



UNIVERSITY of OULU  
OULUN YLIOPISTO

# Oulun yliopiston Avoin tiede ja tutkimus - ATT - politiikkaa ja toiminta- ohjetta valmisteleavan työryhmän väliraportti tutkimusneuvostolle 25.06.2015

**ATT-TYÖRYHMÄN JÄSENET:** SINIKKA ESKELINEN (PJ), MIRKA HINTSANEN, KARI KEINÄNEN, PÄIVI KYTÖMÄKI, RISTO VIRTANEN, MINNA MÄNNIKKÖ, RIITTA KEISKI  
TYÖRYHMÄN SIHTEERI AIJA RYYPÖ

# Tutkimusaineistot - mitä ja kenelle:

## Open Access - Top down vai bottom up

Top down:  
Rahoittajan vaatimukset

Kansainväliset ja kansalliset rahoittajat ovat esittäneet yhä tiukemmin vaatimuksen tutkijoiden tutkimusaineistojen ja tutkimustulosten avaamisesta ja järjestämisestä muiden tutkijoiden ja päättäjien käytettäväksi.

Bottom up:  
Tutkijoiden oikeudet

Yliopisto - Universitas on tutkijoiden yhteisö, jossa tutkimusta tekevät oppineet keräävät tietoa, jäsentävät sitä, luovat uusia kokonaisuuksia ja tarjoavat tuloksensa muiden tutkijoiden arvioitavaksi sitten kun katsovat ajan olevan sopiva.



# Tutkimusaineistot - mitä ja kenelle:

## Peruskäsitteitä:

Top down

Bottom up

Luvat

Oikeudet

Sopimukset

Keksinnöt

Aineistojen hallinnan rahoitus

Aineistojen arviointi ja säilytys



# Tutkimusaineistot - mitä ja kenelle:

## Luvat

Luvanvaraisia tutkimuksia on

- ihmistieteissä kun tutkitaan potilaita tai ala-ikäisiä lapsia ja nuoria
- kun käytetään koe-eläimiä
- kun käytetään geenimuunneltuja organismeja
- kun tutkimuksia tehdään luonnonsuojelualueilla tai ne koskevat uhanalaisia tai rauhoitettuja lajeja.

Eettiset toimikunnat, geenitekniikan laitos ja ympäristöhallinto käsittelevät luvanvaraiset tutkimuskysymykset.

Aikuisten ihmisten haastatteluihin ei yleensä tarvita lupaa, mutta usein luodaan henkilörekisteri ja pitää noudattaa henkilörekisterilakia.

Haastatteluihin voi liittyä tekijänoikeuksia: tekijänoikeuksia voi olla sekä haastateltavalla että haastattelijalla. Haastattelujen käytöstä on aina sovittava. Haastateltavalta on saatava riittävän laajat käyttöoikeudet. Henkilötietolain näkökulmasta laajan suostumuksen saaminen voi olla ongelmallista.



# Tutkimusaineistot - mitä ja kenelle:

## Oikeudet

Aineistojen omistusoikeus on erilainen riippuen siitä, millaisesta aineistosta puhutaan. Fyysiset materiaalit, mm. biologiset materiaalit kuuluvat yliopistolle (vasta-aineet, hiiret, kivet, kemikaalit...), mutta valokuvat usein ylittävät teoskynnyksen, jolloin omistus onkin niiden ottajalla. Samaan kategoriaan menevät esim. piirustukset, itsetehdyt kartat.

Kerättyihin aineistoihin voi liittyä kolmansien osapuolten oikeuksia, esim. tekijänoikeuksia. Oikeudenhaltijoita pitää saada riittävän laajat käyttöoikeudet tutkimusaineistojen käyttöön.

Tietokoneohjelmien ja tietokantojen (suojattu tekijänoikeudella) oikeudet kuuluu tekijälle, jollei oikeuksia ole erikseen siirretty yliopistolle.



# Tutkimusaineistot - mitä ja kenelle:

## Sopimukset

Tekijänoikeudella suojatun aineiston (esim. haastattelut, valokuvat, tietokoneohjelmat ja tietokannat) omistus kuuluu lähtökohtaisesti niiden tekijälle/tekijöille. Näissä tapauksissa pitäisi tehdä vapaaehtoinen oikeuksiensiiirto yliopistolle, jotta aineistoa voidaan käyttää jatkossakin yliopiston tutkimuksiin. On tärkeää selvittää, miten tämä hoidetaan. Jos yliopistolle siirretään kaikki oikeudet aineistoihin, on huolehdittava siitä, että tutkija voi paikkaa vaihtaessaan hyödyntää omia aineistojaan myös myöhemmässä tutkimuksessa.



# Tutkimusaineistot - mitä ja kenelle:

## Sopimukset

Monissa yrityksissä työsopimuksessa tekijänoikeudet siirretään työnantajalle, mutta yliopistoissa tällainen ei ole yleisesti käytössä. Pitää siis tehdä erilliset siirrot projekti/ tietokanta/ tietokoneohjelmakohtaisesti.

Esim. Tekes-rahoitteisissa projekteissa tehdään oikeuksien siirto yliopistolle (allekirjoitettu sopimus) projektin alkaessa, koska rahoittaja vaatii, että kaikkien tulosten omistusoikeus pitää olla yliopistolla (joka voi sitten siirtää oikeuksia esim. yritysosapuolille).



# Tutkimusaineistot - mitä ja kenelle:

## Keksinnöt

Keksintöjen omistusoikeudesta säädetään Korkeakoulukeksintölaissa (yliopistot ja ammattikorkeakoulut) ja Työsuhdekeksintölaissa (tutkimuslaitokset).

