

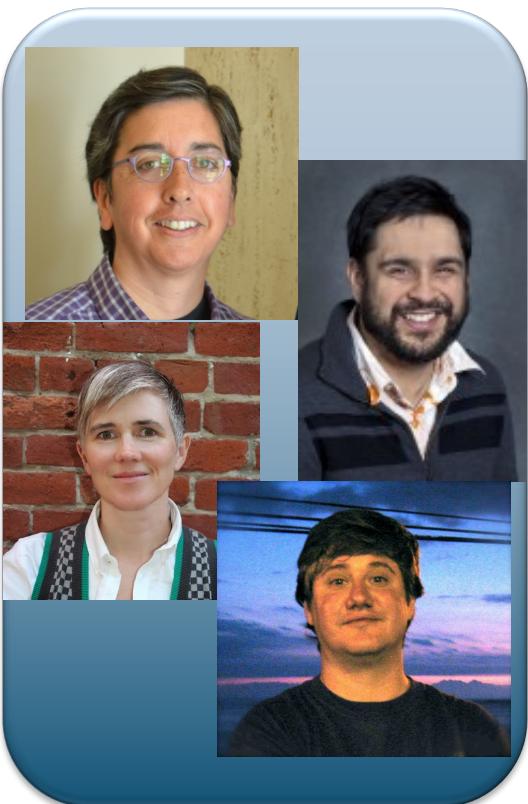


# ESS-DIVE - Environmental System Science Data Infrastructure for a Virtual Ecosystem

**LBNL (Berkeley Lab): Deb Agarwal, Charuleka Varadharajan, Shreyas Cholia, Cory Snavely, Dan Gunter, William Riley**

**UCSB NCEAS (National Center for Ecological Analysis and Synthesis):  
Matt Jones, Amber Budden, and Dave Vieglais**

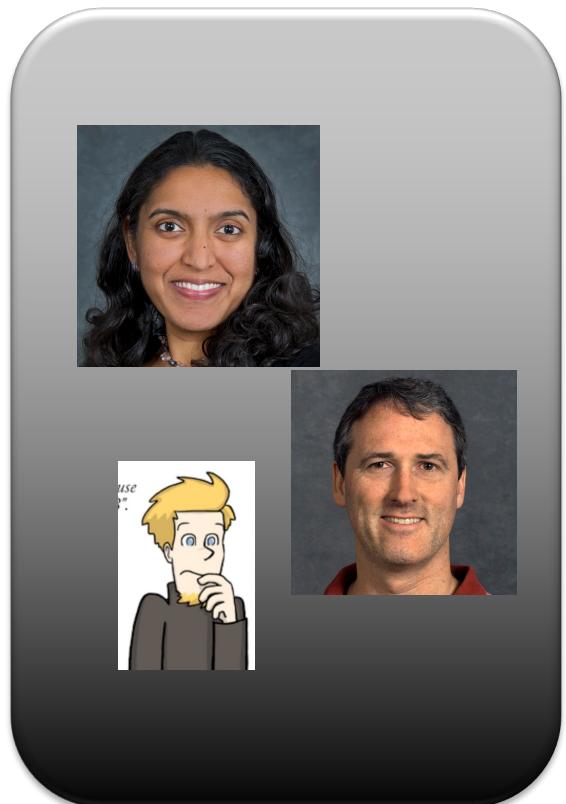
# Our Team is Well Positioned to Build the Archive



Data Scientists



Digital Librarians



Domain Scientists

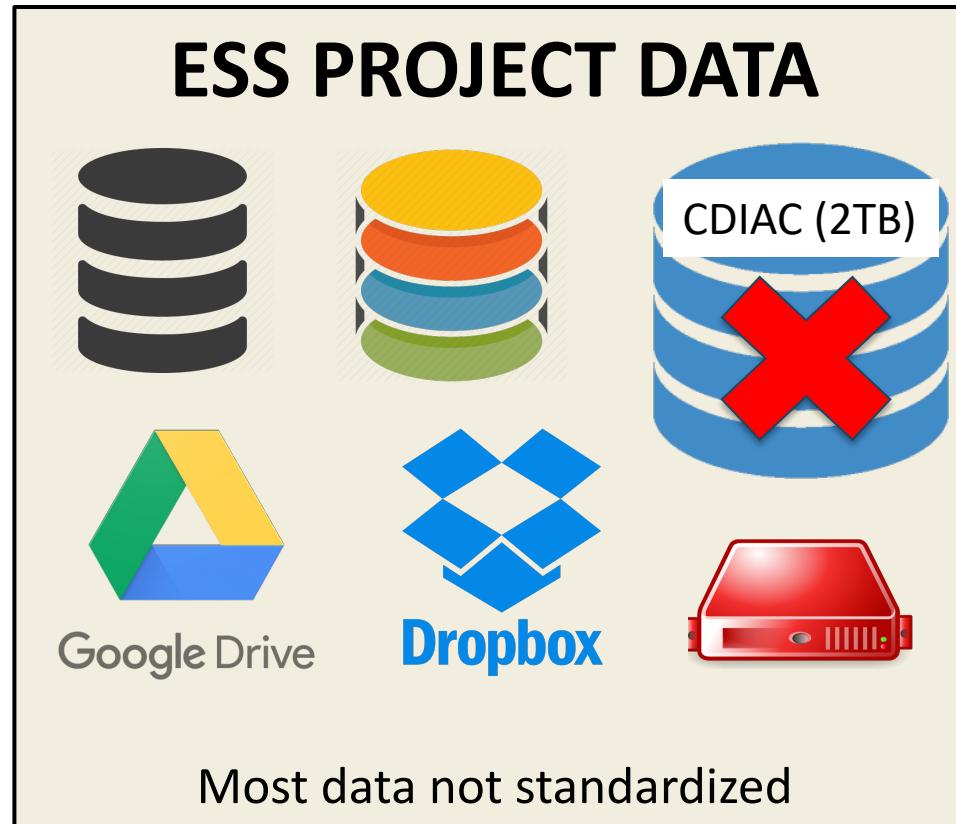
# Long-Term Vision: A Data Archive that Enables ESS Science Through Community Engagement

