

# SciLifeLab Data Management support and services

*Introduction to Data Management Practices for NGL*

NBIS DM Team

data-management@scilifelab.se

<https://doi.org/10.17044/scilifelab.c.6820587>





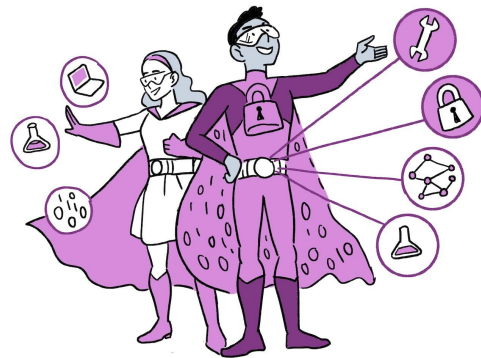
Help **maximize impact** of SciLifeLab and other data

Promoting FAIR, open science, and good data management practises throughout the data lifecycle

Provide services and resources for data management, IT and data sharing

Make bioinformatics support and training easily accessible

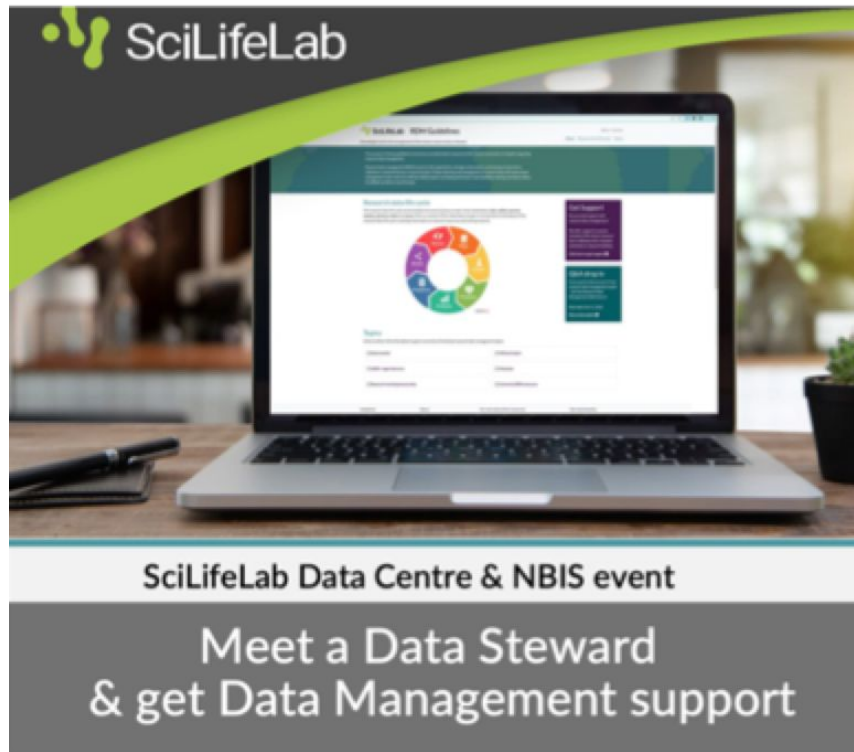
- ❑ **Support**
  - Consultations
  - Hands-on support
  - Drop-ins
- ❑ **Guidelines**
  - SciLifeLab RDM Guidelines
  - RDMkit
- ❑ **Services**
- ❑ **Tools**
- ❑ **Training**



*SciLifeLab*

## Contact us

- [data-guidelines.scilifelab.se](https://data-guidelines.scilifelab.se)
- [data-management@scilifelab.se](mailto:data-management@scilifelab.se)



[scilifelab.se/event/meet-a-data-steward](https://scilifelab.se/event/meet-a-data-steward)

Increase your data management skills and get your IT or RDM questions answered.

Each event takes place in a hybrid format and features a 15 minutes mini-lecture followed by 45 minutes Q&A

## Fall 2023

**19 Sept** - Tools for reproducible research

**24 Oct** - Sharing sensitive personal data from life science research

**21 Nov** - FAIR Data stewardship

**21 Dec** - New services on SciLifeLab Data Platform

The purpose of these guidelines is to serve as an information resource to life science researchers in Sweden regarding research data management.

