

AINEISTONHALLINTA

Koulutukset Turun yliopistossa

*Jukka Rantasaari
Palvelupäällikkö,
Kirjasto*

*Päivi Kanerva,
Informaatikko,
Kirjasto*



- Systemaattisempaa koulutusta aineistohallinnasta vuodesta 2016 lähtien:
- Ensimmäisessä “aallossa” koulutettiin yliopiston ja kirjaston henkilökuntaa siitä, mistä on kyse.
- Luotiin dataopas: <https://utuguides.fi/tutkimusdata>
- Suomen Akatemian haun aikaan aineistohallintatyöpajoja ja yleisinfoja Tutkimusrahoitus-yksikön kanssa yhteistyönä
- Kaksi kurssia tohtorikoulutettaville (UTUGS)
- Näiden lisäksi yksittäisiä koulutuksia tiedekunnille/oppiaineille/tutkimusryhmille tilauksesta. Esimerkkinä väitöskirjatutkijoiden Tutkimusetiikka-kurssi, jossa omat esitykset avoimesta julkaisemisesta ja datasta.
- Koulutusten sanoma muuttunut avoimen tieteen hypetyksestä aineistojen laadukkaaseen hallintaan
- Tulevaisuuden suunnitelmissa aineistohallintakoulutukset opinnäytetöiden ohjaajille

Information Resources and Tools for Research – kurssi tohtorikoulutettaville

- Aineistohallintaan liittyvää koulutusta lisätty kirjaston pitämälle väitöskirjatutkijoiden UTUGS-kurssille (Information resources and tools for research) syksyllä 2016
- Kurssi toteutetaan kokonaan verkko-opetuksena
- Kurssin Open science -osuus sisältää koulutusta niin avoimesta julkaisemisesta (“Test Your Skills..”) kuin aineistohallinnasta. Konkreettisenä tuotoksena tehtävä tuottaa aineistohallintasuunnitelma omasta tutkimuksesta/esimerkkiaineiston pohjalta. Kirjasto kommentoi suunnitelmia.

Topics	Assignments	Opens	Deadline
Introduction	<ul style="list-style-type: none"> • Assignment 1: Introduce yourself • Assignment 2: Test your skills in library services 	7.10.2019	14.10.2019
Building your search 1	<ul style="list-style-type: none"> • Assignment 3: Test your skills in information seeking techniques • Assignment 4: Building your search strategy 	7.10.2019	14.10.2019
Tools for evaluation of publications	<ul style="list-style-type: none"> • Assignment 5: Evaluation of publications 	15.10.2019	21.10.2019
Building your search 2	<ul style="list-style-type: none"> • Assignment 6: Searching in the databases 	22.10.2019	28.10.2019
Managing the references	<ul style="list-style-type: none"> • Assignment 7: Managing the references 	29.10.2019	4.11.2019
Open science	<ul style="list-style-type: none"> • Assignment 8: Test your skills in Open Science 	5.11.2019	11.11.2019
Data management plan	<ul style="list-style-type: none"> • Assignment 9: Data Management Plan 	5.11.2019	18.11.2019
Publishing ethics	<ul style="list-style-type: none"> • Assignment 10: Academic publishing 	19.11.2019	25.11.2019
Keeping up	<ul style="list-style-type: none"> • Assignment 11: Keeping up - discussion 	26.11.2019	2.12.2019
Feedback	<ul style="list-style-type: none"> • Assignment 12: Feedback 	26.11.2019	2.12.2019

Aineistohallintasuunnitelmien kommentointia

- Kommentointi keskittynyt Suomen Akatemian syyskuun haun aikaan.
- Kommentoitavaksi tulevat aineistohallintasuunnitelmat lähetetään yhteiseen palveluosoitteeseen openutu@utu.fi
- Palveluosoitteen takana useita eri alojen asiantuntijoita; kirjasto koordinoi vastaamista/vastuussa kokonaisuudesta
- Kirjastossa suunnitelmia kommentoi 7 henkeä
- Kommentointia tehty s-postitse ja syksystä 2019 enemmän myös keskinäisissä tapaamisissa
- Muu kommentointityö:
 - Väitöskirjatutkijoiden UTUGS-kursseilla harjoitustyönä olleita aineistohallintasuunnitelmia.
 - Muut rahoitushaut

Mitä kommentointi käytännössä tarkoittaa?

