



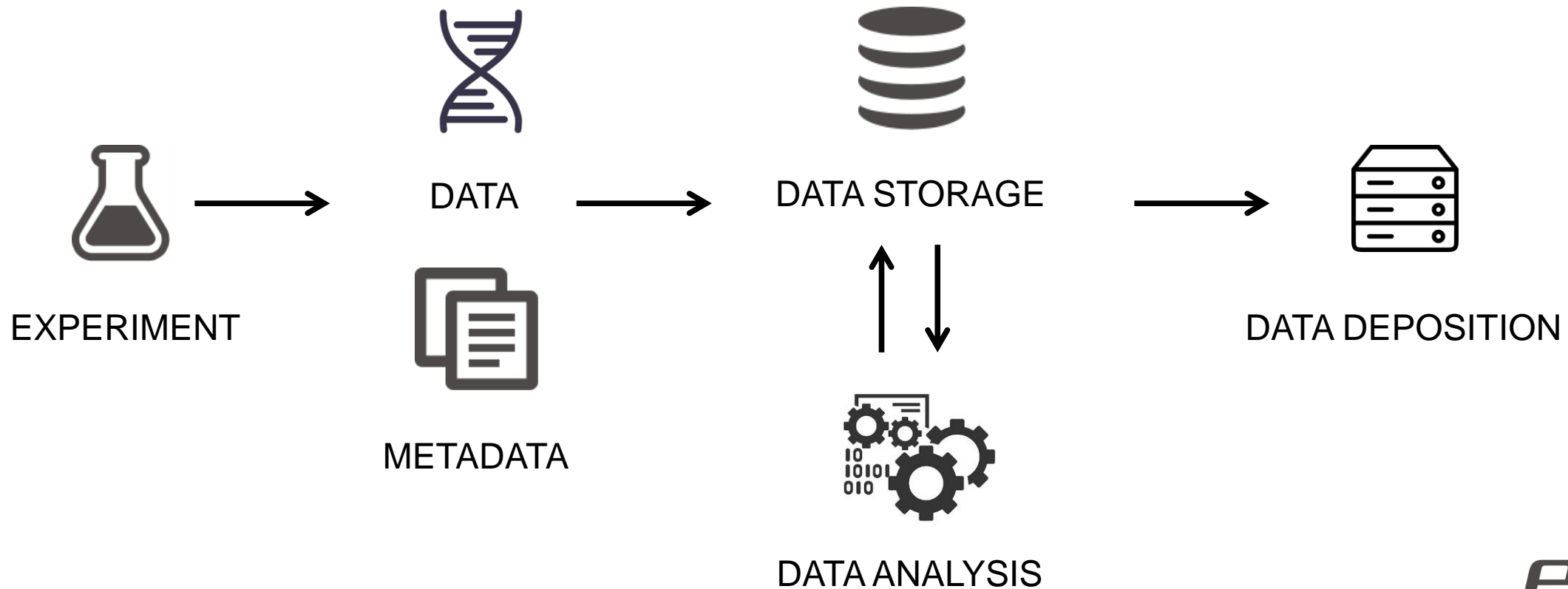
# National storage infrastructures

*NeLS/StoreBioInfo, NIRD, TSD*

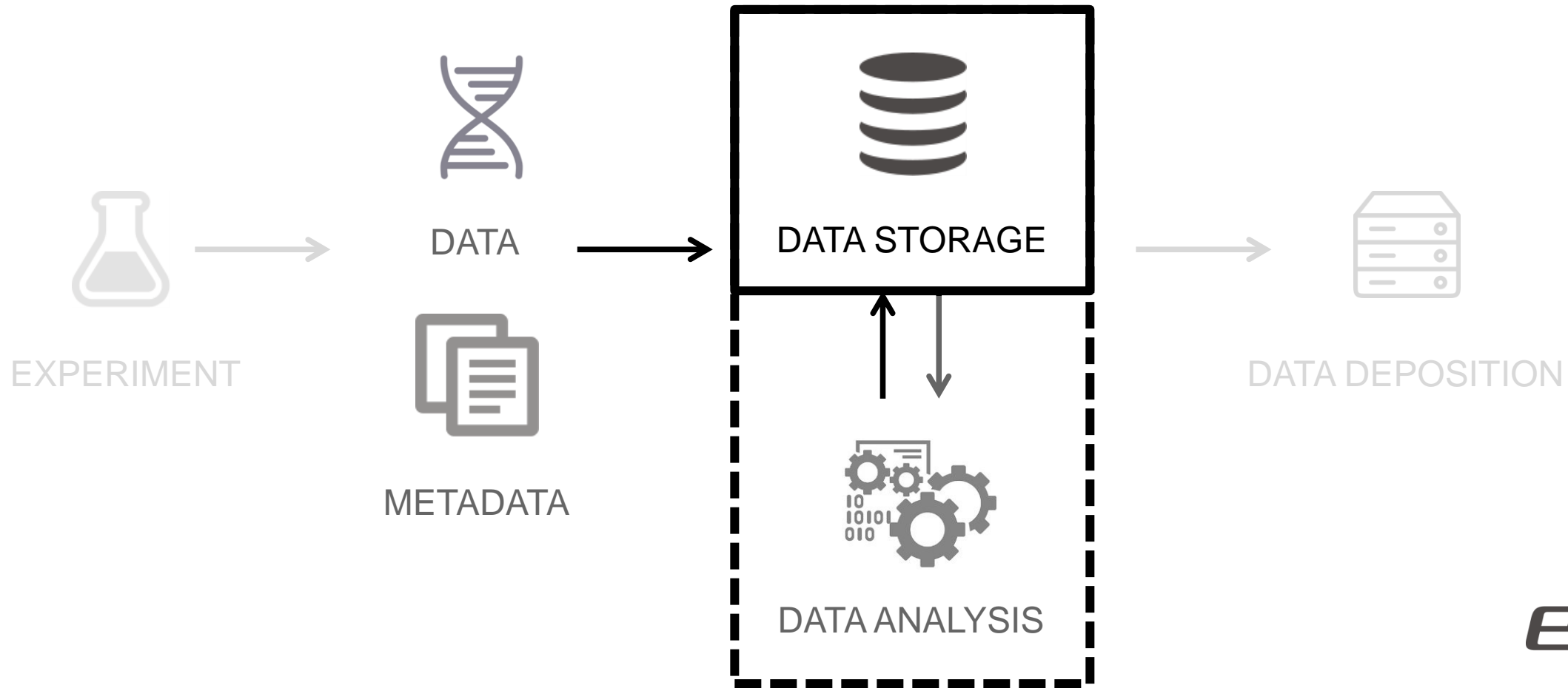


Erik Hjerde  
Arturo Vera-Ponce de Leon  
ELIXIR Norway

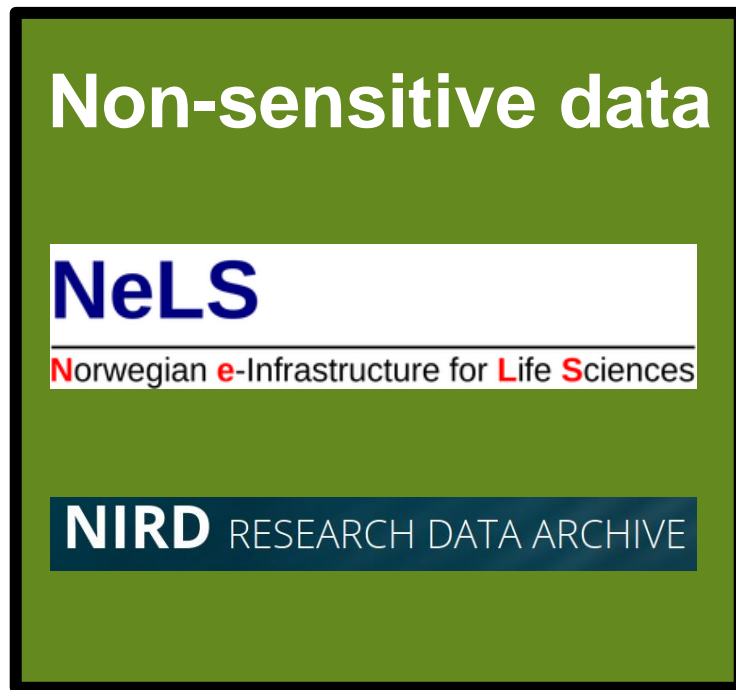
# *Data storage – from the researcher perspective*



# *Data storage – from the researcher perspective*



# *National storage infrastructures. Where to put my data???*



DATA ANALYSIS



# Data storage – General considerations

## The ELIXIR Research Data Management kit

Data life cycle	+
Your role	+
Your domain	+
Your problem	-

Compliance monitoring

Data analysis

Data management plan

Data organisation

Data protection

Data publication

Data quality

**Data storage**

Data transfer

Identifiers

Licensing

