



# 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**

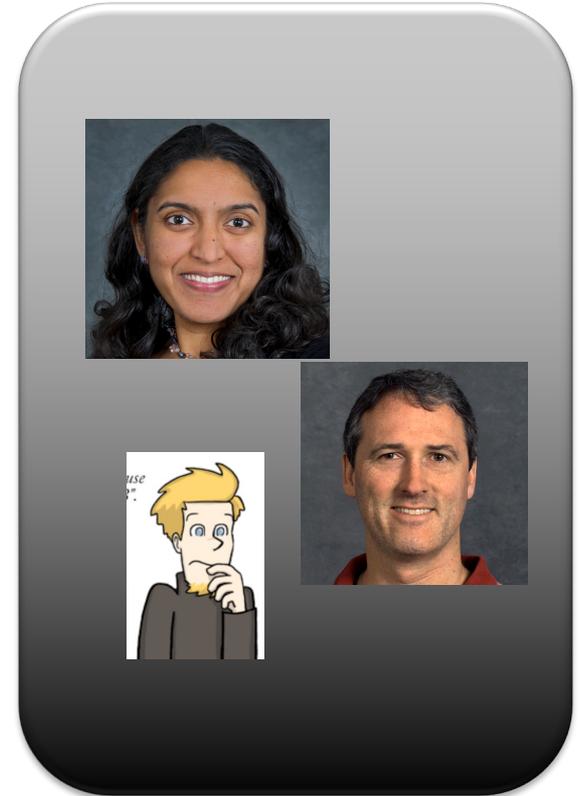
# Our Team is Well Positioned to Build the Archive



**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

**ESS PROJECT DATA**

The diagram illustrates the state of ESS project data storage in 2017. It features a central box titled "ESS PROJECT DATA" containing several icons: a stack of four black disks, a stack of four colored disks (yellow, orange, blue, green), a stack of four blue disks with a white label "CDIAC (2TB)" and a large red "X" over it, the Google Drive logo, the Dropbox logo, and a red server rack icon. Below the icons, the text "Most data not standardized" is written.

Google Drive

Dropbox

Most data not standardized

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

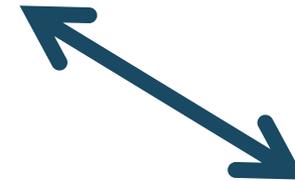
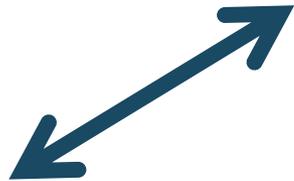
# 2020



# Three-Pronged Approach to Developing ESS-DIVE

---

## User Services



## Data Archiving



## Community Engagement



# 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



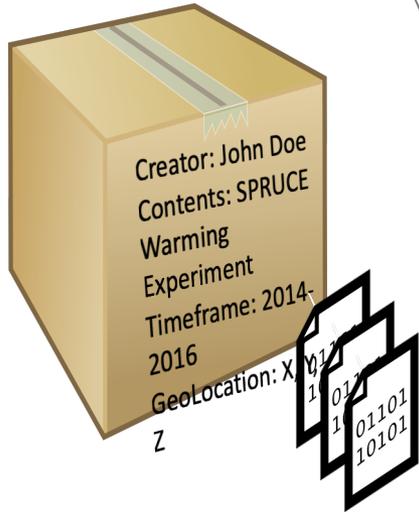
- Leverage existing **digital library** developments and experience
- Experience **migrating archives**
- Data team for several **ESS projects** including: AmeriFlux/FLUXNET, NGEETropics, Watershed Function SFA, etc.
- Activities in **standards and ontologies** such as Ecological Metadata Language (EML)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



