BRIGHT Data Catalog

The Gate to FAIR Research Informatics Platform

Ding He, Research Data Steward, BRIGHT at DTU

Nordic Love Data Week
2026-02-10

For me

I want to introduce our data infrastructure in the context of our needs.

I want share my experiences working with our solutions.

For you

I hope you can have some fruits of thoughts from my experiences for your context.

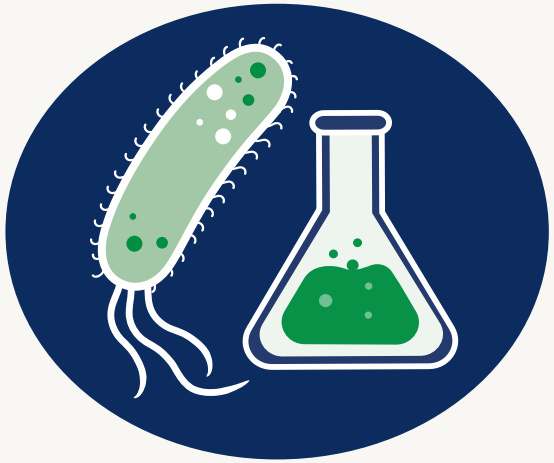
I hope you can bring your thoughts to your institutional environments and find good solutions for your needs.

Agenda

1. BRIGHT/Informatics at DTU
2. Data Catalog & Integrations
3. My thoughts & Discussions

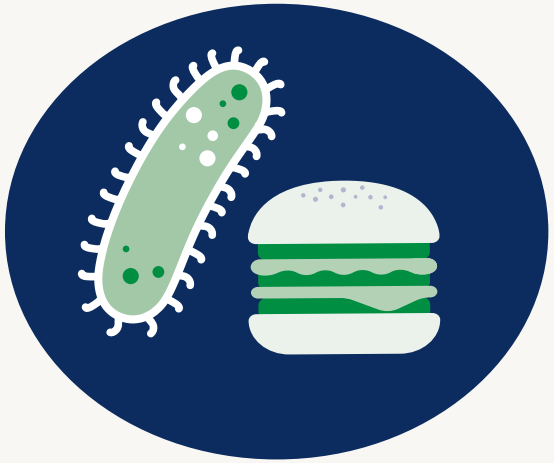
BRIGHT INFORMATICS

Biotechnology Research Institute for Green H Transition



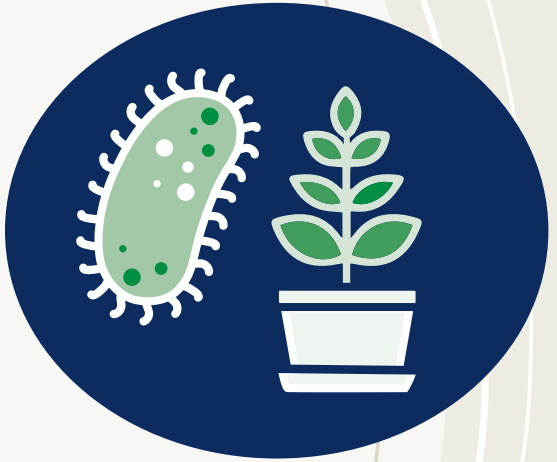
SUSTAINABLE MATERIALS

To establish pipelines for designing and implementing smart, bio-inspired alternatives to unsustainable polymer materials and commodity chemicals



MICROBIAL FOODS

To establish a knowledgebase and framework for the rational design of microbial food products that fulfil the criteria of consumer acceptance, sustainability, scalability, and health

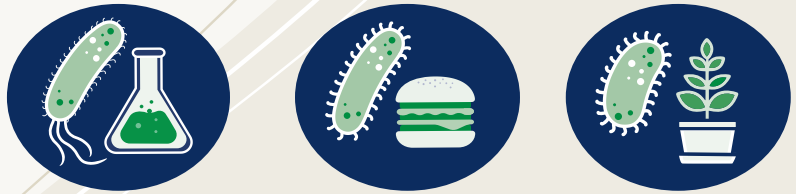


MICROBES FOR NET ZERO AGRICULTURE

To create engineered microbes as novel, effective solutions for net zero agriculture.

