

Aineistojen avaaminen ja avoin julkaiseminen

Tutkimus- ja kehittämisosaaminen 5:
Oppimateriaalit

Johanna Kiviluoto ja Riikka Sinisalo
LAB-ammattikorkeakoulu

Tekstit: CC BY-SA

Kuvat ja kuvat: © tekijät



Oppimistavoitteet

- Opiskelija osaa löytää, tuottaa, hallita ja hyödyntää tutkimusdataa vastuullisesti sekä hahmottaa avoimen ammatillisen ja tieteellisen julkaisemisen kentän ja toimintatavat sekä eri toimijoiden roolit. Opiskelija arvioi lähteiden merkitystä ja analysoi erilaisten aineistojen hyödyllisyyttä.



Sisältö

- Tutkimusaineistojen hallinta (diat 4-25)
- Aineistojen avaaminen (diat 26-43)
- Tulosten julkistaminen ja julkaiseminen (diat 44-62)

Opintojakso on osa Tutkimus- ja kehittämisosaaminen 1–5 -kokonaisuutta, joka on tuotettu yhteistyössä Ammattikorkeakoulujen avoin TKI-toiminta, oppiminen & innovaatioekosysteemi -hankkeessa. Opintojakso on suunnattu ylempää ammattikorkeakoulututkintoa suorittaville.



Osio 1: Tutkimusaineistojen hallinta



Mitä on tutkimusaineisto?

- Tutkimusaineisto koostuu tutkimusdatasta ja sen sisältöä kuvailevista metatiedoista.



Tutkimusdata (1)

- Tutkimusdataa on kaikki aineisto, jota tarvitaan tutkimuksen tulosten perustelemiseen ja toistamiseen sekä kaikki muu muille aineistoon tutustuville mahdollisesti hyödyllinen materiaali.
- Tutkimusaineistoilla tarkoitetaan niitä materiaaleja, jotka tutkimuksen tai hankkeen aikana syntyvät, kuten esimerkiksi
 - kysely- ja haastatteluaineistot
 - erilaiset mittaus- ja havainnointiaineistot
 - video-, kuva-, ääni- ja tekstiaineistot
 - muistiinpanot, päiväkirjat

Lähde: Helsingin yliopiston kirjasto. 2019. Datanhallinnan perusopas: Mitä datasi on? [Viitattu 29.4.2021].

Saatavissa: <https://libraryguides.helsinki.fi/datanhallinta/data>



Tutkimusdata (2)

- Käytännössä dataa voivat olla esimerkiksi
 - tutkimuksen fyysinen tai digitaalinen lähdemateriaali (esimerkiksi arkistomateriaali)
 - tutkimuksen aikana syntynyt data (esimerkiksi mittaustulokset)
 - eri menetelmillä kerätty data (kuten kyselyt, haastattelut, mittaukset, arvioinnit, kuvantamismenetelmät yms.)
 - järjestetyt kokoelmat
 - aineiston koodaus, luettelointi, jäsentely ja kommentointi eri tasoilla
 - analysoidun datan eri versiot
 - fyysiset ja digitaaliset laboratoriopäiväkirjat
 - lähdekoodit ja ohjelmistot
- Tutkijan vastuulla on huolellinen suunnittelu, dokumentaatio sekä aineiston jakaminen parhaaksi katsomallaan tavalla sopimuksia, lakeja ja hyvää etiikkaa noudattaen.

Lähteet:

Helsingin yliopiston kirjasto. 2019. Datanhallinnan perusopas: Mitä datasi on? [Viitattu 29.4.2021]. Saatavissa: <https://libraryguides.helsinki.fi/datanhallinta/data>
ATT-hanke. Avoimen tieteen ja tutkimuksen käsikirja osa 2: tutkijalle ja organisaatiolle. [Viitattu 29.4.2021]. Saatavissa: <https://avointiede.fi/sites/default/files/2020-05/avoimuuden%20k%C3%A4sikirja%20tutkijoille.pdf>

Tutkimusdataa kuvailevat tiedot eli metatiedot (1)

- Jotta aineisto olisi tutkimukseen kelpaavaa, siihen täytyy liittyä tietoja sen alkuperästä. Esimerkiksi datan liitteeksi tarvitaan kuvailevaa ja teknistä tietoa siitä mitä informaatiota se sisältää. Siksi tutkimusdataan liittyy paljon tietoa siitä, miten se on rakenteistettu ja koodattu, miten se on syntynyt ja miten sitä on käsitelty. Nämä tiedot voi tallentaa esimerkiksi tietoarkiston luettelon metatietoihin, koodikirjoihin tai muuhun dokumentaatioon.
- Tutkimusaineistoa kuvaileva metadata on helppointa ja kustannustehokkainta tallentaa vaiheittain aineiston elinkaaren aikana. Puutteellisen metadatan täydentäminen ja korjaaminen vuosia aineistonkeruun jälkeen ei ole välttämättä edes mahdollista tai ainakin maksaa monikymmenkertaisesti sen, mitä maksaisi metadatan tuottaminen aineiston elinkaaren alkuvaiheissa.

Lähteet:

Tietoarkisto. 2017. Aineistonhallinnan käsikirja. Miksi aineistonhallintaa ja jatkokäyttöä? [Viitattu 29.4.2021]. Saatavissa: <https://www.fsd.uta.fi/aineistonhallinta/fi/miksi-aineistonhallintaa-ja-jatkokaytto.html>

ATT-hanke. Avoimen tieteen ja tutkimuksen käsikirja osa 2: tutkijalle ja organisaatiolle. [Viitattu 29.4.2021]. Saatavissa: <https://avointiede.fi/sites/default/files/2020-05/avoimuuden%20k%C3%A4sikirja%20tutkijoille.pdf>



Tutkimusdataa kuvailevat tiedot eli metatiedot (2)

- Tutkimusaineistojen metatietojen perimmäinen käyttötarkoitus on vastata tutkimusaineistoihin liittyviin kysymyksiin. Metatiedoilla kuvataan, mitä tutkimusaineistojen sisällöstä ja olemassaolosta halutaan jatkossa tietää. Tutkimusala ja sen vallitsevat käytänteet vaikuttavat tuotettavan metatiedon määrään. Useilla aloilla on kehitetty juuri kyseisen tutkimusalan aineistojen erityispiirteet huomioivia metatietomalleja. Nämä mallit näkyvät tutkijalle usein aineistojen säilytys- tai hakupalveluiden kautta, jolloin käytettävä palvelu ohjaa tutkijaa metatietojen tuottamisessa. Laadukas metatieto on tutkimusaineiston käyntikortti, jonka tutkija voi antaa tutkimusaineistostaan siitä kiinnostuneille. Tutkijoilla on ratkaisevan tärkeää tietoa omista aineistoistaan, jonka lisäksi tarvitaan tiedonhallinnan ja muita asiantuntijoita säilyvyyden, käytettävyyden ja löydettävyyden varmistamiseksi.

Lähde:

ATT-hanke. Avoimen tieteen ja tutkimuksen käsikirja osa 2: tutkijalle ja organisaatiolle. [Viitattu 29.4.2021]. Saatavissa: <https://avointiede.fi/sites/default/files/2020-05/avoimuuden%20k%C3%A4sikirja%20tutkijoille.pdf>

Tutkimusaineistojen hallinta

- Tutkimusaineistojen hallinnalla tarkoitetaan sitä, että tutkimusaineistot (data) ja niihin liittyvä kuvaileva tieto (metatieto, metadata) on luotu, tallennettu ja järjestetty siten, että aineisto säilyy käyttökuntoisena ja luotettavana ja että tietosuoja ja tietoturva on varmistettu aineiston koko elinkaaren ajan.

Lähde: Tietoarkisto. 2017. Aineistohallinnan käsikirja. Miksi aineistohallintaa ja jatkokäyttöä? [Viitattu 29.4.2021].

Saatavissa: <https://www.fsd.uta.fi/aineistohallinta/fi/miksi-aineistohallintaa-ja-jatkokayttoa.html>

Aineistohallinnan edut (1)

1. Jatkokäytettävyys

- Aineistojen jatkokäyttö on taloudellista ja säästää rajallisia resursseja. Aikaa ja rahaa ei tarvita uuden materiaalin keräämiseen, jos sopiva aineisto on jo saatavilla. Lisäksi laajat kysely- ja haastatteluaineistot sisältävät usein täysin analysoimatonta materiaalia. Jo olemassa olevien aineistojen käytöllä voidaan myös välttää päällekkäiset tiedonkeruut.
- Arkistoidut aineistot mahdollistavat ajalliset vertailut, ja ennen koottujen tutkimusaineistojen valossa voi tarkastella monipuolisia ja vaihtoehtoisia tutkimusongelmia. Lisäksi aineistojen saatavuus voi edistää tutkimusmenetelmien kehittämistä. Kehittyvät menetelmät, ohjelmistot ja teknologia tarjoavat puolestaan uusia mahdollisuuksia tutkimukseen. Arkistoiduista aineistoista on hyötyä myös opetuksessa ja opiskelussa.
- Arvokkaista tutkimusaineistoista ei ole tutkimusyhteisölle ja tulevaisuudessa tehtävälle tutkimukselle hyötyä, jos tiedot aineistoista ovat ensisijaisesti alkuperäisten tutkijoiden mielissä ja häviävät sitä mukaa, kun kyseiset henkilöt siirtyvät muihin tehtäviin.

Lähde: Tietoarkisto. 2017. Aineistohallinnan käsikirja. Miksi aineistohallintaa ja jatkokäyttöä? [Viitattu 29.4.2021].

Saatavissa: <https://www.fsd.uta.fi/aineistohallinta/fi/miksi-aineistohallintaa-ja-jatkokayttoa.html>



Aineistohallinnan edut (2)

2. Rahoittajan vaatimukset

- Aineistohallintasuunnitelma on monen kansallisen ja kansainvälisen tutkimusrahoittajan vaatimus.

