

# TUTKIMUSDATAN HALLINNAN TILANNEKARTOITUKSEN RAPORTTI

---

## Lisenssi

Creative Commons Suomi CC-BY-NC-SA 3.0

Raportin käyttäjälle annetaan oikeus jakaa eli kopioida, levittää, näyttää ja esittää teosta sekä valmistaa muutettuja teoksia seuraavilla ehdoilla:

- 

Opetus- ja kulttuuriministeriö nimetään teoksen tekijäksi (ei kuitenkaan siten, että ilmoitus viittaisi lisenssinantajan tukevan lisenssinsaaajaa tai teoksen käyttötapaa).

- 

Käyttäjä ei saa käyttää teosta kaupallisesti.

- 

Mikäli käyttäjä tekee muutoksia tai käyttää teosta omien teostensa pohjana, tulee johdannaisteos jakaa samalla tai saman kaltaisella lisenssillä



1. Johdanto.....	3
2. Tilannekartoituksen toteutus .....	4
3. Tilannekartoituksen yleistettävyys .....	4
4. Yleisiä huomioita työpajoista .....	5
5. Yliopistojen tutkimusdatan hallinnan tila.....	5
6. Tutkimushankkeiden tutkimusdatan hallinnan tila .....	8
7. Johtopäätöksiä ja kehitysehdotuksia datan hallintaan yliopistojen näkökulmasta.....	10
8. Tilannekartoituksen johtopäätökset OKM:n ja TTA:n kannalta .....	12
9. Tilannekartoituksen johtopäätökset tutkijoiden kannalta .....	12
9.1. Datat julkaiseminen ja jakaminen .....	14
9.1.1. Datat jakamisen haasteita .....	14
9.1.2. Omistajuus- ja tekijänoikeudet .....	15
9.1.3. Datat säilytys .....	15
10. Tutkimusdatan hallinnan haasteet ja mahdollisuudet .....	16
11. Huomioita PAS-työhön.....	17
12. Huomioita metatietotyöhön .....	18
13. Huomioita tietoinfrastruktuurityöhön.....	18
14. Loppusanat.....	20
Lähteet: .....	21

Figure 1 <i>Yliopistojen tutkimusdatan hallinnan tasot.</i> .....	6
Figure 2 <i>Yliopistojen tutkimusdatan hallinnan tila</i> .....	6
Figure 3 <i>Tutkimushankkeiden tutkimusdatan hallinnan tasot</i> .....	9
Figure 4 <i>Tutkimushankkeiden tutkimusdatan hallinnan tila (työpajojen esitykset, N=20)</i> .....	9
Figure 5 <i>Tutkimushankkeiden tutkimusdatan hallinnan profiilit, N=33</i> ....	10
Figure 7 <i>Aineistojen koon suhde aineistojen määrään</i> .....	13
Figure 6 <i>Valveutuneisuuden suhde käytäntöjen vaikuttavuuteen</i> .....	13
Figure 8 <i>Tutkimusdatan hallinnan olemassa olevat palvelut, mahdollisuudet, puutteet ja haasteet</i> .....	17
Figure 9 <i>Tutkimushankkeiden aktiivisuus tutkimusdatan hallinnassa</i> .....	19
Figure 10 <i>Yliopistojen aktiivisuus tutkimusdatan hallinnassa</i> .....	19

# 1. Johdanto

Tutkimusdatan hallinnan tilannekartoitus toteutettiin keväällä 2013 Opetus- ja kulttuuriministeriön (OKM) toimeksiannosta. Tilannekartoituksella haluttiin kerätä ja tuottaa [Tutkimuksen tietoaineistot -hankkeen](#) (TTA) kehitystyössään tarvitsemaa tietoa tämänhetkisistä tutkimusdatan hallinnan käytännöistä, ongelmista sekä tulevaisuuden tarpeista ja suunnitelmista. Selvitystyössä kiinnitettiin huomiota erityisesti yliopistojen ja tutkimushankkeiden tilanteeseen.

Tilannekartoituksen tuloksia hyödyntävät mm. Tiedon saatavuus -työryhmä, TTA-PAS -kustannusanalyysi, Kata-aineistokatalogin kehitys, tietoinfrastruktuuri-työryhmä, metatietotyöryhmä sekä Opetus- ja kulttuuriministeriö.

Tilannekartoituksen tavoitteena on ollut selvittää seuraavia asioita:

1. Tutkimusdatan hallintaan liittyvät haasteet sekä toimivat ratkaisut ja käytännöt eri toimijoiden näkökulmasta (tutkimus, tutkimushallinto, tietohallinto, kirjastot, arkistot)
2. Olemassa olevat tutkimusdatan hallinnan prosessit, palvelut sekä tutkimusinfrastruktuuripalvelut
3. Puuttuvat palvelut ja tarpeet
4. Selventää tutkimusaineistojen määrää säilytyksen eri vaiheissa

## 2. Tilannekartoituksen toteutus

Tilannekartoitusta toteutettiin pääasiassa työpajojen avulla. Työpajoja järjestettiin kevään aikana yhteensä seitsemän. Työpajasarjan lanseeraus toteutettiin Helsingissä Tieteiden talolla. Yliopistoilla toteutettavat työpajat järjestettiin Oulun yliopistolla, Tampereella (Tampereen yliopisto sekä Tampereen teknillinen yliopisto), Turussa (Turun yliopisto sekä Åbo Akademi), Helsingin yliopistolla, Jyväskylän yliopistolla sekä Espoossa Aalto-yliopistolla.

Työpajat koostuivat kolmesta tai neljästä tieteenalakohtaisesta tutkijan esityksestä, joissa luotiin katselmus kyseisen tutkimusryhmän datan hallinnan nykytilaan, aineistojen luonteeseen ja määrään sekä tulevaisuuden tallennus- ja pitkäaikaissäilytystarpeisiin. Esityksissä käytiin myös läpi datan hallinnan ongelmakohtia sekä ajankohtaisia tarpeita ja toiveita. Työpajoihin kutsutut tutkijat pyrittiin valitsemaan siten, että esitykset olisivat mahdollisimman laajasti eri tutkimusaloilta. Lisäksi haluttiin mahdollisuuksien mukaan kutsua työpajoihin [Tutkimusinfrastruktuurien tiekartan](#) 2. hakuvaiheeseen päässeitä tutkimushankkeita (5). Myös yliopistojen suosituksia otettiin huomioon tutkimushankkeita valittaessa.

Tieteenalakohtaisten tutkimusryhmien esitysten jälkeen työpajoissa kuultiin yliopiston tietohallinnon esitys kyseisen yliopiston datan hallinnan nykykäytännöistä ja tulevaisuuden suunnitelmista.

Työpajojen lopuksi järjestettiin paneelikeskustelu, johon osallistui kulloinkin kyseisen yliopiston tutkimus- sekä tietohallinnon edustaja, kirjaston sekä mahdollisesti arkiston edustaja sekä esityksen pitäneet tutkijat. Paneelikeskusteluissa käytiin läpi tutkimusdatan hallintaan liittyviä aihealueita, kuten haasteita tutkimusdatan hallinnassa, yliopiston datapolitiikkaa, datan julkaisua ja jakamista, datan kuvailua sekä omistajuus- ja tekijänoikeuskysymyksiä.

Työpajojen lisäksi tietoa kerättiin kahden sähköisen kyselyn avulla. Ensimmäinen kysely toteutettiin Helsingissä pidetyn aloitustyöpajan taustoituskyselynä ja toinen kysely oli tarkoitettu kaikille työpajoihin osallistuville täytettäväksi ilmoittautumisen yhteydessä. Kyselyissä kartoitettiin tarkemmin mm. datanhallinnan käytäntöjä sekä datan määrää ja laatua, pitkäaikaissäilytystarpeita sekä tutkimusryhmän ongelmia ja tarpeita datanhallinnan suhteen.

Tilannekartoituksen työryhmään kuuluivat mm. Saara Värttö (CSC), Päivi Rauste (CSC), Mari Elisa Kuusniemi (Helsingin yliopiston kirjasto) sekä Jari Suhonen (CSC).

