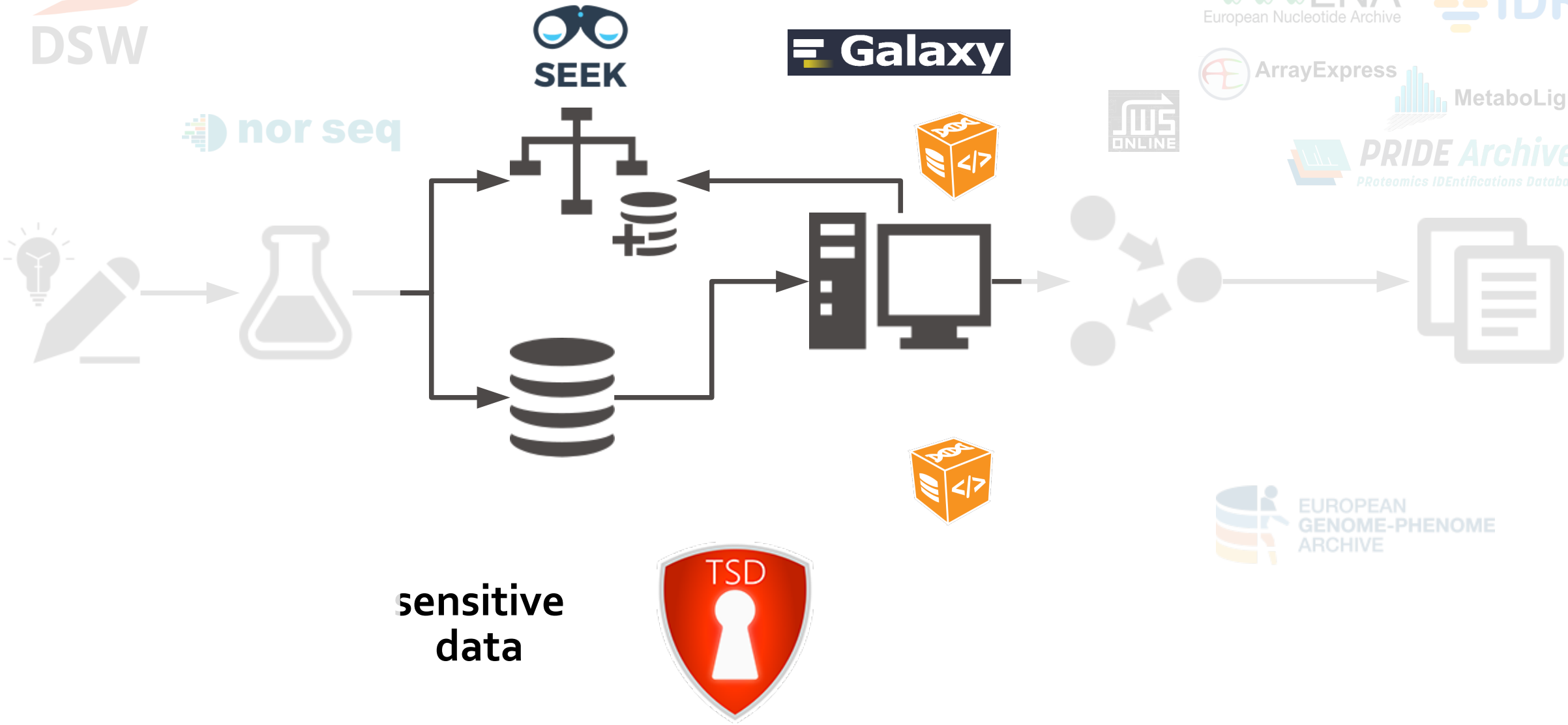
Data Storage Platforms

Korbinian Bösl
Data manager ELIXIR Norway
11 March 2020

NeLS

Norwegian e-Infrastructure for Life Sciences



NeLS

nor seq

Feide ola@uib.no

My Projects

NeLS

Norwegian e-Infrastructure for Life Sciences

Data curation

Tekle, et al. (2018) F1000Res

github.com/elixir-no-nels/nels-core

https://nels.bioinfo.no/

NeLS

nor seq

Feide ola@uib.no

My Projects

NeLS

Norwegian e-Infrastructure for Life Sciences

Flexible work on active data

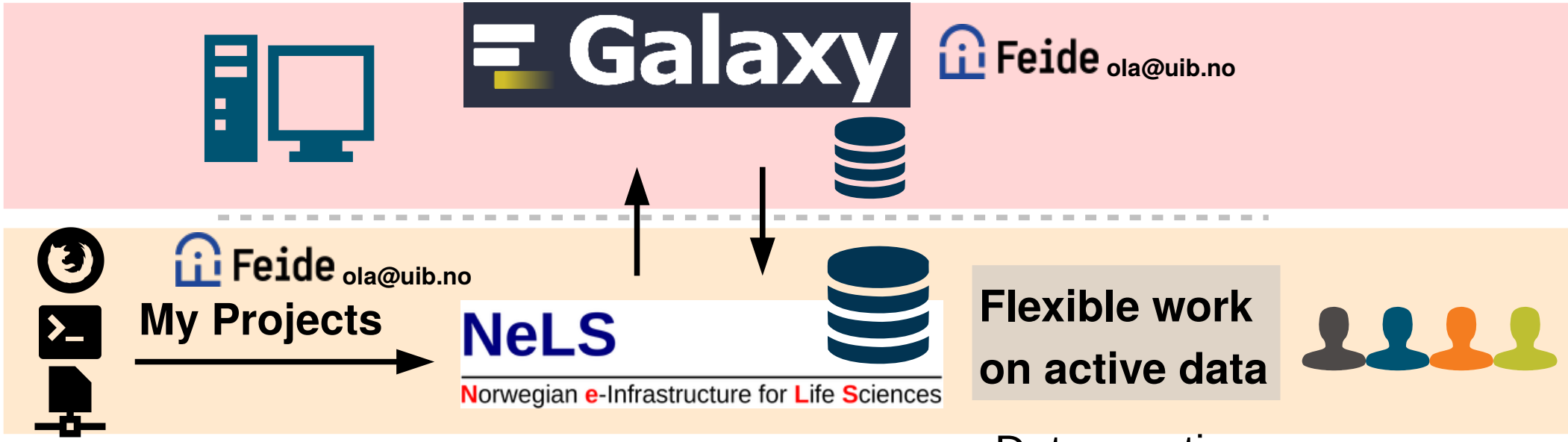
Data curation

Tekle, et al. (2018) F1000Res

