

Marjut Salokannel
OTT, Dosentti, Helsingin yliopisto

**TUTKIMUKSEN PERUSTEET:
TUTKIMUSDATAN JA MUUN PRIMÄÄRIN
TUTKIMUSAINEISTON SAATAVUUTEEN
LIITTYVÄ OIKEUDELLINEN SÄÄNTELY**

EDILEX

Artikkeli
Julkaistu Edilexissä 21.11.2012
www.edilex.fi/lakikirjasto/9054

Marjut Salokannel

OTT, Dosentti, Helsingin yliopisto

**TUTKIMUKSEN PERUSTEET: TUTKIMUSDATAN JA MUUN PRIMÄÄRIN
TUTKIMUSAINEISTON SAATAVUUTEEN LIITTYVÄ OIKEUDELLINEN
SÄÄNTELY**

1 Tieteen vapaus ja itsekritiikki

Tieteellisen tutkimuksen riippumattomuus on Suomessa turvattu perustuslaissa (731/1999), jonka mukaan tieteen, taiteen ja ylimmän opetuksen vapaus on turvattu (16 § 3 mom.). Lain tasolla tämä on vahvistettu yliopistolaisissa (558/2009), jossa todetaan, että yliopistoissa valitsee tutkimuksen, taiteen ja opetuksen vapaus (6 §). Tutkimuksen vapaus turvaa yksittäisen tutkijan oikeutta valita tutkimusaiheensa ja metodinsa sekä oikeutta julkistaa tutkimuksensa tulokset. Tutkimuksen vapauteen kuuluu myös tutkimusyhteisön ja sen toimijoiden vastuu tutkimuksen laadusta ja itsearvioinnista tieteellisten käytäntöjen mukaisesti.¹ Tieteen vapautta koskevalla säännöksellä on kiinteä yhteys sananvapauteen.² Tieteen suuntautumisen tulee toteutua ensisijaisesti tieteellisen yhteisön itsensä harjoittaman tieteen kritiikin eikä ulkopuolisen ohjauksen kautta.³ Tutkimustulosten avoimuus ja julkisuus turvaavat tieteen itsekriittistä funktiota mahdollistamalla tutkimustulosten ulkopuolisen verifioinnin tiedeyhteisön sisäisten kriteerien mukaisesti.⁴

Tutkimustulosten saatavuus ja niiden todentaminen tausta-aineiston avulla on yksi tieteellisen tutkimuksen peruspilareita. Teknologian mukanaan tuoma uusi tutkimuksen paradigma, e-science, on entisestään korostanut avoimen tieteellisen tiedon merkitystä. Tietoa pitää voida hakea, yhdistää ja muokata yli tieteen rajojen, jolloin uusi tieto rakentuu uusista näkökulmista entisen päälle. OECD:n ohjeet tieteellisen tiedon saatavuudesta olivat yksi ensimmäisistä institutionaalisella tasolla annetuista kansainvälisistä julkilausumista avoimen tieteellisen tiedon, tässä tapauksessa tutkimusdatan puolesta. Sen jälkeen vastaavia suosituksia on annettu muun muassa EU:ssa ja yksittäisissä jäsenvaltioissa. Tutkimusrahoittajat edellyttävät laajenevassa määrin niiden rahoittamien tutkimusten tulosten ja niiden taustadatan avointa saatavuutta jatkotutkimuksen käyttöön.⁵ Myös yliopistot edellyttävät tutki-

1 Ks. Tarkemmin *Mäenpää*, O. Yliopistolaki, WSOYPro 2009, Helsinki, s. 76 ss.

2 HE 309/1993 vp.

3 Ibid.

4 Ks. *Mäenpää*, Yliopistolaki, s. 98 ss.

5 Ks. tästä tarkemmin *Salokannel, Marjut*, Julkisin varoin tuotettu tieto tutkimuksen käyttöön, Oikeus 2011:1, s. 71 ss.

mustulosten julkisuutta niin sanotuissa Open Access -päätöksissään, joiden perusteella tutkimusjulkaisujen vertaisarvioitu käsikirjoitus on tallennettava yliopiston avoimeen arkistoon.⁶ Suomessa on myös julkaistu valtioneuvoston periaatepäätös julkisin varoin tuotetun tiedon saatavuudesta⁷ ja opetusministeriön raportti julkisin varoin tuotetun tiedon saamisesta tehokkaammin tutkimuksen käyttöön.⁸

Tiedon avoimeen saatavuuteen liittyy kuitenkin moninaisia ja usein epäselviä juridisia kysymyksiä. Epätietoisuus on ollut omiaan estämään muun muassa tutkimustiedon jatkokäyttöä. Tutkimusdataan kohdistuu toisaalta avointa saatavuutta turvaava sananvapaus ja toisaalta Suomessa poikkeuksellisen laaja tekijänoikeuspohjainen yksityinen omistusoikeus. Tämän lisäksi dataan voi kohdistua erilaisia salassapitoperusteita, kuten yksityisyyden suoja tai liikesalaisuus. Tässä kirjoituksessa tutkitaan, onko tutkimusdatan avoin saatavuus mahdollista toteuttaa eri lainsäädäntöjen ristipaineessa.

2 Tutkimustieto julkishyödykkeenä

Julkisin varoin tuotettua tutkimustietoa pidetään taloustieteessä julkishyödykkeenä, jonka arvo ei vähene käytettäessä vaan tieteellisen tiedon kyseessä ollen pikemminkin kasvaa.⁹ Julkishyödykkeille on ominaista, että markkinat eivät tuota niitä optimaalisella tavalla, minkä vuoksi niiden tuottaminen vaatii yhteiskunnan tukea. Tieteellistä tietoa voidaan pitää jopa globaalina julkishyödykkeenä, koska useimmiten se hyödyttää koko globaalia tutkijayhteisöä. Näin ollen yliopistolla tuotetun tiedon mahdollisimman laaja saatavuus hyödyt-

6 Helsingin yliopiston open access -päätöksestä ks. *Salokannel, M.*, University of Helsinki opens its research vaults: a few words on open access and the new research environment in Finland, *ScieComInfo*, Vol 4, No 23 (2008).

7 Valtioneuvoston periaatepäätös julkisen sektorin digitaalisten tietoaineistojen saatavuuden parantamisesta ja uudelleenkäytönedistämisestä 3.3.2011, www.valtioneuvosto.fi/toiminta/periaatepaatokset/periaatepaatos/fi.jsp?oid=322887&c=0&toid=135471&moid=135472.

8 ”Tieto käyttöön – Tiekartta tutkimuksen sähköisten tietoaineistojen hyödyntämisen edistämiseksi.” Saatavilla osoitteessa www.minedu.fi/export/sites/default/OPM/Julkaisut/2011/liitteet/okm04.pdf?lang=fi.

⁹ Ks. esim. *Salokannel, M.*, Global public goods and private rights: Scientific research and intellectual property rights, *Nordiskt Immateriellt Rättskydd*, No. 4/2003, s. 334 ss., *Artzberger & al.*, Promoting Access to Public Research Data for Scientific, Economic, and Social Development, *Data Science Journal*, Volume 3, 29 November 2004, s. 135.

tää koko yhteiskuntaa ja tiedon saatavuuden kokonaistaloudelliset kerrannaisvaikutukset ovat suuret.

Tieteellistä tietoa sen paremmin kuin muitakaan julkishyödykkeitä ei voida tuottaa markkinoiden avulla, minkä vuoksi niiden tuottaminen edellyttää julkista tukea. Tämä julkinen tuki ei kuitenkaan ole välttämättä taloudellista – vaikka sitäkin varmasti tarvitaan – vaan se voi toteutua myös lainsäädännöllisten toimenpiteiden kautta. Jos lainsäädäntö estää aineiston saataville saattamisen sekä sen hyödyntämisen tutkimuskäyttöön, niin lain muuttaminen näiltä osin mahdollistaa tutkimusaineistojen käytön ilman, että tällä on mitään negatiivista vaikutusta muuhun talouteen vaan pikemminkin päinvastoin: tutkimusaineistojen saattaminen laajempaan käyttöön lisää jatkotutkimusta ja tuo potentiaalisesti uusia kaupallisestikin hyödynnettäviä tuloksia. Tällä puolestaan on kauaskantoisia positiivisia vaikutuksia kansantaloudelle.

3 Tutkimusdatan määrittely

Tässä tutkimuksessa otetaan periaatteelliseksi lähtökohdaksi OECD:n julkisesti rahoitetun tutkimustiedon saatavuutta koskevista periaatteista annettu tutkimusdatan määrittely, jossa tutkimusdata määritellään seuraavasti:

“...as factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings. A research data set constitutes a systematic, partial representation of the subject being investigated.”¹⁰

Tässä tutkimuksessa tarkastellaan ainoastaan tutkijoiden itsensä tuottamaa raakadataa ja muita primäärejä tutkimusaineistoja. Tämän lisäksi tarkastelu rajoitetaan digitaalisessa muodossa olevaan aineistoon. Tämän tutkimuksen ulkopuolelle jäävät muun muassa tutkimuksessa käytetyt julkishallinnon rekisteri- ja tilastoaineistot, joiden käyttöön liittyy oma erityinen juridinen problematiikkansa. Koska tämä tutkimus keskittyy digitaalisessa muo-

¹⁰OECD principles and guidelines for access to research data from public funding, 2007. Saatavilla osoitteessa <http://www.oecd.org/dataoecd/9/61/38500813.pdf>

dossa olevan datan käyttöön liittyviin oikeudellisiin kysymyksiin, siinä ei käsitellä erilaisia näytekokoelmia, vanhoja käsikirjoituksia ja valokuvia tai muuta vastaavaa fyysisessä muodossa olevaa tutkimusmateriaalia. Tähän materiaaliin sisältyvä digitaalisessa muodossa oleva tieto, kuuluu sen sijaan tutkimuksen piiriin. Biologista materiaalia sisältyviin näytekokoelmiin liittyy puolestaan muuta sen käyttöä sääntelevää säännöstöä, johon ei tämän kirjoituksen puitteissa ole mahdollista mennä.

4 Rahoittajat ja tutkimusdatan saatavuus

Tutkimusrahoittajat ovat maailmalla olleet urauurtavassa asemassa tutkimusdatan avaamisessa jatkokäytölle. Julkiselle tutkimusrahoittajalle voidaankin luontevasti katsoa lankeavan rooli oman rahoituksensa tulosten mahdollisimman laajan hyödynnettävyyden takaajina. Yleensä rahoittajat kytkvät datan saatavuuden tutkimustulosten avoimen saatavuuden velvoitteeseen, mikä tarkoittaa sitä, että tutkimusjulkaisujen ja niiden taustadatan täytyy olla avoimesti saatavilla.

4.1 Yhdysvaltalaisten tutkimusrahoittajien käytäntöjä

Maailman suurin lääketieteellisen tutkimuksen rahoittaja the National Institutes of Health (NIH) on ollut yksi uranuurtajista julkisen tutkimustiedon avaamisessa kansalaisten käyttöön. NIH edellyttää, että yli 500 000 dollarin projektit sisältävät tutkimusdatan jakosuunnitelman ja että data talletetaan NIH:in osoittamaan tietokantaan.¹¹ Genomitutkimuksessa¹² data talletetaan täysin avoimeen kansainväliseen geenipankkiin niin, että data on tiedeyhteisön saatavilla mahdollisimman nopeasti ja laajasti.¹³

¹¹ NIH Data Sharing Policy and Implementation Guidance. Osoitteessa http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm#app

¹² Genomidatana pidetään dataa ”involving sequence and related genomic data obtained with advanced sequencing technology (e.g., medical resequencing data, sequence data from non-human species, including microorganisms, transcriptomic and epigenomic data, as well as data needed for interpretation, including associated clinical, other phenotype and metadata, such as supporting study documents and methodologies)”.

¹³ NIH:in tutkimustulosten avoin saatavuus on vahvistettu laissa Consolidated Appropriations Act of 2007 (H.R. 2764).

NIH määrittelee tutkimusdatan tallennetuksi tosiseikkoja sisältäväksi aineistoksi, joka on yleisesti hyväksytty tutkimusyhteisössä välttämättömäksi tutkimustulosten verifioimista varten.

Koko genomien kattavat assosiaatiotutkimukset (*Genome Wide Association Studies*) voivat nykytekniikalla mahdollistaa tutkimuskohteiden tunnistamisen, minkä vuoksi NIH muutti täydelliseen vapaaseen saatavuuteen perustuvaa politiikkaansa GWAS:ien suhteen vuonna 2008. Nyt se edellyttää, että data julkistetaan kahdella eri tasolla.¹⁴

NIH julkisti vuonna 1999 tutkimusvälineiden jakamista koskevan ohjeistuksensa. NIH katsoo, että ainutlaatuisten tutkimusvälineiden käytön rajoittaminen patenteilla voi estää uutta tutkimusta. Tällaisia tutkimusvälineitä ovat muun muassa erilaiset geenitestit. NIH suosittelee tutkimusvälineiden ei-yksinomaista lisensointia ja kehottaa välttämään yksinoikeuksien luovutusta.¹⁵ Jotta Suomessa toimivat tutkijat voisivat ottaa vastaan NIH:in tutkimusrahoitusta, on heidän sitouduttava toimimaan rahoitusehtojen mukaisesti. Tämä koskee muun muassa datan jakamista.

4.2 Eurooppalaisten tutkimusrahoittajien käytäntöjä

4.2.1 EU:n komissio

EU:n komissio julkaisi kesällä 2012 digitaalistrategiaan sisältyvän tiedonannon ja suositukset jäsenmaille tieteellisen tiedon avoimesta saatavuudesta. Komission mukaan julkisesti rahoitetun tutkimuksen tuloksena syntyvien tutkimusjulkaisujen ja datan avointa saatavuutta on edistettävä ja tutkimusjulkaisujen avoin saatavuus on oltava yleisenä periaatteena kaikissa EU:n tutkimusrahoitusohjelmissa. Jäsenvaltioiden on luotava selkeät toimintalinjaukset julkisesti rahoitetun tutkimusdatan levittämiseksi ja avoimeksi saatavuudeksi. Tähän toimintapolitiikan on sisällytettävä konkreettiset päämäärät ja indikaattorit niiden toteuttamisen seuraamiseksi. Niistä on käytävä ilmi toimeenpanosuunnitelma ja vastuunjako, mu-

¹⁴ Ensimmäisellä, open access -tasolla julkistetaan yhteenvetotiedot ja aggregaattigenotyyppidata. Toinen taso muodostuu kontrolloidun saatavuuden takana olevasta yksikkötason datasta (genotyypejä ja fenotyypejä). Tämä data on saatavilla tutkimuslaitosten tutkijoille, jotka tekevät Data Access -pyynnön, jonka NIH:in Data Access Committee tarkistaa.

¹⁵ Principles and guidelines for sharing biomedical research tools, 1999, Osoitteessa www.ott.nih.gov/policy/research_tool.aspx. NIH:llä on myös erikseen ohjeet malliorganismien datan jakamisesta. http://grants.nih.gov/grants/policy/model_organism/index.htm

kaan luettuna asianmukaisten lissensiointiperiaatteiden luominen. Toimintalinjauksiin on sisällytettävä niiden toteuttamisen edellyttävä rahoitussuunnitelma.