3. Osa hyvää tieteellistä käytäntöä

- Yksi tieteellisen tutkimuksen luotettavuuden edellytys on, että tutkimus on suoritettu hyvän tieteellisen käytännön edellyttämällä tavalla. Tutkimusaineistot ovat keskeinen osa tutkimusprosessia ja tutkimushanketta.
- Tutkimuseettisen neuvottelukunnan ohjeiden mukaan hyvään tieteelliseen käytäntöön kuuluu muun muassa, että tutkimuksessa sovelletaan tieteellisen tutkimuksen kriteerien mukaisia ja eettisesti kestäviä tiedonhankinta-, tutkimus- ja arviointimenetelmiä ja että tutkimuksessa syntyneet tietoaineistot tallennetaan tieteelliselle tiedolle asetettujen vaatimusten edellyttämällä tavalla.
- Tutustu hyvään tieteelliseen käytäntöön [Tutkimuseettisen neuvottelukunnan \(TENK\) verkkosivuilla](#) ja katso [video \(1:53\)](#)

Lähde: Tietoarkisto. 2017. Aineistohallinnan käsikirja. Miksi aineistohallintaa ja jatkokäyttöä? [Viitattu 29.4.2021].

Saatavissa: <https://www.fsd.uta.fi/aineistohallinta/fi/miksi-aineistohallintaa-ja-jatkokayttoa.html>



Tutkimuseettinen ennakoarviointi

- Tietyissä ihmistieteellisissä tutkimuksissa edellytetään eettistä ennakoarviointia. Eettisellä ennakoarvioinnilla tarkoitetaan tutkimussuunnitelman arviointia tieteenalakohtaisten eettisten käytänteiden mukaisesti painottuen tutkimuksesta tai sen tuloksista tutkittavalle mahdollisesti koituvan haitan ennakointiin.
- Tutkimuksen eettinen ennakoarviointi on tarkoitettu perustutkinnon jälkeiseen tutkimustyöhön. Perustutkintojen opinnäytteiden ohjaajat vastaavat siitä, että opinnäytetöissä noudatetaan eettisiä periaatteita. Mikäli perustutkinnon opinnäyte suunnitellaan toteutettavaksi tavalla, joka kuuluu ennakoarviointia edellyttävien tutkimusasetelmien ryhmään, opiskelijan tulee hakea ennakoarviointia eettiseltä toimikunnalta yhdessä ohjaajansa kanssa.
- Lue [ohje](#) ennakoarvioinnista ja katso [video \(2:25\)](#)

Lähde: Tutkimuseettinen neuvottelukunta. 2009. Humanistisen, yhteiskuntatieteellisen ja käyttäytymistieteellisen tutkimuksen eettiset periaatteet ja ehdotus eettisen ennakoarvioinnin järjestämiseksi. [viitattu 2.8.2019]. Saatavissa: <https://www.tenk.fi/sites/tenk.fi/files/eettisetperiaatteet.pdf>

Aineistohallintasuunnitelma

- Aineistohallinnan suunnittelu on tärkeä osa tutkimuksen suunnittelua. Sen avulla voidaan varmistaa tutkimusaineistojen järkevä jatkokäyttö. Aineistojen keruun, käytön, säilytyksen ja jatkokäytön suunnittelu selkeyttää myös tutkijan ja tutkimusryhmän omaa toimintaa. Aineistohallinnan suunnittelun tarkoitus on varmistaa, että tutkimusaineistojen suhteen noudatetaan hyvää tieteellistä käytäntöä, että tutkimusaineisto ei vaarannu missään vaiheessa ja että aineiston jatkokäyttö on mahdollista.
- **Aineistohallintasuunnitelmassa** kuvataan miten tutkimusaineisto hankitaan, miten sitä käytetään ja säilytetään tutkimusprojektin aikana ja miten mahdollistetaan aineiston käyttö projektin päätyttyä. Tutkimushankkeen alettua aineistohallintasuunnitelmaa pitää tarkentaa ja olosuhteiden muuttuessa sitä pitää päivittää, mutta tärkeimmät suuntaviivat ja periaatteet täytyy miettiä jo ennen projektin alkua ja viimeistään ennen aineiston keräämistä.
- Katso Research Data Netherlandsin [video \(5:30\)](#) aineistohallintasuunnitelman hyödyistä ja rakenteesta.

Lähde: Tietoarkisto. Aineistohallinnan suunnittelu. Aineistohallinnan käsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. [viitattu 10.5.2021]. Saatavissa: <https://www.fsd.tuni.fi/fi/palvelut/aineistohallinta/aineistohallinnan-suunnittelu/>

Aineisto

- Mitä aineistoa kerätään?
- Miten aineisto kerätään?

Tutkimuskysymykset määrittävät keskeisimmin, mitä aineistoa kerätään. Tyypillisimpiä aineistoja ovat kyselyt, haastattelut, ryhmähaastattelut, kirjoitukset, asiakas- tai asiointitilanteet, viralliset dokumentit, arkistoaineistot, verkkosivustot, rekisteriaineistot ja media-aineistot.

Aineistotyyppi vaikuttaa aineiston keruutapaan. Kvantitatiivisia aineistoja voidaan kerätä haastattelemalla, postikyselyinä, verkkokyselyinä, poimimalla tiedot erilaisista lähdedokumenteista tai mittaamalla. Haastattelut, ryhmähaastattelut ja asiakas- ja asiointitilanteet kerätään yleensä tallentaen ne ääni- tai äänikuvatiedostoiksi. Kirjoitukset kerätään tavallisesti kirjoituspyyntöjen tai -kutsujen avulla sähköpostitse, maapostitse tai tarkoitusta varten luodun verkkosivuston kautta. Virallisista dokumenteista saa nykyisin suuren osan suoraan verkosta, mutta osan dokumenteista saa vain erikseen pyytämällä tai anomalla lupaa niiden tutkimuskäyttöön. Rekisteriaineistojen saaminen edellyttää lupamenettelyä. Ohjeita lupien hakemiseen saa Rekisteritutkimuksen tukikeskuksesta.

Lähde: Tietoarkisto. Aineistonhallinnan suunnittelu. Aineistonhallinnan käsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. [viitattu 10.5.2021]. Saatavissa: <https://www.fsd.tuni.fi/fi/palvelut/aineistonhallinta/aineistonhallinnan-suunnittelu/>

Oikeudet

- Mitkä ovat tutkimusaineiston omistus-, hallinta- ja tekijänoikeudet?
- Kuka antaa käyttöoikeudet aineistoon?
- Miten tutkittavia informoidaan?

Aineistoon voi liittyä tekijänoikeuksia, vaikka suuri osa empiiristä aineistoa jää tekijänoikeuslain suojan ulkopuolelle. Jos aineistoon liittyy tekijänoikeuksia, aineiston tekijänoikeuden haltija määrää aineiston käyttötavat ja aineiston käyttöön on saatava oikeudenhaltijan suostumus. Riippumatta siitä, onko aineisto tekijänoikeudella suojattua vai ei on aineiston tekijöiden määrittäminen tärkeää, sillä arkistoidun aineiston jatkokäyttäjät viittaavat omissa tutkimuksissaan aineistoon ja sen tekijöihin.

Tutkimuksessa tulee aina erikseen sopia aineiston omistamisesta ja käyttöoikeuksista. Aineiston käyttöoikeuksista on hyvä sopia koskien sekä alkuperäistä tutkimusta että sen jälkeistä aikaa. Ennen tutkimussopimuksen laatimista tulee tarkistaa, mitkä ovat tutkimusrahoittajan tutkimusaineistoja koskevat vaatimukset, jotta aineistosta sovitaan rahoittajan edellyttämällä tavalla.

Kun aineistoa kerätään suoraan tutkittavilla, heille annettu informaatio määrittää oleellisesti aineiston käyttöoikeuksia.

Lähde: Tietoarkisto. Aineistonhallinnan suunnittelu. Aineistonhallinnan käsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. [viitattu 10.5.2021]. Saatavissa: <https://www.fsd.tuni.fi/fi/palvelut/aineistonhallinta/aineistonhallinnan-suunnittelu/>

Tietosuoja ja tietoturva (1)

- Miten varmistetaan tietosuoja?
- Minkälaiset oikeudet eri käyttäjäryhmillä on aineistotiedostojen lukemiseen ja käsittelyyn?
- Miten huolehditaan tietoturvasta?
- Miten hoidetaan varmuuskopiointi?

Tutkimuksessa tietosuoja kiteytyy tutkittavia koskevien henkilötietojen käsittelyn suunnitelmallisuuteen ja huolellisuuteen. Henkilötietoja ei tule kerätä ja käsitellä tarpeettomasti ja tiedot tulee suojata niin, etteivät ulkopuoliset pääse niihin käsiksi. Mikäli tutkimusaineiston muodostamiseen tarvitaan henkilötunnusta, tulee sopia erikseen, ketkä ovat oikeutettuja tunnistellisen aineiston käsittelyyn.

Kun osana tutkimusta käsitellään tunnistellisia tutkimusaineistoja, henkilötietolaki edellyttää, että tutkija laatii rekisteriselosteen. Rekisteriseloste lisää henkilötietojen käsittelyn avoimuutta ja se on annettava pyynnöstä tutkittavan nähtäväksi. Verkkopalvelun kautta kerättävän aineiston yhteydessä voidaan käyttää myös tietosuojaselostetta. Tietosuojaseloste on laajennettu rekisteriseloste, joka toteuttaa samalla myös laajemmin henkilötietolain mukaista informointivelvollisuutta. Tutustu tietosuojavaltuutetun laatimiin ohjeisiin. Epäselvyyksien välttämiseksi rekisteriselosteen tulisi sisältää kaikki se tieto, mitä tutkittavalle annetaan aineistosta, sen käsittelystä ja käytöstä.

Tietoturvalla tai tietoturvallisuudella tarkoitetaan tietojen, järjestelmien ja tietoliikenteen suojausta. Sähköisen tutkimusaineiston kopioiminen ja levittäminen on helppoa, samoin esimerkiksi tahaton tuhoaminen tai muuttaminen. Varmuuskopiointi on osa tietoturvaa, mutta lisäksi on varmistettava, että aineiston ulkopuolinen (luvaton) käyttö on estetty.

