StraboExperimental / LAPS Development
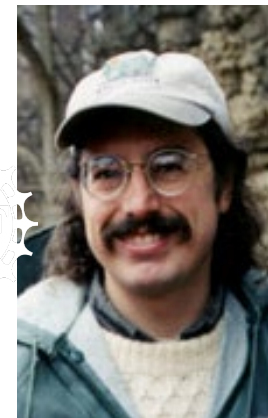
Matej Pec
Ulrich Mok

Andreas Kronenberg
Noah Phillips
Julie Newman
Hannah Cunningham

Basil Tikoff
Jason Ash

Digital Repository for Experimental Data - Goal

Input of Experimental Data

Searching and Downloading Experimental Data
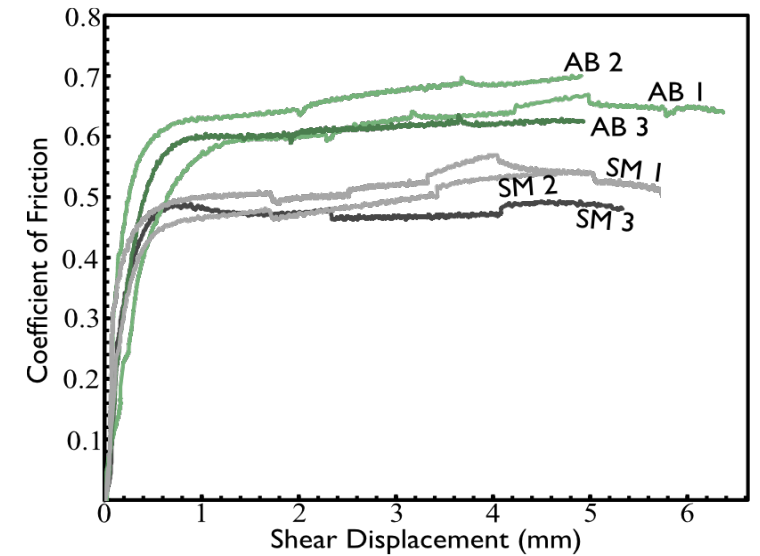
Plotting Experimental Data from Multiple Sources

Digital Data Repositories - Challenges

A. **Technical Challenges**
- Permanent Physical Storage
- Database Maintenance and Longevity
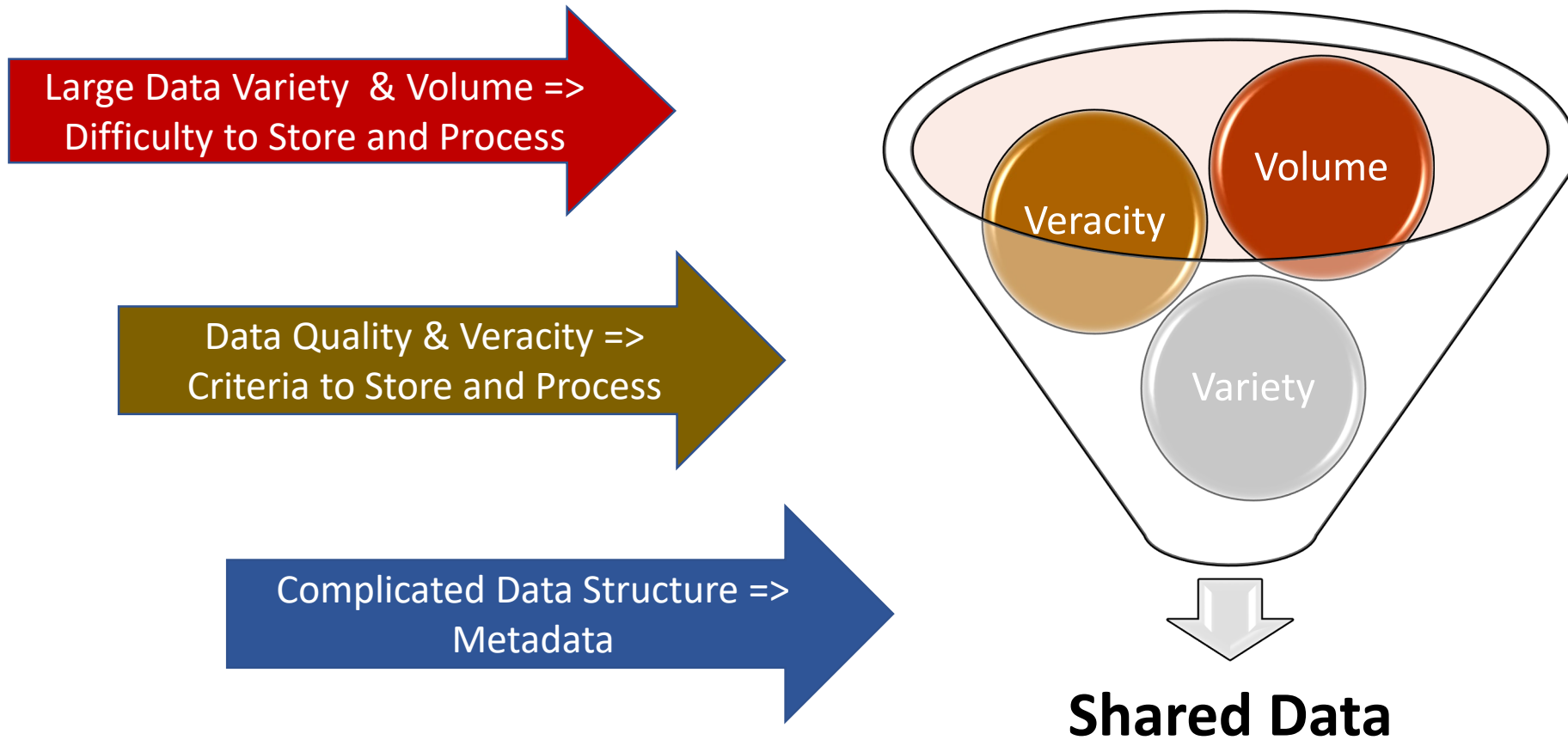- Data Access (Software, API)
- Processing Tools
- Multiple solutions in development (Strabospot, Epos)...

