

Anna-Riikka Pasanen

# **AINEISTONHALLINNAN ERI TOIMIJAROOLEISSA**

Tamperelaisten korkeakoulujen tutkijoiden  
suhtautuminen tutkimusaineistojen avoimuuteen ja  
hyödyntämiseen Tampere3-hankkeen aikaan

Informaatioteknologian ja viestinnän tiedekunta

Pro gradu -tutkielma

Huhtikuu 2021

# TIIVISTELMÄ

Anna-Riikka Pasanen: Aineistonhallinnan eri toimijaroleissa – tamperelaisten korkeakoulujen tutkijoiden suhtautuminen tutkimusaineistojen avoimuuteen ja hyödyntämiseen Tampere3-hankkeen aikaan  
Pro gradu -tutkielma  
Tampereen yliopisto  
Informaatiotutkimuksen maisteriohjelma  
Huhtikuu 2021

---

Tutkielmassani selvitettiin ensiksikin, miten Tampereen yliopiston (TaY), Tampereen teknillisen yliopiston (TTY) ja Tampereen ammattikorkeakoulun (TAMK) tutkijat suhtautuivat Tampere3-hankkeen aikoihin tutkimusaineistojen avoimuuden toteuttamiseen ja tutkimusaineistojen hyödyntämiseen, toiseksi erosivatko heidän aineistonhallinnan näkemyksensä ja käytäntönsä toisistaan ja kolmanneksi vaikuttiko korkeakoulutausta niihin. Aihetta taustoitettiin esimerkiksi avoimen tieteen periaatteilla ja tutkimusaineistojen elinkaarella sekä aiempien tutkimusten havainnoilla aineistojen avoimuuden esteistä ja kannustimista.

Kolmen korkeakoulun intranet-sivuilla keväällä 2018 jaetulla kyselyllä saatiin 16 analysointikelpoista vastausta. Saadut tulokset esitettiin hyödyntämällä viitekehystä, jossa aineistonhallinta jäsennettiin aineiston keruun, normien, tiedeyhteisön, uudelleenkäytön, käytettävyyden ja tiedeinstituutioiden osa-alueisiin. Niihin liitettiin aineistonhallinnalle hahmotetut toimijaroolit: aineiston kerääjä, suojelija, käsittelijä, uudelleenkäyttäjä ja uudelleenkäytön valvoja sekä aineistonhallinnan tuen saaja ja tiedeyhteisön jäsen. Roolijakoon johdatti kirjallisuudesta noussut huomio, jonka mukaan tutkimusta ja aineistonhallintaa koskevat toimijaroolit voivat vaikuttaa aineistonhallinnan vastuunjakoon ja tuloksellisuuteen. Tunnistamalla, millaisiin rooleihin eritaustaiset tutkijat astuvat – onko näkemyksissä tai käyttäytymismalleissa samaa – voidaan osaltaan tukea Tampereen uuden korkeakoulu-yhteisön tutkimusaineistojen hallinnan, ja siten myös tutkimustyön, kehittämistoimien mielekäästä kohdentamista.

Tutkijoiden suhtautumista tutkimusaineistojen avoimuuteen selvitettiin tiedustelemalla heidän kantojaan ensiksikin aineistonhallinnan eri rooleihin liitettyihin syihin jättää aineistoja jakamatta ja toiseksi ehtoihin, joita aineistojen avoimuudelle voisi asettaa. TAMK:n ja TaY:n vastaajat olivat aktiivisimpia ehtojen asettajia uudelleenkäytön valvojan rooliin ja TAMK:n ja TTY:n vastaajat aineiston kerääjän rooliin liittyen. Aineistojensa jakamatta jättämistä TAMK:n ja TTY:n vastaajat perustelivat eniten aineistonhallinnan tuen saajan rooliin ja TAMK:n ja TaY:n vastaajat suojelijan rooliin liittyvillä syillä.

Tilastollisena testinä käytettiin pienten otosten analysointiin sopivaa Fisher-Freeman-Haltonin eksaktia testiä. Sen mukaan korkeakoulutausta toi tilastollisesti merkitseviä eroja vastaajien välille kaikkiin muihin rooleihin paitsi aineiston käsittelijään liittyneissä kysymyksissä. TAMK:n vastaajajoukossa suhtauduttiin muita myönteisemmin aineiston kerääjän rooliin liitettyihin tutkimusaineistojen avoimuuden ehtoihin, yksittäisistä ehdoista jaettuun tekijyyteen. Lisäksi vain heidän parissaan toivottiin aineiston keruun ja säilytyksen kustannusten kattamista aineistojen uudelleenkäytön ehdoksi, mikä liittyi aineistonhallinnan tuen saajan rooliin. TTY:n vastaajajoukko taas torjui muita voimakkaammin sekä aineiston suojelijan että uudelleenkäytön valvojan roolit. TaY:n vastaajien parissa suojelijan roolilla oli vahva sija, ja julkaisuoikeuksien puuttuminen oli heille muita keskeisempi syy rajata aineistojen avoimuutta. Tiedeyhteisön jäsenen roolia tarkasteltaessa vain TaY:n vastaajajoukossa puollettiin väittämää ”muut saavat helposti aineistojani käyttöönsä” ja vain TTY:n vastaajajoukossa pidettiin kyseiseen rooliin liitettyä syytä ”muut eivät tarvitse aineistojani” perusteena jättää aineistoja jakamatta.

Valtaosa vastaajista näki tutkimusaineistoille alkuperäistarkoituksesta erovia hyödyntämistapoja, mutta he olivat valmiimpia jakamaan aineistojaan kuin hyödyntämään itse valmiita aineistoja. Näkemyksiä ja käytäntöjä vertailtaessa taas havaittiin, että vaikka vastaajat korkeakoulutaustasta riippumatta käyttivät vähän metadataa ja olivat huolissaan tutkimusaineistojen tulkintaongelmista ja vain puolet heistä kielsi jyrkästi jättäneensä oman alansa aineistoja käyttämättä vaikeaselkoisuuden vuoksi, enemmistö oli tyytyväisiä aineistonsa hyödyntämisen prosesseihin. Ristiriidan yksi selitys on, että potentiaalisesti haastavia valmiita aineistoja ei harkitakaan käytettäväksi, vaan ne rajataan pois heti tutkimusprosessin alussa eivätkä ne päädy analysoitaviksi asti. Siten onnistuneen aineistonhallinnan yhteyttä tutkimuksen alkuvaiheisiin ei välttämättä aina tunnisteta: metatieto voisi tehdä useammista tutkimusaineistoista ymmärrettäviä.

Avainsanat: avoin tiede, Tampere3, tutkimusaineistot, aineistonhallintasuunnitelmat, kyselytutkimukset

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

# SISÄLLYSLUETTELO

1	JOHDANTO.....	1
2	TEOREETTINEN VIITEKEHYS.....	4
	2.1 Tutkimusaineisto ja sen avoimuus käsitteinä.....	4
	2.2 Tutkimusaineiston synty ja hallinta.....	9
3	KIRJALLISUUSKATSAUS.....	19
	3.1 Keskeiset tutkimukset.....	19
	3.2 Aineistonhallinnan periaatteiden tehokkuus.....	24
	3.3 Eettisten ja oikeudellisten kysymysten vaikutus avoimuuteen.....	27
	3.4 Aineistonhallinnan vastuiden jako.....	30
	3.5 Tutkijat aineistonhallinnan toimijoina.....	35
4	TUTKIMUSASETELMA JA -MENETELMÄ.....	44
	4.1 Tutkimuskysymykset.....	44
	4.2 Määrällinen kyselytutkimus.....	45
	4.3 Kyselyn toteutus.....	47
	4.4 Aineiston analysointi – riippumattomuudesta ja roolijako.....	53
5	TULOKSET.....	67
	5.1 Tavoitetut vastaajat.....	67
	5.2 Tutkimusaineistojen kerääjä – jaossa oleva kunnia.....	71
	5.3 Tutkimusaineistojen suojelija – noudatettavat normit.....	74
	5.4 Tiedeyhteisön jäsen – kohti yhteisiä tavoitteita?.....	78
	5.5 Tutkimusaineiston uudelleenkäyttäjä ja uudelleenkäytön valvoja.....	81
	5.6 Tutkimusaineiston käsittelijä – aineistojen käytettävyys.....	88
	5.7 Aineistonhallinnan tuen saaja – tiedeinstituutioiden vaikutus.....	93
6	YHTEENVETO JA POHDINTA.....	105
	6.1 Aineistojen jakamatta jättäminen ja avoimuuden ehdollisuus.....	105
	6.2 Tutkimusaineistojen (uudelleen)hyödyntämishalut ja -kyvyt.....	112
	6.3 Tutkimuksen jatkonäkymät.....	115
	LÄHTEET.....	116
	LIITTEET.....	I

# 1 JOHDANTO

Tutkielmassani selvitän, miten kyselytutkimukseeni keväällä 2018 osallistuneet tamperelaisissa korkeakouluissa työskennelleet tutkijat näkevät tutkimusaineistojensa hallinnan ja oman roolinsa siinä. Aloittaessani syksyllä 2017 tutkielmani suunnittelun Tampereen kolmea eri korkeakoulua – Tampereen ammattikorkeakoulua (TAMK), Tampereen teknillistä yliopistoa (TTY) ja Tampereen yliopistoa (TaY) – oltiin parhaillaan liittämässä yhteen poikkeuksellisen laajassa Tampere3-hankkeessa (esim. Tolvanen 2017, 2). Uuden korkeakoulu yhteisön oli määrä aloittaa toimintansa 1.1.2019 (Tampere3 2018). Hankkeen kotisivuilla kuvailtiin sen tavoitteita esimerkiksi seuraavasti:

”Uudessa korkeakoulu yhteisössä kohtaavat terveyden, tekniikan ja yhteiskunnan tutkimus, jossa merkityksellisiä ilmiöitä tarkastellaan inhimillisestä, taloudellisesta ja teknologisesta näkökulmasta. Yhdistämme kaikki tutkimuksen tasot perustutkimuksesta aina soveltavaan tutkimukseen ja palvelukonsepteihin. Tampere3 tarjoaa avoimen areenan yhteentörmäyksille: teemme tiedettä, joka ammentaa käytännöstä sekä synnyttää huipputuloksia ja huippusovelluksia.” (Tampere3 2018.)

Oli siis ajankohtaista tutkia seudun korkeakoulujen tutkimustyön eroja ja yhtäläisyyksiä – tässä tutkielmassa tutkimusaineiston hallinnan näkökulmasta. Kyse ei silti ole vain hetkeen tarttumisesta, vaan tutkielma tuo lisänsä aiempaan tutkimukseen. Tamperelaista korkeakoulu ympäristöä minua innosti tutkimaan Carol Tenopirin et al. kansainvälinen kyselytutkimus, jossa nostettiin esiin tarve tutkia lisää tiedeyhteisön erillisiä aineistonhallintakulttuureita. Yksi alan tutkimuksen visioita on, että aineistonjakamiskäytäntöjä kehittämällä saadaan jakamista suosiva kulttuuri vallitsemaan valtaosassa tieteenaloja. (Tenopir et al. 2015, 21.)

Korkeakoulujen tutkimusaineistojen arvoa ja niiden hallinnan tutkimisen tärkeyttä ei voine korostaa liikaa. Graham Pryor toteaa osuvasti, että kalliisti tuotettua aineistoa ei pitäisi kohdella kuin älyllisen kaivoksen jätettä. Vaikka aineisto on sitä, mitä tieteellisessä tutkimuksessa ensivaiheessa syntyy, vain pieni osa siitä tulee näkyväksi. Tieteellisessä julkaisussa laaja tutkimusprosessi esitetään seulo-tussa ja tiivistetyssä muodossa, kun taas suuri aineistomäärä, josta julkaisu on tuotettu, jää piiloon ja suurilta osin saavuttamattomiin. On epätodennäköistä, että julkaisussa saataisiin hyödynnettyä ja tuotua esiin kaikki tutkimusprosessin aikana kerätyn aineiston sisältämä arvo. Kun huomioidaan kaikki se älyllinen työ ja resurssit, joita tutkimusohjelman läpivieminen vaatii, seuraa huomattavaa taloudellista vahinkoa, jos aineisto jätetään hallinnoimatta. (Pryor 2012, 3-4.) On todettu, että tiedettä tarvitaan osaltaan todistamaan, että järjestetyllä ja jaetulla aineistolla on arvoa (Haendel et

al. 2012, 3). Tutkimusaineistojen hallinnan tutkimisella tuetaan osaltaan yhä monimutkaisempien tieteellisten kysymysten ratkaisemista ja eri tiedeyhteisöjen tarpeiden huomioimista, sillä tehokkaat, pitkäaikaiset ja kestävät aineistohallintakäytännöt ovat edellytys haasteista selviytymiselle. Aiemmat tutkimukset ovat antaneet viitteitä siitä, että tutkijat kokevat tarvitsevansa pääsyn muiden tuottamien aineistojen äärelle ja että aineistojen jakamisen puuttuminen voi olla suuri este tieteen kehittymiselle. Aineistohallintaa tutkimalla ja määrittelemällä rakennetaan parempaa ymmärrystä aineiston jakamiseen liittyvästä informaatiokäyttäytymisestä, jakamisen kuluista ja hyödyistä niin tieteelle kuin yhteiskunnallekin, sekä siitä, mikä tekee toimintatavoista tehokkaita. Tällöin osataan myös tarjota tarvittavaa koulutusta ja infrastruktuuria. (Piwowar 2011, 10; Tenopir et al. 2015, 21.)

Käsitykset kullekin tiedeyhteisön toimijalle kuuluvista tutkimusaineistojen hallinnan vastuista voivat olla horjuvia. Tutkijan rooliksi voidaan nähdä pikemminkin aineiston tuottaja tai käyttäjä kuin aineiston hallitsija tai huolehtija. Aineistonkeruuprosessi voi olla vaikea, kallis ja vuosia vievä. Se voi vaatia erilaisten verkostojen luomista ja hyödyntämistä ja innovatiivisilla tutkimusaloilla myös uusien aineistontuottamistekniikojen kehittämistä. Kyseiset toimet vievät tutkijan ajan. Päähuomio ei siis ole ollut aineiston pitkäkestoisessa hallinnassa tai ylläpidossa, varsinkin kun urapalkinnot eivät ole liittyneet aineiston säilytykseen ja hallintaan, vaan tulosten laadukkuuteen. Vaikka aineiston jakamisesta syntyvä arvo saatetaan tiedostaa, aineiston valmistelu jatkokäyttöä varten nähdään yleisesti taakkana, joka vaatii tietoja, joiden ei ole katsottu kuuluvan tutkijan taitoihin, sillä kyse ei ole tutkimuksen teosta. (Pryor 2012, 7.) Kun tutkimuksen elinkaaren aikana syntyvää aineistoa kohdeltiin pitkään vain julkaisemisprosessin sivutuotteena, aineistojen ajateltiin olevan vain tutkijoita ja heidän tutkimuskumppaneitaan varten ja hallittavissa henkilökohtaisia metodologioita käyttäen. Nykyisin aineistoja on alettu arvostaa itsenäisinä tuotteina, sillä ne mahdollistavat kokeiden uusintamisen ja luovat pohjan uudelle tutkimukselle. (Higgins 2012, 18.) Aineistohallintaa käsittelevässä tutkimuksessa on havaittu ristiriita siinä, mitä tutkijat näkevät, että aineistolle pitäisi tehdä ja mitä oikeasti tehdään, ja se on nähty tärkeäksi tutkimuskohteeksi. (Tenopir et al. 2015, 3.)

Tutkijakeskeistä näkökulmaa aineistojen hallintaan on siis syytä painottaa (ks. myös Tenopir et al. 2015, 4), ja tässä tutkielmassa suunnataankin huomio siihen, miten Tampereen korkeakoulujen tutkijat omaksuvat tutkimusaineistojen hallinnan eri rooleja. Tutkijoiden aineistohallinnan näkemykset ja käytännöt yhdistetään tässä tutkimusaineiston kerääjän, suojelijan, uudelleenkäyttäjän ja uudelleenkäytön valvojan sekä aineistojen käsitteijän sekä aineistohallinnan tuen saajan rooleihin.

Toisaalta roolien hahmottamisella ja tutkijoiden asettamisella niihin korostetaan tutkijoiden toimijuutta ja vastuita aineistohallinnassa, toisaalta edistetään aineistohallinnan kehittämistoimien kohdentamista. Roolit heijastavat tutkimusaineiston elinkaaren vaiheita ja liittyvät siihen, mitä tutkijat pitävät tehtävänä ja miten ja miksi he haluavat ohjata aineistojen hallintaa. Tästä tutkimuksesta on mahdollisesti hyötyä tamperelaisten korkeakoulujen tutkimustyön synergiaa ajatellen: miten arvokkaat tutkimusaineistot saadaan mahdollisimman hyvin hyödynnettyä.

## 2 TEOREETTINEN VIITEKEHYS

Keskityn aineistohallintaan akateemisessa ympäristössä, mikä vaatii ensiksi aineistohallinnan eri osa-alueiden käsitteellistä purkamista alkaen siitä, mitä hallittava tutkimusaineisto on ja mitä periaatteita siihen liitetään. Alaluvussa 2.1 huomio on tutkimusaineiston määritelmässä ja aineiston saatavuudessa. Alaluvussa 2.2 käyn läpi tutkimusaineiston syntyä ja hallintaa, jolloin olennaisia käsitteitä ovat tutkimuksen ja tutkimusaineiston elinkaaret, aineistohallintasuunnitelma sekä aineiston jakaminen. Esittelen myös jakamista jäsentävän kuusiosaisen viitekehyksen, jota jatkojalostan luvussa 4 (alaluku 4.4) tämän tutkielman asetelmaan sopivaksi.

### 2.1 Tutkimusaineisto ja sen avoimuus käsitteinä

Perustavanlaatuinen käsite tutkielmassani on *aineisto*. Etenkin englanninkielisessä tutkimuskirjallisuudessa käytetään sanaa ”data”, mutta vaikka se on suomeenkin vakiintunut käsite, käytän tutkielmassani ilmaisua ”aineisto”. Tutkimuskirjallisuudessa esiintyvä käsite ”data management plan” (DMP) kääntyy tässä tutkielmassa ”aineistohallintasuunnitelmaksi” ja vastaavasti pelkkä ”data” kääntyy ”aineistoksi”. Ratkaisua puoltaa ensiksikin, että Suomen Akatemiakin vaatii rahoituksen hakijoilta aineistohallintasuunnitelmaa – ei siis tiedonhallintasuunnitelmaa tai datanhallintasuunnitelmaa. Toiseksi menettelyllä vältetään ristiriidoilta, sillä tieteenalat suhtautuvat eri tavoin käsitteeseen ”data”, ja käyttämässäni kyselylomakkeessa tiedustelen eri alojen tutkijoilta, millaisen aineiston kanssa he työskentelevät. Data on kaiken informaation rakennuspalikka. Se on tulosta mistä tahansa systemaattisesta tutkimuksesta, onpa se sitten havainnoiteja, kokeiluja tai hypoteesien testausta. Vasta tulokinnan ja kontekstoinnin myötä syntyy uutta tietämystä. Siten niin humanistit, luonnontieteilijät kuin yhteiskuntatieteilijätkin tuottavat dataa. Kuitenkin humanististen alojen tutkijat saattavat painottaa olevansa tekemisissä informaation tai tietämyksen kanssa, siinä missä data nähdään luonnontieteisiin kuuluvana asiana. (Pryor 2012, 2-3.) Toisaalta tutkimusaineistoksi kutsutaan myös esimerkiksi empiiristä tutkimusta varten hankittuja käsittelemättömiä tietoja – synonyymi sille on havaintoaineisto (Heikkilä 2014, 13). Tässä tutkielmassa tutkijat tuottavat yhdenmukaisuuden vuoksi datan sijasta (tutkimus)aineistoa, mutta data-käsitteellä on sijansa, jos kyse on vakiintuneesta käsitteestä, kuten avoin data, tai mahdollisten käänkövirheiden välttämisestä.

Tutkimustyössä voi nykyisin nähdä syntyvän tarkemmin sanottuna kahdenlaisia digitaalisia objekteja, joista aineisto muodostuu. Ensiksikin on yksinkertaisia, erillisiä digitaalisia objekteja, kuten yksittäisiä teksti-, kuva- tai äänitiedostoja. Ne ovat itsenäisinä yksinkertaisimpia hallittavia: ne

käsittävät vain yhden tiedoston, joka on yhdessä formaatissa, ja niiden käsittely vaatii vain yhden ohjelmiston. Toiseksi on monimutkaisia digitaalisia objekteja<sup>1</sup>, jotka syntyvät monia digitaalisia objekteja yhdistämällä. Niiden käsittely on haasteellisempaa, sillä se saattaa vaatia ohjelmistopakettien yhdistelmän, ja lisäksi osat pitää liittää käsitteellisesti yhteen, jotta niiden keskinäiset suhteet ovat ymmärrettävissä. (Higgins 2012, 19.)

Tutkimusaineistojen asema ja määritelmä tieteellisessä keskustelussa on muuttunut, minkä taustalla on niin avoimen tieteen – ammattikorkeakoulukontekstissa avoimen tutkimus-, kehittämis- ja innovaatiotoiminnan – periaatteet kuin taloudellisuuskin. Tieteellisten menetelmien on oltava järjestelmällisiä, todennettavissa olevia ja tutkimuksen tuloksia varmentavia. Toisaalta tutkimusalan järjestelmällinen tuntemus on tieteellisen asiantuntijuuden kulmakiviä. Toisaalta tieteen tekemisen avaaminen ja mahdollisimman suuren läpinäkyvyyden turvaaminen rakentavat luottamusta tieteesseen ja tutkimukseen, ja arvostus ja luottamus lisäävät halua hyödyntää tieteellistä tietoa yhteiskunnassa. (TJNK 2018, 5, 8; Päällysaho & Latvanen 2019, 25.) Tutkimuksen laatu paranee osaltaan aineistojen jakamisen kautta, sillä siten muille tutkijoille annetaan mahdollisuus toistaa tutkimus sekä vahvistaa ja laajentaa sitä, sen sijaan, että jaetaan passiivisesti vain tutkimuksen tuloksia. Jos tutkimusaineistoja ei jaeta, miten esimerkiksi uuden metodin esittelevä artikkeli voidaan arvioida riittävän hyvin, kuten tarkistaa, saisivatko muut metodilla saman tuloksen? Aineistojen jakamisen myötä tieteellisestä artikkelista tulee ”omavarainen”: kyse ei ole vain tekstistä, vaan kokonaisuudesta, joka sisältää myös tutkimusaineiston ja esimerkiksi tietokonekoodin (computer code), joiden avulla tulokset on saatu. Koodien jakaminen vähentää uudelleenikäytön teknisiä esteitä tehden aineistoihin tutustumisesta helpompaa. (Poisot et al. 2013, 12, 14.)

Aineistojen jakamisen voi nähdä lisäävän tiedeyhteisön tasa-arvoisuutta. Ensiksikin hyvät kontaktit ovat taanneet osalle tutkijoista paremman pääsyn aineistojen äärelle, kun taas osa on jäänyt

---

<sup>1</sup> Sarah Higgins antaa kolme esimerkkiä monimutkaisista objekteista. Ensiksikin vaikkapa digitoitu kirja koostuu digitoituista kuvista: kustakin kirjan sivusta on otettu valokuva. Objekti on vielä varsin yksinkertainen hallittava, sillä kukin tiedosto on samassa formaatissa. Toista lajia edustaa vaikkapa tekstitiedosto, jossa on upotettu kuva. Objekti on tehty yhdistämällä kaksi formaattia. Kolmantena esimerkkinä on verkkosivu, joka on yhdistelmä eri tiedostoista, joista kukin on eri formaatissa. Sivun koostuu esimerkiksi lähdekoodista (Hypertext markup language, HTML), upotetuista kuvista (Joint Photographic Experts Group, JPEG), linkitetyistä dokumenteista (Portable Document Format, PDF) ja sivun ulkoasun luomisesta (Cascading Style-sheet File, CSS). (Higgins 2012, 19-20.)



oikeiden kontaktien puuttuessa siitä paitsi – aineistojen julkaiseminen vapaan pääsyn säilytyspaikoissa häivyttää nämä rajanvedot. (Poisot et al. 2013, 11-12.) Toiseksi, tutkijan tallennettua itse omia aineistojaan julkisiin arkistoihin hänen työllään on potentiaalia tulla jaetuksi aina kansainvälisellä tasolla asti. Aiempaa useampi nimekäs alan tutkija voi viitata häneen ja hänet saatetaan listata muihinkin julkaisuihin yhdeksi tutkimuksen tekijäksi. Samalla hänen tutkimuksensa voi hyödyntää globaaliakin tiedeyhteisöä. (Tenopir et al. 2015, 4; ks. myös Poisot et al. 2013, 13-14.) Tosin jakamiseen liittyy myös harmaita alueita uudelleenkäytön näkökulmasta. Esimerkiksi jos tutkija ottaa jonkin jaetun aineistokokonaisuuden työnsä pohjaksi, lisää siihen omaa aineistoaan ja julkaisee laajemman aineistokokonaisuuden, alkuperäinen tekijä saa häneltä viittauksen, yhden ainoan, mutta kaikki tulevat viittaukset tulevat uuden tutkijan tuoreelle julkaisulle ja laajemmalle aineistokokonaisuudelle, kuten tietokannalle. Kunkin aineistokokonaisuuden voi nähdä vanhentuvan, jäävän pois käytöstä, heti, kun sen pohjalta rakennettu uusi kokonaisuus julkaistaan. (Moles et al. 2013, 20-21.)

Kolmanneksi aineistojen jakaminen ja aineistojen tunnustaminen tieteellisen tuotannon osaksi voi tuoda tutkijoille viittausten lisäksi enemmän julkaisemismahdollisuuksia. Pitkään todennäköisyyteen saada artikkeli julkaistuksi lehdessä on vaikuttanut sen kuvaamien tieteellisten tulosten merkittävyys – jos tulokset ovat olleet ”negatiivisia”, artikkeli jää helpommin julkaisematta. Aineistojen julkistaminen edistää vähintäänkin tulevien tutkimushankkeiden suunnittelua, sillä sen myötä ollaan entistä tietoisempia niin ”positiivisia” kuin ”negatiivisiakin” tuloksia tuottaneista aineistoista. (Poisot et al. 2013, 13.) Suomessa tiedeviestinnän suosituksissa linjataan, että tieteen julkisuuden ei pidä pohjautua yksinomaan uusien tutkimustulosten uutuuksiarvoon, vaan myös hitaasti edistyvällä perustutkimuksella sekä luonteeltaan uusia tuloksia varmentavalla tutkimuksella on arvokas sijansa tutkimuksessa ja innovaatiokehityksessä. (TJNK 2018, 5-6.)

Kansainvälisellä tasolla käsitteiden hiomisen ja merkityksen tärkeydestä kertoo esimerkiksi yhdysvaltalaisen National Science Foundationin (NSF) 2010-luvun alun ohjeistuksessa ”Proposal Award Policies and Procedures Guide” esitetty linjaus käsitteen ”julkaisu” (publication) korvaamisesta käsitteellä ”tuote” (product). Tarkoituksena oli tehdä selväksi tutkijoille, että tutkimusrahoittajat pitävät tutkimustuotteita arvokkaina, eivätkä vain julkaisuja. Siten jaetut aineistokokonaisuudet olisivat jatkossa ensisijaisten tutkimuskohteiden asemassa. Tuolloin uskottiin yhä useamman rahoittajan omaksuvan kyseisen lähestymistavan. (NSF 2013; Poisot et al. 2013, 14.) Rahoittajien heräämisen taustalla on näkemys siitä, että primääristen tutkimusaineistojen jakaminen ja uudelleenkäyttö muodostavat tieteen kentällä voimavaran. Kuluja kertyy aineistojen keräämisestä laboratorioissa tai

kentällä, laitteiden ylläpidosta ja henkilöstön palkkauksesta. Uuden aineiston tuottaminen, kun vastaavanlaista on jo saatavilla, on turhaa työtä. Resurssien käyttö tehostuu, kun niitä ei ohjata samanaisten aineistokokoelmien tuottamiseen. Tosin tehokkuustavoitteiden toteutuminen edellyttää helppoja tapoja löytää jaetuista aineistoista oikeanlainen tieto eli esimerkiksi huolellisten hakemistojen laatimista. (Piwowar 2011, 1; Pryor 2012, 4; Poisot et al. 2013, 13.)

Suomessa tutkimusaineistojen lisääntynyt arvostaminen itsenäisinä tuotteina näkyy esimerkiksi ”Avoimen tieteen ja tutkimuksen julistuksen 2020–2025” julkaisemisessa. Siinä aineistot tunnustetaan itsenäisiksi tutkimustuotoksiksi, ja tutkimusaineistot ja -menetelmät muodostavat yhdessä yhden neljästä tieteen avoimuuden osa-alueesta. Tutkimusaineistojen pitäisi olla ”niin avoimia kuin mahdollista ja niin suljettuja kuin välttämätöntä”. (Avoimen tieteen koordinaatio 2020.) Aineistoihin liittyvä olennainen käsite onkin *Open Access (OA)* eli *avoin saatavuus*, joka kytkeytyy myös taloudelliseen näkökulmaan. Kun puhutaan avoimesta saatavuudesta, aineiston jakamista on puollettu, kuten sanottua, yleisenä hyvänä sekä tapana lisätä läpinäkyvyyttä ja vahvistaa aiempien tulosten paikansa pitävyyttä kuluja sekä säästötoimena, mutta myös tutkijan eettisenä velvollisuutena – etenkin tutkimuksen ollessa julkisrahoitteinen. OA-toiminnan ensimmäisiä kimmokkeita olikin ajatus siitä, että julkisen rahoituksen turvin tehdyt tieteelliset julkaisut on tehtävä julkisiksi. Julkisrahoitusta saavat tutkijat ovat useimmissa maissa valtion virkamiehiä. Aineistojen tuottaminen voidaan nähdä osana heidän työnkuvaansa, eikä ole ollut pätevää perustetta, miksi he voisivat vaatia siihen omistajuutta. Tekijyyden vaatiminen aineistoihin on oikeutetumpaa, mutta sekään ei estä aineiston jakamista. (Poisot et al. 2013, 12; Tenopir et al. 2015, 4; TJNK 2018, 6.)

Myös Sami Borg ja Arja Kuula käsittelevät avointa saatavuutta todeten sen tarkoittavan sähköisten tutkimusaineistojen kohdalla sitä, että merkittävien julkisrahoitteisten aineistojen olisi lähtökohtaisesti oltava avoimessa käytössä ja että esteetön ja tasavertainen aineistojen käyttömahdollisuus taataan mahdollisimman pienin kokonaiskustannuksin mahdollisimman suurelle käyttäjäkunnalle. Tutkimusaineistojen Open Access ei silti aina tarkoita niiden täysin avointa ja maksutonta (jatko)käyttöä, vaan aineistoihin voidaan liittää käyttörajoituksia ja avoimenkin aineistojen käyttö voi edellyttää käyttölupia, käyttöehtositoumuksia ja käyttäjätunnistusta. Kuitenkin aineistojen rajaamiselle avoimuuden ulkopuolelle tulisi aina olla perusteltu syy. (Borg & Kuula 2007, 9-10, 15-16.)

Käsite *avoin data* (open data) on puolestaan liitettävissä aineiston laajimpaan merkitykseen. Se kuvaa aineistoa, joka on avoimesti saatavilla, saavutettavaa, ymmärrettävää ja vapaasti uudel-

leenkäytettävissä – joskin se liitetään usein hallinnollisen ja akateemisen sektorin aineistoihin. (Wessels et al. 2017, 45-46, 159.) Esimerkiksi Suomessa vuonna 2020 Digi- ja väestötietovirasto (DVV) vastaa Avoindata.fi-palvelusta eli kansallisesta avoimen datan portaalista, jossa kaikki Suomen avoin data on saatavilla samasta paikasta. Hakemisto kattaa ensisijaisesti julkisen hallinnon avoimen datan, ja sen kategorioita ovat muun muassa energia, liikenne, kulttuuri, taide ja vapaa-aika, talous ja rahoitus, tiede ja teknologia sekä rakennettu ympäristö ja infrastruktuuri. Palvelu on tarkoitettu kaikille kansalaisista yrityksiin ja viranomaisiin sekä avoimen datan julkaisemiseen että sen hyödyntämiseen (DVV 2020a; DVV 2020c.) Palvelussa avoin data määritellään digitaalisessa muodossa olevaksi informaatioksi, joka on kaikkien vapaasti käytettävissä mihin tahansa käyttöön maksutta, kunhan sen alkuperäinen lähde mainitaan. Data määritellään digitaalisesti tallennetuksi, merkeistä ja symboleista koostuvaksi koneellisesti luettavissa olevaksi informaatioksi, joka voi muodostaa niin dokumentteja, tietokantoja, kuulemisten transkripteja kuin audiotallenteitakin ja jota jalostamalla voidaan saada aikaan merkityksellistä informaatiota. Avoin merkitsee, että datalla on avoimen hyödyntämisen mahdollistava käyttöluva eli lisenssi. (DVV 2020b.)

Toisaalta puhutaan myös *avoimesta tietämyksestä* (open knowledge), jonka määrittelyyn on osallistunut esimerkiksi Open Knowledge Foundation (OKF). Avoimen tietämyksen käsitteessä avoimeen dataan sisällytetään muukin kuin tieteellisessä tutkimuksessa syntynyt aineisto, eli esimerkiksi kulttuuriaineistot, kuten kirjastojen, arkistojen ja museoiden kokoelmat, sekä liikenteen, ympäristön ja talouden seurannassa kerättävä ja kertyvä data. Kyse on minkä tahansa sisällön, informaation tai datan vapaasta käytöstä, uudelleenkäytöstä ja uudelleenjakelusta niin, että sille ei aseteta mitään oikeudellisia, teknologisia tai sosiaalisia rajoituksia. Avoimen tietämyksen ominaispiirteisiin kuuluu ensiksikin, että aineiston on oltava saatavissa kokonaisuudessaan ja koneluettavassa, muokattavissa ja mieluiten internetistä ilmaiseksi ladattavissa olevassa muodossa. Toiseksi aineiston käyttöehtojen pitää sallia sen uudelleenkäyttö ja -jakelu sekä aineistojen yhdistely, vaikkakin voidaan edellyttää, että aineistoon tehtyjen muutosten pitää olla jäljitettävissä. Kolmas piirre on universaalin osallistumisen huomioiminen: aineiston on oltava samalla tavalla kaikkien hyödynnettävissä, eli mitään ryhmää tai ketään henkilöä ei saa suosia eikä syrjiä, ja aineistoa saa käyttää mihin tarkoitukseen ja millä alalla tahansa. (Wessels et al. 2017, 46-47; Open Knowledge Foundation 2020.) Seuraavaksi pureudutaan kuitenkin tarkemmin nimenomaan tieteellisen aineiston syntyyn.

## 2.2 Tutkimusaineiston synty ja hallinta

Graham Pryor kuvailee *tutkimuksen elinkaaren* olevan käytännössä monikerroksinen kokonaisuus, johon kuuluu erilaisia työkaluja, metodologioita ja toistuvia vaiheita, kuten hypoteesien laatimista, aineistojen tulkintaa, yhdistelyä ja uudelleenkäyttämistä sekä tulosten ja aineistojen julkaisemista. Kaikki alkaa yleensä ideasta tai hypoteesista ja päättyy johonkin tieteelliseen tuotokseen, kuten julkaisuun. Olennaista on, että jokaisessa elinkaaren vaiheessa tutkija kerää, käyttää ja/tai luo aineistoa. Usein alun suunnittelevassa vaiheessa kerätään ja arvioidaan valmista julkistettua aineistoa, josta osa saattaa valikoitua uuden tutkimuksen raaka-aineistoksi. Myöhemmin tutkimusprosessi itsessään johdattaa etsimään lisääaineistoa esimerkiksi vertailuja varten. Toisaalta tietyillä tutkimusaloilla, kuten bioinformatiikassa ja matematiikassa, aiemmat laajat aineistokokonaisuudet ovat perustavanlaatuisen edellytys uusille työprosesseille. Toisaalta hyvä esimerkki uusien kontekstien löytämisestä on sääaiheiset asiakirjat, esimerkiksi laivojen lokikirjat, joita käytetään nykyään ilmastonmuutoksen tutkimiseen, vaikka niitä ei alkujaan laadittu siihen tarkoitukseen. (Pryor 2012, 5-7.)

Sarah Higgins hahmottelee tutkimusentekoprosessissa olevan kolme aineiston luomiseen liittyvää vaihetta. Tutkimusvaiheessa luodaan tai kerätään raaka-aineistoa, jonka varaan tutkimus perustuu. Yhdistämisvaiheessa käytetään esimerkiksi mallintamis- ja tilastoimistekniikoita, jotta aineisto tehdään ymmärrettäväksi. Uudelleenkäyttövaiheessa hyödynnetään kahden edellisen vaiheen aineistoja uuden aineistokokonaisuuden luomiseen, ja kokonaisuutta voidaan käyttää uusien hypoteesien testaamiseen. Läpi tutkimuksen elinkaaren luodaan myös ydintutkimuksen tekoa tukevia dokumentteja, kuten tutkimusta selittäviä dokumentteja, liittyen esimerkiksi työvirtaan, hallintoon, pöytäkirjoihin, ryhmäwikieihin ja tiedotukseen. Aineiston ymmärrettävyyden takaaminen ja uudelleenkäytön edistäminen edellyttää kaiken edellä kuvatun aineiston hallinnointia läpi tutkimuksen elinkaaren. (Higgins 2012, 18-19.)

Yhdistyneessä kuningaskunnassa aineistojen pitkäaikaissäilytykseen erikoistunut Digital Curation Centre (DCC) taas on jäsennellyt *aineiston kuraation elinkaaren* (a data curation lifecycle). Ydinajatuksena on, että toimet kohdistuvat aineistoon itseensä sen omien tarpeiden takia, ei vain tutkimusprosessin mahdollistajana. Aineiston kuraation elinkaaren voi nähdä kulkevan tutkimuksen elinkaaren ohessa ja sen alun olevan heti tutkimuksen käsitteellistämisen vaiheessa. Elinkaariajattelulla pyritään varmistamaan, että kaikki aineiston käsittelyn vaiheet – käsitteellistämisestä ja luomisesta arvonmäärittämiseen ja seulontaan, säilyttämiseen ja ylläpitoon, kuvailuun, esilletuontiin ja

uudelleenkäyttöön – on suunniteltu ja että ne toteutetaan tietyssä järjestyksessä ja tarpeen mukaan toistettaen. (Pryor 2012, 7-9.)

Myös Sami Borg ja Arja Kuula kirjoittavat *aineistojen elinkaaresta*. He liittävät elinkaaren suunnittelun aina ensivaiheesta jatkokäyttöön keskeiseksi osaksi huolellista ja tarkkaa tulosten tallentamista. Heidän hahmotelmassaan julkisrahoitteisen tutkimusaineiston elinkaaresta on kymmenen kohtaa:

1. tutkimussuunnitelma
2. eettinen arviointi ja luvat
3. rahoitushakemus
4. rahoittaja/rahoituspäätös
5. aineiston keruun valmistelu / keruu
6. ensikäyttö
7. aineiston arkistointi ja julkaiseminen
8. jatkokäytön tuki
9. jatkokäyttö
10. julkaisut ja datat (tietokannat).

Tosin aina ei edetä edellä kuvatussa järjestyksessä, vaan esimerkiksi vaiheesta 6 voidaan siirtyä suoraan vaiheeseen 10, ja vaihe 1 kytkeytyy vaiheisiin 3 ja 5. (Borg & Kuula 2007, 12, 26.)

Tutkimusaineistojen hallinnasta puhuttaessa esimerkiksi ”Avoimen tieteen ja tutkimuksen julistus 2020–2025” visioi, että aineistoja hallittaisiin lähitulevaisuudessa tarkoituksenmukaisesti FAIR-periaatteita tavoitellen. (Avoimen tieteen koordinaatio 2020.) Periaatteiden mukaan aineistosta on tarkoitus tehdä löydettävää (findable), saavutettavaa (accessible), yhteentoimivaa (interoperable) ja uudelleenkäytettävää (re-usable). Periaatteet ovat esimerkiksi osa Euroopan unionin vuoden 2020 datastrategiaa, mutta ne ovat olleet keskusteluissa jo pidempään. (Euroopan unionin neuvosto 2016; Euroopan komissio 2020.)

Aineistonhallinnassa mukaan lukien aineiston säilymisessä on olennaista, että keskeiset toimet tehdään oikeaan aikaan aineiston elinkaaren aikana, heti aineiston synnystä alkaen. Pienikin laiminlyönti voi vähentää merkittävästi aineistosta huolehtimisen edellytyksiä, mikä tarkoittaa, että esimerkiksi sähköisistä tiedostoista ei voidakaan enää tehdä käyttökelpoisia ja ymmärrettäviä. Etenkin digitaalisessa ympäristössä aineisto on helposti siirrettävissä, poistettavissa, kopioitavissa ja muokattavissa, mikä tekee aineiston eheydestä, luotettavuudesta ja provenienssista kyseenalaista, mikä

puolestaan vaikuttaa osaltaan uudelleenkäyttömahdollisuuksiinkin. Esimerkiksi metadatan ja dokumentoinnin puuttuminen voi tehdä aineistosta tunnistamatonta ja käyttökelvotonta. (Higgins 2012, 17-18; Volk et al. 2014, 886.)

Ethan P. White et al. jakavat aineistohallintatoimet kolmeen keinoon ja tavoitteeseen. Ensiksikin hyvin dokumentoitua aineistoa on helpompi ymmärtää, toiseksi kunnolla muotoiltua (formatted) aineistoa on helpompi käyttää erilaisillakin ohjelmistoilla ja kolmanneksi vakiintuneissa säilytyspaikoissa ja avoimella lisenssillä oleva tieto on helpommin löydettävissä ja käytettävissä. Aineistohallinnassa voi olla kyse – tai sen voi aloittaa – pienistäkin teoista. Kun heti tutkimusprojektin alussa käytetään muutama tunti sen miettimiseen, miten aineistoa aiotaan dokumentoida, jäsenellä ja jakaa, voidaan ehkäistä tarvetta järjestellä myöhemmässä vaiheessa aineistoa tai kerätä sitä uudelleen käsiin. (White et al. 2013, 8.) Judith Davidson on jäsentänyt laadullisen tutkimuksen tekemisen prosessia ja tutkimusryhmien vuorovaikutusta, ja hän sivuaa kirjassaan ”Qualitative research and complex teams” aineistohallintaa esimerkiksi pohtimalla kokonaisten tutkimusprojektien arkistoinnista ja saavutettavuutta. Hän näkee tutkimuksen tekemisen sekä aineistohallinnan, mukaan lukien aineiston tallentamisen, toisiinsa tiiviisti liittyvinä alueina. Hänen mukaansa muutos kohti toimintamallia, jossa kaikki tutkimusprojektia koskevat materiaalit ja prosessit tallennetaan ja tuodaan saataville julkisiin säilytyspaikkoihin, on tehnyt keskeiseksi sen, että tutkijat ajattelevat projektin suunnittelun alusta asti datajärjestelmiään (data system) sekä tietokantoina että arkistoina. Siinä missä he itse käyttävät rakentamaansa tietokantaa siinä työssä, jota he parhaillaan tekevät, he itse tai muut voivat käyttää tietokantaa tulevaisuudessa uusissa projekteissa. (Davidson 2019, 61-62.)

Vaikka pienillä teoilla ja ajattelutapojen muutoksilla voidaan päästä alkuun, tyypillisesti aineistohallinta vaatii toimiakseen dokumentoituja menettelytapoja ja strategioita. (Higgins 2012, 20.) Tutkimuksenteon mahdollistavalla rahoituksen alueella merkittävänä kehityssuuntana on ollut sekä aineistoihin liittyvien vaatimusten tiukentaminen että rahoittajien odotusten tarkempi huomioon ottaminen. Suuntaus näkyy siinä, että tutkijoita on alettu vaatia liittämään rahoitushakemuksiinsa *aineistohallintasuunnitelma*. Julkinen ja johdonmukainen aineistohallinta ja -jakamissuunnitelma on väline, jolla vähennetään todennäköisyyttä rahoittaa samanlaista tutkimusta; aineiston sijoittaminen arkistoihin julkisesti haettaviksi edistää taas sen uudelleenkäyttöä. Tämä kaikki tukee samalla (julkisten) rahoittajien toiminnan läpinäkyvyyttä ja taloudellisuuspyrkimyksiä. (Donnelly 2012, 85.) Rahoittajien lisäksi osa tieteellisistä aikakauslehdistä vaatii kirjoittajia jakamaan tutkimusaineis-

tonsa toisten tutkijoiden kanssa (Tenopir et al. 2011, 2), mutta se ei ole sama asia kuin aineistonhallintasuunnitelmavaatimus. 2010-luvulle tultaessa oli viitteitä murroksesta koskien tutkimusaineistojen hallinnan suunnitelmallisuutta. Toisaalta valmiina olevan tutkimusaineiston uudelleenkäyttöä arveltiin tuolloin voitavan tukea enemmänkin, sillä rahoittajat olivat olleet taipuvaisia tukemaan uuden aineiston luomista. Toisaalta osan rahoittajista havaittiin jo vaativan tutkijoita tarkistamaan, oliko jo olemassa olevaa aineistoa, jota voitaisiin hyödyntää uuden sijasta, tai tekemään selvityksiä siitä, miten uusi aineisto täydentäisi vanhoja aineistokokonaisuuksia. (Jones 2012, 62-63.)

Pitkään oli ollut harvinaista, että rahoittajat tai julkaisijat olisivat vaatineet tutkijoita jakamaan tutkimusaineistojaan, mutta 2000-luvulla aineistojen jakamista koskevissa toimintapolitiikoissa ja -suosituksissa on vaadittu alaluvussa 2.1 esiteltyä julkisella rahoituksella tuotetun aineiston avointa saatavuutta, mikä on edellyttänyt suunnitelmien laatimista. Esimerkiksi Yhdysvalloissa National Institute of Health (NIH) alkoi vaatia vuodesta 2003 lähtien suurilta tutkimushankkeilta aineistonjakamissuunnitelmaa ja jo aiemmin mainittu NFS alkoi vuodesta 2011 vaatia aineistonjakamissuunnitelmaa puolestaan kaikilta hakijoiltaan. Kyse on kuitenkin myös laajemmasta kuin vain Yhdysvalloissa tapahtuvasta kehityksestä. Esimerkiksi Euroopan unionin alueella ohjeita on antanut Euroopan komissio ja Australiassa Australian Research Council. (Piwowar 2011, 1; Tenopir et al. 2015, 4.) 2010-luvun alussa Yhdistyneessä kuningaskunnassa useimmat tieteenalakohtaiset tutkimusrahoittajat odottivat avustuksen hakijoilta tietoja aineistojen saatavuuden, hallinnan ja pitkänajan säilytyksen suunnittelusta. Suurin osa rahoittajista vaati rahoituksenhakuvaiheessa hakijoilta aineistonhallinta- ja -jakamissuunnitelman, mutta vaatimukset suunnitelmien kattavuudesta vaihtelivat. Vasta muutama toimija pyysi tarkkaa tietoa aineistonhallinnan järjestelystä projektin aikana, kuten varmuusstrategioita sekä tietoja vastuiden ja taitojen jakamisesta tutkimusryhmän jäsenten kesken. Rahoittajat antoivat tyypillisesti jonkinlaista ohjausta suunnitelmien laatimiseen, esimerkiksi koottuja vastauksia kysymyksiin. Kuitenkin tutkimusaineiston jakamiseen ja uudelleen käyttämiseen siirtymistä tukemaan tarvitaan merkittävä kulttuurinen muutos. (Pryor 2012, 5; Jones 2012, 49, 54, 63.)

Mitä tulee aineistonhallintasuunnitelman sisältöön, sen pitäisi käsittää kaikki olennaiset aineistoa koskevat näkökulmat tutkimusprojektin koko elinkaaren ajalta, aina käsitteellistämisen vaiheesta aineiston pitkäaikaiseen säilyttämiseen tai hävittämiseen asti. Suunnitelma voi olla hyvin yleisluontoinen, kuten keskittyä kattamaan projektin päätavoitteet ja asettumisen tieteenalan kehitykseen, tai se voi sisältää yksityiskohtia alkaen siitä, millaisella kameralla aineistoon kuuluvat kuvat on otettu. Kuitenkin olennaista on, että suunnitelma on harkittu ja tiettyyn tarkoitukseen räätälöity, eikä vain

muokattu aiemmista ”menestyneiksi” katsotuista suunnitelmista. Se, että omaksutaan menestyksekkäitä elementtejä toisen projektin suunnitelmasta päämääränä vain rahoittajan vaatimuksiin vastaaminen, tuo mukanaan hyvin rajallisesti hyötyjä ja voi olla lopulta pikemminkin haitallinen lähestymistapa. Sitä vastoin tasapainon etsiminen ja hyvien käytäntöjen punnitseminen on arvokasta. (Donnelly 2012, 87.)

Digital Curation Centre on hahmotellut kolme aineistohallintasuunnitelman tasoa tai versiota. Minimaalinen aineistohallintasuunnitelma (minimal data management plan) laaditaan esimerkiksi rahoitushakemusta tehtäessä, ja siinä keskitytään täyttämään vain rahoituksen saamisen edellyttämät minimivaatimukset. Yksityiskohtaisen suunnitelman laatimiseen ei ole syytä panostaa, jos projektin toteutumisestakaan ei ole vielä takeita. Suunnitelma on tällöin luonteeltaan lupaileva: ”Me teemme näin...” Ydinsuunnitelma (core plan) laaditaan, kun rahoitus on saatu, ja se sisältää enemmän yksityiskohtia. Siinä voidaan jo huomioida kaikki keskeiset aineistohallintakysymykset mukaan lukien aineiston pitkäaikaishallinta ja säilytys, ja se on sävyltään astetta selittävämpi: ”Teemme näin näistä syistä...” Täydellinen suunnitelma (full plan) huomioi vielä pidemmän ajan aineistohallinnan kysymykset. Se liittyy projektin loppupuoleen, ja toimijat, jotka eivät välttämättä ole olleet mukana aineiston keruussa tai luomisessa, saattavat säilyttää ja ylläpitää suunnitelmaa vielä pitkän ajan kuluttua tutkimusprojektin päättymisestä. Aineiston säilyttämisen puolesta puhuu se, että sillä on arvoa muillekin kuin alkuperäisille kerääjille. Kukin suunnitelma rakentuu edeltäjänsä rakentamalle perustalle ja versioita voi olla useita, riippuen tutkimushankkeen kestosta. Mitään ei pidä hakata kiveen, vaan suunnittelussa pitää olla valmis reagoimaan muutoksiin. (Donnelly 2012, 89-90.) Myös Suomen Akatemian rahoituksen käytön yleisissä ehdoissa (2017–2018) suositellaan päivittämään aineistohallintasuunnitelmaa hankkeen keston aikana (Suomen Akatemia 2017a, 26).

Jokainen aineistohallintasuunnitelma on erilainen, sillä niihin vaikuttaa esimerkiksi maa, missä toimitaan, tavoitteleeeko tutkimus julkista vai kaupallista hyötyä ja onko tutkimus toistettavissa vai ainutkertainen. Rahoittajat, arkistot, tutkimusinstituutiot, tutkimusryhmät ja julkaisijat voivat kaikki asettaa omia vaatimuksiaan aineistoille koskien esimerkiksi tiedostomuotoja ja säilytysstandardeja, säilytysaikoja ja -määriä sekä aineiston siirron ja jakamisen aikataulutusta. Toisaalta tutkimusprojektin tutkijoiden henkilökohtaisilla mieltymyksilläänkin, kuten sillä, mitä tietokonejärjestelmiä he suosivat, on merkitystä. Toisaalta on suoranaisia laista nousevia vaatimuksia, kuten vaatimus saat-taa julkisrahoitteiset aineistot saataville tai vastaavasti estää aineistojen julkaiseminen esimerkiksi niin, että yksittäiset henkilöt on mahdollista tunnistaa. Kun laaditaan suunnitelmia ja strategioita,



suurimmassa osassa tutkimusprojekteista on otettava huomioon yksi tai useampi näistä näkökulmista. (Donnelly 2012, 88.)

Toisaalta siinä missä puhutaan aineiston elinkaaresta ja tutkimuksen elinkaaresta, myös aineistonhallintasuunnitelma-käsitteen lisäksi on käsite *tutkimussuunnitelma*. Borgin ja Kuulan mukaan tutkimusaineistojen elinkaari on usein suunniteltavissa varsin pitkälle jo tutkimussuunnitelmaa tehtäessä – vaikkakin vuonna 2007 he olisivat toivoneet tutkimussuunnitelmiin sähköisten tutkimusaineistojen elinkaaren parempaa huomiointia. He kuvailivat tutkimussuunnitelman olevan keskeinen varsinkin yliopistojen julkisrahoitteisessa tutkimuksessa, sillä se saattaa olla ainoa asiakirja, jossa selvitetään aineiston käyttöä koskevat suunnitelmat ja tiedot, kuten luvat ja sopimukset. Tutkimuksen edellytysten ja suunnitelmien valmiuteen jo varhaisessa vaiheessa – esimerkiksi huomioidaanko aineistonkeruun instrumentit ja luvat suunnittelussa – liitetään sitä enemmän vaatimuksia mitä suurempi tutkimushanke on ja mitä keskeisempi rahoittaja sillä on. (Borg & Kuula 2007, 13–14.)

Tiedonjulkistamisen neuvottelukunnan (TJNK) julkaisemissa tiedeviestinnän suosituksissa ”Viesti rohkeasti, vaikuta vastuullisesti” (2018) mainitaan niin tutkimussuunnitelma kuin tutkimuksen elinkaari ja tutkimusaineistotkin *tiedeviestinnän* kontekstissa. Tiedeviestinnässä on pelkistetysti kyse tiedeyhteisöjen sisäisestä ja ulkoisesta tiedonvälityksestä, joka koskee tutkimuksesta saatua tietoa, tutkimustuloksia, metodeja, tieteellistä ajatusmaailmaa sekä tieteenalojen teoreettista pohjaa. Suosituksissa tiedeviestinnän korostetaan olevan osa paitsi tutkimuksen suunnittelua, toteutusta ja tulosten julkistamista, myös tutkimusaineistojen jatkokäyttöä. Suosituksissa linjattiin, että jo tutkimussuunnitelmaa laadittaessa tiedeviestintä – ja tutkimusaineistojen huomiointi siis osana sitä – pitää ottaa osaksi tutkimustyötä. (TJNK 2018, 3, 7.)

Myös Judith Davidson käsittelee samanaikaisesti sekä tutkimusryhmän toiminnan ja viestinnän että aineistonhallinnan suunnittelua näkökulmanaan tutkimusprojektin toteuttaminen. Hän painottaa pelkän tutkimuksen lähteenä käytettävän aineiston hallinnan sijasta kaiken tutkimusryhmän toiminnassa projektin aikana syntyvän tiedon hallintaa ja osoittaa tutkimusryhmien vetäjille useita päätöksenteon osa-alueita hoidettavaksi. Yksi osa-alue on päättää, mitä ryhmän historian ja *toiminnan aktiivisuuden dokumentoinnilta* ja ymmärrettäväksi saattamiselta odotetaan. Toinen osa-alue on *tutkimusmetodien valitseminen*, mikä sisältää niin eri vaihtoehtojen läpikäynnin kuin päätökset siitäkin, missä laajuudessa aineistoja kerätään. Kolmas osa-alue on *projektinhallintajärjestelmän käyttöön ottaminen, lähestymistapana ”digitaalisten työkalujen pakki”*. Davidson nostaa esimerkiksi

aineiston digitaalisen säilyttämisen suunnittelun esille samassa yhteydessä kuin projektin suunnitteluvälineiden sekä viestintävälineiden ja -kanavien valinnan ja sosiaalisen median käytön suunnittelun. Neljäs osa-alue on tutkimussuunnitelman sulauttaminen osaksi *pitkäaikaista aineiston arkistointisuunnitelmaa*. (Davidson 2019, 47-63.) Viimeksi mainittuun sopivasti Graham Pryor toteaa, että puolestaan aineistohallintasuunnitelma toimii tehokkaana apuvälineenä, mikäli aineisto on tarkoitus saattaa lopulta arkistoon säilytettäväksi ja tarvittava säilytyspaikka on jo tiedossa. (Pryor 2012, 5.) Aineistohallinnan suunnittelu ja aineiston säilyttämisen suunnittelu voidaan nähdä toistensa osa-alueina. Pohjimmiltaan molemmat lähestymistavat liittyvät aineiston saatavuuden kysymyksiin, kuten siihen, voiko tutkija päästä käsiksi siihen aineistoon, johon haluaa, kun sitä haluaa, haluamassaan muodossa ja aineiston eheyteen luottaen. (Donnelly 2012, 83-84.)

Koska aineistohallintasuunnitelma-käsitteen sijaan kansainvälisessä kirjallisuudessa käytetään toisinaan käsitettä *aineistonjakamissuunnitelma* (data sharing plan) (esim. Piwowar 2011, 1), on syytä tarkastella myös aineistohallinnan ja aineiston jakamisen välistä suhdetta. *Aineiston jakamisesta* on kyse silloin, kun tutkijat tuovat oman aineistonsa tietoisesti muiden saataville joko tutkimuskäyttöön tai muihin tieteellisiin tarkoituksiin. Aineistoa jaetaan esimerkiksi sisällyttämällä sitä julkaistujen artikkelien yhteyteen, julkaisemalla aineistoa instituution tai tutkijan omilla verkkosivuilla, lähettämällä aineistoja tutkijakollegoille tai luovuttamalla aineistoja arkistoon. (Tenopir et al. 2011, 9; Tenopir et al. 2015, 3.) Aineiston jakamista tapahtuu samankin instituution tai tutkimusryhmän keskuudessa, joten hyvät jakamiskäytännöt nopeuttavat ja helpottavat työskentelyä myös tutkimusorganisaatioiden sisällä. Samalla riippuvaisuus kollegoista, jotka saattavat siirtyä aineistoihin ja tietoihin muualle, vähenee. (Borg & Kuula 2007, 32; White et al. 2013, 2, 8; Enwald 2018, 27; Alakyyry et al. 2018, 27.)

Tässä tutkielmassa keskeisellä sijalla on Benedikt Fecherin, Sascha Friesiken ja Marcel Hebingin laatima viitekehys, jossa akateeminen aineiston jakamisen prosessi jaetaan kuuteen eri kategoriaan. Kategoriat olivat aineiston luovuttaja, tutkimusyhteisö, tutkimusorganisaatio, datainfrastrukturi, normit sekä aineiston vastaanottaja, ja niistä kuhunkin kuuluu alaluokkia. (Fecher et al. 2015, 1.) Seuraavaksi katsotaan kategorioiden sisältöjä tarkemmin.

*Aineiston luovuttajaan* keskittyvä viitekehysten osa-alue kohdistuu yksittäiseen aineistoa keräävään tutkijaan. Sen alaluokkiin kuuluu kontrollin aste eli esimerkiksi tieto siitä, ketkä käyttävät aineistoa, miten sitä käytetään ja tulkitaan ja kenelle ensisijaiset oikeudet kuuluvat. Tarvittavat resurssit, kuten

aineistoa koskeva tietämys, aika, vaiva, taidot ja rahoitus, muodostavat myös oman alaluokkaansa samoin kuin sosiodemografiset tekijät sekä aineiston jakamisen palkitsevuus. (Fecher et al. 2015, 9-12.) *Aineiston vastaanottajan* osa-alue on jaettu kahteen alaluokkaan. Ensimmäinen alaluokka koskee vastaanottajan harjoittamaa tutkimusaineiston uudelleenkäyttöä epäsuotuisan käytön näkökulmasta. Kyse on esimerkiksi kilpailuetujen päätyemisestä muille kuin alkuperäiselle tutkijalle, aineiston väärentelystä tai kaupallisesta väärinkäytöstä tai siitä tehtävistä vääristä tulkinnoista. Toisessa alaluokassa tarkastellaan vastaanottajan organisaatiota. Huomio on siinä, tapahtuuko aineiston uudelleenkäyttö kaupallisen vai julkisen toiminnan piirissä ja miten tällöin huolehditaan aineiston asianmukaisesta suojaamisesta. Pystyykö aineiston uudelleenkäyttäjän organisaatio huolehtimaan aineiston asianmukaisesta turvaamisesta, aina fyysisestä säilyttämisestä henkilökunnan turvallisuuspolitiikkoihin? (Fecher et al. 2015, 16-17.)

*Tutkimusyhteisö*-kategorian alaluokkiin kuuluu ensiksikin aineistonjakamiskulttuuri, joka voi olla niin tieteenalakohtainen kuin aineiston luonteenkin mukainen. Toisena alaluokkana ovat esimerkiksi aineistojen yhteentoimivuutta tukevat metadatta ja tallennusformaatteja koskevat standardit. Kolmas alaluokka on tieteellinen arvo eli näkemys siitä, että aineistojen jakaminen tukee tieteellistä kehitystä, lisää verkostoitumista ja ajatustenvaihtoa sekä parantaa laadunvalvontaa. Neljäs alaluokka koskee julkaisuja, kuten julkaisijan vaatimusten ja kannustimien vaikutusta jakamiseen. (Fecher et al. 2015, 13-14.) *Tutkimusorganisaation* osa-alueeseen kuuluu niin aineiston luovuttajan oma organisaatio kuin hänen rahoittajansakin. Oma organisaatio voi olla vaikkapa yliopisto tai yrityksen kehitysosasto, ja se vaikuttaa muun muassa sisäisten toimintapolitiikkojensa ja organisaatiokulttuurinsa, kuten aineiston jakamiseen kannustamisen ja datainfrastruktuurin välityksellä. Rahoittajat vaikuttavat paitsi toimintapolitiikoillaan, myös rahallisten korvausten kohdistamisella. (Fecher et al. 2015, 12-13.) *Datainfrastruktuurin* osa-alue taas tarkoittaa tekijöitä, jotka liittyvät aineistojen välittämistä ja säilyttämistä varten tarvittavaan tekniseen infrastruktuuriin. Siihen kuuluvat arkkitehtuurin, käytettävyyden ja hallintaohjelmistojen alaluokat. (Fecher et al. 2015, 17-18.)

