

NBIS Data Management Strategy

DOI	10.17044/scilifelab.27604476
Version	2024-11-13
Last updated by	Niclas Jareborg & DM unit
Approval date	2024-11-13
Time for next revision	2025-10

Purpose	1
Scope of NBIS Data Management	1
Governing principles and documents	1
NBIS Data Management Vision & Mission	2
Staff and internal organisation	3
Funding	3
1. Support strategy	3
2. Training & Outreach strategy	6
3. Collaboration strategy	7
Definition of terms	9

Purpose

The purpose of this document is to outline the scope and guiding strategic decisions underlying the operations of the NBIS Data Management unit.

Some of the terms used in this document that might need clarifications are defined in the section Definition of terms.

Scope of NBIS Data Management

As part of the national infrastructure for bioinformatics, the NBIS Data Management unit sees it as its remit to provide data management support to the wider life science community in Sweden. The NBIS Data Management Vision and Mission are based on this scope.

Governing principles and documents

All activities of the NBIS Data Management unit are based on the principles of **Open Science**, **Reproducible Research**, and **FAIR**.

This strategy is then also guided by, and adheres to, the following governing documents:

- [NBIS Strategic Plan](#), including NBIS Vision & Mission
- [SciLifeLab Strategy](#)
- [SciLifeLab Data Policy](#)
- [Swedish Research Council's recommendations on open access to research data](#)
- [Swedish National guidelines for Open Science](#)

- [ELIXIR position paper on FAIR data management in the life sciences](#)
- [ELIXIR Scientific Programme 2024-2028](#)

NBIS Data Management Vision & Mission

Given the scope, guiding principles, and governing documents defined above, the following Vision & Mission statements have been formulated:

Vision

Swedish Life Science researchers apply good data management practices so that the research outputs produced are available to the global research community and to society at large, according to the principles of Open Science, Reproducible Research, and FAIR.

Mission

To work toward the stated vision, the following three areas have been identified:

1. Support

Swedish Life Science researchers shall have access to data management support to be able to manage their research data according to international best practices. To enable this, NBIS Data Management shall:

- Develop **procedures** for providing efficient data management support to research projects throughout the research data lifecycle
- Provide data management **planning** knowledge and resources/tools for research projects
- Provide support and resources/tools for **data submission** to recommended community deposition databases
- Provide support and resources/tools to enable **reproducible research**

2. Training

Swedish Life Science researchers shall have access to training in efficient data management practices that can be applied in their daily work. To enable this, NBIS Data Management shall:

- Provide data management **training** for researchers
- Provide **guidelines and documentation** for data management “best practices”
- Promote awareness of the importance of Open Science, Reproducible Research, FAIR and good data management practices in the Swedish Life Science community

3. Collaboration

Data management practices, solutions and training for Swedish Life Science shall be established in collaboration with NBIS internally, as well as with relevant national and international stakeholders. To enable this, NBIS Data Management shall:

- Collaborate with NBIS staff to develop data management aspects of support procedures and policies
- Contribute to data management activities and capacity building within ELIXIR
- Collaborate with SciLifeLab Data Centre to develop data management policies, procedures and solutions to support researchers
- Interact with other national and international stakeholders to harmonise data management policies, procedures and solutions to support researchers

Staff and internal organisation

Staff

The NBIS Data Management unit currently consists of 11 Data Stewards (DS) and one Data Manager (Line manager). The majority of the DM staff are presently located at BMC, Uppsala University, but also in Stockholm and Gothenburg.

Team organisation

The staff is broadly organised into two functional teams - General data management support ("DM support") and Human data helpdesk. Staff can belong to one or both of these teams. The two teams staff two separate support functions, as well as the internal development needed to continuously improve the operations of those support functions. Most staff are also engaged in the team that performs the training and outreach activities of the unit ("Training & outreach team"). To enable interactions with the rest of NBIS support activities, most Data Stewards are also affiliated to one or more of the NBIS Support units.

Staff skills

To foster collaboration and ensure redundancy, the aim is that staff skill sets are a mix of specialisation in some areas and cover other topics more broadly. Individual staff members are to be able to cover several different support topics, but are not expected to cover all areas. Learning and developing data management skills in this changing area is in most cases done by engaging in tasks that are to be solved, i.e. it is expected that solving tasks where skill development is needed will take longer than when fully proficient. Skill development is also a prioritised reason for engaging in external collaborations.