B. **Societal Challenges - Community Standards**
- Metadata and Vocabulary
- Which data to include (processed-unprocessed)?
- Avoid data Inconsistencies
- Data Quality Criteria
- Copyright, Privacy Issues

Geoscience Data - Problems

Large Data Variety & Volume =>
Difficulty to Store and Process

Data Quality & Veracity =>
Criteria to Store and Process

Complicated Data Structure =>
Metadata

Generated Data

Veracity

Volume

Variety

Shared Data

Geoscience Data - Dilemma

AGU requires that the underlying data and software needed to understand, evaluate, and build upon the reported research be available at the time of peer review and publication.

Specifically:

1. Depositing the data and software in a community accepted, trusted repository

2. Including an [Availability Statement](#) as a separate paragraph in the Open Research section

explaining to the reader where and how to access the data and software

3. Including [citation(s)](#) to the deposited data and software in the Reference Section

Geoscience Data - FAIR

Findable **A**ccessible **I**nteroperable **R**eusable

OSF
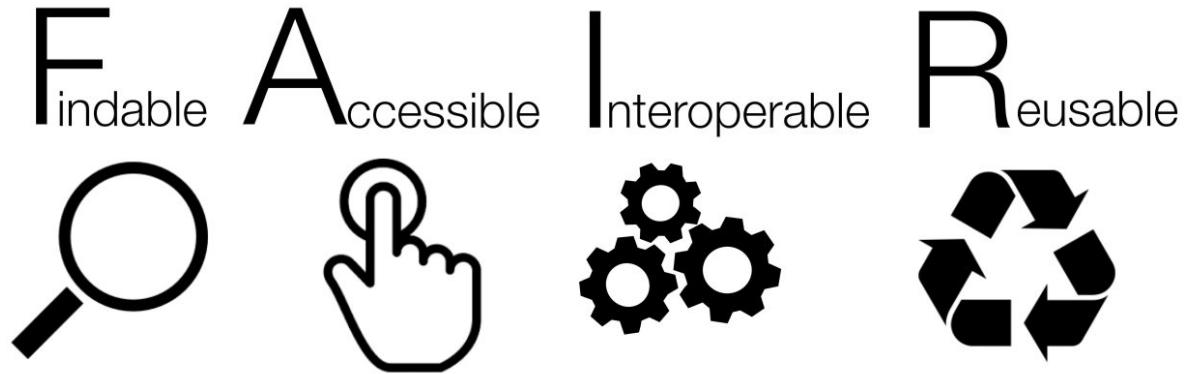
Project Based

5 Gb / Project

DOI* for Project

Rudimentary Search

Texas Data Repository

Dataset Based

10 Gb / Dataset ; 4 Gb / File

DOI* for each File

Rudimentary Search

*DOI=Digital Object

StraboExperimental

Summary

- To improve experimental data usability - structured or semi-structured data sets are a necessity

- FAIR Principles require standard vocabulary

- Experimental variables: use it (as metadata) or loose it.

