



SEINÄJOEN AMMATTIKORKEAKOULU  
SEINÄJOKI UNIVERSITY OF APPLIED SCIENCES

Oskari Saarinen

---

## Tutkimusaineistot Laurean hankkeissa

Opinnäytetyö

Kevät 2022

SeAMK Liiketoiminta- ja kulttuuri

Tradenomi (AMK), Kirjasto- ja tietopalveluala



SEINÄJOEN AMMATTIKORKEAKOULU

## Opinnäytetyön tiivistelmä

Koulutusyksikkö: Liiketoiminta- ja kulttuuri

Tutkinto-ohjelma: Kirjasto- ja tietopalveluala

Tekijä: Oskari Saarinen

Työn nimi: Tutkimusaineistot Laurean hankkeissa

Ohjaaja: Ari Haasio

Vuosi: 2022

Sivumäärä: 59

Liitteiden lukumäärä: 3

---

Opinnäytetyössä tarkasteltiin tutkimusaineiston hallintaa Laurea ammattikorkeakoulun TKI-ympäristössä tapahtuvissa hankkeissa. Tavoitteena oli selvittää tutkimusaineistoja Laurean TKI-hankkeissa sekä miten aineistoa hallitaan. Tavoitteena oli myös selvittää mahdollisia haasteita aineistohallinnassa. Opinnäytetyö tehtiin toimeksiantona Laurean tieto- ja julkaisupalveluille.

Teoreettinen viitekehys on muodostunut tuomalla esiin aiempaa tutkimusta liittyen aineistohallintaan ja tutkimusaineistoon. Teoriaosuudessa keskeisiä teemoja ovat aineistohallinta, tutkimusaineisto, tutkimusaineiston elinkaari ja aineistohallintasuunnitelma. Tutkimusote työssä oli laadullinen ja aineistonkeruussa menetelmänä käytettiin teemahaastatteluita. Haastattelut toteutettiin Teams-alustalla ja niihin osallistui Laurean TKI-hankkeisiin osallistuneita työntekijöitä.

Työn tuloksista selviää, että yleisimpiä aineistoja, joita haastateltavien TKI-hankkeissa syntyy ovat kysely- sekä haastatteluaineistot. Myös muita aineistoja syntyi kuten kokeellista aineistoa sekä työpaja-aineistoja. Aineistojen säilytyksessä ilmeni erilaisia käytäntöjä. Aineistoja säilytettiin esimerkiksi kovalevyllä ja lukkojen takana sekä oman organisaation OneDrivessä, hankkeen verkkolevyllä ja omalla tietokoneella. Jotain aineistoa oli tallennettu tutkimusdata-arkistoon.

---

<sup>1</sup> Asiasanat: aineistohallinta, data, tutkimusaineisto, TKI-toiminta

SEINÄJOKI UNIVERSITY OF APPLIED SCIENCES

## **Thesis abstract**

Faculty: School of Business and Culture

Degree programme: Library and Information services

Author/s: Oskari Saarinen

Title of thesis: Research data in Laurea projects

Supervisor(s): Ari Haasio

Year: 2022

Number of pages: 59

Number of appendices: 3

---

This thesis examined the management of research data in research, development & innovation projects in Laurea University of Applied Sciences. The purpose was to investigate research data and the management of it in Laurea projects. The purpose was to also discover the possible challenges that occurs in the management of research data. The thesis was made as a commission to the information -and publishing services of Laurea.

The theoretical framework consists of a cursory run through on some of the current literature on research data management. Research data and research data management is also examined as well as an examination of research data lifecycle and data management plan. The research method used in this thesis to acquire the empirical data was a qualitative semi-structured interview. The interviews were held in the Microsoft Teams platform and the partakers were Laurea employees who also worked in the Laurea projects.

The results indicate that the most common research data from the interviewees research, development & innovation projects were questionnaires and interviews. Other data was for example empirical data and data from workshops. In the management of research data there were different kinds of practices. Data was stored for example in hard drive and behind a lock and also data was kept in organizational OneDrive, project cloud storage and in personal computers. Some data was stored in research data -archives.

---

<sup>1</sup> Keywords: Research data management, research data, RDI

## SISÄLTÖ

Opinnäytetyön tiivistelmä .....	2
Thesis abstract .....	3
SISÄLTÖ .....	4
Käytetyt termit ja lyhenteet.....	5
1 JOHDANTO .....	7
2 LAUREA AMMATTIKORKEAKOULU .....	9
3 TKI-TOIMINNASTA .....	12
4 AIEMPI TUTKIMUS .....	15
5 TEOREETTINEN VIITEKEHYS .....	18
5.1 Tutkimusaineistot ja tutkimusdata .....	18
5.1.1 Tutkimusaineiston ja tutkimusdatan määrittelyä.....	18
5.1.2 Tutkimusdatan kategorisointi .....	22
5.1.3 TKI-aineistot.....	26
5.2 Aineistohallinta ja tutkimusaineiston elinkaari .....	27
5.2.1 Miksi aineistohallintaa? .....	27
5.2.2 Tutkimusaineiston elinkaari.....	30
5.2.3 Aineistohallintasuunnitelma.....	32
6 TUTKIMUSONGELMA .....	35
7 TUTKIMUSMENETELMÄ JA TOTEUTUS.....	37
7.1 Tutkimusmenetelmä .....	37
7.2 Toteutus .....	38
8 TULOKSET .....	41
9 POHDINTA JA JOHTOPÄÄTÖKSET .....	46
LÄHTEET .....	48
LIITTEET .....	59

## Käytetyt termit ja lyhenteet

<b>ACRL</b>	Association of College & Research Libraries.
<b>Arene ry</b>	Ammattikorkeakoulujen rehtorineuvosto.
<b>CERN</b>	European Organization for Nuclear Research.
<b>DMPtuuli</b>	Aineistonhallinnan työkalu.
<b>ERC</b>	European Research council (Euroopan tiedeneuvosto)
<b>FAIR</b>	Löydettävä (findable), saavutettavaa (accessible), yhteentoimivaa (interoperable) ja uudelleenkäytettävää (re-usable).
<b>Force11</b>	Yhteisö tutkijoita, kirjastotyöntekijöitä, arkistotyöntekijöitä, julkaisijoita ja tutkimusrahoittajia.
<b>GenBank</b>	Kattava tietokanta, joka sisältää julkisesti saatavilla olevia nukleotidi sekvenssejä.
<b>Horizon Europe</b>	EU: tutkimuksen ja innovoinnin puiteohjelma vuosille 2021-2027.
<b>Konkordaatti</b>	Virallinen sopimus (Cambridge Dictionary: hakusana Concordat 2021).
<b>MANTRA</b>	Verkkopohjainen aineistonhallinnan koulutuskurssi. Kurssia ylläpitää Edinburghin yliopiston aineistonhallinnan tukitiimi.
<b>Metadata</b>	Tietoa kuvaileva tieto.
<b>OECD</b>	The Organisation for Economic Co-operation and Development.
<b>OKM</b>	Opetus- ja kulttuuriministeriö.
<b>RDI</b>	Research, development and innovation
<b>STEM</b>	Science, technology, engineering ja mathematics.
<b>TKI</b>	Tutkimus-, kehittämis- ja innovaatiotoiminta.

<b>UNIFI ry</b>	Suomen yliopistojen rehtorineuvosto.
<b>VTI</b>	Vastuullinen tutkimus- ja innovaatiotoiminta.
<b>Zenodo</b>	Data-arkisto tutkimusaineistolle.
<b>3AMK</b>	Kolmen ammattikorkeakoulun, Haaga-Helian, Laurea ja Metropolian strateginen liittouma.

# 1 JOHDANTO

Tässä opinnäytetyössä tarkastellaan tutkimuksen puolella entistä merkittävämpänä asiana pidettyä aineistonhallintaa. Tutkimusaineiston hallinta on muuttunut, kun tieteellinen tutkimus on suuntautunut digitaaliseen ympäristöön entistä enemmän. Taustalla vaikuttavat myös eri toimijoiden vaatimukset ja pyrkimykset. Avoimen tieteen pyrkimys mahdollisimman läpinäkyvään tutkimukseen ja tutkimuksellisten aineiston avoimeen saatavuuteen on esillä erilaisissa keskusteluissa. Tässä opinnäytetyössä ajatellaan tuloksia tutkimusaineistosta erillisenä asiana.

Association of College & Research Libraries huomioi vuonna 2014 tieteellisissä kirjastoissa (academic library) tapahtuvia trendejä. Yksi trendeistä liittyi dataan. Julkaisussa korostui avoin data, datasuunnitelman hallinta ja ”big data” -tutkimus. (ACRL 2014.) Vuoden 2020 julkaisussa ”2020 top trends in academic libraries” ACRL Research Planning and Review Committee -jäsenet tarkastelivat trendejä, jotka vaikuttavat tieteellisissä kirjastoissa. Datapalveluiden kohdalla trendit viittasivat etiikkaan ja kypsyyteen (ACRL 2020).

Data voi olla yrityksille nykypäivänä hyvin arvokasta. Kauppaketjut esimerkiksi hyödyntävät asiakasdataa ja tietoa pyritään hyödyntämään mahdollisimman paljon päätöksenteossa. Korkeakoulukirjastot ovat ottaneet osaksi uutta työnkuvaa aineistonhallinnan viitekehykseen kuuluvia tehtäviä tarjotakseen korkeakouluissa tehtävään tutkimukseen tukea tutkijoille.

Toisessa luvussa kerrotaan Laurea ammattikorkeakoulusta, 3AMK-yhteistyöstä sekä TKI-toiminnasta, eli tutkimus-, kehittämis- ja innovaatiotoiminnasta Laureassa. Kolmannessa luvussa pohditaan hieman TKI-toiminnan määritelmää varsinkin ammattikorkeakoulujen osalta. Neljännessä luvussa tarkastellaan aiempaa tutkimusta aineistonhallinnasta niin suomalaisissa, kuin kansainvälisissä korkeakouluissa.

Teoreettisen viitekehyksen ensimmäinen pääluke koostuu tutkimusaineiston ja tutkimusdatan määrittelyn pyrkimyksestä. Tutkimusdatan kategorisointia on tuotu esille. Myös ammattikorkeakoulujen hankkeissa tuotettuja TKI-aineistoja on tuotu esille ensimmäisen osion kolmannessa luvussa. Toisessa pääluvussa tarkastellaan yleisellä

tasolla mitä on aineistonhallinta sekä katsotaan aineistonhallintaa tutkimusaineiston elinkaaren näkökulmasta. Toisen pääluvun kolmannessa alaluvussa tarkastellaan aineistonhallintasuunnitelmaa, joka on keskeinen osa aineistonhallintaa.

Tutkimusongelman kautta on pyritty selvittämään mitä tutkimusaineistoja Laurean hankkeissa syntyy ja miten niitä hallinnoidaan. Yhden tutkimuskysymyksen kautta on myös haluttu selvittää mahdollisia haasteita TKI-toimijoilla mitä tulee aineistonhallintaan. Tutkimusmenetelmänä empiirisen aineiston hankinnassa on käytetty laadullisia teemahaastatteluita, joihin osallistui Laurea ammattikorkeakoulun TKI-toimijoita. Haastattelut toteutettiin Teams-verkkoalustalla 24.11.2021 – 26.1.2022 välisenä aikana. Nauhoitetut haastattelut litteroitiin ja analysoitiin käyttäen sisällönanalyysiä.

Tämän opinnäytetyön toimeksiantaja on Laurea ammattikorkeakoulun tieto- ja julkaisupalvelut. Työn tekijällä on ollut työsuhde Laurea ammattikorkeakoulun organisaatioon vuonna 2020. Opinnäytetyön tekemisen aikana ei ole ollut työsuhdetta. Tekijä on viestinyt Laurean yhteyshenkilöiden kanssa opinnäytetyön etenemisestä sekä empiirisessä osassa toteutuneiden haastateltavien yhteystietojen saannin osalta.



## 2 LAUREA AMMATTIKORKEAKOULU

Laurea ammattikorkeakoulu toimii Uudellamaalla kuudella eri kampuksella. Nämä ovat Lohja, Otaniemi, Leppävaara, Tikkurila, Hyvinkää ja Porvoo, jossa kampuksella toimii myös Haaga-Helia ammattikorkeakoulun yksikkö. Koulutusalat ovat liiketalous, sosiaali- ja terveysala sekä matkailu-, ravitsemis- ja talousala, joissa on tarjolla 18 AMK-koulutusta, joista 6 on englanninkielistä. Laurea ammattikorkeakoulu (aiemmalta nimeltä Vantaan ammattikorkeakoulu ja Espoo-Vantaan ammattikorkeakoulu) aloitti toimintansa vuonna 1991. Korkeakoulussa on noin 7 750 opiskelijaa ja työskentelee 660 henkilökunnan jäsentä. Korkeakoululla on yli 30 000 alumnia. Laurean liikevaihto on 55 miljoonaa euroa. (Laurea organisaationa, [Viitattu 16.10.2021].)

Laurea ammattikorkeakoulun tutkimusrahoitus vuonna 2019 oli 7 221 000 euroa ja tästä ulkopuolinen tutkimusrahoitus samana vuonna oli lähes 4 500 000 euroa. (Ammattikorkeakoulujen tutkimusrahoitus, [Viitattu 19.3.2021]). Tutkimushenkilökunnan tutkimustyövuosia vuonna 2019 oli 54 ja tutkimushenkilökuntaa 165 henkilöä. (Ammattikorkeakoulujen tutkimushenkilökunnan tutkimustyövuodet, [Viitattu 19.4.2021]; Ammattikorkeakoulujen tutkimushenkilökunta, [Viitattu 19.4.2021].)

Laureassa tutkimus, kehitys- ja innovaatiotoiminnan tavoitteena on osaamisen, ratkaisujen ja uuden liiketoiminnan tuottaminen samalla edistäen tulevaisuuden hyvinvointia, turvallisuutta ja kansainvälistä kilpailuetua. Laurean ammattikorkeakoulun toiminta- ja taloussuunnitelman mukaan korkeakouluyksiköt toteuttavat organisaation TKI-toimintaa tutkimusohjelmissa ja tutkimusohjelmien sekä läpileikkaavien teemojen alueilla tutkimus- ja kehitystyö on käytännönläheistä. Ohjelmissa ja teemojen määrittelyssä ja toiminnassa on huomioitu tulevaisuuden tarpeet ja Laurean vahvuudet. (Laurea-ammattikorkeakoulun toiminta- ja taloussuunnitelma 2021 – 2023, 2020, 7– 8.)

Laurealla oli keväällä 2021 noin 80 TKI-hanketta käynnissä, joissa rahoitus oli ulkopuolista. Laurean tutkimus- ja kehittämistoiminnassa oli samaisena keväänä kolme tutkimusohjelmaa: turvallisuuden tutkimusohjelma, kestävän ja monipuolisen sosiaali- ja terveysalan tutkimusohjelma sekä palveluliiketoiminnan ja kiertotalouden tutkimusohjelma. Läpileikkaavat teemat ovat: yrittäjyys, pedagogiikka ja yhteiskehittäminen, digitalisaatio ja

tiedon hallinta yhteiskunnassa sekä eettinen ennakointiosaaminen. (Tutkimus, kehitys ja innovaatiot, [Viitattu 19.10.2021].)

Laurea ammattikorkeakoulu on osa 3AMK-liittoumaa, jossa on mukana Laurean lisäksi Metropolia ja Haaga-Helia. Rauhala (2021, 17) kertoo, että 3AMK-liittoumasopimuksen mukaan 3AMK-liittouman tavoitteena on vahvistaa jäsenammattikorkeakoulujen strategisten päämäärien toteutumista ja tuottaa niille lisäarvoa valituilla strategisilla alueilla. Hän mainitsee, että liittoumassa tavoitellaan systemaattisesti johdetulla kehitys- ja kokeilutoiminnalla enemmän tuloksia ja vaikuttavuutta ja, että liittouman yksi keskeisistä yhteistyöalueista on kansainvälisesti kilpaillun rahoituksen kasvattaminen tutkimus-, kehittämis- ja innovaatiotoiminnassa.

Laurean viimeisimmässä strategiassa yksi kolmesta toimintaa ja työskentelyä ohjaavista arvoista on avoimuus. Strategiassa sanotaan avoimuuden parantavan T&K-toiminnan laatua ja nopeuttavan tiedon jakamista aluekehitysyhteistyössä (Laurea ammattikorkeakoulun strategia 2030, 2019, 10). Laureassa noudatetaan vastuullisen tutkimuksen- ja innovaatiotoiminnan (VTI) periaatteita. VTI:n periaatteet on huomioitu tutkimus-, kehittämis- ja innovaatiotoiminnassa, oppimisessa ja opetuksessa sekä muussa toimintakulttuurissa. Laurea on koordinoanut ammattikorkeakoulujen yhteisessä Opetus- ja kulttuuriministeriön (OKM) rahoittamassa Ammattikorkeakoulujen avoin tki-toiminta, oppimien ja innovaatioekosysteemi -hankkeessa työpakettia 2 (palveluverkosto) ja työpakettia 3 (tki-integroitu avoin oppiminen). (Vastuullinen tiede ja innovaatiopolitiikka, [Viitattu 20.10.2021].)

Vuonna 2018 julkaistussa 3AMKin Viisas avoimuuspolitiikassa kerrotaan kyseisen asiakirjan kuvaavan TKI-toiminnan avoimeen toimintamalliin liittyvät Haaga-Helian, Laurean ja Metropolian strategisen liittouman (3AMK) yhteiset lupaukset ja niitä toteuttavat periaatteet ja linjaukset. Johtava teema on toiminnan yhteiskunnallinen vaikuttavuus, johon pyritään avoimilla toimintamalleilla. (Viisas avoimuus: 3AMK:n yhteiset avoimuuden lupaukset ja periaatteet 2018, 4, 5.) Lähtökohdaksi asianomaisessa politiikassa on asetettu TKI-infrastruktuurien, ympäristön, aineistojen, menetelmien ja tulosten avaaminen aina kun se on mahdollista. Näin Laureassa syntyvät aineistot, menetelmät ja tulokset pystytään uudelleen käyttämään ja hyödyntämään mahdollisimman laajalti. (Vastuullinen tiede ja

innovaatiopolitiikka, [Viitattu 27.10.2021].). Laurean toiminta- ja taloussuunnitelmassa sanotaankin, että:

TKI-toiminnan on vastattava muuttuvan maailman tarpeisiin riittävän nopeasti. TKI-hankkeista saatavien ideoiden ja tulosten tuotteistaminen sekä palveluiden kehittäminen ja niiden kaupallistaminen on yhä tärkeämpää tulevaisuudessa. Tätä tukee TKI-toiminnan avoimuus, mikä vahvistaa yhteiskunnallista vaikuttavuutta. (Laurea-ammattikorkeakoulun toiminta- ja taloussuunnitelma 2021 – 2023, 7, [Viitattu 19.10.2021].)

### 3 TKI-TOIMINNASTA

Tutkimus- ja kehittämistehtävä (TK/TKI-toiminta) lisättiin ammattikorkeakoulujen lakisääteisiin tehtäviin vuonna 2003. Vuonna 2015 ammattikorkeakoululakiin lisättiin innovaatiotoiminta. Samana vuonna vastuu ammattikorkeakoulujen perusrahoituksesta siirtyi valtiolle. (Laakso, Stormi & Mikkonen 2020, 34). Päällysaho ym. (2021, 177) tuovat esiin sen, että Suomessa ammattikorkeakoulut ovat ainoa instituutio, joille on asetettu lakisääteinen tehtävä innovaatiotoiminnassa.

Arene ry:n eli ammattikorkeakoulujen rehtorineuvoston työryhmän (2017, 6) mukaan ammattikorkeakouluissa tehtävä TKI-toiminta tuottaa yhteistyössä esimerkiksi tutkimus- ja koulutusorganisaatioiden sekä työ- ja elinkeinoelämän ja julkishallinnon kanssa uutta tietoa ja osaamista. Asiaan liittyy myös kumppaneiden toiminnan kehittäminen. Heidän mukaansa kumppanit hyödyntävät ja soveltavat tietoa, joka tapahtuu ammattikorkeakoulujen tuella. Arene ry:n työryhmän selvityksessä (2017, 6) mainitaankin, kuinka ammattikorkeakoulut ovat tulleet olennaiseksi osaksi Suomen tutkimus- ja innovaatiojärjestelmää. Ammattikorkeakouluihin kohdistunut uudistus on Laakson ym. (2020, 34) mukaan luonut kannusteita kehittää TKI-toimintaa lisäämällä ulkopuolista hankerahoitusta, kehittämällä kumppanuusmalleja ja laajentamalla TKI-toimintaa verkostojen kanssa.