Kansallisten toimintalinjausten tuloksena julkisesti rahoitetun tutkimuksen tuloksena syntyvän tutkimusdatan on oltava julkisesti saatavilla, käytettävissä ja jälleen hyödynnettävissä digitaalisten e-infrastruktuurien välityksellä ottaen kuitenkin huomioon muun muassa yksityisyyden suojasta, liikesalaisuuksista ja valtioiden sisäisestä turvallisuudesta aiheutuvia rajoituksia. Tutkimusdatan täytyy olla helposti identifioitavissa ja linkitettävissä muuhun dataan ja julkaisuihin. Datan arviointia ja käyttöä koskevan informaation täytyy olla saatavilla sen yhteydessä. Komissio suosittelee myös, että kansallisten tutkimusrahoittajien ja -organisaatioiden täytyy luoda asianmukaiset mekanismit avustamaan ja palkitsemaan tutkimusdatan jakamista.¹⁶

4.2.2 European Research Council

EU:n komission puitteissa toimiva European Research Council (ERC) jakaa kilpailtua tutkimusrahoitusta EU:n puolesta. Se antoi joulukuussa 2007 oman tutkimustulosten ja taustadatan avoimuutta korostavan suosituksensa.¹⁷ Nyt tutkimusjulkaisujen ja niiden taustadatan avoin saatavuus on rahoituksen ehto. Taustadatan on oltava saatavilla enintään 6 kuukauden kuluessa siihen liittyvän julkaisun julkaisuajankohdasta.

4.2.3 Ison-Britannian tutkimusneuvostot (Research Councils)

Vuonna 2008 kaikki Ison-Britannian tutkimusneuvostot sitoutuivat tutkimustulosten avoimen saatavuuden periaatteeseen. RCUK:n mukaan julkisin varoin rahoitettu tutkimus täytyy olla yleisön saatavilla nyt ja tulevien sukupolvien hyväksi. Tämä tarkoittaa myös tutkimustulosten taustadataa. Julkaisuista täytyy olla joko suora pääsy taustadataan tai maininta

¹⁶ Ks. tarkemmin Commission recommendation of 17.7.2012 on access to and preservation of scientific information, C(2012) 4890 final sekä Communication from the commission to the European parliament, the Council, the European economic and social committee and the committee of the regions, Toward better access to scientific information: Boosting the benefits of public investments in research, COM (2012) 401 final.

¹⁷ ERC Scientific Council Guidelines for Open Access, December 2007.

siitä, miten taustadata on saatavilla. Datan säilytykseen sovelletaan tieteenalakohtaisia säilytysaikoja. Tätä politiikkaa vahvistettiin viimeksi kesällä 2012.¹⁸

5 Tieteelliset julkaisut ja tutkimusdatan saatavuus

Tutkimustulosten avoimuus ja julkisuus turvaavat tieteen itsekriittistä funktiota mahdollistamalla tutkimustulosten verifiointin ulkopuolisten toimesta. Perusoikeustasolla tämä nojautuu tieteen ja sananvapauden periaatteille. Tutkimustulokset on perinteisesti julkaistu tieteellisissä aikakauskirjoissa, jotka tutkimuslaitokset ovat tilanneet kirjastoihinsa ja online-käyttöön. Näin julkaisuista on julkisin varoin maksettu ensiksi tutkijan työstä ja sen jälkeen ostamalla julkaisujen käyttöoikeudet.

Tieteellinen julkaisutoiminta on kuitenkin tällä hetkellä murrosvaiheessa, jossa toimintatavat muuttuvat uuden teknologian tarjoamia uusia hyödyntämistapoja etsien. Tieteellisessä julkaisutoiminnassa erilaiset avoimet julkaisualustat ovat lyömässä itseään läpi ensisijaisina julkaisukanavina. Tämä puolestaan on nostanut esiin piilevän ristiriidan sananvapauden ja toisaalta omistusoikeuspohjaisen tekijänoikeuden välillä. Mahdollinen konflikti ei tässä yhteydessä synny niinkään tutkijoiden oikeuksien kanssa vaan lähinnä tutkimustulosten hyödyntäjien, erityisten tiedekustantajien kanssa, joita esimerkiksi tutkimusjulkaisujen käsikirjoitusten avoin saatavuus yliopistojen arkistoista voi häiritä. Myös tutkimuksessa esiintyvien henkilöiden yksityisyyden suoja voi rajoittaa tutkimustulosten avoimuutta, mistä tarkemmin jäljempänä.

Mikäli eri perusoikeuksien soveltamisessa ilmenee kitkaa, on perusoikeuksien painoarvoja punnittava keskenään kyseisessä tilanteessa. Suomessa on olemassa vakiintunut tulkintaperiaate tilanteisiin, jolloin lainsäädännön soveltamisessa voi perusoikeusnäkökulmasta riippuen syntyä ristiriitainen tulos. Tuomioistuimen on tällöin valittava perusteltavissa olevista lain tulkintavaihtoehdoista sellainen, joka parhaiten edistää perusoikeuksien tarkoituksen toteutumista ja eliminoi perustuslain kanssa ristiriitaisiksi katsottavat vaihtoehdot. Suomes-

¹⁸ Research Councils UK Policy on Access to Research Outputs. Saatavilla osoitteessa <http://www.rcuk.ac.uk/research/Pages/outputs.aspx/>.

¹⁸ HE 309/1993 vp.

sa on näin voimassa perustuslainmukainen tai perusoikeusmyönteinen tulkinta.¹⁹

Tieteen vapaus ja avoimuus on tiedejulkaisemisessa painanut enemmän kuin kustantajien kaupalliset intressit ja suuret kansainväliset tiedekustantajat ovat julkaisupolitiikassaan joutuneet ottamaan huomioon tutkimusrahoittajien ja lainsäätäjien (Yhdysvallat) vaatimukset tutkimustulosten saatavuudesta.

Luonnontieteiden alueilla johtava tieteelliset julkaisut, kuten *Nature* ja *Science*, edellyttävät, että tieteellisen kirjoituksen taustadata on talletettu tietokantaan, josta lukijoiden on mahdollista todentaa tutkimustulokset. Kun *Science* poikkesi tästä käytännöstä ihmisen genomien kartoituksen yhteydessä yksityisen yrityksen hyväksi, asiasta nousi vastalauseiden myrsky tiedemaailmassa. Yhä kasvava määrä julkaisuja edellyttää luonnontieteiden alalla julkaisun taustadatan saatavilla oloa. Datan jakamispolitiikan vahvuuden on myös katsottu korreloivan positiivisesti julkaisun tieteellisen painoarvon mittarien kanssa.²⁰ Yhteiskuntatieteellisillä ja humanistisilla aloilla datan jakamiskäytännöt tieteellisten julkaisujen julkaisuedellytyksenä ovat vasta muotoutumassa. Seuraavassa esitellään yhden keskeisimmän luonnontieteellisen alan julkaisun, *Science*-aikakauskirjan datan julkistamisvaatimukset.

Science-aikakauskirjan datapolitiikka kiteytyy seuraavasti: ”All data necessary to understand, assess, and extend the conclusions of the manuscript must be available to any reader of *Science*.” Lehti kieltää viittaamasta dataan, joka sisältyy julkaisemattomiin kirjoituksiin, mukaan luettuna ne, jotka on jo hyväksytty julkaistaviksi. Lehti kehottaa kirjoittajia julkistamaan datan joko kansainvälisessä etabloituneessa julkisessa tietokannassa, tai mikäli tällaista ei ole saatavilla, lehden omassa verkkolehdessä. Silloin, kun on kyseessä hyvin suuri datan määrä, lehti on tehnyt kirjoittajan kanssa sopimuksen, jossa kirjoittaja sitoutuu tallettamaan datan jonkin instituution verkkosivulle, ja kopio datasta jää Sciencelle. *Science* on laajentamassa datan julkaisemisvelvoitetta kattamaan myös tietokoneohjelmat, joita on käytetty datan luomisessa ja analysoinnissa.²¹

¹⁹ Ibid.

²⁰ Piwowar and Chapman, A review of journal policies for sharing research data, *Nature Proceedings* : hdl:10101/npre.2008.1700.1

²¹ Making Data Maximally Available, editorial, *Science*, vol 331, 11 Feb. 2011.

6 Julkisuuslainsäädäntö

Oikeudellisena perusteena julkisella rahoituksella tuotetun tutkimusdatan saatavuudelle voidaan pitää sananvapauteen perustuvaa kansalaisten oikeutta saada pääsy julkiseen tietoon. Tätä tukee viranomaistoiminnan julkisuutta edistävä viranomaisten asiakirjojen ja tallenteiden julkisuusperiaate. Perusoikeustasolla tieteen vapauteen kuuluu myös tutkijan oikeus julkaista tutkimustuloksensa ja alistaa ne tiedeyhteisön arvioitavaksi.²²

Perustuslain mukaan viranomaisen hallussa olevat asiakirjat ja muut tallenteet ovat julkisia, jollei niiden julkisuutta ole välttämättömien syiden vuoksi lailla erikseen rajoitettu. Jokaisella on oikeus saada tieto julkisesta asiakirjasta ja tallenteesta (12.2 §). Asiakirjojen julkisuuteen kuuluu oleellisesti, että kansalaisilla on myös tosiasiallisesti mahdollisuus saada tieto asiakirjoista. Tämä tarkoittaa muun muassa sitä, ettei tiedonsaantia rajoiteta esimerkiksi taloudellisin perustein.

Tutkimusaineistojen saatavuuden kannalta oleelliseksi viranomaistoiminnan julkisuusperiaatteen ilmentymäksi nousee sananvapauden tukeminen ja julkisten tietovarantojen hyödyntämisen mahdollistaminen.²³ Sananvapaus on julkisuusperiaatteen voimakkain tukijalka. Tutkimuksen kannalta sananvapauden päälle rakentuvat tieteellisen tutkimuksen ja opetuksen vapaus. Toisaalta taas julkisuusperiaate edellyttää, että julkisin varoin tehdyn tutkimuksen tulokset ovat julkisia.²⁴

Asiakirja- ja tallennejulkisuus toimivat Suomen perustuslaissa viranomaistoiminnan avoimuuden varmistajina. Ne antavat kansalaisille mahdollisuuden tiedon hankkimiseen viranomaistoiminnasta. Tallennejulkisuus on pääsääntö ja se luo jokaiselle oikeuden saada tiedon julkisesta asiakirjasta tai tallenteesta. Perustuslaissa vahvistetaan yleisöjulkisuuden periaate, joka antaa kaikille kansalaisille vahvan oikeuden saada tietoa viranomaisten toiminnasta. Tästä periaatteesta voidaan poiketa ainoastaan erittäin tärkeiden syiden perusteella, joita voivat olla muun muassa kansalaisten yksityisyyden suojaan liittyvät syyt.²⁵

²² Ks. esim. PeVL 13/2006 ja *Tuori, K.* Perusoikeudet, WSOYPro 2011, s. 622.

²³ Ks. tarkemmin *Mäenpää, Olli*, *Julkisuusperiaate*, WSOYpro, Helsinki, 2008, s. 3 ss. sekä *Salokannel, Marjut*: *Julkisesta datasta avoimeen dataan: julkisen datan lisensiointi*. ETLA discussion papers, no. 1277, 2012.

²⁴ Ks. *Mäenpää*, *Yliopistolaki* s. 96 ss.

²⁵ HE 309/1993 vp.

Laissa viranomaisten toiminnan julkisuudesta (621/1999) on säädetty viranomaisten hallussa olevien asiakirjojen julkisuudesta. Lain mukaan viranomaisten asiakirjat ovat julkisia, jollei laissa ole erikseen toisin säädetty. Asiakirjat on laissa määritelty hyvin laajasti sisältäen myös audiovisuaaliset tallenteet ja muut ainoastaan tietojenkäsittelyyn avulla saatavilla olevat teokset (5 §). Laissa on yksityiskohtaiset määräykset siitä, milloin asiakirjan katsotaan tulleen julkiseksi.

Lain mukaan jokaisella on oikeus saada tieto viranomaisen hallussa olevasta julkisesta asiakirjasta. Asianosaisella on oikeus saada asiaa käsittelevältä viranomaiselta tieto muunkin kuin julkisen asiakirjan sisällöstä, mikäli se voi vaikuttaa hänen asiansa käsittelyyn (9 ja 11 §).

Laissa on erikseen säädetty yleistä asiakirjojen julkisuusperiaatetta rajoittavista salassapitoperusteista. Poikkeuksen yleiseen asiakirjojen julkisuusperiaatteeseen muodostavat lain mukaan muun muassa tutkimusaineistot. Julkisuuslain 24 §:n 21 kohdan mukaan salassa pidettäviä asiakirjoja ovat

”asiakirjat, jotka koskevat opinnäytetyön tai tieteellisen tutkimuksen suunnitelmaa tai perusaineistoa taikka teknologista tai muuta kehittämistyötä tai niiden arviointia, jollei ole ilmeistä, että tiedon antaminen niistä ei aiheuta opinnäytetyön, tutkimuksen tai kehittämistyön suorittamiselle taikka niiden hyödyntämiselle tai sen asianmukaiselle arvioinnille tai tutkijalle taikka tutkimuksen tai kehittämistyön toimeksiantajalle haittaa.”

Tämän säännöksen tarkoituksena on ollut suojata lähinnä yksityisyyden suojaa ja yrityssalaisuutta silloin, kun se vaarantuisi esimerkiksi haettaessa tutkimusrahoitusta viranomaiselta. Liian varhainen tutkimustulosten julkistaminen voi myös vaarantaa patenttihakemuksen. Pääsääntöisesti tutkimusaineistot olisivat lain mukaan salaisia.

Tämä säännös ei kuitenkaan estä tutkimusaineistojen julkistamista silloin, kun se ei vaaranna tutkijan uraa, tutkimuskohteiden yksityisyyden suojaa tai muuten haittaa tutkimustiedon hyödyntämistä.²⁶ Julkisuuslain 28 § antaa mahdollisuuden tietojen luovuttamiseen tutkimuskäyttöön, ellei tämä vaaranna tutkimuskohteiden henkilötietojen suojaa. Lain mukaan viranomaisen voi, jollei lailla toisin säädetä, antaa yksittäistapauksessa luvan tietojen saa-

²⁶ HE 30/1998 vp.

miseen salassa pidettävästä asiakirjastaan muun muassa tieteellistä tutkimusta varten, jos on ilmeistä, ettei tiedon antaminen loukkaa niitä etuja, joiden suojaksi salassapitovelvollisuus on säädetty. Lupaa harkittaessa on viranomaisen huolehdittava siitä, että tieteellisen tutkimuksen vapaus turvataan.

Toisaalta kuitenkin laissa edellytetään, että jos asiakirjaan sisältyvät tiedot on annettu viranomaiselle sen suostumuksella, jonka etujen suojaamiseksi salassapitovelvollisuus on säädetty, lupaa ei saa antaa vastoin suostumuksessa tiedon käytölle ja luovutukselle asetettuja ehtoja.

Tässä yhteydessä on myös syytä korostaa, että tieteellisen tutkimuksen peruskonventioihin kuuluu, että tutkimustulokset täytyy voida verifioida tutkimusdatan tai muun taustaineiston avulla. Datan täytyy olla saatavilla tulosten todentamista varten. Tutkimusalasta ja datan luonteesta riippuu, missä muodossa ja missä laajuudessa data voi olla ulkopuolisten saatavilla. Tätä tukee tekijänoikeuslain 25 d §:n mukainen säännös, jonka mukaan tekijänoikeus ei rajoita laissa säädettyä oikeutta saada tieto julkisesta asiakirjasta. Näin tutkimusaineistoihin mahdollisesti liittyvä tekijänoikeus saa väistyä sananvapauteen perustuvan tutkimustulosten verifikaation tieltä.

