

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.

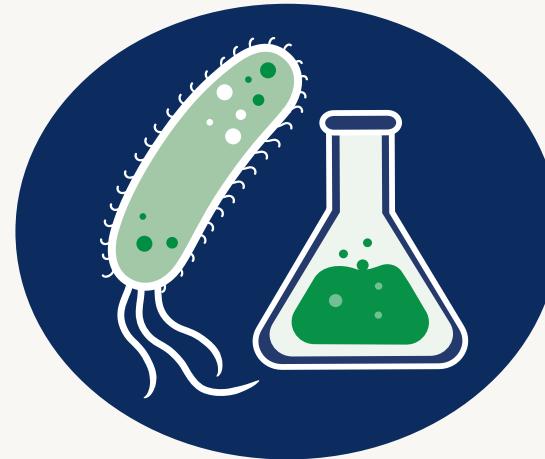
Agenda

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



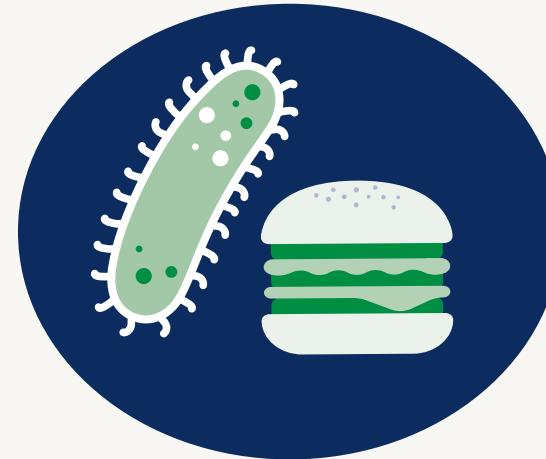
BRIGHT INFORMATICS

Biotechnology Research Institute for Green H Transition



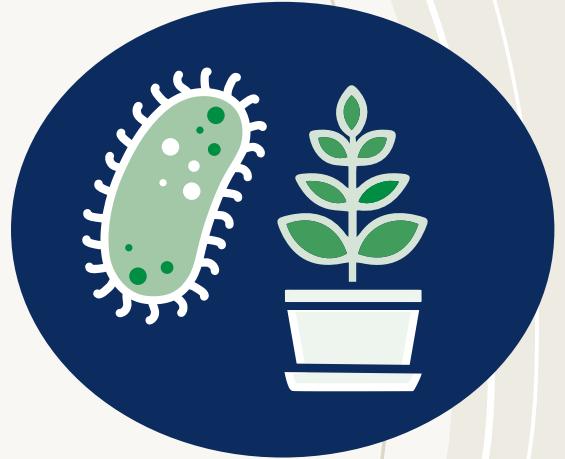
SUSTAINABLE MATERIALS

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

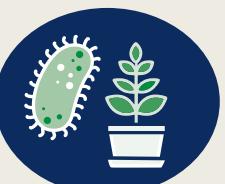
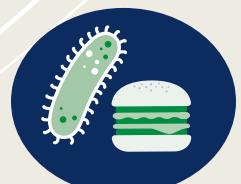


MICROBES FOR NET ZERO AGRICULTURE

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

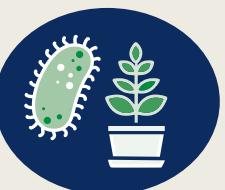
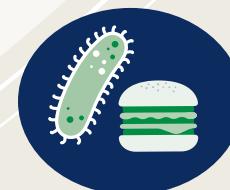
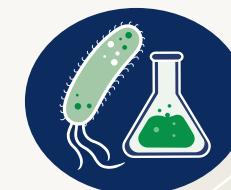


