
Data Steward

Who is a Data Steward?

A Data Steward is a new position supporting research data management within a scientific team. This role is responsible for maintaining the quality, integrity, and access measures of data and metadata in a manner that complies with applicable laws, institutional policies, and individual permissions. The goal of the Data Steward is to ensure that data at all stages of the research cycle (i.e. planning, collection, processing, analysing, preservation, sharing, and reuse of data - see Figure 1) is handled appropriately. This includes overseeing data management processes, managing data quality, and ensuring compliance with data protection and security regulations.



Data Stewards establish connections and open communication channels between researchers, policymakers, software developers, and infrastructure providers (both within and outside the institution) to successfully implement the necessary elements that enable researchers to adopt modern Research Data Management (RDM) practices.

With the increasing demands for research data management and publication, it is advisable to ensure sufficient competence in this area – all researchers don't need to become experts in data organization and management. It is enough if they have access to someone who can help them set up and adhere to proper principles. The goal for scientific

teams is to work with data to ensure their integrity, security, and to fulfill the [FAIR principles](#) to the greatest possible extent.

Role of a Data Steward within a scientific team

A Data Steward serves as a bridge between various parts of the institution. They support and closely collaborate with the main data creators and users (researchers, students, etc.). They provide researchers with advice, ensure that data management adheres to institutional policies, and may also engage in practical work within the project.

Core roles of a Data Steward

- Direct support for research Data Management. Providing assistance and guidance in managing research data throughout its lifecycle.
- Selection of appropriate data and metadata formats. Advising on and helping choose suitable data and descriptive metadata formats to ensure data usability and longevity.
- Assisting in setting up workflows and practices so that created data comply with FAIR (Findable, Accessible, Interoperable, Reusable) principles to the greatest extent possible.
- Promoting the principle of making data "*as open as possible, as closed as necessary*" thus balancing transparency with confidentiality requirements.
- Helping to identify legitimate obstacles to open data publication and determining which data can be made openly accessible.
- Participating in the planning of project data management and the creation of Data Management Plans (DMPs) and maintaining the DMPs up to date throughout the project lifecycle.

While the Data Steward's role does not include direct data analysis, they should assist the team in ensuring that the data can be reproducibly analyzed.

Key competencies

Personal predispositions

- Active knowledge of English including specialized terminology in the research field.
- Basic computer literacy with a very positive attitude towards IT.
- Organizational skills including the ability to algorithmize processes.

Internal environment

- Optimal knowledge of the research topic and the processes involved.
- Knowledge of available tools and procedures for data management storage, backup, established best practices, etc. (may be specific to each workplace).
- Detailed knowledge of [FAIR principles](#) and the ability to implement them in practice.
- Knowledge of "*best practices*" in research data management at Charles University and other related institutions.

