OPEN-SOURCE ONLINE PLATFORM FOR UAS HTP DATA MANAGEMENT

Data to Science Engine (D2SE)

Presenter: Ben Hancock; Geospatial Data Science Lab, Purdue University

OVERVIEW

How can researchers better manage large volumes of UAS HTP data?

Designing an open-source data management platform

Core features of Data to Science platform

Demonstration

DESIGNING AN OPEN-SOURCE DATA MANAGEMENT PLATFORM

GIS day
INSPIRE THE WORLD THROUGH GIS

PURDUE UNIVERSITY

MOTIVATION

Researchers today are capturing large volumes of UAS HTP at a high spatial and temporal resolution

Research can be impeded when we don't know what data we have and how to access it

Collaboration across disciplines and organizations present additional challenges

OBJECTIVE

Build an open-source online platform that research groups can use as a one-stop shop to:

- Upload, storage, and organize UAS data
- Process and visualize the data with interactive tools
- Collaborate with other researchers
- Publish findings to a public central catalog

MAJOR PLATFORM COMPONENTS

Application Instance

Containerized full-stack web application

- Frontend
- Backend
- Database
- Web server

Open-source and self-hosted

Public Central Catalog

Searchable catalog of published datasets from application instances

APPLICATION INSTANCE

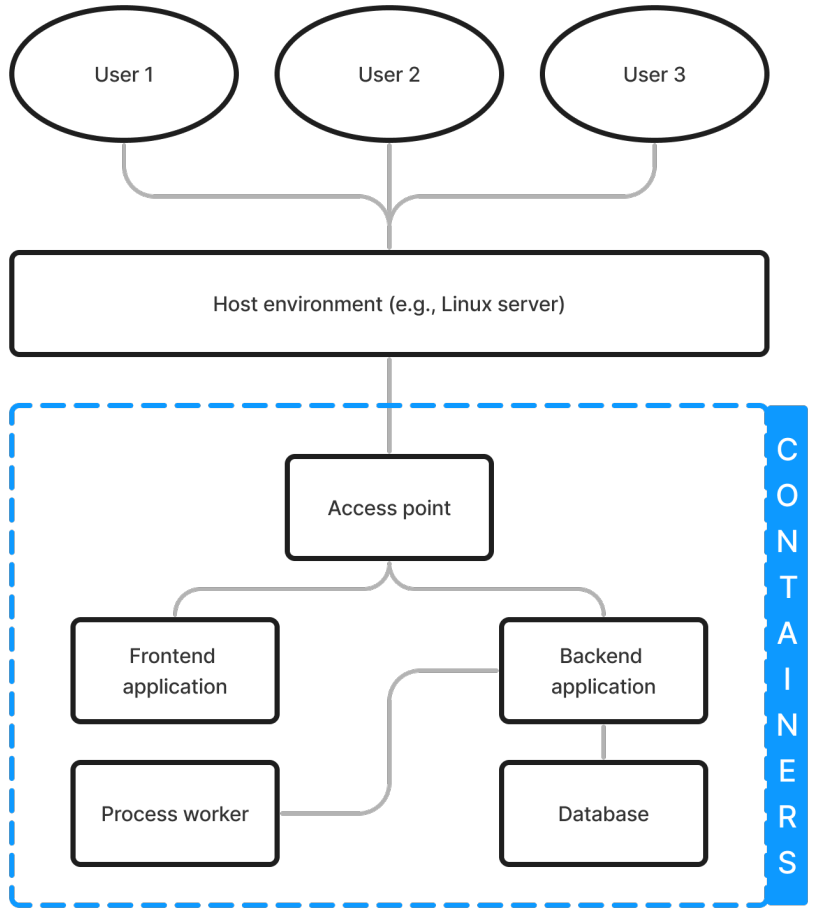
Entire application runs inside containers

One container for each component of the application

Highly portable; hostable from a variety of environments

- Linux, Windows, Kubernetes etc.

