



Add your questions/comments  
to [padlet.com/laulumieli/DMP](https://padlet.com/laulumieli/DMP)

# From Theory to Practice

**Embedding RDM Competencies into  
Data Management Plans**

LIBER Conference 2024,  
July the 5th, Limassol, Cyprus

Jukka Rantasaari

Head of Library Services, University of Turku

Doctoral Researcher, Åbo Akademi University

# Sound Research Data Management (RDM) – What's the Point?

- Maintains credibility and reproducibility of research
- Helps researchers manage data efficiently
- Supports rigorous scientific methods
- Promotes sharing and collaboration
- Contributes to societal trust in scientific findings



Add your questions/comments  
to [padlet.com/laulumieli/DMP](https://padlet.com/laulumieli/DMP)

# Journey of RDM Competencies

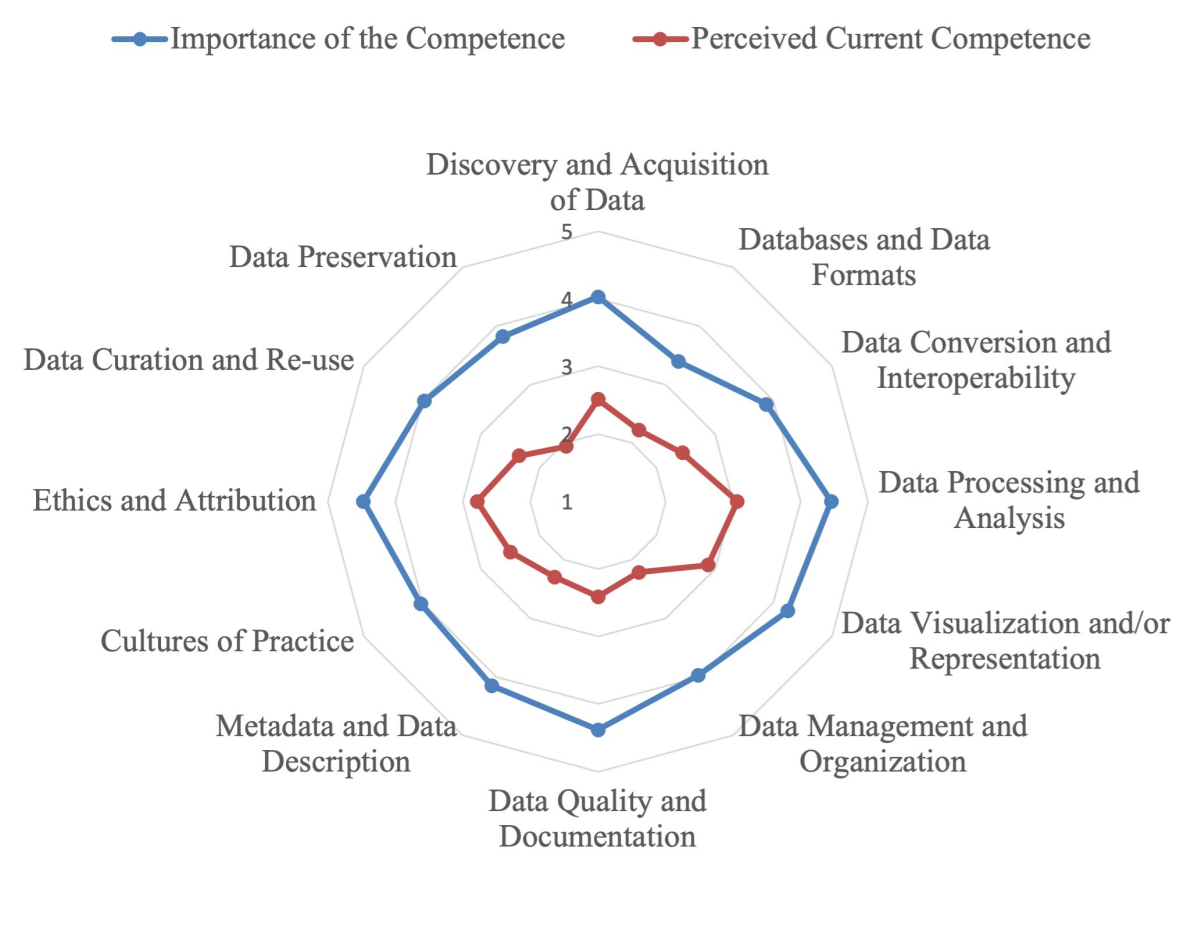
**Starting with Interviews, Progressing Through Training, and  
Assessing Outcomes**



**UNIVERSITY  
OF TURKU**

# Doctoral Students' RDM Skill Gaps

- Comprehensive planning and documentation of research data management (RDM)
- Understanding RDM's importance for research quality
- Legal and ethical aspects of data handling





# Basics of Research Data Management (BRDM) (3 ECTS)

CLINICAL HEALTH SCIENCES	SURVEY RESEARCH	QUALITATIVE RESEARCH	NATURAL SCIENCES	TEACHERS
Introductory Lecture: Background and concepts; Characteristics of a high quality research plan; Course practicalities				Head of library services; Data librarian; Grant writer
Research plan: Objective; design; implementation; expected results	Research plan: Objective; design; implementation; expected results	Research plan: Objective; design; implementation; expected results	Research plan: Objective; design; implementation; expected results	Researchers; Lectors; University teachers
Data management plan (DMP): data life cycle; DMP-tool; roles; resources; metadata; documentation				Self-study module; Data librarian
IPR, agreements and licenses (in Finnish)		IPR, agreements and licenses (in English)		Head of legal affairs unit; Assistant legal counsel; Data librarian
Data privacy: privacy notice; risk analysis; (in Finnish)		Data privacy: privacy notice; risk analysis; anonymization (in English)		Data protection officer; Data archive specialist
RedCap: building a form based database	RedCap: building a survey form	NVivo: organizing and coding data	RedCap: building a form based database	Head of biostatistician team; Lector
Data storage, protection, processing, describing and IT Service solutions				IT system architect
Data preservation, sharing and citing; Open data repositories				Data librarian
A voluntary Q&A Session				Module teachers
DMP: returning and peer-reviewing	DMP: returning and peer-reviewing	DMP: returning and peer-reviewing	DMP: returning and peer-reviewing	Head of library services; Data librarian
a general level feedback on DMPs				Head of library services



# Assessing the Results

## Self-assessment of the learning

- Rating RDM competencies
  - Scale: 1=No competence; 4=very competent
- Giving course and module-based feedback
  - What are the three things you learned?
  - How do they change your practices?

