

# Data Management in ct.qmat

– An overview/introduction –

Florian Goth, Christoph Brüning, Fakher Assaad

Universität Würzburg  
Institut für Theoretische Physik und Astrophysik

01. July 2021



**ct.qmat**

**Complexity and Topology  
in Quantum Matter**



**TECHNISCHE  
UNIVERSITÄT  
DRESDEN**

- Two sites: Würzburg, Dresden
- Diverse groups across theory and experiment
- Established in 2019
- Research: Novel topological materials, quantum magnetism, topological photonics, tailoring topological functionality.

Simulations/Experiment are a cornerstone of good scientific practice and are facilitated by data management solutions. Efficient data management allows to document the workflow leading to published data and facilitates open access (OA) publication forms where all primary material is accessible. In addition, the many locations within the cluster render collaborative tools a necessity. Bringing both requirements together we foresee the need for collaborative data management solutions that use established tools from the community while at the same time leveraging the latest in storage technology.

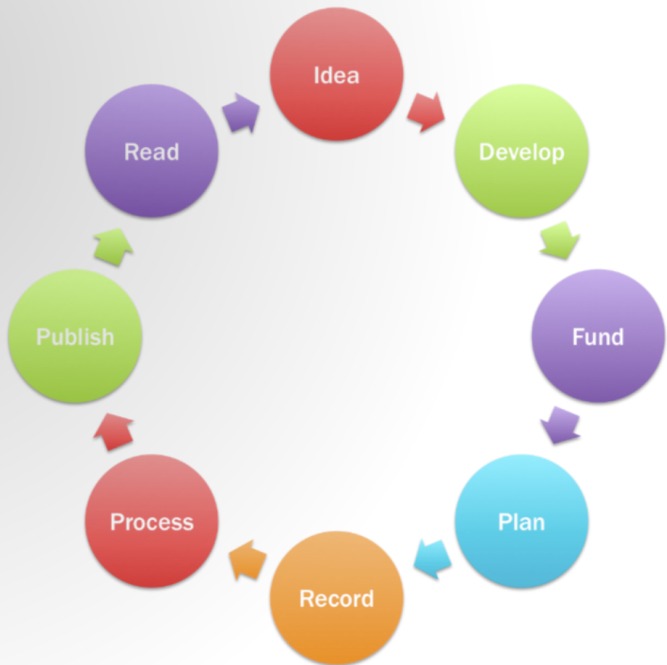
## What is important

- Delocalized
- Collecting Data
- Working on Data

->Collaborative Data Management

## Political/external Pressures

- Open Access Policy of DFG
- Increased pressure for proper data management plans from funders
- Publishers require open access to certain data
- Reproducibility





## What do we have to build

-> Structures/Software that enables:

- collaborative decentralized work on data from various sources
- longterm, flexible storage
- in a way that enables the openness of the research cycle
- makes analysis steps reproducible
- and in a manner that satisfies expectations on data management plans
- open, extensible infrastructure should enable flexible services on this long time scale

## User Needs

Plan/Design Analyze

Publish ReUse/Reproduce

---

Education  
Cooperation  
Representation

Product Layer

Storage Retrieval Analysis SW Dev Organisation

---

Webservices

---

Compute

Storage

## Summary

- The full cycle is available in software on the web -> Delocalized
- Almost integrated
- A researcher can publish projects at his will -> Full control
- Any Service can be executed on top of the basic Infrastructure(Kubernetes + Object Store). -> Flexible

## Status

- gitlab present for a long time and we work on onboarding scientists ✓
- Set up Storage Cluster ✓
- -> Set up Kubernetes
- Connect gitlab
- Add Services: DataVerse, elabFTW, JupyterHub/BinderHub, OSF.io