Lähde: Tietoarkisto. Aineistonhallinnan suunnittelu. Aineistonhallinnan käsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. [viitattu 10.5.2021]. Saatavissa: <https://www.fsd.tuni.fi/fi/palvelut/aineistonhallinta/aineistonhallinnan-suunnittelu/>

Tietosuoja ja tietoturva (2)

- Varmuuskopiointi kannattaa mahdollisuuksien mukaan automatisoida. Automaattinen varmuuskopiointi on tärkeää, sillä varmuuskopion ottaminen unohtuu helposti, jolloin kaikki data saattaa kadota.
- 3-2-1- sääntö on varmuuskopioinnin perusta:
 - vähintään kolme kopiota datasta
 - kahdella eri välineellä
 - yksi kopio on pilvessä
- Suurta tallennuskapasitettia vaativat aineistot kannattaa säilyttää tutkimuksen ajan Opetus- ja kulttuuriministeriön tarjoamassa IDA-säilytyspalvelussa. IDA-säilytyspalvelu on hyödyllinen ja tietoturvallinen ratkaisu myös eri yliopistojen välisille yhteisprojekteille, joissa analysoidaan samaa aineistoa. Palvelu on tarkoitettu Suomen korkeakouluille ja Suomen Akatemian projekteille (akatemiahankkeet, huippuyksiköt, tutkimusohjelmat, tutkimusinfrastruktuurit).

Lähde: Tietoarkisto. Aineistonhallinnan suunnittelu. Aineistonhallinnan käsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. [viitattu 10.5.2021]. Saatavissa: <https://www.fsd.tuni.fi/fi/palvelut/aineistonhallinta/aineistonhallinnan-suunnittelu/>

Tiedostoformaatit ja ohjelmistot

- Mitä ohjelmistoja käytetään aineiston tallentamiseen ja prosessointiin?
- Mitä tiedostoformaatteja ja tallennusmedioita käytetään?

Kvantitatiivisten aineistojen käsittelyyn on olemassa useita tilasto-ohjelmia. Myös laadullisten aineistojen käsittelyyn ja jäsentelyyn on omat ohjelmansa, vaikka esimerkiksi tekstien analyysissä käytetään edelleen usein tavallisia tekstinkäsittelyohjelmia. Valittu ohjelma määrittää käytettävät tiedostoformaatit. Koska ohjelmistot uusiutuvat koko ajan, aineistosta on hyvä tallentaa yksi täysin ohjelmistoriippumaton versio tai ainakin versio, jota useat eri ohjelmistot tukevat.

Aineistotiedostoja voidaan tallentaa ja siirtää optisilla medioilla (esim. CD, DVD, Blu-ray) ja haihtumattoman muistin medioilla (esim. USB-muistitikut). Varmin tapa säilyttää aineistotiedostoja on tallentaa ne magneettisille medioille (esim. kiintolevyt ja magneettiset nauhat).

Lähde: Tietoarkisto. Aineistonhallinnan suunnittelu. Aineistonhallinnan käsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. [viitattu 10.5.2021]. Saatavissa: <https://www.fsd.tuni.fi/fi/palvelut/aineistonhallinta/aineistonhallinnan-suunnittelu/>

Aineiston käsittelyn ja sisällön kuvaus

- Miten varmistetaan aineiston (tekninen) laatu?
- Miten aineiston käsittely dokumentoidaan?
- Miten tallennetaan aineiston keruuta ja sisältöä kuvaava tieto (metadata)?

Aineiston laatuun vaikuttavat tallennusvaiheessa tehtävät sisällölliset ja tekniset ratkaisut. Niitä ovat esimerkiksi tietojen syöttö havaintomatriisiin tai ääni- ja äänikuvatallennuksen tekninen toteutus. Myös aineiston keruun jälkeen tehtävät aineiston käsittelyn ratkaisut vaikuttavat aineiston laatuun. Niitä ovat esimerkiksi aineistotiedostojen nimeäminen ja järjestäminen, puuttuvien tietojen määrittäminen ja muuttujien nimeäminen kvantitatiiviselle aineistolle ja tekstintäminen eli litterointi kvalitatiiviselle aineistolle.

Huolellinen dokumentaatio aineiston keruun vaiheista ja aineiston käsittelystä varmistaa tutkimuksen laadun. Aineiston dokumentointi on elintärkeää myös aineiston pitkäaikaissäilyttämisessä. Aineiston kattava kuvailu eli huolella tehty metadata mahdollistaa aineiston informoidun jatkokäytön ja pitkän elinkaaren.

Lähde: Tietoarkisto. Aineistonhallinnan suunnittelu. Aineistonhallinnan käsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. [viitattu 10.5.2021]. Saatavissa: <https://www.fsd.tuni.fi/fi/palvelut/aineistonhallinta/aineistonhallinnan-suunnittelu/>

Elinkaari

- Mitä aineistolle tapahtuu tutkimuksen päätyttyä?

Tutkimusaineiston käyttöarvo perustuu pitkälti tutkimushankkeen aikana tehtävään aineistohallintatyöhön. Kun se on tehty suunnitelmallisesti, aineisto voi olla hyvin käyttökelpoinen tutkimuksen päätyttyä jatkotutkimuksissa tai menetelmien opetuksessa ja opiskelussa.

Mikäli tutkimusaineisto jää tutkimuksen päätyttyä yksittäiselle tutkijalle, yliopistolaitokselle, tutkimusyksikölle tai organisaatiolle, aineiston haltijan tulee ratkaista aineistohallinnan perusasiat: säilytys, käyttö- ja arkistokopiot, aineistosta tiedottaminen, käyttöehdot ja jakelu.

Kaikkia tutkimusaineistoja ei kannata säilyttää pysyvästi. Arkistointia harkittaessa on arvioitava muun muassa aineiston ainutkertaisuutta, käytettävyyttä, käyttörajoituksia, jatkokäyttöarvoa opetuksessa ja tutkimuksessa sekä arkistoinnin kustannuksia. Arkistoinnin kustannuksia lisää huomattavasti puutteellinen aineistohallinta tutkimuksen aikana, sillä tutkimuksen päätyttyä aineiston perustietojen hankkiminen ja aineiston järjestäminen voi viedä kohtuuttoman paljon aikaa. Aineiston tuhoamisen tulee silti aina olla tietoinen päätös eikä seurausta puutteellisesta tai huolimattomasta aineistohallinnasta.

Lähde: Tietoarkisto. Aineistohallinnan suunnittelu. Aineistohallinnan käsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. [viitattu 10.5.2021]. Saatavissa: <https://www.fsd.tuni.fi/fi/palvelut/aineistohallinta/aineistohallinnan-suunnittelu/>

DMPTuuli ja aineistohallintasuunnitelman hyvät käytänteet

- Aineistohallintasuunnitelma voi tuntua valtavalta urakalta. Se kannattaakin ajatella käytännöllisenä työkaluna, joka nopeuttaa ja helpottaa työskentelyä aineiston keruuvaiheesta analyysiin saakka. Aineistohallintasuunnitelma varmistaa, että kaikki tarvittava tieto on kerätty ja dokumentoitu riittävällä tarkkuudella, kaikki aineiston keruuseen osallistuvat ymmärtävät asiat samalla tavoin ja käyttävät samoja asteikkoja ja määritteitä. Katso [video](#) (4:46, NYU Health Sciences Library) ongelmista, joita aineistohallintasuunnitelmalla olisi voitu välttää.
- Onneksi aineistohallintasuunnitelman kanssa ei tarvitse painiskella yksin, vaan tukea oman suunnitelman laatimiseen saat käyttämällä aineistohallintasuunnittelutyökalua [DMPTuulia](#). DMPTuulista löydät eri rahoittajien [suunnitelmapohjia](#), valmiita [aineistohallintasuunnitelmia](#) sekä ohjeita suunnitelman täyttämiseksi.
- Lue [artikkeli](#) hyvän aineistohallintasuunnitelman laatimisen säännöistä



Aineistohallintasuunnitelman hyvät käytänteet

- Muutamia hyviä käytänteitä aineistohallintasuunnitelman laatimiseen:
 - Informatiivisuus
Suunnitelman lukijoissa on sekä sisällöllisiä että aineistohallinnan asiantuntijoita, kerro riittävästi (muttei liikaa) molemmille lukijakunnille
 - Täsmällisyys
Mieluummin "tutkimuksessa haastatellaan 50 henkilöä" kuin "toteutamme laajan haastattelututkimuksen"
 - Tiiviys
Kerro olennainen tiiviisti
 - Muista, että avattunakin aineisto on sinun keräämäsi ja saat siitä kunnian!



Aineistohallintasuunnitelma ammattikorkeakoulussa (1)

- Ammattikorkeakouluissa tehdään paljon yritysysteistyötä. Kun tutkimusta suunnitellaan, kannattaakin ottaa myös aineistohallinnan näkökulma esille. On sekä tutkimuksen tekijän että toimeksiantajan etu, että asiat on määritelty yhdessä ja molemmat osapuolet tietävät ainakin:
 - mitä dataa kerätään
 - millä menetelmin
 - kenen omistuksessa data on
 - kuka määrää datan jatkokäytöstä
 - missä dataa säilytetään sekä tutkimuksen aikana että sen jälkeen
- Aineistohallintasuunnitelma kattaa nämä kaikki asiat, joten hyvin laadittu aineistohallintasuunnitelma selkeyttää myös toimeksiantajan kanssa käytäviä keskusteluja. Jos tutkimuksen tekeminen edellyttää tutkimuslupaa, ovat aineistohallinnan kysymykset usein osa lupahakemusta, joten aineistohallintasuunnitelman laatiminen jouduttaa myös tätä puolta.