## Assessment of the DMPs by the Teacher

- Rating Data Management Plans (DMPs)
  - [Finnish DMP Evaluation Guidance \(FDEG\)](#)
  - Scale: 0=Poor; 2=Excellent
- Content analysis of RDM best practices applied in DMPs

# Example of the FDEG's Assessment Criteria

## 3. Documentation and metadata

3.1 How will you document your data in order to make the data findable, accessible, interoperable and re-usable for you and others?

### Evaluation guidance 2021

#### Excellent

- Clearly outlines the documentation needed to verification and enable data re-use.
- Lists the metadata standards used for each data type.
- Describes how the documentation protocol is agreed (and documented), if no standard is available for a data type.
- Refers to documentation requirements of a data repositories/archives planned to use.
- Outlines who is/are responsible for the documentation during the data lifecycle (collection, analysis, storing, publishing, etc.)

#### Satisfactory

- Clearly outlines the documentation needed to to verification enable data re-use.
- Indicates how the data will be organised during the project (for example naming conventions, version control strategy and folder structures).
- Mentions common data documentation elements like, a 'readme' text file, file headers, code books, lab notebooks.

#### Poor

- Provides little or no details on the metadata that will accompany the data.
- Provides no information, or only a very vague mention of documentation, without providing any detail or explanation.

# Recommended RDM Best Practices

Category	Definition
<b>Controller</b>	Data controller named
<b>Data table</b>	Data type specific classification included in a DMP
<b>Detailed descriptions</b>	Descriptions of RDM practices are detailed
<b>Funder's policy</b>	Explained funder's or publisher's data sharing policy
<b>Legal basis</b>	Stated legal basis for handling personal data
<b>License for reuse</b>	Named a license for data reuse
<b>Open data</b>	At least part of the data will be opened
<b>Open metadata</b>	Metadata will be opened
<b>Ownership</b>	The ownership is clearly described and justified
<b>Permission</b>	Permission asked for data sharing and reuse
<b>Resources</b>	Evaluated and described resources needed
<b>Secure storing</b>	Used only secure storing of personal data
<b>Storing by types</b>	A data type specific storing platform
<b>Usage rights</b>	Specified collaborators' different rights of use
<b>Why closed</b>	If not shared the data, justification mentioned



# Variations in Data Management Plans (DMPs)



Add your questions/comments  
to [padlet.com/laulumieli/DMP](https://padlet.com/laulumieli/DMP)



UNIVERSITY  
OF TURKU

# Data Table: An Example

Data type and source	File format	Personal or sensitive data	Ownership and agreements	Metadata documentation	Storage during project	Sharing data after the project	Long-term archiving	Estimated size
<b>Lab notes (Data produced)</b>	.doc .txt .pdf	Yes. Subject to IPR check	PI and group	Programme generates metadata by itself	Electronic lab notebook (eLabJournal)	Project team	Discarded after 15 years	< 10 MB
<b>RNA sequences (Data produced)</b>	raw: FASTA, BAM, .xlsx	no	PI	Readme.txt	UTU's network drive and cloud	European Nucleotide Archive	no	< 1 GB
<b>MRI images (Data reused)</b>	DICOM, .nii, .tiff	Yes, record keeper: xx	PI	Readme.txt	Database x at TYKS, backup	NITRC after anonymization	no	< 1 GB
<b>Question-naire forms (Data collected)</b>	Paper forms	Yes, record keeper: xx	PI	Readme.txt	Locked filing cabinets in PI's office	No, metadata shared in Zenodo/Etsin	Discarded 5 years after publication	

# Differences Between DtDMPs and Prose DMPs

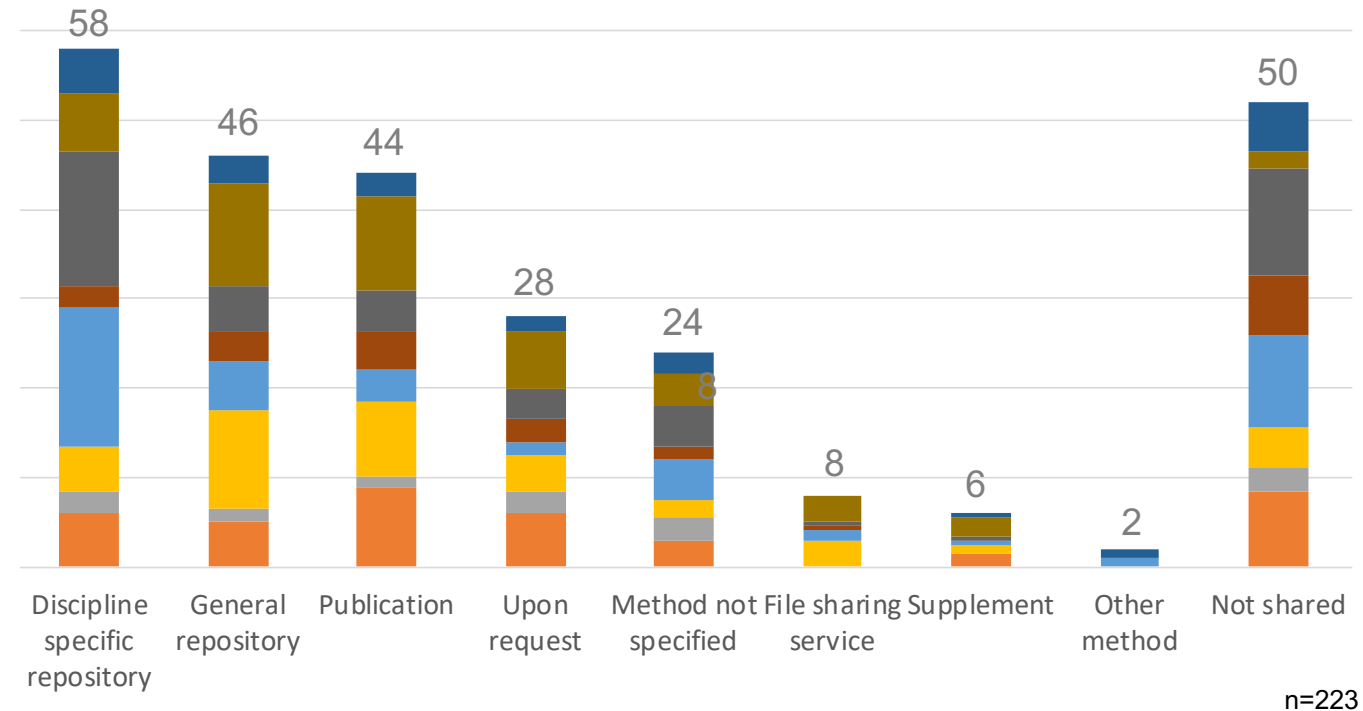
## Differences were significant

- Detailed descriptions
- Usage rights specified
- Permission for sharing
- Data controller identified
- Reason for withholding
- Data type specific storing
- Ownership defined