## 3. Tilannekartoituksen yleistettävyyttä

Tästä tilanneraportista ei voida johtaa suoraan yleistettäviä johtopäätöksiä. Tämä johtuu mm. siitä, että raportti perustuu osallistuneiden antamiin tietoihin työpajoissa ja kyselyissä sekä sähköpostikyselyyn niiden yliopistojen osalta, joihin ei ollut mahdollista järjestää työpajaa aikataulu- ja resurssisyistä. Kartoituksen otanta ei ole tarpeeksi laaja (N=231), jotta tuloksia voitaisiin täysin yleistää. Tutkijoita oli haasteellista tavoittaa työpajojen kautta ja dataintensiiviset tutkimusryhmät ovat yliedustettuina kartoituksessa. Tutkimusryhmiä, jotka eivät ole heränneet datanhallinnan tarpeisiin, oli vaikea tavoittaa. Lisäksi vastauksiin ja

niiden tulkintaan on vaikuttanut se, että tutkimusdatan hallinnan palvelut, aineistonhallinnan periaatteet ja terminologia eivät ole kaikille vastaajille ja työpajaan osallistujille tuttuja.

## 4. Yleisiä huomioita työpajoista


Tilannekartoitukseen osallistuneilla oli positiivinen suhtautuminen työpajoihin. He pitivät hyvänä asiana sitä, että Tutkimuksen tietoaaineistot -hankkeessa ollaan kiinnostuneita yliopistojen asioista ja tilanteesta. Työpajat herättivät paljon keskustelua. Yliopistot arvostivat työpajoja, koska niissä on jopa ensimmäistä kertaa tuotu saman pöydän ääreen hallinnon edustajia sekä tutkijoita. Työpajat ovat tuoneet hallinnolle tutkijoiden tarpeita esiin ja joissakin tapauksissa hallinnon ihmiset ovat heränneet siihen, että heillä voisi olla merkittävä rooli tutkimusdatan hallinnan tuessa.

Työpajan sisältöä pidettiin toisaalta vaikeana. Jotkut tutkijat pitivät tutkimusdatan hallintaa niin haastavana, että oli liian korkea kynnyks osallistua työpajoihin. Datan määrä tutkimushankkeessa voi olla myös niin pieni, että ei ole herätty datanhallinnan tarpeisiin. On vaikea saada tietoa tieteenaloista, joissa tutkimusdatan hallintaa ei varsinaisesti harjoiteta. Joko aihetta pidetään liian vaikeana tai epärelevanttina tai sitten ei haluta raportoida julkisesti ongelmia tutkimusdatan hallinnassa.

## 5. Yliopistojen tutkimusdatan hallinnan tila

Analyysi yliopistojen tutkimusdatan hallinnan tilasta perustuu yliopiston hallinnon edustajien ja tutkijoiden antamaan tietoon työpajassa, heidän vastauksiinsa etukäteiskyselyssä sekä sähköpostikyselyyn niiden yliopistojen osalta, joihin ei ollut mahdollista aikataulu- ja resurssisyistä järjestää työpajaa.

Sähköpostikyselyihin vastanneiden yliopistojen tilannekatsaukset on lisätty mukaan analyysiin.

4. taso		<p>Yliopiston tutkimusdatan hallinnan palvelut ja suunnitelmat vastaavat hyvin toisiaan, yliopistossa</p> <ul style="list-style-type: none"> <li>▪ on yhteinen datapolitiikka</li> <li>▪ on tutkijoille ohjeistus aineistonhallinnan suunnitelman laatimiseen</li> <li>▪ tietohallinto tukee tutkimusryhmiä tutkimusdatan tallennuksen suunnittelussa ja toteutuksessa ja heillä on tarjota tarvittava tallennustila tutkimusryhmille joko yliopiston sisäisestä palvelupaletista tai ulkoisesti, jonkin muun palveluntarjoajan toimittavana</li> <li>▪ hallinnon toimijat suunnittelevat ja toteuttavat yhdessä tutkimusdatan hallinnan tuen ja jokaisen toimijan rooli ja vastuu on selvä</li> <li>▪ on selvää tutkimusdatan omistajuus-tekijänoikeuskysymykset</li> </ul>
---------	---	--

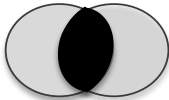
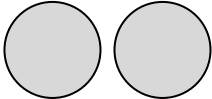
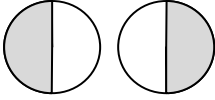
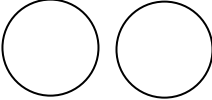
3. taso		Yliopiston tutkimusdatan hallinta on suunnitelmallista ja ainakin osa tutkimusdatan hallintaan tarjotuista palveluista on toteutettu suunnitelmien mukaisesti.
2. taso		Yliopistossa on laadittu tutkimusdatan hallinnan suunnitelmat. Olemassa olevat tutkimusdatan hallinnan palvelut poikkeavat merkittävästi suunnitelluista.
1. taso		Yliopistossa on havahduttu tutkimusdatan hallinnan suunnitteluun ja ensimmäisiin toimenpiteisiin ryhdytty suunnitelman laatimiseksi. Joitakin tutkimusdatan hallintaan <i>soveltuvia</i> palveluja on saatavilla yliopistossa.
0. taso		Yliopistossa ei ole tutkimusdatan hallinnan suunnitelmia. Tutkimusdatan hallinnan palveluita ei ole saatavilla.

Figure 1 Yliopistojen tutkimusdatan hallinnan tasot.

(Pallot kuvaavat suunnitelmia ja palveluja sekä niiden välisiä suhteita)

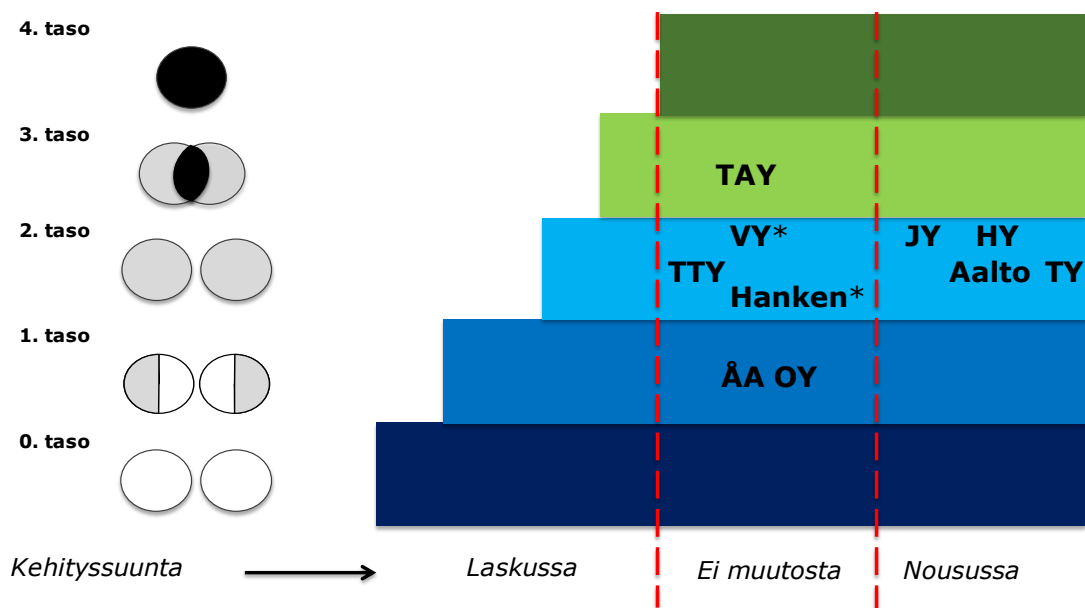


Figure 2 Yliopistojen tutkimusdatan hallinnan tila

\* Yliopiston asettaminen tarkasti kypsyyssportaikolle vaikeaa, sillä tietohallinnon osalta ei tutkimusdatan hallinnan palveluista ja tilanteesta ole tietoa.

Tutkimusdatan hallinnan kypsyyssportaikolla yksikään yliopisto ei yltänyt korkeimmalle kypsyyssportasolle. Millään yliopistolla ei ole yhteistä organisaation laajuista datapolitiikkaa, muutamalla yliopistoilla sellaisen laadinta on asialistalla tai datapolitiikkaluonnos on olemassa, mutta sitä ei ole hyväksytty. Paneelikeskustelujen perusteella yleisenä näkemyksenä yliopistojen parissa on, että koko organisaation kattavan datapolitiikan laadinta on vaikeaa tai

mahdotonta, sillä datanhallinnan tarpeet, käytännöt ja datatyypit vaihtelevasti suuresti laitoksittain ja tutkimusryhmittäin. Yhteinen datapolitiikka jäisi silloin hyvin yleiselle tasolle, eikä sitä siinä tapauksessa pidetty kovin hyödyllisenä.