Ways of working

To work towards the unit vision, the unit applies agile ways of working to identify and prioritise needs and requirements, with regular prioritisation of goals and evaluation of results. On the level of the whole unit, this is done by defining Objectives and Key Results (OKRs) to be achieved during four-month periods. The functional crews define more fine-grained objectives that align to the OKRs.

Funding

Funding for the NBIS Data Management operations and activities are funded through the NBIS grant from the Swedish Research Council, university co-funding as well as from SciLifeLab funding to NBIS. The Human data helpdesk functions and development are also funded through EU project grants and additional sources.

Currently, no user fees are obtained to fund the data management support services to research groups. Motivation for this is to lower the threshold for researchers to apply for data management support, to facilitate compliance with national and EU mandates on Open Science and the FAIR principles. Exceptions for this are extensive support cases, where user fees might be charged.

1. Support strategy

Support to Swedish life science researchers and projects shall be in the form of **services** that provide expert knowledge by Data Stewards, in the form of **consultations** as well as **hands-on support**.

NBIS Data Management will **not** operate technical services (such as hosting IT systems or managing servers) to provide support. Any technical services deemed necessary shall be provided by either other parts of NBIS, or by other external collaborators.

RDM services

The following data management support services have been prioritised to fulfil the support mission, and are divided into two separate support functions - *General data management support*, and *Human data helpdesk*.

General data management support (“DM support”)

Request handling

Support request channels are offered to requestors at the [SciLifeLab RDM Guidelines portal](#), as well as on the NBIS web page ([support request forms](#)). SciLifeLab RDM Guidelines offer a joint entryway for getting data management support without having to know whether NBIS Data Management or SciLifeLab Data Centre should provide the service. This joint support request channel is staffed by both NBIS Data Management staff and SciLifeLab Data Centre staff. Support requests that are to be handled by NBIS Data Management, regardless of communication channel, are all directed to the NBIS support issue handling system (“Redmine”). Data management support requests that originate in NBIS support projects are also handled as separate issues in Redmine. The request handling process is defined as written operating procedures.

Services

Data Management Planning consultations

Provide expert knowledge guidance to researchers and infrastructure staff on establishing data management plans (DMPs). The guidance focuses on the use of the [SciLifeLab Data Stewardship Wizard \(DSW\)](#) tool, including the life science-focused *knowledge models (KMs)* developed by the unit, but other tools in use by Swedish universities are also in scope.

Data deposition consultations and support

Provide expert knowledge guidance to researchers and infrastructure staff on where and how to deposit research data in international public repositories. The guidance focuses on general advice on good practices for sharing data but also includes specific guidance for the [ELIXIR Deposition Databases](#), such as the ENA, ArrayExpress, PRIDE, etc.

Provide *hands-on support* with depositing research data to international public repositories on behalf of the requester. The service primarily caters for deposition of data to the [ELIXIR Deposition Databases](#). Currently, data submission to the following repositories are well supported:

- European Nucleotide Archive
- ArrayExpress
- PRIDE
- BioStudies

Other ELIXIR Deposition Databases are also supported, although lead times might be longer. *Note that support for submission of human data is provided by the Human data helpdesk crew, below.*

Consultations and hands-on support for other international repositories are considered to be out of scope, although exceptions can be made depending on the complexity of the task, especially if no suitable repository is available for the particular data type on the ELIXIR Deposition Database list.

General human data consultations

Provide expert knowledge guidance to researchers and infrastructure staff on how to work with human data, including practical, ethical, legal and information security aspects. *Deposition of human genomic and phenotypic data in repository instances operated by NBIS is handled by the Human data helpdesk function.*

Other general Data Management consultations

Provide expert advice on data management aspects not covered by the services described above, such as: documenting data with metadata, organising files and folders, advice on reproducibility, long-term storage, etc. Depending on the complexity of the request and how well it aligns with the NBIS Data Management strategy, it might be possible to provide hands-on support for some requests.

Human data helpdesk

The Human data helpdesk services operate to support the deposition and access of human data to public repository instances with controlled access, *that are technically operated by NBIS.*

Request handling

Requests to deposit data are managed through a [request form on feqa.nbis.se](https://fega.nbis.se), that is then managed in the NBIS support issue handling system Redmine. The request handling process is defined as written operating procedures.

Services

Federated European Genome-phenome Archive Sweden (FEGA Sweden)

Provide expert knowledge guidance and hands-on support, to researchers on how to describe and deposit data to the Swedish node of the Federated European Genome-phenome Archive. FEQA Sweden currently supports deposition of genomic and phenotypic data.

Operational development

Operation of the service functions require development and continuous improvement. These activities are broadly staffed by the same crews responsible for the operations of the service functions. Development activities are done together with internal and external collaborators whenever relevant (see 3. *Collaboration strategy* below).