Documentation and metadata

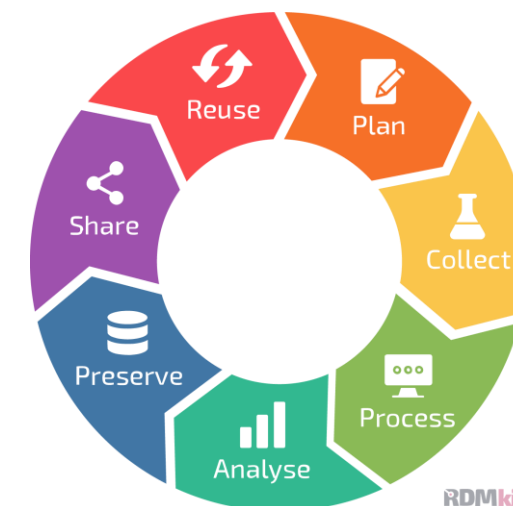
Sensitive data

All tools and resources

- What features do you need in a storage solution when collecting data?
- How do you estimate computational resources for data processing and analysis?
- Where should you store the data after the end of the project?
- Relevant tools and resources

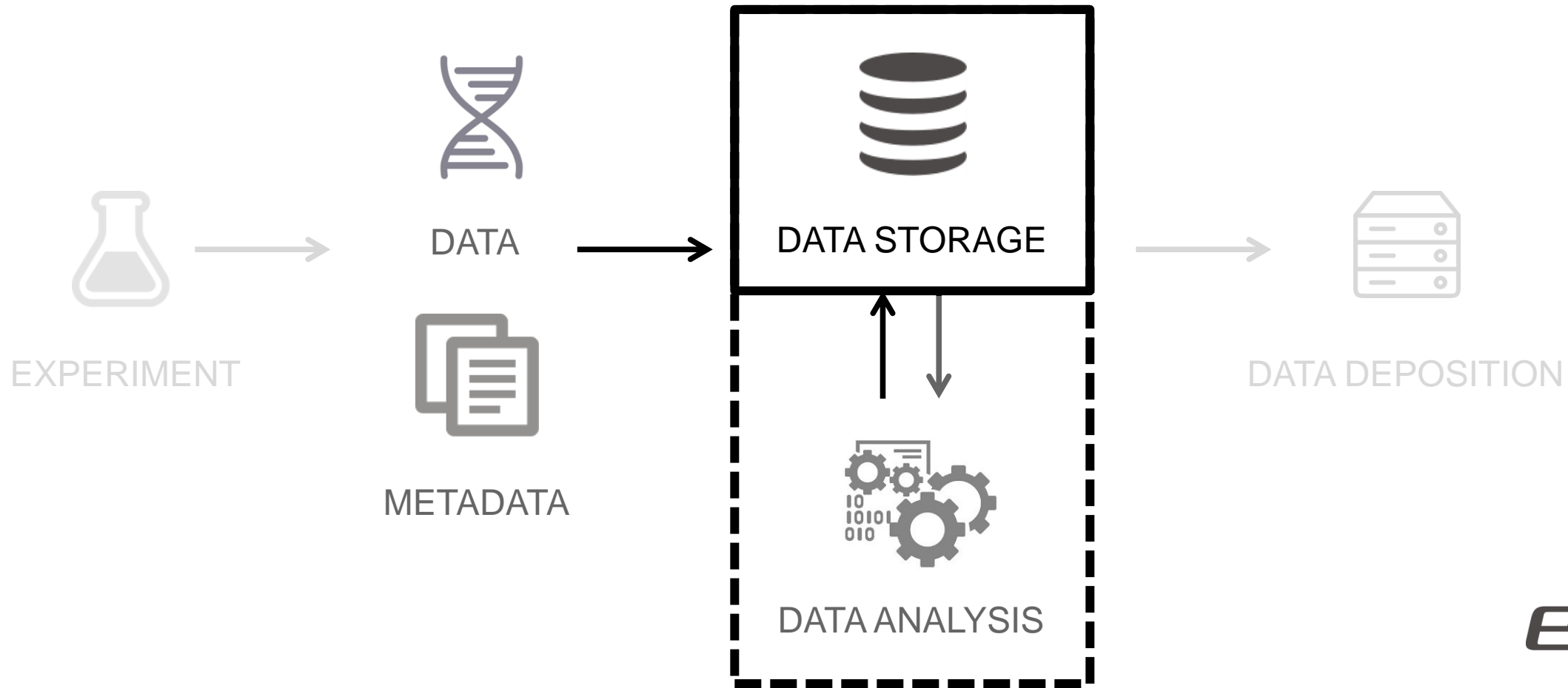
### Relevant tools and resources

Tool or resource ⓘ	Description	Tags	Registry
<a href="#">CERNBox</a>	CERNBox cloud data storage, sharing and synchronization	storage	
<a href="#">CS3</a>	Cloud Storage Services for Synchronization and Sharing (CS3)	storage	
<a href="#">DATAVERSE</a>	Open source research data repository software.	storage researcher data manager IT support	



