OHJE SOSIAALISEN MEDIAN TUTKIMUKSEEN JYVÄSKYLÄN YLIOPISTOSSA

JOHDANTO

Sosiaalinen media on tärkeä tutkimuskohde ja tutkimuksen väline. Aineiston digitaalisen saatavuuden vuoksi tutkimusaineiston muodostaminen voi vaikuttaa helpolta. Sosiaalisen median tutkimuksessa on kuitenkin käytettävä samaa harkintaa tutkimusetiikan, tutkimuksen tarkoituksen ja sen tieteellisen merkittävyyden suhteen kuin tutkimuksessa muutenkin, mutta tutkimusetiikkaa ja lainsäädäntöä joudutaan soveltamaan perinteisistä tutkimusasetelmista poikkeavassa ympäristössä. Vastuu sosiaalisen median tutkimuksen toteutuksesta on tutkijalla itsellään. **Tämän ohjeen tarkoituksena on auttaa tutkijaa sosiaalisen median tutkimukseen liittyvissä erityiskysymyksissä ja helpottaa tutkimuksen tekemistä.** Tässä ohjeessa sosiaalisen median tutkimuksesta ja -aineistoista käytetään ilmaisua sometutkimus ja someaineistot.

Ohjeen lähtökohtana on kansallisen työryhmän julkaisu ”Näkökulmia sosiaalisen median eettiseen ja vastuulliseen tutkimukseen” (Eilo ym., 2026). Kyseisessä dokumentissa on käsitelty sometutkimusta laajemmin, ja tuotu esiin hyvät askelmerkit tutkijalle ja opinnäytetyön tekijälle ratkaisujen pohtimiseksi. Askelmerkit käsittelevät seuraavia asioita: 1. Tutkimuskysymys, 2. Alustan eli palveluntarjoajan ehdot (käyttöehdot ja luvat), 3. Sisällöntuottajat, 4. Henkilötiedot, 5. Suostumus ja informointi, 6. Julkisuus, 7. Riskit, 8. Eettinen ennakkoarviointi, 9. Aineiston säilyttäminen tutkimuksen aikana ja sen jälkeen, 10. Avoimuus ja julkaiseminen.

Tämä JYUn oma sometutkimusohje tarkentaa edellä mainittua dokumenttia. Ohjeen laatimisen keskeisenä motivaattorina ovat olleet aloittelevilta tutkijoilta, opiskelijoilta ja opinnäytetöiden ohjaajilta tulleet sometutkimukseen liittyvät kysymykset. Käytännön neuvontatyössä on havaittu, että sometutkimuksen kysymysten ratkaiseminen edellyttää useiden eri näkökulmien huomioimista. Näkökulmien yhteensovittamiseksi ja yhdenmukaisten käytäntöjen luomiseksi on tärkeää laatia kokonaisuuden huomioiva, kattava ohjeistus. **Tämän ohjeen laatimisen olennaisena tavoitteena on mahdollistaa sometutkimuksen tekeminen tavalla, joka kunnioittaa yhtä lailla tieteellisen tutkimuksen vapautta ja tutkittavien oikeuksia kuin tutkimusetiikkaa ja Suomen lakia.**

Ohjeen laatimiseen ovat osallistuneet Jyväskylän yliopiston aineistonhallinnan, tutkimusetiikan ja lainsäädännön asiantuntijat. Ohjetta ovat sen laatimisvaiheessa kommentoineet sometutkimusta tekevät tutkijat, tietosuojavastaava sekä Jyväskylän yliopiston ihmistieteiden eettinen toimikunta.

Sometutkimus on moniulotteinen ja nopeasti kehittyvä alue. Eri tieteenaloille on syntynyt erilaisia sometutkimuksen traditioita ja käytäntöjä, joita on totuttu noudattamaan. Tämä ohje ei voi vastata tyhjentävästi kaikkiin ala- tai aineistokohtaisiin käytännön kysymyksiin, vaan on tarkoitettu ensisijaisesti juridisiin ja eettisiin kysymyksiin vastaamiseksi.

Konkreettisia tulkintakysymyksiä voi esittää tai lisätukea pyytää seuraavien kanavien kautta.

Aineistonhallinta: researchsupport-osc@jyu.fi, opetus@library.jyu.fi
Lakiasiat: legal@jyu.fi
Tietosuoja: tietosuoja@jyu.fi
Tutkimuseettiset kysymykset: ethics-committee@jyu.fi

Ohjetta päivitetään tarvittaessa. Ohjeen päivittämisestä vastaa Avoimen tiedon keskus.

Päivityshistoria:

Päivämäärä	Mitä päivitetty
25.11.2026	Ohjeen ensimmäinen versio julkaistu

Sometutkimusohjeen työryhmä: Outi Aarresola (työryhmän koollekutsuja, asiantuntija/aineistonhallinta OSC), Visa Hiltunen (juristi, HALLA), Tiia Puputti (asiantuntija/opetustiimi, OSC), Pilvi Hämeenaho (erityisasiantuntija, TIP, eettinen toimikunta).

Sisällys

1. Mitä sometutkimuksella tarkoitetaan tässä ohjeessa 5
2. Sometutkimuksen luonteesta 5
2.1. Julkisen aineiston tutkiminen 6
2.2. Ei-julkisten aineistojen tutkiminen 7
2.3. Uuden aineiston tuottaminen tutkimusta varten 8
3. Henkilötiedot ja tietosuoja 9
3.1. Yleistä 9
3.2. Henkilötietojen käsittelystä informointi, kun tietoja ei kerätä rekisteröidyiltä 11
3.2.1. Rekisteröityjä on mahdollista informoida suoraan 11
3.2.2. Rekisteröityjä ei ole mahdollista informoida suoraan 12
3.2.3. Poikkeukset informoinnista 13
3.3. Someaineistojen tutkimus ja muu tutkimus, jossa henkilötietoja ei kerätä rekisteröidyiltä 14
3.4. Henkilötietojen käsittelystä informointi, kun tiedot kerätään rekisteröidyiltä 14
3.5. Tietosuojaa koskeva vaikutustenarviointi 14
4. Tutkimusetiikka: eettinen ennakkoarviointi ja tutkimuksen riskit 15
5. Tekijänoikeus, kappaleiden valmistaminen ja palveluntarjoajan ehtojen vaikutus tutkimuksen tekemiseen 17
5.1. Yleistä 17
5.2. Kappaleiden valmistaminen eli aineiston tallentaminen tutkimuskäyttöön 18
5.2.1. Yksityinen käyttö 18
5.2.2. Kopioston kopiointilupa 18
5.2.3. Tiedonlouhinta 18
5.2.4. Yhteenveto 19
5.3. Somealustojen käyttöehdot 19
5.3.1. Sallivat käyttöehdot 19
5.3.2. Rajoittavat käyttöehdot 19
6. Sometutkimuksen perusteella tehdyt tutkimusjulkaisut 20

6.1. Yksityisyydensuoja....20
6.2. Eri henkilöryhmät....21
6.3. Tekijänoikeusnäkökulmia....21
7. Aineiston säilyttäminen hankkeen aikana....21
8. Aineistojen käsittely tutkimuksen päätyttyä....22
9. Lähteet....23

1. Mitä sometutkimuksella tarkoitetaan tässä ohjeessa

Tässä ohjeessa sometutkimuksella tarkoitetaan tutkimusta, joka kohdistuu erilaisten digitaalisten palvelujen käyttäjien sähköisesti julkaisemaan sisältöön. Someaineistot ovat usein vuorovaikutteisia ja usein puhutaan arkikielellä somepostauksista. Someaineistot voivat sisältää esimerkiksi kuvia, tekstejä tai videoita. Someaineiston julkisuus vaihtelee – aineisto voi olla saatavilla hyvin rajatulle kohdejoukolle tai kenelle tahansa. Someaineiston lähteenä voivat olla esimerkiksi sisällöntuottajien tilit, keskusteluryhmät, kommenttipalstat tai yritysten ja muiden organisaatioiden sometilit.[1]

2. Sometutkimuksen luonteesta

Onnistuneen, tutkimuseettisesti kestävän ja lainsäädäntöä noudattavan tutkimuksen suorittaminen edellyttää tutkimuksen huolellista suunnittelua ja kokonaisuudenhallintaa. Keskeistä on ymmärtää someaineiston luonne: kaikki somessa julkaistu aineisto ei ole julkista. Someaineiston käyttömahdollisuuksiin ja tutkijalta edellytettäviin menettelyihin vaikuttaa ratkaisevasti se, missä ja millaiselle kohdeyleisölle aineisto on julkaistu ja kuka on julkaisija.

