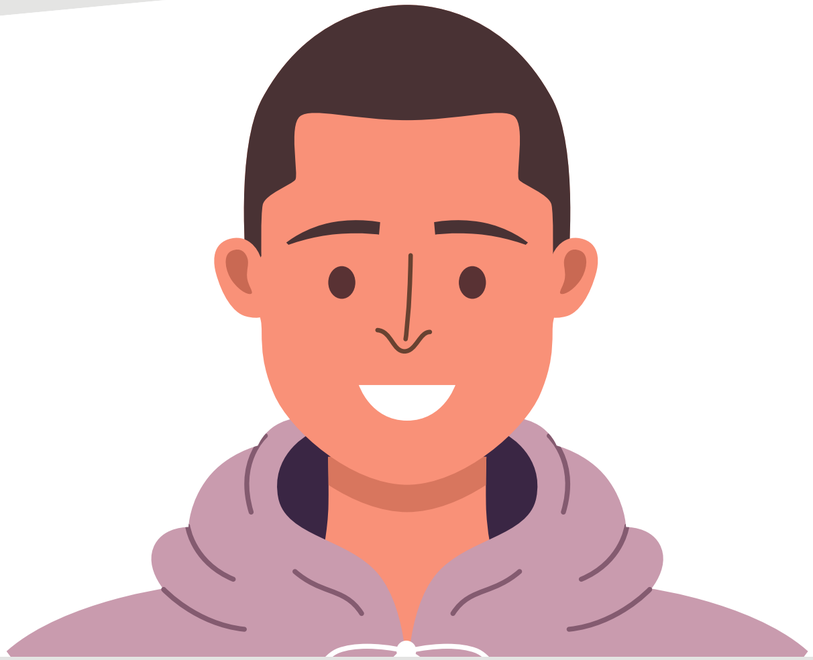
Tomáš Havel
Team/Project Data Steward



“I like when data are truly of high quality. When a researcher opens a dataset and everything fits — formats, metadata, code.”

Challenges: lack of institutional recognition for the technical role, fragmented practices across teams, high workload and time pressure

My goal: make data practically reusable, improve data quality and clarity

What I need: better connection between researchers and IT, space to share good practices

Job Description

I support research teams to ensure their data are high-quality, well-structured and reusable.
I work directly with teams at the faculty or on individual projects, providing technical support for FAIR principles, from data cleaning to repository setup.
I collaborate with IT services and the library to make sure workflows run efficiently and meet required standards.

Example from practice

Validation and Cleaning of Research Data
I contribute to processing large datasets within disciplinary research projects. I help researchers prepare data for sharing, check formats, enrich metadata and create simple scripts to automate repetitive tasks.
Outcome: a dataset that meets FAIR principles, is easy to publish and can be reused for further analysis.

MY PROFILE

PROXIMITY TO RESEARCH
Administration — **Research**

IT SKILLS
MS Office — **Programming**

SCOPE
Team — **Institution**

ROLE FOCUS
Working with data — **Training**

Main Tasks

- **Preparation and validation of data**
- **Cleaning of datasets**
- **Repository setup and management**
- **Workflow automation**
- **Collaboration with research teams on data quality**

Tools and Competencies

Python
R
GitHub
OpenRefine
Power BI

- **FAIR principles**
- **Format validation and pipeline development**
- **Work with disciplinary repositories**
- **Team collaboration and communication with IT**

Impact / Contribution

For researchers:
Helps prepare data that are clean, structured and ready to share, ensures datasets are readable, reliable, and FAIR-compliant.

For Institution:
Supports automation that increases efficiency and sustainability of data workflows, and strengthens compliance with FAIR standards.

Research Data Specialist – technical and practical support for research teams (represents approx. 21% of data stewards in the Czech Republic).

Use Case 2
Team/Project Data Steward

Role of the Data Steward

- **Supports** research teams with data preparation, cleaning and validation.
- **Translates** technical procedures into clear steps for researchers.
- **Automates** repetitive tasks and sets up workflows for data management.
- **Ensures** that data meet FAIR principles and are reusable.
- **Collaborates** with the library, IT, and project teams on repository management.

Competencies

- Knowledge of FAIR principles, metadata and data formats.
- Ability to validate, transform and clean large datasets.
- Experience with versioning and process automation.
- Collaboration with researchers and support in solving technical challenges.

Tools

Python	OpenRefine	OSF
R	Power BI	LINDAT
GitHub	Zenodo	DSpace

Validation, Cleaning and Preparation of Research Data for Sharing and Reuse.

Scenario A research team is completing a project focused on sociodemographic and behavioural data. Their goal is to publish a dataset that meets FAIR principles, and fulfils the requirements of the funder. However, the data come from multiple sources, vary in format and contain incomplete metadata.

The team/project data steward helps the researchers resolve technical and organisational issues. Using Python and OpenRefine, they create scripts for data cleaning and validation. Together with the researchers, they set up a workflow in GitHub for sharing code and versioning files. The steward checks that the dataset descriptions meet funder requirements and that the dataset is stored correctly in the chosen repository.
As a result, the project outputs are clear, reusable and prepared for further research.

Benefits for the Institution

Improves the quality and transparency of data outputs.

Ensures that projects comply with FAIR principles and funder requirements.

Helps standardise data management practices across faculties and teams.

Benefits for Researchers

Simplifies data preparation and publication by reducing technical errors.

Saves time through automation and clear workflows.

Ensures reliability, quality, and trustworthiness of datasets.

Provides direct technical and methodological support within research projects.

Research Data Specialist – technical and practical support for research teams (represents approx. 21% of data stewards in the Czech Republic).