellisten seuran valtuuskunta~~
- Markku Roinila, tietoasiantuntija, Helsingin yliopisto
- ~~Eeva Savolainen, tietoasiantuntija, Aalto-yliopisto~~
- Tuomas Vienola, suunnittelija, Itä-Suomen yliopisto
- ~~Marjut Vuorinen, kehittämisspäällikkö, Terveystieteiden tutkimuskeskus~~
- Joonas Nikkanen, kehityspäällikkö, CSC
- Mia Salmi, tietoarkkitehti, CSC
- ~~Jonne Ritari, kehittämisasiantuntija, Turun yliopisto~~

Asiat:

1. Kokouksen avaus ja työjärjestyksen hyväksyminen

Puheenjohtaja avasi kokouksen klo 12.00. Hyväksyttiin työjärjestys.

2. Edellisen kokouksen muistio - [Liite 1](#)

Hyväksyttiin 9.5.2023 pidetyn kokouksen muistio.

3. Tilannekatsaus (Nikkanen) - [Liite 2](#)

Tutkijan tiedot -profiilityökalu lanseerattiin 25.5. ja tässä vaiheessa lanseeraus oli suunnattu erityisesti organisaatioille keskittyen viestintämateriaaleihin, työkalun käyttökokemuksiin ja tiedonsiirtoihin. Lanseeraus oli osa Tutkijan tiedot yhteisökokousta. Yhteisökokouksia järjestetään tulevana syksynä kolme, joista ensimmäinen on 20.9. jolloin aiheena on viestintä ja motivointi. Seuraava kokous on 1.11. aiheenaan teemahaun mahdollisuudet, ja vuoden viimeinen pidetään 13.12. jolloin käsitellään Luvittamista (rajapintakehitys).

Katsauksessa kesän aikana toteutettuun kehitykseen työryhmälle esiteltiin mm. organisaatietietojen käsittelyn parantamista, Tutkijat-osion hakutulospäivitykseen lisättyjä

filttereitä, sekä käyttäjien mahdollisuutta ilmaista ns. kiinnostuksen kohteet osaksi profiilia. Syksyn ja loppuvuoden kehityksestä nostettiin esiin mm. organisaatioiden kantarakenteen uudistaminen, alayksiköiden historiatieto ja fuusioituneiden organisaatioiden näyttäminen, ja hankehaavi-kehitys.

Ryhmässä nousi kysymys siitä, miten toimitaan sellaisessa tilanteessa, jossa tutkijan profiiliin julkaissut henkilö kuolee. Tähän ei toistaiseksi ole olemassa mitään prosessia, mutta kysymys on tiedostettu ja se koskee laajasti koko kenttää. Muun muassa ORCIDin suunnalta voisi tiedustella, miten siellä ollaan suunniteltu tämä hoitaa.

4. Keskusteluasiat (Nikkanen) - [Liite 3](#)

a. Tietojen poisto ja niiden hallinta

Lähtökohtana esitykselle on se, että teoreettisesti on mahdollista, että organisaatiolta poistetaan esimerkiksi henkilön irtisanouduttua tiedot omasta järjestelmästä ja täten myös tutkimustietovarannosta. Tämä toimisi kuitenkin vastoin OmaData -ajattelua, jonka mukaan Tutkijan tiedot -profiilityökaluun kirjautunut henkilö itse hallitsee siellä olevia tietoja, ja niiden näkyvyyttä palvelussa.

Kokouksessa tehtiin ehdotus tietojen poistosta, jonka mukaan sallitaan organisaatioiden tehdä poistoja lähettämilleen tiedoille, säilötään poistettavat tiedot tutkimustietovarannossa niin, että ne ovat tarvittaessa palautettavissa, sekä siirrytään käyttämään vain .xml muotoisia tiedostoja poistoille.

Työryhmä kävi ehdotuksesta keskustelua. Yksi haastava kysymys tietojen poistossa on se, että jotkut organisaatiosta tulevat tiedot ovat uniikkeja, eikä niitä siis ole esimerkiksi ORCIDissa ollenkaan. Jos tällaisia tietoja nyt poistetaan, profiiliin tehneelle henkilölle ei tule asiasta tietoa, ja tiedot vain häviävät profiilista. Yleisimmin kuitenkin poistetaan virheellisiä tietoja, jonka pitääkin olla organisaatioille mahdollista, onhan tietojen oikeellisuus organisaatioiden vastuulla. Työryhmä olikin sitä mieltä, että kevein perustein tietojen poistoa ei tehdä.

Poistettujen tietojen säilyttämisestä tutkimustietovarannossa keskusteltaessa todettiin, ettei tämä aiheuta haasteita tiedon tallentamisen suhteen, eli kyseessä oleva tietomäärä on suhteellisen pieni. Lisäksi tiedusteltiin, mikäli tutkija voisi itse palauttaa poistettuja tietoja profiilityökalussa. Tämä ei ainakaan toistaiseksi ole mahdollista, vaan nykyisellään tietoja voidaan palauttaa vain tapauskohtaisesti.