# 2023



# Currently: ESS Projects Must Store Their Own Data

2017

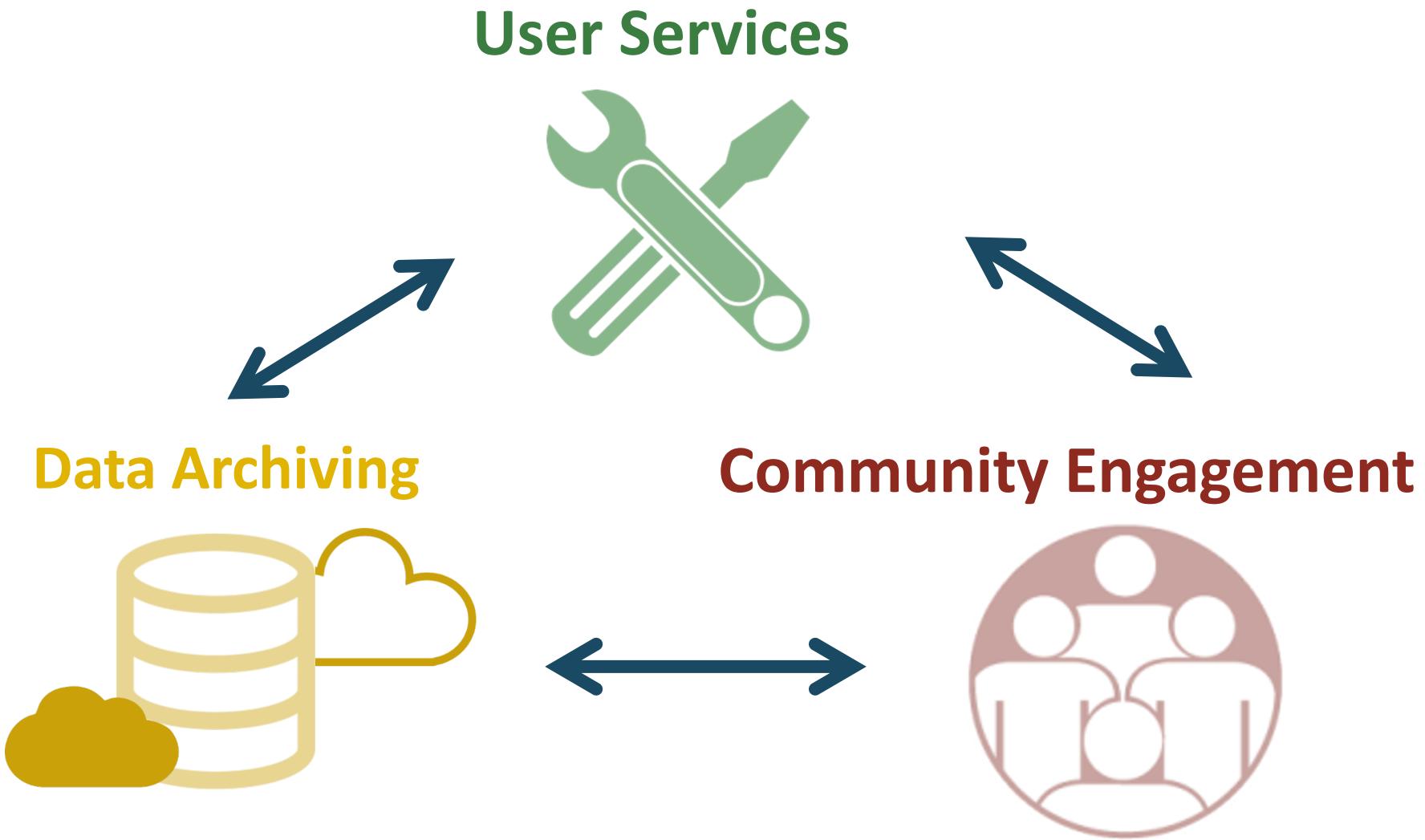


# Near-Term Vision: A Data Archive that Stores All ESS Data and Provides Easy Access to That Data

# 2020



# Three-Pronged Approach to Developing ESS-DIVE



# Innovative Capabilities of ESS-DIVE To Enhance Value of Public Archive to ESS Community

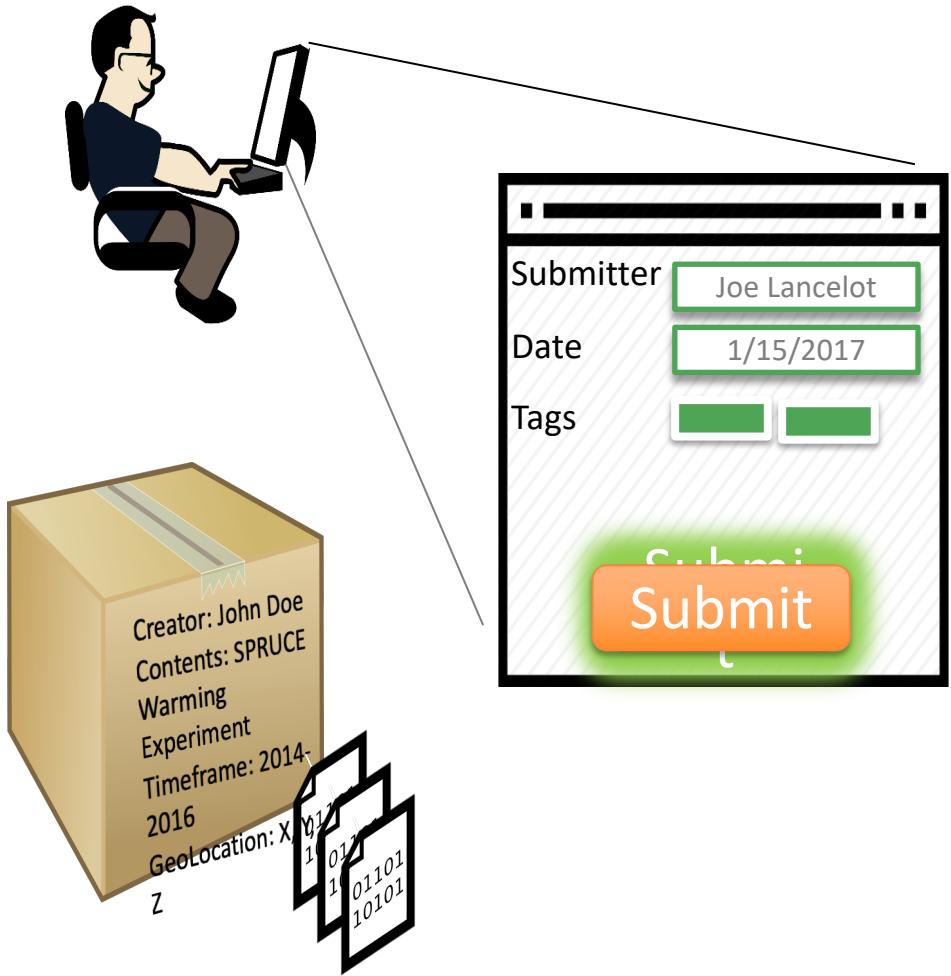


- Provide archiving of **private/project** data packages before public release
- Enable new data tools and capabilities through **community development**
- Understand user needs with **User Experience Research**
- Build **fusion database** to enable integrated data search, download, and visualization

