

# NBIS DM team training Strategy 2025 – 2028

DOI	<a href="https://doi.org/10.17044/scilifelab.27606417.v3">10.17044/scilifelab.27606417.v3</a>
Version	3.1
Last updated by	Elin Kronander & DM unit
Approval date	2025-01-23
Time for next revision	2025-10

<b>Introduction</b>	<b>1</b>
<b>2025-2028 Roadmap</b>	<b>2</b>
<b>2025 Work plan</b>	<b>3</b>
<b>Definitions</b>	<b>4</b>

## Introduction

This document describes the general Data Management Team Training Strategy, with the aims and explicit goals for the team training efforts to be realised during 2025. By the end of 2025, the current training strategy will be evaluated and updated as a continued strategy for the following year. The revised strategy document will be versioned and published along with the previous one(s).

*Mission statement for training* - Swedish Life Science researchers shall have access to training in efficient Data Management practices, concepts and practicals that can be applied in their daily work. The Data Management achievements and efforts at NBIS should be shared with interested collaborators within the ELIXIR network.

Approaches, activities and resources defined as necessary for fulfilling delivery:

- Provide Data Management and Open Science training for researchers
  - Courses
  - Workshops
- Provide guidelines and documentation for Data Management “best practices”
  - SciLifeLab RDM guidelines (collaboration with Data Centre)
  - ELIXIR RDMkit (collaboration with Data Centre)

- Promote awareness of the importance of Open Science, Reproducible Research, and FAIR in the Swedish Life Science community
  - Outreach
  - Community building
  - Networking

## Target audience

The primary target audience for the Data Management training is primarily researchers and staff active in Swedish Life Science research. However, several actors on both national and international level who work in close proximity to, or in collaboration with, the primary target audience can become secondary parts of the NBIS DM training efforts. As NBIS engagement with third-party actors (e.g. health care regions and industry) expands we will broaden our target audience to include these stakeholders.

## 2025-2028 Roadmap

NBIS Data Management training should primarily be targeted to a broad audience in Sweden and aim to interact with researchers early in their careers to increase its impact. The responsibility of educating PhD/PostDocs in entry-level RDM skills should eventually be directed towards the Universities, but the immediate need to raise general awareness can and should be addressed. The course catalogue should be expanded to contain courses targeted at intermediate technical skills first, followed by intermediate RDM skills. There is no aim to provide courses at an advanced level for either skill set during this period.

### General objectives:

- It should be possible for other actors to adapt and deliver our course material themselves
- We should be available to assist with co-delivery or co-creation of teaching material for both national and international actors
- All training material should also be made available as self-study material
- All training material should follow the [ELIXIR FAIR training standards](#), and we should set an example of FAIR training materials to other NBIS teams
- All members of the Data Management team should be encouraged to participate in external (to the DM team) courses to keep themselves upskilled on data management topics.
- All NBIS courses should have adapted integrated data management content that fits the aims and outcomes of that particular course

**Courses** should primarily be targeted towards PhD/Postdocs, PI's, Infrastructure staff and NBIS/Bioinformaticians with a distribution of content in regards to skill as follows:

- [RDM skills](#): Beginner level (80%) & Intermediate level (20%)
- [Technical skills](#): Beginner level (50%) & intermediate level (50%)

**Outreach** should be targeted to a broad audience with a distribution in regards to role as follows:

- 50% - PhD/PostDocs
- 25% - PI's
- 15% - Infrastructure staff
- 10% - National and international stakeholders

### **Collaboration and community building:**

Community building efforts should be targeted towards other national actors in Data management with a focus on Life Science. This work should be done in collaboration with SciLifeLab Data Centre and discussed with SND. As national actors with a large contact area to research communities, NBIS DM training should also act to bridge Swedish DM communities to relevant ELIXIR, RDA, and EOSC communities.

The Data Management and Open Science training efforts should be done in collaboration with and aligned with the SciLifeLab Training Hub and Data Centre to support the training needs of the Life Science and DDLS community.

### **Guidelines:**

The main resources for guidelines and best practices should be SciLifeLab RDM-Guidelines and ELIXIR RDMkit.

## 2025 Work plan

The NBIS Data Management training activities for 2025 should have an emphasis on outreach and community building. We should increase our interaction with national actors in Data Management and promote all training activities that benefit our target audience. In addition, pilot and establish training content targeted towards intermediate skills.