Tilastokeskuksen määritelmän mukaan (Tutkimus- ja kehittämistoiminta, [Viitattu 24.10.2021].): ”tutkimuksella ja kehittämisellä (T&K) tarkoitetaan yleisesti luovaa ja systemaattista toimintaa tiedon lisäämiseksi ja tiedon käyttämistä uusiin sovelluksiin”. Tavoitteena siinä on jotakin olennaisesti uutta. Tutkimuksen ja kehittämisen tavoitteena on uuden tiedon ja tuloksien tuottaminen. Se, että olemassa olevaa tietoa sovelletaan uusien tuotteiden, prosessien tai menettelytapojen kehittämiseksi ei ole kyseisen määritelmän mukaan T&K-toimintaa. Määritelmän mukaan luovuus, ongelmanasettelu, uudenlaisten käsitteiden ja hypoteesien testaaminen on tunnusomaista T&K-toiminnalle. Kuitenkin tuotteiden, prosessien, käytänteiden tai mallien rutiininomainen uudistaminen tai kehittäminen ei ole määritelmän mukaan T&K-toimintaa. Toiminnalle tunnusomaista on epävarmuus. T&K-toimintaan kuuluu sen suunnitelmallinen suorittaminen ja toteutuksen seuranta sekä tarkoituksen määrittely ja suunniteltujen resurssien kohdennus. Määritelmän mukaan on mahdollista, että T&K-toiminta on yhden henkilön tai ryhmän suorittamaa

tavoitteellista toimintaa, mutta on usein organisoitu projektiksi. (Tutkimus- ja kehittämistoiminta, [Viitattu 24.10.2021].)

Tilastokeskuksen määritelmän mukaan tutkimuksen ja kehittämisen voidaan katsoa kattavan perustutkimuksen, joka ei kuulu ammattikorkeakoulujen tehtäviin, soveltavan tutkimuksen ja kehittämistyön. Perustutkimukselle tunnusomaista on uuden tiedon tavoittelu ilman välitöntä käytännön sovellusta. Määritelmän mukaan perustutkimusta on esimerkiksi:” ominaisuuksien, rakenteiden, syy- ja seuraussuhteiden analyysit, joiden tavoitteena on uusien hypoteesien, teorioiden ja lainalaisuuksien muodostaminen, todentaminen ja selittäminen”. (Tutkimus- ja kehittämistoiminta, [Viitattu 24.10.2021].) Kielitoimiston sanakirja määrittää perustutkimuksen: ”uuden tiedon löytämiseen, kuvaamiseen ja selittämiseen pyrkivä tutkimus (joka ei tähtää käytännöllisiin päämääriin) (Kielitoimiston sanakirja: hakusana perustutkimus, [Viitattu 25.10.2021]). Tilastokeskuksen määritelmän mukaan soveltava tutkimus on: ”tutkimusta, jossa tavoitteena on jokin uuden tiedon löytäminen.”. Määritelmässä tuodaan esiin esimerkkinä sovellusten etsiminen perustutkimuksen tuloksille. Kehittämistyö määritellään Tilastokeskuksen mukaan: ”kehittämistyö, jolla tarkoitetaan tutkimuksen tuloksena ja/tai käytännön kokemuksen kautta saadun tiedon käyttämistä uusien tuotteiden, prosessien tai menetelmien aikaansaamiseen tai olemassa olevien olennaiseen parantamiseen.” (Tutkimus- ja kehittämistoiminta, [Viitattu 24.10.2021].)

Tilastokeskuksen määritelmässä (Innovaatiotoiminta, [Viitattu 26.10.2021]) innovaatiotoiminnasta sanotaan näin: ”innovaatiotoiminta käsittää kaikki yrityksen kehittämis-, rahoitus- tai kaupallistamistoimet, jotka tähtäävät tai johtavat innovaatioihin”. Arene ry:n työryhmän (2017, 37) mukaan innovaatiotoiminta on erinomaisesti sovellettavissa ammattikorkeakoulujen ketterään ja käytännönläheiseen toimintakulttuuriin. He lisäävät vielä, että innovaatiotoiminnassa tunnistetaan tutkimus- ja kehittämistoiminnan tuloksista sellaisia aihioita, joilla on kaupallista potentiaalia. Heidän mukaansa näiden aihoiden kaupallista arvoa arvioidaan ja niitä jalostetaan siten, että ne vastaavat paremmin kysyntää.

Ammattikorkeakoulujen vahvuutena TKI-toiminnassa voidaan nähdä sen käytännönläheisyys. Ammattikorkeakouluissa tehtävässä tutkimuksessa on ominaispiirteitä

soveltavan ja tavoitetutkimuksen alueilta ja perustutkimuksesta. Kehittämistoiminnassa löydetään uusia ratkaisuja ja tietoa viedään käytäntöön. Tätä toteutetaan työ- ja elinkeinoelämän kanssa alueellisella, kansallisella ja kansainvälisellä tasolla. Innovaatiotoiminnan tuotoksena on uusia ja paranneltuja tuotteita, palveluita, prosesseja ja menetelmiä. Innovaatiotoiminnan edistämistä toteutetaan esimerkiksi kaupallistamisen, yrittäjyyden ja avoimen innovaatiotoiminnan avulla. (Arenen ry:n työryhmän 2017, 8–9.)

Ammattikorkeakoulukontekstissa puhutaan avoimesta (avoin TKI-toiminta) tutkimus-, kehittämis- ja innovaatiotoiminnasta. Tällä on pyrkimyksenä edistää yhteistyötä työ- ja elinkeinoelämän kanssa sekä edesauttaa, että tieto ja osaamien liikkuvat ammattikorkeakoulujen innovaatioekosysteemeissä. (Päällysaho & Latvanen 2019b, 25). Avoin TKI-toiminta voidaan nähdä vastineena avoimelle tieteelle ja tutkimukselle (Marjamaa & Latvanen 2017, 42). Rissanen (2017, 6) mainitsee avoimen toimintakulttuurin avaavan tutkimus-, kehittämis- ja innovaatiotoiminnassa käytettyjä menetelmiä, aineistoja, tuloksia ja tuotoksia käytettäväksi kaikille niille, jotka niitä haluavat käyttää. Rissanen kuitenkin huomauttaa, että tämä pitäisi tapahtua hyvän tieteellisen käytännön, tutkimusetiikan sekä juridiset reunaehdot huomioiden. Näin, Rissanen huomioi, on mahdollista edistää TKI-toiminnan laatua, luotettavuutta ja näkyvyyttä ja siten vahvistetaan ammattikorkeakoulujen yhteiskunnallista vaikuttavuutta ja innovaatioiden syntyä.

Avointa TKI-toimintaa voidaan katsoa prosessimallin avulla (Päällysaho & Latvanen 2017). Tämä avoimen TKI-toiminnan malli kattaa TKI-hankkeen elinkaaren eri vaiheet. Se huomioi hankkeen valmistelun, toteutuksen, tulosten hallinnan, tulosten avaamisen ja julkaisemisen aina hanketulosten liiketoiminnalliseen hyödyntämiseen.

## 4 AIEMPI TUTKIMUS

Aineistohallinta on laaja alue ja tässä luvussa ei pyritä kaiken kattavaan kirjallisuuskatsaukseen, koska se vaatisi itsessään oman opinnäytetyön. Tässä on pyritty tuomaan esiin joitain relevantteja ja ajankohtaisia tutkimuksia, niin kotimaisia kuin kansainvälisiäkin.

Tampereen yliopiston kyselytutkimuksessa kartoitettiin tutkijoiden näkemyksiä avoimesta julkaisemisesta ja tutkimusaineistojen avaamisesta, joka on osa aineistohallintaa (Mäki ym. 2017). Kyselytutkimuksella, joka oli osa kansainvälistä Data Literacy and Research Data Management -yhteistyöprojektia, kartoitettiin suomalaisten tutkimusorganisaatioiden jäsenten käsityksiä, kokemuksia ja mielipiteitä liittyen tutkimusaineiston jakamiseen ja pitkäaikaissäilytykseen (Enwald 2018, 31–32). Yhdessä tutkimuksessa tohtoriopiskelijoilta ja tiedekunnan jäseniä pyydettiin itsearvioimaan tai arvioimaan tohtoriopiskelijoiden vallitsevat aineistohallinnan kompetenssit sekä pyydettiin arvioimaan näiden kompetenssien tärkeys (Rantasaari 2021).

Yliopiston pro gradu -tutkielmissa on muun muassa keskitytty avoimeen tieteeseen ja tutkimusaineistoihin (Kautonen 2021), tutkimusaineiston hallintaan digitaalisessa toimintaympäristössä (Satama 2021) sekä esimerkiksi kartoitettu tutkijoiden näkemyksiä ja kokemuksia tutkimusaineistojen aineistohallinnasta (Isola 2018).

Seinäjoen ammattikorkeakoulussa haastateltiin TKI-hankkeita vetäviä projektipäälliköitä (Päällyaho ja Latvanen 2016). Tapaustutkimuksessa selvitettiin projektipäälliköiden näkemyksiä hankkeissa syntyneiden aineistojen ja tulosten hallinnasta, avaamisesta ja jatkohyödyntämismahdollisuuksista. Ammattikorkeakouluille laadittiin marras-joulukuussa 2018 verkkokysely. Se oli suunnattu TKI-toiminnan parissa työskentelevälle henkilöstölle, opettajille ja muille asiantuntijoille. Kyselyllä kartoitettiin avoimen TKI-toiminnan kannalta keskeisiä tekijöitä. Näitä olivat esimerkiksi olemassa olevat käytännöt ja asenteet avointa TKI-toimintaa kohtaan. (Jaalamaa ym. 2020; Talvela ym. 2020.) Haastattelututkimuksella selvitettiin ammattikorkeakouluhenkilöstön tuen ja tiedon tarvetta työssään avoimen TKI-hankkeen suunnittelun ja toteuttamisen eri vaiheissa (Ahonen ym. 2020). Samassa selvityksessä kartoitettiin myös tiedon ja tuen tarvetta avoimiin oppimateriaaleihin sekä

opetukseen liittyen. Ahonen ym. (2020) artikkelissa käsiteltiin TKI-työtä tekevien vastauksia, jotka liittyivät TKI-hankkeen suunnitteluun, toteuttamiseen sekä päättämiseen ja raportointiin. Tutkimukset aineistohallinnasta ovat keskittyneet yliopistoihin tai tutkimuspainotteisiin korkeakouluihin, eivätkä niinkään korkeakouluihin, jotka eivät ole tutkimuspainotteisia. Toisaalta, kuten Laakso ym. (2020, 37) toteavat ammattikorkeakoulujen ja yliopistojen rahoitusosuuksiin liittyen, ammattikorkeakouluja saatetaan mieltää edelleen enemmän kouluttajiksi kuin soveltavan tutkimuksen tekijöiksi.

Stamatoplos, Neville ja Henry (2016) tutkivat datanhallinnan toimintaympäristöä korkeakoulussa, joka on Carnegie Classification viitekehyksen mukaan maisteritason korkeakoulu (Master's-level institution) ja ei ole tutkimuspainotteinen korkeakoulu. Se on jossain määrin verrattavissa Suomessa oleviin ammattikorkeakouluihin. Stamatoplos ym. kertovat, että kyseisessä instituutiossa toimivilla tutkijoilla oli kuitenkin samoja huolia ja haasteita aineistohallintaan liittyen, kuin mitä oli erilaisissa tutkimuksissa tullut esille tutkimuspainotteisissa korkeakouluissa tutkijoiden taholta.

Tenopir ym. (2020) kyselytutkimuksessa tarkasteltiin maailmanlaajuisesti tutkijoiden käytäntöjä ja näkemyksiä aineistohallinnasta, sisältäen aineiston tallennuksen, aineiston jakamisen sekä aineiston käytön ja uudelleenkäytön. Tenopir ym. (2015) tutkimuksessa vertailtiin miten datan jakamisen ja uudelleenkäytön näkemykset vertautuvat aikaisempaan Tenopirin ym. (2011) kyselytutkimukseen.

Kartoittavassa katsauksessa pyrittiin kuvaamaan olemassa olevan tutkimuskirjallisuuden volyyymi, aiheet ja metodologiset luonteet liittyen aineistohallintaan akateemisissa instituutioissa (Perrier ym. 2017). Perrierin, Blondalin ja MacDonaldin (2020) metasynteesissä tarkasteltiin akateemisten tutkijoiden näkemyksiä, näkökulmia ja kokemuksia datan jakamisesta ja tutkimusdatan uudelleenkäytöstä.

Kyselytutkimuksella arvioitiin Puolassa toimivien lääketieteen tutkijoiden tietämyksen tasoa aineistohallintaan liittyen sekä heidän suhtautumistaan aineiston jakamiseen (Milewska ym. 2021). Krahen ym. (2020) tutkimuksessa tutkittiin terveys- ja lääketieteen tutkijoiden vallitsevia aineistohallinnan käytäntöjä australialaisessa akateemisessa instituutiossa. Joo ja Peters (2019) tutkivat yhden tutkimusyliopiston tutkijoiden tarpeita dataan liittyvässä



tuessa ja heidän aineistohallinnan käyttäytymistään. Islannissa University of Icelandissa selvitettiin tietämystä ja näkemyksiä aineistohallinnasta, aineistohallinnan metodien käyttöä sekä havaittua tarvetta tuelle, suhteessa osallistujien tutkimusalueeseen (Palssdottir 2021, 322). Reichman ym. (2021, 1415) tutkimuksessa tehtiin yleiskatsaus aineistohallinnan käytännöistä läpi tiedekuntien ja tieteenalakontekstien.

## 5 TEOREETTINEN VIITEKEHYS

### 5.1 Tutkimusaineistot ja tutkimusdata

Tässä luvussa tarkastellaan tutkimusaineiston ja tutkimusdatan määrittelyä. Tutkimusdataa tarkastellaan tarkemmin ja tuodaan ilmi sen vaihtelevuus ja monipuolisuus. Viimeisessä alaluvussa katsotaan, mitä tiedetään ammattikorkeakoulujen projekteissa syntyvistä TKI-aineistoista.

Tutkimusdata on kompleksinen käsite ja sitä on pyritty tässä luvussa määrittelemään aineistohallintaan liittyvän kirjallisuuden kautta. Tutkimusaineisto taas on yläkäsite tutkimusdatalle, mutta niitä käytetään tarkoittaen samaa asiaa. Ammattikorkeakoulujen tehtävä ei ole perustutkimus, joten osa käsitellystä asiasta tässä luvussa ei välttämättä täysin päde ammattikorkeakouluihin, sillä kirjallisuus aineistohallinnasta keskittyy pääosin yliopistopuolella tehtävän tutkimuksen kontekstiin.

#### 5.1.1 Tutkimusaineiston ja tutkimusdatan määrittelyä

Kuusniemi ja Nykyri (2021, 58) tuovat esiin, että sanoja tutkimusdata (research data) ja tutkimusaineisto (research material) voidaan käyttää ristiin toistensa synonyymeinä. He lisäävät, että ne voidaan määritellä ja kääntää eri tavalla riippuen kontekstista. Artikkelissaan he tarkoittavat tutkimusdatalla sitä tietoaaineistoa, joka on kerätty, havaittu, mitattu tai luotu tutkimustulosten todentamiseksi ja se mikä tekee datasta tutkimusdataa, on konteksti, kuten tutkimustarkoitus. Heidän mukaansa minkä tahansa tietoaaineiston (data) on mahdollista olla tutkimusdataa, jos sitä analysoidaan tutkimustarkoituksessa.

Kuusniemi ja Nykyri (2021, 59) kertovat tutkimusaineiston olevan tutkimusdataa laajempi käsite ja pitävän sisällään tutkimusdatan lisäksi tutkimuksessa käytettävät sekundaarilähteet, joita ovat esimerkiksi käytetty kirjallisuus ja menetelmäohjeet. Tutkimusaineistoa ovat myös fyysiset artefaktit ja näytteet heidän mukaansa. Tutkimustuotosta voi heidän mukaansa pitää tutkimusdataa laajempänä käsitteenä ja

tutkimusdata voidaan käsittää yhtenä tutkimustuotostyyppinä eli tutkimusdata olisi generisessä suhteessa tutkimustuotokseen nähden ja partitiivisessa suhteessa tutkimusaineistoon nähden Kuusniemen ja Nykyrin sanoin.

Empiirisessä tutkimuksessa tutkimusdataa tuotetaan, kerätään ja käytetään (Kuusniemi & Nykyri 2021, 59). Coxin ja Verbaanin (2018, 20) mukaan jotkut tutkijat eivät kerää dataa, kun kyseessä on esimerkiksi puhtaasti teoreettinen tieteenala kuten filosofia. Kuusniemi ja Nykyri (2021, 59) huomauttavat, että teoreettisen tutkimuksen datan tunnistaminen saattaa olla hankalaa. Tiedeyhteisössä datan voidaan ajatella olevan olennainen osa tutkimusprosessia. Hoffman (2017, 42) tuo esiin kuinka tutkijat tarvitsevat informaationpalasia testatakseen malleja, ratkoakseen ongelmia, vastakseen tutkimuskysymyksiin ja evaluoidakseen hypoteeseja. Hoffman viittaa tässä datan tarjoavan todisteet, jolla tarkastella väitteiden ja perustelujen pätevyyttä.

Sewell (2020, 17-18) tuo esiin, että data tai suomeksi myös tietoaaineisto on mitä tahansa kerättyä informaatiota, olipa se sitten olemukseltaan numeerista (kvantitatiivista dataa) tai informaatiota, jota ei voi mitata numeroilla (kvalitatiivinen data). Hän mainitsee tutkimusdatan olevan yksinkertaisesti informaatiota, jota käytetään ja tuotetaan tutkimusprosessin aikana. Briney (2015, 4, 6) mainitsee tutkimusdatan määrittelemisen olevan haasteellista, koska se on olemukseltaan heterogeenista. Hän tuo esiin tutkimusalojen kirjavuuden ja kuinka jokin tietty tutkimusalan alue voi käyttää hyvin isoa valikoimaa aineistoa. Briney määrittelee tutkimusdatan yksinkertaisesti seuraavasti: ” kaikkea sitä mihin tehdään analyysia.” Myös Cox, Pinfield ja Smith (2016, 3) mainitsevat tutkimusdatan vaihtelevuuden ja volyymin olevan yksi aineistonhallinnan haasteista sekä datan luonti ja käyttö voi olla hyvin erilaista eri tutkimusaloilla. Borgmanin (2015, 79) mukaan datan moninaisuus ei liity vain tutkimuksen lähestymistapojen laajuuteen ympäri tutkimusalueita, mutta myös niihin moniin tapoihin, joilla yksittäiset havainnot, objektit, todisteet, tekstit, näytteet, henkilöt, eläimet ja muut entiteetit voidaan esittää datana.

Tiedonjulkistamisen neuvottelukunnan ja Tieteellisten seurain valtuuskunnan julkaisussa (2021, 9) huomioidaan, että tutkimusaineistojen ja -menetelmien aihepiirin sanasto voi olla erilaista eri tieteenaloilla. Julkaisussa mainitaan, että lainsäädännössä termit ja niiden käännökset vaihtelevat paljon. Julkaisussa lisätään vielä, että termejä, joita käytetään, on

määritelty eri yhteyksissä toisistaan poikkeavalla tavalla ja kun tieteenalat ovat erilaisia, on yhtenäisen sanaston luominen haastavaa. Tiedonjulkistamisen neuvottelukunnan ja Tieteellisten seurain valtuuskunnan Osalinjauksessa 1: tutkimusdatan avoin saatavuus tutkimusaineisto ja tutkimusdata on määritelty seuraavanlaisesti:

Tutkimusaineisto (research material) on tutkijan tai tutkimusryhmän tutkimusprosessin aikana käyttämä resurssi, eli digitaalisessa, analogisessa tai fyysisessä muodossa olevaa tieteellisen ja taiteellisen tutkimuksen perusaineistoa. Tutkimusaineisto on laajempi käsite kuin tutkimusdata, kattaen esimerkiksi lähdekirjallisuuden (esimerkiksi asiakirja-aineisto) ja näytteet (esimerkiksi verinäytteet, sammalet).

