

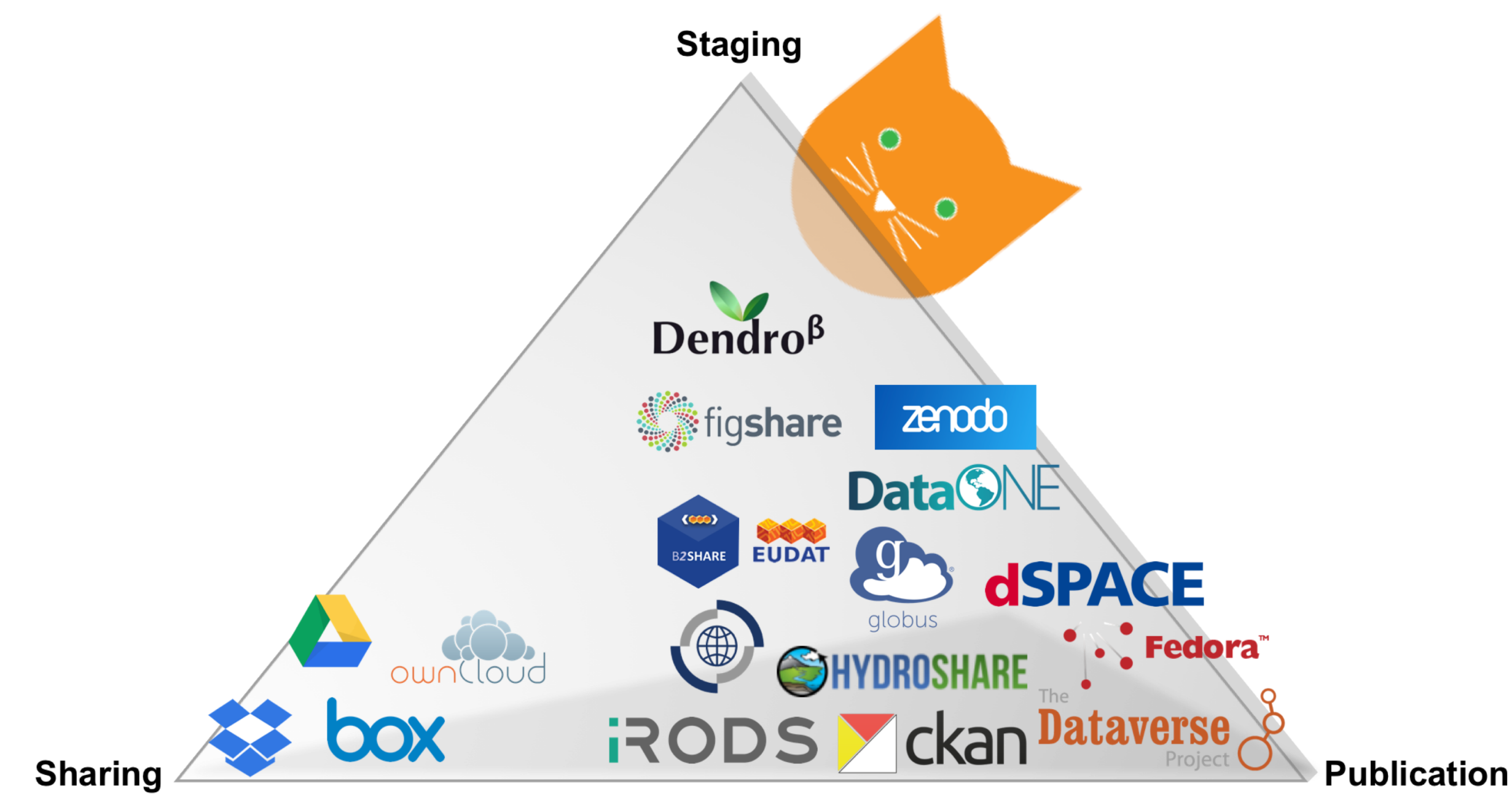
Clowder Open Source Customizable Research Data Management

