DataNet Federation Consortium

Reagan Moore	(UNC)
Mary Whitton	(RENCI-UNC)
Arcot Rajasekar	(UNC)
John Orcutt	(UCSD)
Jon Goodall	(USC)
William Regli	(Drexel)

http://dfc.renci.org

DataNet Federation Consortium

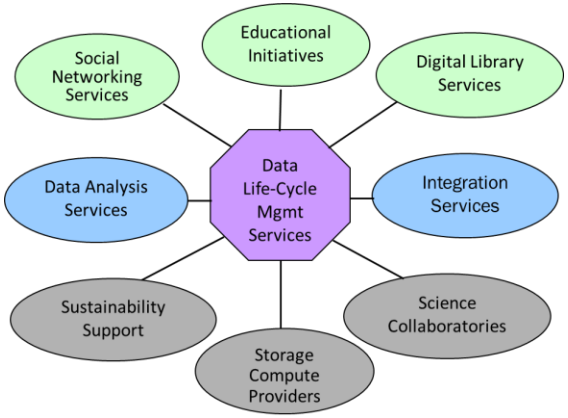
Data Driven Science

- Implement national data grid
  - 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
- Integrate "live" research data into education initiatives
  - Enable student research participation through control policies

Project
Shared Collection
Processing Pipeline
Digital Library
Reference Collection
Federation

Collection Life Cycle

Cyber-infrastructure Partners:
Univ. of North Carolina, Chapel Hill
Univ. of California, San Diego
Arizona State University
Drexel University
Duke University
University of Arizona
University of South Carolina



Science and Engineering Initiatives:
Ocean Observatories Initiative
the iPlant Collaborative
CUAHSI
CIBER-U
Odum Social Science Institute
Temporal Dynamics of Learning Center

i·R·O·D·S
Policy-based data management

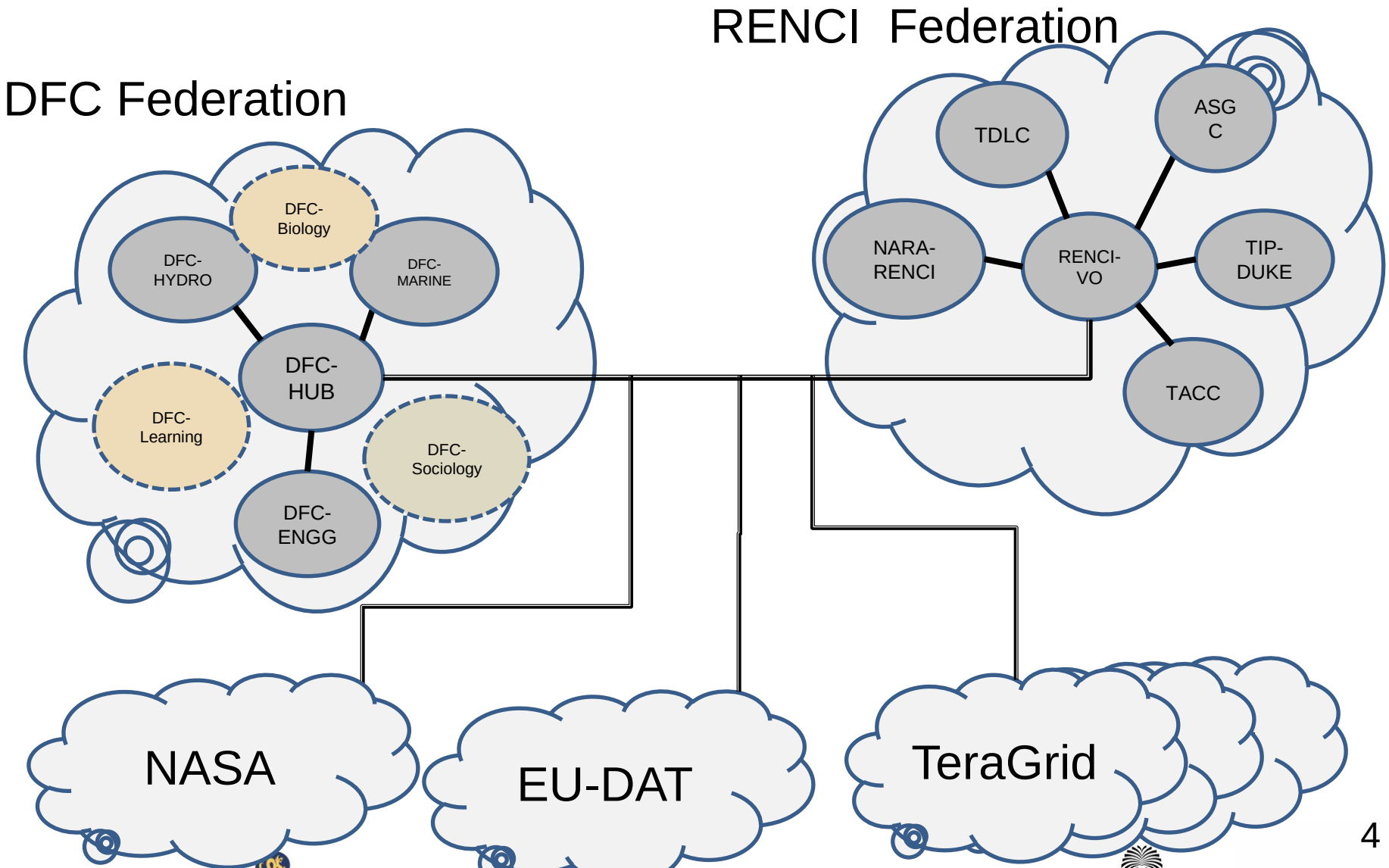
National Science Foundation Cooperative Agreement: OCI-0940841

