

Matkalla

DATA- informaatiolukutaitoiseen TIEDEYHTEISÖÖN

*Avoimen tieteen päivät 19.-20.11.2018 /
Jukka Rantasaari*



**TURUN
YLIOPISTO**

MIKSI?



*Tarve olisi yksiselitteiselle, mutta ei pakottavalle ohjeistukselle, miten laadullinen tieto hallitaan, sen säilyttäminen tutkimuksen aikana ja tutkimuksen jälkeen. Saako säilyttää tutkimuksen jälkeen, miltä osin? Ei vain tekninen vaan myös eettinen näkökulma. Jokainen on omassa päässään miettinyt nämä asiat.”
(Tohtorikoulutettava)*



Suurimmat puutteet osaamisessa lienevät siinä, että kertynyttä dataa ei prosessoida ajoissa ja/tai se kerätään puutteellisin tiedoin, joita ei merkitä tarpeeksi hyvin ylös, jolloin analysointi vaikeutuu.” (Ohjaaja)



- DATAPOLITIikka
- AVOIMEN TIETEEN LINJAUKSET
- KUILU NYKYISEN VS. TAVOITELLUN OSAAMISEN VÄLILLÄ

HYVÄN DATANHALLINNAN VAIKUTUKSET

- * Vähemmän virheitä, tehokkaampi ajankäyttö, helpottaa jakamista
 - Mahdollistaa datan uudelleenkäytön, jakamisen ja avaamisen
- * Parantaa projektin yhteensopivuutta rahoittajien mahdollisten datanhallintasuunnitelman ja jakamisen vaatimusten suhteen
- * Lisää tutkimuksen näkyvyyttä ja vaikuttavuutta
- * Auttaa ammatillisessa ansioitumisessa
- * Auttaa menestymään myös yliopiston ulkopuolisessa työelämässä
- * (datan hankkiminen, arvioiminen ja käsittely käytäntöjen parantamisen, palvelujen kehittämisen ja päätöksenteon apuna)

TARPEET

➤ KOULUTUKSET
➤ PALVELUT

Tavoite:

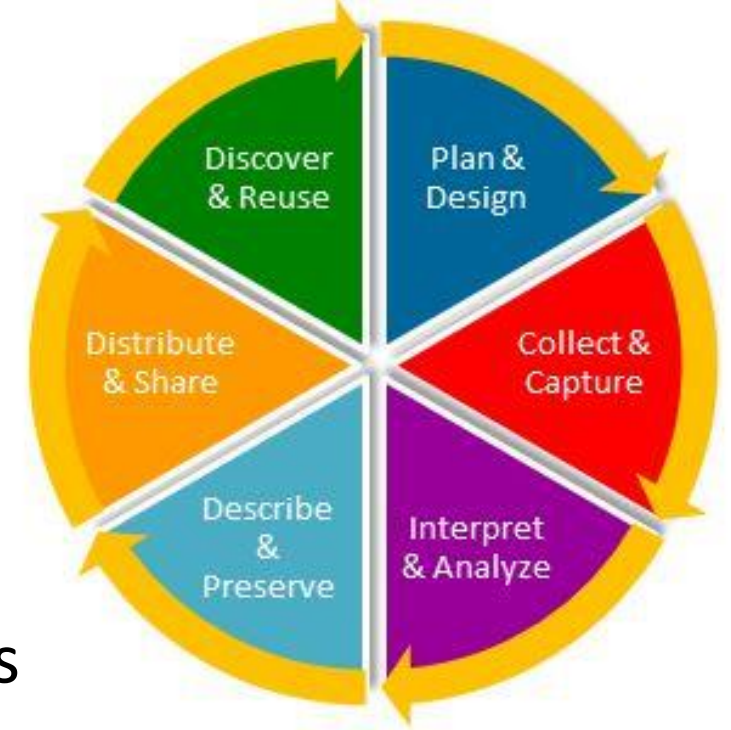
- infra, palvelut ja koulutukset tutkijoiden tarpeisiin

Mitä ovat tutkijoiden tarpeet?