7 Tutkimusdatan arkistointi

Tutkimusdatan ja muiden tutkimusaineistojen pitkäaikaissäilytystä koskevaan sääntelyyn vaikuttaa keskeisesti yliopistolaki, jossa vahvistettiin yliopistojen tieteellinen, taloudellinen ja hallinnollinen autonomia. Tämä tarkoittaa, että yliopiston sisäisestä hallinnosta päättävät laissa säädetyin rajoituksin yliopiston omat, eivätkä valtion yleiset hallintoviranomaiset.²⁷ Myös yliopistojen tutkimusaineistojen pitkäaikaissäilytyksen katsotaan kuuluvan tämän autonomian piiriin. Yliopistolakia koskevassa hallituksen esityksessä todetaan, että yliopistojen toiminnan yhteydessä syntyvien tutkimusaineistojen asianmukainen säilytys on tarpeen turvata. Tämä tehtävä jää kuitenkin yliopistojen itsensä vastuulle.²⁸ Tietohallintolainsäädännön valmistelun yhteydessä vahvistettiin, että yliopistojen tietohallinnon tarpeet liit-

²⁷ Ks. *Tuori*, s. 626, teoksessa *Perusoikeudet*, 2. Painos, WSOYPro, 2011.

²⁸ HE 7/2009 vp, s. 87.

tyvät erityisesti tutkimuksen ja opetuksen tarkoituksenmukaiseen toteuttamiseen sekä kansainvälisen tutkimusyhteistyön ylläpitämiseen, mikä puolestaan kuuluu yliopistojen itsehallinnon piiriin.²⁹

Asiakirjojen arkistointi ja julkisuuslaki ovat kiinteässä yhteydessä keskenään. Julkisuuslain mukaan julkisina pidettävät viranomaisten asiakirjat tulevat arkistolain perusteella arkistoitaviksi. Arkistolainsäädäntöä on sovellettu yliopistoissa ensisijaisesti erilaisten hallinnollisten asiakirjojen arkistointiin ja säilytykseen. Julkisuuslaissa on lähtökohdaksi otettu, että tutkimuksen perusaineisto on salaista. Tämä ei kuitenkaan sellaisenaan estä aineistojen arkistointia, vaan ainoastaan pääsyä arkistoihin. Yleensä salassapitoperuste on myös lakannut olemasta silloin, kun tutkimustulokset on julkaistu, ellei kyseessä ole tutkimuskohteiden yksityisyyden suoja. Toisaalta tutkimusaineistot ovat suojattuja tekijänoikeuslain nojalla, jolloin yksinomainen oikeus määrätä niiden julkistamisesta on oikeudenhaltijalla, joka lähtökohtaisesti on tutkija tai tutkimusryhmä, ellei asianomaisen tutkimuslaitoksen tai rahoittajan kanssa ole muusta sovittu. Mutta niin kuin olemme toistuvasti todenneet, tieteellisen työn peruskonventioihin kuuluu, että tutkimustulokset voidaan verifioida taustadatan avulla.

Käytännössä tämä tarkoittaa sitä, että yliopistot ovat vapaita itse päättämään siitä, miten niiden piirissä syntyvä tutkimusdata ja tutkimusaineistot saatetaan laajempaan käyttöön ja yleisön saataville.

8 Tekijänoikeudet ja tutkimusdata

8.1 Tekijänoikeudet perusoikeudellisesta näkökulmasta

Tekijänoikeudellista suojaa on Suomessa perinteisesti käsitelty yhtenä immateriaalioikeuksista ja asetettu se perusoikeudellisesta näkökulmasta omistusoikeuden piiriin.³⁰ Kun asiaa kuitenkin tarkastellaan laajemmasta, mannereurooppalaiseen *droit d'auteur* -traditioon pohjautuvasta näkökulmasta, tekijänoikeussuojan henkilöoikeudellinen puoli nousee myös esille. Tekijänoikeus yhdistetään tällöin ennen kaikkea taiteen mutta myös tieteen vapautteen ja

²⁹ PeVL 46/2010 vp – HE 246/2010.

³⁰ Ks. HE 309/1993 vp sekä *Länsineva, P.*, kirjassa *Perusoikeudet*, s. 573.

laajemmin sananvapauteen. Esimerkiksi ranskalaisessa tekijänoikeudellisessa keskustelussa on katsottu Kantiin pohjautuen, että tekijä kommunikoi teoksensa välityksellä yleisölle, jolloin tekijänoikeuden tehtävä on suojata teosta tekijänsä henkilön jatkeena.³¹

Tässä katsannossa tekijänoikeudellisessa suojassa olisi nimenomaisesti kysymys sananvapaudesta ja taiteellisen ilmaisun suojasta. Tämä lähestymistapa sopii erityisen hyvin myös tämän tutkimuksen aiheeseen, jossa tarkastellaan toisaalta tutkijoiden tieteellisen työn tulosten oikeudellista asemaa, ja toisaalta tieteellisen työn raaka-aineiden oikeudellista status-ta.

8.2 Kuka omistaa raakadatan?

Kun pohditaan raakadatan ja tutkimusaineistojen avointa saatavuutta, nousee esiin kysymys siitä, kuka määrää tutkimusprojektin tuloksena syntyneen raakadatan käytöstä. Tätä ennen on kuitenkin ratkaistava, voiko esimerkiksi ilmakehämittaustuloksiin tai muihin vastaaviin automaattisen teknisen toiminnan tuloksena syntyneisiin datakokoelmiin ylipäättänsä syntyä omitusoikeus, ja jos voi, mitä se pitää sisällään.

Ylimmän tieteellisen tutkimuksen riippumattomuuden periaatteesta seuraa, että yliopistoilla ei ole työnantajan ominaisuudessa oikeuksia yliopistoilla tehtyihin tutkimuksiin tai niiden taustadataan, ellei tästä ole erikseen sovittu. Tekijänoikeudet kuuluvat lähtökohtaisesti tutkijoille. Asia ei ole näin yksiselitteinen valtion sektoritutkimuslaitosten kohdalla, joissa tekijänoikeudet tutkimusjulkaisuihin – ainakin tekijän attribuutio-oikeuden osalta – kuuluvat useimmiten tekijöille. Tutkimusdata sen sijaan kuulune yleensä tutkimuslaitoksille.

Jos tutkimusprojektissa ei ole erikseen sovittu primääridatan ja tutkimuksen tuloksena mahdollisesti syntyvien aineistojen omistajuudesta ja säilyttämisestä eikä tutkimuksen mahdollisella rahoittajalla ole aineiston omistajuuteen liittyviä ehtoja, lähtökohtana on, että aineiston omistajuus kuuluu sen keränneelle tutkijalle tai tutkimusryhmälle. Jos kyseessä on esimerkiksi näyttekokoelma, omistajuus perustuu fyysisen aineiston omistajuuteen, mutta jos kyseessä on aineettomassa muodossa oleva aineisto, esimerkiksi digitaalisessa muodossa oleva data, omistajuus on tekijänoikeuspohjainen. Seuraavassa tarkastellaan, miten tekijänoikeuslainsäädäntö soveltuu tutkimusdatan omistajuuteen.

³¹ Qu'est-ce-qu'un livre?, *Emmanuel Kant*; textes de Kant et Fichte trad. et présentés par Jocelyn Benoist. Paris: Presses universitaires de France, 1995.

8.3 Primääri tutkimusdata ja muut tutkimusaineistot

8.3.1 Tekijänoikeussuojasta yleensä

Tekijänoikeudella suojataan mannereurooppalaisessa tekijänoikeusjärjestelmässä tekijän suhdetta teokseensa. Teosta suojataan tekijänsä omaperäisen ilmaisun tuloksena. Tekijänoikeus ei suojaa ideoita eikä tosiseikkoja vaan sitä muotoa, jonka tekijä ilmaisulleen antaa. Tekijänoikeus suojaa tekijän teokselle antamaa omaperäistä ilmiä. Niin sanottu tekijänoikeudellinen teoskynnys on kuitenkin hyvin matala. Käytännössä teostaso ylittyy heti, kun teoksesta jotenkin ilmenee sen tekijän omaperäisen ajattelun tulos. Teoksella tarkoitetaan tekijänoikeudellisessa mielessä täten sitä muotoa, jonka tekijä ilmaisulleen antaa.

Sinänsä keinotekoisesta erottelulla tekijänoikeudellisesti suojatun teoksen muodon ja tekijänoikeuden ulkopuolelle jäävän sisällön, teoksen ideoiden ja tosiseikkojen suojan välillä, on perustavanlaatuinen merkitys sananvapauden näkökulmasta. Kukaan ei voi omistaa ideoita tai luonnossa esiintyviä tosiseikkoja, vaan tekijänoikeudellisen suojan synty edellyttää aktiivista tekijän intellektuaalista panosta.

EU:n tuomioistuin on viimeksi tapauksessa *Infopaq* vahvistanut tekijänoikeudellisen teoskynnyksen määritelmän EU:ssa. Tuomioistuimen mukaan tekijänoikeutta voidaan soveltaa ainoastaan sellaiseen aineistoon, joka on omaperäinen siinä mielessä, että se on tekijänsä henkinen luomus.³² Alun perin tekijänoikeudellinen teoskäsite harmonisoitiin tätä määritelmää noudattaen ensin tietokoneohjelmien suojaa koskevan direktiivin (91/250/EY) ja sitten tietokantojen suojaa koskevan direktiivin (96/9/EU) yhteydessä ja viimeksi tekijänoikeudellisten suoja-aikojen harmonisoimista koskevan direktiivin (2006/116/EY) yhteydessä.³³ Teostasoa ei tämän määritelmän mukaisesti ole korkea, mutta tutkimusdatan osalta se kuitenkin rajaa ratkaisevasti suojan aluetta. Lämpökameran mittaustulokset tai laserkeilausten tuloksena syntyneet kartat eivät varmastikaan yllä tähän teostasoon.

Seuraavassa tarkastellaan lähemmin tutkimusdatan tekijänoikeudellista statusta. Kun pohdimme tutkimusdatan tekijänoikeudellista suojaa, on tehtävä ero primäärin tutkimusdatan ja sekundäärin tutkimusaineiston välillä.

³² Euroopan yhteisöjen tuomioistuin, asia C-5/08, *Infopaq International v Danske Dagblades Forening*, tiivistelmä sekä kohta 37.

³³ Ks. tarkemmin *Salokannel*, *Ownership of Rights in Audiovisual Productions*, Kluwer Law International 1997, s. 61 ss.

8.3.2 Primäärin tutkimusdatan tekijänoikeudellinen status

Primääri tutkimusdata on tutkimuksessa käytettävää raakadataa digitaalisessa, kirjallisessa, visuaalisessa tai audiovisuaalisessa muodossa. Kyseessä voi olla havainto- ja mittaustuloksia, simuloituja tuloksia, haastatteluja, kyselyaineistoja, nauhoituksia, kuvia ja rekisteritietoja tai näiden yhdistelmiä. Arvioitaessa tekijänoikeudelliselta kannalta tutkimusdatan statusta, on kysyttävä, onko kyseessä tutkijan omaperäisen ilmaisun tai tutkijan näkemyksen leimaama teos. Näin voi olla esimerkiksi audiovisuaalisen datan ja haastatteluaineistojen kohdalla, jolloin voi syntyä tekijänoikeuslain mukainen elokuvateos tai kirjallinen teos.

Sen sijaan, jos kyseessä on puhdas mittausdata tai vastaava automaattisesti kerätty data, tähän dataan sellaisenaan ei kenelläkään ole tekijänoikeutta. Sen sijaan sitä tietokantaa, mihin data on kerätty, voidaan tietyin edellytyksin suojata tekijänoikeuslain avulla joko kirjallisena teoksena, tekijänoikeuslaissa annetun erityissuojan perusteella, niin sanotun *sui generis* -tietokantasuojan tai luettelosuojan perusteella.

Seuraavassa katsotaan, miten tutkimusdataa sisältävät tietokannat tai aineistokokoelmat voisivat tulla suojatuiksi tekijänoikeuslain mukaisina teoksina tai teoskokoelmina. Sen jälkeen tarkastelemme, miten tekijänoikeuslaissa säädetty erityinen *sui generis* -voisi soveltaa tutkimusdatatietokantoihin. Viimeiseksi tarkastelemme tekijänoikeuslain mukaisen luettelosuojan soveltumista tutkimusdatatietokantoihin.

8.3.2.1 *Tekijänoikeudellisena teoksena suojattava tutkimustietokanta tai muu aineisto*

Jos tutkijan luova panos ilmenee aineiston valinnassa ja järjestämisessä tietokantaan, kyseessä on kirjallisena teoksena suojattava tietokanta. Tietokanta on tietokantadirektiivissä (96/9/EY) määritelty ”teosten, tietojen ja muiden itsenäisten aineistojen kokoelmaksi, jotka on järjestetty järjestelmällisellä tai menetelmällisellä tavalla ja joihin elektronisesti tai muulla tavoin on mahdollistettu pääsy.” Suoja ei ulotu tietokoneohjelmiin, joita on käytetty tietokannan luomiseen tai käyttämiseen (1 artikla).

Suoja kohdistuu tällöin kuitenkin ainoastaan tietokantaan sisältyvän aineiston omaperäiseen valintaan ja järjestelyyn, toisin sanottuna tietokannan rakenteeseen. Tekijänoikeussuoja ei ulotu tietokantaan sisältyviin teoksiin. Tämä vahvistettiin viimeksi Euroopan yhteisön tuomioistuimen toimesta tapauksen *Dataco et al. v Yahoo et al.* yhteydessä. Tietokanto-

jen tekijänoikeussuoja kohdistuu tietokannan rakenteeseen, joka ilmenee tietokannan sisällön omaperäisenä valitsemisena ja järjestämisenä. Mitään muita kriteereitä ei saa käyttää sen arvioinnissa, onko tietokanta tekijänoikeudellisesti suojattu teos vai ei.³⁴ Tuomioistuin toteaa nimenomaisesti, että tekijänoikeudellisesti suojatun tietokannan suojan kriteerit on täydellisesti harmonisoitu direktiivissä, mikä sulkee pois muiden mahdollisten suojan kriteereiden käytön kansallisella tasolla.³⁵

Tutkijalla on tekijänoikeus tietokantaan ja hän voi määrätä tietokannan käytöstä laissa säädettyin tavoin, mutta tekijänoikeus ei kata tietokannassa olevia mahdollisia teoksia tai muuta dataa.

Erityisesti humanistisilla tutkimusalueilla tutkimusaineistot voivat koostua olemassa olevista teoksista, jolloin niihin voisi ehkä soveltua tekijänoikeuslain 5 §:n mukainen kokoomateoksen käsite. Tällöin kokoomateoksen valmistajalle, tässä tapauksessa tutkijalle, syntyisi oikeus kokoomateokseen, mutta tämä oikeus ei rajoittaisi millään tavoin kokoomateokseen sisältyvien yksittäisten teosten käyttöä. Kokoomateoksen käsitettä ei voi soveltaa puhtaaseen tutkimusdataan, joka sinänsä ei ole tekijänoikeudellisesti suojattu.³⁶ EU:n tuomioistuimen *Dataco et al. v Yahoo et al.* päätöksen jälkeen tekijänoikeudellisina teoksina suojattaviin tutkimusaineistotietokantoihin on joka tapauksessa sovellettava ainoastaan tietokantadirektiivin asiaa koskevia säännöksiä.