Aineistohallintasuunnitelma ammattikorkeakoulussa (2)

- Jos opintojen yhteydessä kerättyä aineistoa halutaan saattaa muiden tietoon ja/tai käytettäväksi, puhutaan aineistojen avaamisesta. Tutkimuksen julkaisijat ja rahoittajat edellyttävät yhä useammin, että tutkimuksen tausta-aineistot tuodaan mahdollisimman avoimesti saataville viimeistään tutkimuksen päätyttyä, joten tämä on hyvä huomioida esimerkiksi hanketyössä. Tutkimusaineistojen avoimella saatavuudella ja jaettavuudella edistetään tutkimusaineistojen jatkokäyttöä, josta hyötyvät aineiston tuottajan itsensä lisäksi muut tutkijat ja parhaimmillaan yhteiskunnat maailmanlaajuisesti. Aineistojen avaamisessa suositellaan kuitenkin lisensoijaa, joiden avulla tutkija voi itse määrittellä datansa julkisuuden astetta ja käyttöoikeuksia. On myös tärkeä huomioida rahoittajien ja tutkimusorganisaation ohjeistukset ja vaatimukset, lain ja tutkimusetiikan lisäksi.
- Aineistohallintasuunnitelma on keskeinen dokumentti myös aineistojen avaamista suunniteltaessa.
- Ammattikorkeakoulujen rehtorineuvosto Arene on julkaissut [ammattikorkeakoulujen opinnäytteiden eettiset suositukset](#), joissa suositellaan myös aineistojen avaamista tutkimusetiikan ja juridiikan määräämissä rajoissa, jotta TKI-toiminnan avoimuus, laatu, luotettavuus ja näkyvyys paranisi.



Osa 2: Aineistojen avaaminen



Aineistojen avaaminen

- Tutkimusaineiston säilytystarpeen arviointi
 - Kun tutkimusaineistoa ei enää aktiivisesti käsitellä, arvioidaan sen säilytystarve aineistonhallintasuunnitelmassa esitetyn mukaan.
- Seikkoja, joita aineiston säilytystarvetta arvioidessa kannattaa pohtia
 - rahoittajien ohjeet
 - yhteistyökumppaneiden kanssa tehdyt sopimukset
 - eettiset ja lainsäädännölliset tekijät
 - tutkimusartikkelin (tai muun julkaisun) kustantajan vaatimukset julkaisun pohjana oleva aineiston saatavuudesta ja avoimuudesta

Lähde: Seinäjoen ammattikorkeakoulu. 2019. Säilytystarpeen arviointi. Avoin TKI-toiminta: SeAMK opas. [viitattu 4.9.2019].

Saatavissa: <https://seamk.libguides.com/avoinTKI-toimintaSeAMKopas/datansailyttaminenSeAMKData>

Tutkimusaineiston säilytystarpeen arviointi (1)

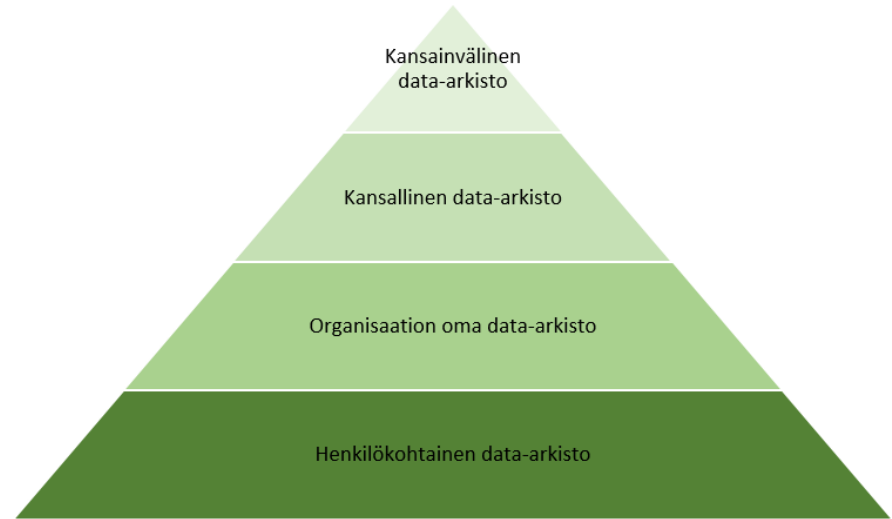
- Edellä mainittujen lisäksi datan säilyttämistä ja avaamista voi arvioida seuraavilla kriteereillä:
 - Aineiston laatu
 - virheettömyys, aitous, väärentämättömyys, sisäinen ristiriidattomuus, kattavuus, ajantasaisuus, oikeellisuus ja käyttökelpoisuus
 - Aineiston arvo
 - tieteellinen, historiallinen, kaupallinen, uutuusarvo, ainutlaatuisuus
 - Aineiston jatkokäytettävyys
 - kaupallinen potentiaali, soveltuvuus uusiin käyttötarkoituksiin tai analyyseihin
 - Kustannukset
 - aineiston säilyttämisen ja toisaalta tuhomaisen kustannukset, uuden vastaavan aineiston keruun kustannukset

Lähde: Seinäjoen ammattikorkeakoulu. 2019. Säilytystarpeen arviointi. Avoin TKI-toiminta: SeAMK opas. [viitattu 4.9.2019].

Saatavissa: <https://seamk.libguides.com/avoinTKI-toimintaSeAMKopas/datansailyttaminenSeAMKData>

Tutkimusaineiston säilytystarpeen arviointi (2)

- Data-aineistoja voidaan avata muiden käyttöön tallentamalla niitä kansallisiin tai kansainvälisiin data-arkistoihin. Tällöin puhutaan datan julkaisemisesta.
 - Tarjolla on sekä yleisiä että tieteenalakohtaisia data-arkistoja. Arkiston valinnassa kannattaa kiinnittää huomiota mm. arkiston ylläpitäjän maineeseen ja luotettavuuteen, turvallisuuteen, säilytysaikaan ja muihin säilytyssehtoihin sekä käytön tilastointiin.
- Vain murto-osa aineistoista on niin merkittäviä, että ne tallennetaan kansainvälisiin data-arkistoihin. Henkilökohtaisissa data-arkistoissakin olevat aineistot kannattaa kuitenkin kuvailla aineistojen hakupalveluihin löydettäväksi, jotta aineistosta kiinnostuneet voivat ottaa yhteyttä aineiston omistajaan. Jo tämä lisää avoimuutta ja tehostaa resurssien käyttöä.



Lähde: Seinäjoen ammattikorkeakoulu. 2019. Säilytystarpeen arviointi. Avoin TKI-toiminta: SeAMK opas. [viitattu 4.9.2019].

Saatavissa: <https://seamk.libguides.com/avoinTKI-toimintaSeAMKopas/datansailyttaminenSeAMKData>

Kuva: <https://seamk.libguides.com/avoinTKI-toimintaSeAMKopas/datansailyttaminenSeAMKData> (CC BY 4.0)

Tutkimusaineiston julkaisemisen edellytykset

(1)

- Tutkimusaineiston julkaisemiseen tulee valmistautua jo tutkimuksen alusta saakka, jotta tarvittavat seikat on huomioitu riittävällä tasolla.

1. Sovi ajoissa datan omistajuudesta

- Jos tutkimuksen aineistohallintasuunnitelma on tehty kattamaan myös datan omistajuuden, voidaan sitä pitää suunnitelman laatimiseen osallistuneiden välillä riittävänä sopimuksena. Jos aineistohallintasuunnitelma ei ole tarpeeksi kattava, tai mukaan on liittynyt uusia toimijoita tai olosuhteet ovat muuttuneet, on sopimuksen tekeminen kannattavaa.
- Kirjallinen sopimus aineiston omistajuudesta kannattaa tehdä jo ennen aineiston keruun aloittamista. Jos käytät jonkun toisen keräämää aineistoa, varmista kirjallisesti, että teillä on aineiston käyttöön tarvittavat oikeudet.
- Sopimus kannattaa tehdä aina kun tutkimuksen vaiheisiin liittyy muita tahoja, esimerkiksi aineiston käsittelijänä tai tutkimuksen toimeksiantajana.
- [Lue lisää sopimuksista >](#)

Lähde: Larmo, K., Kuusniemi, M. E. & Fuchs, S. 2018. Tutkijan muistilista tutkimusdatan julkaisemiseen. Vastuullinen tiede. [Viitattu 10.5.2021].

Saatavissa: <https://vastuullinentiede.fi/fi/jatkokaytto/tutkijan-muistilista-tutkimusdatan-julkaisemiseen>

Tutkimusaineiston julkaisemisen edellytykset (2)

2. Huomioi tutkimusaineiston julkaisupalvelun vaatimukset

- Kannattaa selvittää, missä palveluissa oman alasi aineistoja yleensä julkaistaan ja missä aineistosi saisi parhaan näkyvyyden. Se, että arkisto tarjoaa julkaistulle aineistolle pysyvän tunnisteiden (DOI, URN) helpottaa aineistoon viittaamista.
- Kullakin tutkimusaineistosta sisältävällä palvelulla on erilaiset vaatimukset, joten niiden selvittäminen ja huomiointi alusta saakka helpottaa aineiston julkaisuprosessia. Laadukkaat data-arkistot pitävät huolen siitä, että aineisto on helposti jatkokäytettävissä, ja siksi niillä on vaatimuksia aineiston formaatille ja dokumentaatiolle. Datan pitää olla dokumentoitu siten, että se on ymmärrettävää ja tutkimustulokset ovat datan perusteella helposti todennettavissa. Tässä merkittävässä osassa on tutkimusdataa kuvaavat metatiedot, joita kerätään tutkimuksen edetessä. Metatietoja käsitellään enemmän tutkimusaineistojen hallintaa käsittelevällä oppitunnilla.
- [Esimerkki julkaistusta tutkimusaineistosta >](#)

Lähde: Larmo, K., Kuusniemi, M. E. & Fuchs, S. 2018. Tutkijan muistilista tutkimusdatan julkaisemiseen. Vastuullinen tiede. [Viitattu 10.5.2021].

Saatavissa: <https://vastuullinentiede.fi/fi/jatkokaytto/tutkijan-muistilista-tutkimusdatan-julkaisemiseen>

Tutkimusaineiston julkaisemisen edellytykset (3)

3. Varmista tutkittavien tietosuoja

- Jos aineisto sisältää sensitiivistä informaatiota, kuten arkaluonteisia henkilötietoja, on ennen julkaisua selvitettävä, voiko aineistoa julkaista esimerkiksi anonymisoituna. Anonymisointi tutkimusaineiston osalta tarkoittaa kaikkien niiden tietojen poistamista, jotka mahdollistavat henkilön tunnistamisen. Anonymisointi kannattaa suunnitella huolellisesti ja kirjata ylös tehdyt toimenpiteet perusteluineen. Kun anonymisointi suunnitellaan jo ennen aineiston keruuta, aineiston keruussa voidaan minimoida tunnisteellisen tiedon kerääminen ja suunnitella anonymisoinnin toteutus osaksi tutkimusprosessin kulkua.
- [Lue lisää anonymisoinnista >](#)

Lähde: Larmo, K., Kuusniemi, M. E. & Fuchs, S. 2018. Tutkijan muistilista tutkimusdatan julkaisemiseen. Vastuullinen tiede. [Viitattu 10.5.2021].