https://www.biosustain.dtu.dk/technologies/biofoundry

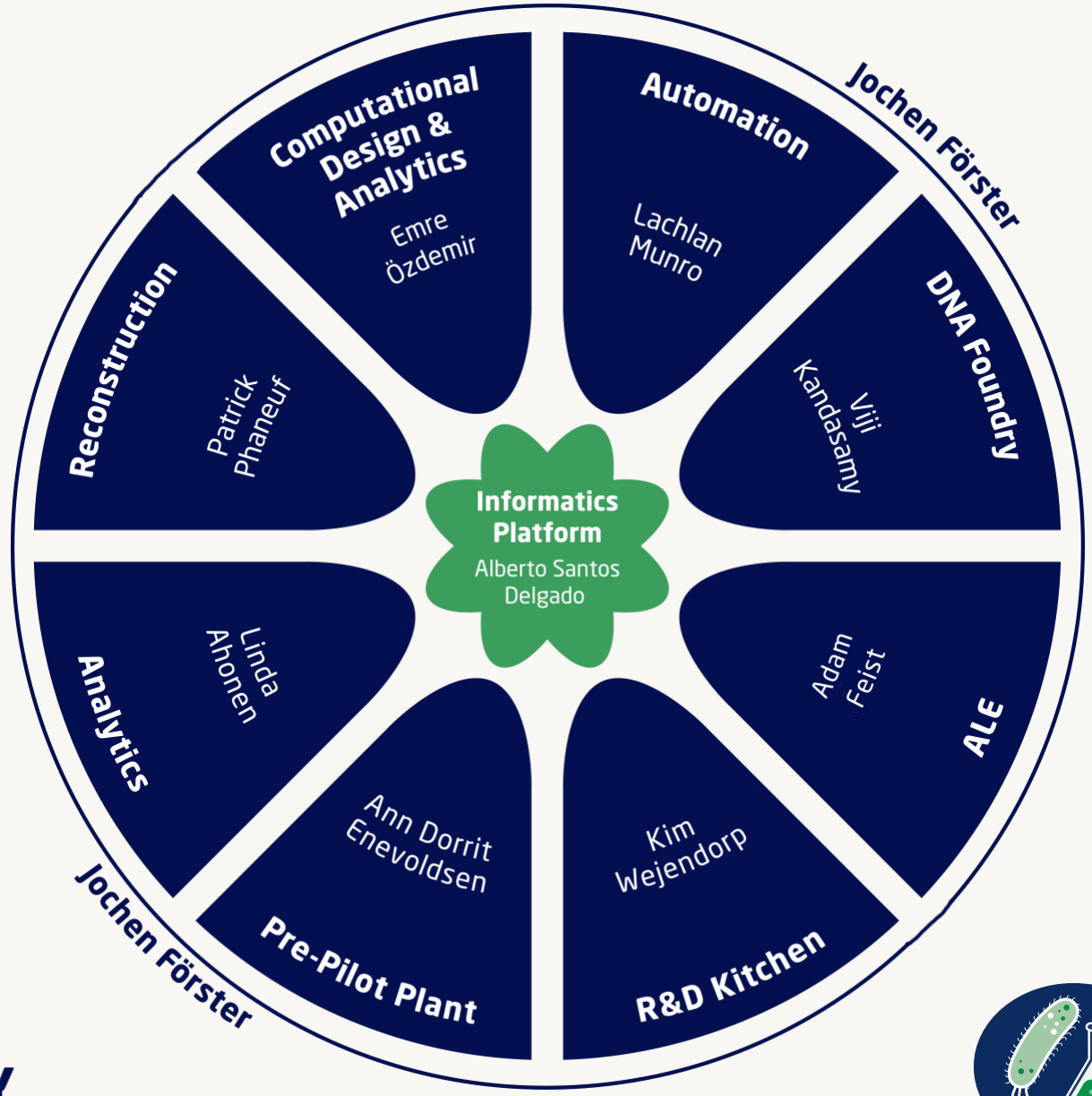
BRIGHT BIOFOUNDRY



https://www.biosustain.dtu.dk/technologies/biofoundry

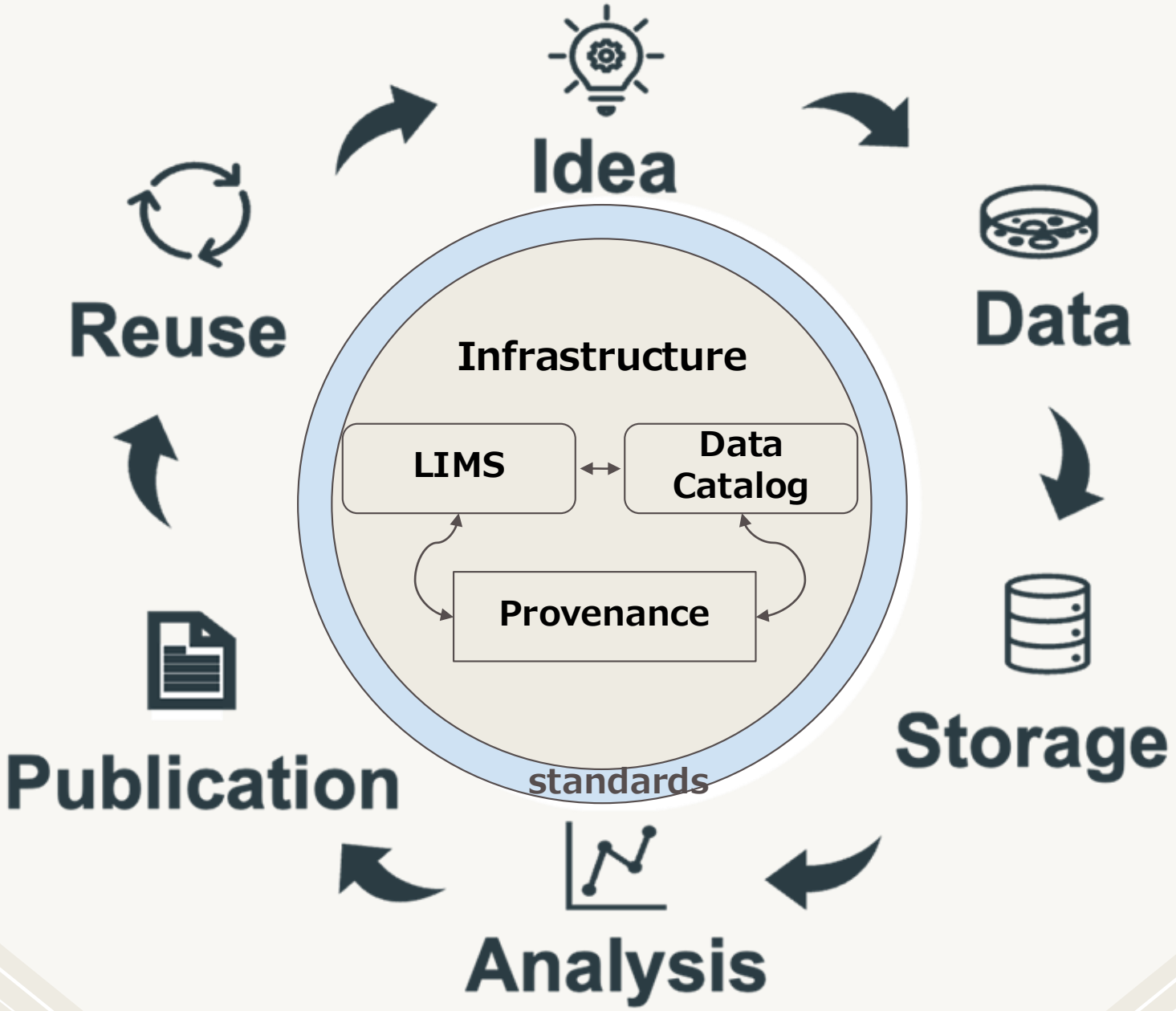
DTU

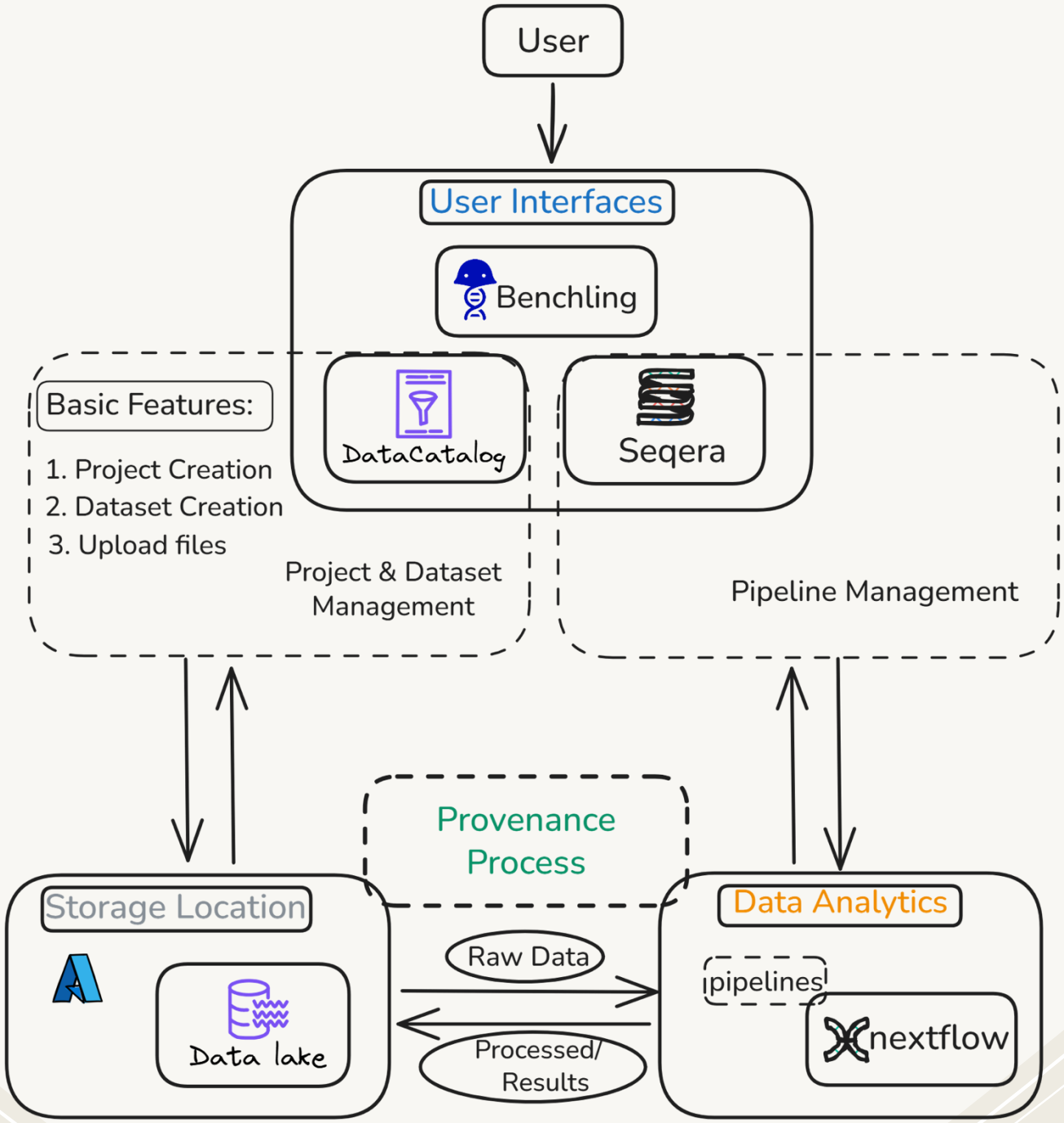
BRiGHT



BRIGHT
BIOFOUNDRY


Idea
Data
Analysis