8.3.2.2 *Sui generis -tietokantasuojan puitteissa suojattavat tutkimusaineistotietokannat*

Tallennettu tai mitattu data sinänsä, raakadata, ei ole suojattu tietokantana, mutta voidaanko sitä suojata *sui generis* -tietokantaoikeudella ja jos voidaan, mitä tämä edellyttää? Tietokantojen suojasta annetun direktiivin mukaan *sui generis* -suojan edellytyksenä on, että tietokannan *sisällön kerääminen, varmistaminen tai esittäminen merkitsee määrällisesti ja/tai laadullisesti huomattavaa investointia*, koko sisällön tai laadullisesti ja/tai määrällisesti arvioituna olennaisen osan kopioinnin ja/tai uudelleenkäytön estämiseksi. (7.1 artikla)

³⁴ Direktiivi 96/9/EY, 15 ja 16 johtolause. Ks. EU:n tuomioistuin tapauksessa C-604/10 Football Dataco and Others v Yahoo! UK Ltd and Others, 1.3.2012, kohta 38.

³⁵ Ibid. kohta 52.

³⁶ HE 170/1997.

Raakadataa sisältävät tutkimustietokannat ovat tyypillisesti sellaisia, joihin kerätään automaattisen mittauslaitteiston avulla tietoja esimerkiksi luonnonilmiöistä (esimerkiksi ilma-kehän kaasuja, maaperämittauksia, säätietoja), liikennetietoja (esimerkiksi automääriä) tai muuta automaattisen mittausmenetelmän avulla syntynyttä raakadataa. Tällöin data syntyy keräämisprosessin tuloksena. Euroopan yhteisöjen tuomioistuimen tulkinnan mukaan tällaisessa tapauksessa ei ole kyseessä tietokantadirektiivin mukainen *sui generis* -tietokanta, koska kyseessä on datan luominen (*creation of the data*). Päätöksessään tapauksessa *Fixtures Marketing v. Veikkaus* tuomioistuin totesi, että direktiivin tarkoittamalla investoinnin käsitteellä tarkoitetaan *voimavaroja, jotka käytetään olemassa olevien aineistojen etsintään ja niiden kokoamiseen tietokantaan*. Tämä käsite ei kata niitä voimavaroja, joita käytetään tietokannan sisällön muodostavien aineistojen luomiseen. Tuomioistuimen mukaan datan synnyttämiseen kohdistuvaa investointia ei voida pitää tietokantasuojan kannalta merkityksellisenä investointina. Ainoastaan tietojen keräämiseen suuntautuva investointi voidaan ottaa lukuun suojan soveltavuutta harkittaessa.³⁷

Sama erottelu koskee toisaalta datan mittauksen ja todentamisen ja toisaalta datan luomisen välistä suhdetta. Jos mittausta ja datan varmistamista voidaan pitää samanaikaisena datan luomisena, tämän tuloksena syntyvää tietokantaa ei voida pitää *sui generis* -suojan mukaan suojattavana tietokantana. Tuomioistuimen mukaan tietokannan sisällön varmistamiseen liittyvänä investointina voidaan pitää *voimavaroja, jotka tietokantaa muodostettaessa ja sen toiminnan aikana käytetään etsittyjen aineistojen paikkaansa pitävyyden tarkastamiseen sen varmistamiseksi, että tietokantaan sisällytettävä tai sisältyvä informaatio on luotettavaa*. Tietokannan sisällön esittämiseen liittyvällä investoinnilla puolestaan tarkoitetaan niitä voimavaroja, *joita kohdistetaan tietokantaan sisältyvien aineistojen järjestelmälliseen tai menetelmälliseen järjestämiseen ja siihen, että mahdollistetaan yksilöllinen pääsy aineistoihin*.³⁸

Tekijänoikeuslain esitöissä todetaan, että raakadataa ei voida pitää direktiivin mukaisen erityissuojan kohteena eikä se ole sellaisenaan suojattua tietokantasuojan perusteella. Esitöissä mainitaan esimerkkinä tietokantasuojan ulkopuolelle jäävistä aineistoista kaukokartoitussatelliitin ja meteorologisen havaintojärjestelmän avulla kerätty data sekä tiedostoksi

³⁷ European Court of Justice, 9th November 2004, Case C-46/02, (*Fixtures Marketing Ltd v Oy Veikkaus Ab*).

³⁸ *Ibid.* 37.

muutettu geneettinen data. Nämä eivät muodostu direktiivin tarkoittamista itsenäisistä aineistoista vaan edustavat luonnossa toisiinsa tietyssä suhteessa olevia asioita.³⁹

Sui generis -tietokantasuojan sisältö

Tekijänoikeuslain mukaan tietokantojen *sui generis* -suoja kohdistuu tietokannan keräämiseen, valmistamiseen tai esittämiseen käytettyyn huomattavaan työmäärään eikä siinä edellytetä minkäänlaista tekijän omaperäistä panosta. Suoja-aika on 15 vuotta tietokannan valmistumisesta ja se jatkuu aina uuden 15 vuoden jakson, kun tietokantaa oleellisesti muutetaan.

Tietokannan käyttöön tarvitaan tietokannan valmistajan lupa. Tietokannan valmistajalla on lain mukaan *yksinomainen oikeus määrätä työn koko sisällöstä tai sen laadullisesti tai määrällisesti arvioiden olennaisesta osasta valmistamalla siitä kappaleita ja saattamalla se yleisön saataviin* (TekijäL 49.2 §).

Laajan tietokantasuojan rajoitukset eivät käy yksiin tekijänoikeudellisesti suojattuihin teoksiin sovellettavien rajoitussäännösten kanssa. *Tietokannan yksityinen käyttö on laissa kielletty*. Tietokannasta saa suoraan lain perusteella *irrottaa laadullisesti tai määrällisesti arvioiden epäoleellisen osan ilman valmistajan suostumusta*. Tietoyhteiskuntadirektiivi olisi antanut mahdollisuuden säätää opetus- ja tutkimuskäyttöä koskevan rajoitussäännöksen tietokantojen suojaan, mutta näin ei kuitenkaan Suomessa tehty.

Rajoitussäännösten suppeus on erityisen ongelmallista Suomessa tehtävän laajaa tietomasojen louhintaa edellyttävän tutkimuksen kannalta. Tiedon louhinta olisi periaatteessa voinut olla mahdollista tietoyhteiskuntadirektiivissä annetun tilapäistä kappaleen valmistamista koskevan poikkeussäännöksen nojalla. Sen mukaan verkossa olevan aineisto selailun pitäisi olla mahdollista edellyttäen, että tilapäistä tallennusta koskevat muut edellytykset täyttyvät.⁴⁰ Valitettavasti tätä rajoitusta ei kuitenkaan sovelleta tietokoneohjelmiin eikä tietokantoihin, jotka ovat EU:n sääntelyssä oman säännösten alaisia.⁴¹ Sen sijaan eikaupalliseen tutkimustarkoitukseen tapahtuva tiedon louhinta olisi mahdollista tietokantadirektiivin mukaan, tätä rajoitussäännöstä ei Suomessa kuitenkaan ole saatettu voimaan.

³⁹ HE 179/1977.

⁴⁰ Direktiivi 2001/29/EY, 33 johtolause.

⁴¹ Ibid. 20 johtolause.

Tietokantoihin sovellettavan poikkeussäännöksen nojalla ei Suomen lain mukaan voida selata tietokantaa kokonaisuudessaan tai tehdä *data mining* -tyyppistä tutkimusta. Monet tieteenalat perustuvat pitkälle datan selaukseen ja louhintaan ja se, ettei tutkimustoiminnassa voida vapaasti selata tietokantoja, rajoittaa huomattavasti tutkimuksen teon edellytyksiä muun muassa genomitutkimuksessa, kielitieteessä, informaatiotutkimuksessa ja muilla vastaavilla aloilla, jotka perustuvat datan analysointiin ja yhdistämiseen uusilla tavoilla. Niissä jäsenvaltioissa, joissa on tieteellisen tutkimuksen hyväksi säädetty tätä koskeva rajoitussäännös, on tiedon louhinta digitaalisesta tietokannasta tieteellistä tutkimusta koskevan poikkeuksen nojalla on mahdollista. Palaamme tähän poikkeussäännökseen lähemmin jäljempänä kappaleen valmistamisoikeuden yhteydessä.

Tietokantojen omistajuuden kannalta on oleellista tietää, millä taholla on oikeus määrätä tietokannan käytöstä. Yleisten tekijänoikeudellisten periaatteiden mukaisesti oikeudet tekijänoikeudelliseen teokseen kuuluvat lähtökohtaisesti sen tekijälle. Tietokantojen osalta laissa on kuitenkin annettu erityissäännös, jonka mukaan oikeudet tietokantaan siirtyvät työsuhteessa työnantajalle, paitsi kun tietokanta on syntynyt yliopisto- tai korkeakoulututkijan itsenäisen tutkimustyön tuloksena (TekijäL 40 b §). Mitä tämä käytännössä yliopistomaailmassa tarkoittaa on epäselvää. Lähtökohtaisesti voitaneen sanoa, että oikeudet tutkimusdataa sisältäviin tietokantoihin ovat korkeakoululaitoksessa tutkijoilla tai tutkimusryhmillä ellei mahdollisissa rahoitussopimuksissa ole muusta sovittu.

8.3.2.3 *Luettelosuojan puitteissa suojattavat tutkimustietokannat*

Mikäli tutkimusdata ei tule suojatuksi tietokantana, voi sen valmistaja silti määrätä Suomessa sen käytöstä laissa annetun luettelosuojan pohjalta. Luettelosuoja on pohjoismaisessa tekijänoikeuslaissa perinteisesti ollut suojamekanismi, jolla alun perin on suojattu puhelin- ja osoiteluetteloja ja muita vastaavia luetteloja, joihin on kerätty suuri määrä tietoja. Vaikka on ilmeistä, että suojalla on alun perin tarkoitettu lähinnä analogisessa muodossa olevia luetteloita, se sopii suojaamaan myös raakadatasta muodostuvia tutkimustietokantoja.

Lain mukaan *luettelon, taulukon, ohjelman tai muun sellaisen työn, jossa on yhdisteltynä suuri määrä tietoja, valmistajalla on yksinomainen oikeus määrätä työn koko sisällöstä tai sen laadullisesti tai määrällisesti arvioiden olennaisesta osasta valmistamalla siitä kappale-*

leita ja saattamalla se yleisön saataviin (TekijäL 49 §). Suoja-aika on myös luetteloiden kohdalla 15 vuotta.

Raakadataa sisältävät tietokannat voivat tulla suojuetuiksi luettelosuojan avulla, jos katsotaan, etteivät *sui generis* -suojan edellytykset toteudu. Toisin sanottuna, jos katsotaan että tutkijat ovat luoneet eivätkä keränneet dataa, ja yksittäisellä datalla ei ole arvoa sinänsä vaan ainoastaan suhteessa toisiinsa, ei *sui generis* -suoja ei sovellu, jolloin tietokantaa suojataan luettelosuojan avulla. Riittävää luettelosuojan toteutumisen kannalta on, että tietokantaan on kerätty suuri määrä tietoja.⁴²

Valtaosa luonnontieteiden alueella kerätystä mittaus- ja havaintodatasta tulee todennäköisesti suojuetuiksi luettelosuojan kautta. Koska kyseessä on pohjoismainen suojuemuoto, tämä tarkoittaa, ettei vastaavaa suojuuta ole muissa maissa vaan suojuu kohdistuu ainoastaan Suomessa valmistettaviin luetteloihin. Käytännössä tällä seikalla ei tutkimusdatan osalta ole muuta merkitystä kuin siinä tapauksessa, että joku ulkomailta käsin luvatta käyttää oleellista osaa luettelosta lainvastaisesti. Tällöin käyttöä ei pidetä tekijänoikeuden loukkauksena vaan korkeintaan sopimusrikkomuksena Suomen ulkopuolella.

Tutkimusdataa suojataan näin ollen lähes aina jonkin tyyppisen tekijänoikeuslain mukaisen suojuen kautta. Dataan voi kohdistua myös muiden tahojen tekijänoikeuksia, mikäli siihen sisältyy esimerkiksi valokuvia, tietokoneohjelmia tai muiden valmistamia tietokantoja. Näiden teosten hyödyntämiseen tutkimuksessa on hankittava tarpeelliset luvat, mikäli laissa ei ole tähän oikeuttavaa rajoitussäännöstä (ks. tarkemmin edellä). Tietokannan tai luettelosuojan haltijan oikeus käyttää tietokantaa ei saa rajoittaa siihen sisältyvien teosten tekijöiden oikeuksia teoksiinsa.

8.4 Sekundääritutkimusaineisto

Sekundäärisestä tutkimusaineistosta voidaan puhua silloin, kun aineisto ei ole tutkijan itsensä keräämää havainto-, mittaus-, haastattelu- tai muuta vastaavaa tietoaineistoa, vaan aineiston muodostaa muiden valmistama aineisto. Näin on usein esimerkiksi historian, kirjallisuuden ja taiteiden tutkimuksessa samoin kuin oikeustieteessä. Myös lääketieteellisessä tutkimuksessa hyödynnettävät potilasasiakirjat ja rekisteritutkimuksen rekisterit ovat toissi-

⁴² Ks. HE 170/1997.

jaisesti hyödynnettävää tutkimusaineistoa siinä mielessä, että aineiston alkuperäinen synty-tarkoitus on ollut muu kuin tutkimus.

Tällöin alkuperäiseen aineistoon voi jo kohdistua tekijänoikeus, ja tutkijan on itsensä hankittava tarpeelliset luvat aineiston hyödyntämiseen tutkimuksessa, mikäli laissa ei ole tähän oikeuttavaa rajoitussäännöstä (ks. rajoitussäännöksistä tarkemmin jäljempänä).

8.5 Tekijänoikeudellisen suojan sisällöstä

Silloin kun tutkimusaineistoa suojataan tekijänoikeudellisena teoksena – lähinnä silloin, kun on kyseessä teoksena suojattava tietokanta – kyseisen teoksen tekijällä on kaikki tekijänoikeuslain mukanaan tuomat yksinoikeudet määrätä teoksen käytöstä. Edellä olemme jo tarkastelleet eräiden lähioikeuksien mukaan suojattavien tietokantojen suojan laajuutta ja jäljempänä tarkastelemme ensiksi tekijänoikeudellisia yksinoikeuksia ja tärkeimpiä rajoitussäännöksiä, minkä jälkeen vuorossa ovat eräät varsinaisille tekijänoikeuksille ja lähioikeuksille yhteiset säännökset (teknisten suojien ja oikeuksien hallinnointijärjestelmien suoja ja tutkimuspoikkeus).

Jos tutkimusdataa suojataan varsinaisen tekijänoikeudellisen suojan avulla, tekijänoikeus antaa tekijälle yksinoikeuden määrätä teoksen käytöstä laissa mainituin tavoin. Tekijän moraaliset oikeudet puolestaan suojaavat tekijän henkilökohtaista suhdetta teokseensa riippumatta siitä, ovatko teoksen taloudelliset hyödyntämisoikeudet luovutettu edelleen.

8.5.1 Tekijän moraaliset oikeudet

Tekijän niin sanotuilla moraalisisilla oikeuksilla suojataan mannereurooppalaisessa tekijänoikeusjärjestelmässä tekijän erityistä suhdetta teokseensa. Moraalisia oikeuksia ei ole *sui generis* -tietokantojen valmistajilla tai luettelosuojan haltijoilla. Sen sijaan tekijänoikeudellisena teoksena suojattavan tietokannan valmistajalla tällainen oikeus ainakin teoreettisesti voi olla.