*Normeja* koskevalla viitekehyksen osa-alueella on kaksi alaluokkaa. Ensiksikin eettiset normit liittyvät erityisesti aineiston kerääjän moraaliseen näkökulmaan. Kyse on esimerkiksi luottamuksellisuudesta ja varsinkin tutkimuksen kohteen huomioinnista, tutkija-tutkittava-suhteesta, sekä sen arvioinnista, mitä vahinkoa tutkimusaineistojen jakamisesta voi seurata. Toinen alaluokka koskee oikeudellisia normeja, joihin kuuluu esimerkiksi aineistojen omistajuus, käyttöoikeudet, yksityisyydensuoja ja tekijänoikeudet. Kyse voi toisaalta olla lainsäädännöstä, toisaalta esimerkiksi aineiston

kerääjän ja tutkimuskohteen välisistä sopimuksista. Omistajuuskysymykset ja käyttöoikeudet koskettavat kaikentyyppisiä aineistoja ja nousevat esiin etenkin siinä vaiheessa, kun aineistoja siirretään säilytykseen. (Fecher et al. 2015, 14-16).

Muun muassa eri tieteenalojen kulttuuriset käytännöt ja päämäärät vaikuttavat siihen, miten tutkimusaineistojen käsittelyä milloinkin lähestytään, kuten painotetaanko niiden hallintaa vai jakamista. Osalla alakohtaisista rahoittajista painopiste on esimerkiksi digitaalisten resurssien säilyvyydessä, kun taas osan huomio on enemmän aineiston jakamisen potentiaalissa. (Pryor 2012, 5.) Tutkimusaineistojen tulevaa saatavuutta ja uudelleenkäyttöä suunniteltaessa huomio kohdistetaan muun muassa aineiston tuottamistilanteisiin. Joillakin aloilla, etenkin yhteiskunnallisissa, humanistisissa ja ympäristötieteellisissä piireissä, voi olla mahdotonta luoda uudelleen olosuhteita, joissa aineisto on tuotettu ja tallennettu, sillä kyseessä voi olla vaikkapa jokin ilmastoon liittyvä ilmiö tai ihmisten käyttäytymistä pitkäkestoisesti tutkiva hanke. Tällöin aineiston pitkäaikaissäilytykselle voidaan antaa enemmän painoarvoa. Aineistojen uudelleenkäyttö ja löydettävyyden puolesta korostuvat olosuhteissa, joissa toivotaan nopeaa ja helppoa aineistonjakamista, kuten lääketieteissä, tai joissa kokeita on helppo toistaa. (Donnelly 2012, 84.) Esimerkiksi biolääketieteellisen alan ja terveysalan rahoittajat antavat enemmän arvoa aineiston jakamiselle kuin aineistonhallinnalle. (Jones 2012, 54.)

Aineistopolitiikkojen retoriikka on alkanut painottua aineistonhallinnan sijasta yhä enemmän aineiston jakamiseen ja uudelleenkäyttöön, joiden nähdään tuovan suuremman arvon kuin pelkkä säilyttäminen. Kielellinen siirtymä aineistonhallinnasta aineiston jakamiseen voimistaa kustannustehokkuuskeskeisyyttä. (Jones 2012, 54, 56, 63.) Kuitenkaan aineistonhallintasuunnitelman keskiössä ei välttämättä ole pyrkimys aineiston jakamiseen. Kyse voi joskus olla pikemminkin aineiston saatavuuden rajoittamisesta tai estämisestä. Suunnitelmassa voidaan esittää, että projektissa tuotettu aineisto on asianmukaisinta tuhota heti projektin päätyttyä, vaikka se ei tuekaan aineiston uudelleenkäyttöä. Toisaalta tutkimusprojektissa, jossa kerätään arkaluontoista aineistoa, saatetaan asettaa tiukkoja rajoja aineiston julkaisemiselle vaikkapa aineistoon liittyvien ihmisten henkilöllisyyden suojelemiseksi ja tutkimuksen tekemisen vaatiman luottamuksen säilyttämiseksi. Tämä eettinen vaatimus on vastoin aineiston jakamisen periaatetta. Hyvin laadittu aineistonhallintasuunnitelma tunnistaa, selittää ja hallinnoi kyseisiä aineiston piirteitä. (Donnelly 2012, 85-86.)

Käsitteestä "aineistonjakamissuunnitelma" syntyy aktiivisempi ja päämääräkeskeisempi mielikuva aineiston käsittelystä ja käytöstä, kun taas termi "hallinta" on avoimempi ilmaisu, joka ei ota kantaa

aineiston käytön tavoitteisiin. Voinee sanoa, että aineistonhallintasuunnitelma antaa tilaa kuunnella aineiston luonnetta. Siinä voi olla kyse niin aineiston jakamisesta kuin säilyttämisestäkin, mutta jakamissuunnitelman päämäärä on yksiselitteisempi. Tosin aineiston jakaminen voi olla tutkijan näkökulmasta aineistonhallintaprosessin päätepiste, vaikka silloin, kun aineistojen avoimuudella pyritään edistämään aineistojen uudelleenkäyttöä ja tutkimusten uusinnettavuutta, kyse on pikemminkin uudesta alusta tai uuden tieteellisen keskustelun avaamisesta.

Jos suunnitelman laatimisen tavoitteena on edistää aineistojen kokonaisvaltaista hyödyntämistä, käsite *aineistonuudelleenkäyttösuunnitelma* voisi olla hedelmällinen ja korostaa aineistojen monikäyttöisyyttä ja pitkäikäisyyttä; vaihtoehtoisesti *aineistonhyödyntämissuunnitelma* korostaisi aineiston asemaa tutkimusprosessissa tai useassakin tutkimusprosessissa.<sup>2</sup> Toisaalta ”uudelleenkäyttö” on myös käyttäjäriippumattomampi käsite kuin ”jakaminen”: kyse voi olla niin siitä, että muut käyttävät suunnitelman laatineen tutkijan aineistoja, kuin siitäkin, että tutkija käyttää itse muiden aineistoja tai palaa omien aiempien aineistojensa äärelle.

Tässä tutkielmassa aineistonhallinta käsitetään laajana yläkäsitteenä, jonka katsotaan pitävän sisällään niin aineistojen keruun, käsittelyn, säilyttämisen, jakamisen kuin uudelleenkäytönkin, eli puhuttaessa mistä tahansa niistä, puhutaan samalla aineistonhallinnasta. Päähuomio on tutkimusaineistojen avoimuuden toteutuksessa sekä tutkimusaineistojen hyödyntämisessä, mukaan lukien uudelleenhödyntämisessä.

---

<sup>2</sup> Vrt. tiedeviestinnän suositusten maininta hyödyntämssuunnitelmasta: ”Tutkimussuunnitelmaa laadittaessa tiedeviestintä nivotaan osaksi tutkimustyötä. Suunnitelma, joka voi olla viestintä-, vuorovaikutus- tai hyödyntämssuunnitelma, arvioidaan rahoituspäätöksiä tehtäessä ja tutkimusta arvioitaessa.” (TJNK 2018, 7.)

## 3 KIRJALLISUUSKATSAUS

Kirjallisuuskatsauksen aluvussa 3.1 esittelen etenkin, millä metodeilla tutkimusaineistojen hallintaa on lähestytty tutkimuksissa sekä Suomessa että kansainvälisesti – iso osa pohjautuu tämän tutkielman lailla tutkijoiden parissa tehtyihin kyselyihin. Sen jälkeen huomio siirtyy tutkimusten antiin koskien aineistojen avoimuutta ja hyödyntämistä. Aluvussa 3.2 nostan esiin, miten aineistonhallinnan periaatteita, joita teoreettisessa viitekehyksessäkin käytiin läpi, on havaittu toteutettavan. Olennainen teema on esimerkiksi tiedeinstituutioilta tulevien velvoitteiden ja suositusten välillä tasapainoilu. Aluvussa 3.3 huomio on itse aineistoissa eli niiden avoimuuteen liittyvissä eettisissä ja oikeudellisissa kysymyksissä. Aluvun 3.4 aiheena on aineistonhallinnan vastuiden, kuten säilytysvastuiden, jakaminen. Aluvussa 3.5 nostetaan esille tutkijanäkökulma aineistonhallintaan, kuten aineistonjakamismäärät, aineistonhallinnan kuormittavuus ja avoimuudesta seuraavat urahyödyt.

### 3.1 Keskeiset tutkimukset

Suomalaiseen toimintaympäristöön keskittyneet Sami Borg ja Arja Kuula Yhteiskunnallisesta tietovarankirjasta (FSD) kirjoittivat vuoden 2007 julkaisussaan ”Julkisrahoitteisen tutkimusdatan avoin saatavuus ja elinkaari. Valmisteluraportti OECD:n datasuosituksen toimeenpanosta Suomessa” OECD:n datajulistuksen ”Declaration on Access to Research Data from Public Funding” merkityksestä tutkimusaineistojen käsittelylle. He selvittivät vuonna 2006 toteutetulla verkkokyselyllä tutkijoiden suhtautumista yleisemmin tutkimusaineistojen avoimuuteen sekä tarkemmin suosituksen toimeenpanoon. Kysely kohdistettiin Suomen yliopistojen humanististen, yhteiskuntatieteellisten ja käyttäytymistieteellisten alojen professoreille. Vastaajia kertyi 150, ja vastausprosentti (28,4 %) mahdollisti tilastollisen tarkastelun. Borg ja Kuula huomauttivat, että sähköisiä aineistoja hyödyntävät tutkijat ja laitokset saattoivat olla vastaajissa yliedustettuna, mikä voi näkyä tuloksissa. Julkaisussa esiteltiin myös alustavasta suhtautumisesta datajulistukseen keväällä 2005 tehty tiedustelu, joka oli kohdistettu 25 suomalaiselle tutkimusorganisaatiolle; puolet vastasi. (Borg & Kuula 2007, 26, 40, 62-63.)

Keskeiseen kirjallisuuteen kuuluu myös Graham Pryorin toimittama teos ”Managing Research Data” vuodelta 2012. Julkaisu sisältää muun muassa Sarah Jonesin artikkelin ”Research data policies: principles, requirements and trends”, jossa Jones vertailee Yhdistyneen kuningaskunnan eri tutkimusrahoittajien toimintapolitiikkoja (Jones 2012, esim. 50–53). Martin Donnelly puolestaan esittelee artikkelissaan ”Data management plans and planning” brittiesimerkkejä aineistonhallinnan suunnittelusta. Havainnollistavaa aineistonjakamiskeskustelua on käyty myös Queen's Universityn

julkaiseman, vertaisarvioidun *Ideas in Ecology and Evolution* -lehden vuoden 2013 erikoisnumerossa ”Data Sharing in Ecology and Evolution”, erityisesti artikkelissa ”Nine simple ways to make it easier to (re)use your data” (White et al. 2013). Kyseessä on siis tieteenalakeskeisempi, erityisesti ekologian aineistoihin keskittyvä, julkaisu.

Yksi tutkielmani keskeisintä lukemistoa on Carol Tenopirin et al. vuonna 2015 julkaisema artikkeli ”Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide”. Sen edeltäjä on Tenopirin et al. artikkeli ”Data Sharing by Scientists: Practices and Perceptions” vuodelta 2011. Molemmissa tieteellistä tutkimusta tekeviä tutkijoita oli tavoiteltu kyselyin. Vuoden 2011 artikkeli pohjautui jaksolla 2009–2010 suoritettuun kyselyyn. Vuoden 2015 artikkeli taas keskittyi jaksolla 2013–2014 tehtyyn kyselyyn sekä sen vertailuun aiempaan: miten tutkijoiden aineiston jakamiseen ja uudelleenkäyttöön liittyneet käytännöt ja mielipiteet olivat muuttuneet kolmessa vuodessa, jona aihepiiriä käsittelevä keskustelu oli myös lisääntynyt? Kyse oli sähköpostikyselystä ”lumipalloeefektiä” hyödyntäen eli kyselyn vastaanottaneet lähettivät kyselyä tiedepiireissä eteenpäin. Tutkijoita tavoiteltiin ympäri maailmaa, mutta enemmistö vastaajista tuli Yhdysvalloista. Näkemysten ja käytäntöjen mahdollisia eroja selittäviksi muuttujiksi oli valittu vastaajien ikäryhmät – 22–39-vuotiaat, 40–49-vuotiaat ja yli 50-vuotiaat – maantieteelliset vaikutusalueet sekä tieteenalat. Tutkimuksessa tehtiin jako ihmiskoeaineistoa käyttävien – mukaan lukien lääketiede/terveystieteet, liikeala / kaupallinen ala, psykologia ja yhteiskuntatieteet – ja muita aineistoja käyttävien välille, ja vertailtiin tieteenaloja toisiinsa myös pareittain. (Tenopir et al. 2015.)

Kontekstia aineistonhallinnalle rakentaa myös jo teoreettisessa viitekehyksessä sivuttu Benedikt Fecherin, Sascha Friesiken ja Marcel Hebingin vuonna 2015 PLoS One:ssa julkaistu artikkeli ”What drives academic data sharing?” Sen pohjana oli ensiksikin 98 systemaattisesti läpikäytyä artikkelia viittauksineen. Ne valittiin käymällä läpi vuosina 2001–2013 julkaistuja artikkeleita ja ottamalla niistä tarkasteluun aineiston jakamista tiedemaailmassa sekä primääritutkijan näkökulmaa käsitelleet. Toiseksi siihen kuului empiirisenä osiona kyselytutkimus, jolla testattiin lehdistä tehtyjä havaintoja. Tulosten tulkinnassa on huomioitava, että kysely ei edustanut tiedemaailman yleistä linjaa, sillä se tavoitti lähinnä saksankielisiä tutkijoita, joille sekundaariaineiston eli muiden keräämän aineiston käyttäminen oli tuttua: kyselyyn vastasi 603 sekundaariaineiston käyttäjää, jotka analysoivat German Institute for Economic Researchin toimeenpaneman German Socio-Economic Panelin aineistoja. Vastaajien keski-ikä oli 37 vuotta ja vastaajista 61 prosenttia oli miehiä. Suurin osa edusti talouden tai sosiologian/yhteiskuntatieteiden alaa. (Fecher et al. 2015, 3-7, 10.)

Veerle Van den Eynden et al. – yhtenä tekijänä oli myös jo aiemmin mainittu Carol Tenopir – julkaisivat vuorostaan vuonna 2016 artikkelin ”Towards open research. Practice, Experiences, barriers and opportunities”, jossa selvitettiin niin ikään aineiston jakamista ja uudelleenkäyttöä koskevia käytäntöjä ja näkemyksiä. Artikkelin pohjautui osittain lähes 600 tutkijaa tavoittaneeseen kyselytutkimukseen, joka keskittyi Wellcome Trust -säätiön rahoittamiin tutkijoihin, mutta sen tuloksia vertailtiin myös toiseen kyselytutkimukseen. Käytäntöjä ja näkemyksiä selittävinä muuttujina olivat vastaajien urakehitys, tutkimusalat sekä asuin- tai työskentelyseudut.

Katja Fält & Annaleena Okuloff laativat puolestaan vuonna 2016 Yhteiskunnallisen tietoarkiston (FSD) julkaiseman selvityksen ”Avoin tiede ja tutkimusaineistot -kysely 2015. Raportti terveystieteiden ja humanististen alojen tutkijoille suunnatun kyselyn tuloksista”. Heidän Internet-kyselynsä liittyi kiinteästi Tietoarkistoon ja sen palveluiden kehittämiseen, Tietoarkiston tieteenalalaajentumiseen. Kysely osoitettiin suomalaisten yliopistojen ja tutkimuslaitosten tutkijoille – tieteenaloina oli historia-alat sekä harkiten terveys- ja lääketieteet, pois lukien biotieteet – sikäli, kun heidän sähköpostiosoitteensa olivat Internetissä saatavilla. Se lähetettiin kaikkiaan 1428 tutkijalle, joista 199 vastasi. Kyselyn tuloksia vertailtiin Borgin ja Kuulan tutkimukseen niiden kysymysten osalta, joita käytettiin molemmissa, mutta raportissa huomautettiin, että kaikki Borgin ja Kuulan edellä mainitun, tuolloin kymmenen vuoden takaisen, selvityksen aiheista eivät olleet enää ajankohtaisia ja esiin oli noussut uusia keskeisiä teemoja, joita ei ollut aiemmin voitu käsitellä. (Fält & Okuloff 2016, 4, 19.) Raportti käsittelee paljolti samoja teemoja, joista itsekin olen kiinnostunut, mutta kyseessä on raportti, jossa esitellään yksinkertaisesti tulokset ilman kirjallisuudella kontekstointia.

Tutkimusaineistot ovat saaneet luonnollisesti myös Tampere3-hankkeessa huomiota, mistä on esimerkkinä Kristiina Tolvasen vuoden 2017 raportti ”Tutkimuksen infrastruktuurien kartoitus Tampere3-korkeakouluissa”. Metodologisesti kyseessä oli haastattelututkimus, ja haastateltaviksi valittiin asiantuntijoita, kuten dekaaneja, joilla olisi infrastruktuurin kannalta olennaisia näkemyksiä. Kyse oli tätä tutkielmaa teknisemmästä lähestymistavasta tutkimuksentekoon ja sen puitteisiin, kuten korkeakoulujen yhteisiin laitteistoihin, eikä kartoituksen pyrkimyksenä ollut tutkimuksellisen näkemyksen luominen. (Tolvanen 2017, 3.) Kuitenkin raportissa omistettiin tutkimusaineistoillekin oma lukunsa, jossa tuotiin esiin, että monessa tapauksessa tietyt infrastruktuurit mahdollistavat aineistojen tuottamisen. Lisäksi käsiteltiin jo mainitun tamperelaisen Yhteiskunnallisen tietoarkiston roolia aineistojen säilyttäjänä ja huomioitiin korkeakoulujen suuri tarve tutkimuksen aktiivivaihetta



palvelevan aineistonhallintainfrastruktuurin kehittämiseksi sekä systemaattiselle ja dokumentoidulle aineistojen tallentamiselle ja aineiston anonymisointipalveluille. (Tolvanen 2017, 14-16.)

Alkuvuodesta 2018, samaan aikaan, kun työstin tätä tutkielmaa, Juuso Ala-Kyyny, Tuija Korhonen ja Markku Roinila esittelivät Signum-lehdessä julkaistussa artikkelissa ”Tutkimusdatan avaamisen esteet: haastattelututkimus Helsingin yliopistossa” helsinkiläistutkijoiden näkemyksiä aineistonhallinnasta. Tutkimuksessa oltiin kiinnostuneita siitä, miten eri tieteenaloilla tunnustetaan arvokas aineisto, miten sitä säilytetään ja miten se on kuvailtu. Tutkimuksen pohjana oli 30 haastattelua. Haastateltavat tulivat Helsingin yliopiston neljältä kampukselta; kampusjako ilmensi myös tieteenaleroja. Katsauksessa päädyttiin haastattelumetodiin, koska kyselytutkimus ei ollut kevättalvella 2017 tuottanut tulosta. (Alakyyny et al. 2018, 25, 26.)

Tähän tutkielmaan kerättiin aineisto keväällä 2018, mutta sen jälkeenkin tutkimusaineistojen hallinnasta on julkaistu useita kiinnostavia ja samanhenkisiä julkaisuja. Esimerkiksi Heidi Enwaldin artikkeli ”Tutkimusaineiston avoin jakaminen – tutkimusorganisaatioiden jäsenten käsityksiä, kokemuksia ja mielipiteitä” julkaistiin Informaatiotutkimus-lehdessä joulukuussa 2018. Enwaldin päämääränä oli kyselytutkimuksen avulla etsiä vastausta siihen, millaisia käsityksiä ja kokemuksia ja mielipiteitä tutkimusorganisaatioissa työskentelevillä on tutkimusaineistojen jakamisesta ja pitkäaikaissäilytyksestä. Hän vertaili mies- ja naistutkijoiden eroja<sup>3</sup>, jatko-opiskelijoiden ja muissa työrooleissa olevien eroja sekä tieteenalakohtaisia eroja. Tutkimus oli osa kansainvälistä Data Literacy and Research Data Management -yhteistyöprojektia, jossa Iso-Britanniasta, Turkista ja Ranskasta olevien tutkijoiden laatima kysely toteutettiin yli 20 eri maassa. Suomessa kysely laitettiin levitykseen kesäheinäkuussa vuonna 2017 Tuuli-projektin verkostoa hyödyntämällä, sillä projektin edustajien katsottiin voivan jakaa kyselyä tehokkaasti eri tutkimusorganisaatioihin. Vastauksia kertyi yli 600, ja Tampereen yliopistosta vastaajia oli ainakin 21. (Enwald 2018, 31-34.)

Seliina Päälylysaho ja Jaana Latvanen Seinäjoen ammattikorkeakoulusta kirjoittivat puolestaan vuonna 2019 Signum-lehden artikkelissaan ”Avoin toimintakulttuuri vahvistuu ammattikorkeakou-

---

<sup>3</sup> Sukupuolen merkitystä aineistonhallintaan tutkitaan, vaikkakin sen merkityksen arvellaan olevan välillinen: sukupuoli vaikuttaa jossain määrin alavalintoihin ja alat, kuten käytetyt aineistot, näkemyksiin. (Enwald 2018, 43.)

luissa”, että useat ammattikorkeakoulut panostivat parhaillaan tutkimusaineistojen elinkaaren hallinnan kehittämistyöhön ja tekivät strategisia valintoja aineistojen avoimuuden edistämiseksi. Toisaalta osa oli rakentanut itse sopivia palveluja aineistojen kuvailemiseksi ja etsimiseksi, toisaalta osa hyödynsi valmiita kansallisia palveluja ja osallistui aktiivisesti niiden parantamiseen. (Päällysaho & Latvanen 2019.) Ammattikorkeakoulukontekstiin sopien, tosin tutkimusaineistojen sijaan laajemmin avoimen dataan liittyen, Mikko Uusitalo kirjoitti vuonna 2017, että Tampereen ammattikorkeakoulussa oli päätetty kokeilla Peppi-tietojärjestelmän avulla, mitä korkeakoulun datan avaaminen kaikkien käyttöön tarkoittaisi ja mahdollistaisi. TAMK:n kokonaisarkkitehtuuriperiaatteissa tiedon linjattiin olevan koko organisaation käytössä olevaa eli yhteiskäyttöistä pääomaa, mutta periaate ulotettiin avoin data -konseptilla organisaation ulkopuolellekin. Ajatuksena esimerkiksi oli, että avoin data tarjoaisi muun muassa ohjelmoinnin opiskelijoille tilaisuuden toteuttaa sovelluksia, joita TAMK:n opiskelijat voivat käyttää. Hankkeen pilotointi tapahtui vuonna 2014. (Uusitalo 2017.)

Tutkielman tuoreinta kirjallisuutta oli teos ”Sustainable Digital Communities” (Sundqvist et al. 2020), joka oli koottu Ruotsin Boråsissa alkuvuonna 2020 pidettyyn iConference 2020 -konferenssiin lähetetyistä vertaisarvioiduista artikkeleista. Yksi niistä on seattlelaista Washingtonin yliopistoa edustaneen Bernadette M. Boscoen artikkeli ”The What of Data: Defining Which Scientific Research Is Appropriate to Share”. Boscoen mielenkiinnon kohteena oli tarkastella, mitä osia tutkimusaineistoista jaetaan, arvioimalla tutkimusaineistojen välittämistä eri konteksteissa. Konteksteissa on kyse etenkin tiedeyhteisöjen normeista. (Boscoe 2020, 687.) Boscoe pohjasi artikkelinsa kolme vuotta kestäneen astronomian alalle kohdistuneen laadullisen tutkimuksen antiin. Hän esimerkiksi haastatteli itse yhteensä 40 astronomia tai muuta alaan toimijaa, kuten ohjelmoijia ja aineistojen säilytyspaikkojen työntekijöitä, ja hyödynsi myös astronomien haastatteluja, joita informaatiotieteilijät olivat keränneet toisen tutkijan johdolla. Lisäksi hän vietti kaksi vuotta erään astronomian tutkimusryhmän seurassa tutkien heidän aineistoihinsa kohdistamia käytäntöjä. (Boscoe 2020.)

Boscoe hyödynsi kontekstuaalisen eheyden viitekehystä eli CI-viitekehystä (contextual integrity, CI)<sup>4</sup>, joka liittyy kontekstisidonnaisia normeja heijasteleviin tietovirtoihin. Boscoe osoitti, miten CI-

---

<sup>4</sup> CI-viitekehys on suunniteltu etenkin henkilötietojen käsittelyyn. Sitä on käytetty muun muassa tilanteissa, joissa arkaluontoisiin aineistoihin on kohdistunut normatiivisia rikkomuksia. (Boscoe 2020, 687-689.)

viitekehys voi lisätä ymmärrystä siitä, mikä osa tutkimusaineistosta jaetaan, sillä sen avulla voidaan jäljittää, miten ihmiset siirtävät tietoa eri konteksteissa. (Boscoe 2020, 687-689.) CI-viitekehys on kuvaileva ja diagnostinen eli epäkohtia määrittelevä. Viitekehyksessä informaatiovirtaa jäsennetään arviointia varten viidellä parametrillä. Ensiksikin on kohde (subject), joka voi olla vaikkapa arkistossa oleva aineisto. Toinen parametri on aineiston lähettäjä ja kolmas aineiston vastaanottaja. Lähettäjät ja vastaanottajat ovat toimijoita, kuten yksittäinen tutkija, toinen tutkimusryhmä tai tiedejulkaisu. Neljäs parametri on informaatiotyyppi eli muoto, jossa tieto on, kuten sähköposti, DOI (digital object identifier), analysoitu aineisto tai artikkeli. Viimeinen parametri on välitysperiaatteet, jotka ovat määriteltävissä informaatiovirtoihin kohdistettaviksi rajoitteiksi. Välitysperiaate voi esimerkiksi olla, että aineistoa jaetaan ainoastaan kahden tutkimusryhmän välillä ja yhteisjulkaisun laatimista varten. CI-viitekehystä käyttämällä voidaan myös paljastaa muuttuvia välitysperiaatteita vertailemalla eri ajoilta olevia parametreja ja tutkimalla normeja. Boscoe perusteli tutkimustaan aineistoja kuratoivien aineistohallinnan ammattilaisten tukemisella. Aineistojen jakamisen lisääntyessä on tärkeää luoda kehittyneempiä ja erikoistuneempia tapoja jakaa aineistoa eri tasoilla, sillä siten löydetään vastauksia myös siihen ongelmaan, mitä aineistoa edes voidaan jakaa ja kenen kanssa. Boscoen yhtenä tavoitteena oli tuoda esiin normien ja aineistojen väliset suhteet siten, että saataisiin tietoa olennaisten infrastruktuurien suunnitteluun. Syvällisempi ymmärrys tutkimusaineiston välittämisen eri yksityiskohdista voi auttaa aineistojen säilytyspaikkojen suunnittelussa ja ylläpidossa. (Boscoe 2020.)

### **3.2 Aineistohallinnan periaatteiden tehokkuus**

Teoreettisessa viitekehyksessä tuotiin esiin, että tiedemaailmassa ja yhteiskunnassa laajemmin on kannustettu tutkimusaineistojen avoimuuteen. Aineistohallintaa varten laadittujen toimintapolitiikkojen suhteen voi tulla ongelmia, jos ei ole pakotteita, saati kannustimia, joilla varmistettaisiin, että rahoituksen saamisen jälkeenkin pysyttäisiin sovituissa aineistohallintasuunnitelmissa. Vaarana on tuuliajolle päätyminen. (Pryor 2012, 5.) Strategian menestyminen riippuu sen toimeenpanosta. (Donnelly 2012, 84.) Esimerkiksi Yhdysvalloissa NSF:n toimia aineiston jakamisen käytännön toteuttamisen suhteen on pidetty epäjohdonmukaisina. (Fecher et al. 2015, 13.) Toisaalta 2010-luvun alussa Sarah Jonesin mukaan Yhdistyneessä kuningaskunnassa valtaosa rahoittajista asetti aineiston julkaisulle aikarajat, osa kuukauden tarkkuudella. (Jones 2012, 54.) Suomen Akatemian rahoituksen käytön yleisissä ehdoissa (2017–2018) kirjoitetaan, että sen rahoittamat tutkimusaineistot “tulee avata mahdollisimman pian tutkimustulosten julkaisemisen jälkeen” (Suomen Akatemia

2017a, 26). Toisaalta rahoittajille on osoitettu, että kun avustuksista ja koulutuksista tehdään päätöksiä, tiedonhallinnan tarpeisiin on tarvetta ohjata tukea. (Haendel et al. 2012, 3.) Toisaalta koska julkaiseminen on keskeisin keino saada akateemisessa maailmassa tunnustusta, aikakauslehtien aineistohallintapolitiikoilla, niiden velvoittavuudella, voi olla enemmän vaikutusta kuin tutkimusrahoittajien tai instituutioiden vastaavilla. (Jones 2012, 62; Fecher et al. 2015, 14; FCRD 2018, 4.)

Tutkimusrahoittajien ja julkaisijoiden suositusten on havaittu vaikuttavan tutkijatason käytäntöihin koskien aineistojen avaamista ja jatkokäyttämistä – niiden on nähty lukeutuvan tutkijoiden pääsyihin jakaa aineistoja. Kuitenkin sen suhteen esiintyy tieteenalakohtaista vaihtelua, ja vaihtelussa ei ole kyse ainoastaan eri tieteenalojen edustajien erilaisesta halusta mukautua vaatimukseen, vaan julkaisijoiden vaatimuksetkaan eivät ole kaikilla aloilla samanlaiset. Esimerkiksi terveystieteiden ja humanististen tieteiden julkaisut eroavat toisistaan selvästi. (Fält & Okuloff 2016, 18; Van den Eynden et al. 2016, 4.) Suomessa rahoittajien, kuten Suomen Akatemian ja Tekesin, samoin kuin julkaisijoiden vaatimukset tunsivat yhteensä yli puolet kaikista Fältin ja Okuloffin tutkimuksen vastaajista, joskin humanistit tunsivat tutkimusaineistojen avoimuutta koskevat rahoittajien ja julkaisijoiden suositukset terveystieteilijöitä paremmin. (Fält & Okuloff 2016, 6.) Enwaldin tutkimuksessa vastaajat jakautuivat tasaisesti kolmeen joukkoon – myöntäviin, kieltäviin ja epätietoisiin – kun heiltä kysyttiin, suositteleeko heidän tutkimuslaitoksensa aktiivisesti aineistojen jakamista open access -periaatteella. Tieteenaloihin, saati sukupuoleen, liittyviä tilastollisesti merkitseviä eroja ei ollut. Sen sijaan asemalla tiedeyhteisössä oli vaikutusta. Esimerkiksi jatko-opiskelijat kokivat harvimmin laitoksensa suosittavan avoimuutta. (Enwald 2018, 34-35, 43.)

Tosin vuonna 2015 Tenopir et al. eivät havainneet merkittäviä tieteenalakohtaisia eroja siinä, millaisena tutkijat näkivät organisaatioidensa osallistumisen aineistohallintaan ja sen tukemiseen. Kuitenkin heidänkin havaintojensa mukaan aineistohallinta- ja jakamispolitiikat ovat yleisesti vaihdelleet instituutioittain ja rahoittajittain, ja toimeenpanokin on ollut vaihtelevaa, vaikka aineiston säilyttämistä, jakamista ja uudelleenkäyttämistä tuetaan laajasti periaatteen tasolla. (Tenopir et al. 2015, 4, 17.)

Borgin ja Kuulan professorikyselyssä kolmen neljäsosan enemmistö vastaajista olisi *suositannut* tutkimusryhmää luovuttamaan tutkimusaineistot tiedeyhteisön avoimeen käyttöön tilanteessa, jossa ryhmä ei itse käyttäisi aineistoa enää aktiivisesti, aineisto olisi julkisrahoitteinen ja sähköisessä muodossa, siihen ei liittyisi tietosuojongelmia eikä tekijänoikeudellisia tai kaupallisia intressejä, sen

toimittaminen tiedeyhteisön avoimeen käyttöön tapahtuisi vähäisin työpanoksin tietopalveluorganisaation kautta ja organisaatio huolehtisi aineiston käyttöluupa-asioista ja informoisi ryhmää aineiston myöhemmästä käytöstä. Vain viidennes olisi *velvoittanut* ryhmää luovuttamaan aineistot. Myös Suomen yliopistojen ja tutkimusrahoittajien ylivoimainen enemmistö oli asiassa samoilla linjoilla. (Borg & Kuula 2007, 42-44.) Toisaalta pakottavien toimien on arveltu voivan lisätä suopeuden sijasta kielteisyyttä aineiston jakamiseen; esimerkiksi rahoittajat voisivat niiden sijaan korostaa valistavalla otteella aineistojen jakamisesta olevan enemmän hyötyä kuin haittaa. (Poisot et al. 2013, 16; Fält & Okuloff 2016, 15.) Toisaalta taas Sarah Jonesin mukaan lehdet, jotka vaatimisen sijasta suosittelevat aineistojen jakamista, edistävät avoimuutta tehottomammin kuin vahvimmat toimintapolitiikat laatineet lehdet. (Jones 2012, 57.) On esitetty, että niin aineistohallintapolitiikkoihin kuin täsmällisempiin aineistonjakamisprosesseihin kuuluisi sisällyttää viittaamiskäytäntöjen ja tekijäkumppanuuden kaltaisia kannustimia. 2010-luvun alkupuolella tutkimusaineistoihin viittaamisen ennakoitiin voivan muodostua tulevien vuosien tieteelliseksi käytännöksi. (Tenopir et al. 2015, 18, 20-21.)

Jotta kannustimet saataisiin vaikuttamaan käytännössä, kehittämistoimeksi on esitetty ensiksikin, että arkistot voisivat asettaa säännöt viittaustandardeista, jotta aineistoja käytettäessä kunnioitettaisiin paitsi aineistojen alkuperäisiä tekijöitä, myös aineistoja säilyttäviä paikkoja. (Tenopir et al. 2015, 20-21.) Toiseksi korkeakouluilla on tehtävänsä siinä, että niin opiskelijat kuin tutkijatkin koulutetaan viittaamaan aineistoihin. Kolmanneksi akateemiset julkaisijat ovat vaikuttavassa asemassa, sillä ne voivat paitsi vaatia tutkijoita viittamaan myös aineistoihin, eikä vain artikkeleihin, joita niiden pohjalta on tehty, myös varmistaa, että aineistoviittaukset on asianmukaisesti tehty. (Poisot et al. 2013, 16; FCRD 2018, 4.) Neljäntenä on rahoittajan ja työpaikkojen merkitys. Toisaalta rahoittajilla on oma roolinsa palkitsemisjärjestelmiä laadittaessa. Esimerkiksi National Science Foundation on antanut tunnustusta aineistojen uudelleenkäytön arvolle kannustamalla rahoituksenhakijoita lisäämään jakamansa aineistokokonaisuudet ansioluettelonsa. Toisaalta akateemisia virkoja jaettaessa voitaisiin huomioida tutkijan panostukset hyödyllisten ja paljon viitattujen aineistokokonaisuuksien tuottamiseen ja jakamiseen. (Tenopir et al. 2015, 20-21; Poisot et al. 2013, 14.) Aineistoihin viittamisen teema oli 2010-luvun lopullakin mitä ajankohtaisin. Esimerkiksi vuonna 2018 Finnish Committee for Research Data (FCRD) tarttui aiheeseen julkaisussaan ”Tracing data: data citation roadmap for Finland”. Kyseinen tiekartta etsi esimerkiksi eri osapuolia ja vallitsevia oloja kartoittamalla sekä tavoitteita asettamalla ratkaisua siihen, miten aineistoihin viittaamisen käytännöt voitaisiin vakiinnuttaa ja laatia yhteentoimivia viittauskäytäntöjä, tieteellisiä julkaisuja vastaavalla tavalla.

Tavoitteena oli saada lisäkannustimia aineistojen hallintaan ja jakamiseen – aineistot olisivat aidosti arvostettu, legitiimi lisä ansioluettelossa – jotta aineistot, tieto, saataisiin virtaamaan vapaasti. (FCRD 2018, 6, 11.) Kun on luotu aineistonjakamisen kulttuuri, aineistonsa jakamatta jättävällä tutkijalla on suuri riski jäädä unohduksiin, sillä aineisto jää tuottamattomaksi, kuten vaille viittauksia. (Poisot et al. 2013, 13-14.)

Jotta aineistonhallinnan eri osa-alueiden strateginen kehittäminen ja suunnitelmien oikea kohdentaminen onnistuu, tarvitaan tietoja tiedemaailman eri ryhmien eroavaisuuksista. Aineistonhallinnan näkemyksissä samoin kuin niissä tapahtuvissa muutoksissa on yleisesti ikäryhmittäistä, alueellista ja tieteenalakohtaista vaihtelua. Muun muassa kulttuuristen ja tieteenalakohtaisten erojen on havaittu olevan ikään liittyviä syvempään juurtuneita. Onkin päätelty, että esimerkiksi aineiston jakaminen voi olla normaalimpaa tietyissä ryhmissä kuin toisissa. Kuitenkin 2010-luvun puolivälissä tutkijoilla oli kansainvälisellä tasolla aiempaa positiivisempia näkemyksiä jaetun aineiston arvosta sekä enemmän halua jakaa aineistoja ja käyttää valmiita aineistoja. (Tenopir et al. 2015, 7, 17, 21.) Myös Suomessa Borgin ja Kuulan tutkimuksessa näkyi jo 2000-luvun ensimmäisellä vuosikymmenellä myönteisyys aineistonjakamista kohtaan: yli kolme neljäsosaa vastaajista suhtautui vähintään melko myönteisesti omien sähköisten aineistojensa avoimeen jatkokäyttöön. (Borg & Kuula 2007, 43.)

### **3.3 Eettisten ja oikeudellisten kysymysten vaikutus avoimuuteen**

Tutkimusaineistojen avoimuus tukee tieteentekoprosessin eettisyyttä, tiedettä, mutta myös aineistoihin itseensä liittyy eettisiä näkökulmia. Toimintapolitiikkatason vaatimukset saada aineistot vapaasti saataville nostattavat aineiston luonteesta riippuen paljon legitiimejä kysymyksiä koskien immateriaalioikeuksia, turvallisuusasioita, tietosuojaa sekä aineistojen mahdollisia väärinkäyttöjä ja -tulkintoja. Esimerkiksi erilaisten aineistojen hyödyntämistä koskevien suostumusten pitäisi perustua riittävään tietoon, mutta asiakirjojen laatimisessa rajoitteena on, että aineistojen tulevat käyttötavat voivat olla tuntemattomia. Toisaalta on työlästä, jos esimerkiksi haastatteluaineistojen säilyttämiseen on kysyttävä lupa jokaiselta haastateltavalta, ja se voi myös supistaa aineistojen kattavuutta. Toisaalta kaikissa tilanteissa ei välttämättä kyetä varmistamaan aineistoissa esiintyvien henkilöiden yksityisyyttä anonymisoinnilla, vaikka joillakin aloilla se voikin olla aineistojen selkeän rakenteen ansiosta sujuvaa. (Borg & Kuula 2007, 37; Tenopir et al. 2015, 4; Alakyynty et al. 2018, 27.)

Ratkaisuksi tietolähteiden suojelemiseksi on esitetty yleistetyemmän aineiston julkaisemista ja sen riskin vähentämistä, että suuret määrät yksityiskohtia muuttaisivat muuten anonyymien tiedon

tunnistettavaksi. (Tenopir et al. 2015, 21.) Rahoittajana Suomen Akatemia on esittänyt, että aineistoilla voi olla erilaisia avoimuuden asteita vaihdellen kaikille täysin avoimesta salassa pidettävään. (Suomen Akatemia 2017a, 26.) Toisaalta jos aineistot siirretään data-arkistoihin, aineistojen uudelleenkäyttöön liitettyjen uhkakuvien hallintakeinoina on rajattu käyttöoikeus tutkimukseen tai opetukseen, lupakäytännöt ja jatkokäyttäjät juridisesti sitovat käyttöehdot. (Borg & Kuula 2007, 51.)

Mitä arkaluontoisempi tutkimuksen aihe on, sitä enemmän tietoturva aineistojen digitaalinen säilytys vaati. Tietoturvan kannalta olennaisia kysymyksiä on, missä palvelimet sijaitsevat, mitä turvakäytäntöjä noudatetaan, millaisia varmuuskopiointipalveluita tarvitaan sekä mitä seurauksia ja vahinkoa tietosuojan mahdollisesta rikkoutumisesta voisi koitua tutkimushankkeeseen osallistuneille. Tutkimusprojektille voidaan laatia muodolliset eettiset asiakirjat, jotka kattavat projektia koskevat niin oikeudelliset kuin sosiaalisetkin normit ja jotka sijoitetaan saataville yhteiseen sähköiseen säilytyspaikkaan niille varattuun osaan. Asiakirjat on laadittava koskettamaan läheisesti kaikkia projektin osa-alueita ja osapuolia niin, että koko projektia tarkastellaan niiden valossa. (Davidson 2019, 53, 64-66.)

Tietosuojariskit ja tutkimuseettiset riskit mietityttivät miltei puolta Borgin ja Kuulan professorikyselyyn vastanneista. Vastaajista kolmannes näki lainsäädäntöön, esimerkiksi tekijänoikeuteen, liittyvät ongelmat vähintään melko tärkeänä syynä sille, että tutkimusaineistoja ei käytetä jatkotutkimuksissa. Lähes puolet nimesi vähintään melko tärkeäksi syyksi aineistojen jatkohyödyntämisen vähäisyydelle sen, että niiden omistus- ja käyttösuhteita ei tehty selväksi aineistojen keräämisvaiheessa. Myös yliopistojen sopimuskäytännöissä on ollut puutteita koskien tutkimusaineistojen omistusoikeuksia. (Borg & Kuula 2007, 32, 51, 70.)

Toisaalta myös kymmenisen vuotta Borgin ja Kuulan tutkimuksen jälkeen Enwaldin tutkimuksessa laillisuuteen ja eettisyyteen liittyvät kysymykset olivat tutkimusaineistojen avoimuutta koskeneista huolista yleisimmät – runsas kolmannes vastaajista nimesi ne. Eettisyyttä ja laillisuutta koskevat teemat huolestuttavat eniten insinööritieteiden ja teknologia-alojen edustajia sekä yhteiskuntatieteilijöitä ja ihmistieteilijöitä. (Enwald 2018, 39-40.) Tenopirin et al. tutkimuksessa havaittiin, että tietyillä tieteenaloilla selvä enemmistö, mukaan lukien yli puolet lääketieteen/terveystieteen edustajista, on taipuvaisia toteamaan, että heillä ei ole aineistojensa jakamiseen oikeuksia. Kolmannes kaikista kyselyn vastaajista piti tärkeänä uudelleenkäytön ehtona, että aineistonkäytölle hankitaan oikeudellinen lupa – tosin yli kaksi viidesosaa torjui tarpeen viidenneksen ollessa epävarma. (Tenopir et al.

2015, 16; S1 Appendix.) Fältin ja Okuloffin kyselyssä terveystieteilijät näkivät samojen keskeisten jatkokäyttöön liittyvien haittojen, kuten tietosuoja- ja tutkimuseettisten riskien lisääntymisen ja tutkittavien informoinnin puutteen, koskevan sekä aineistojen jatkokäyttöä että pelkkää arkistointia. Humanistit kokivat heidän laillaan yhdeksi suurimmista haitoista sen, että tutkittavia ei ole informoitu aineistojen arkistoinnista. (Fält & Okuloff 2016, 14-15.)

Omistajuuskysymykset ovat yksi tutkimusaineistojen jakamista hankaloittava tekijä, mutta juuri aineistojen jakaminen voi myös osaltaan kannustaa selventämään tekijäsuhteita. Aineistojen jakaminen muuttaa yhteistyöprojekteissa tapaa tehdä kunkin osapuolen työpanosta heijastavia päätöksiä siitä, kuka saa tulla merkityksi tekijäksi ja missä järjestyksessä tekijät listataan. Kun aineistoista tehdään julkisia ja viitattavissa olevia, niihin voidaan viitata suoraan ilman kytköstä artikkeleihin, joita niiden pohjalta on tehty. Siten aineiston tuottaneen tutkijan vaikuttavuuden voi nähdä lisääntyvän. Artikkelijulkaisun tekijöiksi voidaan puolestaan nimetä aineiston analysoineet tutkijat. (Poisot et al. 2013, 13.)

Judith Davidson painottaa tekijyyteen liittyen ennakkoinnin merkitystä: tutkimusprojektien ja tutkimusryhmien sisällä on tarpeen laatia kirjallinen linjaus kaikkien projektin puitteissa tehtävien julkaisujen tekijyyksistä. Linjaus tekee eri tutkimusjulkaisuihin liittyvät odotukset läpinäkyvämmiksi, mikä paitsi vähentää ryhmän sisäisiä jännitteitä, myös estää ryhmän jäseniä tekemästä eettisiä rikkomuksia. Tulosten levittämistä koskevaan sopimukseen tulisi dokumentoida päätökset esimerkiksi aineiston omistajuudesta, jos hanke rahoitetaan avustuksilla, ja myös rahoittajalta mahdollisesti tulevista sitovista vaatimuksista on oltava selvillä. Kun tutkimusprojektissa on useita osapuolia, olennaisia ratkaistavia kysymyksiä on, mitä aineiston käyttöä koskevia velvoitteita tai oikeuksia kunkin roolin haltijalle kuulu ja millaista prosessia noudatetaan vaikkapa siinä vaiheessa, kun projektiin osallistunut haluaisi käyttää aineistoa esimerkiksi omassa kirjoittelussaan. Samoin on sovittava, kuka arvioi viime kädessä niin aineistot kuin dokumentit ja muutkin tuotokset ohjeiden noudattamisen varmistamiseksi. Oma kysymyksensä on niin ikään se, ketkä ovat portinvartijoita siinä vaiheessa, kun päätöksiä kyseisistä asioista alun alkaen tehdään. Davidsonin mukaan tärkeä sääntö kullekin tutkimusryhmän jäsenelle projektin aineistoa käytettäessä on keskustella tutkimusryhmän johdon kanssa, mikäli esiin nousee epävarmuuksia. (Davidson 2019, 46-47.)

Toisaalta Davidsonin tuodessa esiin tutkimusryhmän sisäistä päätäntävaltaa Bernadette M. Boscoen tutkimuksessa ilmeni, että voi olla ongelmallistakin, jos esimerkiksi ryhmän johtajalla on suuri mutta



epämuodollinen päätäntävalta, sillä ryhmän jäsenet voivat kaivata vielä ylemmän tason julkaisulinjauksia. (Boscoe 2020, 691.) Toisaalta yleisestä välittömyydestä kertoo, että Tenopirin et al. mukaan tutkijat suhtautuvat suopeasti aineistojen jakamiseen kollegan pyytäessä sitä suoraan, sillä tutkija kokee tulevansa vakuutetuksi siitä, että aineistoa käytetään asianmukaisesti ja alkuperää kunnioittaen – tosin pyynnöt ovat harvassa. (Tenopir et al. 2015, 3.) Martin Donnelly muistuttaa, että aineistohallinnassa inhimilliset haasteet ovat usein esimerkiksi teknologisia haasteita hankalampia, joten kommunikointi on ratkaisevan tärkeää. (Donnelly 2012, 89.) Seuraavaksi tarkastellaankin tarkemmin aineistohallinnan vastuita ja vuorovaikutusta.

### **3.4 Aineistohallinnan vastuiden jako**

Sarah Higgins esittelee kolme tapaa hallita aineistoja tutkimusorganisaatiossa. Ensimmäisessä tavassa aineistohallinta on tutkimusryhmän vastuulla läpi kuraation elinkaaren ja aineisto säilyy heidän hallussaan, kun tutkimus on valmis. Toisessa tavassa tutkimusryhmä hallinnoi ja säilyttää aineistoa, ja yleiskatsauksen aineistohallintaan liittyvistä toimista organisaation sisällä tekee informaatioalan ammattilainen, kuten kirjastonhoitaja. Isoissa organisaatioissa, kuten yliopistoissa, saattaa olla laitostason välikäsi, joka toimii yhteyshenkilönä tutkijan ja informaatioalan ammattilaisen välillä. Kolmannessa tavassa tutkimusryhmä vastaa aineistohallinnasta, kun tutkimus on käynnissä, minkä jälkeen aineisto siirretään huolehtivan tahon haltuun, kuten organisaation kirjastoon tai alakohtaiseen säilytyspaikkaan. (Higgins 2012, 22.)

Graham Pryor nostaa esiin tutkijoiden oman valinnan ja informaatioalan ammattilaisten itseohjautuvuuden. Toisaalta tutkijat voivat olla itseriittoisia, eivätkä tartu informaatioalan ammattilaisten apuun. Toisaalta, vaikka he myöntävät olevansa huolissaan aineiston saatavuuden ja hankinnan esteistä, tutkimustyö on niin kiivastahtista, että ei ole aikaa etsiä apua oman tutkimusryhmän ulkopuolelta. Siksi informaatioalan ammattilaisten on raivattava itse itselleen toimintatilaa: opittava tutkijoiden huolista ja oltava aloitteellisia avun ja tuen tarjoamisessa, mikä on ensi askel yhteyden luomisessa tutkimusyhteisöön. (Pryor 2012, 9-10.) Esimerkiksi kirjaston työntekijöillä voi olla asiantuntemusta myös aineistohallinnasta, kuten aineiston jakamisen standardeista. Kuitenkin tutkimusaineistot ovat olleet tietopalvelun näkökulmasta väliinputoajia: jos tutkija haluaa käyttää valmista aineistoa, voi olla, että hän ei saa keneltäkään apua esimerkiksi sen alkuperän ymmärtämisessä. (Haendel et al. 2012, 3.)

Judith Davidson antaa ymmärtää hyvinkin tiiviin yhteistyön olevan arkistoinnissa tarpeen: jotta suunnittelu tehdään asianmukaisesti, tutkimusryhmän pitäisi ottaa informaatioalan asiantuntija ryhmänsä jäseneksi. Toisaalta hänkin toteaa, että asiantuntijan läsnäolo voi tapahtua myös konsultin roolissa ja arkistokin voi ottaa tehtävän hoitaakseen. (Davidson 2019, 62.) Bernadette M. Boscoe korostaa taas tutkijoiden omaa vastuuta ulkopuolisten asiantuntijoiden informoimisessa mainiten, että tutkijoiden on tärkeää pitää aineistojen hoitaja, kuten arkistonhoitaja, ajan tasalla esimerkiksi eri aineistojen alkuperästä ja niihin liittyvistä sopimuksista: aineistoihin saatetaan liittää erilaisia oletuksia, joita ei ole sanallistettu yksiselitteiseen muotoon. Erot voivat olla tutkimusryhmäkohtaisiakin, sillä yksi ryhmä saattaa jakaa hyvin avoimesti aineistojaan niiden käsittelyyn liittyviä koodoja myöten siinä missä toisen linjana voi olla ehdoton jakamiskielto. Jos sopimuksista ei tehdä täsmällisiä, aineiston vastaanottaja voi olla epätietoinen siitä, miten laajasti saatua aineistoa saa käyttää. Yksi ratkaisumalli onkin, että jos aineisto on jaettu muille vain yhtä tarkoitusta ja yhtä kertaa varten, se voitaisiin suojata rajatun aikaa voimassa olevalla koodilla. (Boscoe 2020, 691.) Myös Helsingin yliopiston tutkijoiden parissa on havaittu alakohtaisesti vaihtuvat aineistohallinnan käytännöt ja henkilösidonnaisuus (Alakyynty et al. 2018, 26-25), ja Enwaldin kyselyyn tullessa avovastauksissa tuotiin esille sitä, että työnantajan ohjeet rajoittavat aineiston jakamista (Enwald 2018, 39; ks. myös Tenopir et al 2011, 2-3).

Tutkimusaineistojen säilyttäminen on merkittävä vastuualue. Judith Davidson pohti vuonna 2019 laadullisen tutkimuksen yhteydessä aineistojen digitaalista varastointia epäillen, että toimistojen paperiarkistokaappien aika alkaa olla ohi. Jotkut projektit nojaavat oman laitoksensa palvelimeen, toiset sijoittavat aineistonsa kaupallisten palveluntarjoajien haltuun, mistä on esimerkkinä Dropbox tai OneDrive. (Davidson 2019, 53.) Tenopirin et al. tutkimuksessa havaittiin, että kaudella 2013–2014 tutkimusaineistoja sijoitettiin eniten esimerkiksi tutkijan omalle tietokoneelle tai organisaation palvelimelle tai paperiversiona toimistoon. Tutkijat nojasivat vielä esimerkiksi pilvivaihtoehtojen markkinoille tulon jälkeenkin paljolti henkilökohtaisiin säilytystapoihin. Tutkimus paljasti, että 2010-luvun puolivälin lähestyessä yli 90 prosenttia kaikista vastaajista säilytti ainakin osan aineistostaan henkilökohtaisella tietokoneellaan. Hieman alle 70 prosenttia valitsi aineistoilleen organisaationsa palvelimen samoin kuin paperimuodon toimistossaan. Osastonsa palvelimella aineistoa säilytti hieman alle puolet vastaajista. Kunkin arkistoon rinnastettavan säilytyspaikan hylkäsi yli kolme viidesosaa kantaa ottaneista. Tosin paperisäilytyksen kohdalla tieteenalakohtaiset erot olivat suuria. Samoin havaittiin, että kasvatustieteilijät sekä myös tekniikan alan edustajat suosivat selvästi

muita aloja enemmän laitoksen palvelinta. (Tenopir et al. 2015, 8, 9, 17, S1 Appendix.) Alakyynty et al. puolestaan havaitsivat, että Helsingin yliopistossa kaikilla tieteenaloilla aineistosta vastasivat lähinnä tutkijat tai tutkimusryhmät itse. Digitaalisia aineistoja säilytettiin vaihtelevan huolellisesti ja niin henkilökohtaiset verkkolevyt, muistitikut ja ulkoiset kovalevyt kuin yliopiston yhteiskäyttöiset verkkolevyt ja pilvipalvelimetkin saivat kannatusta. Pitkäaikaissäilytysratkaisun suunnittelu tai sopiminen olivat sidoksissa esimerkiksi arkistoinstituutioiden valmiuteen ottaa aineistot vastaan. (Alakyynty et al. 2018, 26-27.)

Sillä, miten paljon tutkijat tallentavat aineistoja eri sijainteihin, on osaltaan vaikutusta myös siihen, miten paljon he jakavat aineistoaan. Silloin kun aineistoja tallennetaan esimerkiksi palvelimille, muilla tutkijoilla voi olla pääsy niihin. Tenopirin et al. kyselyaineistossa vajaalle viidennekselle vastaajista säilytyspaikan puuttuminen oli syy jättää aineistoja jakamatta. (Tenopir et al. 2015, 11, S1 Appendix.) Toisaalta julkisessa arkistossa julkaistu tutkimusaineisto voi erota siitä, joka on luotu tutkimusprosessia toteutettaessa eli analyysin tuloksena. Juuri viimeksi mainittua aineistoa saatetaan jakaa esimerkiksi tutkimusryhmän tai yksityisten säilytyspaikkojen kautta tai yliopiston palvelimien tai erilaisten alustojen välityksellä. (Boscoe 2020, 690.) Kuitenkin varsinkin silloin, kun on kyse digitaalisista aineistoista, olisi tärkeää, että rahoitus, henkilökunta sekä laitteisto ja järjestelmät ovat oikea-aikaisesti valmiina toimenpiteiden tekoa varten, mikä vaatii sekä strategisen että taloudellisen suunnittelun tukea. Pienet tutkijaryhmät, joilla ei ole organisaatiotason tukea, voivatkin olla aineistojen tehokkaan hyödyntämisen saralla vaikeuksissa. (Higgins 2012, 20; Bascoe 2020, 688.)

Aineistonhallintaa edistää sekä sovittu säilytysratkaisu, kuten luonteva taho, jolle esimerkiksi tietyn tieteenalan aineistot siirretään, että aineiston siirtoa varten kehitellyt vakiintuneet toimintatavat. (Pryor 2012, 5; Alakyynty et al. 2018, 28.) Borg ja Kuula linjaavat, että tutkimusaineiston avoimuutta tuetaan luomalla niitä säilyttäviä, dokumentoivia ja välittäviä erityisorganisaatioita, jotka ovat luotettava kolmas osapuoli aineiston kokoajien ja jatkokäyttäjien välissä, ja mainitsevat esimerkkinä oman organisaationsa eli Yhteiskuntatieteellisen tietoarkiston. Suomessa on organisoitu tutkimusaineistojen avoimuutta ja hallittua jatkokäyttöä perustamalla erilaisia data-arkistoja tai muita sähköisten aineistojen akateemisia tietovarantoja. Humanistisilla aloilla perinnearkistot, kuten Suomalaisen Kirjallisuuden Seuran kansanrunousarkisto, painottuvat etenkin paperiaineistojen arkistointiin, mutta toisaalta osa niistä on kehittänyt myös aineistojensa digitointia. Kielitieteen alalla taas esimerkiksi Kotimaisten kielten tutkimuskeskuksessa on kattavat sähköiset tutkimusaineistot. (Borg & Kuula 2007, 8, 50.) Judith Davidson kirjoittaa vuonna 2019, että aineistojen arkistointi on ollut

pitkään merkittävää kvantitatiivista tutkimusta tekeville, mutta se on viime vuosina alkanut saada jalansijaa myös laadullista tutkimusta tekevien parissa. Kun niin Yhdysvallat kuin Euroopan unionikin sekä joukko kansallisia ja globaaleja organisaatioita alkoivat vaatia laadullisen aineiston arkistointia, toiminnan tukeminenkin alkoi lisääntyä nopeasti.<sup>5</sup> (Davidson 2019, 61.)

Enwaldin tutkimuksessa kolme neljäsosaa tutkijoista näki säilytyspaikan järjestämisen kuuluvan tutkimuslaitoksen vastuulle, vähän alle puolet kannatti tutkimusorganisaation ulkopuolista ilmaista tallennustilaa, runsas kymmenesosa maksullista ulkopuolista tallennustilaa ja alle kymmenes rahoittajan järjestämää säilytyspaikkaa. Avovastauksissa tutkijat nostivat arkistovaihtoehtoja esille. Tutkimuslaitoksen, rahoittajan ja ulkoisen maksuttoman tallennustilan kannattajia oli etenkin insinööri- ja teknologia-alojen parissa. (Enwald 2018, 40-41.) Fältin ja Okuloffin raportin mukaan silloin, kun tutkimusaineistot päätetään arkistoida, yleisin menettelytapa on valita oman yliopiston tai tutkimuslaitoksen säilytys. Selvä enemmistö heidän vastaajistaan piti sopivimpana arkistoinnina valtakunnallista data-arkistoa, ja yli puolet kannatti myös yliopistojen ja tutkimuslaitosten data-arkistoa. Sen sijaan alle puolet piti kansainvälistä data-arkistoa soveltuvana säilytyspaikkana. Humanistit valitsivat selvästi terveystieteilijöitä useammin erillisen tutkimusaineistojen arkistoinnin palvelut. (Fält & Okuloff 2016, 6–9.) Kymmenisen vuotta aiemmin Borgin ja Kuulan kyselyyn vastanneiden mukaan data-arkistot ovat paras tapa hoitaa tutkimusaineistojen arkistointi ja jatkokäyttöön luovutus. Kahden kolmasosan mukaan oli erittäin tai melko hyvä asia, jos tehtävästä vastaisivat yliopistokohtaiset data-arkistot. Parhaimmaksi tavaksi arvioitiin valtakunnalliset tieteenalakohtaiset data-arkistot, joita miltei puolet piti erittäin hyvänä tapana ja lähes kolmasosa melko hyvänä tapana hoitaa tutkimusaineistojen avoimuutta koskevat käytännön työt. (Borg & Kuula 2007, 49, 71.)

Toisaalta siinä missä arkistoilla on valmiudet ottaa vastaan niin kirjallisessa kuin audiomuodossakin olevaa aineistoa niiden ollessa yleisissä formaateissa, monimutkaisimpien muotojen tullessa

---

<sup>5</sup> USA:ssa esimerkiksi Michiganin yliopiston yhteydessä Inter-University Consortium for Political and Social Research (ICPSR) arkistoi jonkin verran laadullista aineistoa, ja hiljattain Syracusen yliopistossa Center for Qualitative and Multi-Method Inquiry perusti laadulliselle aineistolle arkiston, Qualitative Data Repositoryn (QDR). Irlannissa puolestaan Maynoothin yliopisto perusti laadullisen aineiston arkiston, Irish Qualitative Data Archiven. Toisaalta UK on ollut edelläkävijä laadullisen aineiston arkistoinnissa. Siellä on kehitetty esimerkiksi QualiBank, joka on UK Data Servicen haku- ja selailukäyttöliittymä. (Davidson 2019, 61.)

vastaan arkistoissakin saatetaan joutua pitämään mietintätauko esimerkiksi ohjelmistokielen puutteiden vuoksi. (Davidson 2019, 62.) Toisinaan ongelmia voi tulla myös silloin, jos tutkimusryhmien, julkaisijoiden ja aineiston säilytyspaikkojen normit törmäävät. Esimerkiksi aineistojen säilyttäjinä ja välittäjinä toimivat organisaatiot eivät välttämättä ole järjestäneet aineistoja siinä muodossa, jossa julkaisijat haluavat vastaanottaa tutkijoiden aineistoja. Ristiriitojen takana voi olla myös nopeasti muuttuvat normit. (Boscoe 2020, 687, 689, 692.)

