

Data Support  
University of Helsinki

**RDM Syventävä –  
Resurssit ja vastuut**

**Helsingin yliopiston Datatuki  
Tuija Korhonen**

**RDM Syventävä 2: 25.4.2024**



**RDM Syventävä  
Helsingin yliopiston Datatuki  
Kevät 2024**

# ESITYKSEN SISÄLTÖ

Tässä esityksessä keskitytään tutkimusdatan hallintaan liittyviin

- resursseihin: aika, työmäärä ja raha
- vastuihin: roolit, tehtävät ja työnjako

Nämä ylläolevat asiat on yleensä suunniteltu hyvinkin tarkasti tutkimussuunnitelmassa koko projektia ajatellen.

Datanhallintaan liittyviä resursseja ei usein huomioida tai niitä ei osata arvioida tutkimuksen alkuvaiheessa.

Aineistonhallintaan ja varsinaisiin tutkimuskuluihin kuuluvia kuluja voi olla vaikea erotella.

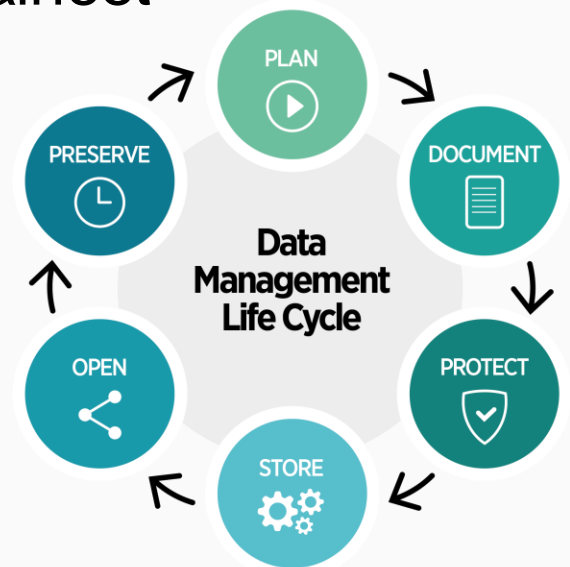
# MIETITÄÄN HETKI DATANHALLINTAA

- Mistä asioista voisi tulla kustannuksia, jotka liittyvät tutkimusdatasi hallinnoimiseen?
- Kuinka paljon työaikaa voisi mennä datan käsittelyyn?
- Mistä nämä kustannukset maksetaan?
- Mitä erilaisia tehtäviä datanhallintaan sisältyy koko projektin elinkaaren aikana?
- Voitko nimetä henkilöt,
  - a) ketkä ovat vastuussa eri tehtävistä?
  - b) ketkä tekevät konkreettisesti työn ruohonjuuritasolla?

# AINEISTONHALLINNAN TYÖNKULKU JA ELINKAARI

Lähtökohta: Mitä olet kirjoittanut aineistohallintasuunnitelmaasi?

- Suunnitelma (DMP) on usein aika yleisellä tasolla.
- Suunnittele aineistohallintasi työnkulku ja elinkaari **konkreettisesti** tasolla, jossa huomioidaan datan käsittelyn kaikki eri vaiheet
- Ota selvää, mitä palveluja **kotiorganisaatiosi** tai käyttämäsi **tutkimusinfrastruktuuri** tarjoaa ja mistä sinun on huolehdittava **itse**.



A collection of pocket watches is shown. The central focus is a gold-colored watch with a white face and a decorative gold case. The watch face has the word "QUARTZ" printed on it. The case is open, revealing a blue dial with gold hands and a gold-colored case with intricate floral patterns. The watch is surrounded by several other pocket watches, some of which are silver-colored and have various designs on their cases. The background is dark and out of focus, emphasizing the watches in the foreground.