- Tohtoriopiskelijoiden ja heidän ohjaajiensa haastattelut:
 - Kuinka hankit, hallitset, tallennat ja säilytät datasi?
 - Millaiset käytännöt, osaamiset ja tarpeet?

TEEMA-ALUEET

- * Projektissa kerättävä data ja sen elinkaari
- * Sopimukset, lupa-asiat
- * Datan versionhallinta, varmuuskopiointi ja tallennus
- * Prosessointi, analysointi ja visualisointi
- * Organisointi, dokumentointi, kuvailu, laadun seuranta
- * Ulkoisen datan hankinta ja käyttäminen
- * Immateriaalioikeudet ja tietoturva
- * Tieteenalakohtaiset käytännöt ja kulttuuri
- * Pitkäaikaissäilytys, uudelleenkäyttö ja jakaminen



The Data Lifecycle, adapted from the Research 360 Project

HAASTAT- TELUT

- Alustavia tuloksia

”

Suurin osa ei dokumentoi datankäsittelyn vaiheita, koska he eivät ajattele, että kukaan käyttäisi dataa meneillään olevan projektin jälkeen.” (Ohjaaja)

”

Opiskelijat eivät tällä hetkellä saa näihin (tutkimussopimusasioihin) oppia.” (Ohjaaja)

Datanhallinnan TÄRKEYS / OSAAMINEN



Yliopistolle olisi saatava tieteellinen piirtäjä, visualisoija. Siten saataisiin hyvälaatuisia kuvia julkaisuihin. Nyt on todella tehotonta, että jokainen opiskelija yrittää itse opiskella jotakin ohjelmaa.” (Biostatistikko)



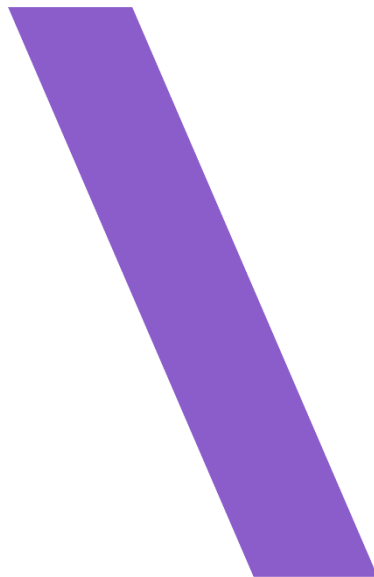
Yhtä tärkeä kuin versiointi on ekstraktointi (eri vaiheiden erottelu). Pitää olla lokit – kuka on käyttänyt dataa. Pystytäänkö todentamaan muutokset, jotka tehty dataan. Tämä ei ole opiskelijoilla hallinnassa.” (Ohjaaja)