General data management support (“DM support”)

DSW Knowledge Models (KMs)

Develop and continuously improve the KMs underlying the templates for DMPs, as well as export templates, in DSW. The KMs have a focus on providing guidance and being relevant for life science research in the Swedish (academic) context.

Data deposition

Develop procedures and tools to facilitate and quality assure submission of research data to international public repositories, with a focus on the ELIXIR Deposition Databases. Order of priority is those repositories suitable for the largest amounts of requests, i.e. ENA for sequencing data,

ArrayExpress for expression data, etc. Over time, support capacity for other repositories will be adjusted if demand increases.

Operational procedures

Establish and continuously improve operating procedures to enable effective and efficient support services.

Human data helpdesk

Operational procedures

Establish and continuously improve operating procedures to enable effective and efficient support to researchers to deposit human data to, and access from, public repository instances with controlled access, *that are (or will be) technically operated by NBIS*. The Federated European Genome-phenome Archive Sweden (FEGA Sweden) node is operational, and the 1+ Million Genome (1+MG) node is being established. This work is carried out in connection with several international projects in which NBIS is engaged, such as the Genomic Data Infrastructure (GDI) and Genome of Europe (GoE) projects, among others. Also, engage relevant national stakeholders, higher education institutions and health care regions, that are controllers of data that will be deposited in the repositories operated by NBIS, to ensure procedures are in place to grant access to data stored in these repositories.

2. Training & Outreach strategy

The strategy for data management training and outreach is captured in a separate document that outlines the strategy and plans in more detail. This section is a brief summary of that strategy document.

- [NBIS DM Unit Training Strategy](#) (DOI: 10.17044/scilifelab.27606417)

Training

Courses

Provide training material and trainer-led courses, that targets a broad audience in Sweden and aims to interact with researchers early in their careers to increase the impact. Focus is to primarily target beginner and intermediate level skills. Courses build on modular training material that is also made publicly available for other actors to adapt and deliver, as well as for self-study. A three day trainer-led beginner course is given twice a year as part of the NBIS course catalogue.

Guidelines

Develop and provide written guidelines and best practices through the [SciLifeLab RDM Guidelines portal](#) and the [ELIXIR RDMkit](#). The SciLifeLab RDM Guidelines portal is operated together with the SciLifeLab Data Centre and focuses on data management practices relevant for life science research in the Swedish context (see 3. *Collaboration strategy* below). It also functions as a joint entry point to request data management support from SciLifeLab and NBIS (see 1. *Support strategy* above). The ELIXIR RDMkit is a joint international effort to capture and harmonise data management best practices that the DM unit contributes content to. Duplication of information in these two resources is kept at a minimum. The SciLifeLab RDM Guidelines refer to the less context-specific content in RDMkit whenever relevant.

Outreach

Promote awareness of the data management principles, services and resources available to the Swedish life science community and infrastructures through outreach events in different formats, such as in-person or online seminars and workshops. Apart from those events planned by the unit, outreach events are also available upon request from users. The outreach activities aim at reaching all Swedish life science researchers, regardless of their physical location. Also, outreach activities aim to engage and promote interaction with data management professionals working at other relevant stakeholder institutions or organisations (see 3. *Collaboration strategy* below).

3. Collaboration strategy

To carry out and develop the operations of the unit, it will engage with the following list of prioritised internal and external stakeholders to fulfil its mission.

NBIS internally

Bioinformatics support projects

Support NBIS *bioinformatics support projects* with data management competence when requested. All bioinformatics support projects are assigned a DS, that can be contacted if needs arise. NBIS project coordinators shall inform bioinformatics support projects users of the availability of these services. Data Stewards also support NBIS bioinformaticians through their affiliation to the different NBIS Support teams.

EU-funded projects

Together with other parts of NBIS, mainly the SysDev unit, engage in EU-funded projects relevant to the mission of the DM unit, that NBIS partake in and contribute to.

SciLifeLab Data Centre

Collaborate with Data Centre to establish data management services and training for researchers and infrastructure units.

SciLifeLab DSW

Together with Data Centre, develop and continuously improve the data management planning guidance and capability of the [SciLifeLab DSW](#). The DM unit is responsible for the content, and Data Centre is responsible for the technical operations.

Data management service development

Together with Data Centre, develop services to enable good data management practices for Swedish researchers and infrastructure units. High priority on services that facilitate data deposition to international repositories. The DM unit mainly provides data management knowledge and competence, and Data Centre provides technical development.