Organisaatiotasolta tai tutkimuksen rahoittajatasolta tulevat toimintapolitiikat tukevat parhaimmillaan sitä, että vaaditut aineistohallinnan toimet toteutetaan sulauttamalla ne tutkimusorganisaation tai yliopiston työvirtaan. (Higgins 2012, 20.) Martin Donnelly kirjoittaa tutkijoille kehitellyistä työkaluista nostaan esiin muun muassa aineistohallintasuunnitelmien laatimiseen ja ylläpitoon suunnitellun ”DMP Onlinen”. Se on tarkoitettu muun muassa auttamaan tutkijoita määrittelemään aineistoon liittyviä rooleja ja vastuita, tunnistamaan riskejä aineistohallinnan siirtymävaiheissa sekä varmistamaan (sähköisen) aineiston turvallisen siirtymisen alkuperäisiltä haltijoilta myöhemmille vastuutahoille. (Donnelly 2012, 94-95.) DMPToolista on myös suomenkielinen versio DMP-Tuuli, jonka pariin myös Suomen Akatemia opasti tutkijoita vuonna 2017 aineistosuunnitelmaohjeissaan (Suomen Akatemia 2017b).

Tutkimuksen rahoittajien keskuudessa yleinen näkemys on ollut, että kaikissa tapauksissa varoja ei kannata sijoittaa aineistohallintaan ja jakamiseen. Rahoittajien parissa on tehty linjauksia aineistojen vähimmäissäilytysajoista samoin kuin listauksia ominaisuuksista, joilla voisi tunnistaa, millaiset aineistot kannattaa säilyttää pitkäaikaisesti, sekä perustettu omia datakeskuksiakin. Sarah Jones kirjoitti vuonna 2012, että aineistohallinnan ja aineiston jakamisen aiheuttamat kulut aletaan huomioida aineistohallintapolitiikoissa. Hänen esittelemässään vertailussa vuonna 2012 Yhdistyneessä kuningaskunnassa vain yksi toimija, Cancer Research UK, otti kannan, että se ei anna lisärahaa kyseisiin toimiin, vaan näki niiden kuuluvan kiinteäksi osaksi itse tutkimusprosessia. (Jones 2012, 55-56.) Martin Donnellyn mukaan yhä useampi rahoittaja on alkanut suhtautua kannustavasti siihen, että tuenhakijat sisällyttävät aineistohallinnan erilliseksi kulueräksi projektinsa budjettiin. (Donnelly 2012, 85.) Suomessa Suomen Akatemia puolestaan katsoi vuonna 2017 tutkimusaineistojen säilyttämisen ja avaamisen kulujen kuuluvan tutkimushankkeen isäntäorganisaation yleiskustannuksiin. Kuitenkin perustelluissa tapauksissa ne oli mahdollista hyväksyä Akatemian tutkimusrahoituksesta katettaviksi tutkimuskuluiksi. (Suomen Akatemia 2017c.) Suomen Akatemia opasti ohjeissaan aineistojen hallinnointi-, luettelointi- ja arkistointipalveluja tarjoavien toimijoiden luo. Sellaisia

olisivat esimerkiksi Tietoarkisto (FSD), joka ottaa vastaan yhteiskunnallisten aineistojen lisäksi humanistisia ja terveystieteellisiä aineistoja, kieliaineistoihin liittynyt FIN-CLARIN-järjestelmä, Avoimen tieteen ja tutkimuksen (ATT) -palvelu sekä kaikille tieteenaloille tarkoitettu European Data Infrastructure (EUDAT) -tallennuspalvelu (Suomen Akatemia 2017c).

Tenopirin et al. artikkelin mukaan yksimielisyys myös tarpeesta saada tutkimusorganisaatioilta rahoitusta aineistohallintaan tutkimusprojektin aikana on lisääntynyt. (Tenopir et al. 2015, 9.) Heidän kyselyaineistonsa mukaan jaksolla 2013–2014 neljännes vastaajista selitti aineistojen jakamatta jättämistä rahoituksen puutteella ja vajaa kuudesosa rahoittajan vaatimusten puutteella. Lisäksi seitsemäsosa piti jakamisen ehtona sitä, että ainakin osa aineistoon liittyvistä kustannuksista, esimerkiksi jakamiskustannuksista, katetaan. (Tenopir et al. 2015, S1 Appendix.) Tosin etenkin vanhempi tutkijapolvi on erottunut nostamalla rahoituksen puutteen keskeiseksi aineistojen jakamisen esteeksi, ja jakamiseen suunnattu lisärahoitus on heille kannustin. Nuorimmat tutkijat eivät ole nähneet rahoitusta suurimpana huolenaiheena. (Tenopir et al. 2015, 10-11; Van den Eynden 2016, 6.) Vaikka rahoituksen puutetta valiteltiin, se ei ollut aivan yhtä yleistä kuin jaksolla 2009–2010. Kehityksen taustalla saattaa olla rahoittajien antamat suuremmat taloudelliset tuet aineistohallintaan, parempi ymmärrys aineistohallinnan ja aineiston jakamisen arvosta tai sen tunnustaminen, että aineiston jakaminen on tarpeellinen osa tutkimuksen elinkaarta, mikäli päämääränä on kehittää tieteellisen tutkimuksen tekemistä. (Tenopir et al. 2015, 18.) Enwaldin tutkimuksessa tiedustelu jakamisen ja säilyttämisen kustannuksien maksajasta puolestaan osoitti, että Suomessa työskentelevistä tutkijoista yli 70 prosentin mukaan kustannukset kuuluisi osoittaa heidän tutkimuslaitokselleen, kahden viidesosan mielestä kansalliselle taholle ja noin kolmanneksen mukaan rahoittajalle. Vain noin kymmenesosan mielestä kustannukset kuuluvat tutkijan tai tutkimusryhmän maksettaviksi. (Enwald 2018, 42.)

### **3.5 Tutkijat aineistohallinnan toimijoina**

Seuraavaksi keskitytään tutkimusaineistojen avoimuuteen ja hyödyntämiseen tutkijoiden näkökulmasta. Ensiksikin kerrotaan, mitä havaintoja tutkimuksissa on tehty aineistojen jakamisesta ja uudelleenkäytön määristä. Toiseksi luodaan silmäys siihen, miten tutkijat suhtautuvat aineistojen työstämiseen jatkokäyttöä varten. Lopuksi käydään läpi, mitä hyötyjä he saavat panostuksista aineistohallintaan. Toisaalta tutkimusten perusteellisuuden ja toisinnettavuuden arvosta käytävä keskustelu

ei ole jäänyt täysin periaatteen ja ihanteen tasolle, vaan todella motivoi tutkijoita toimiin, mutta toisaalta esteitäkin on paljon (Van den Eynden et al. 2016, 4; Alakyynty et al. 2018, 27).

### **Aineistojen jakamis- ja uudelleenkäyttömäärät**

Heidi Enwaldin tutkimuksessa yli kolme neljäsosaa vastaajista olisi halunnut säilyttää tutkimusaineistonsa myös tutkimusprojektin loputtua, vajaa kymmenesosa ei olisi. Tieteenalakohtaisia tilastollisesti merkitseviä eroja ei ollut, mutta esimerkiksi jatko-opiskelijat halusivat säilyttää aineistoja vähemmän kuin muissa asemissa olevat. (Enwald 2018, 40.) Helsingin yliopistossa tutkijat taas näkivät toisaalta yleisesti säilyttämisen arvoisina laajat ja pitkän ajanjakson kattavat niin analogiset kuin digitaalisetkin aineistot sekä kansainvälisten ja kalliisti toteutettujen hankkeiden aineistot. Toisaalta analogiset aineistot katsottiin voitavan hävittää digitoinnin jälkeen. Hävittämiskäytännöissä nähtiin kehitettävää, jotta senkin yhteydessä otetaan huomioon aineistojen arkaluonteisuus. (Alakyynty et al. 2018, 26-27.) Fältin ja Okuloffin tutkimuksen mukaan niin terveystieteilijöillä kuin humanisteillakin aineisto jää tutkimuksen päätyttyä yleisimmin alkuperäisten tutkijoiden haltuun. Kuitenkin siinä missä humanistit hävittivät aineistojaan hyvin harvoin, terveystieteissä neljännes vastaajista nimesi aineistojen hävittämisen vallitsevaksi käytännöksi. Muutosta ei näytä vuosien varrella tapahtuneen. (Fält & Okuloff 2016, 6.)

Kannustus tutkimusaineistojen jakamiseen on siis lisääntynyt periaatteiden tasolla, mutta varsinaisen aineistonjakamistoimintakin, myös arkistointi, on lisääntynyt 2010-luvulla sekä Suomessa että kansainvälisesti. (Tenopir et al. 2015, 1, 8; Fält & Okuloff 2016, 7-8.) Tenopirin et al. mukaan tutkimusjaksolla 2013–2014 vain alle kymmenesosa aineistojen jakamismäärään kantaa ottaneista tutkijoista ei jakanut aineistoaan. (Tenopir et al. 2015, S1 Appendix.) Vaihtelua on, sillä sittemmin Van den Eyndenin et al. tutkimuksessa puolet tutkijoista toi aineistojaan muiden saataville, joko kokonaisina kokonaisuuksina tai osina. Kukin oli jakanut keskimäärin neljä aineistokokonaisuutta viiden edellisvuoden aikana. (Van den Eynden et al. 2016, 4.)

Näkemykset aineiston jakamisesta eroavat selvästi aloittain, mikä ilmeni myös eettisten ja oikeudellisten kysymysten punnitsemisen yhteydessä. Joillakin ihmiskoealoilla ollaan yleisesti haluttomampia jakamaan aineistoja. Kun kyse on laajalle tutkijayleisölle jakamisesta, esimerkiksi lääketieteen/terveystieteiden ja liikealan edustajat ovat ilmaisseet selvästi alhaisempaa halukkuutta jakaa aineistojaan. (Tenopir et al. 2015, 14-15.) Bernadette M. Boscoenkin mukaan tieteenalakohtaisilla normeilla on tärkeä sijansa päätöksentekoprosesseissa, kun arvioidaan, minkä aineistonosan

julkaiseminen on hyväksyttävää. Kuitenkin hän lisää, että vaikka tieteenaloja vertailtaessa esimerkiksi astronomian ala erottuu pitkällä aineistojen jakamisen ja julkistamisen perinteellään, alan käyttäjien tarkempi tutkiminen paljastaa, että senkin sisällä on eroja riippuen esimerkiksi maasta, erikoistumisalasta, projektista, tutkimusryhmästä ja käytetyistä laitteista. (Boscoe 2020, 687, 689; ks. myös Fecher et al. 2015, 13.)

Kun huomio kohdistetaan suomalaistutkijoiden aineistojen saavutettavuuteen, Heidi Enwaldin mukaan suurin osa hänen kyselynsä vastanneista näki käyttämänsä aineiston olevan vapaasti heidän tutkimusryhmänsä käytössä tai erillispyynnöstä muiden tutkijoiden käytettävissä. Etenkin luonnontieteilijöiden, ihmistieteilijöiden ja yhteiskuntatieteilijöiden joukossa oli aineistojaan vapaasti kaikkien käyttöön antavia – tosin yhteiskuntatieteilijöiden keskuudessa oli myös eniten pääsyn aineistoihinsa vain itsellään pitäviä. Vastaajat jakautuivat kahtia aineistojaan huoleti ja mielellään jakaviin ja toista mieltä oleviin; etenkin lääke- ja terveystieteen edustajat kuuluivat jälkimmäiseen joukkoon. Kaikkiaan vajaa kolmannes tutkijoista ei nähnyt aineistojensa jakamisessa ongelmia, vähän vajaa puolet näki. Työrooli tai tieteenala eivät saaneet aikaan tilastollisesti merkitseviä eroja. (Enwald 2018, 36, 42-44.)

Tenopir et al. havaitsivat nuorten tutkijoiden suhtautuvan vanhinta ikäryhmää positiivisemmin ajatukseen aineistojen jakamisesta. Kuitenkin, kun ajatukset olisi pitänyt muuttaa teoiksi, nuoret osallistuivat aineiston jakamiseen vähiten, vanhimmat eniten. Taustalla voi olla, että vanhin tutkijapolvi voi jo haluta jättää työnsä tiedemaailmalle perinnöksi, jonka pitkäikäisyyttä avoimen aineiston kulttuurin edistäminen voi turvata. (Tenopir et al. 2015, 10, 18-19.) Myös Suomessa urakehityksen vaikutus näkyy, sillä jatko-opiskelijoiden parissa kolmannes näki, että muilla ei ollut pääsyä heidän aineistoihinsa, ja he näkivät muita useammin, että oma tutkimusryhmäkään ei saa vapaasti tai pyynnöstäkään käyttää heidän aineistojaan (Enwald 2018, 43.)

Nuorimmat tutkijat näkevät myös muiden tutkijoiden tai instituutioiden aineistojen puutteellisen saatavuuden vanhempia tutkijoita suurempana esteenä sekä tieteen yleiselle kehitykselle että omalle työnteolleen. He myös käyttäisivät vanhempia tutkijoita mieluummin muiden aineistoja, jos ne olisivat helposti saatavilla, mutta olisivat valmiita jakamaan aineistoja itse enemmän, jos saisivat asettaa rajoituksia saatavuudelle. Ilmiössä on nähtävissä, että itsenäisen toiminnan tarve kasvaa iän myötä. Alakohtaisiakin eroja on, sillä esimerkiksi jotkin ihmiskoeaineistoja hyödyntävät alat torjuvat



vahvemmin näkemyksen, jonka mukaan muiden aineistojen tavoittamattomuudesta seuraa isoja esteitä tieteen kehitykselle. (Tenopir et al. 2015, 10, 14, 19.)

Keskeisimpiä aineistojen avoimen jakamisen esteistä saattaa yksinkertaisesti olla, että tutkijat eivät tiedosta sitä arvoa, mikä heidän aineistoillaan voi olla muille tutkijoille – (raaka-)aineiston arvoa alkuperäisen kontekstinsa ulkopuolella pidetään rajallisena – eivätkä he näe tarvetta tehdä aineistostaan julkisia. Tenopirin et al. jakson 2013–2014 kyselyssä neljäsosa piti sitä, että muut eivät tarvitse heidän aineistojaan, syynä jättää aineistojaan jakamatta, kun pari vuotta aiemmin osuus oli ollut kuudesosa. Hieman alle puolet vastaajista taas piti yhteisymmärrystä vastavuoroisesta aineistonjakamisesta ehtona aineistojen jakamiselle, kolmannes torjui sen. (Tenopir et al. 2015, 3-4, S1 Appendix.) Muutamaa vuotta aiemmin oli havaittu, että 85 prosentilla vastaajista oli kiinnostusta käyttää muiden aineistoja, jos ne olisivat helposti saatavilla, mutta vain puolet jakoi aineistojaan ja noin kolmannes piti omia aineistojaan helposti saatavilla olevina. (Tenopir et al. 2011, 9.)

Kun tutkijat päätyvät hyödyntämään valmiita aineistoja, Van den Eyndenin et al. tutkimuksen mukaan tavoitteena on lähinnä taustatiedon ja kontekstin saaminen omaa tutkimusta varten, tutkimuksen validointiin, metodologioiden kehittämiseen ja uusiin analyyseihin. Aineistojen uudelleen käytön määrään vaikutti niin tutkijan uranvaihe, tieteenala kuin tutkimusmetodikin, mutta neljännes heidän kyselynsä vastanneista ei ollut koskaan käyttänyt valmiita aineistoja. Lisäksi havaittiin, että aineistojen vapaa saatavuus ja välitön käsiksi pääsy niihin eivät ole tutkijoille niin merkityksellisiä arvoja kuin se, että aineisto saadaan hyvämaineisesta lähteestä ja korkealaatuisessa ja hyvin dokumentoidussa muodossa. Valmista aineistoa saadaan kollegalta tai suoraan aineiston kerääjältä tai säilytyspaikasta. (Van den Eynden et al. 2016, 5.)

Metodien yhdistäminen voi osaltaan tuoda tutkimusprosessiin sekä aineistojen uudelleenhyödyntämistä että eri aineistokokonaisuuksien yhdistämistä ja uusien kokonaisuuksien rakentamista. Metodien yhdistely voi tarjota syvällisemmän ymmärryksen eri tutkimuskysymyksistä, kuten tuoda uusia työkaluja tai otantoja sekä täydentäviä selvityksiä ilmiöihin. Eri metodien samanaikainen käyttö tarkoittaa usein sekä kvantitatiivisen että kvalitatiivisen aineiston keräämistä. Toisaalta tutkija voi muuttaa laadullisen aineiston numeroiksi tilastollista analyysia varten, toisaalta luoda pohjan kvantitatiiviselle tutkimukselle kvalitatiivisilla metodeilla. Toisaalta on punnittava, käytetäänkö esimerkiksi kvalitatiivista aineistoa määrällisten löytöjen vahvistamiseen vai yksittäisten löytöjen etsimiseen. Metodeihin liittyvänä haasteena on eri tieteenalojen paradigmojen yhteentörmäys. Ensiksikin

jo metodien yhdistely vaatii laajempaa ja perusteellisempaa taustatyötä ja tiedonkeruuta kuin vain yhden metodin käyttäminen, ja toiseksi etenkin erilaisista metodologisista taustoista tulevien aineistokokonaisuuksien sulauttaminen yhteen on monimutkaista. Eri tieteenaloilla onkin toiveita kattavampien viitekehyksien luomisesta metodien yhteensovittamiseksi. Monesti olisi keskeistä, että jos tutkimusryhmässä käytetään erityyppisiä aineistoja, niiden hallintaan tarkoitettut järjestelmät voitaisiin jollakin tavalla sovittaa yhteen. (Davidson 2019, 42-43, 51; Anderson 233-240.)

### **Kokemukset aineiston käsittelystä**

Tenopirin et al. artikkelissa havaittiin kahden eri ajanjakson tutkimustuloksia vertailtaessa merkittäviä muutoksia tutkijoiden tyytyväisyydessä tutkimuksen ja tutkimusaineiston elinkaaren aikana tapahtuviin prosesseihin. Tyytyväisyys omien aineistojen etsimiseen oli vähentynyt, mutta kaikkein tyytymättömmimpiä oltiin prosessiin, joka koski aineistojen tallentamista yli projektin elinkaaren, sekä metadataan liittyviin työkaluihin. (Tenopir et al. 2015, 9.) Helsingin yliopiston tutkijoiden parissa tehty tutkimus 2010-luvun lopulla osoitti, että tutkimusaineiston pitkäaikaissäilytys ei ole kaikille tutkijoille edes määrittelyltään tuttu käsite eikä tutkimusaineiston elinkaariajattelu vaikuta olevan tutkimuskulttuurissa keskeistä. (Alakyynty et al. 2018, 28.)

Vaikka erilaisia vastuunjakoja onkin hahmoteltu, aineistojen avoimuuden toteuttamisen velvoitteet ja vastuut – kuten valinnat ja panostukset, joita aineiston muille ymmärrettäväksi ja käyttökelpoiseksi tekeminen sekä säilytykseen siirtäminen vaativat – ovat 2010-luvulla mietityttäneet tutkijoita sekä Suomessa että maailmalla. (Tenopir et al. 2015, 9; Fält & Okuloff 2016, 14-15; Alakyynty et al. 2018, 27; Bascoe 2020, 688.) Vaikka tutkijat haluaisivatkin edistää aineistojen jatkokäyttöä, esimerkiksi nopea tietotekninen kehitys voi osoittautua ongelmaksi. Kun teknisiin muutoksiin ja aineistojen jakamiseen ei ole suunnitelmallisesti varauduttu, sähköisten aineistojen vioittuneisuus sekä vanhentuneet tallennusmuodot ja sovellusohjelmat muodostuvat avoimuuden esteiksi. Viitteitä on siitä, että muutoksiin varautuminen ei välttämättä ole juurikaan kehittynyt ajanjaksolla 2000-luvulta 2010-luvun lopulle. Aineistojen pitkäaikaissäilytystä varten kaivataan yhä resursseja, niin analogisen aineiston digitoimiseen kuin digitaalisen aineiston siivoamiseenkin, ja tutkijoiden parissa on pelkoa aineistojen tuhoutumisesta. (Borg & Kuula 2007, 35; Alakyynty et al. 2018, 26-28; Enwald 2018, 39.)

Kun tutkijat arvioivat tutkimusaineistojen avoimuuden haittoja, aineistojen jatkokäyttökuntoon saattamisen ongelmat nousevat esiin vuodesta toiseen. Yli puolet Borgin ja Kuulan kyselyn vastaajista piti vähintään melko suurena haittana, jos tutkijoiden aikaa menisi ensiksikin vanhojen

aineistojen muokkaamiseen käyttökelpoisiksi ja myöhemmin kenties jatkokäyttäjien neuvontaan. (Borg & Kuula 2007, 30, 34-35, 70, ks. myös Fält & Okuloff 2016, 14-15 ja Fecher et al. 2015, 11.) Helsingin yliopistossa tutkijat pitivät metadatan muodostamista joko ylimääräisenä työnä, johon ei liikene aikaa, tai jonkun muun kuin tutkijan itsensä tehtävänä, kuten tutkimusavustajalle tai opiskelijalle osoitettavana vastuuna. (Alakyyny et al. 2018, 27-28.) Tenopirin et al. jakson 2013-2014 kyselyssä vajaa kaksi viidesosaa nimesi aineistonjakamattajättämissyyksi ajan puutteen – tieteenala vaikuttaa huomattavasti sen ongelmallisena pitämiseen – vajaa viidennes standardien ja vajaa seitsemäsosa taitojen puutteen. Tosin ajan puutteen merkitys on ollut laskussa. (Tenopir et al. 2015, 8, S1 Appendix.) Toisaalta jatkokäyttäjiltäkin voi kuluu valtavasti aikaa siihen, kun he ovat yhteydessä aineistot tarjonneisiin, milloin sähköpostitse ja puhelimitse, milloin kasvokkain, riittävän metadatan saamiseksi. Sekin koetaan kohtuuttoman tehottomaksi ja kuormittavaksi. (Volk et al. 2014, 886.)

Toisaalta tutkijoiden aika tai motivaatio ei riitä tutkimusaineiston kuvailuun, toisaalta heidän huolensa aineistojen väärinkäytöstä ja väärintulkinnoista on lisääntynyt ilman suurta tieteenalakohontaista vaihtelua. Huolta on siitakin, että omat virheet paljastuisivat aineistoa uudelleen käytettäessä. (Tenopir et al. 2015, 4-3, 7, 16; Fecher et al. 2015, 16-17.) Enwald havaitsi aineiston väärinkäytön ja väärin tulkintojen huolestuttavan jatko-opiskelijoita muita tutkijaryhmiä vähemmän, mikä voi kertoa siitä, että he työskentelevät enemmän itsenäisesti, heille ei ole välttämättä vielä viestitty avoimesta jakamisesta eikä heille ole ehtinyt kertyä aihepiiristä käytännön kokemusta. (Enwald 2018, 40, 43.) Borgin ja Kuulan kyselyyn vastanneista miltei puolet arvioi vähintään melko suureksi haitaksi sen, että heidän aineistojaan ei käytettäisi jatkotutkimuksissa oikein. Ihmistieteissä ongelman nähtiin koskevan niin määrällisten kuin laadullistenkin tutkimusaineistojen, kuten teksti-, kuva- ja äänitiedostojen, tulkintaa. Puuttuvathan aineistoa uudelleen käytettäessä esimerkiksi aineiston keräämisen yhteydessä, inhimillisessä vuorovaikutustilanteessa, tehdyt havainnot. Sekä aineiston keruun että itse aineiston riittävällä kontekstoinnilla olisikin tärkeä rooli ongelman ratkaisemisessa. (Borg & Kuula 2007, 30-31.) Juuri metadatan avulla aineistoa voidaan käyttää ilman jatkuvaa nojaamista sen jakajan tietämykseen, ja nimenomaan metadatatalla tuetaan tutkimuksen läpinäkyvyyttä ja toistettavuutta: tutkijan on ymmärrettävä aineiston alkuperä ymmärtääkseen sen merkittävyyden. *Kuvaileva metadata* liittyy aineiston rakenteeseen, sisältöön, tuottajaan ja sijaintiin – toisin sanoen se vastaa kysymyksiin kuka, mitä, milloin, missä ja miten. *Prosessimetadata* on tietoa esimerkiksi aineiston prosessoinnista ja yhteenvetämisestä ja lisättävissä aineiston yhteyteen vaikkapa kirjallisenä asiakirjana tai koneluettavana tiedostona. (Volk et al. 2014, 887; White et al. 2013, 2-4.) Mikäli

aineisto päätyy arkistoon, arkistokuvailu on yksi niistä arkistollisista hallintatoimenpiteistä, joilla pyritään takaamaan arkistotiedon säilyvyys ja käytettävyys. (Lampela 2016, 8, 6.)

Aineiston keruun ja käsittelyn takana on monesti muu kuin tutkija itse (Borg & Kuula 2007, 31). Kuitenkin on esitetty, että koska tutkija tuntee itse aineistonsa parhaiten, toisinaan paras ja luotettavin ratkaisu olisi, että hän myös kuvailisi aineistonsa itse, yleisesti hyväksytyjä standardeja käyttäen (Haendel et al. 2012, 2; Alakyynty et al. 2018, 28). Tenopirin et al. tutkimuksissa noin puolet vastaajista käytti metadataa. Käytön lisääntyminen vuosien varrella osoittautui vähäiseksi, vaikkakin kehityksessä oli tieteenala- ja standardikohtaisia eroja – tosin Helsingin yliopiston tutkijoiden parissa tehdyssä tutkimuksessa metadatan käyttö oli alasta riippumatta enimmäkseen puutteellista. Joskus yhteisiä ja vakiintuneita standardeja kuvailevalle metadatalle ei ole lainkaan ja tutkijat eivät välttämättä koe osaavansa käyttää tarvittavia aineistohallintajärjestelmiä. Jotta aineiston kuvailun, säilyttämisen ja jakamisen mukanaan tuomasta työstä selviydytään, tutkijoita pitää siis joko kouluttaa tai käyttää apuna aineistohallinnan ammattilaisia. Myös helppokäyttöisten aineistohallintatyökalujen lisäkehittelylle on ollut tilausta. (Tenopir et al. 2015, 3, 9, 15, 17-18, 20; Van den Eynden et al. 2016, 4; Alakyynty et al. 2018, 28.) Toisaalta on havaittu, että metastandardeista ei aina tarjota riittävästi tietoa korkeakouluopintojen aikana. Esimerkiksi tietokantoja, tiedostoformaatteja ja aineistohallinnan periaatteita käsittelevä kurssi voisi lisätä paitsi metadatataitoja, myös aineistojen jakamista ja uusien käyttötarkoitusten löytämistä eri aineistoille. (Volk et al. 2014, 887, 892.)

Eriyiset metadastandardit tuovat yhdenmukaisuutta ja automaattisuutta aineistojen käsittelyyn, mutta tärkeintä on, että metadataa yleensä on liitetty aineistoihin. Hyvän metadatan kirjaaminen ei välttämättä vaadi suurta ajallista uhrausta. Helpoin tapa luoda metadataa on alkaa kuvailla aineistoa jo aineiston suunnittelu- ja keruuvaiheissa, mikä auttaa tutkijoita pysymään organisoituneena laajemminkin, helpottaa aineiston parissa työskentelyä keruun jälkeen ja lopulta tekee myös aineiston julkaisemisesta helpompaa. (White et al. 2013, 2-4.) Tutkijat ovat tottuneet toimimaan ad hoc -hengessä aineistojensa hallinnassa eli ratkomaan tilanteita sitä mukaa, kun ongelmia ja kysymyksiä ilmaantuu, mikä voimistaa tunnetta aineistohallinnasta ylimääräisenä taakkana (Donnelly 2012, 85; Volk et al. 2014, 892.) Kuitenkin Ethan P. White et al. muistuttavat, että hyvät aineiston parissa työskentelyn käytännöt ovat päämäärä sinänsä riippumatta esimerkiksi siitä, onko tutkija aikeissa jakaa aineistoaan. Hyvin dokumentoitu ja standardisoitu aineisto voi auttaa myös aineiston alkuperäistä kerääjää käyttämään ja uudelleenkäyttämään aineistoa, jos hän itse palaa aineiston äärelle kuukausia tai vuosia sen alkuperäisen keruun tai analyysin jälkeen. (White et al. 2013, 8.)

## Aineiston kerääjälle kaavailut urahyödyt käytännössä

Tutkijoiden kokemus aineiston jakamisen vähäisestä palkitsevuudesta saa tutkimuksissa huomiota. (Fecher et al. 2015, 11-12.) Pelkkä jakamisen helpottaminen ei edistä itse jakamista, vaan tutkijoille pitää tarjota myös ”sisäisiä” kannustimia. Tutkijoiden yleisesti mainitsema ja asemaltaan vankistuva huoli on ollut, että toinen tutkija pääsisi hyötymään aineistosta ennen alkuperäistä kerääjää. Myös kilpaileviin instituutioihin suhtaudutaan varauksella. Monilla tieteenaloilla tutkimusaineisto halutaan pitää käytettävissä pitkäaikaissäilytykseen siirrosta huolimatta eli tutkijat voivat pitää aineistoja itsellekin relevanttina pitkiä aikoja. (Borg & Kuula 2007, 70; Tenopir et al. 2015, 8; Fecher et al. 2015, 16-17; Alakyyry et al. 2018, 28.)

Tenopirin et al. kyselyn vastaajista yli kaksi viidesosaa halusi julkistaa omat tuloksensa ennen aineistojen jakamista. Tosin tarve oli keskimääräistä voimakkaampi biologian ja fysiikan aloilla ja selvästi merkityksettömämpi esimerkiksi tietojenkäsittelyn ja informaatiotieteiden parissa. (Tenopir et al. 2015, 14-16, S1 Appendix.) Enwaldin tutkimuksessa tieteellisen etumatkan menettämistä pelkääviä oli etenkin luonnontieteitä edustavien vastaajien parissa, joista hieman yli kolmannes ilmaisi kyseisen huolen. (Enwald 2018, 40.) Boscoe selvitti astronomien parissa, mitä tutkijat pitävät törkeänä aineistonormien rikkomisena, ja vastaus koski häikäilemätöntä kilpailua. (Boscoe 2020, 692.)<sup>6</sup>

Merkittävimpiä aineistojen jakamista ohjaavia tekijöitä on myös tutkijan halu saada tunnustusta aineistojensa jakamisesta. (Tenopir et al. 2015, 7-8.) Esimerkiksi Tenopirin et al. kyselyssä yli neljä viidesosaa halusi aineistojen uudelleenkäytön ehdoksi viittaamisen aineiston kerääjään tai muun tähän kohdistuvan tunnustuksen. Rungas kolmannes halusi kerääjälle jaetun tekijyyden tutkimukseen, joissa aineistoa käytetään, sekä mahdollisuuden kommentoida ja arvioida keräämiinsä aineistoihin perustuvia tuloksia. (Tenopir et al. 2015, S1 Appendix.) Alaluvussa 3.2 esitellyt uudet viittaus- ja julkaisujärjestelmät motivoisivat erityisesti uransa alussa olevia, sillä nuorimmat tutkijat painottavat aineistoihin liittyvän määräysvallan ja kunnian merkitystä muita ikäryhmiä voimakkaammin.

---

<sup>6</sup> Tutkijat ilmaisivat tieteellisten käytäntöjen mahdollistavan tilanteen, jossa tutkimusryhmä sijoittaa keräämänsä aineiston esimerkiksi arkistoon ja laatii sen pohjalta artikkelia. Arkistoitu aineisto on suojattu embargo-ajalla, mutta aika voi päättyä ennen artikkelin valmistumista. Toinen ryhmä voi hyödyntää aineistoa ja julkaista tulokset omassa artikkelissaan ennen alkuperäisen tutkimusryhmän artikkelia. Tutkijat totesivat, että asetelmalle ei voida mitään, mutta toisaalta avautuneeseen tilaisuuteen tarttunut tutkimusryhmä voi saada huonon maineen tiedeyhteisössä. (Boscoe 2020, 692.)

Kyseistä aineistohallintalinjaa selittää nykyiseen akateemiseen järjestelmään syvälle iskostunut kilpailuun kannustava publish or perish -periaate. (Tenopir et al. 2015, 19; Van den Eynden 2016, 5-6.)

Van den Eyndenin et al. tutkimuksessa yhteistyömahdollisuuksien potentiaalinen lisääntyminen nähtiin keskeisenä aineistojen avaamiseen motivoivana tekijänä. Osalla, vaikkakin vain pienellä joukolla, oli kokemustakin jakamisesta seuranneesta yhteistyöstä. (Van den Eynden et al. 2016, 4.) Tenopirin et al. tutkimuksessa noin kolme viidesosaa halusi mahdollisuuden tehdä yhteistyötä projektin kanssa, jossa käytetään hänen aineistojaan. Rungas kaksi viidesosaa halusi itselleen kopiot artikkeleista, jotka pohjautuvat hänen keräämiinsä aineistoihin, samoin kuin listan kaikista tuotoksista, joissa aineistoa on käytetty, sekä hyväksyttäväkseen lausunnon käyttötarkoituksista, joihin hänen aineistojaan käytetään. Vain kolmannes halusi, että aineistoon perustuvia tuloksia ei jaeta ilman aineiston kerääjän hyväksyntää. (Tenopir et al. 2015, S1 Appendix.) Nuoret tutkijat arvostavat muita ikäryhmiä enemmän mahdollisuutta solmia yhteistyösuhteita aineiston jakamisen yhteydessä ja saada kommentoida sitä, mitä muut ovat heidän aineistoillaan tehneet. (Tenopir et al. 2015, 11.)

Aineiston avaamisen riskeinä pidettyjen tekijöiden määrän on havaittu lisääntyneen, ja tietyt esteet tuntuvat myös säilyvän eivätkä näkemykset riskeistä juurikaan vaihtele tieteenalakohtaisesti. (Tenopir et al. 2015, 1, 16.) Fält & Okuloff antoivat kyselyssään kuuden vaihtoehdon listan koskien jatkokäytön mahdollisia haittoja, ja vähintään puolet vastaajista koki kunkin niistä haitaksi. (Fält & Okuloff 2016, 14-15.) Aiemmin Borgin ja Kuulan kyselyssä vastaajista lähemmäs puolet piti isoa osaa avoimuuden haittoiksi esitetystä vaihtoehdoista todellisina. Suoranaisia esteitä jakamiselle oli vastaajien mielestä useita niin, että mikään yksittäinen este ei erottunut ylitse muiden. (Borg & Kuula 2007, 30.) Tosin tutkijoiden huolten on arvioitu olevan sikäli turhia, että hyvin harva kertoo saaneensa huonoja kokemuksia jakamisesta. Toisaalta suurin osa ei ole myöskään kokenut saaneensa käytännössä suoria hyötyjä tutkimusaineistojensa jakamisesta, mutta silti aineistojen avoimuuteen liitetyt edut ovat painaneet suurimmalla osalla vaakakupissa haittoja enemmän. (Borg & Kuula 2007, 41; Van den Eynden et al. 2016, 4.) Kuitenkin, jos tietyillä aineistohallinnan osa-alueilla ei tapahdu muutosta, voi olla syytä miettiä nykyisten aineistohallintapolitiikkojen, -suunnitelmien ja -infrastruktuurien tehokkuutta. Toisaalta on mahdollista, että on saavutettu piste, jonka jälkeen muutosta ei voikaan saada aikaan. Toisaalta on mahdollista, että institutionaaliset vallitsevat puitteet ovat niin vankkoja, että muutosta ei saada aikaiseksi. (Tenopir et al. 2015, 18.)

## 4 TUTKIMUSASETELMA JA -MENETELMÄ

Aiemmissa luvuissa kävi siis ilmi, että tutkimusaineiston elinkaareissa ja tutkimuksen elinkaareissa on monia vaiheita, joihin liittyy aineistojen keräämistä, muokkaamista, siirtämistä ja jakamista. Tutkijat ovat enemmän tai vähemmän halukkaita ryhtymään toimiin, ja heitä ja heidän taustaorganisaatioitaan ohjaavat osin eri intressit. Tässä käyn läpi, miten itse lähestyin aineistohallintaa tamperelaisien tutkijoiden näkökulmasta. Metodini oli määrällinen kyselytutkimus. Määrällisessä kyselyssä on kuusi pääelementtiä: tavoitteenasetanta, tutkimuskysymykset, kysely, otos, aineistonkeruustrategia sekä aineistonanalysointistrategia (Punch 2003, luku 2.4). Alaluvussa 4.1 esittelen tutkimuskysymykset, alaluvussa 4.2 metodini ja alaluvussa 4.3 kyselyn toteutuksen. Viimeisessä alaluvussa 4.4 on vuorossa kyselyaineiston analysoinnin esittely.

### 4.1 Tutkimuskysymykset

Tampere3-hankkeen yhdeksi tavoitteeksi nimettiin kaikkien tutkimuksen tasojen yhdistäminen. Tutkimusaineistojen hallinta on siinä keskeistä, kun otetaan huomioon se sija, jonka tutkimuskirjallisuus antoi aineistoille tutkimuksen elinkaareissa. Monipuolinen käsitys tamperelaistutkijoiden aineistohallinnan näkemyksistä ja käytännöistä auttaa ymmärtämään, millainen lähtökohta korkeakoulujen yhteiselle tutkimusaineistojen hallinnalle on. Rakensin kuvaa yhdistyvien korkeakoulujen tutkijoiden aineistohallinnasta seuraavien tutkimuskysymysten avulla:

1. Miten tamperelaistutkijat suhtautuvat tutkimusaineiston hallinnan eri rooleissa (ks. alaluku 4.4) aineistojen avoimuuden toteuttamiseen sekä aineistojen hyödyntämiseen?
2. Miten tamperelaistutkijoiden näkemykset koskien tutkimusaineistojen hallintaa eroavat heidän käytännöistään?
3. Vaikuttaako tamperelaistutkijoiden edustama korkeakoulu heidän näkemyksiinsä ja käytäntöihinsä koskien tutkimusaineistojen hallintaa? Mitä korkeakoulujen väliset eroavaisuudet tai yhteneväisyydet ovat?

Näkökulmana aineistohallintaan on siis tutkimusaineistojen avoimuuden toteutuminen, aineistojen hyödyntäminen ja tutkijakeskeisyys. Avoimuudella viitataan tässä jaetun aineiston määriin ja jakamisen esteisiin ja ehtoihin; hyödyntämisellä sekä vastaajien omien että toisten aineistojen käsittelyyn, käyttöön ja uudelleenkäyttöön. Tutkimustulokset tukevat osaltaan aineistohallinnan kehittämistä – kuten alan koulutuksen kohdentamista sekä aineistojen monipuolisempaa ja pitkäaikaisempaa hyödyntämistä – ja siten tutkimustyön tehostamista uudessa korkeakoulu yhteisössä.

Tutkimuskysymykset sivuavat tutkijoiden näkemyksiin ja käytäntöihin keskittymisen osalta Carol Tenopirin et al. vuoden 2015 artikkelia, sillä aineiston keruussa käyttämäni kyselylomake nojasi vahvasti kyseisessä tutkimuksessa käytettyyn lomakkeeseen. Tässä käytetyssä lomakkeessa oli kolme osiota: ensimmäisessä pyydettiin tutkijan taustatietoja, toisessa esitettiin tutkijan omaan tutkimukseen liittyviä kysymyksiä ja kolmannessa huomio oli tutkimusaineistojen uudelleenkäytössä. Tenopirin et al. tutkimus vaikutti aineiston keruuseen, mutta ei tapaan, jolla analysoin aineistoni. Esitän tulokseni kuusiosaisen viitekehysten kautta. Viitekehys osa-alueineen jäseni sitä, millaisiin konteksteihin aineistohallinnan näkemykset ja käytännöt liittyivät. Kuten johdannossa kuvailin, tavoitteenani on korostaa tutkijoiden toimijuutta aineistohallinnassa. Siksi yhdistin viitekehysten osa-alueet aineistohallinnan eri rooleihin, joiksi määrittelin aineiston kerääjän, aineiston suojelijan, tiedeyhteisön jäsenen, aineiston uudelleenkäyttäjän ja uudelleenkäytön valvojan sekä aineiston käsitelijän ja aineistohallinnan tuen saajan. Esittelen viitekehystä tarkemmin alaluvussa 4.4.

## **4.2 Määrällinen kyselytutkimus**

Tutkielmani rakentuu aineiston suhteen määrällisen kyselyn varaan. Siinä missä tutkijoiden aineistohallintaa on lähestytty seurantatutkimuksenkin kautta – vähintäänkin on vertailtu ajanjaksoja, jos ei samaa tutkijajoukkoa – mikä näkyi kirjallisuuskatsauksessa, tarkastellessani kolmen yhdistymässä olevan korkeakoulun tutkijoiden aineistohallinnan eroja keskityn yhteen tiettyyn hetkeen ja tiettyyn tapaukseen. Määrällisessä kyselyssä ei ole tarkoitus manipuloida tarkasteltavia muuttujia tai käsitellä niitä tutkimuksellisiin tarkoituksiin. (Punch 2003, luku 2.4.) Kun lähetin kyselylomakkeen vastaajajoukolle, en voinut tutkijana kontrolloida tekijöitä, jotka vaikuttivat vastaajien tapoihin vastata kysymyksiin. Pystyin vain muotoilemaan kysymyksistä mahdollisimman yksiselitteisiä.

Havainto- eli tilastoyksikkö on perusjoukon yksikkö, johon tutkimus kohdistuu. Tutkimuksen otoksen koko ilmastaan usein havaintoyksikköjen määränä. Tässä tutkielmassa havaintoyksikkönä on esimerkiksi yksittäinen tutkija. Havaintoaineisto taas muodostuu havaintoyksikköiden tiedoista. Kyseiset havaintotiedot ovat vastauksia tutkimuskysymyksiin tai mielipiteitä tutkittavista asioista, ja niitä kutsutaan muuttujien arvoiksi. Muuttuja taas on mikä tahansa mitattava ominaisuus tai muu suure, jonka arvot vaihtelevat. Esimerkiksi tässä niin tutkijan korkeakoulu, tutkimusala kuin tyytyväisyys aineistohallintaan ovat muuttujia; ne tekevät eroja yksiköiden välille. (Punch 2003, luku 1.3; Heikkilä 2014, 33, 13, 280; Tähtinen et al. 2020, 16.) Kyselytutkimuksen logiikkana on tutkia, miten yksikön kohdalla näkyy vaihteluita eri muuttujien kohdalla. (Punch 2003, luku 1.1.)



Kun aineiston, kyselyn tulosten, analysoimiseksi valitaan sopivaa tilastollista testiä, tarkasteltavien muuttujien on täytettävä testin vaatimat edellytykset. Testin valintaan vaikuttaa otoskoko, vertailtavien ryhmien lukumäärä, käytetty mitta-asteikko, onko muuttuja jatkuva vai epäjatkuva, miten muuttuja jakautuu (esimerkiksi Gaussin käyrä), ovatko otokset toisistaan riippumattomia ja ovatko havainnot toisistaan riippumattomia. (Heikkilä 2014, 99, 183.) Määrällisessä tutkimuksessa ollaan kiinnostuneita muuttujien välisistä suhteista pikemminkin kuin vain muuttujien kuvailusta. Yksi tyyppinen käsitteellinen viitekehys muuttujien hahmottamiselle on kausaalisuus. Muuttujien välille mietitään lähes automaattisesti mahdollisia syy-seuraussuhteita. ”Syy” ja ”seurauksen” sijaan empiirisessä tutkimuksessa käytetään termejä ”riippuva” ja ”riippumaton”. (Punch 2003, luku 1.3; Heikkilä 2014, 229.)

Kyselyssäni hyödynnettiin mitta-asteikoista ensiksikin *luokittelu- eli nominaaliasteikkoa*, kun kysyttiin esimerkiksi vastaajan korkeakoulua, ammattiasemaa ja käyttämiä aineistotyyppejä. Kyseiseen asteikkoon kuuluvien muuttujien arvoista voidaan todeta vain, mihin luokkaan ne kuuluvat. Luokkia ei voida asettaa mitattavan ominaisuuden mukaan järjestykseen eikä arvoilla voi tehdä laskutoimituksia. *Järjestys- eli ordinaaliasteikkoa* käytettiin vaikkapa kysyttäessä, missä määrin vastaaja on samaa mieltä lomakkeella esitettyjen väittämien kanssa tai paljonko hän jakaa aineistojaan. Sitä käytettäessä muuttujien arvot ovat laitettavissa järjestykseen, joskaan arvot eivät välttämättä ole tasavälein. Vastaajan syntymävuoden tiedustelu voidaan taas tulkita *välimatka- eli intervalliasteikon* käyttämiseksi. (Heikkilä 2014, 81-82.)

Muuttujat ovat luokiteltavissa sisällöllisen ja teknisen jaottelun mukaan. Teknisessä jaottelussa ne jaetaan kolmeen tyyppiin: *jatkuvat, kvantitatiiviset muuttujat* (välimatka-asteikolliset), *diskreetit, kvalitatiiviset muuttujat* (luokittelu- ja järjestysasteikolliset) ja *dikotomiset binäärimuuttujat* (vastausvaihtoehtoina ”kyllä” ja ”ei”). Sisällöllisessä jaottelussa on taas kyse jaosta *tulosmuuttujiin* (selitettävät, riippuvat muuttujat, Y-muuttujat) ja *selittäviin muuttujiin* (riippumattomat muuttujat, X-muuttujat). Jos sisällöllistä jaottelua ei tehdä, muuttujien voi sanoa olevan kausaalisesti samantasoisia; jos jako tehdään, muuttujien voi sanoa olevan kausaalisesti eritasoisia. (Heikkilä 2014, 229.) Koska yhtenä tämän tutkielman tutkimuskysymyksenä on korkeakoulujen välisten yhtäläisyyksien ja erojen selvittäminen, selittävänä muuttujana on ennen kaikkea vastaajien korkeakoulutausta. Yksittäisissä tilanteissa selittävä muuttuja voi olla lisäksi jokin muu, kuten aineiston jakamismäärä.

### 4.3 Kyselyn toteutus

Kyselyn toteuttamisen metodina oli itsetäytettävä kyselylomake, joten en tavannut vastaajia kasvokkain. Aineistonkeruustrategia voi vaikuttaa vahvasti aineiston laatuun, joka taas vaikuttaa suuresti havaintojen uskottavuuteen. Mitä enemmän aineiston keruuseen ja sen hallintaan panostetaan, sitä parempilaatuista aineisto on. Strategiassa teknisluonteisina huomionkohteina olivat kyselyn reliabiliteetti eli luotettavuus sekä validiteetti eli kelpoisuus. Jos samat henkilöt vastaisivat uudella kyselykierroksella kysymyksiin samalla tavalla kuin ensimmäisellä kerralla, aineisto olisi hyvin luotettava. Kelpoisuudessa taas on kyse siitä, edustaako aineisto sitä mitä sen ajatellaan edustavan, eli mittaavatko saadut vastaukset todella niitä muuttujia, joita tutkijana ajattelin niiden mittaavan. Lisäksi oli huomioitava ei-tekniset teemat eli suunniteltava, miten vastausprosenttiin ja vastaajan asenteeseen voi vaikuttaa. (Punch 2003, luku 3.5.) Seuraavassa esittelen kyselylomakkeen laatimis- ja levittämisen prosessin. Tutkimusprosessin näkökulmasta katsottuna tutkimuskysymykset ohjaavat lomaketta – ne antavat yhdessä listan muuttujista, joita lomakkeen tarvitsee mitata – kun taas itse lomake on aineiston keruun työväline. Esillä on pidettävä paitsi lomakkeen ja tutkimuskysymysten välinen suhde ja muuttujien määritelmät, joita jo sivuttiin, myös kyselylomakkeen alkuperä, lomakkeen laatimisvaiheen rajoitteet ja lomakkeen testaaminen. (Punch 2003, luku 3.3.)

#### Kyselylomake

Koska tutkielmassani on kyse aineistojen uudelleen käyttämisestä, teemaan sopivaa oli hyödyntää aiemmin käytettyä kyselylomaketta. Kyselyni pohjana oli Carol Tenopirin et al. tutkimuksessa käytetty lomake, jonka he jakoivat vapaasti käytettäväksi artikkelinsa ”Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide” yhteydessä (Tenopir et al 2015, S2 Appendix). Kyselytutkimusta on tehty paljon monissa maissa, joten erinomaisia kyselylomakkeita on jo olemassa. Suuriin ongelmiin kuuluu niiden löytäminen. Keith F. Punch huomauttaa kyselytutkimuksen laatimisprosessiin kuuluvankin juuri sen pohtiminen, luoko tutkija itse lomakkeen erityisesti omaa tutkimustaan varten vai onko jokin valmis tutkimusinstrumentti tai jokin yhdistelmä paras. Usein menestyksekkäs kysely perustuu lomakkeeseen, joka on sekoitus olemassa olevia mutta ehkä muokattuja osia sekä erityisesti kehitettyjä osia. Myös mittaamistavat tuovat haasteita. Hyvien mittausinstrumenttien laatiminen on työlästä, riippuen eri muuttujien määrästä ja hyödynnettävien vastauskaalojen laajuudesta, joten valmiiden instrumenttien vähintäänkin jonkinasteiseen hyödyntämiseen tulisi jopa kannustaa. Silti kyselylomakkeen sovittaminen paikallisiin oloihin on

usein tarpeen, mutta se ei aina ole suoraviivaista, sillä monikulttuurisessa tutkimuksessa kohdataan jo käännöshaasteitakin. (Punch 2003, luku 3.3.)

Kun muokkasin kyselylomaketta oman tutkielmani tarpeisiin sopivaksi, huomio oli erityisesti lomakkeen ytimekkyydessä, ymmärrettävyydessä ja käännöksen sujuvuudessa. Kun käytössä oli jo aiemmin toteutetussa kyselyssä käytetty lomake, ennakkotestaamisen rooli oli vähäisempi, mutta useamman silmäparin turvin varmistettiin kysymysten toimivuus, sekä se, että kyselyn vastaaminen ei vienyt aikaa noin varttia enempää. Jos vastaaminen vie kauemmin kuin 20–30 minuuttia, se vaikuttaa merkittävästi vastaajien sitoutumiseen ja siten vastausprosenttiin ja edelleen kyselystä saatavan aineiston laatuun (Punch 2003, luku 3.3). Kyselylomakkeen toteutuksessa käytin Googlen kehittämää Google Forms -sovellusta, sillä sen havaittiin mahdollistavan kaikki lomakkeelta vaaditut toiminnot, kuten monivalinnat, ja se oli helppokäyttöinen ja ilmainen vaihtoehto. Valmiin lomakkeen sisältö on tutustuttavissa liitteessä 1.

Tenopirin et al. lomake oli liian pitkä ja raskas pro gradu -tutkielmassa sellaisenaan käytettäväksi. Kysymyksiä oli 21 ja vastausvaihtoehdot olivat usein pitkiäkin virkkeitä. Suuri osa muokkauksistani liittyi karsimiseen ja tiivistämiseen kyselyn keventämiseksi. Lyhyt vastaamaan houkutteleva kysely on parempi kuin liialla pituudellaan vastausprosentin ja siten validiteetin vaarantava. (Punch 2003, luku 3.3.) Toisaalta karsimiseen vaikutti jonkin verran aikataulullisetkin syyt, joiden vuoksi kaikkiin alkuperäisen lomakkeen kohtiin ei ollut mahdollista etsiä yhtä kattavia ja perusteltuja suomalaisia vastineita vastausvaihtoehdoiksi. Valmiiden vaihtoehtojen korvaaminen avovastausvaihtoehdolla taas olisi sekä työllistänyt vastaajia että laajentanut analyysiä liiaksi. Toisaalta karsimisella järjestettiin tilaa sille, että pystyin kyselyä liikaa pidentämättä lisäämään Tampereen oloja kartoittavia vastausvaihtoehtoja, kuten tiedustelun siitä, oliko tutkija käyttänyt naapurikorkeakoulujen aineistoja.

Omassa kyselylomakeversiossani oli 17 kysymystä. Merkittävin muutos oli, että poistin Tenopirin et al. tutkimukseen kuuluneen tutkijoiden organisaatioihin keskittyneen osuuden. Ratkaisua puolsi jo se, että Tampere3-organisaatioiden aineistonhallintakäytännöistä voisi saada tietoa esimerkiksi organisaatioiden dokumenteista – organisaatioita oli tässä vain kolme siinä missä Tenopirin et al. kysely kohdistettiin paljon laajemmalle vastaajajoukolle – joten ei ollut mielekäästä rasittaa tutkijoita osiolla. Toiseksi sen poisto ei vaikuttanut tutkielmani tavoitteenasetantaan: lomakkeen on vastattava tutkimuskysymyksiä ja jäljelle jätettiin juuri tutkijakohtaisia näkemyksiä ja käytäntöjä selvittäneet osat.

Käänsin alkuperäisen lomakkeen englannista suomeksi ja muotoilin vastausvaihtoehtoja Tampereen oloihin sopiviksi esimerkiksi ammattinimikkeiden ja tieteenalavalintojen osalta, missä hyödynsin jonkin verran suomalaisissa kyselyissä käytettyä sanastoa ja vaihtoehtoja. Käännösratkaisujen yhteydessä pohditutti esimerkiksi kysymys, puhuako aineiston saatavuuden kohdalla ”pääsyn estämisestä” vai ”pääsyn puuttumisesta” – valitsin neutraalimmalta kuulostavan jälkimmäisen ilmaisun. Lisäksi käytin lomakkeessa kysymystasolla täsmällisempää ilmaisua ”tutkimusaineisto” pelkän ”aineiston” sijaan. Otin myös vapauksia tyyliellä ilmaisua. Alkuperäisessä lomakkeessa puhutaan esimerkiksi ”instituutioista” ja ”osastoista”. Liitin ne yhteen muotoon ”organisaatiot ja yksiköt”. Samoin yhdistin ”metadatan” ja ”dokumentoinnin”, joista kyseltiin alun perin erikseen. Metadata ja dokumentointi eivät ole toistensa synonyymejä, mutta koska aineiston kuvailu ja järjestäminen ovat vain yksi tutkielmani teemoista, nostin tiivistämisen etusijalle.

Kyselylomaketta muotoiltaessa yhtenä yleisperiaatteena on kysymysten ja vastausvaihtoehtojen helppo ymmärrettävyys eli niiden pitäminen mahdollisimman niin lyhyinä ja selväsanaisina. Olennaista olisi varmistaa, että kukin vaihtoehto ja kysymys sisältää vain yhden ajatuksen, jotta vastaajan ei tarvitse pohtia niitä kauan eikä ajautua ”se riippuu” -tyyliseen ajatteluun. (Punch 2003, luku 4.2.) Lomakkeessani vastausskaala oli enintään kuusiosainen ääripäiden ollessa ”vahvasti samaa mieltä” ja ”vahvasti eri mieltä”. Kun skaala on laaja, vaihtelevuutta saadaan todennäköisesti enemmän, mikä on toivottavaa. Samalla on silti panostettava luotettavuuteen ja merkityksellisyyteen eli siihen, että vastaajat voivat johdonmukaisesti ja itsevarmasti valita heille sopivan vaihtoehdon. Sisäisesti johdonmukaiset ja monikohtaiset arviointiskaalat mahdollistavat määrällisessä tutkimuksessa järjestelmällisen vertailun. (Punch 2003, luku 4.2.) Esimerkiksi lomakkeessani oli monessa kohdassa mahdollista valita ”en osaa sanoa” tai ”en ole varma”, jolloin vastaajaa ei painostettu tekemään valintaa. Toisaalta vastaaminen perustui vapaaehtoisuuteen niin, että kysymyksen saattoi halutesaan ohittaa, mikä oli Tenopiriltä et al. otettu linja. Tosin riskiksi jäi, että johonkin kohtaan ei huomata tai unohdetaan vastata.

Kun kysymysten toimivuus oli tarkistettu ja lomake hiottu mahdollisimman toimivaksi ulkoasultaan ja pituudeltaan, oli aika alkaa panostaa kyselyn onnistuneeseen levittämiseen eli laatia saateviesti ja vastaamisohjeet sekä suunnitella kyselyn esille saattamista muistutuksineen. Sekä vastaajan asenteeseen että vastausprosenttiin vaikuttaminen vaativat suunnittelua. Vastaajien lähestymisen on oltava ammattimaista ja osallistumaan houkuttelevaa. (Punch 2003, luku 3.5.) Koska käytin itse täytettävää kyselylomaketta, saatekirjeessä oli olennaista välittää paitsi riittävät etenemisohjeet,

myös tiedot siitä, mistä tutkimuksessa oli kyse, kuka sen suorittaa ja miksi, miten vastaajat on valittu ja mitä kerätylle informaatiolle tehdään. Kyselyyni vastaaminen oli myös anonyymiä. (Punch 2003, luku 4.4.) Saatekirjeeni on luettavissa liitteestä 2 a.

## **Vastaajien tavoittaminen**

Itsetäytettäviin kyselylomakkeisiin pohjautuvien tutkimusten logiikkana on vastata tutkimuskysymyksiin keräämällä tietoa joltakin ihmisjoukolta. Siksi olennainen pohdittava on, miltä joukolta aineistoa kerätään, miten monta vastaajaa on ja miten heidät valitaan. ”Otos” on tutkimuksessa tekninen termi. Se tarkoittaa pienempää osajoukkoa jostakin isommasta ryhmästä. Tyypillinen tilanne on, että tietystä ihmispopulaatiosta halutaan saada jotakin selville, mutta siitä on mahdollista tutkia vain otosta. Tällöin on punnittava, missä määrin otos esittää koko populaatiota eli perusjoukkoa. Otoksesta tehdyt havainnot halutaan usein yleistää otoksesta takaisin populaatioon, mistä juontaa edustavuus- ja yleistettävyyksymysten tärkeys. Otos on valittava mahdollisuuksien mukaan niin, että kaikilla potentiaalisilla suhteilla, joita muuttujien välillä on, on maksimoitu mahdollisuus tulla havaituksi. Otannan takana olevan strategian olisi noudatettava tutkimuksen ja sen tutkimuskysymysten yleistä logiikkaa. (Punch 2003, luku 3.4.)

Kohteenani on jo tutkielmani nimen mukaisesti tamperelaistutkijat, jolla tarkoitan Tampereen yhdistyissä korkeakouluissa tutkimusta tekeviä henkilöitä. Lähtökohtanani oli kohdistaa kyselyni kolmelle organisaatiolle tasapuolisesti, sillä yhtenä tutkimuskysymyksenäni oli tutkia korkeakoulutautan vaikutusta tutkijoiden vastauksiin. Kohdistin kyselyn nimenomaan tutkijoille, sillä olin rajannut tutkielmani käsittelemään *tutkimusaineistoja* – Tenopirin et al. tutkimuksessa oli huomioitu jossain määrin myös esimerkiksi hallinnon työntekijät. Täsmällisempiä osajoukkoja en kyselyäni kohdentaessani etsinyt. En esimerkiksi karsinut tiettyjä aloja pois, kuten Okuloff ja Fält tekivät vaikkapa biotieteiden suhteen. En tehnyt rajauksia myöskään tutkijan aseman perusteella, siinä missä Kuula ja Borg kohdistivat kyselynsä professoreille. Tampereen korkeakoulut eroavat sekä myöntämiensä tutkintojen että kattamiensa tutkimusalojen osalta, joten toiveissa oli saada esille eri tutkimusalojen näkemyksiä mahdollisimman monipuolisesti – toisaalta tiedostettiin, että sama ala voi yhdistää tutkijoita enemmän kuin sama työpaikka, mutta toisaalta, kuten kirjallisuuskatsauksessa kävi ilmi, myös yksittäisten alojen sisällä voi olla paljon eroavaisuuksia. Otantamenetelmänä oli yksinkertainen satunnaisotanta, jota kutsutaan myös umpimähkäiseksi otannaksi. Siinä jokaisella populaation yksiköllä on yhtä suuri todennäköisyys päätyä otokseen. (Heikkilä 2014, 34.)

Monesti, kun hankitaan aineistoa pienen mittakaavan tutkimukseen, lopputuloksena on jonkinlainen mukavuustekijöihin perustuva suunta tai tietyn joukon luo ajautuminen. Opinnäytetyövaiheelle ominaisesti minullakin oli rajallinen aikataulu, resurssit ja pääsy ihmisten luo. Keith F. Punch huomauttaa, että opiskelijan on usein tyydyttävä pieneen otokseen, jota voidaan hallita projektin puitteissa ja joka on saavutettavissa, mutta muistuttaa, että kuitenkin mille tahansa kyselylle on olemassa teoreettisesti ideaali otos; on mietittävä, minkä otoksen valitsisi ideaalitulanteessa, resurssien ja saatavuuksien ollessa optimit. (Punch 2003.) Esimerkiksi tässä työssä olisin ihannetilanteessa tavoittanut *jokaisen* Tampereen korkeakoulussa työskentelevän tutkimustyötä tekevän henkilön; nyt lähtökohtana oli tavoittaa *mahdollisimman paljon* tutkijoita kolmesta tamperelaisesta korkeakoulusta. Toisaalta myös lopulta tavoitettu otos vaikuttaa siihen, millaiseksi tutkimusasetelma viime kädessä muotoutuu eli millainen analyysi on mahdollista tehdä.

Aloitin vastaajajoukon tavoittamisen suunnittelun lähettämällä kyselylomakkeen levittämistä koskeneen tiedustelun Tampereen yliopiston intranet-sivuilla keväällä 2018 tarjolla olleeseen Tampere3-hankkeen sähköpostiosoitteeseen, jota intranetissä kehoitettiin käyttämään, mikäli esitettävänä olisi hankkeeseen liittyviä kehittämisideoita tai sitä koskevia kysymyksiä. Kyseisestä kanavasta sain vastaukseksi ohjeen ottaa yhteyttä korkeakoulujen viestintäyksikköihin sekä kyseisten yksikköjen sähköpostiosoitteet. Päätin edetä neuvon mukaisesti. Liitin kyselylomakkeeseen johtavan linkin mukaan korkeakoulujen viestintäyksikköihin lähettämiini sähköpostiviesteihin. (Sähköpostiviestit 22.4.2018, 23.4.2018, 25.4.2018a, ks. liite 3.)