Tutkimusdata (research data) on tutkimusaineistoa, joka on kerätty, havaittu, mitattu tai luotu hypoteesien vahvistamiseksi ja tutkimustulosten todentamiseksi. Tutkimusdata voi yleensä olla digitaalisessa muodossa, mutta se voi esiintyä myös analogisessa tai fyysisessä muodossa (esimerkiksi laboratoriapäiväkirjat).

Tieteen termipankki (2018) määrittelee tutkimusaineiston seuraavasti:

Tutkijan tuottama tai tutkimusprosessin aikana käyttämä resurssi eli digitaalisessa, analogisessa tai fyysisessä muodossa oleva tieteellisen tutkimuksen perusaineisto tai perusaineistosta jalostettu aineisto, johon tutkimuksen löydökset ja tulokset perustuvat.

Tutkimusaineistot voivat siis olla aineellisia tai aineettomia. Tutkimukseen kelpaavaksi, aineistossa täytyy olla tietoja ainakin sen alkuperästä. Datan liitteeksi tarvitaan kuvailevaa ja teknistä tietoa, jotta tiedetään, mitä informaatiota se sisältää. Tutkimusdataan liittyy tiedot siitä, miten sitä on rakenteistettu ja koodattu sekä millä tavoin se on syntynyt ja miten sitä on käsitelty. Näitä tietoja voidaan tallentaa metatietoihin, koodikirjoihin tai johonkin muuhun dokumentaatioon. Data muodostaa tämän kokonaisuuden kanssa tutkimusaineiston. (Tieteen termipankki: hakusana Tutkimusaineisto 2018.)

OECD (The Organisation for Economic Co-operation and Development) on määritellyt tutkimusdatan, jossa kontekstina on lähinnä digitaalisen datan jakaminen seuraavasti:

Tutkimusdata on määritetty faktuaalisiksi todisteiksi (numeeriset tulokset, tekstuaaliset todisteet, kuvat ja ääni), joita käytetään primaarisina lähteinä tieteellisessä tutkimuksessa ja jotka ovat yleisesti hyväksytyjä tieteellisessä

yhteisössä tarpeelliseksi katsottuna vahvistamaan tutkimuksen tuloksia. Tutkimuksen datasetti muodostaa systemaattisen osittaisen kuvauksen kohteesta, jota tutkitaan.

Tämä termi ei kata seuraavia: laboratoriopäiväkirjat, alustavat analyysit ja tieteellisten tutkielmien luonnokset, suunnitelmat tutkimuksesta tulevaisuudessa, vertaisarvioinnit, tai henkilökohtainen kommunikaatio kollegan kanssa tai fyysiset objektit (esim. laboratorionäytteet, bakteerikannat ja koe-eläimet kuten hiiret). Pääsyä kaikkiin näihin tutkimuksen tuotoksiin ja tuloksiin on hallittu eri näkökohdilla, kuin mitä on tässä käsitelty. (OECD 2007, 13–14)<sup>1</sup>

Tutkimusdatalle voi löytyä erilaisia pyrkimyksiä määritellä se, niin kuin yllä on tuotu esiin. Cox & Verbaan kertovat (2018, 24), että erilaisissa menettelytapadokumenteissa saatetaan määritellä tutkimusdata melko muodollisesti ja se saattaa olla hyödyllistä, mutta jotkut määritelmät toimivat paremmin tiettyihin tieteenaloihin tai monitieteisiin tutkimusalueisiin. He mainitsevat, että jotkut määritelmät soveltuvat paremmin esimerkiksi luonnontieteen aloihin ja joku määritelmä voi määritellä tutkimusdatan eräänlaisen listauksen kautta, kuin avata dataa konseptuaalisesti. Asiaa voi lähestyä myös yksinkertaisesti kuten Briney (2015, 6) tai niin kuin OECD:n (2007) määritelmässä, myös ottamalla huomioon sen mitä määritelmä ei sisällä. Esimerkiksi Isossa-Britanniassa yhteistyössä eri sidosryhmän jäsenten kanssa on kehitetty konkordaatti Concordat Open Research Data (2016). Konkordaatti tarkoittaa virallista sopimusta (Cambridge Dictionary: hakusana Concordat 2021). Asianomaisessa dokumentissa huomioidaan, että tutkimusdatan tarkoitus on antaa vastauksia tutkimuskysymyksiin ja siinä tuodaan ilmi datan olomuoto, esimerkiksi onko se raakadataa, alkuperäisdataa vai johdannaisdataa.

Sewellin (2020, 18–19) mukaan nk. STEM-aloilla (englanniksi science, technology, engineering ja mathematics) saatetaan kutsua jotain tuotosta dataksi ja usein dataa ajatellaan faktoiksi ja luvuiksi. Sewell korostaa, että ihmistieteissä, taidealoilla ja yhteiskuntatieteissä, informaatiota, jota työstetään ei aina mielletä dataksi. Tämä saattaa Sewellin mukaan olla syynä siihen, että datanhallintaan kannustamiseen ei vastata suotuisasti, mutta informaation hallinnan oppimiseen saattaa olla kiinnostusta. Sewell lisää,

---

<sup>1</sup> Tekijän oma käännös

että kun antaa tukea aineistohallinnassa, saattaa olla hyvä ymmärtää kohdeyleisöä ja huomioida, mikä terminologia parhaiten vetoaa yleisöön.

Informaatiotutkimuksessa käytetty tiedon arvoketju auttaa havainnollistamaan datan muuttumista tiedoksi ja vielä sen jalostumista viisaudeksi. Arvoketju alkaa datasta ja jatkuu informaation, tiedon ja tietämyksen kautta viisauteen. Kun data saa jonkin mielekkään järjestyksen muuttuu se informaatioksi. Informaatio tässä arvoketjussa voi olla joko totta tai epätotta. Informaatio voi olla kirjallista, suullisesti välitettyä ja osa siitä voi olla informaatiota, joka liittyy asioihin ja esineisiin, jota on mahdollista havaita eri aistien avulla. Tiedoksi informaatio muodostuu tulkinnan avulla. (Haasio, Harviainen & Savolainen 2019, 20–21)

### 5.1.2 Tutkimusdatan kategorisointi

Briney (2015, 6) mainitsee, että tietoaineistoa saatetaan määritellä tyypeittäin. Tutkimusdata yhteiskuntatieteissä (social sciences) lasketaan hänen mukaansa yleensä kahteen eri kategoriaan, jotka ovat kvantitatiivinen tai kvalitatiivinen. Du ja Khan (2020, 3) määrittävät kvalitatiivisen datan kokoelmaksi informaatiota, jolla aiotaan kuvailla jotain aihetta pikemminkin kuin mitata sitä. Esimerkkinä he antavat mielipiteet, käsitykset ja näkemykset sekä metodeiksi datan keräämiseen he mainitsevat havainnoinnin, fokusryhmät, haastattelut ja arkistomateriaalin. Kvantitatiivinen data on Du'n ja Khanin mukaan numeraalista ja sitä voidaan käyttää laskennassa ja tilastollisissa testeissä. Kvantitatiivisen datan keräämisen menetelmiksi he listaavat kokeet, kyselyt sekä havainnoinnin ja haastattelut.

Primääridata on Krierin ja Strasserin (2014, 2) mukaan alkuperäistä dataa, jota tutkija kerää tietyssä projektissa ja se saa alkunsa tietyistä kokeista tai havainnosta. He mainitsevat myös, että tutkija saattaa usein käyttää sekundääridataa, jonka joku muu on luonut. Krier ja Strasser antavat sekundääridatasta esimerkiksi väestönlaskentadatan, jota jokin kansallinen organisaatio on voinut kerätä. Essen ym. (2018, 134) huomauttavat, että eri akateemisissa tieteenaloissa on eroja siinä, miten nämä datatyypit ymmärretään. Heidän

mukaansa, kun primääridataa muutetaan, siitä tulee sekundääridataa. He lisäävät, että muutettua primääridataa saatetaan ajatella raakadataksi, jossakin tutkimuksen vaiheessa.

Parland-von Essen ym. (2018, 131–133) tuovat esille datan kategorisoinnin teknisen näkökulman, datan kategorisoinnin modaliteetin tai formaatin mukaan. Heidän mukaansa tutkimusdata voi tulla ihmisen käytettäväksi eri modaliteeteissa korvien tai silmien kautta ja erilaisilla luonnollisilla kielillä tai koodilla, jota voidaan tulkita aisteilla. Modaliteetti ja tekninen formaatti ei ole kuitenkaan heidän mielestään riittävä tapa ryhmittää resursseja kategorioihin, vaikka olisikin relevantti tapa. He lisäävät vielä, että tutkimuksessa pitää erottaa kasvavat tietoaaineistot (englanniksi datasets) joita luodaan ja kerätään systemaattisella tavalla sekä muuttumaton ns. jäädytetty datan julkaiseminen, joka toimii vahvistuksena tietylle tutkimuksen lopputulokselle. Cox ja Verbaan (2018, 21) mainitsevat, että datan rajaaminen formaatin mukaan voisi olla hyödyllistä, mutta se rajoittaisi, jos haluttaisiin esimerkiksi tietää mitä itse dokumentit tai laskentataulukot sisältävät.

Briney (2015, 6) esittää, että tieteellinen tutkimusdata on mahdollista kategorisoida neljään eri ryhmään: Havaintoperäinen (observational data), kokeellinen (experimental data), mallinnus (simulation data) ja koostettu (compiled data). Tässä näkökulmana on datan kategorisointi laadun tai alkuperän mukaan. Myös Parland-von Essen ym. (2018, 133) tuovat esiin samat neljä kategorioita datan kategorisoinnille, kun lähtökohtana on, miten data on luotu tai kerätty tutkimuksen kontekstissa. Parland-von Essen ym. tuovat esiin MANTRan (2017), joka on verkkopohjainen aineistohallinnan koulutuskurssi, esittämän viitteellisen (reference) tai kanonisen (canonical) datan. Kappaleen alapuolella on tarkemmin määritelty näitä kategorioita.

Taulukko 1. Datan kategorisointi alkuperän tai laadun mukaan (Parland-von Essen ym. 2018, 133–134).

Havaintoperäinen (observational data)	Kokeellinen (experimental data)	Mallinnus (simulation data)	Koostettu (compiled data)	Kanoninen tai viitteellinen data
Yleensä kerätään reaaliaikaisesti. Ei voida toisintaa tai saada kiinni uudelleen. Esimerkkejä ovat anturin lukemat, kyselyn tulokset, telemetria ja ihmisen havainnot.	Tyypillisesti kerätään valvotussa ympäristössä laboratorioissa. Esimerkkejä ovat geenisekvenssit, mikroskopia ja spektroskopia.	Sisältää dataa, joka on luotu jäljittelemällä tosielämän toimintoja tai systeemejä käyttämällä matemaattisia malleja. Tämänlaista dataa luodaan yleensä ilmasto- ja talousmalleihin, kemiallisiin reaktioihin tai seismiseen aktiiviteettiin. Se on usein määrällisesti runsaasta ja suhteellisen helppoa toisintaa, jos prosessi on hyvin dokumentoitu	Käsittää valmiiden tietopisteiden käyttöä (data point) eri tietolähteistä luomalla dataa transformaation kautta. Usein toisinnettavissa olevaa, mutta saattaa olla hyvin kallista ja aikavievää. Esimerkkejä ovat 3D -mallit, koostetut tietokannat tai informaation saaminen tekstinlouhinnasta.	On tavallisesti vertaisarvioitua ja usein julkaistua tai/ja kuratoitua kokoelmaa tietoaaineistoja. Esimerkkejä ovat geenisekvenssi tietopankit, kemialliset rakenteet tai väestönlaskennallinen data

Yksi näkökulma kategorisoinnissa on datan luontaiset ominaisuudet (inherent quality), kuten onko data sensitiivistä ja sen takia ei avoimesti saatavilla. Silloin kysymyksiä arvioidaan enemmän eettisin tai lakisääteisillä perusteluilla. Konteksti on sidottuna kulttuurisiin ja yhteiskunnallisiin arvoihin ja argumentteihin. Tästä seuraa tietty tapa kategorisoida dataa ja sen saatavuutta ja se vaikuttaa aineistonhallintaan. (Parland-von Essen ym. 2018, 136.)



Parland-von Essen ym. mainitsevat datan voivan olla täysin avointa tai sen käyttöä voidaan rajoittaa lakisääteisen, sopimusvelvoitteisen tai eettisen syyn vuoksi.

Suomen yliopistojen rehtorineuvosto UNIFI ry:n (2018) julkaisussa on huomioitu, että tutkimusdata ei muodosta yhtä kokonaisuutta. Julkaisussa esitetään kolme tutkimusdatan tyyppiä, jotka ovat kumuloituva data, pysyvä data ja julkinen data. Kumuloituva data, joka syntyy tutkimuksen käynnissä ollessa, ja sen hallinnointia olisi mielekkäintä hallinnoida paikallisesti. Pysyvä data liittyy johonkin tiettyyn tutkimustulokseen ja julkaisun mukaan on tärkeää säilyttää tulosten läpinäkyvyyden ja toistettavuuden toteutumiseksi. Pysyvän datan säilytys on luonteeltaan lähellä arkistointia ja soveltuu kansallisen tason koordinointiin. Julkinen data voi olla joko tutkimuksen raaka-ainetta tai sitten se on yksi sen tuotteista pysyvän datan ohessa. Julkista dataa voi tulla myös tiedeyhteisön ulkopuolelta ja silloin koordinaatiotarpeet mainitaan julkaisussa kohdistuvan enemmän esimerkiksi eettiseen hallintaan. (Suomen yliopistojen rehtorineuvosto UNIFI ry 2018).

Tutkimusdata voi olla myös ei-digitaalista. Tästä Du ja Khan (2020, 3) antavat esimerkkeinä tutkijoiden käsin kirjoitetut päiväkirjat, laboratoriopäiväkirjat ja erilaiset aineistot, joita ei säilytetä tietokonejärjestelmässä. Heidän mukaansa tutkimusprojektiin sisältyy todennäköisesti ei-digitaalista sekä digitaalista dataa tai dataa, joka digitoidaan myöhemmin. Käytännöllisesti katsoen he mainitsevat, että analoginen data on luonnostaan jatkuvaa ja identifioi jokaisen nyanssin asiasta, jota mitataan tai tallennetaan, kun taas digitaalinen data on taipuvainen näytteenottoon (sampling) koodatakseen asiaa, jota mitataan. Du ja Khan luonnehtivat ei digitaalista dataa sellaiseksi, joka on olemassa siitä huolimatta, että mitataanko sitä vai ei.

Force11 on tuonut esille FAIR Data -periaatteet (Force11, [Viitattu 7.12.2021]). FAIR periaatteiden on Wilkinson ym. (2016) mukaan tarkoituksena toimia suosituksena parantamaan tietoaineistojen uudelleenkäyttöä. Periaatteissa korostuu koneiden kyky automaattisesti löytää ja käyttää dataa sekä tukea yksilöiden datan uudelleenkäyttöä (Wilkinson 2016). FAIR tulee englannin kielen sanoista ja tarkoittaa löydettävää (findable), saavutettavaa (accessible), yhteentoimivaa (interoperable) ja uudelleenkäytettävää (re-usable) (fairdata, [Viitattu 1.12.2021]). EU-neuvosto on linjannut FAIR-periaatteiden käytöstä (Council of the European Union 2016).

### 5.1.3 TKI-aineistot

Kun puhutaan tutkimusaineistoista ammattikorkeakoulujen osalta, voidaan sillä tarkoittoa TKI-toiminnassa syntyviä ja prosessoituja aineistoja (Lehto & Suikkanen, 2021.) TKI-hankkeissa Päällysahon ja Latvasen (2019b) mukaan yleisimpiä aineistoja ovat kysely- ja haastatteluaineistot ja ne ovat yleensä digitaalisessa muodossa. Heidän mukaansa hankkeissa syntyy myös mittaus- ja mallinnusaineistoja sekä kuva- ja äänitallenteita. He lisäävät vielä, että työpajoissa syntyy myös aineistoja ja näitä aineistoja kerätään post-it-lappujen tai fläppipapereiden avulla, mutta asianomaisten aineistojen säilytyksessä ja tallentamisessa ei välttämättä ole järjestelmällisyyttä tai niitä ei tallenneta digitaaliseen muotoon.

Hämeen ammattikorkeakoulun aineistonhallinnan verkkosivujen ([Viitattu 15.11.2021]) mukaan TKI-aineistoja voivat olla myös laboratoriopäiväkirjat, lähdekoodit, ohjelmistot tai kenttätöihin perustuvat havainnot. TKI-hankkeissa saattaa myös esiintyä aineettomia tuloksia, joita voivat olla oivallukset, kokemukset ja uusi osaaminen (Päällysaho & Latvanen 2016). Lehto ja Suikkanen (2021) mainitsevat, että ammattikorkeakouluissa tyypillisesti tuotetaan small dataa, joka on big datan vastakohta. Miglanin (2016) määritelmässä small dataa kuvaillaan pieniksi dataseteiksi, jotka voivat vaikuttaa päätöksiin nykyhetkessä sekä sen muodostuvan selvistä ja yksityiskohtaisista datasettien ominaisuuksista, joilla voidaan analysoida vallitsevaa tilannetta. Lehdon ja Suikkasen (2021) mukaan kun katsotaan ”TKI-aineistoja” tukipalveluiden näkökulmasta, ne olisi hyvä nähdä laajana terminä, tällöin sisältäen myös palvelutoiminnan ja koulutuksen pää- tai sivutuotteena syntyvän datan.

Päällysahon ja Latvasen (2019b) mukaan yritysysteistyöhankkeissa tyypillisimpiä aineistoja ovat myös kysely- ja haastatteluaineistot. Näissä saatetaan usein sivuta heidän mukaansa yritykselle strategisesti tärkeitä aiheita tai ne voivat liittyä uusien tuotteiden tai palvelujen kehitystyöhön. Esiin nousee luottamuksellisuus ja aineistojen suojaaminen, vaikka yritysälhtöistä aineistoa olisi kerätty ilman tunnistetietoja. Yritysaineistoissa henkilötietojen suojaaminen pitää huomioida sekä tiedot, jotka liikesalaisuuden näkökulmasta vaativat

suojaamista. Kuvailevien tietojen tallentamisessa saattaa myös esiintyä tietosuojaan liittyviä ongelmia. Yrityksiltä kerätään myös palautteita. (Päällysaho & Latvanen 2019b.)

Laakso ym. (2020, 34) tuovat esiin avoimen tieteen ja tutkimuksen kehityksen näkymisen OKM:n perusrahoituksen uudistamisessa. Uudistetussa ammattikorkeakoulujen rahoitusmallissa vuodesta 2021 alkaen avoimesti saatavilla olevia julkaisuja painotetaan 1,2-kertaiseksi verrattuna sellaisiin julkaisuihin, jotka eivät ole avoimia (Visio 2030 työryhmien raportit, [Viitattu. 21.3.2021]). Laakso ym. (2020, 35) esittävät, että tulevaisuudessa olisi odotettavissa rahoitusmallin uudistuksia, jotka liittyvät ammattikorkeakoulujen avoimiin tutkimusaineistoihin.

## **5.2 Aineistonhallinta ja tutkimusaineiston elinkaari**

Aineistonhallinnan käsikirjan (2021) mukaan aineistonhallinnalla tarkoitetaan sitä, että tutkimusaineistot (data) ja niihin liittyvä kuvaileva tieto (metatieto) on luotu, tallennettu ja järjestetty siten, että aineiston säilyy käyttökuntoisena ja luotettavana ja että tietosuoja ja tietoturva on varmistettu aineiston koko elinkaaren ajan. Jatkokäytöllä käsikirjassa tarkoitetaan toisten kokoaman, johonkin muuhun tarkoitukseen kerätyn aineiston hyödyntämistä. Silloin käytössä on usein erilaisia menetelmiä, kuin ne mitä alkuperäiset tutkijat ovat käyttäneet. (Aineistonhallinnan käsikirja 2021.)