- Kirjastossa kommentoidaan suunnitelmaa kokonaisuutena.
 - Huomioita myös luettavuudesta, pituudesta ja rakenteesta
- Tutkijoiden ”hoksauttamista” – ei anneta valmiita vastauksia, ohjataan oikean tiedon ääreen ja toivotaan, että tutkija osaa soveltaa ohjeita suhteessa omaan tutkimukseensa
- Kommentit usein ”kannattaa vielä miettiä”
”kiinnittäisin enemmän huomiota”
ohjaavaa, suostuttelevaa, maanittelevaa



The Basics of

RESEARCH DATA MANAGEMENT (BRDM)

*Jukka Rantasaari
Palvelupäällikkö,
Kirjasto*



TAUSTAA



DATAPOLITIikka 2016

- TY:n datapolitiikassa kirjastolla päävastuu:
 - Tutkijat saavat koulutusta ja tukea aineistohallintasuunnitelmien laatimiseen sekä datanhallintaan tutkimuksen elinkaaren ajan.
 - Tutkimusdatan ja metadatan hallinta (suunnittelu, säilytyspaikat, tallentaminen)
- Ennen kuin aletaan suunnitella koulutusta, perehdytään nykytilanteeseen, mitä nähdään tärkeänä, missä on kehittämisen varaa ja tarpeita eli kontekstoidaan koulutus

HAASTATTELUT

Haastattelututkimuksen teoreettinen viitekehys

- Carlsonin et al. DIL-projektissa määritellyt datainformaatiolukutaidon 12 kompetenssia
- Haastattelupohjat tdk:lle ja opiskelijoille
- → muokattiin viitekehystä, mm.tärkeyttä mittaavien Likert-kysymysten lisäksi lisättiin kysymys havaitusta tämänhetkisestä osaamisesta.

Haastateltavien valinta

- Tieteenalan tai oppiaineen
 - oletettu dataintensitiivisyys;
 - erityyppiset ja muotoiset datat ja datalähteet
 - “eResearch”-ille tyypillisten piirteiden kuten datan ja tutkimusmenetelmien digitaalisuus, datan ja data-alustojen määrän moninkertaistuminen, ja yhteistyön tarve useiden eri tahojen kuten muiden tutkijoiden ja tutkimuksen tukipalvelujen kanssa tieteenalalla tai oppiaineessa
- Mukaan valittiin myös joitakin haastateltavia oppiaineista, jotka eivät ole ns. “dataintensiivisiä”, esim. teoreettinen fysiikka; HKT

HAASTATELLUT

Faculty	Doctoral Candidate	Supervisor	Biostatistician	Total	Percentage of all interviewees	Doctoral Candidates of Faculty	Percentage of all Doctoral Candidates at UTU
Faculty of Medicine + Turku PET Centre	5	3	4	12	34	637	33
Faculty of Science and Engineering + Finnish Centre for Astronomy with ESO	4	6		10	28	363	19
Turku School of Economics	3	2		5	14	160	8
Faculty of Humanities	1	2		3	9	360	19
Faculty of Social Sciences	1	2		3	9	182	10
Faculty of Education	1	1		2	6	126	7
Faculty of Law				0	0	74	4
				35	100	1902	100