Lähtökohtaisesti keksinnön omistaja on sen tekijä, mutta laki ja/tai sopimus voi antaa jollekin muulle keksinnön omistusoikeuden. Suomessa korkeakoulukeksintölaki antaa yliopistolle tietyissä tapauksissa mahdollisuuden lunastaa keksinnön oikeudet (keksintö on syntynyt Tekesin, Akatemian, EU:n tai yrityksen rahoituksella), muulloin keksinnön oikeudet kuuluvat keksijälle.

Tutkijoiden suuntaan tärkein tieto on se, että heillä on velvollisuus ilmoittaa yliopistolle kaikista syntyneistä keksinnöistä. Sopimuksin voidaan rajoittaa tutkimustulosten ja keksintöjen omistusoikeutta.





# Tutkimusaineistot - mitä ja kenelle:

## Aineistojen hallinnan rahoitus

Uusissa rahoitushakemuksissa edellytetään aineistonhallintasuunnitelmaa, jossa tutkija kuvaa, miten hän säilyttää ja hallinnoi aineistoaan ja miten hän avaa sen yhteiseen käyttöön. Uusissa rahoitushakemuksissa myös anotaan rahoitusta aineistojen säilytykseen ja hallintaan.

Isoissa tutkimushankkeissa, joiden infrastruktuurit ovat kansallisella tai Euroopan tiekartalla, on järjestelmät ja rahoitus tutkimusaineistojen säilytykselle ja hallinnoinnille jo olemassa tai rakenteilla.

Usein ongelmana on, miten rahoitetaan olemassa olevat kokoelmien ja analogisten aineistojen, kuten äänitteiden saattaminen säilytettävään ja hallinnoitavaan tilaan. Vastuu on omistajalla eli yliopistolla/ tutkimuslaitoksella.

Yliopisto voi järjestää asiantuntija-arvion ja sillä perusteella priorisoida rahoituksen säilytettäville aineistoille.



# Tutkimusaineistot - mitä ja kenelle:

## Aineistojen hallinta ja tekijänoikeudet

Yliopisto voi arkistona digitoida aineistoja, vaikka niihin liittyisi tekijänoikeuksia, mutta jos näitä tekijänoikeuden alaisia digitoituja aineistoja halutaan levittää muiden käytettäväksi, ei digitoituja aineistoja voi levittää ilman oikeudenhaltijoiden lupaa. On selvitettävä, miten hankitaan luvat. Esimerkiksi yliopisto voi neuvotella tekijänoikeusjärjestöjen kanssa sopimuslisenssin. Toisaalta aineiston säilyttäjältä edellytetään teknisiä valmiuksia erotella, kenen käytettäväksi aineistot tulevat, jotta oikeudenhaltijat ovat valmiita sopimuslisenssiin.



# Tutkimusaineistot - mitä ja kenelle:

## Aineistojen arviointi ja säilytys

Yliopistoilta edellytetään aktiivista aineistojen hallinnan ohjeistusta: mitä aineistoja säilytetään ja miten säilytyksen ja käytön kustannukset katetaan.

Tähän tarvitaan yliopistojen datapolitiikkaa. OKM on perustanut Avoin tiede ja tutkimushankkeen, jossa rakennetaan Suomeen avoimuuden politiikkaa.

Tästä saa lisätietoa www-sivuilta [www.avointiede.fi](http://www.avointiede.fi)



# Tutkimusaineistot - mitä ja kenelle:

## Open Access - Top down vai bottom up

Top down  
OKM:n ATT - hanke:

Avoin tiede ja tutkimus -hanke on OKM:n vuonna 2014 käynnistämä hanke tiedon saatavuuden ja avoimen tieteen edistämiseksi. Tavoitteena on, että vuoteen 2017 mennessä Suomi nousee yhdeksi johtavista maista tieteen ja tutkimuksen avoimuudessa ja että avoimen tieteen mahdollisuudet hyödynnetään laajasti yhteiskunnassa.

[www.avointiede.fi](http://www.avointiede.fi)

Bottom up:  
TUHA-verkko

Tutkimuksen tuen ja hallinnon verkosto, TUHA on laaja-alainen yhteistyöverkosto, jonka toimintaan voivat osallistua kaikki tutkimuksen tuen ja hallinnon alueella toimivat yliopistoissa, ammattikorkeakouluissa, tutkimuslaitoksissa sekä muissa sidosryhmissä.



# Tutkimusaineistot - mitä ja kenelle:

## Open Access - Top down vai bottom up

Top down  
OKM:n ATT - hanke:

Hankkeen toteutuksesta vastaavat strategiaryhmä sekä asiantuntijaryhmä. Strategiaryhmän puheenjohtaja on johtaja Riitta Maijala OKM:stä. ATT-hanketta rahoittaa OKM useilla milj euroilla. Sen puitteissa toimii useita työryhmiä, mm. tutkimusaineistojen pitkäaikais-säilytystä pohtiva työryhmä.

Bottom up:  
TUHA-verkko

Tuha-verkon verkkosivuja ylläpitää CSC.

<https://confluence.csc.fi/page/viewpage.action?pageId=39061987>

CSC rahoittaa kahta työryhmää, muuten oma-kustanteinen.



# Tutkimusaineistot - mitä ja kenelle:

## Aineistojen arviointi ja säilytys

### Selvitys toimintakulttuurin avoimuudesta korkeakouluissa 2015

<http://avointiede.fi/toimintakulttuuri>

Avoin tiede ja tutkimus -hankkeessa on toteutettu selvitys, jossa on arvioitu suomalaisten korkeakoulujen avoimen toimintakulttuurin nykytilaa. Selvityksen tietovarantona on käytetty joulukuussa 2014 kerättyä aineistoa korkeakoulujen julkisilta verkkosivuilta liittyen avoimeen toimintakulttuuriin. Tätä tietovarantoa on lisäksi täydennetty helmimaaliskuussa 2015 opetus- ja kulttuuriministeriön korkeakouluille lähettämään selvityspyyntöön tulleilla vastauksilla.