User data remains under application instance



DEPLOY WITH DOCKER

Entire application instance can be deployed with a few commands

docker compose build – Creates images of each container's environment

docker compose up – Runs the containers

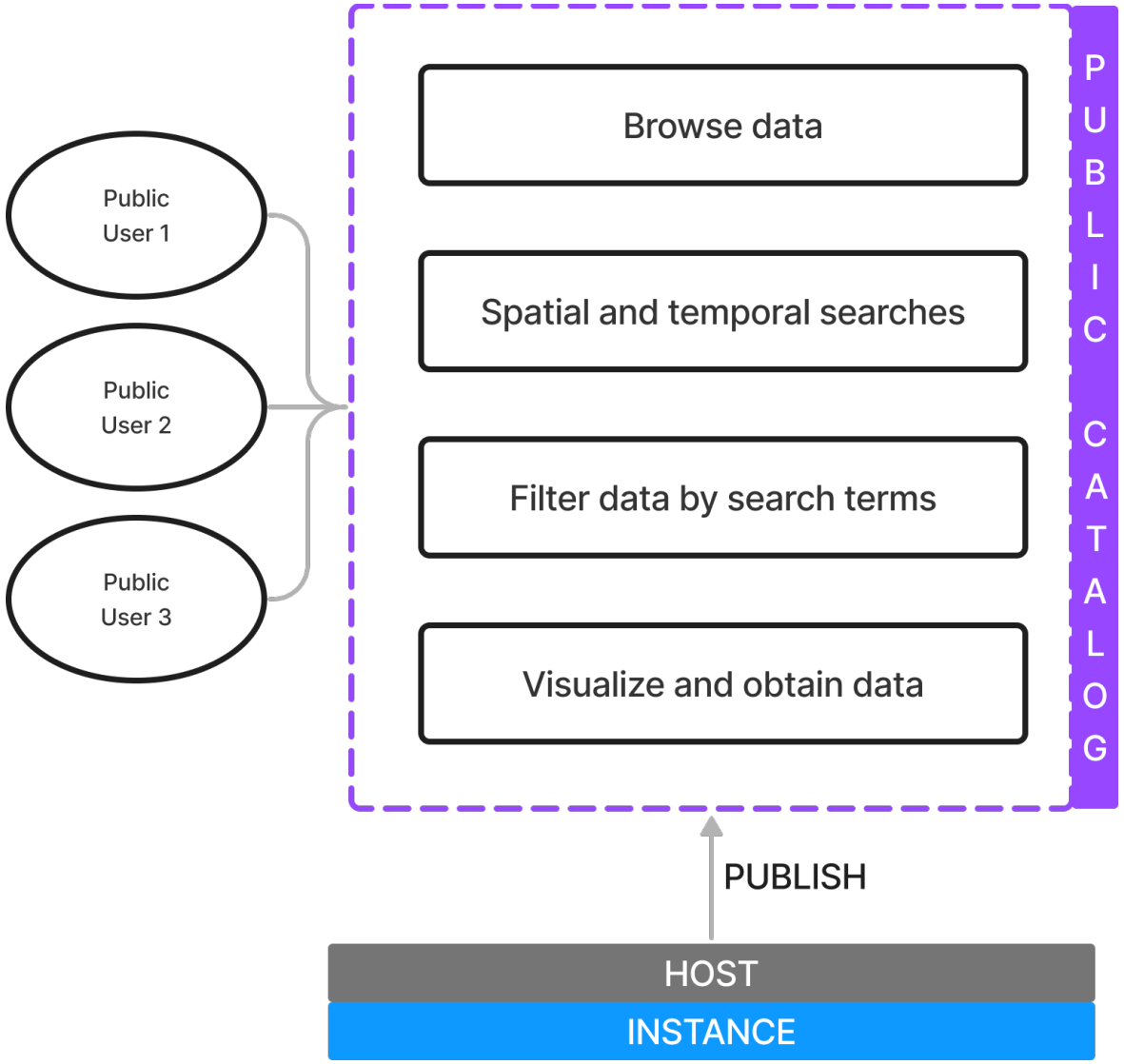
Does not require an in-depth knowledge of the technologies running within containers

