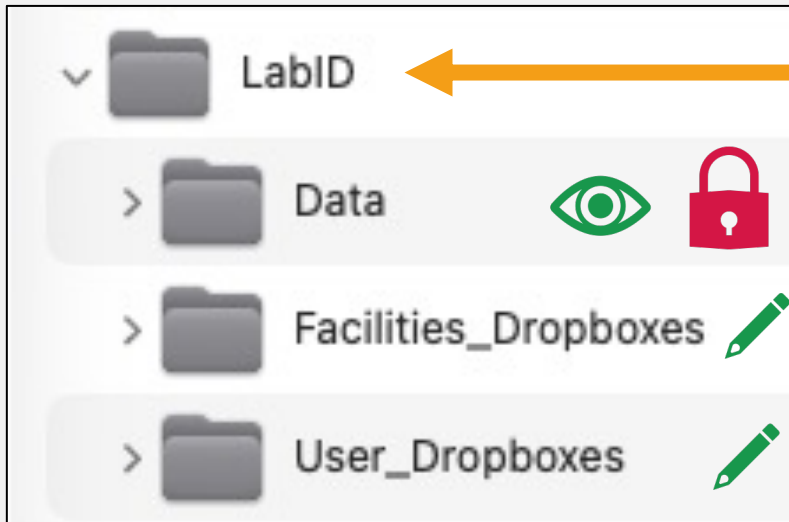


Storage Organization in LabID



Charles Girardot
Multimodal Open Data Integration Support (MODIS)
Data Science Center

LabID Storage Volume



A Storage Volume – a Folder

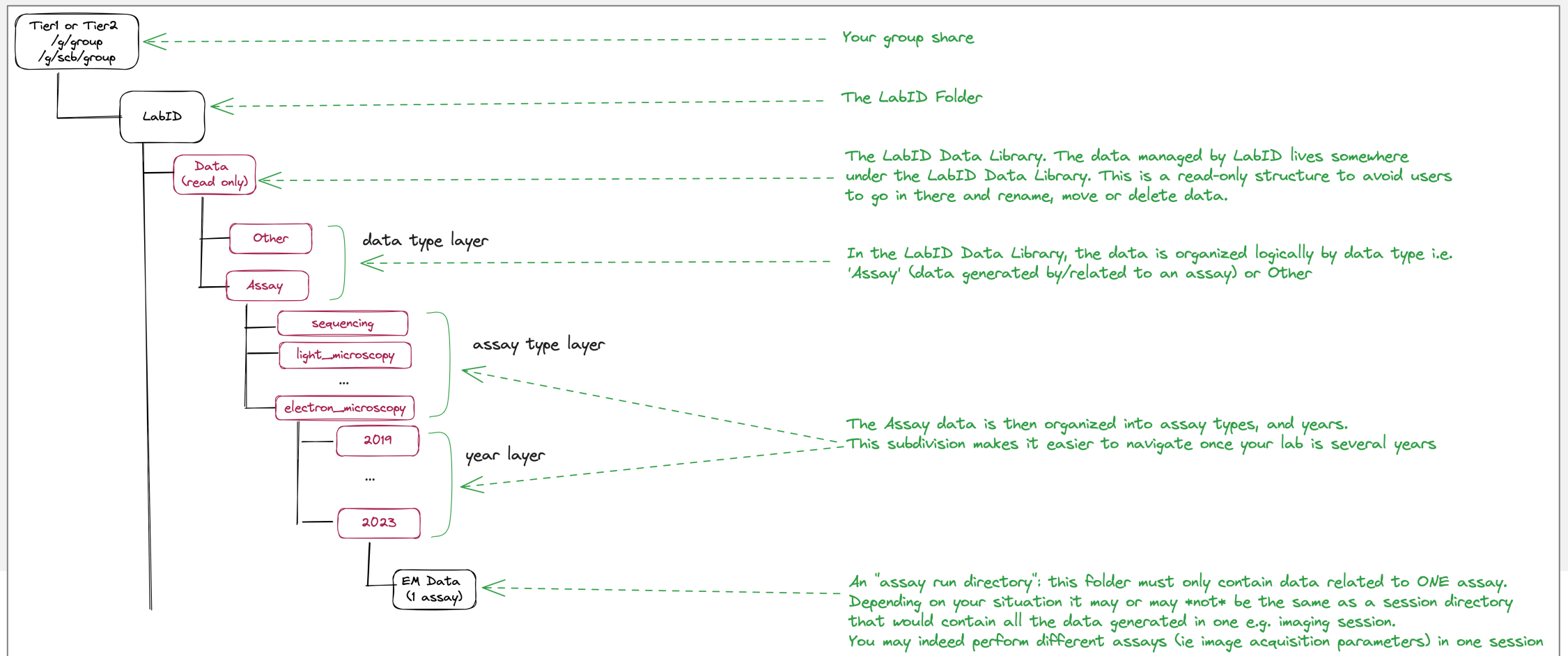
Data registered in LabID lives in the read-only **Data** folder