# ESS-DIVE Leverages Team Activities and Experience



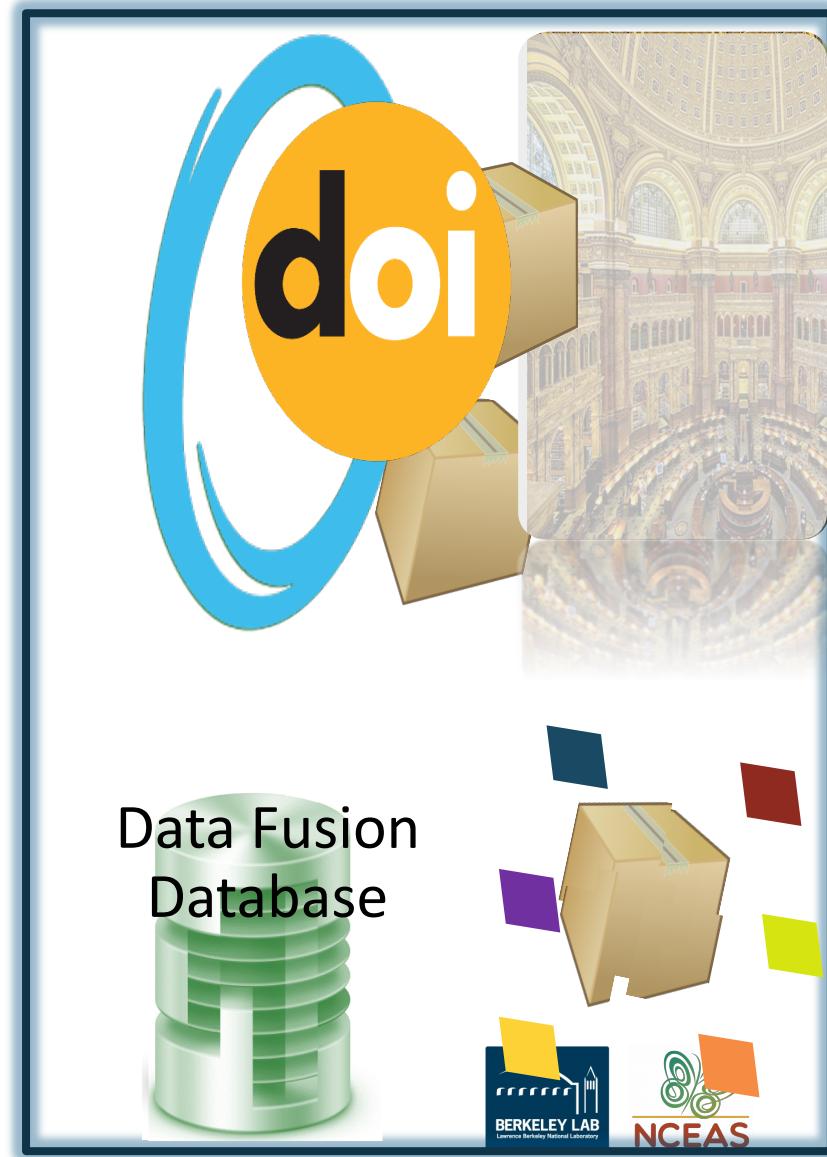
- Leverage existing **digital library** developments and experience
- Experience **migrating archives**
- Data team for several **ESS projects** including: AmeriFlux/FLUXNET, NGEE-Tropics, Watershed Function SFA, etc.
- Activities in **standards and ontologies** such as Ecological Metadata Language (EML)
- Developed initial versions of **software needed** to build ESS-DIVE
- Experience with **large-scale data transfers** and tools to enable transfers



