

Fairdata-verkosto - kokous 5.10.2023

Aika: Torstai 5.10.2023 klo 14-16

Etäkokoussuone: <https://cscfi.zoom.us/j/66393561333?pwd=dUcrzRuTTNmRW9XaEFCNjVGSXhZz09>

Nauhoite: https://video.csc.fi/media/t/0_mya2pjvl

Asialista

1. Kokouksen avaus
2. Katsaus tutkimuksen datan hallinnan viitearkkitehtuurityön etenemiseen, **Eeva Nyrövaara, CSC**
3. Tilannekatsaus: Tutkimusdataan liittyvien organisaatiroolien kehittämisestä CSC-palveluissa, **Heidi Laine, CSC**
4. Katsaus IDAn toimintoihin ja prosesseihin, **Suvi Pousi, CSC**
5. Fairdatassa julkaistujen aineistojen katselu- ja latausmäärät, **Erja Kortelainen, CSC**
6. Fairdata-kehityksen syksyn 2023 painopisteet, **Erja Kortelainen & Juha Lehtonen, CSC**
7. Ajankohtaiset koulutukset ja tapahtumat, **Erja Kortelainen, CSC**
 - a. Näkyvyyttä ja vaikuttavuutta tutkimukselle: datajulkaisut - käytännön kokemuksia tilaisuus 1.11.2023 9:30-12, [ilmoittaudu mukaan!](#)
 - 22.11.2023 klo 14:00-15:30 jatkotapahtuma datajulkaisuista yhteistyössä [eLTER:n](#) kanssa, jossa pidempi alustus mistä datajulkaisuissa on kyse sekä tutkijapuheenvuoro. Ilmoittautuminen tulossa.
 - b. Laatua ja vaikuttavuutta - Miten teet datastasi FAIR -webinaari 2.11.2023 klo 10:00-11:00 [Ilmoittautuminen on auki!](#)
8. Muut asiat
 - a. PAS-palveluiden tuottama tiedostojen validointipalvelu on julkaistu, [ks. uutinen](#).
9. Seuraava kokous: 7.12.2023 klo 14:00-16:00