Moraaliset oikeudet antavat tekijälle oikeuden meriittiin; ne suojaavat tekijän oikeutta tulla mainituksi hyvän tavan mukaisesti aina kun hänen teostaan tai sen osaa käytetään. Oikeus tekijämeriittiin on akateemisessa maailmassa ehkä keskeisin tekijänoikeuden komponentti ja ainoa todellinen tekijänoikeudesta nouseva kannustinmekanismi.

Tämän vuoksi on ensisijaisen tärkeää, että kaikissa eri verkkolevitysyhteyksissä pidetään huolta siitä, että tekijän nimi pysyvästi säilyy teoksen tai sen osan yhteydessä erilaisten pysyvien tunnisteiden avulla. Lisenssiehdoissa on myös velvoitettava jatkokäyttäjät aina ilmoittamaan teoksen lähteen, toisin sanoen tekijän nimen ja sen mistä teoksen löytää (esimerkiksi yliopiston digitaalinen arkisto tai muu pysyvä verkko-osoite). Tekijän nimi on myös aina oltava pysyvästi koodattuna teoksen metatiedoissa.

Tutkimusdatan ollessa kyseessä datan metatiedoissa on oltava pysyvä alkuperämerkintä, josta käy ilmi, mistä data on peräisin ja kuka tai mikä taho mahdollisesti omistaa datan oikeudet. Suositeltavaa on, että aina dataa käytettäessä, sen alkuperä ilmoitetaan. Tekijänoikeustiedot sisältävät yleensä myös datan julkaisupäivämäärän.⁴³ Datan julkaiseminen ja sen siteeraus voivat toimia myös meritoitumistekijänä tutkijauralla.

Moraalisten oikeuksien toinen komponentti, niin sanottu respektioikeus, suojaa tekijän kunniaa ja mainetta. Silloin kun on kyseessä tekijänoikeudellisena teoksena suojattu tutkimusaineisto, respektioikeus suojaa aineiston integriteettiä; se voi turvata datan aitouden ja koskemattomuuden.

Moraalisia oikeuksia ei voi lain mukaan luovuttaa kuin tarkasti rajoitettua tarkoitusta varten. Moraalisten oikeuksien kokonaisluovutuksen sisältävä sopimus on näin ollen mitätön suoraan lain perusteella.

8.5.2 Taloudelliset oikeudet ja niiden rajoitukset

Tekijän taloudelliset oikeudet on laissa jaettu toisaalta tekijän oikeuteen määrätä teoksen kappaleen valmistamisesta ja toisaalta oikeuteen määrätä teoksen saattamisesta yleisön saataville. Tekijän tai hänen oikeudenhaltijansa suostumus vaaditaan silloin, kun teosta käytetään näillä tavoin, ellei lain rajoitussäännöksissä ole käyttöä sallittu. Tekijän määräysoikeus ulottuu lain mukaan teokseen muuttamattomana tai muutettuna, käännöksenä tai muunnellamana, toisessa kirjallisuus- tai taidelajissa taikka toista tekotapaa käyttäen (Tekijäl 2 §).

Ongelmaksi on teknologian kehityksen mukana osoittautunut laajana määritelty teoksen kappaleen valmistamisoikeus, jos sitä lähdetään tulkitsemaan sananmukaisesti niin, että se kattaa kaikki kopiointitapahtumat riippumatta siitä, onko niillä itsenäistä taloudellista mer-

⁴³ Näin on esimerkiksi Creative Common Attribution -lisensseissä, ks. creativecommons.org.

kitystä. Lain mukaan teoksen kappaleen valmistamisena pidetään kaikkea teoksen kopioimista kokonaan tai osittain, suoraan tai välillisesti, tilapäisesti tai pysyvästi sekä millä keinolla ja missä muodossa tahansa (TekijäL 2 §).

Oikeuskirjallisuudessa on suositeltu kappaleenvalmistamisoikeuden normatiivista tulkintaa, jolloin se kattaisi ainoastaan taloudelliselta kannalta merkittävän kopioinnin.⁴⁴ EU:n tuomioistuin on kuitenkin oikeuskäytännössään systemaattisesti todennut, että kappaleenvalmistamisoikeudelle on annettava laaja merkitys ja sen rajoituksia on tulkittava suppeasti.⁴⁵ Seuraavassa tarkastellaan tutkimuskäytön kannalta yhä tärkeämmäksi nousevaa tutkimusdatan ja muun aineiston käyttöön liittyvää aineiston käyttöä, tiedon louhintaa.

8.5.2.1 Tiedon louhinta

Parhaimmillaan uuden teknologian uutta tietoa luovat sovellukset nousevat esille tiedon louhinnassa, jolloin systemaattisesti tuotetaan, analysoidaan ja yhdistetään eri lähteistä uutta tietoa tavalla, joka ei ole aikaisemmin ollut mahdollista. Tutkimusdatan avoin saatavuus muodostaa erinomaisen alustan eri tiedon louhinnan mekanismeille. Yhdysvalloissa tiedon louhinta on useimmiten mahdollista ilman tekijänoikeudellisia rajoitteita tekijänoikeuslakiin sisältyvän *fair use* -periaatteen avulla. Sen sijaan mannereurooppalainen tiukan oikeusperusteinen tekijänoikeuslaki suhtautuu hyvin tiukasti ilman oikeudenhaltijoiden lupaa tapahtuvaan tiedon louhintaan riippumatta siitä, tapahtuuko se tutkimus- vai kaupallisiin tarkoituksiin tai onko sillä mitään vaikutusta teosten varsinaiseen hyödyntämiseen.

EU- tuomioistuin on hiljattain ottanut kantaan tiedon louhintaan kahdessa *Infopaq*-ratkaisussaan, jotka koskivat teoksen kappaleen valmistamisoikeuden soveltamista tiedon louhintaan sanomalehdistä. Näissä tapauksissa tuomioistuin on pohtinut teoksen kappaleen valmistamisoikeuden ulottuvuutta ja sen suhdetta direktiivissä 2001/29/EU säädettyyn pakollista väliaikaista kappaleen valmistamista koskevaan poikkeussäännökseen. Direktiivin mukaan kappaleen valmistamisoikeus ei koske sellaista tilapäistä kappaleen valmistamista, joka on väliaikaista tai satunnaista sekä erottamaton ja välttämätön osa teknistä prosessia ja jonka ainoa tarkoitus on mahdollistaa välittäjän toimesta tapahtuva teoksen tai muun aineis-

⁴⁴ *Van Eechoud et al.*: Harmonizing European Copyright Law, s. 111.

⁴⁵ Viimeksi mm. tapauksissa *Infopaq I* ja *II*.

ton siirto verkossa kolmansien osapuolien välillä, tai sen laillinen käyttö ja jolla ei ole itse- näistä taloudellista merkitystä.

Teoksen skannaamista ja tallentamista digitaalisessa muodossa on näin ollen pidettävä sen kappaleen valmistamisena. Tuomioistuimen mukaan automaattisen tietojen louhinnan tu- loksena voi syntyä tekijänoikeudellisesti suojattuja teoksia, jos kyseessä on 11 sanan sarja, joka ilmentää tekijänsä omaperäistä henkistä luomistyötä. Kansallisen tuomioistuimen on yksittäistapauksessa arvioitava ylittyykö teostasoa tällaisessa tapauksessa.⁴⁶

Paradoksaalista kyllä, *sui generis* -tietokannan valmistajan oikeuteen ei sovelleta direktiivin 5.1 artiklan väliaikaista kappaleenvalmistamista koskevaa rajoitussäännöstä vaan direktii- vin 7 artiklan mukaista *sui generis* -oikeutta koskevaa kopioinnin käsitettä. *Sui generis* - oikeuteen sisältyvä kopiointi (*extraction*) kattaa sekä suojatun tietokannan sisällön laadulli- sesti tai määrällisesti olennaisen osan siirtämisen että sellaisten epäolennaisten osien siirtä- misen, joka jatkuvuutensa ja järjestelmällisyytensä vuoksi johtaa tietokannan sisällön olen- naisen osan uudelleen muodostamiseen.⁴⁷ Se täytyvätkö nämä edellytykset, kun tietoa louhitaan järjestelmällisesti *sui generis* oikeudella suojatusta tietokannasta, on ratkaistava tapauskohtaisesti.

Teknologioiden konvergoitumisen tuloksena on yhä vaikeampi tehdä eroa sen välillä, mil- loin on kyseessä teosten väliaikainen selailu tai kopiointi ja milloin tämä toimi kohdistuu tietokantoihin. Nykyisen sääntelyn mukaan tekijänoikeudet tietoyhteiskunnassa -direktiivin 5.1 artiklan väliaikaista kappaleen valmistamista koskeva rajoitussäännös koskee ainoas- taan mainitun direktiivin mukaisia oikeuksia eikä tietokantojen kopiointia. Epäyhdenmu- kainen sääntely ei ole omiaan lisäämään oikeusvarmuutta digitaalisessa teosten hyödyntä- misessä. Komissio onkin todennut tekijänoikeudellista sääntelyä koskevassa arvioinnissaan, että tekijänoikeudet pitäisi ulottaa koskemaan myös vastaavia tietokantadirektiivin mukai- sia oikeuksia.⁴⁸

Tässä yhteydessä on huomattava, että yksityistä käyttöä koskevaa rajoitussäännöstä ei voi- da soveltaa korkeakoululaitoksessa tehtävään ammattimaiseen tutkimus- ja opetustyöhön.

⁴⁶ *Infopaq International v Danske Dagblades Forening*, Yhteisöjen tuomioistuin 16.7.2009, asia C-5/08.

⁴⁷ *Directmedia Publishing GmbH v Albert-Ludwigs-Universität Freiburg*, Yhteisöjen tuomioistuin 9.10.2008, asia C-304/07.

⁴⁸ Commission staff working paper on the review of the EC legal framework in the field of copyright and related rights, 19.7.2004, SEC(2004) 995.

Jotta tekijänoikeudellisesti suojatun aineiston käyttö ammattimaisessa tutkimustoiminnassa olisi mahdollista, tähän on oltava joko oikeudenhaltijoiden suostumus tai laissa oleva erityinen rajoitusperuste.

Tiedon louhintaa voisi periaatteessa helpottaa myös laissa annettu mahdollisuus siteerata julkistettua teosta. Julkistetusta teoksesta on lupa hyvän tavan mukaisesti ottaa lainauksia tarkoituksen edellyttämässä laajuudessa. Lainauksen lähde eli tekijän nimi ja julkaisija tai julkistuspaikka (esimerkiksi nettisivu), on aina mainittava. Siteeraamisoikeuden soveltaminen tiedon louhinnassa ja tutkimusdatan käytössä voi kuitenkin olla rajoitettua.

Teoksen muunteleminen edellyttää alkuperäisen teoksen tekijän suostumusta (TekijäL 4.1 §). Sen sijaan, jos joku tekee kokonaan uuden teoksen alkuperäisen teoksen pohjalta siten, että alkuperäinen teos ei enää ole tunnistettavissa, hän ei tarvitse tähän alkuperäisen teoksen tekijän lupaa. Tästä niin sanotusta *vapaasta muuntelusta* säädetään laissa seuraavasti: ”jos joku teosta vapaasti muuttaen on saanut aikaan uuden ja itsenäisen teoksen, ei hänen tekijänoikeutensa riipu tekijänoikeudesta alkuperäisteokseen” (TekijäL 4.2 §).

Tutkimusdatan hyödyntämisen osalta tämän säännöksen soveltaminen voisi tulla kysymykseen esimerkiksi silloin, jos kieliteknologian alan tutkija käyttää hyväkseen digitaalisessa muodossa olevaa aineistoa niin, että hän analysoi tiettyjen lauserakenteiden tai sanojen esiintymistä, ja tämän käytön pohjalta syntyy kokonaan uusi itsenäinen teos, kun tutkimuksen tulokset raportoidaan. Tämän tyyppiseen käyttöön voisi soveltua myös lain siteerausta koskeva säännös.

8.5.2.2 *Teknisten suojien ja oikeuksien hallinnointitietojen suoja*

EU:ssa on luotu ylimääräinen ei-tekijänoikeudellinen suojakerros varsinaisten tekijänoikeuksien suojan päälle, mikä antaa oikeudellisen suojan teoksiin asetetuille teknisille suojauksille tai oikeuksien hallinnointitiedoille. Tekijän luvalla teokseen asetettuja teknisiä suojuksia ei saa kiertää tai poistaa.⁴⁹ Nämä määräykset suojaavat tutkimusdataa tapauksissa,

⁴⁹ Lain mukaan suojatun teoksen suojana olevaa tehokasta teknistä toimenpidettä, jonka teoksen tekijä tai joku muu tekijän luvalla teosta yleisön saataviin saattaessaan on teoksen suojaksi asettanut, ei saa kiertää. Tehokkaan teknisen toimenpiteen kiertämisen mahdollistavia tai kiertämistä helpottavia laitteita, tuotteita tai osia ei saa valmistaa tai tuoda maahan yleisölle levittämistä varten tai Suomen alueelle kolmanteen maahan kuljetettavaksi, levittää yleisölle, myydä, vuokrata, mainostaa myytäväksi tai vuokrattavaksi eikä pitää hallussa kaupallisessa tarkoituksessa. Tehokkaan teknisen toimenpiteen kiertämisen mahdollistavia tai sitä helpottavia palveluja ei liioin saa tarjota yleisölle (TekijäL 56 e §).

jolloin data on saatavilla ainoastaan autentikointijärjestelmän takana. Tällaisen suojan kiertäminen tai murtaminen on tekijänoikeuslaissa sanktioitu teko. Yksityinen käyttö ei Suomen lain mukaan ole mahdollista ohi teoksessa olevan teknisen suojauksen.

Vastaavasti teokseen tekijän luvalla asetettuja oikeuksien hallintatietoja ei saa muuttaa tai poistaa.⁵⁰ Tämä säännös suojaa esimerkiksi tutkimusdatan metatietoja, joita ei saa ilman asianmukaista auktorisointia muuttaa tai poistaa.

8.5.3 Tulevaisuuden muutostarpeet: tekijänoikeudellisten teosten tutkimuskäyttö

Kuten olemme nähneet, Suomessa tekijänoikeuslaki antaa eurooppalaisessakin viitekehyydessä erittäin laajan suojan erilaista dataa sisältäville tietokannoille. Lain mukainen luettelosuoja kattaa käytännöllisesti katsottuna kaikki tutkimusdatatietokannat, joissa on vähintään enemmän dataa. Tätä laajaa suojaa tasapainottamassa ei kuitenkaan ole rajoitussäännöksiä, jotka mahdollistaisivat turvallisen tietokantojen tutkimuskäytön. Suuressa osassa muista EU:n jäsenvaltioista on laissa annettu tieteellisen tutkimuksen hyväksi rajoitussäännös, jonka myös niin sanottu tekijänoikeudet tietoyhteiskunnassa -direktiivi mahdollistaa.⁵¹ Koska jäsenmailla ei ollut velvollisuutta säätää direktiivissä olevia rajoitussäännöksiä, Suomi on jättänyt tämän rajoitussäännöksen voimaan saattamatta.

