



# Introduction to NeLS

*The national bioinformatics infrastructure*

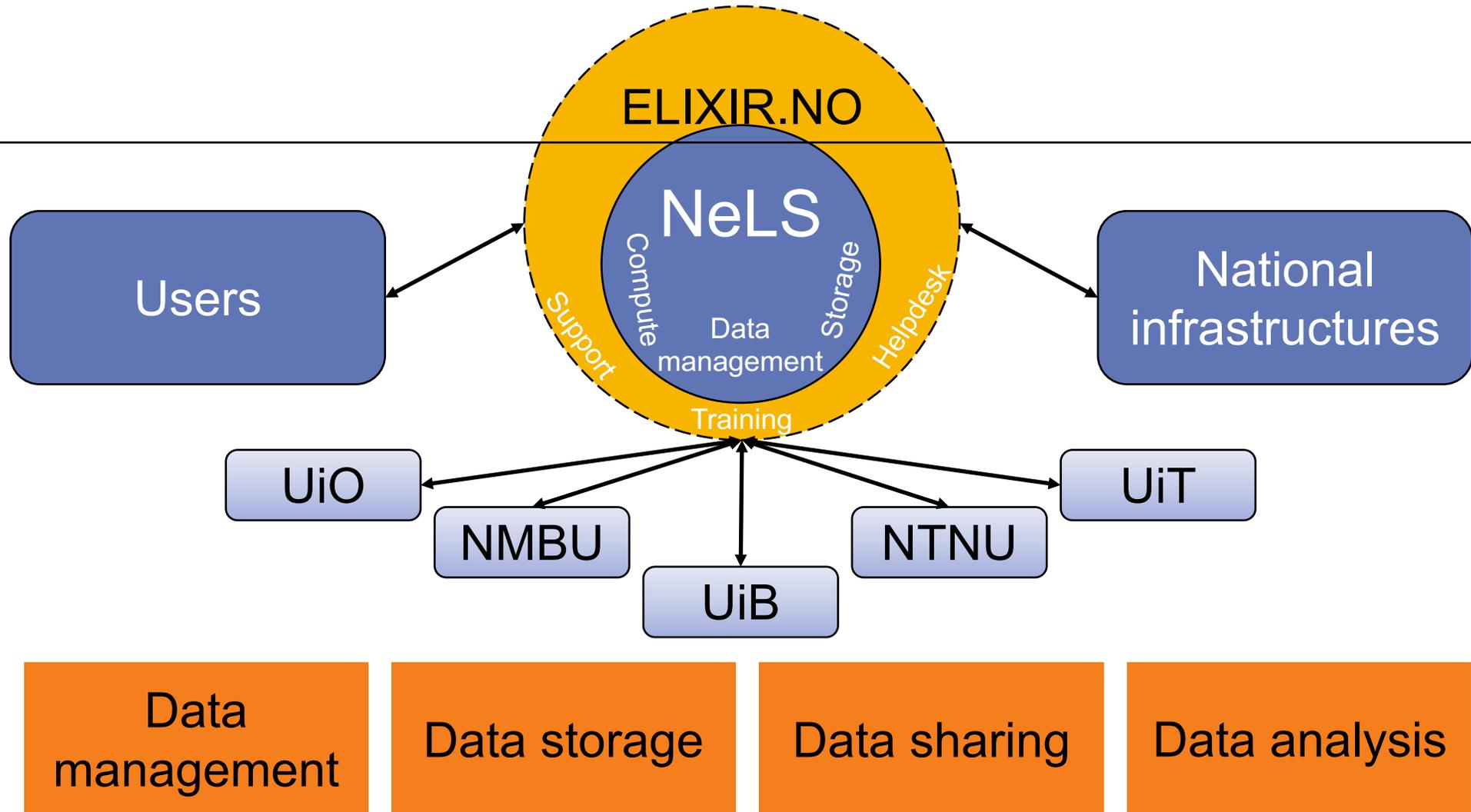


Erik Hjerde

# NeLS - the Norwegian e-Infrastructure for Life Science

International

**National**



# Overview of data management

FAIR data management from data collection to public repository

Sample collection



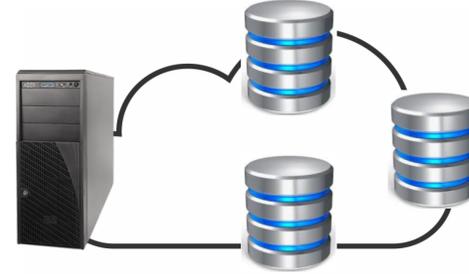
Data production



Data storage/sharing



Data processing



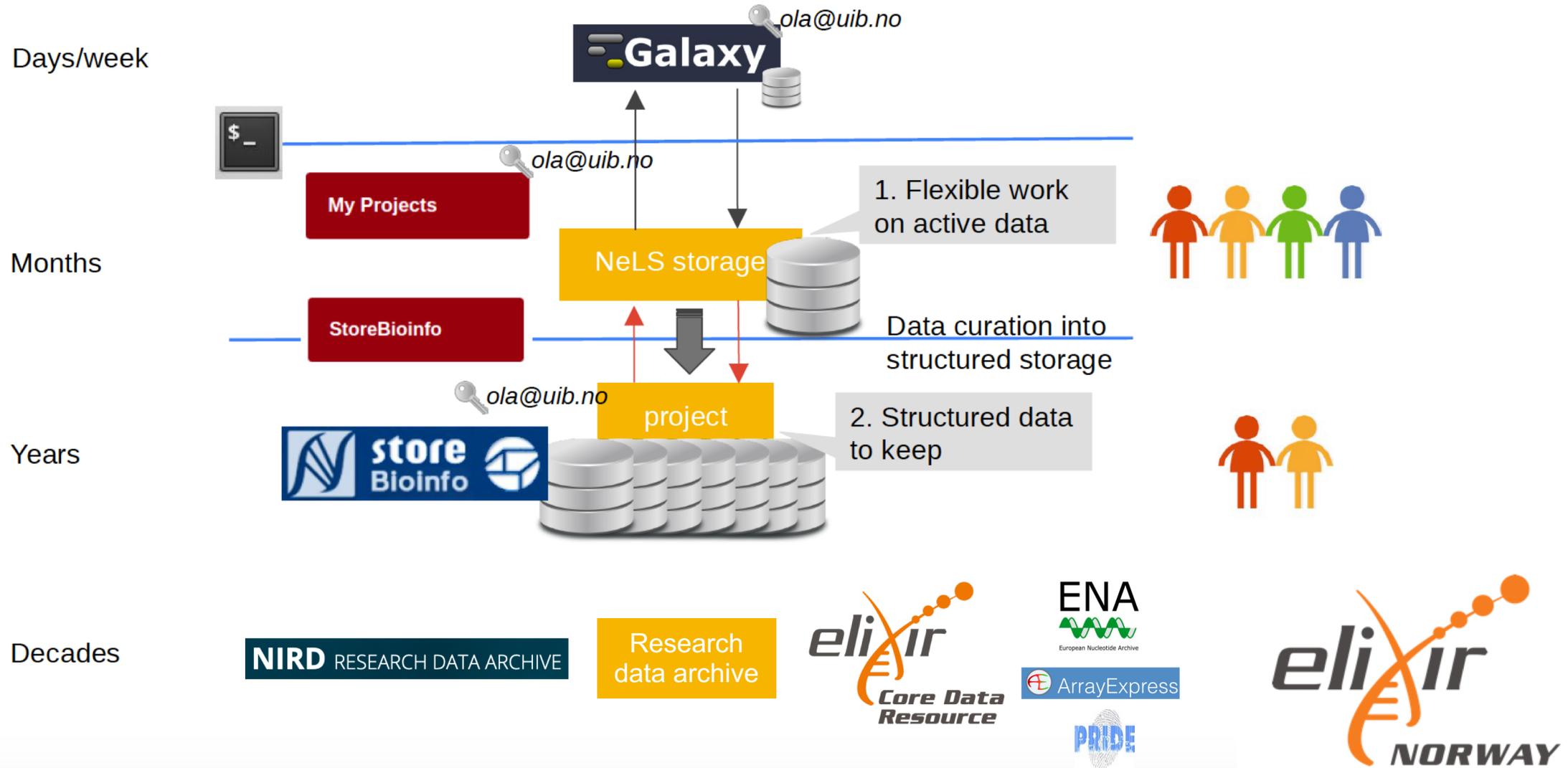
Data archiving



Data management

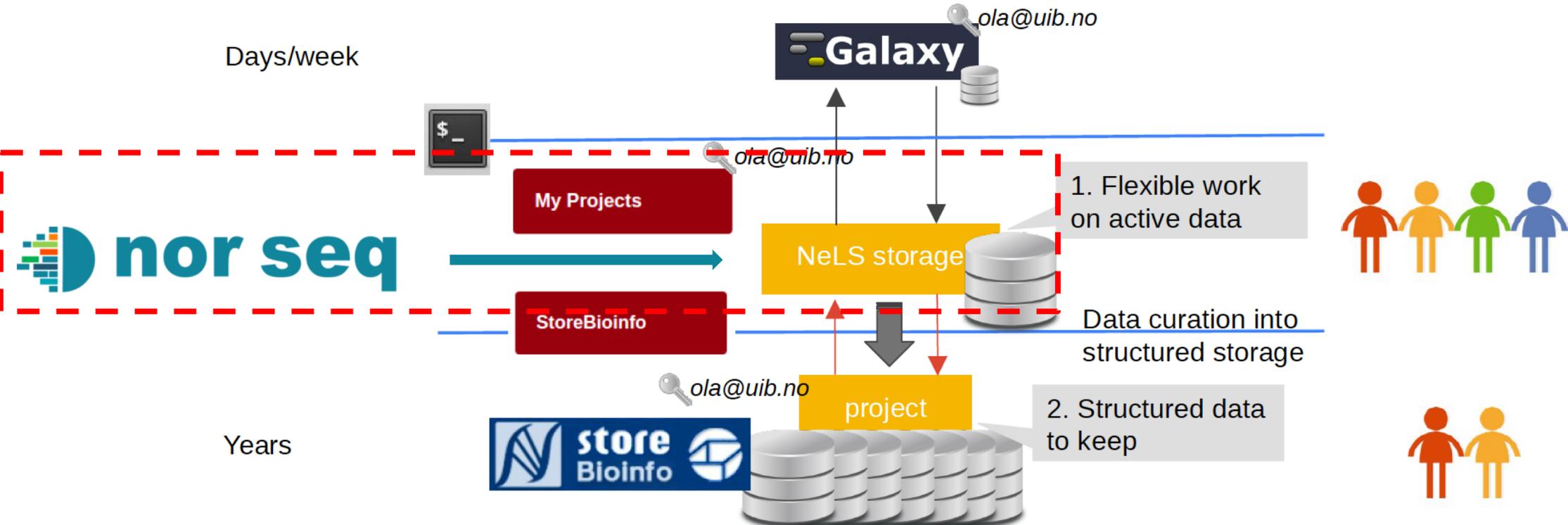


# NeLS architecture



# Data from data providers

Days/week



Years



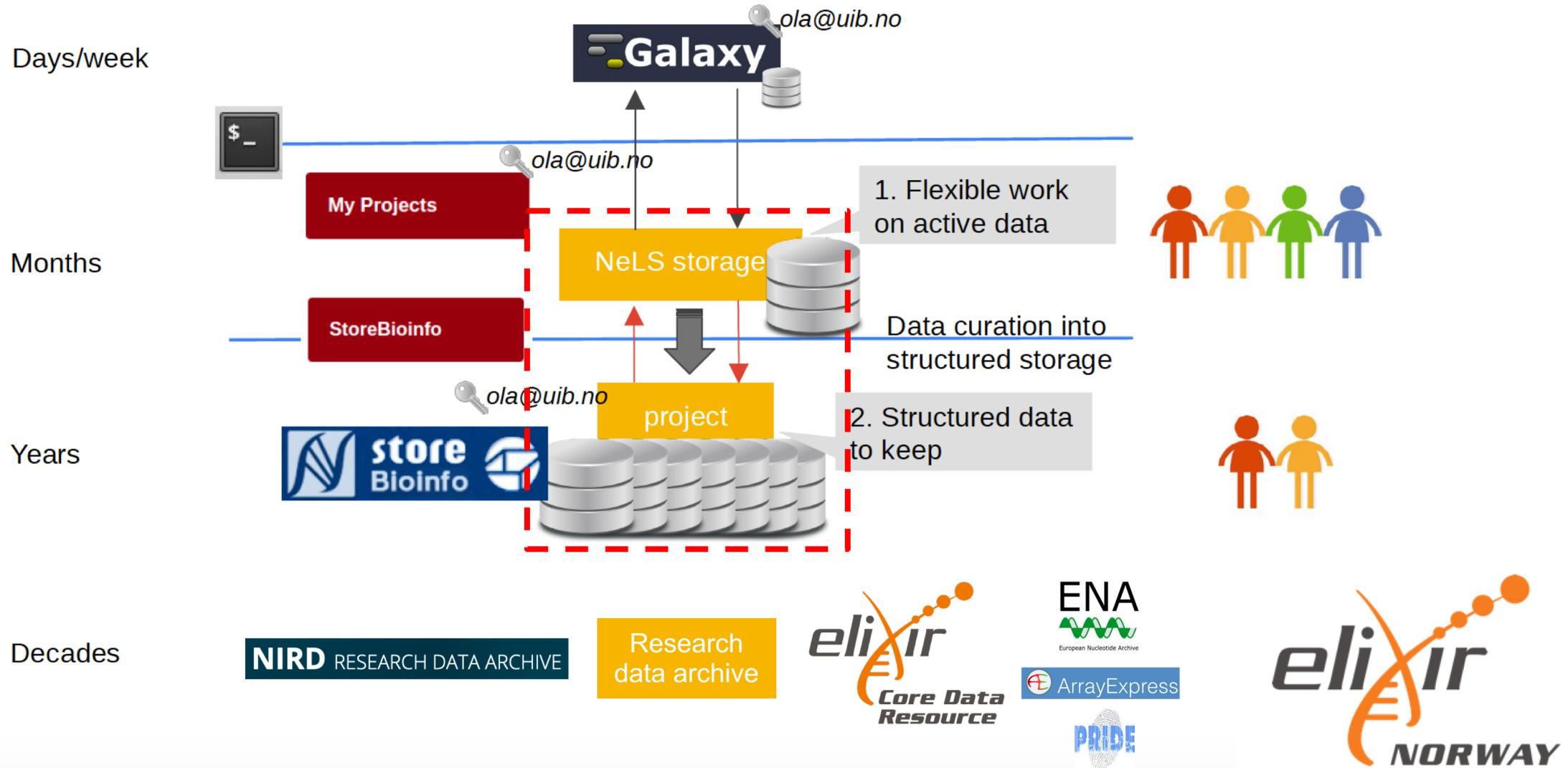
Decades

**NIRD** RESEARCH DATA ARCHIVE

Research data archive



# Data storage in NeLS and StoreBioInfo



# Access to storage and sharing of data in NeLS

Temporary storage – typically during the data analysis and manuscript preparation

Users with permissions can access data in a NeLS project (FEIDE or NeLS ID)

The screenshot displays the NeLS portal interface with a login overlay. The main page content is partially obscured by the overlay, but the following information is visible:

- NeLS** (top navigation bar)
- Welcome !** (left sidebar):
  - This is NeLS portal for the administration of Norwegian e-Infrastructure for Life Sciences. NeLS is one of the packages of ELIXIR Norway.
  - Features**
    - ✓ Federated login using [FEIDE](#)
    - ✓ Sharing of data
    - ✓ Bioinformatics pipelines
    - ✓ API for integration
  - Help (Support)**
    - Send an e-mail to [contact@bioinfo.no](mailto:contact@bioinfo.no) if you have difficulty, question, remark or suggestions. You should include your full name and your affiliation in your request.
- ELIXIR Norway** (middle sidebar):
  - ELIXIR Norway is coordinated by the [University of Bergen](#) and [Universities in Oslo, Trondheim, Tromsø and Ås](#). It receives funding from the [Research Council of Norway](#) through its research infrastructure. NeLS is also supported by the participating institutions.
  - The project aim include:**
    - ✓ To build a Norwegian Node in the pan-European e-Infrastructure [ELIXIR](#), with the Node delivering services and resources to the international community through selected areas
    - ✓ To continue a National help desk serving users and providing services and assistance; and
    - ✓ To provide an e-infrastructure allowing users to easily access, safely store, share, analyse and publish their genomic data
  - Logos for [University of Bergen](#), [University of Oslo](#), [University of Tromsø](#), [University of Ås](#), and [The Research Council of Norway](#).
- Access NeLS** (right sidebar):
  - Log in with Feide**
  - NeLS Portal has requested you to log in with Feide. [NeLS](#)
  - Log in with your Feide account from **Norwegian University of Life Sciences**. [Not your affiliation?](#)
  - Input fields for username and password.
  - [Forgot your username or password?](#)
  - Log in** button.
  - [Help](#) link.
  - [Privacy and cookie information](#) link.

# Access to storage and sharing of data in NeLS

## Access to personal storage and projects

The screenshot displays the NeLS web interface with several overlapping panels. The top navigation bar includes 'NeLS', 'Dashboard', 'Personal', 'Projects', 'StoreBioInfo', and 'Help (Support)'. The user is identified as 'Erik Hjerde (NeLS ID :34)'. A 'System Notice' banner is visible at the top.