Saatavissa: <https://vastuullinentiede.fi/fi/jatkokaytto/tutkijan-muistilista-tutkimusdatan-julkaisemiseen>

Tutkimusaineiston julkaisemisen edellytykset (4)

4. Valitse datalle oikea lisenssi jatkokäyttöä varten

- Jotta tutkimusaineiston käytöstä ei tarvitse sopia jokaisen kanssa erikseen, voidaan aineiston käyttöön liittää valmiit lisenssiehdot. Avoimuutta tukevia lisenssejä on muun muassa [Creative Commons](#) -lisenssit. CC-lisenssillä julkaiseminen tarkoittaa osan oikeuksista tarjoamista käyttäjille sinun määrittelemilläsi ehdoilla. Dataan liittyville algoritmeille, koodeille ja ohjelmistoille sopivat esimerkiksi [MIT](#)- ja [GNU-GPL](#)-lisenssit.
- Monissa data-arkistoissa voidaan tallentaa tietoa eri lisenssiehtojen mukaan, Halutun arkiston käytänteet kannattaakin selvittää jo hyvissä ajoin ennen datan julkaisemista.

Lähde: Larmo, K., Kuusniemi, M. E. & Fuchs, S. 2018. Tutkijan muistilista tutkimusdatan julkaisemiseen. Vastuullinen tiede. [Viitattu 10.5.2021].

Saatavissa: <https://vastuullinentiede.fi/fi/jatkokaytto/tutkijan-muistilista-tutkimusdatan-julkaisemiseen>

Tutkimusaineiston julkaiseminen ja avoimuus

- Tutkimusaineistojen julkaiseminen mahdollistaa tutkimustulosten varmentamisen aineiston avulla, mahdollistaa aineiston jatkokäytön ja tarjoaa aineiston tuottajalle mahdollisuuden meritoitua aineiston kautta. Jotta tutkimusaineiston julkaiseminen on mahdollista, tulee tutkimusaineistolle tehdä aineistohallintasuunnitelma. Aineistohallintasuunnitelmaa ja sen sisältöä on kuvattu Tutkimusaineistojen hallinta -osiossa.

Tutkimusaineiston julkaiseminen ja avoimuus

- Tutkimusaineiston julkaiseminen eroaa jonkin verran esimerkiksi kirjan julkaisemisesta. Tutkimusaineisto voi olla esimerkiksi käytettävissä vain tietyssä tiedostomuodossa tai saatavilla vain tiettyjen ehtojen puitteissa.
- Tutkimusaineistosi voi olla:
 - **Avointa** - Avoin data on dataa, johon kenellä tahansa on pääsy-, käyttö- ja jako-oikeus. Datan pitää olla koneluettavassa muodossa ja kaupallinenkin käyttö on sallittua.
 - **Jaettua** - Jaettu data on usein avointa dataa johon liittyy joitain käyttörajoituksia, esimerkiksi kaupallisen käytön maksullisuus. Jaettua voi olla myös data, johon on pääsy vain tietyillä ryhmillä, esimerkiksi tutkimusalan sisällä.
 - **Suljettua** - Erityisesti henkilötietoja sisältävien tutkimusaineistojen tulee usein olla suljettua. Myös tällaisen datan metatietojen, eli aineiston kuvailutietojen olisi kuitenkin hyvä olla julkisesti saatavilla.



Avoimuuden viiden tähden malli

- Tim Berners-Leen esittämä ns. viiden tähden malli avoimuudelle esittelee, mitä tietoaineiston avaaminen käytännössä edellyttää.



Lähde: Mukaellen Tim Berners-Lee 2009

Lähde: Berners-Lee, T. 2009. Linked data. [viitattu 3.9.2019]. Saatavissa: <https://www.w3.org/DesignIssues/LinkedData.html>

Tutkimusaineistojen hallinnan palvelut

- Opetus- ja kulttuuriministeriö järjestää suomalaisille organisaatioille mahdollisuuden huolehtia tutkimusaineistoistaan Fairdata - palvelukokonaisuuden avulla. Fairdata-palvelut muodostavat yhteentoimivan palvelukokonaisuuden aineistohallintaan, aineistojen kuvailuun ja turvalliseen säilytykseen sekä merkittävien tietoaineistojen pitkäaikaissaatavuuden turvaamiseen.

Mitä tarkoittaa fairdata?

- Kirjainyhdistelmä FAIR viittaa vuonna 2016 julkaistuihin [FAIR periaatteisiin](#), joihin myös Opetus- ja kulttuuriministeriö on sitoutunut.



FAIR -periaatteiden tarkoitus on tehdä tutkimusaineistosta: (1)

– löydettävää (findable)

- aineistolla on pysyvä tunniste, kuten DOI, URN tai Handle ja tunnusten hallinnoinnin tulee olla kestävästi järjestettyä, hallittua ja standardoitua sekä ylläpidettyä niin, ettei tunnus katoa verkosta ja osoittaa aina oikeaan kohteeseen
- aineistolla on kattavat metatiedot
- aineisto kuvaillaan ja rekisteröidään hakupalveluihin
- aineiston pysyvä tunniste käy ilmi metatiedoista

– saavutettavaa (accessible)

- aineisto tai sitä kuvaileva tieto on noudettavissa standardia yhteyskäytäntöä noudattaen
- yhteiskäytäntö on avoin, maksuton ja yleisesti käyttöön otettava ja tukee tarvittaessa tunnistautumista
- metatiedot jäävät saataville, vaikka aineisto on poistettu



FAIR -periaatteiden tarkoitus on tehdä tutkimusaineistosta: (2)

- **yhteentoimivaa** (interoperable)
 - Resurssin ja sen metatiedon sisältö esitetään määrämuotoista, monikäyttöistä, saatavilla olevaa ja jaettavaa kieltä käyttäen, ollen ihmis- ja koneluettavaa sekä järjestelmien välillä siirrettävää
 - Resurssissa ja sen metatiedoissa käytetään avoimia ja koneluettavia sanastoja, ontologioita ja koodistoja, jotka noudattavat FAIR-periaatteita
 - Resurssissa ja sen metatiedoissa on tyyppiteltyjä viittauksia muihin resursseihin, eli esimerkiksi kytkökset aikaisempiin versioihin ja lähteenä käytettyihin aineistoihin ovat esillä
- **uudelleenkäytettävää** (re-usable)
 - Resurssille on annettu runsaasti tarkkoja ja hyödyllisiä metatietoja, joiden avulla voi arvioida myös hyödyllisyyttä
 - Resurssi ja sen metatiedot on julkaistu käyttäen selkeää käyttölisenssiä (kuten koneluettava Creative Commons), jonka ehdot ovat saatavilla
 - Resurssi on aina yhdistettävissä alkuperäänsä ja elinkaareensa
 - Aineisto täyttää oman alansa tiedeyhteisön vaatimukset
- [Lue lisää FAIR periaatteista >](#)

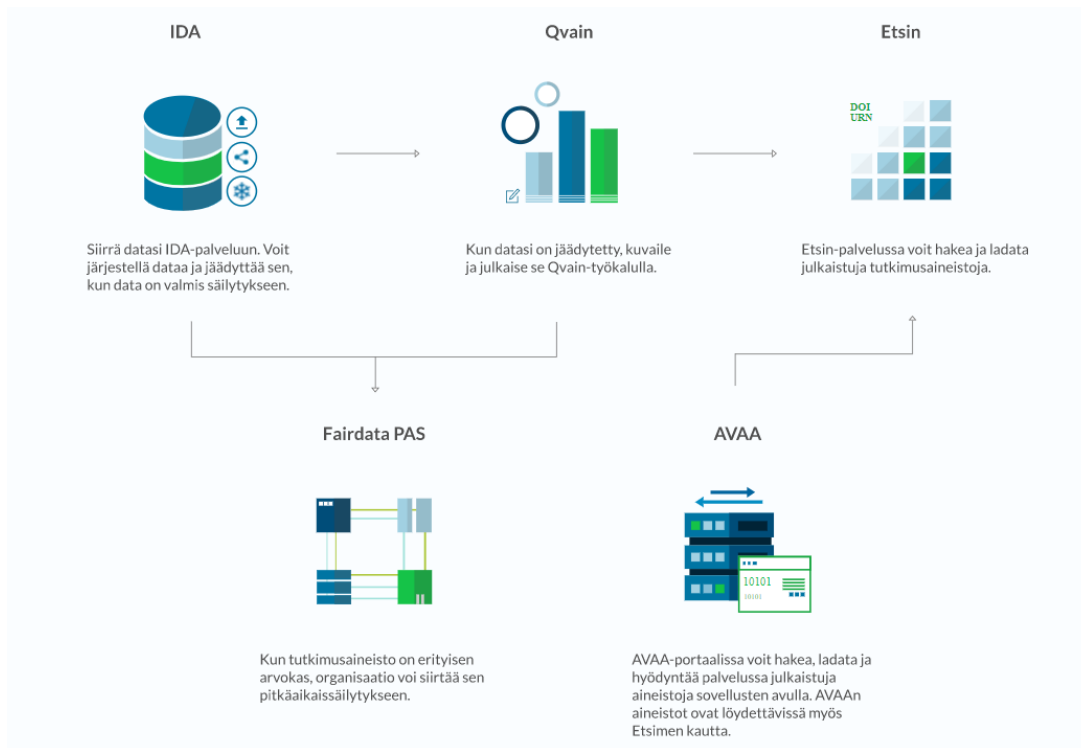


Fairdata palvelut

Fairdata -palvelukokonaisuuteen kuuluvat neljä palvelua, joista kukin huolehtii omalta osaltaan tietoaineistojen säilyttämisen prosessin turvaamisesta. Nämä palvelut ovat