Writable “dropbox” folders allow to **stage** data for registration

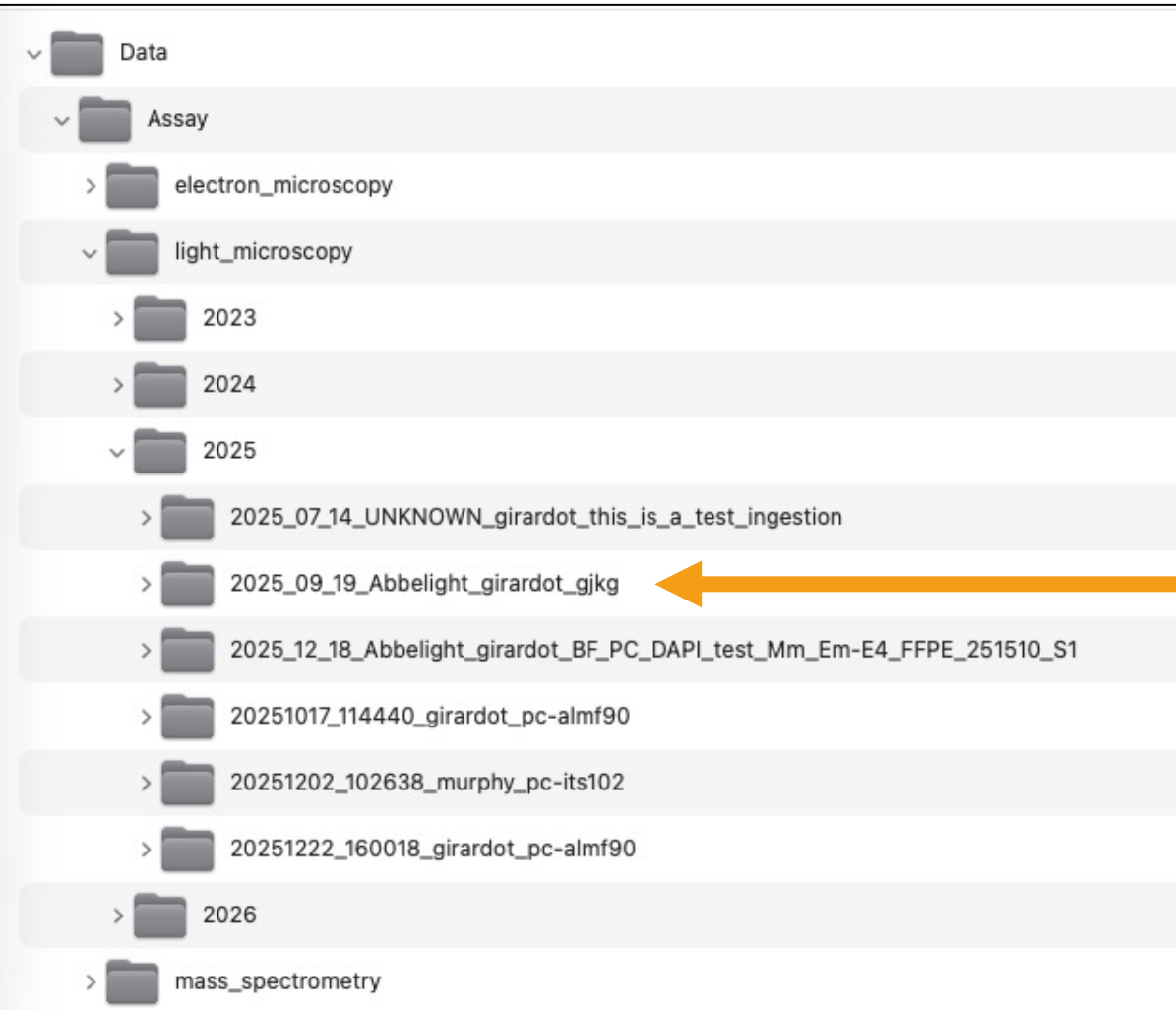
- The data lives on network drives - accessible *locally*
- Group-based **Storage Volumes**
 - Multiple Storage Volume per group possible – one is the **primary** one
- S3 Storage Volume available but only to share a dataset copy stored locally
- Admins create Storage Volumes using the AdminUI