Sometutkimukseen liittyvien kysymysten ratkaisemisessa on olennaista hahmottaa ensin se, **kohdistuuko tutkimus olemassa oleviin ja muutenkin saatavilla oleviin aineistoihin vai tuotetaanko aineistoa tutkimusta varten**. Someaineistot voidaan tietyissä tapauksissa rinnastaa julkisiin tai julkistettuihin[2] tietoihin tai asiakirja-aineistoihin, ja sitä myöten rekisteritutkimukseen. Rekisteritutkimukseen rinnastaminen jäsentää sometutkimukseen liittyviä juridisia ja eettisiä kysymyksiä erityisesti henkilötietojen näkökulmasta. Tämä selkeyttää sometutkimuksen tekemistä ja suunnittelua. Asiaa avataan jäljempänä yksityiskohtaisemmin.

Rekisteritutkimuksella tarkoitetaan tutkimusta, joka kohdistuu olemassa oleviin aineistoihin. Rekisteritutkimus on perinteisesti kohdistunut esimerkiksi viranomaisten hallussa oleviin aineistoihin. Tällöin tärkeitä kysymyksiä ovat olleet aineistojen luovutus-,

[1] Verkkolehtien julkaisijat tulkitsevat pääsääntöisesti, että julkaistun artikkelin kommentit ovat osa artikkelia ja vastaavat mielipidekirjoituksia. Tämä johtuu siitä, että EU:n digipalveluasetuksessa on nimenomaan määritelty, että verkkosanomalehtien kommenttipalstat ovat liitännäisiä pääasialliseen palveluun eivätkä itsenäisiä digipalveluasetuksessa säädeltyjä verkkoalustoja. Tutkimuskäytön ja tämän ohjeen näkökulmasta sanomalehtien kommenttipalstat kuitenkin rinnastuvat julkiseen someaineistoon.

[2] Julkistettu on tekijänoikeuslaissa käytetty käsite. Teos on tekijänoikeuslain 8 §:n mukaan julkistettu, kun se on luvallisesti saatettu yleisön saataviin.

käyttöoikeus- ja salassapitokysymykset sekä mahdollinen henkilötietojen käsittelystä informointi.

Se, että rekisteritutkimusaineistoissa on ollut luonnollisia henkilöitä koskevaa tietoa, ei ole tarkoittanut, että nämä henkilöt varsinaisesti osallistuisivat tutkimukseen. Rekisteriaineistossa mukana olevilta henkilöiltä ei ole myöskään pyydetty suostumusta tutkimukseen osallistumiseen. Aineistossa mukana oleminen ei ole tutkimuseettisessä tarkoituksessa edellyttänyt päätöksentekoa osallistumisesta tai aktiivista osallistumista tiettyyn tutkimukseen.[3]

Tutkimuseettisen neuvottelukunnan eettisen ennakkoarvioinnin ohjeessa (TENK 2019) todetaan pääsääntönä, että eettinen ennakkoarviointi edellytetään silloin, jos poiketaan tietoon perustuvan suostumuksen periaatteista. Edelleen ohjeessa todetaan, että arviointia ei edellytetä julkisten ja julkistettujen tietojen, rekisteri- ja asiakirja-aineistojen ja arkistoaineistojen tutkimukseen (paitsi jos tutkimuksesta aiheutuu turvallisuusuhka tutkittaville tai tutkijoille). Toisin sanoen tutkimuseettisen neuvottelukunnan ohje tarkoittaa, että julkisten ja julkistettujen tietojen, rekisteri- ja asiakirja-aineistojen ja arkistoaineistojen tutkimus ei yleensä edellytä tutkittavan suostumusta. Tässä ohjeessa otetaan kantaa sekä julkisten että ei-julkisten someaineistojen tutkimukseen.

2.1. Julkisen aineiston tutkiminen

Iso osa sosiaalisessa mediassa julkaistusta aineistosta on saatavilla kenelle tahansa. Aineiston julkaisija on tarkoittanut julkaisunsa saataville rajattomalle kohdejoukolle. **Julkisilla sosiaalisen median tileillä tai verkon keskustelupalstoilla olevat tiedot rinnastuvat julkisiin ja julkistettuihin tietoihin tai asiakirja-aineistoihin. Julkisella tarkoitetaan myös sisältöä, joka on alustalla, jonka käyttäjäksi kuka tahansa voi rekisteröityä ja sitä kautta päästä käsiksi julkaisuihin.**[4] Julkisia aineistoja voi lähtökohtaisesti tutkia ilman tutkittavan suostumusta tai eettistä ennakkoarviointia.

Julkisia aineistoja voivat tuottaa yksityishenkilöt tai yhteisöt. Jos tutkit henkilöiden tuottamia julkisia sisältöjä, et tarvitse heidän suostumustaan tutkimukseen. Jos tutkit organisaation tuottamaa julkista sisältöä, et tarvitse tutkimuslupaa. Tarvitset tutkimusluvan tai aineistojen käyttöoikeuden, mikäli pyydät tutkimaltasi organisaatiolta

[3] Joissain tapauksissa henkilötietojen luovuttaminen tutkimuskäyttöön on riippunut siitä, miten ihmisiä on informoitu henkilötietojen luovuttamisesta tai mikäli henkilötietojen käsittelyperusteena on ollut suostumus, millaiset henkilötietojen käsittelyyn liittyvät suostumukset heiltä on kysytty. Henkilötietoja sisältävän aineiston käsittely tietosuojalainsäädännön mukaisesti on kuitenkin eri asia kuin tutkittavan antama tutkimuseettinen osallistumissuostumus.

[4] Julkisella ei kuitenkaan tarkoiteta aineistoja, jotka ovat oikeudettomasti verkossa esimerkiksi tietomurron seurauksena.

erikseen materiaalia, rekrytoit tutkittavia organisaation hallinnoimia kanavia pitkin (esim. tutkimuskutsun välittäminen organisaation intraan tai sähköpostilistalle) tai toteutat tutkimusta organisaation toimitiloissa. Tutkimuslupa haetaan tutkimuksen kohteena olevalta organisaatiolta, heidän ohjeidensa mukaisesti.

Muista, että tutkittavalta saatava eettinen suostumus on eri asia kuin suostumus henkilötietojen käsittelyn oikeudellisena perusteena. Tyypillisesti julkisetkin aineistot sisältävät henkilötietoja. Henkilötietojen käsittelyyn liittyviin vaatimuksiin palataan myöhemmin.

2.2. Ei-julkisten aineistojen tutkiminen

Sosiaalisessa mediassa ja verkon keskusteluryhmissä on myös sisältöä, joka on tarkoitettu selkeästi yksityiseksi tai tietyn kohderyhmän sisäiseksi. Tällaisia aineistoja ei voida katsoa julkisiksi, vaikka ne löytyvätkin verkosta.

Ei-julkista aineisto on, jos se on tarkoitettu vain tietylle erikseen valitulle kohdejoukolle. Valinta voi tapahtua esimerkiksi siten, että sisällön luoja tai ylläpito erikseen hyväksyy seuraajat tai liittymisen keskusteluryhmään. Ei-julkiseksi aineisto on katsottava myös silloin, jos pääsy aineistoon perustuu jonkin yhteisön jäsenyyteen. Tällaisia ovat esimerkiksi yhdistysten, urheiluseurojen tai yritysten jäsenten tai henkilöstön sisäiset keskustelupalstat.

Ei-julkisten someaineistojen käyttö edellyttää aineiston tuottaneiden henkilöiden suostumusta. Jos saat suostumuksen esimerkiksi vain osalta keskusteluryhmään osallistuneista, voit käyttää tutkimukseen vain näiltä osallistujilta saatuja aineistoja. Mikäli poiketaan suostumuksen kysymisestä kokonaan, tulee tutkimukselle suorittaa eettinen ennakkoarviointi.

Vaikka suljetut ryhmät ovat lähtökohtaisesti ei-julkisia, on erityisen suurten, mutta käyttöoikeuksiltaan rajattujen verkkoyhteisöjen kohdalla perusteltua kysyä, onko sisällöntuottajan tai yksittäisen somekäyttäjän julkaisu suljettuun ryhmään saanut sellaisen yleisön, että julkaisu tulisi julkistetuksi tekijänoikeuslaissa ja tutkimuseettisen neuvottelukunnan ohjeessa tarkoitetulla tavalla.[5] Tällöin suuressa suljetussa ryhmässä oleva sisältö voisi rinnastua kohdan 2.1 mukaiseen julkiseen someaineistoon.