- [IDA](#)
Tutkimusdatan säilytyspalvelu, joka mahdollistaa aineiston turvallisen säilytyksen. [Lue lisää >](#)
- [Qvain](#)
Tutkimusaineistojen kuvailutyökalu, joka auttaa oikean ja riittävän kuvailutiedon liittämisen aineistoon. Tarkka metadata mahdollistaa aineiston löydettävyyden ja varmistaa Fairdata-palveluiden yhteentoimivuuden. [Lue lisää >](#)
- [Etsin](#)
Tutkimusaineistojen hakupalvelu, joka sisältää tutkimusaineistoja kuvailevia tietoja eli metatietoja. [Lue lisää >](#)
- [Fairdata-PAS](#)
Fairdata PAS-palvelu huolehtii tutkimusaineistojen pitkäaikaissäilytyksestä. Pitkäaikaissäilytys tarkoittaa datan säilyttämistä ja luettavana pitämistä kymmenien tai satojen vuosien ajan.
- AVAA Avointen aineistojen julkaisualusta
Pääsy aineistoihin ja työkaluihin, joilla aineistoja voidaan hakea, ladata, hyödyntää ja yhdistellä. [Lue lisää >](#)

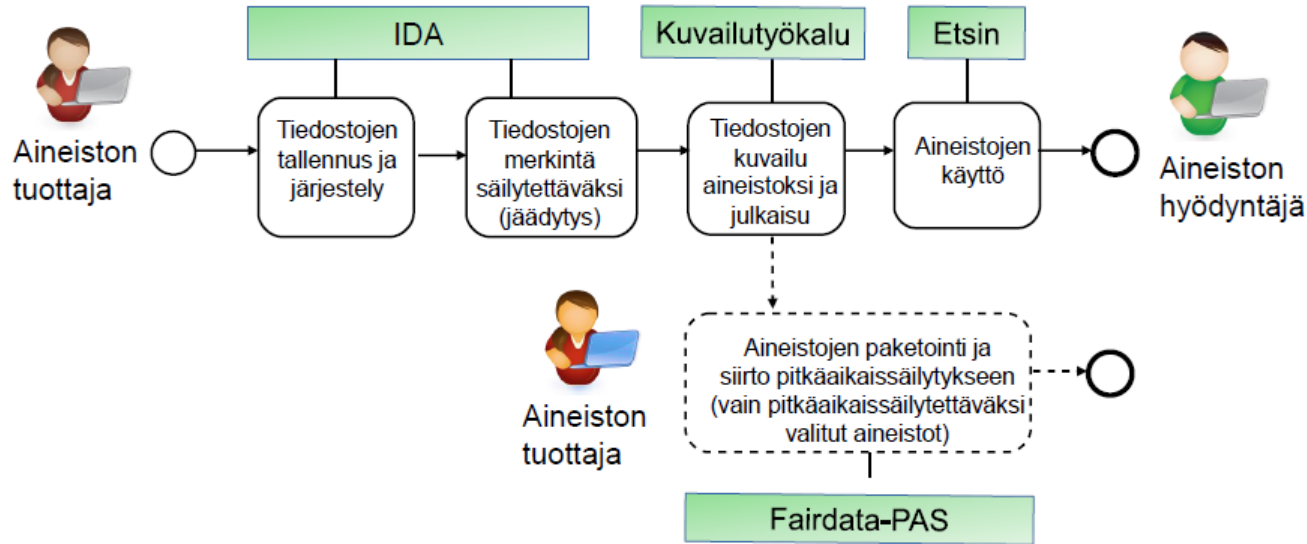
Miten Fairdata-kokonaisuus toimii? (1)



Lähde: Fairdata.fi. 2021. Huolehdi tutkimusaineistostasi. [Viitattu 28.4.2021]. Saatavissa: <https://www.fairdata.fi/#>

Miten Fairdata-kokonaisuus toimii? (2)

- Käytännössä palvelut muodostavat kokonaisuuden, joka näyttäytyy eri toimijoille eri tavoin.



Lähde: Fairdata.fi. 2021. Fairdatan hyödyt. [Viitattu 28.4.2021]. Saatavissa: <https://www.fairdata.fi/tietoa-fairdatasta/fairdatan-hyodyt/>

Valmiit aineistot ja data-aineistoon viittaaminen

- Tutkimusta suunniteltaessa on hyvä huomata, että tutkittavasta aiheesta voi olla jo toisen tutkijan keräämää, uudelleenkäytön mahdollistavaa aineistoa saatavilla. Erityisesti aloittelevalle tutkijalle ja ensimmäistä oppinnytään tekeväille voi olla hyväksi opetella ensin analysoimaan ja tulkitsemaan valmiita aineistoa oman aineiston hankinnan sijaan. Tärkeintä on arvioida, onko aineisto soveltuva omaan tutkimukseen. Valmiita aineistoja käytettäessä tutkimusprosessi voi edetä joko omista ongelmista valmiin aineiston etsimiseen tai päinvastoin: jokin valmis aineisto voi kirvoittaa mielikuvitusta ja saada tutkijan kehittämään uusia tutkimusongelmia. (Saaranen-Kauppinen & Puusniekka 2006)
- Valmiita tutkimusaineistoja voit etsiä mm. seuraavista palveluista:
 - [Tiedejatutkimus.fi](https://www.tiede.fi/tutkimus)
 - [Etsin](https://www.etsin.fi)
 - [FigShare](https://www.figshare.com)
 - [Google Dataset Search](https://www.google.com/datasets)
 - [Zenodo](https://zenodo.org)
 - [Data.europa.eu](https://data.europa.eu)

Lähde: Saaranen-Kauppinen, A. & Puusniekka, A. 2006. KvaliMOTV - Menetelmäopetuksen tietovaranto. Tampere: Yhteiskuntatieteellinen tietoarkisto. [viitattu 5.9.2019].
Saatavissa: https://www.fsd.uta.fi/menetelmaopetus/kvali/L6_6.html

Data-aineistoon viittaaminen

- Kuten kaikkiin tutkimuksessa käytettyihin lähteisiin, myös tutkimusaineistoihin tulee viitata. Niinpä valmista tutkimusaineistoa tai sen osia käytettäessä on viittaus tehtävä aina asianmukaisesti. Jotkut tutkimusaineistopalvelut laativat tallennettuihin aineistoihin malliviittauspohjan, jota voidaan sitten muutella kulloisenkin viittaustyylin mukaiseksi.
- Dataviittauksen pakolliset osat:
 - **Pysyvä tunniste**, jonka avulla data löytyy verkosta (HTTP, URI)
 - **Tekijä(t)** eli tiedot tuottaneet henkilö(t) tai yhteisö(t)
 - **Julkaisuaika**, eli aika, jolloin aineisto on tallennettu arkistoon
 - **Otsikko**, eli aineiston nimi
 - **Organisaatio**, eli arkisto tai tietokanta, jonne aineisto on tallennettu
- Lisäksi valinnaisia tietoja ovat versiotiedot, resurssityypin määrittely (kuten tiedostomuoto), lisenssiehdot, embargo, ORCID (tutkijatunniste)
Esimerkki: Männistö, Ruut (Tampereen yliopisto): Tampereen yliopiston opiskelijoiden henkilökohtainen tiedonhallinta 2017. Versio 1.0 (2019-07-09). Yhteiskuntatieteellinen tietoarkisto [jakaja]. [viitattu 5.9.2019]. Saatavissa: <http://urn.fi/urn:nbn:fi:fsd:T-FSD3249>
- Lue lisää viittaamisesta kansallisista suosituksista [dataviitteiden tekemiseen >](#)
- Lisää esimerkkejä aineistoihin viittaamiseen eri tilanteissa on saatavilla [tietoarkiston sivuilta >](#)



Osa 3: Avoin julkaiseminen



Avoin julkaiseminen



- Avoimella julkaisemisella (open access, OA) tarkoitetaan erilaisten tieteellisten ja tutkimuksellisten julkaisujen kuten artikkelien, artikkelikokoelmien ja tutkimusraporttien vapaata saatavuutta. Julkaisu on avoin silloin, kun se on **Internetin kautta kokonaisuudessaan ilmaiseksi ja esteettömästi** kenen tahansa luettavissa, tulostettavissa ja kopioitavissa.
- **Avoimen julkaisemisen hyödyt**
 - Tutkimusjulkaisujen avoimuus on voimakkaassa kasvussa monestakin syystä. Avoin julkaiseminen mm.
 - lisää tutkimuksen näkyvyyttä, vuorovaikutusta, laatua ja vaikuttavuutta
 - edistää yritysten toimintamahdollisuuksia ja innovaatioita
 - edistää tutkimustietoon perustuvaa päätöksentekoa (aluevaikuttavuus)
 - mahdollisuus tiivistää korkeakoulujen ja elinkeinoelämän keskinäistä suhdetta ja vahvistaa yhteistyötä (mm. innovaatioiden kaupallistaminen)
 - täyttää rahoittajien vaatimukset (mm. Business Finland, Suomen Akatemia ja Horizon 2020 edellyttävät avointa julkaisemista)
 - tuo meriittiä ja näkyvyyttä myös tekijöilleen



Avoimuus lisää myös tasavertaisuutta

- Avoimessa julkaisemisessa on myös tasavertaisuusnäkökulma: perinteisten suurien tiedekustantamojen julkaisemat lehdet ovat kalliita, ja maksumuurin takana olevan yksittäisen artikkelin hinta voi olla useampi kymmenen euroa. Yliopistojen ja ammattikorkeakoulujen kirjastot maksavat vuosittain näistä lisensoiduista lehtipaketeista suuria summia, ja tilausmaksut ovat olleet jatkuvassa kasvussa. Tämä on pakottanut monia kirjastoja supistamaan tarjontaansa. Lisenssien takana olevat aineistot ovat myös käyttöoikeuksiltaan rajattuja - niitä voivat etäkäyttää korkeakoulujen opiskelijat ja henkilöstö, mutta kaikkien muiden on tultava kirjastoon paikan päälle.
- Avoin julkaiseminen tarjoaa mahdollisuuden kenelle tahansa tutustua uusimpaan tutkimukseen, omasta taustasta ja varallisuudesta riippumatta. Tällä on merkitystä myös kehittyvien maiden korkeakouluille ja tutkimuslaitoksille, joilla ei välttämättä ole varaa kalliisiin tilausmaksuihin. Avoin julkaiseminen edistää siten omalta osaltaan tieteen ja tutkimuksen tekemistä.