Link to RDMkit storage: <https://rdmkit.elixir-europe.org/>

# *Data storage – from the researcher perspective*



# Elixir Norway



HOME ABOUT US HELPDESK SERVICES TRAINING NEWS EVENTS CONTACT US COVID-19 RESOURCES

Enter Search... 

## ELIXIR Norway

ELIXIR Norway is a national node of [ELIXIR](#), the pan-European infrastructure for biological information, supporting life science research and its translation to medicine, environment, the bioindustries and society.

ELIXIR Norway is funded by the Research Council of Norway, and builds upon the Norwegian Bioinformatics Platform, which has provided support and tools for Norwegian life science research for nearly two decades.



ELIXIR Norway is proud to be a part of ELIXIR-CONVERGE, to help standardise life science data management across Europe

### News

[Life Science RDM Group event: The Data Steward job role in the Norwegian public sector](#)  
Apr 24, 2024

[First datasets from Norway are now in the Federated EGA Network.](#)  
Mar 4, 2024

[Sweden adds a National Pathogens Portal to the network](#)  
Nov 24, 2023

[More news >>](#)



### Helpdesk

Short- and long-term support with bioinformatics analyses, programming and data management tasks. [Contact the helpdesk.](#)



### Services

Analysis and management of life science data within marine, health, genomics, proteomics, structural analysis and more.



### NeLS e-infrastructure

NeLS, the Norwegian e-Infrastructure for Life Sciences, for analysis, sharing and storage of high throughput genomics data.



### Training

Local and national training in bioinformatics analysis and use of our tools.



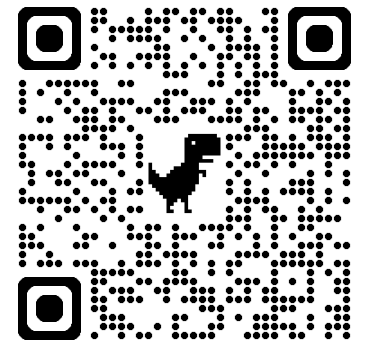
Norges miljø- og biovitenskapelige universitet



The Research Council of Norway

<https://elixir.no/>

Scan me!



**ELIXIR**  
**NORWAY**

# Data storage in NeLS and StoreBioInfo (SBI)

Norwegian e-infrastructure for Life Sciences - developed and operated by Elixir Norway

## NeLS in general

Available to all Norwegian researchers

Enable collaborative projects

Multiple storage layers

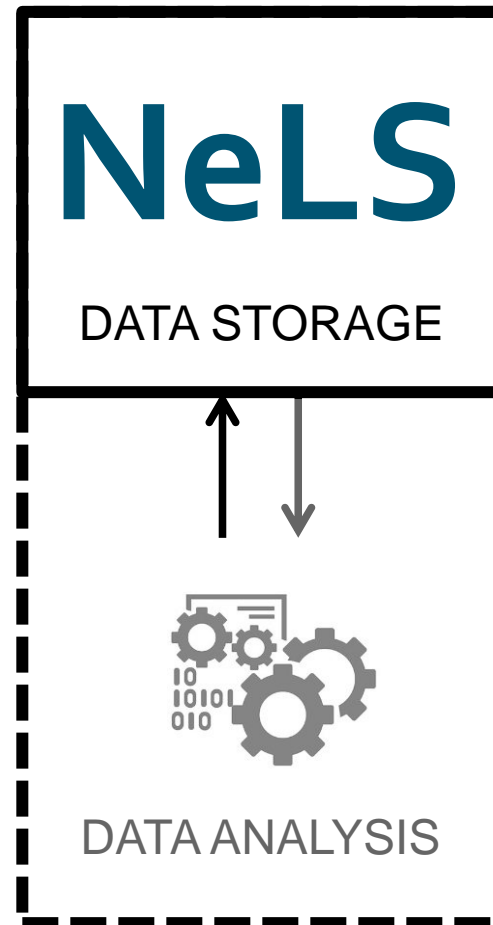
Integrated with compute

Multiple permission roles - shared data

Free of charge\*

Support for data deposition

User support: Elixir Norway helpdesk



## Type of data

Non-sensitive data

Support multiple data types

SEEK integration for metadata



# NeLS

Norwegian e-Infrastructure for Life Sciences



## SERVICES

ELIXIR Norway offers services and infrastructure for life science, including software tools and workflows, databases and storage. In addition ELIXIR Norway provides support through its national helpdesk and offers trainings. Moreover, ELIXIR Norway provides services to the international research community.

## E-infrastructure



NeLS

Galaxy

TSD@UiO

Other sensitive data platforms

The [Norwegian e-infrastructure for life sciences \(NeLS\)](#),  
is a unique national portal for sharing, analysis and  
**midterm storage** of **non-sensitive** molecular life science data.