**NeLS Disk Usage:** A line graph showing disk usage over time from 01-Sep-2016 to 01-May-2018. The usage starts at approximately 74.5 GB and remains relatively stable, ending at about 74.5 GB.

**Account Information:**

Full Name	Erik Hjerde
Email	erik.hjerde@uit.no
Affiliation	University of Tromsø

**My Data:** A list of folders under the 'Personal' view:

- Bio3323
- ELIXIR2-WP4
- Matt
- metagenome
- Test2014

**Projects:** A table listing various projects with their names, roles, and creation dates.

Project Name	Role in Project	Creation Date
<b>Elixir_NO-Documents</b> A project for sharing Elixir.no related documents	Member Add files & folders, Navigate & Download	November 28, 2014
<b>Uio-Test-Project</b> for demo	Member Add files & folders, Navigate & Download	March 19, 2015
<b>UIO_Halvorsen_RNAseq_Mouse_Spleen_T_cells_2018</b> RT#13277._____ Aim: To compare transcriptome of spleen T cells from WT mice and a transgenic mouse model.	Power User Manage File system: Add, Rename, Navigate, Download, Delete all Content	February 22, 2018
<b>Elixir_workshops</b> Folder holding materials/data for Elixir workshops	Member Add files & folders, Navigate & Download	April 15, 2016
<b>UIB-Petersen_neuro_ma_regulation_disorder_2017</b> This is part of a larger consortium project, we are providing neuro samples RNA data (RNA-seq), for 100 cell lines cultivated in our lab at University of Bergen.	Power User Manage File system: Add, Rename, Navigate, Download, Delete all Content	May 11, 2017
<b>maseq-test-data</b> Data, 12 samples, two groups, for testing updated maseq-pipeline	Power User Manage File system: Add, Rename, Navigate, Download, Delete all Content	September 11, 2017
<b>ELIXIR2-WP4</b> Collection of course data and documents	Power User Manage File system: Add, Rename, Navigate, Download, Delete all Content	November 8, 2017

**Data Operations:** A list of recent operations, all completed successfully:

- 5859 : NeLS >> StoreBioInfo
- 5858 : NeLS >> StoreBioInfo
- 5857 : NeLS >> StoreBioInfo
- 5856 : NeLS >> StoreBioInfo
- 5855 : NeLS >> StoreBioInfo
- 5854 : NeLS >> StoreBioInfo
- 5850 : NeLS >> StoreBioInfo

© 2019. Norwegian e-Infrastructure for Life Sciences. All right reserved.

# Data storage in StoreBioInfo

Long term storage – typically for large projects where data is used in several publications

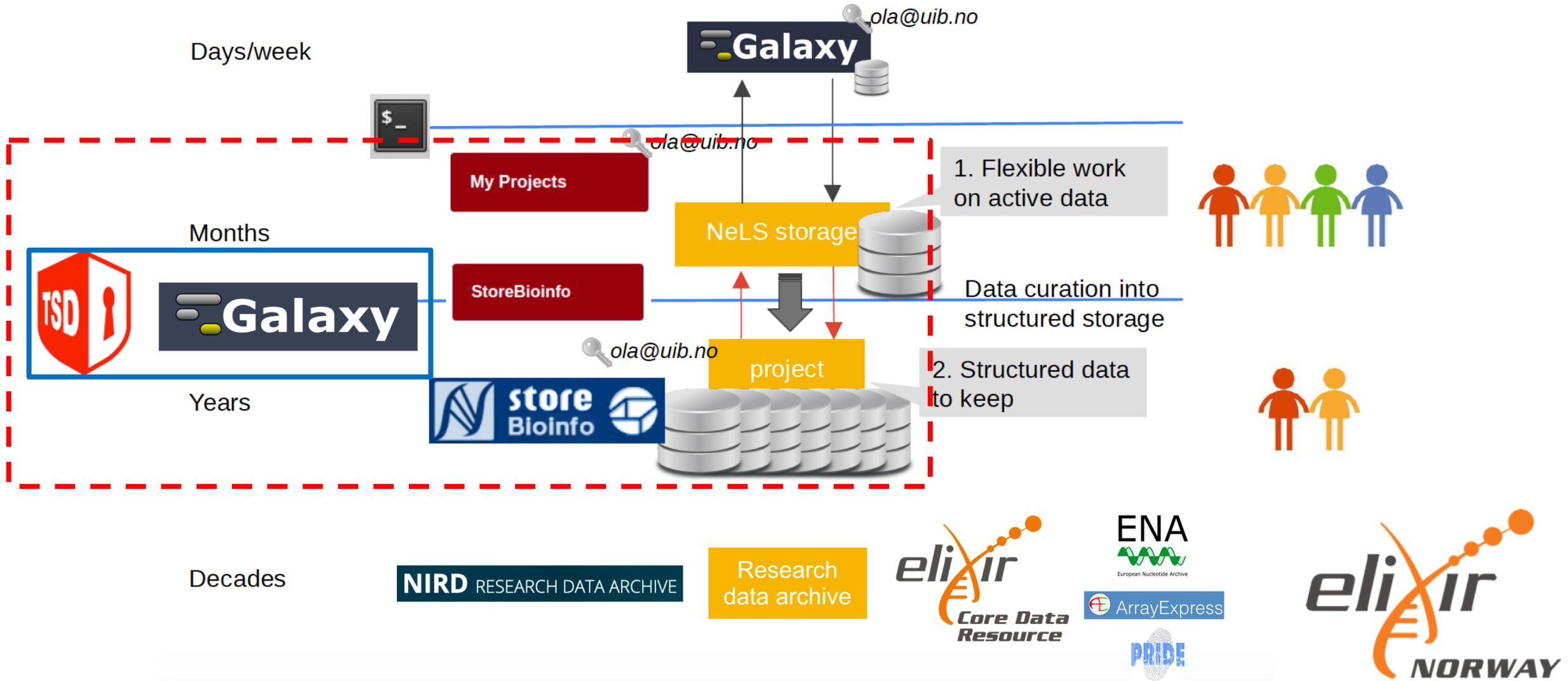
Storage quotas has to be applied for – typically > 1TB data