**RESURSSIT (AIKA, TYÖMÄÄRÄ JA RAHA)**

# AIKA JA TYÖMÄÄRÄ

- Tunnista työnkulussasi ne vaiheet, jotka vievät kallisarvoista aikaa tutkimuksen tekemiseltä
  - Onko dokumentaatio osa päivittäisiä rutiinejasi?
    - Esim. 40 000 digitaalista valokuvaa kerätty projektissa, mutta niiden metadata on puutteellinen; seurauksena suuri ajanhukka projektin lopussa, kun kokoelma pitäisi arkistoida
  - Datun laadunhallinta, konvertiot, siirrot jne.
    - Aiotko litteroida nauhoitetut haastattelut? Transkriointi on työlästä ja laaduntarkistus vie runsaasti aikaa.
    - 100 Gt:n lataus tallennuspaikkaan kohtuullisen internet-yhteyden kautta voi kestää noin 4-5 tuntia; Kuka tarkistaa siirron oikeellisuuden?
    - Omistusoikeudelliset ohjelmistot ja tiedostoformaatit pitää konvertoida avoimen lähdekoodin ohjelmistoille ja formaatteihin arkistointia ja jatkokäyttöä varten
  - Tietosuoja, sopimukset,
    - Tutkimusjuristien vastaaminen kestää 3-5 viikkoa
  - Dokumenttien, sopimusten, tietosuojailmoitusten ja niiden eri käännösten valmistelu
    - Pitää tehdä ENNEN DATANKERUUTA!
  - Aineistonhallinnan suunnittelu? Vie aikaa, mutta lopulta se SÄÄSTÄÄ aikaa

# RAHA

- Tunnista työnkulusta datanhallintaan liittyvät tehtävät, joista aiheutuu kustannuksia
- Tiedätkö
  - ... mitä projektisi budjettiin on kirjattu?
  - ... mihin tehtäviin tarvitaan ulkopuolista rahoitusta ja mitkä voidaan kattaa yleisillä kustannuksilla?
  - ... kuinka paljon eri tehtävät, laitteistot tai tekniset ratkaisut maksavat?
- Tehtäväesimerkkejä:
  - Datat ja prosessien dokumentointi – voitko palkata datamanagerin?
  - Anonymisointi, litterointi, digitointi, jos palvelu ostetaan projektin ulkopuolelta
  - Nimetty juristi ISOILLE projekteille Helsingin yliopistossa – pienet projektit jonottavat tai palkkaavat juristin
  - Kielipalvelut (esim. käännökset tai kielentarkistukset)
  - Suurten datasettien tallennusratkaisut – erityisesti HY:n ulkopuoliset palvelut
  - Ohjelmistolisenssit datan käsittelyyn, jos ei ole tarjolla HY:lla  
(esim. sähköiset labrapäiväkirjat, anonymisointityökalut, dokumentointityökalut) – EHDOTA HANKITTAVAKSI!
  - Teholaskenta HY:n tai CSC:n ulkopuolelta
  - Datat avaaminen ja julkaiseminen? esim. Dryad 120 € / datasetti
  - Pitkäaikaistallennuksen, kuratoinnin jne. kustannukset – myös oikeuden siirto



# RAHA

- HY:n ohjeistus Akatemian rahoitukseen liittyen
  - <https://flamma.helsinki.fi/s/U6b5T>
- Costing tool and checklist: <https://dam.ukdataservice.ac.uk/media/622368/costingtool.pdf>
- Estimating cost RDM tool: <https://www.openaire.eu/how-to-comply-to-h2020-mandates-rdm-costs>



# RESURSSIEN KUVAILU DMP:SSÄ

## Resurssit

- Listaa aineistonhallintaan tarvittavat resurssit, välineet ja tilat (esim. tallennusympäristö, laskentavälineet, laitteisto, datan valmisteluun menevä aika julkaisemista varten, repositoriomaksut)
- Arvio datan jakamiseen, julkaisemisen, pitkäaikaistallentamisen (datan kuratointi) vaatimasta ajasta ja rahasta.
- Kuvaa panostukset asiantuntijoihin, kuten juristit, data steward, ratkaisukonsultoinnin ostaminen tai jos nämä on palkattu projektiin.
- Viittaa budjettiin (rahoittajan vaatimusten mukaisesti).