Jäsenmailla on direktiivin mukaan mahdollisuus säätää rajoitussäännös muun muassa *tekijänoikeudellisten teosten käyttämiseen tieteellisen tutkimuksen hyväksi siltä osin kun käyttö on perusteltua tavoiteltavan ei-kaupallisen tarkoituksen vuoksi ja edellyttäen, että lähde ja tekijän nimi mainitaan, ellei se ole mahdotonta* (artikla 5.3). Tutkimusta pidetään komission mukaan ei-kaupallisena myös silloin, kun se tehdään osittain kolmannen, yksityisen sektorin osapuolen rahoituksella, jos tutkimus itsessään on ei-kaupallista. Tästä esimerkkinä mainitaan, että tutkimus tehdään julkisessa tutkimuslaitoksessa osittain yksityisellä rahoituksella.

⁵⁰Lain mukaan suojatun teoksen kappaleessa olevia taikka teosta yleisölle välitettäessä ilmeneviä sellaisia oikeuksien sähköisiä hallinnointitietoja, joilla tunnistetaan teos, tekijä tai oikeuksien muu haltija taikka jotka ovat tietoja teoksen käyttöehdoista, ei saa poistaa tai muuttaa. Suojatun teoksen kappaletta ei saa levittää yleisölle tai tuoda maahan yleisölle levittämistä varten taikka teosta välittää yleisölle siten, että oikeuksien sähköiset hallinnointitiedot on luvatta poistettu teoksesta tai että niitä on luvatta muutettu (TekijäL 56 f §).

⁵¹Euroopan parlamentin ja neuvoston direktiivi 2001/29/EY, tekijänoikeuden ja lähioikeuksien tiettyjen piirteiden yhdenmukaistamisesta tietoyhteiskunnassa, 22.5.2000.

Lisäksi yksittäisen rajoitussäännöksen toteuttaminen edellyttää, että se täyttää niin sanotun kolmen kohdan testin, jonka mukaan poikkeussäännöstä voidaan soveltaa vain tietyissä erityistapauksissa, jotka eivät ole ristiriidassa teoksen tai muun aineiston tavanomaisen hyödyntämisen kanssa eivätkä kohtuuttomasti haittaa oikeudenhaltijan oikeutettuja etuja (artikla 5.5).

Komission mukaan tämä rajoitussäännös on implementoitu hyvin kapeasti tai tavalla, ettei sen tarkoitus ole toteutunut. Tästä komissio mainitsee esimerkkinä Pohjoismaat.⁵² Komission mukaan nykyinen tilanne ei ole tyydyttävä eikä mahdollista yli rajojen tapahtuvaa verkopohjaista opetus- tai tutkimustoimintaa. Verkkomaailmassa ylikansallisten tutkimusverkostojen luominen muodostaa perustavanlaatuisen osan tiedon vapaasta liikkuvuudesta EU:ssa, eikä jäsenmaiden epäyhtenäisen tekijänoikeudellisen sääntelyn pitäisi olla tälle esteenä.⁵³

Yleiseurooppalaisten tutkimusinfrastruktuurien tehokas hyödyntäminen edellyttää, että tutkijat voivat käsitellä tutkimusinfrastruktuurissa sijaitsevia aineistoja samasta oikeudellisesta asemasta käsin sijaintimaasta riippumatta ja että aineiston käyttöä koko EU:n alueella sääntelevät samat säännökset. Tutkimusdatan on voitava vapaasti liikkua EU:n sisäisten tutkimusinfrastruktuurien sisällä. Toivottavaa olisi, että tämä voitaisiin toteuttaa myös kansainvälisellä tasolla. Esimerkiksi Yhdysvalloissa on yleinen *fair use* -säännös, joka antaa laajan mahdollisuuden teosten käyttämiseen muun muassa tieteellisessä tutkimuksessa. EU:ssa tämän päämäärän saavuttaminen edellyttäisi, että EU:ssa olisi *pakollisesti sovellettava, yksityiskohtainen säännös, joka mahdollistaa tietyin edellytyksin suojatun materiaalin käytämisen tutkimustarkoituksiin*.⁵⁴

⁵²Suomen tekijänoikeuslaissa on annettu yleinen sopimuslisenssisäännös, joka koskee kappaleiden valmistamista julkistetusta teoksesta opetustoiminnassa tai tieteellisessä tutkimuksessa käytettäväksi (TekijäL 14.1 §). Näin valmistettuja teoksen kappaleita voi käyttää *opetustoiminnassa tai tieteellisessä tutkimuksessa yleisölle välittämiseen muulla tavalla kuin radiossa tai televisiossa lähettämällä*. Tämä ei koske kappaleen valmistamista valokopioimalla tai vastaavin menetelmin, josta mistä on olemassa oma, valokopiointia koskevan sopimuslisenssisäännöksen (TekijäL 13 §) nojalla tehty sopimus. Tekijällä on oikeus kieltää teoksensa käyttö tämän sopimuslisenssisäännöksen nojalla paitsi silloin, kun teos on lähetetty radiossa tai televisiossa (TekijäL 14.4 §).

⁵³ Komission tiedonanto 19.10.2009 COM(2009) 532 final.

⁵⁴ EU:n komission rahoittamat ESFRI-tutkimusinfrastruktuurihankkeet ovat tehneet tämännäköisen ehdotuksen Euroopan parlamentille.

9 Patentit

Patenttien vaikutusta julkisesti rahoitetun tutkimuksen tekemiseen on tutkittu paljon, ja tulokset ovat vaihtelevia ja alasta riippuvaisia. Perustutkimustiedon patentoimiseen suhtaudutaan yleensä kielteisesti. Näyttää siltä, että suurin negatiivinen vaikutus patenteilla on biolääketieteessä, jossa erityisesti perustutkimusvälineiden patentointi on estänyt jatkotutkimuksen tekemistä. Patenttiviranomaiset ovat kuitenkin kiristäneet patenttien myöntämiskäytäntöjään, mikä on ollut omiaan helpottamaan tilannetta.

NIH:in biolääketieteellisiä tutkimusvälineitä koskevien rahoitusohjeiden mukaan patentit eivät välttämättä ole paras tapa levittää tutkimustietoa. Ohjeissa todetaan, että jos patentit ovat tarkoituksenmukaisia esimerkiksi tutkimusvälineiden tuotteistamiselle ja levittämiselle tutkimusyhteisölle, ne on lisensoitava tavalla, joka maksimoi niiden mahdollisimman laajan levityksen tutkimusyhteisölle. Euroopassa puolestaan Wellcome Trust kieltää rahoitusehdoissaan kaiken puhtaan genomitiedon patentoimisen ja edellyttää, että se on vapaasti saatavilla julkisessa tietokannassa.

Patentti antaa keksijälle tai muulle patentin haltijalle oikeuden kieltää patentin hyödyntäminen ansiotoiminnassa ilman patentin haltijan lupaa. Patentoitua keksintöä voi hyödyntää ilman lupaa silloin, kun se tapahtuu ansiotoiminnan ulkopuolella. Tässä yhteydessä on muistettava, että korkeakouluissa tapahtuvaa tutkimustoimintaa pidetään ansiotoimintana patenttilain merkityksessä.

Tutkimusdataan perustuvien keksintöjen osalta on oleellista, että patenttilaissa on nimenomaisesti kielletty patentin myöntäminen pelkälle löydölle, tieteelliselle teorialle tai matemaattiselle menetelmälle tai tietojen esittämiselle. Jotta keksintö voitaisiin patentoida, sen täytyy olla uusi. Uutuuden vaatimus on absoluuttinen, missään päin maailmaa ei ole saanut olla vastaavaa keksintöä. Patentoitavan keksinnön täytyy myös olla ihmisen keksinnöllisen toiminnan tulos. Sen täytyy olla vastaus tiettyyn teknologiseen ongelmaan. Löytö ei ole patentoitavissa. Keksinnön täytyy olla toisinnettavissa ja verifioitavissa, ja sen täytyy olla teollisesti hyödynnettävissä. Tämä on tärkeä huomata muun muassa geenien patentoinnin yhteydessä. Jotta geenisekvenssi voitaisiin patentoida, sen täytyy olla eristetty ja sille on täytynyt löytyä teollinen käyttötarkoitus.

Patenttilakiin sisältyy niin sanottu tutkimuspoikkeus, jonka perusteella patentoitua keksintöä voi tutkia uuden tiedon hankkimista varten. Tämä poikkeussäännös ei kuitenkaan Suomen lain mukaan salli patentoidun keksinnön hyödyntämistä laajemmin tutkimuksessa. Keksintöä ei voi käyttää tutkimuksen apuvälineenä Suomen lain tutkimuspoikkeuksen nojalla. Tutkimuspoikkeuksen laajuus vaihtelee eri Euroopan maissa ja uusimmissa laeissa (mm. Belgia ja Sveitsi) on säädetty laajemmasta tutkimuspoikkeuksesta, joka mahdollistaa patentoidun keksinnön hyödyntämisen tutkimuksen apuvälineenä.

Patenttilakiin sisältyy myös niin sanottu *Bolar exception*, jonka mukaan patentti ei koske lääkevalmisteen myyntilupahakemusta varten tarvittavia tutkimuksia, kokeita tai käytännön vaatimuksista aiheutuvia toimia, jotka kohdistuvat kyseisestä lääkevalmistetta koskevaan keksintöön (Patenttilaki 3 §).

Mahdollisesti tutkimustuloksille haettavat patentit on otettava huomioon, kun harkitaan tutkimusdatan ja muiden aineistojen asettamista laajemmin saataville. Tutkimusaineistoja ei pidä julkistaa ennen kuin patenttihakemus on tehty, ja on pidettävä muutenkin huoli siitä, ettei datan julkistaminen vaaranna patentin myöntämistä ja hyödyntämistä.

10 Yksityisyyden suoja ja tutkimusaineistot

Kansalaisten yksityisyyden suoja on Suomessa suojattu perustuslaissa, minkä lisäksi henkilötietojen käsittelystä on säädetty erikseen sitä koskevassa laissa. Julkisuuslaki puolestaan säätelee henkilötietoja sisältävän aineiston luovuttamista viranomaisilta. Yleisen sääntelyn lisäksi useassa erityislaissa (esim. väestörekisterilaki, tilastolaki) on henkilötietojen käsittelyä ja luovuttamista koskevia säännöksiä. Tutkimusaineistojen saatavuuden kannalta on tärkeää erottaa tilanteet, joissa tutkimus itsessään perustuu henkilötietojen keräämiselle ja käsittelylle, ja toisaalta tilanteet, joissa tutkimuksessa hyödynnetään olemassa olevia henkilörekistereitä. Tässä artikkelissa keskitytään niihin kysymyksiin, joita nousee esille tutkijan hyödyntäessä primäärejä henkilötietoja omassa tutkimuksessaan. Rekisteritutkimukseen liittyvä problematiikka ja potilastietojen luovutus ja käyttö jäävät tämän artikkelin ulkopuolelle.

EU:n komissio antoi tammikuussa 2012 ehdotuksen tietosuojaa koskevaksi perusasetukseksi. Sen jälkeen kun asetus on ministerineuvoston ja parlamentin hyväksymä, se on suoraan sovellettavaa oikeutta jäsenvaltioissa.⁵⁵ Tässä tutkimuksessa ei ole ollut mahdollista ottaa huomioon asetusta kokonaisuudessaan, mutta sen tutkimuskäyttöä koskeviin erityissäänöksiin viitataan asianomaisissa kohdissa.

10.1 Henkilötietojen käsittely tutkimuksessa

Suomen perustuslain 10 §:n mukaan jokaisen yksityiselämä, kunnia ja kotirauha on turvattu. Kirjeen, puhelun ja muun luottamuksellisen viestin salaisuus on niin ikään loukkaamaton. Henkilötietojen suojasta säädetään perustuslain mukaan tarkemmin lailla. Perustuslaissa henkilötietojen suojalla tarkoitetaan laajasti yksityisyyden suojan toteutumista henkilötietojen käsittelyssä, rekisteröinnissä ja käyttämisessä.⁵⁶ Tutkimuskohteiden yksityisyyden suoja on näin aina otettava huomioon tutkimusdataa ja muuta aineistoa kerätessä ja käytettäessä. Tästä on tarkemmin säädetty henkilötietolaissa (523/1999), jonka tarkoituksena on perustuslain turvaaman yksityisyyden suojan toteuttaminen henkilötietojen käsittelyssä. Lakia sovelletaan henkilötietojen automaattiseen käsittelyyn sekä myös silloin, kun henkilötiedot muodostavat tai niiden on tarkoitus muodostaa henkilörekiesteri tai sen osa.

Kun tutkimuksessa on tarpeellista käyttää tunnistettavien henkilöiden henkilötietoja, tätä pidetään henkilötietojen käsittelynä henkilötietolain merkityksessä ja käsittelyn on näin ollen täytettävä henkilötietolain määräykset. Henkilötiedon käsite on laissa määritelty laajasti. Henkilötiedolla tarkoitetaan lain mukaan kaikenlaisia luonnollista henkilöä taikka hänen ominaisuuksiaan tai elinolosuhteitaan kuvaavia merkintöjä, jotka voidaan tunnistaa häntä tai hänen perhettään tai hänen kanssaan yhteisessä taloudessa eläviä koskeviksi. (HeitiL 3.1 §) Henkilötiedon käsite kattaa näin ollen tiedot, joiden perusteella henkilö on suoraan tunnistettavissa, sekä tiedot, joita yhdistelemällä muihin tietoihin henkilö voidaan tunnistaa (ns. välillinen tunnistaminen). Sen määrittämiseksi, onko henkilö tunnistettavissa, on EU:n tietosuojadirektiivin mukaan otettava huomioon kaikki kohtuullisesti toteutettavissa

⁵⁵ Proposal for a Council Regulation of the European Parliament and of the Council on the protection of individual with regard to the processing of personal data and on the free movement of such data, COM(2012) 11final.

⁵⁶ HE 309/1993 vp.

olevat keinot, joita joko rekisterinpitäjä tai joku muu voi kyseisen henkilön tunnistamiseksi käyttää (26 artiklan johtolause).⁵⁷

Henkilötietolakia sovelletaan vain henkilötietojen käsittelyyn (HetiL 3.1 § 2 k). Henkilötietojen käsittelyn on oltava lain mukaan huolellista (HetiL 5 §) ja suunnitelmallista (HetiL 6 §). Siinä on noudatettava hyvää tietojenkäsittelytapaa eikä rekisteröidyn yksityiselämän suojaa ja muita yksityisyyttä turvaavia perusoikeuksia saa rajoittaa ilman laissa säädettyä perustetta.

Laki edellyttää, että henkilötietoja voidaan kerätä vain ennalta määritellyä käyttötarkoitusta varten ja tämän käyttötarkoituksen täytyy olla asiallisesti perusteltua rekisterinpitäjän toiminnan kannalta. Henkilötietojen käyttötarkoitus – niiden keräämistapa ja mahdollinen luovutus tai muu käyttö ja säilytys – on määriteltävä ennen kuin henkilötietoja ryhdytään keräämään rekisteriksi (HetiL 7 §). Käsiteltävien henkilötietojen on oltava tarpeellisia käsittelyn tarkoituksen kannalta (HetiL 9 §). Toisin sanottuna tutkimustarkoitukseen kerättävien henkilötietojen tarpeellisuus täytyy voida perustella asianomaisen tutkimusprojektin tarkoituksen mukaisesti.