TULOKSIA

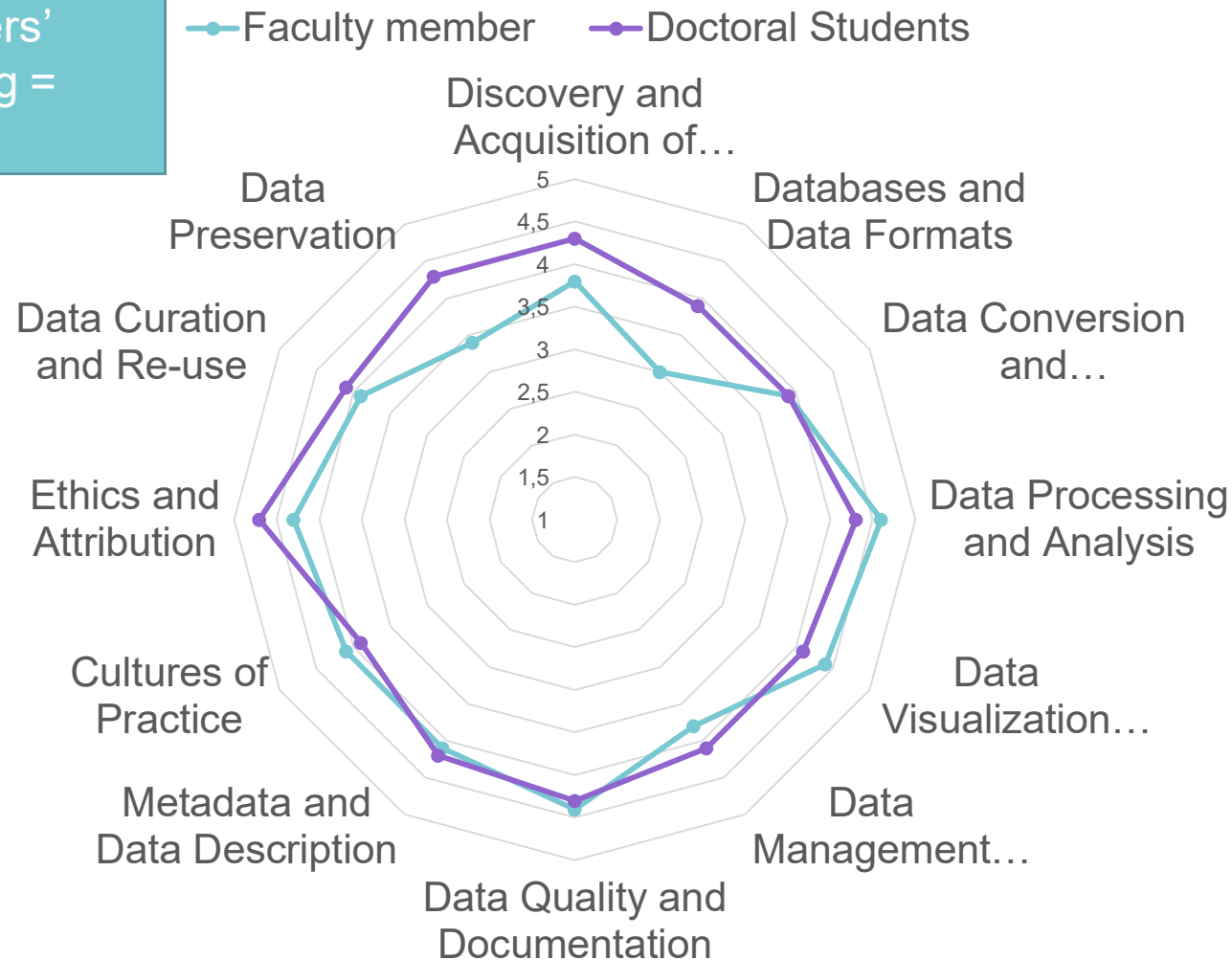


Importance:

Faculty Members Vs. Doctoral Candidates

Faculty members' average ranking = 3.98

Doctoral students' average ranking = 4.17



- Three most important (both groups):
- 1) Ethics and attribution 4.47
 - 2) Data processing and analysis 4.45
 - 3) Data quality and documentation 4.38

p=0.33

Importance Vs. Competence

All interviewees' average ranking for importance = 4.07

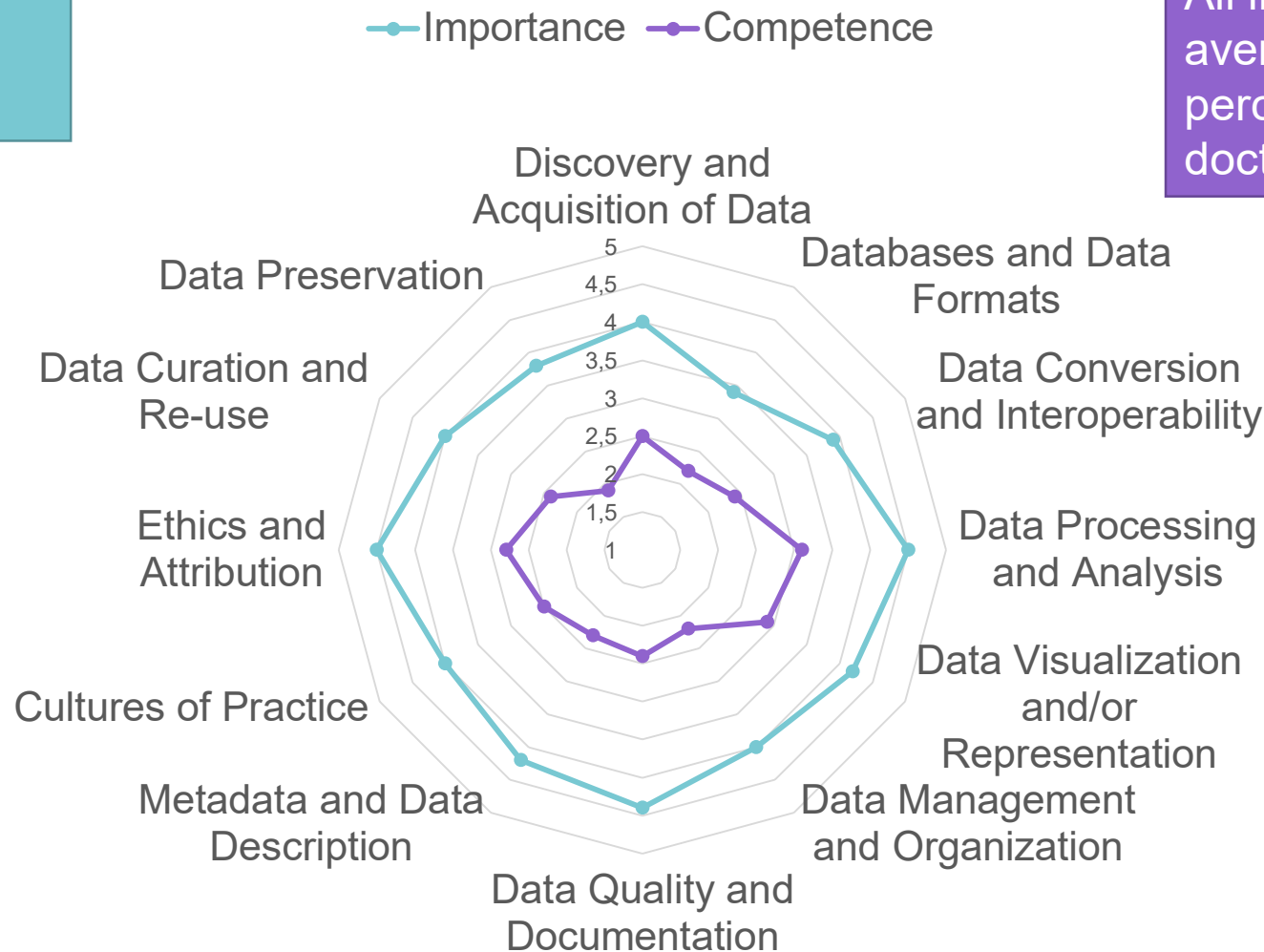
All interviewees' average ranking for perceived competence of doctoral candidates' = 2.47