# AINEISTONHALLINNAN VAATIMAT RESURSSIT

Tasks	Resources
Data management planning	1 week
Agreements (consortium, transfer of rights)	2-4 weeks
Data privacy (GDPR) administration	2-4 weeks
Data documentation and cleaning	1-2 hour/week/person ( ~5% of the project FTE)
Data publishing (include checking the anonymization)	1-2 week(s)/data set (8 main data sets)
Storage space for sensitive data	10 TB = 2 000€/year
Archiving and deleting data	1-2 week(s)/data set (5 unpublished data sets)

Expert help for data management, preservation, and sharing tasks is provided by UH Data Support

DMP, 2019

A collection of pocket watches is shown against a dark background. In the center, a gold watch is open, revealing a white face with black hands and the word "QUARTZ" printed below the 12 o'clock position. The watch's case is decorated with intricate gold filigree. Below it, a silver watch is visible, featuring a detailed engraving of an eagle with spread wings. Other watches are blurred in the background, some with various chain links.

**ROOLIT JA VASTUUT**

# VASTUUT HELSINGIN YLIOPISTON TUTKIMUSDATAPOLITIIKAN MUKAAN

Tutkimusdatapolitiikassa vastuunjako aineistohallinnasta on määritelty eri toimijoiden kesken.