### **5.2.1 Miksi aineistonhallintaa?**

Aineistonhallinnan ajatellaan olevan osa hyvää tieteellistä käytäntöä ja siinä otetaan huomioon koko aineiston elinkaari. Aineistonhallinta tukee tutkimuksen läpinäkyvyyttä, luotettavuutta, toistettavuutta ja todennettavuutta. Tämä liittyy ajatukseen siitä, että voidaan osoittaa mihin aineistoon tutkimuksen tulokset perustuvat (fairdata.fi, [Viitattu 6.10.2021]). Wilkinson ym. (2016) huomauttavat, että hyvä aineistojen hallinta ei ole itsessään päämäärä, vaan se on keskeinen johdin tietämyksen luonnissa ja innovoinnissa sekä

jälkikäteisen datan ja tietämyksen integroinnissa ja uudelleenkäytössä yhteisön osalta, datan julkaisuprosessin jälkeen.

Hey ja Trefethen (2020, 441) kertovat, että Jim Gray esitti 1990-luvulla, että seuraava Big Dataan liittyvä haaste tietokantateknologialle tulisi tieteen puolelta enemmänkin kuin kaupankäynnistä. He huomauttavat, kuinka Jim Gray oli käsittänyt ne teknologiset haasteet, joita erittäin suurten tieteellisten datasettien hallinta ja analyysi aiheuttaisi tieteilijöille. Gray kertoo puheessaan vuonna 2007, että kokeellinen, teoreettinen ja laskennallinen tiede ovat vaikutuksen alaisia datan tulvasta ja neljännestä ”dataintensiivisestä” tieteen paradigmasta, joka oli tuloillaan (Hey, Tansley & Tolle 2009, 33).

Brineyn (2015, 2) mukaan aineistonhallinta oli melko uusi termi vuonna 2015 tieteellisessä tutkimuksessa. Hän mainitsee sen ilmaantuneen 2000-luvun puolivälissä, kun alkoi ilmentyä rahoittajien vaatimuksia aineistonhallintaan ja datan jakamiseen. Resnik ym. (2019, 139) mukaan datan jakamisen voidaan katsoa olevan oleellista tieteen kehityksen kannalta, koska se edistää yhteistyötä, läpinäkyvyyttä, toistettavuutta, kritiikkiä ja uudelleenanalysointia. Datan jakamista voidaan katsoa myös siten, että se säästää aikaa ja resursseja, kun mahdollista kallista datan uusimista ei tarvitse tehdä. Rayn (2014, 2) mukaan se, että löytyy kiinnostusta datan jakamiseen, on vasta ensimmäinen askel sen tekemiseen onnistuneesti. Hän huomauttaa liittyen digitaalisen datan säilytykseen ja hallintaan, että jotta sen voi tallentaa tehokkaasti, että se on löydettävää toisiokäyttäjille ja että sitä voi käyttää luottaen sen oikeellisuuteen ja integriteettiin, edeltää datan jakamista tutkimusryhmien kesken ja pitkäaikaissäilyttämistä uudelleenkäyttöä varten.

Yhtenä esimerkkinä datan jakamisesta voidaan esittää ihmisen koko perimää kartoittava The Human Genome -projektin kautta. Vuonna 1990 alkanut projekti pyrki kansainvälisellä ponnistuksella tunnistamaan ja kartoittamaan yli 20 000 ihmisen perimän geenin sekvenssi ja määrittämään kemiallisten emäsparien sekvenssin, josta muodostuu DNA. Vuonna 2003 päätöksen saanut projekti sai aikaiseksi GenBank -tietopankin, joka kerää DNA-sekvenssejä eri paikoista ympäri maailmaa. Data on julkisesti saatavilla ja sitä louhitaan tutkimustarkoitukseen eri aloilla. (Ray 2014, 2) Yhdysvalloissa yhdessä tutkimuksessa arvioitiin, että 13 miljardin Yhdysvaltain dollariin panostus valtiolta Human Genome-projektiin ja sen seuraajiin on tuottanut taloudellista hyötyä noin biljoona dollaria (RDA

Europe 2014, 5). Tämä on yksi tapa esittää julkisen rahoituksen panostusta, tulosten tasolla. Rayn (2014, 1–2) mukaan GenBank ja muut erittäin näkyvät tutkimusprojektit ovat saaneet huomiota kansallisilta valtionhallinnoilta ja kansainvälisiltä organisaatioilta datan jakamisen mahdollisuuksiin. Hänen mukaansa myös kansainvälinen yhteistyö, jonka avulla ratkaistaan globaalit suuret haasteet, kuten tautien ehkäisy ja hoito sekä ilmastonmuutos, on saanut huomiota yhtä lailla. Kun tutkimusdataa voidaan käyttää uusissa tutkimuksissa, voi sen kautta innovaatioiden synty nopeutua, kun yhteistyö on helpompaa (Borgerud & Borglund 2020).

Aineistonhallinnan relevanssia voidaan tarkastella erilaisista näkökulmista, kuten ”datan tulvasta” (data deluge), joka tulee uusista tieteen tyypeistä tai luottamuskriisistä tutkimuksen rehellisyydestä joillakin aloilla. Yleinen avoimen datan liike on luonut yritystä hallita dataa paremmin ja jakaa sitä enemmän. Monet tutkimuksen rahoittajat ympäri maailman edellyttävät parempaa suunnitelmallisuutta datan hallinnassa ja jakamisessa. (Cox & Verbaan 2018, 5) Esimerkiksi Suomen Akatemian rahoituksella tuotetut tutkimusaineistot ovat oletusarvoisesti avoimesti saatavilla, kuitenkin vaihdellen kaikille täysin avoimesta salassa pidettävään (aineistonhallintasuunnitelma, [Viitattu 23.10.2021]).

Myös artikkeleiden yhteydessä kasvavissa määrin täytyy tutkijoiden julkaista dataa, johon heidän tuloksensa perustuvat (Cox & Verbaan 2018, 5). Tästä esimerkkinä Nature Portfolio (Springer Nature 2022). Yksi tekijä taustalla on ajatus siitä, että julkisen rahoituksen avulla syntyvä tutkimusdata tulisi olla avoimesti saatavilla (Whyte & Tedds 2011) mahdollisimman vähin rajoituksin ja jos eettiset ja lainsäädännölliset velvoitteet sallivat, niin toisiokäyttö tulisi olla mahdollista, jotta dataa ei tarvitsisi luoda uudelleen (Bloemers & Montesanti 2020, 172).

Tieteellisen Nature-lehden artikkelissa vuodelta 2016 kerrotaan tutkimuksen toistettavuudesta (reproducibility). Nature oli laatinut kyselyn, jossa vastaajia oli 1576 tutkijaa. (Baker 2016, 452.) Baker (2016, 452) kertoo, että 70% tutkijoista oli yrittänyt ja epäonnistunut toistamaan toisen tutkijan kokeita ja yli puolet olivat epäonnistuneet toistamaan omia kokeitaan. Hän tuo esiin sen, että vaikka 52% vastaajista oli sitä mieltä, että on olemassa merkittävä ”kriisi” toistettavuudessa, niin alle 31% ajatteli, että epäonnistuminen julkaistujen tulosten toistamisessa tarkoittaa sitä, että tulokset ovat todennäköisesti väärin. Baker vielä lisää, että suurin osa silti luotti julkaistuun kirjallisuuteen.

Aineistohallintapolitiikat ja datan jakamisen politiikat ilmaantuivat vastauksena toistettavuuskriisiin eri tieteenaloilla (Briney 2015, 3). Rahoittajat, ammattikunnat ja tieteelliset aikakausjulkaisut ovat kehittäneet tai hioneet toimintaperiaatteita, joilla edistetään tutkimuksen läpinäkyvyyttä (Resnik ym. 2019, 141).

### **5.2.2 Tutkimusaineiston elinkaari**

Aineistohallinnan käsikirjan (2021) mukaan tutkimusaineiston elinkaari on tavallisesti pidempi kuin sen tutkimushankkeen elinkaari, joka tuotti aineiston. Tutkimusaineiston elinkaarimallit voivat auttaa strukturoimaan tehtävät ja toiminnot, jotka liittyvät aineistojen hallintaan projektin tai organisaation sisällä. (Pouchard, 2015, 180). Pouchardin mukaan ne voivat olla keino viestittää aineistohallinnan erilaisia tehtäviä tavoitellulle yleisölle, kuten tutkijoille, kirjastonhoitajille sekä projektipäälliköille, jotka vastaavat aineiston järjestämisestä projektissa tai laboratoriossa.

Aineiston elinkaaria on laaja valikoima erilaisia, kun jotkut ovat keskittyneet organisaatioon, osa yksittäisiin tutkijoihin ja osa yleisesti ottaen laajempaan yhteisöön (Pouchard 2015, 180). Kowalczykin (2017, 333) mukaan kirjallisuudessa monet aineiston elinkaaret, jotka kuvaavat tutkimusaineiston syntyä, voivat olla geneerisiä, sillä tavalla, että ne liittyvät johonkin tietyn alan piiriin tai ne voivat olla toisaalta kovin tarkoin määriteltyjä, koska ne liittyvät johonkin yhteen tiettyyn laboratorioon tai projektiin.



Kuvio 1. Tutkimusaineiston elinkaari. (Yhteiskuntatieteellinen tietoarkisto (2021). CC-BY.

Kuviossa 1 on esitetty malli tutkimusaineiston elinkaaresta. Siinä on huomioitu se, että vaikka hanke olisi loppunut, aineistoa on mahdollista hyödyntää jatkossa. Kun hanke on päättynyt, kerätty aineisto voidaan luovuttaa arkistoitavaksi ja jatkokäyttöön. Tutkimusaineiston käsittelyä sekä dokumentaatiota tapahtuu tutkimusprojektin aikana mutta myös arkistointivaiheessa. Tutkimussuunnitelmaa laadittaessa kannattaa tutkimusaineiston elinkaaren eri vaiheet suunnitella jo silloin, koska tällöin saatetaan välttyä päällekkäiseltä työltä tai informaation katoamiselta. (Aineistonhallinnan käsikirja 2021)

Aineistonhallinnan käsikirjassa (2021) on mainittu, että tutkimusaineistoa kuvaileva metadata olisi helpointa ja kustannustehokkainta tallentaa vaiheittain aineiston elinkaaren aikana. Siinä mainitaan myös, kuinka pyrkimys puutteellisen metadatan täydentämiseen ja korjaamiseen vuosia aineistonkeruun tapahtumisen jälkeen ei välttämättä ole mahdollista ja saattaisi maksaa monikymmenkertaisesti verrattuna siihen, mitä metadatan tuottaminen aineiston elinkaaren alkuvaiheissa maksaisi. Hyvällä dokumentoinnilla tuotettu metadata mahdollistaa datan tulkinnan ja jatkokäytön (Kuusniemi & Nykyri 2021, 61).

### 5.2.3 Aineistohallintasuunnitelma

Aineistohallintasuunnitelma on dokumentti, jota käytetään kuvaamaan tutkimusaineiston hallinta projektin aikana sekä tutkimusprojektin elinkaaren jälkeen. Suunnitelmassa kuvataan millä tavoin tutkimusaineisto kerätään, miten sitä käytetään ja säilytetään tutkimusprojektin aikana sekä tuodaan esiin, miten aineiston jatkokäyttö mahdollistetaan tutkimusprojektin päätyttyä. (Aineistohallinnan käsikirja 2021). Michener (2015) mainitsee aineistohallintasuunnitelman tyypillisesti kattavan kokonaan tai osittain datan elinkaaren. Aineistohallintasuunnitelmaa voisi ajatella dynaamisena viitekehyksenä, jota ylläpidetään ja muunnellaan tutkimuksen edetessä (European Research Council 2021, 6).

Rahoittajat vaativat nykyään enenemissä määrin aineistohallintasuunnitelmaa, ennen rahoituksen myöntämistä (Sewell 2020, 32-33). Rahoittajien vaatimusten myötä, kasvavissa määrin on kehittynyt jonkin tieteenalan ominaispiirteitä huomioivia tai rahoittajakohtaisia pohjia ja työkaluja näitä suunnitelmia varten (Kvale & Pharo 2021). Kuten Briney (2015, 19) huomioi, rahoittajat investoivat mittavia summia tutkimukseen ja aineistohallintasuunnitelma on yksi tapa varmistaa, että nämä sijoitukset eivät mene hukkaan. Vinesin ym. (2014) artikkelissa tulee ilmi, kuinka julkaistun datan saatavuus laskee 17% vuodessa artikkelijulkaisun jälkeen. Asiaa voi katsoa myös hieman toisesta näkökulmasta. Sewellin (2020, 33) mukaan suunnitelma voi auttaa tutkijaa tuottamaan paremman tutkimusprojektin, kun suunnitelma auttaa huomioimaan kaikki vaiheet sekä yksi asia on myös ajansäästö pitkällä tähtäimellä.

Euroopan tiedeneuvosto (ERC) huomioi FAIR-periaatteiden käytön aineistohallinnan suunnittelussa. Horizon Europe rahoitusohjelman alla (2021 ja eteenpäin) ERC:n projekteissa rahoituksen saajien on jätettävä aineistohallintasuunnitelma, jos projektissa syntyy tutkimusdataa. Tämä data pitää tallentaa luotettuun säilytyspaikkaan ja siihen mahdollistetaan pääsy periaatteen ”Niin avointa kuin mahdollista, niin suljettua kuin välttämätöntä” mukaisesti. (European Research Council 2021, 3.)



ERC:n asiakirjassa (European Research Council 2021, 4–5) ehdotetaan, että aineistohallintasuunnitelma sisältäisi tietoa seuraavista kohdista:

1. Datasetin kuvailu (Dataset description).
2. Standardit ja metadata (Standards and metadata).
3. Nimi ja pysyvä tunniste dataseiteille (Name and persistent identifier for the datasets).
4. Kuratointi ja säilyttämisen menetelmä (Curation and preservation methodology).
5. Datam jakamisen menetelmä (Data sharing methodology).<sup>1</sup>

Yksi jäsenys aineistohallintasuunnitelmalle löytyy esimerkiksi Suomen Akatemian aineistohallintasuunnitelmasta (Aineistohallintasuunnitelma, [Viitattu 30.12.2021]; DMPTuuli [Viitattu 30.12.2021]):

1. Aineiston yleiskuvaus.
2. Eettisten periaatteiden ja lainsäädännön noudattaminen.
3. Dokumentointi ja metatiedot.
4. Tallentaminen ja varmuuskopiointi tutkimushankkeen aikana.
5. Aineiston avaaminen, julkaiseminen ja arkistointi tutkimushankkeen päätyttyä.
6. Aineistohallintaa koskevat vastuut ja resurssit.

ERC:n rahoittamissa projekteissa on jo tehty pakolliseksi avoin pääsy julkaisuihin. Myös tutkimusdataa halutaan julkisesti saataville avoimen tieteen pyrkimyksen mukaisesti. Kuitenkin pitää huomioida, että kaikkea dataa ei voi eikä kannata säilyttää pitkäaikaisesti. Esimerkiksi tällöin on syytä huomioida metadatan säilyminen FAIR-periaatteiden mukaisena ja mainita päätös olla säilyttämättä dataa. Kaikki data ei voi välttämättä olla täysin avointa, kun esimerkiksi yksityisyyteen ja turvallisuuteen liittyvät huolet nousevat esiin. (European Research Council 2021, 3).

DMPTuuli on aineistohallinnan työkalu. Se tarjoaa aineistohallintasuunnitelman pohjia ja opastaa niiden tekemiseen. Sen avulla voi tarkastaa tutkimusrahoittajien ja tutkimusorganisaatioiden vaatimukset. (Data Management Guidelines 2021.) Aineistohallinnan työkaluja (DMP tool) on jo saatavilla maailmanlaajuisesti. Varhaisimmat ovat vuodelta 2010-2011. Suomessa käytettävä DMPTuuli hyödyntää DMPRoadmapin

---

<sup>1</sup> tekijän oma käännös

avoimen lähdekoodin koodikantaa. (Jones ym. 2020, 210.). Konetoteutettava aineistohallintasuunnitelma (Machine-Actionable Data Management Plan) on Cardoso, Proença ja Borbinhan (2020, 118) mukaan konsepti, jonka tarkoituksena on tehdä aineistohallintasuunnitelmasta yhteentoimiva, automatisoitu ja yhä enemmän standardisoitu.

## 6 TUTKIMUSONGELMA

Suomessa yliopistojen tehtäviin kuuluu vapaa tutkimus. Valtion tutkimuslaitoksissa tehtävä tutkimus tukee yhteiskunnallista päätöksentekoa ja elinkeinoelämän uudistamista. Myös ammattikorkeakouluissa tehdään tutkimusta. Ammattikorkeakoululain 4§ (14.11.2014/932) mukaan yksi ammattikorkeakoulun tehtävistä on ammatillisen opetuksen lisäksi soveltavan tutkimus-, kehittämis- ja innovaatiotoiminnan (TKI-toiminta) harjoittaminen. Ammattikorkeakoululaissa (14.11.2014/932) pykälässä 4 sanotaan näin:

Ammattikorkeakoulun tehtävänä on antaa työelämän ja sen kehittämisen vaatimuksiin sekä tutkimukseen, taiteellisiin ja sivistyksellisiin lähtökohtiin perustuvaa korkeakouluopetusta ammatillisiin asiantuntijatehtäviin ja tukea opiskelijan ammatillista kasvua. Ammattikorkeakoulun tehtävänä on lisäksi harjoittaa ammattikorkeakouluopetusta palvelevaa sekä työelämää ja aluekehitystä edistävää ja alueen elinkeinorakennetta uudistavaa soveltavaa tutkimustoimintaa, kehittämis- ja innovaatiotoimintaa sekä taiteellista toimintaa. Tehtäviään hoitaessaan ammattikorkeakoulun tulee tarjota mahdollisuuksia jatkuvaan oppimiseen.

Osassa TKI-hankkeista syntyy tutkimusaineistoa. Päällysaho ja Latvanen (2019b) tuovat esiin sen, että organisaatioilla ei välttämättä ole erityisen paljon tietoa siitä, mitä aineistoja TKI-hankkeissa on syntynyt ja kuinka paljon. Tämä opinnäytetyö yrittää tuottaa asiaan uutta tietoa Laurean kohdalla. Myös Sunikka ja Tärppilä (2020) tuovat esille vuoden 2020 Avoimen tieteen kevätpäivillä, että tutkimusorganisaatioilla on puutteellinen käsitys niistä tutkimusaineistoista, joita heillä on hallussaan. Tutkimusdatan avaamisella on kuitenkin mahdollista saavuttaa taloudellisia, sosiaalisia, poliittisia ja ekologisia etuja ammattikorkeakoululle (Päällysaho ym. 2021).

Tämän opinnäytetyön tavoitteena on tuottaa tietoa siitä, minkälaista tutkimusaineistoa Laurean hankkeissa syntyy ja miten tutkimusaineistoa hallinnoidaan. Alatavoitteena on selvittää mitä aineistonhallintaan liittyviä haasteita hanketoimijoilla esiintyy. Myös tutkimusdatan jakamista on tutkittu, mutta sen tarkastelu ei ole varsinainen tavoite tässä opinnäytetyössä, koska aineistojen avoimuuden tasoa ei varsinaisesti tarkastella, eikä kaikissa hankkeissa välttämättä tuoteta Laurean osalta aineistoja ja opinnäytetyö pitää rajata työn laajuuteen nähden.

Tätä tutkimusta ohjaa siis kolme tutkimuskysymystä:

1. Minkälaista aineistoa Laurean hankkeissa syntyy?
2. Miten tutkimusaineistoa hallinnoidaan?
3. Mitä haasteita aineistohallintaan liittyen esiintyy?

Selvittämällä minkälaista tutkimusaineistoa hankkeissa syntyy ja miten sitä hallinnoidaan edesauttaa paremmin antamaan tukea aineistohallinnan eri osa-alueissa. Ashiqin, Usmanin ja Naeemin (2020) systemaattisessa kirjallisuuskatsauksessa tarkasteltiin kirjallisuutta, joka oli ilmestynyt vuosina 2016–2020 liittyen aineistohallinnan käytäntöihin, haasteisiin, palvelujen/taitojen tarpeisiin ja motivaatiotekijöihin. Katsauksessa ilmeni viisi merkittävää haastetta ja estettä, jotka eniten esiintyivät tarkastetussa kirjallisuudessa. Haasteita olivat datan säilyttämiseen liittyvät ongelmat, immateriaalioikeudelliset huolet, rajallinen organisaation antama tuki, riittämätön ja kokematon tutkimuksen tukihenkilöstö sekä tutkijoiden huoli datan katoamiseen ja väärintulkittamiseen.