Publication

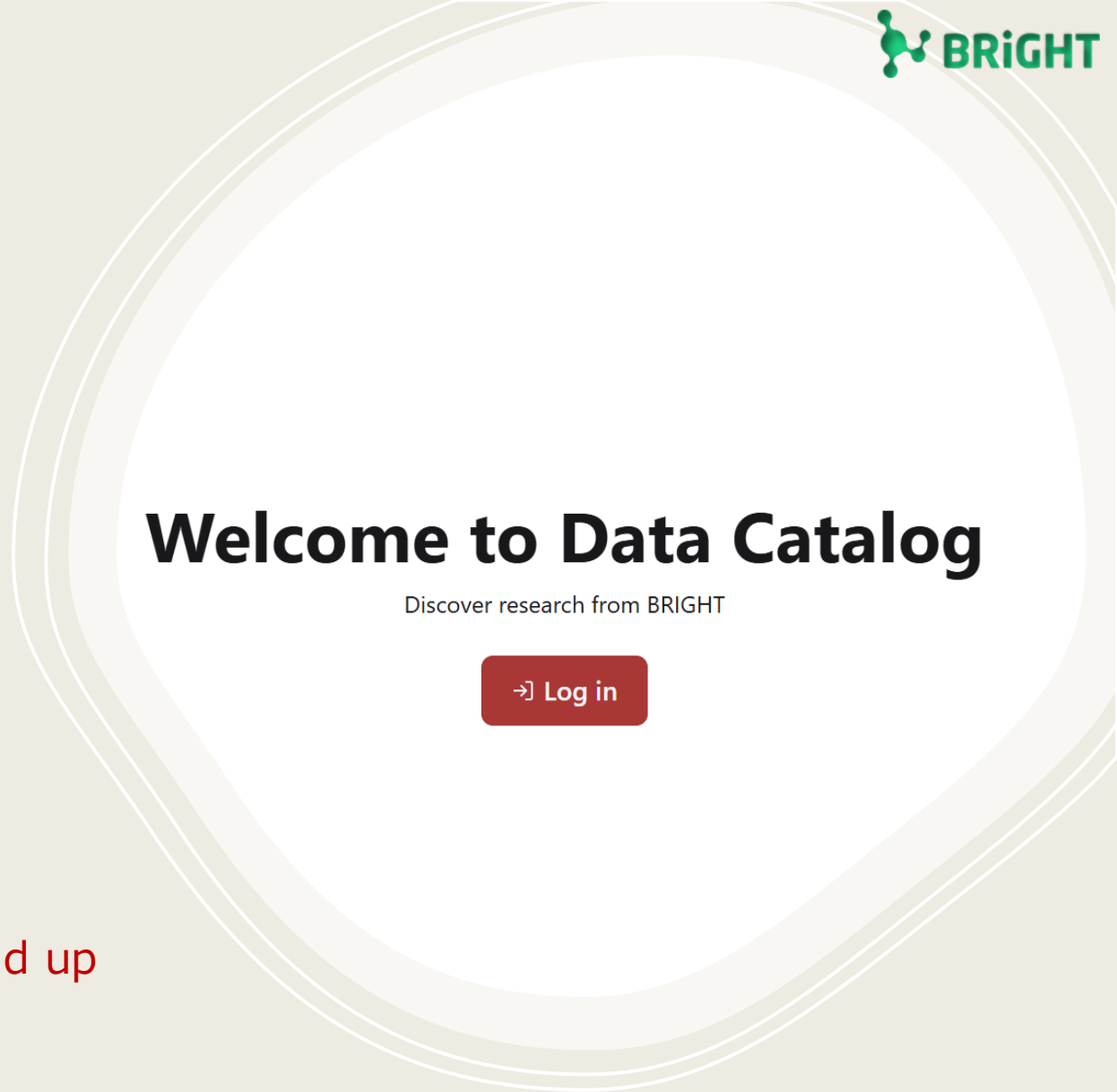

Idea
Data
Storage
Analysis
Publication
Reuse
Infrastructure
LIMS
Data Catalog
Provenance
standards


User
User Interfaces
Benchling
Basic Features:
1. Project Creation
2. Dataset Creation
3. Upload files
DataCatalog
Seqera
Project & Dataset Management
Pipeline Management
Provenance Process
Storage Location
Data lake
Raw Data
Processed/ Results
Data Analytics
pipelines
nextflow

BRIGHT

Data Catalog

Build from the ground up



Three Key Data Assets

Project

- Top-level organizational entity
- Represents a research or analytics unit
- Captures data and rich metadata

