

Community of Practice - Data Citation
Use Case

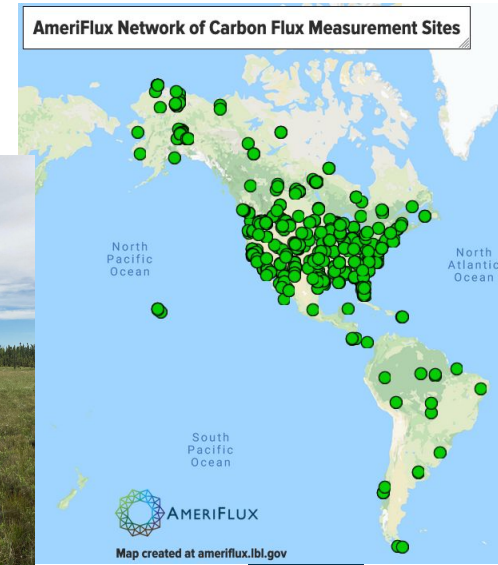
Large Collaborative Earth Science Projects (LCESP)

Deb Agarwal
AmeriFlux/Watershed SFA/NGEE Tropics/ESS-DIVE

LCESP Context

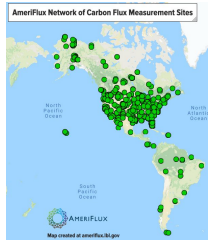
Large collaborations made of up many independent science teams

- Science teams are collecting datasets and publishing them
- 100s of datasets are used together as part of many analyses
- For any particular dataset only a subset of the data is used



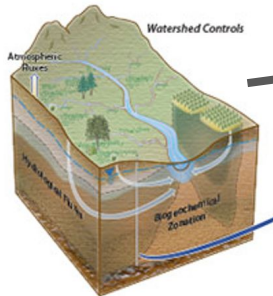
LCESP - Use case

- Each dataset has a DOI and expected citation (typical license is CC-BY-4.0)
- Datasets used in an analysis come from multiple repositories



210 datasets

70 datasets



1 dataset



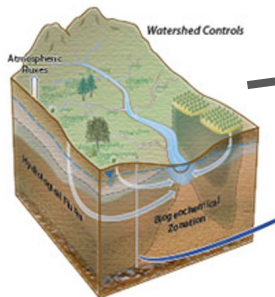
LCESP - Use case detail

- Elements of a dataset do not have PIDs assigned



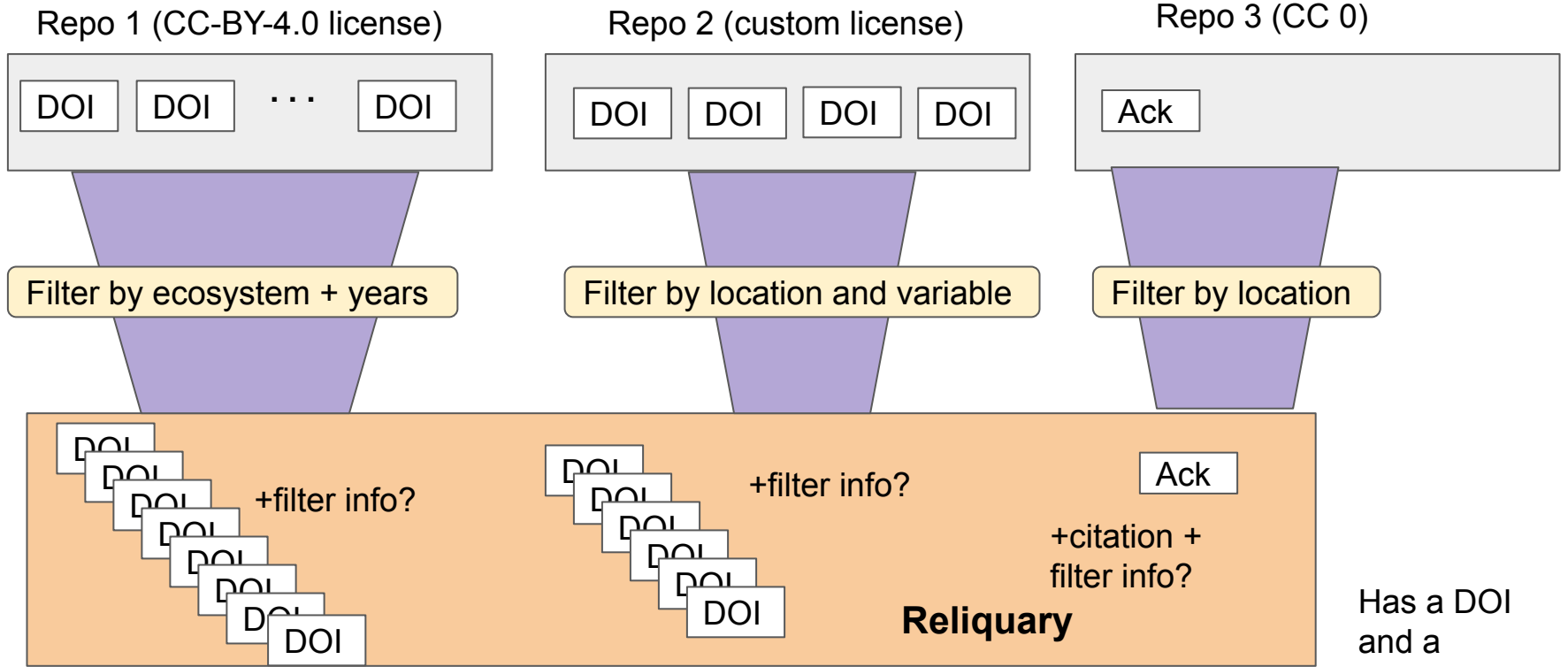
210 datasets (only
2005 - 2010 data)

55 datasets (2008 -
2012 data)



Projects do not have the resources to make a reliquary per user

Use Case Diagram



Characteristics of the use case

- Each user applies different filters
- The data filter can sometimes just be a list of sites
- The set of repositories and number being used varies
- Sometimes the bulk of the data is from a single project/repo but often it is spread across a few

