



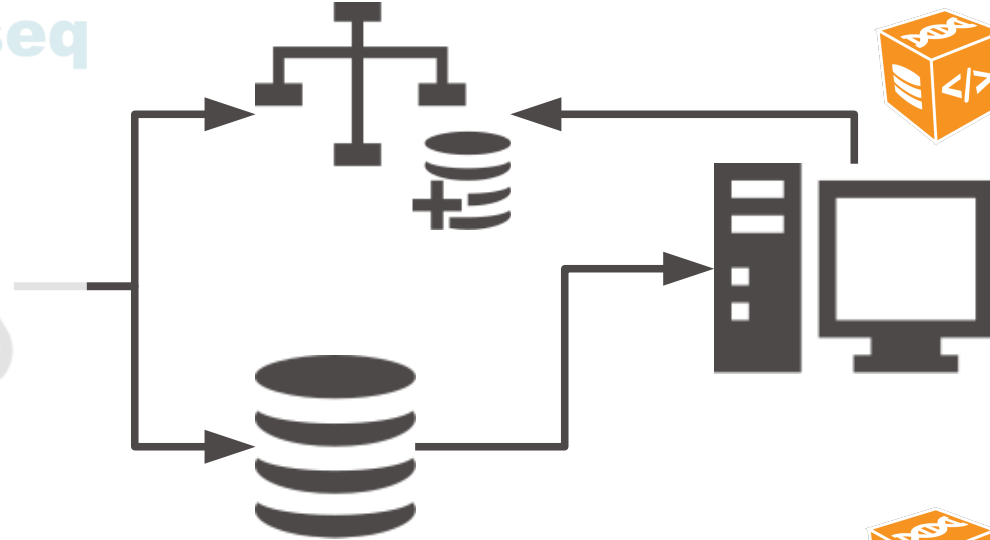
# Data Storage Platforms



Korbinian Bösl  
Data manager ELIXIR Norway  
4 March 2020

# NeLS

Norwegian e-Infrastructure for Life Sciences



sensitive  
data



# NeLS



Feide ola@uib.no



My Projects



NeLS



Norwegian e-Infrastructure for Life Sciences

Data curation

# NeLS



 **Feide** ola@uib.no  
**My Projects**

**NeLS**

Norwegian e-Infrastructure for Life Sciences

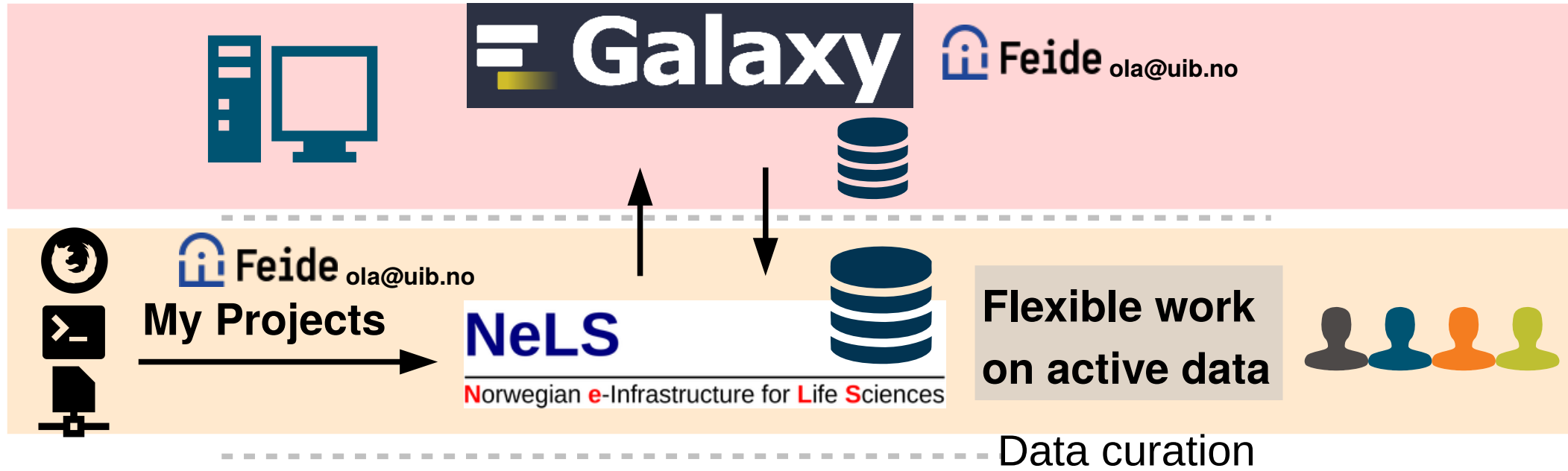


**Flexible work  
on active data**

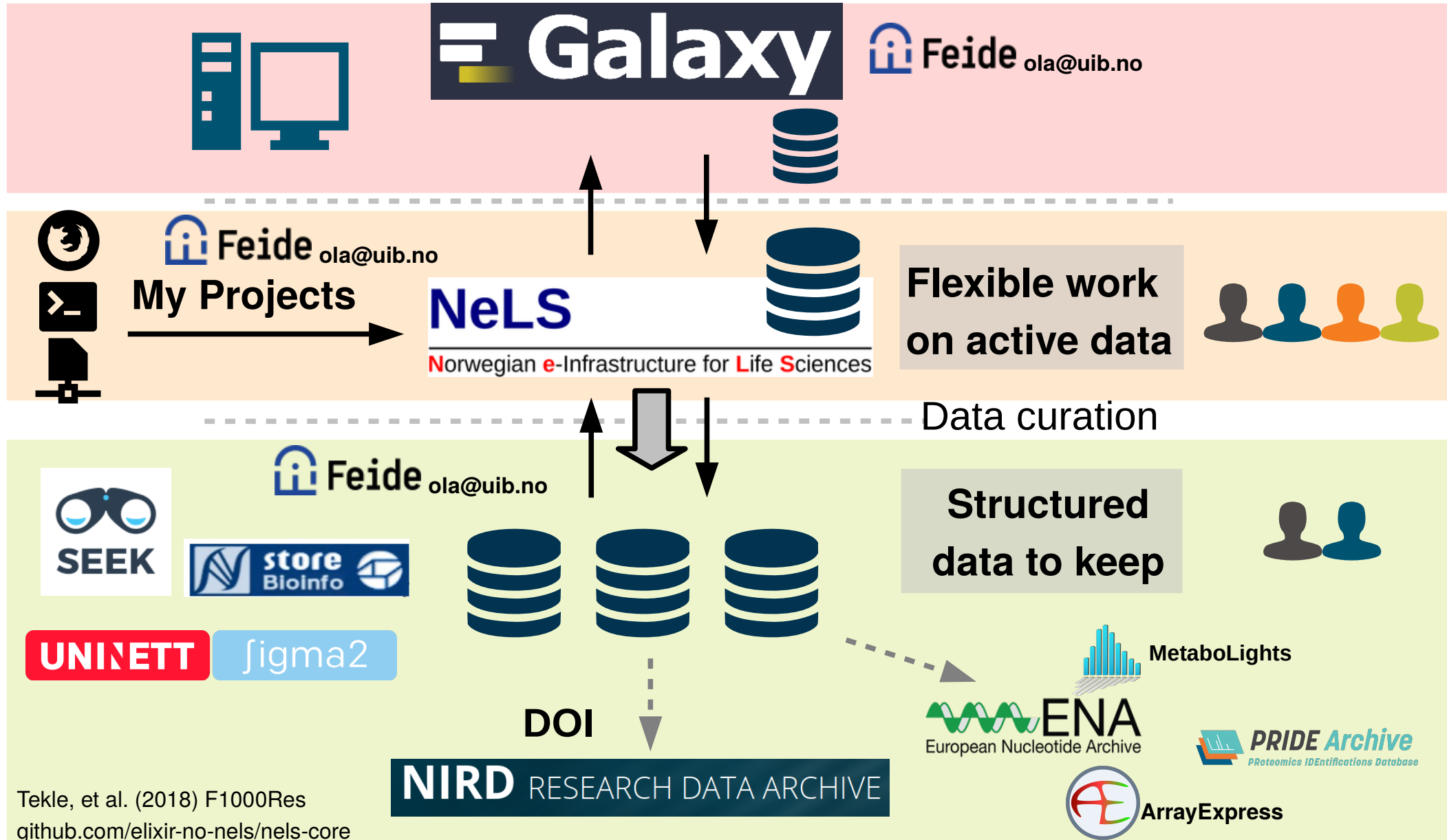


Data curation