Muistio

1. Kokouksen avaus
2. Katsaus tutkimuksen datan hallinnan viitearkkitehtuurityön etenemiseen, **Eeva Nyrövaara, CSC (kalvot)**
 - a. **Yhteenvetoa:**
 - i. Periaatteellinen taso edennyt: arkkitehtuuriperiaatteet tehty, lait ja asetukset, strategiat tunnistettu. Nykytilan analyysia tehty ja tähän on pyydetty kommentteja.
 - ii. Liiketoiminta-arkkitehtuuri: Tämän sisältöjä kerätty mm. tutkimuspalvelupäivillä, tuotoksia työstetty arkkitehtuuriryhmässä.
 - iii. Työ edennyt aikataulussa, nyt ollaan n. puolivälissä.
 - iv. Koitetaan tunnistaa tutkimuksen datanhallinnan vaatimuksia. Keskustellaan, mitkä ovat olennaisia strategioita (mitä halutaan seurata, mihin suuntaan osoittavat). Halutaan asettaa tavoitteet selvemmin.
 - v. Viitearkkitehtuurin nykytilan analyysi:
 1. Analyysissa tarkoituksella korostettu haasteita/puutteita, jotta se herättelisi
 2. Yhtenä huomiona analyysissa: tutkijalle palvelut ei näyttäyty saumattomana tutkimuksen koko kulun ajan
 3. Analyysi keräsi runsaasti kommentteja, mm. yhteisen tavoitetilan hyödyistä ja tarvittavien palvelujen painopisteistä
 - b. **Linkkejä:**
 - i. Yhteistyöfoorumi: <https://wiki.eduuni.fi/x/24KwAQ>
 - ii. Tutkimuksen datanhallinnan viitearkkitehtuurityö: <https://wiki.eduuni.fi/x/3gnzEw>
 - iii. Viitearkkitehtuurin dokumentaatio: <https://wiki.eduuni.fi/x/gg0oF>
 - iv. Nykytilan analyysi, joka oli kommentteilla elo-syyskuussa: <https://wiki.eduuni.fi/x/3TTwFg>
 - v. Saadut kommentit: <https://wiki.eduuni.fi/x/eCyZFw>
3. Tilannekatsaus: Tutkimusdataan liittyvien organisaatiroolien kehittämisestä CSC-palveluissa, **Heidi Laine, CSC (kalvot)**
 - a. **Yhteenvetoa:**
 - i. CSC:llä tehty linjauksia organisaation roolista datanhallinnan ja laskennan palveluissa. Organisaatirooli tunnistetaan ja halutaan mahdollistaa organisaation oikeudet ja vastuut dataan liittyen.
 1. Esimerkiksi käyttäjien sisältöjen hallinnan siirtoa loppukäyttäjältä organisaatiolle ei pakoteta, mutta tämä halutaan mahdollistaa
 2. Läpinäkyvyyttä palvelukäyttöön: Organisaation raporttina MyCSC:ssä tarjolla organisaatoiden edustajille (lisätietoja ks. toukokuun Fairdata-verkoston kokous, agendan kohta 3.)
 3. Aloitettu projektina keväällä 2022, jonka kesto on vuoden 2023 loppuun
 - ii. **Mitä seuraavaksi:**
 1. Linjaus tullaan implementoimaan palveluiden poliittikoihin ja prosesseihin
 2. Yksi konkreettinen mahdollisuus organisaation tuomiseksi mukaan palvelukäyttöön: organisaatirooli mukana CSC:n datanhallinnan ja laskennan palveluja käyttävässä CSC-projektissa varavastuuhenkilönä. Organisaatio voisi näin säilyttää ja hallita palveluissa olevaa dataa vaikka alkuperäinen käyttäjä (tutkija) poistuisi
 3. Tietoturvaan ja tietosuojaan liittyvän palveludokumentaation yhdenmukaisuutta pyritään edistämään
 4. Projektista eteenpäin jatkuvaksi toiminnaksi
4. Katsaus IDAn toimintoihin ja prosesseihin, **Suvi Pousi, CSC (kalvot)**
 - a. Uutta: IDA-yhteyshenkilön mahdollisuus muokata käyttäjien CSC-projektien IDA-säilytystilojen kokoa MyCSC-portaalissa
 - i. **Yhteenvetoa:**
 1. Organisaatiot eivät aiemmin ole voineet vapauttaa MyCSC:ssä CSC-projektien hakemaa käyttämätöntä IDA-säilytystilaa, vaan vapauttamista on pyydetty sähköpostitse CSC:lta. Nyt uutena toimintona organisaation IDA-yhteyshenkilö voi itse vapauttaa projektien käyttämätöntä säilytystilaa muille tarvitsijoille MyCSC-palvelussa.
 - a. MyCSC listaa CSC-projekteittain, paljonko dataa on IDAssa, ja miten paljon sitä on julkaistu, sekä antaa linkit julkaistuihin aineistoihin
 2. Organisaatiokohtaista säilytystilaa voi hakea lisää ottamalla yhteyttä CSC:hen.
 3. **Kysymys:** "Olisiko mahdollista saada näkyviin, milloin käyttäjät ovat viimeksi IDAssa muuttaneet dataa?"
 - a. **Vastaus:** On mahdollista näyttää. Otamme mukaan kehityslistalle