Submitter

Date

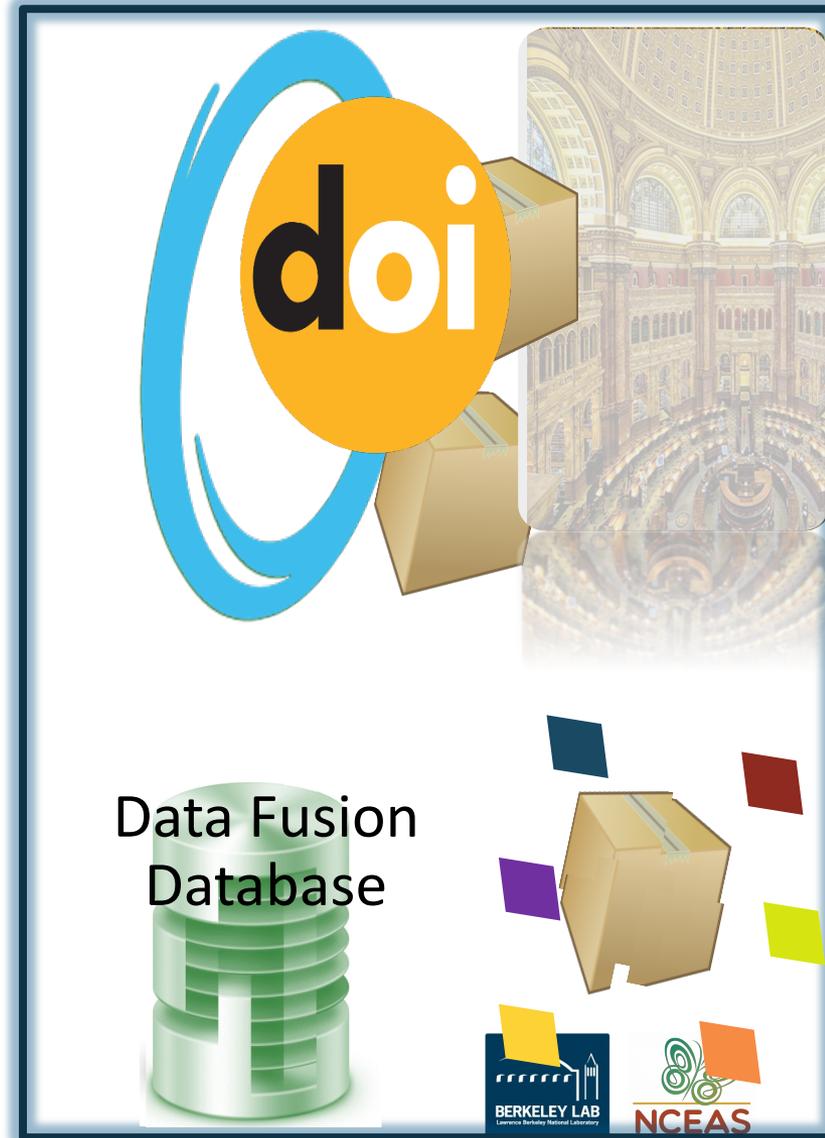
Tags



## ESS-DIVE Data Upload