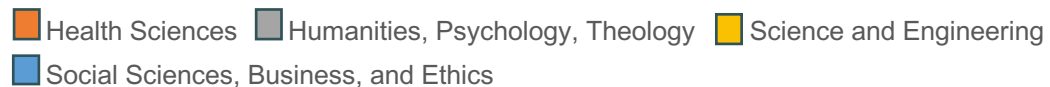
## Differences were insignificant

- Funder's policy
- Legal basis
- License for reuse
- Open data
- Open metadata
- Resources
- Secure storing

# Sharing Methods



## Disciplines



## BRDM Course Tracks



## Discipline-specific Repositories:

- Preferred by 26% overall.
- Dominant in social sciences, business, economics, and survey research.

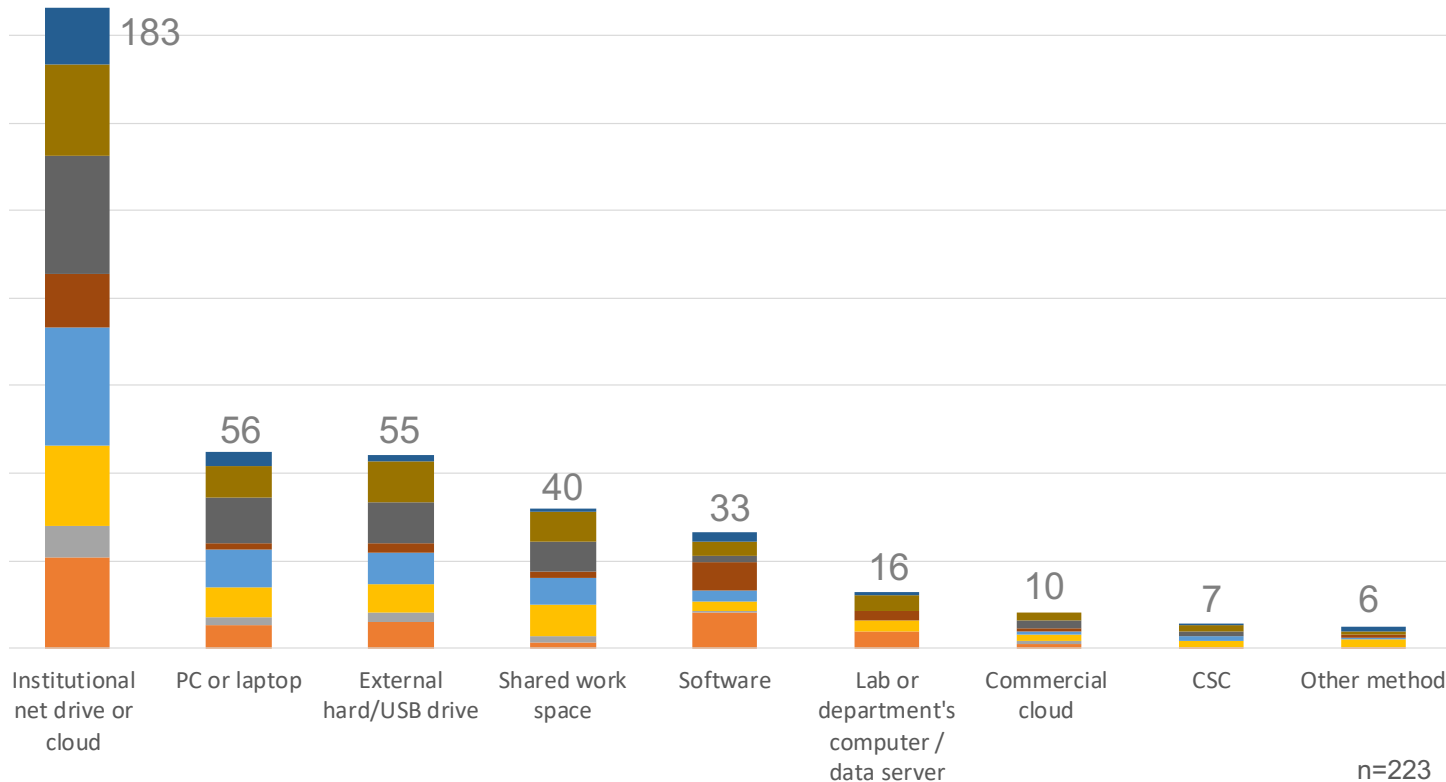
## Generalist Repositories:

- Chosen by ~40% in science and engineering.

## Publication as Sharing Method:

- Common in STEM and natural sciences (30%).
- Less usual in other fields (10%).

