


## Suurimmat juridiset esteet avoimelle tieteelle ja tutkimukselle Suomessa

Avoimen tieteen ja tutkimuksen tavoitteen mukaan tutkimuksen tulokset liikkuvat sujuvasti koko yhteiskunnassa; tutkijalta ja tutkimusryhmiltä toiselle, tutkimusalojen välillä, innovaatiotoimintaan sekä päättäjille ja kansalaisille. Tämän tulee tapahtua tutkimuseettisiä periaatteita noudattaen ja oikeudellista toimintaympäristöä kunnioittaen. Tutkimusympäristön avoimuuteen pyritään aina, kun se on lainsäädännön ja sopimusten puolesta mahdollista. Tällöin tutkimustulosten jatkokäyttöä ei rajoiteta tarpeettomasti, ja käyttöehdot tuodaan selkeästi esille.

Avoim tiede ja tutkimus -hankkeen (ATT) strategiatyössä tunnistettiin yhdeksi keskeisistä esteistä Suomen lainsäädännössä olevat esteet tutkimusjulkaisujen, tutkimusaineistojen ja tutkimusdatan avoimelle saatavuudelle. Tilanteesta pyydettiin tarkempi selvitys ATT-hankkeen Oikeuksien hallinta -työryhmältä, jonka puheenjohtajana toimii tekijänoikeusneuvos Viveca Still opetus- ja kulttuuriministeriöstä.

Selvitys kuvaa keskeiset ongelmat, muutostarpeen sekä ehdotukset ongelmanratkaisuun.

Selvityksen laatineeseen työryhmään kuuluivat:  
*Viveca Still*, opetus- ja kulttuuriministeriö (puheenjohtaja)  
*Pekka Heikkinen*, Kansalliskirjasto  
*Pirjo Kontkanen*, Helsingin yliopisto  
*Mikael Laakso*, Hanken  
*Jonas Lång*, Svenska litteratursällskapet i Finland  
*Seliina Päällysaho*, Seinäjoen ammattikorkeakoulu  
*Maria Rehbinder*, Aalto-yliopisto

<b>Julkaisu</b> Suurimmat juridiset esteet avoimelle tieteelle ja tutkimukselle Suomessa	
<b>Julkaisija</b> Avoin tiede ja tutkimus -hanke	<b>Julkaisuajankohta</b> 17.10.2016
<b>Tekijä</b> Avoin tiede ja tutkimus -hanke, Oikeuksien hallinta -työryhmä	
<b>Lisenssi</b>  Tämä teos on lisensoitu <a href="https://creativecommons.org/licenses/by/4.0/">Creative Commons Nimeä 4.0 Kansainvälinen -lisenssillä</a> .	
<b>Julkaisun jakelu</b> PDF-tiedosto ladattavissa sivuilla <a href="http://avointiede.fi/keskeiset-julkaisut">avointiede.fi/keskeiset-julkaisut</a>	
<b>Yhteystiedot</b> <a href="http://avointiede.fi">http://avointiede.fi</a>	

# Suurimmat juridiset esteet avoimelle tieteelle ja tutkimukselle Suomessa

Tässä kirjoituksessa tarkastellaan lainsäädännössä olevia esteitä tutkimusjulkaisujen, -aineistojen ja -datan avoimelle saatavuudelle Suomessa. Tarkastelun kohteena ovat tekijänoikeuslaki sekä tietosuojalainsäädäntö. Kirjoituksessa esitetään myös esimerkkejä, miten muissa maissa asioita on ratkaistu ja millaisia toimenpiteitä Suomessa voisi olla syytä harkita avoimen saatavuuden edistämiseksi.