Avoimuuden hyötyjä



CC-BY Danny Kingsley & Sarah Brown

Avoimen julkaisemisen reitit (1)

- Kun olet päättänyt julkaista tutkimusjulkaisusi avoimesti, on sinun seuraavaksi mietittävä sopiva julkaisukanava. Avoimeen julkaisemiseen on valittavana kolme erilaista reittiä.
- **Kultainen OA (Gold OA)**
 - Suositeltavin tapa on julkaista artikkelinsa lähtökohtaisesti avoimessa julkaisussa, jolloin puhutaan ns. kultaisesta reitistä. Nämä julkaisut ovat kokonaisuudessaan avoimia, mutta perivät usein kirjoittajilta maksuja, mikä aiheuttaa tutkijalle/tutkimusryhmälle/organisaatioille kustannuksia. Kirjoittajamaksut (APC eli article processing charge) vaihtelevat muutamasta kymmenestä eurosta tuhansiin euroihin. On kuitenkin paljon myös lehtiä, jotka eivät peri kirjoittajamaksua lainkaan. Kirjoittajamaksua tai ei, on artikkeli ilmestymisensä jälkeen ilmaiseksi kenen tahansa käytettävissä kustantajan tai muun välittäjän verkkopalvelussa.



Avoimen julkaisemisen reitit (2)

- **Vihreä OA eli rinnakkaisjulkaiseminen (Green OA)**
 - Julkaisu tai sen osa voidaan kustantajan luvalla julkaista myös organisaation omassa tieteenala- tai organisaatiokohtaisessa julkaisuarkistossa, josta se on vapaasti saatavilla joko heti tai ennalta määrätyn viiveen eli embargoajan jälkeen. Tällöin puhutaan rinnakkaisjulkaisemisesta tai rinnakkaistallentamisesta. Kustantaja usein määrittelee, minkä version artikkelista saa rinnakkaisjulkaista: useimmiten sallittu versio on post-print (myös final draft tai author-accepted version), joka on viimeinen kirjoittajalta kustantajalle lähtenyt ja mahdollisen vertaisarvioinnin jo läpikäynyt korjattu versio, mutta jossa kustantajan taittotyö ei vielä näy. Toisin sanoen tietosisältö on sama kuin lehdessä julkaistavassa versiossa, mutta ulkoasu ja esim. sivunumerointi yleensä jonkin verran poikkeavat. Rinnakkaistallentamisen oikeuksia voi tarkistaa [Sherpa/Romeo palvelusta](#) sekä kustantajien ja lehtien verkkosivuilta.
 - Rinnakkaisjulkaiseminen on ilmaista sekä kirjoittajalle että lukijoille. Suomessa ammattikorkeakoulut käyttävät rinnakkaisjulkaisemisessaan yleisesti Theseusta, ja yliopistoilla on kullakin omat julkaisuarkistonsa.



Avoimen julkaisemisen reitit (3)

- **Hybridi OA**

- Hybridilehti on eri asia kuin kultaisen reitin lehti, jossa kaikki artikkelit ovat avoimesti saatavana. Hybridissä on kyse suurten tiedekustantajien (mm. Elsevier, Springer, SAGE) keksinnöstä, joka mahdollistaa yksittäisen artikkelin ostamisen avoimeksi muuten tilausmaksulliseen lehteen. Tässäkin tapauksessa maksajana on kirjoittaja APC-maksujen noustessa yleensä tuhansiin euroihin. Hybridijulkaisuissa on kyse tuplarahastuksesta (double dipping), sillä kustantaja rahastaa sekä tilausmaksuilla että kirjoittajamaksuilla.

- Katso [video](#) avoimesta julkaisemisesta



Yhteisöpalvelut: avointa, muttei aina luvallista

- Artikkeleita löytää usein avoimesti julkaistuina myös ResearchGaten ja Academia.edun kaltaisista tutkijoiden yhteisöpalveluista. On kuitenkin syytä pitää mielessä, että nämä eivät ole virallisia rinnakkaisjulkaisukanavia, sillä ne eivät mm. täytä pysyvyyden vaatimusta, ja julkaisuoikeuksienkin kanssa on monesti vähän niin ja näin - vielä eilen saatavilla ollut artikkeli saattaakin siis yllättäen olla kustantajan vaatimuksesta yhtäkkiä kadonnut.
- Mikäli Google Scholarissa hakuusi osuu esim. ResearchGatesta löytyvä artikkeli, kannattaa se napata itselleen talteen ja artikkeliin viitatessa merkitä lähdeluettelon tiedot alkuperäisen julkaisun mukaan. Käytä aina myös saatavuustiedoissa artikkelin pysyvää tunnistetta kuten DOI tai URN.



Avoimen julkaisemisen reitit (4)

Kultainen

- kaikki artikkelit heti avoimesti saatavilla
- voivat periä kirjoittajilta maksun julkaisemisesta (APC=article processing charge)
- APC -maksut yleensä maltillisia, paljon myös kirjoittajille täysin maksuttomia julkaisuja

Vihreä

- rinnakkaistallentaminen tai rinnakkaisjulkaiseminen
- organisaatioiden omat julkaisuarkistot (esim. Theseus)
- saatavuudessa usein viive johtuen kustantajan määrittelemästä riiputusajasta (embargo)

Hybridi

- tilausmaksullinen lehti, johon voi ostaa yksittäisiä artikkeleita avoimiksi
- APC –maksut usein tuhansia euroja
- ongelmana tuplarahastus: kustantaja kerää sekä tilausmaksuja että APC -maksuja

Avoimuus, tekijyys ja käyttöoikeudet

- Avoimeen julkaisemiseen kuuluu tärkeänä osana myös sekä tekijänoikeudet että käyttöoikeudet: vapaan saatavuuden lisäksi on myös tärkeää määrittää, mitä vapaasti saatavilla olevalla aineistolla saa tehdä. Etenkin TKI -toiminnassa hankkeiden tuloksia julkaistaan yhä enemmän myös muussa muodossa kuin perinteisinä julkaisuina: näitä voivat olla mm. videot ja podcastit. Hankkeilla voi olla myös esimerkiksi oma blogi, jossa kerrotaan hankkeen kuulumisista ja tuloksista. Tällöin on mietittävä, millä lisenssillä näitä tuotoksia halutaan jakaa. Perinteinen "all rights reserved" -ajattelu istuu huonosti avoimuuteen, minkä vuoksi sen rinnalle on kehitetty vaihtoehtoisia tapoja lisensoida aineistoja.



Creative Commons -lisenssit



- [Creative Commons eli CC](#) on näistä kenties laajimmin käyttöönotettu, ja sen eri versiot ovat käytössä yhä useammassa avoimessa julkaisussa. Creative Commons -lisensseistä tuli Suomessa julkishallinnon standardilisenssi jo vuonna 2014, ja keväällä 2019 EU-komissio päätti että myös sen omat julkistetut dokumentit, valokuvat, videot, raportit, tutkimukset ja tietoaaineistot julkaistaan CC-lisensseillä.
- Tässä kohtaa on hyvä huomata, että avoimuus ei kuitenkaan tarkoita tekijänoikeuksista luopumista: tekijänoikeus ja sen myötä myös oikeus tulla mainituksi (ns. isyysoikeus) säilyy aina tekijällä itsellään. CC-lisenssien avulla voi kuitenkin osan oikeuksista tarjota kenelle tahansa teoksen käyttäjälle, ja nämä oikeudet voi itse määritellä valitsemalla CC-lisensseistä parhaiten sopivan version.



Miten valita sopiva CC -lisenssi? (1)

- CC -lisenssit koostuvat neljästä yhdisteltävissä olevasta kriteeristä:
 - **Nimeä (BY)** Tekijän nimi on mainittava. Teosta käytettäessä on ilmaistava sen tekijä tai nimimerkki. Nimeä BY -lisenssi on aina osa jokaista lisenssiyhdistelmää.
 - **Ei kaupallinen (NC)** Ei kaupalliseen käyttöön. Tekijänoikeudenhaltijalta on pyydettävä lupa, jos teoksen käytön tavoitteena on taloudellinen hyöty.
 - **Ei Muutoksia (ND)** Ei saa muokata. Tekijänoikeudenhaltijalta on pyydettävä lupa teoksen muuntamiseen tai hyödyntämiseen vain osittain.
 - **Ja samoin (SA)** Jaettava samalla lisenssillä. Muunnellun johdannaistyön voi jakaa ainoastaan samalla lisenssillä kuin alkuperäisessä teoksessa.



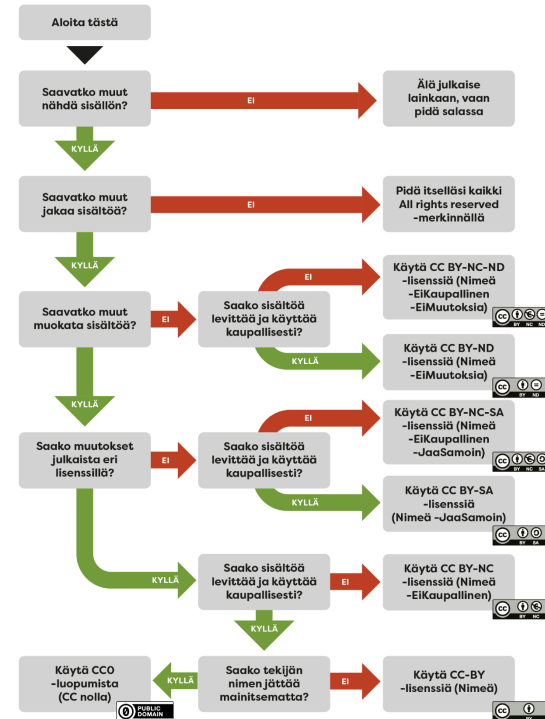
Miten valita sopiva CC -lisenssi? (2)

- Kriteerien yhdistelmästä muodostuu kuusi erilaista lisenssiä, joista valita.
- Valinnassa auttaa myös [CC-lisenssivalitsin](#).

