DataONE
Data Observation Network for Earth

William Michener[1], Amber E Budden[1], Rebecca Koskela[1], Dave Vieglais[2] and the DataONE Team
[1]University of New Mexico, [2]University of Kansas. Contact: wmichene@lternet.edu

UNM KU

Addressing the Earth's environmental problems requires that we change the ways that we do science; harness the enormity of existing scientific data; develop new methods to combine, analyze, and visualize diverse data resources; and create new, long-lasting cyberinfrastructure. DataONE is a federated data network that improves access to, and preserves data about, life on Earth and the environment that sustains it.

DataONE Supports Science by:

- Engaging the relevant science, library, data, and policy communities;
- Facilitating easy, secure, and persistent storage of data;
- Disseminating integrated and user-friendly tools for data discovery, analysis, visualization, and decision-making.

Data Life Cycle

The eight steps of the Data Life Cycle provide the framework for the development, education and outreach activities of DataONE. Tools are designed to meet user needs in all areas of the DLC and training and outreach activities ensure users have access to the information and resources needed for planning, organizing and sharing their data.

plan, collect, assure, describe, preserve, discover, integrate, analyze

Education and Outreach

A suite of resources and training events build out from the DataONE infrastructure. These resources include a Best Practices Database of expert recommendations for data management; a Primer on Data Management; and a Software Tools Database. Designed as a community project, the DataONE Users Group guides the evolution of DataONE.

Investigator Toolkit (ITK)

The Investigator Toolkit provides a user friendly interface for seamless search and retrieval of data held within the DataONE network. The ITK also enables access to customized tools that are familiar to scientists and that can support them in all aspects of the Data Life Cycle.

R, MATLAB, DMPTool, zotero, Kepler, UC3DataUp, ONEMercury

Member Nodes (MN)

Member Nodes encompass a diverse array of institutions that serve as data centers or repositories including libraries, universities, research networks, and governmental and nongovernmental agencies.

LTER, PISCO, DRYAD, USGS science for a changing world, esa, Merritt, KNB, ONEShare, The Cornell Lab

Coordinating Nodes (CN)

Network-wide services enhance interoperability of the Member Nodes and support indexing and replication services. Coordinating Nodes enable scientists to discover networked data wherever they reside and make Member Node data and services more broadly available to the international community.

UNM, UCSB, UT

DataONE Cyberinfrastructure

www.dataone.org

Future Development Objectives

- Progressive integration of semantic capabilities to improve discoverability and reuse of data;
- Provenance tracing to better facilitate determination of the derivation of data and improve reproducibility of analyses;
- Integration with high performance computing facilities such as XSEDE to enable seamless transfer of data objects to computation resources, and archive of computational outcomes;
- Addition of notification and annotation services, so researchers may tag and otherwise annotate data sets to help with content reuse;
- Adapting more investigator tools to address all aspects of the data lifecycle;
- Ongoing expansion of capacity through addition of additional Member Nodes.