

JYUn ohje

aineistohallintasuunnitelman (DMP)

laatimiseen

Tämä organisaatiokohtainen ohje aineistohallintasuunnitelman kirjoittamiseen JYUssa noudattaa yleisen suomalaisen DMP-mallin pohjaa. Esimerkkilauseet on tarkoitettu konkretisoimaan ja herättämään ajatuksia. Älä käytä niitä sellaisenaan, vaan muokkaa niitä oman aineistosi tarpeiden mukaisiksi. Kirjoita vain lauseita, jotka ymmärrät itse.

1. Aineiston yleiskuvaus

1.1 Millaiseen aineistoon tutkimuksesi perustuu? Millaista aineistoa keräät, tuotat tai käytät uudelleen? Missä tiedostomuodossa aineisto on? Esitä myös karkea arvio tuotettavan/kerättävän aineiston koosta.

Toimi näin:

- Kuvaile listauksen tai taulukon avulla, millaisia erilaisia aineistotyyppisiä a) käytät uudelleen (esim. arkisto- ja rekisteriaineistot), b) keräät ja tuotat tutkimuksen aikana sekä c) millaista tulosaineistoa analyysistasi syntyy. Tarkista, että kuvaus vastaa tutkimussuunnitelmaasi.
- Jos keräät ja käsittelet henkilötietoja tai muutoin luottamuksellisia tai salassa pidettäviä tietoja, erittele ne listaan tai taulukkoon arkaluonteisuuden perusteella.
- Käytä mahdollisimman avoimia, standardimuotoisia tiedostomuotoja. Voit hakea sopivaa formaattia [UK Data Servicen valitsimen](#) avulla. Jos aineiston avaaminen ja käsittely edellyttää jonkin tietyn kaupallisen ohjelmiston käyttöä, perustele lyhyesti taulukon alla syys siihen.

- Jos käytät jotain tiettyä laitteistoa aineiston keruuseen tai tuottamiseen, kuvaile tiiviisti, millaista dataa se tuottaa. Jos data noudattaa jotain metadatastandardia, mainitse myös se. Tärkeintä tässä kohdassa on, että ymmärrät, mitä kaikkia aineistotyyppisiä sinulla on ja sen, että aineistotyyppit edellyttävät erilaisia aineistohallinnan toimenpiteitä.

Esimerkki: Aineistotaulukko

Datatyypit	Lähde	Henkilötiedot / arkaluonteinen tai erityisen luottamuksellinen data	Tiedostomuoto	Arvioitu koko ja karttuma
Analysoitu DNA-näyte	Tuotettu DNA-näytteestä	Ei	.xlsx, .csv	2 Gb
Tilastoaineisto X	Valmis aineisto Tietoarkistosta, lisenssi: saatu tutkimuskäyttöön	Ei, anonymisoitu	SPSS (.por, .sav)	
Kysely	Kerätty tutkittavilta	Kyllä, suoria tunnisteita ja terveystietoja	.csv	5 Mt
Haastattelutallenne, video	Kerätty	Kyllä, suoria tunnisteita, ei sensitiivisiä (kasvokuva, ääni)	.avi, .mp4	
Haastattelutranskriptio	Tuotettu	Ei, anonymisoitu	.csv, .txt, .xlsx	>10 Mt
Valokuvat	Tuotettu	Ei	.tif, .jpeg, .gif, .raw	~5 Mt
Hallinnolliset dokumentit	Tutkittavilta kerätyt suostumuslomakkeet, tietosuojailmoitukset, tiedotteet, käsittelysopimukset	Kyllä, suoria tunnisteita	.docx	

Esimerkkejä:

“Arkistomateriaali Kansallisarkistosta on joko digitoitu tai mikrofilmillä. Materiaalin käyttö käsillä olevaan tutkimukseen ei edellytä erillistä lupaa. Käytän myös avointa aineistoa X- ja Y-rekistereistä. Nämä rekisteriaineistot ovat saatavilla avoimesti joko verkossa tai painettuina asiakirjoina.”

“Hyödyntääkseni mikrofilmillä olevaa arkistomateriaalia projektissani otan valokuvat dokumenteista arkistossa tai kirjastossa. Vaihtoehtoisesti skannaan mikrofilmatut sivut USB-laitteelleni, jolta siirrän aineiston ensi tilassa yliopiston verkkoasemalla olevaan aineistokansioon. Muut keskeiset materiaalit kuten lakitekstit ja uudempi kirjallisuus, ovat haettavissa käyttöön kirjastotietokannoista tai tilattavissa suoraan kirjastoista PDF-muodossa.”

”Laadin arkistomateriaalin pohjalta Excel-tietokannan, johon taulukoin tiedot kaikista tutkimushenkilöistäni. Käytän tietokantaa pohjana, josta suodatan analyysiaineiston.”