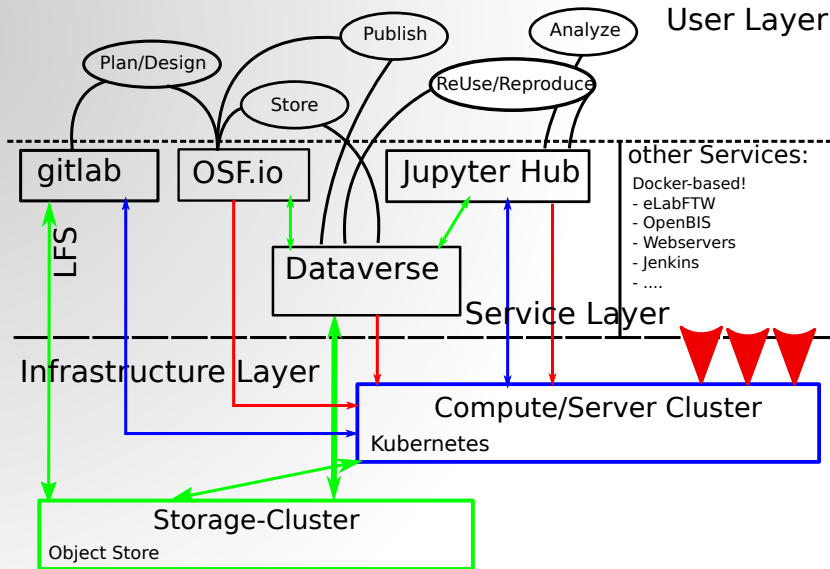


## Notes

- Infrastructure is all opensource -> No vendor lock in.
- The provided tools are all open source, hence researchers can set up open workflows with public tools.
- We educate people about certain tools (software-carpentry.org) -> Behavioural changes.
- In a perfect world workgroups have their whole structure in an ELN and hence the network can be made openly visible.
- Data curators? Maybe for project based research?
- But basically this structure is meant to be live.

### Ideas for the university

- What happens after the lifetime of DFG funded project duration entities?
  - The associated infrastructure could be kept alive in a read-only mode for the forthcoming 10 years -> university guarantees, foundations.
  - open collaborative tools vs. Arbeitsrecht?
  - citable resources, DOIs?
  - Support for this infrastructure!
- 
- Only a platform. Services/analysis need to be done by domain scientists.
  - But Quality of data and quality of scripts is only as good as the knowledge of this scientist.
  - Quality of initial data needs to be kept high -> FDM
  - Web-based research will necessitate the use of scripts. This quality needs to be fostered -> specialized research software engineers.



## Short Explanations

- Object Store: Abstraction Layer for storage.
- Kubernetes: Abstraction Layer for providing services controlled via software.
- gitlab: feature rich version control tool for Software Development.
- DataVerse: Research Data Repository.
- OSF.io: Open Science Framework; Integrated Platform for managing and sharing your research.
- Jupyter Hub: A platform for providing computational environments.

Student Projects supported by VPAL Dataverse (Vice Provost for Advances in Learning)

[Harvard Dataverse](#) > [VPAL Research Dataverse](#) > [VPAL On Request Dataverse](#) > [Student Projects supported by VPAL Dataverse](#) >

## Example Uploaded Dataset

Version 1.0



Williams, Joseph, 2015, "Example Uploaded Dataset", <https://doi.org/10.7910/DVN/8DHGAH>, Harvard Dataverse, V1

[Cite Dataset](#)

[Learn about Data Citation Standards.](#)

Contact Owner

Share

Dataset Metrics

0 Downloads

Description

This dataset is empty; it is a test and example for other users

Subject


Other

Files

Metadata

Terms

Versions

 Export Metadata

Citation Metadata

Dataset Persistent ID

doi:10.7910/DVN/8DHGAH

Publication Date

2015-11-07

Title

Example Uploaded Dataset

Author

Williams, Joseph

Contact

Use email button above to contact.

Williams, Joseph

Description

This dataset is empty; it is a test and example for other users

Subject

Other

Depositor

Williams, Joseph

File Edit View Run Kernel Tabs Settings Help

demo /

Files

- data
- notebooks
- TOSA\_Data
- Myenv
- jupyterlab-vizies.pdf
- jupyterlab.mtd
- Lorenz.ipynb
- Lorenz.py
- matplotlib\_jupyter.mtd

## The Lorenz Differential Equations

Below we start, we import some preliminary libraries. We will also import (below) the accompanying `T Lorenz.py` file, which contains the actual solver and plotting routine.

```
[ ]: %matplotlib inline
free jupyterlab import interactive, fixed
We explore the Lorenz system of differential equations:
```

$$\begin{aligned}\dot{x} &= \sigma(y - x) \\ \dot{y} &= \rho x - y - xz \\ \dot{z} &= -\rho z + xy\end{aligned}$$

Let's change (i.e.  $\sigma, \rho, \gamma$ ) with `ipywidgets` and examine the trajectories.

```
[ ]: free Lorenz import jupyterlab, Lorenz
with interactive as Lorenz, Lorenz.set_defaults(
    sigma=10, rho=28, gamma=8)
w
```

For the default set of parameters, we see the trajectories swirling around two points, called attractors.