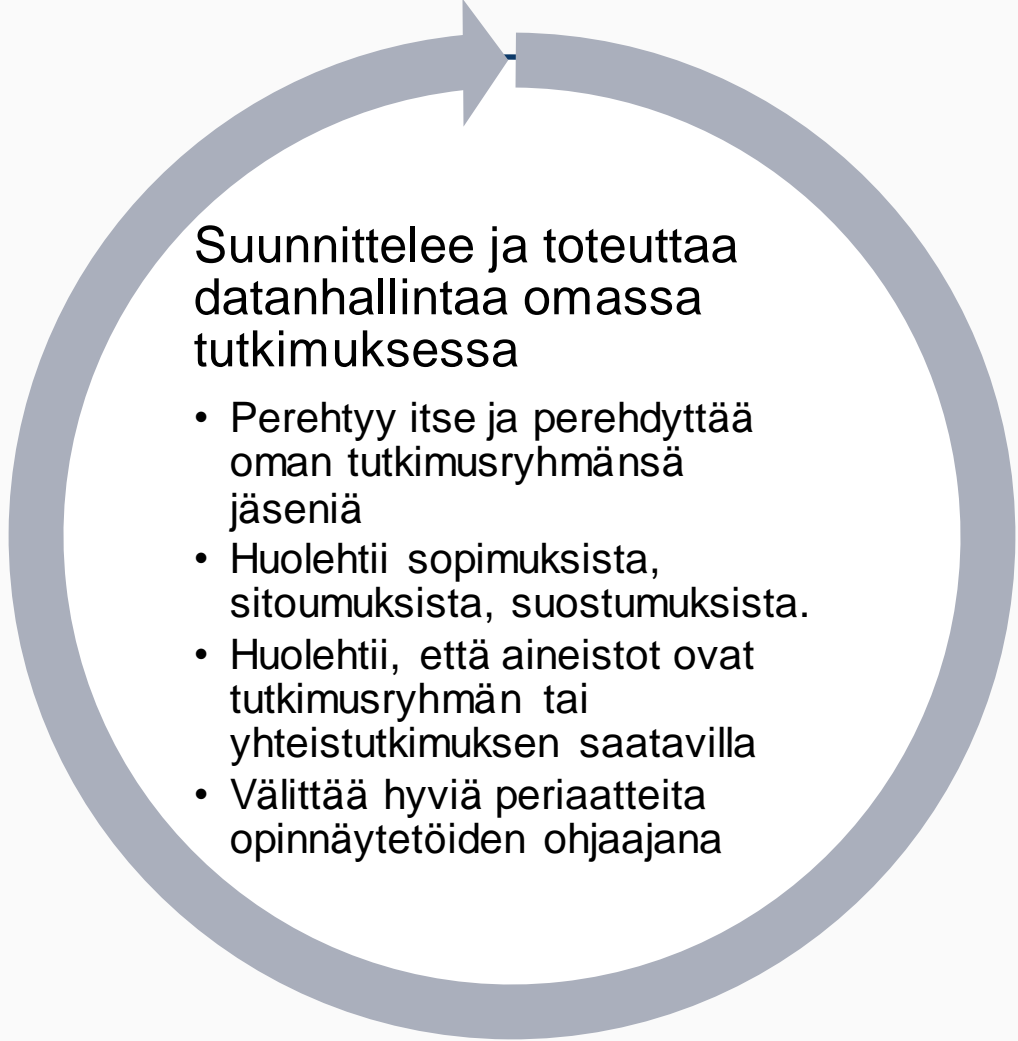
Kerrottu, mitkä asiat ovat

- Tutkijan vastuulla
- Tiedekunnan ja tutkimusta harjoittavan erillislaitoksen vastuulla
- Yliopiston vastuulla

Tutkimusaineistojen hallintaan liittyvät vastuut

<https://www.helsinki.fi/fi/tutkimus/vastuullinen-tiede/avoin-tiede#tutkimusaineistojen-hallintaan-liittyv-t-vastuut--title>

# VASTUUTUTKIJAN/TUTKIJAN TEHTÄVÄT JA VASTUUT



Suunnittelee ja toteuttaa datanhallintaa omassa tutkimuksessa

- Perehtyy itse ja perehdyttää oman tutkimusryhmänsä jäseniä
- Huolehtii sopimuksista, sitoumuksista, suostumuksista.
- Huolehtii, että aineistot ovat tutkimusryhmän tai yhteistutkimuksen saatavilla
- Välittää hyviä periaatteita opinnäytetöiden ohjaajana

# TUTKIJA VASTUUSSA TUTKIMUKSEN AINEISTONHALLINNASTA



**Kuka** on vastuussa projektin/tutkimuksen aineistohallinnasta?

- Tutkija itse omassa tutkimuksessaan
- Vastuututkija vastuussa myös tutkimusryhmän ja koko projektin aineistohallinnasta.
- Muista! Vaikka tehtäviä voi hoitaa eri henkilöt, PI on kuitenkin vastuussa kaikesta!
- **Kokoa** kaikki aineistohallintaan liittyvät tehtävät projektin ja tutkimusdatan elinkaaresta.
- **Määrittele**, kuka on mistäkin tehtävästä vastuussa
  - esim. dokumentointi, varmuuskopiointi, pääsynhallinta, sopimukset
- **Nimeä** vastuulliset henkilöt, aina kun se on mahdollista, jos tehtävät on jaettu tutkimusryhmän kesken
- **Sovi** vastuusta myös silloin, kun tarvitaan projektin ulkopuolista osaamista tai palveluja.

# ERILAISIA AINEISTOHALLINTOTEHTÄVIÄ TUTKIMUKSESSA



## Tehtäviä ja vastuita, esimerkkejä

datan kerääminen  
dokumentointi, metadatan  
tuottaminen  
datan laadunhallinta  
tallentaminen  
varmuuskopiointi  
datan siivoaminen  
datan jakaminen

datan julkaiseminen  
datan arkistointi  
datan tuhoaminen  
tutkittavien informointi  
sopimukset  
tutkimusluvut  
pääsynhallinta dataan  
anonymisointi  
translitterointi



**TUKEA**



# DATATUEN PALVELUT JA TYÖKALUT

## Aineistonhallinnan suunnittelu

- DMP:den kommentointipalvelu
- Verkkotyökalu [DMPTuuli](#) DMP:den kirjoittamiseen

