# Tutkimustietovaranto ja henkilötiedot

## Tutkimustietovarannon sisältämät henkilötiedot

Tutkimustietovarannossa käsitellään julkaisujen ja tutkimusaineistojen tekijätietoja sekä tutkimusrahoitusta hakeneiden henkilöiden nimiä. Henkilötietoja sisältäviä tietoja kerätään ja yhdistellään useista lähdejärjestelmistä. Näitä lähdejärjestelmiä ovat esimerkiksi korkeakoulujen ja tutkimuslaitosten tutkimustietojärjestelmät, julkaisu- ja tutkimusaineistorekisterit sekä tutkimusrahoittajien hankerekisterit. Tutkijat tai muut henkilöt eivät syötä tutkimustietovarantoon manuaalisesti omia tai muiden tietoja.

Paitsi että tutkimustietovarannon tietoja tullaan käyttämään viranomaisten lakisääteisten tehtävien hoitamiseen (esim.OKM), tavoitteena on myös avata eri lähdejärjestelmistä kerätyt tiedot eri palveluiden ja käyttäjien avoimesti hyödynnettäväksi teknisten rajapintojen kautta. Kolmantena tapauksena on tietojen käyttö suljetussa, ei-julkisessa tilassa.

On herännyt kysymys, mitkä ovat ne yleisen tietosuoja-asetuksen mukaiset oikeudelliset perusteet, jolla tutkimustietovarannossa henkilötietoja käsitellään. Yliopistothan saavat käsitellä kirjoilla olevien henkilöiden henkilötietoja lain perusteella ja CSC puolestaan saa tutkimustietovarannossa käsitellä niitä yliopistojen lukuun. Tietosuoja-asetuksen kannalta ongelmakohtia saattaisi olla kolme:

1) kun henkilötietoon liitetään jokin tieto, jota henkilö ei ole itse antanut, esimerkkinä on käytetty affiliaatiota

2) jos osa henkilötiedoista saadaan esimerkiksi säätiöiltä henkilön suostumuksen perusteella, suostumuksen on mahdollistettava henkilötietojen jatkokäyttö ja jakaminen avoimesti (jos sellaista halutaan tehdä)

3) tutkimustietovarannon henkilörekisterin avaaminen harvestoitavaksi teknisen rajapinnan kautta.

## EU:n uusi tietosuoja-asetus

EU:n yleinen tietosuoja-asetus (GDPR) tulee voimaan 25.5.2018. Tarkoituksen on yhtenäistää henkilötietojen käsittelyä jäsenmaissa. Tietosuoja-asetuksen 4 artiklan mukaan henkilötiedoilla tarkoitetaan kaikkia tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön liittyviä tietoja; tunnistettavissa olevana pidetään luonnollista henkilöä, joka voidaan suoraan tai epäsuorasti tunnistaa erityisesti tunnistetietojen, kuten nimen, henkilötunnuksen, sijaintitiedon, verkkotunnistetietojen taikka yhden tai useamman hänelle tunnusomaisen fyysisen, fysiologisen, geneettisen, psyykkisen, taloudellisen, kulttuurillisen tai sosiaalisen tekijän perusteella.

Tietosuoja-asetuksen 4 artiklan 2 kohdan määritelmän mukaisesti henkilötietojen käsittelyllä tarkoitetaan esimerkiksi tietojen keräämistä, tallentamista, järjestämistä, jäsentämistä, säilyttämistä, muokkaamista tai muuttamista, hakua, kyselyä, käyttöä, tietojen luovuttamista siirtämällä, levittämällä tai asettamalla ne muutoin saataville, tietojen yhteensovittamista tai yhdistämistä, rajoittamista, poistamista tai tuhoamista.

Tietosuoja-asetuksen 6 artiklan mukaan käsittely on lainmukaista ainoastaan jos ja vain siltä osin kuin vähintään yksi 6 artiklassa luetelluista edellytyksistä täyttyy. Kaikki 6 artiklan edellytykset ovat yhdenvertaisia, mutta ne antavat poikkeavia mahdollisuuksia henkilötietojen jatkokäsittelylle.