BRIGHT BIOFOUNDRY

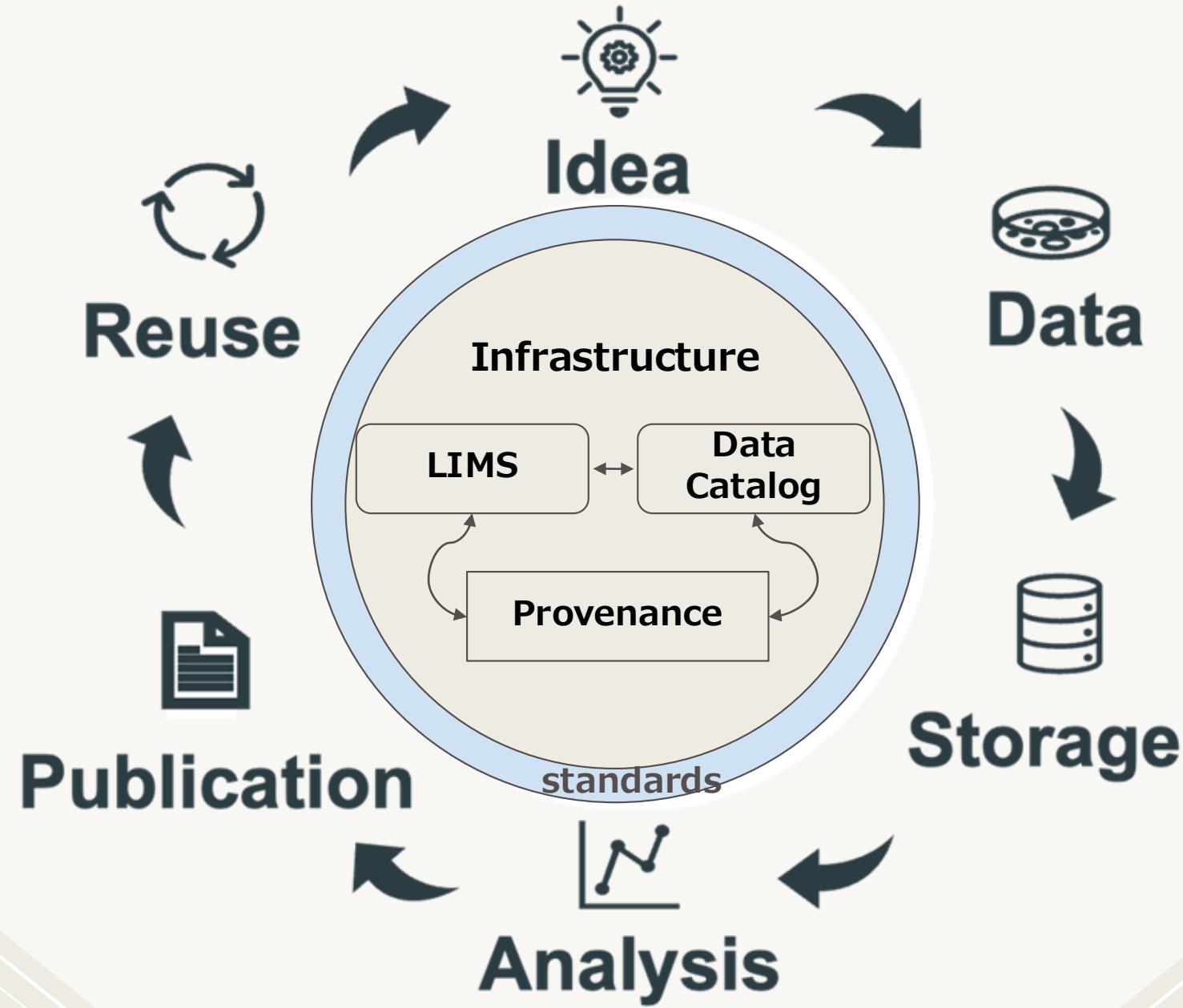


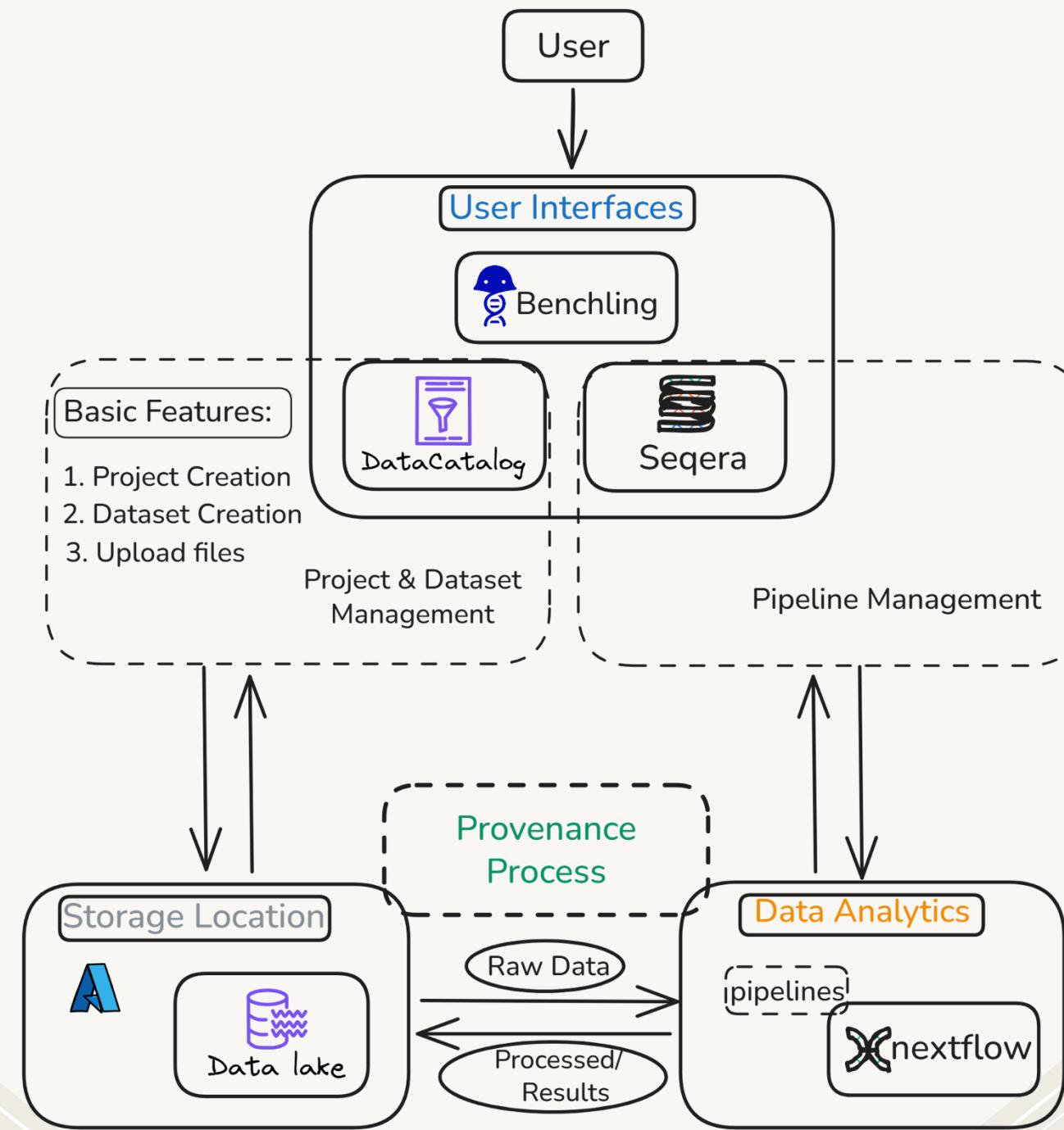


BRIGHT
BIOFOUNDRY









BRIGHT

Data Catalog