The screenshot displays the StoreBioInfo web interface. The top navigation bar includes 'NeLS', 'Dashboard', 'Personal', 'Projects', 'StoreBioInfo', and 'Help (Support)'. The user is identified as 'Erik Hjerde (NeLS ID :34)'. The main content area is divided into three sections:

- NeLS:** Shows the user's location and navigation options for 'Personal' and 'Projects'.
- StoreBioInfo:** Displays a list of projects with columns for 'Name' and 'Creation Date'. The projects listed are:

Name	Creation Date
UiT_bioinformatics_platform_core_storage_facilities_20	October 5, 2018
UNN_Goll_Metagenome_Human_IBS_Faeces_2018	March 19, 2019
UiB- Petersen_neuro_rna_regulation_disorder_2017	May 11, 2017
Elixir-demo	November 24, 2015
- Data Operations:** Shows a list of completed data operations, each with a status of 'Completed successfully' and a completion rate of 100%. The operations are identified by IDs such as 5859, 5858, 5857, and 5856.

# Sensitive data in TSD



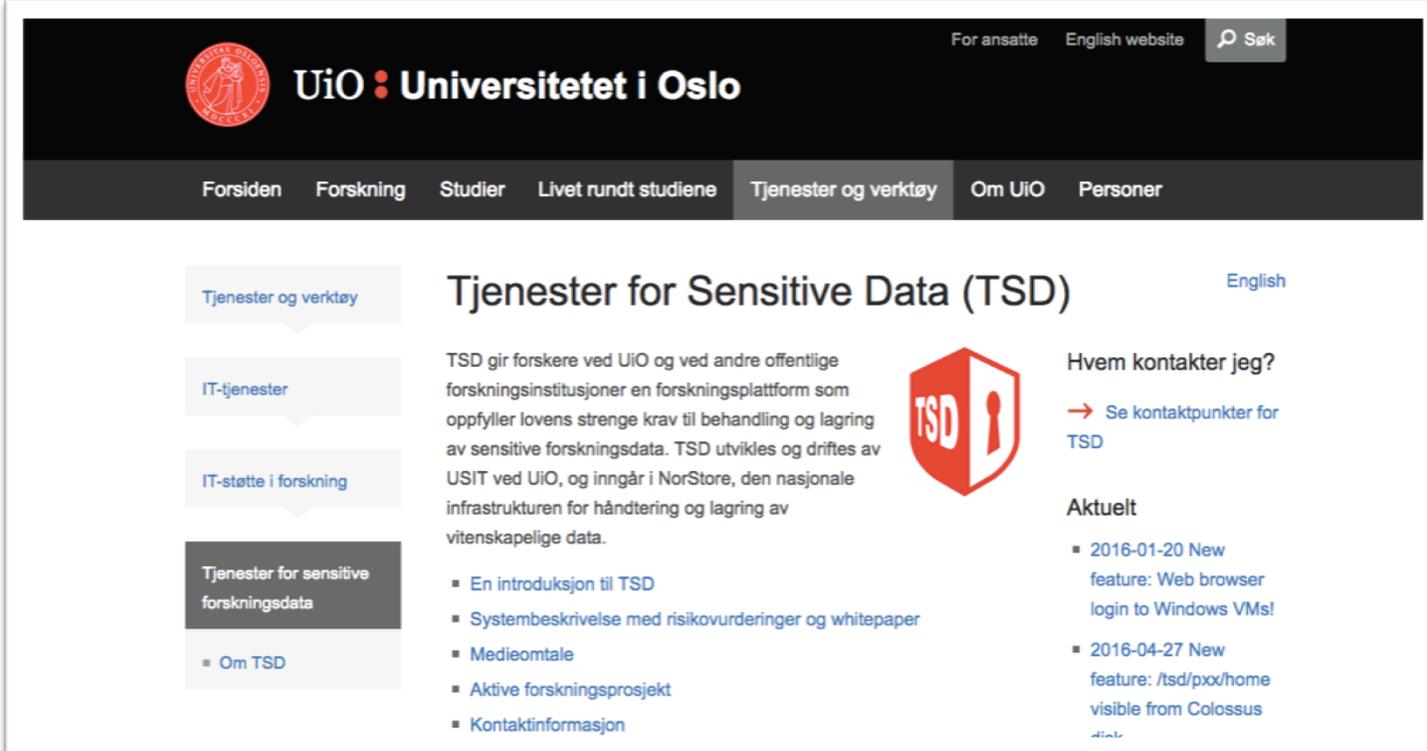
# TSD – Tjenester for Sensitive Data

National infrastructure for computing and data management of sensitive data

Fulfills the requirements by Norwegian law for treatment and storage of sensitive data

TSD is developed and hosted by USIT at the UiO

The NeLS-portal is the only web entry point to get data in and out of TSD



The screenshot shows the TSD page on the UiO website. The header includes the UiO logo and navigation links. The main content area features a sidebar with navigation options, a central text block describing TSD, a red TSD logo, and a right-hand sidebar with contact and news information.

**UiO : Universitetet i Oslo**

For ansatte English website Søk

Forsiden Forskning Studier Livet rundt studiene **Tjenester og verktøy** Om UiO Personer

Tjenester og verktøy

IT-tjenester

IT-støtte i forskning

**Tjenester for sensitive forskningsdata**

- Om TSD

## Tjenester for Sensitive Data (TSD) English

TSD gir forskere ved UiO og ved andre offentlige forskningsinstitusjoner en forskningsplattform som oppfyller lovens strenge krav til behandling og lagring av sensitive forskningsdata. TSD utvikles og driftes av USIT ved UiO, og inngår i NorStore, den nasjonale infrastrukturen for håndtering og lagring av vitenskapelige data.