Aineiston pohjalta on kartoitettu viisi erilaista avoimen toimintakulttuurin kypsyystasoa, ja kukin organisaatio on asetettu tälle kypsyystasoportaikolle.

Tuloksien mukaan korkeimmalla kypsyystasolla ei ole vielä yksikään korkeakoulu. Toiseksi korkeimmalla tasolla ovat Helsingin ja Jyväskylän yliopistot. Kolmannella tasolla on viisi korkeakoulua, neljännellä tasolla 14 korkeakoulua ja alimmalla viidennellä tasolla 9 korkeakoulua.

Selvitys toistetaan vuosittain avoimuuden etenemisen seuraamiseksi ATT-hankkeen ajan aina vuoteen 2017 asti.



# Tutkimusaineistot - mitä ja kenelle:

## Datapolitiikka - Jyväskylän yliopisto

Yliopistoilta edellytetään datapolitiikkaa. Seuraavassa esitellään Helsingin ja Jyväskylän yliopistojen datapolitiikan periaatteet:

Jyväskylän yliopistossa rehtori on hyväksynyt 9.9.2014 Tutkimuksen tietoineistojen hallinnan periaatteet.

Tutkimusaineistojen hallinta ja uudelleenkäytön edistäminen integroidaan tutkimushankkeiden, tutkimusryhmien ja tutkijoiden työhön. Vastuu sovittujen toimintakäytäntöjen toteutumisesta on laitoksen johdolla.

Aineistojen arkistoinnista ja niiden julkaisemisesta avoimeen käyttöön aiheutuvien kustannusten kattamisesta ja vastuista sovitaan yliopiston sisäisissä rahoitusneuvotteluissa viimeistään vuoden 2016 budjettia valmisteltaessa. Tiedekunnat sisällyttävät toiminta- ja taloussuunnitelmiinsa aineistojen hallinnan kustannukset.



# Tutkimusaineistot - mitä ja kenelle:

## Jyväskylän yliopisto

Tutkimusryhmän johtaja sopii yhdessä ryhmän tutkijoiden kanssa, mitä aineistoja tulee tallentaa ja avata muille tutkijoille. Ilman ryhmää toimiva itsenäinen tutkija sopii aineistojen hallinnasta laitoksen johtajan kanssa ja väitöskirjatyöntekijä ohjaajansa kanssa.

Tutkimushankkeissa ja uusia tutkimusympäristöjä rakennettaessa laaditaan ennen tutkimuksen aloittamista aineistohallintasuunnitelma hankkeen vastuullisen johtajan johdolla. Suunnitelmassa tulee ilmetä (Suomen Akatemian ohjeistus) 1. miten tutkimushanke aikoo hankkia ja käyttää tietoaineistoja, 2. mitkä ovat hankkeen käyttämien ja synnyttämien aineistojen omistus- ja käyttöoikeudet, 3. miten tutkimushankkeen tuottamia tietoaineistoja aiotaan säilyttää ja avata muille tutkijoille myöhempää käyttöä varten ja 4. miten aineistot mahdollisesti suojataan.

Yliopistopalvelut vastaa aineistoinfrastruktuurin ja toimintakäytäntöjen kehittämisestä.





# Tutkimusaineistot - mitä ja kenelle:

Jyväskylän yliopisto

## 3.2 Tietoaineistojen omistus- ja käyttöoikeudet

Aineistojen omistus- ja käyttöoikeuksista sovitaan aineistonhallintasuunnitelmaa laadittaessa ennen varsinaisen tutkimustyön käynnistymistä. Osapuolia ovat tilanteesta riippuen yliopisto ja sen yksiköt, yhteistyöorganisaatiot, rahoittajat, tutkimusryhmät, tutkijat ja ihmistieteissä myös tutkimuksen kohteet.

Yliopistolla käytössä olevassa oikeuksiensiirtosopimuksessa sovitaan tulosaineistojen lisäksi myös tutkimusaineistojen omistus- ja käyttöoikeuksista. Yhteisrahoitteisessa tutkimuksessa, tilaajarahoitteisessa tutkimuksessa ja tutkimusryhmien sisäisessä toiminnassa käytetään yliopiston laatimia tutkimusaineistot huomioonottavia sopimus pohjia. Erityisesti aineistojen pitkäaikaissäilytyksessä tukeudutaan kansallisiin palveluihin ja joillakin tutkimusalueilla eurooppalaisiin infrastruktuureihin ja näin mahdollistetaan julkisuustasosta riippuen aineistojen laaja ja avoin käyttöoikeus.



# Tutkimusaineistot - mitä ja kenelle:

Jyväskylän yliopisto

Yliopisto määrittelee tavoitteet ja kohdentaa resurssit aineistojen pitkäaikaissäilytykselle.

Yliopiston arkistosäännön mukaisesti aineistojen säilytysajat määritellään niiden käyttötarpeen sekä oikeudellisen ja tutkimuksellisen merkityksen perusteella.