### **Courses and workshops**

January: FitSM Foundation certification course (*Must*)

February: FAIRification workshop organised with SIMPLER (*Must*)

April: Intro to DM-practices (Updated curriculum) (*Must*)

April: Best Practices for Reference Genome Generation (UGC) (*Must*)

May: Open Science in the Swedish context (*Must*)

October: Training Materials made FAIR by Design (*Could*)

October: Principles and tools for FAIR Research Practices - DDLS Research school (*Must*)

Autumn: Working with human data modular course (e.g GDPR, Anon/pseudonymising tab data, deposit data i FEGA-SE) (*Should*)

Autumn: Introduction to DM practices for infrastructure units in Gothenburg (*Could*)

Autumn: Introduction to DM practices for IBCGO research school in Lund (*Could*)

During 2025: FAIRification workshop organised with ELIXIR's Interoperability Platform (*Should*)

During 2025: Content development for Introduction to DM for PIs (*Should*)

### Outreach and community building

- Roadshow with Data Centre to 2 national sites (Göteborg and Linköping) (*Must*)
- Meet with RDOs/DAUs at Lund, UmU, KTH, and KI (*Must*)
- Participate at/contribute to RDA plenary (*1 Must, 1 Should*)
- Contribute to EOSC Winter school (*Must*)
- Participate in SND-DAU networking events (*Must*)
- [RDA in Sweden event with SND in March](#) (*Must*)
- Organise SciLifeLab Research Data Management webinar series with Data Centre (*Must*)
- Co-organise event during open access week/open science festival (*Could*)
- Checklist of DM considerations that should be mentioned in other NBIS courses (*Should*)
- Participate in GA4GH 13th Plenary activity (*Must*)
- Participate in DDLS research area symposiums (*Could*)
- Provide a seminar in the ICM seminar series (*Could*)
- Organise ELIXIR's Interoperability Platform F2F (*Could*)
- Participate in Mass Spectrometry conference in Umeå (*Could*)
- Participate in INTOXICOM workshop in Uppsala (*Could*)

### RDM Guidelines

SciLifeLab RDM Guidelines:

- Organise contentathons to expand content (with DC) (*Must*)
- Create new 3 new topic pages (*Should*)

ELIXIR RDM ecosystem resources:

- Participate in contentathons (*Should*)
- Update Swedish page on RDMkit (*Must*)

## Definitions

### Role/function

#### PhD/Postdocs

PhDs and postdocs at all stages, working more or less under supervision from a PI.

#### PI's (including staff scientists in research groups)

Project leaders and associated staff with overarching responsibilities in more permanent positions in research groups.

**Infrastructure staff**

Data Scientists and staff scientists working at core facilities, including staff scientists active at national data platforms within the SciLifeLab

**NBIS/Bioinformaticians**

Analytical staff at NBIS, or staff with analytical responsibilities in research groups.

**Other national actors in Data management**

Groups and organisations with particular DM interests, such as University DAU's/SND/particular staff at Data Science Nodes

**Other international actors (ELIXIR)**

Engagement with actors in ELIXIR and CONVERGE activities and other ELIXIR nodes and related tasks.

## RDM skill levels

**Beginner**

New to RDM concepts

**Intermediate**

Users know Open Science and FAIR principles, have experience from DMPs and apply (some) data organisation practices.

**Advanced**

Routinely deposits FAIR data using relevant metadata standards, work Reproducible, applies good data organisation practices. Strives to improve or invent new, tailored RDM processes.

## Technical skill levels

**Beginner**

Point-and-clickers. Users new to terminal and command line interfaces. Users wanting to learn new software skills and/or softwares outside their current toolbox

**Intermediate**

Users with basic knowledge and skills in softwares and scripting but not able to single-handedly solve more complicated tasks. Comfortable, but not versatile, in digital environments and with some scripting/command line experiences

**Advanced**

Users are proficient in scripting and programming. Well-versatile method developers are comfortable in various computer environments.

## Formats of delivery

**Courses** (training events at least 1-full day)

**Workshops** (single hours to half-day events, including hands-on elements)

**Webinars** (online)

**Lectures** (on site)

**Handouts** (electronic and physical information materials)

**Web info** (structured web-based information)

**Self-study material** (persistent online material)

## Target topics and skills

- DMPs
- FAIR principles
- Open science
- Data organisation
- Documentation
  - Lab notes
  - README's
  - Protocols
  - Metadata
- Reproducibility
  - Scripted analysis
  - Version control
  - Workflows
- Data availability and Data publication
- Information security (Confidentiality, Integrity, Availability)
- Ethical and Legal Implications
- Sensitive (human) data publications (FEGA)
- Long-term storage
- Data re-use
- Data Stewardship onboarding / (community building)
- PIDs