Best competencies in:

- 1) Data processing and analysis 3.06
- 2) Data visualization and representation 2.88
- 3) Ethics and attribution 2.79

Weakest competencies in:

- 1) Data preservation 1.94
- 2) Databases and data formats 2.22
- 3) Data management and organization 2.21

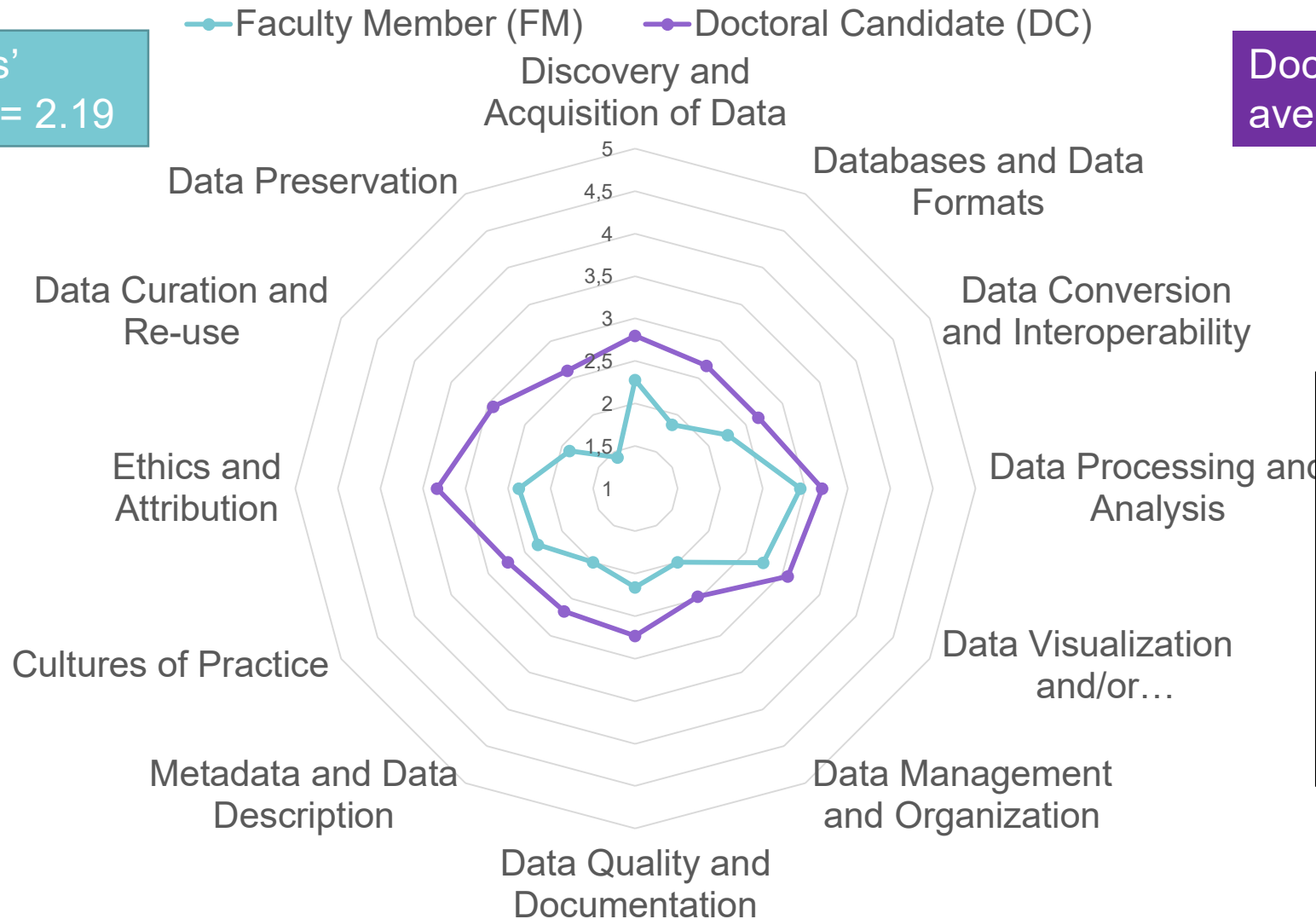


$p < 0.0001$

Perceived competence: Faculty Members Vs. Doctoral Candidates

Faculty members'
average ranking = 2.19

Doctoral candidates'
average ranking = 2.82



Statistically most significant differences in

- 1) Data preservation ($p < 0.00$)
- 2) Data curation and reuse ($p = 0.00$)
- 3) Ethics and attribution ($p = 0.00$)

$p = 0.0018$

KURSSIN SUUNNITTELU



SUUNNITTELU JA RAKENTAMINEN

- Työryhmä
 - Kirjasto; biostatistiikka; tiedekunnat; lakiasiat; tietosuojavastaava; rahoituspalvelut; tutkimusedellytykset; CRC; tutkimuksen IT