“The interactions will be coded with standardized CLASS and ECCOM observation instruments. Parts of the interactions will also be coded and analysed by using conventions used in discourse and conversation analysis. Recall interviews will be transcribed and content coded. Eye-tracking data will be coded according to the teacher fixations and saccades on areas of interests.”

“Data consists of generally accepted formats. Most of the data will be quantified with SPSS program. The data formats for SPSS-data are sav-format and for video and audio recordings .mp4 and .mp3 format. Transcripts and content codings will be in .docx format. Eye-tracking recording are in json and .mp4 format (Tobii Analysis Pro software is needed for analyzing the eye-tracking data). Quantitative data will be processed and analyzed by using SPSS and MPlus statistical programs; Atlas-ti to analyze and handle qualitative data. All other materials, such as questionnaire are saved as .pdf-files. Used software and formats are based on open standards to enable data reuse, interoperability and sharing. The data content and specific methods as well as analyses are described in more detail in the research plan.”

1.2 Miten aineiston yhtenäisyys ja laatu varmistetaan?

Toimi näin:

- Kuvaile tarvittaessa, miten minimoit riskin siitä, että aineistoon tulee virheitä tai hävikkiä esim. kun siirrät sitä laitteelta toiseen (tutkimuslaitteiden kalibrointi, transkription tarkistuttaminen toisella asiantuntijalla, puuttuvien tai virheellisten arvojen tarkistaminen taulukkoaineistosta pistokokein, jne.) Älä jää suotta jumiin tähän kohtaan, sillä kaikissa tutkimuksissa erillinen laadunvarmistus ei aina ole tarpeen. Pääasia on, että päädyt pohtimaan mahdollisia tilanteita, joissa aineisto voi muuntua tahattomasti niin, että sen sisältö vaarantuu.
- Sovi tutkimusryhmän jäsenten kesken yhteiset käytännöt aineiston siirtoon, analyysiin ja muuhun käsittelyyn, jotta vältätte erilaisista kirjaus- ja analyysitavoista aiheutuvat virheet.

Esimerkkejä:

“Projektissa X käytetään ohjelmistoja ja tiedostoformaatteja, jotka ovat avoimien, yleisesti käytössä olevien standardien mukaisia. Näin varmistetaan, että aineisto on saavutettavaa, teknisesti yhteentoimivaa ja uudelleen käytettävää.”

“Projektin nimeää vastuuhenkilön, joka huolehtii siitä, että laadunvarmistustoimenpiteistä huolehditaan.”

“Tiedostojen tallentamisessa käytetään tarkistussummia, joiden avulla varmistetaan se, ettei data muunnu, kun sitä kopioidaan, siirretään ja tallennetaan.”

“Nimetty vastuuhenkilö huolehtii haastattelutranskriptioiden tarkistuksesta.”

“Kun data konvertoidaan tiedostoformaattista toiseen, huolehditaan siitä, että data säilyy alkuperäisessä muodossa.”

2. Eettisten periaatteiden ja lainsäädännön noudattaminen

2.1 Mitä juridisia seikkoja liittyy aineiston hallintaan (esim. EU:n yleinen tietosuoja-asetus ja muu aineiston käsittelyyn liittyvä lainsäädäntö)?

[Perehdy JYUn tutkimuseettisen neuvottelukunnan ohjeisiin mahdollisen eettisen ennakoarvioinnin tarpeesta.](#)

[Perehdy JYUn Tutkijan tietosuojaohjeisiin, jotta osaat suunnitella toimenpiteesi niiden mukaisesti.](#)

[JYUn ohje käytännön tietoturvatyökaluihin](#) (Intranet Unossa).

[JYUn lakipalvelujen yhteystiedot](#)

a. Henkilötiedot

Tunnista, sisältääkö käyttämäsi tai keräämäsi ja käsittelemäsi aineisto henkilötietoja. Jos käsittelet henkilötietoja, tietosuojalainsäädäntö koskee tutkimustasi.

Henkilötietoja ovat kaikki tiedot, joista tutkittava on tunnistettavissa jonkin biologisen, sosiaalisen, kulttuurisen tai taloudellisen piirteen perusteella. Tutkittavat voivat olla tunnistettavissa suoraan (koko nimi, henkilötunnus, kasvokuva, ääni, sormenjälki jne.) tai epäsuorasti (ikä, sukupuoli, kotipaikka, siviilisääty, harvinainen sairaus, vanhempien nimet, poliittiset ja maailmankatsomukselliset mielipiteet jne.) Tiedot voivat olla tunnisteellisia, vaikka sinä tai tutkittava itse eivät pystyisi tekemään tunnistamista – riittää, että joku voisi tietojen perusteella tunnistaa henkilön (esim. aivokuva).

