

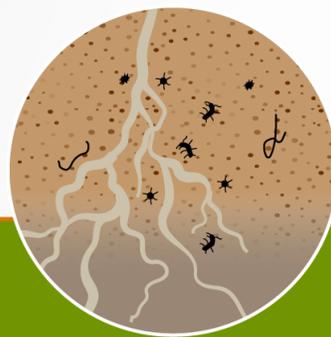
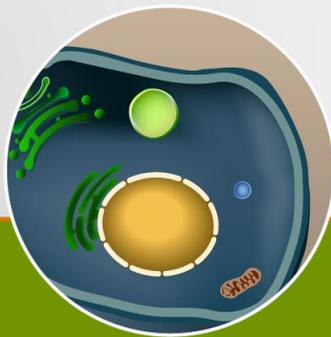
# Associating Data across Collaborative R&D Projects

September 17-18, 2019

Terry Law, Deputy of User Services

David Millard, Project Manager and EMSL Integrated Project Team Lead

Ken Auberry, Senior Data Scientist



SCIENTIFIC INNOVATION THROUGH INTEGRATION

# DOI construct needed to identify contributors and link to originating source



- Why do we care?
- Current structure of a DOI
- Exposing a collaborative project
- Referencing the collaborative project



**DIGITAL  
OBJECT  
IDENTIFIER**

# Current collaborative project dilemma



- Joint program (FICUS) between EMSL and JGI provides access to both facilities with a single proposal award
- Each facility generates and retains their own data in separate repositories
- Need method to assign 'award-level' DOI for increased visibility and impact tracking



# Current state of DOI structure and identified need



## Current state:

- Registry (OSTI) assigns unique prefix to each facility/organization
- Facility assigns unique infix to capture its contribution
- Facility assigns format to infix
- Registry assigns suffix
- DOI resolves to landing page

## Future state:

- Identify common format for infix to provide collaborative context
- Agree on landing page process



OSTI.GOV



Data ID Services



DataCite

FIND, ACCESS, AND REUSE DATA

# Proposed structure for consideration: example



- Prefix identifies collaborative project coordinator
- Infix assigned BY collaborative project coordinator to identify:

Program

Call ID

Data type

Year (e.g., award year, released year, etc.)

Unique system ID

## Example

Prefix assigned  
by registry  
to organization

Suffix assigned  
by registry  
to data

10.25582/FICS.170.proj.2019.12345/1105143

Customizable infix assigned  
by organization

# Proposed structure for consideration: award attributes



- Organization: DOI requestor
- Program: Identified in shared lists
- Program ID: from program coordinator system
- Data Type: shared lists (e.g. proj, data, inst, etc.)
- Year: need to agree on convention (e.g., fiscal or calendar year; 2 or 4 digits (*FYXX*, *FYXXXX*, etc.))
- Unique system ID: from program coordinator system
- Resolving URL → ?

## Example

Prefix assigned by registry to organization

Suffix assigned by registry to data

10.25582/FICS.170.proj.2019.12345/1105143

Customizable infix assigned by organization

# Lists, agreements, and access needed across organizations to enable collaborative infixes

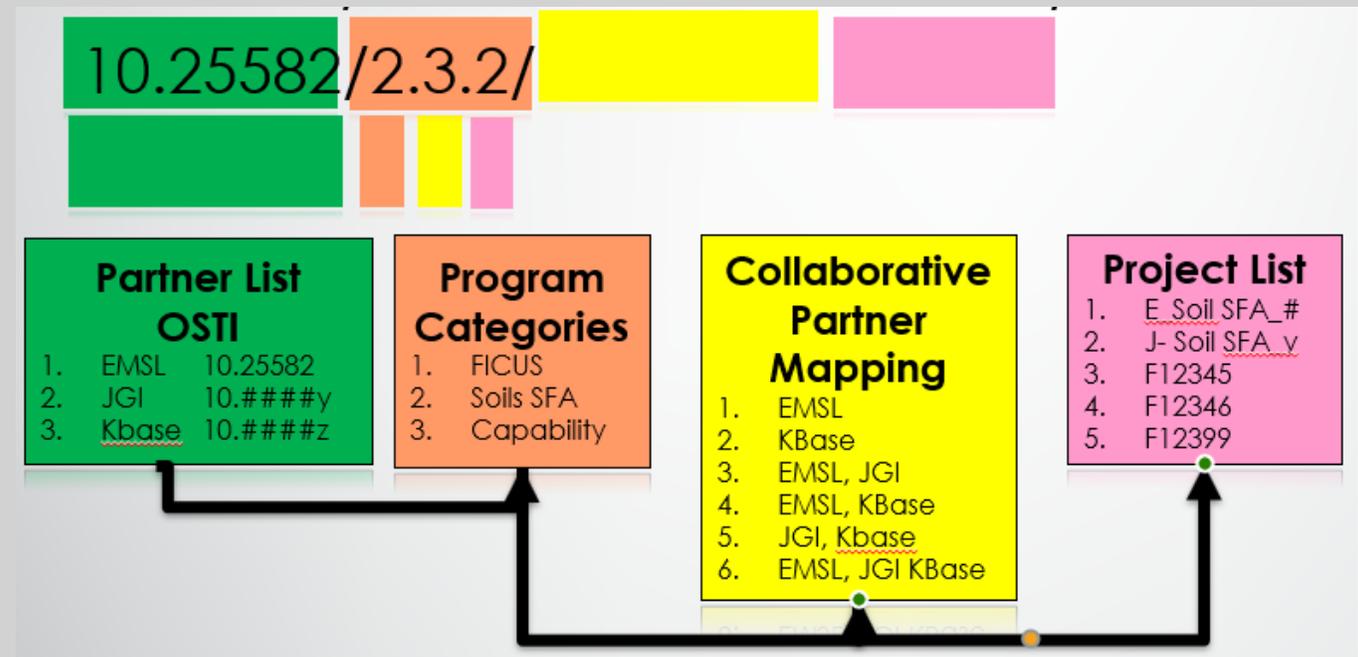


- Agreement on enumeration method
- Provision of enumerated lists and sets
- Requires service to access enumerations and support sort & filter functions
  - Shared enumeration of joint program list
  - Shared enumeration of partner sets
  - Shared enumeration of project lists

## One example of enumerated lists:

- Collaborative program identifier
- Collaborative partner mapping
- Collaborative project identifier

10.25582/FICUS.EMSL-JGI.F12345/[OSTI Suffix]



# Asking workshop attendees to help us define the possibilities



- Is the program POC responsible for resolving the award URL and providing the program landing page?
- Is each contributor only responsible for their contributions?
- How prescriptive does a collaborative program need to be re: providing an infix?
- Do we have enough demand for an external consumer (like data.gov) to provide a landing page at the collaborative program level?

DOE  
Data Explorer