Tutkimusdatan hallinta: tärkeys vs. osaaminen



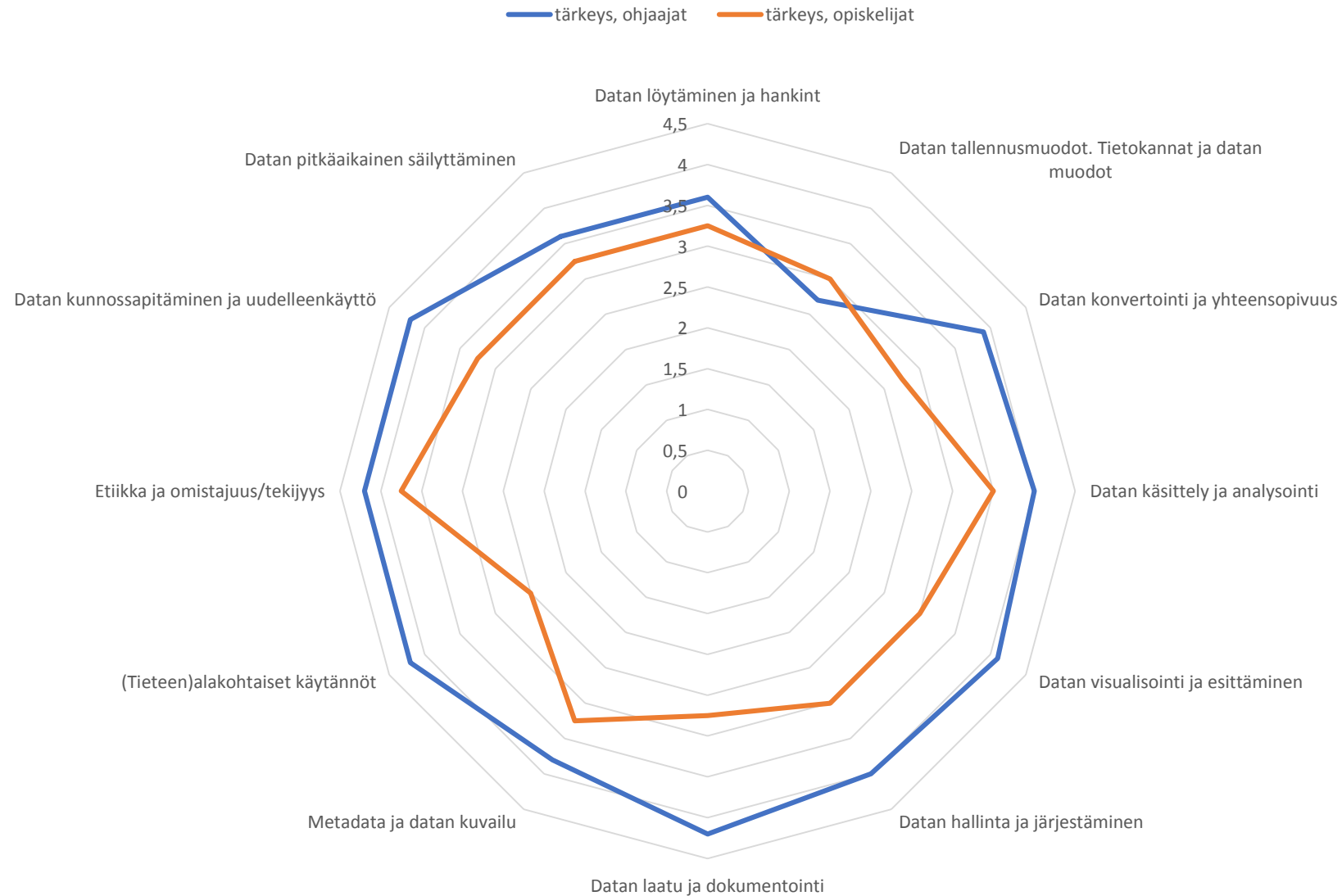
Datanhallinnan TÄRKEYS



Se on varmaan se mindsetti tärkeintä, mitä alkaa pohtimaan, että tohtoriopiskelija ymmärtäisi, että tämä (data) on se arvokas raaka-aine, josta saa tehtyä löydöksiä.

Prosessin jäljitettävyyden on tärkeää, ei sooloilua, tekninen data pitäisi olla jäljitettävissä, toistettavissa, tarkoittaa usein sitä, että datan käsittelyä pitää dokumentoida. Toistettavuus on tieteellinen periaate.” (Ohjaaja)

Tutkimusdatan hallinta: tärkeys ohjaajien / koulutettavien mielestä



tutkimusdatan
hallinnan
PERUSTEET



Kun kulttuurintutkimuksen tieteellisiä käytäntöjä aikoinaan luotin, tällä (kerätyllä tutkimusaineistolla) on ollut vahva rooli suomalaisuuden rakentamisessa, on sisäänrakennettu ajatus siitä, että tuotettavalla aineistolla on arvoa muillekin. On kuitenkin opeteltava erikseen, mitä tämä vaatii metadatalta.” (Ohjaaja)