PUBLIC CENTRAL CATALOG

Publish data from an application instance to a central, public catalog

Only metadata is published to the catalog

Physical dataset will remain within the application instance



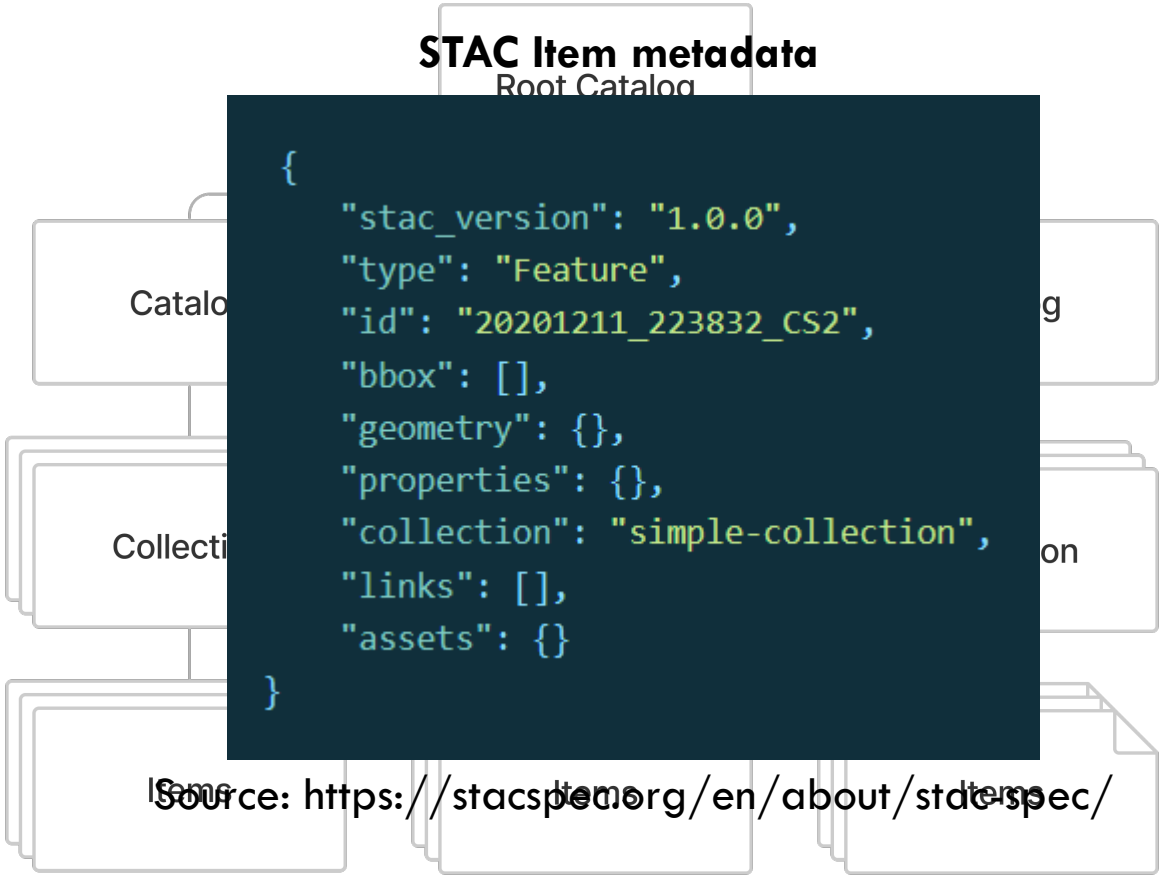
STAC

Centralized catalog built to **S**patio**T**emporal **A**sset **C**atalog (STAC) specification