# Tutkimusaineistot - mitä ja kenelle:

## Helsingin yliopisto

Helsingin yliopiston rehtori on allekirjoittanut datapolitiikan ja toimenpideohjelman helmikuussa 2015

Datapolitiikka - otteita:

5. Helsingin yliopisto tarjoaa tutkijoille ja tutkimusryhmille tutkimusdatainfrastruktuurin, joka käsittää datan hallintaa, käyttöä, löytyvyyttä ja jakamista tukevia työvälineitä, palveluita sekä tallennus-, laskenta- ja käsittelykapasiteettia. Datainfrastruktuuri rakennetaan ja sitä kehitetään yhteistyössä kansallisten ja kansainvälisten toimijoiden kanssa, ottaen huomioon myös näiden tarjoamat palvelut ja infrastruktuurit.

6. Helsingin yliopisto tarjoaa tukea tutkimusaineistoihin liittyvien oikeudellisten kysymysten tunnistamiseen ja ratkaisemiseen. Vastuulliset tutkijat vastaavat siitä, että tutkimuksessa syntyvien aineistojen omistus- ja käyttöoikeuksia koskevat sopimukset tehdään mahdollisimman varhaisessa vaiheessa, soveltuvin osin ennen tutkimushankkeen alkua.



# Tutkimusaineistot - mitä ja kenelle:

## Helsingin yliopisto

7. Rahoitushakemuksiin liitettävissä aineistonhallintasuunnitelmissa on otettava huomioon aineiston ja datan kerääminen, käsittely, omistajuus ja käyttöoikeudet, (pitkäaikais)säilytys, uudelleenkäyttö ja julkaiseminen, suunnitelmallinen tuhoaminen sekä erityisesti niiden vaatimat resurssitarpeet. Tutkimusaineistojen metadatatassa on mainittava tutkimusdatan omistaja ja oikeudelliset rajoitukset tutkimusaineistojen käytölle. Henkilötietojen ja muiden arkaluontoisten aineistojen käsittely ja säilyttäminen on otettava huomioon aineistonhallintasuunnitelmassa.

Toimenpideohjelmassa vastuutahoina ovat: yliopiston johto, viestintä, tutkijat, kirjasto, Kansalliskirjasto, tietotekniikkakeskus, tutkimuspalvelut, tiedekunnat, laitokset, tutkijakoulut, keskusarkisto



# Tutkimusaineistot - mitä ja kenelle:

## Oulun yliopisto

Oulun yliopistoon on perustettu ATT-työryhmä laatimaan Oulun yliopiston datapolitiikkaa ja toimenpideohjelmia.

Tutkimusneuvosto käsitteli asiaa kokouksessaan 26.2.2015 ja päätti esittää työryhmän perustamista valmistelemaan Oulun yliopiston ATT-politiikkaa ja toimintaohjetta yhteistyössä tutkimusneuvoston kanssa. Työryhmän tulisi ottaa huomioon kansainväliset ja kansalliset ohjeet ja ratkaisut.

Tutkimusneuvoston esityksen mukaisesti tutkimusrehtori on nimittänyt 13.3.2015 työryhmään seuraavat jäsenet: Sinikka Eskelinen (pj), Mirka Hintsanen, Kari Keinänen, Riitta Keiski, Päivi Kytömäki, Minna Männikkö, Risto Virtanen. Työryhmän tulee esittää väliraportti kesäkuussa ja tavoitteena on, että ohjeistus on valmiina lokakuussa 2015.



# Tutkimusaineistot - mitä ja kenelle:

Oulun yliopiston datapolitiikkatyöryhmä

Työryhmä on kokoontunut kolme kertaa (8.5.2015, 4.6.2015 ja 17.6.2015)

Työryhmä on tehnyt

- nykytila-analyysin esimerkkien avulla
- ehdotuksen ensisijaisista toimenpiteistä
- tavoitetilan 2017



# Tutkimusaineistot - mitä ja kenelle:

## Oulun yliopisto

Työryhmä on aloittanut esimerkkitapausten avulla tehtävällä nykytilan kartoituksella ja sen perusteella räätälöidään Oulun yliopistolle sopiva datapolitiikka.

Siinä on otettava huomioon Oulun yliopiston monitieteisyys (erityisesti lääketiede ja tekniikka) ja osallistuminen kansallisiin ja kansainvälisiin tiekarttahankeisiin.

Oulun yliopiston tutkimusneuvosto jakaa strategista infrastruktuurirahoitusta esimerkiksi arkistojen digitalisointiin.

Oulun yliopisto on mukana 12 Suomen Akatemian infrastruktuurien tiekarttahankeessa.

Oulun yliopisto osallistuu seuraaviin Euroopan tiekarttahankeisiin  
BBMRI, CLARIN, XFEL, EISCAT, EPOS, INSTRUCT, INFRAFRONTIER, EURO-BIOIMAGING



# Tutkimusaineistot - mitä ja kenelle:

## Oulun yliopisto

Oulun yliopisto on mukana 12 Suomen Akatemian infrastruktuurien tiekarttahankkeessa: Oulun yliopiston koordinoimat infrastruktuurihankkeet:

EISCAT ja EISCAT 3D

INFRAFRONTIER - Euroopan geenimuunneltujen hiirten analysoinnin, säilyttämisen ja jakelun tutkimusinfrastruktuuri

MAX IV Laboratorio

Eurooppalainen röntgen-vapaaelektroni-laser (XFEL)

Infrastruktuurihankkeet, joissa Oulun yliopisto on mukana:

Avoin paikkatiedon tutkimusinfrastruktuuri oGIIR

BBMRI biopankki-infrastruktuuri, Biokeskus Suomi, CLARIN, EPOS, Euro-Biolmaging (EuBI), Instruct, Suomen hila- ja pilvilaskennan infrastruktuuri (FGCI)

Mittag-Leffler Instituutti

Kaikissa näissä syntyy huomattava määrä tutkimusdataa.







