

Miten humanistien aineistot avautuvat?

Fakta Julkaistu 01.12.2015 klo 15:27 Kirjoittaja Jessica Parland-von Essen

Digitaaliset ihmistieteet ovat maassamme kovassa nousussa. Perinteisiä humanistien sähköisiä aineistoja ovat olleet esimerkiksi kielikorpukset. Niitä on Suomessa jo pitkään kehitetty kansainvälisessä CLARIN-yhteistyössä, johon viimein tänä syksynä FIN-CLARIN liittyi myös virallisesti. Hankkeen [Kielipankki](#) uudisti vastikään verkkopalveluaan, josta löytyy paljon aineistoja.

Toinen perinteinen sähköisten aineistojen kategoria on muistiorganisaatioiden aineistot, jotka ovat pitkään olleet hakujärjestelmien ja -toimintojen jatkeita. Tähän ryhmään voidaan luokitella esimerkiksi aineistojen hakupalvelut ja digitaaliarkistot. Niitä ei itseään ole varsinaisesti luotu laskennallisten menetelmien käyttöön. Silti niitä voi käyttää tiedon lisäämiseen digitaalisten ihmistieteiden menetelmien avulla ja ne ovatkin lähteinä usein hyvin semanttisesti rikkaita ja mielenkiintoisia.

Tutkimuksen tietoaaineistoja säilytetään arkistoissa, kuten [Tietoarkistossa](#), joka on korkeatasoinen täyden palvelun talo varsinkin yhteiskuntatieteellisten alojen tutkijoille. Arkistossa tutkijoiden luovuttamaa dataa hoivataan ja sitä jaetaan jatkokäyttöön yhä enenevässä määrin. [Aila-aineistopalvelusta](#) voi helposti ladata itselleen tutkijoiden tuottamia aineistoja, joko avoimesti tai lupaa hakemalla. Koska aineistosta on huolehdittu asianmukaisesti, se on hyvälaatuista ja käytettävää, tutkijalta saatua ja tutkijalle yleensä aika suoraan kelpaavaa.

Maassamme on kuitenkin vielä suuria humanistisen alan tutkimusaineistoja, joita ei ole digitoitu tai julkaistu. Yli sata vuotta tutkijat ovat keränneet esimerkiksi kansanperinneaineistoja ja luoneet erilaisia tietoaaineistoja, joita löytyy monesta eri arkistosta, museosta ja kirjastosta. Näihin on usein panostettu valtavasti työtä ja tieto on usein hyvin rakenteistettua, mutta niiden saatavuus ja käytettävyys on suunnattoman paljon parempaa digitaalisessa muodossa.

AVAA on avointen aineistojen julkaisualue. AVAA oma aineistosi: [ohjeet ja lisätietoa palvelusta](#).



AVAA-palvelussa avattu Kirjojen levinneisyys Suomessa vuoteen 1850 -sovelluksen avulla on mahdollista tarkastella vuosien kirjakartoitustyön tuloksia.

Kuluneen vuoden aikana opetus- ja kulttuuriministeriön AVAA-palvelussa on avattu kaksi tutkijoille tehtyä sovellusta, jotka perustuvat juuri tällaisiin mielenkiintoisiin digitoituihin aineistoihin. Kotimaisten kielten keskus on avannut AVAA-palvelussa Itämerensuomalaisen kielikartaston, Lauri Kettusen suomen murteiden kartaston ja Suomen murteiden sanakirjan levikkikartat. Aineistojen hyödyntämisen tueksi on rakennettu [Kielikartastot-sovellus](#) jolla voi tehdä hakuja eri murreaineistoihin ja tarkastella löydettyjä ilmiöitä kartalla. Aineistot voi myös ladata itselleen tiedostopaketteina ja tarkastella aineistoja rajapinnan kautta. Toinen vastikään AVAA-palvelussa avattu aineisto on [Kirjojen levinneisyys Suomessa vuoteen 1850](#) -sovellus, jonka avulla on mahdollista tarkastella vuosien kirjakartoitustyön tuloksia. Tietokantaan on pääosin Helsingin yliopiston kirkkohistorian laitoksen johdolla tallennettu tietoja vanhoista kirjoista, joita löytyy eri puolelta maamme. Tietoja voi tarkastella hakujen avulla esimerkiksi kartalla tai aikajanalla. Uutta sovelluksessa on, että myös yleisö voi osallistua tiedon keräämiseen luetteloimalla vanhoja kirjoja tietokantaan. Uudet tiedot tallennetaan ja julkaistaan kun asiantuntijat ovat tarkistaneet ne. Sovelluksessa on myös pyritty huomioimaan muu kuin tutkijakäyttö siten, että myös aiheeseen perehtymätön voi tutustua aineistoon.

Haasteena digitaalisten ihmistieteiden puolella on tietojen monimuotoisuus ja käyttötarpeiden erilaisuus. Tutkimustarpeiden on oltava etusijalla, mikä tekee palvelujen käytettävyyden suunnittelun haastavaksi kun palveluja samalla halutaan suunnata suurelle yleisölle ja esimerkiksi opetuskäyttöön sopivaksi. Hyvän aineistohallinnan suunnittelun avulla on kuitenkin mahdollista saavuttaa hyötyjä tehokasta uudelleenkäyttöä mahdollistamalla.