Research data management (RDM) concerns the organisation, storage, preservation, and sharing of data that is collected or analysed during a research project. Proper planning and management of research data will make project management easier and more efficient while projects are being performed. It also facilitates sharing and allows others to validate as well as reuse the data.

## Research data life cycle

The research data life cycle can be divided into several phases as seen in the wheel below; **plan, collect, process, analyse, preserve, share and reuse**. Click on a section of the wheel below to get an introduction to that phase of the research data life cycle, including information on relevant resources and training material.



RDM life cycle from RDMkit licensed under Creative Commons Attribution 4.0 International License.

## Topics

Click on either of the links below to get an overview of individual research data management topics.

Working with human data

Data transfer

FAIR principles

Metadata

### Get Support

Do you need support with research data management?

We offer support to anyone involved in life science research that is affiliated with a Swedish university or research institute.

[Click here to get support](#)

### Meet a Data Steward

Join SciLifeLab Data Centre & NBIS and get data management support. Each event consists of a 15 minutes mini-lecture and a 45 minutes Q&A.

Next date: Sep 19, 2023

[More information on the event page](#)

### Events & Training

Upcoming conferences, webinars, workshops, and training opportunities in Sweden related to data-driven life science can be found on the SciLifeLab Data Platform.

[Go to Events & Training on the Data Platform](#)

### Calls & Jobs

Available positions and funding opportunities within data-driven life science can be found on the SciLifeLab Data Platform.

## [data-guidelines.scilifelab.se](https://data-guidelines.scilifelab.se)

## Contact

[Home](#) / [Contact](#)

The SciLifeLab RDM guidelines are maintained by the SciLifeLab Data Centre and NBIS. We welcome suggestions, contributions, and questions related to the content of this site. This could include, for example, a suggestion for a [topic](#) that is not currently covered. Alternatively, you can request an update or correction to particular pages.

You are also more than welcome to contact us if you want support regarding data management in your project, e.g. regarding writing data management plans, publishing data, working with sensitive human data, or organising and documenting data.

You can fill in the below form to get in touch with us. If you prefer though, you can instead email us at [data-management@scilifelab.se](mailto:data-management@scilifelab.se).

### Contact form

Your name:

Your email:

Type of query:

Request support



Message:

[data-management@scilifelab.se](mailto:data-management@scilifelab.se)

## Research data life cycle

[Home](#) / [Research data life cycle](#) / [Planning phase](#)

On this page:

- [Resources](#)

### Planning

During this phase you plan for which data is needed to answer your research question, and how that data is going to be managed during as well as after the project.

It is wise to write a data management plan (DMP), using either a tool provided by your university or [DS wizard](#) provided by SciLifeLab. The main benefits of writing a DMP are to document how the data is managed and identify possible gaps in the current data management practices. Think of the DMP as a checklist to go through when planning for a new project.

High quality science is often only possible if the infrastructure resources you intend to use get involved already in the planning phase of a project. Consultation and advice regarding data management planning, data generation and data analysis are offered by [NBIS](#) and [SciLifeLab units](#). Many institutes / universities have established [research data offices](#) which provide guidance and support regarding data management questions. There might also be some [policies](#) regarding research data, from funders or universities, that you need to be aware of and adhere to.

Also, some resources have specific application periods and thus needs to be contacted well in advance. If your project includes sensitive (human) data, note that there are ethical and legal issues that you have to consider, such as apply for an ethics approval, establish necessary agreements and report the data processing to your [Data Protection Officer](#). See the page on [Research involving human data](#) for more information.

### Resources

- [RDMkit on Data Management Planning](#)
- [DMP module in course Introduction to data management practices](#)
- [DS Wizard - A short introduction \(video\)](#)
- [How to write a Data management Plan \(video\)](#)





## - The Research Data Management toolkit for Life Sciences

[Home](#)[About](#)[Contribute](#)[Contact](#)[GitHub](#)[Data life cycle](#)[Your role](#)[Your domain](#)[Your tasks](#)[Tool assembly](#)[All tools and resources](#)[All training resources](#)

### Are you working with data in the Life Sciences? Do you feel overwhelmed when you think about Research Data Management?

