



Software to Streamline Sharing of Agricultural Algorithms and Data

David LeBauer, Christophe Schnauffer, Kristina Riemer



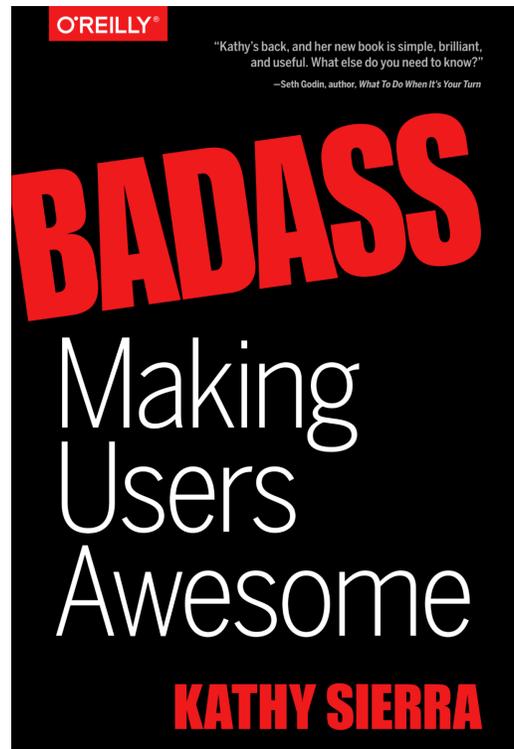
THE UNIVERSITY OF ARIZONA
College of Agriculture
& Life Sciences

Introduction

Challenges

How to make appropriate data deposition commonplace?

Solution



UA Digital Agriculture Group

Mission: Providing solutions for computing in agriculture, so that we can engineer crops and sustainable agricultural landscapes.

Vision: Faster, more collaborative agricultural science and engineering through shared software and data.



Emily Cain



Kristina Riemer



Julian Pistorius



Jorge Barrios



Chris Schnauffer



David LeBauer

Collaborators, Too Many to List:

Tyson Swetnam, Brian Heidorn, CyVerse, Todd Mockler, Nadia Shakoore, Geoff Morris, Vasit Sagan, Robert Pless, Rob Kooper, Max Burnette, Steve Long

Others to be listed as we go ...

Outline: Our Approach in 4 Ts

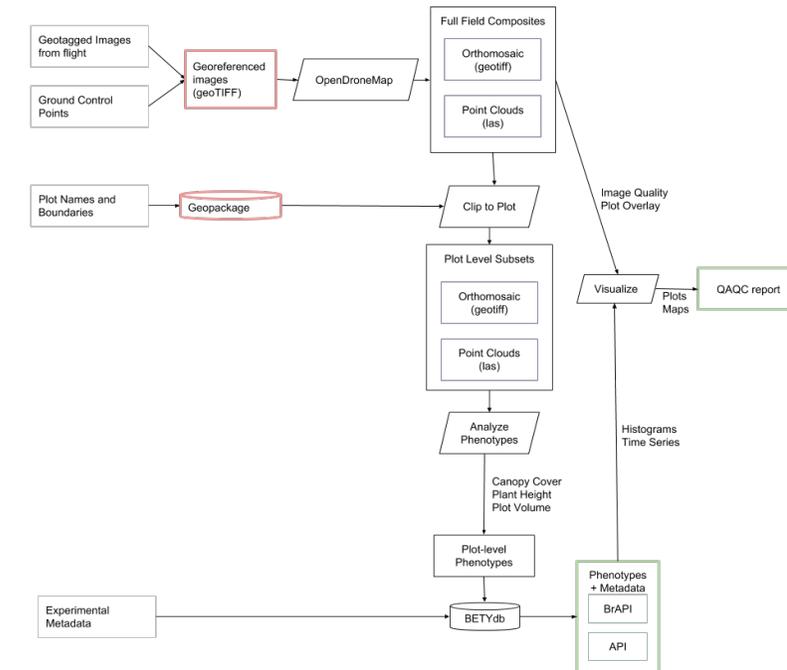
- Tools
- Translators
- Tutorials
- Templates