TAVOITTEET

- * Kurssin suoritettuaan tohtoriopiskelija / postdoc -tutkija ymmärtää, mitä laadukas tutkimussuunnitelma sisältää ja kuinka tärkeää on suunnitella tutkimuksen datanhallinta jo ennen tutkimuksen alkua.
- * Opiskelija tuntee datan elinkaaren vaiheet, erilaiset datatyypit ja -muodot sekä tietoturvan, tietosuojan ja sensitiivisen sisällön vaatimukset datan käsittelylle ja jatkokäytölle.
- * Opiskelija ymmärtää datan arvon paitsi omalle projektille, myös mahdollisille muille käyttäjille tutkimusprojektin aikana ja sen jälkeen ja siten datan organisoimisen, suojauksen ja dokumentoinnin merkityksen koko datan elinkaaren ajan.
- * Lisäksi opiskelija tunnistaa, minkälaisia lupia erilaiset tutkimukset vaativat ja miten ne käytännössä tehdään.
- * Opiskelija ymmärtää immateriaalioikeuksien merkityksen datan jatkokäytön kannalta, hallitsee datan jakamisen ja dataan viittaamisen käytännöt ja standardit sekä tuntee keskeiset yleiset ja tärkeimmät oman tieteenalansa datarepositoryt.
- * Opiskelija ymmärtää datan omistajuuden periaatteet, ja sen mitä eroa on sillä, että data on tutkijan tai yliopiston omistuksessa.
- * Opiskelija oppii käyttämään REDCap -järjestelmää datan keräämiseen ja tutustuu muihin mahdollisiin datanhallintaratkaisuihin.

LINJAT *ja* MODUULIT



*Kaikki, mikä liittyy lain kirjaimeen on epäselvää ja pelottavaa
(Ohjaaja)*



Kaipaisi kyllä ohjeistusta, miten olisi syytä huomioida tulevat käyttäjät, miten olisi syytä dokumentoida ja metadatoittaa. Tiedekunta ei ole tehnyt omaa politiikkaa, miten aineistot olisi syytä kuvata.” (Tohtorikoulutettava)



**TURUN
YLIOPISTO**

TERVEYSTIETEELLINEN TUTKIMUS	KYSELY- JA HAASTATTELUTUTKIMUS	BIOTIETEELLINEN TUTKIMUS	Vastuut
Tutkimussuunnitelman kommentointi, datakappaleen kirjoittaminen, tutkimussuunnitelmaan tarvittavat liitteet	Tutkimussuunnitelman kommentointi, datakappaleen kirjoittaminen, tutkimussuunnitelmaan tarvittavat liitteet	Tutkimussuunnitelman kommentointi, datakappaleen kirjoittaminen, tutkimussuunnitelmaan tarvittavat liitteet	Lehtorit ja yliopisto-opettajat
Immateriaalioikeudet, lupapaperit, lisensointi	Immateriaalioikeudet, lupapaperit, lisensointi	Immateriaalioikeudet, lupapaperit, lisensointi	Lakiasiat, CRC, kirjasto
Datanhallintasuunnitelma	Datanhallintasuunnitelma	Datanhallintasuunnitelma	Kirjasto
Tietosuojaseloste ja vaikutusten arviointi			Tietosuojavastaava, IT-palvelut
RedCap (lomakepohjaisen tietokannan rakentaminen)	RedCapB (kyselytutkimuksen rakentaminen)	RedCap (lomakepohjaisen tietokannan rakentaminen, sähköiset laboratoriotyökalut)	Biostatistikko, lehtori
Datan tallennus, suojaus, käsittely ja kuvailu sekä IT-palveluratkaisut ja tarjonta			IT-palvelut
Datan pitkäaikaistallennus, jakaminen, viittaaminen (kansallinen tiekartta). Datarepositoryt (tieteenalakohtaiset ja yleiset).			Kirjasto, IT-palvelut