Free up to 10Tb

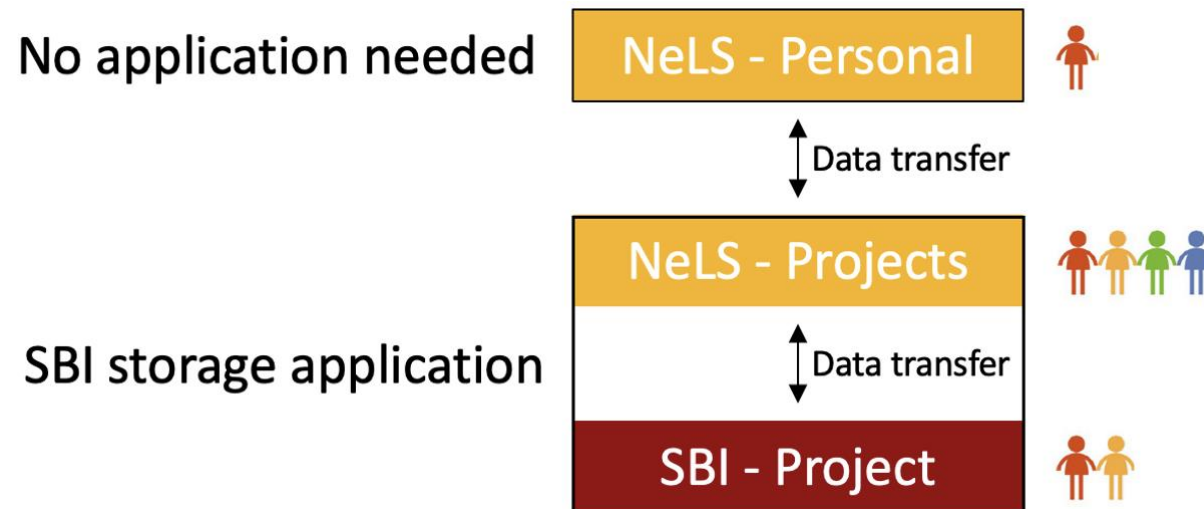


# *Data storage in NeLS/SBI – How to get access*

Storage application needed for projects where data is shared by many users

Access via FEIDE user or NeLS idp can be made for non-FEIDE users

Free storage < 10 TB




Apply for storage: [contact@bioinfo.no](mailto:contact@bioinfo.no)

Wiki for usage: <https://nels-docs.readthedocs.io/en/latest/>


# Login to NeLS using the WebPortal

<https://nels.bioinfo.no/nels-web/#/>

Login

Welcome


This is web portal for Norwegian e-Infrastructure for Life Sciences.



### Helpdesk

Short- and long-term support with bioinformatics analyses, programming and data management tasks. [Contact the helpdesk](#)


Read more . . .



### Services

Analysis and management of life science data within marine, health, genomics, proteomics, structural analysis and more.


Read more . . .



### NeLS e-infrastructure

NeLS, the Norwegian e-Infrastructure for Life Sciences, for analysis, sharing and storage of high throughput genomics data.

Read more . . .



### Training

Local and national training in bioinformatics analysis and use of our tools.

Read more . . .



Scan me!

## Login

You need a valid user credential in order to access NeLS resources. Select your identity provider:

Login with FEiDE Identity

### Norwegian Universities

All Norwegian universities and many other research and education institutions provide a federated login mechanism to allow their users to utilize their own home institution login credentials to access approved online web application services. [Read More . . .](#)

Login with NeLS Identity

### Other Norwegian Users and International Collaborators

For users not having access to a Norwegian university account or other academic institution utilizing FEiDE, we provide our own NeLS identity service for both national and international users.

## Log in with Feide

NeLS

You need to log in via Feide to access NeLS Portal.

Your affiliation



NMBU Norwegian University of Life Sciences  
[Change affiliation](#)



Use work or school account

All users with FEiDE login  
credentials can access the NeLS  
Portal

<https://nels.bioinfo.no/nels-web/#/>

# I don't have FEIDE login credentials (Non-Norwegian collaborators)


- It is also possible to login to NeLS via a NeLS idp. You can apply for a NeLS idp


## NeLS

Norwegian e-Infrastructure for Life Sciences

### Enter your username and password

A service has requested you to authenticate yourself. Please enter your username and password in the form below.

 Username

 Password

[Login](#) [Forgot Password ?](#)

[Apply for a NeLS Account](#) [Help \(Support\)](#)