Aikataulutus, realistinen aikaraja, on olennaista kyselytutkimuksella tehtävässä aineiston keruussa. (Punch 2003, luku 3.5.) Asetin kyselylle pitkän vastausajan, huhtikuun lopusta aina juhannusviikon loppuun eli kesälomakauden tyypilliseen alkuun asti. Pitkä vastausaika oli kuitenkin perusteltu, sillä siinä huomioitiin ennen kaikkea korkeakoulujen tutkimus- ja opetushenkilöstöä työllistävät kevätkiireet, kuten lukuvuoden päättymisen ja pääsykoekauden. Samalla varasin aikaa saada nostettua kyselyni vielä uudestaan esille korkeakoulujen viestintäkanavissa, sillä tiedotteet hautautuvat nopeasti uusien alle. Samoin varauduin levittämisessä tapahtuviin mahdollisiin viiveisiin, mikä osoittautuikin tarpeelliseksi. Kyselystäni tiedotettiin korkeakoulujen intranet-sivuilla kaksi kertaa.

Kyselyni levittämisen ensimmäinen kierros eteni siten, että 25.4.2018 sain vahvistuksen lomakkeeni jakamisesta TTY:n intranetissä ja 8.5.2018 TaY:n ja 17.5.2018 TAMK:n vastaavilla sivuilla. (Sähköpostiviestit 25.4.2018b, 8.5.2018, 17.5.2018b, ks. liite 3.) TaY:n myöhemmän julkaisuajan taustalla on,

että ensimmäinen viestini oli jäänyt viestinnässä huomaamatta, joten vasta toinen viesti johti kyselyn jakamiseen (sähköpostiviestit 7.5.2018, 8.5.2018, ks. liite 3). TAMK:n viive selittyi ensiksikin tutkimuslupaprosessilla. TAMK edellytti vararehtorinsa hyväksymää lupaa, mikäli tutkimustyön kohteena oli korkeakoulun henkilökunta, ja tutkimuksentekoon käytettäisiin sen resursseja, kuten intranetin tiedotepalvelua, sekä henkilökunnan työaika. Toiseksi TAMK:ssa pyydettiin myös englanninkielistä saateviestiä Intranetiä varten. (TAMK 2018; sähköpostiviestit 25.4.2018c, 15.5.2018a, 15.5.2018b, 16.5.2018, ks. liite 3). Kyselyn levittämisen toinen kierros ajoittui kesäkuun alkuun. Viimeinen vastaus tuli 19.6.2018.

Nojauduin kyselyn jakelussa siis ulkopuoliseen apuun sen sijaan, että olisin esimerkiksi tavoitellut suoraan yksittäisiä tutkijoita. Se heikensi mahdollisuuksiani hallita ja seurata kyselyn jakeluprosessia. Minulla oli pääsy ainoastaan oman opiskelupaikkani eli TaY:n intranet-sivuille, mutta sekin oli arvokasta. Sitä kautta huomasin, että vaikka TaY:n viestinnästä ilmoitettiin 11.6.2018, että kyselyni oli nostettu toistamiseen intranetissä esille, niin ei ollutkaan vielä tapahtunut: ilmoitusta ei näkynyt intranetin tiedotevirrassa. Lisätiedusteluni jälkeen ilmoitus julkaistiin toisen kerran 15.6.2018. (Sähköpostiviestit 10.6.2018a, 11.6.2018a, 14.6.2018, 15.6.2018, ks. liite 3.) Kuitenkin TTY:n ja TAMK:n kohdalla vakuutuin siitä, että kyselyäni oli nostettu toistamiseen esille, sillä molemmista korkeakouluista tuli lisää vastauksia samoihin aikoihin, kun olin saanut vahvistukset kyselyni uudelleenjulkaisusta (sähköpostiviestit 10.6.2018b, 10.6.2018c, 11.6.2018a, 11.6.2018b, 11.6.2018c, ks. liite 3).

Ensimmäisellä kyselykierroksella huhti-toukokuussa vastauksia tuli korkeakouluista kolme, TTY:sta ja TaY:sta. Mikä voisi selittää vastausten pienen määrän? Toisaalta ainakin TaY:n intranetissä julkaistiin päivää ennen oman kyselylomakkeeni julkaisemista CSC - Tieteen tietotekniikan keskus Oy:n kysely otsikolla "Millainen tutkimustietovaranto olisi paras?", jonka saateviesti on luettavissa liitteestä 4. Se saattoi potentiaalisen vastaajan silmin olla samansisältöinen kuin kyselyni ja siten vähentää vastaushalukkuutta. Lisäksi esimerkiksi Heidi Enwaldin tutkimukseen liittyvästä kyselystä oli noin vuosi, ja kirjallisuuskatsauksesta muistetaan sen tavoittaneen myös parisen kymmentä tamperelaistutkijaa. Keith F. Punch kirjoittaakin, että monella potentiaalisella vastaajalla voi olla jo kyselyähky, mikä on myös yksi syy lisää panostaa houkuttelevaan lähestymistapaan (Punch 2003, luku 4.3). Olin pyrkinyt tekemään omaa aihettani tutkijoille läheisemmäksi laatimalla saateviestiini paikallisuutta esiintuovan otsikon "Kyselytutkimus tamperelaistutkijoiden aineistohallinnasta", mutta Punchin mainitsema ilmiö saattoi silti konkretisoitua kohdallani.

Toisaalta kiinnitin huomiota siihen, että TAMK:sta ei tullut ensimmäisellä kyselykierroksella vastauksia lainkaan. TAMK:n kehittämisyksiköstä todettiin, että kyselyni kohdennus saattaa hämmentää potentiaalisia vastaajia: TAMK:ssa ei ole puhtaasti tutkimustyötä tekevää henkilöstöä, kuten tutkijanimikkeellä toimivaa väkeä, vaikka tutkimusta tehdäänkin monessa eri yhteydessä, kuten jatko-opinnoissa ja projekteissa (sähköpostiviesti 17.5.2018b, ks. liite 3). Pyytäessäni korkeakoulujen viestintäyksikköjä nostamaan kyselyäni toistamiseen esiin, lähetin TAMK:uun uudelleenmuotoillun saateviestin. (Sähköpostiviesti 10.6.2018b, ks. liite 3.) Nostin sillä kertaa puhuttelussa tutkijoiden rinnalle opetushenkilöstön ja jatko-opiskelijat, ja yhdistin tutkimusaineiston esimerkiksi tuotekehitys- ja innovaatiotyössä (TKI) tuotettavaan aineistoon. Päivitetty saateviesti on luettavissa liitteessä 2 b. Toisella kyselykierroksella vastauksia tuli neljätoista lisää ja nyt joukossa oli muutama vastaus myös TAMK:sta. Ei tosin voida luotettavasti arvioida, oliko TAMK:n vastaukset seurausta uudesta saateviestistä vai ajoituksesta. Intranet osoittautui heikoksi väyläksi tavoittaa korkeakoulujen tutkijoita tai kenties kiinnittää heidän huomionsa. Kyselyyni vastasi 18 vastaajaa – vastauksia analysoitiin 16, mistä kerrotaan lisää alaluvussa 4.4 – joten vastaajajoukko oli pieni, kun huomioi, kuinka isoista organisaatioista Tampereen korkeakouluissa on kyse.

#### **4.4 Aineiston analysointi – riippumattomuudesta ja roolijako**

Vastausten keruun jälkeen tulostin vastaukset Google Formsista Excel-taulukon muodossa ja siirsin ne analyysivaihetta varten IBM SPSS Statistics 26 -tilasto-ohjelmaan (jatkossa SPSS) avattuun datatiedostoon. Tiedonsiirto ei vaatinut vastausten manuaalista kopiointia, mutta vastausvaihtoehdot oli koodattava SPSS-kelpoiseen muotoon.<sup>7</sup> Näppäilyvirheiden riskiä minimoin useilla, varsinaisen siirron lisäksi myös analyysin eri vaiheisiin ajoitetuilla, tarkastuskerroilla. Esimerkiksi vertailin, että vastaukset jakautuivat samalla tavalla Google Formsissa, Excel-taulukossa ja SPSS:ssä. Toisaalta pienten muuttujamäärien käsittely onnistui joskus ilman tilasto-ohjelmaakin (Heikkilä 2014, 119). Toisaalta kopioin aineistosta tarvittaessa osia omiksi pienemmiksi datatiedostoihin SPSS-ohjelmassa. Toimin siten esimerkiksi silloin, kun käytin havaintoyksikköinä vastaajien sijasta valintaker-toja, vastaustapahtumia.

---

<sup>7</sup> Esim. mielipideasteikolliset muuttujat tallensin numeroina (1 = vahvasti samaa mieltä, 2 = osittain samaa mieltä jne.).



Kyselytutkimuksissa joudutaan toisinaan hylkäämään vastauksia esimerkiksi hyvin puutteellisten tietojen tai pilailumielessä annettujen vastausten vuoksi, ja niin kävi tälläkin kertaa. Toisaalta kyselyssäni vastaajat saivat ohittaa kysymyksiä, ja jonkin verran olikin eräkatoa (item nonresponse), mikä tarkoittaa, että joidenkin muuttujien yksittäisiä arvoja puuttui eli joihinkin yksittäisiin kysymyksiin tai kysymysten osiin oli jätetty vastaamatta. (Heikkilä 2014, 42.) Toisaalta hylkäsin kaksi vastaajaa analysoiden vain 16 vastaajaa. Ensiksikin toinen oli vastannut vain kahteen kysymykseen<sup>8</sup> ja toinen vain yhteen<sup>9</sup>. Ne tulkittiin keskenjätetyiksi vastauksiksi, kenties vahingossakin avatuiksi, sillä ne eivät mahdollistaneet syy-seuraussuhteiden selvittämistä. Lomakkeen Taustatiedot-osuuden kohdat, ainakin korkeakoulu ja tutkimusala, olisikin kannattanut laittaa pakollisesti täytettäväksi.

Vastaajat olivat satunnaisesti ohittaneet joitakin kysymyksiä tai alakysymyksiä, ja on mahdotonta tietää, oliko se tarkoituksellista. Todennäköisesti on ollut vahinko, kun vastaaja oli jättänyt esimerkiksi monikohtaisessa kysymyksessä 13 (ks. liite 1) vastaamatta vain yhteen kohtaan. Olennaista on muistaa, että kun puhuttiin  $n$  määrän vastaajista olevan jotakin mieltä, osuus oli laskettu yksittäiseen kysymykseen kantaa ottaneista, esimerkiksi 14 vastaajasta, eikä kaikista 16 vastaajasta. Lisäksi kun analyysissä puhuttiin valinta- tai vastauskerroista, havaintoyksikkönä oli vastaajien sijasta vastustapahtumat. Esimerkiksi 16 vastaajaa teki yhteensä 149 säilytyspaikkaa koskenutta valintaa.

Toisaalta vaikka kaikkiin kohtiin olisikin vastattu, vastauksissa voi olla epäjohdonmukaisuuksia. Esimerkiksi TTY:n vastaajista yksi valitsi vastauksen ”ei lainkaan” joka vaihtoehtoon, kun kysyttiin, missä hän säilyttää aineistoaan, mukaan lukien vaihtoehtoon ”muu”. Näin ollen ei selvinnyt, missä hän aineistojaan säilytti ja oliko kyse huolimattomuusvirheestä. Voi olla, että ristiriidoissa tai kohtien väliin jättämisessä oli myös kyse uupumisesta kyselyyn. Toisaalta kyselylomakkeessa oli tarjolla myös epävarmuutta tai neutraaliutta ilmaisevia vaihtoehtoja. Enimmäkseen lomakkeelle annettiin harkittuja vastauksia päätellen esimerkiksi siitä, että vaihtoehtoa ”en osaa sanoa” käytettiin harvakseltaan, sekä siitä, että lomakkeen Vapaa sana -osiossa annettiin palautetta myös koskien kysymysten asettelua ja tulkintaa.

---

<sup>8</sup> Kysymykset 9 (Vaatiiko ensisijainen rahoittajasi sinua laatimaan aineistohallintasuunnitelman?) ja 10 (Kuinka paljon säilytät tutkimusaineistoasi seuraavissa sijainneissa?).

<sup>9</sup> Kysymys 6 (Mitkä seuraavista käsitteistä kuvaavat parhaiten käyttämäsi tutkimusaineistoa?).

## Pienen otoksen ristiintaulukointi ja riippumattomuustesti

Pieni otos tuottaa usein vääristyneitä jakaumia, ja ääripäät voivat korostua vastauksissa, kun taas suuressa otoksessa vastaukset voivat jakautua tasaisemmin. Otosryhmien välisten erojen pitää olla riittävän kattavia, jotta tuloksia voidaan yleistää – tällöin kyse on *tilastollisesta merkitsevyydestä* (significant difference). Riittävän suurella otoksella vältetään satunnaisvirhe; muutaman kymmenen hengen otoksesta ei pysty tekemään kovin pitkälle meneviä johtopäätöksiä. Esimerkiksi jos kolikkoa heitetään kymmenen kertaa, klaava voidaan kenties saada kahdesti ja kruunu kahdeksan kertaa, mutta jos heittokertoja on tuhat, 50 prosentin todennäköisyys saada kumpi on vain todennäköisesti lähempänä. (Heikkilä 2014, 75; Wetcher-Hendricks 2011, 106-110.) Kyselyssäni vaikkapa aineistojen säilytyspaikkojen suosion kohdalla on huomioitava, että kun kuusi tiedeinstituutiotason säilytyspaikkaa sai TTY:n neljältä vastaajalta 15 ei lainkaan -säilytysmäärän valintakertaa, se tarkoittaa kunkin torjuneen keskimäärin 4,5 säilytyspaikkaa, mutta sama vastaaja saattoi yksin torjua kaikki.

Tilastotieteilijä Jeff Sauro kirjoittaa MeasuringU-sivuston blogikirjoituksessaan “Best Practices for Using Statistics on Small Sample Sizes”, että yhden tutkijan pienenä pitämä otos voi olla toiselle iso. Käyttäjätutkimuksen kontekstiin keskittynyt Sauro sanoo viittaavansa itse pienen otoksen tutkimuksilla sellaisiin, joissa on tarkasteltu 5–30 käyttäjän joukkoa. Hän toteaa määrän olevan kyseisellä alalla tavanomainen. Siitä voidaan havaita jo selvät erot näkemyksissä. Laajan otoksen puuttuminen ei tarkoita, että tilastollista otetta ei voisi käyttää – on keinoja, jotka mahdollistavat pienenkin otoksen varassa toimimisen – mutta on tärkeää tunnistaa siihen liittyvät rajoitteet, kuten suurpiirteisyys. (Sauro 2013.) Toisin sanoen karkeiden erojen selvittämiseen riittää pieni aineisto. (Heikkilä 2014, 40.)

Määrällisessä tutkimuksessa tarkastellaan muuttujien välisiä suhteita. Muuttujien jakaantumisen tarkastelu on vain tulkinnan ensi vaihe, sillä siihen jääminen on akateemisesti ajatellen liian kuvailtavaa ja aikasidonnaista tietoa tuottavaa. Sen sijaan muuttujien välisen suhteen tarkastelu lisää ennustettavuutta. (Punch 2003, luku 2.2; Tähtinen et al. 2020, 36.) Määrällisen kyselyaineiston analysointi on progressiivinen prosessi, jossa aineistoa vedetään yhteen ja tiivistetään, jotta muuttujien suhteista toisiinsa voidaan vetää vakaita johtopäätöksiä. Ennen analysointia aineistoa on valmisteltava esimerkiksi oikolukemalla palautetut lomakkeet. Itse kyselyaineiston analysoiminen on jaoteltavissa kolmeen vaiheeseen. Ensiksi saadut vastaukset tiivistetään ja pelkistetään, mihin kuuluu myös muuttujien esille nostaminen. Toiseksi esitetään muuttujien jakautuminen eli tehdään

kuvailevan tason analyysi. Kolmantena analysoidaan muuttujien väliset suhteet ensin kahdenvälisinä suhteina ja sitten yhteisinä suhteina tutkimuskysymyksen ohjaamana. Systemaattinen taulukojen käyttö auttaa jäsentämään analyysiä ja luomaan rakennetta tulosten esittämiselle. (Punch 2003, luku 4.5.) Tutkimuskysymysten oli ohjattava tutkielmassani analyysiä jo siksi, että moniosainen kyselylomake olisi voinut mahdollistaa suuren määrän selittäviä muuttujia, joiden vaikutusta tutkijoiden aineistohallinnan näkemyksiin ja käytäntöihin olisi voinut tutkia.

Vertailin eri korkeakouluista tulevia vastaajia keskenään kontingenssitauluilla eli ristiintaulukoilla, joilla selvitetään kahden luokitellun muuttujan välistä yhteyttä. Muuttujien saamat arvot, joita vertaillaan keskenään, asettuvat taulukossa vierekkäin. Esitin muuttujat samassa taulukossa niin, että selittävä muuttuja – yleensä korkeakoulujako – oli sarakemuuttuja eli sarakkeille asettuva. Selitettävä muuttuja oli rivimuuttuja eli riveille asettuva. Tosin selittävän ja selitettävän muuttujan erottaminen ei aina, kuten kahden mielipidemuuttujan välistä yhteyttä tutkittaessa, ole mahdollista. Ristiintaulukosta käy ilmi vastausten frekvenssi- ja prosenttijakaumat. Taulukon ruuduissa eli soluissa olevat solufrekvenssit kertovat, montako mainituin ominaisuuksin varustettua vastaajaa aineistossa on. Taulukossa alhaalla on sarake- ja oikeassa reunassa rivisummat. (Heikkilä 2014, 143, 198.)

Koska TaY:sta tuli saman verran vastaajia kuin TTY:sta ja TAMK:sta yhteensä, esitin vertailun helpottamiseksi vastausten jakautumisen taulukossa frekvenssien lisäksi prosentteina. Tilanteessa, jossa vastaajaryhmät ovat eri kokoisia, pelkkien frekvenssien käyttö on tulkintoja vääristävää. (Tähtinen et al. 2020, 172-173.) Kuitenkin on tiedostettava, että prosentit voisivat olla mielekkäämpiä vasta suuresta luvusta laskettuna. Voi olla harhaanjohtavaa esimerkiksi sanoa, että TAMK:n vastaajista 50 prosenttia (2 vastaajaa) ja TTY:n vastaajista 75 prosenttia (3 vastaajaa) on jotakin mieltä, kun käytännössä eroa on yksi vastaaja. Tarja Heikkilän mukaan lukumäärän kertominen on rehellisempää, joskin hän huomauttaa, että niin pienen ryhmän erittely voi olla parempi jättää kokonaan tekemättä. (Heikkilä 2014, 78.) Tässä tutkielmassa pienestä otoksesta rajoitteineen on kuitenkin puhuttu avoimesti, ja taulukkoja käytettiin paljon, jotta tarkat luvut ovat lukijan tarkistettavissa.

Ristiintaulukointi on havainnollinen menetelmä. Kun halutaan tutkia muuttujien välisiä riippuvuuksia, se ei ole aina menetelmistä tehokkain. Ristiintaulukko paljasti, oliko esimerkiksi eri korkeakouluja edustavien vastaajien keskuudessa eroja, kuten miten suuri osuus kunkin korkeakoulun vastaajista oli mitään mieltä. Tilastolliseen lähestymistapaan kuuluu kysyä ja testata, pätevätkö otoksessa havaitut erot myös perusjoukossa. (Heikkilä 2014, 174, 180.) Toisin sanoen vastausten jakautumisen

kuvailun jälkeen oli olennaista selvittää, olivatko mahdolliset erot esimerkiksi eri korkeakoulujen edustajien välillä tilastollisesti merkitseviä eivätkä sattumasta johtuvia. Tutkin, oliko tutkijan edustamalla korkeakoululla vaikutusta hänen näkemyksiinsä ja käytäntöihinsä eli oliko kahden muuttujan – vastaajan korkeakoulun ja esimerkiksi hänen mielipiteensä – välillä tilastollisesti merkitsevää riippuvuutta. Tuloksia katsottaessa on silti erittäin tärkeä tiedostaa, että tilasto-ohjelmia käytettäessä tietoa perusjoukon koosta ja otantasuhteesta ei huomioida millään lailla. (Heikkilä 2014, 42, 280.)

Riippuvuus selviää tilastollisen merkitsevyydestin avulla. Testaamisessa keskeinen käsite on ensiksikin *nollahypoteesi*, joka tarkoittaa, että muuttujien välillä ei ole riippuvuutta. Käsite *merkitsevyytaso* (significance) eli riskitaso puolestaan ilmoittaa, miten suurella todennäköisyydellä havaitun eron tai riippuvuuden taustalla on sattuma – mitä pienempi riski, sitä merkitsevempi tulos. Merkitsevyytason lyhenne on  $p$  (probability), mutta myös  $\alpha$  (alfa) ja SPSS-ohjelman käyttämä lyhenne *Sig.* (significance) viittaavat samaan. Jos  $p$ -arvo on suurempi tai yhtä suuri kuin 0.05, riippuvuus ei ole tilastollisesti merkitsevä. (Heikkilä 2014, 182, 184-185.) Jos muuttujan mittari ei tuota vaihtelevuutta – kaikki vastaajat antavat saman vastauksen – voidaan todeta, että muuttujien välillä ei ole suhdetta. Tosin voi myös olla, että käytetty mittari ei ollut tarpeeksi herkkä tuottamaan vaihtelevuutta vastauksiin.<sup>10</sup> (Punch 2003, luku 4.2.)

Tutkielman analyysissä nollahypoteesi oli, että ei voitu riittävällä varmuudella sanoa selittävän muuttujan (sarakemuuttuja, yleensä tutkijan korkeakoulu) vaikuttaneen selitettävän muuttujan käyttäytymiseen ( $p \geq 0.05$ ). Vaihtoehtoinen hypoteesi oli, että voitiin 95 prosentin todennäköisyydellä sanoa selittävän muuttujan (sarakemuuttuja, yleensä tutkijan korkeakoulu) vaikuttaneen selitettävän muuttujan käyttäytymiseen ( $p < 0.05$ ). Toisin sanoen, jos  $p$ -arvo oli yhtä suuri tai suurempi kuin 0.05, nollahypoteesi jäi voimaan. Jos  $p$ -arvo oli pienempi kuin 0.05, vaihtoehtoinen hypoteesi astui voimaan.  $P$ -arvo saadaan tilastollisella merkitsevyydestillä. (Tähtinen et al. 2020, 40, 42, 121-122; Heikkilä 2014, 182.) Tässä tutkielmassa  $p$ -arvo saatiin käyttämällä *Fisher-Freeman-*

---

<sup>10</sup> Lähtökohtaisesti kysymysten ja vastauskohtien on tuotettava vaihtelevuutta, ja se pitää testata ennen kyselyn jakamista. (Punch 2003, luku 4.2.) Vaihe ei ollut tässä olennainen, sillä lomaketta oli käytetty aiemminkin.

*Haltonin eksaktia testiä eli yleistettyä Fisherin testiä, joka sopii pienen otoskoon, pienten solufrekvenssien sekä monisoluisen ristiintaulukon, esimerkiksi 3 x 6 -taulukon, tilanteeseen. (Esim. sähköpostiviestit 15.10.2020; 17.10.2020, ks. liite 3.)*

Tässä tutkielmassa tyydyttiin vain tulkitsemaan SPSS-ohjelmiston antamia Fisher-Freeman-Haltonin eksaktin testin tuloksia. Ohjelma huolehti laskutoimituksista ja ohjelmiston käyttäjänä puolestani varmistin, että testiä voi käyttää ja sen edellytykset täyttyivät, ja tein johtopäätökset SPSS:n antamien Sig.-arvojen perusteella.<sup>11</sup> (Heikkilä 2014, 201.) Fisher-Freeman-Haltonin eksaktin testin periaatteisiin voi tutustua tarkemmin esimerkiksi G. H. Freemanin ja J. H. Haltonin artikkelin ”Note on an Exact Treatment of Contingency, Goodness of Fit and Other Problems of Significance” parissa (Halton & Freeman 1951). On kuitenkin tarpeen luoda tässäkin katsaus testin taustaan.

Fisher-Freeman-Haltonin eksakti testi liittyy kehityshistorialtaan läheisesti Fisherin eksaktiin testiin ja edelleen Pearsonin khiin neliö -testiin eli  $\chi^2$  -riippumattomuustestiin, jotka sopivat niin ikään ei-numeeristen ristiintaulukoitujen muuttujien analysointiin. Esimerkiksi Tarja Heikkilä kirjoittaa, että kun molemmat muuttujat ovat nominaaliasteikollisia tai toinen on nominaaliasteikollinen ja toinen järjestysasteikollinen, juuri ristiintaulukoinnin ja khiin neliö -testin suorittaminen on väline muuttujien välisten yhteyksien tutkimiseen. (Heikkilä 2014, 191.) Khiin neliö -testi on kolmesta edellä mainitusta menetelmästä yksinkertaisin tapa selvittää sarake- ja rivimuuttujan välinen riippuvuus. Muuttujien mitta-asteikoilla ei merkitystä, sillä khiin neliö -testiä voidaan käyttää kaikkien kohdalla, kunhan sen muut edellytykset täyttyvät. Edellytyksiin kuuluu ensiksikin, että otoksen on oltava satunnaisotos, toiseksi korkeintaan 20 prosenttia odotetuista frekvensseistä saa olla pienempiä kuin viisi ja kolmanneksi jokaisen odotetun frekvenssin on oltava suurempi kuin yksi.<sup>12</sup> (Heikkilä 2014, 200–201; Tähtinen, Laakkonen & Broberg 2020, 167-168.) Keskeistä on huomata, että esimerkiksi

---

<sup>11</sup> Menetelmiä saa käyttää vasta, kun on selvillä menetelmien edellytyksistä, kuten muuttujia koskevista vaatimuksista. Jos ei ole perehtynyt menetelmiin, tilasto-ohjelman käytön vaivattomuus voi johtaa esimerkiksi menetelmien käyttöön väärissä yhteyksissä. Toisaalta Tarja Heikkilä kirjoittaa, että monimuuttujamenetelmiin siirryttäessä tutkijalta vaaditaan jo enemmän metodista osaamista. (Heikkilä 2014, 174, 227.) Niistä ei tässä ollut kyse.

<sup>12</sup> Khiin neliö -testissä tilasto-ohjelma laskee nollahypoteesin edellyttämät odotetut frekvenssit. Odotetut frekvenssit tarkoittavat toteutuneiden eli havaittujen frekvenssien – kyselyn vastaajien vastausten jakautumisen – sijasta solufrekvenssejä tilanteessa, jossa muuttujien välillä ei ole lainkaan riippuvuutta. Sen jälkeen tilasto-ohjelma laskee testisuureen, jonka arvo riippuu odotettujen ja havaittujen eli toteutuneiden frekvenssien välisistä eroista ollen sitä suurempi mitä suurempi ero on. Samoin lasketaan vapausaste, joka perustuu rivien ja sarakkeiden lukumäärään. Ohjelma laskee vapausasteen ja testisuuren pohjalta merkitsevyytason. (Heikkilä 2014, 200.)

khiin neliö -testin yhteydessä pieni p-arvo on edellytys riippuvuudelle, mutta ei voimakkaan riippuvuuden tae. Lisäksi, jos otos ei täytä testin edellytyksiä, khiin neliö -testi johtaa liian helposti nollahypoteesin hylkäämiseen ja virheellisen johtopäätöksen tekemiseen. Edellytysten täyttämiseksi solufrekvenssiä voidaan kasvattaa yhdistelemällä alkuperäisiä luokkia tai jättämällä pois luokat, joissa frekvenssi on liian pieni. (Heikkilä 2014, 201, 207.)

Siinä missä khiin neliö -testissä on otoskoko koskevat vähimmäisvaatimukset, R. A. Fisher teki testistä pieniinkin otoksiin soveltuvan version (Fisher 1922), joka tunnetaan Fisherin eksaktina testinä. Siinä merkitsevyysarvo saadaan ”asymptoottisten todennäköisyysjakaumien käytön sijaan tutkimalla kaikki mahdolliset ristiintaulukon vastaavat solukombinaatiot” (Tähtinen, Laakkonen & Broberg 2020, 167). Menetelmä on laskennallisesti erittäin raskas – esimerkiksi Juhani Tähtinen et al. eivät käy menetelmään liittyviä kaavoja läpi katsauksessaan ”Tilastollisen aineiston käsittelyn ja tulokinnan perusteita” – mutta SPSS tulostaa testin p-arvoineen. Fisherin eksaktin testin tuottama p-arvo tulkitaan samalla tapaa kuin muissakin testeissä. (Tähtinen et al. 2020, 167-168.)

Kuitenkin myös Fisherin eksaktissa testissä on rajoitteensa, joiden vuoksi se ei sovellu tutkielmani tilanteeseen: Fisherin eksakti testi on tarkoitettu 2 x 2 -ristiintaulukoille (Fisher 1922, ks. myös Heikkilä 2014, 208). G. H. Freeman ja J. H. Halton työstivät vuorostaan Fisherin eksaktista testistä isompienkin ristiintaulukoiden, kuten 3 x 5 -taulukon, analysointiin sopivan version. (Freeman & Halton 1951.) Esimerkiksi Graeme D. Ruxton ja Markus Neuhäuser pohtivat vertaisarvoidussa artikkelissaan ”Good practice in testing for an association in contingency tables”, että Fisher-Freeman-Haltonin eksakti testi on yksi sopivimmista vaihtoehtoista 2 x 2 -ristiintaulukkoa suuremmille taulukoille. (Ruxton & Neuhäuser 2010, 1508–1510.)

Aineistoni kohdalla uusien muuttujien luominen, kuten mielipiteiden yhdistäminen, taulukon solujen määrän supistamiseksi ja solufrekvenssien kasvattamiseksi ei olisi tuonut apua. Esimerkiksi mielipiteitä kuvaavia luokkia olisi voitu yhdistää – yhdistetty vaihtoehtojen ”tyytymätön” ja ”erittäin tyytymätön” saamat vastausmäärät sekä vaihtoehtojen ”erittäin tyytyväinen” ja ”tyytyväinen” vastausmäärät ja karsittu vaihtoehto ”en osaa sanoa” – niin, että 3 x 6 -taulukosta olisi tullut 3 x 3 -taulukko. Kuitenkaan pienen otoksen ja kolmen sarakemuuttujan tilanteessa tiivistäminen ei olisi kasvattanut solufrekvenssejä siinä määrin, että khiin neliö -testi olisi mahdollistunut.

Joitakin luokkien yhdistämissä tehtiin, mutta ne esitettiin yhteenvetoina alkuperäisten vaihtoehtojen rinnalla eikä niiden sijasta. Lisäksi tehtiin laajennuksia valitun tilastollisen testin käytön mahdol-

listamiseksi. Esimerkiksi kysymyksen 12 ("Jos tutkimusaineistosi tai osa tutkimusaineistoistasi ei ole muiden saatavilla, miksi ei? Valitse kaikki sopivat vaihtoehdot") yhteydessä otin analyysiin mukaan myös vastaajat, jotka eivät olleet valinneet tiettyä syytä selittämään aineistojensa jakamattomuutta. Analyysissä käytetty vaihtoehto "ei valittu" ei siis ollut vastausvaihtoehtona lomakkeella. Analyysissä tulkittiin, että jos vastaaja ei valinnut vaihtoehtoa, hän ei pitänyt sitä syynä, sen sijaan, että olisi tulkittu, että vastaaja olisi unohtanut ottaa kohtaan kantaa. Vastaajia kysymykseen 12 oli yhteensä 15, kun kaikkiaan heitä oli 16. Siten tulkittiin, että yksi vastaaja ei pitänyt mitään syytä merkityksellisenä. Vaihtoehtona olisi voinut olla "ei mikään näistä".

### **Viitekehys tutkijoiden aineistonhallinnalle toimijarooleineen**

Aineistonhallinnan näkemysten ja käytäntöjen korkeakoulukohtaisen vertailun lisäksi tutkimuskysymyksinäni oli selvittää, miten tamperelaistutkijat suhtautuvat aineistojen avoimuuden toteuttamiseen ja aineistojen hyödyntämiseen sekä miten aineistonhallinnan näkemykset erosivat käytännöistä. Kyselyyn tulleita vastauksia ei olisi ollut mielekästä käydä läpi listamaisesti siinä järjestyksessä, kun kyselylomake eteni, vaan käytin tamperelaistutkijoiden kyselyyni antamien vastausten läpikäynnissä ja esittämisessä viitekehystä, joka pohjautui väljästi Fecherin et al. (2015) alaluvussa 2.2 esiteltyyn viitekehykseen.

Fecherin et al. viitekehys on laadullisen sisällönanalyysin (qualitative content analysis, QCA) tulos. (Fecher et al. 2015, 6.) Koska hyödynsin ja muokkasinkin viitekehystä, QCA-metodi oli tässäkin tutkielmassa läsnä. Toisaalta QCA-metodia käytetään yleensä verbaaliseen aineistoon, kuten vaikkapa lehtiin ja haastatteluihin. (Schreier 2012, luku 1.) Tässä sitä käytettiin tekstimuotoiseen kyselylomakkeeseen. Toisaalta QCA:n perustana on kategorioiden luominen, mistä viitekehyksessäkin oli kyse.<sup>13</sup> (Schreier 2019.) Yksinkertaistetusti QCA:n tavoite on kuvailla systemaattisesti tutkimuksessa käytetyn (laadullisen) aineiston merkitys. Kyse on myös aineiston supistamisesta, mikä tarkoittaa, että analyysissä ei oteta huomioon kaikkea informaatiota, jota vaikkapa lähteenä käytetty haastattelun

---

<sup>13</sup> Toisaalta osa tutkijoista on nähnyt QCA:n hybridimetodin, joka yhdistelee kvalitatiivisia ja kvantitatiivisia elementtejä. Toisaalta osa tutkijoista korostaa QCA-metodin systemaattista luonnetta ja soveltuvuutta suurtenkin aineistomäärien analysoimiseen osan taas painottaessa metodin yhteyttä laadulliseen tutkimukseen. (Schreier 2019.)

litteraatio voi tarjota – tai tässä tapauksessa, mitä kyselylomake voi tuottaa – vaan analyysi rajoitetaan tutkimuskysymysten kannalta olennaisiin näkökulmiin. (Schreier 2012, luku 1.)

QCA-metodi pakottaa valitsemaan analyysiin tietyt näkökulmat ja pysymään niissä siten, että näkökulmien ympärille rakennetaan koodauskehys (coding frame) pää- ja alakategorioineen, jonka kautta aineistoa tarkastellaan (ks. Fecherin et al. viitekehysten kategorioista alaluku 2.2). Tyypillisesti osa analyysin näkökulmista käy ilmi jo tutkimuskysymyksistä – tässä näkökulmana on siis aineistohallinnan toimijaroolit. Kun näkökulma oli selvillä, oli aika tunnistaa, mitä aineistossa, tässä tapauksessa kyselylomakkeessa, sanottiin niistä. QCA-metodissa kuhunkin kategoriaan liittyy tarkennuksia eli alakategorioita. Niitä voidaan juontaa käsitevetoisesti eli nojaamalla siihen, mitä aihepiiristä tiedetään jo ennalta, jolloin aineistoa jäsennetään ”ulkoapäin” tulevilla kategorioilla, tai aineistovetoisesti eli tutkimalla sitä, mitä itse aineistosta on havaittavissa. (Schreier 2012, luku 4.)

Yleisesti ottaen lähestymistapani oli toisaalta käsitevetoinen, sillä otin lähtökohdaksi Fecherin et al. valmiin viitekehysten, jota aloin muokata, mutta toisaalta aineistovetoinen, sillä kyselylomakkeen sisältö vaikutti siihen, millaiseksi analyysissä lopulta käytetty viitekehys muotoutui. Viitekehysten osa-alueet aineistohallinnan rooleineen ovat siis pääkategorioita. Alakategorioita ei kuitenkaan tässä tutkielmassa ole. Esimerkiksi Fecher et al. rakensivat viitekehysten paitsi kyselyaineistonsa, myös kirjallisuuskatsauksesta tekemiensä havaintojen pohjalta, siinä missä tässä tutkielmassa oli kyse hyvin rajallisesta aineistosta. Koska QCA kohdistui aineiston keruun välineeseen, yhteen kyselylomakkeeseen, eikä aineiston keruun antiin, analysoitavaa sisältöä oli vähän ja koodauskehys hyvin suppea. QCA tuki pikemminkin tulosten esittämistä kuin tulosten hankkimista. Lopputuloksena oli aineistohallinnan viitekehys.

Toisaalta otin siis viitekehysten muotoiluprosessissa vaikutteita QCA-metodista. QCA:ssa voidaan nähdä olevan kyse esimerkiksi sosiaalisten tilanteiden ja prosessien rekonstruktiota tukevan rakenteen luomisesta. (Schreier 2019.) Tässä tutkielmassa on kyse tutkijan aineistohallinnan ja sen toimijuuden jäsentämisestä. Toisaalta aineistohallinnan roolien asettaminen sivuaa esimerkiksi palvelumuotoilussa ja liike-elämässä käytettävän persoona-metodin (persona) päämääriä. Persoonien luomisella on tarkoitus antaa kasvot tietyille abstrakteille ihmisryhmille – kuten asiakassegmenteille tai muille erilaisen toiminnan kohderyhmille – esimerkiksi toimivampien palveluiden kehittämiseksi. Ne auttavat tekemään ryhmiä, joilla on samanlaisia palveluntarpeita tai käyttäytymismalleja, hel-



pommin hahmotettavia. Persoonat ovat fiktiivisiä hahmotelmia. Ne voivat tietyissä tilanteissa pohjautua esimerkiksi kerättyyn asiakasdataan, mutta toisinaan olettamuksiinkin. Yksinkertaistettuna joukko ihmisiä tiivistetään yhdeksi hahmoksi, persoonaksi, jolle hahmotetaan tietyt tavoitteet, tietynlainen toimintaympäristö sekä tietyt toimintaresurssit. (Stickdorn et al. 2018, 32, 41, 76-78, 128, 143.) Tässä tutkielmassa aineistohallinnan toimijaroolit olivat tavoitteiltaan kuin persoonia: tarkoitus oli tukea aineistohallinnan kehittämistoimien kohdentamista. Toisaalta katsottiin, miten todelliset tutkijat, kyselyn vastaajat, asettuivat rooleihin, ja toisaalta, mitä sisältöä he niihin tuottivat.

Taulukko 1. Tutkijoiden aineistohallinnan viitekehys aineistohallinnan toimijaroleineen verrattuna Fecherin et al. (2015) laatimaan aineiston jakamisen viitekehukseen.

<b>Aineistohallinnan osa-alueet</b>	<b>Tutkijan toimijarooli aineistohallinnassa</b>	<b>Alkuperäisen Fecherin et al. (2015) viitekehysten osa-alueet</b>
Aineiston keruu	Tutkimusaineiston kerääjä	Aineiston luovuttaja
Normit	Tutkimusaineiston suojelija	Normit
Tiedeyhteisö	Tiedeyhteisön jäsen	Tutkimusyhteisö
Aineiston uudelleenkäyttö	Tutkimusaineiston uudelleenkäytön valvoja Tutkimusaineiston uudelleenkäyttäjä	Aineiston vastaanottaja
Aineiston käytettävyys	Tutkimusaineiston käsittelijä	Datainfrastruktuuri
Tiedeinstituutiot	Aineistohallinnan tuen saaja	Tutkimusorganisaatiot

Taulukosta 1 on nähtävissä, mistä aineistohallinnan osa-alueista ja niihin liittyvistä toimijaroleista viitekehysteni koostuu. Muokkasin Fecherin et al. viitekehystä koskemaan pelkän aineiston jakamisen sijasta laajemmin aineistohallintaa, tutkimusaineistojen hyödyntämistä. Toisaalta pystyin jonkin verran hyödyntämään jäsentelyssä sitä, mitä sisältöjä he mielsivät osa-alueisiin kuuluviksi. Toisaalta uudelleenmuotoilua ohjasivat tutkimuskysymysten ja kyselylomakkeen näkökulmat sekä tutkielman laajuudelle annetut rajat. Liitteestä 5 käy ilmi, miten sovitin kyselylomakkeen kysymykset viitekehykseeni. Viitekehysten avulla saatiin kyselyaineistosta tehokkaasti esille vastaukset tutkimuskysymyksiin, sillä se osoitti, mistä osa-alueista aineistohallinta voi rakentua eli mistä osa-alueista näkemykset sekä käytännöt ja kokemukset nousevat. Päähuomio oli tutkimusaineistojen avoimuuden toteutumisen edellytyksissä sekä tutkimusaineistojen hyödyntämisessä. Kyselyssä esitettiin kysymyksiä esimerkiksi tutkijoiden suhtautumisesta aineistojen avoimuuteen liittyviin ehtoihin ja esteisiin, ja tuloksia analysoitaessa ehdot ja esteet ryhmiteltiin eri osa-alueiden ja roolien mukaisesti. Seuraavassa esittelen viitekehysteni roolijakoineen osa-alue kerrallaan.

Ensimmäisessä viitekehysten osa-alueessa tutkija oli aineiston kerääjä. Osa-alue keskittyi aineistojen avoimuuteen. Mittasin tutkijoiden aineiston kerääjän roolille antamaa painoarvoa esimerkiksi

ottamalla kyselylomakkeella listatuista aineiston jakamisen ehdoista ja esteistä tarkasteluun tutkijan henkilökohtaisiin etuihin liittyvät vaihtoehdot.

Toisessa osa-alueessa tutkija oli tutkimusaineiston suojelija, jolloin lähtökohtana oli, että tutkimusaineistoja hallittaessa niiden avoimuutta säädellään niihin itseensä liittyvistä syistä. En korostanut jakoa eettisiin ja oikeudellisiin normeihin, sillä kyselylomakkeella normeja tarkasteltiin yleisemmin. Syy ”aineistojeni ei kuulu olla julkisia” tiedusteltaessa, miksi aineistoa ei jaeta, voisi olla yhdistettävissä sekä eettiseen että oikeudelliseen näkökulmaan: eettiseen, jos aineiston ei kuulu olla julkinen tietolähteen suojelun vuoksi, ja oikeudelliseen, jos tutkija vaikkapa kokee omistavansa aineiston. Kukin vastaaja tulkitsi kysymyksen omista lähtökohdistaan. Kysely ei myöskään ottanut kantaa siihen, kenen oikeuksia oikeudellisilla luvilla pyritään suojaamaan: tutkijan itsensä, hänen työnantajansa vai tutkimuksen tietolähteiden. Esimerkiksi tietolähteinä käytetyt henkilöt ovat kenties olleet motivoituneita osallistumaan juuri alkuperäiseen tutkimukseen, eivätkä täyttämään toisenlaisten tutkimusten tavoitteita, mistä aineiston uudelleenkäytössä olisi kyse.

Kolmannessa osa-alueessa käytiin läpi tutkimusaineistojen avoimuuden merkitystä asettamalla tutkija tiedeyhteisön jäsenen rooliin. Tarkastelussa oli esimerkiksi tutkijoiden näkemykset vastavuoroisesta aineiston jakamisesta sekä jakamisen laajuudesta. Arvioitiin, miten tärkeänä he pitivät kollegojensa ja Tampereen toisten korkeakoulujen aineistojen hyödyntämistä tutkimuksen kehittymiselle.

Neljännessä, tutkimusaineistojen uudelleenkäyttöä käsittelevässä, osa-alueessa huomio pysyi aineistojen hyödyntämisessä. Ensiksikin tutkittiin, miten monikäyttöisinä tutkijat pitivät aineistoja. Toiseksi tutkijoita tarkasteltiin heidän itse keräämiensä aineistojen uudelleenkäytön valvojina, mutta kunnian sijasta aineiston hyödyntämisen, tutkimuksen tekemisen, näkökulmasta. Tutkijoilta tiedusteltiin heidän suhtautumistaan kuuteen aineistojen uudelleenkäytön seurantaan koskevaan ehtoon. Jaoin ehdot analyysiä varten kahteen ryhmään, joiden avulla pystyttiin hahmottamaan paremmin tutkijoiden halua sitoutua valvontaan. Kuudesta ehdosta kaksi oli *passiivista uudelleenkäytön seuraamista* eli tutkija saisi yhteenvetoja aineistojensa uudelleenkäytöstä. Neljässä muussa ehdossa oli kyse *aktiivisesta uudelleenkäytön seuraamisesta* eli aineiston kerääjä osallistuisi aineiston uudelleenkäyttöön niin tulosten julkaisuprosessiin puuttujana kuin uusien tutkimusten sisällönkin ohjaajana. Kolmanneksi tutkittiin tutkijoiden roolia valmiiden aineistojen uudelleenkäyttäjinä,

kuten heidän kokemuksiinsa valmiiden aineistojen hyödyntämisestä ja motivoituneisuuttaan siihen. Tässäkin huomioitiin sekä tutkijan oma tieteenala että Tampereen korkeakoulut.

Viides osa-alue keskittyi niin ikään tutkimusaineistojen hyödyntämiseen, mutta aineiston käytettävyyden näkökulmasta, jolloin tutkijat olivat aineiston käsittelijöitä. Huomio oli ensiksikin lomakkeen kysymyksissä, jotka selvittivät, miten ymmärrettävinä tutkijat näkivät aineistot. Toiseksi käytiin läpi, miten he suhtautuivat esimerkiksi aineistojen hyödyntämisen ja hallinnan prosesseihin ja välineisiin.

Viimeisessä eli kuudennessa osa-alueessa tutkijat yhdistettiin aineistohallinnan tuen saajan rooliin. Mielenkiinnon kohteena oli tuki, jota tiedeinstituutiot antavat aineistojen avoimuuden ja hyödyntämisen toteuttamiseen, sekä tarjotun tuen että tuen tarpeen näkökulmista. Toisaalta tiedeinstituutioita voivat olla paitsi korkeakoulut, yliopistot ja rahoittajat, myös esimerkiksi tieteelliset seurakuntat ja tiedeakatemit (TJNK 2018, 6), mutta kyselyssä painotuttiin rahoittajiin ja arkistoihin. Osa-alueessa teemana oli ensiksikin yleisemmin tutkijoiden aineistohallintaansa varten saamat puitteet ja vaatimukset. Toisena teemana oli tutkijoiden aineistonsäilytystavat – aiemmassakin tutkimuksessa on nähty tutkijoiden aineistontallennuspaikkavalintojen vaikuttavan heidän aineistojensa avoimuuteen.

Taulukko 2. Säilytyspaikat viitekehyksen aineistohallinnan tuen saaja -osuudessa säilytysvastuualueittain.

Säilytysvastuualue			
	Tiedeinstituutiotason säilytysvastuualue	Yksilökeskeinen säilytysvastuualue	Määrittelemätön säilytysvastuualue
Kyselylomakkeella tarjotut säilytyspaikkavaihtoehdot	<ul style="list-style-type: none"> <li>➤ Organisaation/yksikön palvelin</li> <li>➤ Tutkimushankkeen palvelin</li> <li>➤ Tieteenalakohtainen säilytyspaikka</li> <li>➤ Organisaation oma arkisto</li> <li>➤ Julkaisijan tarjoama säilytys/arkisto</li> <li>➤ Muu (tieto)arkisto</li> </ul>	<ul style="list-style-type: none"> <li>➤ Oma henkilökohtainen tietokone</li> <li>➤ Paperimuoto omassa toimistossa</li> <li>➤ Muistitikku tai ulkoinen kovalevy</li> </ul>	<ul style="list-style-type: none"> <li>➤ Muu</li> </ul>

Lomakkeella esitetyt säilytyspaikkavaihtoehdot jaettiin säilytysvastuualueisiin taulukossa 2 kuvattulla tavalla. Jaottelu on keinotekoinen, sillä esimerkiksi työpaikka voi toki tarjota tutkijan käyttöön ulkoisen kovalevyn. Kuitenkin niin kauan, kun se on tutkijan hallussa, tutkija on vastuussa esimerkiksi sen tallessa pysymisestä. Säilytystä tutkittiin niin säilytyspaikkavalintoja kuin säilytysmääriäkin vertaillen, ja siten arvioitiin, minkä merkityksen tutkijat antavat tiedeinstituutioilta tulevalle tuelle aineistoa säilyttäessään. Osa-alueen viimeisenä teemana oli selvittää rahoittajan ja työpaikan vaikutusta tutkijoiden tyytyväisyyteen aineistohallinnan prosesseihinsa. Tässä osa-alueessa ei siten vain vertailtu korkeakouluja, vaan huomio oli myös korkeakouluissa työnantajana.

Kaikkia lomakkeen kohtia ei voinut sovittaa viitekehysten rajoihin. Esimerkiksi tutkijoiden taustatekijät esittelin omassa yleiskatsausaluvussaan, sillä ne heijastelivat kaikkien viitekehysten osa-alueiden sisältöihin ja tutkijoiden vastauksiin. Kävin läpi, millaisia aineistoja he käyttivät, miten paljon he jakoivat aineistojaan ja miten halukkaita he olivat lisäämään jakamistaan. Samoin esittelin, miten he suhtautuivat yleisesti aineistojen käyttörajoituksiin, siinä missä viitekehysten osa-alueissa nostin esiin esimerkiksi normeihin ja uudelleenkäytön valvontaan liittyvät rajoitukset. Myös esimerkiksi vastaajien ikäjakauma, asemat työyhteisössä ja tieteenalataustat käytiin läpi, mutta lyhyesti, sillä niitä ei käytetty selittävinä muuttujina analyysissä: ne eivät liittyneet tutkimuskysymyksiini. Koska vastaajajoukko oli pieni, ei olisi ollut mielekästäkään vertailla esimerkiksi tutkimusalojen vaikutusta vastauksiin; yksittäistä alaa kohti kun saattoi olla vain yksi vastaaja. Siten tutkijat eivät edusta tässä tutkielmassa esimerkiksi alaansa tai ikäryhmäänsä. Sen sijaan kukin vastaaja edustaa toisaalta itseään ja korkeakouluaan; toisaalta kokonaisuudessaan vastaajajoukko edustaa tamperelaistutkijoita.

Osa kyselylomakkeen kohdista olisi voinut sopia useaan viitekehysten osa-alueeseen. Analysoin esimerkiksi kannanotot väittämään ”käyttäisin toisten tuottamia aineistoja, jos ne olisivat helposti saatavilla” uudelleenkäytön osa-alueessa, sillä painotin vastaajan halua *käyttää* toisten aineistoja. Se olisi voitu analysoida myös tiedeyhteisön yhteydessä, jos olisi painotettu vastaajan halua käyttää *toisten tutkijoiden tuottamia* aineistoja, tai käytettävyyden osa-alueessa, jos olisi keskitytty *saataavuuden helpouteen*. Väittäminen ”aineistoa saatetaan käyttää toisella tavalla kuin mihin se on tarkoitettu” käsiteltiin niin ikään aineistojen uudelleenkäytön osa-alueessa. Sekin olisi voinut sopia tiedeyhteisön osa-alueeseen, jos se olisi yhdistetty ensisijaisesti tieteen kehittämiseen. Samoin se olisi voinut sopia käytettävyyden osa-alueeseen, jos olisi painotettu väittämän rinnastettavuutta esimerkiksi väärintulkintojen mahdollisuuteen ja aineiston huonoon laatuun.

Tiedeinstituutioiden ja käytettävyyden osa-alueet olivat myös lähellä toisiaan. Sijoitin säilytyspaikkavalinnat tiedeinstituutioiden osa-alueeseen samoin kuin tutkijan ajan ja taitojen puutteiden tarkastelun. Ne olisi voitu liittää käytettävyydenkin osa-alueeseen, sillä kyse on myös tutkijoiden panostuksesta aineistojen säilyttämiseen ja jatkokäyttämiseen. Kuitenkin painotin sitä, että taitojen ja ajan puutteen korjaamiseksi juuri tiedeinstituutiot voisivat lisätä tutkijoille tarjottavaa koulutusta sekä muuttaa työajan kohdentamista, mikä tuli ilmi kirjallisuuskatsauksessakin. Myös kannanotot väittämiin ”olen tyytyväinen prosessiin aineistoni tallentamiseksi siksi aikaa, kun projekti on käynnissä” ja ”olen tyytyväinen prosessiin aineistoni tallentamiseksi myös projektin päättymisen jälkeistä aikaa ajatellen” olisi voitu analysoida osana käytettävyyden osa-aluetta, jos olisi painotettu

tutkijoiden tyytyväisyyttä tallentamisen *prosessiin*. Nyt painotin aineiston säilyttämistä, mikä liittyy keskeisesti paitsi säilytyspaikkoihin, myös tiedeinstituutioista rahoittajan aineistohallintasuunnitelmavaatimukseen: teoreettisesta viitekehyksestä (luku 2) muistetaan, että aineistohallinnan suunnittelu ja aineiston säilyttämisen suunnittelu ovat nähtävissä toistensa osa-alueina. Siten ne sisällytettiin tiedeinstituutioiden osa-alueeseen.

Viitekehys ohjasi myös muuttujien valitsemista. Kuten aiemmassa on käynyt ilmi, korkeakoulutausta oli keskeinen selittävä muuttuja. Kuitenkin muutamissa kohdin selittävänä muuttujana käytettiin muitakin, mutta viitekehysten asettamissa rajoissa. Kun selittävä muuttuja oli muu kuin korkeakoulu, kyse oli jonkin viitekehysten osa-alueeseen liittyneen teeman tarkentamisesta. Esimerkiksi aineiston suojelija -osa-alueessa haluttiin saada selville oikeudellisten normien dominoivuus tamperelaistutkijoiden parissa. Luonteva ennako-olettama oli, että jos aineistoja pidetään julkisuuteen kuulumattomina tai julkaisuoikeuksien nähdään puuttuvan, aineistoa ei jaeta, vaikka muut edellytykset täytyisivät, eli vaikka tutkija olisi esimerkiksi jo julkaissut omat tutkimustuloksensa ja hänellä olisi taitoja ja aikaa aineistojensa saataville saattamiseen. Jotta oikeudellisten normien merkitystä pystyttiin täsmentämään, oli perusteltua tarkastella normisidonnaisiksi tulkittujen aineistonajakamattajättämissyiden sekä oikeudellisten lupien merkitystä koskevien näkemysten vaikutusta aineistojen jakamismääriin. Oikeudellisia lupia ja aineiston jakamismääriä ei kuitenkaan käytetty selittävinä eikä selitettävänä muuttujina muissa viitekehysten osa-alueissa. Samoin esimerkiksi tutkijaa aineistohallinnan tuen saajana käsittelevässä osa-alueessa rahoittajalta tulevien vaatimusten vaikutusta selvitettiin ottamalla selittäväksi muuttujaksi ensisijaisen rahoittajan aineistohallintasuunnitelmavaatimus. Muuten aineistohallintasuunnitelmavaatimusta ei käytetty analyysissä selittävänä muuttujana, kuten selvitetty sen vaikutusta vaikkapa tutkijoiden näkemyksiin normeista.

Viitekehys teki siis analyysistä johdonmukaisen ohjailemalla sekä kyselylomakkeen eri osien merkitysten arvioimista että selittävien muuttujien käyttämistä. Sitä hyödyntämällä estettiin analyysin pirstaloituminen ja päällekkäisyydet sekä vedettiin rajoja monitulkinnaisille alueille. Kuitenkin on hyvä tiedostaa, että jaotteluissa oli kyse myös minun tutkijana tekemistäni subjektiivisista valinnoista.

## 5 TULOKSET

Tässä luvussa käyn läpi kyselyni tulokset. Tutkimuksen kohteena oli tamperelaistutkijoiden suhtautuminen tutkimusaineistojen avoimuuden toteutumiseen ja aineistojen hyödyntämiseen – heidän aineistonhallinnan näkemyksensä ja käytäntönsä – ja kyselyaineisto kerättiin Tampereen korkeakouluista alaluvussa 4.3 kuvatulla tavalla. Tulokset esitetään alaluvussa 4.4 läpikäydyn viitekehyksen mukaisesti (ks. myös liite 5). Alaluvussa 5.1 esittelen kyselyllä tavoitettujen tutkijoiden taustat ja perustiedot heidän aineistonhallinnastaan. Sitä seuraavista kuudesta alaluvusta kukin on omistettu yhdelle aineistonhallinnan viitekehyksen osa-alueista.

### 5.1 Tavoitetut vastaajat

Kyselyn 16 vastaajasta puolet on Tampereen yliopistosta (TaY), neljännes Tampereen teknillisestä yliopistosta (TTY) ja neljännes Tampereen ammattikorkeakoulusta (TAMK). Vastaajien enemmistö on naisia. TTY:n vastaajista kaikki ovat miehiä, mutta muissa korkeakouluissa miespuoliset vastaajat ovat vähemmistössä. Korkeakoulutaustan ja sukupuolijakauman välillä on tilastollisesti merkitsevä riippuvuus. Iältään vastaajat edustavat viittä vuosikymmentä 1950-luvulla syntyneistä 1990-luvulla syntyneisiin. Vastaajat asettuvat kuuteen eri asemaan työyhteisössä, mutta puolet on joko tohtoriopiskelijoita tai tutkijatohtoreita. TAMK:n vastaajajoukossa on kolmen eri tutkimusalan edustajia, TTY:n joukossa kahden ja TaY:n parissa kuuden. TTY:n ja TaY:n vastaajilla ei ole yhteistä tieteenalataustaa. Sen sijaan sekä TTY:sta että TAMK:sta on tekniikan, teknologian ja tietojenkäsittelyn alan edustajia ja TAMK:sta ja TaY:sta on molemmista terveysalan ja kasvatustieteiden edustajia. (Liitteet 6 a-d.)

Suomen Akatemia on vastaajien keskuudessa yleisin rahoittaja. Se on mahdollistamassa vastaajista noin kahden kolmasosan työntekoa, TTY:n edustajista jokaisen. Vastaajista seitsemäsosa nimeää tyypilliseksi rahoittajakseen opetus- ja kulttuuriministeriön (OKM), sosiaali- ja terveysministeriön (STM) ja Euroopan unionin (EU tai Euroopan sosiaalirahaston (ESR)). Lisäksi Tekes, Työsuojelurahasto (TSR) ja säätiötaso saavat yksittäisiä mainintoja, ja niin ikään yksi vastaaja sanoo tekevänsä tutkimusta ilman rahoitusta. Toisaalta se, että vastaajat nimeävät vaihtoehdon ”muu” yhteydessä useita rahoittajia, viittaa tutkijoiden vaikeuteen laittaa rahoittajia ensisijaisuusjärjestykseen. Esimerkiksi yksi vastaaja mainitsee muu-valintansa yhtenä tarkennuksena ”Työsuojelurahaston”, vaikka TSR oli lomakkeella vaihtoehdona. Toisaalta kyselylomake olisi voinut kaivata enemmän valmiita rahoittajavaihtoehtoja. (Liite 6 e.)

Tiedusteltaessa, millaisten aineistojen parissa tutkijat työskentelevät, TTY:n vastaajajoukossa käytetään lomakkeella tarjotusta kymmenestä aineistotyyppivaihtoehdosta kahta, kun taas muiden korkeakoulujen vastaajien parissa käytetään niistä viittä. Kaikkien korkeakoulujen vastaajajoukoissa on kokeellisia ja havainnoivia aineistoja hyödyntäviä. Ainakaan kaikki vastaajista eivät nojaa vain yhteen aineistotyyppiin. Eniten useita erilaisia aineistoja käyttäviä tutkijoita on yhteiskuntatieteellisen aineiston valitsevien joukossa ja vähiten kokeellisen aineiston valitsevien parissa. (Liitteet 6 f-g.)

Taulukko 3. Tutkijoiden jakaman tutkimusaineiston määrä.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Kuinka paljon saatat tutkimusaineistoasi muiden saataville? <sup>a, b</sup>	En lainkaan	2	50	2	50	-	-	4	25
	Jonkin verran	1	25	2	50	6	75	9	56,3
	Suurimman osan	1	25	-	-	2	25	3	18,8
	Kaiken	-	-	-	-	-	-	-	-
	En ole varma	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	
Vastausten jakautuminen dikotomisesti <sup>c</sup>	Ei jaa lainkaan	2	50	2	50	-	-	4	25
	Jakaa	2	50	2	50	8	100	12	75

a. p-arvo 0.133, Fisher-Freeman-Haltonin eksakti testi.

b. Keskiarvo 1,94 ja mediaani 2, kun 1 = en lainkaan, 2 = jonkin verran ja 3 = suurimman osan.

c. p-arvo 0.074, Fisher-Freeman-Haltonin eksakti testi.

Taulukossa 3 esitetyt tutkimusaineistojen jakamismäärät kertovat osaltaan tamperelaistutkijoiden aineistojen avoimuuden toteutumisesta. Vastaajista kolme neljäsosaa ilmoittaa jakavansa aineistojaan vähintäänkin jonkin verran, mutta kukaan ei sano jakavansa kaikkea. Neljännes ei jaa aineistojaan lainkaan. TaY:n vastaajajoukko erottuu siten, että he kaikki jakavat aineistojaan, kun taas puolet sekä TTY:n että TAMK:n vastaajista eivät jaa lainkaan. TTY:n vastaajat jakavat vähiten, sillä aineistoa jaetaan heidän joukossaan korkeintaan jonkin verran, kun taas TAMK:n ja TaY:n vastaajien keskuudessa on myös suurimman osan jakavia. Kukaan vastaajista ei ole epävarma jakamismäärästä, mistä päätellen he hallitsevat itse omia aineistojaan tai vähintäänkin tuntevat niihin liittyvät sopimukset.

Taulukko 4. Näkemykset aineiston jakamisesta ja jaetun aineiston avoimuudesta.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Jaan aineistojani muiden kanssa. <sup>a</sup>	Vahvasti samaa mieltä	-	-	-	-	1	12,5	1	6,3
	Osittain samaa mieltä	2	50	2	50	5	62,5	9	56,3
	En samaa enkä eri mieltä	-	-	-	-	1	12,5	1	6,3
	Osittain eri mieltä	-	-	-	-	1	12,5	1	6,3
	Vahvasti eri mieltä	2	50	2	50	-	-	4	25
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	

Muut tarvitsevat luvan käyttääkseen aineistojani. <sup>b</sup>	Vahvasti samaa mieltä	2	50	1	25	3	37,5	6	37,5
	Osittain samaa mieltä	1	25	-	-	4	50	5	31,3
	En samaa enkä eri mieltä	-	-	1	25	-	-	1	6,3
	Osittain eri mieltä	1	25	1	25	1	12,5	3	18,8
	Vahvasti eri mieltä	-	-	1	25	-	-	1	6,3
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	

a. p-arvo 0.385, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.527, Fisher-Freeman-Haltonin eksakti testi.

