

DMP development at SciLifeLab

2020-05-26 - DMPs in Sweden: sharing good practice

Johan Rung, SciLifeLab Data Centre

Niclas Jareborg, NBIS

- Based on the National DMP guidelines
- Fleshed out with more in-depth questions and guidance relevant for Life Science

Data Management Plan

This template is structured according to the [SUHF DMP recommendations \(in Swedish\)](#)

Overview
Data description, and collection or reuse of existing data
Documentation of data and data quality
Data storage and backup
Ethical and Legal aspects
Data publication and archiving
Costs

Overview

Title of the project

Aim of the project

Data storage and backup

Storage

How, where, and for how long will the data be stored **during** the analysis phase of the project?

*Suggestion: Data will be stored at [UPPMAX](#) or in own/in-house cluster during the analysis phase. Note that UPPMAX computing systems are not intended for **long-term** storage of research data.*

Backup

How, where and at what intervals will the data be backed-up?

*Suggestion: Data kept at UPPMAX systems are regularly backed up. Also consider having additional copies elsewhere, e.g. [Swestore](#). **Note that for human data, additional security measures, such as e.g. encryption of the data, should be taken.** Consider submitting the read data to [ENA](#) early in the project as a back-up copy of the "raw" data. (Note that human data should not be submitted to the ENA)*

How will data be recovered in the case of a data loss incident? How will you ensure that the backup copies are intact?

Suggestion: Define and test a defined recovery routine. Regularly check that backed up files are not corrupted.



About

Features

Resources

Get Started

Data Stewardship Wizard

Create Smart Data Management Plans
for FAIR Open Science

Get started

- NL & CZ ELIXIR nodes
- Design aims
 - FAIR DMPs
 - FAIRness evaluation
 - maDMPs

The screenshot shows the 'Data Stewardship Wizard' interface. The top navigation bar includes 'DS Wizard', 'Knowledge Models', 'Questionnaires', 'KM Editor', and 'Storage Costs Evaluator'. The main content area is titled 'My Experiment (Common DSW Knowledge Model, 1.4.0)' and features a 'Save' button. The 'Current Phase' is set to 'Before Submitting the Proposal'. The 'Chapters' list includes: I. Design of experiment (6), II. Data design and planning (5), III. Data Capture/Measurement (3), IV. Data processing and curation (1), V. Data integration (1), VI. Data interpretation (1), and VII. Information and insight (8). The active chapter is 'I. Design of experiment', which contains a question: '1. Is there any pre-existing data?' with a sub-question '1.b.1 Will you be using any pre-existing data (including other people's data)?'. The interface includes a 'Desirable: Before Submitting the Proposal' checkbox and a 'Data Stewardship for Open Science: atg' field.

<https://ds-wizard.org/>

SciLifeLab DSW

- Organization
- Users
- Knowledge Model Editor
- Knowledge Models
- Questionnaires
- How to use

Help >

Niclas Jareborg >

<< Collapse sidebar

Data Stewardship Wizard (DSW) is a tool to aid the creation of data management plans.

If you need help getting started, click on "How to use" in the side menu.

This instance of DSW is managed by SciLifeLab Data Centre. You can contact us at data-management@scilifelab.se if you need support.



Welcome to the SciLifeLab DSW!

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

SciLifeLab DSW

The Fab project (Common SciLifeLab DMP, 1.2.2) (unsaved changes) Discard Save

Organization

Users

Knowledge Model Editor

Knowledge Models

Questionnaires

How to use

Help

Niclas Jareborg

<< Collapse sidebar

Current Phase

Before Submitting the Proposal

Chapters

I. Overview ✓

Description of data - reuse of existing data and/or production of new data ✓

III. Documentation and data quality ✓

IV. Storage and backup ✓

V. Legal and ethical aspects ✓

VI. Accessibility and long-term storage ✓

VII. Responsibilities and resources ✓

VIII. Support from Data Centre ✓

More

Summary Report

1.a.4 Where will the data be stored during the research process?

Desirable: Before Submitting the DMP

a. SNIC center

b. Own/in-house network

c. Other

Clear answer

1.a.4.a.1 Which one?

a. Uppmax

b. PDC

c. NSC

d. LUNARC

e. HPC2N

f. C3SE

Clear answer

- Templates with guiding decision trees
- Tagging / Filtering of questions for e.g. different stakeholders
- Exporting to several different formats
 - “Unstructured” (pdf etc)
 - Export templates
 - “Structured” (json)

Bring Your Own DMP Drop-in

Share with others: [!\[\]\(ab4e2b3fc7e7887b7a72f548aa6f5e60_img.jpg\)](#) [!\[\]\(0a20d1259d5ab849a22cc9906b421113_img.jpg\)](#)

Do you need to create a data management plan?

SciLifeLab Data Centre and NBIS invite to hands-on drop-in consultation workshops for researchers wanting to discuss data management issues and data management plans for their research projects.

- 5 nov 13-16, Earth, Gamma 2, SciLifeLab Solna
- 11 nov 10-13, E10:2309, Navet, BMC, SciLifeLab Uppsala

Please register if you want coffee.