Build from the round up

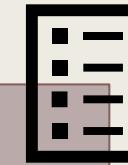
Welcome to Data Catalog

Discover research from BRIGHT

→ Log in

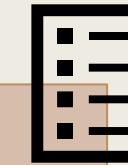
Three Key Data Assets

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

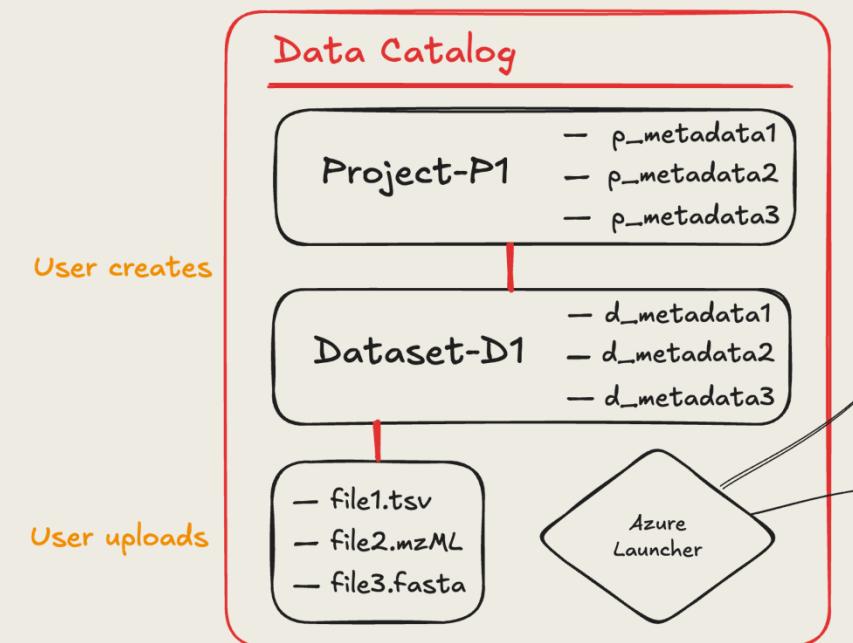
