DataONE - Supporting Scientific Analysis for the Long Term

David A. Vieglais

DataONE, U.S.A.

The Data Observation Network for Earth (DataONE) is a project funded by the National Science Foundation (award #083094) in response to the DataNet solicitation. A primary goal of DataONE is to enable reliable, long-term reuse of data, especially catering for the earth sciences. It does so by providing infrastructure, promoting standards and best practices for data management, and simplifying access to a diversity of content curated by a heterogeneous array of participating data repositories. The core infrastructure operated by DataONE provides the underlying technical basis for several key services including: globally unique, scheme agnostic, resolvable identifiers for all content; replication of content between repositories; logging and content use and usage reporting for contributors and managers; mapping of user identities across multiple accounts while maintaining consistent access control; and extensible discovery services that operate over a variety of metadata formats. The Investigator Toolkit contains libraries, plugins, and applications that provide seamless interaction with core services and participating data repositories, making access to data from disparate repositories straightforward for end users. In this way, content creators and consumers enjoy direct access to a wealth of earth science data through desktop tools such as Excel, R, Morpho, and ONEDrive. These principles, practices, services, and tools combined support the full data lifecycle, and provide a foundation for additional systems and services to be built.