## 7 TUTKIMUSMENETELMÄ JA TOTEUTUS

### 7.1 Tutkimusmenetelmä

Opinnäytetyön empiirinen osuus on toteutettu laadullisia menetelmiä käyttäen. Hirsjärvi, Remes ja Sajavaara (2010, 161) kertovat, että kvalitatiivisessa eli laadullisessa tutkimuksessa kuvataan todellista elämää ja siinä pyrkimyksenä on kohteen kokonaisvaltainen tutkiminen. He huomauttavatkin kvalitatiivisen tutkimuksen pyrkivän löytämään tai paljastamaan tosiasioita enemmänkin kuin totuusväittämien todentamiseen.

Tuomi ja Sarajärvi (2002, 73) tuovat esiin laadullisessa tutkimuksessa käytettyjä aineistonkeruumenetelmiä, kuten haastattelu, kysely, havainnointi ja dokumentteihin perustuva tieto. Vilka (2015, 122) mainitsee myös, että laadullisella tutkimusmenetelmällä tutkimusaineistoa voidaan kerätä eri tavoin, mutta tutkimusaineisto kerätään usein haastatteluina. Hän lisääkin vielä, että esimerkiksi syvähaastattelua ei pitäisi ajatella laadullisena menetelmänä, vaan yhdestä aineiston keräämisen tavoista laadullisessa tutkimusmenetelmässä. Hirsjärvi ym. (2010, 205–206) mukaan haastattelu valitaan esimerkiksi silloin, kun tutkitaan vähän kartoitettua ja tuntematonta aluetta.

Tuomi ja Sarajärvi (2002, 75) tuovat esille haastattelun joustavuuden. Kysymyksiä voidaan toistaa, väärinkäsityksiä oikaista sekä on mahdollista selventää ilmausten sanamuotoja ja tiedonantajan kanssa voidaan käydä keskustelua. Hirsjärvi ym. (2010, 205) ovat samoilla linjoilla haastattelun eduista, koska siinä tilanteen edellyttämällä tavalla aineiston keruu voidaan toteuttaa säädellysti ja vastaajia myötäillen. Haastattelu aineistonkeruunmenetelmänä antaa mahdollisuuden valita henkilöt, joilla on kokemusta tutkittavasta ilmiöstä tai tietoa aiheesta. Kuitenkin aineiston jäädessä niukaksi, jos esimerkiksi tiedonantajilla ei ole kokemusta aiheesta, ongelmaksi saattaa tulla se, mitä tutkimus ylipäätään kertoo ilmiöstä, jota tutkitaan. (Tuomi & Sarajärvi 2002, 76.)

Tutkimushaastattelumuotoja ovat lomakehaastattelu, teemahaastattelu, josta käytetään myös nimeä puolistrukturoitu haastattelu ja avoin- tai syvähaastattelu (Tuomi & Sarajärvi 2002, 76; Vilka 2015, 123). Haastattelu voidaan toteuttaa strukturoituna,

puolistrukturoituna tai avoimena (Metsämuuronen 2011, 116). Tuomi ja Sarajärvi (2002, 76-77) mainitsevat lomakehaastattelun olevan kvantitatiivisen tutkimuksen aineistonkeruumenetelmä, joka voi olla täysin strukturoitu kysymysten ja tutkimuksen toteutuksen osalta. Puolistrukturoidussa haastattelussa tuodaan esille keskeiset teemat tai aiheet, jotka ovat tutkimusongelman selvittämiseksi välttämättömiä. Teemojen käsittelyjärjestyksen ei tarvitse olla järjestelmällinen. Avoin haastattelu ei rakennu kysymysten ja teemojen varaan. Haastattelukertoja voi olla useita ja vuorovaikutus etenee haastateltavan ehdoilla. (Vilkkä 2015, 124, 126). Hirsjärvi ym. (2010, 209–210) mainitsevat avoimelle haastattelulle tyypilliseksi sen, että se on aikaa vievää ja saattaa vaatia perusteellista koulutusta. Edellä mainituista syistä, avointa haastattelua ei valittu tässä opinnäytetyössä haastattelulajiksi.

Haastattelu valittiin ensisijaiseksi aineistonhankintamenetelmäksi, koska se mahdollistaisi myös haastateltavien näkemysten esittämisen paremmin, kuin esimerkiksi kyselyn avoimet vastaukset. Puolistrukturoitu haastattelu antaisi mahdollisuuden avoimemmalle vuorovaikutukselle ja tarvittaessa voisi esittää lisäkysymyksiä tai tarkentavia kysymyksiä. Kyselyä olisi voitu myös käyttää, mutta alhaisella vastausprosentilla ei olisi voitu yleistää tuloksia. Yksi keino on triangulaatio kuten esimerkiksi Reichman ym. (2021), mutta opinnäytetyön laajuuteen nähden opinnäytetyön tekijä päätti jättää kyselyn pois. Tarkoituksena ei ollut myöskään tehdä vertailevaa tutkimusta aikaisempaan aineistonhallinnan käytäntöjen ja osaamisen kartoitukseen 3AMKin osalta (Lahtinen & Merimaa 2019).

## **7.2 Toteutus**

Opinnäytetyön tekijä haki tutkimuslupaa Laurean organisaatiolta lokakuussa ja se tuli hyväksytyksi takaisin. Opinnäytetyön tekijä laati teemahaastattelurungon (liite 1.), jonka tarkastivat sekä toimeksiantaja että opinnäytetyön ohjaaja. Haastattelurungon laatimisessa apuna käytettiin Avoimuuden palveluverkoston tarveselvityksen haastattelututkimuksen runkoa vuodelta 2019 sekä 3AMKin aineistonhallinnan käytäntöjen ja osaamisen tutkimuksen kyselyrunkoa sekä vastausasiakirjaa.

Opinnäytetyön tekijä laati myös tietosuojaselosteen (liite 3.). Tietosuojaseloste sekä haastattelurunko lähetettiin tutkimuslupa-anomuksen yhteydessä. Hyväksytyt tutkimusluvut jälkeen keskusteltiin mahdollisista haastateltavista. Tekijä keskusteli muutaman Laurean organisaation työntekijän kanssa, ketä voisi haastatella empiirisen aineiston keruuta varten.

Laurean hankelistaa (Laurea.fi/hankkeet) käytiin läpi. Hankkeet rajattiin hankelistassa mainitun päättymisajankohdan mukaan, joka oli aikaisintaan 1.1.2021. Vuonna 2020 päättyneitä hankkeita ei otettu mukaan. Hankkeet rajattiin myös alkamispäivän mukaan siten, että hankelistan mukaan 1.1.2021 jälkeen alkaneet hankkeet jätettiin ulkopuolelle. Hankkeet siirrettiin Excel-taulukkoon, jossa ne vielä sekoitettiin. Lista tuli yhteensä 72 hanketta, jotka olivat päättyneet 1.1.2021 jälkeen tai alkaneet 1.1.2021 tai sen jälkeen. Jos listan kautta olisi satunnaisesti valittu hankkeet, olisi saatettu välttää liiallista mielivaltaisuutta otannon kohdalla. Listaa ei kuitenkaan käytetty haastateltavien valitsemisessa.

Koska kaikissa Laurean hankkeissa ei synny dataa eikä ollut tietoa missä hankkeissa dataa syntyy, niin asiasta jouduttiin kysymään TKI-päälliköiltä. Oli harkittu valinta päättää ensisijaisesti olla haastatteleematta sellaisissa hankkeissa työskennelleitä TKI-toimijoita, joissa ei välttämättä synny suoranaisesti tutkimusaineistoa, sillä silloin ei pystyttäisi suoraan välttämättä vastaamaan tämän opinnäytetyön tutkimusongelmaan. Toimeksiantaja selvitti, missä hankkeissa mahdollisesti oli syntynyt dataa ja antoi opinnäytetyöntekijälle yhteystiedot. Opinnäytetyöntekijä otti yhteyttä mahdollisiin haastateltaviin sähköpostitse ja sopi haastatteluun suostuvien kanssa haastattelun ajankohdasta. Haasteltaville lähetettiin tutkimuksen tiedote (liite 2.) ja tietosuojaseloste sekä haastattelurunko, jotta kysymyksiin oli mahdollista tutustua etukäteen.

Haastattelut järjestettiin Teams-verkkoalustalla 24.11.2021 – 26.1.2022 välisenä aikana. Haastateltavilta kysyttiin lupa nauhoittamiseen. Nauhoitukset litteroitiin ja analysoitiin Ruusuvuoren ja Nikanderin (2017, 367) mukaan haastatteluaineiston analyysi aloitetaan haastattelun litteroinnista. Litterointia voidaan määritellä tekstiksi muuttamiseksi. Nämä tekstiksi muutetut ääni- ja kuvatallenteet muodostavat haastattelututkimuksen tutkimusaineiston. Se kuinka tarkasti näitä litteroidaan, riippuu tutkimuskysymyksestä ja

analyysitavasta. Käytettävä aika sekä rahalliset resurssit saattavat myös vaikuttaa litteroinnin laajuuteen. (Ruusuvuori & Nikander 2017, 367, 372.) Ruusuvuori (2010, 425) tuo esiin, että jos kiinnostus haastattelussa on ilmeneviin asiasisältöihin, ei yksityiskohtaista litterointia välttämättä tarvitse tehdä. Tutkimusaineiston analysointiin otettiin sisältöanalyttinen lähestymistapa. Haastattelut litteroitiin sanatarkasti.



## 8 TULOKSET

Haastateltavat olivat olleet töissä organisaatiossa vähintään kymmenen vuotta ja olivat mukana TKI-toiminnassa ja koulutustoiminnassa vaihtelevissa määrin. Nimekkeinä olivat lehtori, yliopettaja ja projektipäällikkö sekä hankeasiantuntija ja haastateltavia oli yhteensä neljä (H1, H2, H3, H4). TKI-työtä oli tehty ainakin kaksi vuotta. Osa oli ollut TKI-hankkeissa koko ajan ollessaan organisaatiossa sekä yksi (H2) oli ollut tutkimuksen kanssa tekemisissä koko työuran. Haastateltavien työajasta noin 50-80 prosenttia meni TKI-toimintaan. Hankkeissa oli edustettuna liiketalouden, sosiaali- ja terveyden sekä matkailu, ravitsemus ja turismin ala.

**Aineistot.** Yleisimmät tutkimusaineistot, joita syntyi, olivat kysely- ja haastatteluaineistot. Näitä aineistoja mainitsi syntyneen kolme (H2, H3, H4). Tutkimusmenetelminä käytettiin sekä kvantitatiivisia että laadullisia menetelmiä. Useampi kuin yksi (H3, H4) haastateltava mainitsi, että hankkeissa saattoi syntyä sekä kysely- että haastatteluaineistoja. Kyselyt olivat enimmäkseen verkkokyselyitä ja haastattelut yksilöhaastatteluita. Kyselyitä tehtiin esimerkiksi yrityksille.

Ilmeni, että on syntynyt myös kokeellisia aineistoja aivotutkimuksesta yhdessä yliopiston kanssa, kun on tutkittu ihmisen päätöksentekoa (H2). Siinä on syntynyt aivodataa, kun on kuvattu koehenkilöiden aivoja. Hankkeissa myös syntyi työpaja-aineistoja, esimerkiksi pienryhmätehtävien tuloksia tai fläpille kirjattuja tuloksia. Yksi haastateltavista (H4) mainitsi, että työpaja-aineistoja syntyi ehkä kaikissa hankkeissa. Näitä on saatettu ottaa talteen digitaaliseen muotoon valokuvaamalla. Hankkeissa kerättiin myös palauteaineistoja. Jotkin haastateltavien (H1, H2) hankkeet olivat niin alkuvaiheessa, että niiden kohdalla ei osattu sanoa tutkimusaineistojen synnystä hankkeessa.

niin, onhan meillä siinäkin, niin kun on kyselyitä sekä laadullista, että tai niin kun haastatteluja, että kyselyitä. Ja sit meillä on tota, niin ja on työpaja työskentelyä. Ja ollaanhan me niitten pohjalta tehty niinkun tämmösiä ammatillisia julkaisuja ja verkkojulkaisuja ja muuta, mutta ne ei tavallaan, se ei se raakadata, ei esimerkiksi ei ole missään julkisessa muitten käytettävissä. (H4.)

**Säilytys.** Aineistoa säilytettiin esimerkiksi nettikyselyalustalla ja palveluntarjoaja säilytti raakadatan. Näiden tuloksia kuitenkin imuroitiin sieltä omiin tietokantoihin. Aineistoa

säilytettiin kovalevyllä ja lukkojen takana. Aineistoa säilytettiin myös organisaation omassa sisäisessä OneDrivessä sekä omalla tietokoneella, hankkeen verkkolevyllä sekä pilvessä. Sisäisessä Teamsissä säilytettiin laadullisia haastatteluita, kun kaikki toimijat olivat tehneet laadullisia haastatteluita, niin oma raakadata säilytettiin omassa sisäisessä Teamsissä. Yksi haastateltava (H2) mainitsi, että kyselyt saattoi tehdä ja analysoida joku toinen organisaatio, joka sitten toteutti tämän aineiston säilytyksen. Ilmeni, että opiskelijoiden tekemiä tuotoksia ja esimerkiksi kyselyjä, kuten palautekyselyjä säilytettiin Google Drivessä (H4).

Ilmeni, että Githubiin oli tallennettu kyselyaineistoja (H1). Githubin verkkosivustolla sanotaan että: ”miljoonat kehittäjät ja yritykset rakentavat, toimittavat ja ylläpitävät ohjelmistojaan GitHubissa — suurin ja kehittynein kehittämisalusta maailmassa”. Yhdessä haastattelussa (H4) ilmeni, että opiskelijat saattaisivat olla mukana tuotosten tuottamisessa hankkeessa. Esimerkiksi yAMK-opiskelijat suorittavat usein yAMK-opinnäytetyönsä kehittämisprojektina.

Data-arkistot kuten yhteiskuntatieteellinen tietoarkisto tai CERNin ylläpitämä Zenodo, jonka taustalla on avoimen tieteen paradigma, eivät olleet kaikille tuttuja. Asiasta kuitenkin toivottiin lisätietoa esimerkiksi mikä olisi sinne hyödyllistä ja minkälaista tietoa julkisesti voisi sinne jakaa huomioiden hankkeiden rajoitteet. Yhdessä haastattelussa (H2) tuli ilmi yhdessä hankkeessa, että kun kyseessä on ollut yhteistyöhanke, niin tutkija toisesta organisaatiosta on hoitanut tiedon hallinnan. Vaikka varsinaista aineiston tallennusta ei itse toteutettukaan, huomioitiin kuitenkin sen suojaus.

**Aineiston hyödyntäminen.** Yksi haastateltava (H3) mainitsi, että aineistoja oli anonymisoitu ja täten aineistoja pystyisi hyödyntämään jatkokäytössä. Aineiston hyödyntämisen kohdalla esiintyi vain yleisiä huomioita, kun kysyttiin miten tutkimuksellista aineistoa voisi hyödyntää jatkokäytössä. Yksi huomio aineiston jatkokäytön suhteen oli, että jos haluaisi tehdä ison koonnin kaikista palautteista, niin meta-analyysin tekeminen näistä saattaisi olla mahdollista. Toinen huomio jatkokäytön suhteen oli, että toteuttamalla tutkimusta samanlaisena, voisi vertailla tuloksia.

Et niistähän yleensä sit kun on kerätty se aineisto, niin niistä on sitten yleensä kirjoitettu vaikka joku raportti tai joku, niin siinä on sit ainakin ne tulokset, joita voi hyödyntää, mutta niitä aineistojen hyödyntämistä, niin sitä ei oo semmoseen

raaka-aineiston datan hyödyntämistä ei oo tavallaan, ehkä toteutettu juurikaan. (H3.)

Mut nää on vähän uus asia vielä niin, tota täs vähän vielä haetaan vielä niitä toimintatapoja ja sitten mikä toimii meillä ja mitä se sit tarkoittaa ja just et mitä se on sit suhteessa näihin tietosuojakysymyksiin ja mitä se on meidän, niin kun sisäisiin mitä me ollaan luvattu, että ei tota paljastu tai että ei oo tunnistettavissa tai että ne yritykset saa niinkun sen yrityksen, joka osallistuu, että saa sen, niin kun ensiarvon, ettei sitä jaeta sitä tietoa ihan vielä kaikille. (H4.)

Tuli ilmi, että aineistoa pystyisi helposti jakamaan, jos esimerkiksi joku tutkija sitä haluaisi katsoa (H2). Yksi haastateltavista (H2) mainitsi, että jotkut lehdet olivat kysyneet artikkelien yhteydessä voisiko niihin liittyvää dataa julkaista. Jos aineiston otanta on ollut pieni, niin sitä ei haluttu julkaista. Jotta dataa pystyisi enemmän avaamaan pitäisi myös huomioida hankkeen eri sidosryhmien halut. Aineistojen avaamisessa tulisi huomioida se, että, jos hankkeessa on mukana yrityksiä, niille olisi mahdollisuus ensi käden hyötyyn syntyneistä aineistoista. Toisaalta, jos hankkeet rahoitetaan julkisilla varoilla, niin voitaisiin myös näkökulmaa avartaa siihen suuntaan, että aineistoista voisi hyötyä myös TKI hankkeiden ulkopuoliset toimijat.

Kahdessa hankkeessa (H3, H4) syntyneitä aineistoja oli tallennettu tutkimusdata-arkistoon. Toinen arkistoista oli Zenodo. Kumpaakaan aineistoa ei tallennettu hankkeessa työskennelleen toimesta. Yhdessä hankkeessa oli ilmennyt, että hankepartneri oli laittanut opiskelijoiden tuotoksia Zenodoon, mutta sitten todettiin, että ne pitää vetää pois sillä ne sisälsivät esimerkiksi ihmisten nimiä (H4). Avoimia tutkimusaineistoja ei ollut hyödynnetty, kun haastateltavilta kysyttiin, oliko julkisesti saatavilla oleva tutkimusaineistoja hyödynnetty. Tuli ilmi, että vaikka aineistoa ei ollutkaan avattu muiden käyttöön, niin esimerkiksi jos projekti ja resurssit olivat loppuneet ja aineistoa ei ollut päästy analysoimaan, niin uusi henkilö saattoi katsoa dataa, kuitenkin kuuluen oman organisaation porukkaan (H2). Yksi haastateltavista (H2) huomioi:

Kyllähän tietysti, jos se on julkista se data, niin se varmaan pitäis olla jalostettu jotenkin, niin kun ymmärrettävämpään muotoon. Et, jos se ois ihan pelkkää raakadataa, niin se voisi olla vaikeeta ulkopuoliselle sitä ymmärtää. (H2.)

**Muita huomioita.** Aineistonhallintasuunnitelmaa mainitsi käyttäneensä kaksi (H2, H3). Vaikka itse ei oltukaan niitä tehty, niin tiedostettiin, että joku muu taho oli ne hoitanut (H1, H4). Joissain hankkeissa oli hyödynnetty DMPTuuli-työkalua (H3). Yksi huomio aineistonhallintasuunnitelmasta oli, että se tuo laadunvarmistusta prosessiin. Metatieto oli osalle haastatelluista tuttu ja osalle ei ja ei tunnistettu, että sitä olisi käytetty erityisemmin. Yksi haastateltavista tiedosti aineistonhallintasuunnitelman olevan myös joillain rahoittajilla vaatimuksena:

Joo kyllähän nyt periaatteessa rahoittajahan aina edellyttää, siis sitä dokumentaatiota siitä mitä se aineisto käsitellään ja muuta, että et on se tuttu ja sen mukaan on sitten toimittu, että. (H2.)

Tiedostettiin että hankkeiden aiheet olivat ajankohtaisia ja mielenkiintoisia ja tietoa haluttiin myös jakaa siten, kun se on tarpeellista (H1). Tässä asiassa huomioiden myös sitä, mitä rahoittaja haluaa viestittävän. Yhdessä haastattelussa (H2), kun keskusteltiin aineistojen avaamisesta, niin tiedostettiin, että siihen pyritään. Mainittiin myös, että dataa kerätään hyvin suuria määriä, mutta huomioitiin, että joka projektilla on esimerkiksi erilaiset tavoitteet. Mainittiin, että tiettyjen kollegoiden kanssa oli käyty keskustelua, että heidän vanhoja aineistojaan analysoitaisiin jostain toisesta näkökulmasta (H2). Asiassa ei ollut kuitenkaan päästy käytäntöön.