# Storing Methods



## Institutional Net Drive or Cloud:

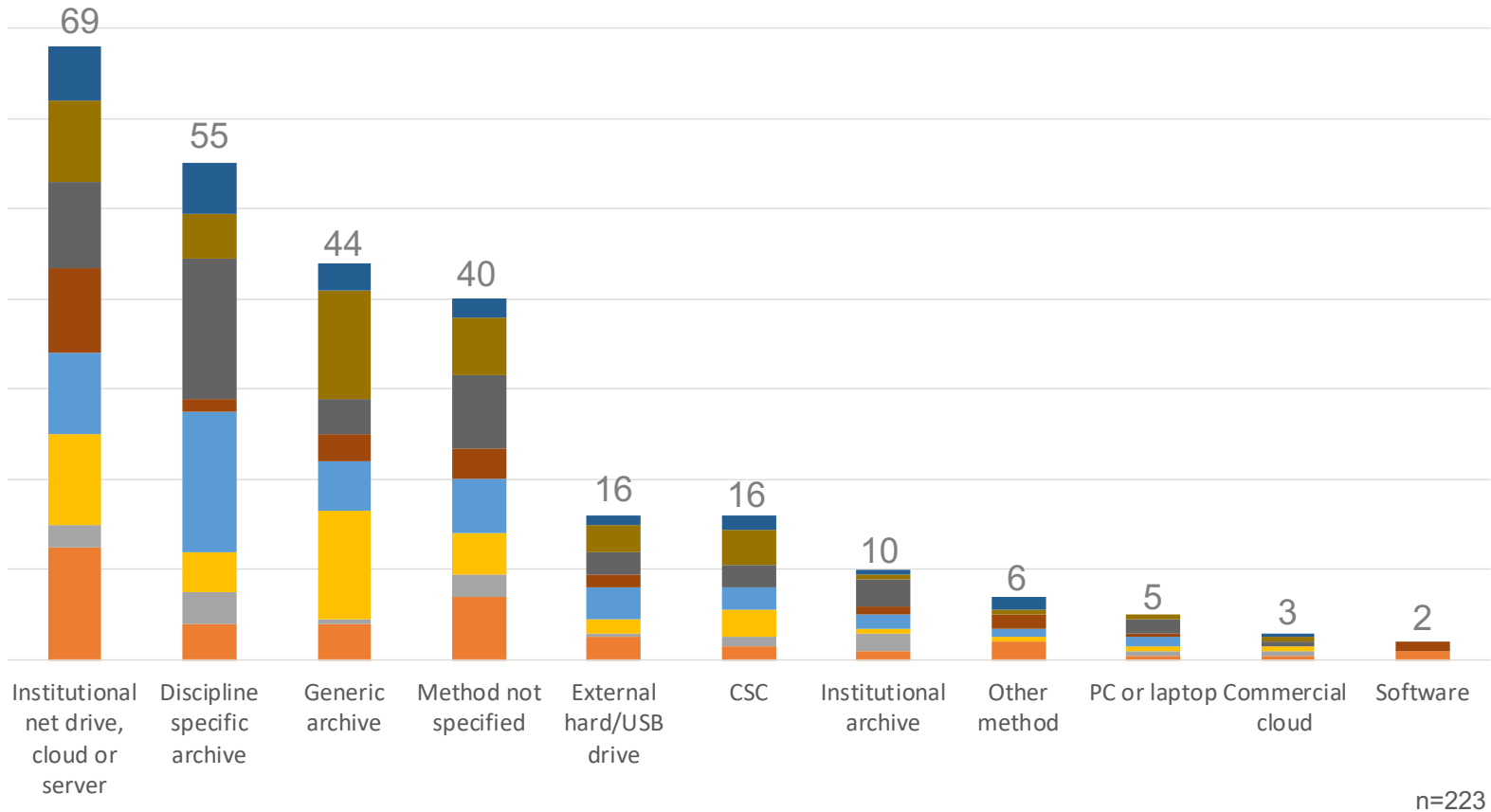
- Utilized in 82% of Data Management Plans (DMPs).
- Backups on other devices also common.

## Discipline/Method-Specific Preferences:

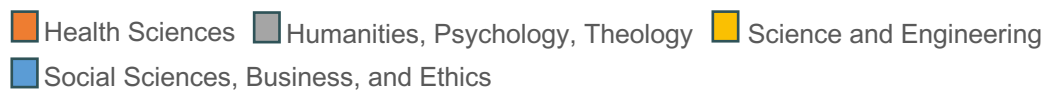
- Natural science researchers often used external drives as a secondary option.
- Health science and clinical method users favoured data collection or analysis software.



# Preserving Venues



## Disciplines



## BRDM Course Tracks



## Institutional Platforms:

- Predominantly used by 40-50% of health sciences and clinical method users.

## Discipline-Specific Archives:

- Chosen by 30-40% of humanities and social sciences researchers.
- Popular among those using qualitative or survey methods.

## Generalist Archives:

- Preferred by nearly 40% of science and engineering researchers.
- Common among natural science method users.

# **Embedding RDM Key Competencies Into Data Management Plans**



Add your questions/comments  
to [padlet.com/laulumieli/DMP](https://padlet.com/laulumieli/DMP)