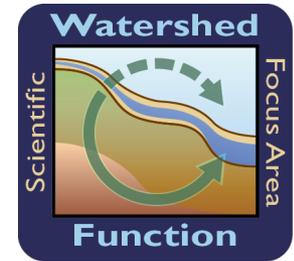
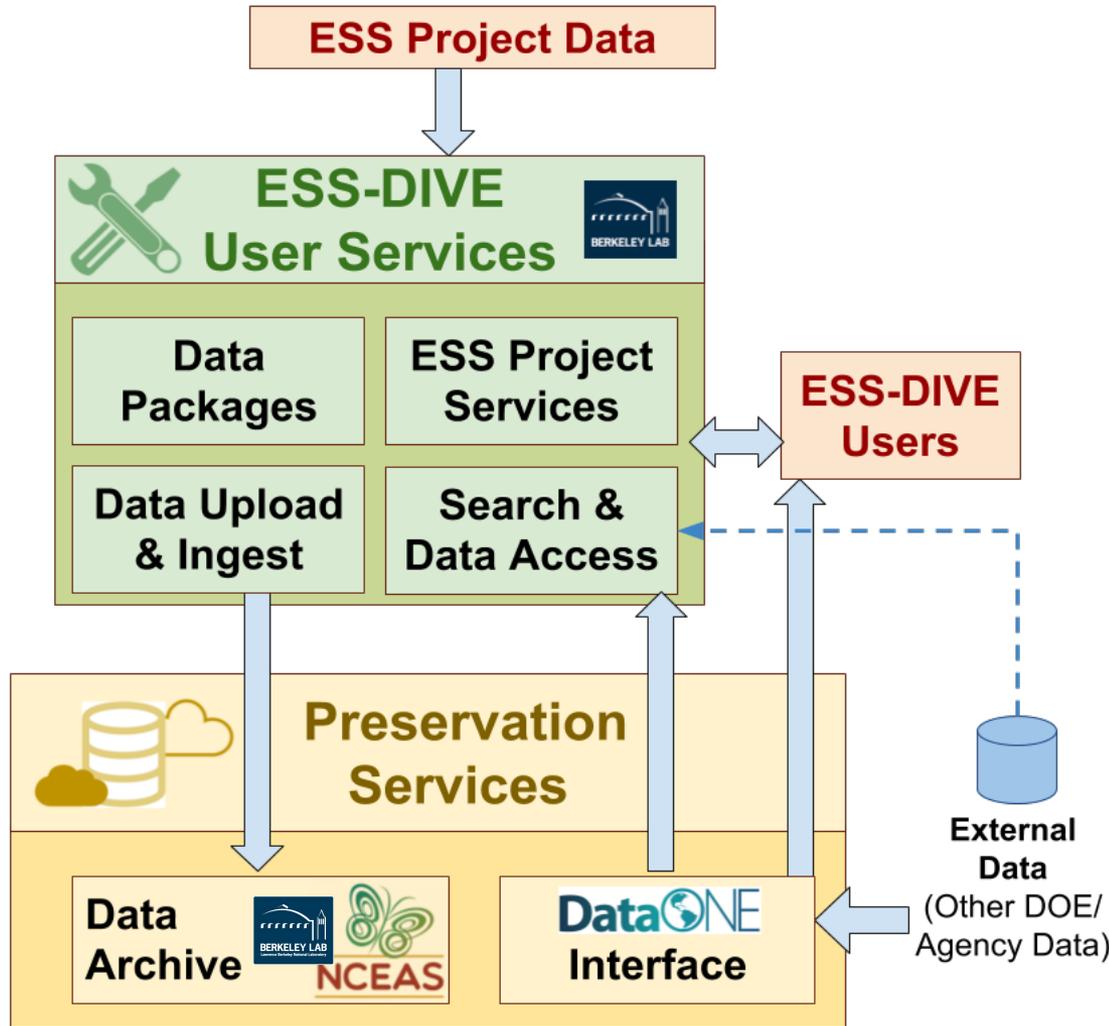
ESS-DIVE Portal