- En introduksjon til TSD
- Systembeskrivelse med risikovurderinger og whitepaper
- Medieomtale
- Aktive forskningsprosjekt
- Kontaktinformasjon

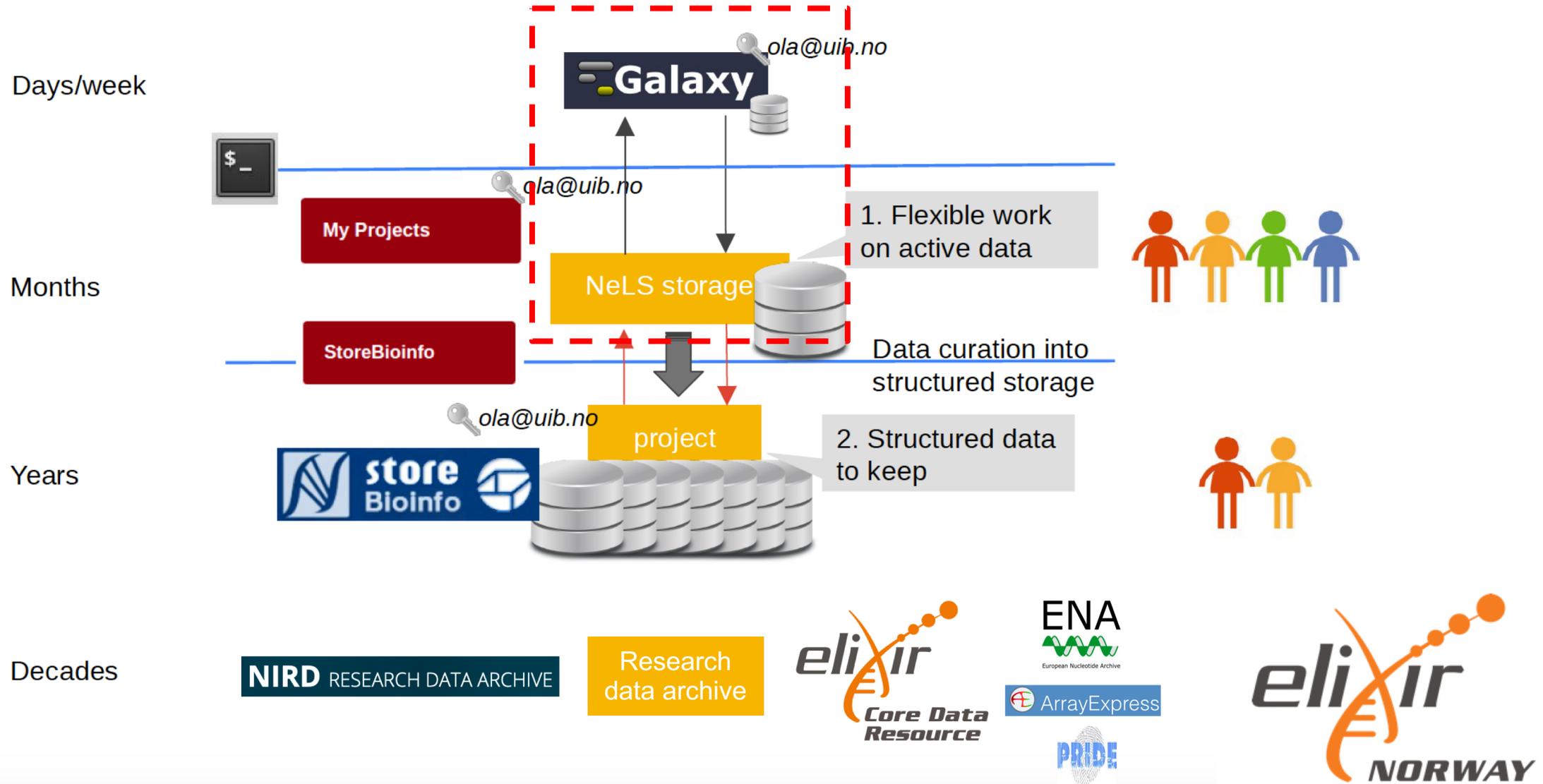
**Hvem kontakter jeg?**

→ [Se kontaktpunkter for TSD](#)

**Aktuelt**

- 2016-01-20 New feature: Web browser login to Windows VMs!
- 2016-04-27 New feature: /tsd/pxx/home visible from Colossus

# Data analysis



# Data analysis in NeLS Galaxy

Each node host a Galaxy server with unique “flavour”

**NeLS galaxy-uit.bioinfo.no**

Tools

search tools

GENERAL GALAXY TOOLS

- Get Data
- Send Data
- Collection Operations
- Text Manipulation
- Filter and Sort
- Join, Subtract and Group
- Convert Formats
- Extract Features
- Metagenomics
- Statistics
- NGS: QC and Manipulation
- NGS: Picard
- NGS: Mapping and Sequence analysis
- NGS: GATK Tools
- Assembly and Validation
- Transcriptomics
- NGS: SAM-tools
- Genome Diversity

Workflows

- All workflows

Welcome to the NeLS Galaxy installation in Tromsø

Galaxy is a web-based platform for data intensive life science research that provides users with a unified, easy-to-use graphical interface to a host of different analysis tools. These tools can be run interactively, one by one, or combined into multi-step workflows that can be executed as a single analysis.

If this is your first time using Galaxy, you might want to have a look at this [Quick Start Guide](#).

Additional documentation and tutorials on using Galaxy can be found [here](#).

This Galaxy server has limitations on disc usage, and you have currently used **117.9 GB** of your total quota of **250.0 GB**. To free up disc space, please move your files to the NeLS Storage after you are finished with them. If you require a larger disc quota, contact the help desk or the local Galaxy administrator.