Kenton McHenry, Klara Nahrstedt, Praveen Kumar, Noah Fahlgren, Luigi Marini

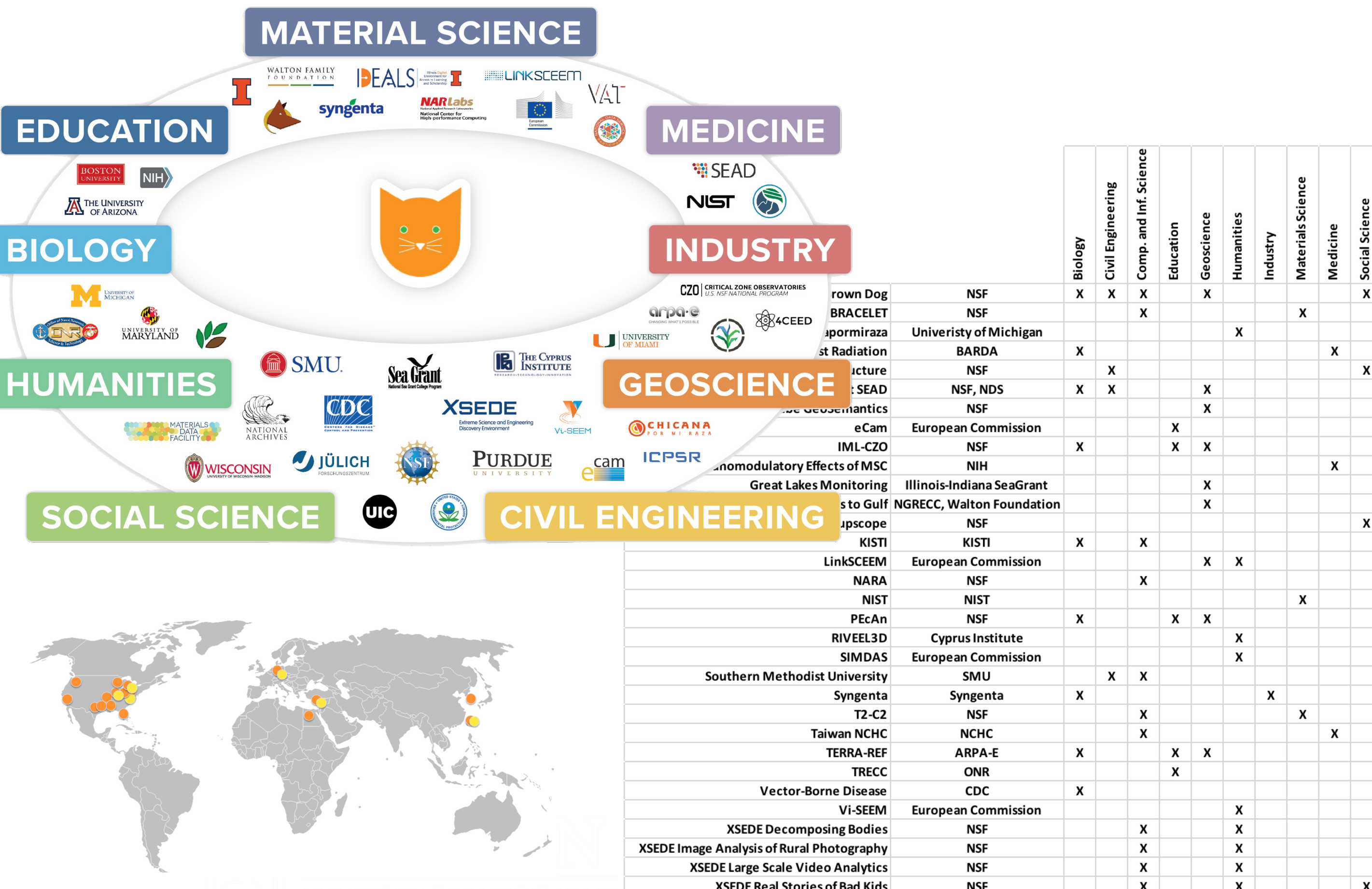
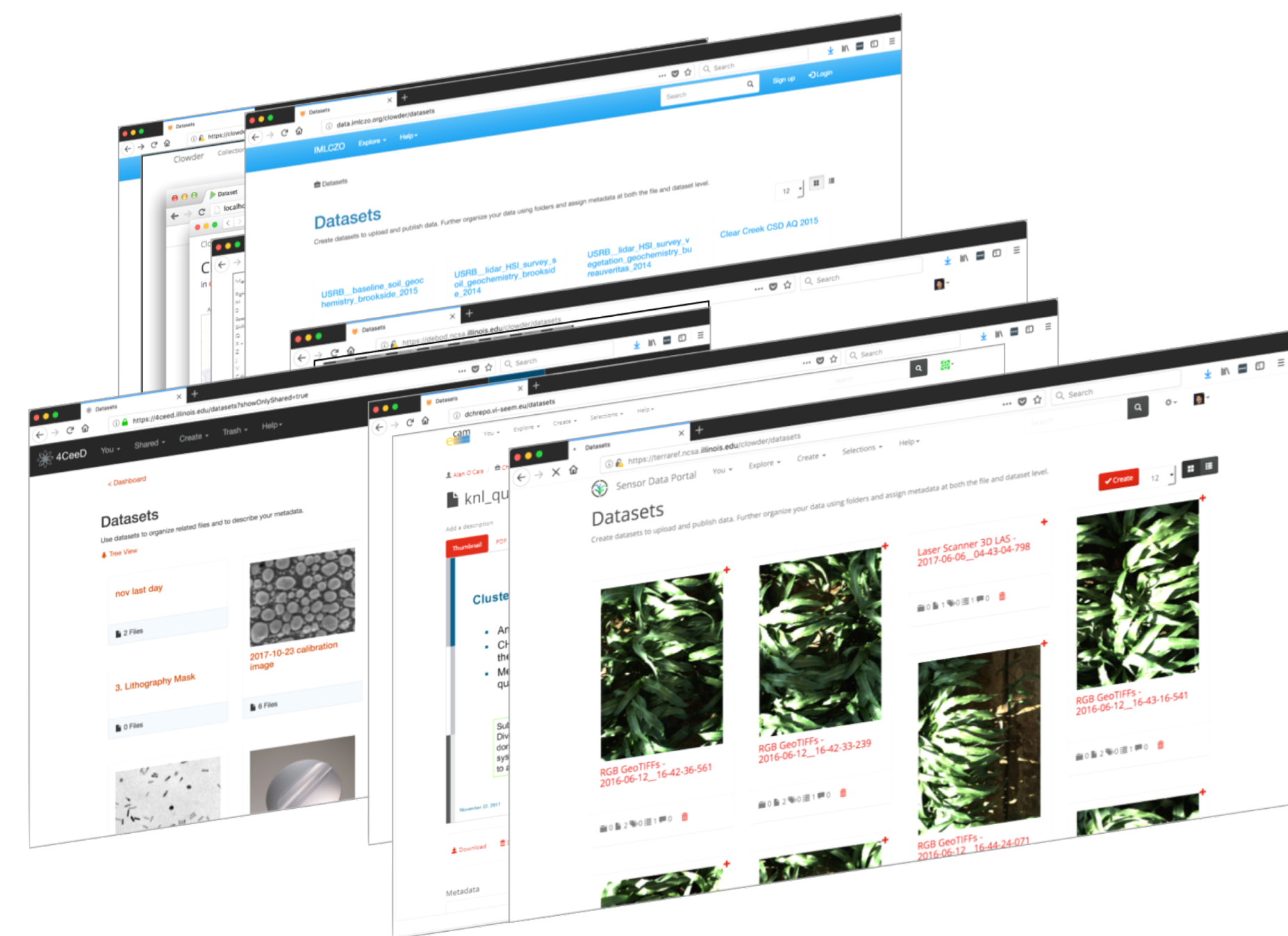