ESS-DIVE Archive



# ESS-DIVE Archive Software Architecture

## Builds on Existing Software

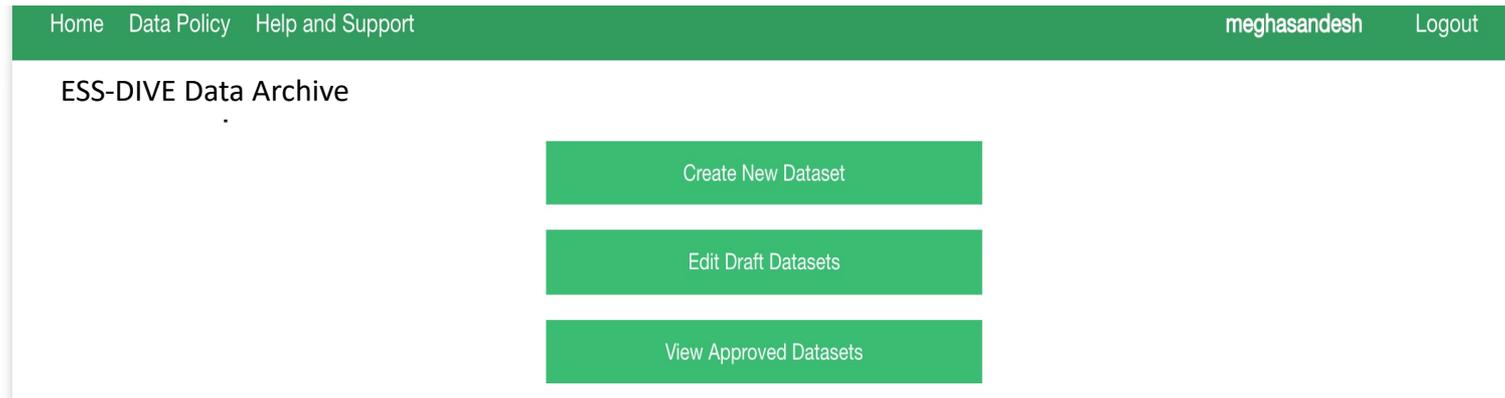


NGEE Tropics

# ESS-DIVE User Portal Starting Point in ESS Developed Software



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



- Advanced data search and visualization (fusion DB)

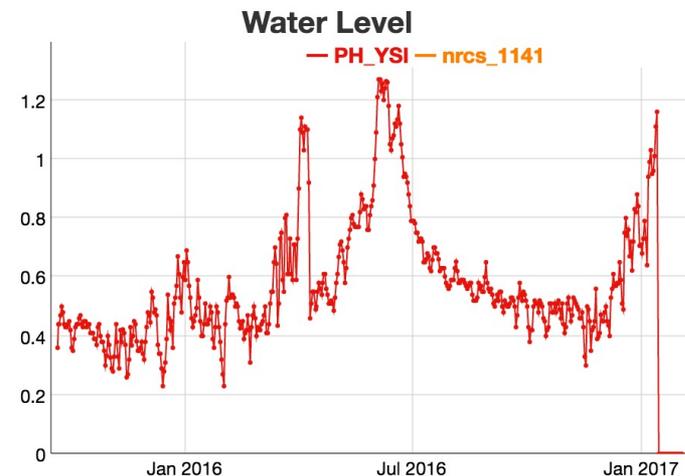
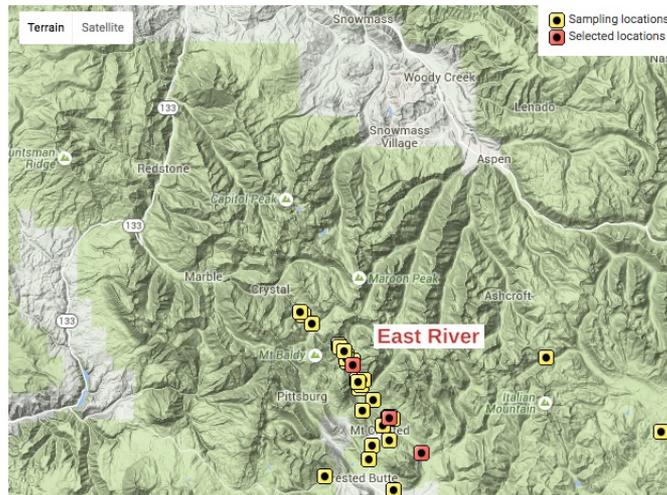
**SITES**  
 East River  Rifle