### Request for NeLS Account

**First Name\***

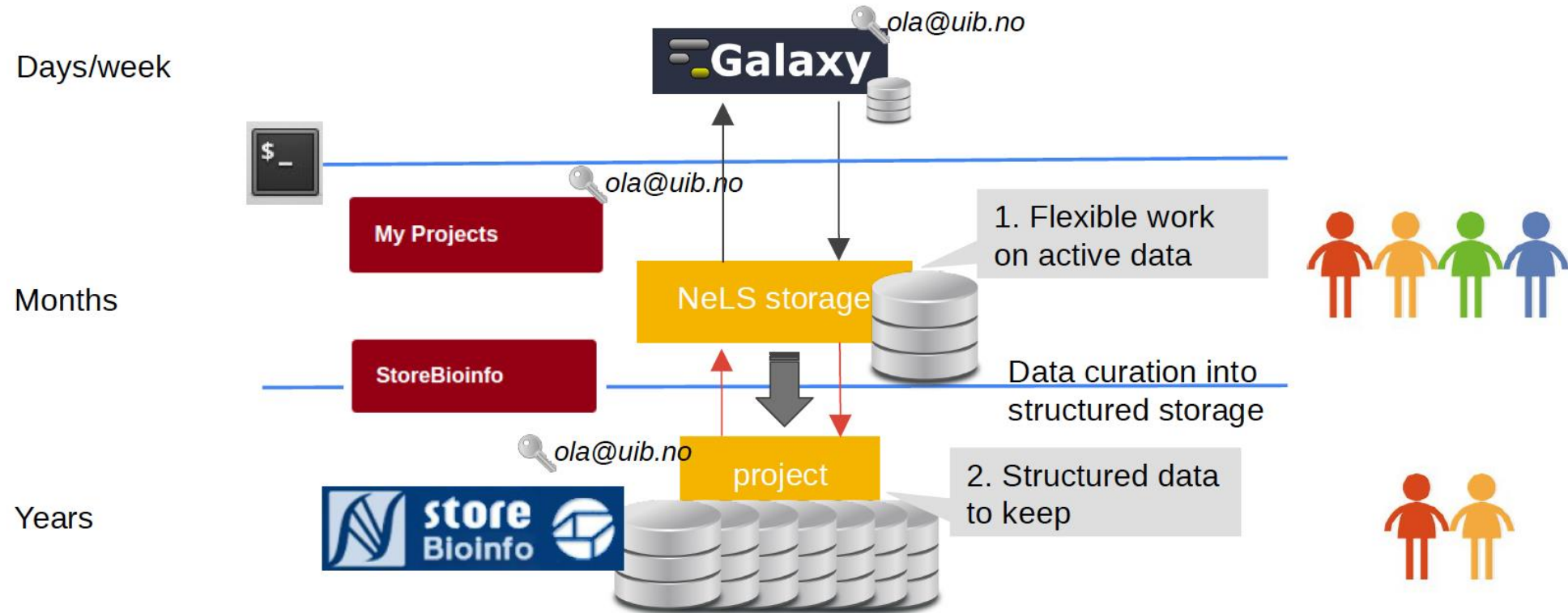
**Last Name\***

**Email\*** (note: email needs to be home institution email address, not personal email)

**Affiliation\***

**Usage Description\***

# NeLS architecture



Decades

**NIRD** RESEARCH DATA ARCHIVE

Research  
data archive

**elixir**  
Core Data  
Resource

**ENA**  
European Nucleotide Archive

**ArrayExpress**

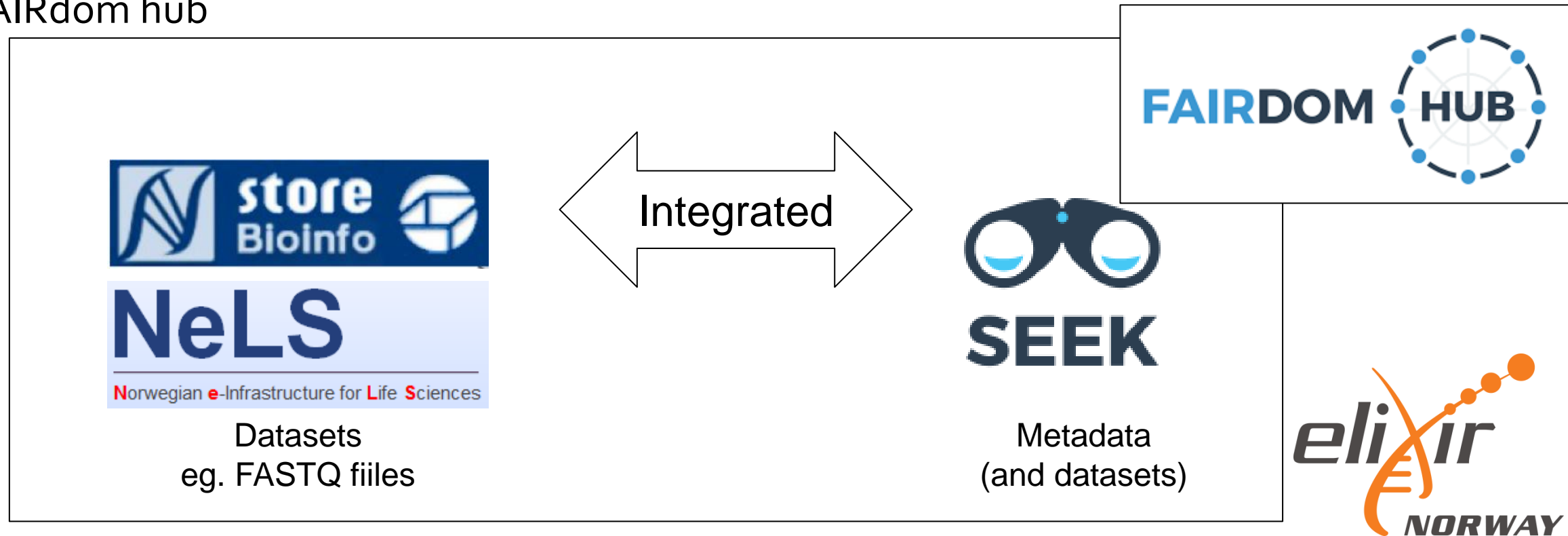
**PRIDE**

**elixir**  
NORWAY

# *SEEK - sharing heterogeneous scientific research datasets, models or simulations, processes and research outcomes*

The SEEK platform is a web-based tool for organising and storing data, and for exploring and annotating data

Norwegian users can link datasets stored in NeLS to a SEEK project using FAIRdom hub



# Data storage in NIRD

National e-Infrastructure for Research Data - owned and operated by UNINETT Sigma2