Jakamismäärät eivät vastaa täysin tutkijoiden kokemusta jakamisesta. Kun heiltä tiedustellaan jakamismäärän sijasta heidän suhtautumistaan väittämään ”jaan aineistojani muiden kanssa”, vahvasti eri mieltä väittämän kanssa olevien osuus vastaa vielä aineistonsa jakamatta jättävien osuutta. Sen sijaan vahvasti samaa mieltä väittämän kanssa olevista kaikki eivät jaa suurinta osaa aineistostaan. Toisaalta vaikka kukaan vastaajista ei jaa määrällisesti kaikkea aineistoaan, yksi kokee silti jakavansa merkittävästi. Vaikka kaikki vastaajat antoivat jakamismäärästään suoran vastauksen, yksi asettuu neutraalille kannalle jakamiskokemuksensa suhteen. (Taulukko 3; Taulukko 4.)

Aineistojen jakaminen ei merkitse automaattisesti niiden täydellistä avoimuutta. Vaikka enemmistö vastaajista jakaa tutkimusaineistojaan, noin kaksi kolmasosaa näkee pääsyn aineistoihinsa olevan luvanvaraista. TaY:n ja TAMK:n vastaajien aineistot ovat vähemmän avoimia, sillä heistä vähintään kolmen neljäsosan enemmistö on väittämän ”muut tarvitsevat luvan käyttääkseen aineistojani” takana, TTY:n vastaajista vähemmistö. Ei tosin ole varmaa, tulkitsevatko vastaajat väittämän viittaavan vastaamishetken olosuhteisiin vai miettivätkö he ihanteellisena pitämäänsä tilannetta. (Taulukko 4.) Toisaalta luvat voivat myös lisätä avoimuutta: aineistonsa jakamatta jättävistä puolet puoltaa vahvasti lupia, mikä voi viitata osan heistä pitävän aineiston antamista muiden käyttöön tietyissä olosuhteissa mahdollisena. Jonkin verran jakavista vajaa puolet kannattaa vahvasti lupia. Sen sijaan suurimman osan jakavat vastaajat eivät asetu vahvasti käyttölupavaatimuksen taakse, mikä voi kertoa siitä, että heidän aineistonsa eivät ole arkaluontoisia. (Liite 6 h.)

Yksi TaY:n tutkijoista mainitsee kyselyn Vapaa sana -osiossa aineistonjakamiskäyttäytymisensä muuttuneen ja jakaa ajattelutapansa väitöskirjaa edeltävään ja seuraavaan aikaan. Siten urakehityksen vaikutus aineistonhallinnan näkemyksiin ja käytäntöihin tiedostetaan. Hän kirjoittaa:

”Aineistojen jakaminen omalla kohdallani on aika uusi juttu. Kerätessäni aineistoja väitöskirjaani varten en juurikaan ajatellut koko asiaa. Nyt aineistojen tallentaminen on selvä asia. Kaikkia aineistoja ei voi kuitenkaan jakaa, eikä kaiken tallentamisessa ole järkeä.”



Itse kyselyssä keskityttiin kuitenkin kartoittamaan menneen ja nykyisen tilanteen vertailun sijasta aineistonhallinnan nykyhetkeä ja tulevaisuuden näkymiä. Taulukko 5 havainnollistaa tutkimusaineiston saatavuuden rajoittamisen vaikutusta vastaajien haluihin jakaa aineistoja:

Taulukko 5. Samanmielisyyss väittämän ”saattaisin todennäköisemmin aineistojani saataville, jos voisin asettaa saatavuudelle ehtoja” kanssa.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Saattaisin todennäköisemmin aineistojani saataville, jos voisin asettaa saatavuudelle ehtoja. <sup>a</sup>	Vahvasti samaa mieltä	-	-	1	25	3	37,5	4	25
	Osittain samaa mieltä	2	50	-	-	3	37,5	5	31,3
	En samaa enkä eri mieltä	1	25	-	-	2	25	3	18,8
	Osittain eri mieltä	1	25	1	25	-	-	2	12,5
	Vahvasti eri mieltä	-	-	2	50	-	-	2	12,5
	En osaa sanoa	-	-	-	-	-	-	-	-
Yhteensä		4		4		8		16	

a. p-arvo 0.154, Fisher-Freeman-Haltonin eksakti testi.

Rajoituksille annetaan arvoa, sillä yli puolet kokee, että lisäehtojen asettaminen aineistojen saatavuudelle voisi motivoida heitä jakamaan enemmän. Vähintään osittain kielteisiä on neljännes ja neutraalejakin on viidennes eli heidän jakamiskäyttäytymistään ohjaavat muut tekijät kuin ehdot. TTY:n vastaajista enemmistö torjuu ajatuksen jakamishalujensa lisääntymisestä. Sen sijaan kolme neljäsosaa TaY:n vastaajista on vähintään osittain myönteisiä eikä yksikään ilmaise suoraan kielteisyttä. (Taulukko 5.)

Taulukko 6. Aineiston jakamismääränsä ilmoittaneiden samanmielisyyss väittämän ”saattaisin todennäköisemmin aineistojani saataville, jos voisin asettaa saatavuudelle ehtoja” kanssa.

		Kuinka paljon saatat tutkimusaineistoasi muiden saataville?							
		En lainkaan		Jonkin verran		Suurimman osan		Yhteensä	
		n	%	n	%	n	%	n	%
Saattaisin todennäköisemmin aineistojani saataville, jos voisin asettaa saatavuudelle ehtoja. <sup>a</sup>	Vahvasti samaa mieltä	-	-	3	33,3	1	33,3	4	25
	Osittain samaa mieltä	1	25	2	22,2	2	66,7	5	31,3
	En samaa enkä eri mieltä	1	25	2	22,2	-	-	3	18,8
	Osittain eri mieltä	-	-	2	22,2	-	-	2	12,5
	Vahvasti eri mieltä	2	50	-	-	-	-	2	12,5
	En osaa sanoa	-	-	-	-	-	-	-	-
Yhteensä		4		9		3		16	

a. p-arvo 0.360, Fisher-Freeman-Haltonin eksakti testi.

Toisaalta voisi olettaa, että jakamisen lisäämishaluihin vaikuttaa se, miten paljon tutkija jakaa jo aineistojaan eli miten paljon hän näkee lisäämisen varaa olevan. Taulukko 6 näyttää, että kaikki jo nyt suurimman osan aineistoistaan jakavista voisivat olla halukkaita lisäämään entisestään jakamis-

taan ehtojenasettamismahdollisuuksien myötä. Jonkin verran aineistojaan jakavistakin yli puolta li- säehdot voisivat motivoida jakamaan enemmän, eikä kukaan heistä ei torjunut ajatusta vahvasti. Sen sijaan aineistonsa jakamatta jättävistä puolet näkee, että ehtojen lisääminen ei vaikuttaisi hei- dän jakamishaluihinsa.

## 5.2 Tutkimusaineistojen kerääjä – jaossa oleva kunnia

”Tieteenalallani aineiston keräämisen toteuttamiseen osallistuu usein tutkimusavus- taja/avustajia tai aineisto kerätään useamman tutkijan voimin. Kun vastasin tästä näkö- kulmasta, jotkin väitteet eivät toimineet kovin hyvin. Se, että aineiston kerääjä/tuottaja voi tarkoittaa monia eri asioita, olisi hyvä huomioida tuloksia tulkittaessa.”

Näin yksi TaY:n tutkijoista kommentoi aineiston kerääjän roolia kyselylomakkeen lopun Vapaa sana -osiossa. Hänen kaltaisiaan voi olla vastaajissa enemmänkin. Tässä alaluvussa tutkin, minkälaisen aineiston kerääjän roolin tamperelaistutkijat ottavat, tarkastelemalla, missä määrin he haluavat tulla yhdistetyiksi keräämiinsä tutkimusaineistoihin ja ajaa omaa etuaan aineiston kerääjänä – tai mahdollisesti aineiston kerääjän etua yleisesti. Huomio on siis kunnia ja eduissa, jotka ohjaavat aineistojen avoimuutta joko estäen sen tai tehden siitä ehdollista.

Taulukko 7. Aineiston kerääjän etuihin liittyvät syyt jättää tutkimusaineistoja jakamatta.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Menettäisin mahdollisuuden hallita aineistojani itse. <sup>a</sup>	Valittu syyksi	-	-	-	-	-	-	-	-
	Ei valittu syyksi	4	100	4	100	8	100	16	100
	Yhteensä	4		4		8		16	
Minun tarvitsee julkaista ensin omat tutkimustulokseni. <sup>b</sup>	Valittu syyksi	1	25	3	75	5	62,5	9	56,3
	Ei valittu syyksi	3	75	1	25	3	37,5	7	43,8
	Yhteensä	4	100	4	100	8	100	16	100
Valinnat yhteensä <sup>c</sup>	Valittu	1	12,5	3	37,5	5	31,3	9	28,1
	Ei valittu	7	87,5	5	62,5	11	68,8	23	71,9
	Yhteensä	8		8		16		32	

a. Fisher-Freeman-Haltonin eksaktia testiä ei voitu suorittaa, sillä rivimuuttujan arvoissa ei ollut vaihtelevuutta.

b. p-arvo 0.530, Fisher-Freeman-Haltonin eksakti testi.

c. p-arvo 0.674, Fisher-Freeman-Haltonin eksakti testi.

Aineiston kerääjän etuihin liitetyistä kahdesta syystä, joiden vuoksi aineistoa jätetään jakamatta, tutkijan halu julkaista ensin omat tutkimustuloksensa on yleisin. Vastaajista yli puolet nimeää sen syyksi – joskin yli 40 prosenttia ei tee niin. TAMK:n vastaajat erottuivat syyn sivuuttami- sessa, sillä heistä vain yksi neljästä pitää sitä olennaisena, kun taas kahden muun korkeakoulun vas- taajien enemmistöt näkevät sillä olevan vaikutusta. Sen sijaan yksikään vastaajista ei sano pitävänsä aineistoaan vain itsellään siksi, että pelkäisi menettävänsä sen hallinnan. (Taulukko 7.) Sitä voi pitää

vakaana kantana, sillä vastaajien oli mahdollista valita kaikki syyt, joita piti vähänkin jakamiskäyttämistä ohjaavina tekijöinä.

Liite 10 a näyttää, että kuudesosa kaikista lomakkeen esittämistä aineistonjakamattajättämisyistä luokiteltiin aineiston kerääjän rooliin liittyviksi. Osa-alueen vaikutus on osuuttaan merkittävämpi, sillä toinen kahdesta syystä, tutkijan tarve julkaista ensin omat tutkimustuloksensa, kerää yksinään viidenneksen kaikista jakamatta jättämistä selittämään tehdyistä valinnoista. TTY:n vastaajien parissa se kattaa keskimääräistä isomman osan, neljänneksen, kaikista syyvalinnoista ollen toinen heidän keskeisimmistä syistä olla tuomatta aineistoja muiden saataville. Sen sijaan TaY:n vastaajien parissa kyseisen kilpailuhenkisen syyn osuus kaikista syyvalinnoista on viidennes ja TAMK:n vastaajien keskuudessa vain seitsemäsosa.

Taulukko 8. Haluaako tutkija asettaa aineistojen kerääjän rooliin liitettyjä ehtoja aineistojen uudelleenkäytölle?

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Jaettu tekijyys (co-authorship) koskien julkaisuja, joissa aineistoa on käytetty <sup>a</sup>	Kyllä	4	100	-	-	1	12,5	5	31,3
	En	-	-	3	75	4	50	7	43,8
	En ole varma	-	-	1	25	3	37,5	4	25
	Yhteensä	4		4		8		16	
Aineiston kerääjälle annetaan tunnustus kaikissa töissä, joissa aineistoa on käytetty. <sup>b</sup>	Kyllä	3	75	3	75	8	100	14	87,5
	En	-	-	1	25	-	-	1	6,3
	En ole varma	1	25	-	-	-	-	1	6,3
	Yhteensä	4		4		8		16	
Aineiston kerääjään viitataan kaikissa töissä, joissa aineistoa on käytetty. <sup>c</sup>	Kyllä	3	75	3	75	5	62,5	11	68,8
	En	-	-	1	25	1	12,5	2	12,5
	En ole varma	1	25	-	-	2	25	3	18,8
	Yhteensä	4		4		8		16	
Yhteensä <sup>d</sup>	Kyllä	10	83,3	6	50	14	58,3	30	62,5
	En	-	-	5	41,7	5	20,8	10	20,8
	En ole varma	2	16,7	1	8,3	5	20,8	8	16,7
	Yhteensä	12		12		24		48	
Vastausten jakautuminen dikotomisesti <sup>e</sup>	Kyllä	10	100	6	54,5	14	73,7	30	75
	En	-	-	5	45,5	5	26,3	10	25
	Yhteensä	10		11		19		40	

a. p-arvo 0.017, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.233, Fisher-Freeman-Haltonin eksakti testi.

c. p-arvo 1.000, Fisher-Freeman-Haltonin eksakti testi.

d. p-arvo 0.128, Fisher-Freeman-Haltonin eksakti testi.

e. p-arvo 0.047, Fisher-Freeman-Haltonin eksakti testi.

Kun huomio on siinä, mitkä aineiston kerääjän kunniaan ja etuihin liittyvät ehdot vuorostaan kannustaisivat tutkijoita avaamaan aineistoaan uudelleenkäytölle, selvä enemmistö, noin kaksi kolmasosaa, haluaa kerääjän saavan viittauksen kaikissa töissä, joissa tämän aineistoja käytetään. Viittaminen ei ole ainoa varteenotettavana pidettävä kunnian jakamisen keino, sillä vieläkin isompi osuus vastaajista, yli neljä viidesosaa, toivoo yleisemmin tunnustusta kerääjälle aina, kun tämän aineistoa käytetään. Kumpikin ehto saa vastaajilta yli puolet enemmän puoltavia kuin torjuvia kantoja. Sen sijaan kolmannes eli huomattavasti harvempi haluaa jaetun tekijyyden koskien julkaisuja, joissa hänen keräämäänsä aineistoa on käytetty. (Taulukko 8.)

Korkeakoulu vaikuttaa tilastollisesti merkitsevästi vastaajien suhtautumiseen aineiston kerääjän etuja ajaviin aineiston avoimuuden ehtoihin kokonaisuudessaan: siinä missä TAMK:n vastaajien enemmistö puoltaa ehtoja eikä kukaan heistä vastusta niitä, muissa korkeakouluissa kannat hajaantuvat. Huomion ollessa vuorostaan yksittäisissä ehdoissa vastaajien korkeakoulutausta vaikuttaa tilastollisesti merkitsevästi heidän suhtautumiseensa jaettuun tekijyyteen. TAMK:n vastaajat erottuvat olemalla yksimielisiä kyseisen ehdon tärkeydestä, kun taas muiden enemmistöt torjuvat sen. Kahteen muuhun ehtoon suhtautumisessa ei ole tilastollisesti merkitseviä eroja, vaan enemmistöt kustakin korkeakoulusta puoltavat niitä. TaY:n vastaajat haluavat vieläpä yksimielisesti aineiston kerääjälle tunnustusta kaikissa töissä, joissa aineistoa on käytetty. (Taulukko 8.)

Liite 11 vahvistaa kuvaa siitä, että aineiston kerääjän roolissa ollaan halukkaita asettamaan ehtoja aineistojen avoimuudelle: rooliin liitettiin neljännes kaikista kyselylomakkeella tarjotuista ehtovaihtoehtojen, mutta ne keräävät kaksi viidesosaa kaikista annetuista vastauksista. Korkeakouluvertailussa rooli on tärkein TTY:n vastaajajoukolle, sillä siihen liittyvät ehdot saavat miltei seitsemän kahdeksasosaa kaikista heidän tekemistään ehtoalinnoista, mutta vain noin kaksi viidesosaa TAMK:n vastaajien kaikista valinnoista ja kolmanneksen TaY:n valinnoista.

Taulukko 9. Samanmielisyyden tarpeesta viitata tutkimusaineistoihin.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
On tärkeää, että aineistoihini viitataan, jos muut tutkijat käyttävät niitä. <sup>a</sup>	Vahvasti samaa mieltä	4	100	2	50	6	75	12	75
	Osittain samaa mieltä	-	-	1	25	1	12,5	2	12,5
	En samaa enkä eri mieltä	-	-	-	-	1	12,5	1	6,3
	Osittain eri mieltä	-	-	-	-	-	-	-	-
	Vahvasti eri mieltä	-	-	1	25	-	-	1	6,3
	En osaa sanoa	-	-	-	-	-	-	-	-
Yhteensä		4		4		8		16	

a. p-arvo 0.633, Fisher-Freeman-Haltonin eksakti testi.

Aineiston kerääjälle osoitettavan kunnian teemaa lähestyttiin myös tiedustelemalla, pitääkö tutkija aineistoihinsakin viittaamista tärkeänä, jos muut tutkijat käyttävät niitä. Seitsemän kahdeksasosaa pitää sitä tärkeänä, ja heistä valtaosa on sitä mieltä vahvasti. Osuus on siten vieläkin suurempi kuin aineiston kerääjään viittaamista edellyttävien osuus. TAMK:n vastaajista kaikki pitävät yksimielisesti erittäin tärkeänä, että heidän aineistonsa saavat viittaukset, TaY:n vastaajista kolme neljäsosaa ja TTY:n vastaajista puolet. (Taulukko 9.) Aineiston kerääjään kohdistettavia viittauksia vaativista kaikki kaipaavat viittausta myös aineistoihin. Kuitenkin aineiston kerääjään viittaamista tarpeettomana jakamisen ehtona pitävästä kaksikosta toinen torjuu vahvasti aineistoonkin viittaamisen tarpeen, mutta toinen suhtautuu asiaan neutraalisti pitäen tarvetta kenties epäolennaisena. Kiinnostavaa on, että kolmikko, joka ei ole varma tarpeesta viitata aina aineiston kerääjään, pitää kuitenkin aineistoihin viittaamista tärkeänä. (Liite 7.) Tulkinnanvaraista on, viittaavatko kyseiset pienet erot vastausten epä johdonmukaisuuteen vai vivahde-eroille annettuun merkitykseen.

### 5.3 Tutkimusaineistojen suojelija – noudatettavat normit

Tässä alaluvussa huomio siirtyy tutkijaan liittyvistä kunniakysymyksistä itse aineistoihin liittyviin normeihin ja niistä mahdollisesti seuraaviin aineiston avoimuuden rajoituksiin. Selvitän, millaisia aineiston suojelijoita vastaajat ovat eli miten he suhtautuvat erilaisiin normeihin niin aineiston jakamatta jättämisen syinä kuin jakamisen ehtoinakin. Kyselylomakkeella tarjotuista normiperusteisista avoimuuden esteistä ”aineistojeni ei kuulu olla julkisia” voi olla osin subjektiivinen mielipide tai sosiaalisista normeista juontava näkemys, kun taas syy ”minulla ei ole oikeuksia julkaista aineistojani” takana voi olla esimerkiksi lainsäädäntöön tai sopimukseen liittyviä objektiivisempia tekijöitä. Toisaalta kyseiset syyt saatetaan tulkita toistensa synonyymeiksi. Ne kulkevat rinnakkain sikäli, että neljä viidesosaa niistä vastaajista, jotka näkevät, että heidän aineistojensa ei kuulu olla julkisia, nimeävät jakamattomuuden syyksi myös julkaisuoikeuksien puuttumisen. Kuitenkaan viidesosa ei siis tee niin. Runsas kolmannes niistä, jotka eivät pidä aineistojen kuulumattomuutta julkisuuteen jakamatta jättämisen syynä, valitsee silti syyksi julkaisuoikeuksien puuttumisen. (Liite 10 c.)

Taulukko 10. Aineiston suojelijan rooliin liittyvät syyt jättää aineistoja jakamatta.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Aineistojeni ei kuulu olla julkisia. <sup>a</sup>	Valittu syyksi	2	50	-	-	3	37,5	5	31,3
	Ei valittu syyksi	2	50	4	100	5	62,5	11	68,8
	Yhteensä	4		4		8		16	
Minulla ei ole oikeuksia julkaista	Valittu syyksi	1	25	-	-	7	87,5	8	50
	Ei valittu syyksi	3	75	4	100	1	12,5	8	50

aineistojani. <sup>b</sup>	Yhteensä	4		4		8		16	
Valinnat yhteensä <sup>c</sup>	Valittu syyksi	3	37,5	-	-	10	62,5	13	41
	Ei valittu syyksi	5	62,5	8	100	6	37,5	19	59
	Yhteensä	8		8		16		32	

a. p-arvo 0.487, Fisher-Freeman-Haltonin eksakti testi. Havaintoyksikkönä vastaajat.

b. p-arvo 0.010, Fisher-Freeman-Haltonin eksakti testi. Havaintoyksikkönä vastaajat.

c. p-arvo 0.014, Fisher-Freeman-Haltonin eksakti testi. Havaintoyksikkönä valintakerrat.

Aineiston suojelijan rooliin liittyvien eli normiperusteisten aineistonjakamattajättämisyiden saamista kannanotoista kaksi viidesosaa liittyy niiden merkityksellisenä pitämiseen. Puolet vastaajista nimeää julkaisu oikeuksien puuttumisen ja kolmannes aineistojensa julkisuuteen kuulumattomuuden syyksi pitää aineistot muiden saavuttamattomissa. Korkeakoulutaustan vaikutus näkemyksiin kyseisistä esteistä kokonaisuudessaan on tilastollisesti merkitsevä: TTY:n vastaajista kukaan ei selitä ratkaisuaan olla tuomatta aineistojaan muiden saataville kummallakaan normikeskeisistä syistä ja myös TAMK:n vastaajajoukko kallistuu kielteiselle kannalle, kun taas TaY:n vastaajien valintakerroista liki kahden kolmasosan enemmistö on syiden vaikutusta puoltavia. Kun avoimuuden esteitä tarkastellaan yksittäin, korkeakoulutausta vaikuttaa tilastollisesti merkitsevästi näkemyksiin julkaisu oikeuksien puuttumisesta. Siinä missä TTY:n kaikki vastaajat ja TAMK:nkin vastaajien enemmistö torjuvat sen vaikutuksen, TaY:n vastaajille se on keskeinen este: se ohjaa seitsemää kahdeksasosaa heistä jättämään aineistoja jakamatta. Sen sijaan aineistojen kuulumattomuus julkisuuteen ei jaa tutkijoita yhtä selvästi, vaan vähintään puolet kunkin korkeakoulun vastaajista torjuu kyseisen syyn vaikutuksen. (Taulukko 10.)

Kyselyssä tarjottiin useita valmiita vaihtoehtoja selittämään aineiston jakamatta jättämistä, mutta kolme vastaajaa muotoilee omankin vastauksen. Kaksi niistä, molemmat TaY:n vastaajilta, liittyy normeihin, joten aineiston suojelijan rooli herättää ajatuksia. Kummankin vastauksen aiheena on havainnointiaineistojen anonymisointi eli tutkimuskohteiden suojelu. Toinen tutkijoista kirjoittaa: "Aineiston anonymisointi on osin mahdotonta." Toinen toteaa: "Anonymisointi pienissä havainnointiaineistoissa käytännössä tuhoaa aineiston", ja pohtii vielä kyselyn lopun Vapaa sana -osiossa: "[...] Haastavinta jakamisessa on se, että havainnointiaineistoja on todella hankala anonymisoida, ne ovat erittäin henkilökohtaisia." He kumpikin nimeävät avoimuuden esteiksi myös julkaisu oikeuksien puuttumisen, mutta eivät aineistojen kuulumattomuutta julkisuuteen. (Liite 10 b.)

Aineiston suojelijan rooliin liitettiin kuudesosa aineistonjakamattajättämisyistä, mutta runsas neljännes kaikista tehdyistä syyvalinnoista liittyy siihen. Siten sen asema aineistojen avoimuutta rajoittaessa on osuuttaan merkittävämpi. TTY:n vastaajien käyttäytymiseen suojelijan rooli ei vaikuta,

mutta sekä TaY:n että TAMK:n vastaajien kaikista syyvalinnoista noin kaksi viidesosaa liittyy siihen. Lisäksi yksittäisistä syistä julkaisu-oikeuksien puuttuminen on TaY:n vastaajien parissa yleisin avoimuuden este kattaen noin neljänneksen kaikista heidän valinnoistaan. Aineistojen kuulumattomuus julkisuuteen taas kerää noin kolmanneksen TAMK:n vastaajien syyvalinnoista ollen toinen heidän keskeisimmistä aineistonjakamattajättämissyistä. (Liite 10 a.) Taulukko 11 näyttää, missä määrin ne, jotka näkevät normisidonnaisten syiden ohjaavan jakamistaan, jakavat lopulta aineistojaan.

Taulukko 11. Aineiston suojelijan rooliin liittyvien syiden aineistojaan jakamatta jättävien aineistonjakamismäärät.<sup>c</sup>

		Avoimuuden este: aineistojeni ei kuulu olla julkisia					
		Valittu syyksi		Ei valittu syyksi		Yhteensä	
		n	%	n	%	n	%
Kuinka paljon saatat tutkimusaineistoasi muiden saataville? <sup>a</sup>	En lainkaan	1	20	3	27,3	4	25
	Jonkin verran	3	60	6	54,5	9	56,3
	Suurimman osan	1	20	2	18,2	3	18,8
	Yhteensä	5		11		16	
		Avoimuuden este: minulla ei ole oikeuksia julkaista aineistojani					
		Valittu syyksi		Ei valittu syyksi		Yhteensä	
		n	%	n	%	n	%
Kuinka paljon saatat tutkimusaineistoasi muiden saataville? <sup>b</sup>	En lainkaan	1	12,5	3	37,5	4	25
	Jonkin verran	6	75	3	37,5	9	56,3
	Suurimman osan	1	12,5	2	25	3	18,8
	Yhteensä	8		8		16	

a. p-arvo 1.000, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.413, Fisher-Freeman-Haltonin eksakti testi.

c. Tässä ei ole tehty suodatusta, joka sulki pois sen, että vaihtoehdon "aineistojeni ei kuulu olla julkisia" olisi valinnut myös vaihtoehdon "minulla ei ole oikeuksia julkaista aineistojani".

Taulukko osoittaa, että aineiston suojelijan rooliin liitettyjen syiden vuoksi aineistojaan jakamatta jättävät eivät erotu muita vähemmän aineistojaan jakavina. Isommat osuudet niistä vastaajista, joiden jakamispäätöksiä kyseiset syyt eivät ohjaa, jättävät aineistojaan jakamatta. Toisaalta on muistettava, että vastaajat käyttävät keskimäärin enemmän kuin yhtä aineistotyyppiä. Eri aineistotyyppien julkaisemiseen voidaan luonnollisesti suhtautua eri tavoin, minkä lisäksi myös tutkimusaiheet vaikuttavat aineistojen sisältöjen arkaluontoisuuteen. Siten ne, jotka jakavat jonkin verran aineistojaan tai suurimman osan aineistoistaan, voivat jakaa sellaisia aineistojensa osia, joihin ei liity normisidonnaisia kysymyksiä. Kuitenkin erityisjärjestelyt voivat mahdollistaa arkaluontoistenkin aineistojen jakamisen, minkä vuoksi on aiheellista tarkastella oikeudellisten lupien vaikutusta aineistojen avoimuuteen. Kyselyssä tutkijoita pyydettiin ottamaan kantaa väittämään "aineiston käytölle on hankittava oikeudelliset luvat" jaetun aineiston uudelleenkäytön edellytyksenä. Näin ollen, jos tutkimusaineistoa jaetaan, voitaisiin valvoa, että sitä hyödynnetään asianmukaisesti.

Taulukko 12. Haluaako tutkija asettaa aineistojen suojelijan rooliin liitetyn ehdon aineistojen uudelleenkäytölle?

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Aineiston käytölle on	Kyllä	3	75	-	-	3	37,5	6	37,5
hankittava oikeudelliset	En	-	-	4	100	1	12,5	5	31,3
luvut. <sup>a</sup>	En ole varma	1	25	-	-	4	50	5	31,3
	Yhteensä	4		4		8		16	

a. p-arvo 0.015, Fisher-Freeman-Haltonin eksakti testi.

Suhtautuminen tarpeeseen hankkia aineiston uudelleenkäytölle oikeudelliset luvat jakaa vastaajat kolmeen varsin samankokoiseen joukkoon: epävarmoinhin, puoltajiin ja vastustajiin. On tulkittavissa, että vastaajista kolmanneksella on hallussaan arkaluontoista tutkimusaineistoa, johon pääsy halutaan rajata oikeudellisin luvuin. Korkeakoulutaustan vaikutus näkemyksiin on tilastollisesti merkitsevää. TTY:n vastaajat erottuvat yksinäisyydellään, sillä heistä kukaan ei näe oikeudellisia lupia tarpeellisina. TAMK:n vastaajat ovat puolestaan oikeudellisten lupien puolesta puhujia, sillä heistä enemmistö kannattaa ehtoa ja vähemmistökin on suoran kielteisyyden sijasta vain epävarma kannastaan. TaY:n vastaajista runsas kolmannes näyttää oikeudellisille luvulle vihreää valoa, mutta puolet on epävarmoja kannastaan. (Taulukko 12.) Aineiston suojelijan rooliin liitettiin vajaa kymmenesosa kaikista aineiston avoimuuden ehdoiksi kyselylomakkeella tarjotuista vaihtoehdoista, ja se myös keräsi noin kymmenesosan kaikista tehdyistä ehtovalinnoista. Sekä TAMK:n että TaY:n vastaajajoukot asettuvat kyseiseen rooliin yhtä vahvasti, kun taas kävi jo ilmi, että TTY:n vastaajat eivät asettaisi lainkaan normeihin liittyviä ehtoja aineistojensa jakamiselle. (Liite 11.)

Taulukko 13. Oikeudellisten lupien tarpeeseen kantaa ottavien näkemykset aineistojensa jakamismääristä.

		Pidätkö ehdon "aineiston käytölle on hankittava oikeudelliset luvat" täyttymistä keskeisenä, ennen kuin muut voivat käyttää tutkimusaineistojasi?					
		Kyllä		En		En ole varma	
		n	%	n	%	n	%
Kuinka paljon saatat	En lainkaan	2	33,3	2	40	-	-
tutkimusaineistoasi	Jonkin verran	4	66,7	3	60	2	40
muiden saataville? <sup>a</sup>	Suurimman osan	-	-	-	-	3	60
	Yhteensä	6		5		5	

a. p-arvo 0.110, Fisher-Freeman-Haltonin eksakti testi.

Taulukko 13 osoittaa, että näkemyserot oikeudellisista luvista eivät tee vastaajien kesken merkittävää eroa aineistonjakamiskäytäntöihin: oikeudellisten lupien puoltajat ja torjujat eivät eroa jakamismääriltään toisistaan. Sitä vastoin näkemyksistään epävarmat jakavat määrällisesti heitä enemmän aineistojaan.



## 5.4 Tiedeyhteisön jäsen – kohti yhteisiä tavoitteita?

Tässä tutkijan roolia tiedeyhteisön jäsenenä tutkivassa alaluvussa tarkastelen ensiksikin, millaisen merkityksen tamperelaisten korkeakoulujen tutkijat antavat tutkimusaineistojen avoimuudelle. Kysely selvitti, miten tärkeänä he pitävät mahdollisuutta käyttää toisten tutkijoiden ja korkeakoulujen keräämiä aineistoja sekä oman alan ja työnteon että tieteen yleisemmän kehityksen näkökulmista. Kyselylomakkeessa ei määritelty tarkemmin ”omaa tieteenalaa”, joten vastaaja voi tulkita sen viittaavan niin oman tutkimusryhmän jäsenten aineistoihin kuin vaikkapa saman alan kansainvälistenkin tutkijapiirien käyttämiin aineistoihin. Toiseksi tarkastelen aineistojen jakamisen yhteisöllisyyttä siitä näkökulmasta, miten helposti tarjolla olevina tutkijat pitävät omia aineistojaan ja miten he suhtautuvat jakamisen vastavuoroisuuteen.

Taulukko 14. Näkemykset oman tieteenalan ja paikallisten aineistojen heikon saavutettavuuden vaikutuksista tieteelle ja tutkimukselle.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Pääsyn puuttuminen toisten oman tieteenalani tutkijoiden tuottamiin aineistoihin on suuri este tieteen kehitykselle. <sup>a</sup>	Vahvasti samaa mieltä	2	50	1	25	-	-	3	18,8
	Osittain samaa mieltä	-	-	-	-	4	50	4	25
	En samaa enkä eri mieltä	1	25	-	-	1	12,5	2	12,5
	Osittain eri mieltä	-	-	1	25	2	25	3	18,8
	Vahvasti eri mieltä	1	25	2	50	1	12,5	4	25
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	
Pääsyn puuttuminen toisten tamperelaisten korkeakoulujen tutkimusaineistoihin on este tutkimuksen tekemiselle. <sup>b</sup>	Vahvasti samaa mieltä	1	25	2	50	-	-	3	18,8
	Osittain samaa mieltä	-	-	-	-	-	-	-	-
	En samaa enkä eri mieltä	1	25	-	-	-	-	1	6,3
	Osittain eri mieltä	-	-	-	-	2	25	2	12,5
	Vahvasti eri mieltä	1	25	2	50	6	75	9	56,3
	En osaa sanoa	1	25	-	-	-	-	1	6,3
	Yhteensä	4		4		8		16	

a. p-arvo 0.159, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.080, Fisher-Freeman-Haltonin eksakti testi.

Runsas kaksiviidesosa vastaajista pitää pääsyn puuttumista oman tieteenalansa tutkijoiden aineistoihin suurena esteenä tieteen kehitykselle, mutta aihe jakaa näkemyksiä, sillä samankokoinen osuus on myös toista mieltä. Toisaalta TAMK:n ja TaY:n vastaajajoukoista molemmista puolet pitää kyseisenlaista tilannetta ongelmana, mutta TTY:n vastaajista enemmistö ei pidä. Toisaalta sekä TTY:n että TAMK:n vastaajien parissa aineistojen avoimuuden rajoittamista tieteen kannalta ongelmallisena pitävät seisovat linjansa takana vahvasti TaY:n vastaajien ollessa kannoiltaan varovaisempia. (Taulukko 14.)

Kun huomio siirretään tutkimusaineistojen avoimuuteen paikallisessa tiedeyhteisössä, vastaajista noin viidenneksen mukaan olisi este tutkimuksen tekemiselle, jos he eivät pystyisi ylittämään Tampereen korkeakoulujaon rajoja aineistojen käytössä. Noin kahden kolmasosan mukaan se ei olisi este. Tosin osittain esteenä pitäviä ole yhtään, joten kun pääsyn puuttumista pidetään esteenä, sitä pidetään sellaisena vahvasti. TTY:n vastaajanelikko jakautuu kahtia siten, että puolet pitää pääsyn puuttumista naapurikorkeakoulujen tutkimusaineistoihin vahvasti esteenä tutkimustyölle ja puolet on jyrkästi toista mieltä. Sitä vastoin kaikki TaY:n vastaajat torjuvat pääsyn puuttumisen haitat. TAMK:n vastaajanelikosta puolet ei asetu kummallekaan kannalle. (Taulukko 14.)

Taulukko 15. Näkemykset omien aineistojen saatavuuden laajuudesta ja vaivattomuudesta.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Muut saavat helposti aineistojani käyttöönsä. <sup>a</sup>	Vahvasti samaa mieltä	-	-	-	-	1	12,5	1	6,3
	Osittain samaa mieltä	-	-	-	-	3	37,5	3	18,8
	En samaa enkä eri mieltä	1	25	-	-	-	-	1	6,3
	Osittain eri mieltä	-	-	3	75	4	50	7	43,8
	Vahvasti eri mieltä	3	75	1	25	-	-	4	25
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	
Olen halukas jakamaan aineistoja laajan tutkijajoukon kesken. <sup>b</sup>	Vahvasti samaa mieltä	2	50	1	25	4	50	7	43,8
	Osittain samaa mieltä	1	25	1	25	3	37,5	5	31,3
	En samaa enkä eri mieltä	-	-	-	-	-	-	-	-
	Osittain eri mieltä	1	25	1	25	1	12,5	3	18,8
	Vahvasti eri mieltä	-	-	1	25	-	-	1	6,3
	En osaa sanoa	-	-	-	-	-	-	-	-
Yhteensä	4		4		8		16		

a. p-arvo 0.017, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.888, Fisher-Freeman-Haltonin eksakti testi.

Taulukko 15 havainnollistaa tutkijoiden näkemyksiä omien aineistojensa saavutettavuuden helpoudesta ja omista panostuksistaan aineistojen avaamiseen. Noin kaksi kolmasosaa vastaajista näkee, että heidän tutkimusaineistonsa eivät ole helposti muiden saavutettavissa. Korkeakoulujen vaikutus vastausten jakautumiseen on tilastollisesti merkitsevää. Ainoastaan TaY:n vastaajien joukossa arvioidaan, että heidän aineistonsa ovat helposti saatavilla – heistä puolet on sitä mieltä ja loputkin vain osittain toista mieltä. Kaikki TTY:n vastaajat ja enemmistö TAMK:n vastaajista torjuvat väittämän aineistojensa helposta saatavuudesta vähintään osittain. Vaikka enemmistö vastaajista ei näe omien tutkimusaineistojensa olevan helposti muiden saatavilla, kolme neljäsosaa vastaajista olisi valmis jakamaan aineistojaan laajan tutkijan joukon kanssa ja lopuistakin suurin osa vastustaa ajatusta vain osittain. Se voi osaltaan viitata ehtojen asettamisen tärkeyteen. Yksikään ei ole epävarma

suhtautumisestaan laajalle joukolle jakamiseen. Kyselyssä ei täsmennetty, koostuuko ”laaja tutkijajoukko” tutkijan oman alan edustajista vai tieteenalarajat ylittävästä joukosta. Korkeakouluista TaY:n vastaajat ovat seitsemän kahdeksasosan enemmistöllä halukkaimpia aineistojen laajaan jakamiseen. Kuitenkin myös TAMK:n vastaajista enemmistö on asiassa myönteisellä linjalla TTY:n vastaajajoukon jakautuessa tasan puoltaviin ja vastustajiin. (Taulukko 15.) Yksi TaY:n vastaajista kommentoi kyselylomakkeen lopussa vielä aineistojen jakamishalujaan tiedeyhteisön rakenteita sivuvaasta näkökulmasta kirjoittaen:

”Joihinkin kysymyksiin on hankala ottaa kantaa sen vuoksi että ei ole määritelty mitä 'muut' (tutkijat) [sic] ovat kyselyssä. Esimerkiksi tutkimuskonsortioiden sisällä jaan jo nyt yliopistojen välillä aineistoja. Tahoilta, joiden kanssa en ole tekemisissä missään muodossa, jakamisen kynnyks on korkeampi.”

Taulukko 16. Haluaako tutkija asettaa tiedeyhteisön jäsenen rooliin liitetyn ehdon aineistojen uudelleenkäytölle?

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Yhteinen ymmärrys	Kyllä	-	-	1	25	4	50	5	31,3
vastavuoroisesta aineiston jakamisesta <sup>a</sup>	En	2	50	3	75	3	37,5	8	50
	En ole varma	2	50	-	-	1	12,5	3	18,8
	Yhteensä	4		4		8		16	

a. p-arvo 0.299, Fisher-Freeman-Haltonin eksakti testi.

Valtaosa siis suhtautuu suopeasti tutkimusaineiston jakamiseen laajalle yleisölle, ja taulukon 16 esittämät luvut viestivät, että vastaajat ovat valmiita edistämään aineistojen saatavuutta tiedeyhteisössä jossain määrin myös muiden toiminnasta riippumattomasti, sillä puolet ei pidä yhteistä ymmärrystä vastavuoroisesta jakamisesta tarpeellisenä aineiston uudelleenkäytön ehtona. Kuitenkin kolmannes odottaa jakamisesta samanlaista vastapalvelusta. Viidennes taas on epävarma ehdon tarpeellisuudesta, mikä saattaa kertoa siitä, että asian merkitystä ei ole entuudestaan pohdittu. TaY:n vastaajajoukosta puolet pitää vastavuoroisuutta arvossa. Sen sijaan TTY:n vastaajista ehdon puoltajia on vähemmistö, ja TAMK:n vastaajista kukaan ei asetu suoraan kannattamaan sitä.

Taulukko 17. Tiedeyhteisön jäsenen rooliin liittyvä syy jättää aineistoja jakamatta.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Muut eivät tarvitse aineistojani <sup>a</sup>	Valittu syyksi	-	-	3	75	-	-	3	18,8
	Ei valittu syyksi	4	100	1	25	8	100	13	81,3
	Yhteensä	4		4		8		16	

a. p-arvo 0.014, Fisher-Freeman-Haltonin eksakti testi.

Kun painopiste siirretään vastavuoroisuudesta yksipuoliseen päätöksentekoon aineiston avoimuuden rajoittamisesta, vastaajista viidennes pitää sitä, että muut eivät tarvitse heidän aineistojaan, syynä jättää aineistoa jakamatta. Korkeakoulutaustan vaikutus valintoihin on tilastollisesti merkitsevä: siinä missä TTY:n vastaajajoukosta kolme neljäsosaa jättää tuomatta aineistoaan muiden saataville, koska arvioi, että muut eivät sitä tarvitse, yksikään TAMK:n ja TaY:n vastaajista ei näe sen ohjaavan heidän jakamispäätöksiään. (Taulukko 17.) Vaikka tiedeyhteisön jäsenen rooliin liitettiin alle kymmenesosa kaikista kyselylomakkeella ehdotetuista aineistonjakamattajättämissyistä, TTY:n vastaajat tekevät rooliin liittyen neljänneksen kaikista syyvalinnoistaan; ”muut eivät tarvitse tutkimusaineistojani” on kilpailuetuun liittyvän syyn rinnalla heidän yleisimmin valitsemansa perustelu. (Liite 10 a.)

Tutkijat, jotka rajaavat aineistoaan muiden ulottumattomiin sillä perusteella, että muut eivät tarvitse heidän aineistojaan, eivät välttämättä näe tarvetta antaa muille mahdollisuutta valita, hyödynnettävätkö heidän keräämiään aineistoja vai eivät. Toisaalta vaatisi lisätutkimusta selvittää, millaisten tekijöiden ja vuoropuhelujen varaan näkemys rakentuu. Toisaalta yksi kolmesta avoimuuden esteeksi vaihtoehdon ”muu” valinneista täsmentää vastaustaan kirjoittaen: ”Ei voisi vähempää kiinnostaa, en minäkään kenenkään raakadataa halua sörkkiä.” Hän on TTY:stä, ja liitteestä 10 b käy ilmi, että hän valitsee jakamatta jättämisen syyksi myös sen, että muilla ei ole tarvetta hänen aineistoilleen. Siten kirjallisen vastauksen voi tulkita täydentävän kyseistä syytä. Hän jatkaa aiheesta myös kyselyn Vapaa sana -osiossa kirjoittaen: ”Ärsyttää, kun alat, joissa kukaan ei todellakaan halua alkaa penkomaan toisten dataa joutuvat alkaa niitä säilömään [...]” Myös toinen TTY:n tutkija kirjoittaa lopuksi: ”Pääasiallinen data on mittalaitteista tulevaa dataa jonka julkistaminen tarkoittaa siitä tehtävän data-analyysin tieteellistä julkaisua. En ymmärrä miten tilanne voisi olla erilainen. Raakadataa ei käytännössä juuri kukaan halua [...]” Näin ollen puolet TTY:n vastaajista antaa ymmärtää yhteisymmärryksen vallitsevan siitä, että ”kukaan” ei halua tutkia tutkijakollegan aineistoja.

## **5.5 Tutkimusaineiston uudelleenkäyttäjä ja uudelleenkäytön valvoja**

Tässä alaluvussa huomio kohdistuu tutkimusaineistojen uudelleenkäyttöön. Ensiksi käydään läpi aineistoon liitettäviä uudelleenkäyttömahdollisuuksia. Toisena teemana on tamperelaistutkijoiden halu kontrolloida itse keräämänsä aineiston uudelleenkäyttöä, eli se, millä keinoin ja missä määrin he ovat valmiita valvomaan jakamiensa aineistojensa hyödyntämistä. Kolmantena tutkitaan, miten he ottavat vuorostaan itse vastaan aineistojen uudelleenkäyttäjän roolin.

Taulukko 18. Näkemykset tutkimusaineistojen monikäyttöisyydestä ja yhdisteltävyydestä.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Olen tyytyväinen kykyihini yhdistää aineistoja eri lähteistä, kun etsin vastauksia tutkimuskysymyksiin. <sup>a</sup>	Vahvasti samaa mieltä	1	25	-	-	4	50	5	31,3
	Osittain samaa mieltä	1	25	3	75	1	12,5	5	31,3
	En samaa enkä eri mieltä	-	-	-	-	1	12,5	1	6,3
	Osittain eri mieltä	-	-	-	-	2	25	2	12,5
	Vahvasti eri mieltä	2	50	1	25	-	-	3	18,8
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	
On oikein koota uusia aineistokokonaisuuksia jaetuista aineistoista. <sup>b</sup>	Vahvasti samaa mieltä	2	50	1	25	6	75	9	56,3
	Osittain samaa mieltä	-	-	2	50	2	25	4	25
	En samaa enkä eri mieltä	2	50	-	-	-	-	2	12,5
	Osittain eri mieltä	-	-	-	-	-	-	-	-
	Vahvasti eri mieltä	-	-	1	25	-	-	1	6,3
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	
Aineistoa saatetaan käyttää toisella tavalla kuin mihin se on tarkoitettu. <sup>c</sup>	Vahvasti samaa mieltä	2	50	3	75	2	25	7	43,8
	Osittain samaa mieltä	-	-	-	-	6	75	6	37,5
	En samaa enkä eri mieltä	-	-	-	-	-	-	-	-
	Osittain eri mieltä	-	-	-	-	-	-	-	-
	Vahvasti eri mieltä	2	50	-	-	-	-	2	12,5
	En osaa sanoa	-	-	1	25	-	-	1	6,3
	Yhteensä	4		4		8		16	

a. p-arvo 0.122, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.066, Fisher-Freeman-Haltonin eksakti testi.

c. p-arvo 0.005, Fisher-Freeman-Haltonin eksakti testi.

Taulukko 18 osoittaa noin kolmen viidesosan enemmistön vastaajista olevan tyytyväinen kykyihinsä yhdistää tutkimusaineistoja eri lähteistä, mutta noin kolmannes näkee tilanteessa pikemminkin parantamisen varaa ja joukossa on voimakastakin tyytymättömyyttä. Kunkin korkeakoulun vastaajista vähintään puolet on tyytyväisiä yhdistelykykyihinsä. TaY:n vastaajajoukko erottuu siten, että heidän parissaan on eniten erittäin tyytyväisiä aineistojen yhdistelijöitä ja erittäin tyytymättömiä ei ole lainkaan. Kysyttäessä suhtautumista eri aineistojen yhdistelyyn noin neljä viidesosaa kaikista vastaajista hyväksyy vähintään jossain määrin uusien aineistokokonaisuuksien kokoamisen jaetuista aineistoista, yli puolet vahvasti. Jälleen vähintään puolet kunkin korkeakoulun vastaajista ovat hyväksyvällä kannalla. Tosin puolet TAMK:n vastaajista suhtautuu asiaan neutraalisti, mikä voi viitata siihen, että he eivät pidä uusien aineistokokonaisuuksien kokoamista keskeisenä osana työtään.

Tutkimusaineistoja pidetään monikäyttöisinä, sillä noin neljä viidesosaa vastaajista arvioi, että aineistoa saatetaan käyttää toisella tavalla kuin mihin se on alkujaan tarkoitettu. Korkeakoulutaustan vaikutus näkemyksiin on tilastollisesti merkitsevä. TaY:n vastaajista kaikki näkevät aineistoille löytyvän poikkeavia käyttötapoja, mutta valtaosa on sitä mieltä vain osittain. Sen sijaan TAMK:n ja TTY:n

parissa osittain samanmielisiä ei ole lainkaan. Lisäksi TAMK:n vastaajista puolella on näkemys, että aineistoa ei käytetä alkuperäistavoista poiketen, ja vain TTY:n vastaajajoukossa on epävarmuutta käyttötapojen monipuolisuudesta. Kuitenkin kaikkien korkeakoulujen vastaajista vähintään puolet näkee aineistoille löytyvän alkuperäisestä käyttötarkoituksesta eroavia käyttötapoja. (Taulukko 18.)

## Tutkijat aineistojensa uudelleenkäytön valvojina

Seuraavassa käydään läpi, missä määrin tutkijat haluavat seurata itse jakamisensa aineistojen käyttöä. Aineiston uudelleenkäytön valvojan rooliin liitettiin puolet kaikista kyselyssä tarjotuista uudelleenkäyttöä koskevista ehtovaihtoehdoista, mutta se saa osakseen noin kaksi viidesosaa kaikista annetuista vastauksista. Kuitenkin TaY:n vastaajajoukossa valvoja on keskeisin ehtojenasettamisrooli – se kerää puolet kaikista heidän tekemistään ehtovalinnoista. TAMK:nkin vastaajajoukon valintakerroista se kattaa miltei kaksi viidesosaa, ja uudelleenkäytön valvojan rooli on heille ehtoja asetettaessa yhtä keskeinen kuin aineiston kerääjän rooli. (Liite 11.)

Taulukko 19. Haluaako tutkija asettaa uudelleenkäytön valvojan rooliin liitetyjä ehtoja aineistojen uudelleenkäytölle?

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
<b>Aktiivinen uudelleenkäytön seuraaminen</b>									
Aineiston kerääjällä on mahdollisuus tehdä yhteistyötä projektin kanssa, jossa aineistoa käytetään. <sup>a</sup>	Kyllä	3	75	-	-	4	50	7	43,8
	En	-	-	3	75	3	37,5	6	37,5
	Epävarma	1	25	1	25	1	12,5	3	18,8
	Yhteensä	4		4		8	1	16	
Tuloksia, jotka perustuvat aineistoon, ei saa levittää missään muodossa ilman aineiston kerääjän hyväksyntää. <sup>b</sup>	Kyllä	1	25	-	-	2	25	3	18,8
	En	3	75	4	100	4	50	11	68,8
	Epävarma	-	-	-	-	2	25	2	12,5
	Yhteensä	4		4		8		16	
Tuloksia, jotka pohjautuvat aineiston hyödyntämiseen, ei voi levittää ilman, että aineiston kerääjällä olisi mahdollisuutta arvioida tuloksia ja tehdä ehdotuksia ja kommentteja. <sup>c</sup>	Kyllä	-	-	-	-	3	37,5	3	18,8
	En	3	75	4	100	4	50	11	68,8
	Epävarma	1	25	-	-	1	12,5	2	12,5
	Yhteensä	4		4		8		16	
Aineiston kerääjä saa ja hyväksyy lausunnon koskien aineiston uudelleenkäytön tavoitteita. <sup>d</sup>	Kyllä	2	50	-	-	4	50	6	37,5
	En	2	50	4	100	4	50	10	62,5
	Epävarma	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	
<b>Passiivinen uudelleenkäytön seuraaminen</b>									
Aineiston kerääjälle pitää antaa kopio artikkeleista, joissa hänen aineistoaan on käytetty. <sup>e</sup>	Kyllä	2	50	-	-	3	37,5	5	31,3
	En	2	50	4	100	5	62,5	11	68,8
	Epävarma	-	-	-	-	-	-	-	-

	Yhteensä	4	100	4	100	8	100	16	100
Aineiston kerääjä saa listan	Kyllä	2	50	-	-	5	62,5	7	43,8
kaikista tuotoksista, joissa hänen	En	2	50	3	75	3	37,5	8	50
aineistojaan on käytetty,	Epävarma	-	-	1	25	-	-	1	6,3
mukaan lukien artikkelit,	Yhteensä	4		4		8		16	
esitykset, opetusmateriaalit jne. <sup>f</sup>									
<b>Yhteenveto aktiivista seuranta-</b>	Kyllä	6	37,5	-	-	13	40,6	19	29,7
<b>vaativien ehtojen keräämistä</b>	En	8	50	15	93,8	15	46,9	38	59,4
<b>vastauskerroista<sup>g</sup></b>	Epävarma	2	12,5	1	6,3	4	12,5	7	10,9
	Yhteensä <sup>j</sup>	16		16		32		64	
<b>Yhteenveto passiivista</b>	Kyllä	4	50	-	-	8	50	12	37,5
<b>seuranta-vaativien ehtojen</b>	En	4	50	7	87,5	8	50	19	59,4
<b>keräämistä vastauskerroista<sup>h</sup></b>	Epävarma	-	-	1	12,5	-	-	1	3,1
	Yhteensä <sup>j</sup>	8		8		16		32	
<b>Yhteenveto kaikkien ehtojen</b>	Kyllä	10	41,7	-	-	21	43,8	31	32,3
<b>keräämistä vastauskerroista<sup>i</sup></b>	En	12	50	22	91,7	23	47,9	57	59,4
	Epävarma	2	8,3	2	8,3	4	8,3	8	8,3
	Yhteensä <sup>j</sup>	24		24		48		96	

a. p-arvo 0.142, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.600, Fisher-Freeman-Haltonin eksakti testi.

c. p-arvo 0.374, Fisher-Freeman-Haltonin eksakti testi.

d. p-arvo 0.287, Fisher-Freeman-Haltonin eksakti testi.

e. p-arvo 0.487, Fisher-Freeman-Haltonin eksakti testi.

f. p-arvo 0.234, Fisher-Freeman-Haltonin eksakti testi.

g. p-arvo 0.008, Fisher-Freeman-Haltonin eksakti testi.

h. p-arvo 0.044, Fisher-Freeman-Haltonin eksakti testi.

i. p-arvo 0.000 [0.000235], Fisher-Freeman-Haltonin eksakti testi.

j. Havaintoyksikkönä valintakerrat.

Taulukko 19 osoittaa enemmistön vastaajista suhtautuvan torjuvasti kuhunkin aineiston uudelleenkäytön seurantatapaan, silloin kun se on aineiston uudelleenkäytölle asetettava ehto. Ehdosta kolme eli puolta vierastetaan erityisen voimakkaasti: ehdon torjuvia vastauksia annetaan vähintään kaksi kertaa enemmän kuin asettamista puoltavia. Niistä kaksi liittyy uudelleenkäytön aktiiviseen seurantaan: enemmistö ei kaipaa mahdollisuutta antaa hyväksyntäänsä keräämänsä aineiston uudelleenkäytölle eikä halua ottaa kantaa aineistoon pohjautuviin tuloksiin. Kolmas laajasti torjuttu ehto on passiivista seuranta-vaativien ehtojen keräämistä vastauskerroista: suurin osa vastaajista ei halua kopioita artikkeleista, joissa hänen aineistojaan on käytetty. Ehdosta eniten kannatusta saa aineiston kerääjälle tarjottava mahdollisuus tehdä yhteistyötä projektin kanssa, jossa aineistoa käytetään, missä on kyse uudelleenkäytön aktiivisesta seurannasta. Se on kuudesta ehdosta ainoa, johon kantaa ottaneista useampi suhtautuu puoltavasti kuin kielteisesti – tosin epävarmoja vastaajia on useita.

Suhtautumisessa yksittäisiin aineiston uudelleenkäytön seurantatapoihin uudelleenkäytön ehtona ei ole tilastollisesti merkitseviä korkeakoulukohtaisia eroja. Kuitenkaan eri korkeakoulujen edustajat eivät suhtaudu samalla tavalla uudelleenkäytön valvontaan: kun tarkastellaan yhteenvetoja

suhtautumisesta passiivisen seurannan ja aktiivisen seurannan tapoihin sekä kaikkiin tapoihin kootusti, korkeakoulukohtaiset erot ovat tilastollisesti merkitseviä. TTY:n vastaajat erottuvat torjuvalla linjallaan: he eivät kannata suoraan mitään ehdoista. TAMK:n ja TaY:n vastaajajoukkojen valintakerat jakautuvat tasan puoltaviin ja vastustaviin suhtautumisessa passiiviseen seurantaan. Kun kyse on aktiivisesta seurannasta, molempien valintakerroista noin kaksi viidesosaa on seurantaa puoltavia ja epävarmojakin vastauksia on jonkin verran. (Taulukko 19.)

Tarkemmin tarkasteltuna TaY:n vastaajajoukon enemmistö ei asetu yhdenkään aktiivisuutta vaativan ehdon taakse, mutta runsas kolme viidesosaa kannattaa passiivista ehtoa, jonka mukaan aineiston kerääjälle toimitetaan lista kaikista tuotoksista, joissa hänen aineistojaan on käytetty. TAMK:n vastaajien enemmistö taas ei kannata kumpaakaan passiivisista ehdoista, mutta puoltaa aktiivisuutta vaativista ehdoista aineiston kerääjälle tarjottavaa mahdollisuutta tehdä yhteistyötä projektin kanssa, jossa tämän aineistoa käytetään – puolet TaY:nkin vastaajista puoltaa sitä. (Taulukko 19.)

Taulukko 20. Uudelleenkäytön valvomiseksi asettavien ehtojen tärkeänä pitämisen osuudet.<sup>a</sup>

	TAMK (n = 4)		TTY (n = 4)		TaY (n = 8)		Yhteensä (n = 16)	
	n	%	n	%	n	%	n	%
<b>Ehtojen (6) täyttämiskaava</b>								
Kaikki tärkeitä	-	-	-	-	1	12,5	1	6,3
Vähintään yksi tärkeä	3	75	-	-	6	75	9	56,3
Yksikään ei tärkeä	-	-	3	75	2	25	5	31,3
Vähintään yksi ei tärkeä	4	100	4	100	6	75	14	87,5
Epävarma vain yhdestä	2	50	2	50	-	-	4	25
Epävarma kahdesta	-	-	-	-	2	25	2	12,5
<b>Aktiivisten ehtojen (4) täyttämiskaava</b>								
Kaikki tärkeitä	-	-	-	-	1	12,5	1	6,3
Vähintään yksi tärkeä	3	75	-	-	6	75	9	56,3
Yksikään ei tärkeä	-	-	3	75	2	25	5	31,3
Vähintään yksi ei tärkeä	3	75	4	100	6	75	13	81,3
Epävarma vain yhdestä	2	50	1	25	-	-	3	18,8
Epävarma kahdesta	-	-	-	-	2	25	2	12,5
<b>Passiivisten ehtojen (2) täyttämiskaava</b>								
Kumpikin ehdoista tärkeä	1	25	-	-	3	37,5	4	25
Vähintään toinen tärkeä	3	75	-	-	5	62,5	8	50
Kumpikaan ei tärkeä	1	25	3	75	3	37,5	7	43,8
Vain toinen ehdoista tärkeä	2	50	-	-	2	25	4	25
Epävarma vain yhdestä ehdosta	-	-	1	25	-	-	1	6,3

a. Kun puhutaan kaikkien tärkeänä pitämisestä, kaikki ehdot on valittu: mistään ei olla epävarmoja. Kun puhutaan kaikkien ei tärkeänä pitämisestä, kaikki ehdot on torjuttu, mistään ei olla epävarmoja, jne.



Taulukko 20 näyttää, että valtaosalla vastaajista on kiinnostusta tietää, missä ja miten heidän keräämiään aineistoja käytetään, mutta he valikoivat seuraamisen tapoja. Yli puolet vastaajista kannattaa vähintään yhden uudelleenkäytön valvomiseksi asetettavan ehdon asettamista. Aktiivisista ehdoista vähintään yhtä kannattaa yli puolet vastaajista ja passiivisista ehdoista vähintään toista kannattaa puolet. Silti valvontaan liittyviin ehtoihin suhtaudutaan myös taulukon 20 mukaan sikäli kielteisvoittoisesti, että isompi joukko kyselyyn vastanneista kannattaa kaikkien ehtojen torjumista kuin kaikkien asettamista: kolmannes vastaajista ei pidä yhtäkään kuudesta ehdosta tärkeänä, kun taas kaikkien kuuden ehdon täyttymistä edellyttää vain kuudestoistaosa, yksi vastaaja. Runsaat kaksi viidesosaa vastaajista ei kaipaisi kumpaakaan passiivisista ehdoista ja noin kolmannes jättäisi kaikkien aktiivisuutta vaativat ehtojen asettamisen väliin.

Jo aiemmin kävi ilmi, että TTY:n vastaajat ovat yksimielisiä ehtojen asettamisen tarpeettomuudesta sikäli, että heitä ei ole joukossa, joka suoraan – joukossa on epävarmuuttakin – kannattaa yhdenkään ehdon täyttymistä. Sekä TAMK:n että TaY:n vastaajien enemmistöt vaativat vähintään yhden ehdon täyttymistä, olipa kyse aktiivisesta tai passiivisesta seurannasta. Toisaalta ainoa kaikkien ehtojen täyttymistä edellyttävä vastaaja on TaY:sta. Toisaalta TAMK:n vastaajat erottuvat ehtojen ystävänä, sillä heitä ei ole kaikki ehdot suoraan torjuvassa neljänneksessä. He suhtautuvat suopeasti myös aineiston kerääjän aktiivisuuteen aineiston uudelleenkäytön valvonnassa, sillä he eivät kuulu vastaajakolmannekseen, joka torjuu kaikki aktiivista valvontaa koskevat ehdot. Muissa korkeakouluissa aktiivisen seurannan tarpeen torjuvia on vähintään neljännes vastaajista. (Taulukko 20.)