The ELIXIR Research Data Management Kit (RDMkit) is an online guide containing good data management practices applicable to research projects from the beginning to the end. Developed and managed by people who work every day with life science data, the RDMkit has guidelines, information, and pointers to help you with problems throughout the data's life cycle. RDMkit supports FAIR data — Findable, Accessible, Interoperable and Reusable — by-design, from the first steps of data management planning to the final steps of depositing data in public archives.

The RDMkit organises information into the six sections displayed below, which are interconnected but can be browsed independently.

### Data life cycle

Start here to get an overview of research data management. Click on a section of the diagram below to get an introduction to that stage of the data management life cycle.



### Your role

Identify your role in research data management, find data management resources relevant for you, and information to help you manage them.

### Your domain

Learn about the data management problems that affect your domain or research community, and the solutions adopted to address them.

<https://rdmkit.elixir-europe.org/>

“The ELIXIR Research Data Management Kit (RDMkit) has been designed to guide life scientists in their efforts to better manage their research data following the FAIR Principles. It is based on the various steps of the data lifecycle, although not all the steps will be relevant to everyone.”

“RDMkit is recommended in the [Horizon Europe Program Guide](#) as the "resource for Data Management guidelines and good practices for the Life Sciences."



# - The Research Data Management toolkit for Life Sciences

[Data management](#)[About](#)[Contribute](#)[GitHub](#)

The Research Data Management toolkit for Life Sciences

Best practices and guidelines to help you make your data FAIR (Findable, Accessible, Interoperable and Reusable)

<https://rdmkit.elixir-europe.org/>

[Data management](#)[About](#)[Contribute](#)[GitHub](#)

## What can we help you find?

## Browse all topics by



### Data life cycle

Start here to get an overview of research data management based on stages in the data life cycle.



### Your role

Identify your role in research data management, find data management resources relevant for you, and information to help you progress in your career path.



### Your domain

Learn about data management tasks that affect domain or research community, and the solution to address them.



### Your tasks

Find guidelines and solutions for tackling common data management tasks.



### Tool assembly

Find concrete combinations of tools and resources assembled into an ecosystem for research data management.



### National resources

Find pointers to country-specific information or national research data management practices.



### All tools and resources

Browse the RDMkit's catalogue of tools and resources for research data management.



### All training resources

Browse all training resources mentioned in RDMkit pages.

## Data management

[Data life cycle](#)[Your role](#)[Your domain](#)[Your tasks](#)[Compliance monitoring](#)[Costs of data management](#)[Data analysis](#)[Data brokering](#)[Data management coordination](#)[Data management plan](#)[Data organisation](#)[Data security](#)[Data sensitivity](#)[Data provenance](#)[Data publication](#)[Data quality](#)[Data storage](#)[Data transfer](#)[Documentation and metadata](#)[Existing data](#)[GDPR compliance](#)[Identifiers](#)

Your tasks

## Data quality

### How do you ensure the quality of research data?

#### Description

Data quality is a term that can be understood in many ways. In enterprise context, it often refers to master data management as defined by the ISO 8000 standards. In science, the quality of data is closely linked to the suitability of the data for (re)use for a particular purpose and it is a key attribute of research data. Data quality affects the reliability of research results and it is a key factor increasing the reusability of data for secondary research. Data quality control can take place at any stage during the research data lifecycle. That said, you should ensure that the necessary procedures are defined during data management planning.

#### Considerations

Quality control is most typically performed during data collection but it should not be neglected in later stages of research data lifecycle. The type of data as well as instruments and processes adopted for data collection in your research will determine the quality assurance measures you can take. Examples of such measures are:

- setup data management working group (DMWG) that includes people who generate data, analyse data and data managers;
- for data collection: DMWG to plan and define data dictionary (including validation rules) before collecting data;
- for metadata collection: DMWG to plan and define metadata data templates;
- use electronic data capture systems;
- automated quality monitoring through tools, pipelines, dashboards;
- training of study participants and researchers, surveyors or other staff involved;
- adopting standards;
- instrument calibrations;
- repeated samples;
- post collection data curation;
- data peer-review.