## 1. Tiedonlouhinta ja tekijänoikeuslaki

Tiedonlouhinta on prosessi, jossa tietoa etsitään koneluettavasta aineistosta yhdistelemällä louhittuja tietoja uudelleen mallien ja suuntausten tunnistamiseksi. Tiedonlouhinta edellyttää, että aineistosta voidaan valmistaa kopioita. Tiedonlouhinta voi edellyttää, että louhittu tai louhittava aineisto muutetaan toiseen muotoon. Lisäksi voi olla tarpeen tallentaa aineistoa tai otteita aineistosta omaksi tietokannaksi tai muuksi koosteeksi.

Tiedonlouhinnan tarjoamat hyödyt ovat laaja-alaisia. Osana EU:n Digital Single Market -strategian toimenpiteitä EU:n tekijänoikeusdirektiiviin on ehdotettu rajoitussäännöstä, joka sallisi tiedonlouhinnan.

Suomen tekijänoikeuslaki edellyttää, että louhintaprosessiin kuuluvaan kopiointiin hankitaan oikeudenhaltijan lupa. Myös aineiston muuntelu saattaa edellyttää oikeudenhaltijan lupaa. Jos aineistoa tai sen osia halutaan tallentaa omaksi tietokannaksi tai muuksi koosteeksi, joka on tarkoitus saattaa yleisön saataviin, tarvitaan myös tekijänoikeudenhaltijan lupa.

Kopioston sopimuslisenssin nojalla saa tutkimusta varten kopioida digitaalisesti julkaisuja tutkimusryhmälle tarkoituksen edellyttämässä laajuudessa, kun se on tutkimuksen edistämisen kannalta olennaista. Tällaisen louhittua sisältöä sisältävän uuden tietokannan tai muun koosteen yleisön saataviin saattaminen tai aineiston pysyvä säilyttäminen ei kuitenkaan ole Kopioston luvalla mahdollista. Muiden tekijänoikeusjärjestöjen kanssa ei ole vastaavanlaista sopimuslisenssiä ja esim. äänitiedostojen käyttäminen edellyttää luvan hankkimista yksittäisiltä oikeudenhaltijoilta.

Kansainväliset kustantajat, kuten esim. Elsevier, ovat valmiita myymään oikeutta tiedonlouhintaan yliopistoille ja tutkimuslaitoksille. Tiedonlouhintalupien ostaminen jokaiselle tutkimushankkeelle nostaa kuitenkin kustantajille maksettavia korvauksia. Esim. viiden suurimman tiedelehtipaketin tilausten kansalliset kokonaiskustannukset olivat vuosina 2012 - 2016 yhteensä lähes 64 miljoonaa euroa. Tämä summa ei sisällä korvauksia tiedonlouhinnasta. Lisäksi kustantajien omat palvelut eivät yleensä salli tutkijan käyttää haluamaansa tiedonlouhintasovellusta, vaan louhinta on mahdollista vain kustantajan tarjoamaa sovellusta käyttäen.

Lähtökohtaisesti kukin kustantaja voi tarjota tiedonlouhinta-oikeutta vain omaan aineistoonsa ja vain siltä osin kuin tästä on sovittu alkuperäisten tekijöiden kanssa. Hyvin suurten datamassojen (ns. big data) tiedonlouhinta ei näin ollen ole yleensä mahdollista. Suomalaiset kustantajat eivät toisaalta yleensä ole sopineet alkuperäisten tekijöiden kanssa tiedonlouhinta-oikeuksista, eivätkä salli tiedonlouhinta.

Tämä kaikki johtaa siihen, että tutkijat eivät voi vapaasti valita tutkimuskohdettaan. Tutkimus rajautuu sellaisiin kohteisiin, joihin on mahdollista saada lupa tiedonlouhintaan. Esimerkiksi eräs sosiaalista mediaa ylläpitävä yritys on antanut luvan erään sosiaalisen median sivuston tiedonlouhintaan tutkimuskäyttöön korvauksetta Helsingin yliopiston tutkijoille. Yritys saa tutkimushankkeen avulla tietoonsa uusia tapoja hyödyntää sivuston dataa ja voi käyttää tätä tietoa uusien liiketoimintamahdollisuuksien kehittämiseen. Samalla tutkijat saavat arvokasta tutkimusdataa.