# NeLS



# NeLS

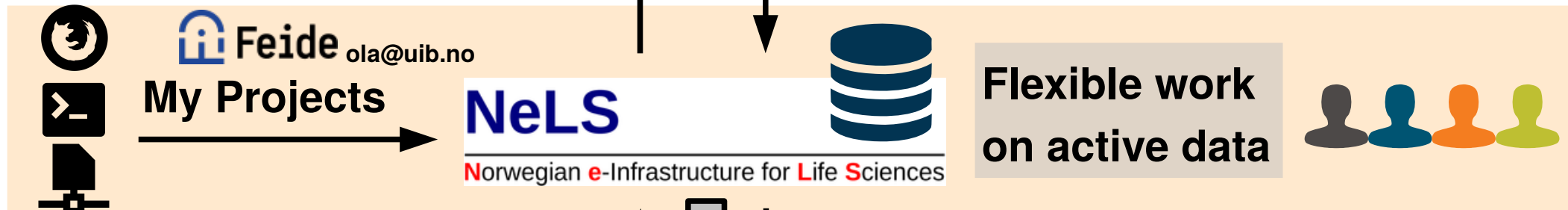


# NeLS

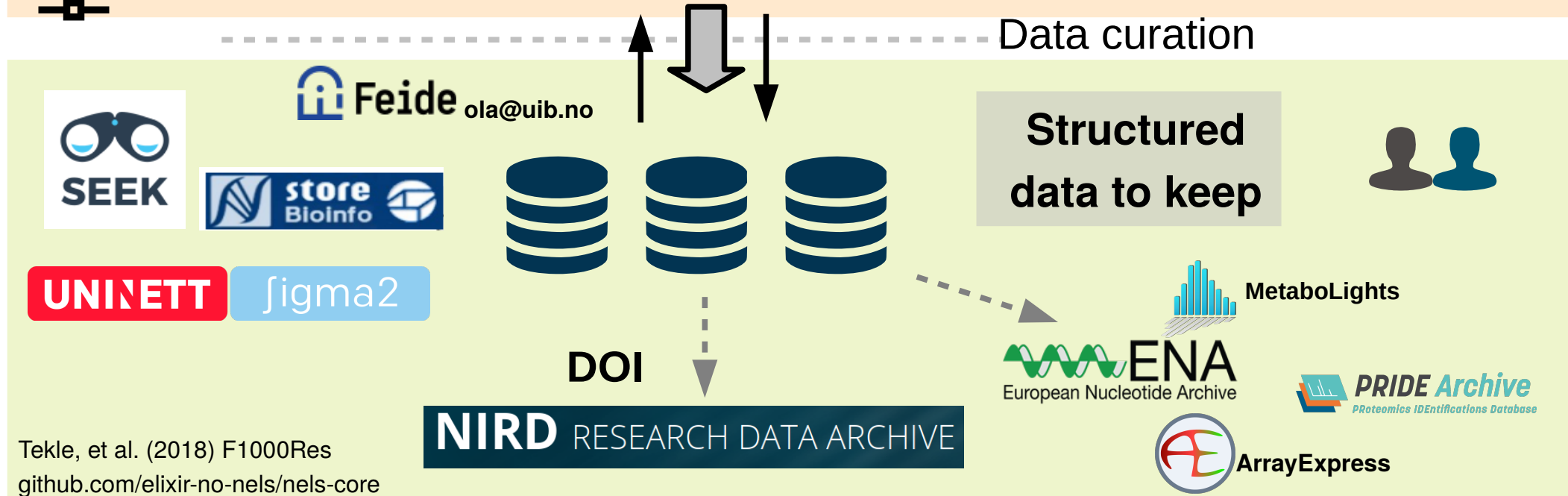
Days - weeks



 **nor seq**  
Months



Years



Decades

<https://nels.bioinfo.no/>

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

# NeLS - Accessibility

small datasets



ola@uib.no



# NeLS - Accessibility

small datasets



ola@uib.no

larger datasets



ssh-key






git


























sFTP

# Data Analysis in NeLS

**Available Pipelines**

Analysis Workflows   

       Search:

Category	Description	Galaxy Nodes
 DNA-Seq		
 Metagenome		
	<div><div>Taxonomic classification of metagenomic sequences (Paired-end)</div><div></div></div>	<div><div>Taxonomic classification of paired shotgun metagenomic sequences based on prediction of reads containing 16S rRNA</div><div> </div></div>
 RNA-Seq		  
	<div><div> Eukaryote RNA-Seq workflows</div><div> </div></div>	
	<div><div> LiceBase RNA-Seq workflows</div><div></div></div>	

# NeLS - Getting Access

## NeLS

Norwegian e-Infrastructure for Life Sciences

Typically > 1TB data

Free for regular sized projects

Allocation through Helpdesk: [contact@bioinfo.no](mailto:contact@bioinfo.no)