UNIVERSITY of OULU  
OULUN YLIOPISTO

# Nykytila-analyysi ja tehtäväkentän kartoitus

# Tutkimusaineistot - mitä ja kenelle:

## Nykytilan analyysi esimerkein: Oppimisympäristö LeaForum:

LeaForumissa säilytettävä data on tallennettu lokaalille palvelimelle, eikä ole yleisen verkon/tietohallinnon piirissä. Datan määrä on tällä hetkellä noin 50 teratavua. Se ei ole nauhavarmennettua, mutta levyvarmennus on käytössä.

Tutkimusdataan pääsy on tällä hetkellä ainoastaan laboratorion toiminnasta vastaavan henkilön kautta. Tutkimusdata pääasiallisesti siirretään ulkoisilla kiintolevyillä tietoturvasyistä.

Learning and Educational Technology Research Unit Let: Aineiston taltioiminen ja säilytys tällä hetkellä:

Aineistoja säilytetään yliopiston/tiedekunnan tarjoamilla jaetuilla verkkolevyillä (suhteellisen kevyt raaka-aineisto (mm. kyselyt, haastattelut) ja niiden jatkokäsittely (mm. Excel- ja SPSS tiedostot, liitteroinnit)

Suuret datamäärät (erityisesti massiiviset videoaineistot) joudutaan miettimään tapauskohtaisesti ja esim. supistamaan ja heikentämään datan tarkkuutta ja hävittämään alkuperäiset. Edelleen suhteellisen suuret datamäärät tallennetaan tämän jälkeen yksittäisille ulkoisille kovalevyille, jotka hankitaan tapauskohtaisesti ja aineiston kanssa työskentelevät kierrättävät keskenään. Käytössä analyysiohjelmia (mm. NVivo & Observer XT Noldus).



# Tutkimusaineistot - mitä ja kenelle:

## Nykytilan analyysi esimerkein: Konenäön tutkimuskeskus(CMV)

Konenäön tutkimuskeskuksen (CMV) aineistot sijaitsevat pääasiassa Oulun yliopiston tietotekniikan, sähkötekniikan ja tietoliikennetekniikan osastojen hallinnoimassa keskitetyssä verkko- ja levypalvelinjärjestelmässä. Järjestelmä tarjoaa käyttäjähallinnan ja pääsyn tallennettuihin tietoihin sekä laskenta- ja varmuuskopiointipalvelun, josta on kuvaus sivulla <http://www.oulu.fi/tietotekniikka/eenet>. CMV:n tutkijoiden tallentaman datan määrä on tällä hetkellä suuruusluokaltaan useita kymmeniä teratavuja (TB) ja se kasvaa koko ajan.

Ryhmän sisällä jokaisella tutkimusryhmän jäsenellä on pääsy yhteiseen verkkohakemistoon, jonne aineistot tallennetaan. Lähtökohtaisesti kaikilla ryhmän jäsenillä on pääsy kaikkiin tietoihin ja kullakin käyttäjällä on vastuu oman aineistonsa sisällöstä, mutta tarvittaessa eri käyttäjien oikeuksia aineiston eri osiin voidaan rajoittaa. Käyttäjäoikeuksia hallinnoivat osaston ylläpitohenkilökunta sekä tutkimusryhmien ja projektien vastuuhenkilöt. Julkisesti jaettava aineisto on nykyisin saatavilla tutkimusryhmän virallisten verkkosivujen kautta, jotka löytyvät osoitteesta <http://www.cse.oulu.fi/CMV/Downloads>.

CMV on myös osa Oulu Bioluminescence Imaging (OBI) -verkostoa, jossa tutkimusryhmän tehtävänä on tarjota kuva-analyysipalvelua eri osapuolille. Tehtävää varten on toteutettu yliopiston infrastruktuurirahoituksella OMERO web-palvelu (<http://cse-pc66.oulu.fi/omero>), jonka kautta ulkopuolisten käyttäjien on mahdollista tallentaa biokuvantamistietoja CMV:n järjestelmään.



# Tutkimusaineistot - mitä ja kenelle:

## Nykytilan analyysi esimerkein: Biologiset DNA-sekvenssiaineistot

- julkaisuissa käytettyä DNA-sekvenssidataa on tallennettu NCBI:n (<http://www.ncbi.nlm.nih.gov/>) ylläpitämään GenBank-tietokantaan, ne ovat sieltä avoimesti saatavilla <https://www.ncbi.nlm.nih.gov/genbank/>
  - tallennettuna myös isoja sekvenssidatoja CSC:n (<https://www.csc.fi/>) palvelimille
  - FinBOL projektissa Kanadassa tuotetut sekvenssiaineistot BOLD datapankkiin <http://www.boldsystems.org/> tulevat pienen viiveen jälkeen avoimesti saataville
  - bakteerien sekvenssidataa tallennettuna European Nucleotide Archive (<http://www.ebi.ac.uk/ena>) tietopankissa ja monia eri aineistoja jaossa ResearchGate (<http://www.researchgate.net/>) kautta
  - muita aineistoja vaihtelevasti exceleissä sekä henkilökohtaisella tietokoneella, yliopiston tietokoneella että yliopiston varmistetuilla levyasemilla
- Yleensä DNA-sekvenssipohjaisissa julkaistuissa tutkimuksissa aineisto päättyy ulkomailta sijaitsevaan tietokantaan ja on sitä kautta vapaasti käytettävissä. Sekvenssiaineistojen tallennuksessa vakiintuneet datapankit ja julkaisusarjojen käytännöt.