Certain areas such as clinical studies, or those involving Next Generation Sequencing have commonly working methods to ensure data quality. Consider familiarizing yourself with data quality standards or established working practices in your field of study.

There are many frameworks proposed in the literature to define and evaluate overall data quality, such as:

On this page

[How do you ensure the quality of research data?](#)[More information](#)[Relevant tools and resources](#)



## Scientific Data Platform

The SciLifeLab Data Platform is a new technical hosting environment for data-centric tools, databases, and support for data-driven life science. The website will function as a hub for the life science community. The platform, funded by the SciLifeLab & Wallenberg National Program for Data-Driven Life Science (DDL), will serve all life science researchers and data-producing facilities in Sweden.

Learn More →

## IT and Software Services

The SciLifeLab Data Centre provides services for IT and data management. We also work to facilitate the communication between SciLifeLab platforms, their users, and the research community, and develop databases and tools. The purpose is to maximize the scientific impact of SciLifeLab generated data and provide a good helpful IT environment for the SciLifeLab platforms and their users.

See All Services →

<https://www.scilifelab.se/data>

## Share data

SciLifeLab Data Centre promotes Open Science and data sharing according to the FAIR principles

Learn More →

## AI model and compute application sharing


SciLifeLab Serve is a service for publishing apps and serving machine learning models. This service is currently under development.

Learn More →

## Pandemic data

SciLifeLab Data Centre maintains the Swedish Pathogens Portal. The aim of the portal is to provide information, guidelines, tools and services to support researchers in utilising Swedish and European infrastructures for data sharing.

Learn More →


**SciLifeLab** Data Platform

About Contact


Services Resources Data Highlights Jobs Funding Calls Events & Training

Hub for data-driven life science research in Sweden

## New Resources Section Launched


### Services

See all services



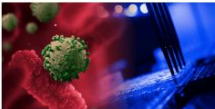
**SciLifeLab Data Repository**

Repository for publishing any kind of research-related data, posters, presentations, and more!




**Hosting on the SciLifeLab Data Platform**

Details about how researchers, research groups, and organisations can apply for hosting via the Platform.



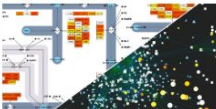
**The Swedish COVID-19 & Pandemic Preparedness Data Portal**

A web portal for resources related to pandemic preparedness research in Sweden




**SciLifeLab RDM Guidelines**

A portal containing information on research data management in Sweden



**Metabolic Atlas**

Metabolic Atlas is a web platform integrating open-source genome scale metabolic models (GEMs) for easy browsing and analysis.



**SciLifeLab Serve**

Tool allowing to host and share R Shiny and Dash apps, as well as trained machine learning models

### What's new?

See all news

August 8, 2023  
Launch of resources section

July 4, 2023  
Glad sommar!

June 1, 2023  
New highlight shows how changes in METTL3 localisation could hold the key to new antiviral strategies against SARS-CoV-2.

June 1, 2023  
New services and SciLifeLab Data Platform Launch materials now available

### About this platform

Read more

The *SciLifeLab Data Platform* is a technical environment offering data-centric tools and databases as well as an overarching website with the goal to support and accelerate data-driven life science research in Sweden. The services hosted on the Platform are available to all life science researchers in Sweden. The platform is maintained by the SciLifeLab Data Centre.

The services hosted on the SciLifeLab Data

## 2 parts:

- A website ([data.scilifelab.se](https://data.scilifelab.se))
- An underlying technical environment for hosting

## Aims to provide:

- Access to & hosting for appropriate services
- Promotion for data-driven life science research
- Resources related to community building and support
- Guidance for researchers

## Scientific Data Platform

The SciLifeLab Data Platform is a new technical hosting environment for data-centric tools, databases, and support for data-driven life science. The website will function as a hub for the life science community. The platform, funded by the SciLifeLab & Wallenberg National Program for Data-Driven Life Science (DDLs), will serve all life science researchers and data-producing facilities in Sweden.

