**Overview**

The goal of this project is to construct a data ingestion pipeline like the following diagram shows:

A diagram of a diagram

Description automatically generated

, where the *DccReporter* and *KOMP\_Reporter* are the two main consumers of the library *Pfs-odata.*

*Pfs-odata* is a Python-based SDK/API Wrapper to help pull data out of PFS (ThermoFisher Platform For Science) is a LIMS system where the Jax  
phenotyping Center stores data collected from their experiments. The main purpose of this library is to provide an easy and flexible way to interact with PFS by hiding the complexity of PFS’s embedded OData RESTful API.

The plan is to make a Python module/library and publishes to public/private PyPI account or docker hub, which will allow users to install via pip or docker. *Pfs-odata* library will allow users to pass in properties/attributes of an entity in PFS they want to select; filtering conditions they would like to apply when searching for certain experiment/samples; orders that they would like to sort the query result as well as create new record and modify existing ones.

The library contains the following part:

1. ***Pfs\_session***
2. ***Models***
3. ***Commons***
4. ***Pfs\_Exceptions***

***Pfs\_session***.***py*** handles all HTTP requests created by users(GET, PUT, POST, DELETE), it contains all functions end users will call when they composed their own scripts/apps.

***Models.py*** contains data model of pfs entities, including Sample, Sample Lot, Assay, Experiment, Strain etc.

***Commons.py*** stores constants and functions that are widely used in ***Pfs\_session.py***.

Finally, ***Pfs\_Exception.py*** is customized exceptions for bugs/issues occurred during the execution of functions in ***Pfs\_session.py***.

GitHub Repository:

<https://github.com/TheJacksonLaboratory/pfs_odata>.

Progress

On week of 08-31-2023, I have implemented functionalities to get data related specific assay, experiment sample and sample lots in json format. It has not been tested yet and more functions are pending implementation.

On 09-11-2023, I have refactored the code and add logs to the code. Also, I created a repository on GitHub, for more info, see <https://github.com/TheJacksonLaboratory/pfs_odata>.

On 09-15-2023, data model of mouse sample and mouse sample lot are created. Next step will be test the GET functions I created and make sure they always work.

On 09-22-2023, end point for jax\_strain has been built. User now can filter the mice based on strain’s name, barcode, jr\_number etc.

Starting week of 10-09-2023, I will start working on generating specimen files while exploring other end points of API for getting data from PFS.