Data Steward Certificate Program

of the University of Vienna

Module Handbook 2023/2024

Table of Contents

OBJECTIVES AND QUALIFICATION PROFILE

Objectives

Research data, as valuable research results, require professional management. Comprehensive research data management is therefore becoming increasingly important. This is what data stewards are responsible for. The role is being established not only in the business world but also at research institutions in Austria and abroad as international best practice in research data management.

Data stewards operate at the interface between scientists and scholars, and research infrastructure and help bridge the gap between these two areas. They support researchers in processing data in a sustainable way and carry out needs assessments and requirements engineering. This certificate program will help the participants to acquire knowledge, expertise and key competences to perform tasks as data stewards in research institutions.

This continuing education and training program links the latest findings on research data management, open science and open research with the tasks of data stewards.

Qualification Profile

Successful participants have profound knowledge of ways to develop innovative services in the field of research data management and increase awareness of open science. They are able to systematically assess researchers' needs and offer tailored support. Together with researchers and software developers, they design innovative workflows for FAIR data management. The acquired competences enable the participants to fulfil the forward-looking role of data stewards at various research institutions.

MODULE OVERVIEW

Data Steward Certificate Program

Semester I and II	ECTS	SWS[1]
Basics of Research Data Management and Open Science	2	2
Basics of IT and Data Science	3	2
FAIR Research Data in the Life Cycle	6	4
Research Data Management Support	2	1
Data Stewardship in Practice: Project Work	2	
Total	15	9

---

[1] SWS = Semesterwochenstunde = Hours per week per semester, see definition of the University of Vienna

Module 1: Basics of Research Data Management and Open Science

Course Dates: 16.-20.10.2023 in Vienna

Workload: 2 ECTS (= 50 hours)

- Classes: 22.5 hours (= 30 lessons x 45 minutes)
- Self-Study: 27.5 hours

Contents:

Introduction to:
- Research Data Management and Open Science
- Role and tasks of data stewards
- Good scientific practice
- Legal and ethical aspects
- Cost estimation and financing models

Learning Outcomes:

This module introduces the basics of research data management and open science. Upon completion, the students:

- Understand the most important concepts of research data management, open science and research as well as data literacy

- Understand the rules of good scientific practice, as well as ethical aspects and the legal framework of data management

- Create budget plans and financing models for the management and permanent storage of research data

- Analyze the role and typical tasks of data stewards

Module 2: Basics of IT and Data Science

Course Dates: 16.-17.11.2023, 14.-15.12.2023 online

Workload: 3 ECTS (= 75 hours)

- Classes: 22.5 hours (= 30 lessons x 45 minutes)
- Self-Study: 52.5 hours

Contents:

- Introduction to
    - Data Science and data-driven research
    - Machine learning
    - Database systems
    - Programming concepts
- Basics of programming:
    - Unix Shell (based on the Carpentries curriculum)
    - Git and GitHub (based on the Carpentries curriculum)
    - Working with Python (based on the Carpentries curriculum)

Learning Outcomes:

This module covers the fundamentals of IT applications and data science in terms of their potential regarding data-driven research. Upon completion, the students will be able to:

- Understand the most important concepts in Data Science, Machine Learning, and Database Systems

- Describe the basic concepts of programming

- Know the basics of structured programming (e.g. data types, control structures, subroutines, functions …)

- Describe the most important programming languages and their different aspects

- Choose a suitable Integrated Development Environment (IDE) for developing own programs

- Work in the Unix Shell using the most important commands for the file system and data files

- Use Git/GitHub for setting up own projects, share them and find other projects

- Develop small programs in Python using basic data types, control structures, functions and software libraries

Module 3: FAIR Research Data in the Life Cycle

Course Dates: 18.-19.01.2024, 15.-16.02.2024, 14.-15.03.2024, 11.-12.04.2024 online

Workload: 6 ECTS (= 150 hours)

- Classes: 45 hours (= 60 lessons x 45 minutes)
- Self-Study: 105 hours

Contents:

Research Data Life Cycle:
- Project management and funding landscape
- Data management plans (DMPs)
- Organizing and structuring data
- Processing and cleansing data
- Data analysis and visualization / tools for data-driven research
- Metadata and documentation of research data
- Persistent identifiers, ontologies, semantic web
- Data security and storage
- Repository management and long-term preservation
- Interoperability and data migration
- Data reuse incl. legal requirements

Domain-Specific Approaches to Data Stewardship
- in the Natural and Life Sciences
- in the Humanities
- in the Social Sciences
- in the Technical Sciences

Learning Outcomes:

This module maps the typical life cycle of research (data) – planning research projects, organizing and structuring, analyzing, sharing, publishing, preserving, and reusing data. Students will analyze research data management practices along this life cycle, and learn to address general and discipline-specific requirements.

Upon completion, the students will be able to:

- Understand the (inter)national funding landscape and funding regulations with regard to research data management. They identify the basics of project management in research, and evaluate data management plans

- Structure, organize and cleanse research data and apply data visualization methods

- Understand various metadata standards and apply these standards in indexing and describing research data

- Identify issues of data security and data migration and apply data security measures

- Analyze repository and data archive management, and assess long-term archiving processes

- Apply various strategies for data search and correctly cite and reuse research data in a legally compliant manner

- Understand disciplinary differences in data stewardship and research data management as well as identify and apply discipline-specific best practices

Module 4: Research Data Management Support

Course Dates: 02.-03.05.2024 online

Workload: 2 ECTS (= 50 hours)

- Classes: 11.25 hours (= 15 lessons x 45 minutes)
- Self-Study: 38.75 hours

Contents:

- Consulting and RDM Support Services
- Basics of Teaching and Knowledge Transfer, Open Education
- RDM Training in Practice
- Needs Assessment
- Requirements Engineering

Learning Outcomes:

In this module, the acquired expertise from the previous three modules is transferred to the future role of a data steward.

Upon completion, the students will be able to:

- Understand the variety of research data management services and analyze their potential in target group-specific application

- Practice various methods of knowledge transfer with a special focus on consulting and training

- Apply methods of needs assessment and requirements engineering

Module 5: Data Stewardship in Practice: Project Work

Course Dates: May-July 2024 online

Workload: 2 ECTS (= 50 hours)

- Self-Study: 50 hours

Contents:

- The final project – consisting of the project deliverable and the accompanying project report – corresponds to a workload of 2 ECTS (50 hours) per person. The students can complete the project either individually or in a team of a maximum of three students.

- The project topics are selected from the various fields in the course list, approved by the project module lead and supervised by teaching staff. The topics should allow the practical application of the skills and knowledge gained from classes in the previous modules.

- Topics include, but are not limited to: Concept for training or training materials addressing different aspects of RDM, explanatory video about RDM, RDM needs assessment of an institution, RDM advising strategy for a (small) university, development and documentation of RDM workflows, …

Learning Outcomes:

This module allows students to work with data stewardship in practice.
Upon completion, the students will be able to:

- Analyze and assess the demand for data stewardship in the community
- Create projects that address the demand for data stewardship in the community
- Analyze the target audience of data stewardship projects
- Carry out a project in the area of data stewardship
- Report on a project in the area of data stewardship