Passiivinen ja aktiivinen uudelleenkäytön valvonta eivät toki ole toisiaan poissulkevia vaihtoehtoja. Kolmanneksesta, joka haluaa aineiston kerääjän saavan lausunnon hyväksyttäväkseen uudelleenkäytön tavoitteista, neljä viidesosaa haluaa myös kopiot artikkeleista, jossa aineistoa on käytetty ja yli kaksi kolmasosaa listan kaikista tuotoksista, joissa aineistoa on käytetty. Tämä kertoo myös siitä, että tutkimusaineiston uudelleenkäytön seurannassa ollaan valmiita johdonmukaisuuteen. (Liite 8.)

## Tutkijat aineistojen uudelleenkäyttäjinä

Taulukko 21. Näkemykset halukkuudesta käyttää valmiita aineistoja ja saatavuuden esteiden vaikutuksista.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Käyttäisin toisten tuottamia aineistoja,	Vahvasti samaa mieltä	1	25	1	25	3	37,5	5	31,3
	Osittain samaa mieltä	-	-	1	25	3	37,5	4	25
	En samaa enkä eri mieltä	2	50	-	-	-	-	2	12,5

jos ne olisivat helposti saatavilla. <sup>a</sup>	Osittain eri mieltä	1	25	1	25	2	25	4	25
	Vahvasti eri mieltä	-	-	1	25	-	-	1	6,3
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	
Pääsyn puuttuminen toisten oman tieteenalani tutkijoiden tuotamiin aineistoihin on rajoittanut kykyäni vastata tieteellisiin kysymyksiin. <sup>b</sup>	Vahvasti samaa mieltä	1	25	1	25	-	-	2	12,5
	Osittain samaa mieltä	-	-	1	25	-	-	1	6,3
	En samaa enkä eri mieltä	1	25	-	-	1	12,5	2	12,5
	Osittain eri mieltä	-	-	1	25	2	25	3	18,8
	Vahvasti eri mieltä	2	50	1	25	5	62,5	8	50
En osaa sanoa	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	

a. p-arvo 0.432, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.474, Fisher-Freeman-Haltonin eksakti testi.

Vastaajat asettuivat vahvasti tutkimusaineistojen uudelleenkäyttäjän rooliin, sillä hieman yli puolet heistä hyödyntäisi valmiita aineistoja, jos ne olisivat helposti saatavilla. Noin kolmannes kallistuu kielteiselle kannalle, mutta enimmäkseen osittain. Korkeakoulukohtaisesti on hajontaa. TaY:n vastaajia helppo saavutettavuus houkuttaisi eniten valmiiden aineistojen pariin: kolme neljäsosaa heistä on sitä mieltä. TTY:n vastaajista puolet suhtautuu lähtökohtaisesti avoimin mielin ajatukseen. TAMK:n vastaajista vain yksi neljästä suhtautuu myönteisesti valmiiden aineistojen käyttöön, joskin heidän joukossaan on kannastaan epävarmojakin, mikä voi kertoa siitä, että aihepiiriä ei ole pohdittu aiemmin, jolloin mielessä ei välttämättä ole käyttötapoja valmiille aineistoille. (Taulukko 21.)

Noin kaksi kolmasosaa vastaajista ei koe pääsyn puuttumisen toisten oman alan tutkijoiden aineistoihin rajoittaneen heidän omaa tutkimustyötään; puolet vastaa niin varmoin mielin. Kuitenkin noin viidennes on törmännyt aineistojen avoimuuden rajaamisesta johtuneisiin esteisiin. Eniten omakohtaista kokemusta oman tutkimustyön vaikeutumisesta on TTY:n vastaajilla eli puolella nelikosta – muistetaan, että TTY:n vastaajista silti kolme neljästä jättää aineistojaan itse jakamatta, koska näkee, että muut eivät tarvitse niitä (ks. taulukko 17). Vaikka TaY:n vastaajista taas moni näki oman alan aineistojen saavutettavuusongelmat riskinä tieteen kehitykselle (ks. taulukko 14), omakohtaisia ongelmia he eivät juuri ilmaise kohdanneensa. (Taulukko 21.)

Taulukko 22. Kokemus ja näkemys Tampereen korkeakoulujen tutkimusaineistojen hyödyntämisestä.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Olen hyödyntänyt toisten tamperelaisten korkeakoulujen tutkimusaineistoja omien tutkimusongelmieni ratkaisemiseksi. <sup>a</sup>	Vahvasti samaa mieltä	-	-	-	-	1	12,5	1	6,3
	Osittain samaa mieltä	-	-	1	25	1	12,5	2	12,5
	En samaa enkä eri mieltä	-	-	-	-	1	12,5	1	6,3
	Osittain eri mieltä	-	-	1	25	-	-	1	6,3
	Vahvasti eri mieltä	4	100	2	50	5	62,5	11	68,8
En osaa sanoa		-	-	-	-	-	-	-	-

	Yhteensä	4	4	8	16				
Olen halukas hyödyntämään toisten tamperelaisten korkeakoulujen tutkimusaineistoja omien tutkimusongelmieni ratkaisemiseksi. <sup>b</sup>	Vahvasti samaa mieltä	2	50	2	50	-	-	4	25
	Osittain samaa mieltä	1	25	-	-	4	50	5	31,3
	En samaa enkä eri mieltä	-	-	1	25	2	25	3	18,8
	Osittain eri mieltä	-	-	-	-	-	-	-	-
	Vahvasti eri mieltä	1	25	1	25	2	25	4	25
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	

a. p-arvo 0.805, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.249, Fisher-Freeman-Haltonin eksakti testi.

Paikallistason tutkimustoimintaa esittelevä taulukko 22 osoittaa, että viidenneksellä vastaajista on jo kokemusta tamperelaisten korkeakoulujen välillä toteutuneesta tutkimusaineistojen vaihdosta, mutta noin kaksi kolmasosaa torjuu vahvasti väittämän. Kun katse suunnataan tulevaisuuteen, yli puolet vastaajista olisi halukas hyödyntämään toisten tamperelaisten korkeakoulujen tutkimusaineistoja omien tutkimusongelmiensa ratkaisemiseksi. Tosin neljännes on vahvasti eri mieltä eikä osittain eri mieltä olevia ole yhtään, joten jos ajatus torjutaan, se torjutaan vahvasti. Lisäksi noin viidennes on kannassaan neutraali, mikä voi tarkoittaa, että heillä ei heti tule mieleen toiminnasta koituvia hyötyjä. Korkeakoulukohtaisesti vähintään kolme viidesosaa torjuu väittämän toisten tamperelaisten korkeakoulujen aineistojen hyödyntämisestä tutkimustyössään. TAMK on ainoa, jonka kaikki vastaajat kieltävät sen jyrkästi, siinä missä TaY:n ja TTY:n vastaajista neljänneksellä on ollut jonkinlaista korkeakoulurajat ylittävää aineistonkäyttöä. Vähintään puolet kunkin korkeakoulun vastaajista suhtautuu myönteisesti mahdollisuuteen hyödyntää naapurikorkeakoulujen aineistoja ja vain neljännekset torjuvasti. Tosin TaY:n vastaajajoukossa ei ole vahvaa myönteisyyttä.

## 5.6 Tutkimusaineiston käsittelijä – aineistojen käytettävyys

On käynyt jo ilmi, että tamperelaistutkijat näkevät tutkimusaineistolla voivan olla useita eri käyttötapoja, ja missä määrin he ovat hyödyntäneet muiden aineistoja ja pitäneet niitä tarpeellisina. Tässä alaluvussa huomio on tutkijoissa aineiston käsittelijöinä eli heidän aineistonkäyttökokemuksissaan. Ensiksi käydään läpi, miten tyytyväisiä tutkijat ovat prosesseihin, joilla he esimerkiksi keräävät, kuvailevat ja analysoivat aineistojaan. Toiseksi syvennetään edellä mainittua tutkimalla, millaista metadastandardien käyttö on vastaajien keskuudessa. Kolmanneksi kerrotaan, miten onnistunutta aineistojen käyttöä vastaajista on eli millaisia kokemuksia heillä on sekä oman tieteenalansa että muiden alojen aineistojen käytöstä.

## Tyytyväisyys tutkimusaineistojen hyödyntämisen prosesseihin

Taulukko 23. Tyytyväisyys tutkimusaineistojen hyödyntämisen prosesseihin ja välineisiin.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Olen tyytyväinen prosessiin, jolla kerään aineistoni. <sup>a</sup>	Vahvasti samaa mieltä	1	25	1	25	3	37,5	5	31,3
	Osittain samaa mieltä	2	50	2	50	4	50	8	50
	En samaa enkä eri mieltä	-	-	-	-	-	-	-	-
	Osittain eri mieltä	-	-	-	-	1	12,5	1	6,3
	Vahvasti eri mieltä	-	-	1	25	-	-	1	6,3
	En osaa sanoa	1	25	-	-	-	-	1	6,3
	Yhteensä	4		4		8		16	
Olen tyytyväinen prosessiin, jolla luetteloin/kuvailen aineistoni. <sup>b</sup>	Vahvasti samaa mieltä	1	25	1	25	2	25	4	25
	Osittain samaa mieltä	3	75	2	50	1	12,5	6	37,5
	En samaa enkä eri mieltä	-	-	-	-	-	-	-	-
	Osittain eri mieltä	-	-	-	-	4	50	4	25
	Vahvasti eri mieltä	-	-	1	25	-	-	1	6,3
	En osaa sanoa	-	-	-	-	1	12,5	1	6,3
	Yhteensä	4		4		8		16	
Olen tyytyväinen prosessiin, jolla etsin omia aineistojani. <sup>c</sup>	Vahvasti samaa mieltä	1	25	-	-	4	50	5	31,3
	Osittain samaa mieltä	2	50	2	50	2	25	6	37,5
	En samaa enkä eri mieltä	-	-	-	-	1	12,5	1	6,3
	Osittain eri mieltä	-	-	1	25	1	12,5	2	12,5
	Vahvasti eri mieltä	1	25	1	25	-	-	2	12,5
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	
Olen tyytyväinen prosessiin, jolla analysoin aineistoni. <sup>d</sup>	Vahvasti samaa mieltä	1	25	-	-	4	50	5	31,3
	Osittain samaa mieltä	3	75	3	75	4	50	10	62,5
	En samaa enkä eri mieltä	-	-	-	-	-	-	-	-
	Osittain eri mieltä	-	-	-	-	-	-	-	-
	Vahvasti eri mieltä	-	-	1	25	-	-	1	6,3
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	
Olen tyytyväinen työkaluihin, joilla tuotetaan meta-tietoja ja aineiston dokumentointi. <sup>e</sup>	Vahvasti samaa mieltä	-	-	-	-	1	12,5	1	6,3
	Osittain samaa mieltä	-	-	1	25	-	-	1	6,3
	En samaa enkä eri mieltä	1	25	-	-	2	25	3	18,8
	Osittain eri mieltä	-	-	1	25	2	25	3	18,8
	Vahvasti eri mieltä	2	50	1	25	-	-	3	18,8
	En osaa sanoa	1	25	1	25	3	37,5	5	31,3
	Yhteensä	4		4		8		16	

a. p-arvo 0.778, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.198, Fisher-Freeman-Haltonin eksakti testi.

c. p-arvo 0.538, Fisher-Freeman-Haltonin eksakti testi.

d. p-arvo 0.319, Fisher-Freeman-Haltonin eksakti testi.

e. p-arvo 0.615, Fisher-Freeman-Haltonin eksakti testi.

Neljä viidesosaa vastaajista on tyytyväisiä tutkimusaineistojen keruuprosessiin. Sitä vastoin vain kahdeksasosa vastaajista on tyytyväisiä työkaluihin, joilla metatietoja ja aineiston dokumentointia

tuotetaan. Yli kolmannes on tyytymättömiä työkaluihin ja puolet kantaa ottaneista ei joko osaa ottaa kantaa tai ole kumpaakaan mieltä. Siitä huolimatta yli kolme viidesosaa on tyytyväisiä aineistonsa luettelointi- tai kuvailuprosessiin samoin kuin aineistojensa etsimisprosessiin, joka voi olla tulkittavissa joko uusien aineistojen etsimiseen tai hakujen tekemiseen omista aineistoista. Kun siirrytään aineiston analysointiprosessiin, tyytyväisten osuus on peräti yli yhdeksän kymmenestä, ja heistä puolet on erittäin tyytyväisiä. (Taulukko 23.)

Kaikkia korkeakouluja yhdistää, että missään vastaajien enemmistö ei ole tyytyväinen metadatan ja aineiston dokumentoinnin työkaluihin; TAMK:n vastaajajoukosta yksikään ei ilmaise tyytyväisyyttään. Jokaisessa korkeakoulussa on myös kannastaan epävarmoja. Sen sijaan kaikkien korkeakoulujen vastaajista vähintään kolme neljäsosaa on tyytyväisiä aineistonsa keräämiseen samoin kuin analysointiin. Kuitenkin siinä missä kaikki TAMK:n vastaajat ja enemmistö TTY:nkin vastaajista ovat vähintään osittain tyytyväisiä myös aineistojen kuvailu- ja luettelointiprosessiin, TaY:n vastaajista runsas kolmannes on tyytyväinen. Tyytyväisyydessä aineistojen etsimiseenkin on pieniä eroja niin, että puolet TTY:n vastaajista on tyytymättömiä, kun muissa enemmistö on tyytyväisiä. (Taulukko 23.)

Taulukko 24. Aineistojen käsittelijät metadastandardien käyttäjinä.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Mitä metadatastandardeja käytät tällä hetkellä, kun kuvailit tutkimusaineistoasi – vai käytätkö mitään? <sup>a</sup>	Organisaationi käyttämää metadataohjeistusta	1	25	-	-	-	-	1	6,3
	Laboratorioni/yksikköni käyttämää metadataohjeistusta	-	-	1	25	-	-	1	6,3
	En mitään	3	75	3	75	6	75	12	75
	En osaa sanoa	-	-	-	-	1	12,5	1	6,3
	Muu (Artikkeleissa lehden ohjeiden ja standardien mukaan)	-	-	-	-	1	12,5	1	6,3
	Yhteensä	4		4		8		16	

a. p-arvo 0.715, Fisher-Freeman-Haltonin eksakti testi.

Tyytyväisyys metadatan ja dokumentoinnin työkaluihin on siis alhaista, ja taulukosta 24 käy ilmi, että kolme neljäsosaa vastaajista ei käytä mitään metadatastandardeja tutkimusaineistojensa kuvailuun. Yksi vastaaja ei osaa sanoa metadatastandardien käyttämisestä mitään – kenties on joko epätietoisuutta siitä, mitä metadatastandardit ovat tai mikä on käytettyjen standardien tausta. Kokonaisvastaajajoukon linjaa mukaillen kunkin korkeakoulun vastaajista kolme neljäsosaa ei käytä mitään metadatastandardia. Metadatastandardien käyttäjiä ei puolestaan yhdistä korkeakoulutausta. TaY:n vastaajajoukossa vaihtoehto ”muu” sai kannatusta, ja vastaaja tarkentaa, että artikkeleissa aineisto kuvataan lehden ohjeiden ja standardien mukaisesti. Syntyy käsitys, että hän

käyttäisi standardeja vain julkaisujen yhteydessä, mutta voi olla, että hän halusi vain täsmentää vastaustaan. Kukaan ei valinnut useita vaihtoehtoja kuvaamaan käyttämiään standardeja.

Taulukko 25. Käytettävyyteen liittyvä aineiston jakamatta jättämisen syy.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Standardit puuttuvat <sup>a</sup>	Valittu syyksi	-	-	-	-	2	25	2	12,5
	Ei valittu syyksi	4	100	4	100	6	75	14	87,5
	Yhteensä	4		4		8		16	

a. p-arvo 0.467, Fisher-Freeman-Haltonin eksakti testi.

Erilaisten standardien puuttuminen erottuu kuitenkin vähäisenä tekijänä tehtäessä päätöksiä aineiston avoimuudesta. Vain kahdeksasosa vastaajista – kaikki TaY:sta – näkee standardien puuttumisen estävän aineistojen jakamista. (Taulukko 25.) Aineiston käsittelijän roolissa nimetään vain neljä prosenttia kaikista aineiston jakamatta jättämistä selittämään annetuista syistä – TaY:n vastaajajoukko on siis yksin osuuden takana. (Liite 10 a.)

## Tutkimusaineistojen ymmärrettävyys

”[...] datan esittämiselle on tieteenalalla yleisesti käytössä [sic] olevat standardit. Suurin osa julkaisemattomasta datasta on epämääräistä tai epäonnistunutta, ja siitä saatu hyöty olisi kyseenalaista tai jopa haitallista koska sen todenperäisyydestä emme pystyisi vastaamaan. Vain julkaistu data on omasta mielestäni tarpeeksi varmaksi koettua jotta se voidaan jakaa. Ja se jaetaan julkaisujen muodossa.”

Näin kirjoittaa yksi TTY:n vastaajista kyselylomakkeen Vapaa sana -osiossa antaen ymmärtää, että aineistojen jakamatta jättämisellä ehkäistään väärän tiedon leviäminen.

Taulukko 26. Tutkijoiden näkemykset tutkimusaineistojen käytettävyydestä.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Aineistoa	Vahvasti samaa mieltä	1	25	3	75	-	-	4	25
saatetaan tulkita	Osittain samaa mieltä	1	25	1	25	6	75	8	50
väärin sen moni-	En samaa enkä eri mieltä	-	-	-	-	1	12,5	1	6,3
mutkaisuuden	Osittain eri mieltä	1	25	-	-	1	12,5	2	12,5
vuoksi. <sup>a</sup>	Vahvasti eri mieltä	1	25	-	-	-	-	1	6,3
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		4		8		16	
Aineistoa	Vahvasti samaa mieltä	1	25	3	75	-	-	4	25
saatetaan tulkita	Osittain samaa mieltä	1	25	1	25	2	25	4	25
väärin sen	En samaa enkä eri mieltä	-	-	-	-	1	12,5	1	6,3
huonon laadun	Osittain eri mieltä	-	-	-	-	3	37,5	3	18,8
vuoksi. <sup>b</sup>	Vahvasti eri mieltä	2	50	-	-	1	12,5	3	18,8
	En osaa sanoa	-	-	-	-	1	12,5	1	6,3

Yhteensä	4	4	8	16
----------	---	---	---	----

a. p-arvo 0.052, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.180, Fisher-Freeman-Haltonin eksakti testi.

Kolme neljäsosaa vastaajista on samaa mieltä – joko vahvasti tai osittain – siitä, että aineistoa saatetaan tulkita väärin sen monimutkaisuuden vuoksi. Puolet pitää aineistojen huonoa laatuakin ongelmallisena onnistuneita johtopäätöksiä ajatellen. Kaikki TTY:n vastaajat ja puolet TAMK:n vastaajajoukosta näkevät sekä aineistojen monimutkaisuuden että huonon laadun voivat aiheuttaa väärintulkintoja. Sen sijaan TaY:n vastaajista vain neljäsosa pitää huonoa laatua ongelmana tulkintoja tehtäessä; monimutkaisuudesta niin ajattelee kolminkertainen määrä. (Taulukko 26.) Seuraavaksi esitetään, miten aineiston ymmärrettävyys vaikuttaa tutkijoiden käytännön ratkaisuihin.

Taulukko 27. Näkemykset valmiiden tutkimusaineistojen vaikeaselkoisuuden vaikutuksista omaan tutkimustyöhön.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Olen itse jättänyt toisten oman tieteenalani tutkijoiden aineistoja käyttämättä niiden vaikeaselkoisuuden vuoksi. <sup>a</sup>	Vahvasti samaa mieltä	-	-	3	75	-	-	3	18,8
	Osittain samaa mieltä	-	-	-	-	-	-	-	-
	En samaa enkä eri mieltä	1	25	-	-	2	25	3	18,8
	Osittain eri mieltä	-	-	-	-	2	25	2	12,5
	Vahvasti eri mieltä	3	75	1	25	4	50	8	50
	En osaa sanoa	-	-	-	-	-	-	-	-
Yhteensä		4		4		8		16	
Olen pitänyt muiden tieteenalojen tutkimusaineistojen vaikeaa ymmärrettävyyttä esteenä tutkimuksen tekemiselle. <sup>b</sup>	Vahvasti samaa mieltä	-	-	2	50	-	-	2	12,5
	Osittain samaa mieltä	1	25	-	-	1	12,5	2	12,5
	En samaa enkä eri mieltä	-	-	-	-	-	-	-	-
	Osittain eri mieltä	-	-	-	-	1	12,5	1	6,3
	Vahvasti eri mieltä	3	75	2	50	6	75	11	68,8
	En osaa sanoa	-	-	-	-	-	-	-	-
Yhteensä		4		4		8		16	

a. p-arvo 0.064, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.326, Fisher-Freeman-Haltonin eksakti testi.

Noin viidenneksellä vastaajista on vahvaa kokemusta toisten oman tieteenalansa tutkijoiden aineistojen käyttämättä jättämisestä niiden vaikeaselkoisuuden vuoksi, joten omien kollegoidenkin tuottamat aineistot voivat olla hankalasti hyödynnettäviä. Toisaalta noin kolme viidesosaa vastaajista kieltää toimineensa niin, puolet peräti tiukasti, mutta linjauksen takana voi myös olla, että vastaaja ei ole edes koettanut käyttää kollegansa aineistoja, jolloin vastauksesta ei välttämättä käy ilmi väittämässä esitetyn syyn merkitys. Noin viidennes asettuu neutraalille kannalle. Korkeakouluvertailussa TTY:n tutkijoista kolmella neljäsosalla on vahvaa kokemusta edellä kuvatusta tilanteesta, kun taas kukaan TAMK:n ja TaY:n vastaajista ei suoraan myönnä niin tapahtuneen. (Taulukko 27.)

Kun huomio siirretään muiden kuin oman alan tutkimusaineistojen hyödyntämiseen, neljännes vastaajista on kokenut aineistojen vaikean ymmärrettävyyden olleen esteenä tutkimuksen tekemiselle. Kuitenkin noin kaksi kolmasosaa kieltää jyrkästi, että heillä olisi sellaisesta kokemuksta. Kustakin korkeakoulusta vähintään puolet vastaajista torjuu vahvasti ajatuksen, että he olisivat pitäneet muiden tieteenalojen aineistojen vaikeaa ymmärrettävyyttä esteenä tutkimuksen tekemiselle, mutta toisaalta TTY:n vastaajanelikosta toinen puoli on pitänyt sitä suuresti esteenä. (Taulukko 27.)

## **5.7 Aineistonhallinnan tuen saaja – tiedeinstituutioiden vaikutus**

”[...] kaikkia velvoittavat säännöt tehdään sellaisten alojen mukaan, joissa datojen jako on yleistä ja mielekästä. Sama kuin hankintasäännöksissäkin, tutkimuslaitteiden hankinnat täytyy kilpailuttaa vanhainkodin vuoden vessapaperien kilpailuttamiseen tehdyillä säännöillä, vaikka kyse on luonteeltaan täysin erilaisista hankinnoista.”

Näillä sanoilla TTY:a edustanut vastaaja ilmaisee ihmettelevään sävyyn kyselylomakkeen Vapaa sana -osiossa näkemyksensä ylhäältä päin tulevasta yksipuolisesta aineistonhallinnan ohjailusta, joskin kommentoimatta tarkemmin, mistä ohjailu on lähtöisin. Tässä alaluvussa käsitellään tutkijoiden taustaorganisaatioiden, tiedeinstituutioiden, vaikutusta aineistonhallintaan. Ensiksi esitetään, mikä vaikutus niihin liittyvillä ehdoilla ja esteillä on aineistojen avoimuuteen ja miten tietoisia vastaajat ovat rahoittajilta tulevasta aineistonhallintasuunnitelmavaatimuksista. Toiseksi keskitytään siihen, missä määrin vastaajat hyödyntävät tiedeinstituutiotason säilytyspaikkoja verrattuna aineiston omassa hallussa pitämiseen. Lopuksi esitellään taustatoimijoista työpaikan ja rahoittajan vaikutus tutkijoiden tyytyväisyyteen niin lyhytkestoista kuin pitkäkestoistakin aineistonhallintaa kohtaan.

### **Aika ja taidot sekä raha ja rahoittaja**

Liite 10 a havainnollistaa, että tutkimusaineiston tai osan siitä jakamatta jättämiselle annetuista vastausvaihtoehdoista kaksi viidesosaa liittyy tiedeinstituutioihin eli aineistonhallinnan tuen saajan rooliin, mutta osa-alue kerää vain noin kolmanneksen kaikista syyvalinnoista. Tosin TaY:n vastaajajoukon valinnoista osuus on neljännes siinä missä TTY:n ja TAMK:n parissa osuus on korkeampi, noin kaksi viidesosaa. Yksittäisistä syistä taitojen puute on aineistojen julkisuuteen kuulumattomuuden rinnalla TAMK:n vastaajien yleisimmin valitsema aineistonjakamattajättämissyy, sillä se kattaa noin kolmanneksen kaikista nelikon tekemistä valinnoista.



Taulukko 28. Aineistonhallinnan tuen saajan rooliin liittyvät syyt olla jakamatta aineistoa.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Rahoitus puuttuu <sup>a</sup>	Valittu syyksi	-	-	1	25	2	25	3	18,8
	Ei valittu	4	100	3	75	6	7	13	81,3
	Yhteensä	4		4		8		16	
Minulla ei ole aikaa <sup>b</sup>	Valittu syyksi	-	-	2	50	3	37,5	5	31,3
	Ei valittu	4	100	2	50	5	62,5	11	68,8
	Yhteensä	4		4		8		16	
Aineistoilleni ei ole sopivaa säilytyspaikkaa <sup>c</sup>	Valittu syyksi	-	-	-	-	-	-	-	-
	Ei valittu	4	100	4	100	8	100	16	100
	Yhteensä	4		4		8		16	
Rahoittaja ei edellytä aineistojeni jakamista <sup>d</sup>	Valittu syyksi	1	25	1	25	-	-	2	12,5
	Ei valittu	3	75	3	75	8	100	14	87,5
	Yhteensä	4		4		8		16	
Minulla on puutteelliset taidot <sup>e</sup>	Valittu syyksi	2	50	1	25	2	25	5	31,3
	Ei valittu	2	50	3	75	6	75	11	68,8
	Yhteensä	4		4		8		16	
Valinnat yhteensä <sup>f</sup>	Valittu syyksi	3	15	5	25	7	17,5	15	18,8
	Ei valittu syyksi	17	85	15	75	33	82,5	65	81,3
	Yhteensä	20		20		40		80	

a. p-arvo 0.771, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.487, Fisher-Freeman-Haltonin eksakti testi.

c. Fisher-Freeman-Haltonin eksaktia testiä ei voitu suorittaa, sillä rivimuuttujan arvoissa ei ollut vaihtelevuutta.

d. p-arvo 0.233, Fisher-Freeman-Haltonin eksakti testi.

e. p-arvo 0.795, Fisher-Freeman-Haltonin eksakti testi.

f. p-arvo 0.752, Fisher-Freeman-Haltonin eksakti testi.

Taulukosta 28 käy ilmi, että taitojen riittämättömyys sekä ajan puute ovat yleisimmät aineistonhallinnan tuen saajan – tai tarvitsijan – rooliin liitetyt syyt jättää aineistoja jakamatta sikäli, että ne kumpikin saavat vastaajista kolmanneksen taakse. Tosin ajanpuutteen kohdalla TAMK:n vastaajat erottuvat muista siten, että vain he eivät näe sen vaikuttavan aineistonjakamiskäyttäytymiseensä. Viidestä syystä vain taitojen puute saa mainintoja kaikista kolmesta korkeakoulusta. Vaihtoehtoista säilytyspaikan puuttuminen puolestaan erottuu siten, että kukaan kantansa ilmaiseista vastaajista ei pidä sitä syynä olla tuomatta aineistojaan muiden saataville. Jälleen on syytä muistaa, että vastaaja olisi voinut halutessaan valita jokaisen kyselylomakkeella tarjotun aineistonjakamattajättämisyyn. Siten säilytyspaikan puuttumisen yksimielinen torjunta kiinnittää huomion. Noin viidennes vastaajista näkee rahoituksen puutteen johtavan aineiston jakamatta jättämiseen, kun taas rahoittajan vaatimusten puuttumisen näkee syyksi kahdeksasosa. TaY:n vastaajat ovat ainoita, jotka eivät mainitse rahoittajan vaatimusten puuttumista jakamattomuuden syyksi ja TAMK:n vastaajat puolestaan ainoita, jotka eivät näe rahoituksen puuttumista syynä.

Taulukko 29. Haluaako tutkija asettaa aineistonhallinnan tuen saajan rooliin liittyvän ehdon aineiston uudelleenkäytölle?

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Ainakin osa aineiston keräämis- ja säilytyskustannuksista on katettava. <sup>a</sup>	Kyllä	3	75	-	-	-	-	3	18,8
	En	1	25	4	100	5	62,5	10	62,5
	En ole varma	-	-	-	-	3	37,5	3	18,8
	Yhteensä	4		4		8		16	

a. p-arvo 0.025, Fisher-Freeman-Haltonin eksakti testi.

Kyselyssä aineistojen avoimuuden ehdoista vain yksi liittyi tiedeinstituutioihin, tarkemmin sanottuna rahoittajiin. Miltei kaksi kolmasosaa vastaajista ei sido aineistojen jakamista keräämis- ja säilytyskustannusten jakamiseen, mutta kustannuksiin osallistumista koskevat näkemykset vaihtelevat tilastollisesti merkitsevästi korkeakoulukohtaisesti: TAMK:n vastaajien enemmistö pitää keskeisenä, että aineiston keräämis- ja säilytyskustannuksia katetaan, siinä missä TaY:n ja TTY:n vastaajista kukaan ei edellytä sitä. (Taulukko 29.) Lisäksi liitteestä 11 käy ilmi, että kyseiset kolme valintaa kattavat kymmenesosan kaikista TAMK:n vastaajajoukon aineistojen uudelleenkäyttöä koskevien ehtojen valinnoista. Toisaalta voi olla, että TAMK:n vastaajien aineiston keräämis- ja säilytyskustannuksia myös korvataan, sillä kukaan heistä ei nimennyt rahoituksen puutetta jakamatta jättämisen syyksi. Toisaalta vaikka neljännekset TTY:n ja TaY:n vastaajista pitävät rahoituksen puutetta yhtenä syynä aineiston jakamatta jättämiselle, kummankaan edustajat eivät suoralta kädeltä vaatisi säilytys- ja keruukustannuksien jakamista – tosin TaY:n riveissä on torjuijen lisäksi linjastaan epävarmoja.

Taulukko 30. Aineiston keräämis- ja säilytyskustannuksiin osallistumisvaatimuksen vaikutus rahoitusta koskevien aineistonjakamattajättämissyiden valitsemiseen.

Aineistojen jakamatta jättämisen syy		Aineiston uudelleenkäytön ehto: ainakin osa aineiston keräämis- ja säilytyskustannuksista on katettava					
		Kyllä		En		En ole varma	
		n	%	n	%	n	%
Rahoitus puuttuu <sup>a</sup>	Valittu syyksi	-	-	1	10	2	66,7
	Ei valittu syyksi	3	100	9	90	1	33,3
	Yhteensä	3		10		3	
Rahoittaja ei edellytä aineistojeni jakamista <sup>b</sup>	Valittu syyksi	1	33,3	1	10	-	-
	Ei valittu syyksi	2	66,7	9	90	3	100
	Yhteensä	3		10		3	

a. p-arvo 0.143, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.625, Fisher-Freeman-Haltonin eksakti testi.

Myös taulukko 30 havainnollistaa, että aineistojaan rahoituksen puuttumisen vuoksi jakamatta jättävät eivät edellytä kustannuksiin osallistumista ehdoksi aineistojen jakamiselle. Toisaalta jos oletetaan rahoittajan todella vaativan tehokkaasti aineistojen jakamista niiltä vastaajilta, jotka eivät

nimeä rahoittajien vaatimusten puuttumista syyksi olla tuomatta aineistoja muiden saataville, heistäkään juuri kukaan ei aseta kustannuksiin osallistumista uudelleenkäytön ehdoksi.

Taulukko 31. Tietoisuus ensisijaisen rahoittajan aineistohallintasuunnitelmavaatimuksesta.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Vaatiiko ensisijainen rahoittajasi sinua laatimaan aineistohallintasuunnitelman? <sup>a</sup>	Kyllä	1	50	1	25	6	75	8	57,1
	Ei	1	50	2	50	1	12,5	4	28,6
	En tiedä	-	-	1	25	1	12,5	2	14,3
	Yhteensä	2		4		8		14	

a. p-arvo 0.422, Fisher-Freeman-Haltonin eksakti testi.

Aineistohallintasuunnitelma kuuluu yhtä useammin rahoittajien vaatimuksiin. Kyselyn vastaajat kokevat olevansa ensisijaisten rahoittajiensa aineistohallintasuunnitelmavaatimuksista hyvin perillä, sillä yli puolet tietää rahoittajansa vaativan suunnitelman ja runsas neljännes tietää, että sitä ei vaadita. Tosin epätietoisuutta asiasta on seitsemäsosalla kysymykseen kantaa ottaneista, jotka edustavat TTY:a ja TaY:a, ja TAMK:n vastaajista puolet ei vastaa kysymykseen. TaY:n vastaajista kolme neljäsosaa tietää rahoittajan vaativan heiltä suunnitelman. TAMK:n vastaajista se vaaditaan toiselta kysymykseen kantaa ottaneista ja TTY:n vastaajistakin vain yksi neljästä on varmuudella suunnitelmavaatimuksen piirissä. (Taulukko 31.) Vastauksissa saattaa olla epä johdonmukaisuutta, sillä aineistohallintasuunnitelmia vaativan Suomen Akatemian rahoittamista tutkijoista viidenes vastaa, että ensisijainen rahoittaja ei vaadi kyseistä suunnitelmaa – jollei heidän senhetkinen ensisijainen rahoittajansa ollut muu kuin tyypillinen rahoittaja. (Liite 9 a.)

### Organisaatio- vai yksilökeskeinen aineiston säilytys?

Seuraavaksi esitetään, missä määrin vastaajat nojaavat tiedeinstituutioiden tukeen tutkimusaineistoja säilyttäessään. Ensin tarkastellaan vastaajien säilytyspaikkakohtaisia valintoja. Sitten huomio on heidän aineiston säilytysmäärissään eli siinä, suosivatko he keskitettyä vai hajanaista säilyttämistä. Vastaajat tekivät yhteensä 149 säilytysmäärä- ja samalla epäsuorasti kyllä vai ei -kannanottoa säilytyspaikkojen käytöstä; 11 valintakertaa jätettiin käyttämättä (ks. liite 9 d). Seuraavassa taulukko 32 esittää vastaajien valintojen jakautumisen yksittäisten säilytyspaikkojen kesken.

Taulukko 32. Yksittäiset säilytyspaikkavalinnat.

Säilytyspaikka	Valinnat	Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
<b>Tiedeinstituutiotason säilytyspaikat (6)</b>									
Organisaationi tai yksikköni palvelin <sup>a</sup>		n = 4		n = 4		n = 8		n = 16	
	Ei lainkaan	1	25	1	25	2	25	4	25
	Jonkin verran	-	-	2	50	1	12,5	3	18,8
	Suurin osa	2	50	1	25	-	-	3	18,8
	Kaikki	1	25	-	-	5	62,5	6	37,5
	Yht. valittu säilytyspaikaksi	3	75	3	75	6	75	12	75
Tutkimushankkeen palvelin <sup>b</sup>		n = 3		n = 4		n = 8		n = 15	
	Ei lainkaan	3	100	3	75	5	62,5	11	73,3
	Jonkin verran	-	-	1	25	2	25	3	20
	Suurin osa	-	-	-	-	1	12,5	1	6,7
	Kaikki	-	-	-	-	-	-	-	-
	Yht. valittu säilytyspaikaksi	-	-	1	25	3	37,5	4	26,7
Tieteenalakohtainen säilytyspaikka/arkisto <sup>c</sup>		n = 4		n = 4		n = 8		n = 16	
	Ei lainkaan	3	75	4	100	6	75	13	81,3
	Jonkin verran	-	-	-	-	2	25	2	12,5
	Suurin osa	1	25	-	-	-	-	1	6,3
	Kaikki	-	-	-	-	-	-	-	-
	Yht. valittu säilytyspaikaksi	1	25	-	-	2	25	3	18,3
Julkaisijan säilytys/arkisto <sup>d</sup>		n = 3		n = 4		n = 8		n = 15	
	Ei lainkaan	3	100	3	75	8	100	14	93,3
	Jonkin verran	-	-	1	25	-	-	1	6,7
	Suurin osa	-	-	-	-	-	-	-	-
	Kaikki	-	-	-	-	-	-	-	-
	Yht. valittu säilytyspaikaksi	-	-	1	25	-	-	1	6,7
Organisaationi oma arkisto <sup>e</sup>		n = 3		n = 4		n = 8		n = 15	
	Ei lainkaan	2	66,7	4	100	8	100	14	93,3
	Jonkin verran	1	33,3	-	-	-	-	1	6,7
	Suurin osa	-	-	-	-	-	-	-	-
	Kaikki	-	-	-	-	-	-	-	-
	Yht. valittu säilytyspaikaksi	1	33,3	-	-	-	-	1	6,7
Muu tietoarkisto tai arkisto <sup>f</sup>		n = 3		n = 4		n = 8		n = 15	
	Ei lainkaan	3	100	3	75	8	100	14	93,3
	Jonkin verran	-	-	-	-	-	-	-	-
	Suurin osa	-	-	1	25	-	-	1	6,7
	Kaikki	-	-	-	-	-	-	-	-
	Yht. valittu säilytyspaikaksi	-	-	1	25	-	-	1	6,7
<b>Yksilökeskeiset säilytyspaikat (3)</b>									
Oma henkilökohtainen tietokone <sup>g</sup>		n = 4		n = 4		n = 8		n = 16	
	Ei lainkaan	1	25	1	25	6	75	8	50
	Jonkin verran	1	25	1	25	-	-	2	12,5
	Suurin osa	1	25	1	25	1	12,5	3	18,8
	Kaikki	1	25	1	25	1	12,5	3	18,8

	Yht. valittu säilytyspaikaksi	3	75	3	75	2	25	8	50
<b>Muistitikku tai ulkoinen kovalevy<sup>h</sup></b>		n = 3		n = 4		n = 8		n = 15	
	Ei lainkaan	1	33,3	2	50	3	37,5	6	40
	Jonkin verran	1	33,3	1	25	4	50	6	40
	Suurin osa	1	33,3	1	25	-	-	2	13,3
	Kaikki	-	-	-	-	1	12,5	1	6,7
	Yht. valittu säilytyspaikaksi	2	66,7	2	50	5	62,5	9	60
<b>Paperimuoto omassa toimistossa<sup>i</sup></b>		n = 4		n = 4		n = 8		n = 16	
	Ei lainkaan	2	50	1	25	1	12,5	4	25
	Jonkin verran	1	25	3	75	5	62,5	9	56,3
	Suurin osa	1	25	-	-	1	12,5	2	12,5
	Kaikki	-	-	-	-	1	12,5	1	6,25
	Yht. valittu säilytyspaikaksi	2	50	3	75	7	87,5	12	75
<b>Määrittelemätön säilytyspaikka</b>		n = 2		n = 3		n = 5		n = 10	
<b>Muu<sup>j</sup></b>	Ei lainkaan	2	100	3	100	2	40	7	70
	Jonkin verran	-	-	-	-	-	-	-	-
	Suurin osa	-	-	-	-	3	60	3	30
	Kaikki	-	-	-	-	-	-	-	-
	Yht. valittu säilytyspaikaksi	-	-	-	-	3	60	3	30

- a. p-arvo 0.110, Fisher-Freeman-Haltonin eksakti testi.  
b. p-arvo 1.000, Fisher-Freeman-Haltonin eksakti testi.  
c. p-arvo 0.405, Fisher-Freeman-Haltonin eksakti testi.  
d. p-arvo 0.467, Fisher-Freeman-Haltonin eksakti testi.  
e. p-arvo 0.200, Fisher-Freeman-Haltonin eksakti testi.  
f. p-arvo 0.467, Fisher-Freeman-Haltonin eksakti testi.  
g. p-arvo 0.464, Fisher-Freeman-Haltonin eksakti testi.  
h. p-arvo 0.808, Fisher-Freeman-Haltonin eksakti testi.  
i. p-arvo 0.679, Fisher-Freeman-Haltonin eksakti testi.  
j. p-arvo 0.208, Fisher-Freeman-Haltonin eksakti testi.

Vastaajat eivät torju yhtenä rintamana mitään kyselyssä tiedustellusta kymmenestä säilytyspaikka- vaihtoehdosta, mutta vain kolme paikkaa saa heidän enemmistönsä suosion. Niistä taas ainoastaan yksi, oman organisaation tai yksikön palvelin, on tiedeinstituutiotason säilytyspaikka – mutta se on kaikkein suosituin yksittäinen paikka niin tiedeinstituutiotason vaihtoehtoja kuin kaikkia kokonaisuutenakin katsottaessa. Vastaajista kolme neljäsosaa käyttää kyseisenlaisia palvelimia. (Taulukko 32.) Toisaalta kaikki organisaation tai yksikön palvelimen torjuvista vastaajistakaan eivät välttämättä nojaa täysin yksilökeskeisiin säilytyspaikkoihin, sillä paitsi että organisaation oma arkisto saa heidän parissaan kannatusta, myös vaihtoehto ”muu” nousee esiin. (Liite 9 b.) Siten vastaajista vähintään 13 eli neljä viidesosaa käyttää ainakin yhtä tiedeinstituutiotason tarjoamaa säilytyspaikkaa, mutta noin viidesosa saattaa säilyttää aineistoaan vain yksilökeskeisiksi luokiteltavin keinoin.

Tiedeinstituutiotason säilytyspaikat erottuvat myös epäsuosituimpina vaihtoehtoina, sillä arkistot ovat käyttäjämäärien perusteella harvimminkin valittuja aineiston säilytyspaikkoja, vaikkakin ne saavat joka korkeakoulusta yksittäisiä hyödyntäjiä. Tieteenalakohtainen arkisto on arkistovaihtoehtoista

suosituin, sillä sellaisen palveluja hyödyntää noin viidesosa vastaajista, mutta kolme muuta arkistoa – julkaisijan säilytys / julkaisijan arkisto, organisaation oma arkisto ja muu tietoarkisto tai arkisto – jakavat viimeisen sijan. (Taulukko 32.) Liitteestä 9 d on laskettavissa, että kyseisten kolmen arkiston yhteenlaskettu osuus kaikista tehdyistä säilytyspaikkavalinnoista on kuusi prosenttia, ja kaikkien neljänkin arkistovaihtoehdon yhteisosuus on noin kymmenesosa.

TaY:n ja TTY:n vastaajajoukkojen kaikista valintakerroista viidennes kohdistuu oman organisaation tai yksikön palvelimeen, TAMK:n kohdalla neljännes. (Liite 9 d.) Sen sijaan vain TAMK:n vastaajajoukon parissa on tiedeinstituutiotason paikoista organisaation omaa arkistoa hyödyntäviä, ja vain TTY:n joukossa on julkaisijan tarjoamaa säilytyspaikkaa sekä muuta tietoarkistoa tai arkistoa käyttäviä. Lisäksi vain TAMK:n vastaajat jättävät hyödyntämättä tutkimushankkeen palvelinta ja TTY:n vastaajat tieteenalakohtaista säilytyspaikkaa. (Taulukko 32.) Kaikista säilytyspaikkavalinnoista niin kokonaisuudessaan kuin korkeakoulukohtaisestikin noin kaksi viidesosaa kohdistuu tiedeinstituutiotason säilytyspaikkoihin. Yksilökeskeisiin kohdistuu noin puolet kaikista säilytyspaikkavalinnoista, vaikka niitä oli puolet vähemmän kuin tiedeinstituutiotasoisia. TAMK:n ja TTY:n vastaajajoukot kallistuvat hieman TaY:n vastaajia voimakkaammin yksilökeskeiseen säilyttämiseen. (Liitteet 9 d ja 9 e.)

Säilytyspaikkakohtaisia säilytysmääriä tarkasteltaessa neljä säilytyspaikkaa erottuu sellaisina, joissa osa tutkijoista säilyttää *kaiken* aineistonsa. Vain yksi niistä, organisaation tai yksikön palvelin, on tiedeinstituutiotason säilytyspaikka, mutta se erottuu jälleen kaikkein suosituimpana. Se on paikoista ainoa, jossa enemmistö siihen kantaa ottavista säilyttää joko suurimman osan aineistostaan tai kaiken. (Taulukko 32.) Siihen kohdistuu yli puolet vastaajien kaikista kaikki-säilytysmäärävalinnoista ja miltei viidennes suurin osa -valinnoista. (Liite 9 d.) Lisäksi *kaiken* aineistonsa organisaation/yksikön palvelimella säilyttävät eivät säilytä missään muualla keskitetysti *kaikkea* aineistoaan, mikä korostaa kyseisen säilytyspaikan merkittävyyttä. (Liite 9 c.) Arkistot erottuvat epätyypillisimpinä säilytyspaikkoina myös säilytysmäärien näkökulmasta, sillä kukaan vastaajista ei säilytä kaikkea aineistoaan missään neljästä arkistovaihtoehdosta, joskin niiden lisäksi samalla tavalla erottuu tutkimushankkeen palvelin.<sup>14</sup> (Taulukko 32.)

---

<sup>14</sup> Liitteestä 9 d on laskettavissa, että kyseinen viisikko kerää yhdessäkin vajaan viidenneksen suurin osa -valinnoista mutta noin neljänneksen jonkin verran -vaihtoehdon valinnoista.

Tiedeinstituutiotason säilytysmääriä koskevassa korkeakouluvertailussa paljastuu, että TaY:n vastaajista vähintään viisi kahdeksasta säilyttää kaiken aineistonsa keskitetysti.<sup>15</sup> TTY:n vastaajista keskitetysti aineistonsa säilyttää neljannes ja TAMK:n vastaajista korkeintaan puolet. Kukaan TTY:n vastaajista ei säilytä kaikkea aineistoaan yleisesti suosituimmassa vaihtoehdossa eli oman organisaation tai yksikön palvelimella. (Taulukko 32.) Jatkoselvityksen aihe olisi, miten tietty aineisto tai osuus siitä valikoituu tiettyyn säilytykseen. Yksi TaY:n vastaajista ottaakin valintoihinsa kantaa kyselylomakkeen Vapaa sana -osiossa kertoen käyttävänsä ulkoista kovalevyä varmuuskopiooidessaan osaa aineistostaan. Kukaan vastaajista ei ollut epävarma säilytysmäärästään.

Taulukko 33. Säilytysmäärävalinnat säilytysvastuualueittain ja yhteensä (havaintoyksikkönä valintakerrat).

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Tiedeinstituutiotason säilytyspaikat <sup>a</sup>	Ei lainkaan	15	75	18	75	37	77,1	70	76,1
	Jonkin verran	1	5	4	16,7	5	10,4	10	10,9
	Suurin osa	3	15	2	8,3	1	2,1	6	6,5
	Kaikki	1	5	-	-	5	10,4	6	6,5
	Epävarma	-	-	-	-	-	-	-	-
	Yhteensä	20		24		48		92	
Yksilökeskeiset säilytyspaikat <sup>b</sup>	Ei lainkaan	4	36,4	4	33,3	10	41,7	18	38,3
	Jonkin verran	3	27,3	5	41,7	9	37,5	17	36,2
	Suurin osa	3	27,3	2	16,7	2	8,3	7	14,9
	Kaikki	1	9,1	1	8,3	3	12,5	5	10,6
	Epävarma	-	-	-	-	-	-	-	-
	Yhteensä	11		12		24		47	
Määrittelemättömät säilytyspaikat <sup>c</sup>	Ei lainkaan	2	100	3	100	2	40	7	70
	Jonkin verran	-	-	-	-	-	-	-	-
	Suurin osa	-	-	-	-	3	60	3	30
	Kaikki	-	-	-	-	-	-	-	-
	Epävarma	-	-	-	-	-	-	-	-
	Yhteensä	2		3		5		10	
Säilytysmäärät yhteensä <sup>d</sup>	Ei lainkaan	21	63,6	25	64,1	49	63,6	95	63,8
	Jonkin verran	4	12,1	9	23,1	14	18,2	27	18,1
	Suurin osa	6	18,2	4	10,3	6	7,8	16	10,7
	Kaikki	2	6,1	1	2,6	8	10,4	11	7,4
	Epävarma	-	-	-	-	-	-	-	-
	Yhteensä	33		39		77		149	

a. p-arvo 0.233, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.889, Fisher-Freeman-Haltonin eksakti testi.

<sup>15</sup> Viisi heistä valitsee saman säilytyspaikan (organisaation tai yksikön palvelin) kohdalla määrän "kaikki". (Taulukko 32.)

- c. p-arvo 0.208, Fisher-Freeman-Haltonin eksakti testi.
- d. p-arvo 0.494, Fisher-Freeman-Haltonin eksakti testi.

Mitä hajanaisempaa säilyttäminen on, sitä epäsuositumpaa tiedeinstituutiotason säilyttäminen on. Tiedeinstituutiotason säilytyspaikkoja käyttävät painottuvat säilytysmäärävalinnoillaan keskitettyyn säilyttämiseen: tiedeinstituutiotason säilytysvastuualueen hyödyntämistä ilmaisevista valintakerroista (22) yli puolet liittyy vähintään suurimman osan aineistosta säilyttämiseen, mutta yksilökeskeisen säilytysvastuualueen hyödyntämistä ilmaisevista valintakerroista (29) noin kaksi viidesosaa. (Taulukko 33, ks. myös liite 9 d.) Tiedeinstituutiotason vastuualue kerää hieman yli puolet kaikista kaikki-säilytysmäärävalinnoista. Se ja yksilökeskeinen vastuualue ovat tasoissa, kun mukaan laskeetaan myös suurimman osan aineistostaan tietyssä paikassa säilyttävien osuus. Kun kyse on vain aineiston säilyttämisestä suurimmaksi osaksi tietyssä paikassa, yksilökeskeinen säilytysalue vie vuorostaan niukan voiton. Kaikista jonkin verran -vastauksista enää vain runsas kolmannes kohdistuu tiedeinstituutiotason säilytysvastuualueeseen loppujen kohdistuessa yksilökeskeiseen. (Liite 9 d.)

Korkeakouluista TaY:n vastaajajoukko erottuu säilytysmääriltään tiedeinstituutiokeskeisimpänä, sillä heidän tekemistään kaikki-säilytysmäärävalinnoista noin kolme viidesosaa liittyy kyseisen vastuualueen säilytyspaikkoihin. TAMK:n vastaajajoukossa vastaavat valinnat jakautuvat tasan tiedeinstituutiotason ja yksilökeskeisen säilytyksen kesken. Sitä vastoin TTY:n vastaajien parissa kaikkea aineistoa ei säilytetä lainkaan tiedeinstituutiotasoisesti, vaan täysin keskitetylle säilyttämiselle valitaan yksilökeskeinen säilytystapa – kuitenkin puolet heidän suurin osa -säilytysmäärän valinnoistaan kohdistuu tiedeinstituutiotason säilytyspalveluihin. Hajanaisessa säilyttämisessä tiedeinstituutiotason säilytysvastuun asema on yksilökeskeistä heikompi kaikissa korkeakouluissa: siihen kohdistuu vain neljännes TAMK:n vastaajien jonkin verran -valinnoista, runsas kolmannes TaY:n vastaavista valinnoista ja vajaa puolet TTY:nkin valinnoista. (Liite 9 d.)

Kyselyssä tiedusteltiin myös vastaajien näkemyksiä aineiston keskitetystä säilyttämisestä keskusarkistonkaltaisessa paikassa sekä rajoitusten asettamisen että jakamismäärän näkökulmista. Aiemmassa kävi ilmi, että arkistosäilytyksellä ei ole vankkaa sijaa vastaajien keskuudessa, ja seuraavan taulukon luvut kertovat vuorostaan, miten halukkaita tutkijat olisivat niiden käyttämiseen.



Taulukko 34. Näkemykset kaiken ja osan aineiston keskitetystä sijoittamisesta muiden saataville rajoituksetta.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Olen halukas sijoittamaan kaiken aineistoni keskusarkistoon (tms.) saataville ilman rajoituksia. <sup>a</sup>	Vahvasti samaa mieltä	-	-	1	25	-	-	1	6,3
	Osittain samaa mieltä	1	25	-	-	2	25	3	18,8
	En samaa enkä eri mieltä	-	-	-	-	-	-	-	-
	Osittain eri mieltä	1	25	-	-	1	12,5	2	12,5
	Vahvasti eri mieltä	2	50	3	75	5	62,5	10	62,5
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4	100	4	100	8	100	16	100
Olen halukas sijoittamaan ainakin osan aineistostani keskusarkistoon (tms.) saataville ilman rajoituksia. <sup>b</sup>	Vahvasti samaa mieltä	2	50	1	25	4	50	7	43,8
	Osittain samaa mieltä	-	-	-	-	3	37,5	3	18,8
	En samaa enkä eri mieltä	-	-	-	-	-	-	-	-
	Osittain eri mieltä	1	25	2	50	-	-	3	18,8
	Vahvasti eri mieltä	1	25	1	25	1	12,5	3	18,8
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4	100	4	100	8	100	16	100

a. p-arvo 0.807, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.317, Fisher-Freeman-Haltonin eksakti testi.

Tutkimusaineiston jakamiseen suhtaudutaan kielteisesti, kun puhutaan *kaiken* aineiston sijoittamisesta vaikkapa keskusarkistoon kaikkien saataville ilman rajoituksia. Kolme neljäsosaa vastaajista vastustaa sitä ja heistä valtaosa vahvasti. Neljännes suhtautuu suopeasti ajatukseen, mutta enemmistö heistä vain osittain. Vähintään puolet kunkin korkeakoulun vastaajajoukosta vastustaa ajatusta tiukasti. Vastaajien myönteinen suhtautuminen arkistosäilyttämistä kohtaan lisääntyy, kun siirrytään tiedustelemaan näkemyksiä vain osan aineiston saataville saattamisesta rajoituksetta. Silloin neljänneksen sijaan noin kaksi kolmasosaa vastaajista kannattaa aineiston antamista keskitettyyn säilytykseen ja noin kolmannes vastustaa, vain vajaa viidennes vahvasti. Tosin korkeakoulukohmainen vertailu paljastaa, että suurin muutos tapahtuu TaY:n vastaajajoukossa, josta seitsemän kahdeksasosaa on siirtynyt kannattamaan ajatusta, kun taas TTY:n ja TAMK:n vastaajista vähintään puolet on yhä kielteisellä kannalla. (Taulukko 34.)

## Organisaatioiden vaikutus tyytyväisyyteen aineistonhallintaan kohtaan

Taulukko 35. Korkeakoulutaustan vaikutus tyytyväisyyteen tutkimusaineistojen hallintaa kohtaan.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Olen tyytyväinen prosessiin aineistoni tallentamiseksi siksi aikaa, kun projekti on käynnissä	Vahvasti samaa mieltä	3	75	1	25	5	62,5	9	56,3
	Osittain samaa mieltä	-	-	2	50	3	37,5	5	31,3
	En samaa enkä eri mieltä	-	-	-	-	-	-	-	-
	Osittain eri mieltä	-	-	-	-	-	-	-	-
	Vahvasti eri mieltä	-	-	1	25	-	-	1	6,3

(lyhytkestoinen aineistohallinta). <sup>a</sup>	En osaa sanoa	1	25	-	-	-	-	1	6,3
	Yhteensä	4		4		8		16	
Olen tyytyväinen prosessiin aineistoni tallentamiseksi myös projektin päättymisen jälkeistä aikaa ajatellen (pitkäkestoinen aineistohallinta). <sup>b</sup>	Vahvasti samaa mieltä	-	-	-	-	1	12,5	1	6,7
	Osittain samaa mieltä	1	25	2	66,7	1	12,5	4	26,7
	En samaa enkä eri mieltä	1	25	-	-	1	12,5	2	13,3
	Osittain eri mieltä	1	25	-	-	5	62,5	6	40
	Vahvasti eri mieltä	-	-	1	33,3	-	-	1	6,7
	En osaa sanoa	1	25	-	-	-	-	1	6,7
	Yhteensä	4		3		8		15	

a. p-arvo 0.148, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.194, Fisher-Freeman-Haltonin eksakti testi.

Kaikista vastaajista miltei yhdeksän kymmenestä on tyytyväisiä lyhytkestoiseen aineistohallintaansa, yli puolet hyvin tyytyväisiä. Myös kunkin korkeakoulun vastaajista vähintään kolme neljästä on tyytyväisiä. Pitkäkestoisesta aineistohallinnasta puhuttaessa tyytyväisten joukko supistuu kolmannekseen, hyvin tyytyväisten osuus yhteen viidestoistaosaan. Neutraaleja tai epävarmoja vastaajia on viidennes, mikä voi viitata siihen, että pitkäkestoista aineistohallintaa ei ole pohdittu tai sen määritelmä ei ole tuttu. Ainoastaan TTY:n vastaajien enemmistö, kaksi kolmasosaa, on tyytyväisiä ja hekin vain osittain. Tyytymättömimmät vastaajat löytyvät TaY:n riveistä, sillä heistä viisi kahdeksasosaa ilmaisee tyytymättömyytensä. (Taulukko 35.)

Taulukko 36. Ensisijaisen rahoittajan aineistohallintasuunnitelmavaatimuksen vaikutus tyytyväisyyteen tutkimusaineistojen hallintaa kohtaan.

		Aineistohallintasuunnitelmavaatimus							
		Kyllä		Ei		En tiedä		Yhteensä	
		n	%	n	%	n	%	n	%
Olen tyytyväinen prosessiin aineistoni tallentamiseksi siksi aikaa, kun projekti on käynnissä (lyhytkestoinen aineistohallinta). <sup>a</sup>	Vahvasti samaa mieltä	6	75	1	25	1	50	8	57,1
	Osittain samaa mieltä	2	25	2	50	1	50	5	35,7
	En samaa enkä eri mieltä	-	-	-	-	-	-	-	-
	Osittain eri mieltä	-	-	-	-	-	-	-	-
	Vahvasti eri mieltä	-	-	1	25	-	-	1	7,1
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	8		4		2		14	
Olen tyytyväinen prosessiin aineistoni tallentamiseksi myös projektin päättymisen jälkeistä aikaa ajatellen (pitkäkestoinen aineistohallinta). <sup>b</sup>	Vahvasti samaa mieltä	1	12,5	-	-	-	-	1	7,7
	Osittain samaa mieltä	3	37,5	-	-	1	50	4	30,8
	En samaa enkä eri mieltä	-	-	2	66,7	-	-	2	15,4
	Osittain eri mieltä	4	50	-	-	1	50	5	38,5
	Vahvasti eri mieltä	-	-	1	33,3	-	-	1	7,7
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	8		3		2		13	

a. p-arvo 0.487, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.066, Fisher-Freeman-Haltonin eksakti testi.

Sekä rahoittajan aineistohallintasuunnitelmavaatimuksen piirissä olevat että sen ulkopuolella olevat vastaajat ovat tyytyväisempiä lyhytkestoiseen aineistohallintaansa kuin pitkäkestoiseen. Ne, joilta ensisijainen rahoittaja ei vaadi aineistohallintasuunnitelmaa, eivät ilmaise tyytyväisyyttä enää pitkäkestoiseen aineistohallintaansa. Sen sijaan aineistohallintasuunnitelmavaatimuksen piirissä olevista puolet on tyytyväisiä myös aineistojen tallentamiseen projektin jälkeistä aikaa ajatellen eikä kukaan heistä ole erittäin tyytymätön. Silti aineistohallintasuunnitelma ei takaa täyttä tyytyväisyyttä – olettaen, että suunnitelmavaatimuksella on jo ollut vaikutusta vastaajien toimintaan eli että osittaista tyytymättömyyttä ilmaisevat ovat jo laatineet ja toteuttaneet heiltä vaaditun suunnitelman. Aineistohallintasuunnitelman piirissä olevilla on selvä käsitys tyytyväisyydestään, kun taas vaatimuksen ulkopuolella olevien vastaajien parissa on myös neutraalia suhtautumista. Kuitenkin on muistettava, että tässä aineistohallintasuunnitelmavaatimuksen ulkopuolelle määritellyiltä vastaajilta voi jokin muu taho kuin rahoittaja vaatia suunnitelman, tai he voivat toimia itseohjautuvasti. Aineistohallintasuunnitelmavaatimus selittäväenä muuttujana ei saanut aikaan tilastollisesti merkitseviä eroja tutkijoiden näkemyksiin. (Taulukko 36.)

## 6 YHTEENVETO JA POHDINTA

Tämän tutkimuksen tavoitteena oli tuottaa sekä aineistohallinnan että tutkimustyön tukemiseksi lisätietoa tutkimusaineistojen avoimuuden toteutumisesta ja niiden hyödyntämisestä Tampereen yhdistyvissä korkeakouluissa. Tutkimuksessa selvitettiin, missä suhteessa korkeakoulujen tutkijat muistuttavat aineistohallinnan toimijoina toisiaan, mikä tukee aineistohallinnan mahdollisten kehittämistoimien kohdentamista. Tutkijoiden sovittaminen aineistohallinnan eri rooleihin tarjosi keinon vertailla jäsennellysti eritaustaisia tutkijoita: missä rooleissa näkemykset ovat jyrkän kielteisiä ja mihin suhtaudutaan myötämielisesti. Fisher-Freeman-Haltonin eksaktilla testillä saatujen p-arvojen mukaan korkeakoulutausta vaikutti tutkimuksen aineistossa tilastollisesti merkitsevästi selitettävien muuttujien arvoihin kaikkien muiden roolien osa-alueissa paitsi aineistojen käsittelijää tarkastelleessa. Vaikka suuressa osassa aineistohallinnan näkemyksiä ja käytäntöjä havaitut korkeakoulukohtaiset erot eivät olleet tilastollisesti merkitseviä, ne olivat silti todellisia.