Yleisesti ottaen suljettujen yhteisöjen sisältöjen katsomiseen julkiseksi someaineistoksi on suhtauduttava varovaisuudella. Tässä ohjeessa ei voida määritellä sellaista suljetun

[5] TENK 2019, 16. Ennakkoarvioinnin kriteeriä a) täsmennetään ohjeissa niin, että ”Eettisen toimikunnan arviointia ei edellytetä julkisten ja julkistettujen tietojen, rekisteri- ja asiakirja-aineistojen ja arkistoaineistojen tutkimukseen.” Tekijänoikeuslain 8 § puolestaan määrittelee teoksen julkistetuksi, kun se on saatettu yleisön saataviin luvallisesti.

ryhmän käyttäjien vähimmäismäärää, mikä väistämättä tarkoittaisi aineiston muuttumista julkiseksi. Arvioon julkisuudesta voivat vaikuttaa ainakin tutkimuksen aihealue ja tutkimusta varten kerättävä sisältö, tutkittava alusta sekä kulttuuriset ja maantieteelliset seikat. Tutkimusasetelman ja aineiston julkisuusluonteen arviointiin voit pyytää JYUn asiantuntijoiden apua (ks. tukiosoitteet johdannossa).

Mikäli olet päätymässä siihen, että suuressa suljetussa ryhmässä oleva sisältö on julkista eikä sen tutkiminen ilman tekijöiden suostumusta edellytä ennakkoarviointia, tulee tällainen ratkaisu dokumentoida huolellisesti. Dokumentointi on tärkeää, jotta voit myöhemmin perustella pääsäännöstä poikkeavan ratkaisusi olevan hyvän tieteellisen käytännön mukainen.

Tutkimuseettisestä näkökulmasta suljettujen yhteisöjen tutkimiseen voi liittyä muitakin kysymyksiä. Jos tutkija joutuu perustelemaan liittymispyynnön suljettuun ryhmään muulla kuin tutkimustarkoituksella, rinnastuu tämä harhautukseen (tutkittaville ei anneta riittävää tietoa tutkimuksesta). Tällaisessa tilanteessa eettisen toimikunnan ennakkoarvio on välttämätön.

2.3. Uuden aineiston tuottaminen tutkimusta varten

Jos olet itse tutkijana esittämässä julkisella keskustelupalstalla kysymyksiä, niin silloin keräät uutta aineistoa eikä kyse ole enää rekisteritutkimukseen rinnastettavasta tutkimuksesta. Tällöin sinun on informoitava tutkittavia tutkimuksesta ja henkilöti etojen käsittelystä sekä hankittava tutkittavien suostumus osallistumiseen kuten muussakin tutkimuksessa. Tällaisessa asetelmassa tulee käyttää selkeää aikarajausta, ja kertoa, milloin tutkija osallistuu keskusteluun ja miltä ajalta aineistoa kerätään. Näin osallistujat voivat valita, osallistuvatko keskusteluun tutkimuksen aikana vai eivät.

Tiedota valitsemallasi alustalla tai keskustelussa hyvissä ajoin tutkimuksesi aikataulusta. Kerro tutkimuksestasi ja henkilötietojen käsittelystä kirjallisesti, ja ole valmis myös vastaamaan kysymyksiin. Erilaiset alustat tarjoavat mahdollisuuden kerätä aineistoa myös vuorovaikutteisesti, joko kirjoittaen tai puhuen. Voit esimerkiksi luoda oman tutkimuskanavan, jolle kutsut osallistujia. Kerro kaikille osallistujille aina selkeästi tutkimuksesi tarkoitus ja tavoite, tallennatko viestit ja keskustelut ja miten käsittelet osallistujien henkilötietoja.

Kerätessäsi uutta aineistoa somessa sinun tulee pystyä erottamaan tutkimukseen liittyvä toimintasi muusta somessa olemisesta. Varmista, että esimerkiksi seuraamaasi keskusteluun osallistuva tietää sinun olevan paikalla tutkijana. Käytä siis selkeää profiilia ja pidä tutkimuksesta kertovat tiedotteet helposti löydettävissä. Voit kiinnittää

tutkimuksesta kertovan tiedotteen viestiketjun alkuun tai kertoa tutkimuksestasi kanavan kuvauksessa.

3. Henkilötiedot ja tietosuoja

3.1. Yleistä

Henkilötiedoilla tarkoitetaan kaikkia tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön liittyviä tietoja.[6] Niin some- kuin muussakin tutkimuksessa on ennen tutkimuksen aloittamista arvioitava, sisältääkö tutkimus henkilötietojen käsittelyä. Someaineistojen tutkimuksessa on hyvin todennäköistä, että käsittelet henkilötietoja.

Henkilötietojen käsittelyä tapahtuu varmasti, jos tallennat itsellesi kopiona esimerkiksi omalla nimellä tehdyn postauksen tai kuvia, joissa esiintyy tunnistettavia ihmisiä, tai muita tunnistettavaan yksilöön liittyviä tietoja. Vaikka tutkisit ilmiöitä, voi esimerkiksi muistiinpanojen tai tekemiesi analyysien kautta aineistoon tallentua henkilötietoja. Henkilötiedon kriteerien täyttymiseen riittää sekin, että henkilö on mahdollisuus tunnistaa välillisesti aineistossa olevien tietojen tai saatavilla olevien lisätietojen kautta (esimerkiksi syöttämällä hakukoneeseen keskustelijan nimimerkki, ks. myös suorista lainauksista osio 4 tutkimusetiikka).

Anonyymin tekstipohjaisen keskustelupalstan tutkiminen voi teoriassa onnistua ilman henkilötietojen käsittelyä. Anonyymiydestä on kuitenkin hankala varmistua etukäteen. Jos osa keskusteluun osallistujista käyttää yleisesti tiedossa olevaa tai muuten henkilöllisyyden selvittämisen mahdollistavaa nimimerkkiä, ei tutkimusta käytännössä voi toteuttaa ilman henkilötietojen käsittelyä. Ennalta ei voi myöskään tietää, mitä kaikkea keskustelussa tulee ilmi. Tutkittaessa julkisia keskustelupalstoja (ilman, että tutkija

[6] Henkilötietojen määritelmä on tietosuoja-asetuksen 4 artiklan 1 kohdassa: Tässä asetuksessa tarkoitetaan ’henkilötiedoilla’ kaikkia tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, jäljempänä ’rekisteröity’, liittyviä tietoja; tunnistettavissa olevana pidetään luonnollista henkilöä, joka voidaan suoraan tai epäsuorasti tunnistaa erityisesti tunnistetietojen, kuten nimen, henkilötunnuksen, sijaintitiedon, verkkotunnistetietojen taikka yhden tai useamman hänelle tunnusomaisen fyysisen, fysiologisen, geneettisen, psyykkisen, taloudellisen, kulttuurillisen tai sosiaalisen tekijän perusteella, Tunnistettavuutta on avattu johdantolauseessa 22 seuraavasti: ”Jotta voidaan määrittää, onko luonnollinen henkilö tunnistettavissa, olisi otettava huomioon kaikki keinot, joita joko rekisterinpitäjä tai muu henkilö voi kohtuullisen todennäköisesti käyttää mainitun luonnollisen henkilön tunnistamiseen suoraan tai välillisesti, kuten kyseisen henkilön erottaminen muista. Jotta voidaan varmistaa, voidaanko keinoja kohtuullisen todennäköisesti käyttää luonnollisen henkilön tunnistamiseen, olisi otettava huomioon kaikki objektiiviset tekijät, kuten tunnistamisesta aiheutuvat kulut ja tunnistamiseen tarvittava aika sekä käsittelyajankohtana käytettävissä oleva teknologia ja tekninen kehitys.”

ohjailee tai osallistuu keskusteluun) ei ole mahdollista vaikuttaa siihen, mitä tietoja aineisto tulee sisältämään.

Asiaa kannattaa ajatella myös tutkimuksen toteuttamiskelpoisuuden kannalta. Se, että tutkimus sisältäisi henkilötietojen käsittelyä, ei estä tutkimuksen suorittamista. **On turvallisempaa ja helpompaa lähteä siitä, että käsittelet henkilötietoja, niin tutkimuksen aikana ei tule yllätyksiä ja tarpeita rajata kerättäviä tietoja.** Tutkimusta suunnitellessa ja suorittaessa on kuitenkin muistettava, että henkilötietoja ei saa käsitellä laajemmin kuin tutkimuksen toteuttaminen edellyttää.

