

HMC Survey Taskforce – Key Approaches & Insights

Markus Kubin^{1,*}, Lucas Kulla², Christine Lemster³, Stanislav Malinovskii³, Andreas Schmidt⁴, Sangeetha Shankar⁵, Karl-Uwe Stucky⁴

¹ Helmholtz-Zentrum Berlin für Materialien und Energie; ² Deutsches Krebsforschungszentrum (DKFZ) Heidelberg; ³ GEOMAR Helmholtz Centre for Ocean Research Kiel; ⁴ Karlsruhe Institute of Technology; ⁵ Institute of Transportation Systems, German Aerospace Center (DLR)

contact: markus.kubin (at) helmholtz-berlin.de

Overview & Mission

- **Monitoring:** HMC's Survey Taskforce monitors human factors regarding (meta)data management practices, gaps, and needs across Helmholtz.¹
- **Objective:** Inform HMC's strategy to improve the FAIR data ecosystem.
- **Target Group (2024):** Helmholtz staff working on data management tasks.²

Deliverables & Outputs

- The Taskforce has completed its following work packages (D1-D5).

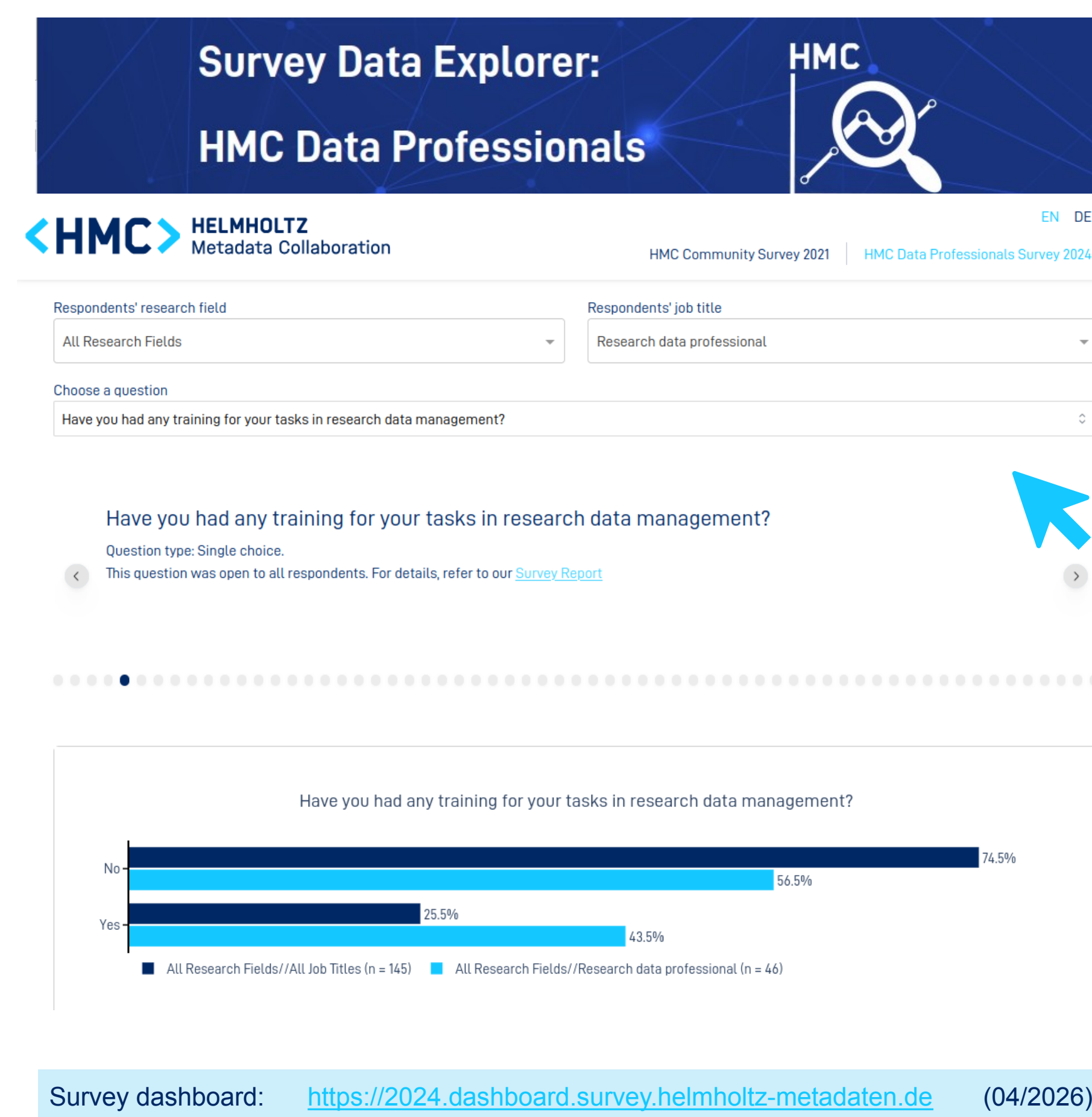
Deliverable	Description	Finalized
D1: Data Analysis	Analysis results shared within HMC. Analysis published.	11/2024 04/2025
D2: Workshop	Interactive workshop with 50 attendees.	12/2024
D3: Report	Comprehensive survey report.	04/2025
D4: Public Dataset	Interoperable survey dataset to be published at GESIS.	2026 (submitted, under review)
D5: Survey Dashboard	Interactive visualization of survey data for all 6 research fields & professions.	11/2025
+ Outreach	Presentations; posters; workshops. Social media campaign (8 weeks).	2024-2026 03/2026

Key Approaches & Methodology

- **Hypothesis-Driven Design:** Survey design started with hypotheses (e.g., training rates, FAIR alignment) leading to targeted questions.²
- **FAIR-Aligned:** FAIRaware questionnaire for self-assessment adopted.³
- **Reusable Workflows:** Data analysis pipelines from the HMC Community Survey 2021 were adapted and reused.²
- **LLM Usage:** Large Language Models (LLMs) were applied for data cleaning (saving several weeks of work).²
- **Open Science:** Analysis code & survey data prepared for open access.

Survey Dashboard & Live Demo

- **Get interactive:** Explore survey data by research field & professional role.



Survey dashboard: <https://2024.dashboard.survey.helmholtz-metadaten.de> (04/2026)



Selected Key Findings

- **Resources:** Majority reports lack of resources (personnel, time, funding).
- **Training:** Minority received formal training; high interest in training.²
- **FAIR Practices:** Metadata practices strong in Findability, but less pronounced in Interoperability and Reusability.²
- **Policies:** Majority aware of center policies, but minority feels that sufficient assistance for implementation is provided.²

Discussion & Ideas on Future Directions

- **Suggestions:** (1) Annually recurrent survey for longitudinal analysis. (2) Development of fully automated analysis pipeline. (3) Combined question catalog with 2021 survey for broader target groups.
- **Your Feedback:** (1) What topics should the next survey cover? (2) How can HMC's survey activities better serve your needs? (3) What strategic measures should HMC prioritize?

[1] Helmholtz Metadata Collaboration; A survey on research data management practices among researchers in the Helmholtz Association (2022). https://doi.org/10.3289/HMC_publ_05

[2] Helmholtz Metadata Collaboration; A survey on the status quo, gaps and needs among research data professionals in the Helmholtz Association (2025). https://doi.org/10.3289/HMC_publ_08

[3] Data Archiving and Networked Service. FAIR-Aware: Your first step towards your FAIR dataset (2021). URL: <https://fairaware.dans.knaw.nl/> (accessed April 23, 2026)