Outline: Our Approach in 4 Ts

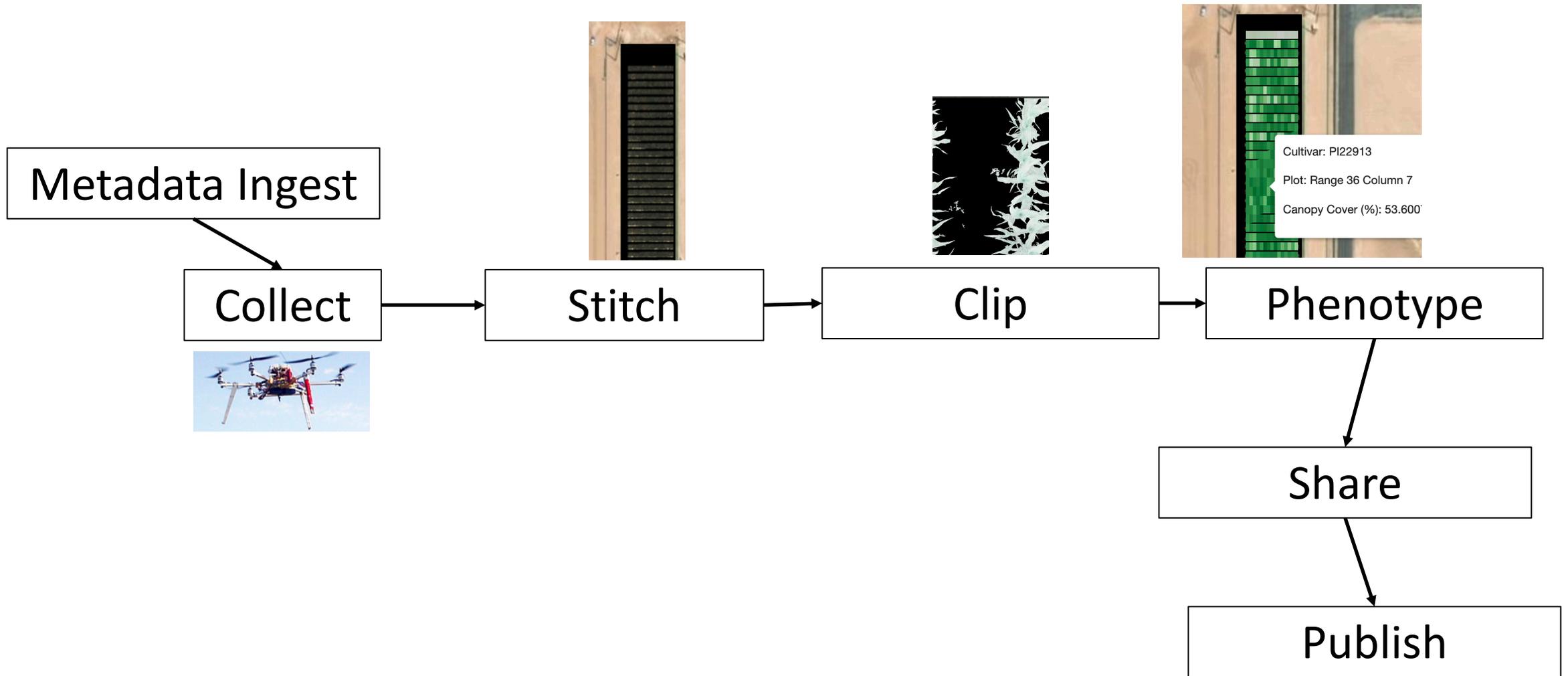
- **Tools**
- **Translators**
- **Tutorials**
- **Templates**



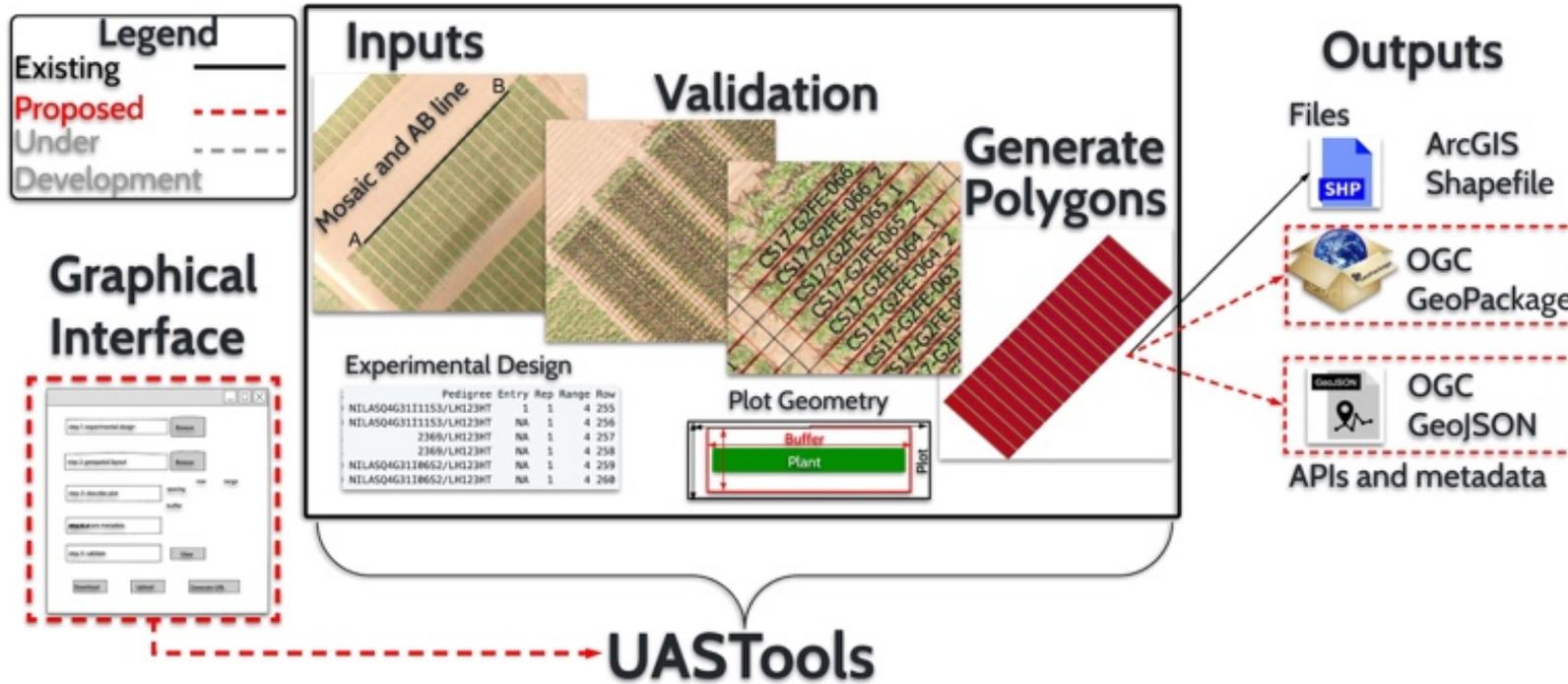
High Throughput Phenomics (images → phenotypes) Pipelines