Files



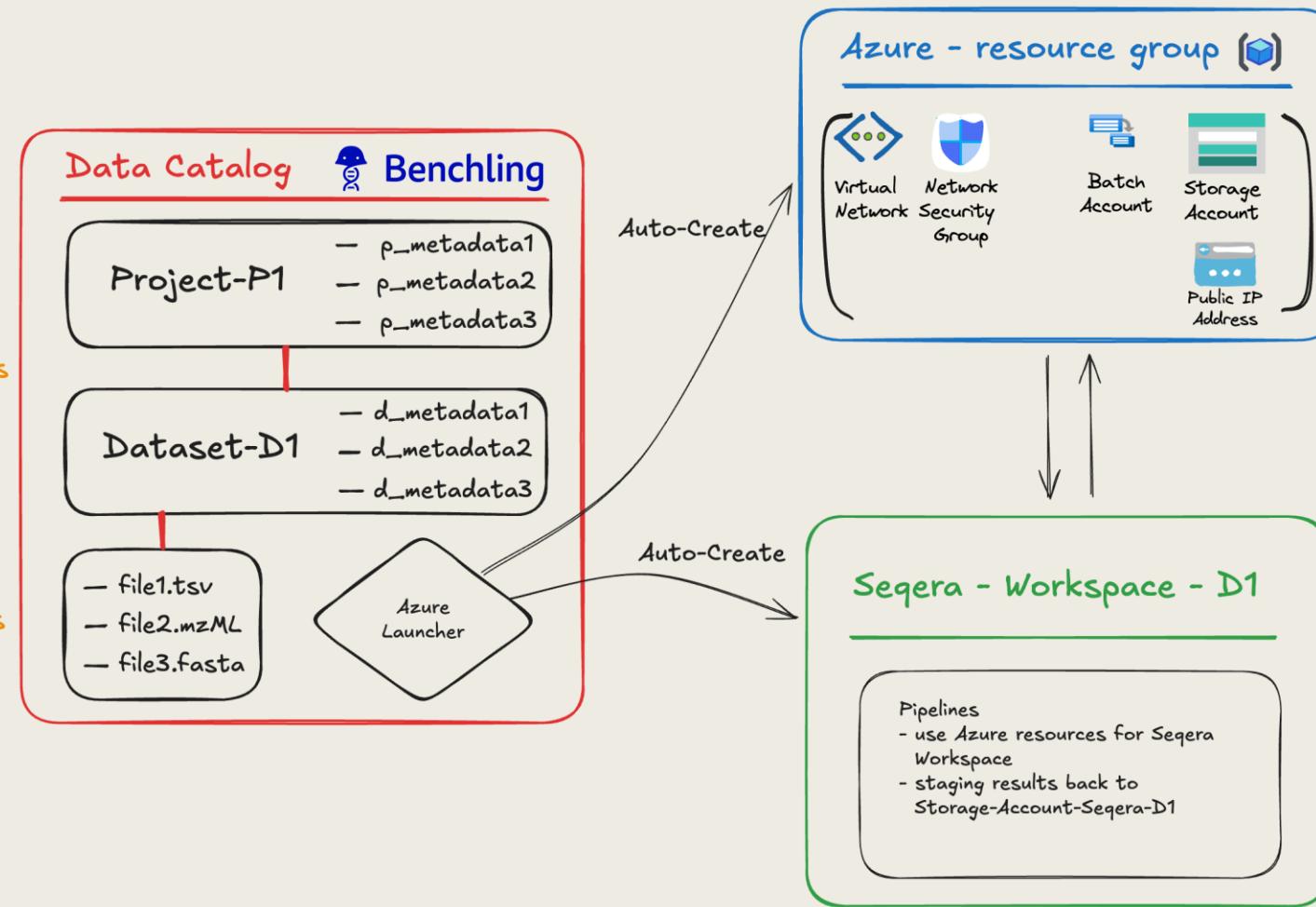
- The smallest data entity in the system
- Each file belongs to a dataset
- Store in the cloud