### 6.1 Aineistojen jakamatta jättäminen ja avoimuuden ehdollisuus

Kyselyn perusteella tamperelaistutkijoiden tutkimusaineistojen saatavuus oli isolta osin rajoitettua: yli kaksi kolmasosaa vastaajista arvioi, että muut eivät saa heidän aineistojaan helposti käyttöönsä ja samansuuruinen osuus näki muiden tarvitsevan luvan heidän aineistojensa käytölle. Olettaen, että he tulkitsivat kysymysten tarkoittavan vallinnutta olosuhdetta, lupamenettely oli heidän keskuudessaan voimissaan. Toisaalta vallinneet rajoitukset eivät välttämättä olleet tutkijoille mieluisimpia, sillä enemmistö koki lisäehtojen voivan lisätä heidän jakamishalujaan. Siten tutkijoille annettavan vaikutusvallan lisäämisellä aineistojen avoimuuden rajoja määriteltäessä voi olla vaikutusta avoimuuden käytännön toteutukseen. Aineistojaan jo jakavat olivat innokkaimpia lisäämään jakamistaan lisäehtojen myötä – siitä päätellen aineistojen jakamista ei haluta myöskään vähentää nykyisestä. Lisähuomiota on kohdistettava siihen, mikä pitää lainkaan aineistojaan jakavat kannustimien vaikutuspiirin ulkopuolella. Aineiston jakamatta jättäminen on monen tekijän summa, sillä kun tutkijoilta tiedusteltiin, miksi he eivät julkaise tutkimusaineistojaan tai julkaisevat vain osan, kahdestatoista tarjotusta vaihtoehdosta kymmentä pidettiin relevanttina vaikkakin vaihtelevin osuuksin. Esteet voivat olla hyvinkin yksilökeskeisiä, jolloin niiden purkamiseen on vaikeampaa kohdentaa toimia.

Tutkittaessa tamperelaistutkijoiden suhtautumista aineistojensa avoimuuteen keskeisellä sijalla oli ensiksikin aineistohallinnan eri toimijarooleihin liittyneet syyt jättää aineistoja jakamatta ja toiseksi

halut asettaa jakamiselle ehtoja. Kummankin kohdalla oli havaittavissa, että osa rooleista yhdisti enemmän toisaalta TAMK:n ja TaY:n vastaajia ja toisaalta TAMK:n ja TTY:n vastaajia, mutta ei yhtä paljon TaY:n ja TTY:n vastaajia. Linja voi mukaila vastaajien tieteenalataustaa, sillä TTY:n ja TaY:n vastaajilla ei ollut yhteistä taustaa. TaY:n ja TTY:n vastaajiakin yhdisti silti suhtautuminen joihinkin yksittäisiin ehtoihin sekä haluttomuus asettaa tiettyihin rooleihin liittyviä ehtoja.

Taulukko 37. Yhteenvedo aineistonhallinnan rooleista avoimuuden esteitä osoitettaessa ja ehtoja asetettaessa.

Korkeakoulu	Sija	Ehtoja asetettaessa korostuneet roolit	% (valintakerroista)	Sija	Esteitä osoitettaessa korostuneet roolit	% (valintakerroista)
TAMK	1.	Kerääjä	38,4	1.	Suojelija	43
	1.	(Uudelleenkäytön valvoja) <sup>b</sup>	38,4	1.	Tuen saaja	43
	2.	Suojelija	11,5	2.	Kerääjä	14
	2.	Tuen saaja	11,5			
TTY	1.	Kerääjä	86	1.	Tuen saaja	42
	2.	Tiedeyhteisön jäsen	14	2.	Kerääjä	25
				2.	Tiedeyhteisön jäsen	25
			-	(Määrittelemätön) <sup>a, b</sup>	8,3	
TaY	1.	(Uudelleenkäytön valvoja) <sup>b</sup>	49,7	1.	Suojelija	39
	2.	Aineiston kerääjä	33,3	2.	Tuen saaja	27
	3.	Suojelija	7	3.	Kerääjä	19
				4.	(Käsittelijä) <sup>b</sup>	7,7
			-	(Määrittelemätön) <sup>a, b</sup>	7,7	

<sup>a</sup>Viittaa estevaihtoehdon ”muu” saamiin valintakertoihin. Sen täsmennyksistä osa liittyi kerääjän, osa tiedeyhteisön jäsenen rooleihin, mutta olivat jossain määrin tulkittavissa valmiiden estevaihtoehtojen täsmennyksiksi.

<sup>b</sup>Sulkeissa olevat roolit liittyvät vain esteisiin tai vain ehtoihin.

Taulukko 37 kokoaa yhteen, että TAMK- ja TaY-taustaiset vastaajat asettuivat useampiin aineistonhallinnan rooleihin kuin TTY:n vastaajat, etenkin ehtoja nimettäessä. Muistetaan, että TTY:n vastaajista enemmistöä lisäehtojen asettaminen ei myöskään motivoinut lisäämään aineistojensa avoimuutta. Siten mikäli TTY-taustaisten tutkijoiden aineistojen avoimuutta pyritään lisäämään, esteiden purkamisella voi olla suurempi merkitys kuin ehtojen asettamisella. Taulukko kiteyttää myös, että tuen saajan roolia lukuun ottamatta samat toimijaroolit yleensä toistuivat sekä esteistä että ehdoista puhuttaessa, mutta eri painotuksin – samat roolit eivät olleet kärkisijalla ehtoja asetettaessa kuin esteitä nimettäessä. Kun tiettyyn rooliin liittyen ollaan halukkaita asettamaan ehtoja, mutta rooliin liittyvät esteet eivät nouse yhtä voimakkaasti esiin, se voi viitata kyseisen roolin kehittämisen olevan avoimuuden näkökulmasta suoraan jakamiseen kannustavaa; esteen poistamisen jälkeen saatetaan tarvita vielä kannustimia.

## **Esteet**

Esteitä aineistojen avoimuudelle osoitettaessa TAMK:n ja TaY:n vastaajien parissa korostui aineiston suojelijan rooli. Korkeakoulutausta toi rooliin suhtautumiseen tilastollisesti merkitsevän eron. TTY:n vastaajajoukko erottui olemalla yksimielisesti sitä mieltä, että siihen liittyvät syyt eivät vaikuttaneet heidän päätöksiinsä jättää aineistoja jakamatta, siinä missä TaY:n vastaajajoukolle ne olivat keskeisiä. Linjaero voi kertoa tutkimusaineistojen luonteesta tai tiedeyhteisön kulttuurista koskien esimerkiksi aineistojen käyttöoikeuksia. Tuloksen perusteella TAMK- ja TaY-taustaisten parissa voi olla TTY-taustaisten joukkoa hedelmällisempää esimerkiksi koulutuksin varmistaa, että avoimuuden esteiden minimoimiseksi on tehty kaikki mahdollinen jo tutkimuksen suunnitteluvaiheissa, kuten aineiston keruun ja uudelleenkäytön sopimuksia laadittaessa.

Aineiston suojelijan rooliin liitetyistä normikeskeisistä avoimuuden esteistä jo yksin julkaisuoikeuksien puuttuminen keräsi noin viidenneksen kaikista annetuista perusteluista olla tuomatta aineistoja muiden saataville, ja sen valitsi noin puolet vastaajista. Niin ikään Sami Borgin ja Arja Kuulan tutkimuksessa miltei puolet vastaajista nosti esimerkiksi aineistojen omistus- ja käyttösuhteiden epäselvyydet aineiston jatkokäytön esteeksi. Siten ratkaisuja oikeudellisiin kysymyksiin ei ole välttämättä tullut viime vuosina lisää ainakaan niin, että ne olisivat vaikuttaneet tämän tutkielman vastaajien aineistonjakamishalukkuuteen. Se voi viitata siihen, että kaikkia esteitä ei voi täysin poistaa ehtojaan lisäämällä. Normisyihin vedoten aineistojaan muiden saataville tuomatta jättävät eivät erottuneet aineistojaan muita vähemmän jakavina, mikä kertoo muiden syiden vahvasta vaikutuksesta aineistojen avoimuuden toteuttamisessa.

Aineistohallinnan tuen saajan rooliin liittyneet aineistonjakamattajättämisytyt olivat TAMK:n ja TTY:n vastaajien parissa keskeisimmällä sijalla, kun taas TaY:n vastaajajoukossa sen merkitys jäi taka-alalle – tosin tilastollisesti merkitsevää eroa ei ollut. TTY:n kohdalla ajan puute ja TAMK:n kohdalla taitojen puute olivat aineistojen avoimuuden toteutumisen esteinä, mikä kertoo siitä, että tiedeintituutioiden myöntämällä lisäresursseilla voisi olla vaikutusta etenkin heidän käytäntöihinsä. Kaikkiin tutkijoista vain kolmannes nimesi ajan ja taitojen puutteen syyksi sivuuttaa jakaminen. Sen sijaan esimerkiksi Borgin ja Kuulan kyselyn vastaajista yli puolet piti haittana sitä, että tutkijoiden työaikaa menisi aineistohallinnan käytännön toimiin, ja sama tema puhutti muissakin tutkimuksissa. Myönteisesti ajatellen ainakin osa niistä tutkijoista, jotka eivät valinneet taitojen ja ajan puutetta avoimuuden esteiksi, saattoivat nähdä, että heillä oli molempia riittävästi ja aineistojen mahdollinen

jakamatta jättäminen johtui jostakin muusta. Vaihtoehtoisesti kyseinen toimintatapa johtui niin vahvasti jostakin muusta, että ajan ja taitojen puutteeseen ei nähty enää tarvetta ottaa kantaakaan.

Tamperelaistutkijat eivät kokeneet olevansa laiminlyötyjä aineistojen avoimuuden rahoittamisen suhteen, sillä vain pieni osa nimesi rahoituksen puutteen syyksi jättää aineistojen jakaminen väliin. Vieläkin pienempi osa valitsi syyksi rahoittajan vaatimusten puutteen. Se voi kertoa vaatimusten olevan tehokkaita tai rahavirtojen kohdentamista merkityksettömmämpiä. Osa vastaajista, vaikkakin pieni, ei tiennyt rahoittajansa aineistohallintasuunnitelmavaatimuksesta, minkä lisäksi muutama ohitti kysymyksen. Epätietoisuus voi viitata toisaalta siihen, että kyseinen vastaajien osajoukko ei nähnyt aineistohallintaa heille kuuluvana vastuualueena, toisaalta tiukasti rajattuihin vastuualueisiin esimerkiksi tutkimusryhmän sisällä tai ryhmän sisäisen viestinnän suppeuteen.

Aineistohallinnan tuen saajan rooliin liittyen yksi tekijä yhdisti tutkijoita korkeakoulusta riippumatta: he eivät selittäneet tutkimusaineistojen jakamatta jättämistä säilytyspaikan puutteella. Siten he saivat säilytyspaikkojen osalta jo merkittävästi tukea tai vaihtoehtoisesti eivät nähneet säilytyksellä olevan vaikutusta jakamistoimintaansa. Tutkimuskirjallisuuden mukaan tiedossa oleva luonteva toimija, jonka haltuun aineistot siirretään, tukee aineistohallintaa. Vielä 2010-luvun alkupuolella kansainvälisissä tutkimuksissa havaittiin tutkijoiden turvautuvan henkilökohtaisiin säilytystapoihin. Toisaalta miltei kaikki kyselyyni vastanneet hyödynsivät tiedeinstituutiotason säilytystä, etenkin suurempien aineistomäärien kohdalla. Oman organisaation tai yksikön palvelimien asema oli merkittävin. Tiedeinstituutiotason säilyttäminen oli silti sitä vähäisempää mitä pienemmästä aineistomäärästä oli kyse, ja lisäksi arkistoissakin säilytettiin aineistoja pienehköjä määriä. TTY:n vastaajajoukko sivuutti aineistojen keskitetyimmässä säilyttämisessä tiedeinstituutiotason vaihtoehdot, millä voi olla merkittävä vaikutus heidän käytäntöihinsä, kuten aineistojen saavutettavuuteen.

Aiemmissa kyselytutkimuksissa tutkijat näkivät juuri arkistot oikeina toimijoina vastaamaan aineistojen avoimuudesta, ja myös Borg ja Kuula nostivat esiin, että arkistot voivat olla keskeisessä asemassa, kun kehitetään tapoja huomioida tutkimusaineistojen alkuperä ja hallita uudelleenkäyttöön liitettyjä uhkakuvia. Kuitenkaan kyselyyni vastanneita arkistot eivät olleet juuri houkutelleet tarkasteltiinpa sitten säilytyspaikkavalintoja tai tutkimusaineistojen säilytysmääriä. He luottivat omaan organisaatioonsa enemmän kuin ulkopuoliseen tukeen – Heidi Enwaldinkin tutkimuksessa enemmistö näki säilytysvastuun kuuluvan tutkimuslaitoksille. Näin ollen tamperelaistutkijoiden joukkoa ei ollut vielä saatu täysin arkistojen kaltaisten toimijoiden vaikutuspiiriin, varsinkaan isoista säilytysmääristä

puhuttaessa. Voi olla, että arkistoihin siirrettävät aineistojen osat on valikoitu tarkoin. Tutkijat eivät näe aineistojaan homogeenisenä kokonaisuutena, vaan erilaisia aineistoja tai aineistokokonaisuuden eri osia kohdellaan eri tavoin. Kiinnostavaa olisi selvittää, onko eroja niin eri projekteissa kuin yhdenkin tutkimusprojektin sisällä syntyneiden aineistokokonaisuuksien kohtelussa.

Siitä päättäminen, luovuttaako tutkimusaineistoaan ilman rajoituksia keskusarkiston kaltaiseen paikkaan vai ei, ei herättänyt yksittäisessä vastaajassa ristiriitaisia ajatuksia, sillä neutraaleja tai ”en osaa sanoa” -vastauksia ei annettu tällöin lainkaan. Kysely ei tuonut suoraa vastausta siihen, miksi myönteisyys vain osankin aineiston jakamiseen rajoituksitta keskusarkiston kaltaiseen paikkaan jäi vastaajajoukossa noin kolmen viidesosan tasolle – tosin Fältin ja Okuloffin mukaan sekä aineistojen jatkokäyttöön että pelkkään arkistointiin liitetään samoja haittoja. Halutaanko aineistoille asettaa rajoituksia vai vaikuttaako säilytyspaikka, eli se, millaisen keskusarkiston kautta aineistoa jaetaan, siihen, miten voimakkaan positiivisesti ajatukseen aineistojen saataville saattamisesta viime kädessä suhtaudutaan? Tutkijoilla oli vahvat näkemykset. Se, että vähintään puolet TAMK:n ja TTY:n vastaajista suhtautuivat kielteisesti arkistokaltaiseen säilytyspaikkaan aineiston säilytysmäärästä riippumatta, osoitti, että aineistonjakamishalujen kannalta merkitystä on muullakin kuin vain mahdollisuudella vaikuttaa jakamismäärään.

Koska aineistojen jakaminen oli arkistojen käytön vähäisyydestä huolimatta varsin yleistä, tampere-laistutkijat vaikuttavat jakavan tutkimusaineistojaan yksityisemmin. Kirjalliskatsauksesta muistetaan, että palvelimen käyttö, mikä oli vastaajien parissa yleistä, voi edistää aineistojen avoimuutta, saatavilla oloa, organisaation sisällä. Kuitenkin aineisto, joka jaetaan arkistojen ulkopuolella, voi aiempien tutkimusten mukaan olla luonteeltaan tutkimusprosessin aikana analyysin tuloksena luotua eikä jatkokäyttöä varten käsiteltyä.

Kirjallisuudessa korostettiin kilpailullisten tekijöiden ohjaavan tutkijoiden aineistonhallintaa, etenkin aineistojen jakamishaluja. Kuitenkaan aineiston kerääjän rooli ei ollut missään Tampereen korkeakouluista keskeisimmällä sijalla, kun aineistojen avoimuudelle osoitettiin esteitä. Rooliin liitettiin yhtäältä kilpailullinen syy jättää aineistoja jakamatta eli halu julkaista itse ensin omat tulokset. Kyselyni vastaajat jakautuivat suhtautumisessaan sen vaikutuksiin varsin tasan puoltaviin ja torjuviin. Kuitenkin on huomioitava, että kyse on vain väliaikaisesta avoimuuden esteestä; se voisi olla rinnastettavissa ehtoihinkin. Sen merkitystä voidaan lieventää ja vaikutusaikaa lyhentää kehittä-



mällä avoimuuden palkitsevuutta, kuten tekemällä avoimuudesta nopeammin palkitsevaa, ja tuomalla jakamisen monipuolisia hyötyjä esille. Toisaalta aineiston kerääjän roolissa kaikkia vastaajia yhdisti se, että kukaan heistä ei jättänyt aineistojaan jakamatta aineistojensa hallinnan menettämisen pelosta. Linja voi kertoa paitsi omistushalujen puuttumisesta, myös luottamuksesta aineistonhallinnan menetelmiin, joiden avulla hallintaa ei menetetä. Tosin on myös mahdollista, että tutkijat mielsivät hallinnan menettämisessä olevan kyse aineistonhallinnan velvollisuuksista – niistä ei välttämättä haluta pitää kiinni.

Tiedeyhteisön jäsenen rooli ei korostunut minkään korkeakoulun kohdalla osoitettaessa syytä olla tuomatta aineistoja toisten saataville. Kuitenkin TTY:n vastaajat erottuivat tilastollisesti merkittävästi jättämällä aineistoja jakamatta vedoten siihen, että muut eivät tarvitse niitä. Se oli heille keskeisimpiä yksittäisiä avoimuuden esteitä. Näin ollen heidän keskuudessaan voitaisiin aineistonhallintaa suunniteltaessa lisätä keskustelua tutkimusaineistoihin liittyvistä jatkokäytön visioista ja selvittää tarkemmin, miten perusteltu näkemys omien aineistojen käyttökelvottomuudesta muille on.

## **Ehdot**

Kun huomio siirrettiin aineistojen avoimuuden esteistä ehtojen asettamiseen aineistojen avoimuudelle, TAMK:n ja TTY:n vastaajien keskuudessa halu tukea aineiston kerääjän asemaa ja kunniaa eli aineiston kerääjän rooli nousi keskeisimmälle sijalle. Kuitenkin TAMK:n vastaajien roolille antaman osuuden suuruus oli lähempänä TaY:n vastaajien roolille antamaa osuutta kaikista valinnoista.

Ehtoja asetettaessa aineiston kerääjän rooli oli ainoa kaikkien korkeakoulujen vastaajia yhdistävä. Viittauskäytäntöjen kehittämiseen on nähty viime vuosina syytä panostaa, ja myös tässä kyselyssä selvän enemmistön keskuudessa viittausten ja vielä vahvemmin tunnustuksen saaminen oli keskeinen avoimuuteen motivoiva tekijä, kun katsottiin ehtojen kokonaisvalintamäärien lisäksi yksittäisten vastaajien määriä. Kunnian jakamisen käytännöissä voi silti olla yhteensovittavaa, mikäli eri korkeakoulujen tutkijoille laaditaan aiheesta yhteisiä ohjeita tai suunnitellaan yhteisiä tutkimushankkeita. Korkeakoulutausta vaikutti tilastollisesti merkittävästi siihen, miten suhtauduttiin jaettuun tekijyyteen koskien julkaisuja, joissa tutkijan aineistoa on käytetty – TAMK:n vastaajat olivat muita myönteisempiä.

Uudelleenkäytön valvojan roolin saama keskeisin asema ehtoja asetettaessa yhdisti TAMK:n ja TaY:n joukkoja. Korkeakoulutaustan vaikutus suhtautumiseen uudelleenkäytön valvontaan, niin aktiivi-

seen kuin passiiviseenkin, oli tilastollisesti merkitsevä. TTY:n vastaajat eivät nähneet lainkaan tarvetta asettua valvojan rooliin ehtojen asettamisen merkeissä, mutta sen sijaan TAMK:n ja TaY:n vastaajajoukoista vähintään noin kaksi viidesosaa puolsi uudelleenkäytön seurantaan vaikkakin valikoidusti. Suhtautumisessa yksittäisiin ehtoihin tilastollisesti merkitseviä eroja ei ollut. Kuitenkin vähintään puolet TAMK:n ja TaY:n vastaajista suhtautui suopeasti yhteistyöprojektien ja uuden tutkimuksen tavoitteiden ohjailun kaltaiseen valvontaan, mikä tukee tutkimuskirjallisuudessa esille nousutta näkemystä siitä, että yhteistyömahdollisuudet voivat kannustaa tutkijoita jakamaan aineistoja. Kirjallisuudessa nostettiin esiin, että tutkijat pitävät aineistonhallintaa, mukaan lukien uudelleenkäyttäjien neuvontaa, raskaanakin velvoitteena. Koska vastaajistani isolla osalla oli haluja aktiiviseenkin uudelleenkäytön valvontaan, he eivät pitäneet aineistonhallintaa ainakaan siltä osin ylimääräisenä taakkana.

On siis huomioitava, että TaY- ja TAMK-taustaiset tutkijat voivat olla TTY-taustaisia motivoituneempia sitoutumaan aineiston hyödyntämisen pitkäaikaisempaan seurantaan. Kyseinen osa-alue voi olla hedelmällistä pitää vahvasti mukana tutkimuksen ja sen elinkaaren suunnittelussa. Se voi ruokkia tutkimusprojekteissa halua käyttää valmiita aineistoja ja pohtia aineistoihin liittyviä yhteistyömahdollisuuksia, ja siten edistää aineistojen avoimuuden kulttuuria. Aineiston käyttöön liittyvien vuorovaikutuskanavien toimivuuden kehittämiseksi voi olla tilausta. TTY-taustaisten tutkijoiden suhteen osa-alue vaikuttaisi olevan kehityskohde: onko torjuvassa suhtautumisessa uudelleenkäytön valvontaan kyse vain haluttomuudesta seurata muiden tutkimustyötä vai heijasteleeko se esimerkiksi näkemystä tutkimuksen elinkaaresta? Menettävätkö tutkimusaineistot kiinnostavuutensa sen jälkeen, kun tulokset on julkaistu?

Minkään korkeakoulun edustajat eivät olleet ehtojen asettamisessa aktiivisimmillaan niiden liittyessä aineiston suojelijan rooliin. Silti kyseisessä roolissa korkeakoulutausta sai aikaan tilastollisesti merkitsevän eron suhtautumistavoissa aineiston uudelleenkäytön normikeskeiseen ehtoon: TTY:n vastaajajoukko hylkäsi yksimielisesti oikeudellisten lupien tarpeen TAMK:n enemmistön puoltaessa tarvetta ja TaY:n vastaajista puolen ollessa epävarmoja siitä. Aineiston suojelijan rooliin panostaminen ei siten tukisi TTY:n vastaajajoukon aineistonhallinnan kehittämistä – he eivät nähneet siihen liittyviä aineistojen avoimuuden esteitä eivätkä myöskään syytä asettaa ehtoja.

Sekä TTY:n että TaY:n vastaajat torjuivat yksimielisesti tarpeen asettaa aineistonhallinnan tuen saajan rooliin liittyneitä aineiston avoimuuden ehtoja, vaikka yhdistivät rooliin eniten avoimuuden

esteitä. Ainoa aineistonhallinnan tuen saajan rooliin liittynyt tilastollisesti merkitsevä ero koski ehtojen asettamista: aineiston keräämis- ja säilytyskustannuksiin osallistuminen aineiston avoimuuden ehtona sai kannatusta vain TAMK:n vastaajien parissa.

## **6.2 Tutkimusaineistojen (uudelleen)hyödyntämishalut ja -kyvyt**

Tutkimuskirjallisuudessa arveltiin, että aineistoja jätetään merkittävästi jakamatta, koska niillä ei nähdä olevan arvoa alkuperäisen kontekstin ulkopuolella. Nyt neljä viidestä vastaajasta näki kuitenkin aineistoja voitavan käyttää toisella tapaa kuin mihin ne on alkujaan tarkoitettu. Kyselyssä ei kontrolloitu, miten väittämä ”aineistoa saatetaan käyttää toisella tavalla kuin mihin se on tarkoitettu” tulkittiin. Ensiksikin vastaaja saattoi miettiä niin omien aineistojensa uudelleenkäyttöä kuin itseäänkin muiden keräämien aineistojen käyttäjänä. Toiseksi väittämä voitiin lukea joko neutraalina toteamuksena tai kielteisenä näkymänä riippuen siitä, mietittiinkö aineiston käytön monipuolisuutta vai tahallisia tai tahattomia väärinkäyttömahdollisuuksia. Korkeakoulujen välillä oli tilastollisesti merkitseviä eroja näkemyksissä aineistojen monikäyttöisyydestä.

Kun huomio oli aineiston uudelleenhyödyntämisen tavoitteissa, tutkijat eivät suhtautuneet eri aineistojen yhdistelyyn toimintana kielteisesti, sillä yli neljä viidesosaa hyväksyi uusien kokonaisuuk- sien kokoamisen jaetuista aineistoista. Tosin kun huomio siirtyi hyödyntämisen mahdollisuuksiin, kolmasosa ilmaisi tyytymättömyyttä kykyihinsä yhdistellä aineistoja eri lähteistä, mikä on siten kehityskohde Tampereen korkeakouluissa: taitojen puute ei koske vain aineistojen avoimeksi saattamista, vaan myös tutkijoiden tapoja hyödyntää aineistoja.

Mielenkiintoinen ero aineistonhallinnan näkemyksissä ja käytännöissä koski tutkijoiden intoa jakaa aineistoja ja käyttää jaettuja aineistoja, mikä havaittiin myös tutkimuskirjallisuudessa. Carol Tenopirin et al. tutkimuksessa omia aineistojaan jakoi noin puolet ja kolmannes piti niitä helposti saatavilla olevina, mutta yli neljä viidesosaa oli halukas käyttämään helposti saavutettavia aineistoja. Tässä kyselyssä hieman yli neljännes vastaajista piti aineistojaan helposti saatavilla olevina – tosin siinä havaittiin tilastollisesti merkitsevä korkeakoulukohtainen ero, sillä myönteisesti suhtautuvan osuuden takana oli vain TaY:n vastaajia. Sitä vastoin vähän yli puolet olisi käyttänyt toisten tutkijoiden aineistoja, jos ne olisivat helposti saatavilla, noin kolmanneksen asettuessa kielteiselle kannalle. Toisaalta muistetaan, että huomattava enemmistö, kolme neljäsosaa, jakoi vähintään jonkin verran aineistojaan, ja samansuuruinen osuus ilmaisi suoraan olevansa valmis jakamaan aineistojaan laajan tutkijajoukon kanssa. Näin ollen, jos katsotaan jaetun aineiston helpon saavutettavuuden sijasta

vain jakamismääriä ja halua jakaa aineistoja laajalle tutkijajoukolle, kyselyyn vastanneet olivatkin jonkin verran haluttomampia käyttämään muiden tuottamia aineistoja kuin jakamaan aineistoaan. Toisaalta TTY:n vastaajien keskuudessa osuudet menivät tasan: puolet suhtautui avoimin mielin muiden tuottamien aineistojen käyttöön ja puolet oli halukkaita jakamaan aineistoja laajalle tutkijajoukolle. Kuitenkin TaY:n vastaajista yli neljä viidesosaa oli valmis jakamaan aineistojaan laajalle yleisölle, mutta kolme neljäsosaa oli valmis käyttämään muiden tuottamia aineistoja niiden olleessa helposti saatavilla. TAMK:n vastaajista kolme neljästä oli valmis jakamaan aineistojaan laajalle tutkijajoukolle, mutta aineistojen helppo saavutettavuus olisi houkuttanut vain yhtä neljästä käyttämään toisten aineistoja.

Tamperelaistutkijoiden joukko jakoi siis tutkimusaineistojaan jossain määrin ilman hyödyntämishalua. Kävi ilmi, että enemmistö ei odottanut muilta vastavuoroista aineiston jakamista. Kenties vastaavasti itseltäkään ei odotettu valmiiden aineistojen hyödyntäjäyttä, vaikka uudelleenkäyttöön ei suhtauduttukaan kielteisesti ja omien aineistojen uudelleenkäyttöä oltiin valmiita tietyissä rajoissa seuraamaan ja siihen myös osallistumaan – valmiita ainakin ajatuksen tasolla. Tutkijat saattoivat nähdä, että aineistojen uudelleenkäyttö ei kosketa juuri heitä itseään käytännön tasolla. Tällöin aineistojen jakamisella nähtiin olevan arvoa tai merkitystä sinänsä, ja jaettuja aineistoja ei nähty tutkimustoiminnan kohteena. Kyseistä näkemystä tukee se, että vajaa viidennes vastaajista oli kokenut pääsyn puuttumisen toisten oman alan tutkijoiden aineistoihin rajoittaneen heidän kykyään ratkaista tutkimusongelmia, mutta jonkin verran isompi osuus, runsaat kaksi viidesosaa, näki, että pääsyn puuttuminen heidän omien tieteenalojensa aineistoihin rajoittaa silti merkittävästi tieteen kehittymistä. Lisäksi TaY:n vastaajista neljänneksellä oli kokemusta toisten tamperelaisten korkeakoulujen aineistojen hyödyntämisestä tutkimustyössään, mutta koska heidän keskuudessaan pääsyn puuttumista toisten tamperelaisten korkeakoulujen aineistoihin ei pidetty tutkimuksen esteenä, ilmeisesti aiempaa kokemusta ei pidetty ratkaisevan tärkeänä tehdylle tutkimukselle; kenties käytetty aineisto oli ollut vain tutkimusta täydentävässä osassa, mikä tutkimuskirjallisuudenkin mukaan oli tyypillinen tapa käyttää valmiita tutkimusaineistoja. Toisaalta korkeakoulujen yhdistymisen kynnyksellä enemmistöllä kunkin korkeakoulun vastaajista ei ollut kokemusta naapurikorkeakoulujen aineistojen käytöstä, mutta vain neljännes suhtautui jyrkän kielteisesti niiden käyttämiseen tutkimusongelmiensa ratkaisemisessa. Siten valmiiden aineistojen käyttämiseen on vähintäänkin orastavaa myönteisyyttä, mutta tavoitteellisuus toiminnan toteuttamiseksi voi puuttua.

Tarkasteltaessa tutkijoita aineiston käsittelijöinä huomio kiinnittyi puolestaan ristiriitaan, jossa tyytyväisyys aineistonkäyttö- ja -hallintaprosesseihin, esimerkiksi analysointiin ja etsintään, oli suurta huolimatta ensiksikin metadatan ja metadatastandardien vähäisestä käytöstä, toiseksi korkeasta tyytymättömyydestä metadatan ja dokumentoinnin työkaluihin sekä kolmanneksi siitä, että aineistojen käytettävyydessä oli ongelmia. Käytettävyyden ongelmiin viittasi se, että vähintään puolet piti aineiston monimutkaisuutta ja huonolaatuisuutta tutkimuksen riskitekijöinä. Toimintatapojen ja tyytyväisyyden välinen ristiriita voi selittyä paitsi sillä, että potentiaalisesti haastavia valmiita aineistoja ei edes harkita käytettävän, myös sillä, että tietyt tulkinnallisesti vaikeina pidetyt aineistot rajataan pois jo varhaisessa vaiheessa tutkimusprosessia eli ne eivät päädy esimerkiksi analysoitaviksi asti. Siten onnistuneen aineistonhallinnan yhteyttä tutkimuksen alkuvaiheisiin, tutkimuskysymysten asetteluun asti, ei välttämättä aina tunnusteta: järjestelmällisempi kuvailu voisi tehdä useammista aineistoista ymmärrettäviä. Nyt noin viidenneksellä vastaajista oli vahvaa kokemusta siitä, että he olivat jättäneet oman alansa tutkijoiden aineistoja käyttämättä niiden vaikeaselkoisuuden vuoksi – ja vain puolet kielsi jyrkästi toimineensa niin. Vain kahdeksasosa oli pitänyt vahvasti muiden alojen aineistojen vaikeaa ymmärrettävyyttä esteenä tutkimuksen teolle. Viimeksi mainitun taustalla voi luonnollisesti olla, että toisten alojen aineistoja ei pidetä vaikeasti ymmärrettävinä, mutta hyvinkin myös se, että muiden alojen aineistoja ei ole käytettykään, jolloin ei ole merkitystä, ovatko ne vaikeaselkoisia vai eivät. Edellä kuvatulla ristiriidalla ja sen mahdollisilla selityksillä voi olla vaikutusta myös tutkimusaineistojen uudelleenkäyttöhaluihin.

Tutkimustulokseni vahvistavat siis osaltaan aiempien tutkimusten luomaa käsitystä siitä, että metadatastandardien ja metadatan käyttö ei ole tutkijoille läheistä. Ne tukevat myös aiemmassa tutkimuksessa toistunutta viestiä siitä, että metadatakäytäntöjen omaksuminen vaatii johdonmukaisia perehdyttämistoimia. Vastaajajoukon parissa arkistojen vähäinen hyödyntäminen ja aineistojen yleinen keskittämätön säilyttäminen saattoivat myös vaikuttaa metadatan käytön vähyyteen ja kuvailun mahdolliseen vapaamuotoisuuteen. Koska aineiston hyödyntämisen prosessit olivat tutkijoista onnistuneita, vaikka esimerkiksi metadatan tuottamiseen tarkoitettuihin työkaluihin oltiin tyytymättömiä, voi olla, että heillä oli aineiston etsinnälle, kuvailulle ja luetteloinnille muut vapaamuotoisemmat käytännöt ja järjestelmät – tai työkalujen mahdollisesta hankaluudesta tai huonoudesta huolimatta lopputuloksiin oltiin tyytyväisiä. Kirjallisuuden perusteella tutkijoiden aika tai motivaatio ei riitä aineiston kuvailuun, mutta vastaajieni enemmistön tyytyväisyys aineistojensa kuvailun ja

luetteloinnin prosessiin viittaa heidän harjoittaneen vapaamuotoista aineiston kuvailua. Lisäksi heidän voi nähdä pitäneen käytäntöjään toimivina, vaikka ne eivät välttämättä noudattaneetkaan standardeja. Kuitenkin tutkimusaineistojen monimutkaisuus ja huono laatu ovat tekijöitä, joihin liittyviä huolia ei saa sivuuttaa.

### **6.3 Tutkimuksen jatkonäkymät**

Tutkijoiden vastauksia analysoitaessa kyselylomakkeesta löytyi kehityskohteita. Esimerkiksi tulkinanvaraista oli, ajattelivatko vastaajat osan väittämistä viittaavan vastaamishetkenä vallinneeseen vai ihanteelliseen tilanteeseen. Tosin kaikkia kysymyksiä ei olisi ollut realistista muotoilla täysin yksiselitteisiksi, sillä kysymysten lukeminen ei saa vaatia liikaa vastaajan aikaa. Kuitenkin selventäviä lisäkysymyksiä olisi, kokeeko vastaaja jo voivansa asettaa ehtoja aineistojensa avoimuudelle riittävästi, miten isossa tutkimusryhmässä hän työskentelee ja kenelle hän näkee aineistoa koskevien oikeuksien kuuluvan, jolloin saadaan tietoa siitä, miten itsenäisiä päätöksiä hän pystyy tekemään aineistonhallinnasta.

Tulosten perusteella tutkijoiden motivointi tutkimusaineistojen uudelleenkäyttäjiksi sekä valmiiden tutkimusaineistojen valikointitavat tutkimusten suunnitteluvaiheessa ovat aiheita, joiden tarkastelua voisi jatkossa syventää. Kun osalla aineistojaan jakavista vastaajista ei välttämättä ollut itsellään mielenkiintoa käyttää muiden tuottamia aineistoja, herää kysymys, miksi tutkijat haluavat itse koskemattomia aineistoja. Huomio voitaisiin kohdistaa aineiston kerääjän ja uudelleenkäyttäjän väliseen vuorovaikutukseen, kuten monien ilmaisemien yhteistyötoiveiden toteutumisen edellytyksiin. Tutkimuksen painopiste olisi tällöin voimakkaammin aineistojen uudelleenkäyttäjien näkemyksissä. Nyt vain pieni osa kyselystä keskittyi vastaajien omiin kokemuksiin aineistojen uudelleenkäytöstä. Tässä tutkimuksessa rajoitteena olikin, että kyselylomaketta ei laadittu tulosten esittämisessä käytettyä viitekehystä vastaavaksi, vaan tulokset tulkittiin viitekehystä apuna käyttäen ja myös näkemys viitekehyksestä tarkentui vasta tuloksia analysoitaessa.

Aihepiirin tutkimisen kannalta lupaavaa on, että kyselyyn vastanneet jättivät varsin vähän epävarmoja vastauksia ja noin kolmannes vastaajista jätti kommentin myös Vapaa sana -osioon. Se viittaa siihen, että tutkijat olivat pohtineet aineistoihin liittyviä teemoja ja muodostaneet niistä mielipiteitä ja olivat motivoituneita osallistumaan tutkimusaineistojen käsittelyn kehittämiseen. Tosin tuloksia tulkittaessa pitää huomioida, että vastaajajoukko saattoi koostua aihepiiristä tavallista enemmän kiinnostuneista.

# LÄHTEET

## Internet-sivut

Avoimen tieteen kansallinen koordinaatio, Tieteellisten seurain valtuuskunta. Avointiede.fi. Vierailtu 1.12.2020.

DVV 2020a: *Avoindata.fi*-palvelu. Vierailtu 1.12.2020.

DVV 2020b: Digi- ja väestötietovirasto, ”Mitä on avoin data?” *Avoindata.fi*-palvelu. Haettu 1.12.2020 osoitteesta: [<https://www.avoindata.fi/fi/opas/mita-on-avoin-data>].

DVV 2020c: Digi- ja väestötietovirasto, ”Palvelukuvaus”. *Avoindata.fi* -palvelu. Haettu 1.12.2020 osoitteesta: [<https://www.avoindata.fi/fi/opas/palvelukuvaus>].

Open Knowledge Foundation 2020: Open Knowledge Foundation, ”Open Definition 2.1”. Haettu 14.2.2021 osoitteesta: [<https://opendefinition.org/od/2.1/en/>].

TAMK 2018: TAMK:n verkkosivut, ”Tutkimusluvut”. Haettu 11.6.2018 osoitteesta: [<http://www.tamk.fi/web/tamk/tutkimusluvut>].

Tampere3 2018: Tampere3:n verkkosivut, ”Tampere3: Hyvinvoivan yhteiskunnan kehittäjä, moottori ja tulkki”. Haettu 27.4.2018 osoitteesta: [<https://www.tampere3.fi/tietoa-meista>].

## Kirjallisuus, artikkelit ja raportit

Alakyyny et al. 2018: Ala-Kyyny, Juuso, Korhonen, Tuija, & Roinila, Markku, ”Tutkimusdatan avaamisen esteet: haastattelututkimus Helsingin yliopistossa”. *Signum* 49(4), 2017, s. 25-29. Haettu 6.3.2021 osoitteesta: [<https://doi.org/10.25033/sig.69198>].

Anderson 2016: Anderson, Valerie, R., ”Introduction to Mixed Methods Approaches”. Teoksessa Jason, Leonard A. & Glenwick, David S. (toim.), *Handbook of Methodological Approaches to Community-Based Research: Qualitative, Quantitative and Mixed Methods*. Oxford University Press, New York 2016, s. 233-241.

Avoimen tieteen koordinaatio 2020: Avoimen tieteen koordinaatio, Tieteellisten seurain valtuuskunta, *Avoimen tieteen ja tutkimuksen julistus 2020–2025*. Vastuullisen tieteen julkaisusarja 1:2020. Tiedonjulkistamisen neuvottelukunta ja Tieteellisten seurain valtuuskunta, Helsinki 2020. Haettu 6.3.2021 osoitteesta: [<https://doi.org/10.23847/isbn.9789525995213>].

Borg & Kuula 2007: Borg, Sami & Kuula, Arja, *Julkisrahoitteisen tutkimusdatan avoin saatavuus ja elinkaari. Valmisteluraportti OECD:n datasuosituksen toimeenpanosta Suomessa*. Yhteiskuntatieteellisen tietoarkiston julkaisuja 6/2007. Haettu 6.3.2021 osoitteesta:

[<https://www.fsd.tuni.fi/fi/tietoarkisto/julkaisut/julkaisusarja/fsdjs06/>].

- Boscoe 2020: Boscoe, Bernadette M., "The What of Data: Defining Which Scientific Research Is Appropriate to Share". Teoksessa Sundqvist, Anneli, Berget, Gerd, Nolin, Jan & Skjerdingsstad, Kjell Ivar (toim.) 2020, s. 687-694.
- Davidson 2019: Davidson, Judith, *Qualitative research and complex teams*, Oxford University Press, New York 2019.
- Donnelly 2012: Donnelly, Martin, "Data management plans and planning", Teoksessa Pryor, Graham (toim.), *Managing Research Data*. Facet Publishing, London 2012, s. 83-103.
- Enwald 2018: Enwald, Heidi, "Tutkimusaineiston avoin jakaminen – tutkimusorganisaatioiden jäsenten käsityksiä, kokemuksia ja mielipiteitä". *Informaatiotutkimus* 37(4), 2018, s. 26-47. Haettu 4.3.2021 osoitteesta: [<https://doi.org/10.23978/inf.77411>].
- Euroopan komissio 2020: Euroopan komissio, *Komission tiedonanto Euroopan parlamentille, neuvostolle, Euroopan talous- ja sosiaalikomitealle ja alueiden komitealle. Euroopan datastrategia*. Bryssel 19.2.2020. Haettu 4.3.2021 osoitteesta: [<https://eur-lex.europa.eu/legal-content/FI/TXT/HTML/?uri=CELEX:52020DC0066&from=EN>].
- Euroopan unionin neuvosto 2016: *Council Conclusions on the Transition towards an Open Science System (27.5.2016)*, Bryssel 27.5.2016. Haettu 4.3.2021 osoitteesta: [<https://data.consilium.europa.eu/doc/document/ST-9526-2016-INIT/en/pdf>].
- FCRD 2018: Finnish Committee for Research Data & Laine, Heidi (toim.), *Tracing data: data citation roadmap for Finland*. Finnish Committee for Research Data, Helsinki 2018. Haettu 15.3.2021 osoitteesta: [<http://urn.fi/URN:NBN:fi-fe201804106446>].
- Fecher et al. 2015: Fecher, Benedikt, Friesike, Sascha & Hebing, Marcel, "What Drives Academic Data Sharing?" *PLoS ONE* 10(2), 2015. Haettu 15.3.2021 osoitteesta: [<https://doi.org/10.1371/journal.pone.0118053>].
- Fisher 1922: Fisher, R. A., "On the Interpretation of  $\chi^2$  from Contingency Tables, and the Calculation of P". *Journal of the Royal Statistical Society* 85(1), 1922, s. 87-94.
- Freeman & Halton 1951: Freeman, G. H. & Halton, J. H., "Note on an Exact Treatment of Contingency, Goodness of Fit and Other Problems of Significance". *Biometrika* 38(1-2), 1951, s. 141-149.
- Fält & Okuloff 2016: Fält, Katja & Okuloff, Annaleena, *Avoin tiede ja tutkimusaineistot -kysely 2015: raportti terveystieteiden ja humanististen alojen tutkijoille suunnatun kyselyn tuloksista*.



Yhteiskuntatieteellinen tietoaarkisto, Tampere 2016.

- Haendel et al. 2012: Haendel, Melissa A. & Vasilevsky, Nicole A. & Wirz, Jacqueline, A., "Dealing with Data: A Case Study on Information and Data Management Literacy". *PLoS Biology* 10(5), 2012, s. 1-4. Haettu 4.3.2021 osoitteesta: [<https://doi.org/10.1371/journal.pbio.1001339>].
- Heikkilä 2014: Heikkilä, Tarja, *Tilastollinen tutkimus*. 9. uudistettu painos vuodelta 2017. Edita Publishing Oy, Helsinki 2014.
- Higgins 2012: Higgins, Sarah, "The lifecycle of data management". Teoksessa Pryor, Graham (toim.), *Managing Research Data*. Facet Publishing, London 2012, s. 17-46.
- Jones 2012: Jones, Sarah, "Research data policies: principles, requirements and trends". Teoksessa Pryor, Graham (toim.), *Managing Research Data*. Facet Publishing, London 2012, s. 47-82.
- Lampela 2016: Lampela, Anssi, *Johdatus suomalaisen arkistokuvailun historiaan*. TRIM Research Reports 20. Tampereen yliopisto 2016.
- Moles et al. 2013: Moles, Angela & Dickie, John B. & Flores-Moreno, Habacuc, "A response to Poisot et al.: Publishing your dataset is not always virtuous". *Ideas in Ecology and Evolution* 6(2), 2013, s. 20–22.
- NSF 2013: National Science Foundation, *Proposal Award Policies and Procedures Guide*. Haettu 2.12.2020 osoitteesta: [<https://www.nsf.gov/pubs/policydocs/pappguide/nsf13001/gppg-print.pdf>].
- Piwowar 2011: Piwowar, Heather A., "Who Shares? Who Doesn't? Factors Associated with Openly Archiving Raw Research Data". *Plos ONE* 6(7), 2011, s. 1-13. Haettu 15.3.2021 osoitteesta: [<https://doi.org/10.1371/journal.pone.0018657>].
- Poisot et al. 2013: Poisot, Timothée E. & Mounce, Ross & Gravel, Dominique, "Moving toward a sustainable ecological science: don't let data go to waste!" *Ideas in Ecology and Evolution* 6(2), 2013, s. 11–19.
- Pryor 2012: Pryor, Graham, "Why manage research data?" Teoksessa Pryor, Graham (toim.), *Managing Research Data*. Facet Publishing, London 2012, s. 1-16. doi:10.4033/iee.2013.6b.14.f
- Punch 2003: Punch, Keith F., *Survey Research: The Basics*. SAGE Publications Ltd, London 2003.
- Päällysaho & Latvanen 2019: Päällysaho, Seliina & Latvanen, Jaana, "Avoin toimintakulttuuri vahvistuu ammattikorkeakouluissa". *Signum* 51(2), 2019, s. 25–29. Haettu 4.3.2021 osoitteesta: [<https://doi.org/10.25033/sig.83527>].
- Ruxton & Neuhäuser 2010: Ruxton, Graeme D. & Neuhäuser, Markus, "Good practice in testing for an association in contingency tables". *Behavioral ecology and sociobiology* 64(9), 2010, s.

1505-1513. Haettu 15.3.2021 osoitteesta: [<https://doi-org.libproxy.tuni.fi/10.1007/s00265-010-1014-0>].

Sauro 2013: Sauro, Jeff, "Best Practices for Using Statistics on Small Sample Sizes". *Measuring U* [<https://measuringu.com>] 13.8.2013. Haettu 4.3.2021 osoitteesta: [<https://measuringu.com/small-n/>].

Schreier 2012: Schreier, Margrit, *Qualitative Content Analysis in Practice*, SAGE, Los Angeles 2012.

Schreier 2019: Schreier Margrit & Stamann, Christoph & Janssen, Markus & Dahl, Thomas & Whittal, Amanda, "Qualitative Content Analysis: Conceptualizations and Challenges in Research Practice—Introduction to the FQS Special Issue 'Qualitative Content Analysis I'". *Forum: Qualitative Social Research* 20(3), 2019.

Stickdorn et. al 2018: Stickdorn, Marc & Hormess, Markus Edgar & Lawrence, Adam & Schneider, Jacob, *This is service design doing*. O'Reilly Media, Incorporated, Sebastopol (CA) 2018.

Sundqvist et al. 2020: Sundqvist, Anneli & Berget, Gerd & Nolin, Jan & Skjerdingsstad, Kjell Ivar (toim.), *Sustainable Digital Communities: 15th International Conference, IConference 2020, Borås, Sweden, March 23-26, 2020, Proceedings*. Springer International Publishing AG, Cham 2020. Haettu 23.5.2020 osoitteesta: [<https://doi.org/10.1007/978-3-030-43687-2>].

Suomen Akatemia 2017a: Suomen Akatemia, *Rahoituksen yleiset ehdot ja ohjeet 2017–2018*. AKA/4/02.04.10/2017. Haettu 4.3.2021 osoitteesta: [[https://www.aka.fi/globalassets/10rahoitus/rahoituksen-kayton-ehdot/yleiset\\_ehdot\\_270917.pdf](https://www.aka.fi/globalassets/10rahoitus/rahoituksen-kayton-ehdot/yleiset_ehdot_270917.pdf)].

Suomen Akatemia 2017b: Suomen Akatemia, *Yksityiskohtainen aineistohallintasuunnitelmaohje*. Haettu 5.12.2017 osoitteesta (viimeksi muokattu 21.6.2017): [<http://www.aka.fi/fi/rahoitus/nain-haet/hakuohjeet/akatemian-aineistohallintasuunnitelmaohje-dmptuulissa/>].

Suomen Akatemia 2017c: *Avoim tiede: avoin julkaiseminen ja aineistojen avaaminen*. Haettu 10.12.2017 (viimeksi muokattu 28.8.2017) osoitteesta: [<http://www.aka.fi/fi/rahoitus/hyvatutkimustapa/tieteen-avoimuus/>].

Tenopir et al. 2011: Tenopir, Carol & Allard, Suzie & Douglass, Kimberly & Umur Aydinoglu, Arsev & Wu, Lei & Read, Eleanor & Manoff, Maribeth & Frame, Mike, "Data Sharing by Scientists: Practices and Perceptions". *PLoS ONE* 6(6), 2011. Haettu 15.3.2021 osoitteesta: [<https://doi.org/10.1371/journal.pone.0021101>].

Tenopir et al. 2015: Tenopir, Carol & Dalton, Elizabeth D. & Allard, Suzie & Frame, Mike & Pjesivac, Ivanka & Birch, Ben & Pollock, Danielle & Dorsett, Kristina, "Changes in Data Sharing and Data

- Reuse Practices and Perceptions among Scientists Worldwide”. *PLoS ONE* 10(8), 2015. Haettu 16.3.2021 osoitteesta: [<https://doi.org/10.1371/journal.pone.0134826>].
- TJNK 2018: Tiedonjulkistamisen neuvottelukunta, *Viesti rohkeasti, vaikuta vastuullisesti. Tiedeviestinnän suositukset*. Tiedonjulkistamisen neuvottelukunta 2018. Haettu 29.11.2020 osoitteesta: [[https://www.tjnk.fi/sites/tjnk.fi/files/tiedeviestinnan\\_suosituks\\_2018.pdf](https://www.tjnk.fi/sites/tjnk.fi/files/tiedeviestinnan_suosituks_2018.pdf)].
- Tolvanen 2017: Tolvanen, Kristiina, *Tutkimuksen infrastruktuurien kartoitus Tampere3-korkeakouluissa. Raportti, marraskuu 2017*. Haettu 27.4.2018 osoitteesta: [[https://www.tampere3.fi/filebank/4782-Tutkimusinfrastruktuurit\\_Tampere3ssa\\_Kartoitus\\_marraskuu\\_2017\\_RA-PORTTI.pdf](https://www.tampere3.fi/filebank/4782-Tutkimusinfrastruktuurit_Tampere3ssa_Kartoitus_marraskuu_2017_RA-PORTTI.pdf)].
- Tähtinen et al. 2020: Tähtinen, Juhani & Laakkonen, Eero & Broberg, Mari, *Tilastollisen aineiston käsittelyn ja tulkinnan perusteita*. Turun yliopiston kasvatustieteiden tiedekunnan julkaisuja C: 22, Turun yliopiston kasvatustieteiden laitos 2020.
- Uusitalo 2017: Uusitalo, Mikko, ”Avoin data TAMKissa”. Teoksessa Kenttälä-Koivumäki, Tiina & Saarinen, Susanna (koostaneet julkaisun), *TAMK-konferenssi - TAMK Conference 2017. Learning and working together*. Tampereen ammattikorkeakoulun julkaisuja, Tampereen ammattikorkeakoulu, Tampere 2017, s. 167-171. Haettu 4.3.2021 osoitteesta: [<http://julkaisut.tamk.fi/PDF-tiedostot-web/Muut/TAMK-Conference2017.pdf>].
- Van den Eynden et al. 2016: Van den Eynden, Veerle & Knight, Gareth & Vlad, Anca & Radler, Barry & Tenopir, Carol & Leon, David & Manista, Frank & Whitworth, Jimmy & Corti, Louisa, *Towards open research. Practice, Experiences, Barriers and Opportunities. Survey of Wellcome researchers and their attitudes to open research*. Wellcome Trust, Journal contribution, 2016.
- Volk et al. 2014: Volk, Carol J. & Lucero, Yasmin & Barnas, Katie, ”Why is data sharing in collaborative natural resource efforts so hard and what can we do to improve it?” *Environmental Management* 53(5), 2014, s. 883-893. DOI 10.1007/s00267-014-0258-2
- Wessels et al. 2017: Wessels, Bridgette & Finn, Rachel L. & Wadwha, Kush & Sveinsdottir, Thordis, *Open Data and the Knowledge Society*. Amsterdam University Press, Amsterdam 2017. Haettu 4.3.2021 osoitteesta: [<http://library.oapen.org/handle/20.500.12657/31743>].
- Wetcher-Hendricks 2011: Wetcher-Hendricks, Debra, *Analyzing quantitative data: an introduction for social researchers*. John Wiley & Sons, Inc., Hoboken, New Jersey 2011.
- White et al. 2013: White, Ethan P. & Baldrige, Elita & Brym, Zachary T. & Locey, Kenneth J. & McGlenn, Daniel J. & Supp, Sarah R., ”Nine simple ways to make it easier to (re)use your data”. *Ideas in Ecology and Evolution* 6(2), 2013, s. 1-10. doi:10.4033/iee.2013.6b.6.f

# LIITE 1. Kyselylomake.

## KYSELYTUTKIMUS TAMPERELAISTUTKIJOIDEN AINEISTONHALLINNASTA

[saateviesti ks. liite 2 a]

### TAUSTATIEDOT

1. Syntymävuosi [alaspöytävalikko]

2. Sukupuoli

- Mies
- Nainen
- En halua vastata
- Muu

3. Työpaikka

- Tampereen ammattikorkeakoulu
- Tampereen teknillinen yliopisto
- Tampereen yliopisto
- Muu

4. Mikä seuraavista kuvaa parhaiten ensisijaista tutkimusalaasi?

- Maatalous ja luonnonvarat
- Liikeala ja hallinto
- Kasvatustieteet
- Tekniikka, teknologia ja tietojenkäsittely
- Fysiikka, kemia ja matematiikka
- Informaatiotieteet
- Lääketiede ja sosiaali- ja terveysala
- Psykologia
- Yhteiskuntatieteet ja humanistiset tieteet
- Kulttuuriala
- Matkailu-, ravitsemis- ja talousala
- Muu \_\_\_\_\_

5. Nykyinen asemasi työyhteisössä

- Professori
- Apulaisprofessori
- Lehtori
- Tuntiopettaja
- Tutkijatohtori/postdoc
- Tohtoriopiskelija
- Tutkinto-opiskelija
- Muu \_\_\_\_\_

## OMA TUTKIMUS JA OMAT TUTKIMUSAINEISTOT

Seuraava kysymysosio koskee tutkimustasi ja tutkimusaineistoasi.

6. Mitkä seuraavasti käsitteistä kuvaavat parhaiten käyttämäsi tutkimusaineistoa (valitse kaikki käyttämäsi aineistotyytit)? [monivalinta]

- Abioottinen tutkimus (maaperä, mikroilmasto jne.)
- Bioottinen tutkimus
- Tietomalliaineisto
- Kokeellinen aineisto (jonkin verran manipulointia)
- Havainnoiva aineisto (ei manipulointia)
- Haastatteluaineisto
- Kaukokartoitettu abioottinen aineisto
- Kaukokartoitettu bioottinen aineisto
- Yhteiskuntatieteellinen tutkimus
- Muu \_\_\_\_\_

7. Mitä metadastandardeja käytät tällä hetkellä, kun kuvaillet tutkimusaineistoa – vai käytätkö mitään? [monivalinta]

- Organisaationi käyttämää metadataohjeistusta
- Laboratorioni/yksikköni käyttämää metadataohjeistusta
- En mitään
- En osaa sanoa
- Muu \_\_\_\_\_

8. Mikä seuraavista kuvaa parhaiten tutkimuksesi tyypillistä rahoittajaa?

- Suomen Akatemia
- Tekes
- TSR
- Muu \_\_\_\_\_

9. Vaatiiko ensisijainen rahoittajasi sinua laatimaan aineistonhallintasuunnitelman?

- Kyllä
- Ei
- En tiedä

10. Kuinka paljon säilytät tutkimusaineistoasi seuraavissa sijaissa? Valitse kunkin kohdalla parhaiten kuvaava vaihtoehto.

	En lainkaan	Jonkin verran	Suurimman osan	Kaiken	En ole varma
Organisaationi tai yksikköni palvelimella	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tutkimushankkeen palvelimella	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Omalla henkilökohtaisella tietokoneellani	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Muistitikulla tai ulkoisella kovalevyllä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Paperimuodossa omassa toimistossani	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tieteenalakohtaisessa säilytyspaikassa/arkistossa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Julkaisijan säilyttämänä / julkaisijan arkistossa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Organisaationi omassa arkistossa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Muussa tietoarkistossa tai arkistossa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Muu	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11. Kuinka paljon saatat tutkimusaineistoasi muiden saataville?

- En lainkaan
- Jonkin verran
- Suurimman osan
- Kaiken

12. Jos tutkimusaineistosi tai osa tutkimusaineistoistasi ei ole muiden saatavilla, miksi ei? Valitse kaikki sopivat vaihtoehdot. [monivalinta]

- Rahoitus puuttuu.
- Standardit puuttuvat.
- Muut eivät tarvitse aineistoja.
- Ei ole aikaa saattaa aineistoja saataville.
- Aineistoilleni ei ole sopivaa säilytyspaikkaa.
- Aineistojeni ei kuulu olla julkisia.
- Minulla ei ole oikeuksia julkaista aineistoja.
- Rahoittaja ei edellytä aineistojeni jakamista.
- Menettäisin mahdollisuuden hallita aineistoja itse.
- Minun tarvitsee julkaista ensin omat tutkimustulokseni.
- Minulla on puutteelliset taidot aineistojeni julkiseksi saattamiseksi.
- Muu \_\_\_\_\_

13. Seuraavat väittämät koskevat sitä, miten keräät ja käytät tutkimusaineistoa. Kerro, missä määrin olet samaa mieltä seuraavien väittämien kanssa, jotka päättävät tämän lauseen: Olen tyytyväinen...

	Vahvasti samaa mieltä	Osittain samaa mieltä	En samaa enkä eri mieltä	Osittain eri mieltä	Vahvasti eri mieltä	En osaa sanoa
...prosessiin, jolla kerään aineistoni.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...prosessiin, jolla luetteloin/kuvailen aineistoni.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...prosessiin aineistoni tallentamiseksi siksi aikaa, kun projekti on käynnissä (lyhytkestoinen aineistonhallinta).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...prosessiin aineistoni tallentamiseksi myös projektin päättymisen jälkeistä aikaa ajatellen (pitkäkestoinen aineistonhallinta).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...prosessiin, jolla etsin omia aineistoja.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...prosessiin liittyen aineistoni analysoimiseen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...työkaluihin, joilla laaditaan metadata ja dokumentointi.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

14. Seuraavat väittämät liittyvät tutkimusaineiston jakamiseen ja saatavuuteen. Kerro, missä määrin olet samaa mieltä väittämien kanssa:

	Vahvasti samaa mieltä	Osittain samaa mieltä	En samaa enkä eri mieltä	Osittain eri mieltä	Vahvasti eri mieltä	En osaa sanoa
Jaan aineistoani muiden kanssa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Muut tarvitsevat luvan käyttääkseen aineistoani.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Muut saavat helposti aineistoani käyttöönsä.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## NÄKEMYKSET TUTKIMUSAINEISTON KÄYTÖSTÄ JA UUELLEENKÄYTÖSTÄ TUTKIMUKSESSA

15. Seuraavassa on kyse mahdollisista ehdoista, joita tutkimusaineistojen käyttämiseen liittyy. Kerro kunkin ehdon kohdalla, pidätkö sen täyttymistä keskeisenä, ennen kuin muut voivat käyttää tutkimusaineistojasi.

	Kyllä	En	En ole varma
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aineiston kerääjälle <i>annetaan tunnustus</i> kaikissa töissä, joissa aineistoa on käytetty.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aineiston kerääjään <i>viitataan</i> kaikissa töissä, joissa aineistoa on käytetty.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aineiston kerääjällä on mahdollisuus tehdä yhteistyötä projektin kanssa, jossa aineistoa käytetään.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tuloksia, jotka perustuvat aineistoon, ei saa levittää missään muodossa ilman aineiston kerääjän hyväksyntää.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tuloksia, jotka pohjautuvat aineiston hyödyntämiseen, ei voi levittää ilman, että aineiston kerääjällä olisi mahdollisuutta arvioida tuloksia ja tehdä ehdotuksia ja kommentteja.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aineiston kerääjälle pitää antaa kopio artikkeleista, joissa hänen aineistoaan on käytetty.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aineiston kerääjä saa listan kaikista tuotoksista, joissa hänen aineistojaan on käytetty, mukaan lukien artikkelit, esitykset, opetusmateriaalit jne.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aineiston käytölle on hankittava oikeudelliset luvat.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yhteinen ymmärrys vastavuoroisesta aineiston jakamisesta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aineiston kerääjä saa ja hyväksyy lausunnon koskien aineiston uudelleenikäytön tavoitteita.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ainakin osa aineiston keräämis- ja säilytyskustannuksista on katettava.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

16. Seuraavat väittämät liittyvät tieteellisen tutkimusaineiston käyttöön. Kerro, missä määrin olet samaa mieltä kunkin väittämän kanssa. Ensimmäiset väittämät koskevat aineistoa yleisesti, seuraavat oman tieteenalasi aineistoja ja viimeiseksi tiedustellaan suhtautumistasi toisten tamperelaisten korkeakoulujen aineistoihin.