## ESS-DIVE Data Upload

ESS-DIVE  
Portal

ESS-DIVE  
Archive

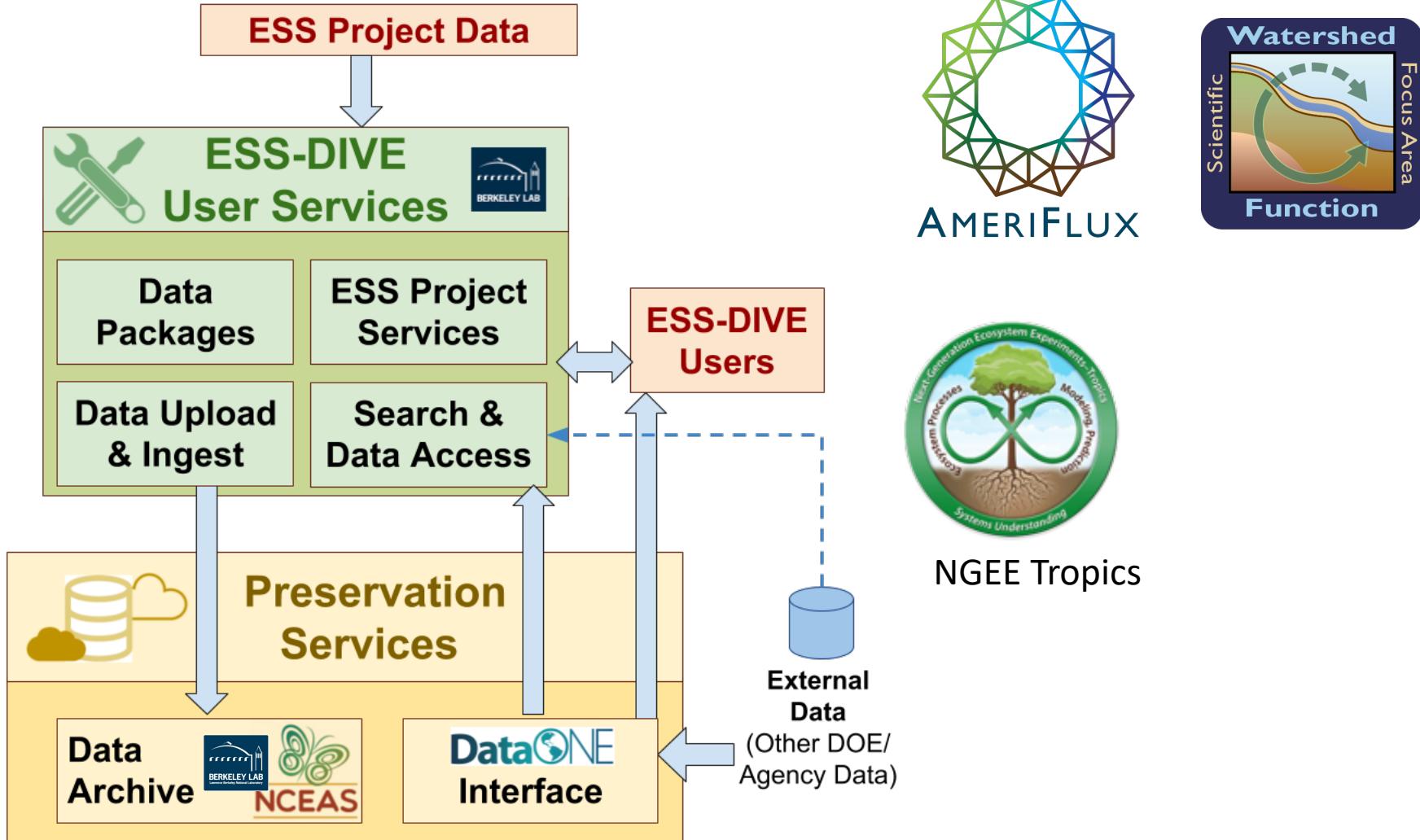


U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science

# ESS-DIVE Archive Software Architecture

## Builds on Existing Software



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science

# ESS-DIVE User Portal Starting Point in ESS

## Developed Software



- Hosted at Berkeley Lab
- Portal for data upload, get DOIs, catalog search/access

Home Data Policy Help and Support      meghasandesh Logout

ESS-DIVE Data Archive

Create New Dataset

Edit Draft Datasets

View Approved Datasets

- Advanced data search and visualization (fusion DB)

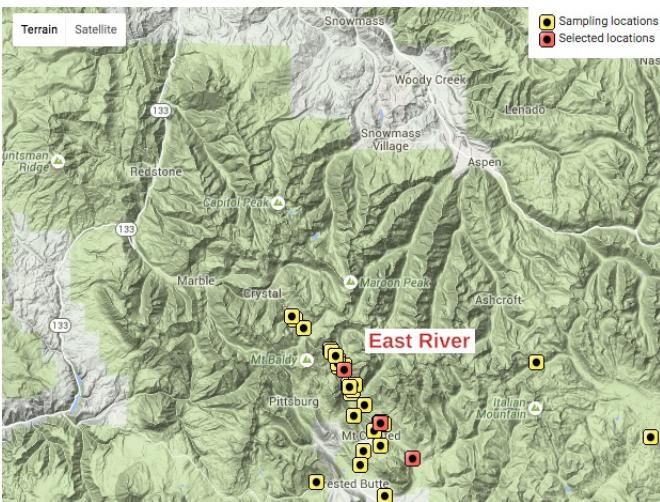
SITES  
 East River  Rifle

LOCATIONS  
Avery, BCC\_ISCO, PH\_ISCO

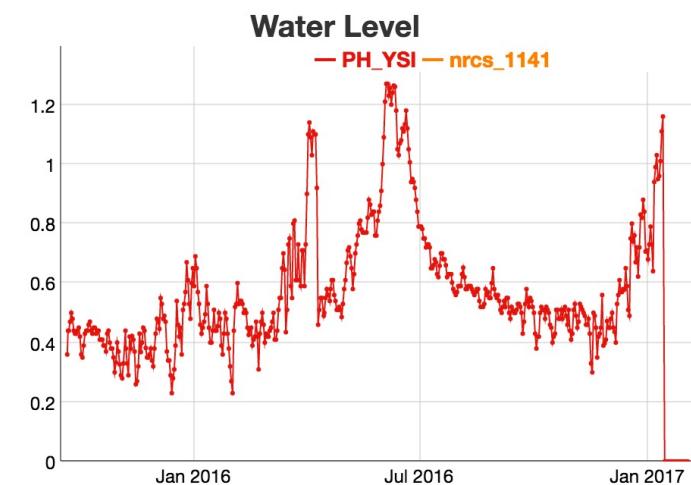
PARAMETERS  
Water Temperature, Nitrate (NO<sub>3</sub>)