github.com/elixir-no-nels/nels-core

https://nels.bioinfo.no/

NeLS

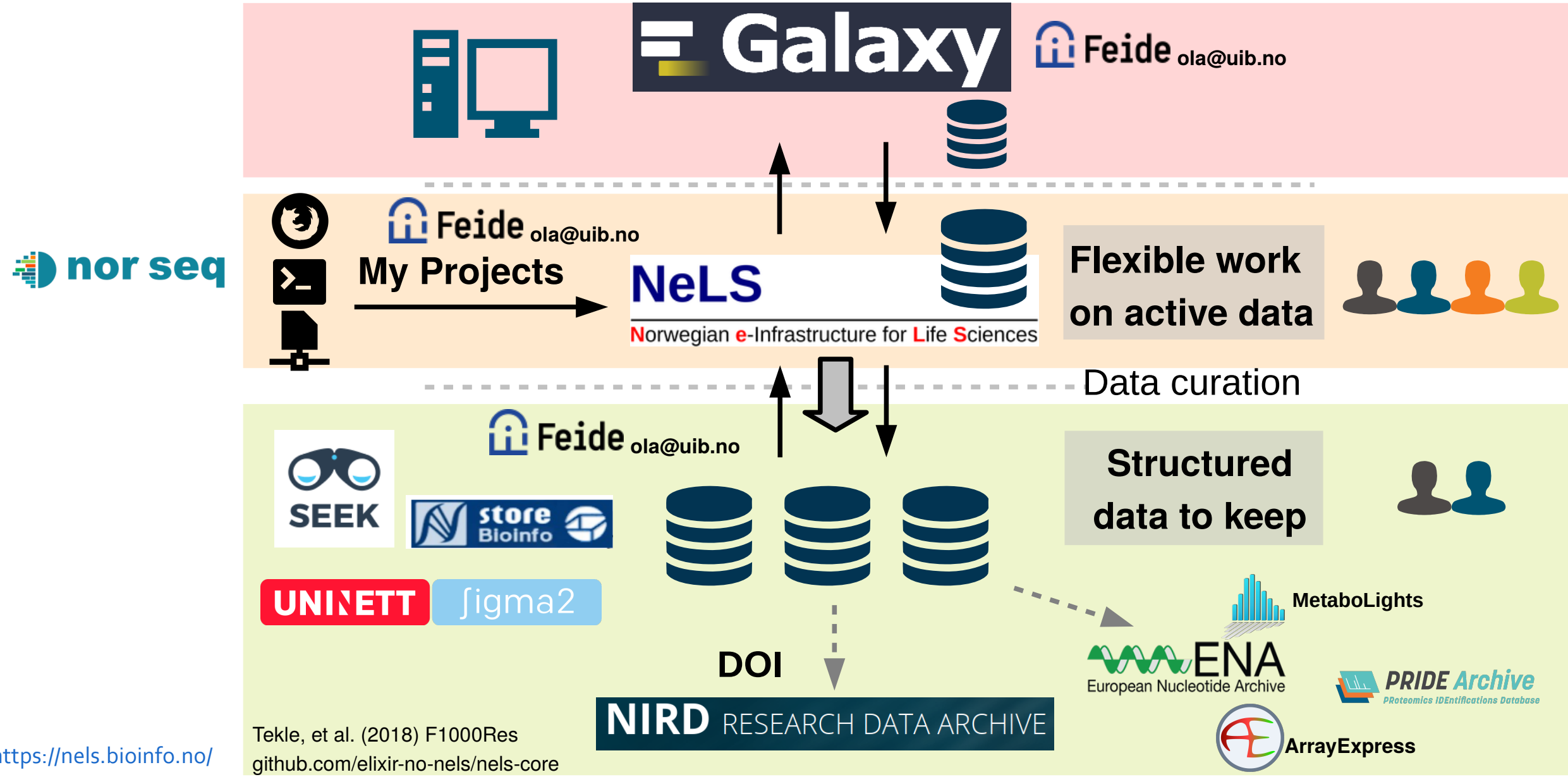


nor seq

Tekle, et al. (2018) F1000Res
github.com/elixir-no-nels/nels-core

https://nels.bioinfo.no/

NeLS



Tekle, et al. (2018) F1000Res
github.com/elixir-no-nels/nels-core

https://nels.bioinfo.no/

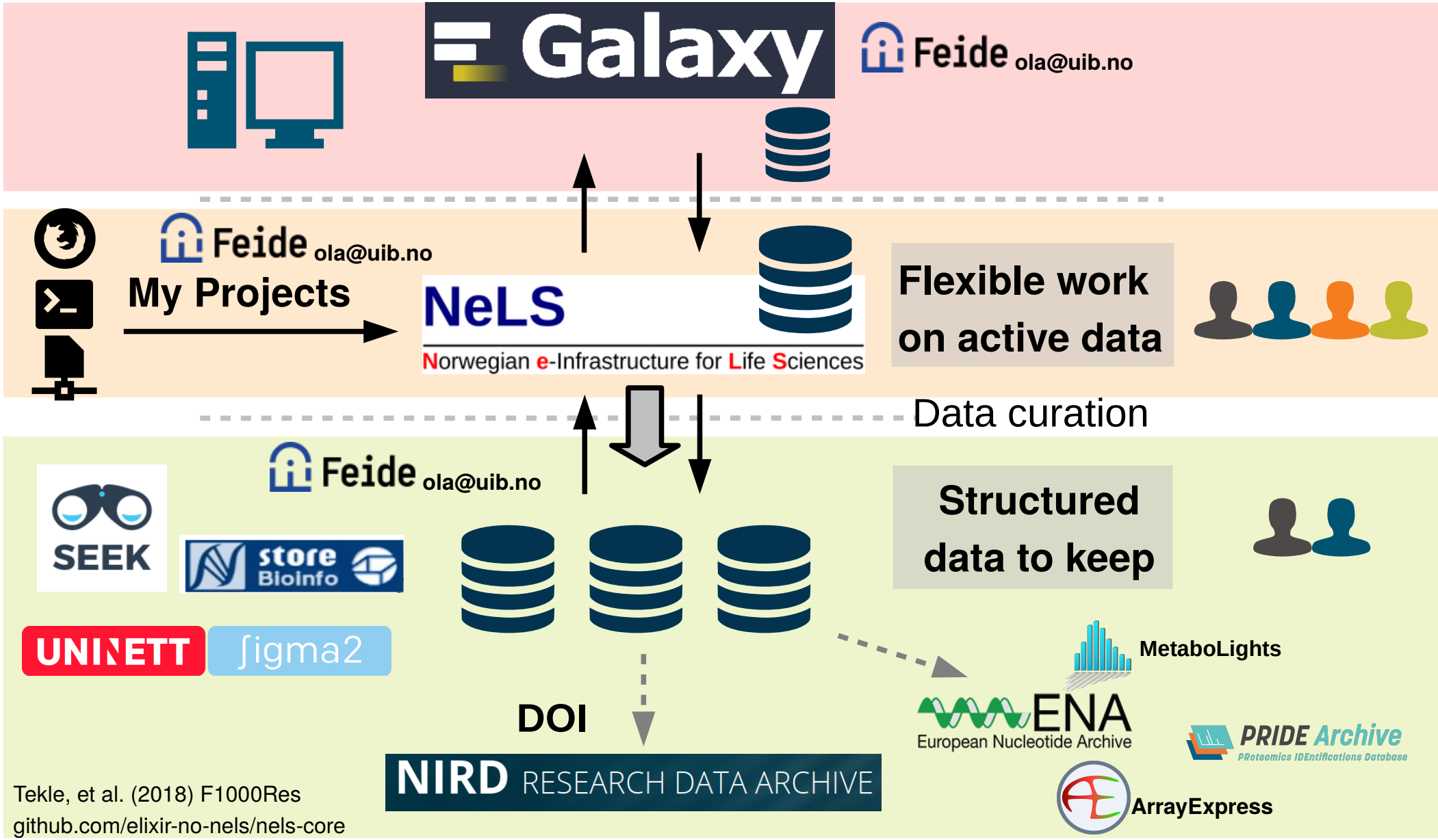
NeLS

Days - weeks

Months

Years

Decades



Tekle, et al. (2018) F1000Res

github.com/elixir-no-nels/nels-core

https://nels.bioinfo.no/

NeLS - Accessibility

small datasets

Feide

ola@uib.no

NeLS - Accessibility

small datasets