Tiedonlouhinta lienee USA:ssa sallittua tietyissä rajoissa tekijänoikeuslain fair use -säännöksen nojalla. Säännöksen soveltamisessa otetaan huomioon käytön tarkoitus, eli erityisesti se, tavoitellaanko käytöllä taloudellista hyötyä. Huomioon otetaan myös hyödynnetyn aineiston luonne ja laajuus sekä käytön vaikutuksia tekijänoikeuden alaisen teoksen kaupalliseen arvoon. Esimerkiksi Google Books, jossa laajoja kirja-aineistoja on digitoitu ja johon liittyy Google Books:n mahdollistama tiedonlouhinta, toisintamatta käyttäjille teoksia kokonaisuudessaan, on pidetty fair use -säännöksen nojalla sallittuna.

Japanin, Ison-Britannian ja Irlannin tekijänoikeuslainsäädännössä on erityisiä rajoitussäännöksiä, jotka sallivat tiedonlouhinnan.

Vastaavanlaista sääntelyä voisi harkita myös Suomessa. Lähtökohtana tulisi olla, että on oikeus louhia laillisesti hankittua aineistoa tekijöiden oikeutetut intressit huomioiden. Tulisi ymmärtää, että tieteellinen tutkimuskäyttö ei kilpaile oikeudenhaltijoiden kaupallisten liiketoimintaintressien kanssa. Tutkimuksen toteuttamismahdollisuudet eivät saisi riippua siitä, saadaanko louhintalupa vai ei. On muistettava, että tutkimusrahoittajat ja tiedekustantajat edellyttävät yhä enemmän tutkimusaineistojen julkista saatavuutta.

EU:ssa meneillään olevan tekijänoikeuslainsäädännön uudistuksen yhteydessä Suomen tulisi edistää mahdollisuuksia tiedonlouhintaan myös rajat ylittävissä tilanteissa, sillä yhä enenevästi tutkimustyötä tehdään laajemmissa verkostoissa, jolloin on tärkeää, ettei tiedonlouhintaan kohdistu kansallisia esteitä.

## 2. Julkaisujen rinnakkaistallentaminen ja tekijänoikeuslaki

Edistääkseen julkaisujen avointa saatavuutta korkeakoulut ja tutkimuslaitokset ovat alkaneet yhä enemmän tallentaa julkaisuja ylläpitämiinsä julkaisuarkistoihin. Tällainen tallentaminen edellyttää kuitenkin tekijänoikeuden haltijoiden, tavallisesti kustantajien lupaa.

Keskeinen ongelma on kustantajien ja tutkimusrahoittajien välisten vaatimusten ristiriita. Tutkimusrahoittajat edellyttävät yhä enemmän tutkimusjulkaisujen pysyvää avointa saatavuutta viimeistään asettamassaan määräajassa. Kustantajat eivät yleensä salli rinnakkaistallentamista näissä määräajoissa.

Tutkijoiden meritoituminen ja korkeakoulujen laadun arviointi perustuvat tieteellisiin julkaisuihin. Kansainväliset tieteelliset kustantajat eivät välttämättä salli suoraan, ilman korvausta tällaista rahoittajien edellyttämää avointa julkaisemista, vaan korkeakoulujen ja tutkimuslaitosten on maksettava kustakin artikkelista ns. Article Processing Charge eli APC-maksu, joka on tyypillisesti 1500 - 3000 euroa artikkelilta. Tämä maksu veloitetaan normaalien tilausmaksujen lisäksi. Heikon neuvotteluasemansa vuoksi tutkijoiden ja korkeakoulujen on yleensä taivuttava kustantajien tarjoomiin epäedullisiin sopimuksiin.