Direct upload through Nourseq facilities

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



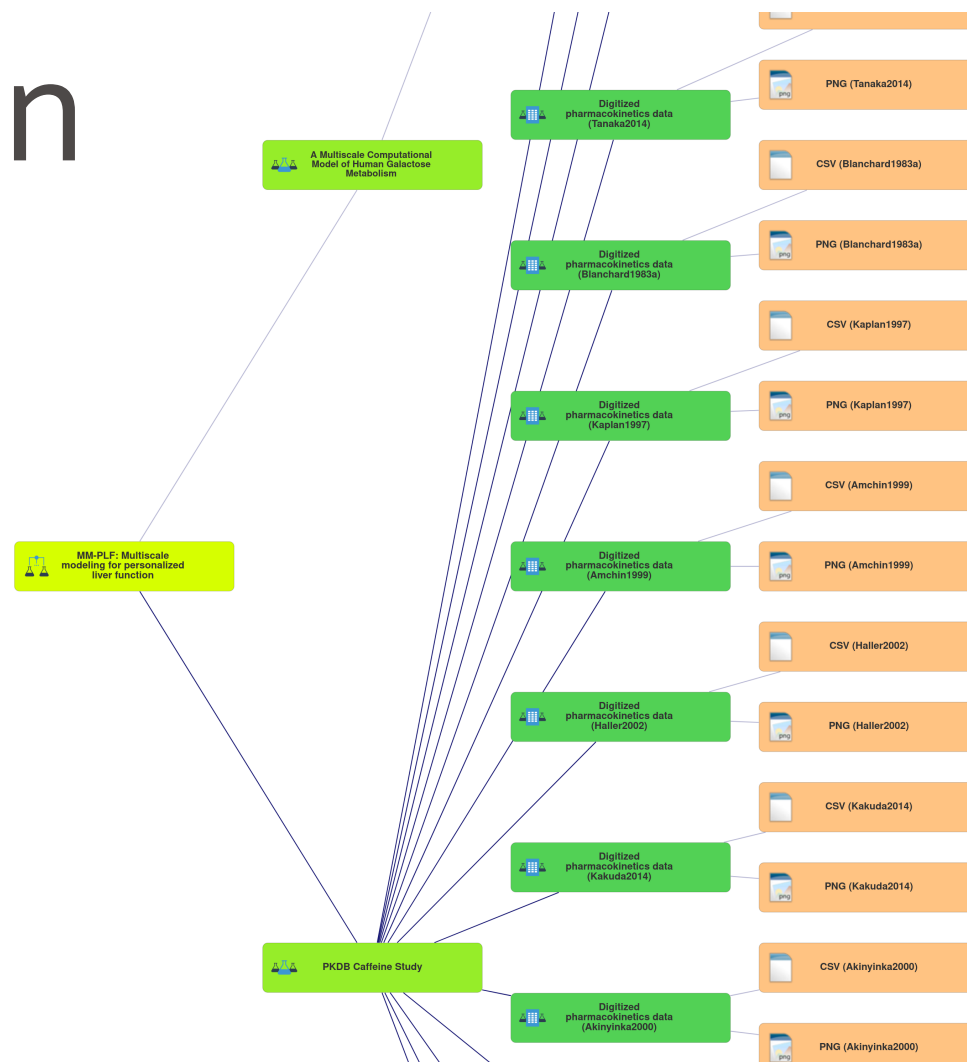
# SEEK





# SEEK

# Investigation Study Assay



## RightField



File Edit Sheet Help

	A	B	C	D	E	F
1	# This is an excel template...					
2	# Use this template for ...					
3	# Click the Metadata Ex...					
4	# Field names (in blu...					
5	# CLICK HERE for the F...					
6						
7	<b>SERIES</b>					
8	# This section describes ...					
9						
10	<b>title</b>					
11	<b>summary</b>					
12	<b>summary</b>					
13	<b>overall design</b>					
14	<b>contributor</b>					
15	<b>contributor (SEEK ID)</b>					
16	<b>SEEK Project</b>	Project				
17	<b>Experiment Class (a...</b>	transcriptomics				
18	<b>Experiment Design t...</b>	ExperimentDesignT...				
19	<b>Technology type</b>	microarray				
20	<b>quality control type</b>	QualityControlDesc...				
21						
22	<b>SAMPLES</b>					
23	# The Sample name...					
24	# CLICK HERE to find t...					
25						
26	<b>Sample name</b>	<b>title</b>	<b>CEL file</b>	<b>source name</b>	<b>organism</b>	<b>characteristics:...</b>
27	<b>SAMPLE 1</b>				organism	
28	<b>SAMPLE 2</b>				organism	
29	<b>SAMPLE 3</b>				organism	
30	<b>SAMPLE 4</b>				organism	
31	<b>SAMPLE 5</b>				organism	
32	<b>SAMPLE 6</b>				organism	
33	<b>SAMPLE 7</b>				organism	
34	<b>SAMPLE 8</b>				organism	
35	<b>SAMPLE 9</b>				organism	
36	<b>SAMPLE X</b>				organism	
37						
38						
39	<b>PROTOCOLS</b>					
40	# This section includes pr...					
41	# Protocols which are ap...					
42						
43	<b>growth protocol</b>					
44	<b>treatment protocol</b>					
45	<b>extract protocol</b>					
46	<b>label protocol</b>					

Selected cells: B17:B17

ONTOLOGY HIERARCHIES

MGEDOntology.owl x

JERMOntology x

- ExperimentClassType
  - fluxomics
  - genomics
  - interactomics
  - metabolomics
  - proteomics
  - reactomics
  - single\_cell
  - transcriptomics
- InformaticsAnalysisType
  - ModelAnalysisType
- CultureGrowth
  - FactorsStudied
    - concentration
    - expression

TYPE OF ALLOWED VALUES

☐ Free text

☐ Direct subclasses

☒ Subclasses

☐ Instances

☐ Direct instances

ALLOWED VALUES

- Comparative genomic hybridization
- RNAi
- gene expression profiling
- methylation profiling
- microRNA profiling
- tiling path

Apply

Metadata Template Matrix Template Metadata Example Matrix Example 1 Matrix Example 2