TIME PERIOD  
01-2014 TO 05-2016

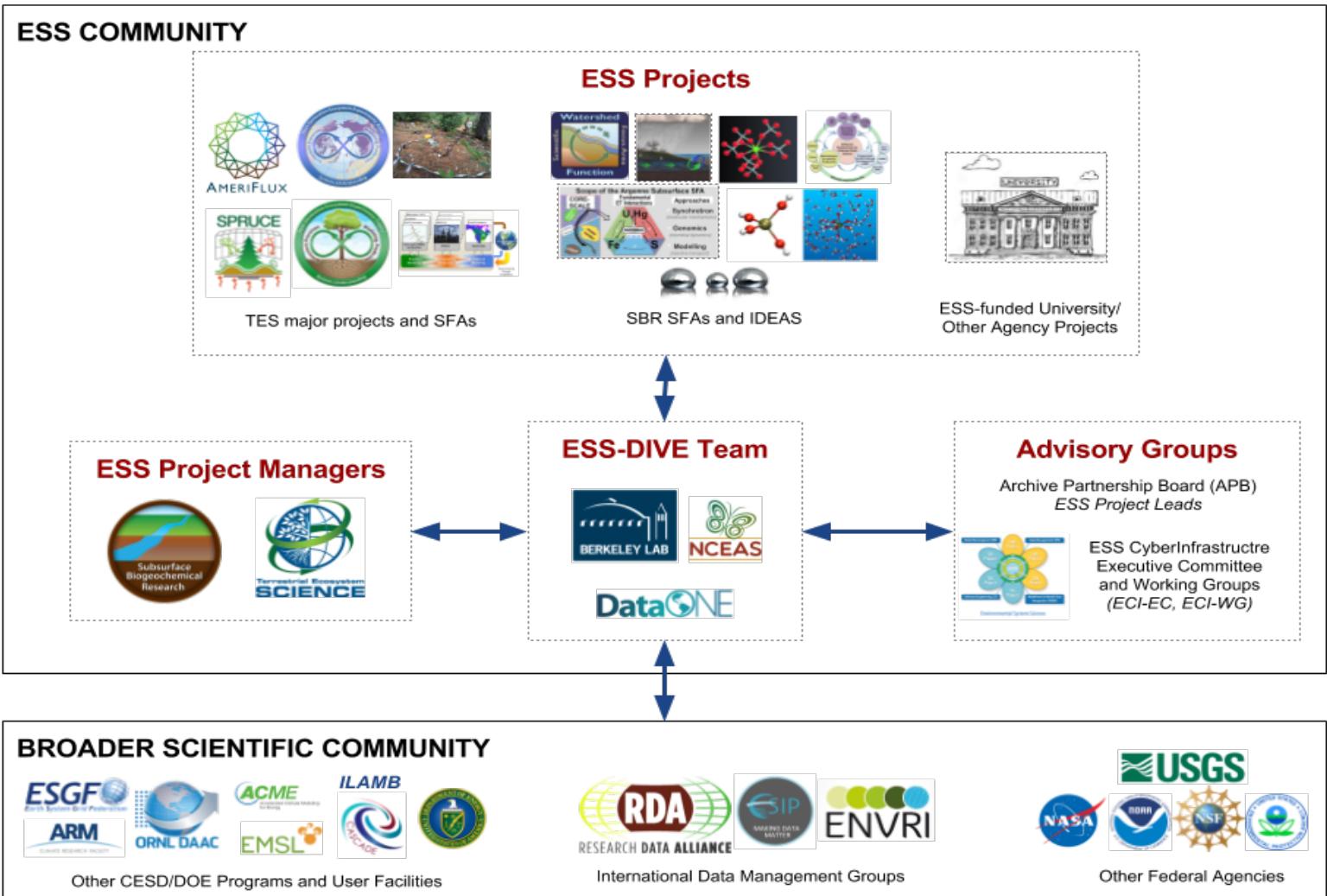
**SUBMIT**



The map displays a terrain-based view of the East River area in Colorado. It shows various towns like Snowmass, Woody Creek, and Aspen. Numerous sampling