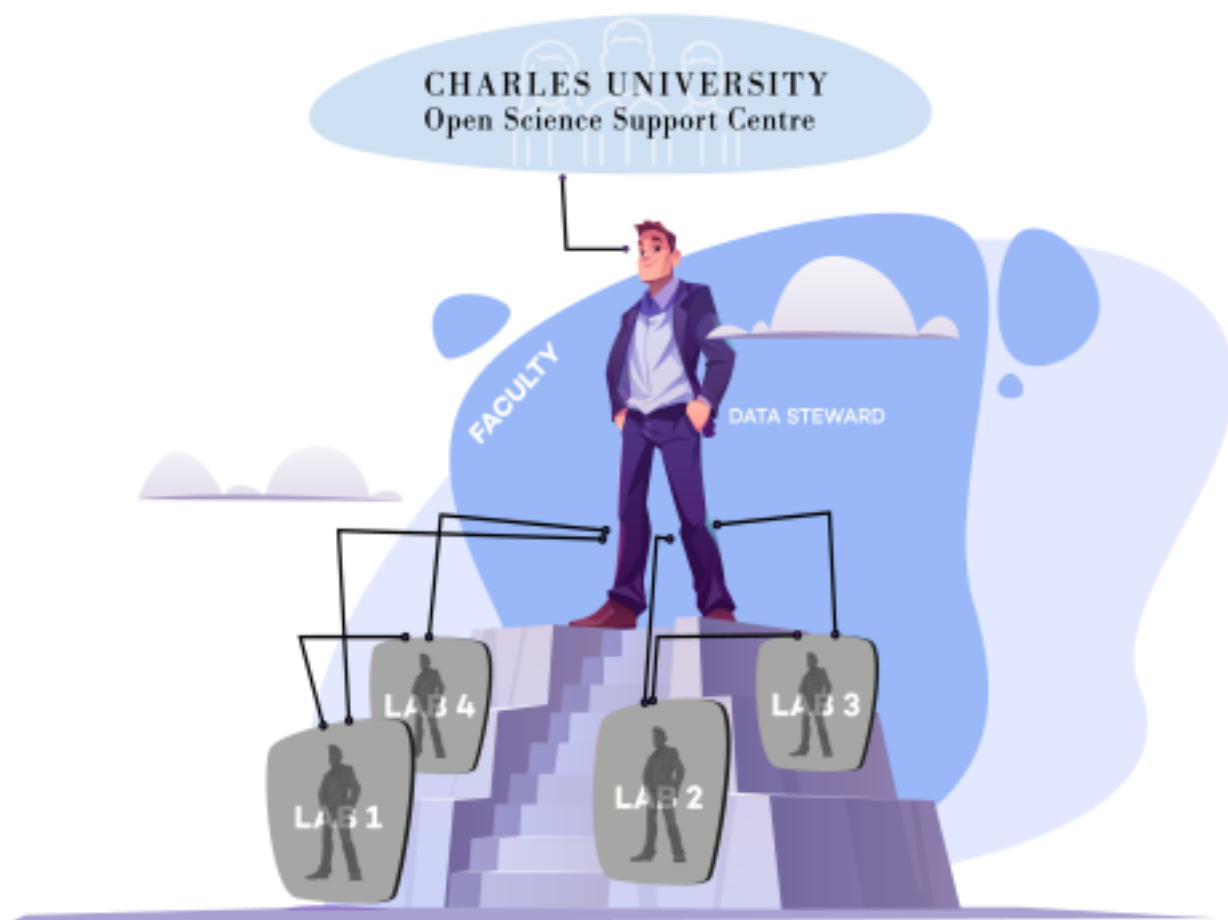
Data Stewards in the structure of Charles University

The formal position of Data Steward is new at Charles University. However, this may rather formalize activities that already occur as part of established best practices at many workplaces. The introduction of this role highlights the importance of data management and helps to more precisely define specific responsibilities.

For the role of Data Steward, two factors are particularly important:

1. Knowledge of current trends, requirements, and possibilities.
2. The ability to implement them into everyday practice.

We therefore assume that individual Data Stewards should be as close as possible to specific research teams and should understand (to some extent) the ongoing research in the context of data. The [Central Library of Charles University](#) should become an appropriate platform for communication among Data Stewards across the University, especially the [Open Science Support Centre](#). The entire scheme of integrating Data Stewards into the structure of CU is summarized in Figure 2. The Central Library can provide initial training for new Data Stewards at a general level (introduction to FAIR principles, examples of good and bad practices in data management, principles of creating data management plans, data opening, and sharing) and subsequently help Data Stewards to follow current trends and provide them with systematic and ad hoc support.



Given that the education of Data Stewards is not yet formalized and it cannot be expected that there are already "ready-made" experts. We recommend utilizing existing team members (e.g., PhD students) who can expand their current qualifications in this direction (with the support of the Central Library of Charles University) and simultaneously effectively use their knowledge of the environment.

Contact

Data Stewards at Charles University

Individual faculties at Charles University have their own Data Stewards, who are available to the faculty researchers or academic staff. A list of their contacts is available on the website of the Center for Open Science.

Data Stewards Community in the Czech Republic

If you are involved in data management or are interested in this topic, you can join the emerging Data Stewards community in the Czech Republic. This initiative was launched alongside the emerging EOSC project, which includes, among other things, building a community around data management. Discord has been chosen as the communication channel. You can join the group through this link: <https://discord.gg/eMpzXFRaPn>.

Open Science Support Centre

For questions regarding Data Stewards and data management issues, you can also contact the Open Science Support Centre.

Email: openscience@cuni.cz