Henkilötietojen käsittely edellyttää laillisen käsittelyperusteen olemassaoloa. Tyypillisimmin käsittelyperuste yliopistojen suorittamassa tieteellisessä tutkimuksessa on yleinen etu [Tietosuoja-asetus Artikla 6 (1) e ja tietosuojalaki 4 § 1 mom 3 kohta]. Mikäli tieteellisen tutkimuksen kriteerit eivät täyty (tyypillisimmin näin voi tapahtua kandidaatin tutkielman kohdalla, useimmiten pro gradu -tutkielma tulkitaan tieteelliseksi tutkimukseksi), mahdollisia muita käsittelyperusteita ovat rekisteröidyn suostumus [Tietosuoja-asetus Artikla 6 (1) a] ja oikeutettu etu (Tietosuoja-asetus Artikla 6 (1) f]. Oikeutetun edun käyttäminen käsittelyperusteena vaatii niin kutsutun tasapainotestin suorittamista.[7]

Tietosuoja-asetuksessa käytetään termejä rekisterinpitäjä ja rekisteröity. Rekisterinpitäjä, eli tutkija tai organisaatio, vastaa henkilötietojen keräämisestä ja niiden käsittelystä. Rekisteröity taas viittaa niihin ihmisiin, joiden tietoja käsitellään. Tieteellisen tutkimuksen kontekstissa tunnistettavissa olevat tutkittavat tai aineistossa esiintyvät henkilöt ovat rekisteröityjä.

Jos tutkimuksesi aineistossa on erityisiä henkilötietoryhmiä, kuten terveystietoja[8], voit edelleen tehdä tutkimusta, mutta käsittelyssä pitää huomioida seuraavat seikat: Kun käsitellään erityisiä henkilötietoryhmiä, tarvitaan käsittelyyn lisäksi peruste sille (tietosuoja-asetuksessa on säädetty lähtökohtaisesta kiellosta käsitellä erityisiä henkilötietoryhmiä). Näitä perusteita löytyy tietosuoja-asetuksen 9 artiklasta ja tietosuojalain 6 pykälästä. **Tietosuojalain 6 §:n 1 momentti 7 kohta mahdollistaa erityisten henkilötietoryhmien käsittelyn tieteelliseen tutkimukseen.** Sometutkimuksen kohdalla erityisten henkilötietojen käsittelyn lisäperusteena voi käyttää myös sitä, että käsittely koskee rekisteröidyn nimenomaisesti julkisiksi saattamia tietoja (tietosuoja-asetuksen 9 artiklan 2 kohdan e alakohta).

---

[7] Tasapainotestillä verrataan henkilötietojen käsittelystä vastaavan tahon (rekisterinpitäjä) etuja niiden henkilöiden etuihin, joihin käsittely kohdistuu (rekisteröidyt).

[8] Erityiset henkilötietoryhmät on lueteltu tietosuoja-asetuksen 9 artiklan 1 kohdassa.

Ei-tieteellisissä töissä, joissa käsittelyperusteena on suostumus, tarvitaan erityisten henkilötietoryhmien käsittelyyn nimenomainen suostumus. Oikeutetun edun perusteella ei voida käsitellä erityisiä henkilötietoryhmiä.

Henkilötietojen käsittely edellyttää lähtökohtaisesti sitä, että aineistossa esiintyviä henkilöitä eli rekisteröityjä informoidaan heidän henkilötietojensa käsittelystä. Informointivelvoitteet eroavat hieman riippuen siitä, kerätäänkö tietoja henkilöiltä itseltään vai ei.

3.2. Henkilötietojen käsittelystä informointi, kun tietoja ei kerätä rekisteröidyiltä

Julkisten someaineistojen tutkimuksessa tietoja ei tyypillisesti kerätä henkilöiltä itseltään. Tällöin informointivelvoitteet ovat tietosuoja-asetuksen 14 artiklan mukaiset, eli informointi voi tapahtua viimeistään kuukauden[9] kuluessa henkilötietojen saamisesta[10]. Käytännössä tämä tarkoittaa, että julkisten someaineistojen tutkimuksessa tutkijan tulee miettiä, miten hän ajoittaa tietojen keräämisen ja informoinnin. Mikäli aineiston kerääminen kestää alle kuukauden, voi informoinnin tehdä aineiston keräämisen päätyttyä. Jos aineiston kerääminen kestää yli kuukauden, on suositeltavaa informoida henkilötietojen käsittelystä ennen aineiston keräämisen alkua. Tutkijan on pohdittava, miten informointi vaikuttaa aineiston sisältöön.

Informoinnin keskeinen väline on tietosuojailmoitus. Asetelma, jossa henkilötietoja ei kerätä rekisteröidyiltä itseltään vaan ne päätyvät osaksi tutkimusaineistoa ilman henkilön omaa myötävaikutusta, eroaa merkittävästi esimerkiksi perinteisistä haastattelututkimuksista. Tällaista tutkimusta varten on laadittu oma tietosuojailmoitusmallinsa, johon on yhdistetty myös tutkittavan tiedotteessa normaalisti kerrottavia asioita soveltuvin osin. Ajatuksena on, että yhdistetyn tiedotteen ja tietosuojailmoituksen tulee olla ymmärrettävä lukijalle, joka kuulee tutkimuksesta ensimmäistä kertaa. Tiedote-tietosuojailmoituksen laatimisen lisäksi keskeistä on arvioida, miten se voidaan saattaa rekisteröidyille tiedoksi. Valittu ratkaisu tulee kirjata tietosuojailmoitukseen.

3.2.1. Rekisteröityjä on mahdollista informoida suoraan

Jos tutkimuksesi kohteena on rajoitettu joukko henkilöitä, joiden yhteystiedot ovat saatavilla, tulee tietosuojailmoitus toimittaa näille henkilöille suoraan.

---

[9] Tietosuoja-asetuksen 14 artiklan 3a kohdan mukaisen kuukauden aikarajan laskemisessa merkityksellinen on se ajankohta milloin tutkija henkilötiedot kerää, ei se, sillä milloin tiedot on alun perin julkaistu.
[10] Henkilötietojen käsittelystä on kuitenkin informoitava jo aiemmin, mikäli tietoja käytetään viestintään rekisteröityjen kanssa ja rekisteröityyn otetaan yhteyttä TAI silloin kun henkilötietoja luovutetaan edelleen.

Jos esimerkiksi tutkit jonkin somevaikuttajan tai muun yksittäisen henkilön somekanavaa tai -tiliä, lähetä hänelle tietosuojailmoitus henkilökohtaisesti. Voit käyttää esimerkiksi sähköpostia tai tutkimasi alustan viestitoiminnallisuutta. Tietosuoja-asetus käyttää sanamuotoa ”toimitettava rekisteröidylle”. Arvioi tilanteen mukaan, miten voit hoitaa henkilöiden informoinnin siten, että informointi tosiasiassa tavoittaa heidät.

3.2.2. Rekisteröityjä ei ole mahdollista informoida suoraan

Jos on mahdotonta tai kohtuuttoman vaivan takana saada henkilöihin yhteys ja toimittaa tietosuojailmoitus henkilökohtaisesti, mahdollistaa tietosuoja-asetuksen 14 (5) b artikla suorasta informointivelvoitteesta poikkeamisen. Yksilöllinen informointi voi olla kohtuuttoman vaivalloista, mikäli rekisteröityjä on paljon, heidän yhteystietonsa eivät ole saatavilla tai henkilötiedot on julkaistu kauan aikaa sitten. Rekisterinpitäjän on tietosuoja-asetuksen 14 artiklan 5 b kohdan mukaan toteutettava asianmukaiset toimenpiteet rekisteröidyn oikeuksien ja vapauksien sekä oikeutettujen etujen suojaamiseksi. Ensisijainen toimenpide näiden takaamiseksi on julkaista tietosuojailmoitus ja pyrkiä varmistamaan, että rekisteröidyt voivat saada tietosuojailmoituksen olemassaolosta tiedon.

Rekisteröityjen julkisen informoinnin tulisi tapahtua tarkoituksenmukaisia keinoja käyttäen. **Jos tutkit jonkin julkisen somekanavan kommentteja tai keskustelupalstan viestejä, toimiva keino on linkittää julkaistu tietosuojailmoitus kommentteihin, viestiketjuun tai tehdä kokonaan oma keskustelunavaus palstalle.**[11]

Tietosuojailmoitus voidaan julkaista esimerkiksi tutkimushankkeen verkkosivuilla. Opiskelija voi julkaista tietosuojailmoituksen omalla JYU-verkkosivullaan. Tietosuojailmoitus on syytä pitää saatavilla verkossa aineiston keräämisen ajan ja silloin, kun tutkimustuloksia julkaistaan. Tutkimustulosten julkaisu saattaa herättää kysymyksiä tutkimukseen liittyvästä henkilötietojen käsittelystä, jolloin tutkimuksen tietosuojailmoituksen suora verkkosaatavuus on sekä aineistossa esiintyvien henkilöiden että tutkijan etu.

[11] On hieman tulkinnanvaraista ja tutkimusasetelmasta riippuvaa, voidaanko tietosuojailmoituksen julkaisemista ja siihen linkittämistä katsoa tietojen toimittamiseksi rekisteröidyille tietosuoja-asetuksen 14 (1) artiklan tarkoittamalla tavalla yksiöllisesti vai onko kyse 14 (5)b artiklan mukaiseen poikkeukseen vetoamisesta (yleisluontoinen/julkinen informointi). Lähtökohtana on kuitenkin aina se, että tiedot henkilötietojen käsittelystä on pyrittävä toimittamaan rekisteröidylle, oli keinona sitten niiden lähettäminen suoraan rekisteröidylle tai julkisten kanavien käyttäminen rekisteröidyn tavoittamiseksi,.