Active curation involves recording data and metadata as close to the source as practical and driving that acquisition through the deployment of capabilities that help data producers manage their research.

Think "Smart Opensource Dropbox"

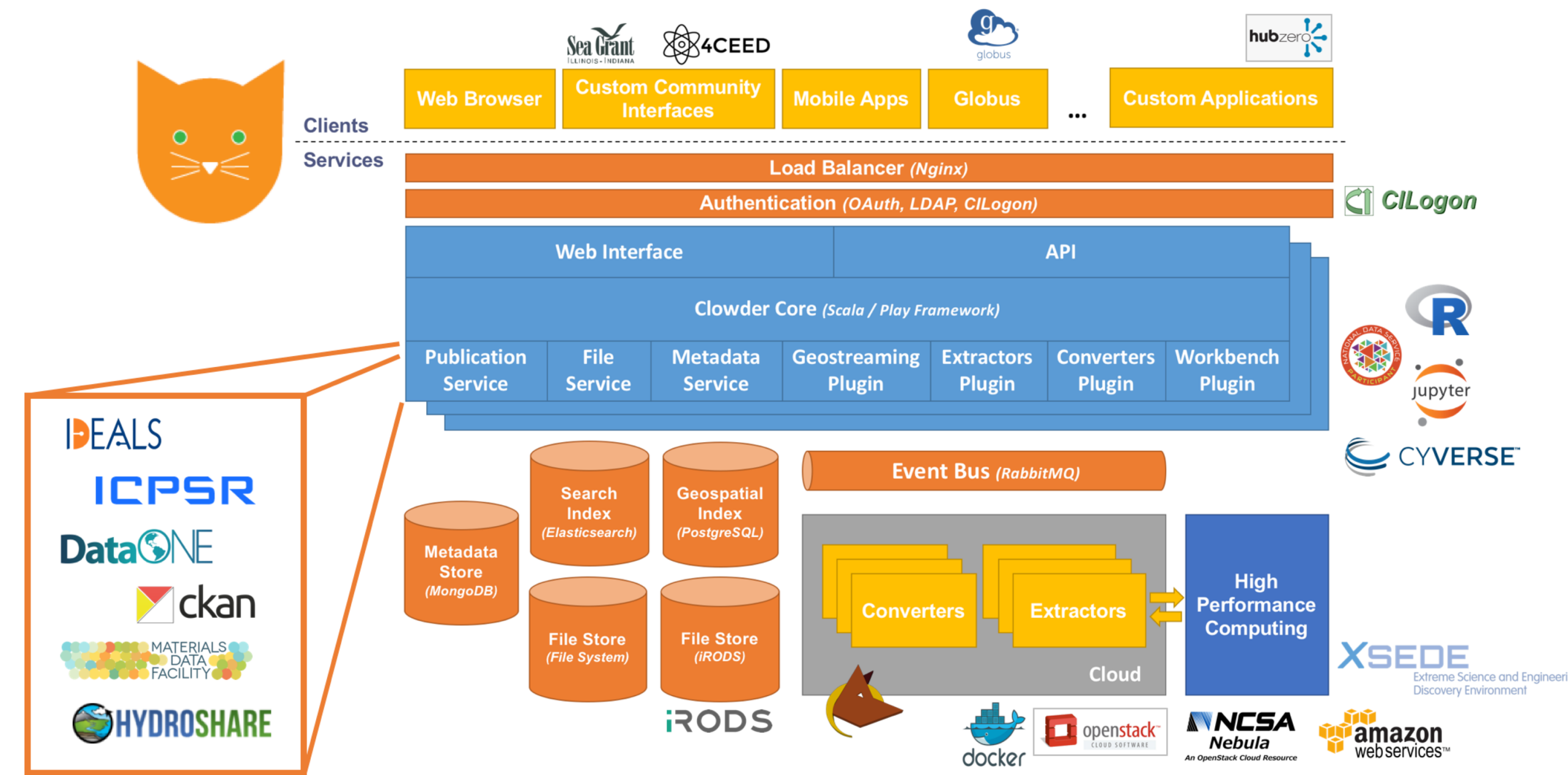


Open Source Community



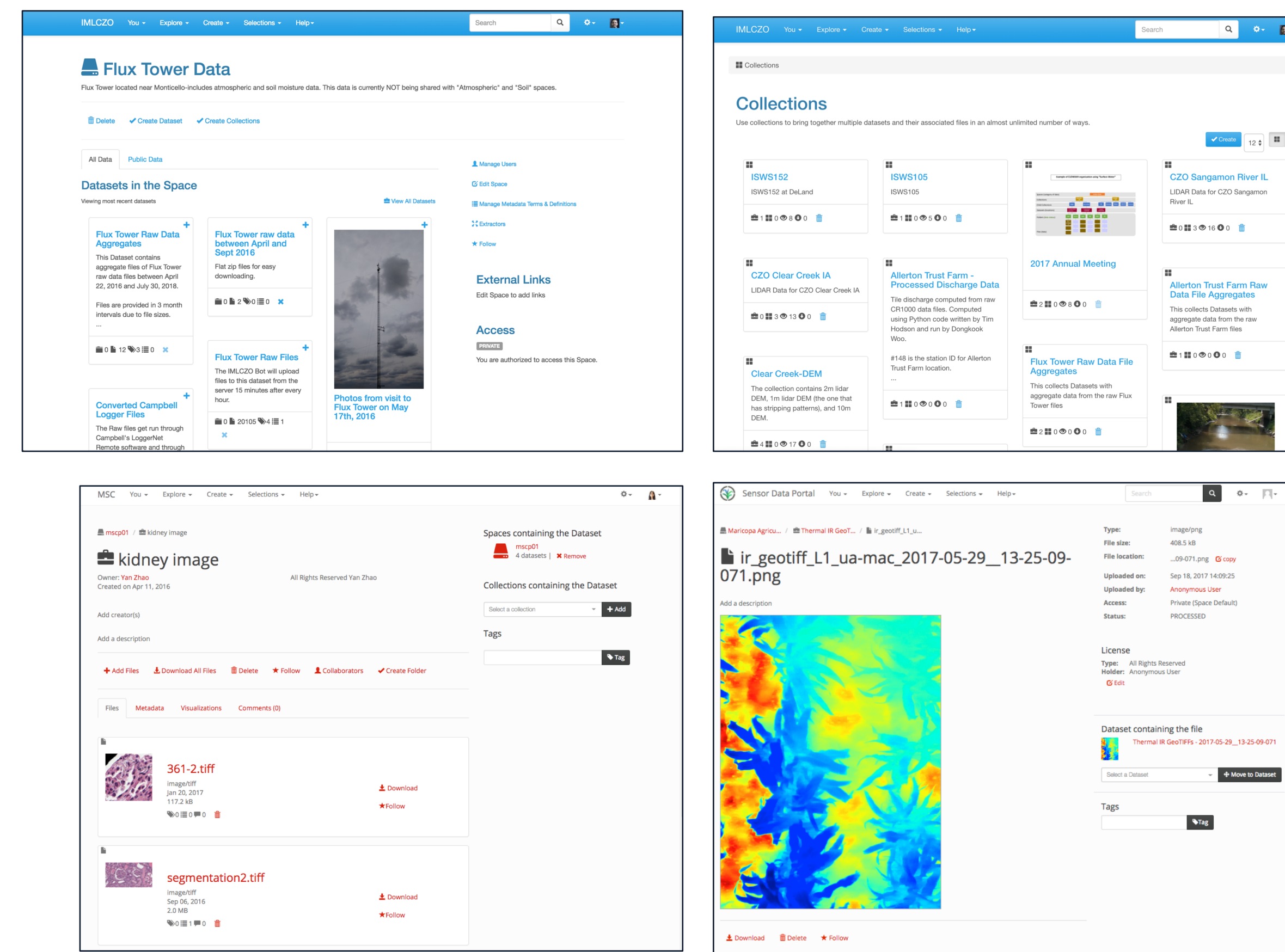
Architecture

Clowder is built from the ground up to be extensible and support new use cases and domains.

