

# NÄKÖKULMIA AVOIMEEN DATAAN - YLIOPISTON ROOLI TOIMIVAMMAN ARJEN JA PAREMPIEN PALVELUJEN KEHITTÄMISESSÄ KANSALAISILLE

Turun yliopistossa on meneillään All Eyes on Data -keskustelu-sarja, jonka puitteissa käsitellään datanhallintaa, analysointia ja hyödyntämistä monista eri näkökulmista. Yliopistoissa tuotettujen tutkimusaineistojen hyödyntäminen liiketoiminnan kehittämisessä nousi teemaksi viime vuoden joulukuun avoimessa tilaisuudessa, johon osallistui noin 80 henkeä korkeakouluista, yrityksistä ja kuntasektorilta. Tilaisuudessa haastettiin tutkijoita pohtimaan, millainen rooli heidän tutkimusaineistollaan voisi olla toimivamman arjen ja laadukkaampien palvelujen ja tuotteiden kehittämisessä kansalaisille.

## Yliopistot raakadatan tuottajina yhteiskunnan eri toimijoille

**Y**hteiskunnan ja talouden eri sektoreiden toimijat hyödyntävät ja analysoivat erilaista rutiininomaisesti keräämäänsä dataa kehittääkseen toimintaansa vastaamaan yhä kompleksisempiin haasteisiin ja säilyttääkseen kilpailukykyänsä globaaleilla markkinoilla. Organisaation itse keräämää dataa voidaan rikastaa yhdistelemällä sitä esimerkiksi eri lähteistä saatavaan avoimeen dataan. Datan hyödyntämiseen voidaan soveltaa big data -analytiikkaa ja avoimen lähdekoodin ohjelmistoja. Vaikka nykyään on saa-

tavilla suuria määriä dataa, kaikki data ei välttämättä ole relevanttia eikä datan prosessoiminen ole järkevää ainoastaan siitä syystä, että data on saatavilla.

Dataan liittyvät juridiset ja eettiset kysymykset ovat tällä hetkellä keskiössä datan hyödyntämistä ja avaamista koskevassa keskustelussa yliopistoissa ja muissa korkeakouluissa. Yliopistoissa tuotetun datan hyödyntäminen vaatii selkeyttä datan omistajuuden ja käyttöoikeuksien suhteen. Tutkimusaineistoon liittyvistä juridisista kysymyksistä onkin tärkeää sopia ennen datankeruun aloittamista. Vasta näin voidaan luoda yliopistoissa tuotetusta datasta eri tavoin laadukasta ja toimivaa

arkea sekä erilaisia kansalaisille suunnattuja palveluita. Liiketoimintaa yliopisto voi tehdä ainoastaan sellaisella datalla, johon yliopistolla on oikeudet.

Data tarjoaa uusia mahdollisuuksia kehittää liiketoimintaa. Datan hyödyntämiseen kytkeytyvässä muutoksessa ei kuitenkaan ole kyse vain teknologiasta: murroksen kohteena ovat myös yritysten liiketoimintamallit, osaamistarpeet, toimintatavat, kulttuuri ja johtaminen. Datan hyödyntäminen liiketoiminnan raaka-aineena tarkoittaa kehittyneimmillään sitä, että yritys lähtee toiminnassaan liikkeelle asiakkaiden tarpeista, jotka ovat moninaisemmat kuin yrityksen palvelut ja/tai tuotteet. Tässä kehitystyössä voidaan hyödyntää myös yliopistoissa syntyvää dataa eli tutkimusaineistoja. Tutkimusaineistojen hyödyntämiseen liittyviä haasteita voidaan ratkoa tarjoamalla tutkijoille tukipalveluja.