**UNIVERSITY  
OF TURKU**

# Assessing the Results

## Self-assessment of the learning

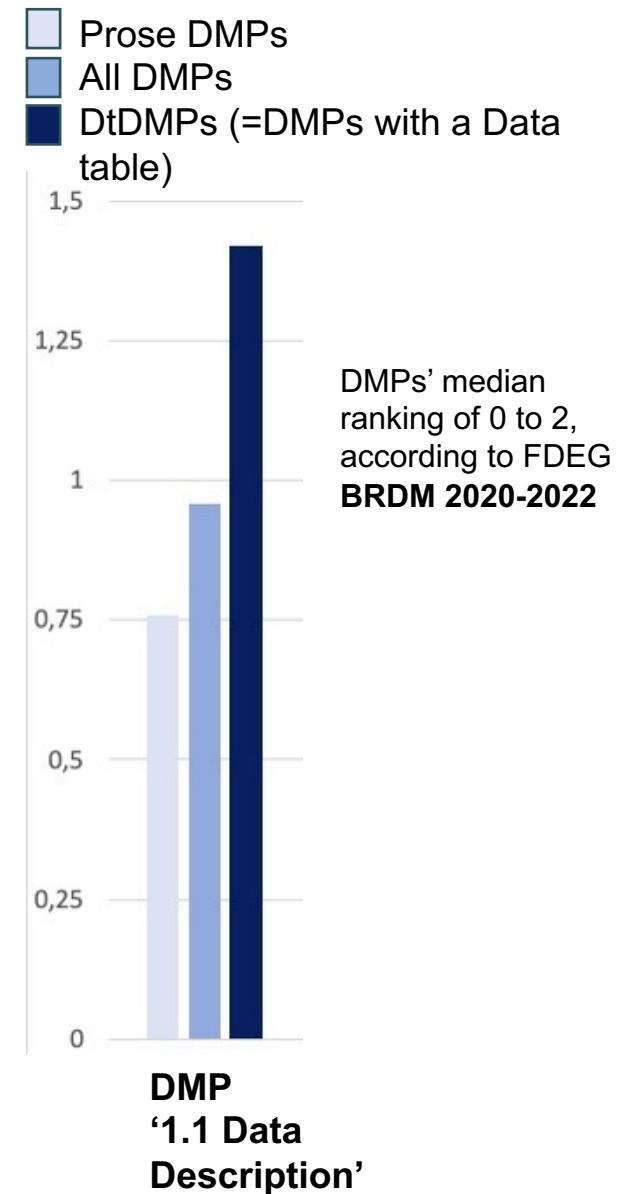
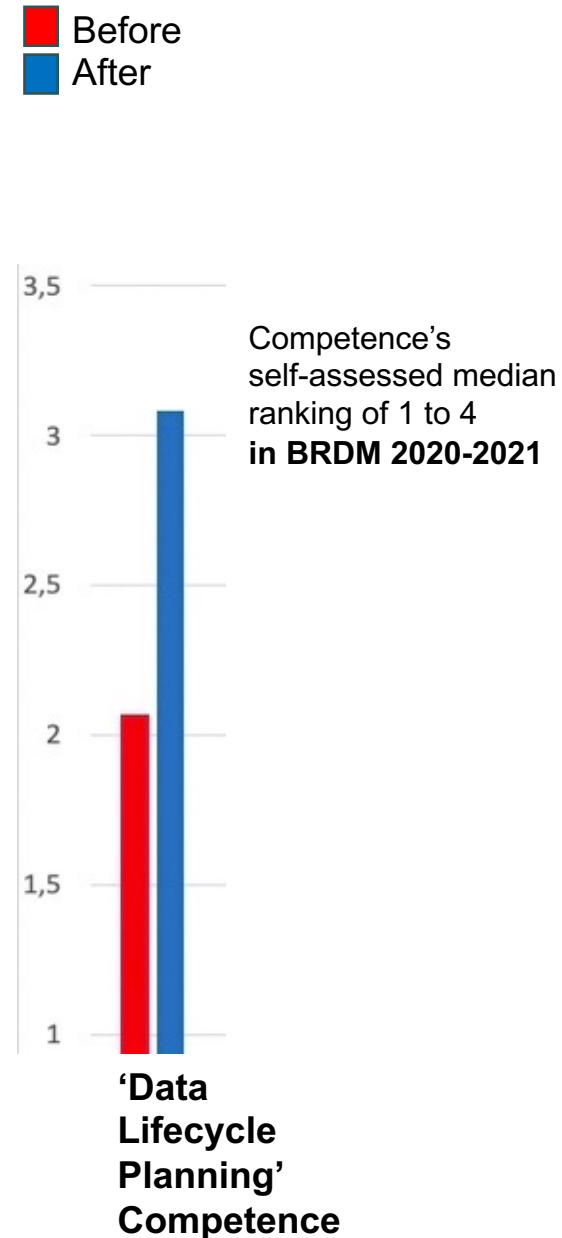
- Rating RDM competencies
  - Scale: 1=No competence; 4=very competent
- Giving course and module-based feedback
  - What are the three things you learned?
  - How do they change your practices?

## Assessment of the DMPs by the Teacher

- Rating Data Management Plans (DMPs)
  - [Finnish DMP Evaluation Guidance \(FDEG\)](#)
  - Scale: 0=Poor; 2=Excellent
- Content analysis of RDM best practices applied in DMPs

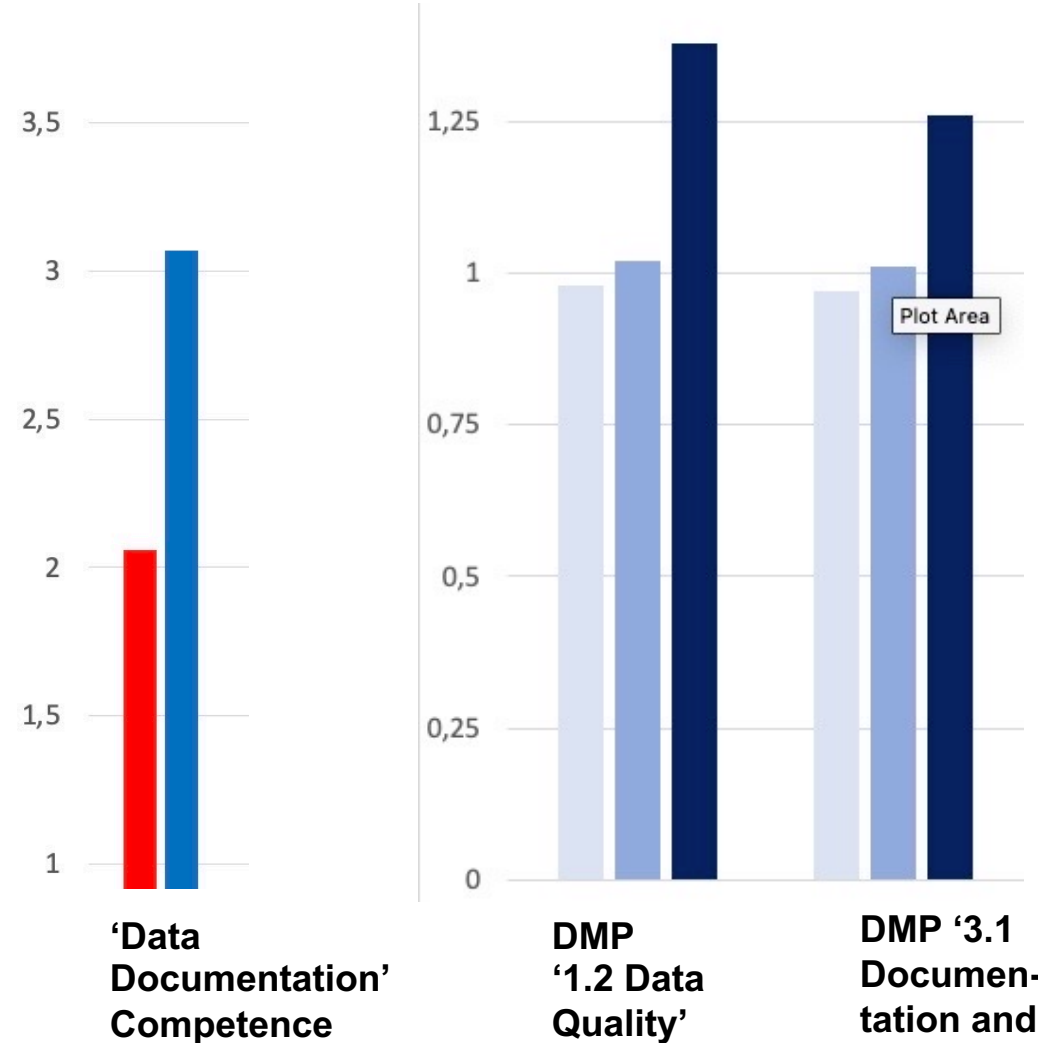
# Strengthening 'Data Lifecycle Planning' Proficiency

- **Self-assessment and Feedback**
  - Enhanced research quality via detailed DMPs using DMP-Tuuli, meeting ethical and funder standards.
- **DMP Assessment**
  - DMPs with Data Tables (= DtDMPs) outperformed prose DMPs in quality.
  - DtDMPs contained more detailed RDM descriptions, aligning with FAIR principles.
- **Next Steps**
  - Incorporating a data table in all DMPs to enhance data lifecycle descriptions.