Learn More →

## IT and Software Services

The SciLifeLab Data Centre provides services for IT and data management. We also work to facilitate the communication between SciLifeLab platforms, their users, and the research community, and develop databases and tools. The purpose is to maximize the scientific impact of SciLifeLab generated data and provide a good helpful IT environment for the SciLifeLab platforms and their users.

See All Services →

<https://www.scilifelab.se/data>

## Share data

SciLifeLab Data Centre promotes Open Science and data sharing according to the FAIR principles

Learn More →

## AI model and compute application sharing

SciLifeLab Serve is a service for publishing apps and serving machine learning models. This service is currently under development.

Learn More →

## Pandemic data

SciLifeLab Data Centre maintains the Swedish Pathogens Portal. The aim of the portal is to provide information, guidelines, tools and services to support researchers in utilising Swedish and European infrastructures for data sharing.

Learn More →

## IT and Software Services

We have a number of services available, and some things are in development.

### Systems for operations

#### ORDER PORTAL

A web based system for unit communication with prospective users. We can provide you with an instance configured according to your own specific requirements. Try out a demo instance at [orderportal.scilifelab.se](https://orderportal.scilifelab.se), contact us or Per Kraulis directly.

#### PUBLICATIONS DATABASE

The Data Centre built and maintains the [SciLifeLab Publications database](#) where the results of research assisted by SciLifeLab units can be found. A [second database](#) was set up for publications of researchers and fellows affiliated with SciLifeLab.

#### DATA MANAGEMENT PLANS

We encourage SciLifeLab units to support users with data management plans. We provide templates and a software system to manage such plans, the [SciLifeLab instance of the Data Stewardship Wizard](#).

#### TOOL SUPPORTING REPRODUCIBLE METHODS

SciLifeLab Data Centre recently purchased a Premium license for [protocols.io](https://protocols.io). Current accounts using your SciLifeLab email addresses have automatically been upgraded to a Premium account. For other accounts, to upgrade and take advantage of the SciLifeLab license, please email [datacentre@scilifelab.se](mailto:datacentre@scilifelab.se).

#### SURVEY TOOLS

The online survey tool [Typeform](#) is offered to the units and administration. It allows to collect questionnaire responses from users/researchers. Email us for access.

#### DATA DELIVERY SYSTEM

The [Data Delivery System \(bds\)](#) is a cloud-based system for the delivery of data from SciLifeLab platforms to their users.