### Haasteita datan jakamisessa ja hyödyntämisessä

**D**atan hyödyntämiseen tutkimuksessa ja liiketoiminnan kehittämisessä liittyy useita haasteita. Dataa kertyy useista eri lähteistä ja data on usein tallennettu eri formaateilla eri tallennuspaikkoihin. Mikä tahansa data on tuotos, joka syntyy datankeruun ja käsittelyn kautta. Siten se on riippuvainen näistä toiminnoista. Ei myöskään riitä, että datan hyödyntämiseen on halua; datan pitää sisältää hyödyllistä informaatiota. Datan metadata ja kuvailu

onkin hyödyntämisen ja löydettävyyden kannalta olennaista.

Keskeisen haasteen muodostaa myös datan laatu. Tutkimusaineistojen hyödyntämisen osalta tärkeintä on laadukas tutkimusprosessinaikainen datanhallinta, joka kattaa aineistonkeruun, käsittelyn ja analysoinnin. Datan hyödyntäminen vaatii hyvin pohdittua suunnitelmaa datan hallinnasta, joka edellyttää esimerkiksi strategisia ratkai-

**”Tohtorikoulutettavat olivat valmiita jakamaan dataansa, mutta koska sopimuksia, lupia, datan dokumentointia ja kuvailua ei oltu tehty tästä näkökulmasta, jakaminen ei useimmiten ollut mahdollista”**

suja aineiston avaamisesta ja arkistoinnista. Suunnitelman toteuttamiseen tarvitaan puolestaan taitoja.

Turun yliopiston kirjasto haastateli tutkimuksen palveluista vastaavan palvelupäällikkö Jukka Rantasaaren johdolla yli 30 tohtorikoulutettavaa, ohjaajaa ja biostatistikkoa. Kirjaston lisäksi mukana haastatteluissa oli myös Tutkimuksen IT -yksikön päällikkö Meri-Tuulia Kaarakainen. Haastatteluissa kysyttiin haastateltavien näkemyksiä mm. datan elinkaaren eri vaiheiden tärkeydestä sekä heidän arviotaan tohtorikoulutettavien osaamisen tasosta em. vaiheissa. Alustavien



tulosten perusteella voidaan sanoa, että tohtorikoulutettavien taitojen kehittymistä tulee tukea tutkimusdatan elinkaaren hallinnan kaikissa vaiheissa, jotta tutkimusdatan hyödyntäminen myös meneillään olevan projektin jälkeen ja ulkopuolella on mahdollista. Tuloksia on jo hyödynnetty tänä keväänä aloitetussa monialaisen työryhmän suunnittelemassa ja toteuttamassa Tutkimusdatan hallinnan perusteet -koulutuksessa, joka on suunnattu tohtorikoulutettaville ja post doc -tutkijoille.

Tohtorikoulutettavat olivat valmiita jakamaan dataansa, mutta koska sopimuksia, lupia, datan dokumentointia ja kuvailua ei oltu tehty tästä näkökulmasta, jakaminen ei useimmiten ollut mahdollista. Lisäksi, usein projektissa kerätyn datan omistajuus ei ole yksittäisellä tohtorikoulutettavalla vaan data on osa suurempaa datakokonaisuutta, jonka omistajana on projekti. Usein dataa halutaan myös suojella, koska dataan liittyy kilpailuetua tai koska avaamisen pelätään vahingoittavan datan jatkokäyttöä ja hyödyntämistä kaupallisessa tarkoituksessa. Tallennuspaikkojen tuntemus on niin ikään puutteellista.

### Monialaista tukea tutkijoille

Turun yliopistossa useat eri palveluyksiköt toimivat yhteistyössä mahdollistaakseen tutkijoille laadukkaat ja ajantasaiset dataan liittyvät tukipalvelut. Palveluja halutaan kehittää tutkijalähtöisesti.

Vuonna 2018 lanseerattu All Eyes on Data -keskustelusarja on hyvä esimerkki tavasta tukea tutkijoita. Keskustelusarjaa toteutettavat yhteistyössä yliopiston avoimen tieteen asiantuntijat ja Digitaaliset tulevaisuudet -kokonaisuus, joka on yksi Turun yliopiston strategian temaattisista tutkimusalueista.

