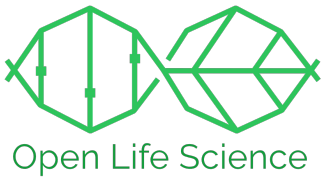
Improving Documentation for Agricultural Research Software

Team Members: Kristina Riemer, Emily Cain, David LeBauer



DIgital Agriculture Group

We develop software and enable data intensive agriculture research

Goals:

- make software and data more accessible
- enable contributors

OLS Focus: Drone Processing Pipeline

@DigitalAgGroup

https://osf.io/xdkcy/

Initial Open Science Goals

- **Improve documentation to:**
  - Increase use of research software
  - Identify user needs
  - Engage with and encourage contributors

Getting User Feedback

Helpful:

- Feedback & communication
  - In person
  - Slack
- Availability of mentor / experienced user

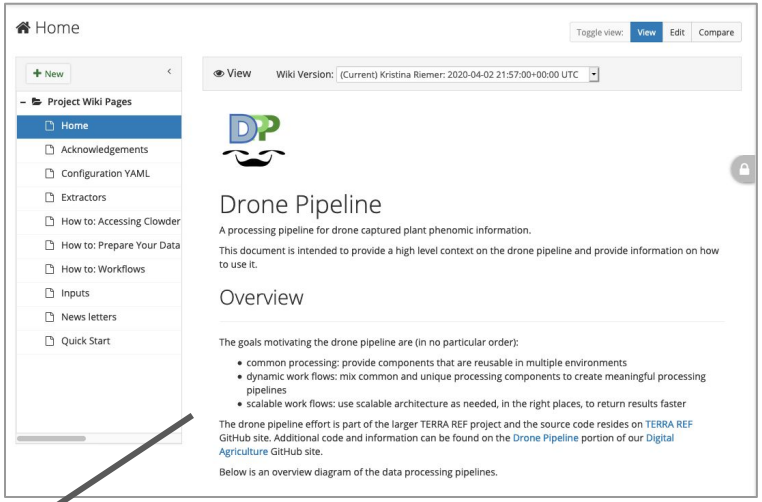
Barriers:

- Technical skills
- Data management

Future:

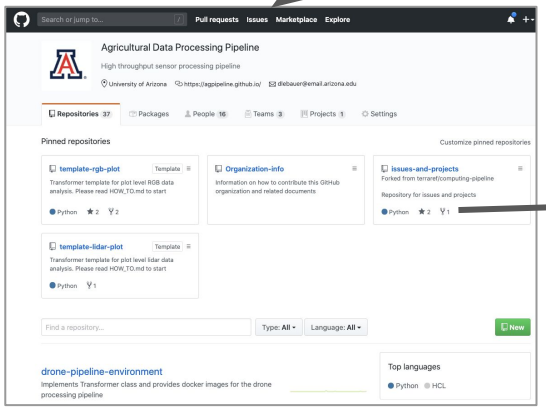
- More advanced users with data
- Possible developers
- Pipeline & documentation needs to be more automated and user-friendly, but how?

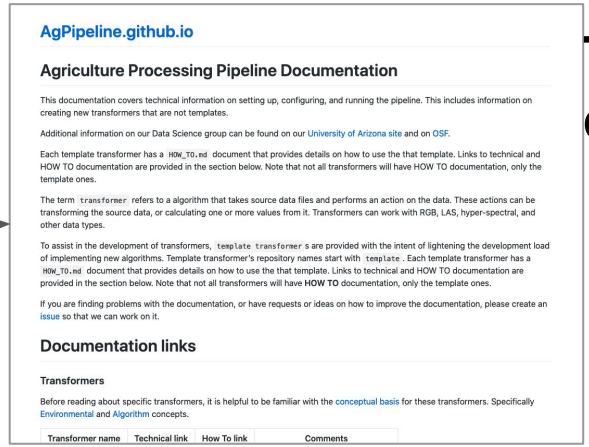
OLS Outcomes



Main website:
- Personas
- Tools list

GitHub:
- Code
- Community health files





Technical documentation

Accomplishments

- Identified audience: users and contributors
- Improved website, documentation from their perspective
- Added community health files, personas
- Built relationships and understanding within our team

Challenges

- Inertia
- Organizing materials intuitively
- Choosing project to improve

Next steps

- Add personas and tools to main website
- Have new users “test” pipeline documentation

Next level of open science is building open, inclusive, and welcoming communities

Acknowledgements

- Pipeline developers Chris Schnaufer & Julian Pistorius
- Mentor Katrin Leinweber
- OLS leaders Yo, Malvika, and Bérénice
- Everyone in cohort 1!
- All of the amazing speakers

Open Canvas

project : Increasing Ag Pipeline Usability (through documentation & software)

Problem

Overall: ag researchers need flexible, scalable, open software to process sensing data; need to be able to add new algorithms; need to be able to use the data that it generates

OLS-specific: identifying, engaging with, and getting contributions from potential users of the AgPipeline and the data it generates

Solution

Overall: AgPipeline including: templates + workflow + databases & APIs

OLS: Improve documentation of datasets and software so they are useful to others and encourage contributions

Unique Value Proposition

We are developing and maintaining a pipeline for processing agricultural data that doesn't already exist, which should be able to be used by many researchers to accelerate research in this domain and will be extended to others.
Our software and data is open, automated, scalable, extensible. We facilitate best practices in software development and data curation.

Key Metrics

- feedback from 2 contributors + Chris
- 5 contributions to AgPipeline from outside of our group by July 1
- 2 new tutorials for data access
- clear landing page for all agpipeline related stuff

User Profiles

Target audience and early adopters

- Early adopters = researchers working "on the fringes" of some of these projects already
- Target audience = ag researchers already working with HTP & image data
- IT staff who deploy pipeline at their own institution

User Channels

1. Funded projects that we are collaborating on
2. Landing page, twitter, youtube tutorials
3. Conferences, workshops, etc

Resources Required

- Guides for documentation best practices
- People hours to improve our documentation to those standards (requires people with technical experience)
- guidance and feedback
  - from Chris and Julian and other devs
  - from end users (e.g. Michele, Emmanuel, Sateesh, ...)
- Github, Zenhub, Slack, CyVerse, OSF
- High Performance Computing
- Servers for Databases, APIs

Contributor Profiles

Contribution types and ideal contributors

- Algorithm contributors - scientists / engineers who want to analyze their data
- Core software development
Contributions = improvements to our documentation that reduce pain points and make products easier to use

Contributor Channels

1. Funded projects that we are collaborating on
2. Interested collaborators at conferences e.g. Phenome, SPIE
3. People we are proposing to work with
4. GitHub

Project Execution

Product

Community

See next slide for instructions!