Päätös 4.1. Edetään kokouksessa esitellyn ehdotuksen mukaisesti: sallitaan organisaatioiden tehdä poistoja lähettämilleen tiedoille, säilötään poistettavat tiedot tutkimustietovarannossa niin, että ne ovat tarvittaessa palautettavissa, sekä siirrytään käyttämään vain .xml muotoisia tiedostoja poistoille.

b. Aktiviteettien deduplikointi

Nykyisin aktiviteettitietoja voidaan tuoda profiilityökaluun sekä kotiorganisaatiosta että ORCID-palvelusta, jolloin käyttäjälle saatetaan listata useita samoja aktiviteetteja, jotka ovat duplikaatteja keskenään. Koska aktiviteeteilla ei julkaisujen tavoin ole tiettyihin sääntöihin liittyviä de-duplikointimenetelmiä, niiden de-duplikoinnin mahdollisuudet ovat rajatut. Kokouksessa esiteltiin ehdotus, jonka mukaan ORCID-palvelusta siirrettävät tiedot (pl. Julkaisut ja profiilitiedot) menevät de-duplikoinnin kautta ja niitä verrataan kotiorganisaatioista siirtyviin tietoihin sillä oletuksella, että tiedot kotiorganisaatioista ovat aina parempilaatuisia/kattavampia. De-duplikointi perustuu siirrettävän tiedon nimeen/otsikkoon ja tarkastelusta karsitaan mahdollisesti pois ns. yleiset sanat ja merkit. Koska näin tarjottava tieto on kahdesta lähteestä tulleiden tietojen yhdistely, käyttäjälle ei tarjota mahdollisuutta valita kumman tiedon (ORCID vai kotiorganisaatio) ottaa profiiliin. Keskusteltiin myös siitä, että ORCIDista tulevan tiedon voisi jopa tiputtaa kokonaan pois duplikaattitilanteessa, koska se on tietoa, joka menee kotiorganisaatiosta ORCIDiin. Jos tieto on ainoastaan ORCIDissa, se on eri asia, koska silloin ei ole duplikaattia.

Päätös 4.2. Toteutetaan de-duplikointi osaksi profiilityökalua kokouksessa ehdotetun mukaisesti. ORCID poistetaan lähteenä, mikäli sieltä tuleva tieto on duplikaatti kotiorganisaation kanssa. Ratkotaan kokouksessa esiteltyjä avoimia kysymyksiä asian edetessä.

c. Asiantuntijalausunnot aktiviteettikoodistossa

Asiantuntijalausunnoista aktiviteettikoodistossa on keskusteltu työryhmässä 30.3. ja 9.5. On todettu, että nykyinen aktiviteettikoodisto ei tue kaikilta osin asiantuntijoiden/tutkijoiden asiantuntijalausuntoja. Työryhmälle on valmisteltu ehdotus lisäyksestä aktiviteettikoodistoon. Ehdotuksessa muutokset kohdistuvat luokkaan 4 – konsultointi siten, että nykyistä määritelmää laajennetaan koskemaan myös kutsuttuna tehtävää asiantuntijapalvelua julkiselle tai yksityiselle sektorille. Koodin nimi muutetaan muotoon 4 – konsultointi ja lausunnot. Ehdotuksessa ylätasoa alle lisätään tehtävä 4.2, mikä tarkoittaa sitä, että nämä 2. tason tehtävät tulevat myös määritellä, jotta ne voidaan erottaa toisistaan.

Päätös 4.3. Edetään kokouksessa ehdotetun mukaisesti: nykyistä luokan 4 määritelmää laajennetaan koskemaan myös kutsuttuna tehtävää asiantuntijapalvelua julkiselle tai yksityiselle sektorille. Koodin nimi muutetaan muotoon konsultointi ja lausunnot. Tietomallityökaluun tuotetaan ehdotuksen mukainen luonnos, ja mikäli 2. tason määritelmiä pitää tarkentaa, tuodaan ne seuraavaan työryhmän kokoukseen.

d. Aktiviteettikoodiston jatkotyöstö ja tarkennukset luokitteluihin

Aktiviteettikoodiston jatkotyöstöä ja luokitteluihin esitettyjä tarkennuksia esittelevä kalvo annettiin työryhmälle tiedoksi.

Päätös 4.4. Aktiviteettikoodiston jatkotyöstö ja tarkennukset luokitteluihin otetaan seuraavaan työryhmän kokoukseen lyhyesti.

5. Muut asiat

Markku Roinilan jäädessä vuorotteluvapaalle asiantuntija Tiina Käkelä osallistuu loppuvuoden kokouksiin.

6. Seuraavat kokoukset:

- a. [Työryhmän kokous 24.10.2023](#)
- b. [Työryhmän kokous 12.12.2023](#)

7. Kokouksen päättäminen 13.35