Kuvaile tässä kohdassa, mitä henkilötietoja keräät, ja miten huolehdit niiden suojaamisesta käytännössä sekä pseudonymisoi tai anonymisoi tiedot. Sinun ei

tarvitse toistaa tiedotetta ja tietosuojailmoitusta; voit tässä keskittyä kuvaamaan käytännön toimenpiteet, joilla suojaat tiedot ja tunnistat, että henkilötietojen käsittelylle tulee määritellä mahdollisimman tarkka elinkaari, jossa on alku ja loppu.

Esimerkkejä:

“Koska lähdemateriaali [tarkenna tarvittaessa, mitkä osat aineistosta] on saatavilla julkisissa arkistoissa ja kirjastoissa ja kohdehenkilöt ovat kuolleet, arkistoaineiston käsittelyyn ei liity juridisia, oikeuksiin liittyviä tai tutkimuseettisiä kysymyksiä.”

“Tutkimuksessa ei käsitellä henkilötietoja.”

“Tutkimuksessa käsitellään henkilötietoja. Tutkittavilta kerätään suoria ja epäsuoria tunnisteita. Kerättävä aineisto sisältää mm. henkilöiden nimet, yhteystiedot, asuinpaikan, tietoa koulutustaustasta sekä psykologisia ja taloudellisia tietoja.”

“Aineisto kerätään suoraan osallistujilta ja potilasrekistereistä osallistujilta saadun suostumuksen mukaisesti. Osallistujilta on pyydetty tietosuoja-asetuksen mukainen nimenomainen suostumus tietojen käsittelyyn. Suostumuslomake sisältää omat sarakkeensa kotona tehdyille haastattelulle, terveyskeskuksessa tehdyille arvioinneille sekä potilasrekisteritiedoille.”

b. Sisältääkö aineisto erityisiin henkilötietoryhmiin kuuluvia henkilötietoja (ns. sensitiiviset/arkaluonteiset henkilötiedot)?

Jos käsittelet erityisiä henkilötietoryhmiä, kuvaile, miten huolehdit niiden suojaamisesta ja pseudonymisoinnista. Mainitse oikeusperuste tietojen käsittelylle (yleinen etu vai nimenomainen suostumus).

Erityisiä henkilötietoryhmiä ovat seuraavat:

1. rotu tai etninen alkuperä
2. poliittiset mielipiteet

3. uskonnollinen tai filosofinen maailmankatsomus
4. ammattiliiton jäsenyys
5. bionäytteen analyysissa tuotettu geneettinen data
6. biometrinen data, jota käsitellään henkilön tunnistamiseksi
7. terveystiedot
8. seksuaalista suuntautumista tai käyttäytymistä koskevat tiedot.

Joa käsittelet näitä tietoja, niiden käsittelyn on perustuttava eurooppalaisen tietosuojasetuksen 9. artiklaan. Käsittelyn oikeusperusteeksi suositellaan aina ensisijaisesti yleisen edun mukaista tieteellistä tutkimustarkoitusta. Jos se ei syystä tai toisesta sovellu, perusteeksi valitaan tutkittavien nimenomainen suostumus.

Esimerkki:

“Tutkimuksessa kerätään ja käsitellään erityisiä henkilötietoryhmiä. Kerätyt henkilötiedot sisältävät tietoa mm. nimistä, asuinpaikasta ja terveydestä.””Tutkimuksessa kerättävät erityiset henkilötiedot anonymisoidaan analyysin päättyessä karkeistamalla tiedot sairauksista sairaustyypeittäin. Nimet ja sähköpostiosoitteet poistetaan tutkimuksen päättyessä.”

c. Tiedot rikostuomioista ja rikoksista

Tunnista, sisältävätkö keräämäsi henkilötiedot tietoja rikostuomioista ja rikoksista, sekä oikeusperuste tietojen käsittelylle. Rikostietojen käsittely edellyttää tutkittavien suostumusta.

Esimerkki:

“Tutkimus sisältää tietoja rikostuomioista ja rikoksista, joita käsitellään tutkittavien suostumuksella.”

d. Henkilötietojen käsittely ja siihen liittyvät roolit

Tunnista ja kuvaile roolit henkilötietoja ja eri henkilötietoryhmiä sisältävien aineiston osien käsittelyssä.

Rekisterinpitäjä tarkoittaa luonnollista tai oikeushenkilöä, joka määrittää yksin tai yhdessä toisen tahon kanssa tarkoituksen ja tavan henkilötietojen käsittelylle. Kun kaksi tai useampi rekisterinpitäjä määrittää yhdessä tietojen käsittelyn tarkoituksen, ne toimivat **yhteisrekisterinpitäjinä** (esim. konsortiohankkeet, joissa on mukana useita tutkimusorganisaatioita).

Huolehdi tarvittaessa tietosuojasetuksen edellyttämien aineistokäsittelysopimusten (data processing agreement, DPA) tekemisestä, jos tietojen käsittely edellyttää niiden siirtämistä ulkopuoliselle **henkilötietojen käsittelijälle** (esim. analyysipalvelu, transkriptiopalvelut). Esimerkiksi projektissa työskentelevä tutkimusavustaja ei ole käsittelijä, vaan käsittelijä on tutkimuksen ulkopuolinen taho. Käsittelijä käsittelee tietoja rekisterinpitäjän lukuun, eikä päätä itse tietojen käsittelyn tarkoituksesta.