## NIRD in general

Available to all users with account

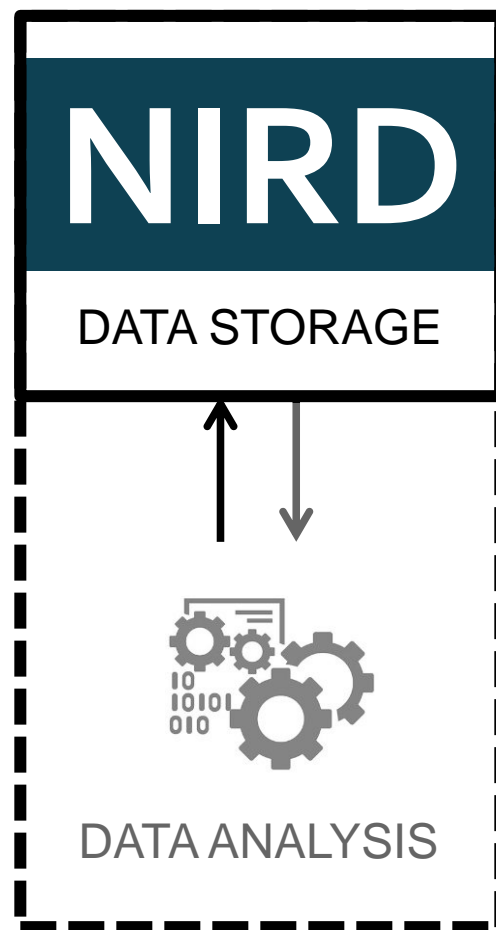
Enable collaborative projects

Integrated with compute

Multiple permission roles - shared data

Free of charge\*

User support: Metacenter support

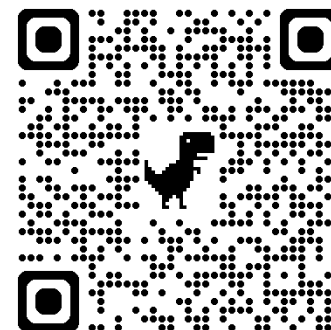


## Type of data

Non-sensitive data

Support multiple data types

All type of scientific data\*



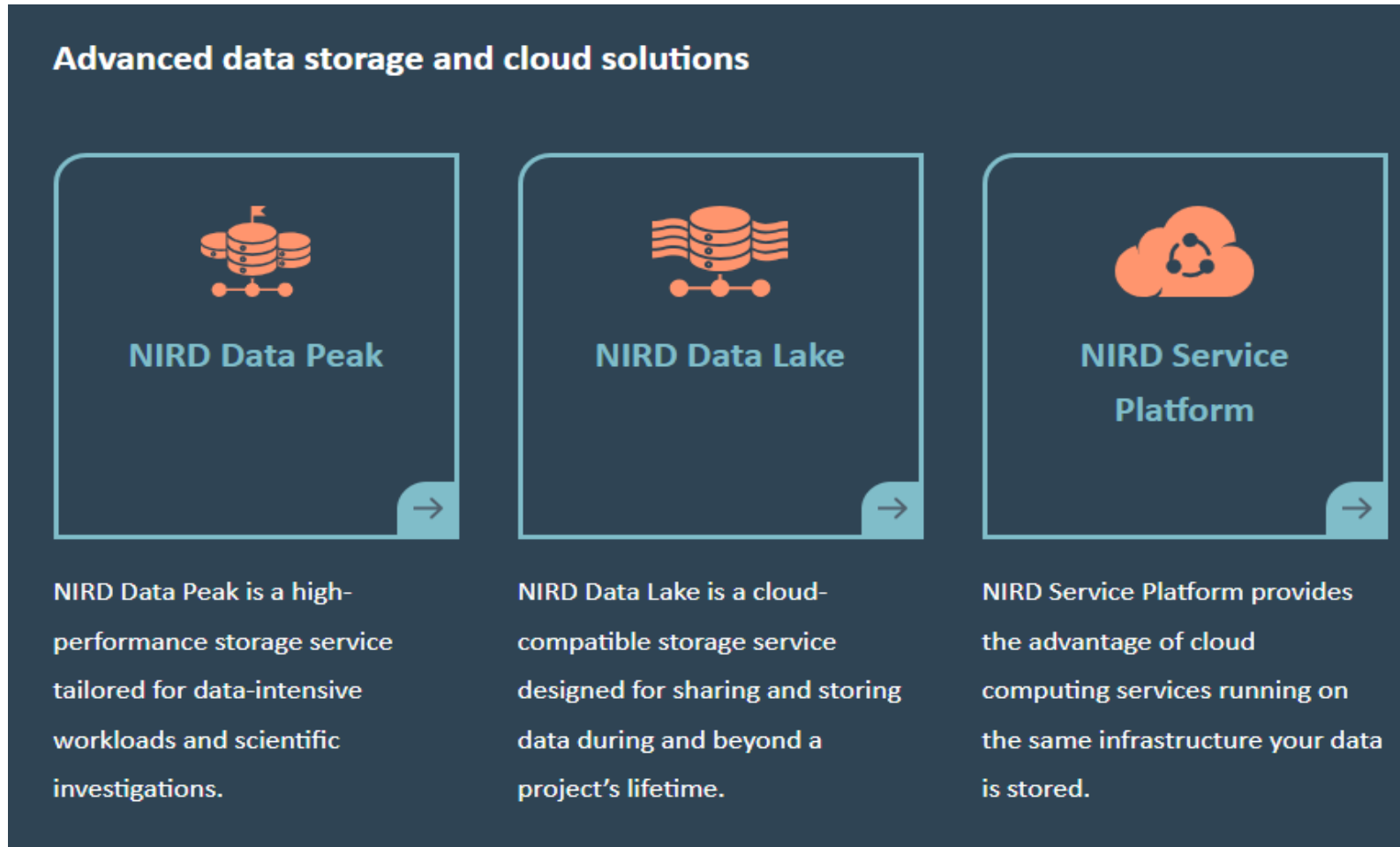
Scan me!





# Data storage in NIRD

National e-Infrastructure for Research Data - owned and operated by UNINETT Sigma2



intensive I/O operations

Long term-storage

NIRDTolKit

# Data storage in NIRD – How to get access

Apply for a UNINETT SIGMA2 user

Apply for a storage quota - Free storage < 10 TB

## Four price categories

The model introduces four different categories of contributors.

- Category A  
Large projects with funding from the Norwegian Research Council of 15 MNOK or more or storage needs above 10 TiB, paying for operational expenses
- Category B  
Non-commercial projects needing Dedicated Resources, paying for capital and operational expenses
- Category C  
Commercial research and industry which will pay the full cost price
- Category 0  
Non-contributing projects which are smaller projects not in any of the above categories

## The process explained

- 1 Submit user application
- 2 Signature by project management
- 3 Review by Sigma2
- 4 Set a password and log in

Scan me!



<https://www.sigma2.no/apply-e-infrastructure-resources>



# Access data storage in NIRD

The NIRD storage can be accessed via the command line using SSH

Data import/export via command line tools such as SCP or SFTP or other file transfer tools (e.g. WinSCP)

```
(base) auve@L004538:~$ NIRD
Last login: Wed Nov  6 20:22:33 2024 from 158.36.102.145
Welcome to nird.sigma2.no!

Documentation: https://documentation.sigma2.no/
Support email: support@metacenter.no
Request resources:
https://www.sigma2.no/content/apply-e-infrastructure-resources/
~~~~~
Latest news from: https://opslog.sigma2.no

o 2024-11-25: Reduced capacity due to NTNU cooling upgrade week 48
o 2024-11-05: Slowness on Saga
o 2024-10-30: Saga is partially down and very slow.
o 2024-10-28: NIRD datalake full
o 2024-10-21: Issues with login to Fram
~~~~~
Module command is not yet stable. Please report any issue encountered.
~~~~~
(miniforge3) [auve@login0-nird-lmd: ~]$ █
```



# All data storage in NIRD can be accessed in all the Sigma2 HPC (e.g. SAGA)

[🏠](#) / [NIRD](#) / NIRD mounts on clusters

## NIRD mounts on clusters

The NIRD project storage areas, namely NIRD Data Peak (TS) and NIRD Data Lake (DL) are mounted on the login nodes of Betzy, Fram, and Saga. One can directly access the NIRD project area from the login nodes of the aforementioned compute clusters.

The path to Data Peak (TS) project areas is

```
/nird/datapeak/NSxxxxK
```

while the path to Data Lake (DL) project is

```
/nird/datalake/NSxxxxK
```

where NSxxxxK is the ID of the project.