A Simple Drone Pipeline



Metadata Ingest: Plot Boundaries

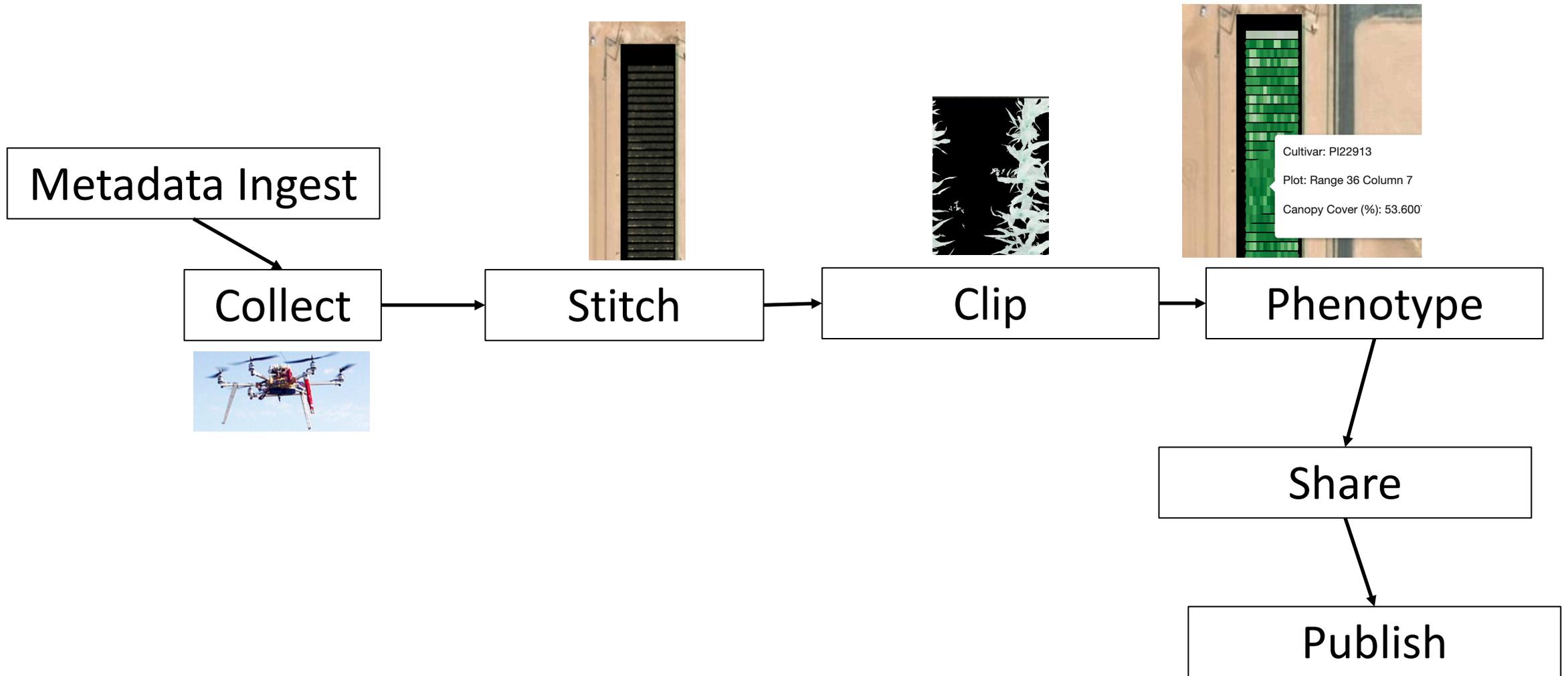


Works ✓

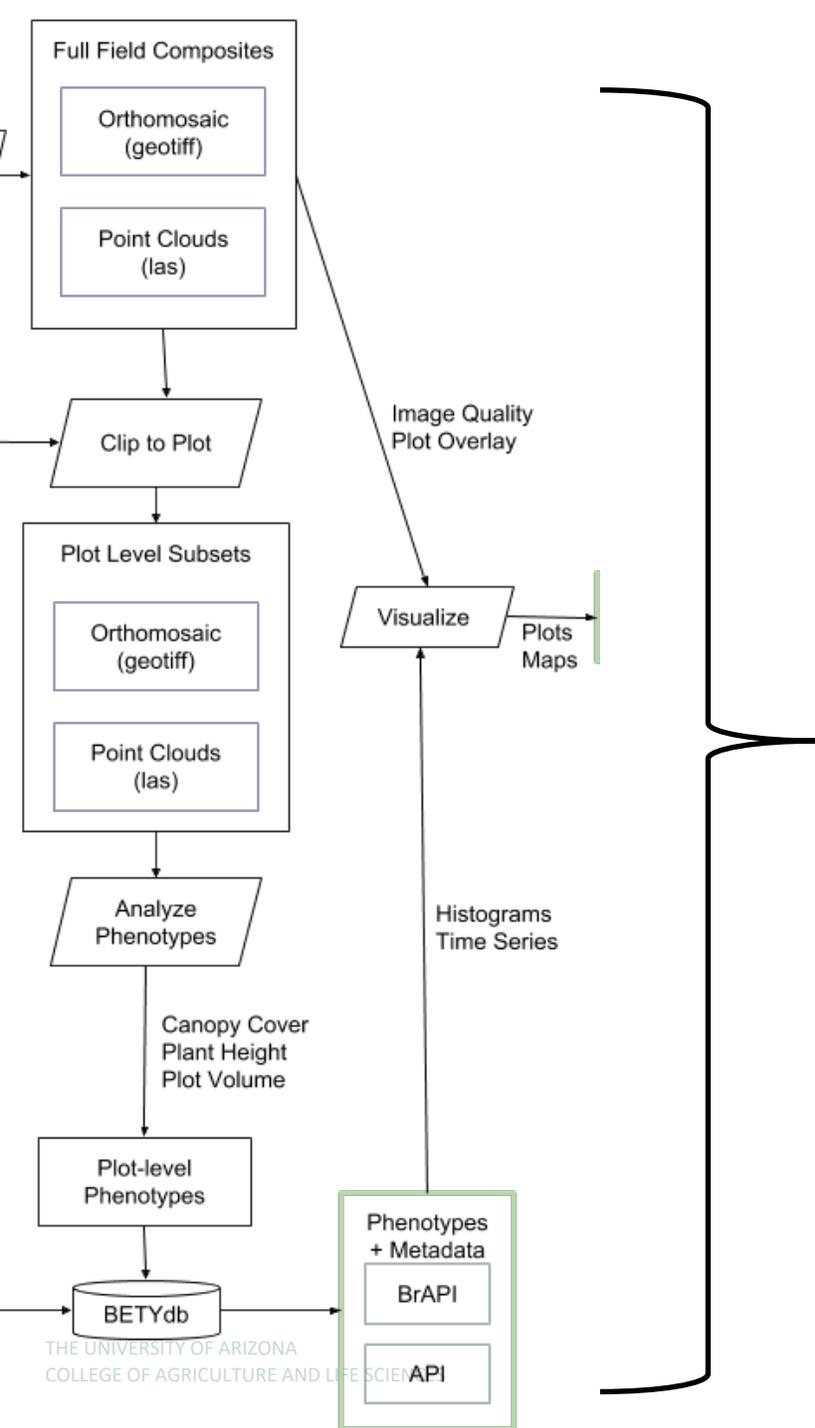
Open Source ✓*

Standard API ✓*

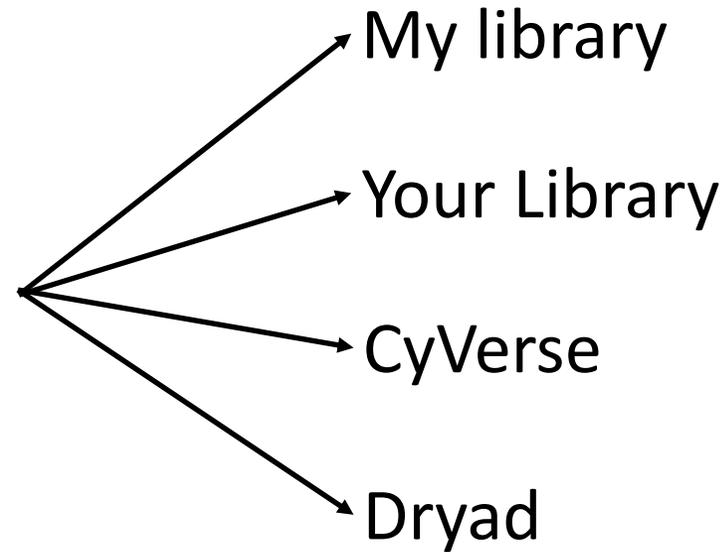
A Simple Drone Pipeline



Data Publication



Publishable Object



Data Publication



Clowder

Open Source Data Management for Long Tail Data

A customizable and scalable data management system to support any data format and multiple research domains. *Data logs in the clouds.*

V1.8.0 RELEASED!

JOIN US ON SLACK!

Create New Publication Request

Name

My Publication

Creator(s): 

One or more creators are required

Edit Metadata

SELECT REPOSITORY



Note: The links in this page redirect to the live objects.

Add metadata

Select field

No metadata available for this resource

Select Repository

SUBMIT TO
REPOSITORY >

Candidate Repositories

The results below are based on an analysis of the dataset's properties and metadata and the preferences you specified.

Inter-university Consortium for Political and Social Research

Match: all requirements are satisfied.