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

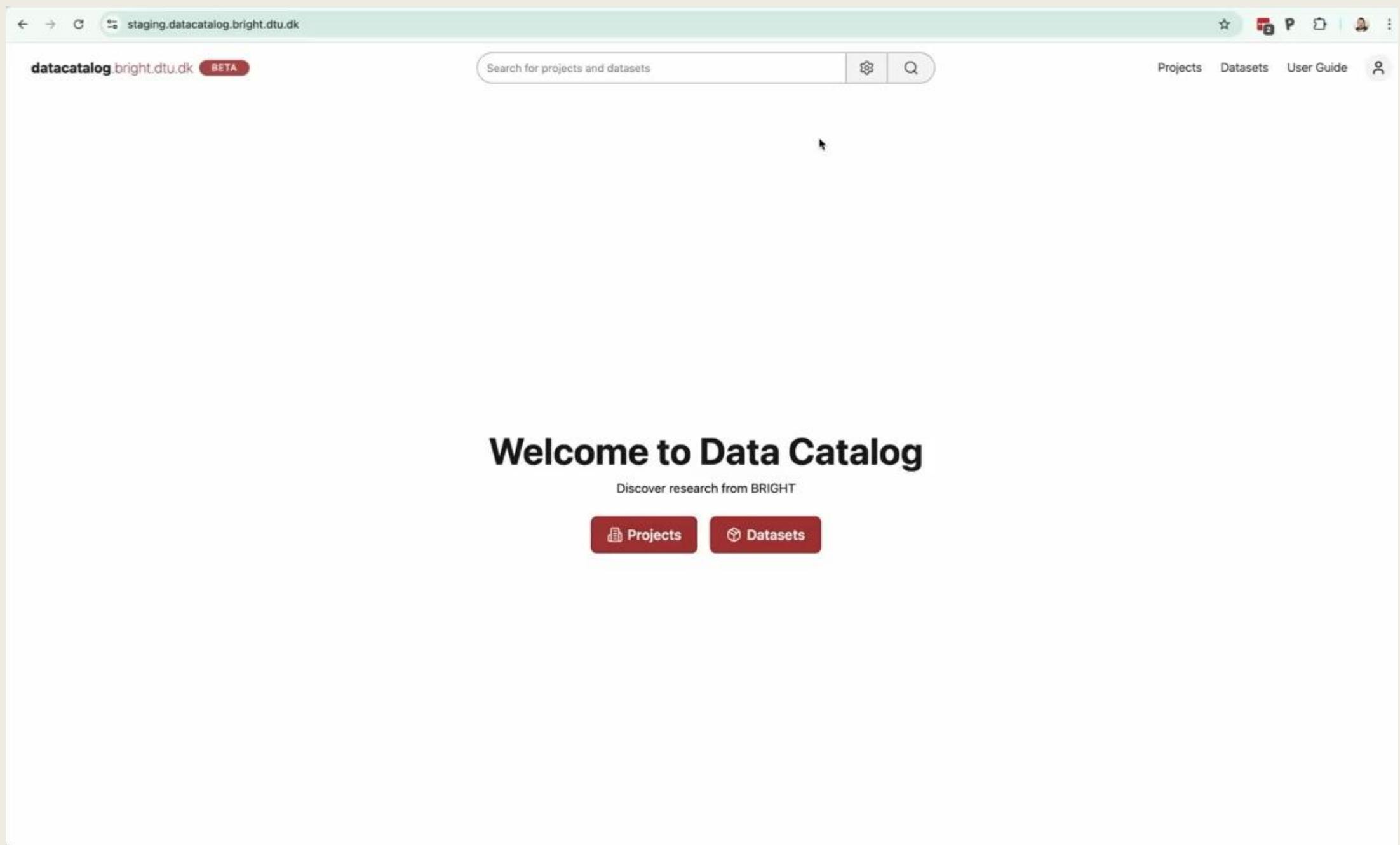
Welcome to Data Catalog

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.



The screenshot shows a web browser window for the Data Catalog. The URL in the address bar is `staging.datacatalog.bright.dtu.dk`. The page header includes the DTU logo, a search bar with placeholder text "Search for projects and datasets", and a gear icon. On the right side of the header are links for "Projects", "Datasets", "User Guide", and a user profile icon. The main content area features a large "Welcome to Data Catalog" heading, a subtitle "Discover research from BRIGHT", and two prominent red buttons labeled "Projects" and "Datasets".

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

The screenshot shows the 'ColonyTracker 2.0' project page on the datacatalog.bright.dtu.dk platform. At the top, there's a search bar with placeholder text 'Search for projects and datasets', a gear icon, and a magnifying glass icon. To the right are links for 'Projects', 'Datasets', 'User Guide', and a user profile icon.

The main content area features a large heading 'ColonyTracker 2.0' and a detailed description of the platform's purpose:

ColonyTracker is an advanced digital platform designed to revolutionize the way microbiologists, laboratory technicians, researchers, and students monitor, document, and analyze microbial colony growth across diverse experimental settings. Born from the need to streamline traditional laboratory record-keeping while enhancing data accuracy and accessibility, ColonyTracker bridges the gap between conventional petri dish observation and modern computational analysis, providing users with a centralized hub for all aspects of colony management.

To the right of the description is a red 'Archive' button with a trash can icon.

Below the description, there's a navigation bar with icons for 'Metadata', 'Datasets', 'Permissions', 'DMP', and 'History'. The 'DMP' tab is active.

The main content area is divided into two sections: 'Overview' and 'Sections'.

Overview: A progress bar shows '0%'. A callout box states: 'You have answered 0 out of 17 questions needed to create a full initial DMP'.

Sections: A note says: 'Each section can be saved and returned to at a later time. You don't have to fill in all the questionnaires at once.' Below are two sections: 'Project metadata (0/2)' and 'Funding (0/1)'. Each section has a dropdown arrow icon.

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 available resources.

novo nordisk fonden

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