locations are marked with yellow squares, while selected locations are marked with red squares. A legend in the top right corner identifies these symbols.



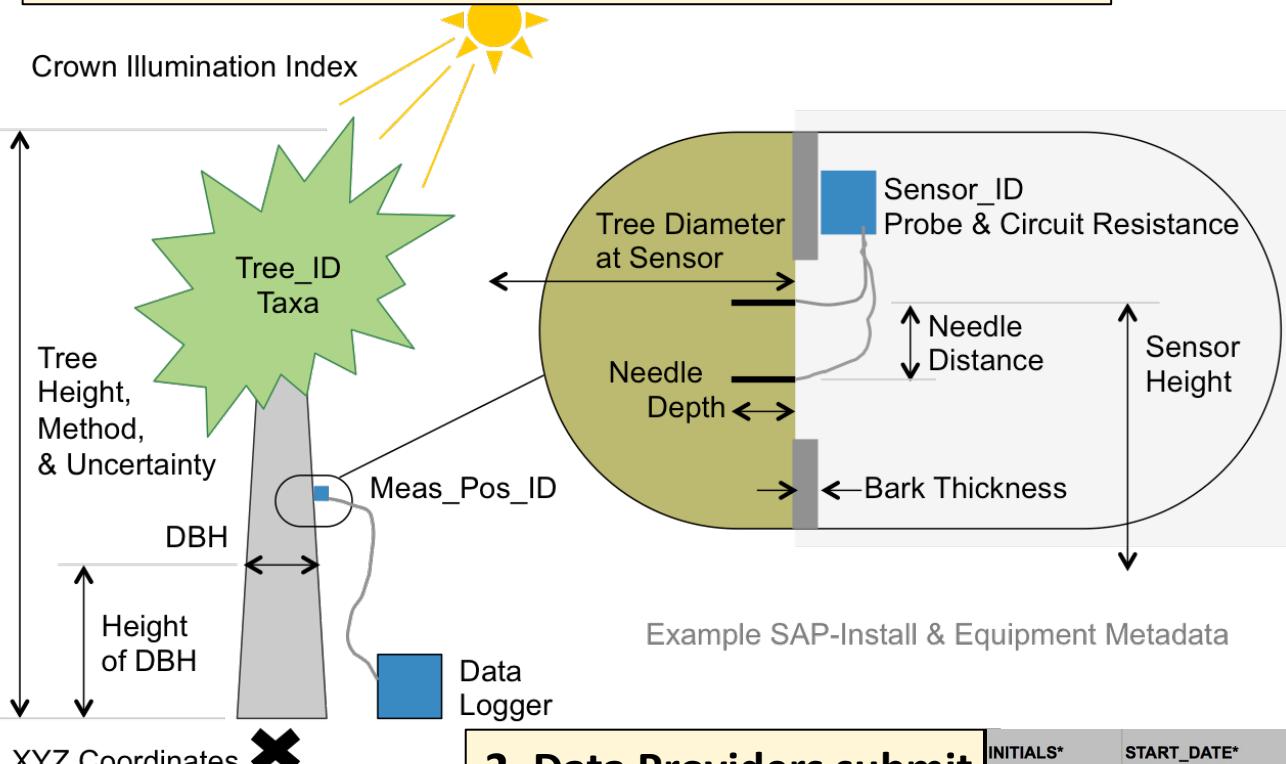
# ESS-DIVE Is a Partnership With the ESS Community