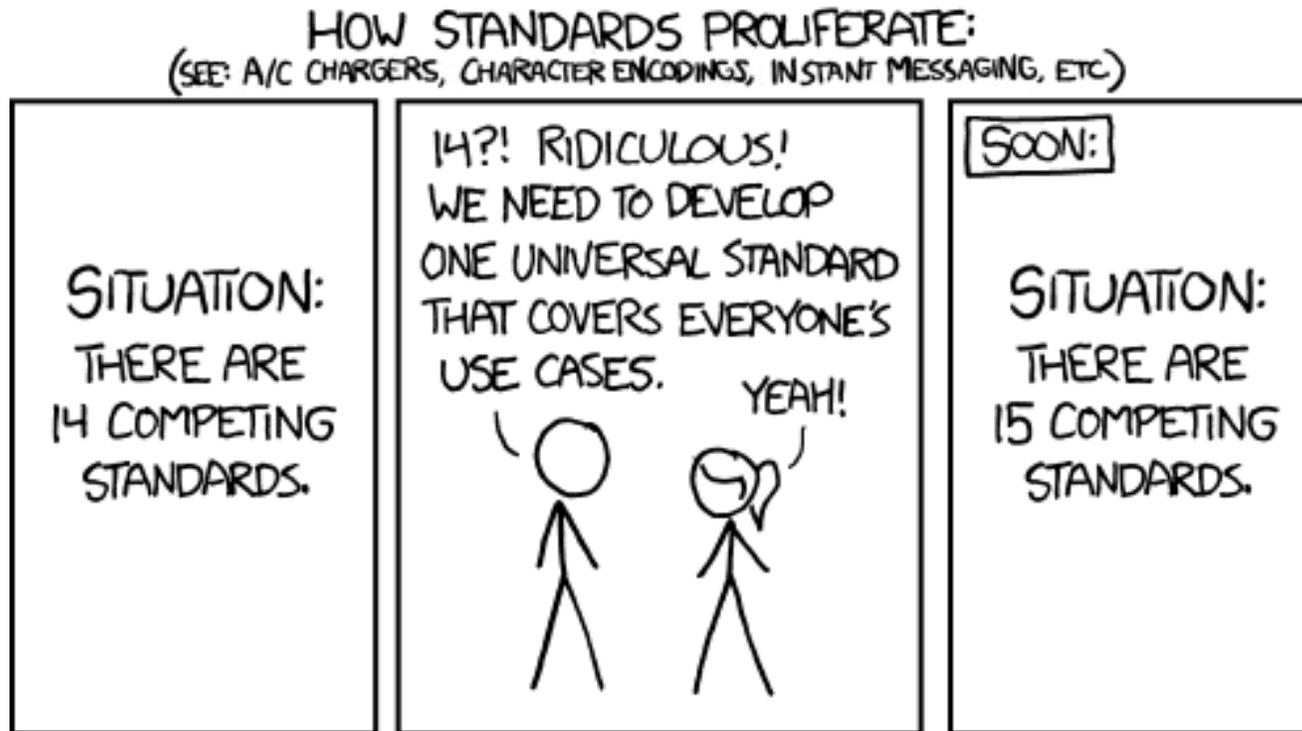
terraref.org/clowder/

THE UNIVERSITY OF ARIZONA
COLLEGE OF AGRICULTURE AND LIFE SCIENCES

Outline: Our Approach in 4 Ts

- Tools
- **Translators**
- Tutorials
- Templates

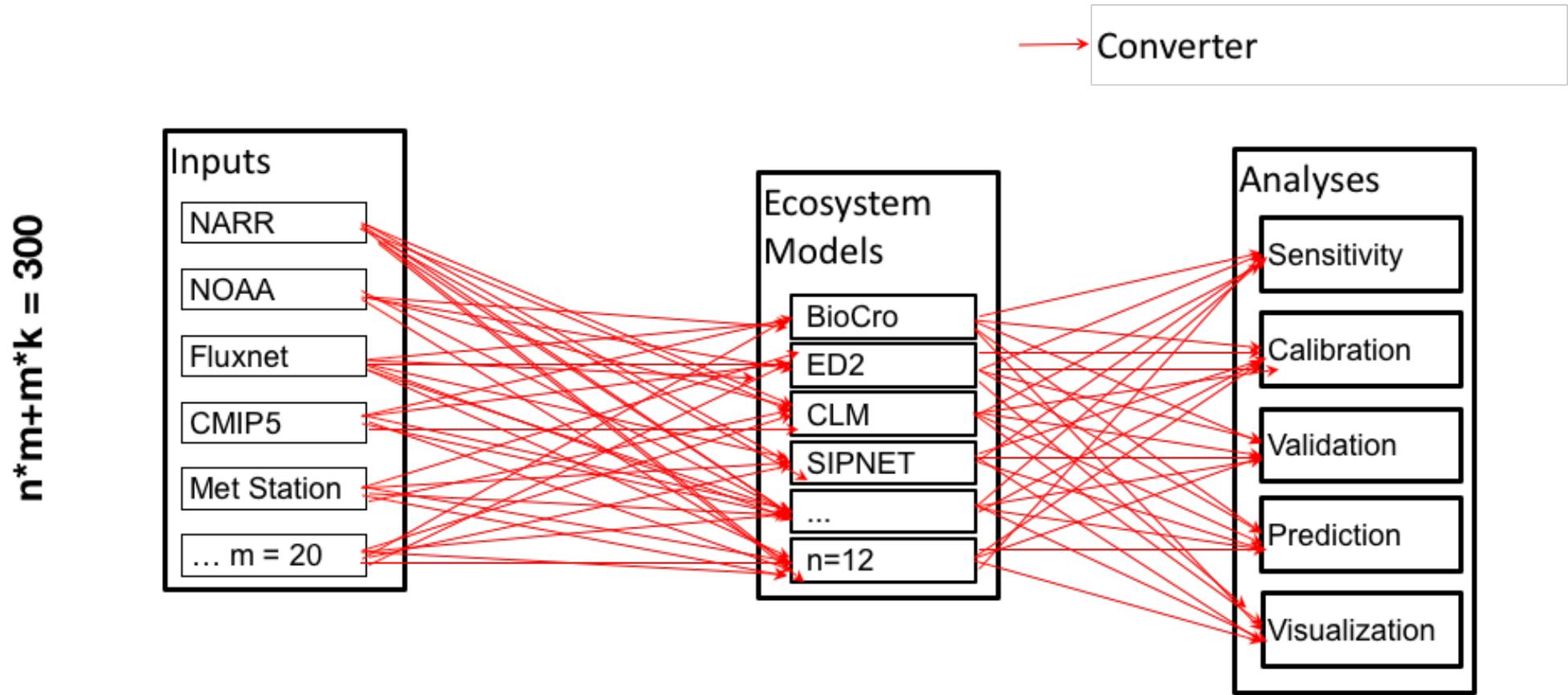
Translators