D·I·C·E
The iSchool at Drexel
College of Information Science and Technology
THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL
UCSD
UNIVERSITY OF SOUTH CAROLINA
College of Engineering and Computing
renci

Collaboration Environment

- Manage data and workflows for reproducible research
  - Enforce domain policies
  - Capture workflow process provenance
  - Interact with remote data resources
  - Capture copies of all research data and metadata
  - Orchestrate data intensive and compute workflows
- Enable research project to manage research results
  - Create / share / publish / archive / repurpose
  - Reference collections

Federation of Federations



D·I·C·E

The iSchool at Drexel
College of Information Science and Technology

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

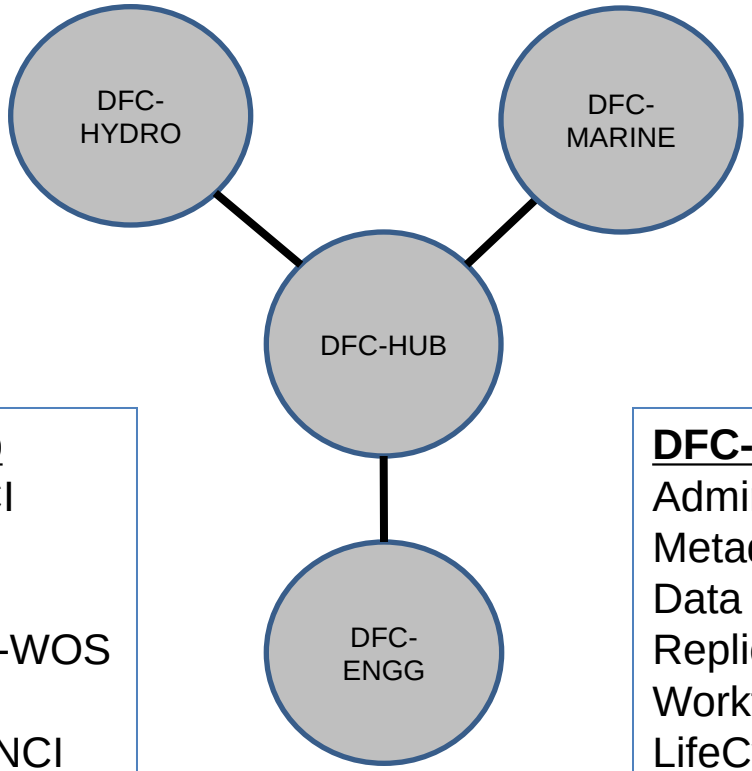
UCSD

UNIVERSITY OF SOUTH CAROLINA
College of Engineering and Computing

renci

DFC Grid Phase-1
(current federation)

DFC-Hydrology
Administration: RENCI
Metadata: RENCI
Data Resc: USC
Data Resc: NCDC
Replica Resc: RENCI
Workflow Resc: ALL
LifeCycle Engine: RENCI
Message Hub: RENCI



DFC-Marine
Administration: UCSD
Metadata: UCSD
Data Resc: UCSD
Replica Resc: RENCI
Workflow Resc: ALL
LifeCycle Engine: UCSD
Message Hub: UCSD

DFC-Federation Hub
Administration: RENCI
Metadata: RENCI
Data Resc: RENCI
Replica Resc: RENCI-WOS
Workflow Resc: ALL
LifeCycle Engine: RENCI
Message Hub: RENCI

DFC-Engineering
Administration: Drexel
Metadata: Drexel
Data Resc: Drexel
Replica Resc: RENCI
Workflow Resc: ALL
LifeCycle Engine: RENCI
Message Hub: Drexel

Data Grid

- Build on capabilities of iRODS data grid
  - Soft links: Register remote data file into a collection
  - Active object: Register a workflow into a collection
  - Policies: Validate assessment criteria
    Automate administrative tasks
  - Federation: Manage federation policies

Applications

- Provide access to NOAA climatic data record archive for Ocean Observatories Initiative data streams
  - Interface to OOI stream service
- Automate hydrology workflow analyses
  - Access to NOAA, USGS repositories
- Build an engineering digital library
  - Support registry for engineering data formats

Extensible Environment

- Federate additional data management systems
  - Incorporate Pluggable Authentication Modules
  - Support data analysis libraries – netCDF
  - Incorporate semantic mediation – HIVE
- Apply to NSF infrastructure projects
  - EarthCube – NSF Geosciences
  - XSEDE

DFC DataNet FEDERATION CONSORTIUM

Reagan W. Moore

rwmoore@renci.org

http://irods.diceresearch.org

***NSF* OCI-0940841 "DataNet Federation Consortium**

D·I·C·E

i·R·O·D·S

rexel

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

UCSD|D

NSF

THE NATIONAL ARCHIVES ARCHIVES.GOV

renci