

Fairdata-verkosto - IDA-työpaja 24.3.2021

Aika: Keskiyö 24.3.2021 klo 13-15

Etäkokoushuone: <https://cscfi.zoom.us/j/64392772680?pwd=NUt5bzh6VW96bTdLbUhwK1J0eU9hQT09>

Nauhoite: <https://kannu.csc.fi/s/bFFxk6yTRSXAqA>

Mentimeter-kyselyn vastaukset: [IDA-työpaja_mentimeter_kyselyn_vastaukset.pdf](#)

Wikimerkintä: <https://wiki.eduuni.fi/x/RcOhCw>

Ennakkomateriaali

Alla ennakkomateriaaliksi PowerPoint-esitys, johon toivomme osallistujien tutustuvan ennen työpajaa. Esitys sisältää uuteen IDAn kapasiteetin hallintamalliin liittyviä ehdotuksia, joita käsittelemme työpajassa. Mikäli esityksen sisällöstä herää kysymyksiä tai kommentteja, pyydämme lähettämään niitä jo ennen työpajaa osoitteeseen fairdata@csc.fi.

CSC laatii työpajan jälkeen organisaatioiden kommenttien perusteella ehdotuksen IDAn uudeksi hallintamalliksi ja lähettää sen Fairdata-verkostolle. Hallintamalliehdotusta voi kommentoida ennen kuin CSC tekee asiasta päätöksen. Tiedotamme tästä kommentointimahdollisuudesta lisää myöhemmin.

[2021-03-24 IDA-työpajan ennakkomateriaali](#)

Asialista

IDAn tallennuskapasiteetin jako eri organisaatioiden kesken.

- OKM on viime vuosina vastannut organisaatioiden IDA-tila-anomuksiin ja erillismyöntöjä koskeviin hakemuksiin. Nyt IDA-tilojen hallintamallia tulee muuttaa, sillä OKM ei ole jatkossa mukana prosessissa.
- IDAn kapasiteetin jakoperusteisiin halutaan kommentteja IDAa käyttävien organisaatioiden edustajilta.

Työpajan kulku ja keskustelumuistio

- **Työpajan avaus (Kuisma Lehtonen, CSC)**
 - Ministeriön palveluohjaus on mutunut niin, että jatkossa ministeriö ei vastaa organisaatioiden IDA-tila-anomuksiin ja erillismyöntöjä koskeviin anomuksiin. OKM on pyytänyt CSC:tä vastaamaan näistä vastaisuudessa. Toivomme teiltä keskustelua ja kommentteja tästä aiheesta.
- **Ennakkomateriaalien läpikäyminen (Suvi Pousi, CSC): [2021-03-24_IDA-työpaja_kapasiteetin_hallintamalli_tyopajaversio.pptx](#)**
 - **Kommentteja IDAn kapasiteetin hallintamalliin**
 - Miksi organisaatiokohtaista allokointia tehdään etukäteen? Onko allokointi päällekkäistä hallintaa projektin tilan hallinnalle? Voisiko olla niin, että käyttöoikeuksia haetaan ja tilan kulutusta seurataan ja jos alkaa tapahtumaan jotakin ylläpitäjän näkökulmasta hämmäntävää, niin sitten pohditaan mitä tehdään? Voisiko siis olla yksi "pool", josta käyttäjät hakevat tilaa projekteille suoraan?
 - CSC: IDAssa organisaatio on myöntänyt tilan projekteille CSC:n sijaan ja organisaatiolla on myös oikeus poistaa projekti ja sen data palvelusta (<https://www.fairdata.fi/ida-tutkimusdatan-sailytyspalvelun-kayttoehdot/>). Huomioimme kuitenkin kommentin ja pohdimme voiko asiaa suoraviivaistaa.
 - **Kommentteja organisaation ja IDA-yhteyshenkilöiden rooliin palvelussa sekä IDAa käyttävien projektien hallintaa**
 - Onko mietitty organisaatioiden ja yhteyshenkilöiden roolia? Voisiko CSC myöntää oikeudet projekteille suoraan?
 - CSC: Haluaako organisaation datatuki olla miettimässä, mikä on sopiva palvelu IDA-tilaa hakevalle projektille?
 - Ammattikorkeakoulu puolella yhteisiä ohjeita on tuloillaan, jolloin tarve tähän voi vähentyä. Ollaan menossa siihen suuntaan, että oletetaan datan tallentajien tekemän aineistonhallintasuunnitelman jo projektin alkuvaiheessa. Täten yhteyshenkilöiden rooli voi olla pienentymässä.
 - Orpodatojen suhteen toivotaan nykyistä parempaa politiikkaa. Nykyisillä käytännöillä arkistodatat menevät muualle, jos projekti päättyy.
 - IDAn käyttötilastointiin on oltu tyytyväisiä, sitä hyödynnetään jatkossakin mielellään.
 - Lisääkö CSC usein projektien tila-allokointeja IDAssa? Jos projekti on saanut tilan ja pyytää lisää kun se on täynnä? Kyllä, näitä tulee joskus mutta verrattain harvoin ja ne menevät CSC:n asiakaspalvelun kautta organisaation IDA-yhteyshenkilön hyväksyminä.
 - **Kommentteja tallennustilojen kokojen tarkistamiseen**
 - Kommentoitu, että olisi hyödyllistä, jos anottua tilaa voisi myös pienentää dynaamisesti.
 - Kommentoitu, että automatisoitu tallennustilan skaalausprosessi olisi hyvä ja tästä politiikasta selkeä tiedotus projektien omistajille. Manuaalinen projektien läpikäynti on hyvin työlästä.
 - Ehdotettu tilan pienentämistä suoraan jos esim. kahteen vuoteen ei ole käytetty tilaa. CSC voisi hoitaa tämän.
 - **Tulevien lakimuutosten mahdollinen vaikutus IDAn käyttöön**
 - Tiedoksi: Avoimen datan direktiivin täytäntöönpanoa koskevaan lausuntopyyntöön tulleet lausunnot nähtävillä: <https://www.lausuntopalvelu.fi/FI/Proposal/Participation?proposalId=209d1bca-2280-4e87-ab0f-ef045ff40cd4>
 - **Aikataulusta**
 - CSC:n koostama hallintamalliehdotus tulee jakoon vasta pääsiäisen jälkeen, kommentointi menee siis huhtikuun puolelle.
- **Kysymyksiä ja vastauksia mentimeterin avustuksella**
 - Menti-kyselyn vastaukset: [IDA-työpaja_mentimeter_kyselyn_vastaukset.pdf](#)
 - **Mentistä herännyttä keskustelua**