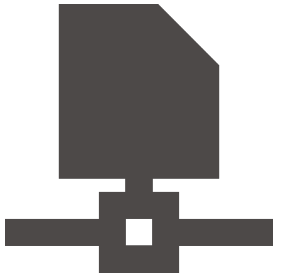
Feide

ola@uib.no

larger datasets

ssh-key

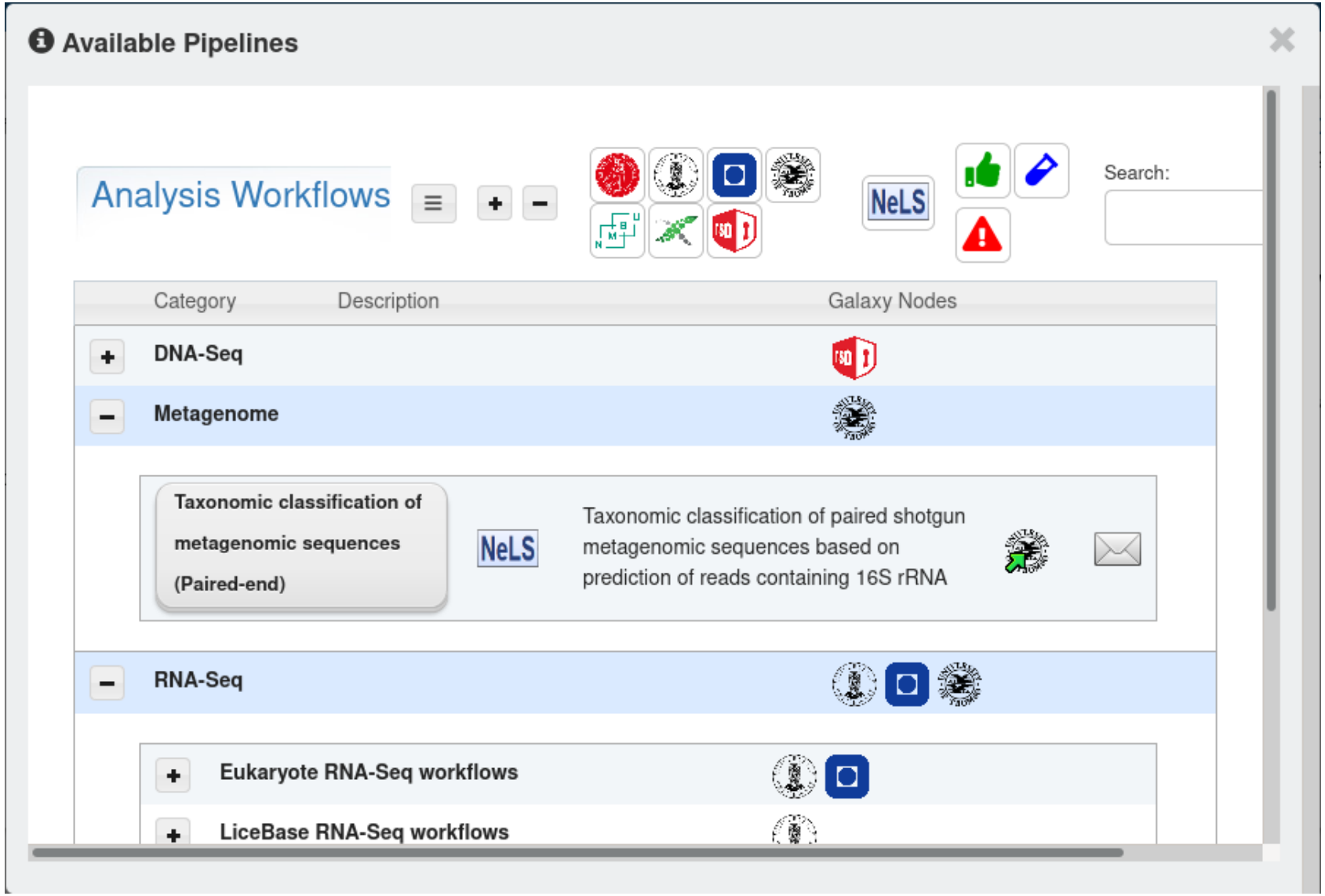
git



sFTP

https://nels.bioinfo.no/

Data Analysis in NeLS



https://nels.bioinfo.no/

NeLS - Getting Access

NeLS

Norwegian e-Infrastructure for Life Sciences

Typically > 1TB data

Free for regular sized projects

Allocation through Helpdesk: contact@bioinfo.no

Direct upload through Norseq facilities

https://nels.bioinfo.no/pages/user-terms.xhtml

SEEK

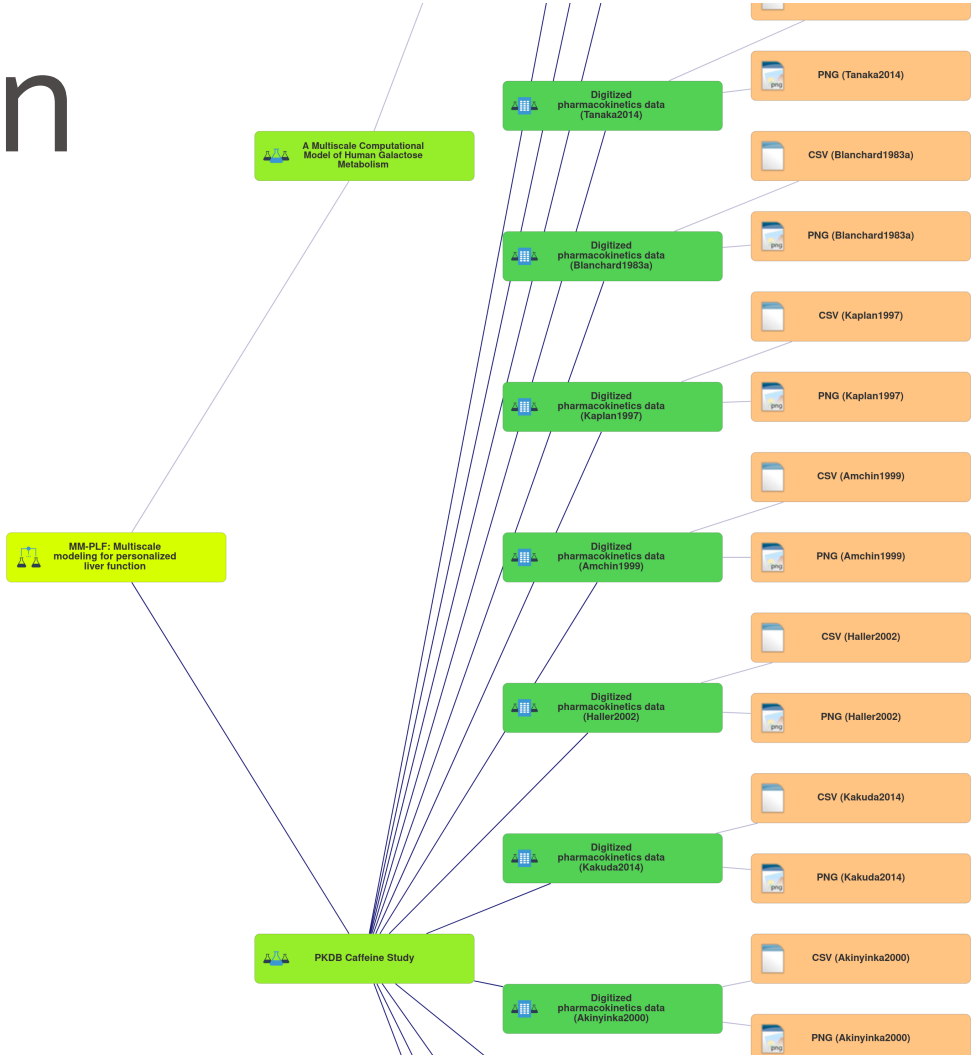
https://seek4science.org/
https://fairdomhub.org/