- RDM kompetenssit
 - geneerisiä, joiden toteutustapa riippuu oppiaineesta ja datan tyypeistä ja muodoista
 - RDM yleiskurssi, jossa kolme tutkimuslinjaa ja seitsemän moduulia
 - hyödynnettiin olemassaolevia kv. datanhallinnan koulutushankkeita kuten [DIL-projektin](#) Cornellin yliopiston kurssia ja [New England Collaborative Data Management Curriculumia](#).
- Esittely UTUGS:n johtoryhmälle
 - 3 op:n vapaaehtoinen kurssi UTUGS:n yhteisten opintojen ryhmään.
- 5-6 kokoontumista syksyn ja alkukevään aikana

INTRODUCTORY LECTURE
MODULE 1: RESEARCH PLAN,
MODULE 2: DMP
Data management and organization
Cultures of practice

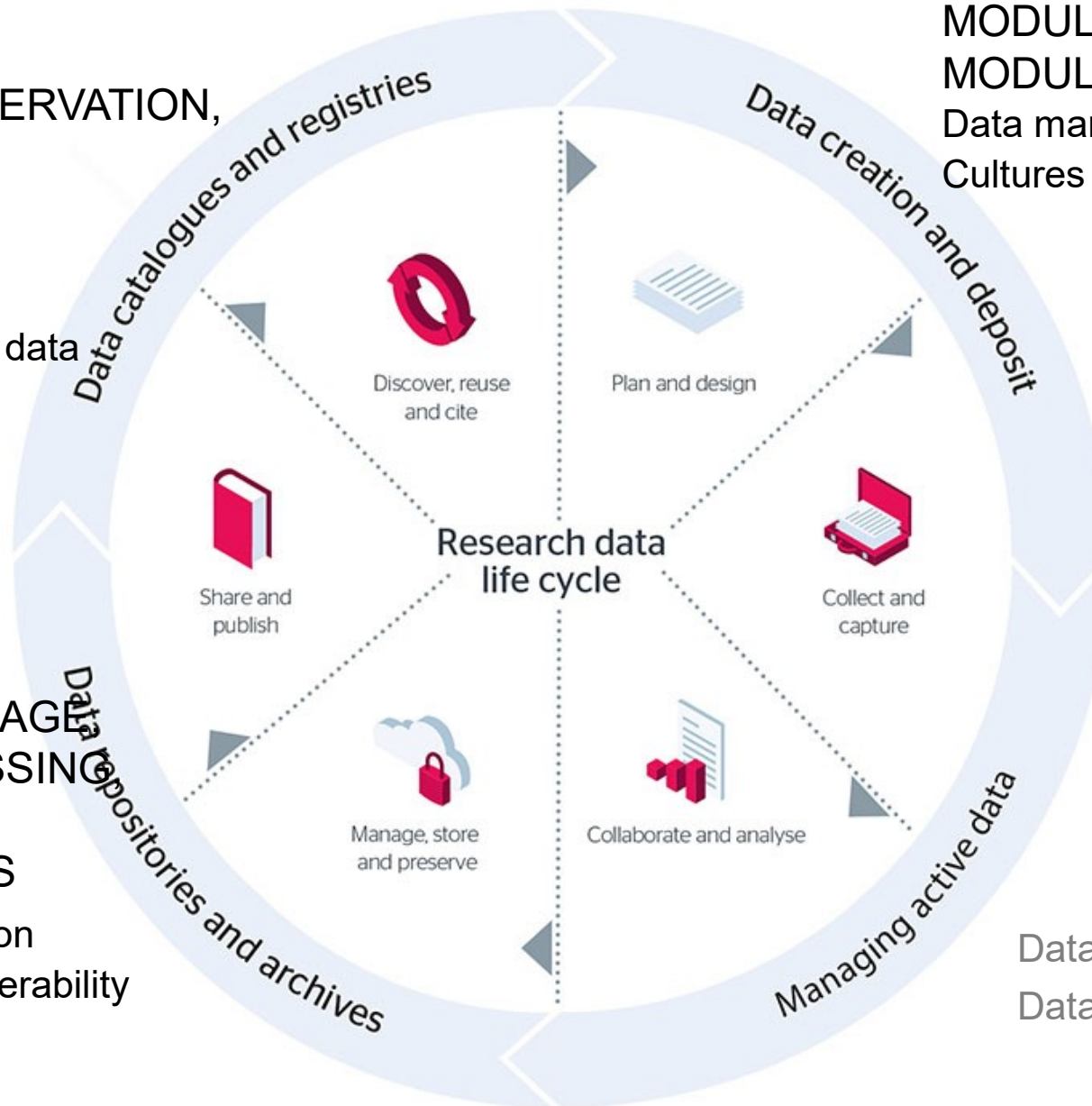
MODULE 3: IPR, PERMITS, LICEN
MODULE 4: PRIVACY NOTICE,
RISK ANALYSIS
Ethics and attribution