Lähde: [ImagOA-opas](#) (CC BY)



Lisenssin valinta



Kuvien käyttö avoimissa julkaisuissa

- Tekijänoikeudet on huomioitava myös silloin, kun haluat käyttää julkaisuissasi muuta kuin itse tekemääsi kuvitusta. Internet on täynnä kuvia ja erilaisia kuvapankkeja, mikä ei kuitenkaan automaattisesti tee kuvista vapaasti käytettäviä: kiinnitä siis aina huomiota kuvan käyttöä mahdollisesti rajoittaviin tekijänoikeusmerkintöihin.
- Turvallisinta on pelata varman päälle ja etsiä lähtökohtaisesti CC-lisenssoituja kuvia. [Googlen kuvahaussa](#) voit rajata hakutuloksia käyttöoikeuksien mukaan (Työkalut -> Käyttöoikeudet). Voit myös etsiä suoraan CC-lisenssoituja kuvia esimerkiksi [CC Search](#) -hakukoneella.
- ImagOA -oppaaseen on koottu [kattava listaus avoimista kotimaisista ja ulkomaisista kuva-aineistoista](#), joista voi etsiä sopivia kuvia omassa julkaisussa käytettäväksi. Kuvapankit sisältävät myös paljon CC-lisenssoituja kuvia, mutta kannattaa kuitenkin aina tutustua kunkin palvelun ja yksittäisen kuvan käyttöoikeuksiin tarkasti, jotta määritellyt oikeudet riittävät omaan käyttötarkoitukseesi. Kiinnitä siis huomio lisenssien yhteensopivuuteen: jos olet julkaisemassa aineistoa esimerkiksi CC BY -lisenssillä, voit käyttää muiden kuvia, joissa on lisenssinä joko tuo sama CC BY tai sitten CC0/Public Domain.
- Creative Commons -lisenssoituja aineistoja käytettäessä on aina mainittava tekijä, lisenssi ja verkko-osoite, josta se on otettu. Tiedot voi lisätä kuvan yhteyteen kuvatekstinä, tai useamman kuvan ollessa kyseessä laatia kuvista erillinen kuvaluettelo julkaisun loppuun lähdeluettelon tapaan.
- Tutustu tarkemmin kuvien käyttöön ja CC -lisensseihin Aalto yliopiston ylläpitämästä [ImagOA -oppaasta](#). Katso myös video [Creative Commons licences explained](#) (5:32).



Predatorit eli saalistajajulkaisut

Varo saalistajia!

- Avoin julkaiseminen edistää tiedettä ja tutkimusta monin tavoin. Kuitenkin myös sillä on valitettavasti kääntöpuolensa. Tieteellinen kustannus- ja julkaisutoiminta on isoa ja kannattavaa liiketoimintaa, ja kun se yhdistetään yhä kasvavaan vaatimukseen avoimuudesta, on muodostunut valmis markkinarako myös vähemmän rehellisille toimijoille. Näistä puhutaan yleisesti saalistavina open access -kustantajina (predatory open access publishing), ja ne hyödyntävät toiminnassaan kultaista OA-julkaisemista. Saalistajakustantajien liiketoimintamalli on yksinkertainen: ne perivät kirjoittajilta kohtuuttomia APC-maksuja ja julkaisevat niitä vastaan oikeastaan mitä tahansa.



Predatorit eli saalistajajulkaisut (1)

- Epäilyttävän OA-julkaisijan tunnusmerkkejä
 - sekä kustantajien että heidän julkaisemiensa lehtien nimet ovat usein harhaanjohtavia, eivätkä nimet vastaa sisältöä
 - kustantajan instituution takaa paljastuu usein yksityisosoite (esim hotmail tai gmail)
 - puutteellinen vertaisarviointijärjestelmä, tai ei järjestelmää ollenkaan
 - kalastelevat artikkeleita lähettämällä viestejä suoraan tutkijoiden sähköpostiin
- Mikäli siis saat sähköpostiisi viestiä kustantajalta, joka on kiinnostunut julkaisemaan esimerkiksi konferenssiesityksesi, ole tarkkana. Useimmat huomaa huijaukseksi ja roskapostiksi jo pelkän viestin tyylin ja kieliasun perusteella, mutta osa saattaa olla hyvinkin vakuuttavia. Epäilyn herätessä kannattaa kustantaja googlettaa, sillä vältettävistä ja epämääräisistä OA-kustantajista löytyy helposti tietoa.
- Lue tunnusmerkeistä lisää Anna-Sofia Ruthin artikkelista [Kuinka välttää saalistajajulkaisuja?](#)



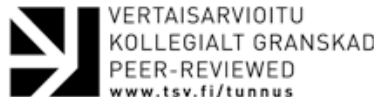
Predatorit eli saalistajajulkaisut (2)

- Seuraavissa videoissa perehdytään vielä tarkemmin sekä saalistajajulkaisuihin että niiden tunnistamiseen.
 - Video 1: [What are predatory publishers](#) (1:18)
 - Video 2: [Identifying predatory publishers](#) (2:14)
 - Video 3: [8 ways to identify a questionable open access journal](#) (3:46)



Miten löydät luotettavia OA -julkaisuja

- Kotimaiset avoimet tiedejulkaisut
 - Kotimaisia avoimia tiedejulkaisuja löydät kätevimmin [Journal.fi - Suomalaiset tiedelehdet verkossa](#) -sivustolta, jota ylläpitää Tieteellisten seurain valtuuskunta TSV. Sivustolla on tällä hetkellä noin 75 suomalaista tieteellistä lehteä ja vuosikirjaa eri aloilta.
 - Suomessa on käytössä myös oma tunnus vertaisarvioituille tiedejulkaisuille. Tämä [vertaisarviointitunnus](#) on tarkoitettu tiedekustantajille, jotka haluavat merkitä julkaisemansa vertaisarvioitua kirjat ja artikkelit, ja tunnuksen myöntää TSV. Käyttöoikeuden myöntäminen edellyttää mm., että tieteelliset julkaisut läpikäyvät vähintään kahden riippumattoman asiantuntijan suorittaman ennakoarvioinnin.



Kansainväliset avoimet tiedejulkaisut (1)

- Suomen rajojen ulkopuolelle siirryttäessä avoimesti saatavilla olevien tiedejulkaisujen määrä kasvaa moninkertaiseksi, onhan englantia yleisimmin tiedejulkaisemisessa käytetty kieli. Samalla kasvaa myös erilaisten hakukoneiden ja tietokantojen määrä. Seuraavassa tutustutaan niistä kahteen yleisesti käytettyyn.
 - **Google Scholar**
 - Scholar on akateeminen Google, jonka hakutulokset keskittyvät tieteellisten kustantajien, tutkimuslaitosten ja korkeakoulujen avoimiin julkaisuihin. Sitä kannattaa hyödyntää niin kotimaisten kuin kansainvälistenkin julkaisujen etsimisessä. On tosin hyvä muistaa, että suomenkielisillä hakusanoilla tulokset painottuvat enemmän opinnäytetöihin: tämä johtuu yksinkertaisesti siitä, että niitä on avoimesti eniten saatavilla. Kannattaakin siis suosia mieluummin englanninkielisiä hakuja aina kuin mahdollista.
 - Katso tästä, miten teet ja rajaat hakuja Google Scholarissa: [Searching with Google Scholar](#) (6:41)



Kansainväliset avoimet tiedejulkaisut (2)

- **DOAJ - Directory of Open Access Journals**

- DOAJin kautta on haettavissa ja luettavissa tällä hetkellä yli 13 700 korkeatasoisen tiedejulkaisun sisällöt. Julkaisuja on 130 maasta, ja ne kattavat kaikki tieteenalat.
- Katso tästä, miten teet ja rajaat hakuja DOAJ - tietokannassa: [Directory of Open Access Journal \(DOAJ\)](#) (3:27)



Avoimet julkaisuarkistot

- Kuten aiemmin on jo mainittu, on rinnakkaisjulkaiseminen yksi väylä avoimeen julkaisemiseen. Yleensä yliopistoilla, korkeakouluilla ja muilla tutkimusorganisaatioilla on omat, organisaatiokohtaiset julkaisuarkistonsa (open repository). Vaikka rinnakkaisjulkaisuja tarttuu haaviin myös Google Scholarin avulla, voi joskus olla tarpeen etsiä suoraan jonkin tietyn organisaation julkaisuja.
 - [OpenDOAR - Directory of Open Access Repositories](#) on kansainvälinen julkaisuarkistojen hakemisto, jonka avulla voit joko paikallistaa jonkin tietyn organisaation arkiston, tai hakea arkistoja esim. tieteenalan, maan tai julkaisutyypin perusteella.
 - Kotimaisten ammattikorkeakoulujen opinnäytetyöt ja muut julkaisut löytyvät keskitetysti [Theseuksesta](#).



Alternative Access: vaihtoehtoisia väyliä artikkeleihin

- Avoimen julkaisemisen jatkuvasta lisääntymisestä huolimatta valtaosa tieteellisistä artikkeleista on edelleen maksumuurien takana. Jos löysit kiinnostavan artikkelin, mutta sitä ei löydy oman korkeakoulusi kirjaston tietokannoista, kokeile josko artikkelista olisi jostain saatavilla final draft -versio. [Google Scholar](#) on tässäkin mainio apu, mutta voit kokeilla myös oheisia, artikkelin DOI-numeroa hyödyntäviä palveluja.
- Unpaywall ja DOAI paikantavat artikkelista mahdollisesti saatavilla olevat ilmaisversiot. Open Access Button toimii kuten edelliset, mutta lisäksi sitä kautta voit lähettää OAButton -tiimille pyynnön artikkelista jonka toivoisit olevan ilmaiseksi saatavilla.
 - [Unpaywall](#)
 - [DOAI](#)
 - [Open Access Button](#)
- Nämä ja muut vaihtoehtoiset tavat päästä lukemaan maksullisia tutkimusjulkaisuja löytyvät koottuna FinElibin [Alternative Access -sivulta](#).