SEEK

Investigation

Study

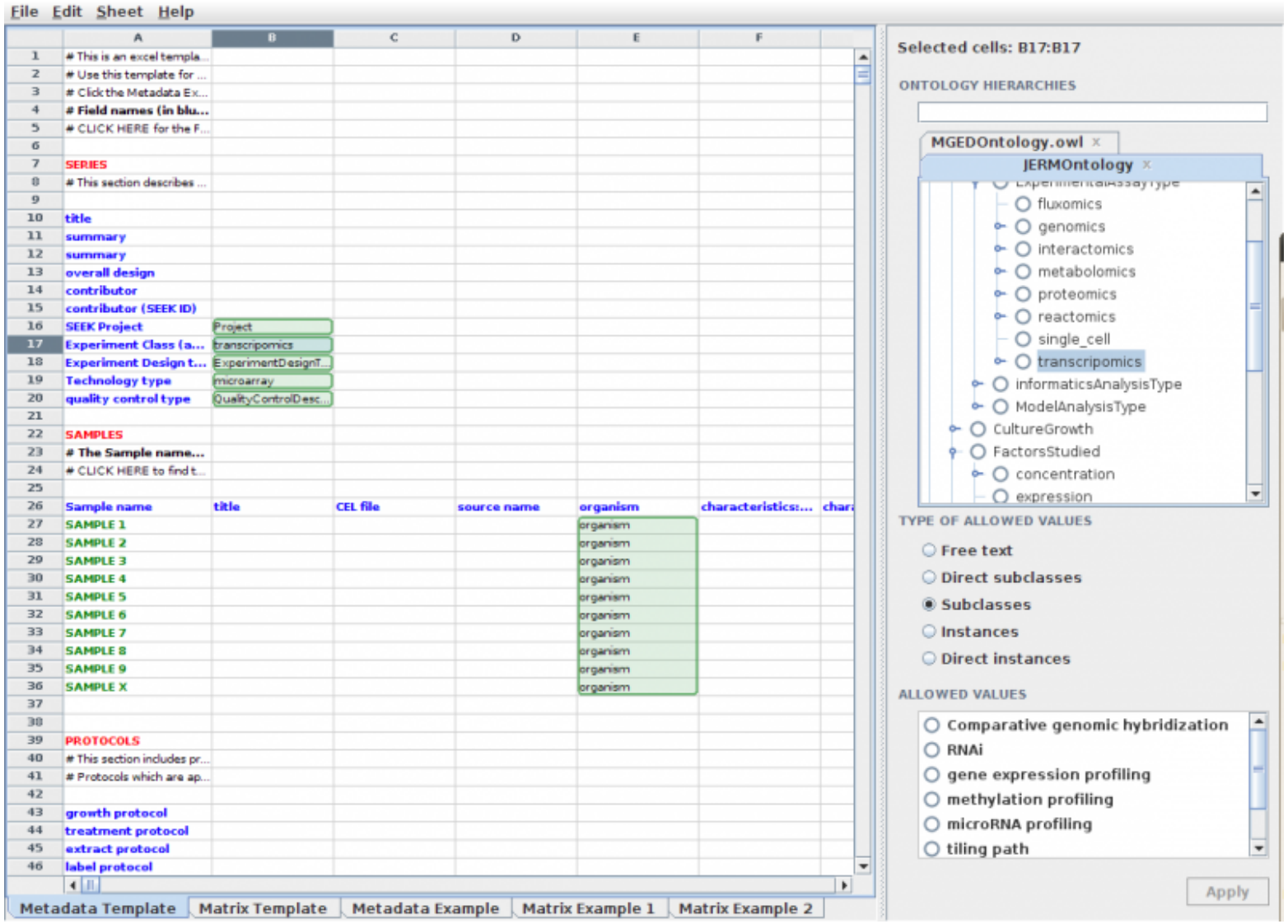
Assay



https://seek4science.org/
https://fairdomhub.org/

SEEK

RightField



https://seek4science.org/
https://rightfield.org.uk/

NeLS
Norwegian e-Infrastructure for Life Sciences

↔

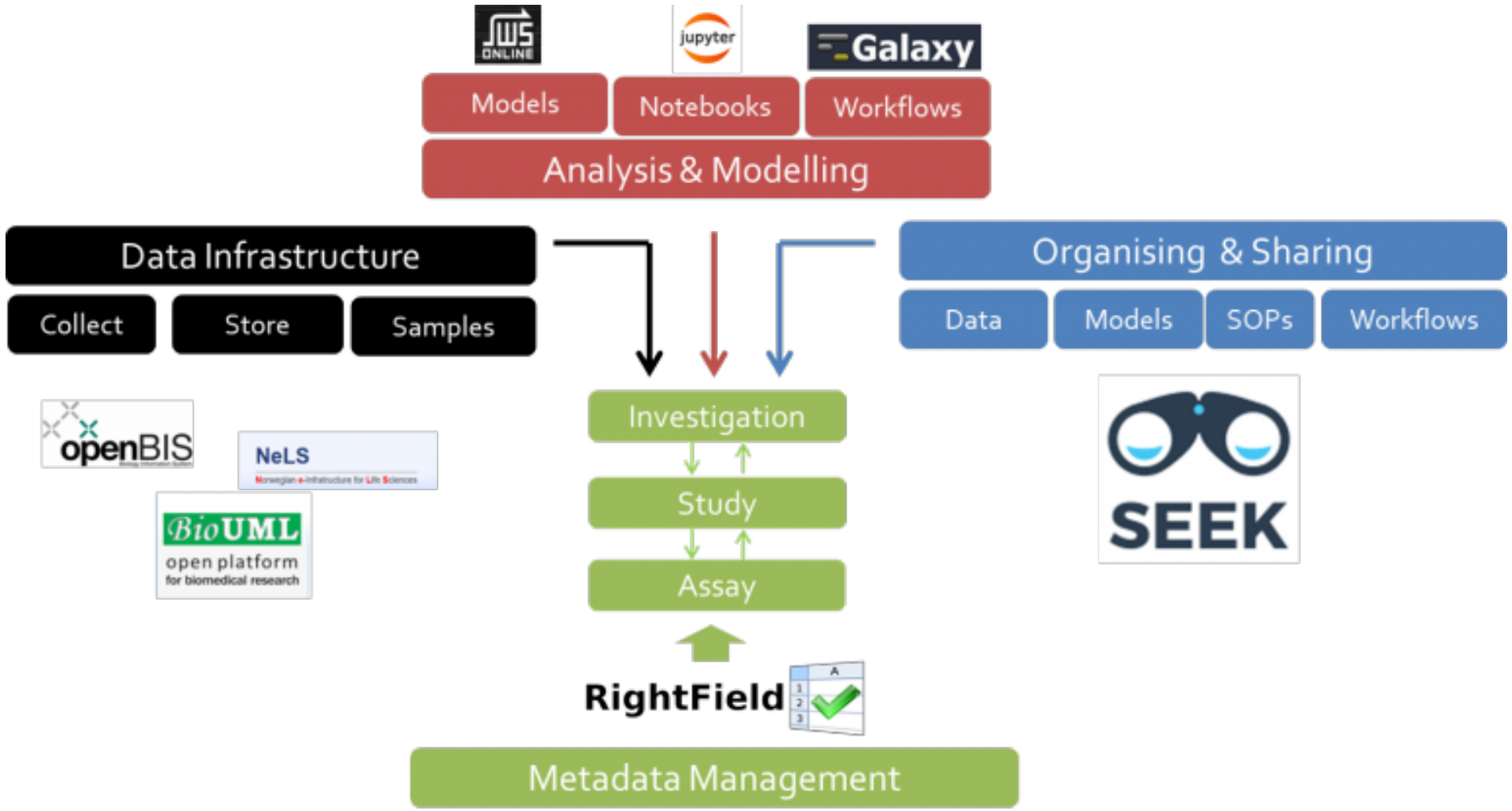
SEEK

Integration

Reference data in NeLS in SEEK

Access data in NeLS through SEEK

Save Metadata to SEEK through NeLS


JWS ONLINE
jupyter
Galaxy
Models
Notebooks
Workflows
Analysis & Modelling
Data Infrastructure
Collect
Store
Samples
openBIS
NeLS
Norwegian e-Infrastructure for Life Sciences
BioUML
open platform
for biomedical research
Investigation
Study
Assay
RightField
Metadata Management
Organising & Sharing
Data
Models
SOPs
Workflows
SEEK

National HPC infrastructure

UNINETT Sigma2

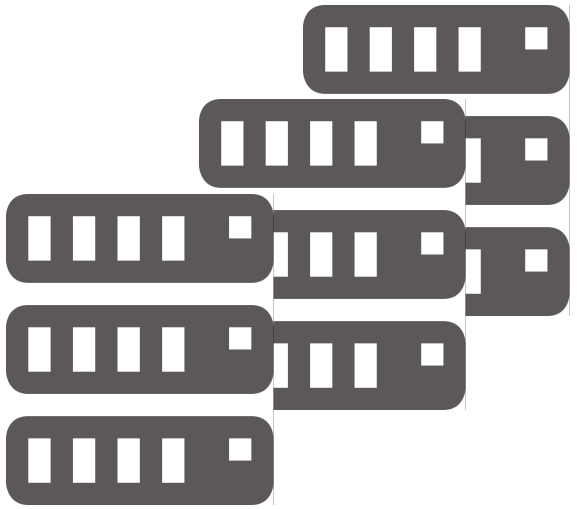
- Advanced Storage Needs
- Advanced Computing Resources
- Free Analysis Toolkit (Rstudio, Jupyternotebook)