Storage Volume Organization

- Acquisition oriented organization as opposed to project based
- Raw Data goes in **Assay** subfolder – sub-organized by technology & year (of registration)
- Derived Data in **Other** subfolder – sub-organized by year (of registration)



The “Run Folder” concept



- Run Directory contains the data generated during an instrument run (imaging session)
- It can contain the data for multiple assays

Data is imported using the registration wizard

NB: **Import datasets**: for derived data only (no assay context)



v26.9.0

Import datasets ^

Import raw datasets

Import datasets

Your dropboxes

These are run folders
to import

DATASET LOADER

Start Select Data Assay details Build datasets Verify Assign samples

Select Datasets Choose your dropbox → Choose the run folder → Select datasets

/g/funcgen/LabID/User_Dropboxes/girardot

Run Folders Computational Support Genom...

- 2021-08-05_SMF_GW_emb2-4_nB/
- 231213_allelicPanel_option1_onedirpersample/
- 231213_allelicPanel_option1_original_cleaned/
- bigwigs/
- E214_HKdoubleKI_Dy634-dUTP/
- E215_HKvarious_AF647-dUTP/
- fvincent_DKTCW_Microscopy/
- Manuel_Demo_20210826-130300/
- Matt_2022_12_22/
- Matt_2023_02_02/
- Matt_2023_02_14/
- Matt_2023_02_22/
- Matt_2023_03_10/

