

DataNet Federation Consortium

Reagan W. Moore (UNC-CH, PI)

Arcot Rajasekar (UNC-CH, co-PI)

Jonathan Goodall (USC, co-PI)

William Regli (Drexel, co-PI)

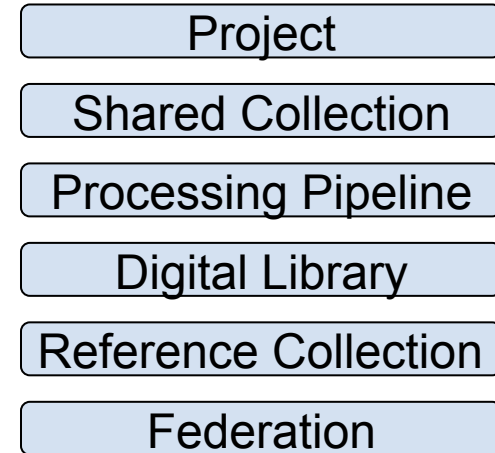
John Orcutt (UCSD, co-PI)

Stan Ahalt (RENCI)

Mary Whitton (UNC-CH, Project Manager)

DataNet Federation Consortium

Data Driven Science

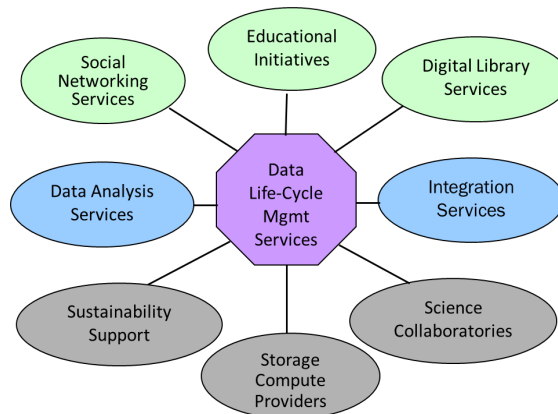


Collection Life Cycle

- Implement national data infrastructure
 - 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

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



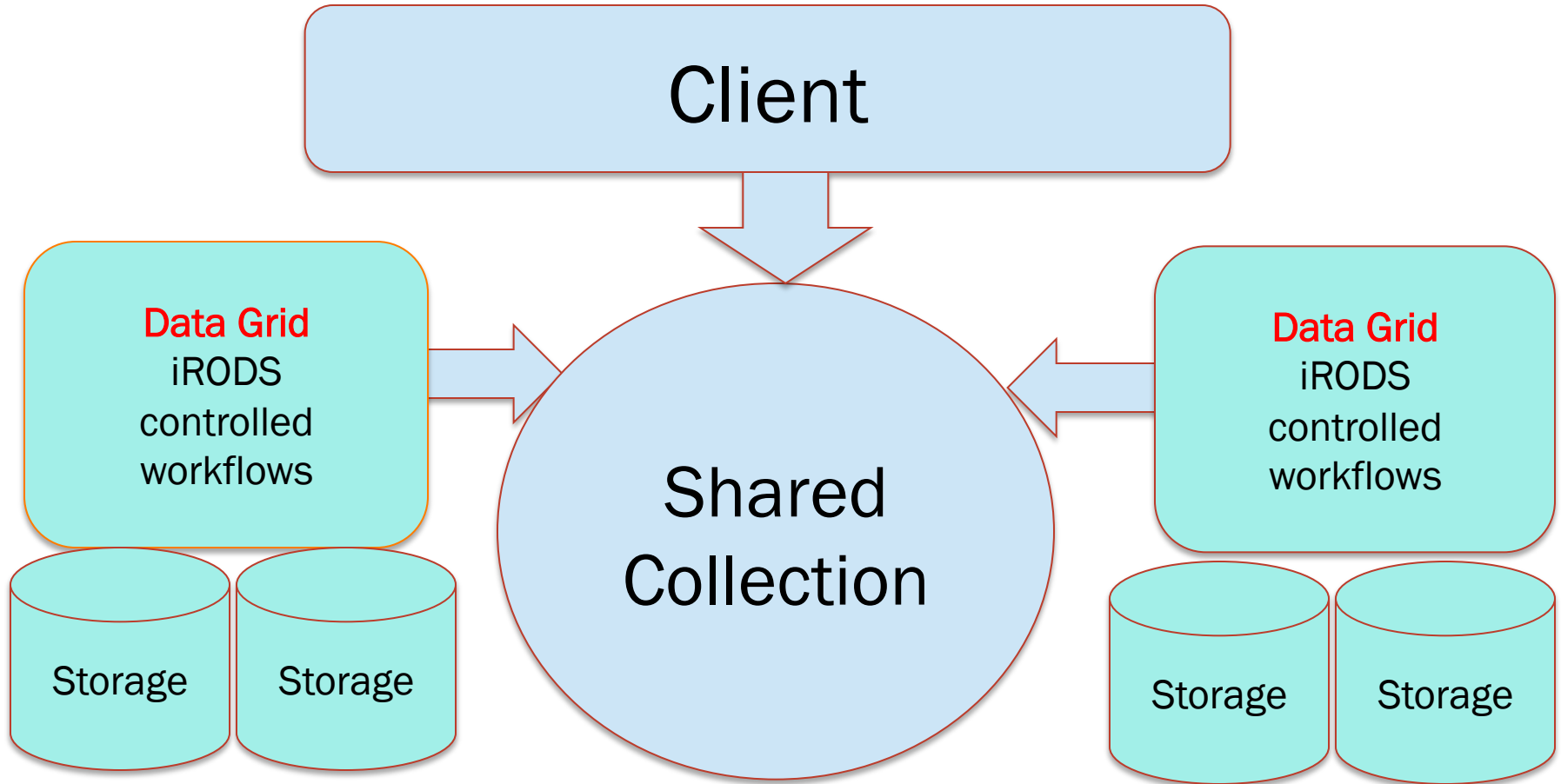
Build National Infrastructure Through Federation