MODULE 5: DATA COLLECTING
AND ORGANIZING
Databases and formats
Data quality and documentation

Data processing and analysis
Data visualization and representation

MODULE 7: DATA PRESERVATION,
SHARING AND CITING,
DATA REPOSITORIES
Data preservation
Data curation and reuse
Discovery and acquisition of data

MODULE 6: DATA STORAGE
PROTECTION, PROCESSING
DESCRIBING,
IT SERVICE SOLUTIONS
Metadata and data description
Data conversion and interoperability



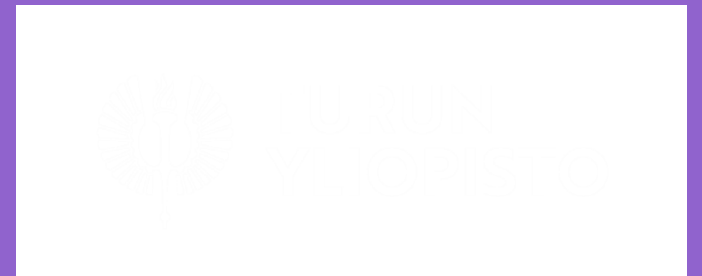
DMP WORKSHOP

OSAAMISTAVOITTEET

- Opiskelija tietää, mitä tarkoitetaan käsitteillä tutkimusaineisto, data ja datanhallinta ja ymmärtää, miksi datanhallinnan suunnittelu on syytä aloittaa jo ennen tutkimuksen alkua.
- Opiskelija ymmärtää, mitä laadukas tutkimussuunnitelma sisältää, mitä ja miten tutkimusaineistosta (datasta ja materiaalista) on syytä kirjata tutkimussuunnitelmaan.
- Opiskelija tuntee datan elinkaaren vaiheet, erilaiset datatyypit ja -muodot sekä tietoturvan, tietosuojan ja sensitiivisen sisällön vaatimukset datan keruulle, käsittelylle ja jatkokäytölle.
- Opiskelija ymmärtää datan arvon paitsi meneillään olevalle projektille, myös itselle ja mahdollisille muille käyttäjille tutkimusprojektin jälkeen ja siten datan dokumentoinnin, kuvailun ja suojauksen merkityksen koko datan elinkaaren ajan.
- 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 datarepositoriot.
- Opiskelija ymmärtää datan omistajuuden periaatteet, ja sen mitä eroa on sillä, että data on tutkijan tai yliopiston omistuksessa.
- Opiskelija oppii käyttämään tietokantaa datan keräämiseen, organisointiin, dokumentointiin ja kuvailuun ja tutustuu muihin mahdollisiin datanhallintaratkaisuihin.

2019 TOTEUTUS

Maaliskuu – Toukokuu 2019



HEALTH SCIENCES	SURVEYS AND INTERVIEWS	NATURAL SCIENCES	Responsibility
Research plan: - Commenting - Describing the research data - Supplements	Research plan: - Commenting - Describing the research data - Supplements	Research plan: - Commenting - Describing the research data - Supplements	Lectors and University Teachers
IPR rights issues, permits and licences	IPR rights issues, permits and licences	IPR rights issues, permits and licences	Legal Affairs, CRC, Library
Data management plan (DMP)	Data management plan (DMP)	Data management plan (DMP)	Library
Privacy notice and risk analysis -			Head of IT Services for Researchers, Data Protection Officer
RedCap (building form based database)	RedCap (building survey form)	RedCap (building form based database, electronic laboratory tools)	Biostatistician
Data storage, protection, processing, describing and IT Service solutions			IT Services
Data preservation, sharing and citing (national citation standard). General and discipline specific open data repositories			Library, IT Services