**K**eskustelusarja pitää sisällään keskustelevia tilaisuuksia, joita järjestetään ympäri yliopistokampusta. Kussakin tilaisuudessa on oma teemansa, joita yliopiston tutkijat voivat myös ehdottaa toteuttaviksi. Tilaisuuksissa puhujina on usein eri alojen asiantuntijoita ja tutkijoita. Esimerkiksi Datasta liiketoimintaa -tilaisuuden puhujina olivat Meri-Tuulia Kaarakainen, innovaatioasiamies Marjo Pihlavisto (Turun yliopisto), datakonsultti Olli Lindroos (Solita) ja data-analyytikko Mikko Tukiainen (Auria Biobank). Seuraava tilaisuus järjestetään toukokuussa 2019 ja teemana on laadullisen tutkimusaineiston hallinta ja avoimuus.

Keskustelusarjan lisäksi Turun yliopisto tarjoaa tutkijalle koulutusta ja tukea datanhallinnan osaamisessa ja kaupallistamisessa sekä kehittää datainfrastruktuuria datan koko elinkaaren hallintaa ajatellen.

Kirjaston kokoamassa tutkimusdata-oppaassa sekä aineistohallintakoulutuksissa tuetaan tutkijoita datanhallintasuunnitelman tekemisessä ja esimerkiksi aineiston avaamiseen ja arkistointiin liittyvissä kysymyksissä.

Parhaillaan käynnissä on myös edellä mainittu Tutkimusdatan hallinnan perusteet -kurssi, jossa käydään läpi datan elinkaaren vaiheet, erilaiset datatyypit ja -muodot sekä tietoturvan, tietosuojan ja sensitiivisen sisällön vaatimukset datan käsittelylle ja jatkokäytölle.

Innovaatiot, yrittäjyys ja koulutusvienti -yksikössä annetaan neuvoa, tukea ja koulutusta siitä, miten tutkimustoiminnan tuloksena syntyneitä aineistoja kannattaa kaupallistaa ja hyödyntää muulla tavoin. Oma roolinsa on myös avoimen tieteen edistämistä koordinoivalla Tutkimusedellytykset-yksiköllä, Lakiasiat-yksiköllä ja

tietosuojavastaavalla.

Turun yliopistossa tallennusmäärät kasvavat tällä hetkellä noin 50 prosentin vuosivauhtia. Tutkimuksen muutuksessa entistä dataintensiivisempään suuntaan on myös Tutkimuksen IT-yksikön rooli yhä tärkeämpi. Yksikön vastuulla on kehittää kattavaa datainfrastruktuuria, jonka avulla mahdollistetaan modernit datanhallinnan palvelut, kuten tutkimuksen aikaisen datan hallinta. Yksikkö tarjoaa myös datan hyödyntämiseen liittyvää ohjelmistokehityspalvelua ja neuvontaa tutkimuksen jälkeiseen datan hallintaan.

### Lähteet:

All Eyes on Data -keskustelusarjan materiaalit ja tiedotus:  
<http://digitalfutures.fi/digital-futures-all-eyes-on-data/>

Tutkimusdata ja Turun yliopiston datapolitiikka: <https://utuguides.fi/tutkimusdata>

Innovaatiot, yrittäjyys ja koulutusvienti -yksikön sivut:  
<https://www.utu.fi/fi/yritysyhteistyö/palvelut>

Tutkimuksen IT:n ohjelmistokehityspalvelut: <http://researchit.utu.fi/>

### Tietoa kirjoittajista

PÄIVI KANERVA  
 Informaatikko  
 Turun yliopiston kirjasto  
[paivi.kanerva@utu.fi](mailto:paivi.kanerva@utu.fi)

ERIKA LILJA  
 Kehittämisasiantuntija  
 Tutkimusedellytykset  
 Turun yliopisto  
[emolli@utu.fi](mailto:emolli@utu.fi)

ANNE-MARIE TUUKKA  
 Tutkimuskoordinaattori  
 Tutkimusedellytykset  
 Turun yliopisto  
[anne-marie.tuikka@utu.fi](mailto:anne-marie.tuikka@utu.fi)