- Data storage requirements forces laboratories to re-assess their workflows: use standard framework instead of insular solution

- StraboExperimental:

  - simplify metadata entry by use of templates and repositories

  - implementing workflow logic can further simplify user entry

Laboratory Workflow (LAPS)

- Local Solution
- Ease of Use
- Adapt to Community Standards as it develops
- Data is Structured
- Can contain Multiple Datasets
- Open Source
- Software + User Readable File Format
- Browser Based Forms (if needed)
- Single Data Container for Data + Metadata
- Reusable Templates

Laboratory Workflow (LAPS)

Combine Experimental Information + Data
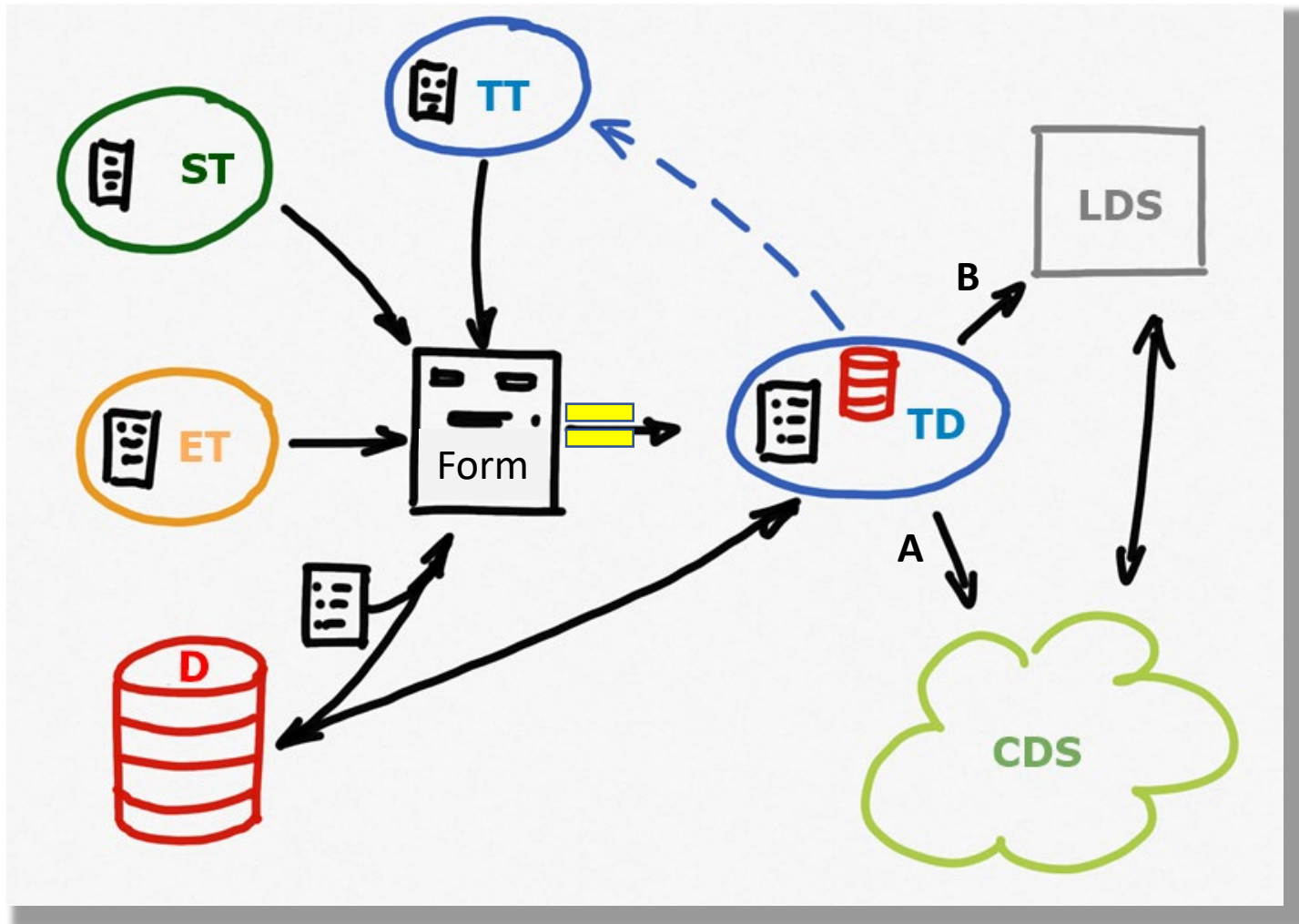
ST: System Template
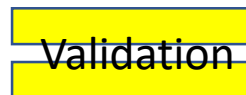ET: Experimental Template
    Sample Template
D: Data