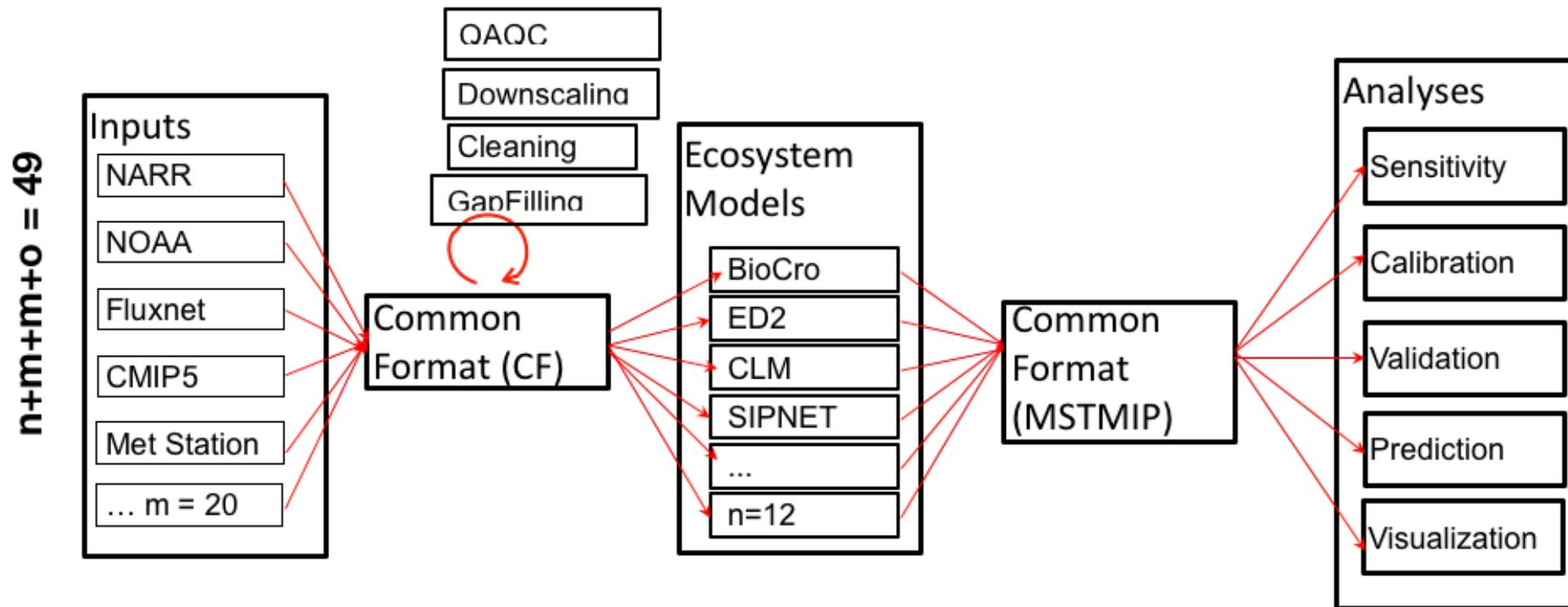
Alternative to a 'new'
standard:

Combine, translate among
existing standards.

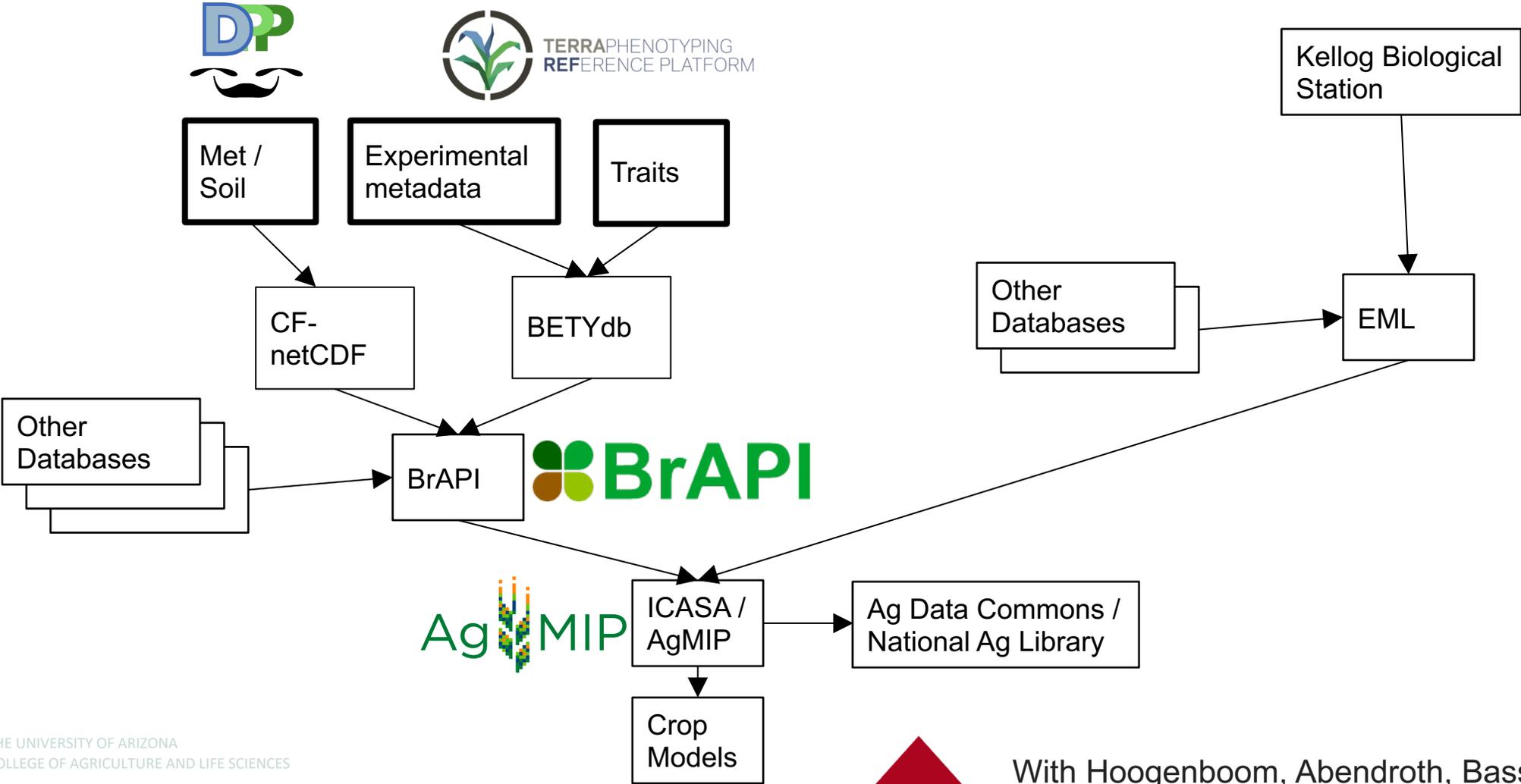
Ecosystem Models Without Standards



PEcAn + Standards: benefits of interoperability



Agricultural Research Data Network (ARDN)