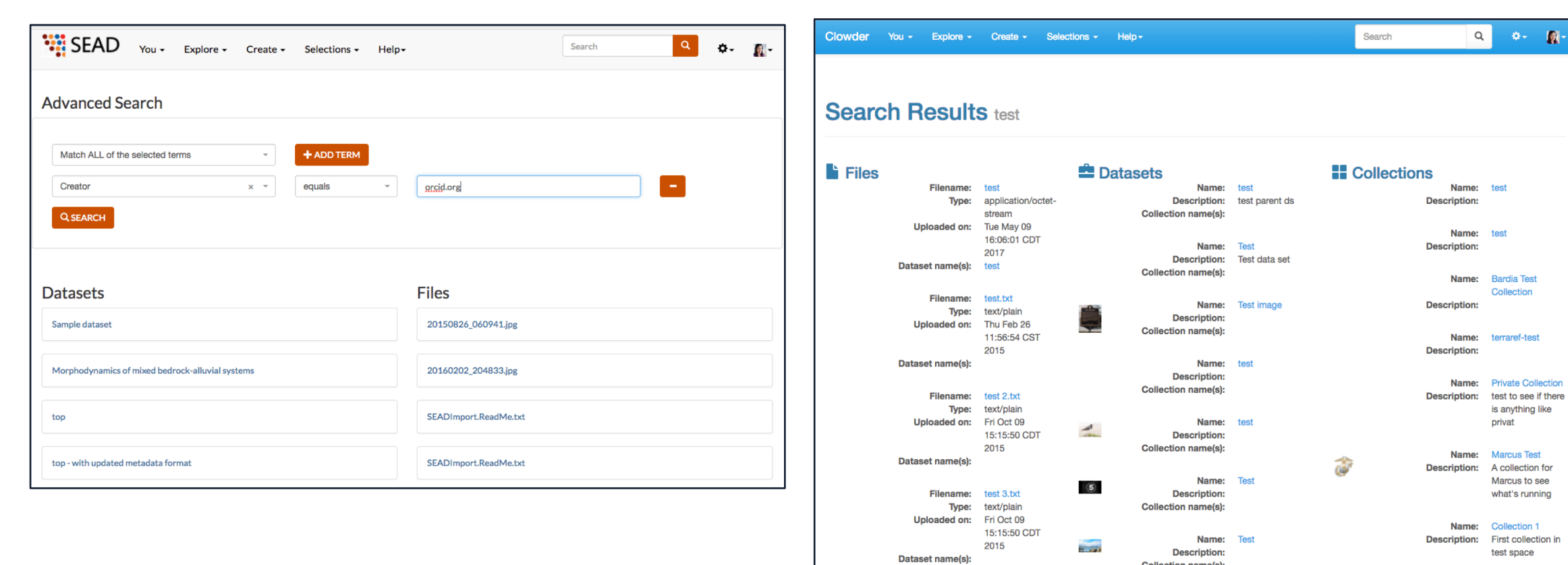


- The scalability of the Clowder system has been proven with one instance (TERRA-REF) having 1 PB of data and close to 40 million files.
- The Brown Dog project runs 60 different extractors as services in a docker swarm, which consists of 30 machines. Extractors are scaled elastically based on the number of requests and the swarm runs around 150 instances of these extractors.

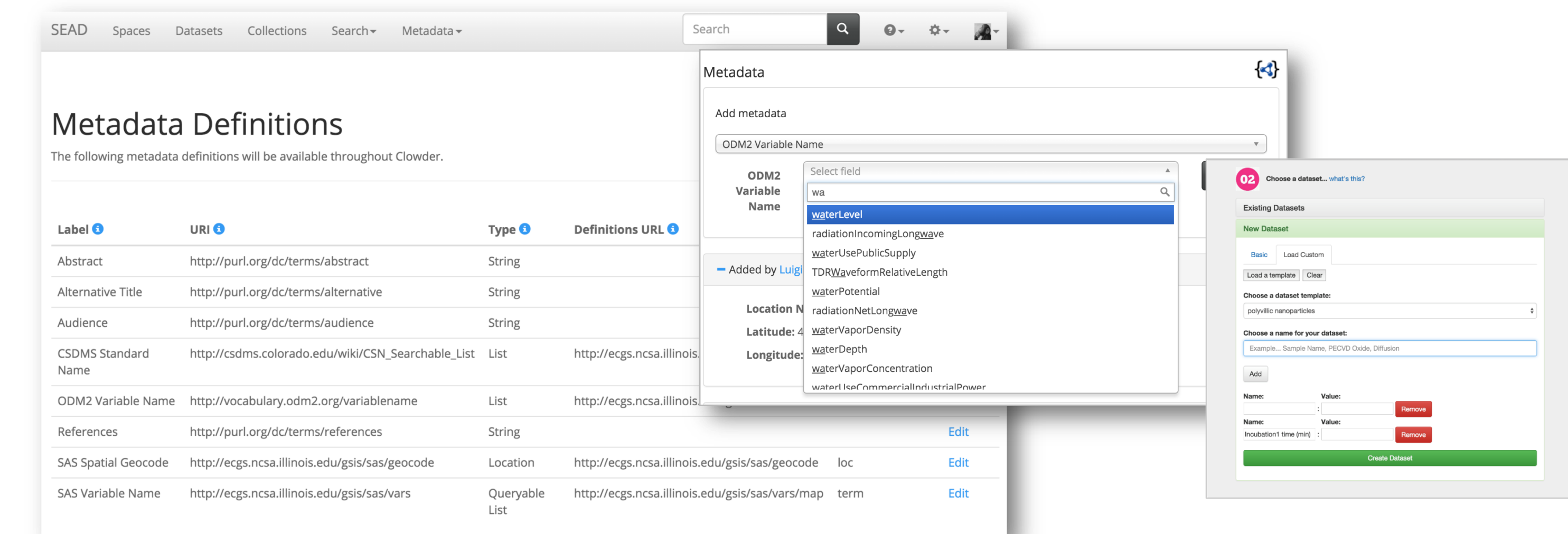
Organize Data Spaces, Collections, Datasets, Files



