

# Digitaaliset infrastruktuurit tutkimuksen palveluksessa

POSTED ON [24.11.2016](#) BY [SAMU KYTÖJOKI](#)



## VERKKARI HELSINGIN YLIOPISTON KIRJASTO

Krakovassa järjestettiin 28.-30.9.2016 *Digital Infrastructures for Research 2016* -kokous. Kokouksen tarkoituksena oli kehittää tutkimusdatan ja muun tutkimuksessa tuotetun aineiston säilytettävyyttä ja jaettavuutta erilaisten tutkimusyhteisöjen ja niiden käytössä olevien repositorioiden kesken. Tämä edellyttää erilaisten palveluntuottajien välistä yhteistyötä sekä niiden digitaalisten infrastruktuurien sovittamista yhteen. Tämä ei ole ainoastaan tekninen haaste, vaan siihen liittyy myös poliittisia ja taloudellisia ratkaisuja. Nopeasti muuttuvien ja kasvavien digitaalisten rakenteiden maailmassa on löydettävä toimivia ja kestäviä ratkaisuja siihen, miten tutkimusdataa voidaan paitsi julkaista myös löytää, saada käyttöön, niveltää yhteen ja käyttää uusissa yhteyksissä.

Kokouksen keskiössä olivat luonnollisesti tutkimusdatan ja muun aineiston kansainväliset jättiläiset, joiden asiantuntijat harvoin kokoontuvat yhteen etsimään ratkaisuja EU-rahoitteisten monenkirjavien projektien yhteisiin kysymyksiin. [OpenAIRE](#) tarjoaa tutkijoille tietoa avoimen tallentamisen mahdollisuuksista, repositorioille mahdollisuuden tehdä aineistonsa näkyviksi, tutkimusyhteisöille ja rahoittajille edellytykset tutkimusten tulosten monitorointiin. OpenAIREssa on säilötyinä monenlaisia julkaisuja ja monenlaista tutkimusdataa useilla maailmankielillä. [Zenodo](#) on CERNin ja OpenAIREn ylläpitämä avoin repositorio.

## Koko euroopan laajuista yhteistyötä

[EUDAT](#):n visiona puolestaan on mahdollistaa Euroopan eri tieteenalojen tutkijoille ja muille toimijoille säilyttää, löytää, saada käsiinsä ja käsitellä tietoja luotettavassa ympäristössä osana yhteistä infrastruktuuria (CDI) yhteistyöverkostossa, jossa toimii useita erilaisia repositorioita. EUDAT-verkoston kuuluu 35 organisaatiota 15 maasta. Kokouksessa esiteltiin EUDAT:n B2FIND-palvelua, jonka avulla käyttäjät löytävät tutkimusdataa ja muita aineistoja. B2FIND on rakennettu siten, että ensin on koottu runsaasti erilaisten aineistojen metadatatietueita, sitten käännetty ja kartoitettu ne yhtenäiseen muotoon. Kolmanneksi ne on ladattu datasetteinä luetteloon, josta niitä on vaivatonta hakea. Vaativilla teknisillä ratkaisuilla homogenisoidun materiaalin näkyväksi tekeminen palvelee ratkaisevasti EUDAT:ssa julkaistujen aineistojen näkyvyyttä.



**GÉANT** ("jättiläinen" ranskaksi) puolestaan palvelee erilaisia Euroopassa vaikuttavia e-infrastruktuureja toimivan ja kustannuksiltaan edullisen verkoston, "palvelupilven" luomiseksi. Sen tarkoituksena on parantaa kansainvälistä tutkimusyhteistyötä eri tieteenaloilla. **EGI:n** luonnontieteiden erilaisille aloille kohdistuvat palvelut ovat räätälöidymiä: EGI tuottaa teknisiä palveluita tutkimusdatan välittämiseksi erilaisten paikallisten tutkimusympäristöjen ja infrastruktuurien välillä "yhdistetyn pilven" (*federated cloud*) avulla. Tätä pilvipalvelua käyttää nyt myös **DARIAH**, taiteiden ja ihmistieteiden digitaalista tutkimusta palveleva verkosto.

Tämän hetken tärkeimpänä tavoitteena näyttää olevan erilaisten EU-rahoituksella Eurooppaan luotujen infrastruktuurien saattaminen täysin keskenään toimiviksi. Tutkijoiden ja muiden käyttäjien tulisi päästä samoilla tunnuksilla kaikkiin e-infrastruktuureihin. edelleen e-infrastruktuurien tarjoamia palveluita tulee voida yhdistää tiedeyhteisöjen kulloisiakin tarpeita vastaaviksi. E-infrastruktuurien sisältämät palvelut tehdään ajan myötä näkyviksi yhtenä kokonaisuutena, jonka sisältöä voi selata. Tällä hetkellä fragmentoituneessa tilassa olevia palveluita on pyritty yhdistämään. Esimerkiksi GÉANT on tehnyt **eduGAIN:n** ja **AARC-projektin** teknisesti yhteensopiviksi, ja EGI ja EUDAT ovat yhteistyössään sovittaneet EGI:n tietojenkäsittelypalvelut yhteen EUDAT:n datanhallintaprosessien kanssa. Tarkoituksena on tarjota loppukäyttäjille saumaton pääsy EGI:n ja EUDAT:n yhtenäiseen infrastruktuuriin ja palveluihin.

## **Käytännön esimerkki digitaalisten ihmistieteiden piiristä**

Rotterdamin Erasmus-yliopistossa tutkijana vaikuttava **Stef Scagiola** puhui suullisen historian narratiivisten aineistojen merkityksestä tutkimukselle sekä niiden säilyttämisen ja jatkokäytön vmukanaan tuomista haasteista. Puolen vuosikymmenen takaiset kasettinauhurilla tehdyt kohisevat haastattelut ovat digiaikana vaihtuneet internetissä saatailla oleviin teknisesti korkeatasoisiin äänitteisiin. Aineistossa eivät välity ainoastaan lauseet ja ajatukset vaan myös elävä ääni ja tunteet. Scagiola itse on projektissaan haastatellut toisen maailmansodan hollantilaisia veteraaneja ja koonnut haastatteluista laajan aineiston. Hän kysyikin esityksessään, kuinka tekniset innovaatiot voivat tukea monimuotoisen suullisen datan hyödyntämistä ja siten uudistaa poikkitieteellistä tutkimusta.

Vaikka EU:n datanhallintaprojektien seuraaminen on peruskirjastolaiselle vaikeaa ja ehkä turhauttavaakin, prosessien keskeiset toimijat ja niiden tarjoamat palvelut sekä niiden saattaminen toimivaksi ja mahdollisimman hyvin loppukäyttäjiä palvelevaksi kokonaisuudeksi on melko hyvin hahmotettavissa. Hankkeiden tekniset yksityiskohdat voivat jäädä perehtyneempien asiaksi. Hyvää perustietoa saa vastikään ilmestyneestä EU:n tiedotteesta [E-infrastructures: making Europe the best place for research and innovation](#).

Teksti

***Matti Myllykoski***

*Johtava tietoasiantuntija*