Yleiseksi vaikutelmaksi työpajoista jäi, että yliopistoissa hallinnon toimijat eivät keskustele yhdessä tutkimusdatan hallinnan teemaan kuuluvista asioista ja eivät suunnittele ja toteuta yhdessä tutkimusdatan hallinnan tukea. Roolijako ja vastuualueet hallinnon eri toimijoiden välillä tutkimusdatan hallinnan tuessa pitäisi olla selkeästi määritelty. Järjestetty työpaja näytti toimivan eräissä yliopistoissa ensimmäisenä foorumina dialogiin tutkimusdatasta hallinnon toimijoiden välillä. Osa työpajaan osallistuneista on jo raportoinut, että toimiin on jo ryhdytty työpajassa identifioitujen puutteiden osalta ja dialogia eri toimijoiden välillä on jatkettu. Toisaalta Helsingin yliopistolla on toiminut jo n. vuoden ajan verkostoryhmä, johon kuuluvat Tietotekniikkakeskus, kirjasto, lakiasiat, arkisto ja Kansalliskirjasto.

Työpajat toimivat myös ainutlaatuisena foorumina hallinnon ja tutkijoiden välillä. Tutkijoiden esitykset tarjosivat hyödyllisen näkymän tutkijan päivittäiseen työhön ja siinä kohtaamiin datanhallinnan ongelmiin.

Tutkimusdatan hallinnan monisyisyys ja tuoreus kävi ilmi, kun lähetimme sähköpostikyselyn tutkimusdatan hallinnan tilasta sellaisille yliopistoille, joissa tällä kertaa ei pystytty työpajaa järjestämään. Viidestä yliopistosta kaksi vastasi kyselyyn ja muut ilmoittivat, että heillä ei joko ole vielä valmiuksia vastata kyselyyn tai että siihen vastaaminen vaatisi keskustelua hallinnon eri toimijoiden välillä ja annetussa aikataulussa se ei ollut mahdollista.

Yliopistoista Tampereen yliopisto asettuu tutkimusdatan hallinnan kypsyysportaikossa toiseksi korkeimmalle eli kolmannelle tasolle. Yliopiston tutkimusdatan hallinta on suunnitelmallista ja ainakin *osa* tutkimusdatan hallintaan tarjotuista palveluista on toteutettu suunnitelmien mukaisesti. Myös muissa kartoitukseen osallistuneissa yliopistoissa tutkimusdatan hallintaa on suunniteltu jossain määrin ja joitain tutkimusdatan hallintaan tarvittavia palveluita on joko suunnitteilla tai toteutettu. Yliopistoista, jotka sijoittuivat toiselle tasolle, koko yliopiston kattavat tutkimusdatan palvelut puuttuvat. Ominaista koko yliopistokentälle on, että kokonaisvaltainen näkemys tutkimusdatan hallintaan puuttuu.


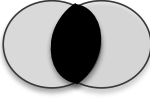
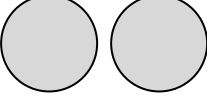

Yliopistoilla ei ole erityisesti tutkimusdatan säilytykselle kokonaisvaltaista, organisaation kattavaa palvelua, johon tärkeät tutkimusdatat säilytettäisiin. Myös ohjeistus datan säilytykseen, ymmärrettävyyden säilytykseen ja jatkokäytön mahdollistamiseen puuttuu. Yleensä palvelut ovat laitos- tai tutkimushankekohtaisia ja dataa säilytetään tutkimusryhmän tai tutkimushankkeen omissa järjestelmissä ja jopa tutkijan työasemalla.

Keskitettyjen, organisaation kattavien säilytys- ja tallennusjärjestelmien puuttuminen täytyy ottaa huomioon TTA-PAS pitkäaikaissäilytysratkaisun suunnittelussa. Näyttää siltä, että tutkimuskentällä ei ole samaan tapaan järjestelmiä, joita voisi liittää suoraan PAS-ratkaisuun kuin on [Kansallisen digitaalisen kirjaston](#) (KDK) toimialueen PAS-ratkaisua hyödyntävillä arkisto- kirjasto- ja museo-organisaatioilla.

## 6. Tutkimushankkeiden tutkimusdatan hallinnan tila

Raportissa esitetty analyysi perustuu työpajassa pidettyihin tutkijoiden esityksiin tutkimusdatan hallinnan nykytilasta sekä etukäteiskyselyn vastauksiin. Tulokset eivät kuvaa koko tieteenalan tutkimusdatan hallinnan kypsyystasoa, vaan ainoastaan arviota datan hallinnasta työpajassa esitellyn tutkimushankkeen tai kyselyyn vastanneen tutkimusryhmän osalta.

Työpajaan osallistuneet tutkimushankkeet ja tutkimusryhmät ovat yleisesti ottaen keskimääräistä valveutuneempia tutkimusdatan hallinnan osalta tai tutkimusdatan hallinnan tarpeet ovat tulleet heille vastikään ajankohtaisiksi.

4. taso		<p>Tutkimushankkeen tutkimusdatan hallinnan palvelut ja suunnitelmat vastaavat hyvin toisiaan, tutkimusryhmässä</p> <ul style="list-style-type: none"> <li>▪ on yhteinen datapolitiikka</li> <li>▪ on tutkijoille ohjeistus aineistonhallinnan suunnitelman laatimiseen</li> <li>▪ tietohallinto tukee tutkimusryhmiä tutkimusdatan tallennuksen suunnittelussa ja toteutuksessa ja heillä on tarjota tarvittava tallennustila tutkimusryhmille joko yliopiston sisäisestä palvelupaletista tai ulkoisesti, jonkin muun palveluntarjoajan toimittavana</li> <li>▪ hallinnon toimijat suunnittelevat ja toteuttavat yhdessä tutkimusdatan hallinnan tuen ja jokaisen toimijan rooli ja vastuu on selvä</li> <li>▪ on selvää tutkimusdatan omistajuus-tekijänoikeuskysymykset</li> </ul>
3. taso		<p>Tutkimushankkeen tutkimusdatan hallinta on suunnitelmallista ja ainakin <i>osa</i> tutkimusdatan hallintaan tarjotuista palveluista on toteutettu suunnitelmien mukaisesti.</p>
2. taso		<p>Tutkimushankkeessa on laadittu tutkimusdatan hallinnan suunnitelmat. Olemassa olevat tutkimusdatan hallinnan palvelut poikkeavat merkittävästi suunnitelluista.</p>
1. taso		<p>Tutkimushankkeessa on havahduttu tutkimusdatan hallinnan suunnitteluun ja ensimmäisiin toimenpiteisiin ryhdytty suunnitelman laatimiseksi. Joitakin</p>



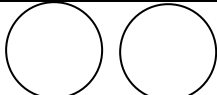
		tutkimusdatan hallintaan <i>soveltuvia</i> palveluja on saatavilla tutkimushankkeessa.
0. taso		Tutkimushankkeessa ei ole tutkimusdatan hallinnan suunnitelmia. Tutkimusdatan hallinnan palveluita ei ole saatavilla.

Figure 3 Tutkimushankkeiden tutkimusdatan hallinnan tasot

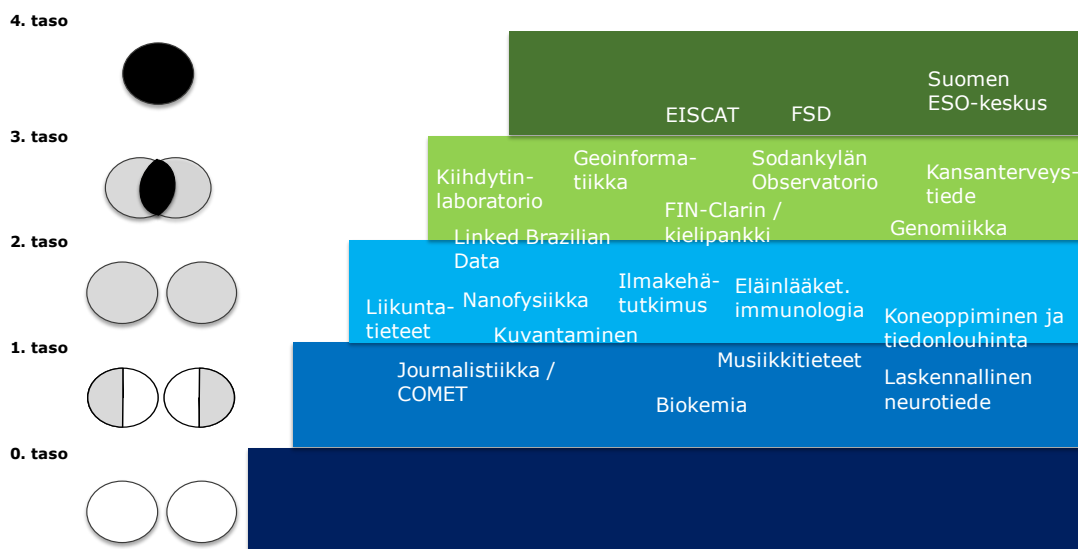


Figure 4 Tutkimushankkeiden tutkimusdatan hallinnan tila (työpajojen esitykset, N=20)

Työpajaan oli vaikea saada esityksiä ja osallistujia kattavasti kaikilta tutkimusdatan hallinnan kypsyystason luokilta. Tutkimusdatan hallinnan palvelut, aineistonhallinnan periaatteet ja terminologia eivät ole kaikille vastaajille ja työpajaan osallistujille tuttuja, mikä on vaikuttanut vastauksiin ja vaikeuttanut niiden tulkintaan.