Dataset

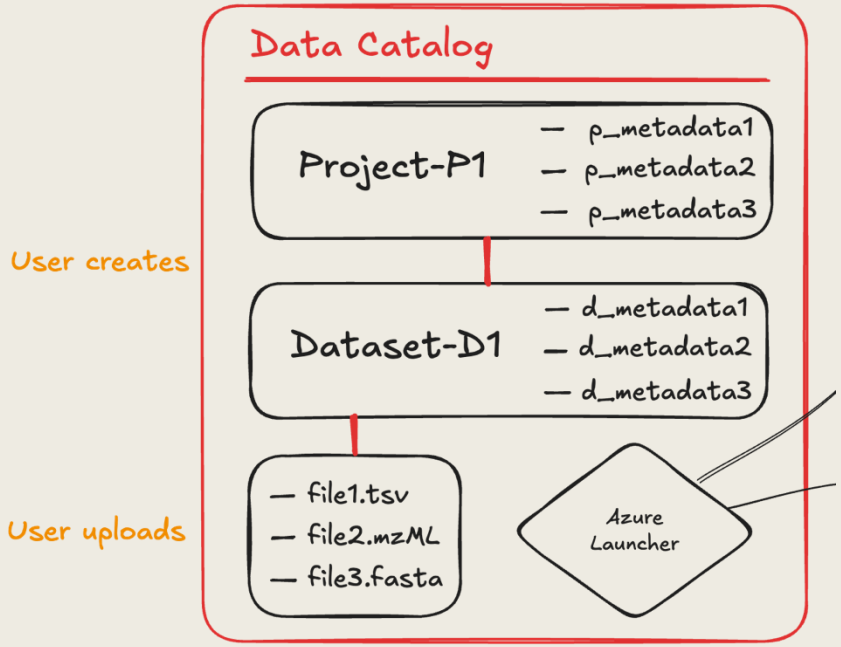
- Collection of related files and metadata
- Describes content, purpose, and properties of the dataset itself

Files

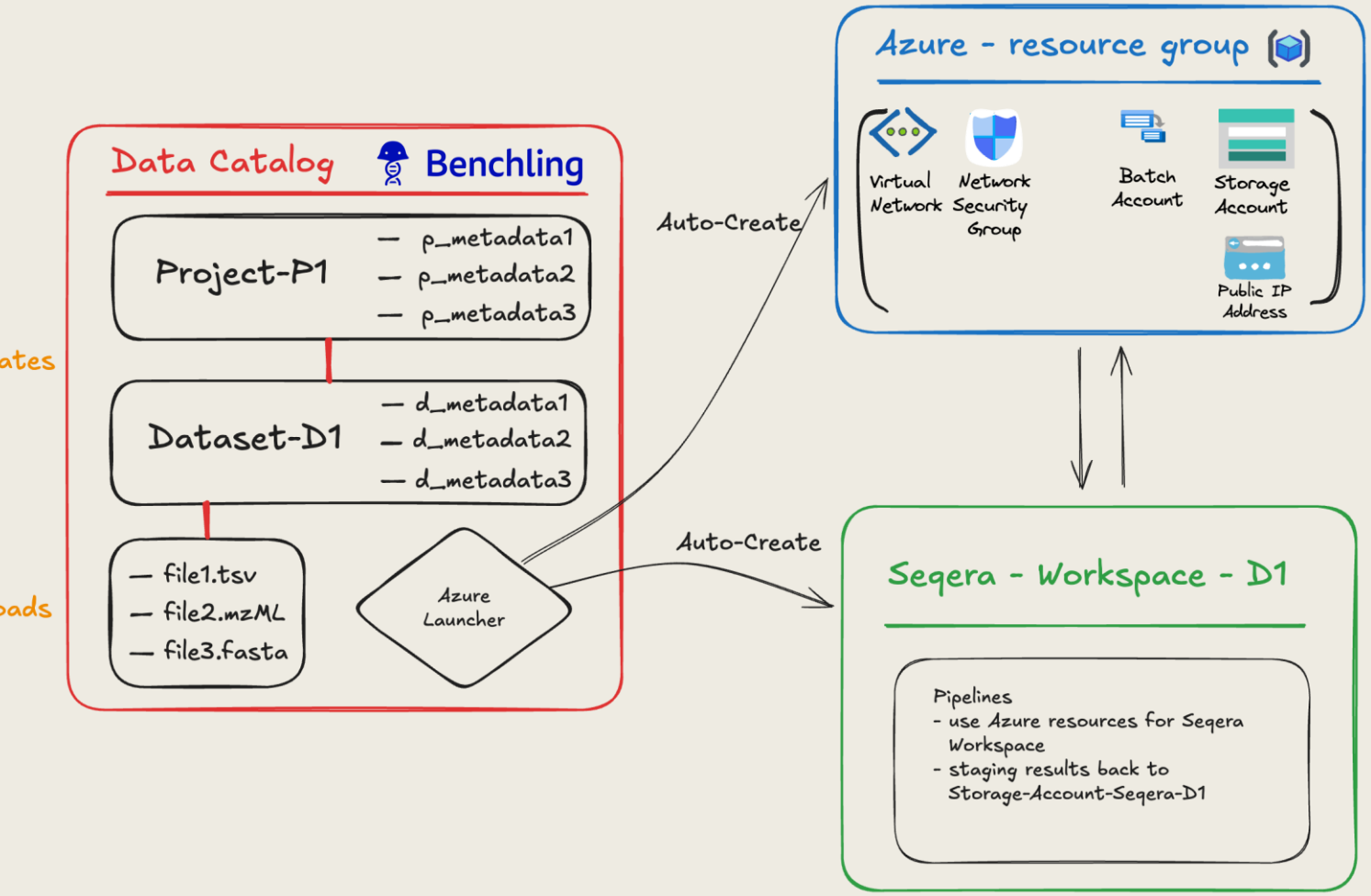
- The smallest data entity in the system
- Each file belongs to a dataset
- Store in the cloud