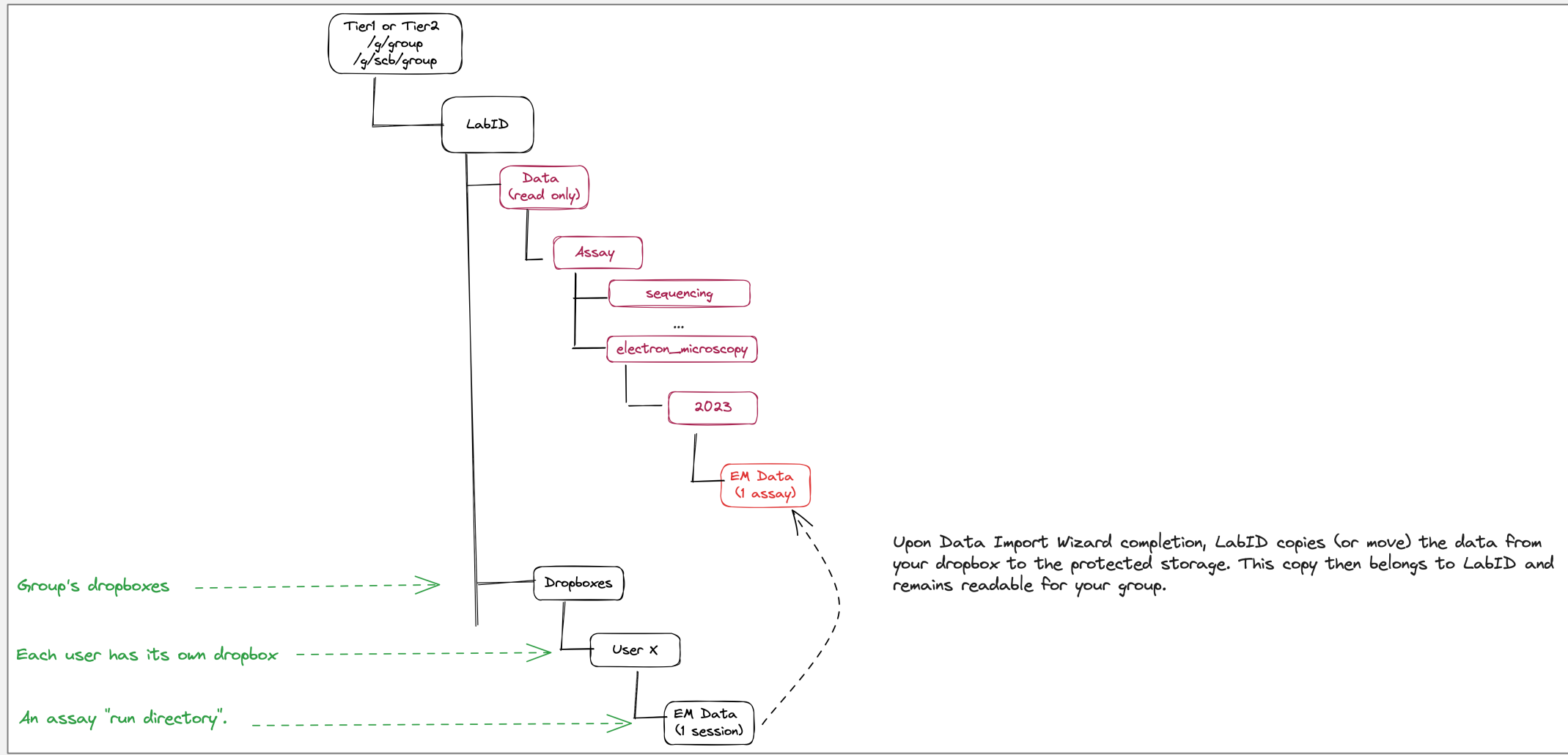
Transfer Strategy -- select -- Please select a transfer strategy