# Tutkimusaineistot - mitä ja kenelle:

## Nykytilan analyysi esimerkein: Lajiaineistot, kokoelmat

- digitoidut biologiset kokoelma-aineistot (Eläinmuseo, Kasvimuseo) tallennettu Specify-ohjelmalla joka on Oulussa käytössä ja tietohallinnon palvelimella varmistustallennettuna. Digitoidut tiedot jaossa ulkomailla sijaitsevan Global Biodiversity Facilityn (GBIF) kautta
  - digitointiaste vaihtelee 0-50% kokoelmasta riippuen
  - kasvitieteellisellä puutarhalla omassa hallinnassa oleva kokoelmatietokanta
  - meneillään oleva geneettisen materiaalin kryo-pitkäaikaissäilytys -hanke
- muissa biologisissa kenttäaineistoissa ja niiden säilyttämisessä huomattavaa vaihtelua. Monet vanhat aineistot arkistoina eikä tallennettuna digitaalisesti. Aineistoja on saatettu digitaaliseen muotoon biodiversiteettiyksikössä ja kenttäasemilla. Joidenkin koordinoitujen hankkeiden aineistoja tallennettu ympäristöhallinnon ja sektorilaitosten tietokantoihin. Uusimpien verkostohankkeiden myötä aineistojen tallennus usein ulkomailla jossa datan hallinta-, käyttö- ja jakopolitiikat määritetty (esim. AVA, EVA, NutNet). Tarkka säilytysmenetelmä ei tiedossa. Verkostossa on siis säännöt hallinnoimisesta.

Kenttäaineistojen tallennusta on edistänyt tiekartalla ollut LTSER-infrahanke, mutta tämä ei ole edennyt toivotussa laajuudessa. Meneillään oleva kokoelmien FinBIF SA-infrahanke tukee kokoelmien digitoimista ja tallentamisjärjestelmien kehitystä.



# Tutkimusaineistot - mitä ja kenelle:

## Nykytilan analyysi esimerkein: suomen kielen tutkimusaineistot

### Suomen kielen tutkimusaineistojen tallentaminen, nykytila ja säilytyskäytännöt

#### Oulun nauhoitearkisto (ONA)

Oulun nauhoitearkistossa (ONA) säilytetään opiskelijoiden ja henkilökunnan nauhoittamia tallenteita. Kokoelmassa on yhteensä yli 7 000 tuntia erilaisia nauhoitteita, joista osa on kopioitu Suomen kielen nauhoitearkistosta (SKNA).

#### Muu arkisto

Suomen kielen arkisto sisältää nauhoitteiden lisäksi sanasto- ja nimistökokotelmat sekä korpuukset.

#### Nykytilanne

Oulun nauhoitearkiston materiaali on kokonaan digitoitu ja tullaan siirtämään osaksi Kielipankin LAT-alustaa vuoden 2015 aikana. Osia aineistosta on myös litteroitu, lisäksi ICLFI-korpus on morfologisesti koodattu sekä lemmatisoitu. Aineistojen ylläpitäminen vaatii jatkuvaa päivittämistä, suomen kielen asiantuntemusta sekä perehtymistä teknisiin välineisiin. Ylläpito sisältyy yhden lehtorin työsuunnitelmaan.

#### Aineistojen säilyttäminen

Oulun nauhoitearkisto on kokonaisuudessaan tallennettu useille ulkoisille kiintolevyille sekä tietokoneille varmuuskopioina. Lisäksi ONA tullaan siirtämään CSC:n LAT-alustan ohella myös Oulun yliopiston sisäisille verkkolevyille varmuussäilytykseen. Muuta materiaalia säilytetään suomen kielen oppiaineen arkistotiloissa luetteloituna. ICLFI-korpus on varmuuskopioitu Kielipankin KORP-palvelun lisäksi yliopiston sisäisille säilytysasemille.



# Tutkimusaineistot - mitä ja kenelle:

## Nykytilan analyysi esimerkein: saamelainen kulttuuriarkisto

Saamelainen kulttuuriarkisto toimii Oulun yliopiston Giellagas-instituutissa ja on hallinnollisesti osa Oulun yliopiston keskusarkistoa. Giellagas-instituutti alkoi kehittämään arkistoa 2009. Arkiston kokoelmat ovat sekä analogisessa että digitaalisessa muodossa. Analogimateriaalia säilytetään Oulun nauhoitearkiston nauhamakasiinissa (HU136). Varsinainen arkistotila ja digitointihuone sijaitsevat Giellagas-instituutin yhteydessä (HU324, HU330). Digitaalinen aineisto on kopioituna useille paikallisille tallenteille (kovalevyt, osittain CD- ja DVD-levyt). Vuoden 2015 aikana koko materiaali tallennetaan LTO-nauhoille.

Kokoelma jakautuu seuraaviin osiin: Saamen kieltä ja kulttuuri koskevat äänitteet, videomateriaalit, valokuvat, kirjalliset dokumentit ja Giellagas-instituutin julkaisut ja opinnäytteet. Arkiston omaa kokoelmaa täydentävät muiden arkistojen saamenkieliset aineistot. Kokoelman erityisosa ovat Puhutun saamen kielen korpus (työ aloitettu 2014 yhteistyönä Fin Clarin Kielipankin kanssa) ja saame vieraana kielenä korpus (työ aloitettu 2015) sekä Áillohaš-kokoelma. Arkiston oma kokoelma on luetteloitu ja kuvailtu paikalliseen tietokantaan. Arkisto on liittymässä vuoden 2015 aikana myös Oulun yliopiston kirjaston ylläpitämään Oula-tietokantaan, joka tarjoaa näkyvyyden Finna-järjestelmässä sekä Arkistolaitoksen Arkistojen portti-portaaliin. Kokoelmatiedot löytyvät Saamelaisen kulttuuriarkiston kotisivuilta ([http://www oulu.fi/giellagasinstituutti/saamelainen\\_kulttuuriarkisto](http://www oulu.fi/giellagasinstituutti/saamelainen_kulttuuriarkisto)).