On mahdollista, että tietosuojailmoituksen linkkaaminen tutkittavan sisällön julkaisijalle ei välttämättä johda siihen, että tietosuojailmoitus tavoittaisi sivustolla esiintyviä henkilöitä. Jos tutkit esimerkiksi organisaation sometilin mainoksia, niin mainoksissa esiintyvien henkilöiden tavoittaminen voi joskus olla huomattavan epätodennäköistä. Jos arvioit, että tietosuojailmoituksen lähettäminen organisaatiolle tai siihen linkkaaminen organisaation sometilin kommenttikentässä ei paranna rekisteröityjen informointia, riittää tietosuojailmoituksen julkaisu.

3.2.3. Poikkeukset informoinnista

Jos on todennäköistä, että henkilötietojen käsittelystä informointi estää tieteellisen tutkimuksen tarkoitusten saavuttamisen tai vaikeuttaa sitä suuresti, on informointivelvollisuudesta mahdollisuus poiketa. Tyypillisesti tällainen tilanne on käsillä esimerkiksi silloin, kun tutkija tiedostaa henkilötietojen käsittelystä informoinnin vaikuttavan keskustelupalstalla käytävän keskustelun sisältöön. Henkilötietojen käsittelystä informointi siis johtaisi erilaisiin tutkimustuloksiin kuin se, jos käsittelystä ei informoitaisi. Tällaisissa tapauksissa vaihtoehtona ovat tietosuojailmoituksen julkaiseminen ilman, että sitä pyritään aktiivisesti toimittamaan rekisteröidyille, tai tietosuojailmoituksen toimittaminen rekisteröidyille sen jälkeen, kun se ei enää haittaa tutkimuksen suorittamista. Valitussa ratkaisussa on löydettävä tasapaino tutkimuksen vapauden ja rekisteröityjen tietosuojaoikeuksien välillä.

Joissain tapauksissa, esimerkiksi tutkimusaiheen ja tutkimuskohteen aiheuttaman turvallisuusuhan vuoksi, tietosuojailmoitus voidaan jättää julkaisematta. Esimerkkinä tällaisesta tapauksesta on väkivaltaisiin ääriryhmiin kohdistuva tutkimus. Tällöin on erityisen huolellisesti arvioitava ja tietosuojailmoitukseen kirjattava, mitkä ovat tietosuoja-asetuksen 14 artiklan 5 b -kohdan edellyttämät asianmukaiset toimenpiteet. Normaalien suojatoimien lisäksi (pseudonymisointi, turvalliset käsittely-ympäristöt, säilytysaikojen rajoittaminen, kerättävien tietojen minimointi) tulee tällaiselle tutkimukselle pyytää eettinen ennakkoarviointi. Tutkijan on erityisen huolellisesti pohdittava sitä, miten paljon henkilötietoja aineisto sisältää ja miten aineistossa esiintyvien yksityisyyttä on mahdollista suojata tutkimusaineiston keräämisessä, analysoinnissa ja tulosten raportoinnissa.

Vaikka tietosuojailmoituksen toimittamisesta rekisteröidyille poikettaisiinkin, **on tietosuojailmoitus kuitenkin aina laadittava**. Tietosuojailmoituksen laatiminen ja siihen kirjatut menettelytavat tietosuojailmoituksen toimittamisesta tai toimittamatta jättämisestä rekisteröidyille ovat tärkeitä keinoja osoittaa henkilötietojen käsittelyä koskevan lainsäädännön noudattaminen.

3.3. Someaineistojen tutkimus ja muu tutkimus, jossa henkilötietoja ei kerätä rekisteröidyiltä

Henkilötietojen käsittelyyn soveltuvan lainsäädännön näkökulmasta sometutkimus ei eroa muusta tutkimuksesta, jossa henkilötietoja kerätään muualta kuin rekisteröidyiltä itseltään. Henkilötietojen määritelmä ja tietosuoja-asetuksen 14 artiklan mukaiset informointivelvoitteet ovat samanlaiset riippumatta siitä keräätkö tai vastaanotatko henkilötietoja sisältävän tutkimusaineiston somesta vai jostain muualta. Muista siis laatia tietosuojailmoitus ja huolehtia rekisteröityjen informoinnista (tai perustella informoinnista poikkeaminen) myös muiden kuin someaineistojen kohdalla! Tätä ohjeistusta ja laadittua tietosuojailmoitusmallia voi hyödyntää myös muihin tilanteisiin, joissa henkilötietoja ei kerätä suoraan rekisteröidyiltä.

3.4. Henkilötietojen käsittelystä informointi, kun tiedot kerätään rekisteröidyiltä

Mikäli tutkimus suoritetaan aiemmin kohdassa 2.3 kuvatulla tavalla (uuden aineiston tuottaminen tutkimusta varten) tai tutkitaan ei-julkisia someaineistoja tutkittavan suostumuksella, kerätään tiedot tutkittavilta itseltään ja tutkittavien informointi on tietosuoja-asetuksen 13 artiklan mukainen. Tällöin henkilötietojen käsittelystä on informoitava silloin, kun tiedot vastaanotetaan. Näitä 13 artiklan mukaisia tapauksia yhdistää se, että tutkittavat päättävät aktiivisesti itse osallistumisestaan tutkimukseen. Näissä tapauksissa käytetään **normaalia tutkittavan tiedotetta ja tietosuojailmoitusta, ei sometutkimuksen tietosuojailmoitusta**.

Jos henkilötiedot kerätään rekisteröidyltä, ei informointivelvoitteista ole mahdollista poiketa.

3.5. Tietosuojaa koskeva vaikutustenarviointi

Mikäli henkilötietojen käsittely aiheuttaa rekisteröidyille (tutkittaville) korkean riskin, täytyy suorittaa tietosuojaa koskeva vaikutustenarviointi ennen henkilötietojen käsittelyn alkua. Vaikutustenarvioinnin tarpeeseen vaikuttavat esimerkiksi tietojen laajamittainen käsittely, tietokokonaisuuksien yhdistäminen, uuden teknologian käyttö, lasten tai muuten haavoittuvassa asemassa olevien henkilötietojen käsittely ja kerättävien tietojen arkaluontoisuus. Ellei tietosuojailmoitusta julkaista, kasvaa todennäköisyys sille, että tarvitaan tietosuojaa koskeva vaikutustenarviointi.

Rekisteröidyn asemaan ja hänelle aiheutuvaan riskiin vaikuttaa myös tutkimuksen toteuttamistapa ja aineiston laatu. Suoria tunnistetietoja sisältävän aineiston kohdalla

riskit ovat luonnollisesti suuremmat kuin niissä tapauksissa, joissa pieni osa aineistoa tuottaneista tai aineistossa esiintyvistä henkilöistä saattaa olla välillisesti tunnistettavissa.

Yliopiston tietosuojaohjeista löytyvän alkukartoituslomakkeen avulla voit arvioida, tarvitaanko vaikutustenarviointia.

4. Tutkimusetiikka: eettinen ennakkoarviointi ja tutkimuksen riskit

Kaikessa tutkimuksessa, mukaan lukien sometutkimus, tulee noudattaa eettisiä ohjeita ja hyvää tieteellistä käytäntöä (TENK 2023). Jos someaineisto ei ole julkista (ks. kohta 2), on tutkimuksen lähtökohtana tutkittavilta saatava tietoon perustuva suostumus. Erityistä eettistä harkintaa tulee noudattaa, jos tutkimuksen kohteena on alaikäisiä, vajaakykyisiä tai muihin haavoittuviin ryhmiin kuuluvia.

Sosiaalinen media on myös kansainvälinen ja kulttuurisesti moninainen ympäristö. On muistettava, että esimerkiksi suomenkielisillä alustoilla toimivat sisällöntuottajat eivät välttämättä ymmärrä suomea täydellisesti. Huomioi kielellinen saavutettavuus laatiessasi informointiasiakirjoja.

Myös julkisten someaineistojen kohdalla tulee huomioida someaineistojen henkilökohtaisuus ja sensitiivisyys suhteessa vaikeisiin aiheisiin. Julkisenkin keskustelun käyttäminen tutkimusaineistona voi vaikuttaa negatiivisesti niihin, joiden kirjoituksia tai sisältöjä käytät, jos aihe on heille vaikea. Tällaisia aiheita voivat olla esimerkiksi syrjintään, kiusaamiseen, mielenterveyden ongelmiin, väkivaltakokemuksiin tai köyhyyteen liittyvät teemat. Voit olla etukäteen yhteydessä esimerkiksi keskustelupalstan ylläpitäjään, joka voi antaa lisätietoja kyseisen yhteisön säännöistä tai toiveista.