The object returned by `interactive` is a `Widget` object and it has attributes that contain the current result and arguments:

```
[ ]: w.x_t = w.result
[ ]: w.observe
[ ]: After interacting with the system, we can take the result and perform further computations. In this case, we compute the average positions in x, y and z.
[ ]: xyz_avg = w.x_t.mean(axis=0)
[ ]: xyz_avg.shape
Creating Histograms of the average positions (across different trajectories) show that, on average, the trajectories swirl about the attractors.
```

```
[ ]: free matplotlib import pyplot as plt
plt.hist(xyz_avg[:,0],10)
plt.title('Average x(t)')
plt.hist(xyz_avg[:,1],10)
plt.title('Average y(t)')
```

JupyterLab Reference

Search Docs

### GETTING STARTED

- Overview
- Installation
- Starting JupyterLab
- Reporting an Issue
- Frequently Asked Questions (FAQ)
- JupyterLab Changelog

### USER GUIDE

The JupyterLab Interface

- JupyterLab UI/UX
- Working with Files
- Text Editor
- Notebooks
- Code Consoler
- Terminals
- Managing Kernels and Terminals
- Command Palette
- Documents and Kernels
- File and Output Formats
- Extensions
- JupyterLab on Kubernetes
- Exporting Notebooks

### DEVELOPER GUIDE


- General Guidelines
- Extension Developer Guide
- Customizing Extension

Docs > JupyterLab Documentation

[Jupyter](#) | [Edit on GitHub](#)

## JupyterLab Documentation

JupyterLab is the next generation web-based user interface for Project Jupyter. Try it on Binder. JupyterLab follows the [Jupyter Community Guides](#).



### Getting Started

- Overview
- Installation
- Starting JupyterLab
- Reporting an Issue
- Frequently Asked Questions (FAQ)
- JupyterLab Changelog

### User Guide

- The JupyterLab Interface



# JHU Electronic Lab Notebook Template

Public 14 ...

Contributors: Reid Boehm, Chen Chiu, Dave Fearon

Affiliated Institutions: [Johns Hopkins University](#)

Date created: 2017-03-15 08:36 PM | Last Updated: 2018-04-13 06:09 PM

Category: Project

Description: A template for researchers to use to structure an OSF project as an Electronic Lab Notebook

License: CC0 1.0 Universal

## Wiki

## JHU Electronic Lab Notebook Template

by

[JHU Data Management Services](#)

JHU DMS created this template to get you started using OSF as a Lab Notebook. Each component within uses the wiki space to provide a brief explanation about suggested content and in some cases links to other helpful resources.

Unlike some electronic lab notebooks, this design works mainly as a "binder" for gathering and access...

[Read More](#)

## Files

Name

Modified

- JHU Electronic Lab Notebook Template
  - OSF Storage (United States)
  - Data Management Plan
  - Literature
  - Protocols
  - Standard Operating Procedures
  - Materials and Methods
  - Notes and Documentation
  - Analysis
  - Results
  - Manuscripts, Reports, Visual Presentations

## Citation

### Components

#### Data Management Plan

Boehm, Chiu & Fearon

A space to include your data management plan

#### Literature

Boehm, Chiu & Fearon

A space to collect associated literature that relates to your experiment.

#### Protocols

Boehm, Chiu & Fearon

A space to include the documentation of the governing regulations and processes of the experiment/project

#### Standard Operating Procedures

Boehm, Chiu & Fearon

A space to store the standard operating procedures specific to your lab or experiment

#### Materials and Methods

Boehm, Chiu & Fearon

A space to include information about the materials/equipment and the methods used to carry out the experiment/project

#### Notes and Documentation

Boehm, Chiu & Fearon

A space to keep notes from each work session and documentation about the data and related materials produced

#### Analysis

Boehm, Chiu & Fearon

A space to keep materials related to analysis of the data collected

#### Facile



# Lab Example

Public | 18 | ...

Contributors: Ian Sullivan, Courtney K. Soderberg, Brandon Thorpe

Affiliated institutions: Center For Open Science

Date created: 2017-06-06 06:00 PM | Last Updated: 2019-12-18 10:28 PM

Category: Project

Description: This project demonstrates one way the OSF can be used to create a space for a lab to share materials and research.

License: CC0 1.0 Universal

## Wiki

This is an example project showing how the OSF might be used by a lab to create a shared lab space, share lab standards and resources, and collate the work that is being done by individuals/groups within the lab.