Jotta henkilötietoja voidaan kerätä tutkimusta varten, toisin sanottuna tutkimuksessa tapahtuvaa henkilötietojen käsittelyä varten, tähän on saatava tutkimuskohteen suostumus. Suostumuksen on oltava lain mukaan vapaaehtoinen, yksilöity ja tietoinen tahdon ilmaisu, jolla rekisteröity hyväksyy henkilötietojensa käsittelyn (HetiL 8 §). Mikäli kyseessä on lääketieteellinen tutkimus, suostumuksen on lähtökohtaisesti oltava kirjallinen.⁵⁸ Suostumusta pyydetessä on tutkimuskohteelle selkeästi ilmoitettava tietojenkäsittelyn tarkoitus, henkilötietojen mahdollinen säilytys ja hävittäminen sekä mahdollinen aikomus luovuttaa tietoja kolmannelle taholle, joista viimeisenä mainittu edellyttää nimenomaista suostumusta. Jos henkilötietoja on tarkoitus käyttää muuhun tarkoitukseen tai toista tutkimushanketta varten, tähän on myös saatava suostumus.

Henkilötietoja ei saa säilyttää enempää kuin on tarpeen käyttötarkoituksen toteutumisen kannalta. Jäljempänä tarkastellaan, mitä tämä tarkoittaa tutkimuksen kannalta.

⁵⁷ Euroopan parlamentin ja neuvoston direktiivi 95/46/EY yksilöiden suojelusta henkilötietojen käsittelyssä ja näiden tietojen vapaasta liikkuvuudesta, 3 artikla.

⁵⁸ Laki lääketieteellisestä tutkimuksesta (488/99) 6 §.

10.2 Rekisterinpitäjän velvollisuudet

Se, joka tai jonka lukuun henkilötietoja käsitellään, on lain tarkoittama rekisterinpitäjä, joka on vastuussa siitä, että henkilötietojen käsittely tapahtuu lain mukaisesti. Yliopistoissa ja korkeakouluissa tapahtuvassa tieteellisessä tutkimuksessa rekisterinpitäjä on tutkija tai tutkimusryhmän johtaja tapauksesta riippuen. Itsenäisessä perustutkimuksessa yliopistolle tai korkeakoululle ei synny rekisterinpitäjän asemaa.

Lain mukaan rekisterinpitäjän on laadittava henkilörekisteristä rekisteriseloste, joka on pidettävä saatavilla. Tutkimuksessa tämä tarkoittanee lähinnä sitä, että rekisteriselosteen on oltava rekisteröityjen ja tutkimukseen tutustuvien henkilöiden saatavilla sekä pyydettyessä myös muidenkin ulottuvilla (HetiL 10 §).

Laissa mainittujen arkaluonteisten tietojen käsittely on pääsääntöisesti kielletty.⁵⁹ Se on kuitenkin mahdollista asianomaisen henkilön suostumuksella (HetiL 12.1 §). Henkilötunnuksen käsittelyä koskee vastaava sääntely. Henkilötunnuksia saa käsitellä vain henkilön nimenomaisella suostumuksella tai laissa nimenomaisesti säädettyissä tapauksissa. (HetiL 13 §)

Tietosuoja-asetusehdotuksen mukaan arkaluonteisia tietoja ja henkilötunnusta voi käsitellä tutkimuskäyttöön, mikäli tunnistetiedot on de-identifioitu niin, että tunnistetiedot säilytetään eri paikassa kuin henkilötiedot (tietosuoja-asetusehdotus 83 §).

Lähtökohtana on, että suostumuksen perusteella tutkimustarkoitukseen kerättyjä tietoja voidaan käsitellä ainoastaan suostumuksessa mainituin tavoin.⁶⁰ Jos suostumuksen perusteella hankittuja henkilötietoja halutaan yhdistää ilman suostumusta kerättyihin rekisteritietoihin, yhdistämiseen täytyy olla tutkimuskohteiden lupa. Toisin sanottuna silloin, kun ke-

⁵⁹ Arkaluonteisina tietoina pidetään henkilötietolain 11 §:n mukaan henkilötietoja, jotka kuvaavat tai on tarkoitettu kuvaamaan:

- 1) rotua tai etnistä alkuperää;
- 2) henkilön yhteiskunnallista, poliittista tai uskonnollista vakaumusta tai ammattiliittoon kuulumista;
- 3) rikollista tekoa, rangaistusta tai muuta rikoksen seuraamusta;
- 4) henkilön terveydentilaa, sairautta tai vammaisuutta taikka häneen kohdistettuja hoitotoimenpiteitä tai niihin verrattavia toimia;
- 5) henkilön seksuaalista suuntautumista tai käyttäytymistä; taikka
- 6) henkilön sosiaalihuollon tarvetta tai hänen saamiaan sosiaalihuollon palveluja, tukitoimia ja muita sosiaalihuollon etuuksia.

⁶⁰ Ks. myös Mäenpää, Yliopistolaki, s. 90 viranomaisen oikeudesta luovuttaa tutkimustarkoituksiin suostumuksen perusteella saatua tietoa.

rätään tietoja tutkimuskohteilta, tietojen käsittelyyn ja yhdistämiseen rekisteritietoihin tähän täytyy olla tutkimuskohteiden lupa. Tämän lisäksi täytyy hankkia asianomaiselta rekisterinpitäjältä lain edellyttämät luvat rekisteritietojen tutkimuskäyttöön ja yhdistämiseen suostumuksen perusteella kerättyihin henkilötietoihin.⁶¹

10.3 Henkilötietojen anonymisointi

Tutkimuksen kannalta on oleellista, että tietosuoja koskevia periaatteita ei sovelleta tietoihin, jotka on tehty anonymiksi siten, ettei rekisteröity enää ole tunnistettavissa. Direktiivin mukaan anonymit tiedot ovat tietoja, jotka koskevat sellaista luonnollista henkilöä, jota rekisterinpitäjä tai muu henkilö ei voi tunnistaa ottaen huomioon kaikki kohtuullisesti toteutettavissa olevat keinot, joita joko rekisterinpitäjä tai joku muu voi käyttää kyseisen henkilön tunnistamiseksi.⁶²

Tiedot voivat olla tällöin joko kokonaan anonymisoituja, jolloin niitä ei voi enää palauttaa alkuperäisiin tunnistettaviin henkilöihin. Voidaan myös käyttää peitenimiä – pseudonyymejä, jolloin tunnisteen on koodattu ja avain koodauksen purkamiseen säilytetään erillään koodatuista tiedoista. Näin voidaan tehdä esimerkiksi kliinisessä lääketieteellisessä tutkimuksessa, jos lääkeyritykselle luovutetaan koodatut tiedot, mutta tutkijalla säilyy koodiavain siltä varalta, että potilaat on pystyttävä uudelleen tunnistamaan, esimerkiksi vaarallisten sivuvaikutusten vuoksi. EU:n ehdotuksessa tietosuoja-asetukseksi asetetaan henkilötietojen käyttö tieteelliseen tutkimukseen ilman tutkimuskohteiden suostumusta mahdolliseksi ainoastaan edellytyksellä, että tunnistetiedot säilytetään eri paikassa, kun muu data.⁶³

10.4 Henkilötietolaki ja tutkimusaineistojen pitkäaikaissäilytys

Kun henkilötietoja kerätään tutkimusta varten, on samassa yhteydessä, kun hankitaan tutkimuskohteiden suostumus heidän henkilötietojensa käsittelyyn, pyydetävä lupa heidän henkilötietojensa säilyttämiseen, mikäli tämä on tarpeen tutkimuksen tarkoituksen toteut-

⁶¹ Ks. tarkemmin tietosuojavaltuutetun toimiston julkaisema Rekisteritutkimuksen tietosuojaopas tutkijoille ja tietopyyntöjä käsitteleville viranomaisille, 27.7.2010, osoitteessa www.tietosuoja.fi.

⁶² Artikla 29:n tietosuojatyöryhmän lausunto henkilötietojen käsitteestä 4/2007, 20.6.2007, 01248/07/FI, WP 136, s. 21.

⁶³ Ks. ehdotus tietosuoja-asetukseksi, 83 artikla.

tamisen kannalta. Erityisen tärkeää tämä on pitkäaikaisissa seurantatutkimuksissa. Mikäli nimenomaista lupaa henkilötietojen pitkäaikaissäilytykseen ei ole hankittu, henkilötiedot on tuhottava heti, kun niiden säilyttäminen ei enää ole välttämätöntä tutkimuksen tarkoituksen toteutumisen kannalta (HetiL 34 §).

Henkilörekisteri, joka on tieteellisen tutkimuksen kannalta tai muusta syystä merkityksellinen, voidaan lain mukaan siirtää korkeakoulun tai tutkimustyötä lakisääteisenä tehtävänä suorittavan laitoksen tai viranomaisen arkistoon, jos kansallisarkisto on antanut siihen luvan. Tietosuojavaltuutetulle on varattava mahdollisuus antaa lausunto luvan myöntämisestä (HetiL 35 §).

Tutkimusaineistojen pitkäaikaissäilytysratkaisuja harkittaessa on pidettävä huoli siitä, ettei tämä vaaranna tutkimusaineistoon sisältyvien henkilöiden yksityisyyden suojaa. Jos säilytettävässä aineistossa ei ole tunnistetietoja, aineisto voidaan säilyttää henkilötietolain sitä estämättä.

10.5 Henkilötietolaki ja tutkimusaineistojen saataville saattaminen

Aineistojen saatavuus tutkimustulosten verifioimista varten on aina järjestettävä tieteellisen tutkimuksen eettisten peruseriaatteiden mukaisesti. Jotta voitaisiin puhua tieteellisen metodin tuottamasta tutkimustuloksesta, tuloksen on oltava verifioitavissa tutkimuksen tausta-aineistojen avulla. Jos nämä aineistot sisältävät henkilötietolain mukaisia tunnistetietoja, aineistojen verifioimisprosessiin sovelletaan henkilötietolakia. Yleensä tutkimuskohteelta saadun luvan yhteydessä pitäisi mainita, että henkilötietoja tullaan käsittelemään tutkimustulosten verifioimista varten, jolloin tämä ei ole ongelma. Tämän lisäksi voidaan edellyttää, että aineistoon perehtyvä henkilö allekirjoittaa salassapitosopimuksen.

Jos taas aineisto saatetaan laajemmin saataville tai julkistetaan, aineistoon sisältyvät tunnistetiedot on poistettava tai anonymisoitava. Muussa tapauksessa rekisteröitävän kanssa tehtävässä sopimuksessa olisi täytynyt sopia aineiston jatkokäytöstä. Se, miten tiettyä tutkimusaineistoa voidaan käyttää jatkossa, riippuu myös aineiston laadusta ja sen käyttöön mahdollisesti sovellettavasta muusta lainsäädännöstä.

10.6 Henkilötietojen siirtäminen EU:n ulkopuolisiin maihin

Monikansallisissa tutkimusryhmissä on otettava huomioon, että henkilötietojen siirtoa EU:n ulkopuolelle on rajoitettu. Tämä on mahdollista, jos tällaisessa maassa on voimassa samantasoinen tietosuoja-asetus kuin EU:ssa. Tämän lisäksi siirtäminen on mahdollista tietyissä muissa tapauksissa, kun voidaan taata riittävä tietosuojan taso. Henkilötietojen siirtäminen EU:n ulkopuolelle on joka tapauksessa mahdollista kyseisten henkilöiden suostumuksella. Jos monikansallisessa tutkimusryhmässä henkilötietoja on tarkoitus käsitellä myös EU:n ulkopuolella sijaitsevien tutkijoiden toimesta, tähän pitää hankkia lupa alkuperäisen suostumuksen yhteydessä.⁶⁴

11 Tutkimusdatan ja tutkimusaineistojen lisensiointi

Tutkimusdatan käyttöön saattamista edeltää usea päätös ja toimenpide ennen kuin se on laajemman yleisön tehokkaasti hyödynnettävissä. Sen jälkeen, kun päätös tutkimusdatan käyttöön saattamisesta on tehty tai kun rahoittaja tätä edellyttää, datan on oltava sellaisessa teknisessä muodossa, että ulkopuoliset pystyvät sitä lukemaan. Se on varustettava tunnistamisen varmistavilla metatiedoilla. Tutkijan on myös päätettävä, mikä on luontevin paikka datan tallettamiselle. Yleensä se on joko kansainvälinen tai kotimainen tieteenalakohtainen repositorio. Jos tieteellinen aikakauskirja tai tutkimusrahoittaja edellyttää datan saattamista yleisön saataville, ne myös yleensä osoittavat sen paikan, minne data tallennetaan. Tällä datarepositoriolla tai tietokannalla on yleensä vahvistettu politiikka sen suhteen, miten yleisö voi dataa käyttää. Esimerkiksi GenBank:iin talletettava geenidata on periaatteessa avoimesti yleisön saatavilla ympäri vuorokauden.⁶⁵

Jos dataan sisältyy henkilötietolainsäädännön mukaan salassa pidettävää tietoa, datan käyttö on mahdollista ainoastaan siihen tarkoitukseen, mihin on saatu tutkimuskohteiden suostumus, ellei laista muuta johdu.

Näiden kahden ääripään välille mahtuu suuri joukko tapauksia, jolloin data on periaatteessa avoimesti saatavilla, mutta sen käyttöä on rajoitettu eri tavoin käyttöehdoissa. Seuraavassa

⁶⁴ Ks. myös tietosuoja-asetusehdotus artikkelit 40–45.

⁶⁵ Ks. tarkemmin GenBank Data Usage, www.ncbi.nlm.nih.gov/genbank/.

tarkastellaan erilaisia tapoja lisensoida tutkimusdatan käyttöä kansainvälisten esimerkkien valossa.

Genomitutkimus on ollut uranuurtajana tutkimusdatan avoimessa jakamisessa. Ihmisen genomin kartoittamisen yhteydessä kansallisten tutkimusrahoittajien konsortio HUGO päätti pitkällisen harkinnan jälkeen saattaa kaiken syntyvän genomitiedon avoimesti verkkoon sitä mukaan kuin dataa syntyi.

Genomitutkimuksen ja luonnontieteiden puolella tutkimusdataa ei perinteisesti varsinaisesti lisensoida vaan se asetetaan vapaasti verkkoon saataville. Jos halutaan nimenomaisesti pätevästi luopua oikeuksista, tämä voidaan tehdä Creative Commons Zero (CC0) -lisenssin avulla.

Koska tutkimusdata sijaitsee yleensä tietokannoissa, lisensointi tarkoittaa käytännössä tietokannan käyttöoikeuden lisensointia. Creative Commons -lisenssit ovat lähtöisin USA:sta, jossa ei ole EU-pohjaista *sui generis* -tietokantasuojaa, ja Yhdysvalloissa erityisesti akateemiset piirit suhtautuvat lähtökohtaisesti erittäin torjuvasti tämän kaltaiseen suojamuotoon. Tästä puolestaan johtuu, että tietokantoja koskevan *sui generis* -suojan sisällyttäminen Creative Commons -lisensseihin on ollut erittäin vaikeaa. Euroopan ulkopuoliset maat ovat pelänneet, että tietokantasuoja eurooppalaisten lisenssien välityksellä ikään kuin takaovesta pääsee myös niiden järjestelmään. Kansainvälinen Creative Commons suosittelee CC0-lisenssin käyttöä, kun halutaan luopua kaikista oikeuksista.⁶⁶ Creative Commonsin piirissä viimeistellään uutta CC4.0-lisenssiä, joka nimenomaisesti sisältäisi myös tietokantaoikeudet.