Search



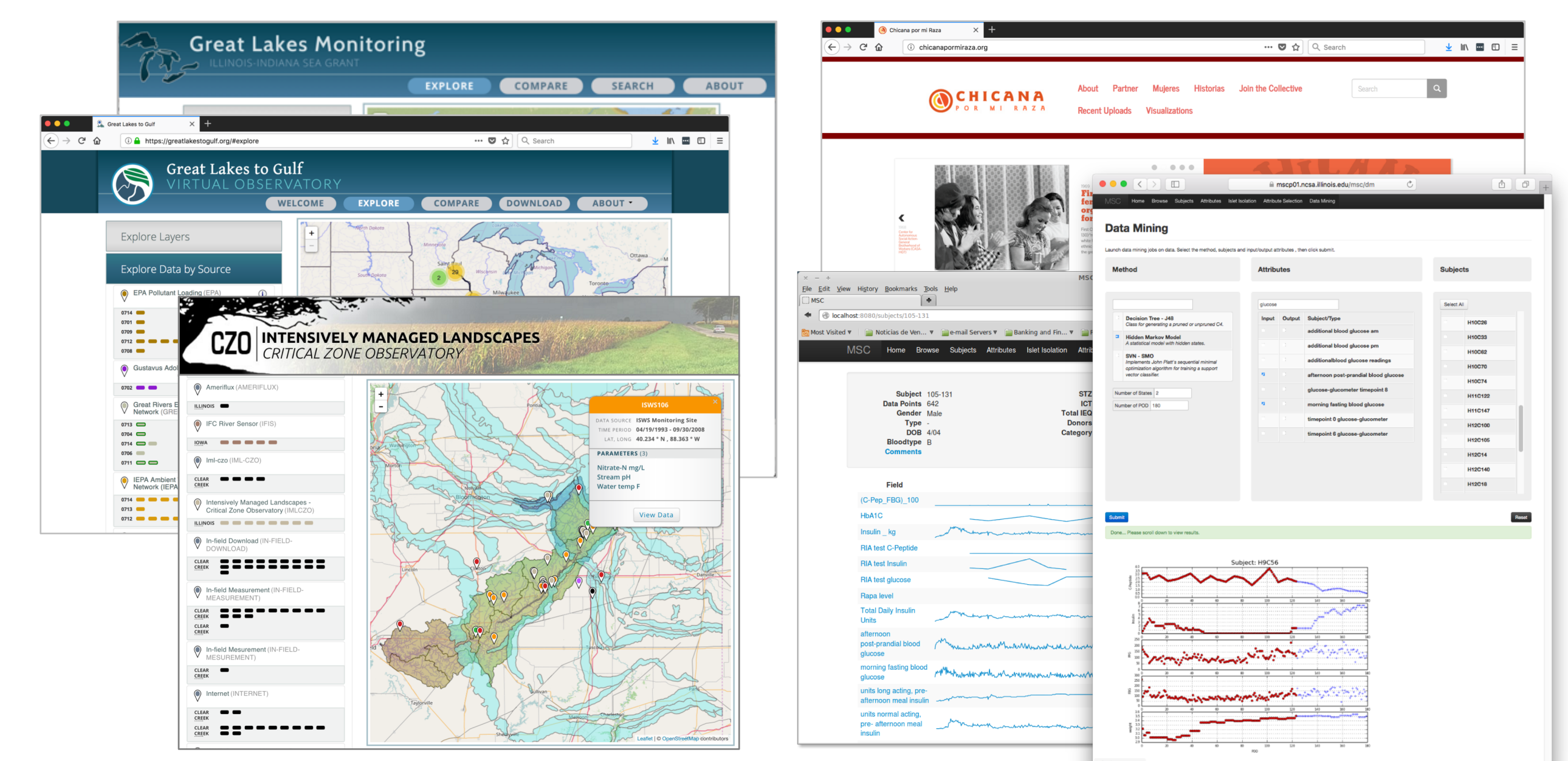
Flexible Metadata



Extensible Extraction & Visualization



Custom Dashboards



Clowder has been funded by: NSF BCS-0941268, EAR-331906, ACI-1261582, ACI-1443013, OCI-0940824, OCI-0525308, OAC-1835834; NIH 1P01AI089556-01A1, Illinois-Indiana SeaGrant DW92329201, European Commission RI-261600, XSEDE OCI-1053575, ARPA-E DE-AR0000594.