

Kansalliskirjaston FAIR-työ: periaatteista kohti käytäntöjä

Liisa Näpärä

Kansalliskirjasto, Helsingin yliopisto

liisa.napara@helsinki.fi

<https://orcid.org/0000-0002-6473-5896>

Matias Frosterus

Kansalliskirjasto, Helsingin yliopisto

matias.frosterus@helsinki.fi

<https://orcid.org/0000-0002-8355-0256>

Asiasanat: kansalliskirjastot, palvelut, data

Viittaaminen: Näpärä, L., & Frosterus, M. (2022). Kansalliskirjaston FAIR-työ: periaatteista kohti käytäntöjä. *Informaatiotutkimus*, 41(2–3), 120–123. <https://doi.org/10.23978/inf.122649>

Artikkeli on lisensoitu Creative Commons Nimeä-EiKaupallinen-JaaSamoin 4.0 Kansainvälinen -lisenssillä

Kansalliskirjaston erilaiset datat, järjestelmät, palvelut ja yksiköt muodostavat mutkikkaan ja toisistaan riippuvaisen prosessin, jossa kokonaisuuden yhteentoimivuus on avainasemassa. Tässä paperissa kerrotaan Kansalliskirjaston FAIR-työstä vuosien 2021–2022 aikana.

FAIR-periaatteet muodostuvat neljästä dataan liitettävästä ominaisuudesta, joita ovat löydettävä (findable), saavutettava (accessible), yhteentoimiva (interoperable) ja uudelleenkäytettävä (re-usable). Ne on luotu ensisijaisesti tutkimusdatan jakamiseen. (David ym., 2020.) Perusperiaatteena on tarjota dataa niin avoimesti kuin mahdollista, mutta niin suljetusti kuin on välttämätöntä. Myös Kansalliskirjaston strategian 2021–2030 mukaan ”Kirjaston tuottamaa dataa hallitaan FAIR-periaatteiden mukaisesti ja ohjeistamme käyttäjiä datan vastuulliseen käyttöön.”

Ymmärtääkseen paremmin FAIR-periaatteiden soveltamista Kansalliskirjastossa toteutettiin vuonna 2021 yksikkökohtaiset (n=16) haastattelut siitä, miten kukin yksikkö kokee FAIR-periaatteiden toteutumisen omasta näkökulmastaan. Yksiköistä mukana olevien henkilöiden määrä vaihteli, mutta useimmiten haastatteluissa oli mukana 5–15 henkilöä. Kukin haastattelu käytiin FAIR-projektiryhmään kuuluvien henkilöiden johdolla. Tuloksena oli yksiköiden välillä erilaista ymmärrystä kustakin FAIR-periaatteesta ja osa koki, etteivät ne koske niin paljon sellaista työtä, jossa ei toimita digitaalisten aineistojen ja datan kanssa. Haasteeksi todettiin FAIR-periaatteiden abstraktius ja järjestelmien suuret riippuvuudet toisistaan ja liian vähäiset resurssit edistää FAIR-periaatteita konkreettisesti. Tulosten ja muun arvioinnin perusteella vuodelle 2022 valittiin neljä pilottiyksikköä, jotka pohtivat, miten FAIR-periaatteita voisi edistää laajemmin koko Kansalliskirjaston tasolla konkretisoimalla niitä. Vuoden 2022 aikana on esimerkiksi pidetty pilottiyksiköiden kesken työpaja, jossa on tutkailtu FAIR-periaatteiden näkymistä arkityössä, pohdittu, mitä on saatu viime aikoina aikaan ja millaista ohjeistusta tarvitaan vielä lisää ja kenelle niitä pitäisi osoittaa.

Tähän mennessä tehty FAIR-työ on vasta alkanut kirkastaa FAIR-periaatteiden merkitystä erilaisille datoille ja järjestelmille. Toisaalta edelleen on haaste linkittää erilaiset datojen ja järjestelmien kehitys johonkin yksittäiseen FAIR-periaatteeseen, eikä FAIR-periaatteiden mukainen ajattelu näy vielä arkityössä, vaikka niihin liittyvää työtä aktiivisesti tehdäänkin eri yksiköiden arjessa. Esimerkiksi periaatteiden edellyttämien koneluettavan metatiedon tai pysyvien tunnisteiden käytön lisääminen tuottavat kysymyksiä siitä, miten ne voidaan konkreettisesti toteuttaa kaikille Kansalliskirjaston erilaisille datoille ja yhteentoimivasti eri järjestelmiin eri palveluiden toiminnasta riippuvaisessa organisaatiokokonaisuudessa.