# Example: User Experience Research Applied to Metadata Standards and Tools



## 1. Sapflow metadata reporting templates developed with community input



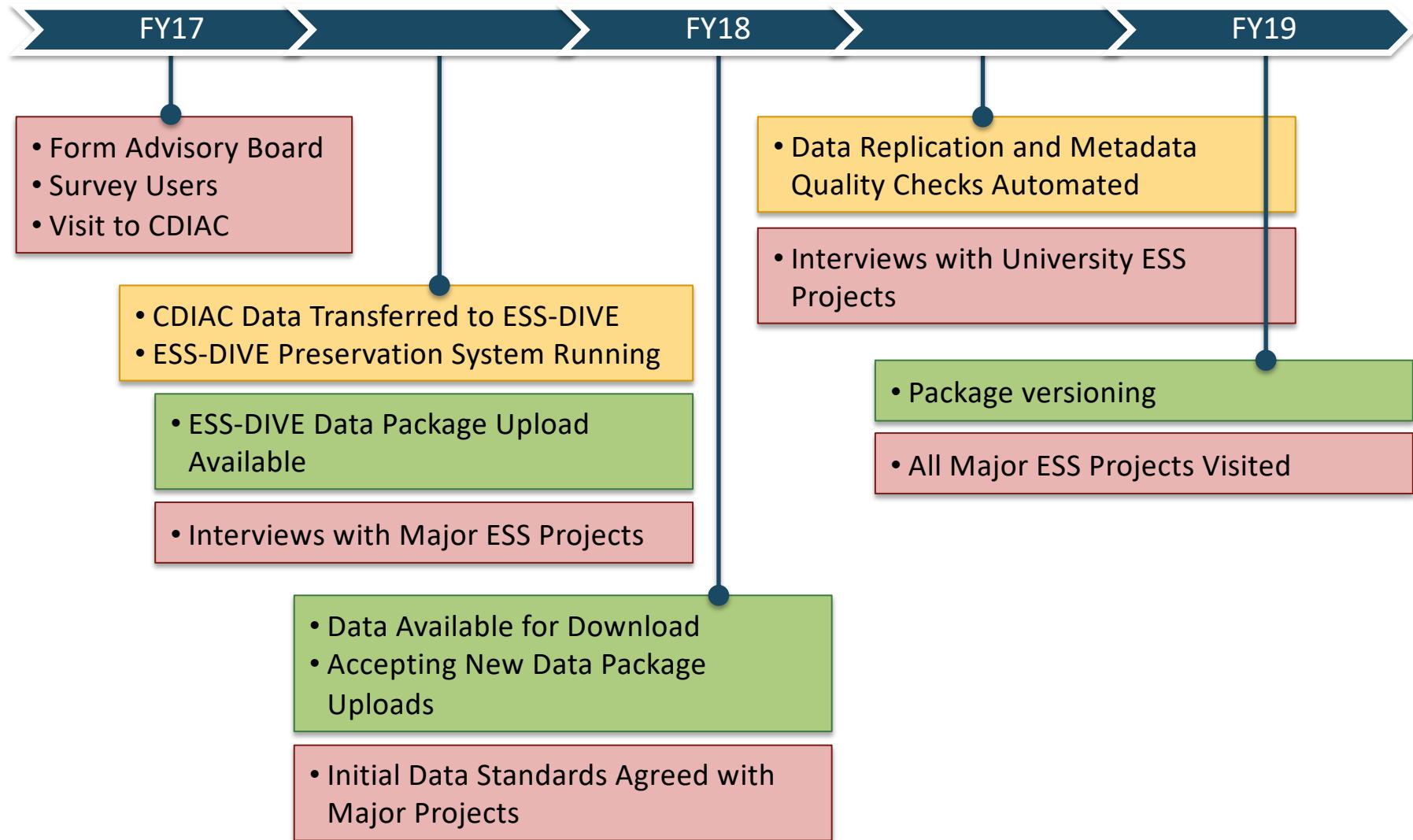
## 2. Data Providers submit data to archive with standard templates

INITIALS*	START_DATE*	END_DATE*	EVENT_TYPE*	EVENT_DESCRIP*
Initials of person entering info	Start date of event	End date of event; if single day event, report same day as start date	Type of event being reported	Short description of even
Abbrev. from General_Info	YYYYMMDD	YYYYMMDD	LIST	free text