Select individual datasets, i.e. files and/or folders

- 2021-08-05_SMF_GW_emb2-4_nB
 - emb2-4

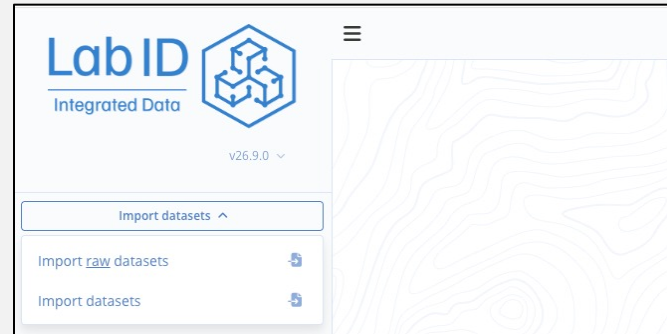
Continue

Data Ingestion Using the Dropbox

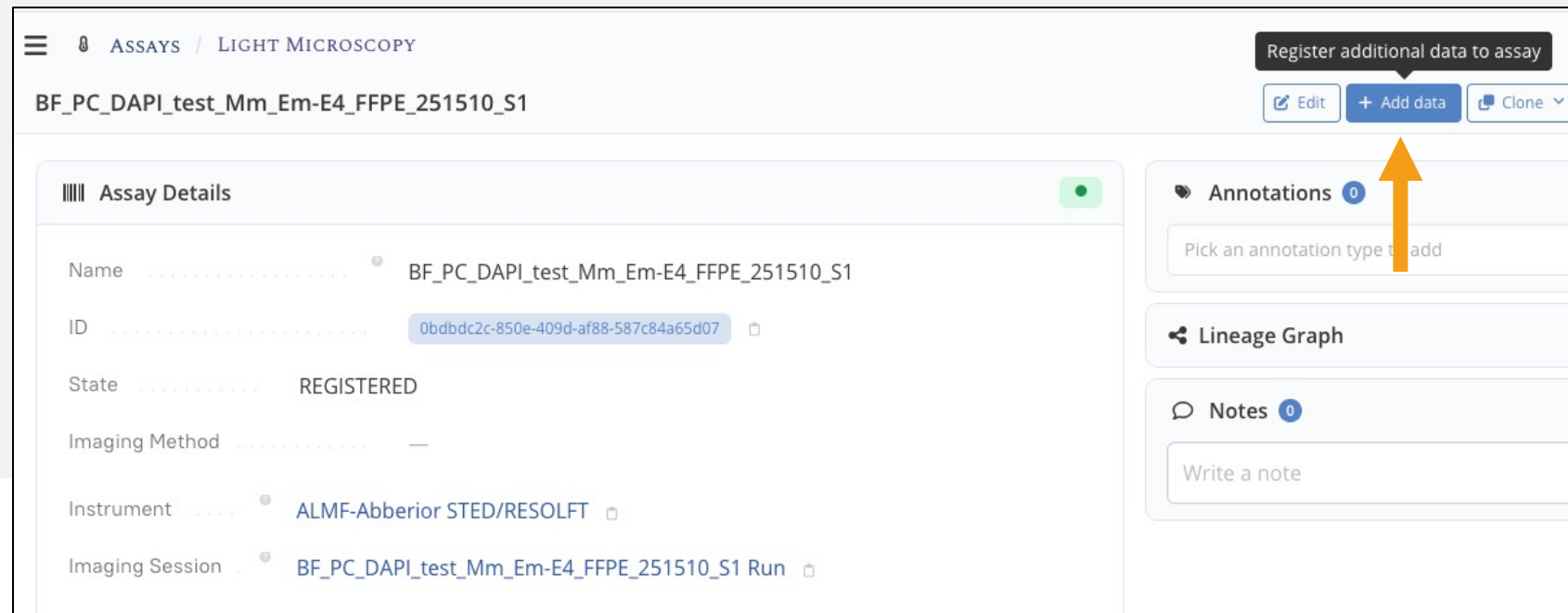


Loading additional data in existing run folders

- Importing Data from global Import Datasets, always end up in a new run folder



- Import data in an existing run directory using the **Add Data** from the Assay Page



Thank you

Multimodal Open Data Integration Support (MODIS) part of EMBL Data-Science Center



Charles Girardot

Head of Multimodal Open Data Integration Support

ORCID: [0000-0003-4301-3920](https://orcid.org/0000-0003-4301-3920)



Laurent Thomas

Postdoctoral Fellow

ORCID: [0000-0001-7686-3249](https://orcid.org/0000-0001-7686-3249)



Md Nayeem Reza

Senior Software Engineer

ORCID: [0000-0003-2068-5812](https://orcid.org/0000-0003-2068-5812)



Matthias Monfort

Web Developer / Workflow Management

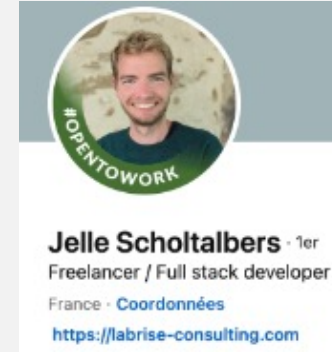


Marco Wetter

Full Stack Developer

Contact : modis@embl.de

Chat: <https://chat.embl.org/embl/channels/labid>



Join LabID on Slack