Three primary data models: Catalog, Collection, and Item

Involved community, large collection of tools and extensions

STAC CATALOG STRUCTURE

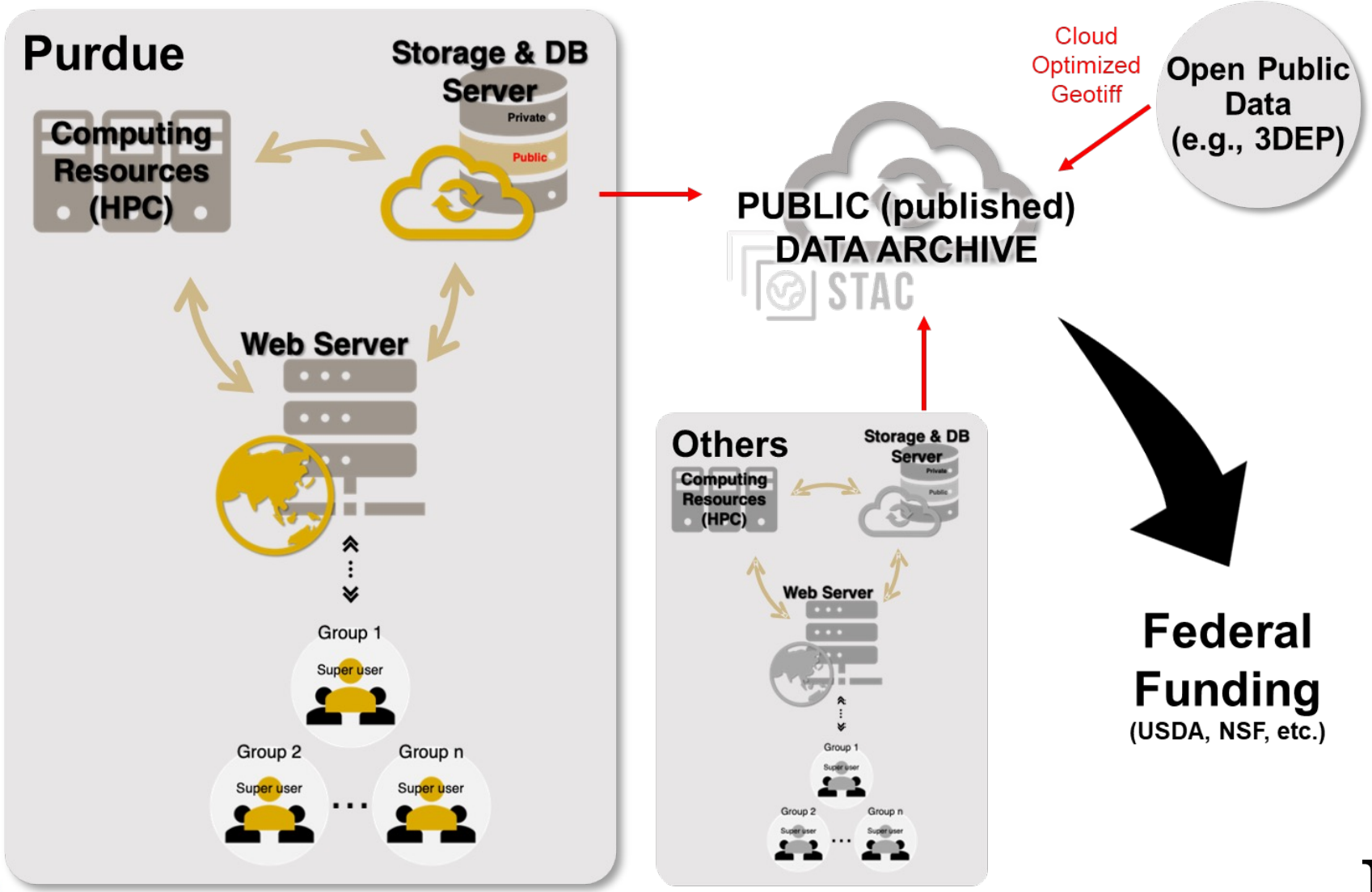


Source: https://stacspec.org/en/about/stac-spec/

GOAL

"To make UAS HTP data FAIR and an online platform SCALABLE"