# Enhancements in 'Data Documentation'

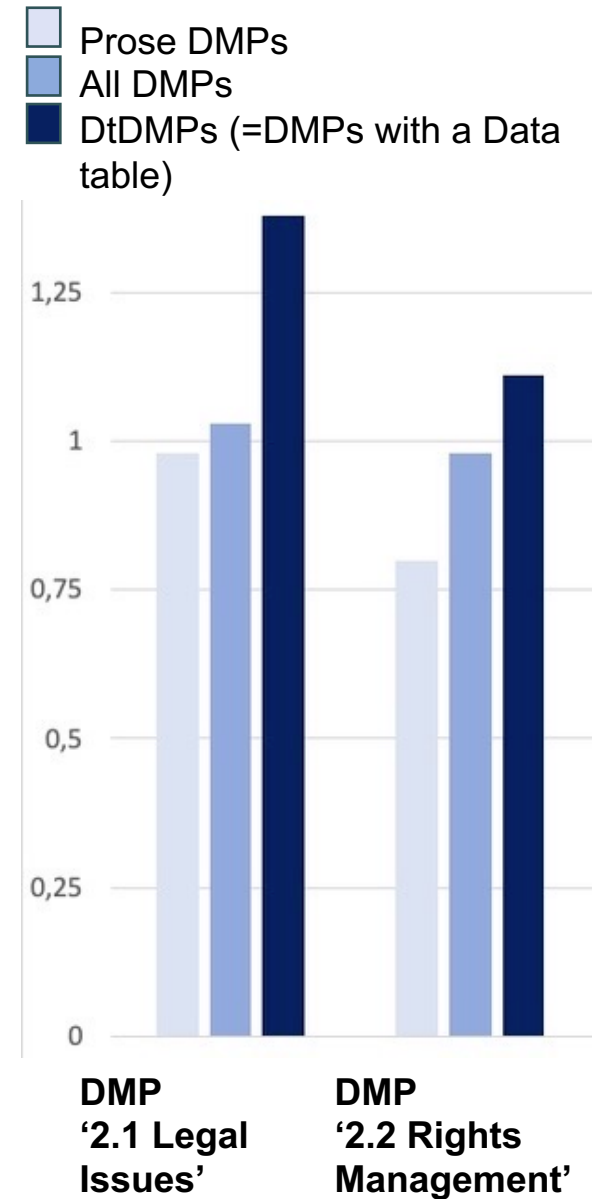
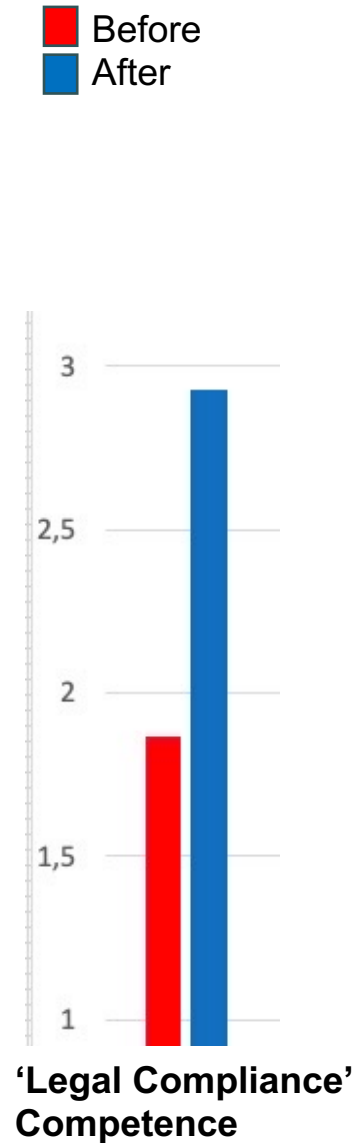
- **Self-assessment and Feedback**
  - Highlighted importance of detailed, clear records for data sharing, reuse, and preservation.
- **DMP Assessment**
  - DtDMPs scored highly: 1.38 for Data Quality, 1.26 for Documentation.
  - Structured descriptions in data tables boost documentation quality and FAIR compliance.
- **Next Steps**
  - Improve prose DMPs to enhance comprehensive and FAIR data documentation.