Kartoituksessa tarkasteltujen tutkimushankkeiden osalta hankkeet jakautuivat tasaisemmin tutkimusdatan hallinnan kypsyysportaitoille kuin yliopistot, eli hankkeita oli kaikilla kypsyystasoilla – lukuun ottamatta kaikkein alinta tasoa. Korkeimmalle kypsyystasolle asettuivat pääosin tutkimushankkeet, jotka ovat osa laajaa kansainvälistä tutkimushanketta tai infrastruktuuria. Näissä tutkimushankkeissa tutkimusdatan hallinta on hyvin suunniteltua ja toteutettua ja tutkimusdatan hallinta on mietitty raakadatan tallennuksesta pitkäaikaissäilytykseen asti tutkimusdatan hyödyntämisen mahdollistamiseksi tuleville tutkijasukupolville. Tällaisia ovat muun muassa Suomen ESO-keskus ja EISCAT\_3D. Monet tutkimusinfrastruktuurit asettuvat kypsyysportaiton toiseksi ylimmälle tasolle; tutkimusdatan hallinta on hyvin suunniteltua ja toteutettua, mutta datan hallintaan ei ole kattavia palveluita koko elinkaaren aikana.

Yleisesti ottaen voidaan sanoa, että hajontaa löytyy temaattisen keskittymän sisällä paljon. Humanististen tieteiden ja tekniikan temaattinen keskittymä jäi portaitossa keskimäärin toiseksi alimmalle tasolle eli eniten tutkimusdatan hallinnassa on kehittämistä näillä alueilla.

Tilannekartoituksessa luotua kypsyystasomallia voitaisiin hyödyntää jatkossa kehityksen seuraamiseen, kun yliopistojen tutkimusdatan hallinnan tilaa



kartoitetaan seuraavan kerran. Kypsyystasomallia voitaisiin hyödyntää myös tutkimushankkeiden datan hallinnan käytäntöjen kypsyyden arvioimiseen.

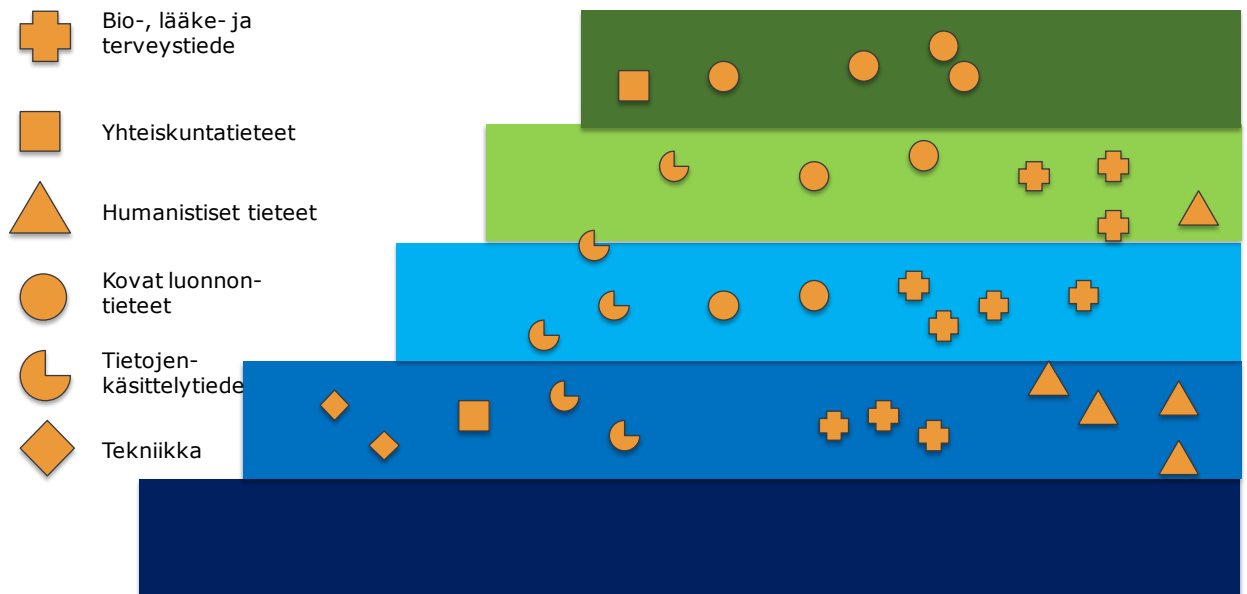


Figure 5 Tutkimushankkeiden tutkimusdatan hallinnan profiilit, N=33

(Työpajojen esitykset, kyselyyn vastanneet sekä esitykset muissa kevään TTA-tilaisuuksissa)

## 7. Johtopäätöksiä ja kehitysehdotuksia datan hallintaan yliopistojen näkökulmasta

Työpajojen paneelikeskusteluissa tuotiin esille, että tutkimusdata on yliopiston tärkeintä pääomaa, johon kaikki muu toiminta pohjautuu. Tämän vuoksi yliopistojen tulisi selkeyttää prosessejaan ja luoda tiiviimpää yhteistyötä eri toimijoiden kanssa, jotta tutkimusdatasta huolehdittaisiin sen arvon edellyttämällä tavalla.

Yksi nykytilan haasteista on tarpeellisen dialogin puuttuminen tutkimushallinnon ja tutkijoiden väliltä. Tietoisuus tutkijoiden työn käytännön haasteista datanhallintaan liittyen puuttuu. Yliopistojen olisi hyödyllistä kehittää sisäinen verkostoitumismalli, jossa tutkimushallinto, tietohallinto, tutkijat sekä kirjaston ja arkiston toimijat saataisiin yhdessä kehittämään datanhallinnan käytäntöjä.

Yliopistoissa suhtaudutaan datan hallinnan kehittämiseen positiivisesti. Yhteinen datapolitiikka kuitenkin yliopistoilta puuttuu. Olisi kuitenkin hyvin tärkeää, että yliopistot sitoutuisivat datanhallinnan käytäntöjen yhtenäistämiseen ja aineistojen keruun, käytön, säilytyksen ja jatkokäytön suunnittelun tukemiseen. Datanhallinnan käytännöt pitäisi olla selvillä tutkijalle jo tutkimuksen alkuvaiheessa, jolloin tutkimusdatan avoimuus ja jatkokäytön mahdollisuus

varmistettaisiin. Yhteiset käytännöt sekä ohjeistus datanhallintasuunnitelman tekemiseen pitäisi löytyä yliopistolta.

Tutkijoille pitäisi pyrkiä tarjoamaan tarvittavat infrastruktuuri- ja tukipalvelut järkevän datapolitiikan toteuttamisen mahdollistamiseksi. Räätelöityjä palveluja on tällä hetkellä vain vähän saatavilla eikä niille ole välttämättä laadittu lainkaan jatkuvuussuunnitelmia. Tämän tilannekartoituksen mukaan tutkimusryhmät hyödyntävät datan hallintaan ennemminkin tutkimusinfrastruktuureja kuin yliopistojen tai TTA:n palveluja.

Tutkimusinfrastruktuureissa datan hallinnan käytännöt voivat olla verrattain kehittyneitä. Haasteena ovat väliinpuotoajat, eli ne tutkimusryhmät tai yksittäiset tutkijat, joita tutkimusinfrastruktuurit eivät palvele ja joille yliopiston oma infrastruktuuri on riittämätön datan hallinnan täysimääräiseen toteuttamiseen. Tutkimusinfrastruktuurien, TTA:n ja yliopistojen kesken tulisi käydä syventävää dialogia siitä, miten infrastruktuurien keskinäistä symbioosia voitaisiin kehittää.

Yliopistot toivoivat myös ohjeistusta tutkimusdatan omistajuus- ja tekijänoikeusasioissa sekä sopimusperusteita näihin. Näiden luomisessa voisi TTA /OKM toimia apuna.