**Haasteet.** Yksi haaste mikä tuli esiin, mutta on enemmänkin projektinhallintaan liittyvää, liittyi tiedon jakamiseen hankkeessa sopimusten puitteissa, joita oli tehty hankkeessa. Yksi näkökulma haasteisiin aineistonhallinnassa oli se, että jotkut lehdet edellyttävät avoimuutta. Ja kuitenkin, kun on sitouduttu, että anonymiteetti säilyy niin, niin dataa ei voitu julkaista. Esiin tuotiin kuitenkin se, että anonymisoimalla sitä voisi julkaista.

Käytäntöjen muuttuminen ja vaatimusten lisääntyminen, joita on pitänyt opetella, oli yksi huomio (H3). Yhdeksi haasteeksi ilmaantui tietosuojaselosteet. Sote-alaan liittyvissä hankkeissa tietosuojakäytännöt ovat olleet tarkempia kuin esimerkiksi liiketalouteen liittyvissä. Toisessa haastattelussa (H4) ilmeni myös se, että sotealan hankkeessa oli otettava tarkemmin tietosuojaan liittyvät kysymykset huomioon.

Haastateltavilta kysyttiin, onko jotain erityistä aineistonhallinnasta tai tutkimusaineistoista mitä he haluaisivat mainita. Yksi huomio oli, että kysymyksen aiheista kokonaiskuva oli puutteellinen (H1). Toinen haastateltava huomioi aineistonhallinnasta ja tutkimusprosessista:

No ei mitään, mun mielestä se on niin kun aika, Laureassa ainakin niin kun hyvällä tasolla toi ja siihen saa, niin kun hyvin apua meiltä ja tota noin sellailta ihan hyvä asia sinänsä, että siihen kiinnitetään niin kun huomiota, että se on kuitenkin aika semmonen niin kuin, niin kun sen ite tutkimusprosessin näkökulmasta se on aika perifeerinen asia siis, sanotaanko näin, et helposti niin kun se unohtuu ja monesti sitte niin kuin, et jossei tämmöstä niin kuin tavallaan rahoittajalta sitä dokumentaatiota edellytetä, miten se tapahtuu, niin tota. (H2.)

Yksi haastateltavista (H4) huomioi, että tärkeää olisi myös oppia mitä aineistoa löytyy käytettäväksi:

nyt oon tietysti tärkeää, että opitaan tallentamaan sitä dataa ja aineistoa sun muuta, mutta sitten toisaalta varmaan vähintään yhtä tärkeää ois oppia, että mitä sieltä löytyy, et oisko siellä itelle jotain käyttöä, et voitaisko sitä, et ei pelkästään, että jaetaan ja että se on nyt sitten johonkin tallennettu ja josta kukaan ei sitä koskaan löydä tai muuta, mutta onks siellä jotain sellasta, mikä vois itselle olla kiinnostavaa hankkeessa. (H4.)

Yhdessä haastattelussa (H4) tuli esiin, että esimerkiksi liikesalaisuuden näkökulmasta, julkaisuissa kirjoitettiin, siten, että esimerkiksi asianomainen hotelli, josta kirjoitettiin, ei mainittu nimellä, että hotelli vastasi jotain ja joltain saarelta ja ne tekevät jollain tavalla. Tuli esiin, että harvoin hanke jää täysin ilman ideaa, mutta IPR:n (intellectual property rights, immateriaalioikeudet) kannalta hankkeet tuottivat enemmänkin yleistä tietämystä ja oivallusta (H2). Kolme keksintöilmoitusta oli tehty, mutta ne oli tehty projektien ulkopuolella ja eivät olleet johtaneet patentteihin.

## 9 POHDINTA JA JOHTOPÄÄTÖKSET

Opinnäytetyön tekijä ei ole ollut mukana TKI-hankkeissa eikä ole opinnäytetyötä tehdessä ollut Laurean organisaatioon työsuhteessa, joten tarkastelukulma tässä on ulkopuolisen näkökulmasta. Myöskään tekijällä ei ole kokemusta TKI-toiminnasta eikä aineistohallinnan kokonaisuudesta, joten tekijältä on saattanut jäädä huomaamatta jotain olennaista. Otanta on hyvin pieni suhteessa vuosittain aloitettavien hankkeiden määrään eikä edusta kaikkia Laurean koulutusaloja, joten tuloksia ei voi yleistää. Tuloksia voi mahdollisesti hyödyntää laajamittaisemmassa samantapaisessa tutkimuksellisessa työssä tai selvityksessä. Opinnäytetyön toimeksiantaja saa uutta tietoa työn kautta, mutta ei todennäköisesti riittävästi esimerkiksi toimintatapojen muutokseen. Vastaisuudessa vastaavanlaisen selvityksen tekijän kannatta olla tekemisissä jo alkuun aineistohallinnan kanssa omassa organisaatiossaan.

Työn teoreettinen puoli oli jokseenkin onnistunut, mutta empiirinen osa jäi kaipaamaan enemmän haastateltavia. Työssä olisi voinut kuitenkin olla enemmän kriittisiä näkökulmia. Empiirinen osa ei siis onnistunut, sillä alkuperäinen tarkoitus oli saada kymmenen haastateltavaa, vaikka silloinkaan tuloksia ei olisi voinut yleistää. Noin puolet haastattelukutsun saaneista eivät vastanneet sähköpostikutsuun. Niukan aineiston takia suorita johtopäätöksiä on hankala tehdä. Haastateltavat eivät osanneet vastata kaikkiin kysymyksiin. Aihe oli aluksi hieman haastava, koska se ei liittynyt varsinaisesti opintosuunnitelmassa käsiteltäviin asioihin. Työn tekijä joutui perehtymään työn aiheeseen, joka toisaalta vei paljon aikaa. Työn loppupuolella opinnäytetyöhön tarkoitettu aika oli jo käytetty, vaikka opinnäytetyö ei ollut valmis. Tämä vaikeutti työn edistymistä työn loppuvaiheessa.

Aikaisemmista selvityksistä oli jo tiedossa, että esimerkiksi kysely- ja haastatteluaineistoja syntyy sekä työpaja-aineistoja. Ehkä uutta tietoa on se, että yhteistyössä yliopiston kanssa tehdyssä hankkeessa syntyi kokeellista aineistoa ja tässä tapauksessa aivodataa. Yksi haastateltavista mainitsi kokonaiskuvan aineistohallinnan ja tutkimusaineistojen kokonaisuudesta olevan puutteellinen, kuitenkin toinen haastateltava mainitsi aineistohallinnan olevan hyvällä tasolla ja siinä koettiin saavan apua. Jos esimerkiksi TKI-työtä tekevällä on taustaa yliopistotutkimuksesta, voi nämä asiat olla selkeämmät jo

entuudestaan. Myös se, että tutkija tai projektityöntekijä on tietoinen avoimista tutkimusaineistoista ja tietää, mistä löytää niitä, on relevantti pohdinnan aihe. Sitähän varten aineistoja halutaan avoimiksi ja julkiseen käyttöön, jotta niistä voi hyötyä mahdollisimman moni ja saattaa esimerkiksi mahdollistaa innovaatioiden syntyä, kun pyörää ei tarvitse uudelleen keksiä. Hankkeiden rahoittajilla saattaa olla erilaisia vaatimuksia tutkimusaineistojen avoimuuden suhteen ja esimerkiksi aineistonhallintasuunnitelmien suhteen. Pitää kuitenkin myös miettiä kriittisesti tutkimusaineiston saattamista avoimeksi. Esimerkiksi kuinka paljon aikaa vievää toimintaa hyvin toteutettu aineistonhallinta on suhteessa aineistojen avoimuuden kautta saatuun yhteiskunnalliseen hyötyyn ja vaikuttavuuteen. Ja mikä on tutkijan tai tutkijan omistusoikeus tutkimuksessa syntyneisiin tutkimusaineistoihin. Kysymyksiä on varmasti monia ja tässä työssä ei ole käsitelty esimerkiksi oikeudellista näkökulmaa eikä eettistä.

Ammattikorkeakoulujen TKI-toiminta on enemmän käytännönläheistä esimerkiksi yritysten ja yhteistyökumppanien rajapinnassa toteutettavaa kehittämistyötä, josta esimerkkinä YAMK-opiskelijoiden toteuttamat kehittämispainotteiset opinnäytetyöt. Ammattikorkeakouluissa perustutkimuksen toteuttaminen ei ole sen pääasiallisia tehtäviä. Puhtaasti tieteen tekemisen kontekstista ei välttämättä kannata ottaa suoria käytäntöjä omakseen, vaan kehittää ammattikorkeakoulujen ekosysteemiin sopivimpia käytäntöjä. Näitä on esimerkiksi pohdittu Avoin TKIO -verkostossa, joka perustettiin Ammattikorkeakoulujen Avoin TKI, oppiminen ja innovaatioekosysteemi-hankkeessa. Kyseinen hanke oli OKM:n rahoittama ja Laurea oli mukana vastaamassa työpakettien toteutuksesta yhdessä hankkeen koordinoijan Seinäjoen ammattikorkeakoulun kanssa.

TKI-toiminnan aineistohallinta ja TKI-toiminnassa syntyvät tutkimusaineistot varmasti puhuttavat vielä tulevaisuudessa. Esimerkiksi sen miten ammattikorkeakoulut huomioivat TKI-aineistojen avoimuuden tulevaisuudessa. Kysymykseksi jää miten tutkimusaineistojen ja -menetelmien avoimuutta voidaan tavoitteellisesti toteuttaa Avoimen tieteen koordinaation, Tieteellisten seurain valtuuskunnan Osalinjaus 1: Tutkimusdatan avoin saatavuus julkaisun mukaisesti. Osalinjauksessa päämäärä on, että: "Tutkimusaineistot ja -menetelmät ovat niin avoimia kuin mahdollista ja niin suljettuja kuin välttämätöntä". Aineistojen hallinnassa pyritään toteuttamaan jo aiemmin työssä mainittuja FAIR-periaatteita tavoitteellisesti.

## LÄHTEET

- ACRL Research Planning and Review Committee. 2020. 2020 top trends in academic libraries: A review of the trends and issues affecting academic libraries in higher education. [Verkkolehtiartikkeli]. College & research libraries news 81(6), 270. Saatavana: <https://doi.org/10.5860/crln.81.6.270>
- ACRL Research Planning and Review Committee. 2014. Top trends in academic libraries: A review of the trends and issues affecting academic libraries in higher education. [Verkkolehtiartikkeli]. College & research libraries news 75(6), 294-302. [Viitattu 13.4.2022]. Saatavana: <https://doi.org/10.5860/crln.75.6.9137>
- Ahonen, M., Latva, M., Päälyysaho, S., & Kärki, A. 2020. Avoimen tki-toiminnan osaaminen ammattikorkeakouluissa. [Verkkolehtiartikkeli]. Modus 3d journal (1). [Viitattu 29.11.2021]. Saatavana: <https://modus3djournal.fi/artikkelit/>
- Aineistonhallinnan käsikirja. 2021. [Verkkojulkaisu]. Tampere: Yhteiskuntatieteellinen tietoarkisto. [Viitattu 3.12.2021]. Saatavana: <https://www.fsd.tuni.fi/fi/palvelut/aineistonhallinta/>
- Aineistonhallintasuunnitelma. ei päiväystä. [Verkkosivu]. Suomen Akatemia. [Viitattu 23.10.2021]. Saatavana: <https://www.aka.fi/tutkimusrahoitus/hae-rahoitusta/nain-haet-rahoitusta/ohjehakemisto/aineistonhallinta/aineistonhallintasuunnitelma/>
- Ammattikorkeakoulujen tutkimusrahoitus. Ei päiväystä. [Excel-tilaus]. Vipunen - opetushallinnon tilastopalvelu. [Viitattu 19.3.2021]. Saatavana: <https://vipunen.fi/fi- layouts/15/xlviewer.aspx?id=/fi-fi/Raportit/Ammattikorkeakoulujen%20tutkimusrahoitus%20-%20amk.xlsb>
- Ammattikorkeakoulujen tutkimushenkilökunta. Ei päiväystä. [Excel-tilaus]. Vipunen - opetushallinnon tilastopalvelu. [Viitattu 19.3.2021]. Saatavana: <https://vipunen.fi/fi- layouts/15/xlviewer.aspx?id=/fi-fi/Raportit/Ammattikorkeakoulujen%20tutkimushenkilokunta%20-%20amk.xlsb>
- Ammattikorkeakoulujen tutkimushenkilökunnan tutkimustyövuodet. Ei päiväystä. [Excel-tilaus]. Vipunen - opetushallinnon tilastopalvelu. [Viitattu 19.3.2021]. Saatavana: <https://vipunen.fi/fi- layouts/15/xlviewer.aspx?id=/fi-fi/Raportit/Ammattikorkeakoulujen%20tutkimusty%C3%B6vuodet%20-%20amk.xlsb>
- Arene ry:n työryhmä. 2017. Innovaatioita, kehittämistoimintaa ja tutkimusta: kaikki kirjaimet käytössä ammattikorkeakoulujen TKI-toiminnassa. [Verkkojulkaisu]. [Viitattu 21.10.2021]. Saatavana: <https://www.arene.fi/julkaisut/raportit/tki-rake-selvitys/>
- Ashiq, M., Usmani, M. H., & Naeem, M. 2020. A systematic literature review on research data management practices and services. [Verkkolehtiartikkeli]. Global knowledge,



memory and communication, ahead-of-print(ahead-of-print). [Viitattu 6.1.2022].  
Saataavana: <http://dx.doi.org/10.1108/GKMC-07-2020-0103>

Avoim kulma: 3AMK avoimuuspolitiikka. päivitetty 29.4.2021. [Verkkosivu].  
libguides.laurea.fi. [Viitattu 20.10.2021]. Saataavana:  
[https://libguides.laurea.fi/avoinkulma/viisasavoimuus?\\_ga=2.52153833.1169754909.1634548054-2146067773.1634548054](https://libguides.laurea.fi/avoinkulma/viisasavoimuus?_ga=2.52153833.1169754909.1634548054-2146067773.1634548054)

Baker, M. 2016. 1,500 scientists lift the lid on reproducibility. [Verkkolehtiartikkeli]. Nature  
533, 452–454. [Viitattu 30.11.2021]. Saataavana: <https://doi.org/10.1038/533452a>

Bloemers, M., & Montesanti, A. 2020. The FAIR funding model: providing a framework for  
research funders to drive the transition toward fair data management and stewardship  
practices. [Verkkolehtiartikkeli]. Data intelligence 2(1–2), 171–180. [Viitattu 3.12.2021].  
Saataavana: [https://doi.org/10.1162/dint\\_a\\_00039](https://doi.org/10.1162/dint_a_00039)

Borgerud, C., & Borglund, E. 2020. Open research data, an archival challenge?  
[Verkkolehtiartikkeli]. Archival science 20(3), 279–302. [Viitattu 18.11.2021].  
Saataavana: <https://doi.org/10.1007/s10502-020-09330-3>

Borgman, C.L. 2015. Big data, little data, no data: scholarship in the networked world.  
[Verkkokirja]. Cambridge, Massachusetts: MIT Press. [Viitattu 22.11.2021]. Saataavana:  
ProQuest Ebook Central -palvelusta. Vaatii käyttöoikeuden.

Briney, K. 2015. Data management for researchers: organize, maintain and share your  
data for research success. Exeter: Pelagic Publishing, UK.

Cambridge dictionary: hakusana Concordat. 2021. [Verkkosivu]. Cambridge University  
Press. [Viitattu 31.12.2021]. Saataavana:  
<https://dictionary.cambridge.org/dictionary/english/concordat>

Cardoso, J., Proença, D., & Borbinha, J. (2020). Machine-Actionable Data Management  
Plans: A Knowledge Retrieval Approach to Automate the Assessment of Funders'  
Requirements. Teoksessa Advances in Information Retrieval (ss. 118–125). Springer  
International Publishing. [https://doi.org/10.1007/978-3-030-45442-5\\_15](https://doi.org/10.1007/978-3-030-45442-5_15)

Cardoso, J., Proença, D., & Borbinha, J. 2020. Machine-actionable data management  
plans: a knowledge retrieval approach to automate the assessment of funders'  
requirements. Teoksessa: J. M. Jose, E. Yilmaz, J. Magalhães, P. Castells, N. Ferro,  
M. J. Silva, & F. Martins (toim.) Springer International Publishing: 42nd European  
Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14–17, 2020,  
Proceedings, Part II. Advances in information retrieval 118–125. [Viitattu 5.1.2022].  
Saataavana: [https://doi.org/10.1007/978-3-030-45442-5\\_15](https://doi.org/10.1007/978-3-030-45442-5_15)

Concordat on open research data. 2016. [Verkkojulkaisu]. UK research and innovation.  
[Viitattu 9.12.2021]. Saataavana: <https://www.ukri.org/wp->

[content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf#:~:text=The%20Concordat%20on%20Open%20Research%20Data%20has%20been,and%20with%20due%20regard%20to%20the%20costs%20involved.](https://content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf#:~:text=The%20Concordat%20on%20Open%20Research%20Data%20has%20been,and%20with%20due%20regard%20to%20the%20costs%20involved.)

Council of the European Union. 2016. The transition towards an Open Science system - Council conclusions. [Verkkajulkaisu]. [Viitattu 1.12.2021]. Saatavana: <https://data.consilium.europa.eu/doc/document/ST-9526-2016-INIT/en/pdf>

Cox, A., & Verbaan, E. 2018. Exploring research data management. [Verkkokirja]. London: Facet. [Viitattu 22.11.2021]. Saatavana: Ebsco eBook collection -palvelusta. Vaatii käyttöoikeude.

Cox, A. M., Pinfield, S., & Smith, J. 2016. Moving a brick building: UK libraries coping with research data management as a 'wicked' problem. [Verkkolehtiartikkeli]. Journal of Librarianship and Information Science 48(1), 3–17. [Viitattu 15.11.2021]. Saatavana: <https://doi.org/10.1177/0961000614533717>

Data management guidelines. 2021. [Verkkajulkaisu]. Tampere: Yhteiskuntatieteellinen tietoaarkisto. [Viitattu 30.12.2021]. Saatavana: <https://www.fsd.tuni.fi/en/services/data-management-guidelines/>

DMPTuuli. ei päiväystä. Public DMP templates. [Verkkosivu]. [Viitattu 30.12.2021]. Saatavana: [https://dmptuuli.fi/public\\_templates](https://dmptuuli.fi/public_templates)

Du, Y., & Khan, H.R. 2020. Data science for librarians. Santa Barbara, California: ABC-CLIO, LLC. [Viitattu 16.11.2021]. Saatavana EBSCO-host Ebooks tietokannasta. Vaatii käyttöoikeuden.

Enwald, H. 2018. Tutkimusaineiston avoin jakaminen – tutkimusorganisaatioiden jäsenten käsityksiä, kokemuksia ja mielipiteitä. [Verkkolehtiartikkeli]. Informaatiotutkimus 37(4), 26–47. [Viitattu 26.11.2021]. Saatavana: <https://doi.org/10.23978/inf.77411>

European research council. 2021. Open research data and data management plans: information for ERC grantees by the ERC scientific council. [Verkkajulkaisu]. [Viitattu 30.12.2021]. Saatavana: [https://erc.europa.eu/sites/default/files/document/file/ERC\\_info\\_document-Open\\_Research\\_Data\\_and\\_Data\\_Management\\_Plans.pdf](https://erc.europa.eu/sites/default/files/document/file/ERC_info_document-Open_Research_Data_and_Data_Management_Plans.pdf)

Fairdata.fi. Ei päiväystä. [Verkkosivu]. FAIR-periaatteet. [Viitattu 1.12.2021]. Saatavana: <https://www.fairdata.fi/tietoa-fairdatasta/fair-periaatteet/>

Force11. ei päiväystä. [Verkkosivu]. Guiding principles for findable, accessible, interoperable and re-usable data publishing. [Viitattu 7.12.2021]. Saatavana: <https://www.force11.org/fairprinciples>

Haasio, A., Harviainen, J.T., & Savolainen, R. 2019. Johdatus tiedonhankintatutkimukseen. Helsinki: Avain.