Outline: Our Approach in 4 Ts

- Tools
- Translators
- **Tutorials**
- Templates

Tutorials: Self Guided, Live Coding Webinars, Videos

 <https://terraref.org/tutorials>

terraref.github.io/tutorials/index.html

- 1 Overview
 - 1.1 Pre-requisites
 - 1.2 Ways of Accessing Data
 - 1.3 Other Resources
- Section 1: Vignettes
 - 2 Vignettes Introduction
 - 3 Accessing trait data in R
 - 4 Accessing weather data in R
 - 5 Retrieve source RGB image files
 - 6 Combining trait, weather, and image ...
- Section 2: Tutorials
 - 7 Accessing Trait Data in R
 - 8 Accessing meteorological data

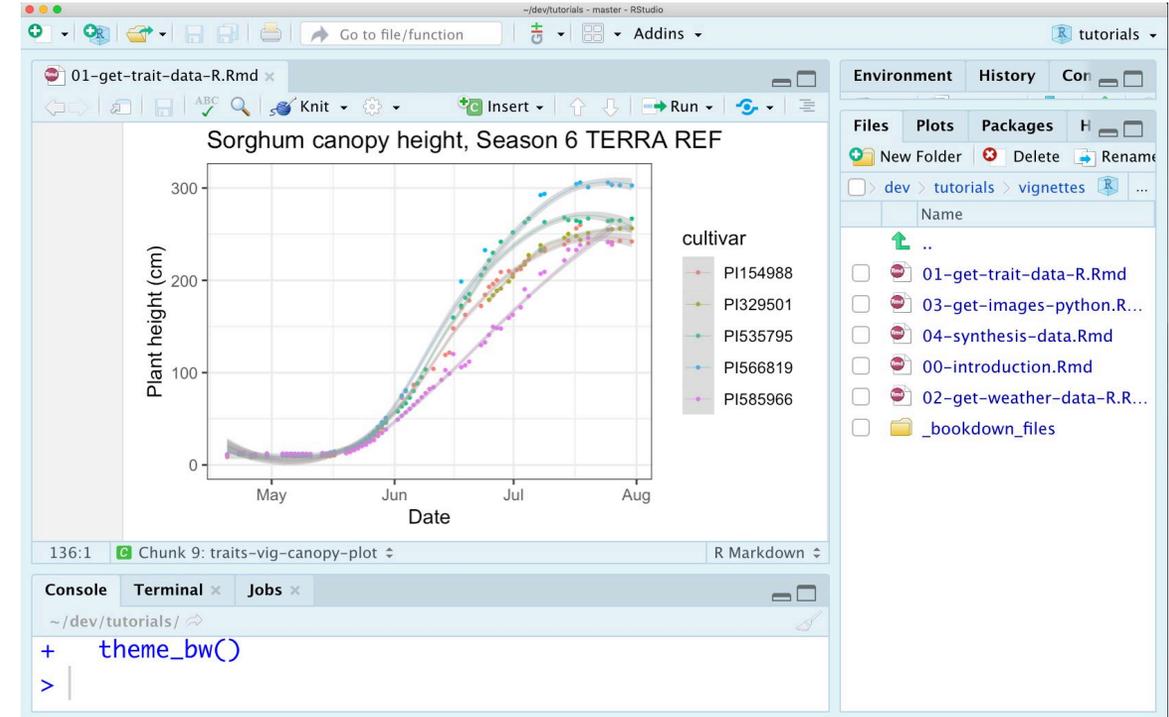
TERRA REF Tutorials

David LeBauer and others

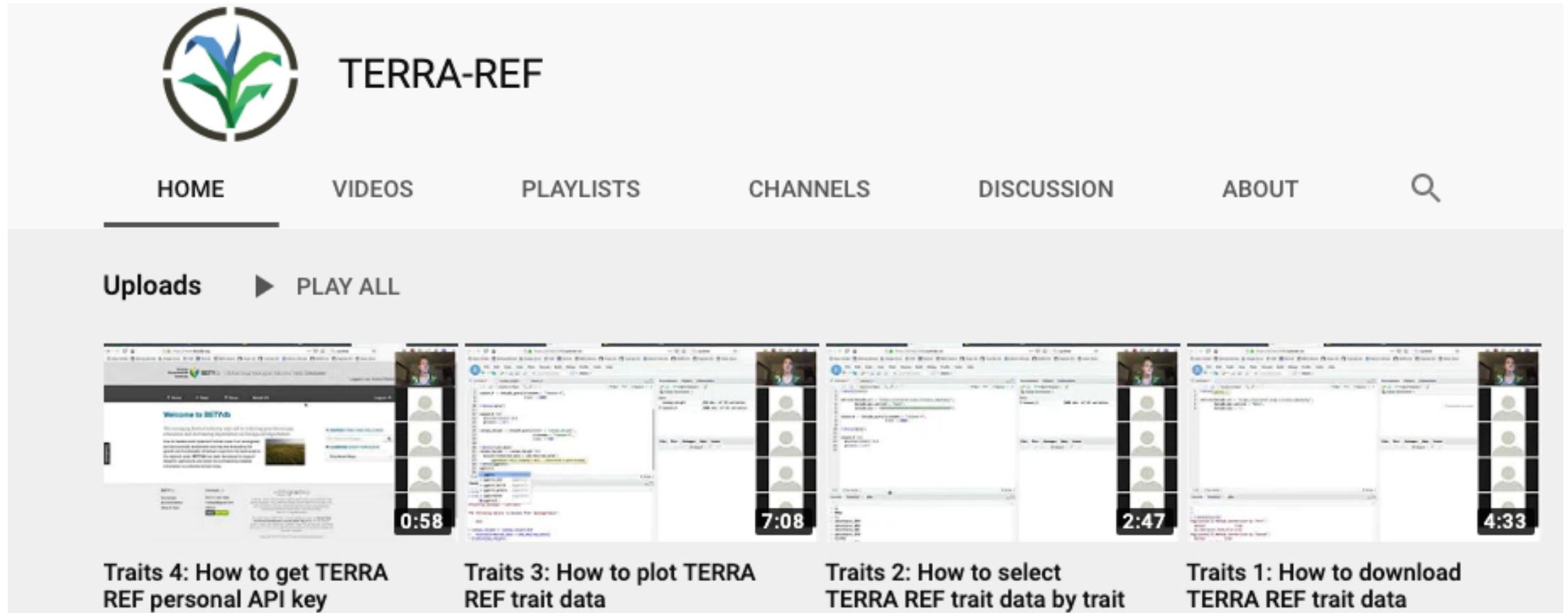
2020-01-06

Chapter 1 Overview

Kristina Riemer, UA



YouTube Videos



The screenshot shows the YouTube channel page for TERRA-REF. At the top left is the TERRA-REF logo, a stylized plant with blue and green leaves inside a circle. To its right is the text "TERRA-REF". Below the logo and name is a navigation bar with the following items: HOME (underlined), VIDEOS, PLAYLISTS, CHANNELS, DISCUSSION, ABOUT, and a search icon. Below the navigation bar is a section titled "Uploads" with a play button icon and the text "PLAY ALL". Below this are four video thumbnails, each with a duration timer in the bottom right corner. The first thumbnail shows a "Welcome to BETMD" page with a duration of 0:58. The second thumbnail shows a terminal window with a duration of 7:08. The third thumbnail shows a terminal window with a duration of 2:47. The fourth thumbnail shows a terminal window with a duration of 4:33. Below each thumbnail is a caption describing the video's content.

Uploads ▶ PLAY ALL

Traits 4: How to get TERRA REF personal API key (0:58)

Traits 3: How to plot TERRA REF trait data (7:08)

Traits 2: How to select TERRA REF trait data by trait (2:47)

Traits 1: How to download TERRA REF trait data (4:33)