Someaineistojen kohdalla on tärkeää muistaa, ettei aineisto ole alun perin syntynyt tutkimuskäyttöön. Somekeskustelut voivat olla hyvin hetkellisiä ja niiden tyyli riippuvainen valitusta alustasta. Tutkimukseen valittu aineisto tulee kontekstoida someviestinnän ja tutkitun alustan erityispiirteet huomioiden. Vinoumien tunnistaminen ehkäisee esimerkiksi kärkevien ja kärjistettyjen näkökulmien toisintamisen tai niiden yleistämisen irti alkuperäisestä tarkoituksesta (tietty aihe tietyllä hetkellä tietyllä keskustelufoorumilla).

Myös sometutkimus saattaa edellyttää eettistä ennakkoarviointia, jos yksikin kriteeri TENKin määrittelemistä kuudesta ennakkoarviointia edellyttävästä tutkimusasetelmasta täyttyy (TENK 2019, 16: Ihmiseen kohdistuvan tutkimuksen eettiset periaatteet ja ihmistieteiden eettinen ennakkoarviointi Suomessa).

Sometutkimuksessa tyypillisimmät ennakkoarviointia edellyttävät kriteerit ovat

- Tutkimukseen osallistumisessa poiketaan tietoon perustuvan suostumuksen periaatteesta
- Tutkimus kohdistuu alle 15-vuotiaisiin ilman huoltajan erillistä suostumusta tai informointia, jonka perusteella huoltajalla olisi mahdollisuus kieltää lasta osallistumasta tutkimukseen
- Tutkimuksessa on riski aiheuttaa tutkittaville tai heidän läheisilleen normaalin arkielämän rajat ylittävää henkistä haittaa
- Tutkimuksen toteuttaminen voi merkitä turvallisuusuhkaa tutkittaville tai tutkijalle tai heidän läheisilleen.

Poikkeaminen tietoon perustuvasta suostumuksesta: Sometutkimuksessa selkein ennakkoarviointia edellyttävä kriteeri on, jos tutkija haluaisi tutkia **ei-julkisia someaineistoja** ilman aineiston tuottajien suostumusta tai tutkia alle 15-vuotiaita ilman heidän ja heidän huoltajansa suostumusta. Jos tietoon perustuvan suostumuksen periaatteesta poiketaan, tulee sille olla vahva tutkimuksellinen peruste ja tutkimukselle tulee pyytää eettinen ennakkoarviointi. Usein selkeintä on rakentaa tutkimusasetelma siten, että sisällöntuottajilta tai muilta tutkittavilta voidaan kysyä suostumus tai vaihtoehtoisesti käyttää julkisia someaineistoja.

Riski aiheuttaa haittaa:
Pohdi huolellisesti, millaisia riskejä tai haittoja tutkittaville voi koitua, jos esimerkiksi uudelleen julkaiset heidän tuottamiaan sisältöjä avoimesti julkaistavassa opinnäytteessä tai tutkimusjulkaisussa. Älä levitä sisältöjä, joita ei ole tarkoitettu julkisiksi.

Monesti tutkimuksessa käytetään suoria aineistositaatteja. Someaineistojen kohdalla sitaattien käyttöä tulee pohtia suhteessa tutkittavien tunnistamiseen. Ennakoi, kuinka paljon esimerkiksi sitaatit voivat lisätä mahdollisuutta, että tutkittava tulee tunnistettavaksi. Aineiston pseudonymisoinnista huolimatta yksittäinen sitaatti on helppo paikantaa tietylle nimimerkille tiettyyn keskusteluun. Kontekstistaan irrotetut sitaatit voivat johtaa vääriin käsityksiin sisällön alkuperäisestä tuottajasta ja siitä, mitä hän alun perin tarkoitti. Pohdi vastaavia vaikutuksia myös aineistossa esiin tulevien organisaatioiden tai yritysten näkökulmasta. Voiko somepostausten uusintaminen aiheuttaa esimerkiksi mainehaittaa, jota alkuperäinen keskustelu ei ole tuottanut?

Haittaa voi koitua myös läheisille ja sivullisille henkilöille. Esimerkiksi rikosepäilyjen ja -tuomioiden uudelleenkäsittely tutkimuksessa voi olla uudelleen traumatisoivaa uhreille ja kuormittavaa kaikkien osallisten läheisille. Varo myös levittämästä ja uusintamasta vihapuhetta sitaateilla, jotka eivät ole tutkimuksen kannalta välttämättömiä.

Mahdollisia riskejä tulee arvioida myös tutkijan ja hänen lähipiirinsä näkökulmasta. Esimerkiksi, jos olet sukupuolentutkija ja tutkit verkossa esiintyvää naisvihaa, voit joutua maalittamisen kohteeksi. Riskiä voi pienentää jättämällä tutkijan suorat tunniste- ja yhteystiedot pois tiedotteesta (yhteydenotot hankkeen yhteiseen sähköpostiin tai vastaava käytäntö).

Pohdi siis huolellisesti, voiko sinulle itsellesi, aineistossa esiintyville henkilöille tai kenelle tahansa sivulliselle koitua harmia siitä, että somessa julkaistuja postauksia analysoidaan ja nostetaan esille ja tuodaan ne uuden yleisön tietoon.

5. Tekijänoikeus, kappaleiden valmistaminen ja palveluntarjoajan ehtojen vaikutus tutkimuksen tekemiseen

5.1. Yleistä

Kirjalliset ja taiteelliset teokset (kuten kuvat, videot ja tekstit) ovat tekijänoikeudella tai tekijänoikeuden lähioikeudella suojattuja. Teos saa tekijänoikeussuojaa, mikäli se ylittää ns. teoskynnyksen eli ilmentää tekijänsä luovaa panosta ja omaperäisiä valintoja. Esimerkiksi someen postatut kuvat ja tekijänsä luovaa panosta ilmentävät kertomukset ovat tekijänoikeudella suojattua sisältöä. Joissain yksinkertaisissa kirjallisissa ilmauksissa voi olla, että teoskynnys ei ylity.[12] Kerättävän aineiston tekijänoikeudellista asemaa on arvioitava ennen aineiston keruun aloittamista. Tehtyjä rajauksia on noudatettava tutkimuksen suorittamisen ajan.

Tekijänoikeuden haltijalla on oikeus päättää kappaleiden valmistamisesta ja teoksen saattamisesta yleisön saataviin. Aineiston kerääminen somesta tutkimusta varten tarkoittaa mitä todennäköisimmin aineiston tallentamista eli kappaleiden valmistamista someaineistosta, koska analyysiä varten halutaan useimmiten tallentaa vaikkapa kuvakaappaus postauksesta tai kopioida keskustelu tallennettavaksi tiedostoksi.

Tutkijalla täytyy olla peruste valmistaa kappaleita tekijänoikeudella suojatusta aineistosta. Näitä käsitellään seuraavissa alaluvuissa. On hyvä huomioida, että kappaleiden valmistaminen, eli tallentaminen on eri asia kuin niiden edelleen välittäminen tai julkaiseminen.

---

[12] Teoskynnyksestä Harenko et al. s. 22–23: ” lähes kaikki kirjallisessa ilmenemismuodossa olevat luomistyön tuotteet saavat lain 1 §:n nojalla suojaa, kunhan tuotteesta ilmenee laatijan itsenäisyys ja omaperäisyys. Jos ilmaisu kuitenkin kuvaa vain lyhyesti ja ilman mitään luovaa panosta esimerkiksi päiväntapahtumaa, kuten lauseessa ”valtion ensi vuoden budjetin loppusumma on viisi prosenttia alempi kuin kuluvana vuonna”, se jäisi todennäköisesti omaperäisyyden puuttumisen takia vaille tekijänoikeussuojaa.”

5.2. Kappaleiden valmistaminen eli aineiston tallentaminen tutkimuskäyttöön

5.2.1. Yksityinen käyttö

Tekijänoikeuslain 12 § mahdollistaa kappaleen valmistamisen julkistetusta teoksesta yksityistä käyttöä varten. Kuten aiemmin on todettu, teos on 8 §:n mukaan julkistettu, kun se on saatettu yleisön saataviin. Opiskelukäyttö on tekijänoikeuslain 12 §:ssä tarkoitettua yksityistä käyttöä, mutta ammattikäyttö ei ole. Kappaleiden valmistaminen työsuhteessa tai muuta vastiketta vastaan tapahtuvan tutkimuksen mahdollistamiseksi ei siis lähtökohtaisesti ole sallittua tekijänoikeuslain 12 §:n nojalla. Opiskelijan on kuitenkin mahdollista valmistaa kappaleita someaineistoista opintoja tai opinnäytetöitä varten tekijänoikeuslain 12 §:n nojalla, mikäli teokset ovat julkistettuja eli tutkitaan julkisia someaineistoja.