Esimerkki:

“Yliopisto A toimii henkilötietojen rekisterinpitäjänä. Tietojen käsittelijöinä rekisterinpitäjän lukuun toimivat verkkokyselypalvelu Webropol, yliopisto B, jonka laboratorio huolehtii datanäytteiden analyysistä, sekä Tietoarkisto, jossa aineisto tullaan avaamaan anonymisoituna tutkimuksen päättymisen jälkeen. Käsittelijöiden kanssa on tehty tietosuojasetuksen mukainen aineistokäsittelysopimus.”

e. Mistä henkilötiedot on saatu?

Kuvaile, mistä lähteistä keräät käsittelemäsi henkilötiedot. Tiedot voidaan saada joko **suoraan tutkittavilta** (haastattelut, mittaukset, kyselyt, havainnointitutkimukset jne.), tai **välillisesti** rekistereistä, arkistoista ja kirjastoista.

Esimerkki:

“Käsiteltävät tiedot saadaan suoraan tutkittavilta heiltä saatavan suostumuksen nojalla.”

f. Henkilötietojen suojaustoimet

Kuvaile, mitä hallinnollisia ja teknisiä tietoturvatavoimia tulet toteuttamaan, jotta suojaat henkilötiedot tutkimuksen kaikissa vaiheissa. Tunnista tietojen suojaan kohdistuvan riskin taso (mitä arkaluonteisempia ja mitä enemmän tietoja on, sitä suurempi riski) ja minimoi ne. Pyri toteuttamaan henkilötietojen käsittelyssä [minimointiperiaatetta](#). Sen mukaisesti keräät tietoja vain tutkimustarkoitustasi varten tarvittavan määrän, [pseudonymisoi](#) ne heti kun mahdollista, ja [anonymisoi](#) tai hävität tunnisteet mahdollisimman pian, mikäli se on aineiston luonne huomioiden mahdollista. Jollet voi pseudonymisoida tietoja, perustele syy siihen (esimerkiksi jo julkistetut tiedot, joita kerätään sanomalehtiarkistosta). Suojaustoimet voivat olla mm. seuraavia:

- tutkimusprojektin jäsenten perehtyminen JYUn tietosuoja- ja tietoturvaohjeisiin (Tietosuoja tieteellisessä tutkimuksessa -itseopiskelimuoduli Moodlessa, henkilöstön tietoturvakoulutukset, Tutkijan tietosuojaohjeet, henkilöstön tietoturvaohjeet
- asianmukaiset tietoturvatavoimet käytettyjen tallennuspaikkojen mukaisesti
- toimenpiteet, joilla estät asiattoman pääsyn henkilötietoihin
- tarvittaessa erillisen vaikutustenarvioinnin (Data protection impact assessment, DPIA) tekeminen ([tarkista, tarvitseeko tutkimuksesi DPIA:n](#))
- tutkimukselle määritellään vastuullinen johtaja ja asianmukainen tutkimussuunnitelma
- henkilötietojen säilytyksen rajoittaminen tietosuoja-asetuksen mukaisesti: määritä henkilötietojen säilytykselle ajallinen päätepiste, jonka jälkeen aineisto joko anonymisoidaan tai hävitetään tai, jos henkilötietoja on tarpeen säilyttää jatkotoimia tai tulosten verifiointia varten tutkimuksen jälkeen, määrääjat tietojen säilytyksen tarpeellisuuden arvioinnille
- huolehdi siitä, ettei tutkittavia voi tunnistaa tutkimusjulkaisuista

- tietosuojasetuksen edellyttämät sopimukset esim. aineiston käsittelystä on tehty hyvissä ajoin.

Esimerkkejä:

“Aineisto pseudonymisoidaan analyysivaiheessa. Pseudonyymiä aineistoa säilytetään [määrittele missä; esim. JYUn Nextcloudissa projektin ryhmäkansiossa/JYUn S:-asemalle tilattavassa projektikansiossa]. Aineistoa ei ole mahdollista anonymisoida, sillä kyseessä on seurantatutkimus. Henkilötietojen suojaustoimia ovat...”.

“Kaikki suorat henkilötunnisteet pseudonymisoidaan koodilla. Koodiavainta tullaan säilyttämään JYUn Nextcloudissa pääsyvalvotussa kansiossa erillään pseudonyymistä aineistosta.”

”Tiedot pseudonymisoidaan keruun jälkeen, ja pseudonymisointiavainta säilytetään erillään pseudonyymistä aineistosta. Tunnisteelliset tiedot hävitetään heti siinä vaiheessa, kun ne eivät enää ole välttämättömiä tutkimuksen tarkoitusta varten.”