Työpajoissa mainittiin useaan otteeseen tarvetta datan hallinnan käytäntöjen liittämistä osaksi tutkijan koulutusohjelmaa. Koulutusta toivottiin mm. metatietojen tallentamisesta, datan hallinnan käytännöistä ja prosesseista, tekijänoikeus-, omistajuus- sekä sopimusasioista. Lisäksi tutkijat halusivat koulutusta ja tietoa olemassa olevista datan hallinnan palveluista. Koulutustarpeiden lisäksi myös erityiselle datamanagerin toimen perustamiselle yliopistossa nähtiin tilausta.

Kirjastot esiintyivät työpajojen paneelikeskusteluissa aktiivisesti ja näkivät kirjastoille soveltuvan laajemman roolin yliopiston datan hallinnan tukemisessa. Kirjaston vahvuusalueita ovat koulutuksen ja neuvonnan lisäksi mm. tietojärjestelmien sisältöjen järjestäminen ja kuvaaminen sekä järjestelmistä tehtävien tiedonhakujen suunnittelu ja toteutus.

Esimerkiksi Turun yliopiston kirjasto vastaa syksyllä 2013 käyttöön otettavan tutkimustietojärjestelmän osalta tutkijoiden kouluttamisesta ja ohjauksesta sekä julkaisutietojen poimimisesta tietokannoista, tietojen tarkastamisesta ja osittain tietojen syöttämisestä.

Helsingin yliopiston kirjaston datatyöryhmä on laatinut ehdotuksen [kirjaston roolista](#) datanhallinnassa. Kyseistä roolia on myös Turun yliopisto esittänyt tämän työpajan jälkeen heidän kirjastonsa rooliksi.

## 8. Tilannekartoituksen johtopäätökset OKM:n ja TTA:n kannalta

Monessa työpajan kommentissa kävi ilmi, että TTA:lle on tilausta, sen palveluita, ohjeistusta ja toimenpiteitä odotetaan. Kansallisen tason suuntalinjoille, ratkaisuille, käytännöille ja kannustimille on tarvetta.

Järkevässä muodossa toteutettu tutkimusdatan hallinta tarvitsee paljon sekä rahallisia että henkilöresursseja. Tutkimusdatan hallintaan pitäisi saada vastuutahot ja resursointi määriteltyä. Pitää miettiä, miten kasvavan datamäärän kanssa toimitaan, samoin tutkimusdatan kuvailuun tarvitaan resursseja.

Yliopistot toivoivat valmiiksi mietittyjä suosituksia toimintokäytännöistä ja sopimusmalleista, ettei kaikkea tarvitsisi miettiä alusta asti itse. Tämä auttaisi myös käytäntöjen yhtenäistämässä yliopistotasolla.

Tutkijat toivovat lyhyttä ja ytimekästä ohjeistusta tutkimusdatan hallintaan sekä pohjaa [aineistonhallintasuunnitelman](#) eli datanhallintasuunnitelman laatimiseen. Monessa yliopistossa ei olla vielä havahduttu datanhallintasuunnitelman laatimisen tärkeyteen, joten tähänkin ulkopuolelta tulevalle ohjeistukselle olisi tarvetta. Työpajoissa kaivattiin myös kansallista järjestelmää tutkimusdatan verifiointiin. Myös kaikille tieteenaloille yhteiselle [REMS](#)-palvelulle nähtiin tilausta. REMS-palvelun kautta tutkijat voivat hakea käyttöoikeuksia tutkimuksen tietoaaineistoihin, ja aineistojen oikeudenhaltijat voivat käsitellä saapuneita hakemuksia sekä hallinnoida käyttöoikeuksia.

Jo yliopistoja koskevassa osiossa todettiin, että koulutus- ja ohjeistustarpeet datanhallintaan liittyen ovat merkittävät. TTA/OKM voisi olla osin mukana kehittämässä näitä koulutuksia sekä luomassa suosituksia yliopistojen tutkimushallinnolle omistajuus- ja tekijänoikeusasioiden käytänteisiin ja sopimukseen.

## 9. Tilannekartoituksen johtopäätökset tutkijoiden kannalta

Datan hallinnan käytännöt ja valveutuneisuus vaihtelivat paljon tutkimushankkeittain. Yleisesti ottaen tilannekartoituksen perusteella voidaan sanoa, että tutkijat haluavat keskittyä tutkimuksen tekemiseen ja datanhallintaan käytetty aika koetaan usein olevan pois tutkimuksen teosta. Näiden käytäntöjen suotaisiinkin tulevan mahdollisimman valmiina ja helppokäyttöisinä. Asioiden selvittäminen koetaan työläänä ja byrokratiaa halutaan välttää. Tutkimusdatan hallinta tulisi pyrkiä tekemään toimivaksi ja tutkimusta helpottavaksi tukipalveluksi. Kaikkeen tarvitaan selkeä, käytännönläheinen ohjeistus, ennemmin järkeviä ohjeita kuin tiukkoja sääntöjä.

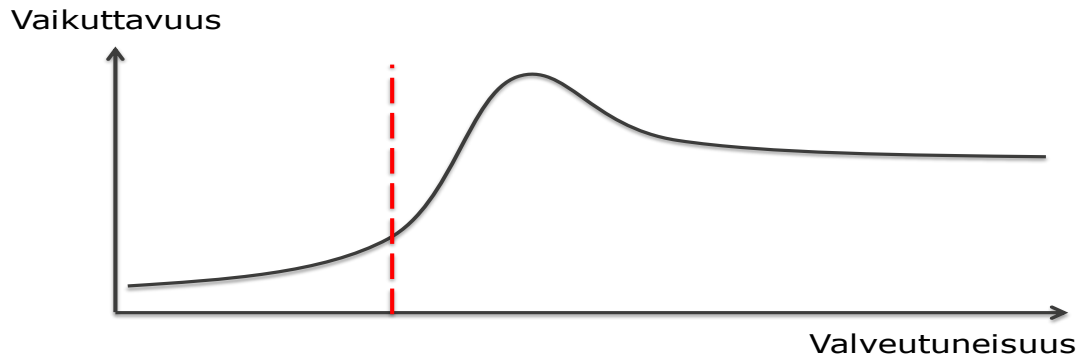


Figure 6 *Valveutuneisuuden suhde käytäntöjen vaikuttavuuteen*

- Tutkijat kokevat aineistojen hallinnan pakonomaiseksi veloitteeksi, johon mieluummin sovelletaan nopeita ratkaisuja kuin vaikeaselkoisia käytäntöjä.
- Tuloillaan on voimakkaan edistymisen ajanjakso, missä toimivilla ratkaisuilla voidaan nopeasti saavuttaa merkittäviä muutoksia toimintakulttuurissa ja sen vaikuttavuudessa.

Miten tutkijoita tulisi kannustaa datan hallintaan? Tutkijoiden käytännön kokemuksen mukaan tutkijan maine kompensoi huonosti datan keräämistä ja hallinnointia. Datan hallinnointia pitää selkeyttää, luoda toimivia infrastruktuureja sekä parantaa ohjeistusta. Tutkimusdatan hallintaan tarvitaan lisää kannustimia. Lisäksi meritoitumismalleja tulee kehittää. Nykyään tutkija meritoituu lähinnä julkaisujen perusteella. Pitää kehittää keinoja, joilla tutkija saa tiedeyhteisöltä tunnustusta hyvin dokumentoidusta aineistosta. **Meriittiä** pitää saada myös datan kautta, esim. uuteen julkaisuun / tutkimukseen pitää saada automaattisesti alkuperäisen datan tuottaneen tutkijan nimi.

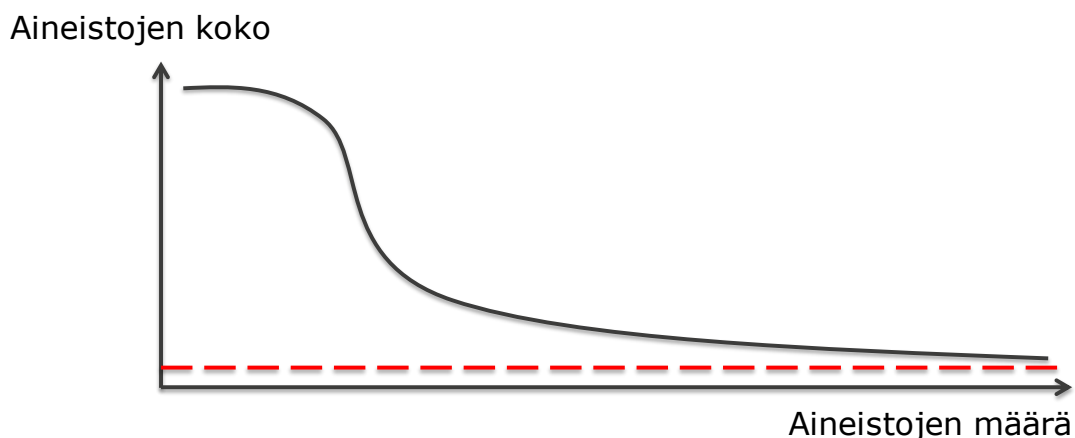


Figure 7 *Aineistojen koon suhde aineistojen määrään*

- Suurikokoiset aineistot dominoivat palvelujen kehitystä. Merkittävä osa tutkimusaineistoista on kuitenkin pienikokoisia.