## Getting started

Welcome to the lab! This wiki contains the steps to take at key moments in the life of a lab like when someone joins the lab, when you start a new experiment or when the p...

[Read More](#)

## Files

Filter		i
Name	Modified	
Lab Example		
OSF Storage (United States)		
New Experiment Template		
Lab Documents		
Lab Meetings		
Research		
OSF video for Cambridge ELN pilot		

## Citation

### Components

#### New Experiment Template

Sullivan, Soderberg & Thorpe

Start all new experiments by forking this section.

#### Lab Documents

Sullivan, Soderberg & Thorpe

Internal documentation and lab reference material should go here.

#### Lab Meetings

Sullivan, Soderberg & Thorpe

Discussion notes, articles, and other internal musings.

#### Research

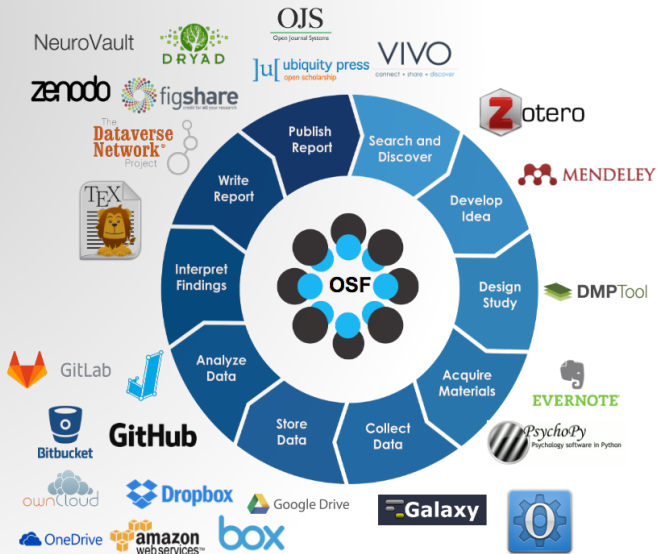
Sullivan, Soderberg & Thorpe

Link all your experiment projects here.

#### OSF video for Cambridge ELN pilot

Soderberg







## Demo Add-Ons

Contributors: Chris Seto, Brian A. Nosek, Johanna Cohoon, Alex Schiller, Joshua M Carp, Robert Liebowitz, Jeffrey R. Spies, Sam Portnow, Harry Rybacki, Denise Holman, Melissa Lewis, Jacob Rosenberg, Lyndsy Simon, Steven Loria, Nan Chen, Tanesha Hudson, Austin Macdonald, Andrew Sallans, Larry Banner, Nicole Pfeiffer, Matthew Spitzer, Fitz Elliott

Affiliated institutions: Center For Open Science

Date created: 2014-02-26 06:03 PM | Last Updated: 2019-05-10 03:41 PM

Category: 🧩 Other

### Wiki

The Open Science Framework (<http://osf.io/>) develops add-on connections to services which make it easier to keep doing the things that you already do and to identify new tools that will make your life easier and your research better. Simply go to "Settings", select an "Add-On", enter the required information for that service, and then you'll be connected.

As an example, giant files no longer have...

[Read More](#)

### Files

Download as zip

Filter



Name ^ v

Modified ^ v

🧩 Demo Add-Ons

+ 📁 Dropbox: demofiles

+ 📁 GitHub: AndrewSallans/demofiles (master)

+ 📁 Google Drive: Demo Files

+ 📁 OSF Storage (United States)

+ 📁 Amazon S3: osfdemofiles (US Standard)

### Citation

#### Tags

CSV DOCK IPYNB JPG PDF PNG PY R SAV XLSX

#### Recent Activity

- 👤 Brian A. Nosek changed permissions for Demo Add-Ons 2019-05-10 03:41 PM
- 👤 Brian A. Nosek reordered contributors for Demo Add-Ons 2019-05-10 03:41 PM
- 👤 Brian A. Nosek added Fitz Elliott as contributor(s) to Demo Add-Ons 2019-05-10 03:40 PM
- 👤 Brian A. Nosek removed Fabian von Feilitzsch as contributor(s) from Demo Add-Ons 2018-09-17 10:48 PM
- 👤 Andrew Sallans copied ExampleCSV.csv in GitHub to ExampleCSV.csv in Amazon S3 in Demo Add-Ons 2016-07-12 10:38 PM
- 👤 Andrew Sallans added Center For Open Science affiliation to Demo Add-Ons 2016-04-28 02:24 PM

< 1 2 3 4 ... 33 >