**LOCATIONS**  
Avery, BCC\_ISCO, PH\_ISCO

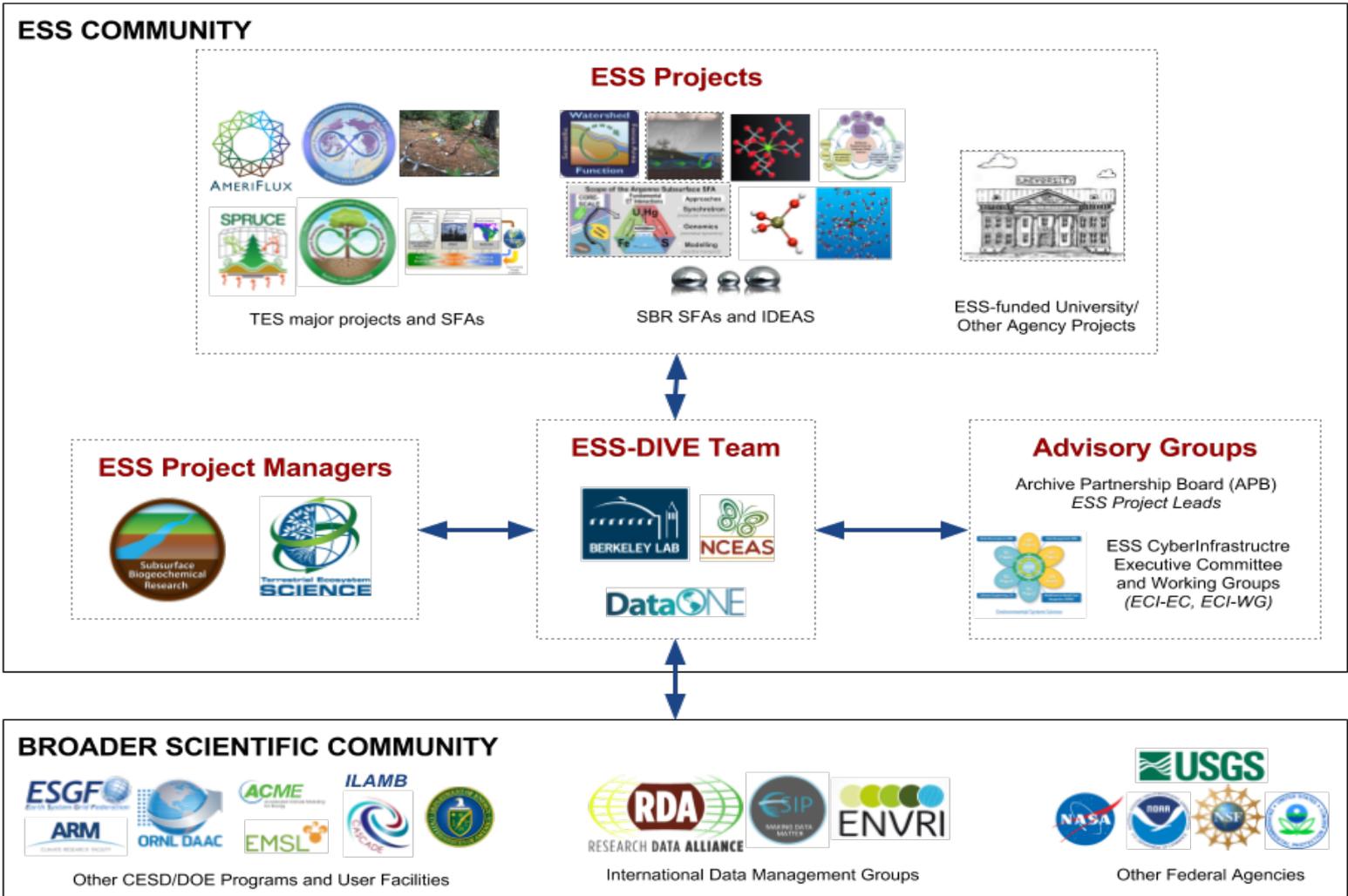
**PARAMETERS**  
Water Temperature, Nitrate (NO3)