Saksassa ja Hollannissa on muutettu tekijänoikeuslakia siten, että rinnakkaistallentaminen on mahdollista ilman korvausta laissa säädetyn määräajan jälkeen. Ranskassa valmistellaan samankaltaista säädöstä. Vastaavanlaista sääntelyä voisi harkita myös Suomessa, jotta esim. tutkimusrahoittajien ehtoja voidaan varmasti noudattaa ilman merkittäviä lisäkustannuksia korkeakouluille ja tutkimuslaitoksille.

## 3. Henkilötietojen käyttäminen

Tietosuoja on tekijänoikeuden ohella keskeinen haaste tutkimustiedon avoimelle hyödyntämiselle. Sääntely on parhaillaan murrosvaiheessa, kun EU:n uusi tietosuoja-asetus tuli voimaan toukokuussa 2016 ja sitä ryhdytään soveltamaan jäsenmaissa toukokuussa 2018. Vaikka asetus on jäsenmaissa suoraan sovellettavaa oikeutta, sisältää se runsaasti direktiivinomaista sääntelyä, eli kohtia, joista jäsenmaat voivat kansallisesti säätää erikseen.

Sisällöllisesti uusi asetus ei merkitse kovin suurta muutosta, vaan sääntelyn rakenne ja peruskäsitteet säilyvät ennallaan. Avoimen tieteen ja tutkimuksen näkökulmasta uusi tietosuoja-asetus sisältää kuitenkin osin vastakkaisia arvoja. Asetustekstissä painotetaan sananvapauden ja tiedon vapaan liikkuvuuden merkitystä, toisaalta koko tietosuojuudistuksen keskeisenä tavoitteena oli nimenomaan korostaa jokaisen yksilön omia mahdollisuuksia päättää itseään koskevan datan käytöstä.

Avoimen tieteen ja tiedon avoimen liikkuvuuden kannalta asetuksessa on myös ongelmakohtia. Henkilötietojen suojaa koskevat peruskäsitteet eli henkilötieto ja henkilötiedon käsittely on määritelty hyvin laajasti. Henkilötietoja ovat kaikki tiedot, jotka voidaan yksilöidä jotakuta yksittäistä henkilöä koskeviksi, henkilötietojen käsittelyä taas on kaikki henkilötietojen "automaattinen käsittely". Henkilötietoja saa käsitellä vain laissa yksilöidyillä edellytyksillä, ja arkaluonteisten henkilötietojen käsittely on muutamia, asetuksen 9 artiklassa mainittuja poikkeuksia lukuun ottamatta, lähtökohtaisesti kokonaan kielletty.

Tietosuojan osalta tilanne on toisaalta tekijänoikeutta huomattavasti ongelmallisempi myös sen vuoksi, että mitään järjestöä, jonka kanssa tietojen käsittelystä voisi kollektiivisesti sopia, ei ole. Tietojen käsittelystä on lähtökohtaisesti sovittava kaikkien niiden yksittäisten henkilöiden kanssa, joiden henkilötietoja aineistoihin sisältyy.

Asetus lähtee lisäksi siitä, että henkilö voi pätevästi antaa suostumuksen tietojensa käyttöön vain johonkin tiettyyn käyttötarkoitukseen. Niin sanotun avoimen suostumuksen antaminen ei siis ole sallittua. Tämä johtaa suuriin ongelmiin materiaalin jatkokäytön kannalta ja vaikeuttaa tutkimuksen toistettavuutta.

Tietynlaista liikkumavaraa lain soveltamiseen tähän asti on antanut se, että käsittelyyn on mm. yleisen edun perusteella voinut hakea oikeusministeriön alaisen tietosuojalautakunnan lupaa. Uuden EU-asetuksen myötä tietosuojalautakunta tullaan kuitenkin lakkauttamaan.<sup>1</sup>

Uuteen asetuksen sisältyy artiklan 17 oikeus tietojen poistamiseen ("oikeus tulla unohdetuksi"). Tämän kohdan nojalla yksityisellä henkilöllä on oikeus vaatia tietonsa poistettavaksi mm. sillä perusteella, etteivät henkilötiedot enää ole tarpeen sitä tarkoitusta varten, jota varten ne alkujaan kerättiin, tai jos käsittelylle ei enää ole olemassa perusteltua syytä. Nykyisen kansallisen lainsäädännön mukaan rekisteröidyllä on vain ollut oikeus vaatia itseään koskevan virheellisen tiedon korjaamista.