BRIGHT Data Catalog offers

- **Metadata management** layer that enables the **registration**, **discovery**, **access**, and **governance** of Projects, Datasets, and Files.
- As the **gateway** between users and the underlying research data infrastructure exposing **metadata records** for all assets (Projects, Datasets, and Files).
- **Searching, browsing, and filtering** of data assets using metadata and keywords.
- Provides **fine-grained access control** mechanisms to manage visibility and permissions.



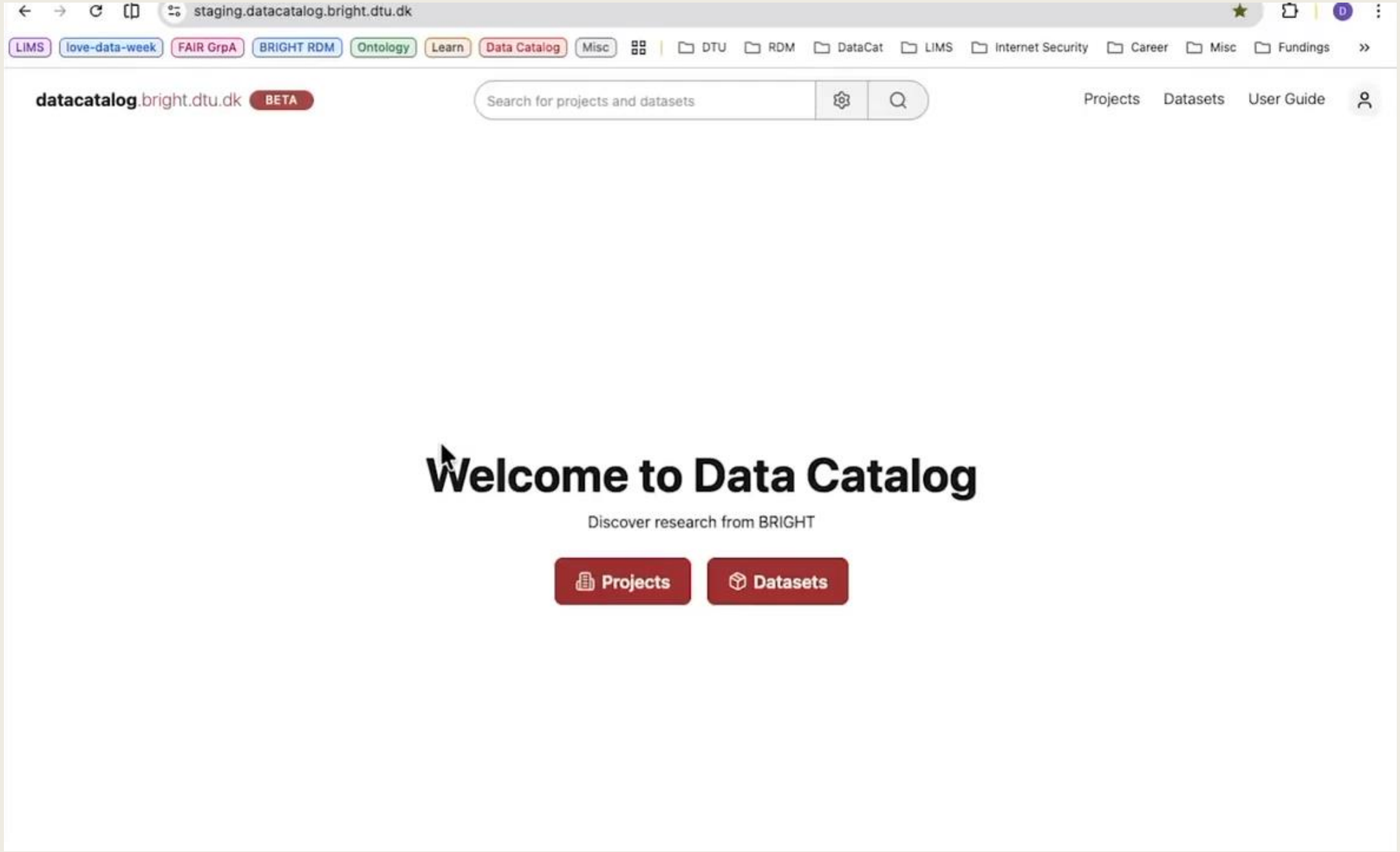
BRIGHT Data Catalog integrates with



- **Azure** cloud solutions provides **scalable computation** and **storage**.
- **Seqera** platform enables highly **customisable** and **reproducible computational pipelines**.
- **LIMS (Benchling)** links **wet-lab** to **drylab**.