https://www.sigma2.no/

National HPC infrastructure

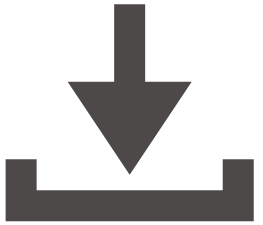
UNINETT Sigma2

NeLS
Norwegian e-Infrastructure for Life Sciences



jupyter RStudio

NIRD toolkit



NIRD archive

Saga

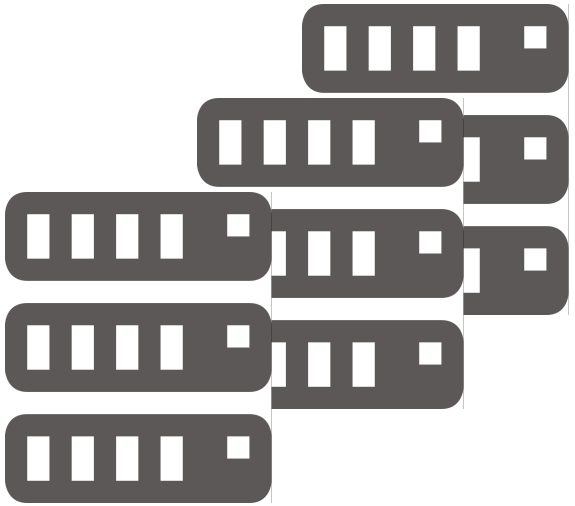
NIRD

https://www.sigma2.no/

HPC infrastructure

UNINETT Sigma2 Saga

244 nodes – 9824 cores – 32 GPUs



Total 75TB RAM – 186-3066GB/node

89TB + 60 TB NVMe+SSD

1PB file system

https://documentation.sigma2.no/quick/saga.html

Storage infrastructure - NIRD

UNINETT Sigma2

Geo replicated

Nightly snapshots

12 PB (→ 36 PB next 4 yrs)

https://documentation.sigma2.no/storage/nird.html

Sigma2 - Getting Access

UNINETT Sigma2

Bi-annual calls (January & July) – Reapplication (!)

User contribution model (projects > 15 mill. NOK)

Application for small grants continually

https://www.sigma2.no/how-apply-user-account
https://www.sigma2.no/data-storage#get-access