**NeLS**

Norwegian e-Infrastructure for Life Sciences



# Integration



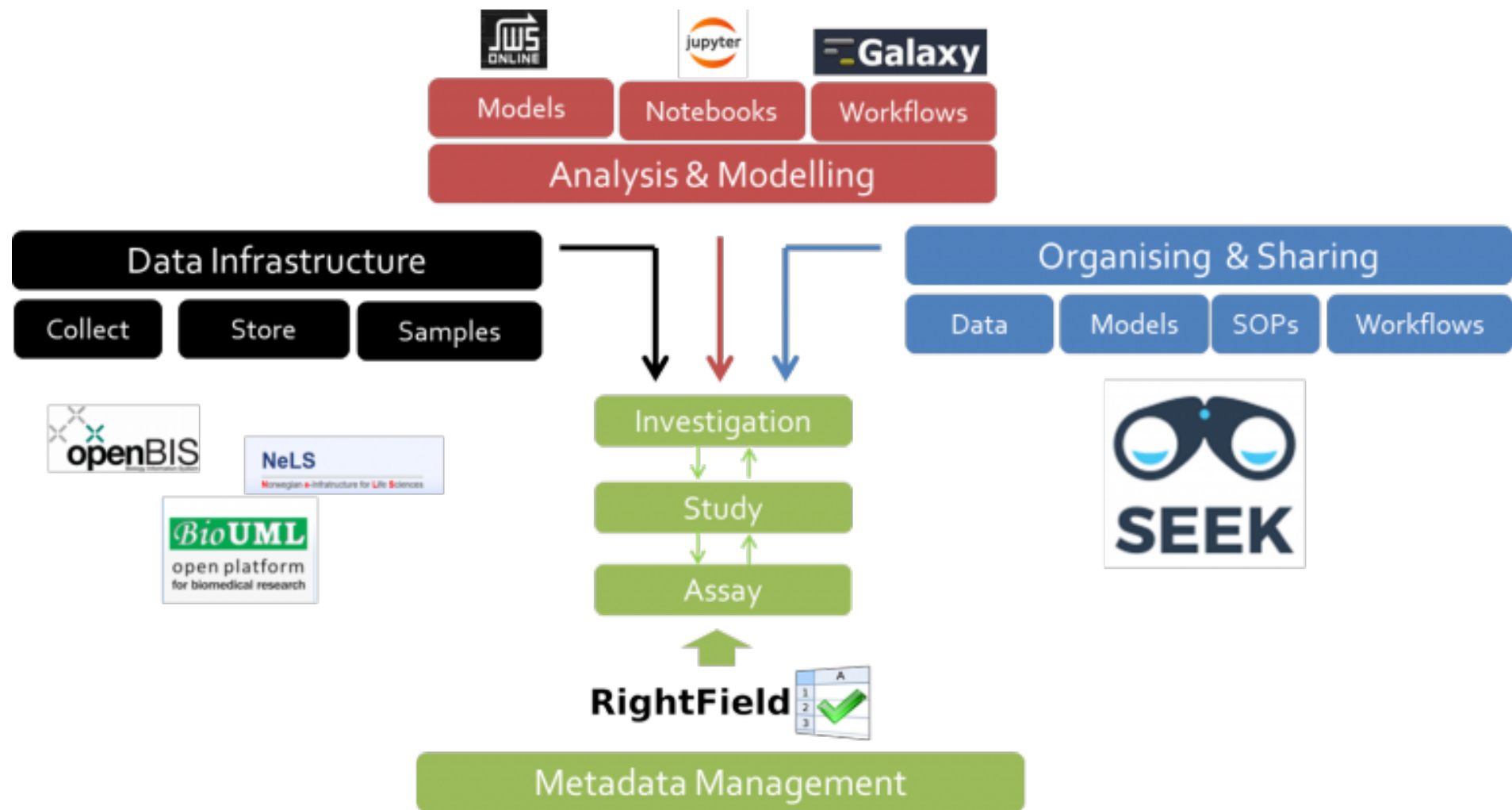
**Reference data in NeLS in SEEK**



**Access data in NeLS through SEEK**



**Save Metadata to SEEK through NeLS**







# BILLY – Tape Backup

↓ Storage on tape – min 4MB/file – 3 copies

🌐 Geo-replication

📺 1 TB 250NOK/yr in 50TB chunks → [bs.uib.no](https://bs.uib.no)