SciLifeLab RDM Guidelines

Together with Data Centre develop and continuously improve the [SciLifeLab RDM Guidelines portal](#) as a knowledge resource hub for the management of life science research data in Sweden, as well as a joint entryway for requesters to apply for data management support. The DM unit mainly focuses on the guideline and information content, and Data Centre mainly on the technical development and operations of the web resource.

Outreach activities

Together with Data Centre, arrange and perform outreach activities regarding data management, in the form of seminars, site visits and workshops that can be accessed by a broad range of Swedish life science researchers and infrastructure units, regardless of where they are located.

ELIXIR

Engage and contribute to relevant activities in ELIXIR to harmonise with, and leverage, data management knowledge, capabilities and standards developed internationally.

ELIXIR Communities

Research Data Management Community

Use the ELIXIR RDM Community as the main interaction point with the ELIXIR formal structures regarding data management. The NBIS Data manager co-leads the Community. The DM unit to engage with and contribute to activities in the Community, as a means to enable data management support and training according to international best practices, as well as for a means for professional development.

Federated Human Data Community

Aim to be represented at community meetings and contribute actively to building a robust, federated data ecosystem in Europe.

ELIXIR Platforms

Engage with and contribute to activities in the ELIXIR Platforms that are pertinent to the mission of the DM unit. Primarily, those of the *Data Platform* and *Interoperability Platform*.

ELIXIR Focus Groups

Engage with and contribute to activities in the ELIXIR Focus Groups that are pertinent to the mission of the DM unit. Primarily, those of the *Research Data Alliance (RDA) Activities* and the *EOSC Focus Group*.

EU-funded projects coordinated by ELIXIR

Engage in relevant EU-funded projects coordinated by ELIXIR where this aligns with the NBIS data management strategy. In particular, projects that aim to establish capability to share human sensitive data under controlled access, such as the Genomics Data Infrastructure (GDI), Genome of Europe (GoE), etc.

Other stakeholders

SciLifeLab infrastructure platforms

Engage with other SciLifeLab infrastructure platforms to enable FAIRification of data produced at the platforms, e.g. Genomics and Clinical Genomics, to facilitate the flow of data to international repositories such as ENA and FEGA Sweden.

Swedish National Data service (SND)

Engage in knowledge sharing forums coordinated by the SND to interact with other relevant Swedish stakeholders with the aim to harmonise and leverage data management solutions nationally. Also, collaborate to make data deposited to the FEGA Sweden node visible in researchdata.se.

Swedish Higher Education Institutions

Engage with data management support functions at Swedish higher education institutions to harmonise the support and guidance for researchers at the different institutions. Also, to set up processes for granting access to personal data stored in FEGA Sweden where the institution is the controller of the data.

Research Data Alliance (RDA)

Engage with and contribute to the work in RDA to influence and ensure alignment with international data management best practices. This engagement is done in both the context of the ELIXIR Research Data Alliance (RDA) Activities Focus Group (see above), and directly by NBIS as an organisation. An NBIS Data Steward serves as a liaison and current co-chair of the RDA Life Sciences Data Infrastructures Interest Group and as a member of the Technical Advisory Board (one of four governance bodies of the RDA).

European Open Science Cloud (EOSC)

Engage with and contribute to the work in EOSC to influence and ensure alignment with international data management best practices. This engagement is done in the context of the ELIXIR EOSC Focus Group (see above), as well as representatives of Uppsala University in EOSC task forces, expert groups and in national coordination activities. Together with other units at NBIS, the DM unit positions its services to contribute to the objectives outlined in the EOSC Strategic Research and Innovation Agenda (EOSC SRIA) and the build-up phase of the EOSC federation.

Definition of terms

Consultation

Written or oral guidance to a requester on how the requester can solve specific tasks or needs

Deposition database

See **International public repository**

International public repository

An online internationally accessible platform for the management, sharing, integration, archiving and dissemination of research 'products', e.g. the [European Nucleotide Archive](https://www.europeannucleotidearchive.org/)

Hands-on support ("Support")

Hands-on help by a DS to solve specific tasks or needs of a requester

Knowledge model (KM)

A template determining the structure of a questionnaire for e.g. a data management plan, see <https://dsw-appendix.scilifelab.se/knowledge-models/>

Objectives and Key Results (OKRs)

OKRs is a goal-setting framework used to define measurable goals and track their outcomes. An Objective is what is to be achieved. Key Results benchmark and monitor how to get to the

Objective. ([OKR description on Wikipedia](#))

Requester

A person or group of people that request assistance with solving task or need