6 artikla: Käsittelyn lainmukaisuus

1. Käsittely on lainmukaista ainoastaan jos ja vain siltä osin kuin vähintään yksi seuraavista edellytyksistä täyttyy:

1. rekisteröity on antanut suostumuksensa henkilötietojensa käsittelyyn yhtä tai useampaa erityistä tarkoitusta varten
2. käsittely on tarpeen sellaisen sopimuksen täytäntöön panemiseksi, jossa rekisteröity on osapuolena, tai sopimuksen tekemistä edeltävien toimenpiteiden toteuttamiseksi rekisteröidyn pyynnöstä;
3. käsittely on tarpeen rekisterinpitäjän lakisääteisen velvoitteen noudattamiseksi;
4. käsittely on tarpeen rekisteröidyn tai toisen luonnollisen henkilön elintärkeiden etujen suojaamiseksi;
5. käsittely on tarpeen yleistä etua koskevan tehtävän suorittamiseksi tai rekisterinpitäjälle kuuluvan julkisen vallan käyttämiseksi;
6. käsittely on tarpeen rekisterinpitäjän tai kolmannen osapuolen oikeutettujen etujen toteuttamiseksi, paitsi milloin henkilötietojen suojaa edellyttävät rekisteröidyn edut tai perusoikeudet ja -vapaudet syrjäyttävät tällaiset edut, erityisesti jos rekisteröity on lapsi.

Ensimmäisen alakohdan f alakohtaa ei sovelleta tietojenkäsittelyyn, jota viranomaiset suorittavat tehtäviensä yhteydessä.

2. Jäsenvaltiot voivat pitää voimassa tai ottaa käyttöön yksityiskohtaisempia säännöksiä tässä asetuksessa vahvistettujen sääntöjen soveltamisen mukauttamiseksi sellaisessa käsittelyssä, joka tehdään 1 kohdan c ja e alakohdan noudattamiseksi määrittämällä täsmällisemmin tietojenkäsittely- ja muita toimenpiteitä koskevat erityiset vaatimukset, joilla varmistetaan laillinen ja asianmukainen tietojenkäsittely muun muassa muissa erityisissä käsittelytilanteissa siten kuin IX luvussa säädetään.

3. Edellä olevan 1 kohdan c ja e alakohdassa tarkoitetun käsittelyn perustasta on säädettävä joko

1. unionin oikeudessa; tai
2. rekisterinpitäjään sovellettavassa jäsenvaltion lainsäädännössä.

Käsittelyn tarkoitus määritellään kyseisessä käsittelyn oikeusperusteessa tai, 1 kohdan e alakohdassa tarkoitetussa käsittelyssä, sen on oltava tarpeen yleistä etua koskevan tehtävän suorittamiseksi tai rekisterinpitäjälle kuuluvan julkisen vallan käyttämiseksi. Kyseinen käsittelyn oikeusperuste voi sisältää erityisiä säännöksiä, joilla mukautetaan tämän asetuksen sääntöjen soveltamista, muun muassa: yleisiä edellytyksiä, jotka koskevat rekisterinpitäjän suorittaman tietojenkäsittelyn lainmukaisuutta; käsiteltävien tietojen tyyppiä; asianomaisia rekisteröityjä, yhteisöjä joille ja tarkoituksia joihin henkilötietoja voidaan luovuttaa; käyttötarkoitussidonnaisuutta; säilytysaikoja; sekä käsittelytoimia ja -menettelyjä, mukaan lukien laillisen ja asianmukaisen tietojenkäsittelyn varmistamiseen tarkoitetut toimenpiteet, kuten toimenpiteet muita IX luvussa esitettyjä erityisiä tietojenkäsittelytilanteita varten. Unionin oikeuden tai jäsenvaltion lainsäädännön on täytettävä yleisen edun mukainen tavoite ja oltava oikeasuhteinen sillä tavoiteltuun oikeutettuun päämäärään nähden.

4. Jos käsittely tapahtuu muuta kuin sitä tarkoitusta varten, jonka vuoksi tiedot on kerätty, eikä käsittely perustu rekisteröidyn suostumukseen eikä unionin oikeuteen tai jäsenvaltion lainsäädäntöön, joka muodostaa demokraattisessa yhteiskunnassa välttämättömän ja oikeasuhteisen toimenpiteen 23 artiklan 1 kohdassa tarkoitettujen tavoitteiden turvaamiseksi, rekisterinpitäjän on otettava huomioon muun muassa seuraavat asiat varmistaakseen, että muuhun tarkoitukseen tapahtuva käsittely on yhteensopivaa sen tarkoituksen kanssa, jota varten tiedot alun perin kerättiin:

1. henkilötietojen keruun tarkoitusten ja aiotun myöhemmän käsittelyn tarkoitusten väliset yhteydet;
2. henkilötietojen keruun asiayhteys erityisesti rekisteröityjen ja rekisterinpitäjän välisen suhteen osalta;
3. henkilötietojen luonne, erityisesti se, käsitelläänkö erityisiä henkilötietojen ryhmiä 9 artiklan mukaisesti tai rikostuomioihin ja rikkomuksiin liittyviä henkilötietoja 10 artiklan mukaisesti;
4. aiotun myöhemmän käsittelyn mahdolliset seuraukset rekisteröidyille;
5. asianmukaisten suojatoimien, kuten salaamisen tai pseudonymisoinnin, olemassaolo

Toisaalta tietosuoja-asetuksen 4 johdantokappaleessa todetaan, että henkilötietojen käsittely olisi suunniteltava niin, että se palvelee ihmistä. Oikeus henkilötietojen suojaan ei ole absoluuttinen; sitä on tarkasteltava suhteessa sen tehtävään yhteiskunnassa ja sen on suhteellisuusperiaatteen mukaisesti oltava oikeassa suhteessa muihin perusoikeuksiin.

Yliopistot saavat käsitellä kirjoilla olevien henkilötietoja lain perusteella ja CSC puolestaan saa tutkimustietovarannossa käsitellä niitä yliopistojen lukuun, mutta tietosuoja-asetuksen kannalta ongelmakohtia käsittelyperusteen löytämisen suhteen saattaa olla kolme:

1) kun henkilötietoon liitetään jokin tieto, jota henkilö ei ole itse antanut, esimerkkinä on käytetty affiliaatiota

2) jos osa henkilötiedoista saadaan esimerkiksi säätiöiltä henkilön suostumuksen perusteella, suostumuksen on mahdollistettava henkilötietojen jatkokäyttö ja jakaminen avoimesti

3) tutkimustietovarannon henkilörekisterin avaaminen harvestoitavaksi. Onko meillä oikeus antaa ja onko vastaanottajalla oikeus vastaanottaa?

EU:n alueella on yli tuhat erilaista tieteen viitetietoja keräävää palvelua (Ne voivat tietysti kerätä myös esim. julkaisuja, tutkimusaineistoja tms.). Viitetietoja harvestoidaan ristiin rastiin, myös kolmansiin maihin, eikä tämä toistaiseksi ole muodostunut ongelmaksi.

## Muu tekijätietojen käsittelyyn vaikuttava lainsäädäntö

Tietosuoja ei ole ainoa säädös, joka kohdistuu henkilötietoihin. Tekijän nimeen kohdistuu myös tekijänoikeuslain 3 §:n suoja, eli kun teoksesta valmistetaan kappale tai teos kokonaan tai osittain saatetaan yleisön saataviin, on tekijä ilmoitettava sillä tavoin kuin hyvä tapa vaatii. Tieteessä hyvään tapaan saattaisi kuulua muun muassa tekijän affiliaation ilmoittaminen viitetiedoissa. Tekijänoikeuslaki on silloin erityislain suhteessa tietosuojalakiin ja saa etusijan soveltamisjärjestyksessä.

## Muut EU-alueen palvelut

Koska kaikkien EU-maiden olisi toimittava samojen ohjeiden mukaan, myös tutkimustietovarannossa tulee pyrkiä samaan linjaan. EU:n alueella on yli tuhat erilaista tieteen viitetietoja keräävää palvelua. (Ne voivat tietysti kerätä myös esim. julkaisuja, tutkimusaineistoja tms.). Viitetietoja harvestoidaan ristiin rastiin, myös kolmansiin maihin, eikä tämä ole vielä muodostunut ongelmaksi.

#### OpenAIRE

Keskeinen palvelu Euroopassa on EU:n rahoittama OpenAIRE-palvelu, jonka tavoitteena on edistää tutkimustulosten avointa saatavuutta. Se kokoaa eri maista noin 2 700 julkaisu- ja muusta arkistosta yli 22 miljoonan julkaisun ja noin 0,5 miljoonan tutkimusaineiston metatiedot. OpenAIREn tietosisältö on haravoitavissa kokonaan tai osittain koneluettavan rajapinnan kautta (<http://api.openaire.eu/>). Lainsäädäntönäkökulmaa OpenAIREssa käsitelty ainoastaan tekijänoikeuksien osalta ([OpenAIRE legal study](http://webdoc.sub.gwdg.de/univerlag/2013/legalstudy.pdf)), mutta tekijätietojen luovuttamista ei ole nostettu ongelmana.