- Käytännöt vaihtelevat voimakkaammin tutkimusaineistojen ”pitkässä hännässä”.

## 9.1. Datan julkaiseminen ja jakaminen

Tilannekartoituksen perusteella dataan pääsy on tutkijoiden mielestä usein hankalaa. Toisaalta osa tutkimusryhmistä käyttää tutkimuksessaan pääosin muualla tuotettua ja jaettua dataa. Ongelmia esiintyy myös tutkimuksessa tarvittavan julkisen datan saamisessa. Esimerkiksi Tilastokeskuksen aineistot ovat maksullisia ja KAVA:n tekijänoikeudet estävät analysoitavan aineiston tallentamisen.

Tutkimusdatan tallennuksessa koetaan olevan ongelmia. Syitä on erilaisia. Tutkijalla ei mahdollisesti ole aikaa aineistojen jälkihoitoon, koska seuraava projekti odottaa jo työstämistä. Tutkimusraporttiin tai artikkeliin käytetty aineisto voi olla epämääräisessä, viimeistelemättömässä muodossa, jota ei kehdata luovuttaa eteenpäin. Usein ei ole myöskään datan tallennukseen selkeää ohjeistusta ja mallinnusta, eikä tietoteknistä tukea saatavissa. Ei ole myöskään riittävästi kannusteita eikä veloitteita datan tallennukseen. Referoitu artikkeli on tutkijan tulevaisuuden kannalta tärkeämpi kuin datan huolehtiminen muiden käyttöön.

Datan jakamisen edellytykset ja prosessit vaihtelevat paljon projekti- ja hankekohtaisesti. Osa tutkimushankkeista jakaa kaiken datansa ennalta sovittujen pelisääntöjen mukaan kansallisessa tai kansainvälisessä palvelussa tai omilla verkkosivuillaan. Dataa jaetaan myös tieteelliseen julkaisuun linkitettyinä.

### 9.1.1. Datan jakamisen haasteita

Datan kuvailu ja dokumentointi on usein työlästä, eikä puutteellisesti dokumentoitua dataa kannata jakaa. Tutkimusaineiston kuvailuun toivotaankin tutkijoiden piirissä nykyistä parempaa rahoitusta ja toimivaa rahoitusmallia. Datan jalostusaste lisää sen uudelleenkäytön potentiaalia ja sen jakamisesta saatavaa hyötyä.

Tutkija voi olla myös omasta datastaan niin ”mustasukkainen”, että sitä ei haluta antaa muiden käsiin. Tämä vaihtelee paljon tutkimusaloittain. Toisaalta arvellaan, että oma data ei kiinnosta muita.

Työpajoissa kysyttiin myös, kenen etu on, että data on avoimesti saatavilla, onko se käyttäjän, tutkijan vai tieteen etu? Kaikki hyötyvät datan avoimuudesta. Käyttäjä saa olemassa olevaa tietoa hyödynnettäväkseen, eikä kaikkea tietoa tarvitsee tuottaa joka kerta uudestaan. Tämä sama asia hyödyttää myös tutkijaa omassa työssä, mitä enemmän aineistoja on saatavilla, sitä nopeampaa ja kustannustehokkaampaa tutkimuksen tekeminen on. Lisäksi datan kautta meritoitumista tulee kehittää, jolloin tutkijoille koituu suurempi hyöty datan jakamisesta. Viittaukset eivät saisi olla riippuvaisia yksittäisestä tutkijasta. Viittaus pitäisi syntyä automaattisesti (kuten Harvardin [Dataverse](#)-sovelluksessa, jossa viittaus alkuperäiseen dataan muodostuu automaattisesti). Viittauksen automaattinen generointi tuo meriittiä ja lisää motivaatiota. Yksi keino saada paljon viittauksia on sisällyttää metatietoon teksti, jota on helppo siteerata. Näin saa paljon viittauksia.

Kolmantena hyötyjänä datan avoimuudesta on tietysti tiede. Tutkimuksen menetit ja tulokset pitäisi olla tarkasti dokumentoituna ja avoimesti kaikkien käytössä, jolloin tutkimukset ovat entistä paremmin verifioitavissa, toistettavissa ja hyödynnettävissä jatkokäyttöön. Datan tulisi myös olla avoimesti louhittavissa. Tiedonlouhinnalla tarkoitetaan menetelmiä, joilla pyritään oleellisen tiedon löytämiseen suurista tietomassoista.

Yhtenä kannustimena datan julkaisuun voisi olla päätös, että julkaisut eivät hyväksyisi tutkimuksia, joiden aineisto ei ole saatavilla. Lisäksi voisi kehittää uuden tutkijan ansioluettelon mallin, jossa julkaistut tutkimusaineistot ovat mukana.

### 9.1.2. Omistajuus- ja tekijänoikeudet

Tietosuoja-asiat tulee myös ottaa huomioon dataa jaettaessa. Epäselvät tilanteet niissä eivät ole harvinaisia. Omistus-, hallinta- ja tekijänoikeusasiat olisi hyvä määritellä jo tutkimuksen alkuvaiheessa, yhtenä osana datanhallintasuunnitelman tekoa.

Usein käytäntönä on määritellä aineistolle "embargo", eli aika, esimerkiksi yksi vuosi tutkimuksen valmistumisesta, jona sitä ei saa julkaista. Tämän jälkeen data vapautetaan koko tutkimusyhteisölle. USA:ssa Presidentti Obaman hallitus määräsi 2/2013 julkisen rahoituksen tutkimukset avoimeksi vuoden sisällä artikkelin valmistumisesta. [Expanding Public Access to the Results of Federally Funded Research](#)

[Creative Commons Suomi](#) tarjoaa tietoa avoimista lisensseistä ja joustavan mahdollisuuden julkaista materiaaleja turvallisesti ja laillisesti.

### 9.1.3. Datan säilytys

Tilannekartoituksen työpajoissa kävi ilmi, että tutkimusdatan tallennuskäytännöt ovat hyvin vaihtelevia. Dataa tallennetaan mm. tutkimusryhmän omille kovalevyille, eikä tutkimuksen valmistumisen jälkeen kukaan huolehdi datan säilymisestä mahdollista jatkokäyttöä tai tarkistuslaskelmia varten. Toisaalta taas monet kansainväliset tutkimushankkeet, kuten EISCAT\_3D ovat varautuneet datan pitkäänkin säilytykseen ja datan käyttöoikeuksista ja jakamisesta on sovittu kirjallisesti.

Datan säilytyksessä ja jakamisessa on havaittavissa tutkimuskulttuurin muutos, viittausten määrän lisääntyminen edesauttaa tätä kehitystä. Työpajoissa peräänkuulutettiin myös tallennusjärjestelmien helppokäyttöisyyttä ja luotettavuutta. Järjestelmien käyttöön pitäisi saada opastusta ja niiden käyttö ei saisi tulla liian kalliiksi tutkijalle. Lisäksi datan tulisi olla tallennuksen jälkeenkin helposti tutkijan käytettävissä, mikäli tarvetta jatkokäyttöön tulee.

Työpajoissa esiteltiin TTA:n kehittämää [IDA-tallennuspalvelua](#), joka on turvallinen ja helppokäyttöinen säilytyspalvelu datalle ja siihen liittyvälle metatiedolle.

## 10. Tutkimusdatan hallinnan haasteet ja mahdollisuudet

Seuraavassa taulukossa on listattu olemassa olevia palveluita, mahdollisuuksia, puutteita ja haasteita tutkimusdatan hallintaan liittyen.