Findable | Accessible | Interoperable | Reusable



GIS day
INSPIRE THE WORLD THROUGH GIS

PURDUE UNIVERSITY

CORE FEATURES OF DATA TO SCIENCE PLATFORM

GIS day
INSPIRE THE WORLD THROUGH GIS

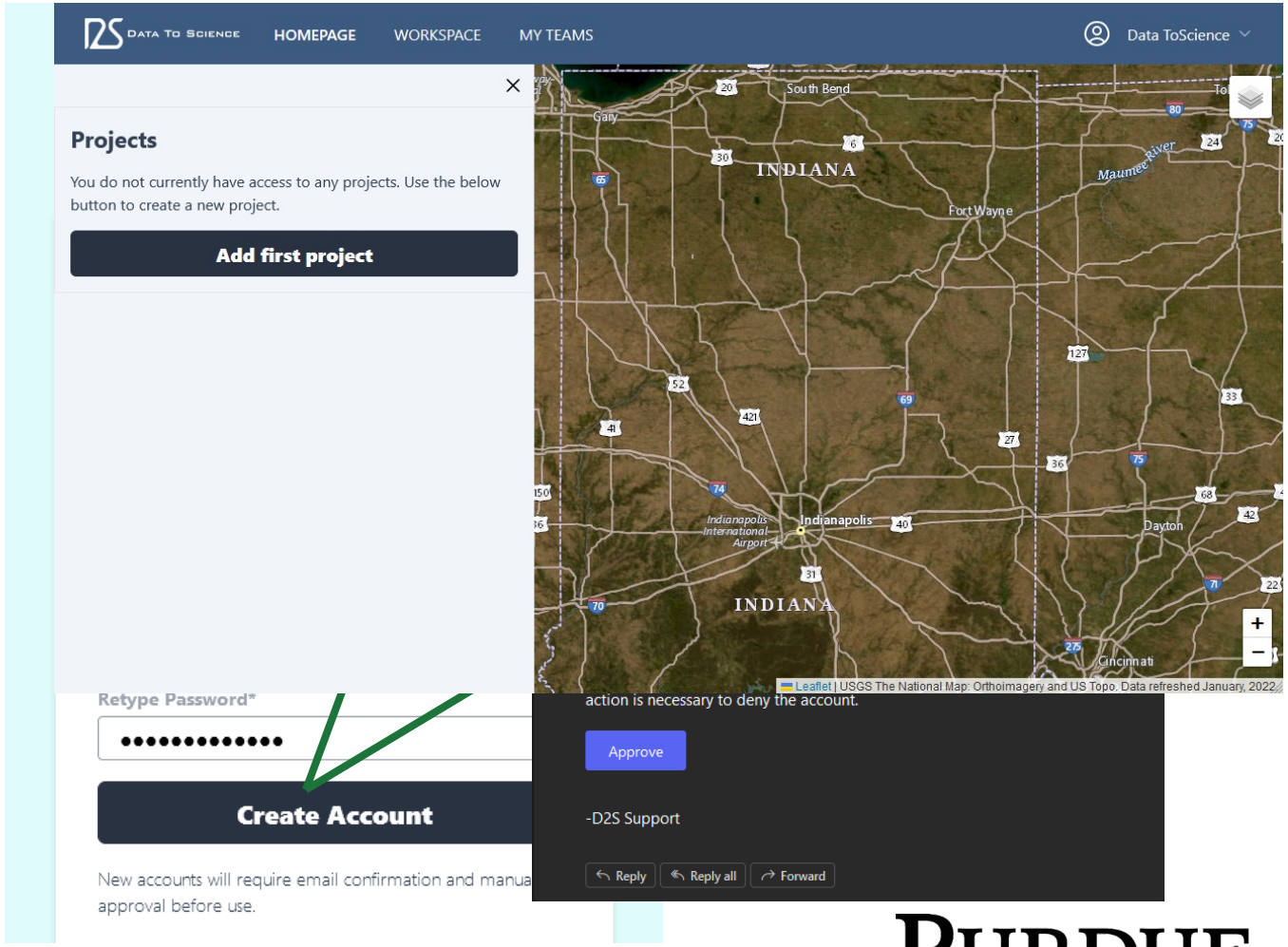
PURDUE UNIVERSITY

ACCOUNT MANAGEMENT

Account required to access web application

User verification required by admins

Email confirmation, password recovery, profiles



GIS day
INSPIRE THE WORLD THROUGH GIS

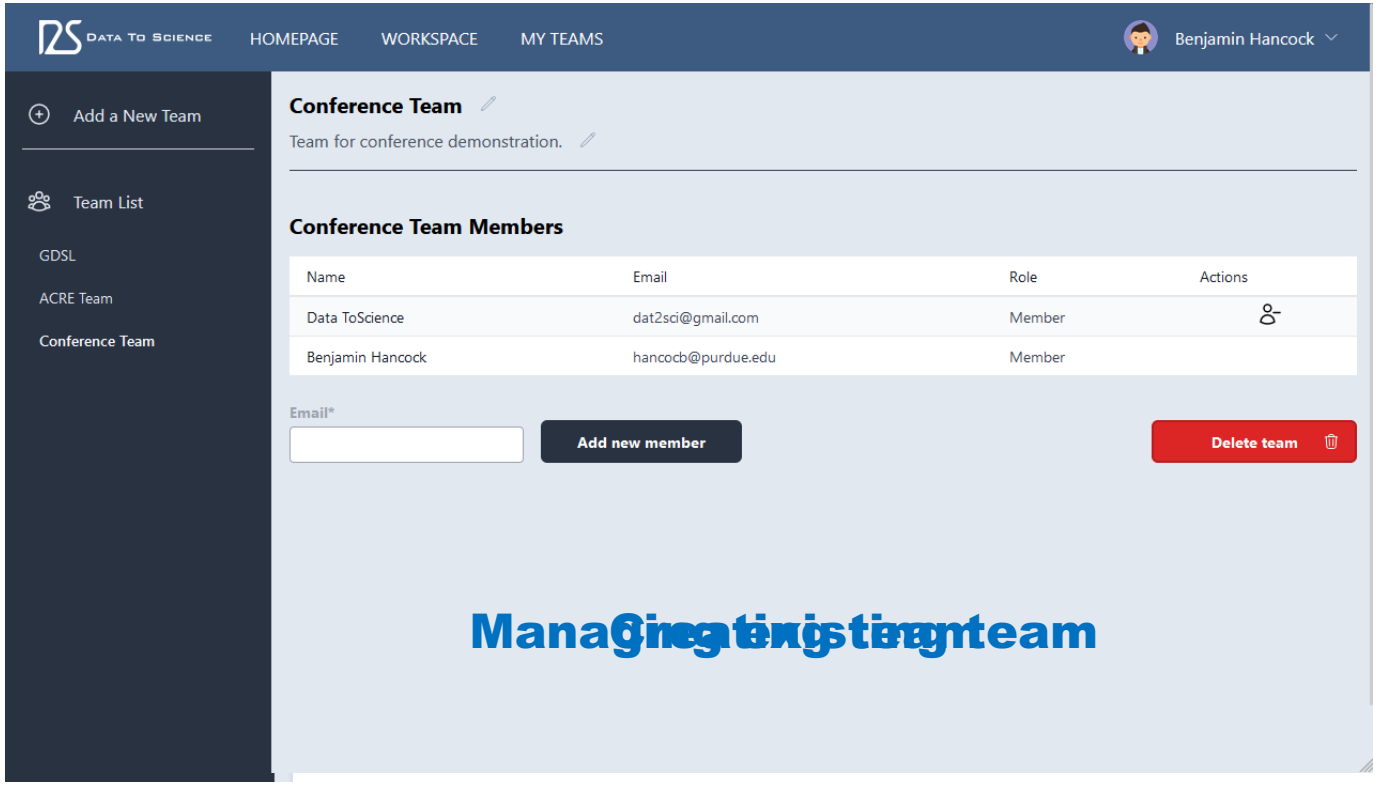
PURDUE UNIVERSITY

TEAMS

Create or join existing team