H1. Laurean työntekijä. Teams-haastattelu.

H2. Laurean työntekijä. Teams-haastattelu.

H3. Laurean työntekijä. Teams-haastattelu.

H4. Laurean työntekijä. Teams-haastattelu.

Hanna, I. 2018. Tampereen yliopiston tutkijoiden näkemyksiä tutkimusaineistojen hallinnasta. [Verkkajulkaisu]. Tampere: Tampereen yliopisto. Viestintätieteiden laitos. Informaatiotutkimus ja interaktiivinen media. Pro gradu -tutkielma. [Viitattu 19.10.2021]. Saatavana: <http://urn.fi/URN:NBN:fi:uta-201805291814>

Hey, T., Tansley, S., & Tolle, K. 2009. Jim Gray on eScience: a transformed scientific method. Teoksessa: T. Hey, S. Tansley & K. Tolle (ed.). The fourth paradigm: data-intensive scientific discovery. Redmond, Washington: Microsoft research 19-33. [Viitattu 1.11.2021]. Saatavana: <https://www.microsoft.com/en-us/research/publication/fourth-paradigm-data-intensive-scientific-discovery/>

Hey, T., & Trefethen, A. 2020. The fourth paradigm 10 years on. [Verkkolehtiartikkeli]. Informatik spektrum 42(6), 441–447. [Viitattu 2.12.2021]. Saatavana: <https://doi.org/10.1007/s00287-019-01215-9>

Hirsjärvi, S., Remes, P. & Sajavaara P. 2010. Tutki ja kirjoita. 15.-16. painos. Helsinki: Tammi.

Hoffman, J.P. 2017. Principles of data management and presentation. [Verkkokirja]. Berkeley, CA: University of California Press. [Viitattu 23.11.2021]. Saatavana: Ebsco eBook Collection -palvelusta. Vaatii käyttöoikeuden.

Hämeen ammattikorkeakoulu. ei päiväystä. [Verkkosivu]. Aineistohallinta. [Viitattu 15.11.2021]. Saatavana: <https://www.hamk.fi/tutkimus/aineistohallinta/>

Jaalamaa, K., Talvela, J., Hyypä, H., Ahlavo, M., Lahtinen, H., Kärki, A., & Päälyssaho, S. 2020. TKI-hankkeiden tuloksia toivotaan nykyistä paremmin jakoon. [Verkkolehtiartikkeli]. UAS journal 10 (1). [Viitattu 2.11.2021]. Saatavana: [https://uasjournal.fi/1-2020/tki-hankkeiden-tuloksia-jakoon/?related\\_post\\_from=6031#](https://uasjournal.fi/1-2020/tki-hankkeiden-tuloksia-jakoon/?related_post_from=6031#)

Jones, S., Pergl, R., Hooft, R., Miksa, T., Samors, R., Ungvari, J., Davis, R. I., & Lee, T. 2020. Data management planning: how requirements and solutions are beginning to converge. [Verkkolehtiartikkeli]. Data intelligence 2(1–2), 208–219. [Viitattu 29.12.2021]. Saatavana: [https://doi.org/10.1162/dint\\_a\\_00043](https://doi.org/10.1162/dint_a_00043)

Joo, S., & Peters, C. 2020. User needs assessment for research data services in a research university. [Verkkolehtiartikkeli]. *Journal of librarianship and information science* 52(3), 633–646. [Viitattu 1.11.2021]. Saatavana: Sage journals -tietokannasta. Vaatii käyttöoikeuden.

Kautonen, T. 2021. Avoin tiede ja tutkimusaineistot: tutkijoiden näkemyksiä ja käytäntöjä. [Verkkojulkaisu]. Tampere: Tampereen yliopisto. Informaatioteknologian ja viestinnän tiedekunta. Informaatiotutkimuksen maisteriohjelma. Pro gradu -tutkielma. [Viitattu 19.10.2021] Saatavana: <http://urn.fi/URN:NBN:fi:tuni-202105034286>

Kielitoimiston sanakirja: hakusana Perustutkimus. ei päiväystä. [Verkkosivu]. Kotimaisten kielten keskus. [Viitattu 25.10.2021]. Saatavana: <https://www.kielitoimistonsanakirja.fi/#/fulltext/perustutkimus?order=entry&field=&baseforms=false>

Kowalczyk, S.T. 2017. Modelling the research data lifecycle. [Verkkolehtiartikkeli]. *International journal of digital curation* 12(2), 331–361. [Viitattu 7.12.2021]. Saatavana: <https://doi.org/10.2218/ijdc.v12i2.429>

Krahe, M. A., Toohey, J., Wolski, M., Scuffham, P. A., & Reilly, S. 2020. Research data management in practice: Results from a cross-sectional survey of health and medical researchers from an academic institution in Australia. [Verkkolehtiartikkeli]. *Health information management journal* 49(2–3), 108–116. [Viitattu 5.1.2022]. Saatavana: <https://doi.org/10.1177/1833358319831318>

Krier, L., Strasser, C.A. 2014. Data management for libraries: a LITA guide. [Verkkokirja]. Chicago, Illinois: ALA TechSource 2014. Saatavana ProQuest Ebook -palvelun kautta. Vaatii käyttöoikeuden.

Kuusniemi, M. E., & Nykyri, S. 2021. Tutkimusdata tieteellisenä julkaisuna. [Verkkolehtiartikkeli]. *Informaatiotutkimus* 40(2), 57–71. [Viitattu 12.11.2021]. Saatavana: <https://doi.org/10.23978/inf.109094>

Kvale, L., & Pharo, N. 2021. Understanding the data management plan as a boundary object through a multi-stakeholder perspective. [Verkkolehtiartikkeli]. *International journal of digital curation* 16(1), 16. <https://doi.org/10.2218/ijdc.v16i1.746>

Laakso, A., Stormi, I. & Mikkonen, A. 2020. Ammattikorkeakoulujen TKI-toiminta Suomen kilpailukyvyn edistäjänä. [Verkkolehtiartikkeli]. *Tieteessä tapahtuu* 38 (5), 33–39. [Viitattu 21.3.2021]. Saatavana: <https://journal.fi/tt/article/view/99571>

Laurea ammattikorkeakoulun strategia 2030. 2019. [Verkkojulkaisu]. Laurea ammattikorkeakoulu. [Viitattu 27.10.2021]. Saatavana: <https://www.laurea.fi/tietoa-meista/strategia-2030/>

Laurea-ammattikorkeakoulun toiminta- ja taloussuunnitelma 2021 – 2023. 2020. Laurea-ammattikorkeakoulu Oy:n hallitus hyväksynyt 8.12.2020. Saatavana: Laurean taloushallinnosta.

Laurea organisaationa. ei päiväystä. [Verkkosivu]. Laurea.fi. [Viitattu 15.10.2021]. Saatavana: <https://www.laurea.fi/tietoa-meista/organisaationa/>

L 14.11.2014/932. Ammattikorkeakoululaki.

Lehto, A. & Suikkanen, E. 2021. Ammattikorkeakoulun TKI-aineistot haltuun datakatalogilla. [Verkkolehtiartikkeli]. Kreodi (3). [Viitattu 26.11.2021]. Saatavana: <http://urn.fi/URN:NBN:fi-fe2021060634287>

MANTRA. 2017. Research data explained. [Sähköinen tietoaaineisto]. Zenodo. <https://doi.org/10.5281/zenodo.1035218>

Marjamaa, M., & Latvanen, J. 2017. Ammattikorkeakoulujen avoimen TKI-toiminnan aineistohallinta – onko sitä? Teoksessa: A. Laakkonen, T. Laakkonen, M. Marjamaa & N. Montonen (toim.) Kohti avointa julkaisemista, opetusta ja TKI-toimintaa ammattikorkeakouluissa. [Verkkajulkaisu]. Vantaa: Laurea- ammattikorkeakoulun julkaisusarja 82, 41-57. [Viitattu 7.1.2022]. Saatavana: <http://urn.fi/URN:ISBN:978-951-799-466-8>

Merimaa, M., & Lahtinen, H. 2019. TKI-aineistojen avaruus — 3amkin aineistohallinnan käytännöt ja osaaminen. [Verkkolehtiartikkeli]. Signum, 51(3), 43–50. [Viitattu 14.2.2021]. Saatavana: <https://doi.org/10.25033/sig.87065>

Metsämuuronen, J. 2011. Laadullisen tutkimuksen käsikirja. [Verkkokirja]. Helsinki: International Methelp. [Viitattu 4.10.2021]. Saatavana. Ellibs-ekirjakokoelmasta. Vaatii käyttöoikeuden.

Michener, W.K. 2015. Ten simple rules for creating a good data management plan. [Verkkolehtiartikkeli]. PLoS comput biol 11(10): e1004525. [Viitattu 29.12.2021]. Saatavana: <https://doi.org/10.1371/journal.pcbi.1004525>

Miglani, S. 2016. Big data and small data: What's the Difference? [Verkkolehtiartikkeli]. Dataversity. [Viitattu 2.5.2022]. Saatavana: <https://www.dataversity.net/big-data-small-data/#>

Milewska, A., Wiśniewska, N., Cimoszko, P., & Rusakow, J. 2021. A survey of medical researchers indicates poor awareness of research data management processes and a role for data librarians. [Verkkolehtiartikkeli]. Health information & libraries journal 00, 1-10. [Viitattu 29.10.2021]. Saatavana: <https://doi.org/10.1111/hir.12403>

- Mäki, K., Huuskonen, S., Leppänen, S., & Toikko, T. 2018 Tampereen yliopiston avoin tiede -kysely 2017 [sähköinen tietoaaineisto]. Versio 1.0 (2018-06-26).  
Yhteiskuntatieteellinen tietoaarkisto. Saatavana: <http://urn.fi/urn:nbn:fi:fsd:T-FSD3221>
- OECD. 2007. OECD Principles and Guidelines for Access to Research Data from Public Funding. [Verkkójulkaisu]. Paris: OECD Publishing. [Viitattu 24.11.2021]. Saatavana: <https://doi.org/10.1787/9789264034020-en-fr>.
- Palsdottir, A. 2021. Data literacy and management of research data: a prerequisite for the sharing of research data. [Verkkolehtiartikkeli]. Aslib journal of information management 73 (2), 322-341. [Viitattu 8.11.2021]. Saatavana: Emerald insight -tietokannasta. Vaatii käyttöoikeuden.
- Parland-von Essen, J., Fält, K., Maalick, Z., Alonen, M., & Gonzalez, E. 2018. Supporting FAIR data: Categorization of research data as a tool in data management. [Verkkolehtiartikkeli]. Informaatiotutkimus 37(4), 127-144. [Viitattu 16.11.2021]. Saatavana: <https://doi.org/10.23978/inf.77419>
- Perrier, L., Blondal, E., & MacDonald, H. 2020. The views, perspectives, and experiences of academic researchers with data sharing and reuse: a meta-synthesis. [Verkkolehtiartikkeli]. PLOS one, 15(2). [Viitattu 2.11.2021]. Saatavana: <https://doi.org/10.1371/journal.pone.0229182>
- Perrier, L., Blondal, E., Ayala, A. P., Dearborn, D., Kenny, T., Lightfoot, D., Reka, R., Thuna, M., Trimble, L., & MacDonald, H. 2017. Research data management in academic institutions: A scoping review. [Verkkolehtiartikkeli]. PLOS one 12(5). [Viitattu 2.11.2021]. Saatavana: <https://doi.org/10.1371/journal.pone.0178261>
- Pouchard, L. 2015. Revisiting the data lifecycle with big data curation. [Verkkolehtiartikkeli]. International journal of digital curation 10(2), 176–192. [Viitattu 6.12.2021]. Saatavana: <https://doi.org/10.2218/ijdc.v10i2.342>
- Päällysaho, S., Kärki, A., Hyyppä, & Laitila, K. 19.05.2021. Tulevaisuuden keskiössä avoin TKI-toiminta ja oppiminen. [Verkkolehtiartikkeli]. UAS journal, 11 (2). [Viitattu 3.1.2021]. Saatavana: <https://uasjournal.fi/2-2021/tulevaisuuden-keskiossa-avoin-tki-toiminta-ja-oppiminen/>
- Päällysaho, S., & Latvanen, J. 2019a. Avoin toimintakulttuuri vahvistuu ammattikorkeakouluissa. [Verkkolehtiartikkeli]. Signum 51(2), 25–29. [Viitattu 7.1.2022]. Saatavana: <https://doi.org/10.25033/sig.83527>
- Päällysaho, S. & Latvanen, J. 2019b. Data käyttöön: näkökulmia ammattikorkeakoulujen datanhallintaan. [Verkkolehtiartikkeli]. UAS journal, 9(2). [Viitattu 28.10.2021]. Saatavana: <https://uasjournal.fi/2-2019/ammattikorkeakoulujen-datanhallinta/>

- Päällysaho, S., Latvanen, J., Lehto, A., Riihimaa, J., Lahti, P., Kärki, A., & Puhakka-Tarvainen, H. 2021. Key Aspects of Open Data in Finnish RDI Cooperation between Higher Education and Businesses. [Verkkolehtiartikkeli]. *Data intelligence* 3(1), 176–188. [Viitattu 26.10.2021]. Saatavana: [https://doi.org/10.1162/dint\\_a\\_00065](https://doi.org/10.1162/dint_a_00065)
- Päällysaho, S & Latvanen, J. 8.3.2017. Avoimen TKI-toiminnan prosessimalli. [Verkkolehtiartikkeli]. *SeAMK verkkolehti*. [Viitattu 27.4.2022]. Saatavana: <https://lehti.seamk.fi/muut-artikkelit/avoimen-tki-toiminnan-prosessimalli/>
- Päällysaho, S., & Latvanen J. 2016. Avoimuus TKI-hankkeiden aineistohallinnassa: Case SeAMK. [Verkkojulkaisu]. Seinäjoen ammattikorkeakoulu. [Viitattu 20.10.2021]. Saatavana: <https://urn.fi/URN:NBN:fi:amk-2016120118572>
- Rantasaari, J. 2021. Doctoral students' educational needs in research data management: Perceived importance and current competencies. *International journal of digital curation* 16(1). [Viitattu 6.1.2022]. Saatavana: <https://doi.org/10.2218/ijdc.v16i1.684>
- Rauhala, P. 2021. 3AMK-liittouma korkeakoulupoliittisena innovaationa: Selvitys Haaga-Helian, Laurea ja Metropolian strategisesta yhteistyöstä ja sen tuloksista 2017–2020. [Verkkojulkaisu]. Helsinki: Metropolia ammattikorkeakoulu. *Metropolia ammattikorkeakoulun julkaisuja Taito-sarja* 75. [Viitattu 27.10.2021]. Saatavana: <https://www.theseus.fi/handle/10024/429937>
- Ray, J.M. 2014. *Research data management*. [Verkkokirja]. West Lafayette, Indiana: Purdue University Press. [Viitattu 29.11.2021]. Saatavana: Ebsco eBook collection - palvelusta. Vaatii käyttöoikeuden
- RDA Europe. 2014. *The data harvest: how sharing research data can yield knowledge, jobs and growth*. [Verkkojulkaisu]. [Viitattu 2.12.2021]. Saatavana: <https://rd-alliance.org/data-harvest-report-sharing-data-knowledge-jobs-and-growth.html>
- Reichmann, S., Klebel, T., Hasani-Mavriqi, I., & Ross-Hellauer, T. 2021. Between administration and research: understanding data management practices in an institutional context. [Verkkolehtiartikkeli]. *Journal of the association for information science and technology* 72(11), 1415–1431. [Viitattu 28.1.2022]. Saatavana: <https://doi.org/10.1002/asi.24492>
- Resnik, D. B., Morales, M., Landrum, R., Shi, M., Minnier, J., Vasilevsky, N. A., & Champieux, R. E. 2019. Effect of impact factor and discipline on journal data sharing policies. [Verkkolehtiartikkeli]. *Accountability in research: policies & quality assurance* 26(3), 139–156. [Viitattu 30.11.2021]. Saatavana: <https://doi.org/10.1080/08989621.2019.1591277>
- Rissanen, R. 2017. Ammattikorkeakoulut avoimen tieteen, tutkimuksen ja innovaatioiden edistäjinä. Teoksessa: A. Laakkonen, T. Laakkonen, M. Marjamaa & N. Montonen (toim.) *Kohti avointa julkaisemista, opetusta ja TKI-toimintaa ammattikorkeakouluissa*.

[Verkkojulkaisu]. Vantaa: Laurea- ammattikorkeakoulun julkaisusarja 82, 6-7. [Viitattu 7.1.2022]. Saatavana: <http://urn.fi/URN:ISBN:978-951-799-466-8>

Ruusuvuori, J., & Nikander, P. 2017. Haastatteluaineiston litterointi. Teoksessa: M. Hyvärinen, P. Nikander & J. Ruusuvuori (toim.) Tutkimushaastattelun käsikirja. [Verkkokirja]. Tampere: Vastapaino, 367-380. [Viitattu 4.1.2020]. Saatavana Ellibs-e-kirjakokoelmasta. Vaatii käyttöoikeuden.

Ruusuvuori, J. 2010. Litteroijan muistilista. Teoksessa: J. Ruusuvuori, P. Nikander & M. Hyvärinen (toim.) Haastattelun analyysi. [Verkkokirja]. Tampere: Vastapaino, 424-431. [Viitattu 7.1.2022]. Saatavana Ellibs-e-kirjakokoelmasta. Vaatii käyttöoikeuden.

Satama, M. 2021. "Metadata on mahdoton, metadata pelottaa". Tutkijat ja tiedonhallinnan ammattilaiset tutkimusaineistojen hallinnan äärellä. [Verkkojulkaisu]. Joensuu, Itä-Suomen yliopisto. Yhteiskuntatieteiden ja kauppatieteiden tiedekunta, Asiakirjahallinnan ja arkistoalan maisteriohjelma. Pro gradu -tutkielma. [Viitattu 19.10.2021]. Saatavana: <http://urn.fi/urn:nbn:fi:uef-20210438>

Sewell, C. 2020. The no-nonsense guide to research support and scholarly communication. London: Facet Publishing.

Springer Nature. 2022. Reporting standards and availability of data, materials, code and protocols. [Verkkosivu]. [Viitattu 5.1.2022]. Saatavana: <https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards>

Stamatoplos, A., Neville, T., & Henry, D. 2016. Analyzing the data management environment in a master's-level institution. [Verkkolehtiartikkeli]. The journal of academic librarianship 42 (2), 154–160. [Viitattu. 23.9.2021]. Saatavana: <https://doi.org/10.1016/j.acalib.2015.11.004>

Sunikka, A., & Näppilä, T. 2020. Tutkimusaineistojen hallinta osana tutkimuksen elinkaarta. Avoin Tiede/Datan avoimuuden asiantuntijaryhmä/Aineistot organisaatioissa –työryhmän loppuraportti, 13.5.2020. Avoimen tieteen kevätpäivät 2020, Helsinki (online). Zenodo. <https://doi.org/10.5281/zenodo.3820215>

Suomen yliopistojen rehtorineuvosto UNIFI ry. 2018. Avoin tiede ja data: toimenpideohjelma suomalaiselle tiedeyhteisölle. [Verkkojulkaisu]. [Viitattu 1.12.2021]. Saatavana: <http://www.unifi.fi/hankkeet/avoin-tiede-ja-data/>

Talvela, J., Jaalamaa, K., Hyypä, H., Ahlavo, M., Päällysaho, S., Lahtinen, H., & Kärki, A. 2020. Avoin TKI ammattikorkeakouluissa: tekijöiden näkemyksiä. [Verkkolehtiartikkeli]. UAS journal 10 (1). [Viitattu 2.11.2021]. Saatavana: [https://uasjournal.fi/1-2020/avoin-tki-nakemyksia/?related\\_post\\_from=6033](https://uasjournal.fi/1-2020/avoin-tki-nakemyksia/?related_post_from=6033)

Tenopir, C., Rice, N. M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., & Sandusky, R. J. 2020. Data sharing, management, use, and reuse: practices and



perceptions of scientists worldwide. [Verkkolehtiartikkeli]. PLOS ONE, 15(3). [Viitattu 29.10.2021]. Saatavana: <https://doi.org/10.1371/journal.pone.0229003>

Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D., & Dorsett, K. 2015. Changes in data sharing and data reuse practices and perceptions among scientists worldwide. [Verkkolehtiartikkeli]. PLOS one, 10(8). [Viitattu 1.11.2021]. Saatavana: <https://doi.org/10.1371/journal.pone.0134826>

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., & Frame, M. 2011. Data sharing by scientists: practices and perceptions. [Verkkolehtiartikkeli]. PLOS one, 6(6). [Viitattu 1.11.2021]. Saatavana: <https://doi.org/10.1371/journal.pone.0021101>

Tiedonjulkistamisen neuvottelukunta ja Tieteellisten seurain valtuuskunta. 2021. Vastuullisen tieteen julkaisusarja 5 (3). Tutkimusaineistojen ja -menetelmien avoimuus. Korkeakoulu- ja tutkimusyhteisön kansallinen linjaus ja toimenpideohjelma 2021–2025. Osalinjaus 1: Tutkimusdatan avoin saatavuus. [Verkkojulkaisu]. [Viitattu 15.11.2021]. Saatavana: <https://doi.org/10.23847/isbn.9789525995466>

Tieteen termipankki. Päivitetty 18.11.2018. [Verkkosivu]. Avoin tiede: tutkimusaineisto. [Viitattu 27.10.2021]. Saatavana: [https://tieteentermipankki.fi/wiki/Avoin\\_tiede:tutkimusaineisto](https://tieteentermipankki.fi/wiki/Avoin_tiede:tutkimusaineisto)

Tutkimus- ja kehittämistoiminta. [ei päiväystä]. Tilastokeskus: Helsinki. [Viitattu 25.10.2021]. Saatavana: [http://www2.tilastokeskus.fi/meta/kas/t\\_ktoiminta.html#tab2](http://www2.tilastokeskus.fi/meta/kas/t_ktoiminta.html#tab2)

Tutkimus, kehitys ja innovaatiot. ei päiväystä. [Verkkosivu]. Laurea.fi. [Viitattu 19.10.2021]. Saatavana: <https://www.laurea.fi/tki/>

Tuomi, J. & Sarajärvi. 2002. Laadullinen tutkimus ja sisällönanalyysi. Jyväskylä Gummerrus.