<p>Olemassa olevat palvelut ja toimivat käytännöt</p>	<ul style="list-style-type: none"> <li>▪ Tampereen yliopistolla kaikille tutkijoille aineistohallinnan koulutusta</li> <li>▪ Helsingin yliopiston kirjaston sisäinen dataosaaja- koulutus</li> <li>▪ Helsingin yliopiston Tietotekniikkakeskuksen neuvontapalvelut</li> <li>▪ Helsingin yliopiston tutkimuksen tuen lakipalvelut</li> <li>▪ Suomen Akatemian suositus säilyttää tutkimusprojektin data FSD:ssä ja FIN-Clarinissa</li> <li>▪ TTA suosittelee IDA-tallennuspalvelun käyttöä</li> <li>▪ Tutkimusinfrastruktuureissa omia temaattisen keskittymän datan hallinnan käytäntöjä ja palveluita</li> </ul>
<p>Mahdollisuudet</p>	<ul style="list-style-type: none"> <li>▪ Yliopistojen hallinnon toimijoiden kattava verkostoituminen ja yhteistyö tutkimusdatan hallinnassa</li> <li>▪ Esim. Turun yliopistossa hallinnon toimijat ovat lähteneet kehittämään datan hallintaa osaksi työpajan innoittamana</li> <li>▪ TTA:lle tilausta; sen palveluita, ohjeistusta ja toimenpiteitä odotetaan</li> <li>▪ Tutkimusinfrastruktuurien ja tulevan tutkimuksen tietoinfrastruktuurin yhteistyö, toisiaan täydentävät palveluelementit</li> <li>▪ Ulkomaisten hyvien käytäntöjen hyödyntäminen</li> <li>▪ Data ymmärretään yliopiston keskeisenä pääomana, tutkimusdatan hallinnan kehittämiseen on herätty ja siihen suhtaudutaan positiivisesti</li> <li>▪ Datamanagerin osaamisen korostuminen, uuden ammattiryhmän kehittyminen</li> </ul>
<p>Puutteet ja puuttuvat palvelut</p>	<ul style="list-style-type: none"> <li>▪ Yliopiston yhteinen datapolitiikka</li> <li>▪ Metatietomalli- ja standardit eivät käytössä</li> <li>▪ Vaivaton tapa ja paikka tallennukseen</li> <li>▪ Datatien hallinnan palvelut ja prosessit yliopistoissa</li> </ul>



	<ul style="list-style-type: none"> <li>▪ Peruskoulutus datanhallintaan ja ohjeistus datanhallintasuunnitelman tekoon</li> <li>▪ Ohjeistus sopimuskäytännöistä, valmiita sopimusmalleja yliopistojen käyttöön</li> <li>▪ Yliopistoissa dialogi puuttuu hallinnon ja tutkijoiden välillä, ei tunneta tutkijan työn käytännön haasteita datanhallintaan liittyen</li> </ul>
Haasteet	<ul style="list-style-type: none"> <li>▪ Useassa yliopistossa yhteisen datapolitiikan luomista pidetään mahdottomana, sillä laitosten ja tutkimusryhmien tila ja tarpeet vaihtelevat</li> <li>▪ Infrastruktuurien välisen yhteentoimivuuden kehittäminen</li> <li>▪ Viestinnän ja markkinoinnin haasteet, miten tavoittaa tutkijat</li> <li>▪ Kannustimien puuttuminen tutkimusdatan hallintaan</li> <li>▪ Pitkäaikaissäilytysratkaisun suunnittelu ja toteuttaminen</li> </ul>

Figure 8 *Tutkimusdatan hallinnan olemassa olevat palvelut, mahdollisuudet, puutteet ja haasteet*

## 11. Huomioita PAS-työhön

PAS-järjestelmällä tarkoitetaan laitteisto- ja ohjelmistokokonaisuutta, jolla tutkimusaineistojen pitkäaikaissäilytys toteutetaan. Pitkäaikaissäilytys tarkoittaa digitaalisen informaation luotettavaa säilyttämistä kymmenien tai jopa satojen vuosien ajan.

Tutkijat näkevät itsensä pitkäaikaissäilytysratkaisun ensisijaisina hyötyjinä. PAS ymmärretään lähinnä pitkäkestoisena tallennusalustana, mutta pitkäaikaissäilytyksen merkitystä laajassa mittakaavassa ei ole tiedostettu eikä sen edellyttämiä muutoksia tiedonhallinnan toimintakulttuurissa. Joitakin PAS:iin verrattavissa olevia kansainvälisiä palveluja on jo olemassa. Kansallinen PAS tulisi profiloida selkeästi ja määritellä sen erityisarvo suhteessa tiedealakohtaisiin ratkaisuihin.

PAS:in potentiaalisilta käyttäjiltä puuttuvat tuotantojärjestelmät, joihin PAS-ratkaisu olisi integroitavissa. Aloilla, jotka tyypillisesti tuottavat vain vähän dataa, ei ole olemassa käytäntöjä PAS:ia tukevan tiedon tuottamiseen.

Aineistot usein kilpailevat pääsystä pitkäaikaissäilytykseen. Tästä syystä valintakriteerien tulisi sisältää myös ehdot aineiston poistolle järjestelmästä.

## 12. Huomioita metatietotyöhön

Tutkimusdatan metatieto on kuvailevaa ja määrittävää tietoa tutkimusaineiston ominaisuuksiin ja käyttöön liittyen. Laadukkaat metatiedot parantavat aineistojen löydettävyyttä ja käytettävyyttä.

Metadatatietojen tallennukseen tutkijat toivoivat helppoja ratkaisuja, jotka eivät vie liikaa aikaa. Datan kuvailuun kaivattiin myös toimivia kannustimia sekä mietintää siitä, kuka maksaa lisätyöstä aiheutuvat kustannukset. Datan kuvailua pidettiin monella tavalla haasteena sekä työläänä toteuttaa. Työpajoissa esiteltiin monia ratkaisuja, joissa suuri osa metadatatista syntyi automaattisesti, joko mittalaitteen tai datanhallintatyökalun toimesta.

Tutkijat kaipaavat ennen kaikkea ohjeistusta metatietotyön toteuttamiseen. Tarve on suurin yhteistyöhankkeissa, joissa metatietojen rooli korostuu. Metatietomallit syntyvät kansainvälisessä yhteistyössä, joka takaa niiden laajan käyttöönoton ja hyödynnettävyyden.

Tutkimusaineistojen metatiedolle on olemassa erilaisia malleja, suosituksia, formaatteja ja standardeja. Tähän mennessä TTA-metatietotyöryhmä on listannut yli 30 erilaista formaattia. Suurin osa niistä on kansainvälisiä ja kehitetty yksittäisen tieteenalan tarpeeseen tai tietyn aineistotyyppin kuvailuun. Ne eroavat toisistaan muun muassa siinä, miten yksityiskohtaisesti tutkimusaineiston voi niiden avulla kuvailla ja siinä, kuinka vakiintuneesti niitä ylläpidetään. ([Minimimetatietomalli](#))

Yhteinen metatietopalvelu, jossa on selkeät ja yksiselitteiset säännöt, kannustaa metatietojen tallennukseen. [Kata](#)-aineistokatalogi on tutkimusaineistojen kuvailuja eli metatietoja keräävä palvelu, jota toteutetaan osana TTA-hanketta. Kata on nyt pilotointivaiheessa ja OKM julkistaa palvelun tämän vuoden kuluessa. Kata tarjoaa teknisen apuvälineen metatietojen tuottamiseen ja levittämiseen sekä helpottaa tutkimuskäyttöön saatavilla olevien aineistojen löytämistä ja käyttämistä uusissa tutkimuksissa.

## 13. Huomioita tietoinfrastruktuurityöhön

Infrastruktuurityötä ajatellen tilannekartoitusaaineisto käytiin läpi huomioiden se, missä määrin yliopistot ja tutkimushankkeet ovat olleet aktiivisia datan hallinnan eri osa-alueilla. Analyysissä tutkimushankkeelle tai yliopistolle annettiin piste aina sen mukaan ovatko he kehittäneet käytäntöjä kyseisellä osa-alueella (1) vai eivät (0). Tämän jälkeen pisteet laskettiin yhteen ja suhteutettiin aineiston kokoon (yliopistojen tai tutkimushankkeiden määrä). Näin saadusta havaintoaineistoista piirrettiin kuvaajat, jotka on esitetty alla.