**TIME PERIOD**  
01-2014 TO 05-2016

**SUBMIT** [Reset all options](#)



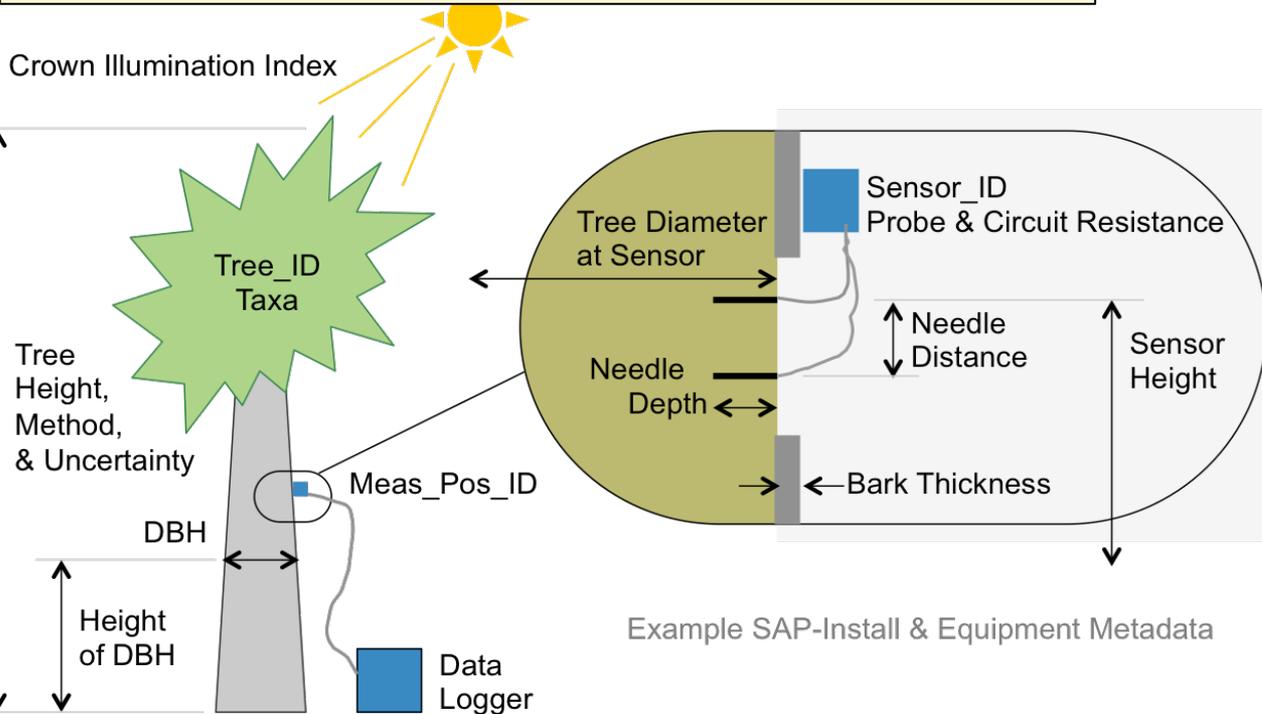
# ESS-DIVE Is a Partnership With the ESS Community