Valtioneuvosto. 23.4.2020. [Verkkojulkaisu]. Kestävän ja kehittyvän yhteiskunnan ratkaisuja tuottava Suomi. [Viitattu 18.3.2021]. Saatavana: <https://minedu.fi/tki-tiekartta>

Viisas avoimuus: 3AMK:n yhteiset avoimuuden lupaukset ja periaatteet. 2018. [Verkkojulkaisu]. Haaga-Helion, Laurean ja Metropolian strateginen liitto (3AMK). [Viitattu 20.10.2021]. Saatavana: [https://libguides.laurea.fi/avoinkulma/viisasavoimuus?\\_ga=2.52153833.1169754909.1634548054-2146067773.1634548054](https://libguides.laurea.fi/avoinkulma/viisasavoimuus?_ga=2.52153833.1169754909.1634548054-2146067773.1634548054)

Vilka, H. 2015. Tutki ja kehitä. 4. uudistettu painos. Jyväskylä: PS-kustannus.

Vines, T. H., Albert, A. Y. K., Andrew, R. L., Débarre, F., Bock, D. G., Franklin, M. T., Gilbert, K. J., Moore, J.-S., Renaut, S., & Rennison, D. J. 2014. The availability of research data declines rapidly with article age. [Verkkolehtiartikkeli]. Current biology

24(1), 94–97. [Viitattu 31.12.2021]. Saatavana:  
<https://doi.org/10.1016/j.cub.2013.11.014>

Visio 2030 työryhmien raportit. Ei päiväystä. Opetus- ja kulttuuriministeriö.  
[Verkkajulkaisu]. [Viitattu 21.3.2021]. Saatavana: <https://minedu.fi/korkeakoulutuksen-ja-tutkimuksen-visio-2030> Julkaisu löytyy nimellä ”temaattisten valmisteluryhmien raportit”.

Whyte, A., & Tedds, J. 2011. Making the case for research data management.  
[Verkkajulkaisu]. DCC Briefing Papers. Edinburgh: Digital Curation Centre. [Viitattu 3.12.2021]. Saatavana: <https://www.dcc.ac.uk/guidance/briefing-papers/making-case-rdm>

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. 2016. The FAIR guiding principles for scientific data management and stewardship. [Verkkolehtiartikkeli]. Scientific data 3(1), 160018. [Viitattu 26.1.2022]. Saatavana: <https://doi.org/10.1038/sdata.2016.18>

Yhteiskuntatieteellinen tietoarkisto. 2021. Tutkimusaineiston elinkaari. [Kuva]. Tampere: Yhteiskuntatieteellinen tietoarkisto. [Viitattu 24.4.2022]. Saatavana: <https://www.fsd.tuni.fi/fi/palvelut/aineistonhallinta/miksi-aineistonhallintaa-ja-jatkokayttoa/> CC-BY

## LIITTEET

Liite 1. Teemahaastattelurunko

Liite 2. Tiedote tutkimuksesta

Liite 3. Tutkimuksen tietosuojaseloste

## **Liite 1. Teemahaastattelurunko**

### **Teemahaastattelurunko – tutkimusaineistot Laurean hankkeissa**

#### **Tausta**

Kauanko olet ollut organisaatiossa?

Mikä on asemasi organisaatiossa?

Miten kauan olet tehnyt TKI-työtä?

Paljonko työajasta menee TKI-toimintaan?

Minkälaisien hankkeiden kanssa olet ollut tekemisissä?

#### **Tutkimusaineistot ja aineistohallinta**

Millaisia TKI-aineistoja hankkeissa, joissa olet ollut mukana, on tuotettu?

Minkä alan hankkeita on ollut kyseessä?

Mikä on ollut hankkeiden laajuus? esim. kansainvälinen tai kansallinen hanke tai hankkeen kesto ja kuinka monta organisaatiota on ollut mukana?

Minkälaista tutkimusaineistoa hankkeessa on tuotettu, johon hankkeen tai tutkimuksen tulokset perustuvat?

-onko tutkimusaineisto kvantitatiivista vai kvalitatiivista?

Millä tavoin keräät tutkimusaineistoa?

Millä tavoin tutkimuksen aikana syntynyt tutkimusaineisto on järjestetty?

Miten tutkimusaineistoa säilytetään tutkimuksen aikana?

Miten tutkimusaineistoa säilytetään tutkimuksen jälkeen?

Miten tietosuoja ja tietoturva huomioidaan aineiston säilytyksessä?

Minne tutkimusaineistoa on tallennettu?

Mitä asioita pitäisi ottaa huomioon mahdollisessa aineiston jatkokäytössä?

Miten aineistoa voisi hyödyntää jatkossa?

Miten aineistoa on kuvailtu?

Mitä aineistoa kuvailevia tietoja lisäät tutkimusaineistoon?

Miten varmistetaan, että aineistosta löytyy eri versioita?

Onko aineistoa tallennettu data-arkistoon hankkeen päätyttyä tai jos hanke ei ole vielä päättynyt voisiko aineistoa hankkeen päätyttyä tallentaa data-arkistoon?

Onko osa aineistosta hyödynnettävissä jatkokäytössä?

Millaisia ongelmia mahdollisessa aineiston jatkokäytössä saattaisi esiintyä?

Oletko avannut aineistoa muiden käyttöön?

Oletko hyödyntänyt valmiita tutkimusaineistoja? Esimerkiksi yhteiskuntatieteellisen tietoarkiston tai Zenodon aineistoja.

Oletko käyttänyt aineistonhallintasuunnitelmaa hankkeessa?

Onko rahoittaja vaatinut aineistonhallintasuunnitelmaa tai tutkimusaineiston jakamista?

Onko jokin osa-alue aineistonhallinnassa, joka erityisesti on ongelmallinen?

Jos hankkeet ovat olleet yhteishankkeita, miten se on vaikuttanut aineiston hallintaan?

Mitä tietosuojan ja -turvaan liittyviä haasteita tutkimusaineiston kannalta esiintyy?

Onko jotain erityistä aineistonhallinnasta tai tutkimusaineistosta mitä haluaisit mainita?

## Liite 2. Tiedote tutkimuksesta



Tiedote tutkimuksesta ja tutkimuksen tietosuojaseloste

# TIEDOTE TUTKIMUKSESTA

## Tutkimusaineistot Laurean hankkeissa.

### 1. Pyyntö osallistua tutkimukseen

Teitä pyydetään mukaan tutkimukseen, jossa tutkitaan minkälaista tutkimusaineistoa Laurean TKI-hankkeissa syntyy ja miten aineistoa hallinnoidaan. Tämä tiedote kuvaa tutkimusta ja teidän osuuttanne siinä.

### 2. Vapaaehtoisuus

Tutkimukseen osallistuminen on täysin vapaaehtoista. Kieltäytyminen ei vaikuta oikeuksiinne ja kohteluunne yhteisön työntekijänä.

Voitte myös keskeyttää tutkimuksen koska tahansa syytä ilmoittamatta. Mikäli keskeytätte tutkimuksen tai peruutatte suostumuksen, teistä keskeyttämiseen ja suostumuksen peruuttamiseen mennessä kerättyjä tietoja ja näytteitä voidaan käyttää osana tutkimusaineistoa.

### 3. Tutkimuksen tarkoitus

Tämän tutkimuksen tarkoituksena on kartoittaa, minkälaisia tki-aineistoja Laurea hankkeissa tuotetaan ja miten niitä hallinnoidaan sekä mitä haasteita aineistonhallinnassa esiintyy. Aineistonhankinta tehdään yksilöhaastatteluin.

### 4. Tutkimuksen toteuttajat

Tutkimuksen toteuttaa opinnäytetyön tekijä Seinäjoen ammattikorkeakoulun opiskelija Oskari Saarinen kirjaston- ja tietopalvelualan tutkinto-ohjelmasta. Toimeksiantaja on Laurea ammattikorkeakoulu.

### 5. Tutkimusmenetelmät ja toimenpiteet

Tutkimus toteutetaan siten, että tki-hankkeissa työskenteleviä haastatellaan yksilöhaastatteluilla. Haastattelun arvioitu kesto on 30-60 minuuttia.

**6. Kustannukset ja niiden korvaaminen**

Tutkimukseen osallistuminen ei maksa teille mitään. Osallistumisesta ei myöskään makseta erillistä korvausta.

**7. Tutkimustuloksista tiedottaminen**

Tutkimustulokset esitetään opinnäytetyössä, joka julkaistaan avoimesti Theseus-tietokannassa.

**8. Tutkimuksen päätyminen**

Myös tutkimuksen suorittaja voi keskeyttää tutkimuksen.

**9. Lisätiedot**

Pyydämme teitä tarvittaessa esittämään tutkimukseen liittyviä kysymyksiä tutkijalle/tutkimuksesta vastaavalle henkilölle.

**10. Tutkijoiden yhteystiedot**

Opinnäytetyöntekijä  
Nimi: Oskari Saarinen  
Puh.  
Sähköposti:

Opinnäytetyön ohjaaja  
Titteli:  
Nimi:  
Seinäjoen ammattikorkeakoulu / Liiketoiminta ja kulttuuri  
Puh.  
Sähköposti:

Suostun henkilötietojen keräämiseen ja osallistun tutkimukseen

\_\_\_\_\_

Allekirjoitus: \_\_\_\_\_

Nimenselvennys: \_\_\_\_\_

### Liite 3. Tutkimuksen tietosuojaseloste



Tiedote tutkimuksesta ja tutkimuksen tietosuojaseloste

## TUTKIMUKSEN TIETOSUOJASELOSTE

Tässä tutkimuksessa käsitellään teitä koskevia henkilötietoja voimassa olevan tietosuojalainsäädännön (EU:n yleinen tietosuoja-astus, 679/2016, ja voimassa oleva kansallinen lainsäädäntö) mukaisesti. Seuraavassa kuvataan henkilötietojen käsittelyyn liittyvät asiat.

### 1. Tutkimuksen rekisterinpitäjä

Rekisterinpitäjällä tarkoitetaan tahoa, joka yksin tai yhdessä toisten kanssa määrittelee henkilötietojen käsittelyn tarkoitukset ja keinot. Rekisterinpitäjä voi olla korkeakoulu, toimeksiantaja, muu yhteistyötaho, opinnäytetyöntekijä tai jotkut edellä mainituista yhdessä (esim. korkeakoulu ja opinnäytetyöntekijä yhdessä). Opinnäytetöissä rekisterinpitäjänä on pääsääntöisesti opinnäytetyöntekijä.

Tässä tutkimuksessa henkilötietojen rekisterinpitäjä on

- |                     |                                     |                       |
|---------------------|-------------------------------------|-----------------------|
| Korkeakoulu         | <input type="checkbox"/>            |                       |
| Toimeksiantaja      | <input type="checkbox"/>            | Toimeksiantajan nimi: |
| Muu yhteistyötaho   | <input type="checkbox"/>            | Yhteistyötahon nimi:  |
| Opinnäytetyöntekijä | <input checked="" type="checkbox"/> |                       |

### 2. Voitte kysyä lisätietoja henkilötietojenne käsittelystä rekisterinpitäjän yhteyshenkilöltä

Rekisterinpitäjän yhteyshenkilön nimi: Oskari Saarinen  
 Organisaatio:  
 Puh.  
 Sähköposti:



### **3. Tutkimuksessa teistä kerätään seuraavia henkilötietoja**

Henkilötiedolla tarkoitetaan luonnollista henkilöä, hänen ominaisuuksiaan tai elinolosuhteitaan kuvaavia tietoja, jotka voidaan tunnistaa häntä, hänen perhettään tai hänen kanssaan yhteisessä taloudessa asuvia koskeviksi. Henkilötiedon määritelmään kuuluvat sellaiset henkilötiedot, joiden avulla henkilö voidaan tunnistaa suoraan, mutta myös sellaiset tiedot, joiden avulla henkilö voidaan tunnistaa välillisesti tai epäsuorasti.

Henkilötietojen käsittely on oikeutettua ainoastaan silloin, kun se on tutkimukselle välttämätöntä. Kerättävät henkilötiedot on minimoitava, niitä ei saa kerätä tarpeettomasti tai varmuuden vuoksi.

Suostumuslomakkeesta kerätään nimi. Tutkimuksessa tehdään haastatteluita, jotka nauhoitetaan. Nauhoitukset kirjoitetaan puhtaaksi, kuitenkin siten, että puhujaa ei voi tunnistaa niistä.

Teillä ei ole sopimukseen tai lakisääteiseen tehtävään perustuvaa velvollisuutta toimittaa henkilötietojanne vaan osallistuminen on täysin vapaaehtoista.

### **4. Tutkimuksessa kerätään henkilötietojanne myös seuraavista lähteistä**

Tutkimuksessa ei kerätä henkilötietojanne muista lähteistä.

### **5. Henkilötietojenne suojausperiaatteet**

Suostumuslomakkeet säilytetään käyttäjätunnuksella ja salasanalla suojatulla tietokoneella. Pääsy lomakkeisiin on vain opinnäytetyön tekijällä.

Haastattelut tehdään pääsääntöisesti verkossa Teams alustalla. Haastatteluiden nauhoitukset säilytetään käyttäjätunnuksella ja salasanalla suojatulla tietokoneella. Toinen versio säilytetään korkeakoulun verkkolevyllä, joka on suojattu käyttäjätunnuksella ja salasanalla. Pääsy nauhoituksiin on vain opinnäytetyön tekijällä.

### **6. Henkilötietojenne käsittelyn tarkoitus**

Tutkimuksen tarkoituksena on kartoittaa, minkälaisia tki-aineistoja Laurea hankkeissa tuotetaan ja miten niitä hallinnoidaan sekä mitä haasteita aineistonhallinnassa esiintyy. Aineistonhankinta tehdään haastatteluin.

### **7. Henkilötietojenne käsittelyperuste**

Henkilötietojen käsittelyn oikeusperuste on säädetty tietosuojasetuksessa.

Henkilötietojen käsittelylle pitää aina olla tietosuojasetuksen tai tietosuojalain mukainen käsittelyperuste. Jos perustetta ei ole, henkilötietoja ei saa käsitellä. Vain käsittelyn tarkoituksen kannalta tarpeellisia tietoja saa käsitellä.

Tässä opinnäytetyössä tietojen käsittelyperuste on suostumus.

## **8. Tutkimuksen kesto aika (henkilötietojenne käsittelyaika)**

Henkilötiedot tuhoetaan sen jälkeen, kun opinnäytetyö on hyväksytty ja muutoksenhaku aika on päättynyt.

## **9. Mitä henkilötiedoillenne tapahtuu tutkimuksen päätyttyä?**

Henkilötiedot tuhoetaan tutkimuksen päätyttyä.

## **10. Tietojen luovuttaminen tutkimusrekisteristä**

Tietoja ei luovuteta tutkimusryhmän ulkopuolelle.

## **11. Henkilötietojenne mahdollinen siirto EU:n tai ETA-alueen ulkopuolelle**

Tietojanne ei siirretä EU:n tai ETA-alueen ulkopuolelle.

## **12. Rekisteröitynä teillä on oikeus**

Koska henkilötietojanne käsitellään tässä tutkimuksessa, niin olette rekisteröity tutkimuksen aikana muodostuvassa henkilörekisterissä. Rekisteröitynä teillä on oikeus:

- saada informaatiota henkilötietojen käsittelystä
- tarkastaa itseänne koskevat tiedot
- oikaista tietojanne
- poistaa tietonne (esim. jos peruutatte antamanne suostumuksen)
- peruuttaa antamanne henkilötietojen käsittelyä koskeva suostumus
- rajoittaa tietojenne käsittelyä
- rekisterinpitäjän ilmoitusvelvollisuus henkilötietojen oikaisusta, poistosta tai käsittelyn rajoittamisesta
- siirtää tietonne järjestelmästä toiseen
- sallia automaattinen päätöksenteko nimenomaisella suostumuksellanne
- tehdä valitus tietosuojavaltuutetun toimistoon, jos katsotte, että henkilötietojanne on käsitelty tietosuojalainsäädännön vastaisesti

Jos henkilötietojen käsittely tutkimuksessa ei edellytä rekisteröidyn tunnistamista ilman lisätietoja eikä rekisterinpitäjä pysty tunnistamaan rekisteröityä, niin oikeutta tietojen tarkastamiseen, oikaisuun, poistoon, käsittelyn rajoittamiseen, ilmoitusvelvollisuuteen ja siirtämiseen ei sovelleta.

Voitte käyttää oikeuksianne ottamalla yhteyttä rekisterinpitäjään.

### **Tutkimuksessa kerättyjä henkilötietoja ei käytetä profilointiin tai automaattiseen päätöksentekoon**

### **Henkilötietojen käsittely aineistoa analysoitaessa ja tutkimuksen tuloksia raportoitaessa**

Teistä kerättyä tietoa ja tutkimusaineistoa käsitellään luottamuksellisesti lainsäädännön edellyttämällä tavalla. Yksittäisille tutkittavalle annetaan tunnuskoodi ja häntä koskevat tiedot säilytetään koodattuina tutkimusaineistossa. Aineisto analysoidaan koodattuna ja tulokset raportoidaan ryhmätasolla, jolloin yksittäinen henkilö ei ole tunnistettavissa ilman koodiavainta. Koodiavainta, jonka avulla yksittäisen tutkittavan tiedot ja tulokset voidaan tunnistaa, säilyttävät opiskelija Oskari Saarinen siihen asti, kunnes opinnäytetyö on hyväksytty ja muutoksenhaku-aika on päättynyt. Tietoja ei anneta tutkimuksen ulkopuolisille henkilöille. Lopulliset tutkimustulokset raportoidaan ryhmätasolla eikä yksittäisten tutkittavien tunnistaminen ole mahdollista.

Tutkimusaineistoa ja tutkimuksen yhteydessä kerättyjä näytteitä säilytetään opiskelijan tietokoneella ja korkeakoulun verkkolevyllä ja henkilötiedot hävitetään sen jälkeen, kun opinnäytetyö on hyväksytty ja muutoksenhaku-aika on päättynyt. Opinnäytetyön on tarkoitus valmistua kevään 2022 aikana. Suostumuslomakkeet, haastatteluiden nauhoitukset ja nauhoitusten litteroinnit tuhotaan pysyvästi.

Kerättyjä tietoja ei käytetä tutkimuksiin myöhemmin.

Aineistoa ei käytetä jatkotutkimuksissa.