4. **Kommentti:** "Tälle valtavan iso peukutus! Meillä vapautui melkein 30% allokoidusta IDA-kapasiteetista tämän ansiosta."
 - b. Uutta: IDAn komentorivityökalujen parannukset ja tutoriaalivideot
 - i. **Yhteenvetoa:**
 1. IDAan voi viedä dataa kahdella eri tavalla: selainkäyttöisen käyttöliittymän kautta sekä IDAn komentorivityökaluilla
 - a. Komentorivityökalut ladattavissa käyttäjän omalle koneelle. Ne ovat myös valmiiksi saatavilla CSC:n Mahti ja Puhti -supertietokoneilla.
 2. Komentorivityökaluihin tullut uusia toimintoja, joiden avulla käyttäjän on helpompi seurata datansiirron onnistumista ja jatkaa keskeytyneitä siirtoja
 3. Työkalujen käytön avuksi on julkaistu tutoriaalivideoita. Myös IDAn komentorivityökalujen käyttöä voi harjoitella Fairdatan Demo-ympäristössä, johon saa pääsyn pyytämällä (fairdata@csc.fi)
 4. Uutinen aiheesta: <https://www.fairdata.fi/uutta-idan-komentorivityokaluissa/>
 - c. Datan poiston käytännöistä CSC-projektin vastuuhenkilön puuttuessa
 - i. **Yhteenvetoa:**
 1. Nykyinen linjaus CSC:n datanhallinnan palveluissa: Projektin vastuuhenkilön puuttuessa CSC-projekti sulkeutuu, projektiin liittyvä data poistetaan 90 päivän kuluttua CSC-projektin sulkemisesta
 - a. Ilmoitamme organisaation IDA-yhteyshenkilöille aina ennen datan poistamista, jotta uusi vastuuhenkilö voidaan tarvittaessa nimetä
 2. Seuraavaksi: CSC-projektin hallinnan ja palvelun mallin kehitys siten, että erityisesti julkaistuja dataa ei poistettaisi. Mallin kehitystyön ollessa vielä kesken toistaiseksi julkaistuja dataa ei poisteta.
5. Fairdatassa julkaistujen aineistojen katselu- ja latausmäärät, **Erja Kortelainen, CSC (kalvot)**
- a. **Yhteenvetoa:**
 - i. Etsimessä kerätään tietoa aineistojen katselu- ja latausmääristä. Tieto tullaan lisäämään Etsimeen aineiston kuvailusivulle todennäköisesti 2024 vuoden alussa.
 - ii. Tietoa aineistojen katselu- ja latausmäärien keräämisestä <https://www.fairdata.fi/aineistojen-katselu-ja-latausmaarat/>
 - iii. Jos organisaatioilla on tarvetta saada määrämuotoista tietoa katselu- ja latausmääristä (COUNTER SUSHI API), pyydämme olemaan yhteydessä fairdata@csc.fi (ks. <https://www.projectcounter.org/code-of-practice-rd-sections/foreword/>)
 1. **Kysymys latausmääremetriikoista:** onko vuositason raporteilla kuluvan kalenterivuoden vai viimeisen 12 kuukauden lataus/katselumäärät? Ohjeesta ja raporteilta tämä ei käy ilmi.
 - a. **Vastaus:** Vuositason raportti on kuluvan vuoden alusta (year to date). Muokkamme raportin otsikointia, jotta tämä on selkeämpi.
6. Fairdata-kehityksen syksyn 2023 painopisteet, **Erja Kortelainen & Juha Lehtonen, CSC** (ks. kalvot kohdasta 5)
- a. Etsin - UI-uudistus
 - i. **Yhteenvetoa:**
 1. Ulkoasullinen uudistus tehty. Samalla lisättiin datan tyyppi näkyviin aineiston etusivulle.
 2. Tulossa: Toimijoiden ja asiasanojen kautta hakeminen (hae aineistoja, joissa sama toimija / asiasana)
 - b. Metax -rajapintauudistus
 - i. **Yhteenvetoa:**
 1. Käynnissä parhaillaan. Tuotantoon ennen kesää 2024. API-rajapintaan tulee joitain kenttämuutoksia sekä jotkut API-yhteyspisteet saattavat muuttua. Tuki vanhan rajapinnan käytölle kuitenkin jatkuu edelleen, kunnes toisin ilmoitetaan. CSC on yhteydessä API-rajapinnan tiedossa oleviin käyttäjiin.
 - c. Qvain -pääkäyttäjäröolin suunnittelu
 - i. **Yhteenvetoa:**
 1. Sisäänkirjautuneen käyttäjän julkaistujen aineistojen listaus tulossa osaksi Etsimen UI:ta (ollut ennen Qvaimessa omana aineistolistanan). Muutoksen avulla tuleva Qvaimen organisaatiokohtainen pääkäyttäjä voi helposti hakea ja selailla niitä lukuisia aineistoja, joihin hänellä on editointioikeus.