- Noussut esiin sensitiivisen datan tallennukselle tarve
 - Organisaatioista tarkennettu, että sensitiivistä tässä tapauksessa on esimerkiksi genomidata, tai vaikka koulualan kyselytutkimukset tai haastatteluvideot, joissa on tunnisteita. Ne tallentaa johonkin, ja esim. genomidataa voi myös jakaa kontrolloidusti. Lisäksi esim. haastatteluaineistot, joita halutaan tarkastella ennen anonymisointia. Sensitiivistä myös esim. taidealan esitysten videoiteja, joissa esiintyy yleisöä jos lupia ei ole syystä tai toisesta pystytty kysymään.
 - CSC: Sensitiiviselle genomidatalle tulossa Federated EGA (EGA: <https://ega-archive.org/>) eli ratkaisu, jossa data itsessään jää suomeen CSC:n turvallisille palvelimille, mutta metatieto menee EGAn järjestelmään.
 - CSC: Muulle kuin genomidatalle tulossa SD-palvelukokonaisuus (SD=sensitive data), joka mahdollistaa arkaluonteisen datan käsittelemisen tietoturvalisessä ympäristössä. Palvelua käytetään selaimen kautta, eli suhteellisen helppokäyttöinen. Palveluun pääsee kaksivaiheisen tunnistautumisen kautta ja ympäristöön submittoitua dataa ei voi esim. ladata omalle koneelle. Tarkka käyttöpolitiikka vielä työn alla, eli esim. kuinka pitkään dataa voi säilyttää palvelussa. Pyrimme järjestämään palveluista webinaarin ennen kesää. Niitä on jo esitelty FUCION ja AAPAn jäsenille, lisäksi olemme menossa Arenen TKI-valiokunnan kokoukseen huhtikuun puolivälissä, mutta "suurelle yleisölle" viestintä tosiaan vasta tulossa. Ottakaa yhteyttä heidi.laine@csc.fi niin voin kertoa lisätietoja.
 - Tarkoittavatko "arkaluonteiset" henkilötiedot tietosuojalain mielessä erityisten henkilötietoryhmien käsittelyä? Ts. ne joiden käsittely on lähtökohtaisesti kiellettyä? SD-palveluiden kohdalla ei tulla ottamaan kantaa tarkalleen tähän asiaan tai luokittelemaan dataa. Datan omistaja on lopulta vastuussa datasta ja sen tarpeista.
 - IDAn käyttöehdot: "Käyttäjä sitoutuu olemaan tallentamatta palveluun tietoja, jotka koskevat EU:n yleisen tietosuoja-asetuksen 9 artiklassa lueteltuja "erityisiä henkilötietoryhmiä" (<https://www.privacy-regulation.eu/en/9.htm>). Lähde: <https://www.fairdata.fi/ida-tutkimusdatan-sailytyspalvelun-kayttoehdot/>
- Kommentoitu, että onnistumistarinat voisivat helpottaa IDAn käyttöä datan avaamisessa organisaatioissa.
 - CSC: Yksi esimerkki on Metsähovin case: <https://www.fairdata.fi/avoimuuden-edistajat-palkinto-fairdata-palveluja-hyodyntavalle-metsahovin-radiotutkimusasemalle/>
- Tiedoksi avoimen datan asiantuntijaryhmän omistajuus ja lisensointi -alatyöryhmän kalvosetti: Aineiston avaamisen esteet - Omistajuus ja lisensointi: <https://doi.org/10.5281/zenodo.2708815>
- Pitkässä juoksussa käyttämätön allokoitu tila hankaloittaa IDAn kapasiteetin riittävyyden arviointia.
- Tietoa erillismyönnoistä
 - <https://www.fairdata.fi/wp-content/uploads/2021/01/IDA-erillismyönnot-2012-.pdf>
 - <https://www.fairdata.fi/ida/idan-kayttajaksi/> (sivun lopussa tietoa erillismyönnoistä)
- Pikaohjeet säilytystilan itsenäiseen hakemiseen: <https://www.fairdata.fi/ida/hae-sailytystilaa/#pikaohje>