TSD – National sensitive data service

Maximum classification

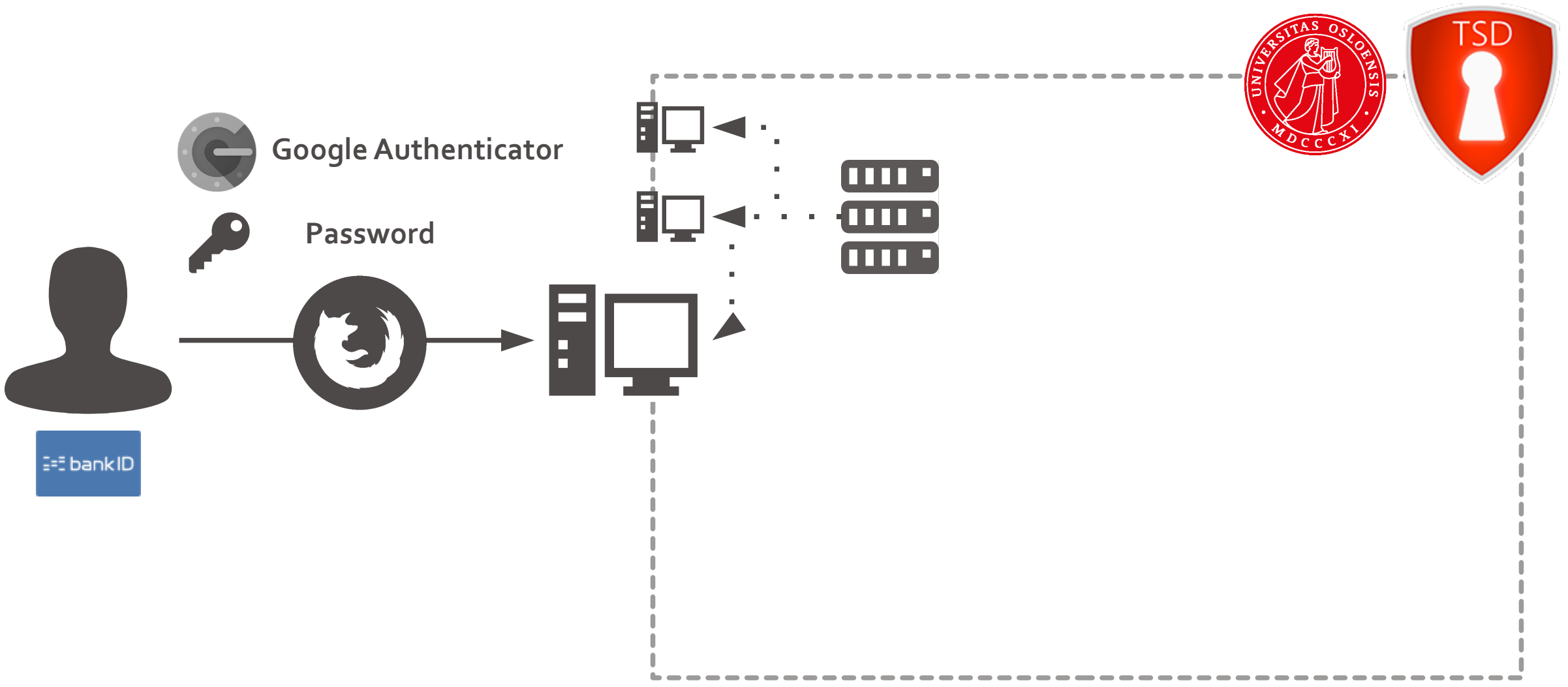
2-factor authentication

Attached HPC resources

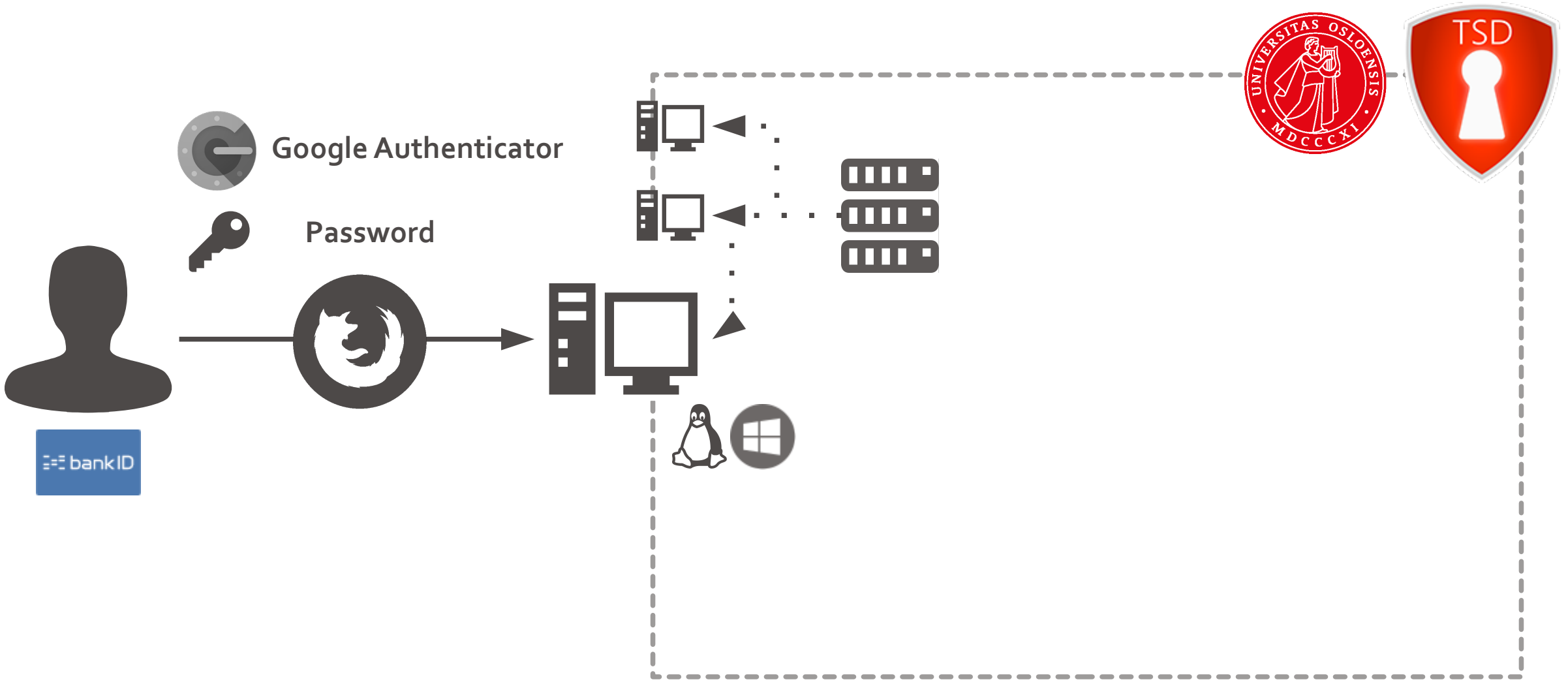
https://www.uio.no/english/services/it/research/sensitive-data/index.html