# TSD – National sensitive data service



Maximum classification



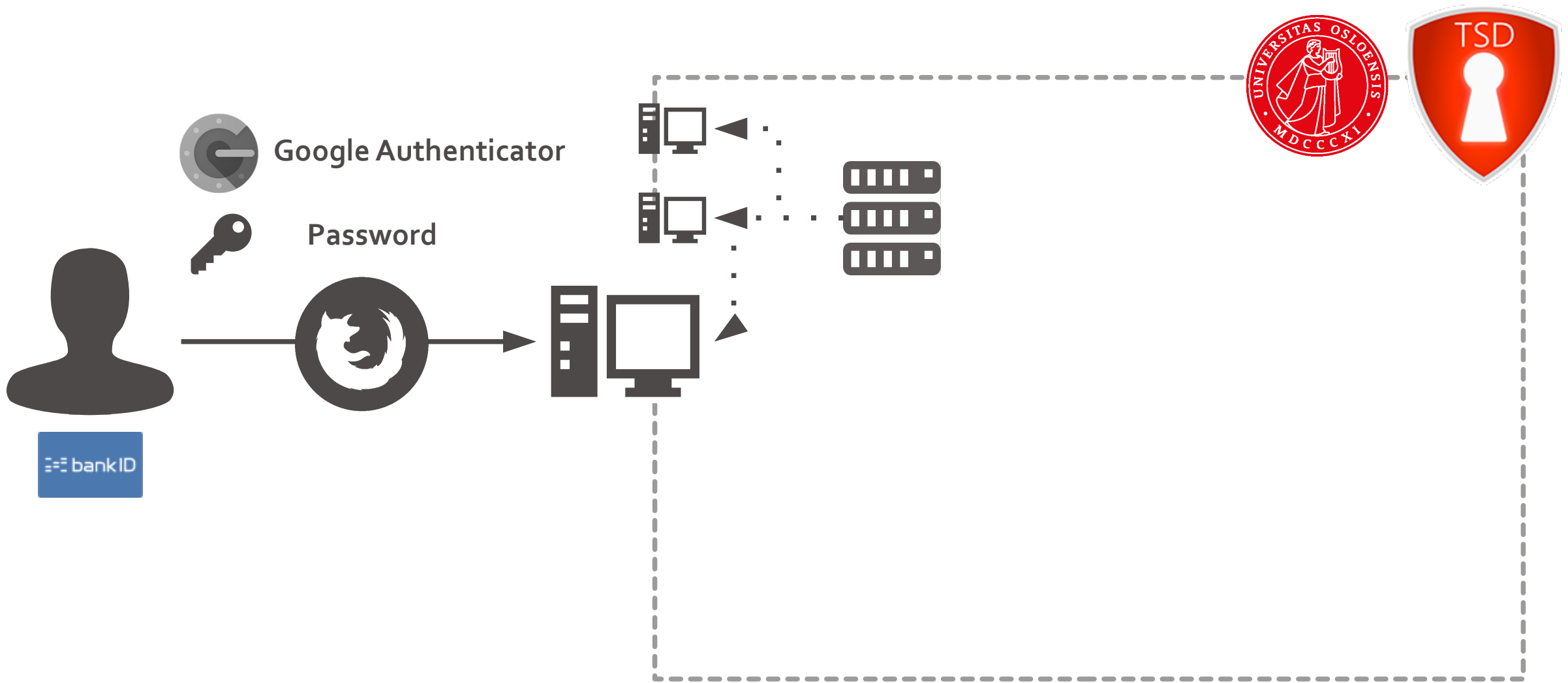
2-factor authentication

Attached HPC resources

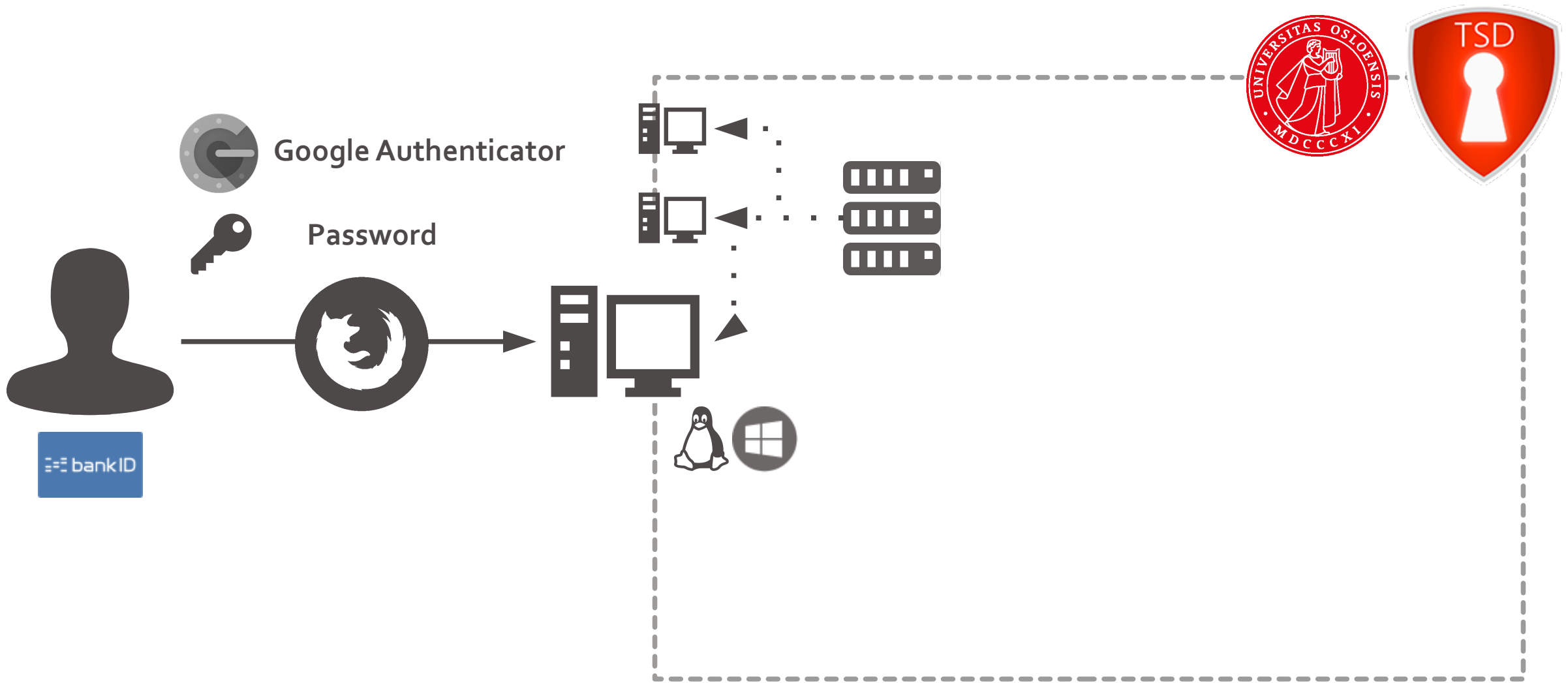
# TSD – National sensitive data service



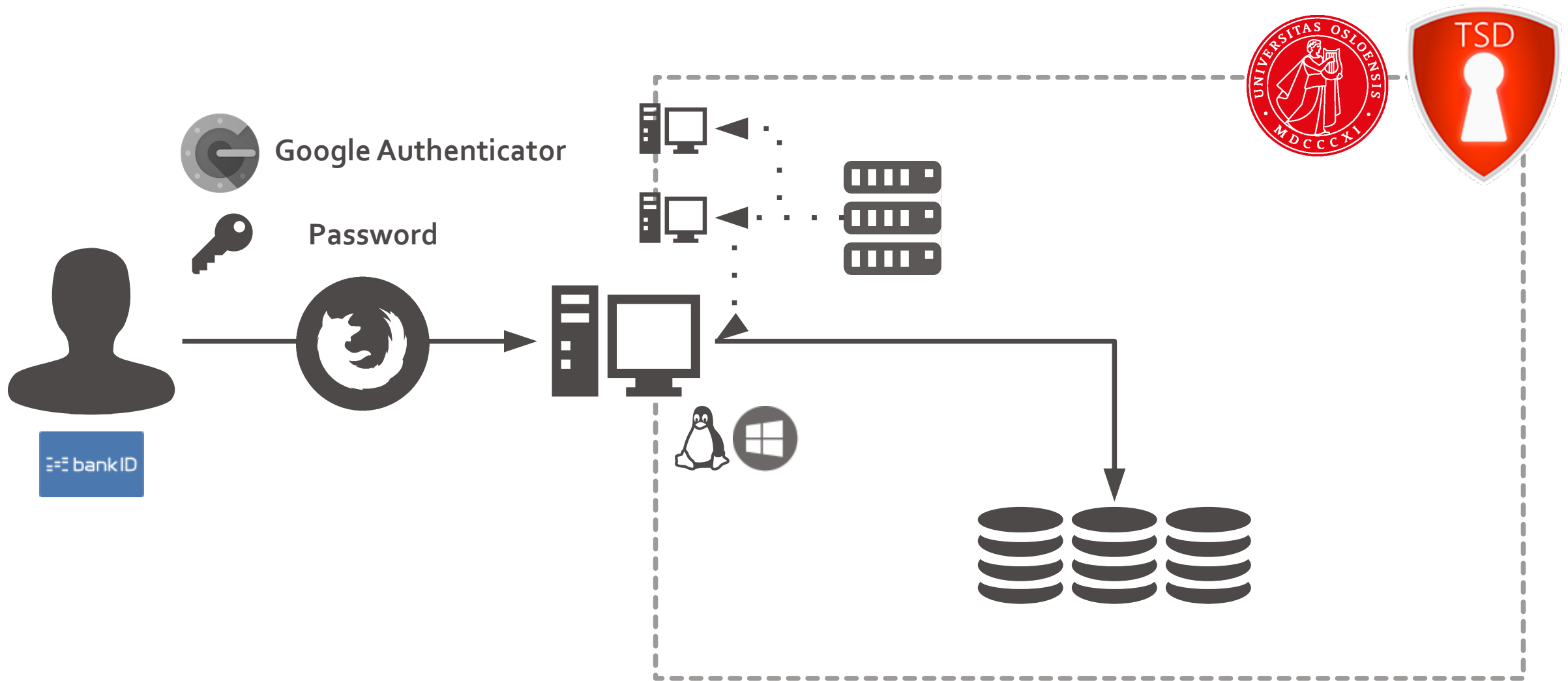
# TSD – National sensitive data service



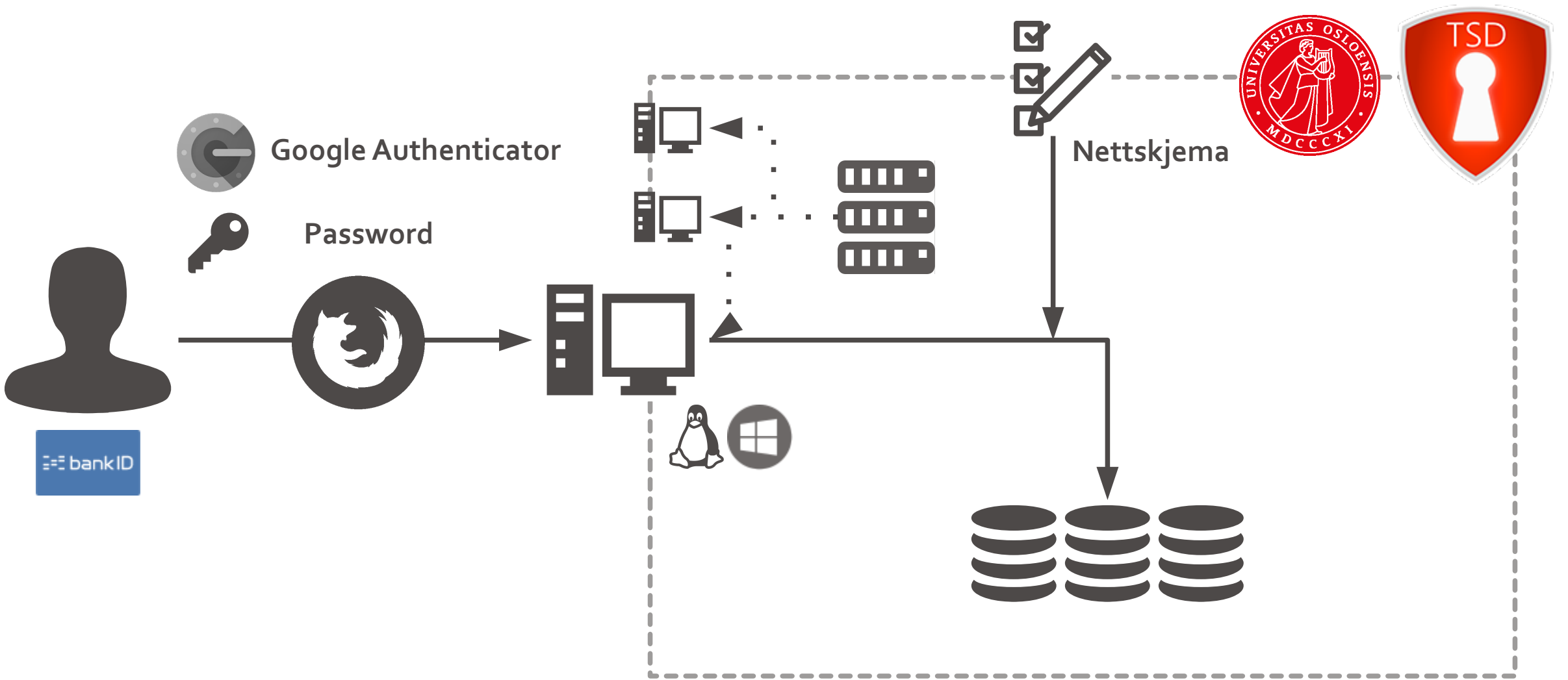
# TSD – National sensitive data service



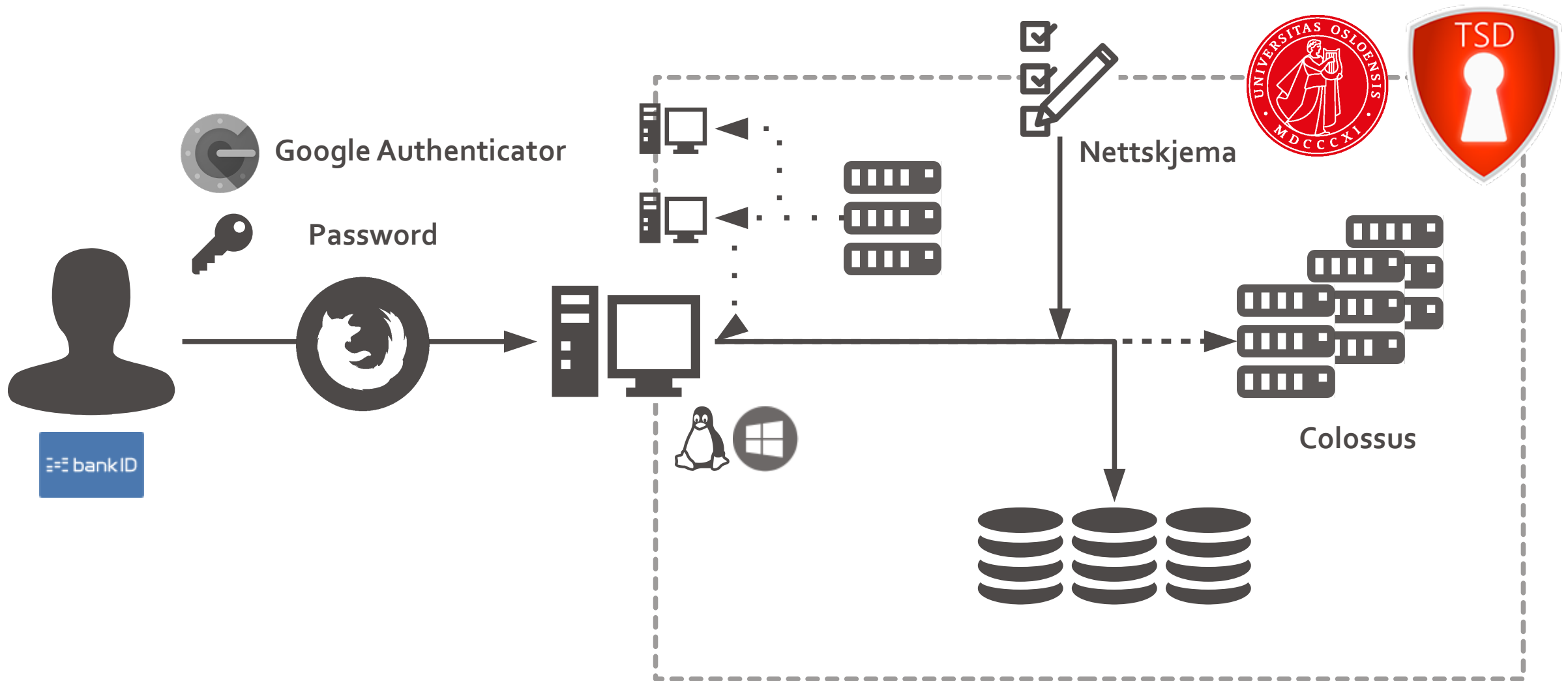
# TSD – National sensitive data service



The diagram illustrates a security architecture for a university system. On the left, a user (represented by a person icon) is associated with a blue box labeled "bankID". The user provides input to a central system (represented by a computer monitor and tower icon) via a "Password" (key icon) and "Google Authenticator" (circular icon). The central system is connected to a network of servers (represented by three server racks) and a database (represented by three stacked cylinders). A dashed line connects the central system to a "Nettskjema" (checklist icon) and a "TSD" (red shield icon with a keyhole). The "Nettskjema" is also connected to a red circular seal of "UNIVERSITAS OSLOENSIS" (University of Oslo) and a red shield icon with a keyhole. The "TSD" is connected to the database. The database is also connected to the central system via a dashed line.



# 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**