LÄHTEET



Tohtoriopiskelijoiden osaaminen vaihtelee paljon, kokonaiskuva datan elinkaaresta, sen hallinnasta ja dokumentoinnista pitkäaikaissäilytyksen näkökulmasta puuttuu.” (Ohjaaja)



Eettisen toimikunnan lausunto oli yksi avaintekijä, joka pani ajattelemaan monia näkökulmia. Niillä opiskelijoilla, joilla tämä jää välistä, nämä kysymykset jäävät pohtimatta. Voisi kuitenkin olla kevennetty menettely, nyt oli turhan byrokraattinen. Samat kysymykset siis jokaiselle mietittäväksi, samalla tavalla kuin tutkimussuunnitelmavaiheessa.” (Tohtorikoulutettava)



- Calzada Prado, J. & Marzal, M.** (2013). Incorporating Data Literacy into Information Literacy Programs: Core Competencies and Contents. *Libri*, 63(2), pp. 123-134. Retrieved 26 Sep. 2017, from doi:10.1515/libri-2013-0010
- Carlson, Jake, Michael Fosmire, C.C. Miller, and Megan Sapp Nelson.** 2011. "Determining data information literacy needs: A study of students and research faculty." *Portal: Libraries & the Academy* 11(2): 629-657. <http://dx.doi.org/10.1353/pla.2011.0022>
- Carlson, Jake and Marianne Stowell Bracke.** 2015. "Planting the seeds for data literacy: lessons learned from a student-centered education program." *International Journal of Digital Curation* 10(1): 95-110. <http://dx.doi.org/10.2218/ijdc.v10i1.348>
- Data Information Literacy Toolkit.** <http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1058&context=dilsymposium>
- Data Literacy (DIL) Section Pages** (2015). <https://nml.gov/data/guides/data-literacy>
- European commission** (2017). Providing researchers with the skills and competencies they need to practice Open Science: Open Science Skills Working Group Report https://ec.europa.eu/research/openscience/pdf/os_skills_wgreport_final.pdf#view=fit&pagemode=none
- European University Association** (2017). Towards Full Open Access in 2020: Aims and recommendations for university leaders and National Rectors' Conferences. <http://www.eua.be/Libraries/publications-homepage-list/towards-full-open-access-in-2020-aims-and-recommendations-for-university-leaders-and-national-rectors-conferences>
- Piorun, Mary, Donna Kafel, Tracey Leger-Hornby, Siamak Najafi, Elaine Martin, Paul Colombo, and Nancy LaPelle.** 2012. "Teaching Research Data Management: An Undergraduate/Graduate Curriculum." *Journal of eScience Librarianship* 1(1): e1003. <http://dx.doi.org/10.7191/jeslib.2012.1003>
- Qin, Jian and John D'Ignazio.** 2010b. "Lessons learned from a two-year experience in science data literacy education." *International Association of Scientific and Technological University Libraries, 31st Annual Conference. Paper 5.* <http://docs.lib.purdue.edu/iatul2010/conf/day2/5>
- Sapp Nelson, Megan R..** 2017. "A Pilot Competency Matrix for Data Management Skills: A Step toward the Development of Systematic Data Information Literacy Programs." *Journal of eScience Librarianship* 6(1): e1096. <https://doi.org/10.7191/jeslib.2017.1096>
- Sapp Nelson, Megan.** 2015a. "Data Management Outreach to Junior Faculty Members: A Case Study." *Journal of eScience Librarianship* 4(1): e1076. <http://dx.doi.org/10.7191/jeslib.2015.1076>
- Sapp Nelson, Megan R.,** "Pilot Data Information Literacy Competencies Matrix Scaffolded Across Undergraduate, Graduate and Data Steward Levels" (2016). *Libraries Faculty and Staff Scholarship and Research. Paper 136.* http://docs.lib.purdue.edu/lib_fsdocs/136



**TURUN
YLIOPISTO**

BRDM-työryhmä

jukka.rantasaari@utu.fi