5.2.2. Kopioston kopiointilupa

Tutkijalla on mahdollisuus tehdä kopioita aineistosta tutkimustaan varten Kopioston kopiointiluvan nojalla. Kopioston kopiointiluvalla kopioita saavat valmistaa korkeakoulun opiskelijat, opettajat, tutkijat, luennoijat ja muu henkilökunta. Kopioita voi tutkimusta varten tehdä myös digitaalisesta aineistosta. Sometutkimuksen kannalta merkittävin rajoitus kuitenkin on, että Kopioston kopiointilupa ei mahdollista kopioiden valmistamista yksityishenkilöiden tekemistä teoksista, joita nämä jakavat sosiaalisessa mediassa, keskustelupalstoilla, blogeissa tai kuvanjakopalveluissa. Yksityishenkilöä ei ole kopiointiluvassa määritelty, mutta sometutkimuksen kontekstissa Kopioston kopiointiluvalla voidaan perustella kappaleiden valmistaminen organisaatioiden, julkisuuden henkilöiden ja ammattimaisten sisällöntuottajien somessa olevista teoksista. Kopioston kopiointiluvan ehdot löytyvät Kopioston verkkosivuilta.

5.2.3. Tiedonlouhinta

Tekijänoikeuslain 13 b § 2 momentti mahdollistaa kappaleiden valmistamisen tieteellisessä tutkimuksessa tapahtuvaa tekstin- ja tiedonlouhintaa varten. Tekstin- ja tiedonlouhinta on tekstin, äänten, kuvien tai datan automaattista laskennallista analyysiä. Aineiston analysointi tekoälyllä katsotaan tiedonlouhinnaksi.

Tiedonlouhintapoikkeuksesta hyötyvät muutkin kuin työsuhteessa olevat yliopistoyhteisön jäsenet eli opiskelijat, emeritukset/emeritat ja apurahatutkijat. Tutkimusta on kuitenkin suoritettava JYU-affilioituneena ja sen on täytettävä tieteellisen tutkimuksen kriteerit. Edellytyksenä on myös, että aineistoon on laillinen pääsy. Tämä tarkoittaa avoimesti verkossa olevaa sisältöä tai maksumuurin tai kirjautumisen takana olevaa sisältöä, johon on hankittu pääsyoikeus. Valmistettuja kappaleita saa säilyttää tieteellistä tutkimusta varten, mutta ei julkaista avoimesti. Tutkijoiden oikeudesta valmistaa kappaleita ja käyttää niitä tiedonlouhintaan ei voida poiketa sopimuksin (ks. kohta 5.3 somealustojen käyttöehdot). Se, että aineistosta voi valmistaa kappaleita nimenomaan tekstin- ja tiedonlouhintaa varten, voi ohjata tutkimuksen teossa käytettäviä menetelmiä.

Lainsäädännön näkökulmasta selkeimmin someaineiston kerääminen siis onnistuu sillä perusteella, että kopiot aineistosta valmistetaan tiedonlouhintaa varten. Nämä mahdollisuudet ja rajoitteet tulee ottaa huomioon jo tutkimussuunnitelmassa.

5.2.4. Yhteenveto

Jos mikään kappaleen valmistamisen mahdollistama poikkeus ei sovellu tutkimukseesi, laadi tutkimusasetelmasi siten, että se ei sisällä kappaleiden valmistamista tekijänoikeudella suojatuista teoksista. Toinen vaihtoehto on pyytää tekijänoikeuden haltijoilta lupa aineiston käyttöön. Silloin on luonnollisesti asianmukaista informoida heitä tutkimuksesta ja henkilötietojen käsittelystä.

5.3. Somealustojen käyttöehdot

Somealustat ottavat omissa käyttöehdoissaan mielellään kantaa siihen, kuka omistaa alustalle ladatut teokset ja miten niitä on sallittua jatkokäyttää. Aineistojen käyttöoikeudet voivat olla joko sallivia tai rajattuja. Käyttöehdoilla ei kuitenkaan voida rajoittaa sellaista toimintaa, joka on lainsäädännön perusteella sallittua.

5.3.1. Sallivat käyttöehdot

On mahdollista, että alustan käyttöehdoissa suoraan sallitaan alustalla julkaistun sisällön käyttö akateemisiin tutkimuksiin. Tällaiset käyttöehdot ovat luonnollisesti tutkijan etu. Jos käyttöehdot ovat tutkimusmyönteiset, ei ole tarpeen etsiä muita tekijänoikeuslaista tai sopimuksista tulevia perusteita oikeuttamaan tutkimusaineiston keräämistä.

5.3.2. Rajoittavat käyttöehdot

On mahdollista, että sosiaalisen median alustat pyrkivät omissa käyttöehdoissaan rajoittamaan aineiston käyttöä tai tiedonlouhintaa. Tiedonlouhintaoikeus kuitenkin perustuu EU-direktiiviin, joten kaikkien EU:n jäsenvaltioiden on tullut mahdollistaa tiedonlouhinta kansallisella lainsäädännöllään. Myös oikeus valmistaa kappaleita yksityiseen käyttöön on lakisääteinen (Tekijänoikeuslaki 12 §). Alustan käyttöehdot eivät voi rajoittaa sellaista toimintaa, jota pakottava lainsäädäntö sallii.

Rajoittavien käyttöehtojen vaikutukset ovat luultavammin käytännöllisiä kuin juridisia. Palveluntarjoaja voi tarjota tiedot käyttöön oman rajapintapalvelunsa kautta, jolloin se voi määrittää tämän palvelun sisällön rajoituksineen. Vastaavasti palveluntarjoaja saattaa sulkea tarjoamansa rajapinnan, mikäli se havaitsee käyttöoikeuden saaneen käyttäjän hankkivan tai hyödyntävän aineistoa myös muuten. On epätodennäköistä, että palveluntarjoaja ryhtyisi varsinaisiin oikeustoimiin alustalta tekstin- tai tiedonlouhintaa EU-lainsäädännön mahdollistamalla tavalla suorittavaa tutkijaa vastaan. Suurimmillaan riski on varmasti silloin, jos tutkija tiedonlouhintaoikeuteensa vedoten suorittaa sellaista toimintaa, jota alustan tarjoaja itse tarjoaa maksullisena palveluna.

Tutkijan kannattaa arvioida aineiston keruutaan ja tutkimusmenetelmiään tilanteen mukaan. Jos tutkimus on mahdollista toteuttaa alustan käyttöehtojen mukaisesti tai sen tarjoamien rajapintojen kautta, näin kannattaa ilman muuta toimia. Jos käyttöehtojen mukaan ei ole mahdollista toimia, on hyvä arvioida käyttöehtojen noudattamatta jättämisestä aiheutuvat riskit ja perustella tekemänsä ratkaisut jo aineiston keräämisvaiheessa.

6. Sometutkimuksen perusteella tehdyt tutkimusjulkaisut

6.1. Yksityisyydensuoja

Mieti tutkittavien yksityisyyden suojaa julkaisuissa jo tutkimuksen suunnitteluvaiheessa. Sometutkimusta varten laaditussa tietosuojailmoitusmallissa pyydetään ottamaan kantaa tutkittavien tunnistettavuuteen myös julkaisuvaiheessa.

Jos tutkimuskohteenasi on ollut ei-julkinen someaineisto, älä julkaise aineistoon sisältyviä valokuvia, joista niiden kohteet ovat tunnistettavissa. Myös julkisen aineiston kohdalla valokuvien julkaisuun tunnisteellisena tulee olla vahvat tutkimusasetelmasta ja aiheesta tulevat perusteet. Erityisesti tämä koskee sellaisia tilanteita, joissa kuvan on julkaissut yksityishenkilö ja siinä esiintyy muita henkilöitä kuin julkaisija itse. Tarvittaessa

julkaisemista voi perustella tietosuojalain 27 §:llä, joka mahdollistaa henkilötietojen käsittelyn akateemisen, taiteellisen tai kirjallisen ilmaisun tarkoituksia varten.

6.2. Eri henkilöryhmät

Tutkimusta julkaistessa on tärkeää huomioida, ovatko tutkimasi henkilöt olleet somessa julkisuuden henkilöinä, yksityishenkilöinä vai jonkin organisaation edustajina. Jos tutkit esimerkiksi tunnetun poliitikon virallista sometiliä, ei ole tarpeellista pyrkiä anonymisointiin. Myös tutkimusaihe vaikuttaa julkaisemiseen: jos poliitikko osallistuu keskusteluun esimerkiksi harrastuksensa vuoksi, tutkimuskohteena on harrastajien välinen vuorovaikutus eikä tutkimusjulkaisussa muutenkaan mainita harrastajia nimeltä, ei julkisuuden henkilön mainitseminen nimeltä ole perusteltua. Jos taas tutkit puoluejohtajien julkisuuskuvan hallintaa, on selvää, ettei anonymisointi ole tarpeellista tai edes mahdollista.