Team members gain access to a team's project data

Only team owner can add and remove members, and delete the team



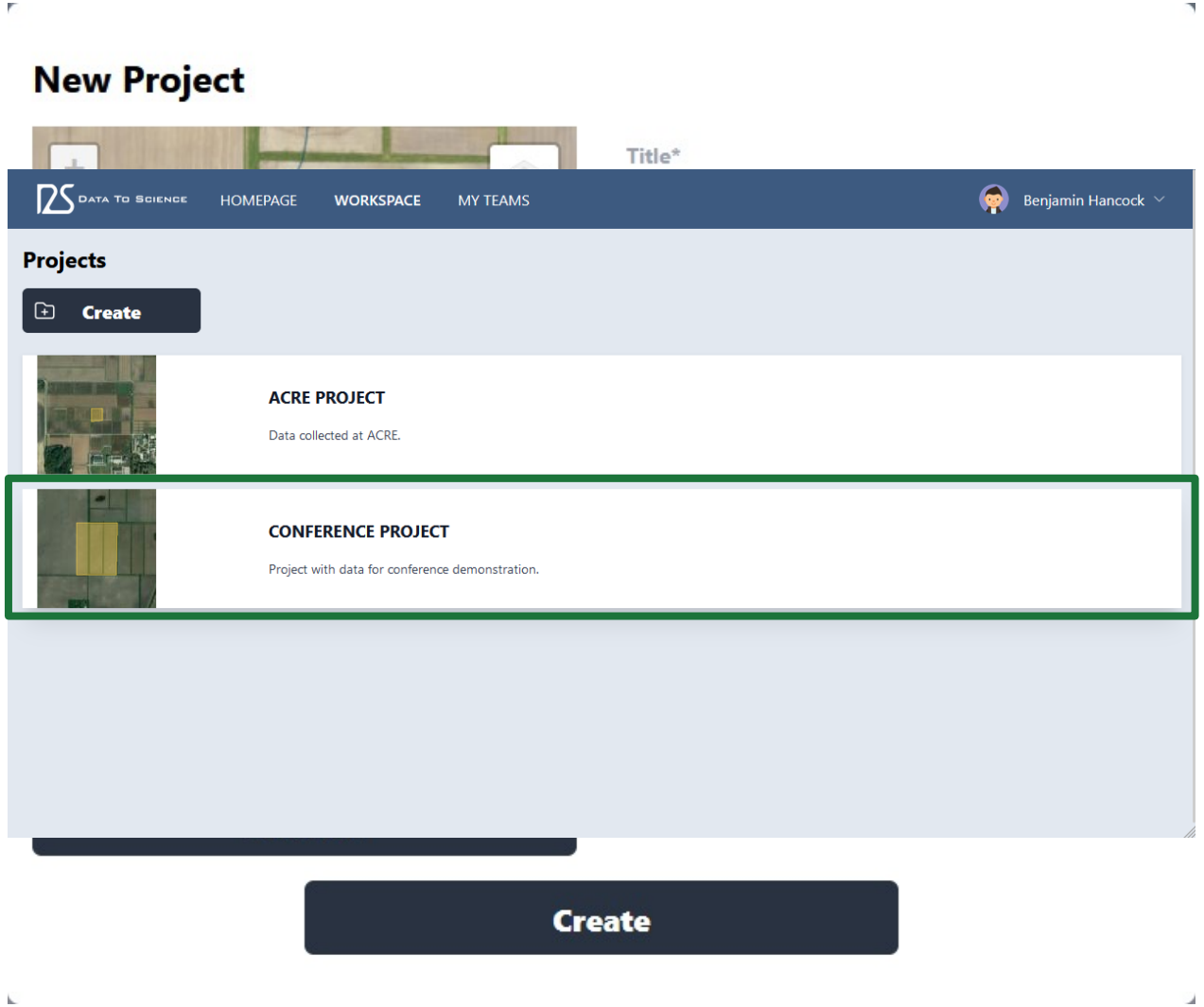
Managing existing team

PROJECTS

Anyone can create a project

Projects can be associated with an existing team

Must provide title, description, and field boundary



FLIGHTS

Multiple flights can be added to a single project

Upload raw data and data products

- Ortho GeoTIFF
- DSM GeoTIFF
- Point Cloud
- Other

DATA FORMATS OPTIMIZED FOR WEB