 2. Pääkäyttäjän toimintoja ja näkymiä testattu käyttäjillä. Testitulosten analyysin perusteella suunnitellaan muutokset jo suunniteltuihin käyttöliittymiin.
 - ii. **Kysymys:** Onko haussa organisaatio=julkaisija-tieto?
 1. **Vastaus:**
 - a. Julkaisija on osa aineiston metadattaa (toimijatieto). Etsimen hakutoiminnossa organisaatio -filterin kautta filterointi tapahtuu timenomaan toimijatietoon sidotun organisaation perusteella.
 - b. Pääkäyttäjä-organisaatio määritellään kuitenkin ihan erikseen (oletuksena se organisaatio, johon kuvailun alkuperäinen tekijä affilioituu. Tätä voi kuitenkin vaihtaa.
 - d. IDA - organisaatirooli CSC-projekteissa ja julkaistun datan hallinnassa
 - i. **Kysymys:** Ooko tätä pääkäyttäjäresurssia IDA/Qvain yms. ylläpitoon olemassa organisaatioissa riittävästi?
 1. **Kommentteja osallistujilta:**
 - a. Merkittävästi helpompaa auttaa kuratoimaan näitä aineistoja, jotka ovat organisaation hallinnassa
 - b. Organisaatiossa resursointi ok. Tilanne voidaan ylläpitää, jos kysyntä kasvaa hallitusti ja hyödyt pysyvät ilmeisinä (eli peruste resursoinnille on selvä).
 - e. Fairdata PAS-palvelu
 - i. Hallintaliittymän UI-uudistus
 - ii. Hallintaliittymän autorisoinnin uudistaminen
 - iii. Aineiston noutotoiminto hallintaliittymään
 - iv. Paketointipalvelun laaja uudistaminen vastaten käyttöönotoissa esiin tulleisiin tarpeisiin
 - v. **Yhteenvetoa:**
 1. Sekä Fairdata PAS-palvelun hallintaliittymän ulkoasua että säilytysviennin prosessia uudistettu. Uudessa prosessissa ensin tunnistetaan että aineisto on PAS-kelpoisessa kunnossa, sitten ehdotetaan PAS:iin vietäväksi, validoidaan ja hyväksytään.
 2. Autorisointi ollut aiemmin organisaatiokohtaista eli organisaatiolla ollut vain yksi taho, jolla on ollut oikeuksia PAS-aineistoihin. Tästä eteenpäin pitkäaikaissäilytettävien aineistojen autorisointi on säilytyspäättökohtaista. Tällä tarkemmalla autorisoinnilla on myös mahdollista mahdollistaa aineistojen nouto PAS-palvelusta tietyille roolille.
 3. PAS-palvelusta noutamisen prosessi on kaksivaiheinen: ensin autorisoitu käyttäjä pyytää hallintaliittymässä PAS-palvelua koostamaan jakelupaketin aineistosta taustatoimintona. Kun jakelupaketti on valmis, on autorisoidulle käyttäjälle tarjolla latauspaketti. Huom: aineistojen nouto PAS-palvelusta ei tarkoita että pitkäaikaissäilytys päättyisi. Se on vain kopion noutamista säilytyksessä olevasta aineistosta.
 4. Paketointipalvelun uudistaminen ei näy hallintaliittymän käyttäjille suoraan, mutta näkyy sellaisille organisaatioille, jotka eivät käytä PAS:n hallintaliittymää vaan paketoivat aineistonsa itse.
7. Ajankohtaiset koulutukset ja tapahtumat, **Erja Kortelainen, CSC**

- a. Näkyvyyttä ja vaikuttavuutta tutkimukselle: datajulkaisut - käytännön kokemuksia tilaisuus 1.11.2023 9:30-12, [ilmoittaudu mukaan!](#)
 - 22.11.2023 klo 14:00-15:30 jatkotapahtuma datajulkaisuista yhteistyössä [eLTER:n](#) kanssa, jossa pidempi alustus mistä datajulkaisuissa on kyse sekä tutkijapuheenvuoro. Ilmoittautuminen tulossa.
 - b. Laatua ja vaikuttavuutta - Miten teet datastasi FAIR -webinaari 2.11.2023 klo 10:00-11:00 [Ilmoittautuminen on auki!](#)
8. Muut asiat
- a. PAS-palveluiden tuottama tiedostojen validointipalvelu on julkaistu, [ks. uutinen](#).
9. Seuraava kokous: 7.12.2023 klo 14:00-16:00

Kaikki kommentit, kehitysehdotukset, puheenaiheet ym. otetaan mielellään vastaan! Kysymyksiä ja ehdotuksia voi lähettää osoitteeseen fairdata@csc.fi!