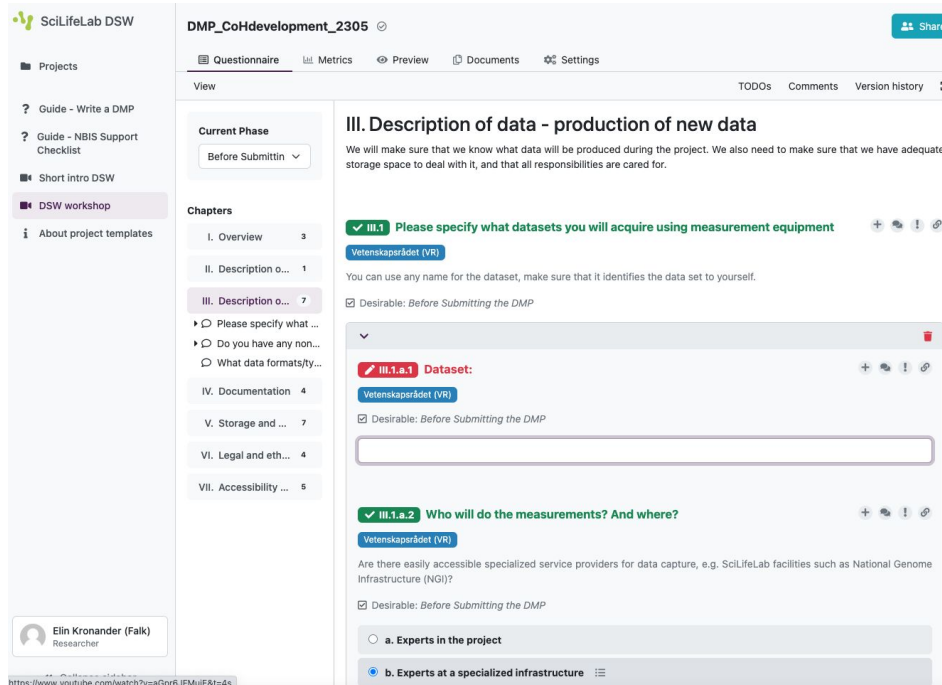
<https://www.scilifelab.se/data>

Systems for operations

Collaboration and communication systems

## DSW - Data Stewardship Wizard

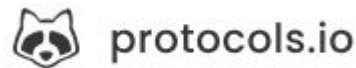
<https://dsw.scilifelab.se/>




- ❑ Life Science specific guidance
- ❑ ‘Bite-sized’ questions
- ❑ Most questions with answer options



An online platform for the creation, management, and sharing of research protocols or methods



## Create

Every new protocol starts out private. You can then choose to keep it private, share it with others, or to make it public.



## Run

Run protocols as checklists. Once your experiment is complete, the changes you made are recorded in your File Manager. You can keep your experiments private or share them with others.



## Publish

Receive a DOI after you publish a protocol. After publication, your protocol is archived with CLOCKSS to ensure longterm preservation of the knowledge.



# Sharing and reusing data

## Scientific Data Platform

The SciLifeLab Data Platform is a new technical hosting environment for data-centric tools, databases, and support for data-driven life science. The website will function as a hub for the life science community. The platform, funded by the SciLifeLab & Wallenberg National Program for Data-Driven Life Science (DDLS), will serve all life science researchers and data-producing facilities in Sweden.

[Learn More](#) →

## IT and Software Services

The SciLifeLab Data Centre provides services for IT and data management. We also work to facilitate the communication between SciLifeLab platforms, their users, and the research community, and develop databases and tools. The purpose is to maximize the scientific impact of SciLifeLab generated data and provide a good helpful IT environment for the SciLifeLab platforms and their users.

[See All Services](#) →

## Share data

SciLifeLab Data Centre promotes Open Science and data sharing according to the FAIR principles

[Learn More](#) →

## AI model and compute application sharing

SciLifeLab Serve is a service for publishing apps and serving machine learning models. This service is currently under development.

[Learn More](#) →

## Pandemic data

SciLifeLab Data Centre maintains the Swedish Pathogens Portal. The aim of the portal is to provide information, guidelines, tools and services to support researchers in utilising Swedish and European infrastructures for data sharing.

[Learn More](#) →

A national repository for storing and sharing personally identifiable information from Swedish biomedical research projects



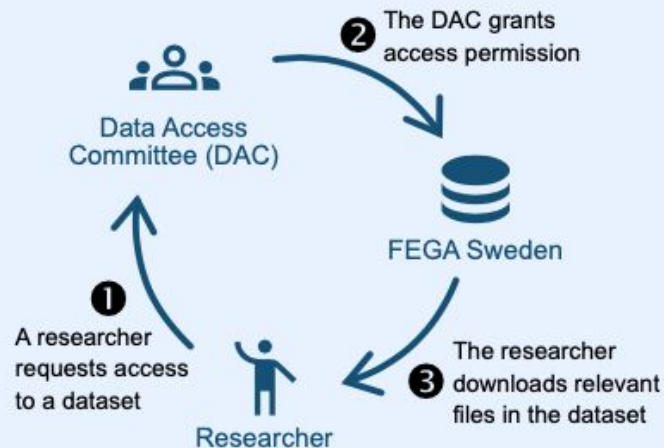
## Finding data

Anonymous information about deposited datasets is searchable through the international web portal [ega-archive.org](http://ega-archive.org).

## Submitting data

- Various types of biomedical omics data are welcome.
- Sensitive data is stored with strong encryption in highly secure data centers in Sweden.
- Datasets are shared under controlled access in accordance with the research project's ethical permit.

## Accessing data



# SciLifeLab Data Repository



- Institutional instance of [FigShare](#)
- Publish general research data without costs for Swedish researchers
- Available for researchers affiliated to all Swedish universities and institutes working within SciLifeLabs' research areas.
- Possible to apply for a license or embargo and set up a metadata record-only.

## Digital Research Hub

FAIRDOM



A (meta)**data**  
**management** platform  
*under development*