”Tiedot pseudonymisoidaan keruun jälkeen, ja pseudonymisointiavainta säilytetään erillään pseudonyymistä aineistosta. Tunnisteelliset tiedot hävitetään heti siinä vaiheessa, kun ne eivät enää ole välttämättömiä tutkimuksen tarkoitusta varten.”

“Kaikki henkilötunnukset ja muut tiedot, joista yksittäisen haastateltavan voi tunnistaa, tullaan poistamaan heti siinä vaiheessa, kun niitä ei enää tarvita tutkimustarkoituksen toteuttamista varten. Yksityiskohtainen anonymisointisuunnitelma on kuvailtu tietosuojailmoituksessa.”

g. Milloin tarvitset vaikutustenarvioinnin (DPIA)?

Selvitä, tarvitseeko tutkimuksesi erillisen vaikutustenarvioinnin (DPIA) tekemistä. DPIA auttaa arvioimaan käsiteltävien henkilötietojen suojaan kohdistuvia riskejä. Tietosuojalainsäädännön mukaisesti DPIA on tehtävä siinä tapauksessa, että tietojen

suojaan ja tutkittavien vapautteen kohdistuu merkittävä niitä uhkaava riski. [Tutustu Tutkijan tietosuojaohjeeseen.](#)

Korkea riski täyttyy, jos kaksi tai useampia seuraavista kriteereistä täyttyy:

- suuri määrä tutkittavia
- suuri määrä henkilötietoja tutkittavaa henkilöä kohden
- erityisten henkilötietojen kuten terveystietojen käsittely
- tietojen kerääminen haavoittuvilta ihmisryhmiltä (esim. lapset)
- automaattinen päätöksenteko (esim. luottopäätökset)
- tutkittavan systemaattinen seuranta.

Lisäksi DPIA edellytetään erityisesti seuraavissa tapauksissa:

- henkilötietojen käsittelyssä hyödynnetään uusia teknologioita
- rikosseuraamustietojen tai erityisten henkilötietoryhmien (rotu tai etninen alkuperä, poliittinen tai filosofinen katsomus, ammattiliiton jäsenyys, geneettiset ja biometriset tiedot, terveystiedot, seksuaalista suuntautumista tai käyttäytymistä koskevat tiedot) laajamittainen käsittely
- henkilön luonteenomaisten piirteiden systemaattinen ja laajamittainen arviointi hyödyntäen automaattista prosessointia (mukaan lukien profilointi), kun on mahdollista, että tietojen pohjalta tehdään päätöksiä jotka aiheuttavat juridisia toimenpiteitä henkilöä kohtaan
- julkisen paikan laajamittainen systemaattinen monitorointi
- Aina, kun tutkimuksessa on perusteltu syy poiketa tutkittavan oikeudesta tulla informoiduksi häntä koskevien tietojen käsittelystä.

Jos vaikutustenarviointi osoittaa, että tutkittavien tietoihin ja vapautteen kohdistuu merkittävä riski, arviointi tulee toimittaa kansalliselle tietosuojavaltuutetulle.

Lisätietoja saat JYUn tietosuojavastaavalta: tietosuoja@jyu.fi.

Esimerkki:

“Projekti aiheuttaa korkean riskin tutkittavien oikeuksille ja vapaudelle, koska tutkimus käsittelee haavoittuvassa asemassa olevien ihmisryhmien (lapset) henkilötietoja. Ennen tietojen keruun aloittamista tutkimukselle toteutetaan tietosuojalainsäädännön edellyttämä vaikutustenarviointi.”

2.2 Miten hallinnoit käyttämäsi, tuottamasi ja jakamasi aineiston oikeuksia?

Kuvaile tässä osiossa, mitä oikeuksia, lisenssejä ja sopimuksia käyttämäsi, keräämäsi ja käsittelemäsi aineistoon liittyy. Tutustu hyvissä ajoin [JYUn sopimusmallipohjiin](#) ja [kevyeen mallipohjaan, jolla tutkijat voivat sopia keskenään tuottamansa aineiston tekijyydestä ja käyttöoikeuksista](#). Tämänhetkisen lainsäädännön tulkinnan mukaan tutkimusdataa ei pääsääntöisesti koske tekijänoikeus, pois lukien esim. valokuvat ja omaperäiset tietokannat. Aineiston omistajuus on samoin asia, joka on vielä kansallisesti määriteltävänä. Siksi on tärkeää sopia mahdollisimman aikaisessa vaiheessa kirjallisesti siitä, kuka/ketkä/mitkä kaikki tahot hallinnoivat oikeuksia aineiston käyttöön ja jakeluun.