```
(BASICS)[auve@login-4: ~]$ ls /nird/datapeak/
acl@NS10016K NS10082K NS11026K NS11105K NS5001K NS8044K NS8302K NS9020K NS9059K NS9188K NS9419K NS9561K NS9623K NS9748K NS9819K NS9879K
files-ns9012kNS10022K NS10086K NS11027K NS12002K NS5011K NS8045K NS8402K NS9021K NS9061K NS9191K NS9461K NS9564K NS9630K NS9754K NS9825K NS9880K
framNS10023K NS10088K NS11028K NS2345K NS5023K NS8052K NS8403K NS9029K NS9063K NS9206K NS9478K NS9576K NS9635K NS9765K NS9828K NS9883K
migratorNS10024K NS10091K NS11030K NS2506K NS6000K NS8060K NS8999K NS9033K NS9064K NS9207K NS9481K NS9588K NS9640K NS9769K NS9829K NS9889K
migrator-betzyNS1002K NS10100K NS11032K NS2615K NS8001K NS8061K NS9000K NS9034K NS9065K NS9252K NS9486K NS9593K NS9641K NS9772K NS9832K NS9897K
mmbbackup.unsupported.PROJECT1NS1003K NS1012K NS11043K NS2649K NS8002K NS8065K NS9001K NS9035K NS9066K NS9279K NS9493K NS9595K NS9644K NS9777K NS9835K NS9898K
mmbbackup.unsupported.protect1_tsminst1NS10044K NS1014K NS11048K NS2806K NS8006K NS8075K NS9003K NS9036K NS9067K NS9295K NS9518K NS9599K NS9648K NS9781K NS9841K NS9970K
NS0000KNS10047K NS1015K NS11051K NS2875K NS8014K NS8085K NS9005K NS9039K NS9073K NS9305K NS9525K NS9600K NS9672K NS9787K NS9853K NS9985K
NS10003KNS1004K NS1018K NS11054K NS2900K NS8015K NS8087K NS9008K NS9042K NS9081K NS9312K NS9529K NS9603K NS9689K NS9794K NS9864K NS9988K
NS10005KNS10054K NS11002K NS11082K NS2993K NS8021K NS8095K NS9010K NS9043K NS9083K NS9320K NS9530K NS9607K NS9691K NS9806K NS9868K NS9989K
NS10006KNS10055K NS11007K NS11083K NS4605K NS8028K NS8112K NS9013K NS9048K NS9114K NS9370K NS9531K NS9608K NS9696K NS9809K NS9869K NS9990K
NS1000KNS10064K NS11009K NS11084K NS4654K NS8029K NS8121K NS9014K NS9052K NS9121K NS9378K NS9540K NS9610K NS9699K NS9813K NS9873K NS9992K
NS10013KNS10068K NS11016K NS11085K NS4659K NS8035K NS8122K NS9015K NS9054K NS9122K NS9403K NS9559K NS9611K NS9715K NS9817K NS9874K NS9997K
NS10014KNS10069K NS11022K NS11096K NS4704K NS8039K NS8128K NS9019K NS9057K NS9181K NS9408K NS9560K NS9616K NS9745K NS9818K NS9876K NS9999K
(BASICS)[auve@login-4: ~]$
```

[https://documentation.sigma2.no/files\\_storage/nird/mounts\\_lmd.html](https://documentation.sigma2.no/files_storage/nird/mounts_lmd.html)



# NIRD ToolKit

## What is the NIRD Service Platform

NIRD Service Platform is a platform to run cloud services, such as web services, domain and community-specific portals, and tools for data visualization, pre/post-processing, data discovery and data sharing.

The service platform offers dedicated resources for data sharing and computing with large datasets. Here you can effortlessly process and share data with your preferred tools, such as

- Jupyter Notebook
- Jupyter Hub
- Spark
- R-Studio

## NIRD Toolkit

Filter packages

### deep-learning-tools



GPU enabled Jupyter notebook

[Read more](#)

[Install](#)

### desktop-vnc



VNC Ubuntu

[Read more](#)

[Install](#)