Voimaan tulleen asetuksen artiklan 17 kohdassa 3 todetaan, että *oikeutta vaatia tietoja poistettavaksi ei ole, jos käsittely on tarpeen sananvapautta ja tiedonvälityksen vapautta koskevan oikeuden käyttämiseksi tai yleisen edun mukaisia arkistointitarkoituksia tai tieteellisiä/historiallisia tutkimustarkoituksia varten.*

Jälkimmäinen em. kohdista soveltuu laajasti myös tieteellisten kirjastojen ja tieteellistä tutkimusta palvelevien arkistojen ja museoiden toimintaan. Muotoilu jättää kuitenkin avoimeksi sen, miten "yleisen edun mukainen arkistointitarkoitus" on määriteltävä. Myös EU-tuomioistuimen ja jäsenmaiden kansallisten tuomioistuinten viimeaikaisessa oikeuskäytännössä "oikeuteen tulla unohdetuksi" on viitattu ja sitä on tulkittu vaihtelevasti.<sup>2</sup>

Asetuksen lähtökohdiana on, että yksittäisten kansalaisten (eli rekisteröityjen) oikeudet ja vapaudet tulee täysimääräisesti turvata myös silloin, kun kyse on yleisen edun mukaisesta arkistoinnista tai tutkimuksesta. Henkilötietoja käsiteltäessä on noudatettava tietojen minimoinnin periaatetta. Paitsi että tiedot tulee suojata teknisesti - eli tulee huolehtia riittävästä tietoturvasta - ne tulee pseudonymisoida eli poistaa julkistettavasta aineistosta suorat tunnisteet. Mahdollisuuksien mukaan aineisto tulee tehdä kokonaan tunnistamattomiksi eli anonymisoida, jolloin tietoihin ei enää sovelleta henkilötietolakia lainkaan. Pseudonymisointi ja anonymisointi voivat kuitenkin olennaisesti vaikeuttaa aineiston jatkokäyttöä tutkimuksessa ja historiallisessa arkistoinnissa. Kansallisessa lainsäädännössä pitäisi etsiä ratkaisuja, jotka mahdollistavat aineistojen jatkokäytön tutkimuksessa ilman tarpeetonta pseudonymisointia ja anonymisointia. Tämä kansallinen säännös mahdollistaisi osaltaan akateemisen sananvapauden toteutumisen artiklojen 85 ja 89 mukaisella tavalla.

Avoimuuden edellytykset ovat jossain määrin erilaisia muistiorganisaatioiden eli kirjastojen, arkistojen ja museoiden sekä toisaalta tutkijoiden näkökulmasta. Muistiorganisaatioiden kannalta olennaisia ovat mahdollisuudet koota, luetteloida ja julkistaa henkilötietoja tutkijoiden käytettäväksi. Historiaa ei voi tallentaa käyttämättä niiden henkilöiden tietoja, jotka ovat olleet historiaan osallisina toimijoina.

Tutkijoilla tietosuoja tulee vastaan niin aineistojen keräämis-, käyttö- ja arkistointivaiheessa kuin tutkimustulosten julkistamisvaiheessa. Nykyisin tutkimusrahoittajat vaativat yhä laajemmin, että tutkimusaineistoja avataan, että tutkimus on toistettavissa ja että aineistoja voidaan käyttää uusissa tutkimushankkeissa.