TD: Test Dataset
TT: Test Template
LDS: Local Data System
CDS: Cloud Data System

Meta Data

Data Set (e.g., Time Series)

Validation

Laboratory Workflow

Templates simplify user input
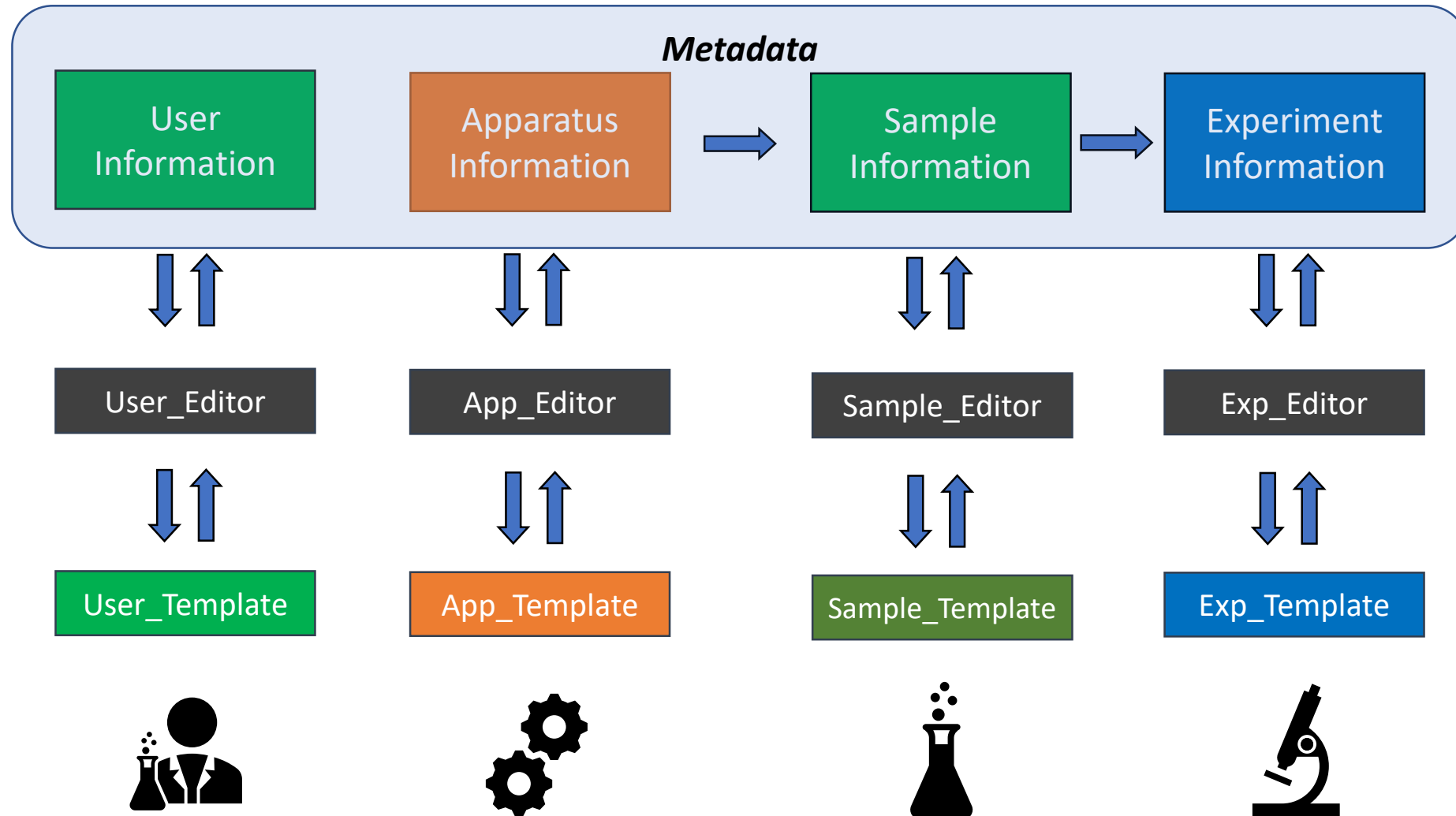
Laboratory Workflow - Demo

http://verve.mit.edu/laps

Editors

Many ways to modify and edit Experimental Data and Metadata

Laboratory Data Workflow - Scenario

Table Headers - Schema