## 3. Climate modeler writes R script to read metadata templates and data files

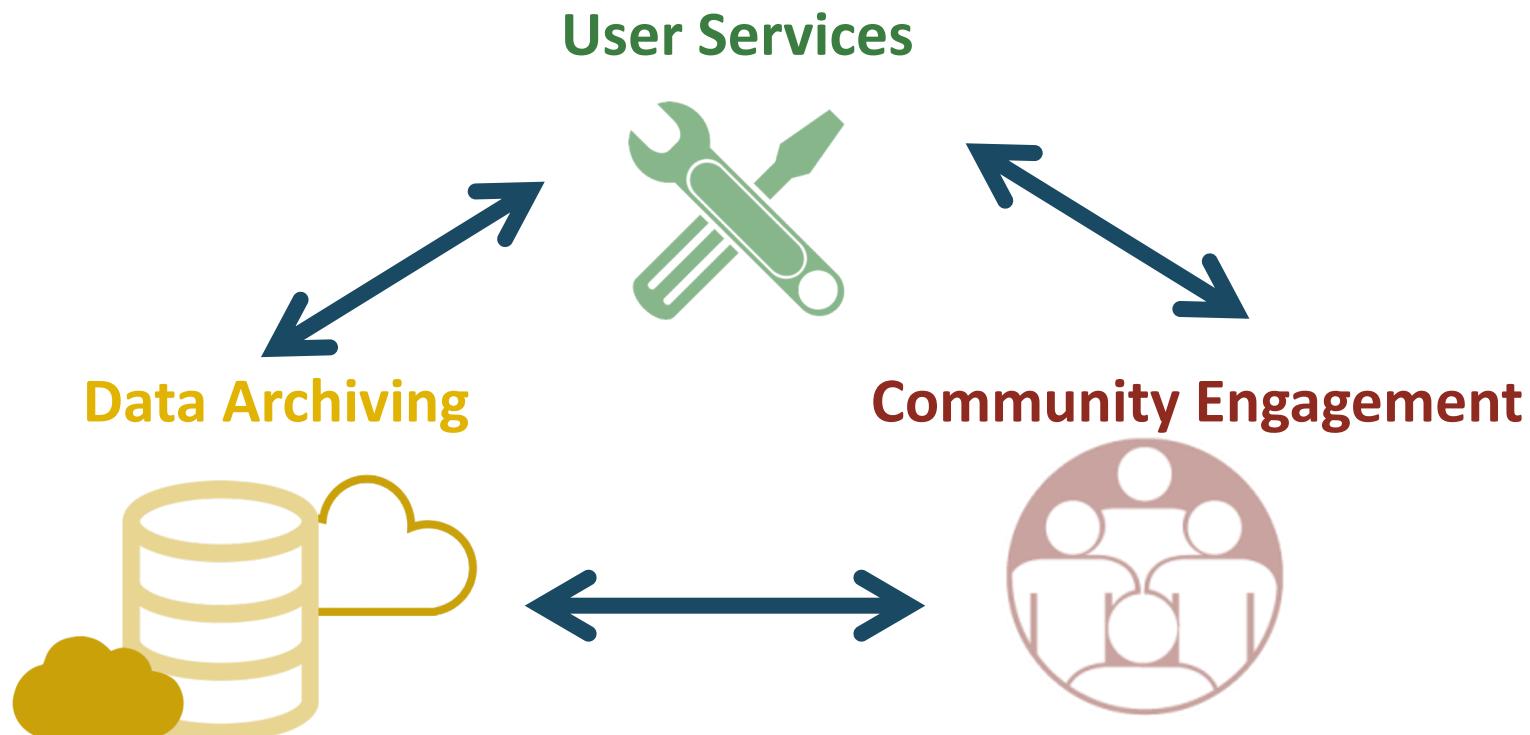
```
## ENSO 2015-2016
# Read in R objects generated from the "ENSO_data_collecting.R" script
# This script reads data and metadata pertaining to meteorology, sap flux, and soil moisture for Manaus site
## B. Christofferson (broadley@anl.gov)
## December 2016
#
library(rchronon)
library(fields)
rm(list=ls())
#
## Paths and filenames
rootdir <- "C:/Users/296342/Documents/"
path.enso <- paste(rootdir, "Data/ENSO/2015-2016/ENSO", sep="")
path.sap <- paste(rootdir, "Data/ENSO/2015-2016/ENSO/SapFlux", sep="")
path.mon <- paste(rootdir, "Data/ENSO/2015-2016/ENSO/mon", sep="")
path.mon.sapIC <- paste(path.mon, "Manaus_20160728/Gratner (Sap velocity)", sep="")
path.mon.sapIC <- paste(path.mon, "Manaus_20160728/ICT (Sap velocity)", sep="")
#
## Load the R object containing the dataframes and lists containing data and metadata
load(file=paste(path.enso, "ENSO.RData", sep=""))
## (Top Level) ::
```

# Project Timeline



# Berkeley Lab + NCEAS poised to implement ESS-DIVE

- Initial versions of software elements available
- Digital library experience
- Immersed in the ESS Community



---

# Questions