

DataNet Federation Consortium

Data Driven Science

- Implement national cyberinfrastructure
 - Federate existing discipline-specific data management systems to enable national research collaborations
- Enable collaborative research on shared data collections
 - Manage collection life cycle as the user community broadens
- Enable reproducible research
 - Manage data collections, workflows, and data flows

Cyber-infrastructure Partners:

Univ. of North Carolina, Chapel Hill
Univ. of California, San Diego
Drexel University
University of Arizona
University of Virginia
Arizona State University

Science and Engineering Initiatives:

Dataverse
Science Observatory Network - SciON
Temporal Dynamics of Learning
Center
HIVE
Cyverse
Hydroshare

Federated Systems

DFC
SEAD
TerraPop
DataONE

National Cyberinfrastructure

- Federation mechanisms
 - Shared name spaces
 - Tightly coupled systems – user names, file names
 - Shared services
 - Loosely coupled systems – independent name space
 - Access, discover, apply service
 - Shared nothing
 - Asynchronous interactions – post to message bus
- Policy-based interaction management
 - Control user interactions
 - Control collection properties
 - Control technology interactions

Research Collaboration Infrastructure

- Discovery Environment (Cyverse)
 - Shared collections – iRODS
 - Workflow execution – Condor
 - Application virtualization – Docker
- Exploring migration of services to storage location
 - HIVE
 - Bitcurator
 - Virus scan
 - Indexing/auditing
 - GABBS
 - SciON sensor data
 - Syndicate
 - Big Data Hub

Policy Sets

- NSF Data Management Plans
 - 38 tasks to be automated with policies
 - Computer actionable rules controlling computer executable procedures
- Protected data management
 - 51 tasks to be automated
 - Identify PII, PCI, PHI
 - Encryption, access approval flags, access controls
- ISO 16363 trustworthiness assessment
 - 133 tasks to be automated

Sustainability

- iRODS policy-based data management
 - iRODS Consortium
 - Membership based support for open source software
- Southern Region Big Data Hub
 - Proposing DFC federation hub as infrastructure prototype for southern region
- Publications
 - <http://datafed.org>
 - <https://dfcweb.datafed.org/idrop-web2/home/link?irodsURI=irods%3A%2F%2Fren2.renci.org%3A1237%2Fdfcmain%2Fhome%2FDFC-public%2FPolicy-course>