#### NARCIS

Alankomaiden **NARCIS** kokoaa julkaisujen, aineistojen ja tutkimushankkeiden metatietoja ja ne ovat sen portaalissa avoimesti saatavilla: <https://www.narcis.nl/>

NARCISilmoittaa käyttöehdoissaan seuraavaa (https://www.narcis.nl/terms/Language/en):

Downloading of the information in NARCIS, or copying it in any other fashion, is permitted. One exception applies to this regulation. Re-use of information in the sections "Persons" and "Organizations" is restricted. Due to contractual and legal reasons information on persons and organizations presented in these sections may not be crawled, or copied in any other fashion. Re-use of information in the sections "Persons" and "Organizations" is allowed solely and in very small measure for personal exercise, study or use.

#### SwePub

Det bibliografiska SwePub-datat är fritt för åtkomst via [OAI PMH, SRU och Xsearch](http://www.kb.se/libris/SwePub/Technical-information/). <http://swepub.kb.se/help.jsp> Ruotsin Swecrisiin kerätään ORCID-numero, sähköpostiosoite, nimet, sukupuolitieto ja personnummer (henkilötunnus), mutta julkisesti näytetään vain etu- ja sukunimi.

#### Forskningsdatabasen (Tanska)

Verkkosivuilla ei muita viittauksia tietosuojaan kuin: "We do not gather personal data." Tämä viittaa kuitenkin sivuilla kävijöistä kerättyihin tietoihin. Tietokannan tiedot ovat vapaasti haravoitavissa.

#### Vlaams Academisch Bibliografisch Bestand voor de Sociale en Humane Wetenschappen VABB-SHW (Belgia)

Ei tietoa, miten henkilötietoja käsitellään tai onko harvestoitavissa.

#### RIAN.ie Open Access, Irlanti

Sisältää selkeät käyttöehdot <http://rian.ie/en/static/Legal>, : The IUA authorises you to view and download the materials on the Site, and to access materials using metadata links on this site, solely for your non-commercial use.

#### Opin Visindi (Islanti)

Terms of use ja muu ohjeistava materiaali toistaiseksi vain islanniksi. Englanninkieliset sivut työnalla. Ei tietoa henkilötietojen käsittelystä tai harvestoinnista.

## Johtopäätöksiä

Tiedeyhteisössä on yleisenä käytäntönä erityisesti julkaisuihin liitettyjen tekijätietojen levittäminen, eikä asiaa ole koettu ongelmaksi. Tutkimusorganisaatioiden keräämät julkaisutiedot tekijätietoineen ovat pääsääntöisesti jo verkossa julkisesti saatavilla. Tiedot löytyvät yleensä sekä julkaisun kustantajan palvelussa että kansainvälisistä viitetietokannoista (esim. CrossRef, Web of Science, Scopus, Google Scholar, PubMed) tai kotimaisista kirjastoviitetietokannoista (esim. Arto, Melinda). Useista näistä tietokannoista tiedot ovat myös siirrettävissä mihin tahansa muuhun järjestelmään koneluettavan rajapinnan avulla.

Lisäksi useimmat tietoja tuottavat organisaatiot julkaisevat omien tutkijoidensa julkaisutiedot verkkosivuillaan (esim. HY:n TUHAT-järjestelmä <https://tuhat.halvi.helsinki.fi/portal/fi/> )

Toisaalta yleinen tietosuoja-asetus ja kansallisen tietosuojalain valmistelutyö antavat tilaisuuden pohtia myös tiedeyhteisön käytänteitä ja punnita tietosuojan tarvetta tieteen vapautta ja siihen kuuluvaa julkisuusvaatimusta vasten.

Euroopan unionin alueella **OpenAIREn** käytäntö on hyvä lähtökohta, jota soveltamalla voitaisiin luottaa siihen, ettei erityissäännöstä asiasta tarvita. Palveluiden osalta tilannetta kuitenkin seurataan jatkuvasti ja tiedot päivitetään tutkimustietovaranto-wikisivuilla olevaan kansainväliseen benchmark-katsaukseen <https://wiki.eduuni.fi/x/sC9iAg>.

Tilannetta tarkastellaan uudelleen vuoden päästä ja otetaan huomioon tapahtunut kehitys muissa palveluissa.