TSD – National sensitive data service

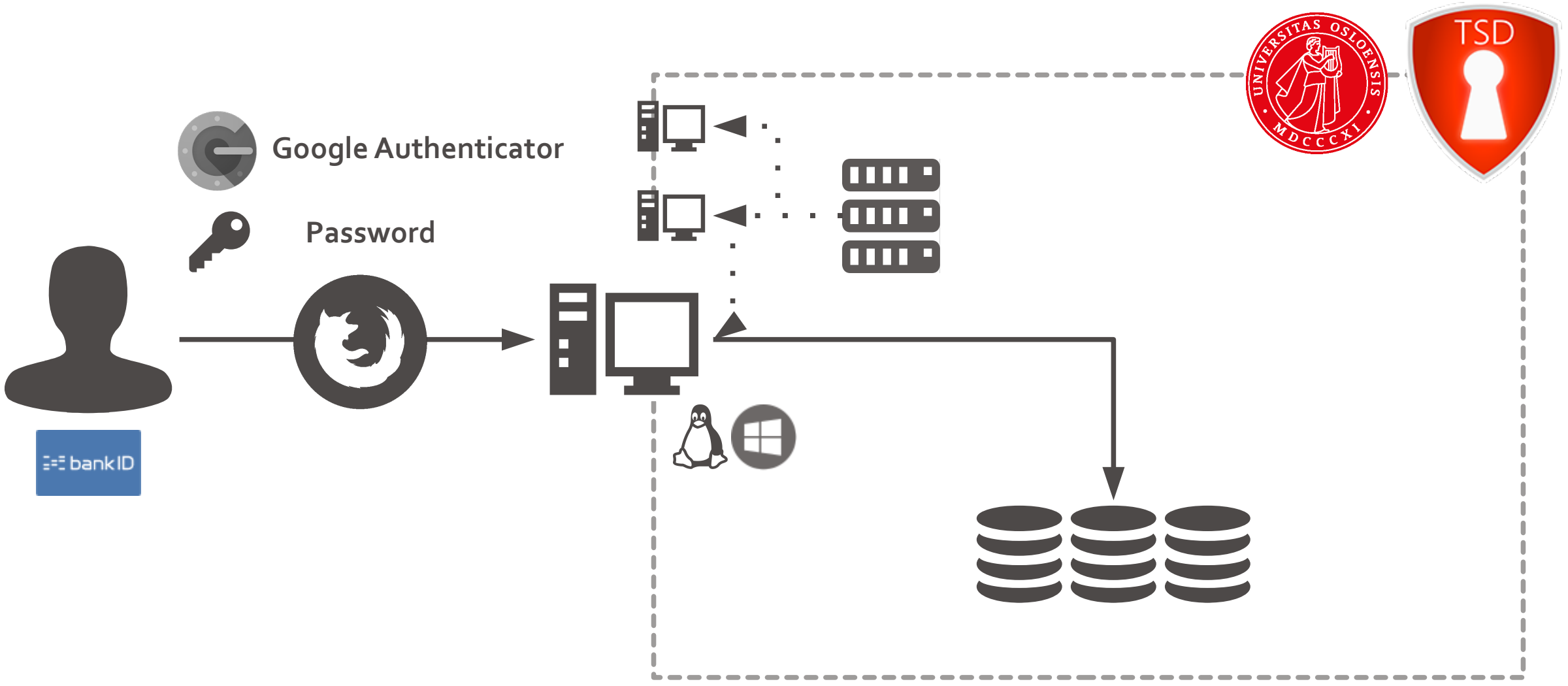
TSD – National sensitive data service



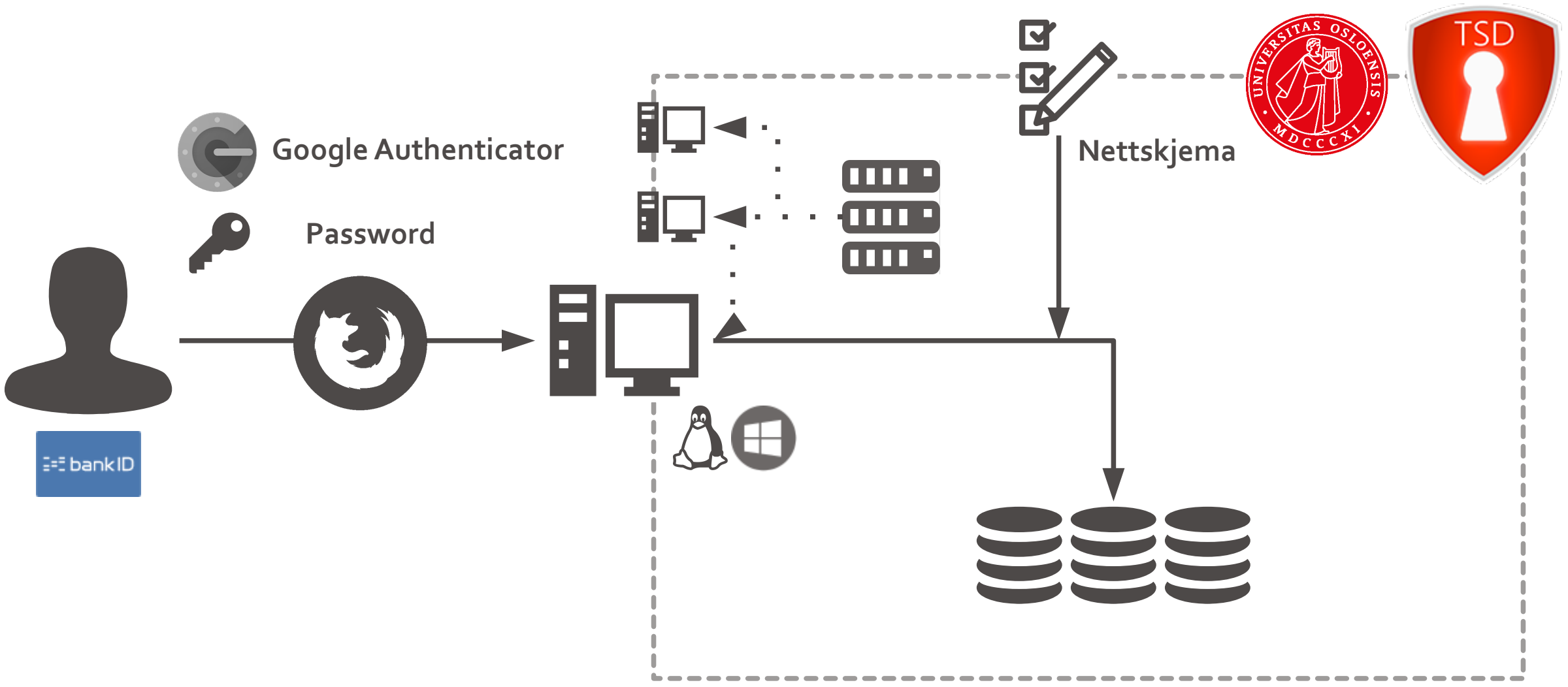
TSD – National sensitive data service



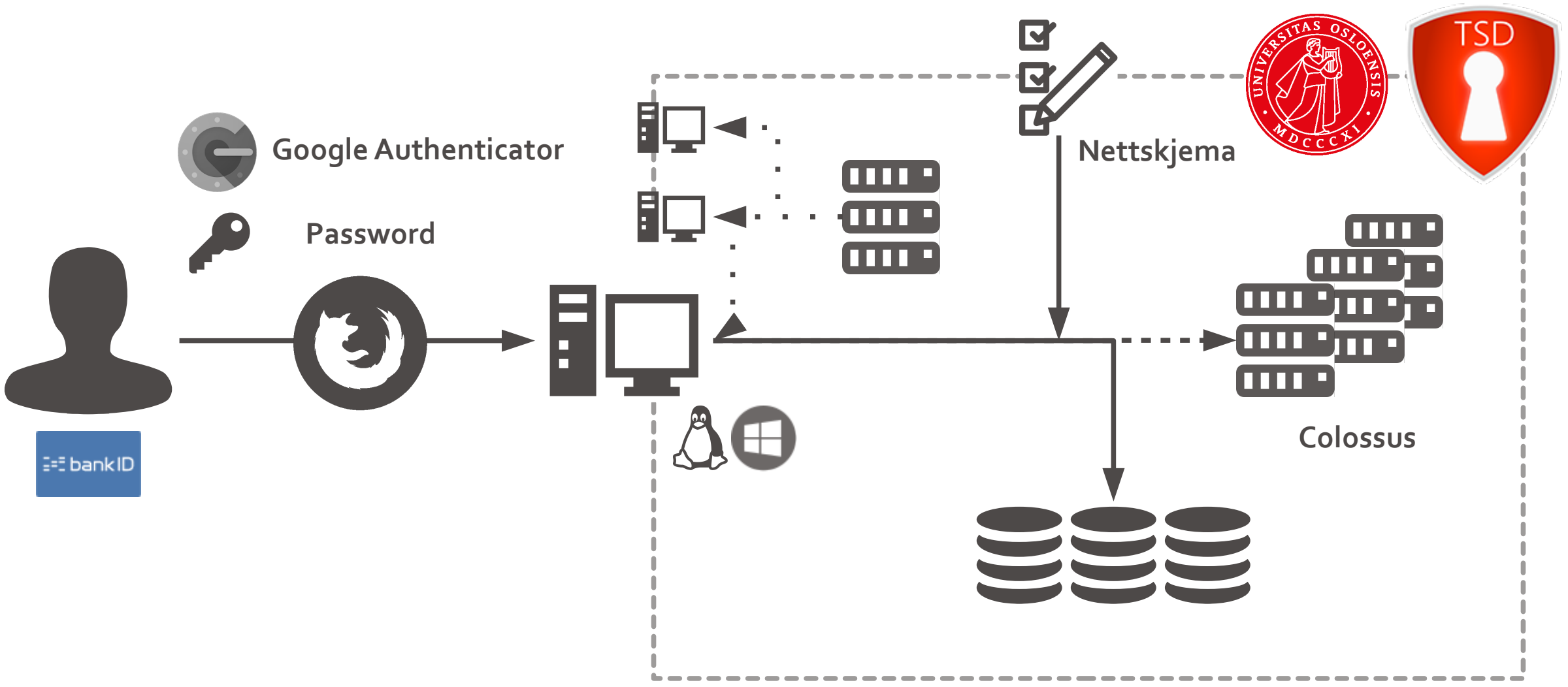
TSD – National sensitive data service



TSD – National sensitive data service



TSD – National sensitive data service



TSD – Getting Access

Subsidized access through Sigma2 calls

User contribution model (non UiO: min 15kNOK/a)

Free storage quotas through NeLS

https://www.uio.no/english/services/it/research/sensitive-data/access/prices/
https://nels.bioinfo.no/