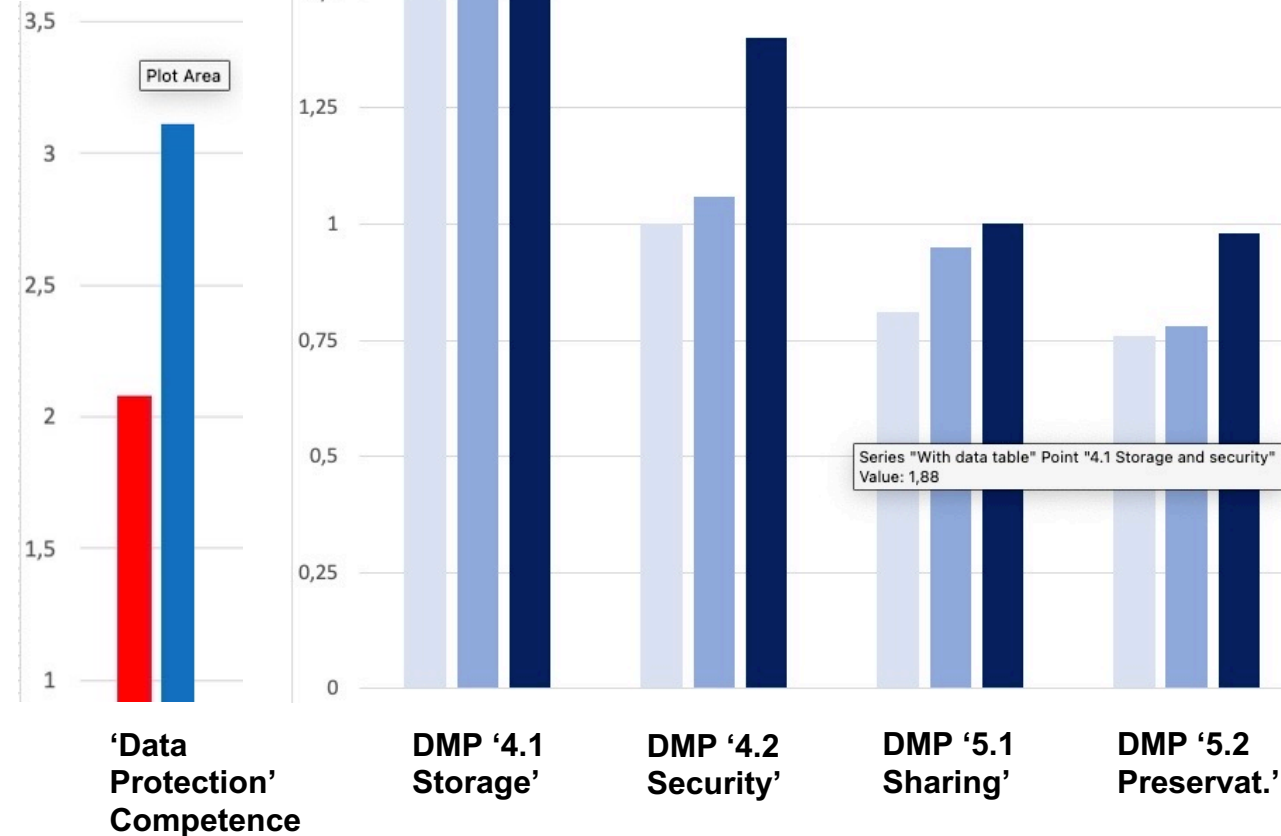
# Advancements in 'Legal Compliance'

- **Self-assessment and Feedback**
  - Gained detailed knowledge of IP rights, GDPR, anonymization, licensing, and consent.
- **DMP Assessment**
  - DtDMPs scored higher on legal issues (1.38) and rights management (1.11) compared to prose DMPs.
  - Demonstrated better handling of legal rights and personal data; formal data sharing methods enhanced license naming.
- **Next Steps**
  - Enhance inclusion of data sharing policies, permissions, and ownership justifications in prose DMPs.



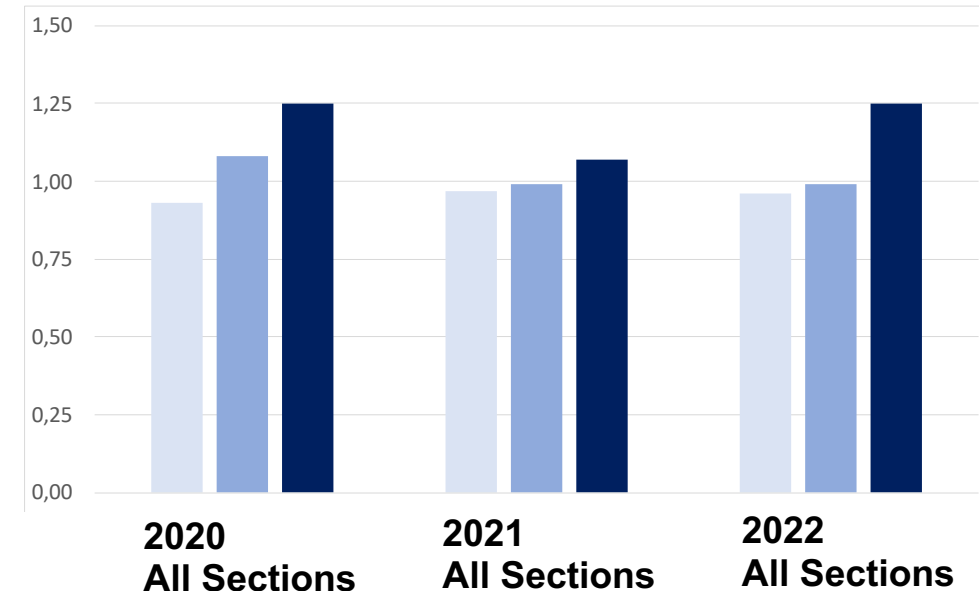
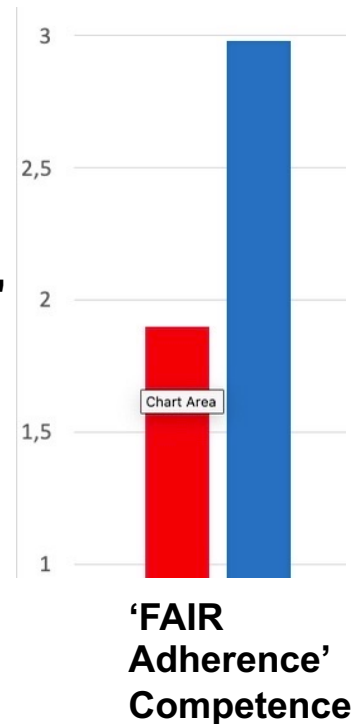
# Enhancing 'Data Protection Strategies'

- **Self-assessment and Feedback**
  - Participants learned about secure data storage, metadata importance, and applying FAIR principles.
- **DMP Assessment**
  - DtDMPs scored higher, demonstrating better expertise in secure data handling.
  - Stressed importance of encryption, choosing right storage solutions, and integrating FAIR principles.
- **Next Steps**
  - Address gaps in adherence to funders' policies, data rights management, and permissions handling.