Toimi näin:

- Jos käytät olemassa olevaa primääriaineistoa esimerkiksi rekistereistä, arkistoista, kirjastoista, yksityisiltä tuottajilta tai sosiaalisen median palveluista, selvitä tässä, että olet tietoinen käyttöoikeuksistasi ja lisenssistä, jolla olet saanut aineiston käyttöösi. Kuvaile tarvittaessa, miten noudatat aineiston tarjoajan ehtoja (tekijänoikeudet, käyttörajoitukset jne.)
- Kuvaile, kuka toimii aineiston oikeudenhaltijan roolissa. Ulkopuolista rahoitusta nauttivissa tutkimushankkeissa tuotettujen aineistojen oikeudet siirtyvät tutkijan kotiorganisaatiolle hankesopimuksen mukaisesti. Myös tällöin on hyvä kirjata sopimukseen tutkijoiden oikeudet aineiston käyttöön ja jakamiseen.
- Kuvaile, että kaikissa tapauksissa aineiston hallinnointiin, käyttöön ja jakeluun liittyvistä oikeuksista on sovittu kirjallisesti. Sopimus selventää kaikkien osapuolten oikeudet ja roolit, ja on edellytys aineiston tulevalle lisensoinnille ja siten

(jatko)käytettävyydelle. Kun oikeuksista on sovittu, aineiston on mahdollista täyttää JYUn ja rahoittajien edellytykset FAIR-periaatteiden mukaisista tutkimusaineistoista.

- JYU suosittelee aineiston lisensointiin yleisesti käytössä olevia lisenssejä kuten Creative Commons -lisenssiä tai Data Commons -lisenssiä. [Tarkista Creative Commonsin lisenssivalitsimella sopiva lisenssi.](#)

Esimerkki:

“Aineiston oikeuksista ja lisensoinnista tullaan sopimaan kirjallisesti ennen aineiston keruun aloittamista. Kun aineisto avataan soveltuvassa digitaalisessa data-arkistossa tutkimuksen päättyessä, se tullaan lisensoimaan Creative Commons 4.0 Nimeä Kansainvälinen -lisenssillä.”

3. Dokumentointi ja metadata

Miten dokumentoit eli kuvailet aineistoasi tutkimuksen aikana, jotta varmistat, että se on mahdollisimman löydettävää, saavutettavaa, yhteentoimivaa ja käyttökelpoista (findable, accessible, interoperable, reusable eli FAIR) sekä itsellesi että muille?

Muistilista:

- Kaikkien uusien Jyväskylän yliopistossa toteutettavissa tutkimusprojekteissa syntyvien aineistojen peruskuvailutiedot kirjataan tutkimustietojärjestelmä Converikseen. Näin varmistat tutkijana, että tieto aineistostasi pysyy tutkimuksen aikana ajantasaisena ja järjestyksessä. Converiksen avulla myös julkaiset metatiedot sekä mahdollisuuden mukaan itse tutkimusaineistosi. Kun pyydät metatietojen julkaisemista Converiksessa tutkimuksesi lopulla, Avoimen tiedon keskus julkaisee ne

puolestasi yliopiston julkaisuarkisto JYXissä. JYXissä metatiedot varustetaan pysyvällä tunnisteella (DOI ja URN), ja ne saavat pysyvän laskeutumissivun. Näin toteutat helposti rahoittajiesi edellytyksiä aineiston metatietojen avaamisesta välittömästi projektin päättyessä. Julkaistut metatiedot siirtyvät automaattisesti kansallisiin aineistokuvailuindekseihin kuten METAXiin, ja ne tulevat näkyville Tiedejatutkimus.fi-portaalissa.

- In addition, describe your planned procedures for methodological 'What, Why, Where, When, and by Whom' of what you do to and with the data (field and lab notebooks, embedded SPSS metadata, working instructions, protocols, README files, field-specific standard metadata, etc.)

Esimerkkejä:

"The research group uses hard-bound laboratory log books as the primary documentation of daily research activities. In these log book, each scientist records the experiments done, the materials used and identifies the names of datafiles and where they are stored. Additionally, README.txt files will be used to store structural (naming conventions and purpose of folders and files) and administrative (e.g., when the files and folders were created, file types and other technical information, and who can access it) metadata. Every main folder and sub-folder will include a README.txt file to serve as a guide on the data and which will be updated to any changes made within its respective folder. Any additions, deletions or changes to naming conventions or purposes are immediately updated to README.txt."

"The project shall name a team member responsible for monitoring and enforcing these data documentation rules. These documentation and metadata and their publication will ensure FAIRness of the data."

"When anyone saves data on the project's shared network storage space (provided and technically administered by University of Jyväskylä, including backups, access control and security), that person shall also update and maintain this documentation:

The storage root as well as every sub-folder shall include a README.txt file. Each README.txt file describes the current naming conventions and purpose of all folders and files in that level of storage organization. Any additions, deletions or changes to naming conventions or purposes are immediately updated to README.txt. [about “master data files”: if you use version control software, or otherwise use a method to consistently distinguish work copies from master data files, you should mention and describe that here briefly.]