ILMOITTAUTUNEET / ALOITTANEET / SUORITTANEET

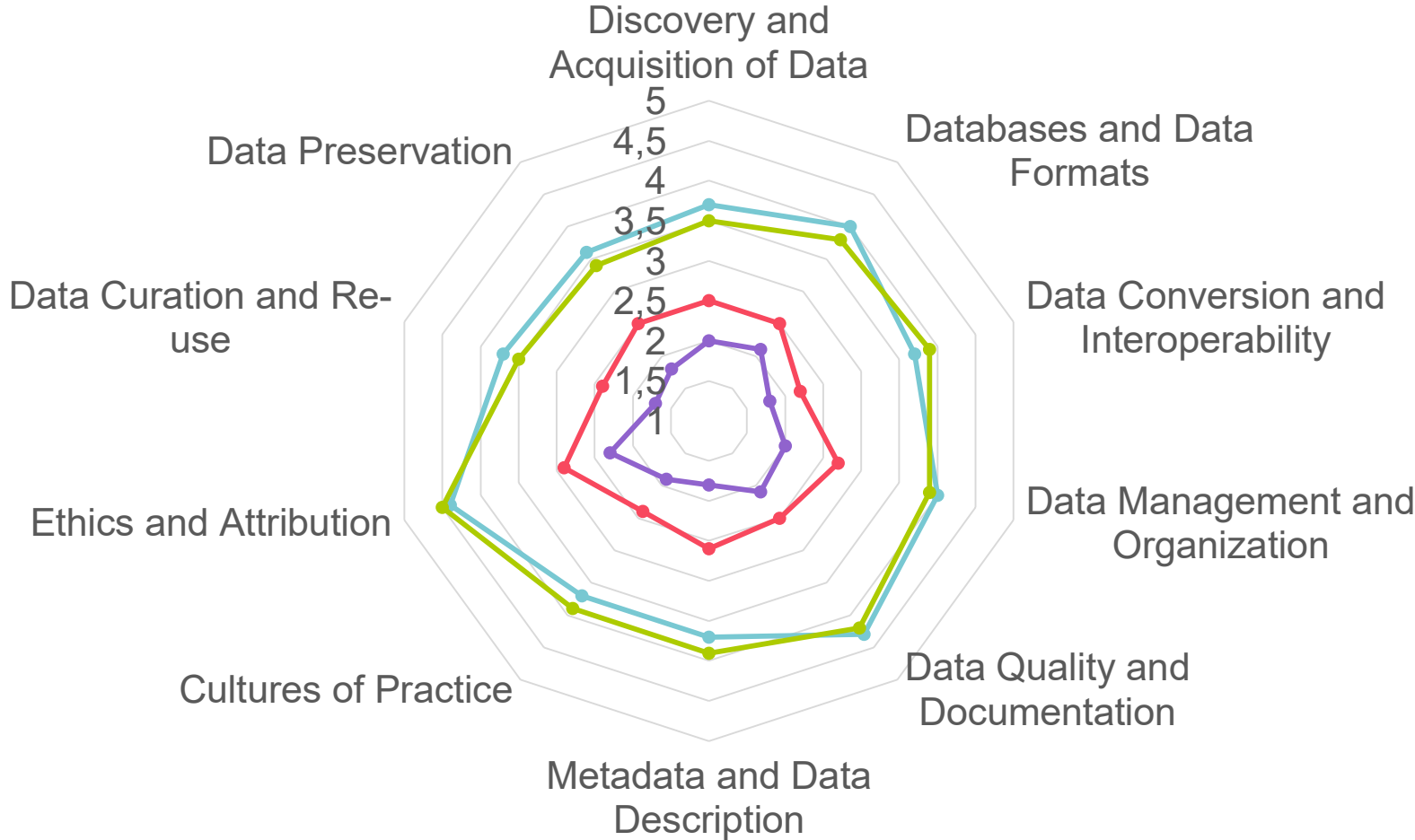
Tohtorikoulutettavat

- 55 ilmoittautunutta
- 36 aloitti kurssin
- 27 suoritti kurssin loppuun (75 %)

Lisäksi post doc –tutkijoita, jotka suorittivat haluamiaan moduuleja

Importance and competence: before and after

— Importance, before — Competence, before — Importance, after — Competence, after



Importance, before: 3.9
Importance, after: 3.9
Competence, before: 2
Competence, after: 2.5