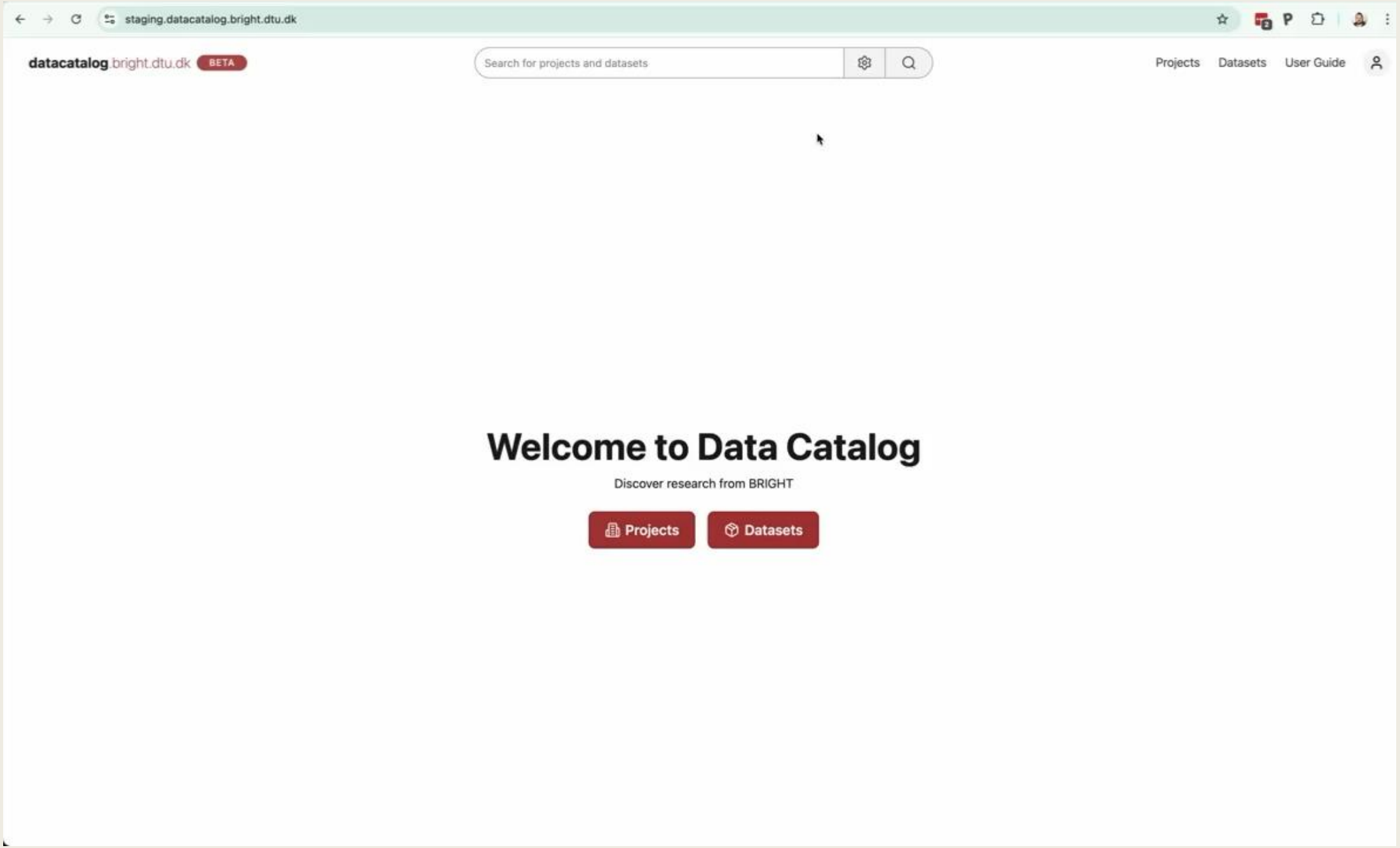
An example collaborative scenario

- Researchers: Ding and Henry
- Research project: E. coli MS-based Proteomics
- Ding finished a Data-Dependent Acquisition (DDA) proteomics with a mass spectrometry, and stored the generated dataset in the Data Catalog. He shared this result with his data scientist team member, Henry.