Date and time

2019-11-05 13:00 - 2019-11-05 16:00

Location

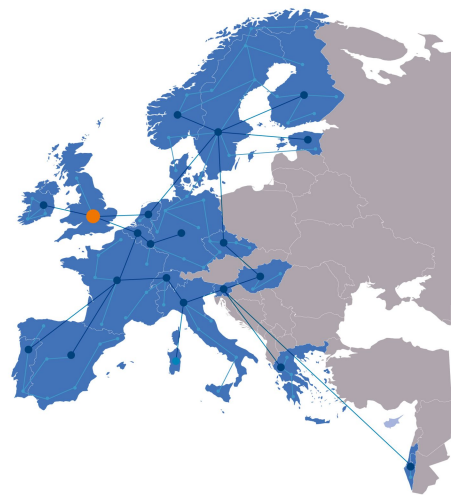
Nov 5 Earth, Gamma 2, SciLifeLab Solna & Nov 11 E10:2309, Navet, BMC, SciLifeLab Uppsala

ELIXIR-CONVERGE

(INFRADEV-03-2019-2020)

Aiming to connect and align ELIXIR Nodes to deliver sustainable FAIR life-science data management services

- All ELIXIR Nodes involved, individuals from **34** different organisations across Europe
- 3 years funding (2020-2023), €5 million
- *Establish a **Data Management Expert Network** for Life Science data*
- *Establish a **Data Management Toolbox***
- *Establish a **Data Management Training** portfolio*

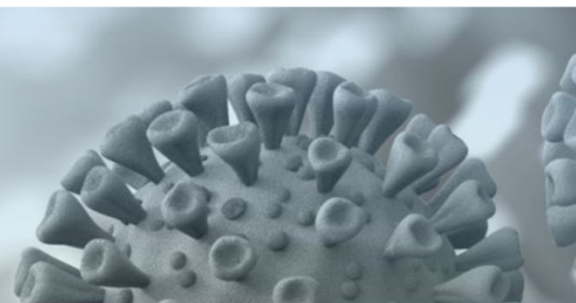




[About](#) [Data Hubs](#) [Other resources](#) [Our partners](#) [Submit data](#)

[Sequences](#) [Expression](#) [Proteins](#) [Structures](#) [Literature](#)

Accelerating research through data sharing



Sequences →

Raw and assembled sequences related to the COVID-19 outbreak, including outbreak isolates and records relating to coronavirus biology. Includes extensive sampling information.

Expression data →

Gene and protein expression data of human genes implicated in the virus infection of the host cells. Identifying cell types and genes with highest expression in SARS-CoV-2 infections.

About this portal

The COVID-19 Data Portal was launched in April 2020 to bring together relevant datasets for sharing and analysis in an effort to accelerate coronavirus research. It enables researchers to upload, access and analyse COVID-19 related reference data and specialist datasets as part of the wider European COVID-19 Data Platform.

SUPPORT & FEEDBACK

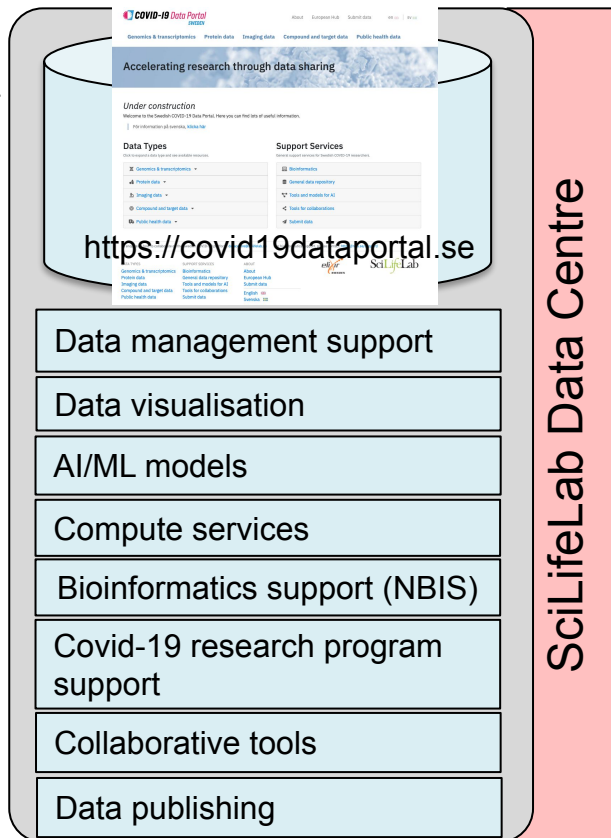
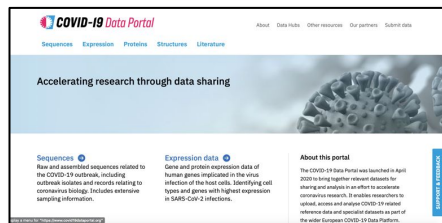
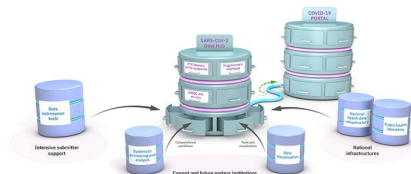
display a menu for "https://www.covid19dataportal.org"

Archive	Data type
European Nucleotide Archive (ENA)	Genetic (non-human) sequences
Expression Atlas	Processed gene expression data
PDB	Protein structures
EMDB	Electron microscopy density maps
ChEMBL	Compounds
Uniprot	Protein sequences
PRIDE	Raw and processed mass spectrometry data
IntAct	Molecular interactions
ComplexPortal	Protein complexes
BioSamples	Register samples

SciLifeLab Covid-19 Data Portal

Virus analyses
Antibody analyses
Epidemiological studies

External relations
Industry
Health care



SciLifeLab Covid-19 research program

Technology platforms

- Genomics
- Imaging
- Chemical Biology
- Proteomics
- Diagnostics Development
- Drug Discovery
- Bioinformatics