Tweets by @elixirnorway

ELIXIR Norway Retweeted

Inge Jonassen @ingejonassen

Inspiring days at @elixirnorway all hands meeting in Oslo - stepping up Data Management efforts nationally and in @ELIXIREurope

Embed View on Twitter

History

search datasets

Copy of 'test Exercise I'

3 shown, 1 deleted, 12 hidden

8.5 KB

16: Concatenate datasets on data 15 and data 14

2: file2.txt

1: file1.txt

Tools and Workflows

Contact Support

ELIXIR.NO aims to develop tools and workflows for specific data analysis tasks, based on input from large user groups and help-desk users in Norway and also toward ELIXIR internationally. To see a list of analysis workflows available on this Galaxy server, click the "Shared Data" menu at the top of the page and select [Published Workflows](#). Visit the [NeLS Portal](#) to see the full list of [analysis pipelines](#) available in all the NeLS Galaxies.

Other NeLS Galaxy installations and resources

The Galaxy installations at each of the participating universities have differing focus and may offer different sets of tools and functionalities.

University of Oslo

University of Bergen

Norwegian University of Science and Technology

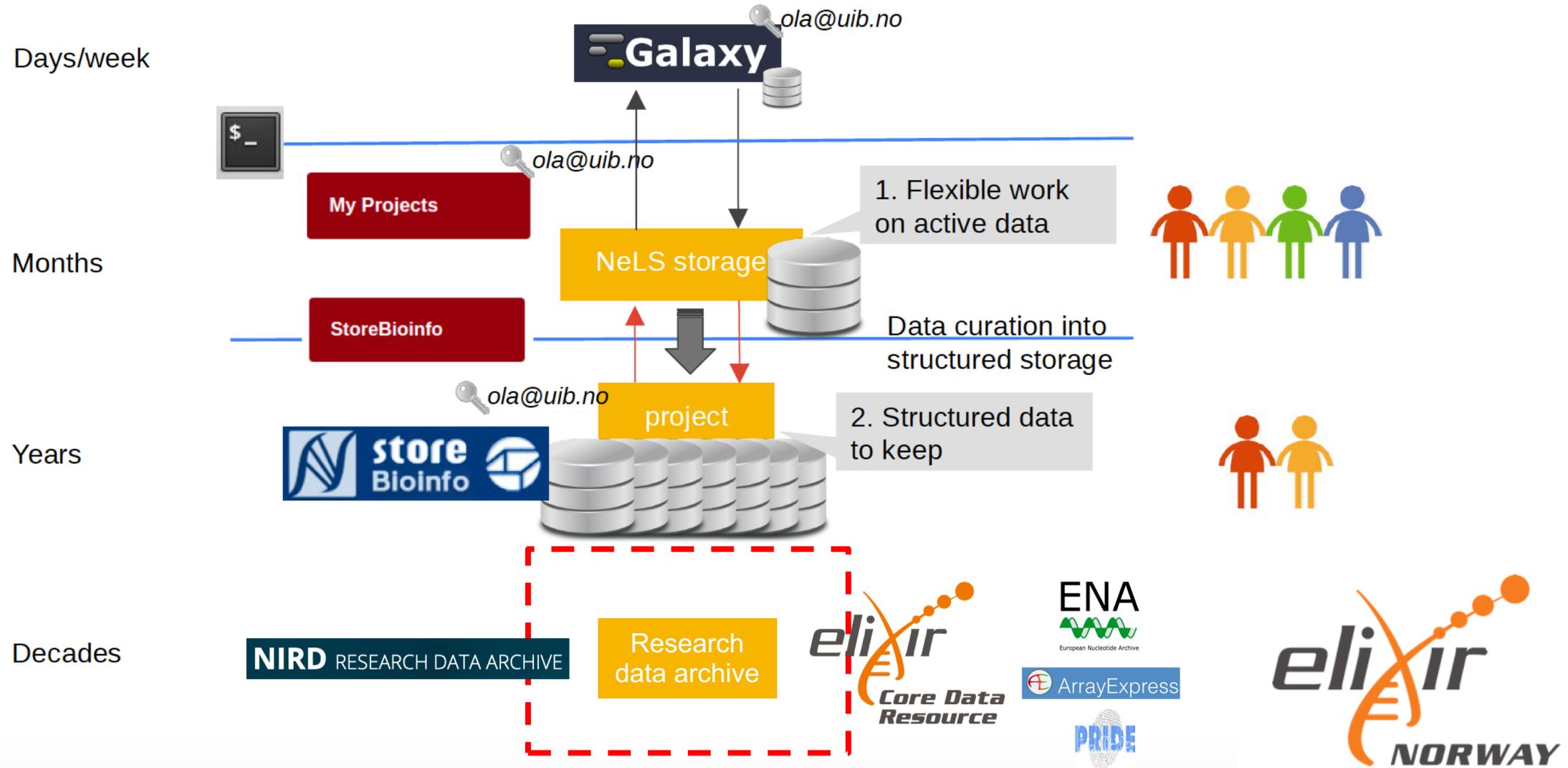
Norwegian Univ. of Life Sciences

lifeportal

NeLS NeLS Portal



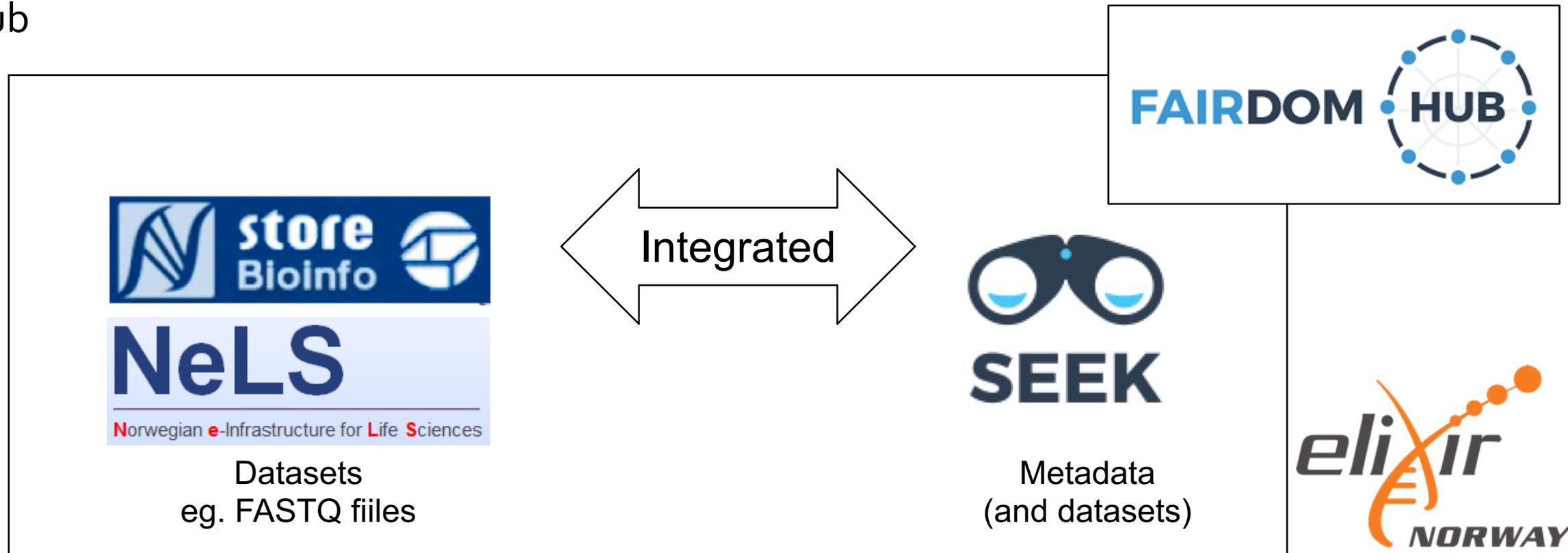
# Data archiving in data repositories



# 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

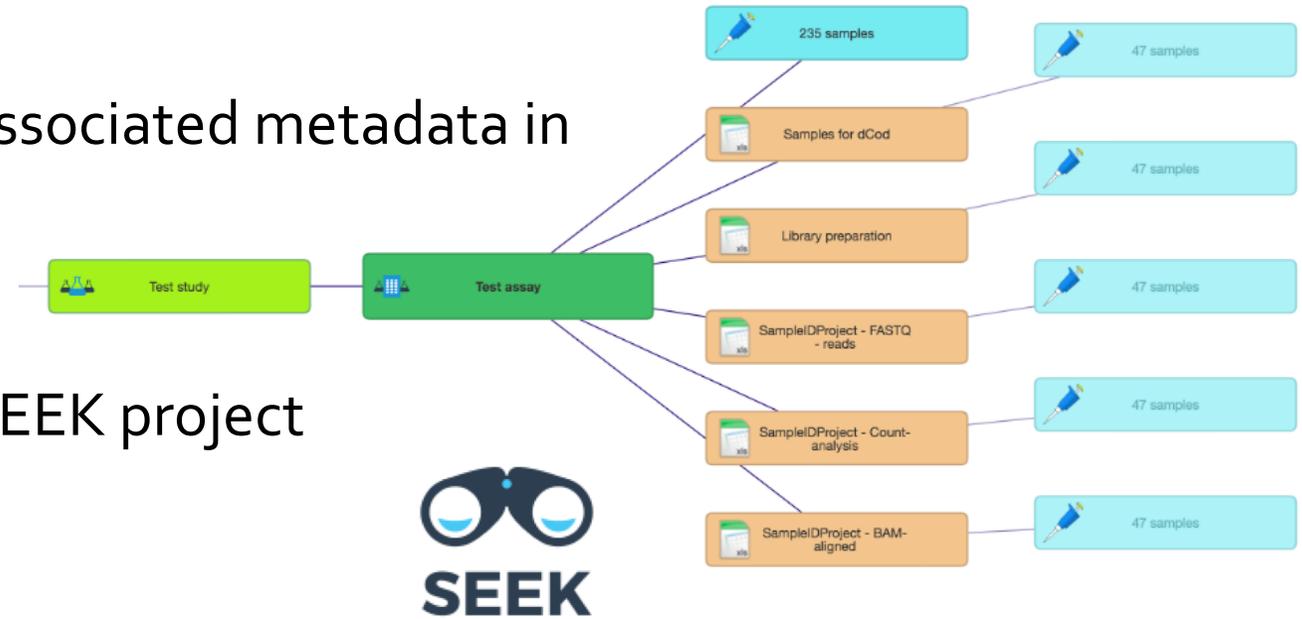
Norwegian users can link datasets stored in NeLS to a SEEK project using FAIRdom hub



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

Typically, organise the project with associated metadata in SEEK

Link datasets stored in NeLS to the SEEK project



StoreBioinfo History

You are at : /StoreBioinfo/seek\_pilot1/SampleIDProject/reads

Name	Size
10_S18_R1_001.fastq.gz	1.3 MB
10_S18_R2_001.fastq.gz	1.3 MB
11_S21_R1_001.fastq.gz	1.3 MB
11_S21_R2_001.fastq.gz	1.3 MB
12_S3_R1_001.fastq.gz	1.3 MB
12_S3_R2_001.fastq.gz	1.3 MB
13_S7_R1_001.fastq.gz	1.3 MB
13_S7_R2_001.fastq.gz	1.3 MB
14_S11_R1_001.fastq.gz	1.3 MB
14_S11_R2_001.fastq.gz	1.3 MB

(1 of 10) Rows per page: 10

Upload Metadata Download Metadata Remove Metadata