Yksittäisten esimerkkien esiintuominen FAIR-työn kautta kuitenkin auttaa sisäisten datojen ja järjestelmien välisessä kehittämisessä. Yksityiskohdat konkretisoivat abstrakteja FAIR-periaatteita, ja lopulta olisi tärkeää ymmärtää, miten jokin toiminto edistää datan löydettävyyttä, saavutettavuutta, yhteentoimivuutta ja uudelleenkäytettävyyttä. Riippuvuudet toisista yksiköistä ja toisten ihmisten työstä eri palveluissa ovat suhteellisen suuria, joten tiedonvaihdon tarpeesta yksiköiden ja organisaatioiden yli on puhuttu paljon FAIR-kokonaisuuden edistämiseksi. Laajemmassa kuvassa on myös ymmärrettävä, että pienet yksityiskohtaiset kehitysasiat ovat osa isompaa kokonaisuutta ja FAIR-työtä, jota Kansalliskirjasto tekee eri tasoilla ja pitkällä aikataulilla.

Kansalliskirjastossa tehdään monenlaista erilaista perustyötä, joilla edistetään aineistojen muuttamista dataksi, jolloin ne voidaan paremmin ymmärtää FAIR-periaatteiden sovellutuksen kohteena. Esimerkiksi Kansalliskirjaston tehtävänä oleva digitointi ja kuvailutyö edistävät aineiston FAIR-periaatteiden toteuttamista muuttamalla analogista dataksi ja sitä kautta myös digitaalisesti löydettäväksi. Saavutettavuutta puolestaan edistetään lisensoimalla tutkimuskäyttöön enemmän dataa ja yhteentoimivuutta kehittämällä eri järjestelmiä, kuten julkaisuarkistoja, Finnaa ja digi.kansalliskirjasto.fi-palvelua. Uudelleenkäyttöä syntyy välillisesti datan määrän kasvattamisen myötä, käytettävyydeltään toimivista palveluista ja viestinnästä niistä. Tällöin tutkimusdatan luominen Kansalliskirjaston datasta on itse asiassa jo datan uudelleenkäyttöä.

Kansalliskirjaston data onkin useimmiten sellaista dataa, jota tarjotaan tutkimukselle digitoinneista tai metadataa kansallisbibliografiasta – ei niinkään tutkimuksessa syntynyttä dataa, johon FAIR-periaatteet on alun perin tarkoitettu. Toisaalta myös tutkimuksessa syntynyt data on kiinnostava Kansalliskirjaston näkökulmasta, koska sillä voidaan joissakin tapauksissa rikastaa olemassa olevia datoja ja edistää niiden käytettävyyttä esimerkiksi löydettävyyden osalta. On kuitenkin huomioitava, että läheskään kaikki Kansalliskirjaston aineisto ei ole digitaalista tai tekijänoikeudesta vapaata, eikä näin ollen verkossa avoimesti saatavilla ja näiden osalta voidaan tehdä FAIR-työtä annettujen reunaehtojen, kuten lainsäädännön ja resurssien puitteissa. Näin voidaan löytää keinoja tuoda dataa niin avoimeksi kuin mahdollista.

Kansalliskirjaston keskeneräisen FAIR-työn on tarkoitus kasvattaa ymmärrystä sekä periaatteista itsestään että niiden sisällön soveltamisessa, joskin yksiköiden tuottamien datojen ja järjestelmien verkosto on varsin monimutkainen. FAIR-työn lopullisena tavoitteena on luoda konkreettinen suunnitelma FAIR-periaatteiden soveltamisesta Kansalliskirjaston datan hallintaan, johon kokonaisuudessaan kuuluu niin järjestelmät, palvelut kuin

yksikötkin. Työn on tarkoitus jatkua edelleen myös vuoden 2022 jälkeen osana projektisuunnittelua, ihmisten arkityötä ja hyvien käytäntöjen jakamista yhteistyössä muiden FAIR-periaatteiden edistämään sitoutuneiden organisaatioiden kanssa. Toistaiseksi työssä on vielä enemmän kysymyksiä kuin vastauksia ja vuoropuhelua tarvitaan sekä Kansalliskirjaston sisäisesti että ulkoisten kumppanien kanssa FAIR-periaatteiden konkretisoimiseksi ihmisten arkityössä.

Lähteet

David, R., Mabile, L., Specht, A., Stryeck, S., Thomsen, M., Yahia, M., . . . Alliance – SHaring Reward and Credit (SHARC) Interest Group, T. R. D. (2020). FAIRness Literacy: The Achilles' Heel of Applying FAIR Principles. *Data Science Journal*, 19(1), 32. <http://doi.org/10.5334/dsj-2020-032>

Kansalliskirjaston strategia 2021–2030 (2020). <https://urn.fi/URN:NBN:fi-fe2020040210186>