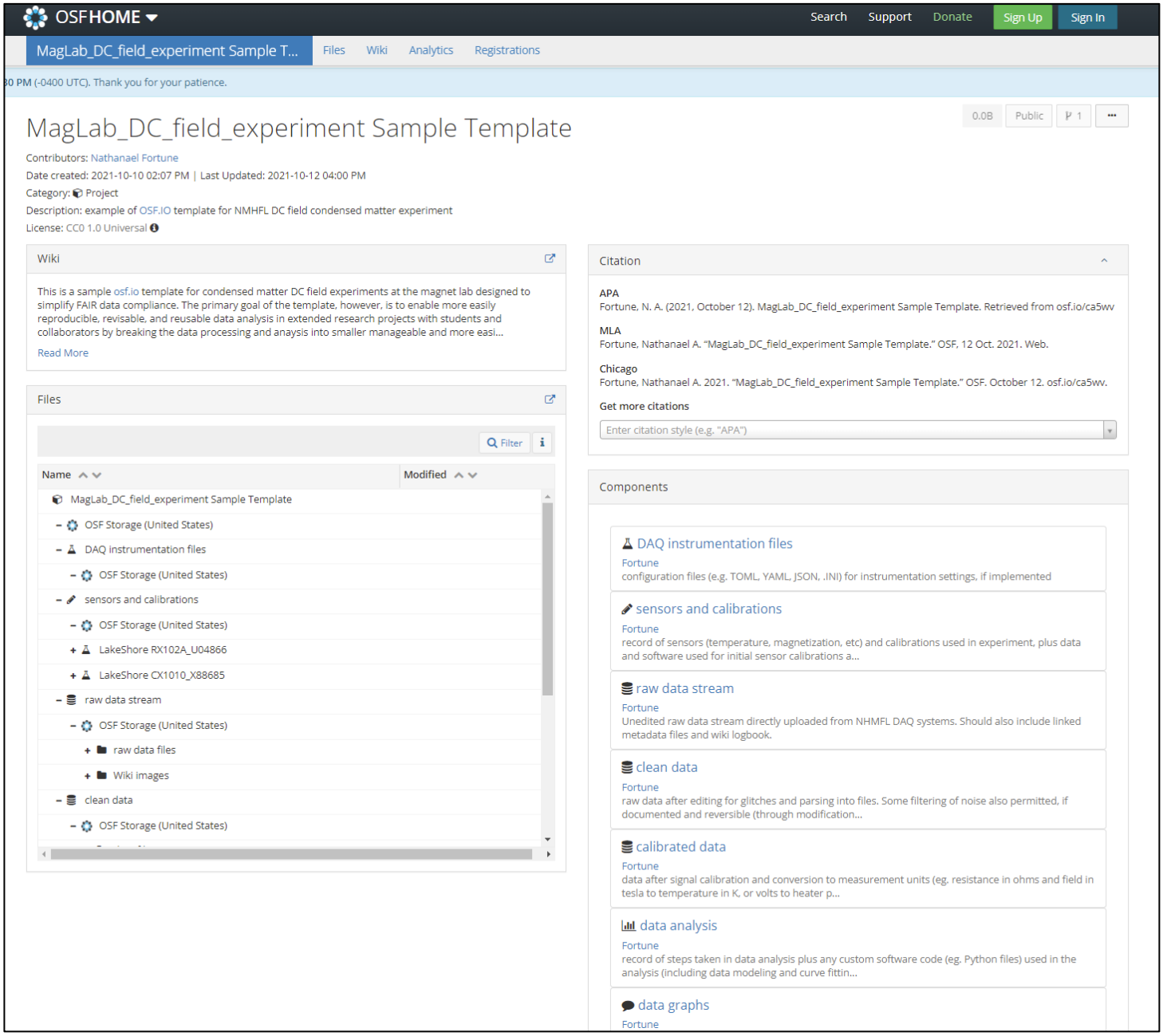
https://osf.io/ca5wv/?view_only=93b85aea12ca4e50b396239a53aa4bb9



and Oct 4, 2021 9:30 PM (-0400 UTC). Thank you for your patience.

Home

Toggle view: View Compare

Menu

- Project Wiki Pages
  - Home
- Component Wiki Pages

View Wiki Version: (Current) Nathanael Fortune: 2021-10-12 20:00:00+00:00 UTC

This is a sample osf.io template for condensed matter DC field experiments at the magnet lab designed to simplify FAIR data compliance. The primary goal of the template, however, is to enable more easily reproducible, revisable, and reusable data analysis in extended research projects with students and collaborators by breaking the data processing and anaysis into smaller manageable and more easily documentable steps. This also requires traceability of samples and calibrations.

In my opinion, three small initial steps, three slightly larger follow up steps, and one still later large step could be taken to greatly simplify (and better automate) this process if osf.io were to be adopted as the lab's standard data storage tool.

First, the data files generated by magnet-lab and user-supplied data acquisition programs could be modestly modified to include identifying column names, instrument units (eg. V), gains (eg. 1 V/uV or 10 nA/V), and, optionally, measurement units (eg. uV or nA) . There should be an easy way to enter this information at the time of data collection (e.g. two additional text fields beneath the ones that ask for the user to supply a name for each column of collected data and the multiplicative gain to be applied to each column).

In a follow up step, the lab could build automatic uploading of these data files to the experimenters designated osf.io website, provided it followed some standard format for the data storage location.

Second, the general purpose plain text comment fields provided in these programs could be split off from the file headers and instead saved as Github-flavored Markdown format text file the format used by OSF.io for wiki pages. The data file and markdown file could share the same file header.

In a follow up step, this too would be automatically uploaded to the experimenter's osf.io experiment file, allowing the user to add additional comments and information, photos, a and graphs without needing to modify the header of the data file accompanying the wiki.

Third, the lab could arrange for the automatic upload of the following experimental metadata (as a standalone wiki file for the entire experiment, such as the wiki you are reading now), once the experimenter enters the sample specific information. There would of course need to be an option for multiple samples and sensors

1. experimenter ID (e.g. ORCID ID, request as part of application for magnet time)
2. magnet system ID (with link to field calibration)
3. cryostat/probe ID (with link to drawings)
4. probe temperature sensor ID (with link to database of calibrations and calibration history)

In a follow up step, the lab would provide free sample, sensor, and instrument tracking databases so that critical sample information and user supplied sensor calibrations could be automatically added to the experimental wiki.

Finally, in a step requiring the largest lift, the lab could ultimately establish and disseminate a lab-standard HDF5 data file format (and naming convention?) in which the above metadata (such as units) plus instrumentation settings are included as file *attributes*. To make this useful, the lab should also create how-to guides on how to access data and metadata stored in these lab-standard HDF5 files using standard data analysis and acquisition apps such as IGOR Pro, Python, and LabVIEW.

PS: The most common question mag lab user committee members have had besides how to generate shareable data is when to share data. Here is the guidance from the Magnet Lab, courtesy Lissa Anderson during the morning FAIR data workshop session:

> MagLab and NSF policies regarding when data are expected to be made publicly available are outlined in our data management plan.
> https://nationalmaglab.org/about/fair-data
>
> In sum, the NSF expects that data be made publicly available within 1 year of publication of an associated manuscript. If an associated manuscript is not published, the MagLab plan states that data should be made publicly available within 3 years of the last time the project was assigned magnet time, but concessions are made for issues regarding IP or publication in process. Please refer to the DMP in the link above.