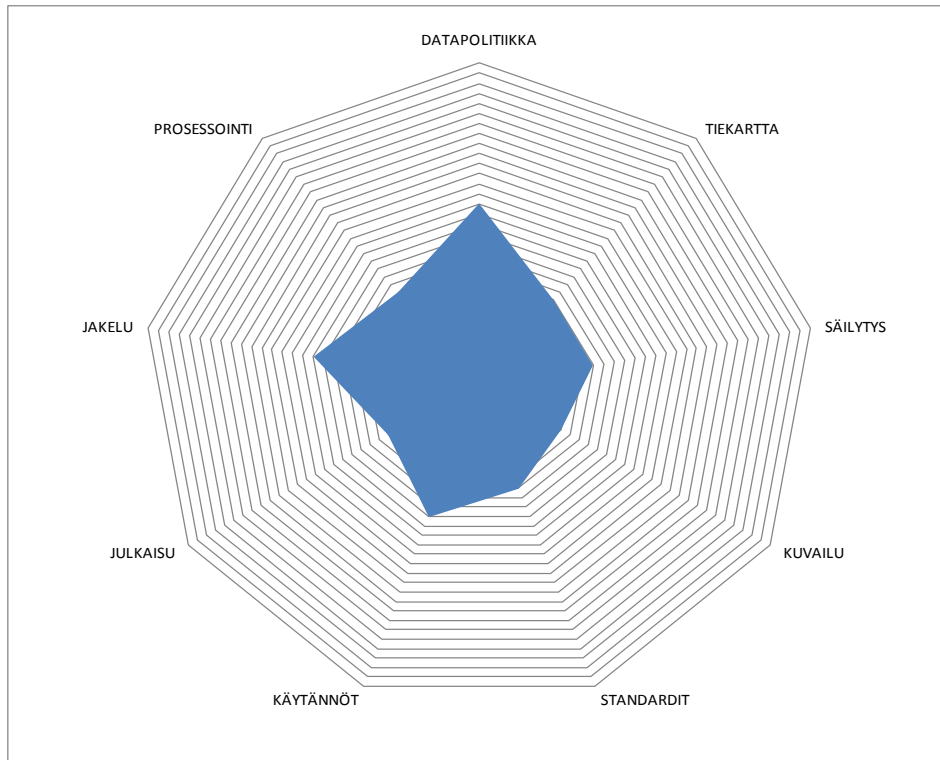


Figure 9 *Tutkimushankkeiden aktiivisuus tutkimusdatan hallinnassa*

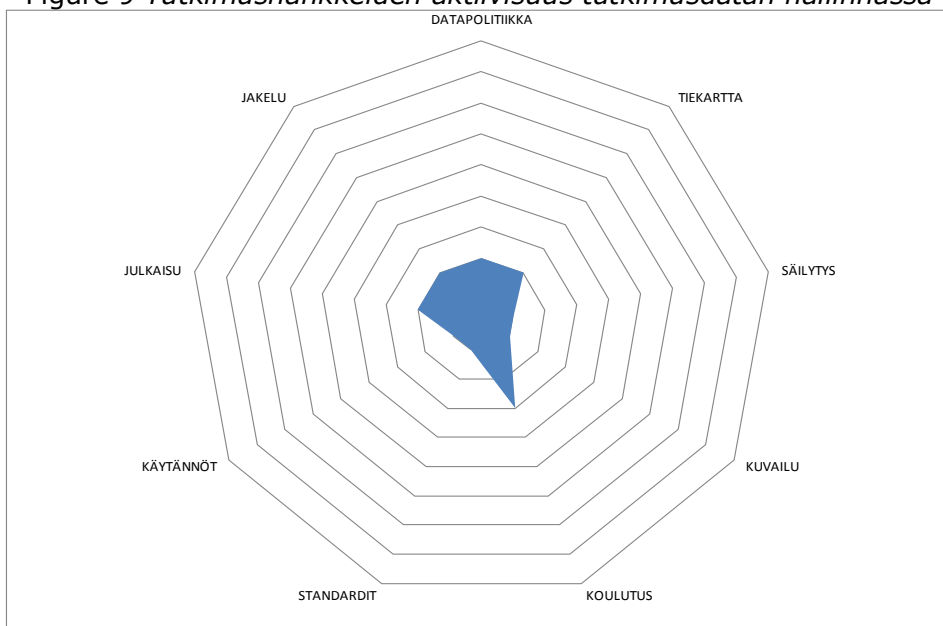


Figure 10 *Yliopistojen aktiivisuus tutkimusdatan hallinnassa*

Kaavioista käy hyvin ilmi se kuinka vallitseva tila on vielä monelta osin määrittelemättä ja oikein kohdistetuilla toimenpiteillä voidaan merkittävästi vaikuttaa infrastruktuurin rakenteeseen. Olennaista on varhainen vuorovaikutus infrastruktuurin käyttäjien ja kehittäjien kesken sen varmistamiseksi, että toimenpiteet saavat riittävästi huomiota ja tukea osakseen.

Kuvaajia voi arvioida myös toisinpäin eli siten, missä määrin toimenpiteet aiheuttaisivat häiriötä infrastruktuurin nykytilaan nähden. Tässäkin tapauksessa tuloksia voi pitää rohkaisevina. Merkittävältä osalta toimijoita puuttuvat käytännöt tutkimusdatan hallintaan ja niilläkin joilla jo on olemassa olevia käytäntöjä, käytännöt eivät ole vielä kypsiä.

## 14. Loppusanat

Tehdyn tilannekartoituksen kannanotoissa tuli ilmi, että tutkimusdata on yliopistojen tärkeintä pääomaa. Tästä pääomasta huolehtiakseen yliopistojen tulisi selkeyttää prosessejaan ja luoda sisäisiä verkostoitumismalleja, jossa eri toimijat saataisiin yhdessä kehittämään datanhallinnan käytäntöjä. Olisi hyvin tärkeää, että yliopistot sitoutuisivat datanhallinnan käytäntöjen yhtenäistämiseen ja tutkijoiden tukemiseen niiden käyttämisessä.

Järkevässä muodossa toteutettu tutkimusdatan hallinta tarvitsee paljon sekä taloudellisia, että henkilöresursseja, joten resursointi sekä vastuutahot pitäisi saada määriteltyä. Myös koulutus- ja ohjeistustarpeet datanhallintaan liittyen ovat merkittävät. Yliopistot kaipaavat apua koulutusten suunnitteluun sekä suosituksia toimintokäytännöistä ja sopimusmalleista mm. omistajuus- ja tekijänoikeusasioihin liittyen.

Kun tutkijoiden valveutuneisuus kasvaa ja he sisäistävät datan hallinnasta saatavat hyödyt, investoinnit infrastruktuurin kehittämiseen alkavat tuottamaan tulosta. Toimivilla ratkaisuilla, jotka on suunniteltu tutkijoiden tarpeisiin, voidaan nopeasti saada aikaan merkittäviä muutoksia tutkimuksen toimintakulttuurissa ja sen kehittämisessä datan jakamiselle otollisemmaksi maaperäksi.

## Lähteet:

Suomen Akatemia (2013) Tutkimusinfrastruktuurien tiekartta, 26.8.2013.  
<http://www.aka.fi/fi/A/Ohjelmat-ja-yhteistyo/Tutkimusinfrastruktuuri/Tiekartta/>

Creative Commons Suomi. 26.8.2013. <http://creativecommons.fi/>

CSC – Tieteen tietotekniikan keskus. REMS. Tietosuojaseloste. 26.8.2013.  
<https://rems.csc.fi/web/guest/tietosuojaseloste>

Harvard University. Harvard Dataverse Network. 27.8.2013.  
<http://dvn.iq.harvard.edu/dvn/>

Helsingin yliopiston kirjasto (2012) Helsingin yliopiston kirjaston tutkimusdataprojekti, loppuraportti versio 1.2.  
[https://helda.helsinki.fi/bitstream/handle/10138/36222/Helsingin\\_yliopiston\\_kirjaston\\_tutkimusdataprojekti\\_versio1.2.pdf?sequence=2](https://helda.helsinki.fi/bitstream/handle/10138/36222/Helsingin_yliopiston_kirjaston_tutkimusdataprojekti_versio1.2.pdf?sequence=2)

The White House (2013) Expanding Public Access to the Results of Federally Funded Research. 10.9.2013.  
<http://www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research>

Tutkimuksen tietoaaineistot. IDA-tallennuspalvelu. 10.9.2013.  
<http://www.csc.fi/sivut/ida>

Tutkimuksen tietoaaineistot. Kata-aineistokatalogi. 10.9.2013.  
[http://www.csc.fi/csc/ajankohtaista/uutiset/kata\\_pilotit](http://www.csc.fi/csc/ajankohtaista/uutiset/kata_pilotit)

Tutkimuksen tietoaaineistot. Minimimetatietomalli (2012)  
<http://www.csc.fi/csc/sivut/e-infra/tta>

Yhteiskuntatieteellinen tietoarkisto. Tutkimusaineistojen tiedonhallinnan käsikirja 21.8.2013.  
<http://www.fsd.uta.fi/tiedonhallinta/osa2.html#tiedonhallintasuunnitelma>