6.3. Tekijänoikeusnäkökulmia

Tekijänoikeuslaissa säädetty sitaattioikeus mahdollistaa sen, että tutkimusjulkaisuissa voit tekijänoikeussuojan estämättä esittää otteita alkuperäisaineistosta. Sitaattioikeus voi kohdistua esimerkiksi tekstinpätkään tai kuvaan. Sitaatin ottamisen tulee olla perusteltua tutkimuksen aiheen ja julkaisun kannalta. Usein sitaatti tarvitaan tutkimuksen analyysin ymmärtämiseksi. Sitaattioikeus perustuu tekijänoikeuslain 22 §:ään. Tekijänoikeuslain 25 §:llä on vielä mahdollistettu erikseen tekstiin liittyvien taideteosten kuvien käyttäminen arvostelevassa tai tieteellisessä esityksessä (ns. kuvasitaatti).

Tekijänoikeudellista sitaattioikeutta ei tarvita, mikäli sitaatin kohteena oleva asia ei lainkaan saa tekijänoikeussuojaa (tekijänoikeus ei rajoita tekijänoikeussuojan ulkopuolisen sisällön käyttöä), kuten hyvin geneeriset lyhyet ilmaukset. Tällöin ei ole tarvetta pohtia myöskään sitaattioikeuden käytön edellytyksiä. Täytyy kuitenkin muistaa, että tekijänoikeudellisen sitaattioikeuden käytön edellytys on se, että teos on julkistettu. Ei-julkistettujen someaineistojen kohdalla tutkimusjulkaisuun ei siis voi ottaa otteita tekijänoikeudella suojatusta teoksesta sitaattioikeuden nojalla.

7. Aineiston säilyttäminen hankkeen aikana

Henkilötietoja sisältävä aineisto tulee säilyttää tietoturvallisessa ympäristössä sekä vain sen ajan, kuin henkilötietoja on tarpeen käsitellä.

Aineistojen säilytystapa määritellään tutkimushankkeen aineistonhallintasuunnitelmassa. Aineistoa, jonka rekisterinpitäjä on JYU tai joiden säilytyksestä JYU muuten vastaa, tulee säilyttää JYU:n ohjeistuksessa määritellyissä käsittely-ympäristöissä. Sama koskee aineistoja, joiden henkilötietojen käsittelijänä toimii JYU (ellei rekisterinpitäjän kanssa ole toisin sovittu). Rekisterinpitäjinä toimivat opiskelijat tai itsenäiset JYU-affilioituneet tutkijat noudattavat ohjeistusta mahdollisuuksien mukaan ja vastaavat tutkimuksensa ja välineidensä tietoturvallisuudesta.

8. Aineistojen käsittely tutkimuksen päätyttyä

Aineistojen käsittelyllä tutkimuksen päätyttyä tarkoitetaan jatkotutkimuksen tekemistä, aineistojen arkistointia tai julkaisemista tutkimuskäyttöön. Aineistojen käsittelyä tutkimuksen tuloksia julkaistessa on käyty läpi jo edellä. Jotta kerättyjä tutkimusaineistoja olisi mahdollista säilyttää hankkeen päättymisen jälkeen tutkimuskäyttöön, tämä on useimmiten suunniteltava ennen aineistonkeruun aloittamista. Syynä tähän ovat someaineistojen ominaispiirteet, jotka ovat haastavia määräajattoman säilyttämisen tai avoimesti julkaisemisen kannalta.

Aineistojen avoin julkaiseminen internetissä kenen tahansa saataville on mahdollista vain poikkeuksellisesti, ellei siihen ole kysytty osallistujien suostumusta. Tähän perusteena ovat sekä tekijänoikeuslainsäädännöstä että tietosuojalainsäädännöstä tulevat rajoitteet. Avoin julkaiseminen olisi mahdollista ilman suostumusta siinä tapauksessa, että aineisto olisi anonymisoitua eikä siinä olisi mukana tekijänoikeudella suojattuja teoksia tai tällaisten teosten muunnelmia. **Aineistojen arkistointi, eli luovuttaminen arkistoon säilytettäväksi ja uudelleen käytettäväksi, voi kuitenkin olla mahdollista.**

Tekijänoikeuslain 13 b § 2 momentti mahdollistaa tiedonlouhintaa varten valmistettujen kappaleiden säilyttämisen tieteellistä tutkimusta varten, mukaan lukien tutkimustulosten todentamiseen myöhemmin.[13] Edellytyksenä on, että teoksen kappaleet ovat vain siihen oikeutettujen saatavilla. Tyypillisin tapa aineistojen arkistoinnissa on aineiston säilyttäminen käyttöoikeusrajoitettuna eli aineiston luovuttaminen vain tutkimuskäyttöön tapauskohtaisen harkinnan perusteella, mikä täyttää 13 b § 2 momentin edellytyksen. Käytännössä siis tekijänoikeuden suojaaman someaineiston arkistointiin varmin tapa on suorittaa tutkimusta tiedonlouhinnan menetelmin ja perustella myös arkistointivaiheessa, että tutkimuksen suorittamista varten someaineistosta valmistetut kappaleet säilytetään myöhemmän, tieteellisen tutkimuksen edellyttämän tiedonlouhinnan mahdollistamiseksi.

[13] Pykälän perustelut sisältäneessä hallituksen esityksessä (HE 43/2022 vp s. 78–79) on nimenomaan todettu, että tiedonlouhinta-aineiston pysyvä säilyttäminen on sallittua.

Henkilötiedoin arkistointi on mahdollista esimerkiksi JYUn tutkimusaineistoarkistoon, mikäli aineisto on tutkimuksellisesti merkittävä ja läpäisee arvonmäärityksen (ks. tutkimusaineistojen arkistointi). Ensisijainen oikeusperuste arkistointiin on yleinen etu. Arkistointisuunnitelmia varten ole yhteydessä aineistonhallinnan palveluihin researchsupport-osc@jyu.fi jo tutkimusta suunniteltaessa. Arkistointikäytäntöjä kehitetään jatkuvasti ja pyritään löytämään ratkaisuja erilaisten aineistojen säilyttämiseen ja avaamiseen.

Mikäli itse aineistoa ei ole lopulta mahdollista arkistoida tai julkaista avoimesti, on mahdollista silti julkaista kattava lähdeluettelo, tiedonhakuun käytetyt menetelmät ja koodit sekä muu menetelmällinen tieto tarkkuudella, joka mahdollistaa aineiston kokoamisen uudelleen alkuperäislähteistä – mikäli aineisto sosiaalisen median tapauksessa alkuperäislähteistä edelleen löytyy.

9. Lähteet

Eilo, H., Falin, P., Harju, A., Huuskonen, S., Kohonen, I., Laine, K., & Satama, M. (2026). Näkökulmia sosiaalisen median eettiseen ja vastuulliseen tutkimukseen. Zenodo. https://doi.org/10.5281/zenodo.18386110

Harenko, K., Kivistö, M., Niiranen, V., & Tarkela, P. (2024). *Tekijänoikeus* (3., uud. p.). Alma Talent.

Hyvä tieteellinen käytäntö ja sen loukkausepäilyjen käsitteleminen Suomessa, Tutkimuseettisen neuvottelukunnan HTK-ohje 2023. Tutkimuseettisen neuvottelukunnan julkaisuja 2/2023. https://tenk.fi/sites/default/files/2023-03/HTK-ohje_2023.pdf

Ihmiseen kohdistuvan tutkimuksen eettiset periaatteet ja ihmistieteiden eettinen ennakkoarviointi Suomessa, Tutkimuseettisen neuvottelukunnan ohje 2019. Tutkimuseettisen neuvottelukunnan julkaisuja 3/2019. https://tenk.fi/sites/default/files/2021-01/Ihmistieteiden_eettisen_ennakkoarvioinnin_ohje_2020.pdf

Lainsäädäntö

Euroopan Parlamentin Ja Neuvoston asetus (EU) 2016/679, annettu 27 päivänä huhtikuuta 2016, luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta (yleinen tietosuoja-asetus) https://eur-lex.europa.eu/legal-content/FI/TXT/HTML/?uri=CELEX:02016R0679-20160504

Tietosuojalaki (1050/2018)
https://www.finlex.fi/fi/lainsaadanto/2018/1050

Tekijänoikeuslaki (404/1961) https://www.finlex.fi/fi/lainsaadanto/1961/404