Each master data file [or folder depending on type of data] will be clearly named as such [you can enter your naming convention here], and each will always be accompanied, in the same folder, by two files:

1) ABOUT_filename.txt/ABOUT_foldername.txt which contains a link to an initialized metadata entry in the University's Converis research information system. The preliminary entry can remain private and unpublished, just for your own use – but if and when you want to publish your data, it is easy because the metadata entry is already there.]

2) METHODS_filename.txt which/METHODS_foldername.txt contains link to, or text of, description of methods used to obtain, process and document the data, in detail sufficient to allow replication and understanding and usage of the data for other purposes.

Both the ABOUT and the METHODS files, and the preliminary metadata entry in Converis, will be created immediately as the data gathering begins and amended and updated throughout the research project as data is building up. These will then be readily available to convert to machine-readable formats upon archival or publication of the metadata in the University's repository JYX, and the eventual publication of the data in [name chosen repository].”

4. Tallentaminen ja varmuuskopiointi tutkimushankkeen aikana

4.1. Minne aineistosi tallennetaan, ja miten se varmuuskopioidaan?

Kuvaile, minne tallennat aineistosi tutkimuksen aikana, ja miten huolehdit varmuuskopioista. JYUssa ensisijaiset tallennuspaikat ovat Nextcloud-pilvitallennuspalvelu, arkaluonteisille aineistoille soveltuva CollabRoom-pilvitallennuspalvelu, yliopiston S:- ja U:-verkkoasemat sekä koodin tallentamiseen tarkoitettu GitLab. Tutkimusmoniviestin soveltuu ei-arkaluonteisen videomateriaalin jakamiseen tutkimuksen aikana.

Näin otat henkilökohtaisen Nextcloud-tilisi käyttöön tai haet Nextcloud-oikeuksia: [Nextcloudin käyttöönotto – Digipalvelut \(jyu.fi\)](#).

Nextcloud-ryhmäkansiot ovat suositeltu ratkaisu aina silloin, kun aineistoa kokoa ja työstää useampi kuin yksi henkilö. Ryhmäkansioissa aineistot pysyvät yliopiston hallussa siinäkin tapauksessa, että kaikki kansioita hallinnoivat yksittäiset henkilöt lähtevät yliopistolta. Tutustu ryhmäkansioiden mahdollisuuksiin ja tilausohjeisiin: [Miten ryhmäkansioita saa käyttöön? – Digipalvelut \(jyu.fi\)](#).

CollabRoom-ohjeet (Intranet Unossa): [CollabRoom - Tietoa ohjelmasta – Intranet Uno \(jyu.fi\)](#).

Esimerkkejä:

"Data (except sensitive data) of the project will always be stored in Nextcloud cloud storage service provided and managed by the Digital Services of the University of Jyväskylä. Cloud service is also ideal solution for sharing research data between collaborators."

[IF project handles sensitive data:] "Sensitive data is stored and operated in

university's highly secure CollabRoom environment that has been tailored especially for this task."

"The University's systems will take automatic backups of the data to prevent catastrophic loss of data. In addition, manual backups of master data files will be taken regularly [insert suitable time interval] and always before any major file-format or data conversions. Non-sensitive video data will be stored in the University's Researchvideo.jyu.fi system."

4.2 Kuka valvoo pääsyä aineistoon, ja miten suojattua pääsyä aineistoon valvotaan?

Esimerkkejä:

"Right to access the data is controlled by responsible researcher [PI, or someone specially dedicated to this], and technical access control is provided by the Digital Services of University of Jyväskylä."

"[Non-sensitive] data will be available to all researchers of the project via Nextcloud cloud storage."

"[In case of sensitive data] PI acts as administrator of project's CollabRoom space and will define who will need and get access to the sensitive data in that workspace."

"Access to the data will be documented and PI will be at any point able to tell who has access to what data."

5. Aineiston avaaminen, julkaiseminen ja arkistointi tutkimushankkeen päätyttyä

5.1 Mikä osa aineistosta voidaan asettaa avoimesti saataville tai julkaista? Missä ja milloin aineisto tai siihen liittyvät metatiedot asetetaan saataville?

Muistilista:

- Converikseen kirjaamasi metatiedot voidaan julkaista välittömästi tutkimuksen päättyessä tai heti siinä vaiheessa, kun niihin ei enää tule muutoksia. Voit siis mainita tässä osiossa aluksi tämän.
- Jos aineistosi ei luonteensa vuoksi sovellu julkaistavaksi avoimesti saataville, kuvaile tässä kohdassa syyt siihen (esim. tietosuoja tai liikesalaisuus). JYUn tutkimusdatapolitiikan mukaan aineistot tulee pyrkiä saattamaan avoimesti saataville aina, kun mahdollista. **Jollet voi avata aineistoa esimerkiksi tietosuojasyistä, metatietojen avaaminen Converiksen kautta JYXiin kattaa yliopiston ja rahoittajien avoimen saatavuuden edellytykset.**
- Kuvaile lyhyesti, mitä osia aineistosta säilytät määräajan (esim. 1-5 vuotta) tutkimustulosten verifiointia varten.
- Jos aiot avata aineiston jossain data-arkistossa, mainitse se myös tässä.
- Voit käyttää apuna seuraavaa jäsenystä:

A) Minkä osan aineistosta hävität tutkimuksen lopulla

B) Minkä osan aineistosta säilytät tulosten verifiointia varten (esim. 1-5 vuotta)

C) Minkä osan aineistosta avaat uudelleenkäyttöä varten jossain soveltuvassa, kuratoidussa data-arkistossa kuten Dryadissa, Tietoarkistossa tai JYXissä (näissä kuratoitu säilytysaika on noin 25 vuotta).

Esimerkkejä:

“Metadata entries of the data will be published immediately in the University's JYX publication repository when they are considered sufficiently complete, even if the data itself is not yet public. (For description of creation and curation of metadata entries please see Section 3 of this plan.) The metadata will then be searchable in e.g. the national Etsin and Research.fi metadata catalogues.”

“Dataset(s) themselves, complete with full description of methods, will be published in [JYX / a certified, field-specific repository / Zenodo].”

“Sensitive parts of the data will be anonymized and published along with other data.”

“Sensitive parts of the data cannot be anonymized and thus cannot be openly published.

Sensitive data will be stored in JYX, only the metadata will be public, and access to the data can be requested and granted on certain conditions [describe the procedure and terms on which access and right to use the data can be granted] .”

“Sensitive parts of the data cannot be stored and will be disposed of after the project is finished.”

5.2. Missä pitkällä aikavälillä arvokas data säilytetään ja kuinka pitkään?

Pitkäaikaissäilytys tarkoittaa, että aineisto säilytetään niin pitkään kuin on tarpeellista - jopa useiden vuosikymmenten tai -satojen ajan. Jyväskylän yliopistolla ainiestojen arviointikäytäntö kansalliseen Fairdata-PAS-pitkäaikaissäilytykseen on parhaillaan kehitteillä, eikä prosessi vielä ole toiminnassa. Lisätietoa pitkäaikaissäilytykseen siirtämisen arviointiprosessin käyttöönotosta on tulossa vuoden 2022 aikana.

Esimerkki:

“University of Jyväskylä will store all data archived in JYX indefinitely (for minimum of 10 years), in the format originally deposited in, but does no special packaging or

continuous curation to guarantee long-term integrity and usability. Data with long-term value will be proposed to national Fairdata-PAS (system for true long-term research data storage). National and university-level policies for determining eligibility for Tutkimus-PAS are currently being developed, but not available yet.”

6. Aineistonhallintaa koskevat vastuut ja resurssit

6.1 Kuka (esim. tehtävä ja laitos) vastaa aineistonhallinnasta?

Kerro tässä lyhyesti, miten edellisissä vastauksissa kuvatut tehtävät ja vastuut on jaettu. Jos vastaat aineiston hallinnasta yksin, esimerkiksi laatiessasi itsenäisesti väitöstutkimusta, riittää kun kerrot, että toimit itse vastaavana henkilönä.

Esimerkkejä:

“Responsibilities for specific issues have been described in earlier sections.” [Make sure this is so. If not, describe them here.]

“The project will name a team member responsible for monitoring and implementing data documentation and data management at both consortium sites. The PI and the SL will have the main responsibility of data management, monitoring it and granting access to the data. The costs associated with storing and sharing research data are regarded as overheads for the project’s host organizations (JYU and X). The preparation of the data will be supported by the Open Science Centre of JYU.”

6.2 Mitä resursseja aineistonhallinta edellyttää, jotta voit varmistaa, että aineisto voidaan avata ja säilyttää FAIR-periaatteiden (ks. edellä) mukaan?

Arvioi resurssit (esim. raha ja aika), joita tarvitaan aineiston hallintaan, säilyttämiseen ja jakamiseen. Pohdi, mitä täydentäviä tietoteknisiä tiloja ja resursseja tarvitaan ja mitkä niiden kustannukset ovat. Jos aineistojen käsittely ja hallinta edellyttää erityisiä resursseja, esimerkiksi lisälaskentatehoa, ostettua analyysi- tai litterointipalvelua tms., voit eritellä ne tähän. Jos olet kuvaillut tarvittavat resurssit jo edellisissä kohdissa, riittää, kun viittaat aiempiin kohtiin.

Esimerkkejä:

“Final preparation of the metadata and publishing it will require [best estimate of time] of work. Publishing the metadata will be supported by Open Science Centre of University of Jyväskylä and publishing the research data by [name repository]. The chosen repository will describe the data in a standard metadata format and store them in a file format suitable for reuse.”

“Anonymization of sensitive data will require [best estimate of time and effort].”