

The Role of Academic Libraries in DataONE: Engaging in E-Science through Project Partnership

Robert J. Sandusky¹, Patricia Cruse², John Kunze², Suzie Allard³, and Carol Tenopir³
¹University of Illinois at Chicago, ²California Digital Library, ³University of Tennessee, Knoxville



DataONE: Observation Network for Earth Addresses:

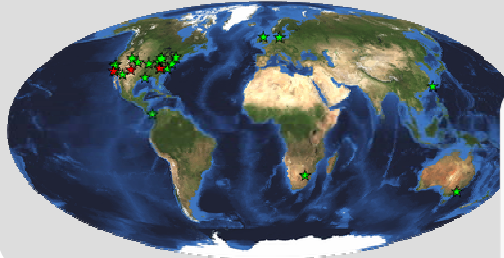
Data Loss: preserve at-risk (orphaned, poorly curated) biological, ecological, environmental data acquired from individual scientists and projects.

Data Dispersion: facilitate global discovery and access through a single, easy-to-use search portal.

Data Deluge: provide tools to empower scientists and organizations to easily and effectively manage, analyze, and synthesize data.

Data Practices: create an informatics-literate workforce through outreach and training

- On-line / in-person certificate programs
- Library-based instruction and outreach
- Exemplar data management plans, best practices guides
- On-line videos, podcasts
- Investigator Toolkit

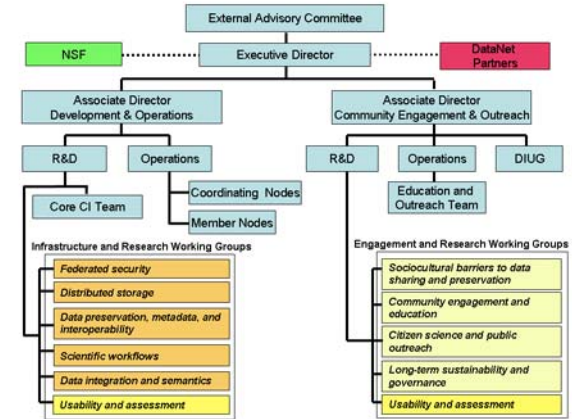


DataONE CyberInfrastructure is globally distributed: a flexible, scalable, and sustainable network initially configured with three Coordinating Nodes (red stars) and dozens of Member Nodes (green stars).

Standards compliant, secure data and metadata replication; data curation services supporting the complete data lifecycle; Investigator Toolkit integrated into existing applications for scientists.

Community Engagement

DataONE is a virtual organization comprised of infrastructure, community engagement, and research working groups; an External Advisory Committee; and an International Users' Group

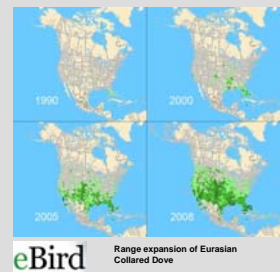


Use Case Scenario: Forecasting the Spread of an Invasive Species across North America

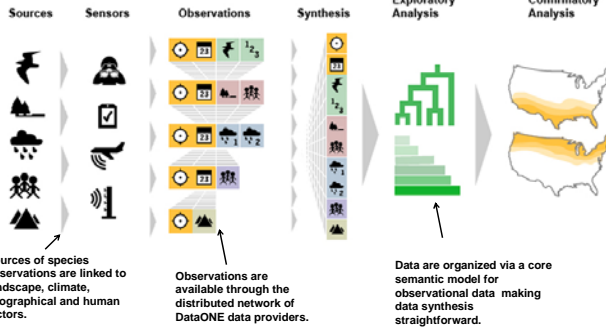
The Eurasian Collared Dove (*Streptopelia decaocto*) was introduced in the Bahamas in 1988. Since then it has spread across North America. There is concern that the Eurasian Collared Dove could compete with native dove species (i.e. Mourning Dove, *Zenaida macrura*, or White-winged Dove, *Zenaida asiatica*) both of which are economically beneficial. An analyst would like to predict how the invasive dove will spread over the next 20 years, and how it might impact the ranges of the other dove species.

Exploratory analysis techniques identify the factors that best predict species occurrence and drive hypotheses generation and predictive analysis.

1. First, the analyst searches DataONE to discover and access data on the distribution of the dove species. She finds that a continent-wide network of citizen scientists has gathered information on the occurrence of Eurasian Collared Dove since it was introduced.



2. Analysis workflows are used to first explore and then predict patterns of species occurrence.



3. Third, accurate long-range forecasting models for each species are presented. Predicted Ranges of 3 Dove species in 2025.



Note: maps are examples of possible outcomes, and not actual representations of range.

Role of Librarians and Libraries

Librarians and library educators have been involved from the earliest stages of the project – there are additional future opportunities:

- Project conception and planning
- Proposal writing
- Membership on all key teams now and in the future:
 - Core Cyberinfrastructure
 - Education, Community Engagement, and Outreach
 - DataONE Management Team
 - External Advisory Committee (future)
 - International Users Group (future)
 - Ten Research / Working Groups (future)

Library-based services:

- Library-based instruction and outreach
- LIS school course development
- Continuing education programs for librarians
- Data-oriented reference services
- Local, sustained data curation services:
 - Institutional Repositories configured as DataONE Member Nodes
 - Local data curation consultation with scientists

