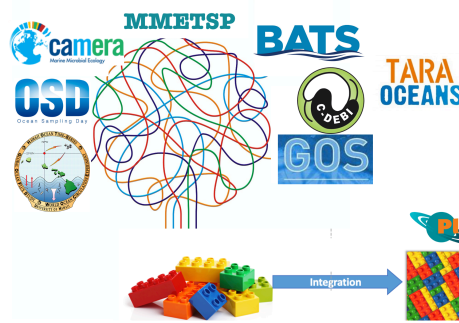# CSSI Element: Accelerating Comparative Metagenomics through an Ocean Cloud Commons

PI: Bonnie Hurwitz Co-PI: John Hartman; University of Arizona
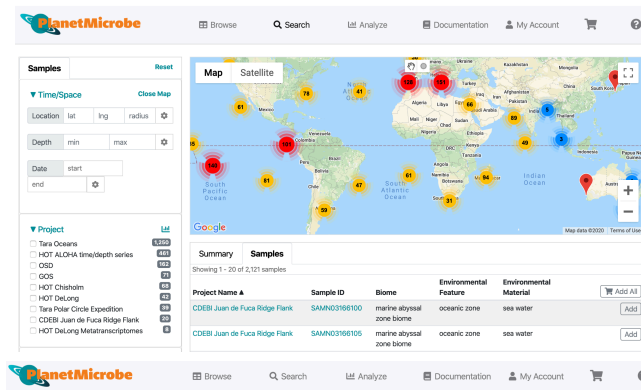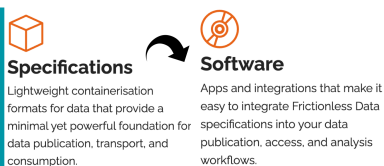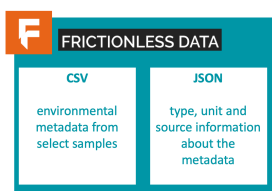
Award #: 1640775

Goals

Data Repository

Global ocean 'omics datasets are spread across diverse projects, with varied metadata and semantics

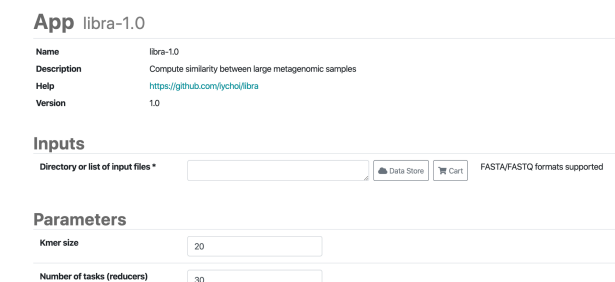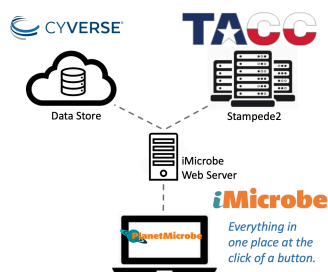## Data containers for dataset interoperability

Planet Microbe uses frictionless data packages to make datasets interoperable, as well as shareable between systems.

FRICTIONLESS DATA

CSV	JSON
environmental metadata from select samples	type, unit and source information about the metadata

Specifications
Lightweight containerisation formats for data that provide a minimal yet powerful foundation for data publication, transport, and consumption.

Software
Apps and integrations that make it easy to integrate Frictionless Data specifications into your data publication, access, and analysis workflows.
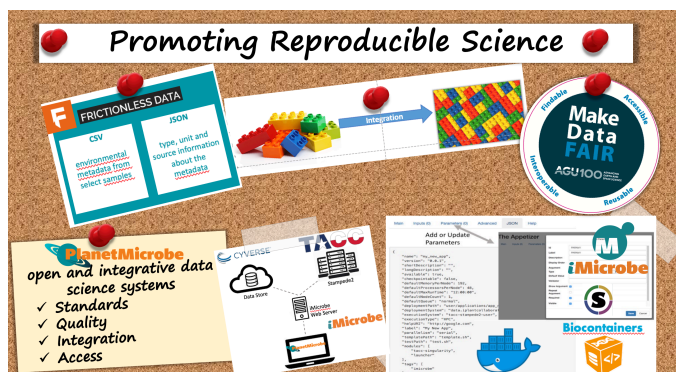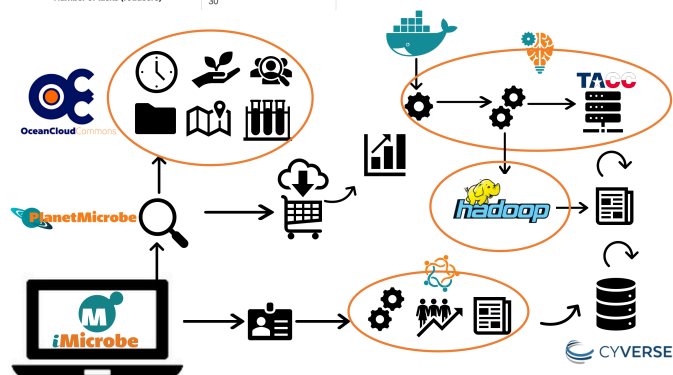

Scalable Analytics

Cloud Architecture
- Leverages CyVerse & TACC
  - Security
  - Scalability
  - Lower cost of development/maintenance
- Apps (tool containerization)
  - Docker & Singularity
  - Build once, run everywhere

Everything in one place at the click of a button.

- > 2000 global marine samples
- Distributed & linked datasets
- Interoperable across projects
- > 25 containerized tools & pipelines
- CI-enabled supercomputing
- User dashboard and data

Promoting Reproducible Science

Alise Ponsero, PhD, MS
Post-Doctoral Fellow

Matt Bomhoff
Sr. Developer

Ken Youens-Clark, MS
Sr. Developer

Kai Blumberg
Graduate Student

http://planetmicrobe.org

http://imicrobe.us

@hurwitzlab