Uploaded GeoTIFFs (.tif) converted to Cloud Optimized GeoTIFF (COG) format

- Remains single file – no need to generate static tiles at different zoom levels
- HTTP GET range requests can be used to fetch parts of file
- Integrate easily with dynamic tile servers, such as TiTiler, for improved mapping performance

Uploaded Point Clouds (.las) converted to Entwine Point Tile (EPT) format

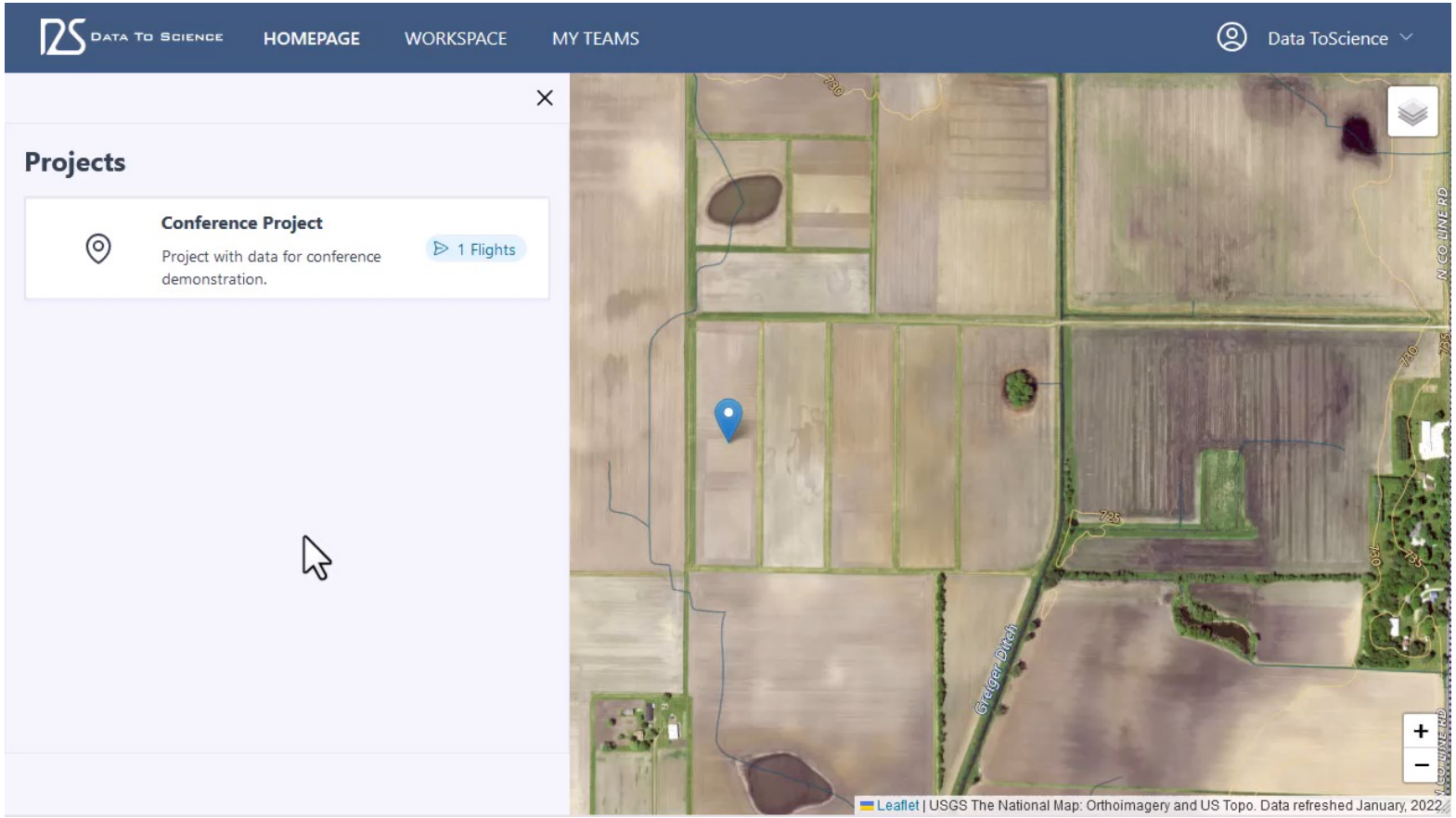
- Combination of JSON metadata and binary point data
- EPT can be rendered in real-time in web browsers with Potree