	Vahvasti samaa mieltä	Osittain samaa mieltä	En samaa enkä eri mieltä	Osittain eri mieltä	Vahvasti eri mieltä	En osaa sanoa
Aineistoa saatetaan tulkita väärin sen monimutkaisuuden vuoksi.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aineistoa saatetaan tulkita väärin sen huonon laadun vuoksi.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aineistoa saatetaan käyttää toisella tavalla kuin mihin se on tarkoitettu.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Olen itse jättänyt toisten oman tieteenalani tutkijoiden aineistoja käyttämättä niiden vaikeaselkoisuuden vuoksi.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pääsyn puuttuminen toisten oman tieteenalani tutkijoiden tuottamiin aineistoihin on suuri este tieteen kehitykselle.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pääsyn puuttuminen toisten oman tieteenalani tutkijoiden tuottamiin aineistoihin on rajoittanut kykyäni vastata tieteellisiin kysymyksiin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pääsyn puuttuminen toisten tamperelaisten korkeakoulujen tutkimusaineistoihin on este	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

tutkimuksen tekemiselle.						
Olen hyödyntänyt toisten tamperelaisten korkeakoulujen tutkimusaineistoja omien tutkimusongelmieni ratkaisemiseksi.	0	0	0	0	0	0
Olen halukas hyödyntämään toisten tamperelaisten korkeakoulujen tutkimusaineistoja omien tutkimusongelmieni ratkaisemiseksi.	0	0	0	0	0	0
Olen pitänyt muiden tieteenalojen tutkimusaineistojen vaikeaa ymmärrettävyyttä esteenä tutkimuksen tekemiselle.	0	0	0	0	0	0

17. Seuraavat väittämät liittyvät tieteellisen tiedon jakamiseen. Kerro, missä määrin olet samaa mieltä kunkin väittämän kanssa.

	Vahvasti samaa mieltä	Osittain samaa mieltä	En samaa enkä eri mieltä	Osittain eri mieltä	Vahvasti eri mieltä	En osaa sanoa
Käyttäisin toisten tuottamia aineistoja, jos ne olisivat helposti saatavilla.	0	0	0	0	0	0
Olen halukas jakamaan aineistoja laajan tutkijajoukon kesken.	0	0	0	0	0	0
Olen halukas sijoittamaan ainakin osan aineistostani keskusarkistoon (tms.) saataville ilman rajoituksia.	0	0	0	0	0	0
Olen halukas sijoittamaan kaikki aineistoni keskusarkistoon (tms.) saataville ilman rajoituksia.	0	0	0	0	0	0
Saattaisin todennäköisemmin aineistojani saataville, jos voisin asettaa saatavuudelle ehtoja.	0	0	0	0	0	0
On tärkeää, että aineistoihini viitataan, jos muut tutkijat käyttävät niitä.	0	0	0	0	0	0
On oikein koota uusia aineistokokonaisuuksia jaetuista aineistoista.	0	0	0	0	0	0
Olen tyytyväinen kykyihini yhdistää aineistoja eri lähteistä, kun etsin vastauksia tutkimuskysymyksiin.	0	0	0	0	0	0

#### VAPAA SANA

Tässä voit halutessasi vielä täydentää vastauksiasi ja/tai antaa palautetta kyselystä.

#### KIITOKSET

Kiitos osallistumisestasi ja tuestasi tutkimukselle! Valmis tutkielma julkaistaan aikanaan Tampereen yliopiston avoimessa TamPub-julkaisuarkistossa.



## LIITE 2. Saatekirjeet.

### *Liite 2 a. Alkuperäinen saatekirje korkeakoulujen intranet-sivuja varten.*

Hyvä tamperelaisessa korkeakoulussa työskentelevä tutkija,

Oletko miettinyt ja tiedostanut, miten käsittelet tutkimusaineistojasi ja miten suhtaudut muiden tuottamiin tutkimusaineistoihin? Entä miten kollegasi tamperelaisissa korkeakouluissa toimivat? Arvostaisin, jos voisit käyttää hetken vastataksesi aihetta kartoittavaan kyselyyni.

Laadin Tampereen yliopistossa tieto- ja asiakirjahallinnan opintosuunnassa pro gradu -tutkielmaa tampere-laistutkijoiden aineistonhallinnasta. Tutkimukseni tavoitteena on selvittää teidän näkemyksiänne ja käytän-töjänne tutkimusaineistojen käsittelyn, jakamisen ja käytön suhteen, sekä Tampere3-hengessä vertailla kolmen eri korkeakoulun tutkijoiden linjauksia. Näin rakentuu kuva, miten samanlaisia tai erilaisia seudun yh-distyvät korkeakoulut ovat tutkimusaineiston hallinnan näkökulmasta.

Kyselyyn vastaaminen vie noin 15 minuuttia. Varsinaisia osa-alueita on kolme ja kysymyksiä seitsemän-toista. Kyselylomake pohjautuu suurilta osin Carol Tenopirin (et al.) kansainvälisessä kyselytutkimuksessa "Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide" (PLoS ONE, 2015) käytettyyn lomakkeeseen.

Kyselyyn pääset vastaamaan tämän linkin kautta: <https://goo.gl/forms/E8zjXjtJlxULUxs2>

Vastausaika on juhannusviikon loppuun (24.6.) asti. Vastaukset käsitellään luottamuksellisesti.

Vastauksista kiittäen,  
Anna-Riikka Pasanen, FM  
Tampereen yliopisto  
Viestintätieteiden tiedekunta  
Informaatiotutkimus ja interaktiivinen media  
Pasanen.Anna-Riikka.J@student.uta.fi

Työni ohjaajana on yliopistonlehtori Pekka Henttonen (Pekka.Henttonen@uta.fi).

Dear researchers of Tampere University of Applied Sciences,

have you paid attention on how you manage your research data and how you deal with research data produced by others? What about your colleagues in other universities of Tampere – how do they work with their data? I would appreciate if you could spend a few minutes to answer my survey that has its focus on those questions. The survey however is only in Finnish.

I'm a student in the University of Tampere, majoring in information studies. This survey is for my master's thesis in which I study the data management behaviour among researchers who work in the Tampere region. The purpose of my research is first to find out what kind of views and practices regarding the managing, sharing and use of research data you have. Secondly, in the spirit of Tampere3, I aim to compare how similar or diverse you researchers are with your data management views and practices in these three universities that are joining their forces.

Answering the questionnaire takes about 15 minutes. There are three main topics and seventeen questions. The questionnaire is based largely on the form used by Carol Tenopir et al. in their international survey

"Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide" (PLOS ONE, 2015).

You have time to reply until the 24th of June. The answers will be kept confidential and answering is anonymous. My work is supervised by university lecturer Pekka Henttonen (Pekka.Henttonen@uta.fi).

Anna-Riikka Pasanen  
University of Tampere  
The Faculty of Communication Sciences (COMS)  
Information Studies and Interactive Media  
Pasanen.Anna-Riikka.J@student.uta.fi

***Liite 2 b. TAMK:n intranet-sivuja varten muokattu saatekirje.***

Hyvä tamperelaisessa korkeakoulussa työskentelevä tutkija / jatko-opiskelija / opetushenkilöstöön kuuluva,

Oletko miettinyt ja tiedostanut, miten käsittelet esimerkiksi TKI-toiminnassa syntyviä aineistoja ja miten suhtaudut muiden tuottamiin aineistoihin? Entä miten kollegasi tamperelaisissa korkeakouluissa toimivat? Arvostaisin, jos voisit käyttää hetken vastataksesi aihetta kartoittavaan kyselyyni.

Laadin Tampereen yliopistossa tieto- ja asiakirjahallinnan opintosuunnassa pro gradu -tutkielmaa tampere-laistutkijoiden aineistohallinnasta. Tutkimukseni tavoitteena on selvittää teidän näkemyksiänne ja käytän-töjänne tutkimusaineistojen käsittelyn, jakamisen ja käytön suhteen, sekä Tampere3-hengessä vertailla kol-men eri korkeakoulun tutkijoiden linjauksia. Näin rakentuu kuva, miten samanlaisia tai erilaisia seudun yh-distyvät korkeakoulut ovat tutkimusaineiston hallinnan näkökulmasta.

Kyselyyn vastaaminen vie noin 15 minuuttia. Varsinaisia osa-alueita on kolme ja kysymyksiä seitsemän-toista. Kyselylomake pohjautuu suurilta osin Carol Tenopirin (et al.) kansainvälisessä kyselytutkimuksessa "Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide" (PLOS ONE, 2015) käytettyyn lomakkeeseen.

Kyselyyn pääset vastaamaan tämän linkin kautta: <https://goo.gl/forms/E8zjXjtJlxULUxs2>

Vastausaikaa on juhannusviikon loppuun (24.6.) asti. Vastaukset käsitellään luottamuksellisesti.

Vastauksista kiittäen,  
Anna-Riikka Pasanen, FM  
Tampereen yliopisto  
Viestintätieteiden tiedekunta  
Informaatiotutkimus ja interaktiivinen media  
Pasanen.Anna-Riikka.J@student.uta.fi

Työni ohjaajana on yliopistonlehtori Pekka Henttonen (Pekka.Henttonen@uta.fi).

Dear teachers/researchers of Tampere University of Applied Sciences,

have you paid attention on how you manage your research and development data and how you deal with data produced by others? What about your colleagues in other universities of Tampere – how do they work with their data? I would appreciate if you could spend a few minutes to answer my survey that has its focus on those questions. The survey however is only in Finnish.

I'm a student in the University of Tampere, majoring in information studies. This survey is for my master's thesis in which I study the data management behaviour among researchers who work in the Tampere region. The purpose of my research is first to find out what kind of views and practices you have regarding the managing, sharing and use of research data. Secondly, in the spirit of Tampere3, I aim to compare how similar or diverse you researchers are with your data management views and practices in these three universities that are joining their forces.

Answering the questionnaire takes about 15 minutes. There are three main topics and seventeen questions. The questionnaire is based largely on the form used by Carol Tenopir et al. in their international survey "Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide" (PLoS ONE, 2015). Here is the link to my survey: <https://goo.gl/forms/E8zjXjtTJlxULUxs2>

You have time to reply until the 24th of June. The answers will be kept confidential and answering is anonymous. My work is supervised by university lecturer Pekka Henttonen (Pekka.Henttonen@uta.fi).

Best regards,

Anna-Riikka Pasanen  
University of Tampere  
The Faculty of Communication Sciences (COMS)  
Information Studies and Interactive Media  
Pasanen.Anna-Riikka.J@student.uta.fi

## **LIITE 3. Lista tutkielman tekemistä koskeneesta sähköpostinvaihdosta 2018–2020.**

Kaikki viestit ovat Anna-Riikka Pasasen hallussa.

TAMK:n kehittämysyksikkö = kehittamisyksikko@tamk.fi

TAMK:n viestintäyksikkö = viestinta@tamk.fi

TaY:n viestintäyksikkö/-palvelut = viestinta@uta.fi

TTY:n viestintäyksikkö/-palvelut = viestinta@tut.fi

Sähköpostiviesti 22.4.2018: Anna-Riikka Pasasen viesti Tampere3-hankkeen kehittämisideoita keränneeseen sähköpostiosoitteeseen tampere3@tut.fi.

Sähköpostiviesti 23.4.2018: TTY:n tiedottaja Saara Riihimäen viesti Anna-Riikka Pasaselle [vastaus Anna-Riikka Pasasen tampere3@tut.fi-osoitteeseen lähettämään viestiin].

Sähköpostiviesti 25.4.2018a: Anna-Riikka Pasasen viesti TAMK:n, TTY:n ja TaY:n viestintäyksiköihin.

Sähköpostiviesti 25.4.2018b: TTY:n tiedottaja Saara Riihimäen viesti Anna-Riikka Pasaselle.

Sähköpostiviesti 25.4.2018c: TAMK:n viestinnän (viestinta@tamk.fi) viesti Anna-Riikka Pasaselle.

Sähköpostiviesti 7.5.2018: Anna-Riikka Pasasen viesti TaY:n viestintäpalveluille (viestinta@uta.fi).

Sähköpostiviesti 8.5.2018: Kyösti Koskelan [TaY:n viestintäpalvelut] viesti Anna-Riikka Pasaselle.

Sähköpostiviesti 10.5.2018: Anna-Riikka Pasasen viesti TAMK:n kehittämysyksikölle.

Sähköpostiviesti 15.5.2018a: TAMK:n johdon assistentti Heidi Moksénin viesti Anna-Riikka Pasaselle.

Sähköpostiviesti 15.5.2018b: TAMK:n kehittämysyksikön viesti Anna-Riikka Pasaselle.

Sähköpostiviesti 16.5.2018: Anna-Riikka Pasasen viesti TAMK:n kehittämysyksikölle.

Sähköpostiviesti 17.5.2018a: TAMK:n kehittämysyksikön viesti Anna-Riikka Pasaselle.

Sähköpostiviesti 17.5.2018b: TAMK:n kehittämysyksikön viesti Anna-Riikka Pasaselle.

Sähköpostiviesti 10.6.2018a: Anna-Riikka Pasasen viesti TaY:n viestintäpalveluille.

Sähköpostiviesti 10.6.2018b: Anna-Riikka Pasasen viesti TAMK:n kehittämysyksikölle.

Sähköpostiviesti 10.6.2018c: Anna-Riikka Pasasen sähköpostiviesti TTY:n viestintäyksikköön.

Sähköpostiviesti 11.6.2018a: TaY:n viestintäpalvelujen viesti Anna-Riikka Pasaselle.

Sähköpostiviesti 11.6.2018b: TAMK:n kehittämysyksikön viesti Anna-Riikka Pasaselle.

Sähköpostiviesti 11.6.2018c: Marika Vapaala-Riikosen (TTY:n viestintä) viesti Anna-Riikka Pasaselle.

Sähköpostiviesti 14.6.2018: Anna-Riikka Pasasen viesti TaY:n viestintäpalveluille.

Sähköpostiviesti 15.6.2018: TaY:n viestintäpalvelujen viesti Anna-Riikka Pasaselle.

Sähköpostiviesti 15.10.2020: Anna-Riikka Pasasen viesti Tampereen yliopiston tilastotieteen lehtori Jarkko Isotalolle.

Sähköpostiviesti 17.10.2020: Tampereen yliopiston tilastotieteen lehtori Jarkko Isotalon viesti Anna-Riikka Pasaselle.

# LIITE 4. Tampereen yliopiston intranet-sivuilla 7.5.2018 jaettu tiedote CSC:n tutkimustietovarantoja koskeneesta kyselystä.

## Kysely tutkijoille: Millainen tutkimustietovaranto olisi paras?

julkaistu 07.05.2018 12.12, Viestintä TaY

CSC - Tieteen tietotekniikan keskus Oy haluaa ajatuksianne ja kehitysideoitanne tutkimustietovarannon tietosisällön ja käyttäjäportaalin suunnittelun tueksi. Voit halutessasi vastata vain osaan kysymyksistä. Olemme erityisen kiinnostuneita tunnistamaan niitä tutkimuksen hallintotyöhön liittyviä ongelmakohtia, joissa tietovirtojen automatisoinnista ja eri lähteissä olevien tietojen uudelleenkäyttämisestä olisi erityisesti hyötyä.

CSC - Tieteen tietotekniikan keskus Oy toteuttaa valtakunnallista tutkimustietovarantoa opetus- ja kulttuuriministeriön toimeksiannosta. Tutkimustietovaranto kokoaa yhteen esimerkiksi julkaisuja, tutkimusaineistoja, tutkimushankkeita, tutkimusinfrastruktuureita, tutkijoita ja tutkimusryhmiä koskevia metatietoja. Tietovirtojen automatisointi ja järjeistäminen esimerkiksi rahoittajien, tutkijoiden kotiorganisaatioiden ja tutkijan käyttämien muiden palveluiden välillä keventävät hallinnollista työtaakkaa.

Kysely on anonyymi.

Huom. Kaikki kysymykset ovat vapaaehtoisia. Voit vastata hautessasi vain osaan kysymyksistä. Toivomme vastauksia 8.6.2018 mennessä.

[Kysely löytyy CSC - Tieteen tietotekniikkakeskuksen sivuilta esittelytekstin jälkeen.](#)

Lisätietoja:

tutkimustietovaranto (at) csc.fi

Walter Rydman

walter.rydman (at) csc.fi

+358 50 381 8521

Takaisin listaukseen

### **Kommentit**

---

Ei kommentteja vielä. Ole ensimmäinen.

## LIITE 5. Kyselylomakkeen sovittaminen viitekehykseen.

### Viitekehyksen osa-alueet

0. Taustoitus
1. Aineiston keruu – aineiston kerääjä
2. Normit – aineiston suojelija
3. Tiedeyhteisö – tiedeyhteisön jäsen
4. Uudelleenkäyttö – uudelleenkäyttäjä ja uudelleenkäytön valvoja
5. Aineiston käytettävyys – aineiston käsittelijä
6. Tiedeinstituutiot – aineistonhallinnan tuen saaja

Kyselylomake	Viitekehyksen osa-alueet							
	0.	1.	2.	3.	4.	5.	6.	
<b>1. Syntymävuosi</b>	x							
<b>2. Sukupuoli</b>	x							
<b>3. Työpaikka</b>	x						x	
<b>4. Ensisijainen tutkimusala</b>	x							
<b>5. Nykyinen asema työyhteisössä</b>	x							
<b>6. Tutkimusaineistotyyppi</b>	x							
<b>7. Mitä metadastandardeja käytät tällä hetkellä, kun kuvaillet tutkimusaineistoasi – vai käytätkö mitään?</b>						x		
<b>8. Mikä seuraavista kuvaa parhaiten tutkimuksesi tyyppillistä rahoittajaa?</b>	x							
<b>9. Vaatiiko ensisijainen rahoittajasi sinua laatimaan aineistonhallintasuunnitelman?</b>							x	
<b>10. Kuinka paljon säilytät tutkimusaineistoasi seuraavissa tilanteissa? [10 vaihtoehtoa]</b>							x	
<b>11. Kuinka paljon saatat tutkimusaineistoasi muiden saataville?</b>	x							
<b>12. Jos tutkimusaineistosi tai osa tutkimusaineistoistasi ei ole muiden saatavilla, miksi ei?</b>								
<i>Rahoitus puuttuu.</i>							x	
<i>Standardit puuttuvat.</i>						x		
<i>Muut eivät tarvitse aineistoja</i>				x				
<i>Ei ole aikaa saattaa aineistoja saataville.</i>							x	
<i>Aineistoilleni ei ole sopivaa säilytyspaikkaa.</i>							x	
<i>Aineistojeni ei kuulu olla julkisia.</i>			x					
<i>Minulla ei ole oikeuksia julkaista aineistoja</i>			x					
<i>Rahoittaja ei edellytä aineistojeni jakamista.</i>							x	
<i>Menettäisin mahdollisuuden hallita aineistoja itse.</i>		x						
<i>Minun tarvitsee julkaista ensin omat tutkimustulokseni.</i>		x						
<i>Minulla on puutteelliset taidot aineistojeni julkiseksi saattamiseksi.</i>							x	
<i>Muu</i>	Ei ennalta määritelty							
<b>13. Seuraavat väittämät koskevat sitä, miten keräät ja käytät tutkimusaineistoa. Kerro, missä määrin olet samaa mieltä seuraavien väittämien kanssa, jotka päättävät tämän lauseen: Olen tyytyväinen...</b>								
<i>...prosessiin, jolla kerään aineistoni.</i>						x		
<i>...prosessiin, jolla luetteloin/kuvailen aineistoni.</i>						x		

<i>...prosessiin aineistoni tallentamiseksi siksi aikaa, kun projekti on käynnissä (lyhytkestoinen aineistohallinta).</i>								X
<i>...prosessiin aineistoni tallentamiseksi myös projektin päättymisen jälkeistä aikaa ajatellen (pitkäkestoinen aineistohallinta).</i>								X
<i>...prosessiin, jolla etsin omia aineistojani.</i>							X	
<i>...prosessiin, jolla analysoin aineistoni.</i>							X	
<i>...työkaluihin, joilla tuotetaan metatietoja ja aineiston dokumentointi.</i>							X	
<b>14. Seuraavat väittämät liittyvät tutkimusaineiston jakamiseen ja saatavuuteen. Kerro, missä määrin olet samaa mieltä väittämien kanssa:</b>								
<i>Jaan aineistojani muiden kanssa.</i>	X							
<i>Muut tarvitsevat luvan käyttääkseen aineistojani.</i>	X							
<i>Muut saavat helposti aineistojani käyttöönsä.</i>				X				
<b>15. Seuraavassa on kyse mahdollisista ehdoista, joita tutkimusaineistojen käyttämiseen liittyy. Kerro kunkin ehdon kohdalla, pidätkö sen täyttymistä keskeisenä, ennen kuin muut voivat käyttää tutkimusaineistojasi.</b>								
<i>Jaettu tekijyys (co-authorship) koskien julkaisuja, joissa aineistoa on käytetty</i>		X						
<i>Aineiston kerääjälle annetaan tunnustus kaikissa töissä, joissa aineistoa on käytetty</i>		X						
<i>Aineiston kerääjään viitataan kaikissa töissä, joissa aineistoa on käytetty</i>		X						
<i>Aineiston kerääjällä on mahdollisuus tehdä yhteistyötä projektin kanssa, jossa aineistoa käytetään</i>						X		
<i>Tuloksia, jotka perustuvat aineistoon, ei saa levittää missään muodossa ilman aineiston kerääjän hyväksyntää</i>						X		
<i>Tuloksia, jotka pohjautuvat aineiston hyödyntämiseen, ei voi levittää ilman, että aineiston kerääjällä olisi mahdollisuutta arvioida tuloksia ja tehdä ehdotuksia ja kommentteja</i>						X		
<i>Aineiston kerääjälle pitää antaa kopio artikkeleista, joissa hänen aineistoaan on käytetty</i>						X		
<i>Aineiston kerääjä saa listan kaikista tuotoksista, joissa hänen aineistojaan on käytetty, mukaan lukien artikkelit, esitykset, opetusmateriaalit jne.</i>						X		
<i>Aineiston käytölle on hankittava oikeudelliset luvat</i>			X					
<i>Yhteinen ymmärrys vastavuoroisesta aineiston jakamisesta</i>				X				
<i>Aineiston kerääjä saa ja hyväksyy lausunnon koskien aineiston uudelleenkäytön tavoitteita</i>					X			
<i>Ainakin osa aineiston keräämis- ja säilytyskustannuksista on katettava</i>								X
<b>16. Seuraavat väittämät liittyvät tieteellisen tutkimusaineiston käyttöön. Kerro, missä määrin olet samaa mieltä kunkin väittämän kanssa. Ensimmäiset väittämät koskevat aineistoa yleisesti, seuraavat oman tieteenalasi aineistoja ja viimeiseksi tiedustellaan suhtautumistasi toisten tamperelaisten korkeakoulujen aineistoihin.</b>								
<i>Aineistoa saatetaan tulkita väärin sen monimutkaisuuden vuoksi.</i>							X	
<i>Aineistoa saatetaan tulkita väärin sen huonon laadun vuoksi.</i>							X	

<i>Aineistoa saatetaan käyttää toisella tavalla kuin mihin se on tarkoitettu.</i>					x		
<i>Olen itse jättänyt toisten oman tieteenalani tutkijoiden aineistoja käyttämättä niiden vaikeaselkoisuuden vuoksi.</i>						x	
<i>Pääsyn puuttuminen toisten oman tieteenalani tutkijoiden tuottamiin aineistoihin on suuri este tieteen kehitykselle.</i>				x			
<i>Pääsyn puuttuminen toisten oman tieteenalani tutkijoiden tuottamiin aineistoihin on rajoittanut kykyäni vastata tieteellisiin kysymyksiin.</i>					x		
<i>Pääsyn puuttuminen toisten tamperelaisten korkeakoulujen tutkimusaineistoihin on este tutkimuksen tekemiselle.</i>				x			
<i>Olen hyödyntänyt toisten tamperelaisten korkeakoulujen tutkimusaineistoja omien tutkimusongelmieni ratkaisemiseksi.</i>					x		
<i>Olen halukas hyödyntämään toisten tamperelaisten korkeakoulujen tutkimusaineistoja omien tutkimusongelmieni ratkaisemiseksi.</i>					x		
<i>Olen pitänyt muiden tieteenalojen tutkimusaineistojen vaikeaa ymmärrettävyyttä esteenä tutkimuksen tekemiselle.</i>						x	
<b>17. Seuraavat väittämät liittyvät tieteellisen tiedon jakamiseen. Kerro, missä määrin olet samaa mieltä kunkin väittämän kanssa.</b>							
<i>Käyttäisin toisten tuottamia aineistoja, jos ne olisivat helposti saatavilla</i>					x		
<i>Olen halukas jakamaan aineistoja laajan tutkijajoukon kesken</i>				x			
<i>Olen halukas sijoittamaan ainakin osan aineistostani keskusarkistoon (tms.) saataville ilman rajoituksia.</i>							x
<i>Olen halukas sijoittamaan kaiken aineistoni keskusarkistoon (tms.) saataville ilman rajoituksia.</i>							x
<i>Saattaisin todennäköisemmin aineistojani saataville, jos voisin asettaa saatavuudelle ehtoja.</i>	x						
<i>On tärkeää, että aineistoihini viitataan, jos muut tutkijat käyttävät niitä.</i>		x					
<i>On oikein koota uusia aineistokokonaisuuksia jaetuista aineistoista.</i>					x		
<i>Olen tyytyväinen kykyihini yhdistää aineistoja eri lähteistä, kun etsin vastauksia tutkimuskysymyksiin.</i>					x		
<b>18. Vapaa sana</b>	Ei ennalta määritelty						



## LIITE 6. Tavoitetut vastaajat -lisätaulukot.

### Liite 6 a. Vastaajien sukupuolijakauma.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Sukupuoli <sup>a</sup>	Nainen	3	75	-	-	7	87,5	10	62,5
	Mies	1	25	4	100	1	12,5	6	37,5
	Yhteensä	4		4		8		16	

a. p-arvo 0.022, Fisher-Freeman-Haltonin eksakti testi.

### Liite 6 b. Vastaajien ikäjakauma.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Syntymäaika	1953–1959	2	50	-	-	2	25	4	25
	1960–1969	2	50	-	-	-	-	2	12,5
	1970–1979	-	-	2	50	1	12,5	3	18,8
	1980–1989	-	-	1	25	3	37,5	4	25
	1990–1991	-	-	1	25	1	12,5	2	12,5
	Ei vastattu	-	-	-	-	1	12,5	1	6,3
	Yhteensä	4		4		8		16	

### Liite 6 c. Vastaajien asema työyhteisössä.<sup>a</sup>

		n	%
Nykyinen asema työyhteisössä	Professori	2	12,5
	Apulaisprofessori	1	6,3
	Lehtori	2	12,5
	Tutkijatohtori/postdoc	3	18,8
	Tohtoriopiskelija	5	31,3
	Muu	3	18,8
	Yliopettaja		
	Yhteensä	16	

a. Taulukossa ei ole käytetty korkeakoulukohtaista jaottelua, sillä pienen otoksen vuoksi esimerkiksi aseman ja tutkimusalan yhdistämällä vastaajille luvattu anonymiteetti voisi vaarantua.

### Liite 6 d. Vastaajien esisijainen tutkimusala.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Mikä seuraavista kuvaa parhaiten ensisijaista tutkimusalaasi? <sup>a</sup>	Liikeala ja hallinto	-	-	-	-	1	12,5	1	6,3
	Tekniikka, teknologia ja tietojenkäsittely	1	25	1	25	-	-	2	12,5
	Fysiikka, kemia ja matematiikka	-	-	3	75	-	-	3	18,8
	Informaatiotieteet	-	-	-	-	2	25	2	12,5
	Lääketeide ja sosiaali- ja terveysala	2	50	-	-	1	12,5	3	18,8
	Kasvatustieteet	1	25	-	-	1	12,5	2	12,5

Yhteiskuntatieteet ja humanistiset tieteet	-	-	-	-	2	25	2	12,5
Muu Solu- ja kudosteknologia <sup>b</sup>	-	-	-	-	1	12,5	1	6,3
<b>Yhteensä</b>	<b>4</b>		<b>4</b>		<b>8</b>		<b>16</b>	

a. p-arvo 0.072, Fisher-Freeman-Haltonin eksakti testi.

b. Solu- ja kudosteknologia on osa TaY:n lääketieteen ja biotieteiden tiedekuntaa (2018). Tässä sitä ei liitetty lääketiede ja sosiaali- ja terveystieteiden, vaan kunnioitettiin sitä, että vastaaja ei itse identifioitunut ko. vaihtoehtoon.

### Liite 6 e. Tyypillinen rahoittaja.

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Tyypillinen rahoittaja <sup>a</sup>	Suomen Akatemia	-	-	4	100	5	62,5	9	64,3
	Tekes	-	-	-	-	1	12,5	1	7,1
	TSR	-	-	-	-	-	-	-	-
	Muu	2	100	-	-	2	25	4	28,6
	ESR, STM, Työsuojelurahasto	1	25	-	-	-	-	1	6,3
	okm perusrahoitus [sic], säätiöt	-	-	-	-	1	12,5	1	6,3
	OKM tai rahoittamaton STM, EU	1	25	-	-	-	-	1	6,3
<b>Yhteensä</b>		<b>2<sup>b</sup></b>		<b>4</b>		<b>8</b>		<b>14</b>	

a. p-arvo 0.097, Fisher-Freeman-Haltonin eksakti testi.

b. 2 TAMK:n vastaajaa ohitti kysymyksen.

### Liite 6 f. Valitut aineistotyytit.

		Korkeakoulu							
		TAMK (n = 4)		TTY (n = 4)		TaY (n = 8)		Yhteensä (n = 16)	
		n	%	n	%	n	%	n	%
Kokeellinen aineisto (jonkin verran manipuloitu) <sup>a</sup>	Valittu	1	25	3	75	1	12,5	5	31,3
	Ei valittu	3	75	1	25	7	87,5	11	68,8
Havainnoiva aineisto (ei manipuloitu) <sup>b</sup>	Valittu	2	50	2	50	4	50	8	50
	Ei valittu	2	50	2	50	4	50	8	50
Haastatteluaineisto <sup>c</sup>	Valittu	3	75	-	-	5	62,5	8	50
	Ei valittu	1	25	4	100	3	37,5	8	50
Yhteiskuntatieteellinen aineisto <sup>d</sup>	Valittu	1	25	-	-	5	62,5	6	37,5
	Ei valittu	3	75	4	100	3	37,5	10	62,5
Muu aineisto <sup>e,f</sup>	Valittu	1	25	-	-	2	25	3	18,8
	Kyselyaineisto	1	25	-	-	1	12,5	2	12,5
	Videoitu aineisto	-	-	-	-	1	12,5	1	6,3
	aidoista vuorovaikutustilanteista siitä tehdyt litteraatiot								
	Ei valittu	3	75	4	100	6	75	13	81,3
Yhteensä	Valittu-valinnat	8		5		17		30	
	Ei valittu -valinnat	12		15		23		50	
	Keskimäärin valittu-valintoja	2		1,25		2,13		1,88	

a. p-arvo 0.139, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 1.000, Fisher-Freeman-Haltonin eksakti testi.

c. p-arvo 0.143, Fisher-Freeman-Haltonin eksakti testi.

d. p-arvo 0.182, Fisher-Freeman-Haltonin eksakti testi.

e. p-arvo 0.771, Fisher-Freeman-Haltonin eksakti testi.

f. Vastausten perusteella aineistotyyppinä olisi pitänyt lomakkeella täsmentää. Esim. vaihtoehdon ”muu” valinneista vastaajista kaksi tarkentaa käyttävänsä kyselyaineistoa ja yksi täsmentää hyödyntävänsä videoitua aineistoa aidoista vuorovaikutustilanteista ja niistä tehtyjä litteraatioita. Voi olla, että yhteiskuntatieteellisen aineiston valinneiden joukossakin on kyselyaineistojen käyttäjiä ilman, että he ovat valinneet vaihtoehdon ”muu”. Tulee myös muistaa, että lomakkeen vaihtoehdoissa on kyse aineiston jakamisesta joihinkin päätyyppeihin. Toisaalta yksittäinen vastaaja saattaa tulkita aineistotyyppit eri tavoin, toisaalta yksittäinen aineistotyyppi voi jakautua useisiin spesifimpiin aineistoihin.

### **Liite 6 g. Kunkin aineistotyyppin rinnalla käytetyt aineistotyyppit.**

Aineistotyyppivalinta	Rinnalla käytetyt aineistotyyppit	Korkeakoulu			
		TAMK (n = 4)	TTY (n = 4)	TaY (n = 8)	Yht. (n = 16)
Kokeellinen aineisto		n = 1	n = 3	n = 1	n = 5
	Havainnoiva	1	1	-	2
	Haastattelu	1	-	-	1
	Yhteiskuntatieteellinen	-	-	-	-
	Muu	-	-	-	-
	Lisävalinnat yhteensä	2	1	-	3
	Lisävalinnat keskimäärin käyttäjäryhmän vastaajalla	2	0,33	-	0,6
	Lisäaineistotyyppit yhteensä	2	1	-	2
Havainnoiva aineisto		n = 2	n = 2	n = 4	n = 8
	Kokeellinen	1	1	-	2
	Haastattelu	2	-	4	6
	Yhteiskuntatieteellinen	1	-	3	4
	Muu	-	-	1	1
	Lisävalinnat yhteensä	4	1	8	13
	Lisävalinnat keskimäärin käyttäjäryhmän vastaajalla	2	0,5	2	1,63
	Lisäaineistotyyppit yhteensä	3	1	3	4
Haastatteluaineisto		n = 3	-	n = 5	n = 8
	Kokeellinen	1	-	-	1
	Havainnoiva	2	-	4	6
	Yhteiskuntatieteellinen	1	-	4	5
	Muu	-	-	1	1
	Lisävalinnat yhteensä	4	-	9	13
	Lisävalinnat keskimäärin käyttäjäryhmän vastaajalla	1,3	-	1,8	1,63
	Lisäaineistotyyppit yhteensä	3	-	2	4
Yhteiskuntatieteellinen aineisto		n = 1	-	n = 5	n = 6
	Kokeellinen	-	-	-	-
	Havainnoiva	1	-	3	4
	Haastattelu	1	-	4	5
	Muu	-	-	1	1

Lisävalinnat yhteensä	2	-	8	10
Lisävalinnat keskimäärin käyttäjäryhmän vastaajalla	2	-	1,6	1,67
Lisäaineistotyyppit yhteensä	2	-	3	3
Muu aineisto	n = 1	-	n = 2	n = 3
Kokeellinen	-	-	-	-
Havainnoiva	-	-	1	1
Haastattelu	-	-	1	1
Yhteiskuntatieteellinen	-	-	1	1
Lisävalinnat yhteensä	-	-	3	3
Lisävalinnat keskimäärin käyttäjäryhmän vastaajalla	-	-	1,5	1
Lisäaineistotyyppit yhteensä	-	-	3	3

**Liite 6 h. Aineiston jakamismääränsä ilmoittaneiden samanmielisyyss väittämän ”muut tarvitsevat luvan käyttääkseen aineistojani” kanssa.**

		Kuinka paljon saatat tutkimusaineistoasi muiden saataville?							
		En lainkaan		Jonkin verran		Suurimman osan		Yhteensä	
		n	%	n	%	n	%	n	%
Muut tarvitsevat luvan käyttääkseen aineistojani. <sup>a</sup>	Vahvasti samaa mieltä	2	50	4	44,4	-	-	6	37,5
	Osittain samaa mieltä	-	-	3	33,3	2	66,7	5	31,3
	En samaa enkä eri mieltä	1	25	-	-	-	-	1	6,3
	Osittain eri mieltä	-	-	2	22,2	1	33,3	3	18,8
	Vahvasti eri mieltä	1	25	-	-	-	-	1	6,3
	En osaa sanoa	-	-	-	-	-	-	-	-
	Yhteensä	4		9		3		16	

a. p-arvo 0.183, Fisher-Freeman-Haltonin eksakti testi.

## LIITE 7. Aineiston kerääjä -lisätaulukko.

### *Aineistoon ja aineiston kerääjään viittaamisen saaman kannatuksen vertaileminen.*

		Pidätkö ehdon "Aineiston kerääjään viitataan kaikissa töissä, joissa aineistoa on käytetty" täyttymistä keskeisenä ennen kuin muut voivat käyttää tutkimusaineistojasi?							
		Kyllä		En		En ole varma		Yhteensä	
		n	%	n	%	n	%	n	%
On tärkeää, että aineistoihini viitataan, jos muut tutkijat käyttävät niitä. <sup>a</sup>	Vahvasti samaa mieltä	10	90,9	-	-	2	66,7	12	75
	Osittain samaa mieltä	1	9,1	-	-	1	33,3	2	12,5
	En samaa enkä eri mieltä	-	-	1	50	-	-	1	6,3
	Osittain eri mieltä	-	-	-	-	-	-	-	-
	Vahvasti eri mieltä	-	-	1	50	-	-	1	6,3
	En osaa sanoa	-	-	-	-	-	-	-	-
Yhteensä		11		2		3		16	

a. p-arvo 0.033, Fisher-Freeman-Haltonin eksakti testi.

## LIITE 8. Uudelleenkäytön valvoja -lisätaulukko.

**Uudelleenkäytön passiiviseen valvontaan liittyviä ehtoja koskevien valintojen vertailu  
aktiiviseen valvontaan liittyviä ehtoja koskeviin valintoihin.**

		Aineiston kerääjälle pitää antaa kopia artikkeleista, joissa hänen aineistoaan on käytetty						Aineiston kerääjä saa listan kaikista tuotoksista, joissa hänen aineistojaan on käytetty.									
		Kyllä		En		Epävarma		Yht.		Kyllä		En		Epävarma		Yht.	
		n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Aineiston kerääjällä on mahdollisuus tehdä yhteistyötä projektin kanssa, jossa aineistoa käytetään	Kyllä	4	80	3	27,3	-	-	7	43,8	5	71,4	2	25	-	-	7	43,8
	En	1	20	5	45,5	-	-	6	37,5	1	14,3	4	50	1	100	6	37,5
	Epävarma	-	-	3	27,3	-	-	3	18,8	1	14,3	2	25	-	-	3	18,8
	Yht.	5		11		-	-	16		7		8		1		16	
Tuloksia, jotka perustuvat aineistoon, ei saa levittää missään muodossa ilman aineiston kerääjän hyväksyntää	Kyllä	2	40	1	9,1	-	-	3	18,8	1	14,3	2	25	-	-	3	18,8
	En	2	40	9	81,8	-	-	11	68,8	4	57,1	6	75	1	100	11	68,8
	Epävarma	1	20	1	9,1	-	-	2	12,5	2	28,6	-	-	-	-	2	12,5
	Yht.	5		11		-	-	16		7		8		1		16	
Tuloksia, jotka pohjautuvat aineiston hyödyntämiseen, ei voi levittää ilman, että aineiston kerääjällä olisi mahdollisuutta arvioida tuloksia ja tehdä ehdotuksia ja kommentteja	Kyllä	1	20	2	18,2	-	-	3	18,8	3	42,9	-	-	-	-	3	18,8
	En	2	40	9	81,8	-	-	11	68,8	3	42,9	7	87,5	1	100	11	68,8
	Epävarma	2	40	-	-	-	-	2	12,5	1	14,3	1	12,5	-	-	2	12,5
	Yht.	5		11		-	-	16		7		8		1		16	
Aineiston kerääjä saa ja hyväksyy lausunnon koskien aineiston uudelleenkäytön tavoitteita	Kyllä	4	80	2	18,2	-	-	6	37,5	5	71,4	1	12,5	-	-	6	37,5
	En	1	20	9	81,8	-	-	10	62,5	2	28,6	7	87,5	1	100	10	62,5
	Epävarma	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Yht.	5		11		-	-	16		7		8		1		16	

## LIITE 9. Aineistohallinnan tuen saaja -lisätaulukot.

### Liite 9 a. Tyypillinen rahoittaja ja aineistohallintasuunnitelmavaatimus.

		Mikä seuraavista kuvaa parhaiten tutkimuksesi tyypillistä rahoittajaa?							
		Suomen Akatemia		Tekes		Muu		Yhteensä	
		n	%	n	%	n	%	n	%
Vaatiiko ensisijainen rahoittajasi sinua laatimaan aineiston-hallintasuunnitelman? <sup>a</sup>	Kyllä	6	66,7	-	-	2	66,7	8	61,5
	Ei	2	22,2	-	-	1	33,3	3	23,1
	En tiedä	1	11,1	1	100	-	-	2	15,4
	Yhteensä	9		1		3		13	

a. p-arvo 0.491, Fisher-Freeman-Haltonin eksakti testi.

### Liite 9 b. Organisaationsa tai yksikkönsä palvelinta käyttämättömien muut valinnat.

Ei lainkaan organisaationsa tai yksikkönsä palvelinta käyttävät (n = 4)

Muut valintoja saaneet säilytyspaikkavaihtoehdot		n
Organisaationi omassa arkistossa	Jonkin verran	1
	Suurimman osan	-
	Kaiken	-
	Yhteensä	1
Muussa tietoarkistossa tai arkistossa	Jonkin verran	-
	Suurimman osan	-
	Kaiken	-
	Yhteensä	-
Omalla henkilökohtaisella tietokoneellani	Jonkin verran	1
	Suurimman osan	-
	Kaiken	1
	Yhteensä	2
Muistitikulla tai ulkoisella kovalevyllä	Jonkin verran	-
	Suurimman osan	1
	Kaiken	1
	Yhteensä	2
Paperimuodossa omassa toimistossani	Jonkin verran	1
	Suurimman osan	1
	Kaiken	1
	Yhteensä	3
Muu	Jonkin verran	-
	Suurimman osan	2
	Kaiken	-
	Yhteensä	2

**Liite 9 c. Kaiken aineistonsa tietyssä paikassa säilyttävien muut säilytysmäärävalinnat.<sup>a</sup>**

		Kaikki aineisto organisaation tai yksikön palvelimella (n = 6)		Kaikki aineisto omalla tietokoneella (n = 3)		Kaikki aineisto muistitikulla tai ulkoisella kovalevyllä (n = 1)		Kaikki aineisto paperi- muodossa (n = 1)	
<b>Muut säilytyspaikka- vaihtoehdot</b>		n	%	n	%	n	%	n	%
Organisaationi tai yksikköni palvelin	Jonkin verran			1	50	-	-	-	-
	Suurin osa			1	50	-	-	-	-
	Kaikki			-	-	-	-	-	-
	Yhteensä			2		-	-	-	-
Tutkimushankkeen palvelin	Jonkin verran	2	66,7	-	-	-	-	-	-
	Suurin osa	1	33,3	-	-	-	-	-	-
	Kaikki	-	-	-	-	-	-	-	-
	Yhteensä	3		-	-	-	-	-	-
Tieteenalakohtainen säilytyspaikka/ arkisto	Jonkin verran	2	100	-	-	-	-	-	-
	Suurin osa	-	-	-	-	-	-	-	-
	Kaikki	-	-	-	-	-	-	-	-
	Yhteensä	2		-	-	-	-	-	-
Julkaisijan säilytys / julkaisijan arkisto	Jonkin verran	-	-	-	-	-	-	-	-
	Suurin osa	-	-	-	-	-	-	-	-
	Kaikki	-	-	-	-	-	-	-	-
	Yhteensä	-	-	-	-	-	-	-	-
Organisaationi oma arkisto	Jonkin verran	-	-	-	-	-	-	-	-
	Suurin osa	-	-	-	-	-	-	-	-
	Kaikki	-	-	-	-	-	-	-	-
	Yhteensä	-	-	-	-	-	-	-	-
Muu tietoarkisto tai arkisto	Jonkin verran	-	-	-	-	-	-	-	-
	Suurin osa	-	-	-	-	-	-	-	-
	Kaikki	-	-	-	-	-	-	-	-
	Yhteensä	-	-	-	-	-	-	-	-
Oma henkilökohtainen tietokone	Jonkin verran	-	-			-	-	-	-
	Suurin osa	1	100			-	-	-	-
	Kaikki	-	-			-	-	1	100
	Yhteensä	1				-	-	1	
Muistitikku tai ulkoinen kovalevy	Jonkin verran	4	100	1	50			-	-
	Suurin osa	-	-	1	50			-	-
	Kaikki	-	-	-	-			-	-
	Yhteensä	4		2				-	-
Paperimuodossa omassa toimistossani	Jonkin verran	4	100	1	50	-	-		
	Suurin osa	-	-	-	-	1	100		
	Kaikki	-	-	1	50	-	-		
	Yhteensä	4		2		1			
Muu	Jonkin verran	-	-	-	-	-	-	-	-
	Suurin osa	1	100	1	100	1	100	1	100
	Kaikki	-	-	-	-	-	-	-	-
	Yhteensä	1		1		1		1	

<sup>a</sup> SPSS:n Recode into Different Variables -toiminnolla selittäväksi muuttujaksi kaiken aineistonsa organisaation/yksikön palvelimella, omalla tietokoneellaan, muistitikulla tai ulkoisella kovalevyllä ja/tai paperimuodossa säilyttävät.



**Liite 9 d. Säilytysmäärävalintakertojen jakautuminen säilytyspaikoittain.**

Esitystapa: Prosenttiosuudet ovat osuudet kunkin säilytysmäärän kokonaisvalintakertamäärästä, joka on esitetty taulukossa ylimmäisenä. (TAMK: ei lainkaan 21, jonkin verran 4, suurin osa 6, kaikki 2, positiiviset säilytysmäärävalinnat yhteensä 12; TTY: ei lainkaan 25, jonkin verran 9, suurin osa 4, kaikki 1, positiiviset säilytysmäärävalinnat yhteensä 13; TaY: ei lainkaan 49, jonkin verran 14, suurin osa 6, kaikki 8, positiiviset säilytysmäärävalinnat yhteensä 28; YHTEENSÄ: ei lainkaan 95, jonkin verran 27, suurin osa 16, kaikki 11, positiiviset säilytysmäärävalinnat yhteensä 54).

	Säilytysmäärät	Korkeakoulu							
		TAMK		TTY		TaY		Yht.	
		n	%	n	%	n	%	n	%
<b>Kaikki säilytyspaikkavalinnat</b>	<b>Yht. valinnat<sup>e</sup></b>	<b>33<sup>a</sup></b>		<b>39<sup>b</sup></b>		<b>77<sup>c</sup></b>		<b>149<sup>d</sup></b>	
	<b>Ei lainkaan</b>	<b>21</b>	<b>64</b>	<b>25</b>	<b>64</b>	<b>49</b>	<b>64</b>	<b>95</b>	<b>64</b>
	<b>Jonkin verran</b>	<b>4</b>	<b>12</b>	<b>9</b>	<b>23</b>	<b>14</b>	<b>18</b>	<b>27</b>	<b>18</b>
	<b>Suurin osa</b>	<b>6</b>	<b>18</b>	<b>4</b>	<b>10</b>	<b>6</b>	<b>8</b>	<b>16</b>	<b>11</b>
	<b>Kaikki</b>	<b>2</b>	<b>6</b>	<b>1</b>	<b>3</b>	<b>8</b>	<b>10</b>	<b>11</b>	<b>7</b>
	<b>Yht. valittu säilytyspaikaksi</b>	<b>12</b>	<b>36</b>	<b>14</b>	<b>36</b>	<b>28</b>	<b>36</b>	<b>54</b>	<b>36</b>
<b>Yhteenveto tiedeinstituutiotason säilytyspaikkavalinnoista (6 säilytyspaikkaa)</b>	Yht. valinnat	20		24		48		92	
	Ei lainkaan	15	71	18	72	37	76	70	74
	Jonkin verran	1	25	4	44	5	36	10	37
	Suurin osa	3	50	2	50	1	17	6	38
	Kaikki	1	50	0	0	5	62,5	6	55
	Yht. valittu säilytyspaikaksi	5	42	6	43	11	39	22	41
Organisaationi tai yksikköni palvelin <sup>a</sup>	Ei lainkaan	1	5	1	4	2	4	4	4
	Jonkin verran	-	-	2	22	1	7	3	11
	Suurin osa	2	33	1	25	-	-	3	19
	Kaikki	1	50	-	-	5	62,5	6	55
	Yht. valittu säilytyspaikaksi	3	25	3	21	6	21	12	22
Tutkimushankkeen palvelin <sup>b</sup>	Ei lainkaan	3	14	3	12	5	10	11	12
	Jonkin verran	-	-	1	11	2	14	3	11
	Suurin osa	-	-	-	-	1	17	1	6
	Kaikki	-	-	-	-	-	-	-	-
	Yht. valittu säilytyspaikaksi	-	-	1	7	3	11	4	7
Tieteenalakohtainen säilytyspaikka/arkisto <sup>c</sup>	Ei lainkaan	3	14	4	16	6	12	13	14
	Jonkin verran	-	-	-	-	2	14	2	7
	Suurin osa	1	17	-	-	-	-	1	6
	Kaikki	-	-	-	-	-	-	-	-
	Yht. valittu säilytyspaikaksi	1	8	-	-	2	7	3	6
Julkaisijan säilytys / julkaisijan arkisto <sup>d</sup>	Ei lainkaan	3	14	3	12	8	16	14	15
	Jonkin verran	-	-	1	11	-	-	1	4
	Suurin osa	-	-	-	-	-	-	-	-
	Kaikki	-	-	-	-	-	-	-	-
	Yht. valittu säilytyspaikaksi	-	-	1	7	-	-	1	2
Organisaationi oma arkisto <sup>e</sup>	Ei lainkaan	2	10	4	16	8	16	14	15
	Jonkin verran	1	25	-	-	-	-	1	4
	Suurin osa	-	-	-	-	-	-	-	-

	Kaikki	-	-	-	-	-	-	-	-
	Yht. valittu säilytyspaikaksi	1	8	-	-	-	-	1	2
Muu tietoaarkisto tai arkisto <sup>f</sup>	Ei lainkaan	3	14	3	12	8	16	14	15
	Jonkin verran	-	-	-	-	-	-	-	-
	Suurin osa	-	-	1	25	-	-	1	6
	Kaikki	-	-	-	-	-	-	-	-
	Yht. valittu säilytyspaikaksi	-	-	1	7	-	-	1	2
<b>Yhteenveto yksilökeskeisistä säilytyspaikkavalinnoista (3 säilytyspaikkaa)</b>	Yht. valinnat	11		12		24		47	
	Ei lainkaan	4	19	4	16	10	20	18	19
	Jonkin verran	3	75	5	56	9	64	17	63
	Suurin osa	3	50	2	50	2	33,3	7	44
	Kaikki	1	50	1	100	3	37,5	5	45
	Yht. valittu säilytyspaikaksi	7	58	8	57	14	50	29	54
Oma henkilökohtainen tietokone <sup>g</sup>	Ei lainkaan	1	5	1	4	6	12	8	8
	Jonkin verran	1	25	1	11	-	-	2	7
	Suurin osa	1	17	1	25	1	17	3	19
	Kaikki	1	50	1	100	1	12,5	3	27
	Yht. valittu säilytyspaikaksi	3	25	3	21	2	7	8	15
Muistitikku tai ulkoinen kovalevy <sup>h</sup>	Ei lainkaan	1	5	2	8	3	6	6	6
	Jonkin verran	1	25	1	11	4	29	6	22
	Suurin osa	1	17	1	25	-	-	2	13
	Kaikki	-	-	-	-	1	12,5	1	9
	Yht. valittu säilytyspaikaksi	2	17	2	14	5	18	9	17
Paperimuoto omassa toimistossa <sup>i</sup>	Ei lainkaan	2	10	1	4	1	2	4	4
	Jonkin verran	1	25	3	33	5	36	9	33
	Suurin osa	1	17	-	-	1	17	2	13
	Kaikki	-	-	-	-	1	12,5	1	9
	Yht. valittu säilytyspaikaksi	2	17	3	21	7	25	12	22
<b>Yhteenveto määrittelemättömistä valinnoista / säilytyspaikka "muu"</b>	Yht. valinnat	2		3		5		9	
	Ei lainkaan	2	10	3	12	2	4	7	7
	Jonkin verran	-	-	-	-	-	-	-	-
	Suurin osa	-	-	-	-	3	50	3	19
	Kaikki	-	-	-	-	-	-	-	-
	Yht. valittu säilytyspaikaksi	-	-	-	-	3	11	3	6

<sup>a</sup> Keskimäärin 3,3 vastaajaa.

<sup>b</sup> Keskimäärin 3,9 vastaajaa.

<sup>c</sup> Keskimäärin 7,7 vastaajaa.

<sup>d</sup> Keskimäärin 14,9 vastaajaa.

<sup>e</sup> TAMK:n vastaajat ohittivat yhteensä seitsemän ja TaY:n vastaajat kolme valintakertaa ja TTY:n vastaajat yhden valintakerran.

**Liite 9 e. Säilytyspaikkavaihtoehtojen hylkääminen ja hyödyntäminen säilytysvastuualueittain.**

		Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
Tiedeinstituutiotason säilytysvastuualueen valintakerrat <sup>a</sup>	Valittu	5	25	6	25	11	22,9	22	23,9
	Ei valittu	15	75	18	75	37	77,1	70	76,1
	Yhteensä	20		24		48		92	
Yksilökeskeisen säilytysvastuualueen valintakerrat <sup>b</sup>	Valittu	7	63,6	8	66,7	14	58,3	29	61,7
	Ei valittu	4	36,4	4	33,3	10	41,7	18	38,3
	Yhteensä	11		12		24		47	
Määrittelemättömän säilytysvastuualueen valintakerrat <sup>c</sup>	Valittu	-	-	-	-	3	60	3	30
	Ei valittu	2	100	3	100	2	40	7	70
	Yhteensä	2		3		5		10	
Yhteensä	Valittu	12	36	14	39	28	36	54	36
	Ei valittu	28	64	25	64	49	64	95	64
	Yhteensä	33 <sup>d</sup>		39 <sup>e</sup>		77 <sup>f</sup>		149 <sup>g</sup>	

a. p-arvo 1.000, Fisher-Freeman-Haltonin eksakti testi.

b. p-arvo 0.925, Fisher-Freeman-Haltonin eksakti testi.

c. p-arvo 0.208, Fisher-Freeman-Haltonin eksakti testi.

d. 7 valintakertaa jätetty käyttämättä.

e. 1 valintakerta jätetty käyttämättä.

f. 3 valintakertaa jätetty käyttämättä.

g. 11 valintakertaa jätetty käyttämättä.

## LIITE 10. Yhteenveto aineiston jakamatta jättämisen syitä koskevien kannanottojen jakautumisesta.

Liite 10 a. Yhteenveto aineiston jakamatta jättämistä selittämään valituista syistä (kysymys 12).

Rooli	Syy	Korkeakoulu							
		TAMK		TTY		TaY		Yhteensä	
		n	%	n	%	n	%	n	%
<b>Aineiston kerääjä (16,7 % syistä)</b>		1	14,3	3	25	5	19,2	9	20
	Minun tarvitsee julkaista ensin omat tutkimustulokseni	1	14	3	25	5	19,2	9	20
	Menettäisin mahdollisuuden hallita aineistojani itse	-	-	-	-	-	-	-	-
<b>Aineiston suojelija (16,7 % syistä)</b>		3	42,9	-	-	10	38,5	13	28,9
	Minulla ei ole oikeuksia julkaista aineistojani	1	14	-	-	7	26,9	8	18
	Aineistojeni ei kuulu olla julkisia	2	29	-	-	3	11,5	5	11
<b>Tiedeyhteisön jäsen (8,3 % syistä)</b>		-	-	3	25	-	-	3	7
	Muut eivät tarvitse tutkimusaineistojani	-	-	3	25	-	-	3	7
<b>Aineiston käsitte- lijä (8,3 % syistä)</b>		-	-	-	-	2	7,7		4
	Standardit puuttuvat	-	-	-	-	2	7,7	2	4
<b>Tuen saaja (41,7 % syistä)</b>		3	42,9	5	41,7	7	26,9	15	33,3
	Minulla on puutteelliset taidot	2	29	1	8,3	2	7,7	5	11
	Aineistoilleni ei ole sopivaa säilytyspaikkaa	-	-	-	-	-	-	-	-
	Minulla ei ole aikaa saattaa aineistojani saataville	-	-	2	16,7	3	11,5	5	11
	Rahoitus puuttuu.	-	-	1	8,3	2	7,7	3	7
	Rahoittaja ei edellytä aineistojeni jakamista	1	14	1	8,3	-	-	2	4
<b>Määrittelemätön (8,3 % syistä)</b>		-	-	1	8,3	2	7,7	3	7
	Muu	-	-	1	8,3	2	7,7	3	7
<b>YHTEENSÄ</b>		7		12		26		45	

**Liite 10 b. Vaihtoehdon ”muu” kysymykseen 12 valinneiden näkemykset muista aineiston jaksamatta jättämisen syistä.**

		Muu		
		”Aineiston anonymisointi on osin mahdotonta”	”Anonymisointi pienissä havainnoi- tiaineistoissa käytännössä tuhoaa aineiston”	”Ei voisi vähem- pää kiinnostaa, en minäkään kenenkään raakadataa halua sörkkiä”
		n = 1	n = 1	n = 1
Minun tarvitsee julkaista ensin omat tutkimustulokseni	Valittu syyksi Ei valittu syyksi	- 1	- 1	1 -
Menettäisin mahdollisuuden hallita aineistojani itse	Valittu syyksi Ei valittu syyksi	- 1	- 1	- 1
Aineistojeni ei kuulu olla julkisia	Valittu syyksi Ei valittu syyksi	- 1	- 1	- 1
Minulla ei ole oikeuksia julkaista aineistojani	Valittu syyksi Ei valittu syyksi	1 -	1 -	- 1
Muut eivät tarvitse aineistojani	Valittu syyksi Ei valittu syyksi	- 1	- 1	1 -
Standardit puuttuvat	Valittu syyksi Ei valittu syyksi	- 1	- 1	- 1
Rahoitus puuttuu.	Valittu syyksi Ei valittu syyksi	1 -	- 1	- 1
Minulla ei ole aikaa saattaa aineistojani saataville	Valittu syyksi Ei valittu syyksi	1 -	- 1	1 -
Aineistoilleni ei ole sopivaa säilytyspaikkaa	Valittu syyksi Ei valittu syyksi	- 1	- 1	- 1
Rahoittaja ei edellytä aineistojeni jakamista	Valittu syyksi Ei valittu syyksi	- 1	- 1	- 1
Minulla on puutteelliset taidot	Valittu syyksi Ei valittu syyksi	- 1	- 1	1 -

**Liite 10 c. Aineiston jaksamatta jättämisen normiperusteisten syiden välinen suhde.**

		Aineistojeni ei kuulu olla julkisia					
		Valittu syyksi		Ei valittu syyksi		Yhteensä	
		n	%	n	%	n	%
Minulla ei ole oikeuksia julkaista aineistojani <sup>a</sup>	Valittu syyksi	4	80	4	36,4	8	50
	Ei valittu syyksi	1	20	7	63,6	8	50
	Yhteensä	5		11		16	

a. p-arvo 0.282, Fisher-Freeman-Haltonin eksakti testi.

## LIITE 11. Yhteenveto uudelleenkäytön ehtojen asettamista puoltavien kantojen jakautumisesta.

Jakamisen ehdot yksittäin ja rooleittain	Korkeakoulu							
	TAMK		TTY		TaY		Yhteensä	
	n	%	n	%	n	%	n	%
<b>Aineiston kerääjä (25 % ehdoista)</b>	10	38,4	6	86	14	33,3	30	40,1
Jaettu tekijyys (co-authorship) koskien julkaisuja, joissa aineistoa on käytetty	4	15,4	-	-	1	2,4	5	6,7
Aineiston kerääjälle annetaan tunnustus kaikissa töissä, joissa aineistoa on käytetty	3	11,5	3	43	8	19	14	18,7
Aineiston kerääjään viitataan kaikissa töissä, joissa aineistoa on käytetty	3	11,5	3	43	5	11,9	11	14,7
<b>Tutkija aineiston suojelijana (8 % ehdoista)</b>	3	11,5	-	-	3	7	6	8
Aineiston käytölle on hankittava oikeudelliset luvat	3	11,5	-	-	3	7	6	8
<b>Tutkija tiedeyhteisön jäsenenä (8 % ehdoista)</b>	-	-	1	14	4	9,5	5	6,7
Yhteinen ymmärrys vastavuoroisesta aineiston jakamisesta	-	-	1	14	4	9,5	5	6,7
<b>Tutkija uudelleenkäytön valvojana (50 % ehdoista)</b>	10	38,4	-	-	21	49,7	31	41,3
Aineiston kerääjällä on mahdollisuus tehdä yhteistyötä projektin kanssa, jossa aineistoa käytetään	3	11,5	-	-	4	9,5	7	9,3
Tuloksia, jotka perustuvat aineistoon, ei saa levittää missään muodossa ilman aineiston kerääjän hyväksyntää	1	3,8	-	-	2	4,8	3	4
Tuloksia, jotka pohjautuvat aineiston hyödyntämiseen, ei voi levittää ilman, että aineiston kerääjällä olisi mahdollisuutta arvioida tuloksia ja tehdä ehdotuksia ja kommentteja	-	-	-	-	3	7	3	4
Aineiston kerääjälle pitää antaa kopio artikkeleista, joissa hänen aineistoaan on käytetty	2	7,7	-	-	3	7	5	6,7
Aineiston kerääjä saa listan kaikista tuotoksista, joissa hänen aineistojaan on käytetty, mukaan lukien artikkelit, esitykset, opetusmateriaalit jne.	2	7,7	-	-	5	11,9	7	9,3
Aineiston kerääjä saa ja hyväksyy lausunnon koskien aineiston uudelleenkäytön tavoitteita	2	7,7	-	-	4	9,5	6	8
<b>Tutkija tuen saajana (8 % ehdoista)</b>	3	11,5	-	-	-	-	3	4
Ainakin osa aineiston keräämis- ja säilytyskustannuksista on katettava	3	11,5	-	-	-	-	3	4
<b>Yhteensä</b>	26		7		42		75	
<b>Keskimäärin ehtoja vastaajaa kohden<sup>a</sup></b>	6,5		1,75		5,25		4,69	

a. TAMK:ssa 4 vastaajaa, TTY:ssä 4 vastaajaa, TaY:ssa 8 vastaajaa, yhteensä 16.