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



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



Example SAP-Install & Equipment Metadata

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

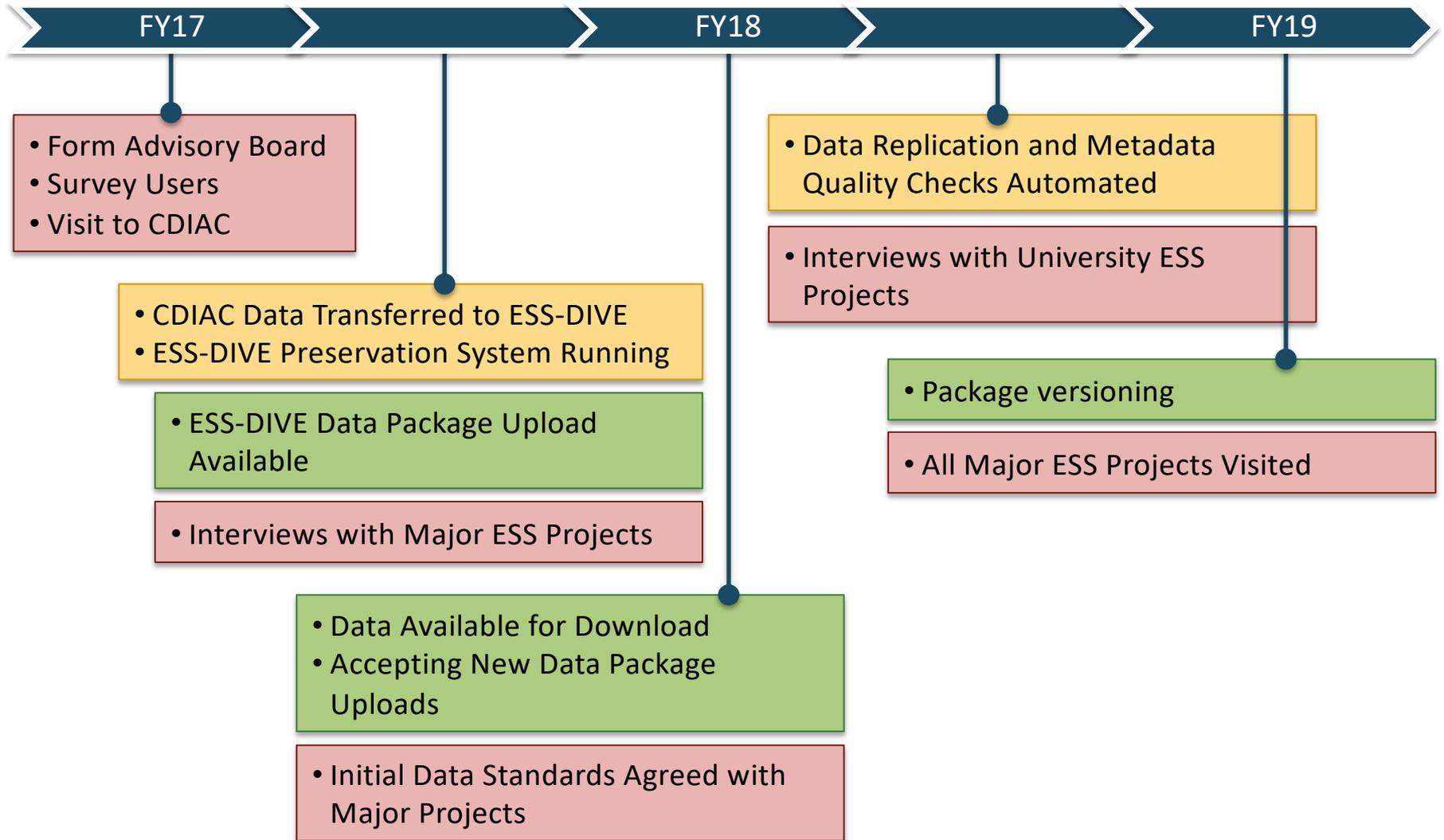
```

RStudio
ENSO_plotting.R
1 ## ENSO 2015-2016
2 ## Read in R object generated from the "ENSO_data_collating.R" script
3 ## containing data and metadata pertaining to meteorology, sap flux, and soil moisture for Manaus site
4 ## Makes plots of sap flux data
5 ## B. Christoffersen (bradley@lanl.gov)
6 ## December 2016
7
8 library(chron)
9 library(fields)
10 rm(list=ls())
11
12 ## Paths and filenames
13 rootdir <- "C:/Users/296342/Documents/"
14 path_ens0 <- paste(rootdir, "Data/NEE-Tropics/ENSO/", sep="")
15 path_man <- paste(path_ens0, "Data/20161211/20162302/Data/IR-Me2-Manaus/", sep="")
16 path_man_sapfl <- paste(path_man, "Manaus_20160720/Granier (Sap velocity)", sep="")
17 path_man_sapICT <- paste(path_man, "Manaus_20160720/ICT (Sap velocity)", sep="")
18
19 ## Load the R object containing the dataframes and lists containing data and metadata
20 load(file=paste(path_ens0, "ENSO.RData", sep=""))
21
22 ## ----
23 (Top Level)
  