### jupyter



Apache Spark Scala R VS Code

[Read more](#)

[Install](#)

### jupyterhub



Jupyter notebook server

[Read more](#)

[Install](#)

### minio



File sharing Cloud storage

[Read more](#)

[Install](#)

### rstudio



R web IDE

[Read more](#)

[Install](#)

### spark



Distributed computing

[Read more](#)

[Install](#)

<https://apps.sigma2.no/>

# NIRD ToolKit: Run Rstudio to analyze data directly on NIRD



NIRD Toolkit

Applications

Arturo Vera Ponce De Leon

Log out

rmemo

Running

> Overview

> Modify

> Resources

## Overview

URL(s) [rmemo.nirdmemo.sigma2.no](https://rmemo.nirdmemo.sigma2.no)  
[shiny-rmemo.nirdmemo.sigma2.no](https://shiny-rmemo.nirdmemo.sigma2.no)  
Contact email [arturo.vera.ponce.de.leon@nmbu.no](mailto:arturo.vera.ponce.de.leon@nmbu.no)  
Projectspace nirdmemo-ns9864k  
Resource usage 2.3 CPU, 8.21 GB RAM

### rstudio



RStudio is an integrated development e...  
1.3.1

### ▼ Mounted volumes

Secret

ConfigMap

Temporary storage

Persistent storage

Volume name

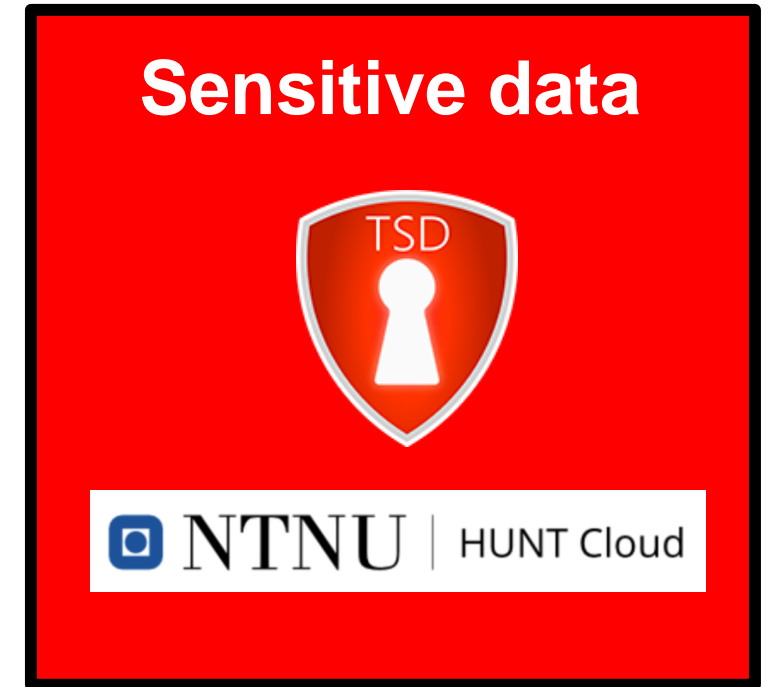
oidconfig

Mount path

/conf



# *Sensitive data storage*



# Data storage in TSD

National service for sensitive data - developed and operated by UiO

## TSD in general

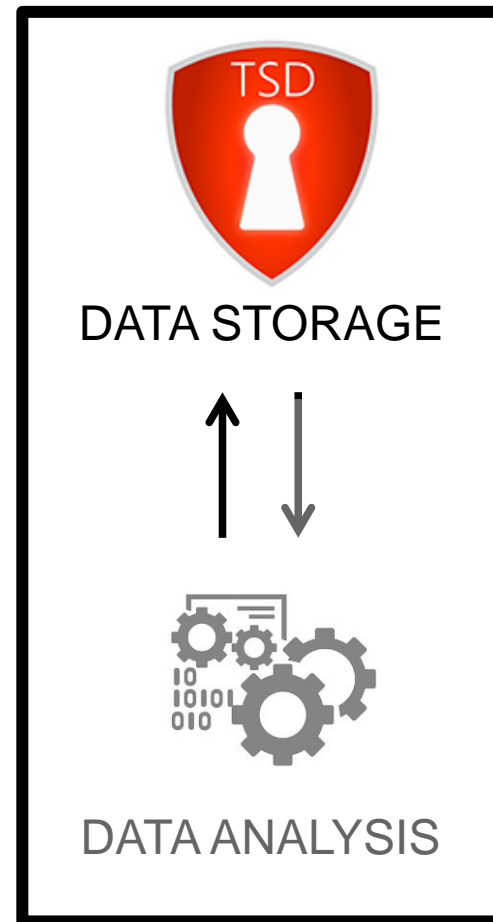
Available to all users with account

Enable collaborative projects

Integrated with compute

Multiple permission roles - shared data

User support: Tech and admin support



## Type of data

Sensitive data

Support multiple data types

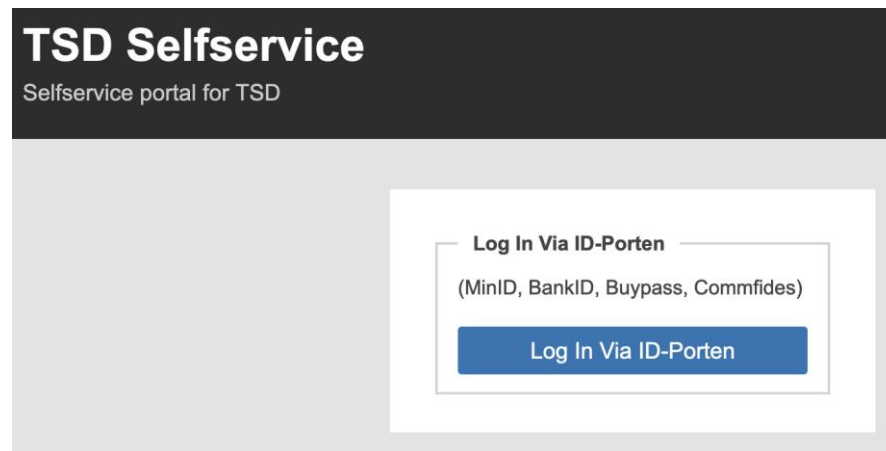


# Data storage in TSD – How to get access

Apply for a project – need to document ethical approval (e.g. from REC)

Generate a TSD user – require 2-factor authentication

Payment models for storage



Apply for a TSD project through ELIXIR.NO:  
ELIXIR Norway has a quota in TSD that, and  
users can get subsidised quota through  
ELIXIR in TSD

Generate user account: <https://selfservice.tsd.usit.no/>

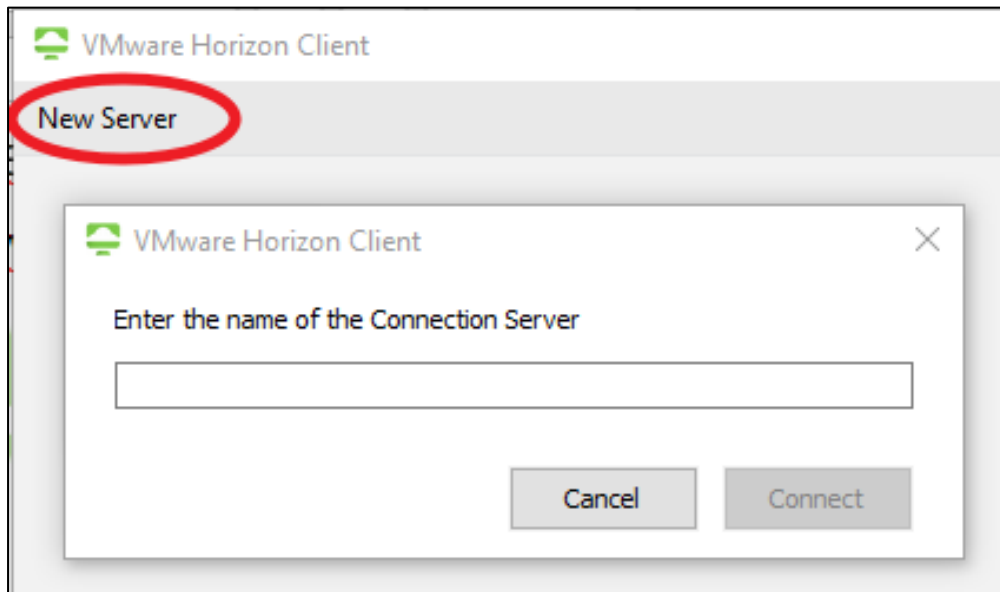
Apply for project: <https://www.uio.no/tjenester/it/forskning/sensitiv/hjelp/start/registrer.html>

Wiki for usage: <https://www.uio.no/tjenester/it/forskning/sensitiv/hjelp/start/index.html>

# Access to data storage in TSD

The TSD storage can be accessed via VMware Horizon + 2-factor authentication

Import/export data via the web file upload service



## TSD Data Portal

### Welcome to the TSD Data Portal

To use this service you should already hold a valid TSD user account.



Import Files



Export Files

# Where should I store and share my life science data?

	Sensitive data	Direct transfer from data producer	Web portal access*	Command line access	Shared projects	Integration for metadata	Support for data deposition	Free of charge
NeLS/SBI	✗	✓	✓	✓	✓	✓	✓	✓
NIRD	✗	✓	✗	✓	✓	✗	✗	✓
TSD	✓	✓	✗	✓	✓	✗	✗	✗

\*Direct access to the data and project management through the web browser

# Example - Data flow/handle using ELIXIR Norway