Asetuksen 85 artiklan mukaan jäsenvaltioilla on velvollisuus sovittaa kansallisessa lainsäädännössä yhteen oikeus sananvapauteen ja oikeus henkilötietojen suojaan. Artiklan mukaan henkilötietojen käsittelylle akateemisen ilmaisun tarkoituksia varten on säädettävä tarvittavat vapautukset tai poikkeukset. Jo nykyisen kansallisen lain nojalla historiallinen tai tieteellinen tutkimus on sellainen erityistarkoitus, jota varten henkilötietoja on voinut käsitellä hieman tavanomaista väljemmin edellytyksin (14§).

Yleisen edun mukaisia arkistointitarkoituksia tai tieteellisen/historiallisen tutkimuksen tarkoituksia koskevista poikkeuksista säädetään artiklassa 89. Sen nojalla jäsenvaltiot voivat säätää kansallisessa

<sup>1</sup> Kansalliskirjasto haki syksyllä 2016 tietosuojalautakunnalta lupaa muistiorganisaatioiden yhteisen nimitietopalvelun perustamiseen. Lautakunta myönsi luvan kesäkuussa 2016, mutta vain määräaikaisena tietosuoja-asetuksen soveltamisen määräaikaan eli 25.5.2018 saakka.

<sup>2</sup> *Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja González*, 2014, jossa EU tuomioistuin katsoi, että henkilöllä oli oikeus vaatia itseään koskevat hakutulokset poistettavaksi internetin hakupalvelusta. Keväällä 2016 belgialainen vetoomustuomioistuin katsoi, että oikeus tulla unohdetuksi ulottuu myös digitoituihin lehtiartikkeleihin itsessään. Toisaalta kesäkuussa 2016 Ranskan korkein oikeus totesi, että oikeus tulla unohdetuksi ei soveltunut, kun artiklan sisältämät faktat olivat sinällään tosia. Pääsy artikkeliin on olennainen osa lehdistön vapauden *käyttämistä*.- Tapausten arviointia vaikeuttaa myös häilyvä rajanveto tietosuojan ja toisen perusoikeuden eli yksityisyyden suojan välillä.

lainsäädännössään asetuksen poikkeuksista, *jos asetuksen soveltaminen todennäköisesti estäisi tai suuresti vaikeuttaisi arkistoinnin tai tutkimuksen toteuttamista.*

Olisikin erittäin olennaista, että asetuksen soveltamiseen liittyvässä kansallisessa lainsäädännössä asetuksen mahdollistama kansallinen liikkumavara tulee mahdollisimman laajasti hyödynnettyä artiklojen 85 ja 89 osalta. Olisi tärkeää ottaa huomioon mm. seuraavat seikat: (1) rekisteröityjen oikeuksia (tietojen oikaisu, poisto, oikeus tulla unohdetuksi, siirto, vastustaminen) rajoitetaan yleisen edun mukaisia arkistointi- ja tutkimustarkoituksia varten, (2) varmistetaan, että yleisen edun mukaisiin arkistointi- ja tutkimustarkoituksiin luetaan yleistä etua palvelevien kirjastojen, arkistojen ja museoiden lisäksi tutkimusinfrastruktuurit, jotka tallentavat ihmisiä koskevia aineistoja, (3) määritetään konkreettiset vaatimukset asetuksen artiklan 89 mukaisille teknisille ja organisatorisille toimenpiteille ja (4) luodaan periaatteet tietojen minimoinnille, pseudonymisoinnille ja anonymisoinnille tavalla, jossa säilytetään tasapaino tieteen vapauden ja avoimuuden toteuttamisessa sekä tutkittavien henkilöiden yksityisyyden suojaamisessa.

Tutkimuksen tekeminen edellyttää, että yleistä etua palvelevat arkistot, kirjastot ja museot sekä tutkimusinfrastruktuurit toimivat saumattomana kokonaisuutena. Kansallisen tietosuojalainsäädännön tulee mahdollistaa tämä.