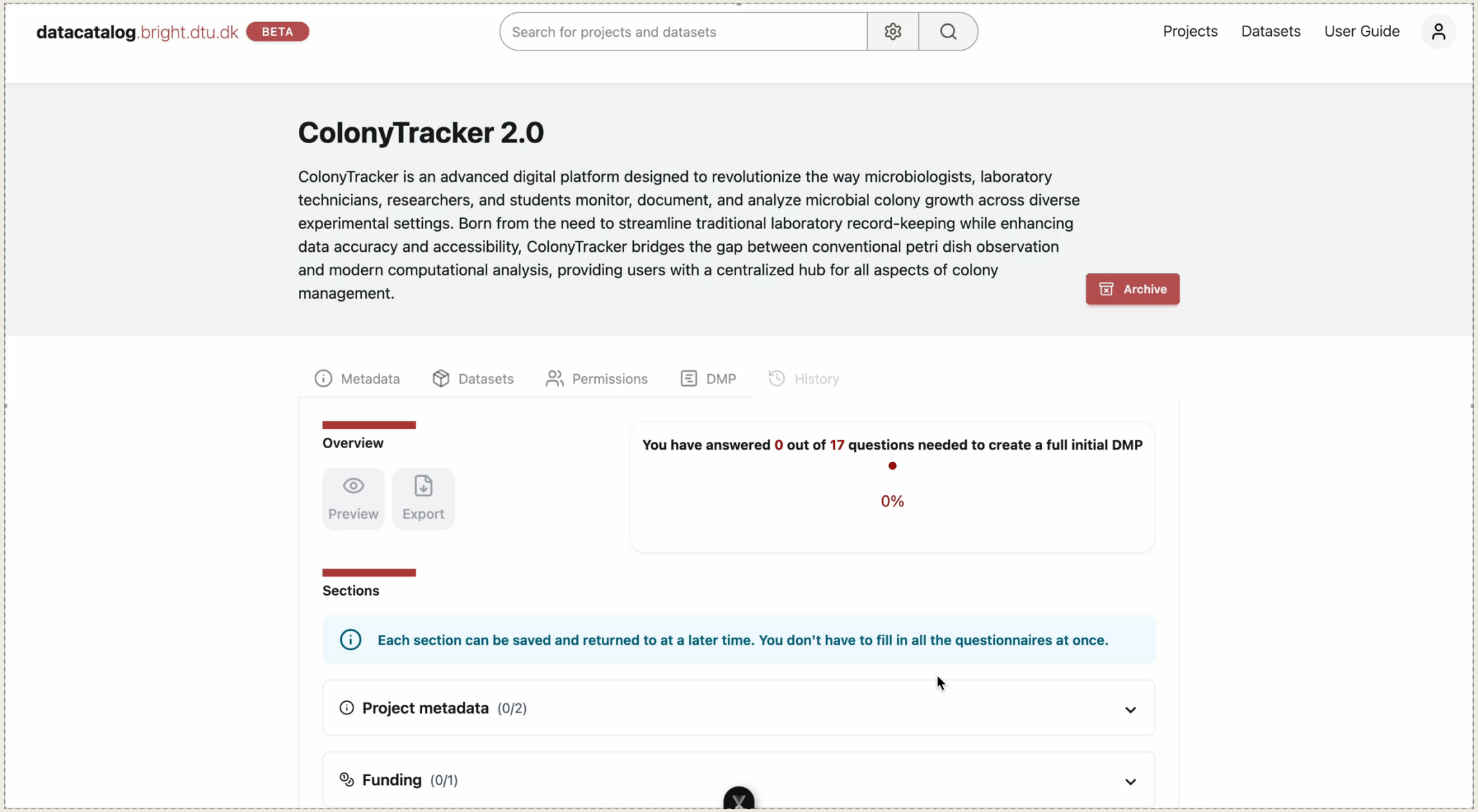

staging.datacatalog.bright.dtu.dk
LIMS
love-data-week
FAIR GrpA
BRIGHT RDM
Ontology
Learn
Data Catalog
Misc
DTU
RDM
DataCat
LIMS
Internet Security
Career
Misc
Fundings
datacatalog.bright.dtu.dk
BETA
Search for projects and datasets
Projects
Datasets
User Guide
Welcome to Data Catalog
Discover research from BRIGHT
Projects
Datasets

An example collaborative scenario

- Researchers: Ding and Henry
- Research project: E. coli MS-based Proteomics
- Ding finished a Data-Dependent Acquisition (DDA) proteomics with a mass spectrometry, and stored the generated dataset in the Data Catalog. He shared this result with his data scientist team member, Henry.
- Henry gained access to this dataset from Ding, performed reproducible and standardized bioinformatics Nextflow pipelines in the Seqera platform.
- Henry can now share his pipeline results with Ding and continue the next phase of the research activities.


staging.datacatalog.bright.dtu.dk
datacatalog.bright.dtu.dk
BETA
Search for projects and datasets
Projects
Datasets
User Guide
Welcome to Data Catalog
Discover research from BRIGHT
Projects
Datasets

Build-in Data Management Plan



My thoughts (questions?)

- Build from scratch or adopt existing solutions?
- Features first or needs first?
- Who needs a solution?
- What issue needs a solution?
- What is the scope?
- Work with users to prioritize requirements.
- Plan strategically based on avaliable resources.

novo nordisk **fonden**

DTU Technical University of Denmark

Software Development

Pasquale Colaianni
Team Lead

Dominik Maszczyk
Senior software engineer

Karina Reimer-Hansen
Software engineer

Marek Fijolek
Software engineer

Data Management

Konstantinos Tsirigos
Team Lead

Atieh Gharib
Data engineer

Ding He
Research Data Steward

Fatemeh Ghanami
Data engineer

Mia Ontiveros
LIMS admin

Vasileia Vagena
Research assistant

Data Science

Alberto Pallejà Caro
Team Lead

Henry Emanuel Webel
Senior data scientist

Juliana Assis
Senior Data scientist

Sebastian Schulz
Senior Data scientist

Edir Castro
Data scientist

Alberto Santos Delgado
Platform Lead / Senior Researcher

Informatics Platform