VISUALIZING PROJECT DATA

Projects, flights, and data products displayed in left pane

COGs are ingested by local TiTiler server and rendered on map

User specific symbology props

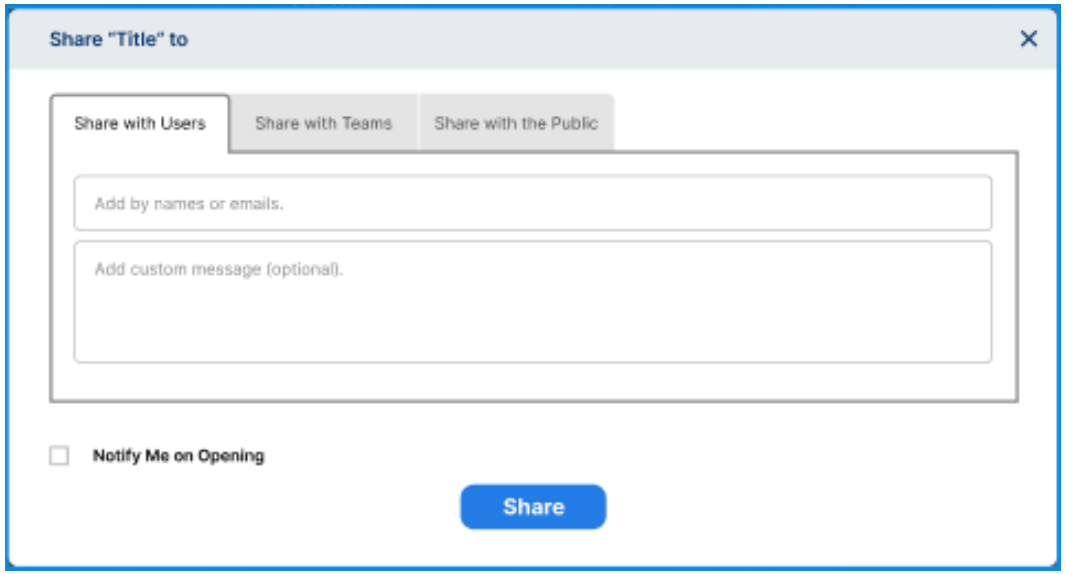


SHARING MAPS

Team members have access to the same data products, but not the same symbology settings

Share specific symbology settings for a data product with other users

Share visualization with non-account holders



DEMONSTRATION

GIS day
INSPIRE THE WORLD THROUGH GIS

PURDUE
UNIVERSITY

ACKNOWLEDGEMENTS

This research was supported [in part] by the intramural research program of the U.S. Department of Agriculture, National Institute of Food and Agriculture, Grant (USDA-NIFA award 2022-70412-38454).

The findings and conclusions in this preliminary presentation have not been formally disseminated by the U. S. Department of Agriculture and Should not be construed to represent any agency determination or policy.

Purdue Plant Science 2.0 Initiative

Purdue Digital Forestry

GIS day
INSPIRE THE WORLD THROUGH GIS

PURDUE UNIVERSITY

THANK YOU

Development Team

Jinha Jung
Project Lead, jinha@purdue.edu

Minyoung Jung
Project Manager, jung411@purdue.edu

Cheryl Zhenyu Qian
Lead UI/UX Designer

Ben Hancock
Senior Web Developer, hancocb@purdue.edu

Na Zhuo
UI/UX Designer

GIS day
INSPIRE THE WORLD THROUGH GIS

PURDUE UNIVERSITY