## Koulutukset

- [Kurssit ja työpajat](#)
- Webinaarit ja muut tapahtumat

## Tutkimusaineiston pitkäaikaissäilytys

- Konsultointia sopivan säilytyspaikan löytämiseksi
- [Datapankki](#)
- [Data-arkisto](#)



## Datan kerääminen

- Dokumentointi ja metadata -[opas](#)
- Työkalu sensitiivisen datan keräämiseen: [REDCap](#)

## Avaa aineistosi PID:llä

- Konsultointia sopivaan avaamis- ja julkaisemisratkaisuun
- DOI:n tutkimusinfrastruktuureille [DataDOI](#)

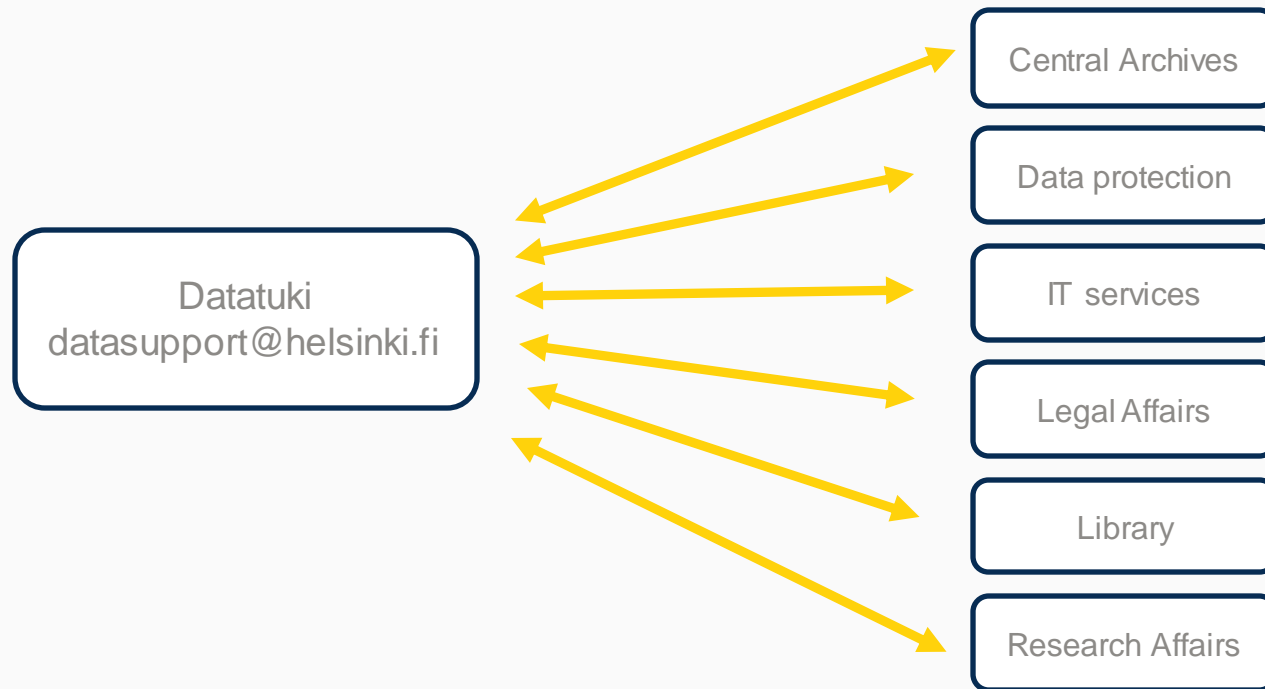
## Lakipalvelut

Tietosuojapalvelut  
Konsultointi eettisissä asioissa

## Tallennusratkaisut

- Konsultointia sopivasta tallennuspaikasta
- Sensitiivisen datan tallennus- [UMPIO](#)
- Tukea [CSC](#):n palveluihin

# HELSINGIN YLIOPISTON DATATUKI



Lisätietoa: <https://www.helsinki.fi/fi/tutkimus/palvelut-tutkijoille/datatuki>

# AINEISTONHALLINNAN OHJE

Aineistohallinnan verkkosivu

<https://www.helsinki.fi/fi/tutkimus/palvelut-tutkijoille/datatuki/aineistohallinnan-suunnittelu>

Voi olla avuksi pitää verkkosivu auki toisessa ikkunassa samalla kun kirjoittaa aineistohallintasuunnitelmaa [DMPTuuli](#):ssa.