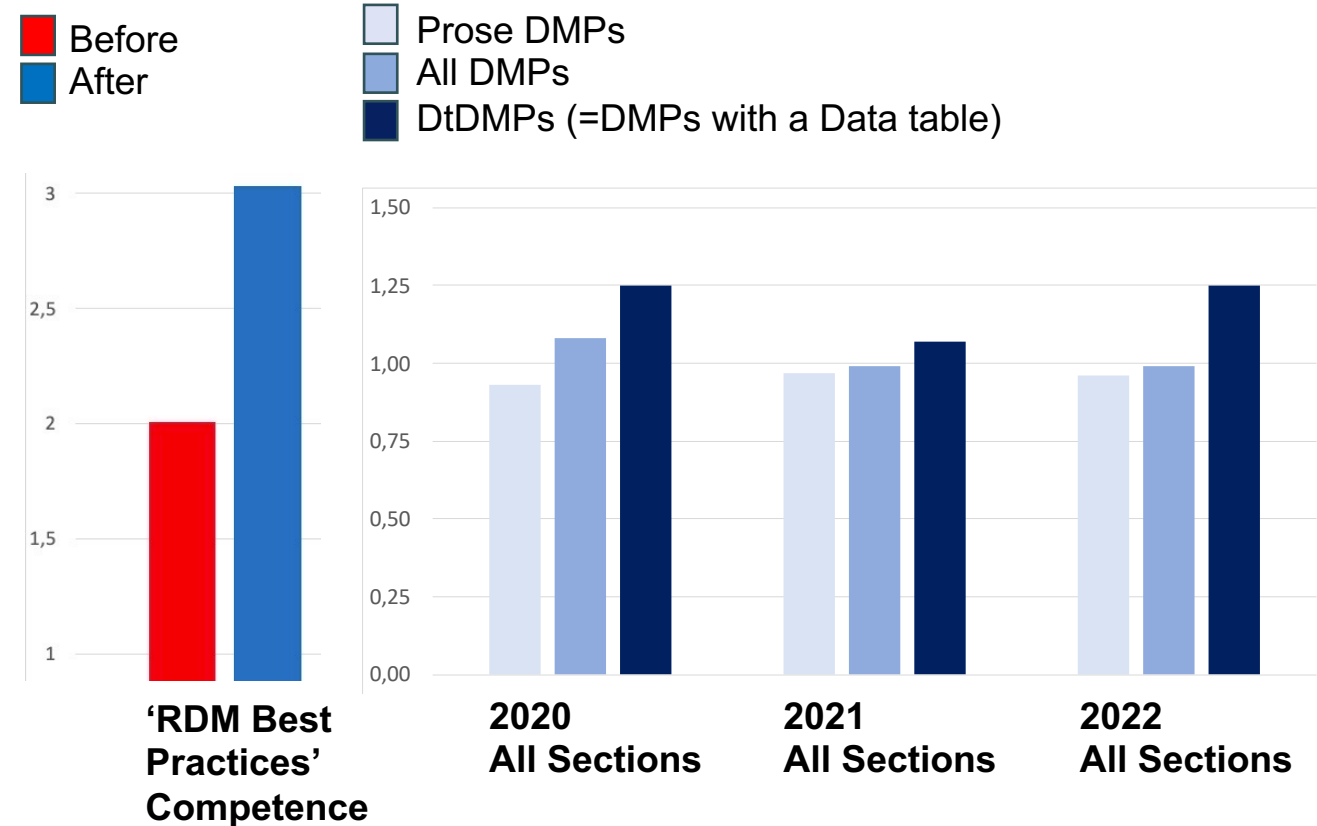
# Advancements in 'FAIR Adherence'

- **Self-assessment and Feedback**
  - Emphasized learning on metadata, data preservation, and open data practices.
- **DMP Assessment**
  - DtDMPs vs. Prose DMPs: Median FAIR scores from 1.07 to 1.25 vs. 0.93 to 0.97 across years.
  - Strong adoption of 'Findable' and 'Accessible' principles in DtDMPs, with detailed descriptions enhancing FAIR principles
- **Next Steps**
  - Focus on clarifying data sharing permissions and documenting data processing in prose DMPs.



# Enhanced Proficiency in 'RDM Best Practices'

- **Self-assessment and Feedback**
  - Enhanced skills in structured data management, robust documentation, and secure storage.
- **DMP Assessment**
  - DtDMPs outperformed prose in adherence to RDM best practices.
  - Stronger formal data sharing methods and clearer responsibilities distribution.
- **Next Steps**
  - Address the gap in budgeting RDM in DMPs by incorporating workload and cost estimates for RDM activities.



# Differences in Relation to Previous DMP Content Analyses

- Enhanced use of the institutional network drives and cloud services.
- Trend towards formal data sharing channels.
- Lower intent for data sharing
- Division of labour for data management tasks.
- Better identification and description of different data types and their needed actions through a data table.
- DMPs frequently fail to mention funders' or publishers' data sharing policies



# Recommendations

- **Researchers:** Adopt DMP evaluation criteria, such as FDEG, and structured data tables to enhance management and clarity of research data.
- **Institutions:** Use the evaluation criteria to assess DMP strengths/weaknesses, customizing support and training for researchers.
- **Educators:** Integrate the evaluation criteria into RDM training to standardize DMP evaluations and improve training impact.
- **Funders & Publishers:** Implement the requirements of the evaluation criteria for DMPs to ensure transparency and high data management standards.
- **All Stakeholders:** Promote evaluation framework adoption to standardize and elevate research data management practices.

# References

- Aalto, S., Ahokas, M., Friman, J., Fuchs, S., Korhonen, T., Kuusniemi, M. E., Laakso, K., Lennes, M., Manninen, S., Ojanen, M., Rantasaari, J., Virtanen, M. E., Xu, Q. (2021) Finnish DMP evaluation guidance + General Finnish DMP guidance. [Working paper]. Zenodo. <https://doi.org/10.5281/zenodo.4729831>
- Rantasaari, J. (2021). Doctoral students' educational needs in research data management: Perceived importance and current competencies. *International Journal of Digital Curation*, 16(1), 36. <https://doi.org/10.2218/IJDC.V16I1.684>
- Rantasaari, J. (2022). Doctoral students' research data management competencies based on the quality of their data management plans. In *Proceedings of the IATUL Conferences* (Vol. 42, pp. 1–9). Perth. Retrieved from <https://docs.lib.purdue.edu/iatul/2022/clr/4>
- Rantasaari, J. (2022). Multi-stakeholder research data management training as a tool to improve the quality, integrity, reliability and reproducibility of research. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, 32(1), 1–54. <https://doi.org/10.53377/LQ.11726>
- Rantasaari, J. (2024). Assessing quality variations in early career researchers' data management plans. *International Journal of Digital Curation*, 18(1), 29. <https://doi.org/10.2218/IJDC.V18I1.873>
- Rantasaari, J., & Kokkinen, H. (2019). Closing the skills gap: The basics of the research data management (BRDM) course: Case University of Turku. In *Proceedings of the IATUL Conferences* (pp. 1–8). Perth: 2019 IATUL Proceedings. Retrieved from <https://docs.lib.purdue.edu/iatul/2019/fair/5/>



**UNIVERSITY  
OF TURKU**

**Thank you!**

[jukka.rantasaari@utu.fi](mailto:jukka.rantasaari@utu.fi)