# Tutkimusaineistot - mitä ja kenelle:

## Nykytilan analyysi esimerkein: Pohjois-Suomen syntymäkohortti

Pohjois-Suomen syntymäkohorttitutkimuksen projektikeskus (Lääketieteellinen tiedekunta, Oulun yliopisto) vastaa kohorttiaineistojen ylläpidosta ja aineiston käytön koordinoinnista. Sähköisesti tallennetun aineiston laajuus on tällä hetkellä n. 30Tb. Aineisto on tallennettu Oulun yliopiston Lääketieteellisen tiedekunnan palvelimille. Aineiston varmistuksesta on huolehdittu. Pitkittäisessä syntymäkohorttitutkimuksessa tavoitteena on seurata tutkittavia koko heidän elämänsä ajan. Näin ollen tutkimusaineiston määrä tulee kasvamaan eikä tutkimustulosten säilytysaikaa ole määritelty. Aineiston sisältöön pääsee tutustumaan rekisteröitymällä kohortin tutkijaksi ja sen jälkeen aineistoa voi hakea tutkimuskäyttöön yksikön internet sivuilla (<http://www oulu.fi/nfbc/node/18136>) kuvattujen ohjeiden mukaan. Kyseessä on ns. arkaluonteiden tietoaaineisto, jonka käyttöä ohjaavat mm. henkilötietolaki sekä tutkittavien tutkimukselle antama suostumus. Aineistosta tehtyä tutkimusta seurataan ja tilastoidaan osana käyttö lupaprosessia (julkaisut, väitöskirjat).





# Tutkimusaineistot - mitä ja kenelle:

## Tehtäväkenttää ja ratkaistavia kysymyksiä:

Yliopistolla pitäisi olla kokonaiskuva siitä, mitä tutkimusaineistoja yliopistolla on, kuka ne omistaa ja mitkä ovat niiden käyttäjäoikeudet

Yliopiston pitäisi linjata, miten kaikki tutkimusaineistojen avoimuuden nimessä suunniteltavat toimet resursoidaan ja rahoitetaan

- tutkimusaineistoihin liittyvä hallinnollinen työ näyttäisi olevan lisääntymässä; resurssointi
- tutkimusaineistojen arviointi ja niiden säilytyksen ja avoimuuden priorisointi; onko aineistojen avoimuus tulevaisuudessa yliopistolle rasite vai kenties pääomaa
- miten järjestetään ATT-tavoitteiden mukainen tutkimusaineistojen elinkaari yliopistoissa (tutkijoiden näkökulma vs hallinnon näkökulma)
- mikä on yliopiston osuus tutkimusaineiston koko elinkaaresta ja miten hallinnollinen käsittely/tukipalvelu yliopistoihin tulisi rakentaa näinä niukkojen resurssien aikoina



# Tutkimusaineistot - mitä ja kenelle:

## Tehtäväkenttää ja ratkaistavia kysymyksiä :

- tutkimusaineistojen talteenotto; kenelle yliopistolla kuuluu?
- Vastuutahot: yliopiston arkisto, kirjasto, tietohallinto, tutkimuksesta vastaavat tutkijat (esim. Oulun yliopiston omat säädökset eli Arkistonmuodostussuunnitelma määrittelee että tutkimusaineistot kuuluvat yliopistolle ellei rahoittajan kanssa ole sovittu muuta)
- kuvailu/metatietojen kiinnittäminen; kuka neuvoo ja avustaa tutkijoita? (ATT-työryhmien esitykset esim. arkaluontoisten tutkimustietojen viitearkkitehtuurista)
- aineistojen pitkäaikaissäilytys; jääkö yliopistolle? kuinka pian voidaan/pitää siirtää valtakunnallisiin tai kansainvälisiin tietovarastoihin? Tässäkin on vasta perustettu ATT-työryhmä
- Palvelutuotanto: arkisto, kirjasto, tietohallinto, lakipalvelut, CSC, kansallisarkisto, kansalliskirjasto ja muut kansalliset ja kansainväliset toimijat
- Horisontti- ja ESFRI- rahoitus pelastajana





UNIVERSITY of OULU  
OULUN YLIOPISTO

# Työryhmän tavoitteet ja toimenpide-ehdotukset

# Tutkimusaineistot - mitä ja kenelle:

## Ratkaistavat kysymykset:

1. Miten selvitetään mitä tutkimusaineistoja yliopistolla on?
  - Luodaanko portaali, johon tutkijat rekisteröivät tutkimusaineistonsa (aihe, laajuus, vastuutaho, rahoittaja, avoimuus ja saatavuus)? (vrt OKM:n suunnittelema infrastruktuuriportaali)
2. Miten arvioidaan tutkimusaineistojen 'arvo' ja miten priorisoidaan niiden säilytyksen ja avaamisen rahoitus?
  - Omistajan eli yliopiston rooli ja vastuutaho?
3. Kansallisten ja kansainvälisten palveluntarjoajien antamat mahdollisuudet
  - Kansallinen tiekartta, FIRI, CSC, ESFRI, Horisonttihankeet
  - Kuka auttaa tutkijaa? Bottom up vai top down?
4. Tiedotus ja koulutus



# Tutkimusaineistot - mitä ja kenelle:

## Tutkimusaineistorekisterit:

1. Miten selvitetään mitä tutkimusaineistoja yliopistolla on?

- CSC:n hakupalvelu ETSIN
- Helsingin yliopiston, Aallon ja Lapin yliopiston käyttämä Tuhat-tietokanta
- Työryhmä ei suosittele Oulun yliopistolle oman palvelujärjestelmän rakentamista