KOKO KURSSIN PALAUTTEITA

PLUSA

- pitäisi olla kaikille pakollinen, koska kaikilla ei ole tutkimusryhmää jolta katsoa mallia
- erinomainen johdatus aiheeseen
- tarpeellinen kaikille tutkijoille
- hyvät luennoitsijat
- tutkimussuunnitelma ja REDCap hyviä, muut eivät niin kolahtaneet, vaikkakin tärkeitä aiheita
- hyvä ja looginen rakenne
- suomenkieli
- aineistohallintasuunnitelma -moduuli ja REDCap parhaat
- monipuolinen katsaus monimutkaiseen asiaan
- sain tietoa työkaluista, joita yliopisto tarjoaa,
- sain tietoa omistajuudesta
- tiedän, mistä etsin jatkossa syventävää tietoa
- erityisesti REDCapin käytön oppiminen, mutta myös muut hyödyllisiä
- kolme erillistä linjaa
- todella paljon asiaa, mutta kaikki toki tarpeellista
- hyödyllinen, tiedän, mistä löydän lisätietoa, ja mitä työkaluja käytän
- oikein hyvä kurssi, olen suositellut muillekin
- todella hyödyllinen kurssi
- hyvä kokonaisuus, kellonaika 14-17 hankala

KOKO KURSSIN PALAUTTEITA

KEHITTÄMISKOhteita

- Kokonaiskuva selkeämmäksi koko kurssille. Yleiskuva jokaisen moduulin alussa. Kuva "Life cycle of research data" on kiva nähdä kurssin alussa, sillä
- ymmärtää, mitä kaikki siihen datanhalintaan kuuluu. Ja jokainen moduuli olisi voinut jäsentää kuvan avulla.
- Moodlessa voisi lukea jokaiseen moduulin otsikon edessä "Moduuli 1: XY-aihe". Helpottaisi oikean kohdan löytämistä.
- aikatauluissa epäselvyyttä
- pohdintoja olisi hyvä voida jatkaa esim. datamanagerin kanssa
- harjoitustehtävät eivät aina tuntuneet hyödyllisiltä
- voi olla vaikea lähteä soveltamaan omaan aiheeseen, koska olisi pitänyt mennä syvemmälle aiheeseen
- materiaalit aikaisemmin jakoon
- DMP:n koostaminen täysin vieraasta asiasta vaikeaa
- jos yliopistolle määritellään omia käytänteitä ja standardeja, olisi hyvä olla nettisivu, johon nämä on koottu linkkeineen
- olisi voinut etukäteen avata terveystieteellisen ja survey-linjojen eroa. Em. painottui nyt liikaa lääketieteeseen
- jatkossa myös laadullisen datanhallintaan liittyvä linja
- jos mahdollista, kaikki opetus omissa ryhmissään, koska kaikilla on omia erikoiskysymyksiä
- lisää asioiden konkretisoimista ja harjoitustehtäviä

POIMINTOJA MODUULIEN PALAUTTEISTA

I learned the impact of
GDPR on data privacy

In the future, I
hope it would be
possible for each
to make their own
research plan and
DMP

I will begin to plan
my data
management in the
beginning of a
research project

I learned the
importance of
agreements, permits
and IPR



2020 TOTEUTUS

Helmikuu – Toukokuu 2020



MUUTOKSET

- Neljä linjaa, kaksi suomen-, kaksi englanninkielistä
- Aloitustuento: johdanto kurssin tematiikkaan, käsitteet, hyvä tutkimussuunnitelma, käytännön toimintatavat kurssilla
- Määriteltiin moduulikohtaiset osaamistavoitteet, vuorovaikutus tunneilla, ennakko- ja jälkitehtävät
- Opiskelijoille mahdollisuus tehdä omaa tutkimussuunnitelmaa ja datanhallintasuunnitelmaa kurssin ajan
- Otetaan yht. 35 tohtorikoulutettavaa + 5 post doc -tutkijaa / linja TY:sta ja ÅA:sta
- Åbo Akademiilta neljä opettajaa

HEALTH SCIENCES	SURVEYS	INTERVIEWS	NATURAL SCIENCES	Responsible
Introductory lecture				
Research plan: - Commenting - Describing the research data - Supplements	Research plan: - Commenting - Describing the research data - Supplements	Research plan: - Commenting - Describing the research data - Supplements	Research plan: - Commenting - Describing the research data - Supplements	Lectors and university teachers
Data management plan (DMP)	Data management plan (DMP)	Data management plan (DMP)	Data management plan (DMP)	Library
IPR rights issues, permits and licences	IPR rights issues, permits and licences	IPR rights issues, permits and licences	IPR rights issues, permits and licences	Legal affairs, CRC, library
Privacy notice and risk analysis				Head of IT services for researchers, Data Protection Officer,
RedCap (building form based database)	RedCap (building survey form)	NVIVO? Something else?	RedCap (building form based database, electronic laboratory tools)	Biostatistician, lector
Data storage, protection, processing, describing and IT Service solutions				IT Services
Data preservation, sharing and citing (national citation standard). General and discipline specific open data repositories				Library, IT Services
DMP workshop				Library + All others



Kiitos!

Jukka Rantasaari, Päivi Kanerva