RESEARCH / RESEARCH ENVIRONMENT /

## RESEARCH DATA MANAGEMENT

We provide assistance in research data management through out the research life cycle including data organization, storing and sharing. On this page you can find guidance to the above mentioned topics and how to answer the general questions of a Data Management Plan (DMP).

- ↓ Data policy and support services
- ↓ Data management planning
- ↓ 1. Research data
- ↓ 2. Ethical and Legal Compliance
- ↓ 3. Documentation and metadata
- ↓ 4. Storing data and access control
- ↓ 5. Opening data and long-term preservation after the research project
- ↓ 6. Data management responsibilities and resources

### Data policy and support services

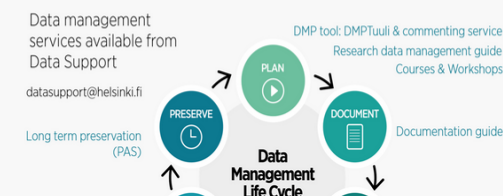


- Research data policy →
- Research data service catalogue →
- Data Support at the University of Helsinki →
- Courses & workshops →

Every member of the University of Helsinki community is responsible for good data management. The University provides research data infrastructures that includes tools and services for supporting the management, use, findability and sharing of data as well as with the capacity for storage, preservation, computing and processing.

Data Support at the University of Helsinki assists researchers in the management of research data. Data Support is a network of experts from the university library, IT Services, Central Archives, Research Affairs, Personnel Services, and Legal Affairs. You can contact us by email: [datasupport@helsinki.fi](mailto:datasupport@helsinki.fi)

On this page you can find University of Helsinki guidance for research data management.



# SSH-ALOJEN AINEISTONHALLINNAN KÄSIKIRJAT

<https://www.helsinki.fi/fi/helsingin-yliopiston-humanistis-yhteiskuntatieteellinen-instituutti/uutiset/apua-aineistohallintaan-ensimmaiset-ssh-alojen-tutkimusaineistokasikirjat-julki>

<https://www.helsinki.fi/fi/helsingin-yliopiston-humanistis-yhteiskuntatieteellinen-instituutti/ssh-alojen-aineistohallinnan-rdm-kasikirjat-2024>

## SSH-ALOJEN AINEISTONHALLINNAN (RDM) KÄSIKIRJAT 2024

Aineistohallinnan käsikirjat aineistotyyppien mukaan.



### Kyselyaineistot

Tässä oppaassa käsitellään kyselyillä kerättyä aineistoa.

Lue lisää →



### Somedata

Tämä opas käsittelee sosiaalisen median alustoilta tutkimuskäyttöön kerättävää sisältöä (somedata) sekä muita digitaalisia jalanjälkiä.

Lue lisää →



### Tutkimuskäyttöön jalostettavat tekstiaineistot

Tämä opas käsittelee tutkimuskäyttöön jalostettaviin tekstiaineistoihin liittyviä tutkimusaineistohallinnan kysymyksiä.

Lue lisää →



### Rekisteriaineistot

Tämä opas käsittelee viranomaisten hallinnon ja suunnittelun tarpeisiin keräämiä yksilötasoisia rekisteriaineistoja.

Lue lisää →



### Video- ja äänitallenneaineistot

Tässä oppaassa video- ja äänitallenneaineistoilla käsitellään tutkimusta varten kuvattuja videoita ja äänitallenteita.

Lue lisää →

Data Support  
University of Helsinki

**KIITOS!**

Helsingin yliopiston Datatuki  
Tuija Korhonen

RDM Syventävä 2: 25.4.2024



RDM Syventävä  
Helsingin yliopiston Datatuki  
Kevät 2024