# Tutkimusaineistot - mitä ja kenelle:

## Tutkimusaineistorekisterit:

Etsin ([etsin.avointiede.fi](http://etsin.avointiede.fi)) on CSC:n ylläpitämä tutkimusaineistojen hakupalvelu. Palvelussa voi kuvailla itse tuottamansa aineiston ja saada sille pysyvän tunnisteiden jolla viitata aineistoon. Etsin kerää tutkimusaineistojen metatietoja myös sovitusta ulkopuolisista lähteistä. Kertyneistä kuvailutiedoista muodostuu kansallinen tutkimusaineistojen luettelo. Tätä luetteloa voi kuka vain selata ja tehdä siihen hakuja. Itse aineistoa ei tallenneta Etsimeen, mutta kuvailija voi halutessaan tarjota linkin aineistoon.



# Tutkimusaineistot - mitä ja kenelle:

## Tutkimusaineistorekisterit:

Tutkimusaineiston tiedot saa Etsimeen useaa kautta:

1. Tutkija syöttää itse aineistoa kuvailevat tiedot Etsin-palvelussa. Tutkimusorganisaatio voi ohjeistaa tutkijoitaan toimimaan näin. Organisaatiotason toimijalle voidaan myöntää ylläpito-oikeudet organisaatioon liitettyihin aineistoihin.
2. Tutkija voi tallentaa aineistojensa tiedot tutkimusorganisaation omaan palveluun. Tällainen voi olla esimerkiksi tutkimustietojärjestelmä, josta kuvailutiedot voidaan haravoida Etsimeen. Organisaatiota kehoitetaan ehdottamaan niiden toiminnassa syntyvien kuvailutietojen haravointia Etsimeen takaamaan aineistojen löydettävyyden.
3. Tutkija voi tallentaa aineistonsa tieteenalakohtaiseen tai yleiseen tietoarkistoon, josta kuvailutiedot haravoidaan Etsimeen. Näin aineiston säilyy arkistossa ja on samanaikaisesti löydettävissä Etsimestä. Näin toimitaan jo muun muassa Yhteiskuntatieteellisen tietoarkiston ja Kielipankin kanssa.



# Tutkimusaineistot - mitä ja kenelle:

## Tutkimusaineistojen talletus:

Oulun yliopiston oma arkisto:

Olisi harkittava mahdollisuutta muodostaa isompia tutkimusarkistokokonaisuuksia yliopiston arkiston alaisuudessa ja resursoida jatkuvuus riippumattomaksi jonkin oppiaineen tai professuurin tai yleensä jonkun työsuhteen tai projektin jatkumisesta.

CSC:n palvelut IDA

Kansallisarkisto, Suomen yhteiskuntatieteellinen tietoarkisto

ESFRI-infrastruktuurit





# Tutkimusaineistot - mitä ja kenelle:

## Ratkaistavat kysymykset:

2. Miten arvioidaan tutkimusaineistojen 'arvo' ja miten priorisoidaan niiden säilytyksen ja avaamisen rahoitus?

- Omistajan eli yliopiston rooli ja vastuutaho?
- Tutkimusryhmän johtajan rooli ja vastuu?
- Rahoittajan vastuu
- Arvon määrittämisen kriteerit: potentiaalinen käyttäjäkunta, aineiston tekninen laatu, tieteellinen arvo, avoimuus, aineiston laajuus



# Tutkimusaineistot - mitä ja kenelle:

## Ratkaistavat kysymykset:

### 3. Kansallisten ja kansainvälisten palveluntarjoajien antamat mahdollisuudet

- Kansallinen tiekartta, FIRI, CSC, ESFRI, Horisonttihankeet
- Kuka auttaa tutkijaa? Bottom up vai top down?



# Tutkimusaineistot - mitä ja kenelle:

## Viestejä tutkijoilta; tiedotustarve:

4. Tiedotus ja koulutus www-sivujen, tietoiskujen ja koulutustapahtumien muodossa tulee keskeiseksi osaksi toimenpideohjelmia.

- Jonkinlainen käsikirja / perusohjeet aineistojen talletuksesta. Tutkijoiden keskuudessa on yllättävän paljon epäselvyyttä ja suoranaista tiedon puutetta liittyen aineistojen säilytykseen ja arkistointiin.
- Monilla aloilla kerätään nykyään tutkimusaineistoja esimerkiksi haastatellen, mutta lupien kysyminen arkistointia varten ja ylipäätään koko ajatus arkistoinnista ei edes juolahda tutkijoiden mieleen.
- Tutkimusaineistojen tallettamisen ja avaamisen tärkeydestä pitäisi tietenkin myös tiedottaa, ja tehdä selväksi kaikkien alojen tutkijoille että sen pohtiminen on oleellinen osa tutkimusprosessia.
- Monilla tutkijoilla ei myöskään ole tietoa siitä, minne aineistot ylipäätään voisi tallettaa.
- Resurssoinnin tarve



# Tutkimusaineistot - mitä ja kenelle:

Tavoitetila 2017:

- Tutkimusaineistojen rekisteröinti ja metatietojen tallennus järjestetty
- Tutkimusaineistojen tallennus ja hallinnointi otettu osaksi tulosneuvotteluja
- Tutkimusaineistojen avaaminen: perusteet ja päätöksentekomekanismit luotu
- Julkaisujen rinnakkaistallentaminen järjestetty

Porkkanoita:

Esimerkiksi rekisteröidyt tutkimusaineistot voitaisiin ottaa tulosrahoituksen perustaksi.