- Ocean Observatories Initiative, National Climatic Data Center
 - Data grid for oceanography, sensor control, real-time data streams, archive
- CUAHSI, UNC Institute for the Environment, National Climatic Data Center
 - Data grid for hydrology, watershed modeling workflow integration
- CIBER-U (Engineering design, undergraduate education)
 - Digital Library, OOI sensor documents
- the iPlant Collaborative
 - Data grid for plant biology, federation with existing biology resources
- Odum Social Science Research Institute
 - DataVerse federation, data archive
- Temporal Dynamics of Learning Center
 - Data grid for cognitive science

Enabling Tools

- Data grid
 - Build shared name spaces for users, files, resources, metadata, rules, procedures
- Soft links
 - Register data from external data management system, accessed through its protocol
- Federated data grids
 - Cross-register users between data management systems
- Workflow integration
 - Register workflows into data grid for storage side procedures
 - Integrate data management workflows with external workflows

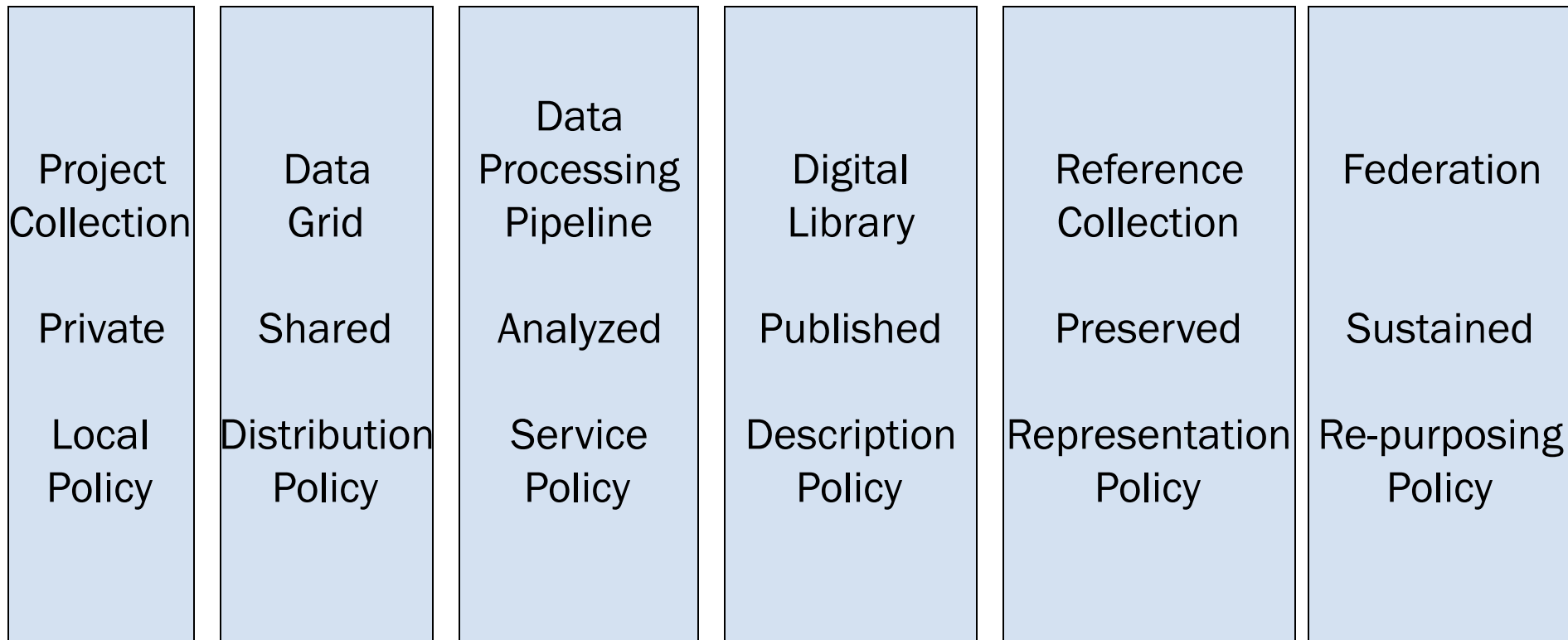
Policy-based Data Management



Consensus on Policies and Procedures controls the shared data within the federation

Community-Based Collection Life Cycle

Each life cycle stage re-purposes the original collection



Stages correspond to addition of new policies to support a broader community

The evolution of policies quantifies how impact is broadened

Education

- Develop policies and procedures to make “live” collections accessible by students
 - Support classification, categorization, feature detection algorithms
- Integrate with student digital libraries
 - UNC-CH School of Information and Library Science
LifeTime Library
 - Students build their own personal reference collection