Euroopassa Open Knowledge Foundation on kehittänyt kolme eri lisenssityyppiä erityisesti tietokantojen lisensointia varten.⁶⁷ OKF:n lisenssit koskevat ainoastaan tietokantojen käyttöä. Ne eivät ulotu tietokantaan sisältyvään materiaaliin. Jos tämä materiaali on tekijänoikeudellisesti suojattua, se on lisensoitava erikseen.

EU:n komissio sallii tietosuojasetusehdotuksessaan henkilötietoja sisältävän datan käytön tutkimustarkoituksiin ilman suotumusta, jos datan käsittely tutkimustarkoitukseen ei ole mahdollista tunnisteettoman tiedon avulla. Tällöin dataan sisältyvät tunnistetiedot pitää

⁶⁶ Ks. <http://creativecommons.org/licenses/>.

⁶⁷ Ks. tarkemmin <http://www.opendatacommons.org/licenses/>.

säilyttää erillään varsinaisesta datasta. Näiden edellytysten täytyessä asianomaista dataa voi käyttää tutkimustarkoituksiin koko EU:n alueella.⁶⁸

Jos tutkimusdatan saatavuus halutaan, esimerkiksi henkilötietojen suojaan liittyvistä syistä, rajoittaa ainoastaan tietyille tutkijoille, voidaan data tallettaa palomuurin taakse järjestelmään, johon on pääsy ainoastaan autentikoinnin kautta. Tällöin voidaan tarkasti kontrolloida, kuka käyttää dataa ja kuinka paljon. Käyttölisenssi pitää tässä tapauksessa laatia aina datan erityislaatua silmällä pitäen. Creative Commons -lisenssit eivät tässä yhteydessä ole mahdollisia, sillä alkuperäinen data on ainoastaan rajoitetusti saatavilla eikä vapaasti edelleen levitettävissä.

Sensitiiviseen mikrodataan voidaan järjestää pääsy data enclave -tyyppisessä teknisesti suojatussa ympäristössä, johon yksittäinen tutkija hakee käyttö lupaa ja saa tutkimusta varten käyttöön kokonaisaineistoja käsittävää mikrodataa, joka voi sisältää sensitiivisiä tietoja. Yksi esimerkki tästä on Chicagon yliopiston NORC Data Enclave. Sensitiivisten rekisteriaineistojen online-käyttöraatkaisuja on tehty myös Euroopassa, esimerkiksi Hollannissa ja Tanskassa. Suljetuissa online-ratkaisuissa dataa ei voi yhdistää muihin aineistoihin, vaan datan käsittely tapahtuu ainoastaan online-ympäristön sisällä.

Kaikista rajoitetuin käyttö on kyseessä silloin, kun data sijaitsee tietyssä paikassa eikä siihen saa pääsyä kuin hakemalla lupa ja menemällä paikan päälle. Tämä voi olla ainoa keino esimerkiksi silloin, jos tutkimuskohteena on lapsia sisältävä audiovisuaalinen aineisto, jonka tutkimustulokset halutaan verifioida.

12 Lopuksi: Avoimempaan tutkimusdataan

Tieteellinen tutkimus lepää perustuslaissa turvatuilla sananvapauden ja tieteellisen tutkimuksen riippumattomuuden periaatteilla. Tieteellinen tieto rakentuu olemassa olevan tiedon päälle, ja uusi tieto edellyttää olemassa olevan tiedon kriittistä tarkastelua. Tutkimustulosten julkisuus ja avoin saatavuus on turvattu perustuslaissa. Jos jokin muu oikeus, kuten niin ikään perusoikeuteen nojautuva tekijänoikeus, on ristiriidassa tutkimuksen julkisuuden kanssa, tutkimuksen julkisuuden painoarvon on katsottava olevan tieteellisen tutkimuksen

⁶⁸ Ks. 83 artikla ehdotuksesta yleiseksi tietosuojasetukseksi.

johtavana periaatteena voimakkaampi kuin omistusoikeuteen pohjautuvan tekijänoikeuden.⁶⁹

Tieteellisen työn tekijöille, tutkijoille, tekijänoikeuden tärkein aspekti on attribuutio-oikeus, oikeus tulla tunnustetuksi tekijänä. Näin he saavat asianmukaisen meriitin, kun heidän teoksiaan käytetään. Silloin kun tämä on mahdollista tutkimusdatan kohdalla, se mahdollisuus on turvattava. Aina kuitenkin näin ei ole, esimerkiksi institutionaalisten mittaustietokantojen kohdalla. Näissäkin tapauksissa voidaan datan kerännyt laitos tai muut instituutiot mainita lähteenä. Tutkijoille on myös oleellista, että heille turvataan oikeus käyttää ensimmäisenä keräämäänsä dataa.

Kun tutkimusdatan avoimuutta tarkastellaan perusoikeudellisesta näkökulmasta, lähtökohdaksi on otettava tieteellisen vapauden turvaaminen, jonka yksi pilareista on mahdollisuus verifioida tutkimustulokset taustadatan avulla. Digitaalinen teknologia on tehnyt taustadatan saatavuuden ja edelleen hyödyntämisen muussa tutkimuksessa mahdolliseksi aivan erittakaavassa kuin aikaisemmin. Kuitenkin tämä vaihtelee tieteenaloittain. Tämän vuoksi taustadatan saatavuuden parametrit on myös määriteltävä tieteenalakohtaisesti. Tällöin myös tekijänoikeudellisen sääntelyn vaikutusta on helpompi tarkastella.

Luonnon- ja bio-lääketieteiden alueella ja muilla sellaisilla alueilla, joissa tutkimusdata koostuu pitkälle erilaisista automaattisesti kerätyistä mittaustuloksista, vakiintunut kansainvälinen käytäntö edellyttää datan julkista saatavuutta, ellei siihen liity yksityisyyden suojaan liittyviä rajoitteita. Suomen tekijänoikeuslainsäädännön näkökulmasta mittaus- ja vastaavaa dataa sisältäviin tietokantoihin kohdistuisi lähinnä luettelosuojaa. Tällaisissa tapauksissa kysymys siitä, omistaako luettelosuojalla suojatun tietokannan oikeudet tutkijat vai heidän työnantajansa, yliopisto, ei liene oleellista datan keränneelle tutkijaryhmälle, ja oikeudet voivat olla yliopistolla. Tällöin datan jatkohyödyntäminen mahdollistuu, kun yliopisto voi yksiselitteisesti sallia sen avoimen saatavuuden, mahdollisesti tietyn ajankohdan jälkeen. Vastaavasti tekijänoikeuslaissa oleva tieteellistä tutkimusta koskeva rajoitussäännös takaisi oikeusvarmuuden datan jatkohyödyntäjille ja toisaalta taas edellyttämällä lähteen mainitsemista, turvaisi tutkijoiden oikeuden tieteelliseen meriittiin.

Toisessa ääripäässä ovat monet yhteiskuntatieteen ja etenkin humanististen tieteiden alueet, joissa tutkijoiden data voi koostua heidän keräämästään ja tietyllä tavalla järjestämään ai-

⁶⁹ Ks. esim. Nature- ja Science-lehtien datapolitiikasta edellä luku 5.

neistosta. Aineistoon voi sisältyä tekijänoikeuden ja henkilötietolainsäädännön suojaamaa materiaalia. Tietokantaa suojataan tällöin tutkijan panoksesta riippuen joko tekijänoikeudellisenä teoksena, *sui generis* -tietokantasuojan tai luettelosuojan avulla. Tässä tapauksessa on myös ilmeistä, että oikeudet tietokantaan eivät siirry lain perusteella työnantajalle. Tällöin tutkijan on huolehdittava siitä, että tekijänoikeuslain mukaiset luvat tietokantaan sisältyvän aineiston tutkimuskäyttöön ja henkilötietolainsäädännön edellyttämät tutkimusluvut on hankittu. Tutkijan on myös varmistettava, että aineistoon on pääsy tutkimustulosten verifiointia varten. Aineiston myöhempi hyödyntäminen tutkimuksessa riippuu sen laadusta ja siihen liittyvistä alkuperäisten teosten oikeuksista. Tässäkin tapauksessa sekä alkuperäisen tutkijan että mahdollisten myöhempien tutkijoiden oikeusturva paranisi huomattavasti, jos laissa olisi annettu tieteellisen tutkimuksen hyväksi rajoitussäännös, joka mahdollistaisi suojattujen teosten käytön tieteelliseen tutkimukseen edellytyksellä, että niiden lähde on mainittu alan käytännön mukaisesti.

Julkisin varoin tuotettu tutkimusdata on julkishyödyke, jonka lähtökohtaisesti täytyisi olla kaikkien saatavilla. Tieteen ja tutkimuksen etu edellyttää, että tutkimusdata on tiedeyhteisön käytettävissä tutkimustulosten verifiointia ja jatkotutkimusta varten. Tämä ei kuitenkaan saa loukata tutkimuskohteiden yksityisyyden suojaa. Tutkijoiden oikeuksien suojaamisessa on keskeistä, että datan lähde asianmukaisesti mainitaan. Samoin tutkijoilla pitäisi olla oikeus ensimmäisenä julkaista keräämäänsä dataan pohjautuvat tulokset. Tämä tutkimus osoittaa, että tutkimusdatan avoin saatavuus lain mukaisilla reunaehdoilla on toteutettavissa ilman kenenkään oikeuksien loukkaamista. Tekijänoikeuslakiin sisällytettävä rajoitussäännös tutkimuksen hyväksi antaisi myös Suomessa työskenteleville tutkijoille tarpeellisen taustatuen kansainvälisessä tutkimusympäristössä.

Lähteet

Kirjallisuuslähteet

Artzberger & al: Promoting Access to Public Research Data for Scientific, Economic, and Social Development, *Data Science Journal*, Volume 3, 29 November 2004, s. 135.

Benoist, Jocelyn: *Qu'est-ce-qu'un livre?*; Emmanuel Kant; textes de Kant et Fichte trad. et présentés par Jocelyn Benoist. Presses universitaires de France. Paris 1995.

Van Eechoud, Mireille; Hugenholtz, Bernt et al. (toim.): Harmonizing European Copyright Law. Kluwer Law International 2009.

Mäenpää, Olli: Yliopistolaki, WSOYPro 2009.

Mäenpää, Olli: Julkisuusperiaate. WSOYpro. Helsinki 2008, s. 3.

Lämsineva, Pekka: Omaisuuden suoja, teoksessa: Perusoikeudet. Werner Söderström lakiti-eto WSLT, 2011, ss. 549 ss.

Making Data Maximally Available. Pääkirjoitus, *Science*, vol 331, 11 Feb. 2011.

Piwowar & Chapman: A review of journal policies for sharing research data, *Nature Proceedings* : hdl:10101/npre.2008.1700.1

Salokannel, Marjut: Julkisesta datasta avoimeen dataan: julkisen datan lisensiointi. ETLA discussion papers, no. 1277, 2012.

Salokannel, Marjut, University of Helsinki opens its research vaults: a few words on open access and the new research environment in Finland, *ScieComInfo*, Vol 4, No 23 (2008).

Salokannel, Marjut: Julkisin varoin tuotettu tieto tutkimuksen käyttöön. *Oikeus* 2011:1, s. 71 ss.

Salokannel, Marjut: Ownership of Rights in Audiovisual Productions. Kluwer Law International 1997.

Salokannel, Marjut: Global public goods and private rights: Scientific research and intellectual property rights. *Nordiskt Immateriellt Rättskydd*, No. 4/2003, s. 334 ss.

Tieto käyttöön – Tiekartta tutkimuksen sähköisten tietoaaineistojen hyödyntämisen edistämiseksi. Opetus- ja kulttuuriministeriön julkaisuja 2011:4. Osoitteessa www.minedu.fi/export/sites/default/OPM/Julkaisut/2011/liitteet/okm04.pdf?lang=fi.

Tuori, Kaarlo: Sivistrykselliset oikeudet, teoksessa: Perusoikeudet. Werner Söderström lakitieto WSLT 2011, s. 605 ss.

Virallislähteet

Artikla 29:n tietosuojatyöryhmän lausunto henkilötietojen käsitteestä 4/2007. 20.6.2007, 01248/07/FI, WP 136. Osoitteessa http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_fi.pdf.

Infopaq International v Danske Dagblades Forening, Euroopan yhteisöjen tuomioistuin, 16.7.2009, asia C-5/08.

Directmedia Publishing GmbH v Albert-Ludwigs-Universität Freiburg, Euroopan yhteisöjen tuomioistuin 9.10.2008, asia C-304/07.

Football Dataco and Others v Yahoo! UK Ltd and Others. Euroopan unionin tuomioistuin 1.3.2012, asia C-604/10.

Fixtures Marketing Ltd v Oy Veikkaus Ab. Euroopan yhteisöjen tuomioistuin, 9.11.2004, asia C-46/02.

Commission recommendation of 17.7.2012 on access to and preservation of scientific information, C(2012) 4890 final .

Communication from the commission to the European parliament, the Council, the European economic and social committee and the committee of the regions, Toward better access to scientific information: Boosting the benefits of public investments in research, COM (2012) 401 final.

Komission tiedonanto 19.10.2009 COM(2009) 532 final.

Commission staff working paper on the review of the EC legal framework in the field of copyright and related rights, 19.7.2004, SEC(2004) 995.

ERC Scientific Council Guidelines for Open Access, December 2007.

GenBank Data Usage. Osoitteessa www.ncbi.nlm.nih.gov/genbank.

HE 7/2009 vp. Hallituksen esitys eduskunnalle yliopistolaiksi ja siihen liittyviksi laeiksi

HE 30/1998 vp. Hallituksen esitys Eduskunnalle laiksi viranomaisten toiminnan julkisuudesta ja siihen liittyviksi laeiksi

HE 170/1997 vp. Hallituksen esitys eduskunnalle laeiksi tekijänoikeuslain ja rikoslain 49 luvun 1 §:n muuttamisesta

HE 309/1993 vp. Hallituksen esitys Eduskunnalle perustuslakien perusoikeussäännösten muuttamisesta

NIH Data Sharing Policy and Implementation Guidance. Osoitteessa http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm#app

OECD principles and guidelines for access to research data from public funding, 2007. Osoitteessa <http://www.oecd.org/dataoecd/9/61/38500813.pdf>

PeVL 46/2010 vp – HE 246/2010: Hallituksen esitys laeiksi julkisen hallinnon tietohallinnon ohjauksesta sekä viranomaisten toiminnan julkisuudesta annetun lain 18 ja 36 §:n muuttamisesta

PeVL 13/2006.- LA 28/2006 vp: Lakialoite laiksi kansainvälisten suhteiden ja Euroopan unionin asioiden tutkimuslaitoksesta sekä laiksi eduskunnan virkamiehistä annetun lain ja valtion maksuperustelain 1 ja 10 §:n muuttamisesta

Principles and guidelines for sharing biomedical research tools, 1999.

Osoitteessa www.ott.nih.gov/policy/research_tool.aspx.

Proposal for a Council Regulation of the European Parliament and of the Council on the protection of individual with regard to the processing of personal data and on the free movement of such data, COM(2012) 11 final.

Research Councils UK Policy on Access to Research Outputs. Osoitteessa <http://www.rcuk.ac.uk/research/Pages/outputs.aspx/>.

Valtioneuvoston periaatepäätös julkisen sektorin digitaalisten tietoaineistojen saatavuuden parantamisesta ja uudelleenkäytön edistämisestä 3.3.2011. Osoitteessa www.valtioneuvosto.fi/toiminta/periaatepaatokset/periaatepaatos/fi.jsp?oid=322887&c=0&toid=135471&moid=135472.