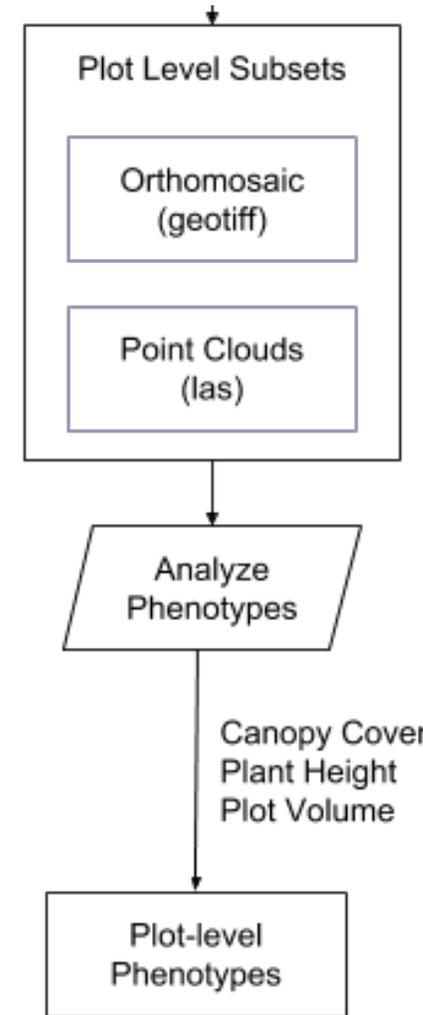
Outline: Our Approach in 4 Ts

- Tools
- Translators
- Tutorials
- **Templates**

'Transformer' libraries

```
def calculate(data, metadata)
  do something
  return data, metadata
```

v1: github.com/terraref
v2: github.com/agPipeline



With Kooper, Burnette, Schnauffer

Templates

Guide users to best practices

- Standard formats and vocabularies
- Best practices
- Testing
- Documentation

Minimize Overhead

- Quickly integrate a new algorithm into a pipeline
- Versioning
- Sharing

Example template

```
def calculate(pxarray: np.ndarray):  
    """Calculates one or more values from plot-level RGB data  
    Arguments:  
        pxarray: Array of RGB data for a single plot  
    Return:  
        Returns one or more calculated values  
    """  
  
    # ALGORITHM: replace the following lines with your algorithm  
    channel_size = pxarray[:, :, 1].size  
  
    # RETURN: replace the following return with your calculated values.  
    return channel_size
```

github.com/AgPipeline/template-rgb-plot
algorithm_rgb.py

A Template for Plot level RGB data



1. **Setup:** Click the `Use this template` button in GitHub to make a copy of this repository (or run `git clone`)
2. **Definitions:** Fill in and modify the definitions in the `algorithm_rgb.py` file
3. **Algorithm:** Replace the code in the `calculate` function with your algorithm
4. **Test:** Run the `testing.py` script to run your algorithm and validate the results
5. **Generate:** Run `generate.py` to create a Dockerfile
6. **Docker:** Create a Docker image for your algorithm and publish it
7. **Finishing:** Finish up your development efforts

Conclusions

Reusable software and data will accelerate science and engineering

Tools, Translators, Templates, and Tutorials enable open, synthetic science

More Information

David LeBauer

dlebauer@arizona.edu

datascience.cals.arizona.edu @digitalagggroup