```

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

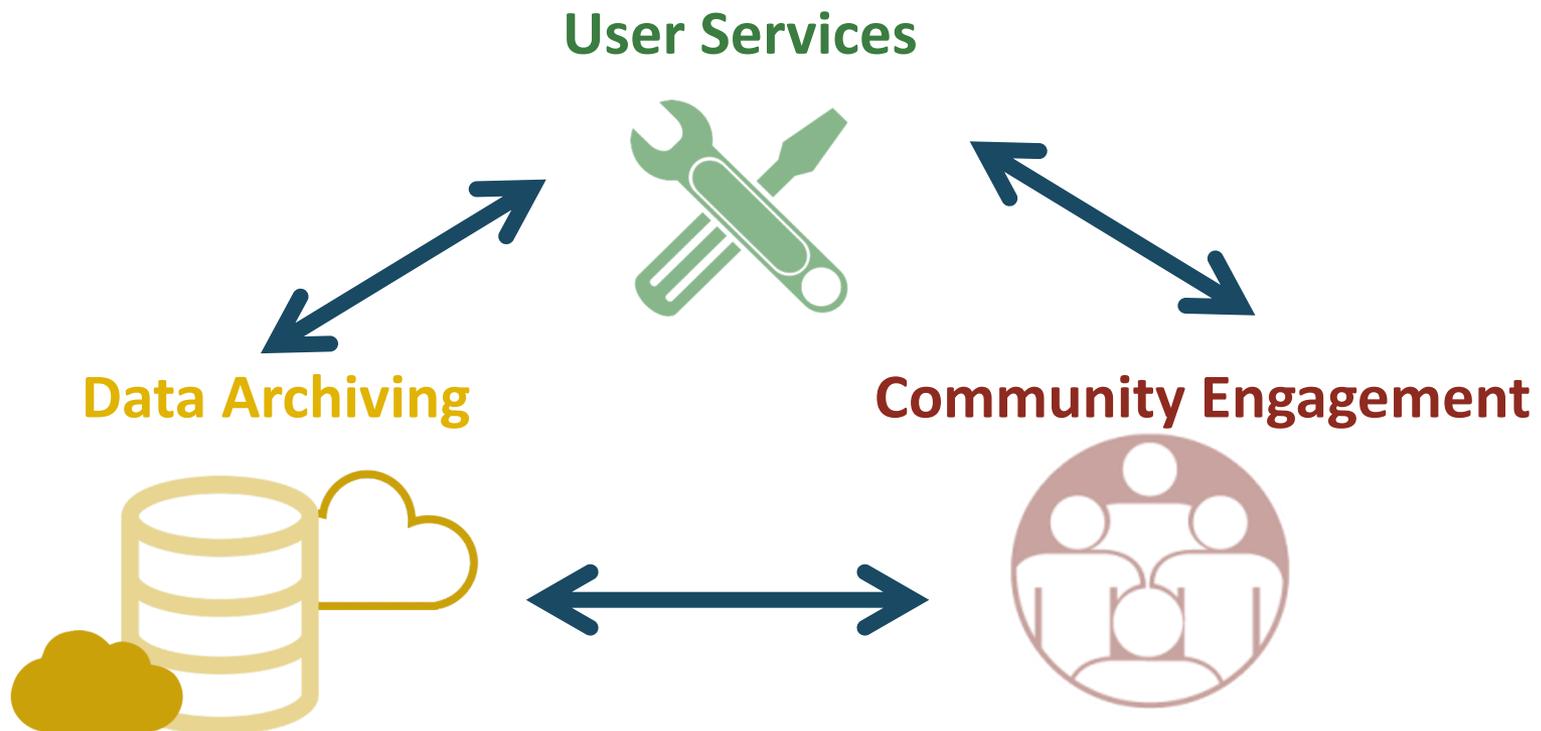
INITIALS*	START_DATE*	END_DATE*	EVENT_TYPE*	EVENT_DESCRIP*
Initials of person entering info	Start date of event	End date of event; if single day event, report same day as start date	Type of event being reported	Short description of event
Abbrev. from General_Info	YYYYMMDD	YYYYMMDD	LIST	free text

# 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