

EPSCoR

iUTAH



Cyberinfrastructure

**Innovative Urban
Transitions and Aridregion
Hydro-sustainability**

Cyberinfrastructure Overview

- Synthesis of diverse data collection and modeling requires a facility with storage, networking, computational, and human resources
- Existing resources are inadequate and spread across Utah institutions
- The *iUTAH Modeling and Data Federation* is a distributed facility that will increase capacity for data collection, organization, management, sharing, synthesis to higher level products, and integration with models

Cyberinfrastructure Team

- Component Faculty/Researchers
 - Jeff Horsburgh (USU), Eric Hawley (USU), Matt Lorimer (USU), Amber Jones, (USU), Stephanie Reeder (USU), Steve Corbato (UU)
- Component Student Assistants
 - None yet – but several planned
- Component Partners (data users, data providers, model stakeholders)
 - Higher education institutions, Utah Education Network
 - State Agencies – Utah Water Rights, Water Resources, Water Quality, Automated Geographic Reference Center
 - Federal Agencies – USGS, USBR, NRCS
 - Existing NSF CI projects – CI-WATER, CUAHSI HIS, DataONE

Cyberinfrastructure Goals

- Goal 1: Increase capacity for data collection, organization, management, sharing, synthesis to higher level products
- Goal 2: Increase capacity for integration of data and models

Cyberinfrastructure Objectives

- Goal 1: Increase capacity for data collection, organization, management, sharing, synthesis to higher level products
 - Objective 1.1: Develop infrastructure to support data collection and management activities of iUTAH facilities and researchers
 - Objective 1.2: Identify and prioritize external datasets needed by iUTAH researchers
 - Objective 1.3: Enable iUTAH researchers to share and access data using standard formats, protocols, and services

Cyberinfrastructure Objectives

- Goal 2: Increase capacity for integration of data and models
 - Objective 2.1: Support iUTAH participants in discovering and accessing iUTAH and relevant external data
 - Objective 2.2: Identify and prioritize modeling needs and models to be used by iUTAH researchers

Cyberinfrastructure Activities

- Goal 1: Increase capacity for data collection, organization, management, sharing, synthesis to higher level products
 - Objective 1.1: Develop infrastructure to support data collection and management activities of iUTAH researchers
 - Activity 1.1.1: Design and deploy a virtual server architecture to host the iUTAH MDF
 - Activity 1.1.2: Assist research facilities with telemetry system design
 - Activity 1.1.3: Develop databases, web services, and software cyberinfrastructure for managing datasets from iUTAH facilities
 - Objective 1.2: Identify and prioritize external datasets needed by iUTAH researchers
 - Activity 1.2.1: Develop and conduct a survey to identify planned and existing data needed by iUTAH researchers
 - Activity 1.2.2: Begin development of relationships with existing agencies, data providers, and existing CI projects

Cyberinfrastructure Activities

- Goal 1: Increase capacity for data collection, organization, management, sharing, synthesis to higher level products
- Objective 1.3: Enable iUTAH researchers to share and access data using standard formats, protocols, and services
 - Activity 1.3.1: Select/develop standard data and metadata formats for iUTAH
 - Activity 1.3.2: Select/develop standard data access services/mechanisms for iUTAH

Cyberinfrastructure Activities

- Goal 2: Increase capacity for integration of data and models
 - Objective 2.1: Support iUTAH participants in discovering and accessing iUTAH and relevant external data
 - Activity 2.2.1: Activity 2.2.2: Develop a searchable metadata catalog for iUTAH data resources to support data discovery and retrieval
 - Objective 2.2: Identify and prioritize modeling needs and models to be used by iUTAH researchers
 - Activity 2.3.1: Develop and conduct a survey to identify planned and existing data needed by iUTAH researchers
 - Activity 2.3.2: Coordinate to leverage CI-WATER EPSCoR Track 2 modeling services and results

Cyberinfrastructure Activities Out Years

- Continued development of priority data services
- Federating data discovery across external data resources and catalogs (Utah AGRC, CUAHSI HIS, DataONE)
- CI Support for Decision Theaters – data and visualization services
- Support for formal data publication and archival via CUAHSI HIS and/or DataONE
- Leverage CI-WATER Track 2 results to provide better linkages to models and computational resources
- Leverage the work of the NSF-funded HydroShare project and work closely with the Education and Outreach team to create enhanced functionality for data/resource sharing and collaboration within an online collaboration environment

Cyberinfrastructure Outputs

- Goal 1: Increase capacity for data collection, organization, management, sharing, synthesis to higher level products
 - Successful hire or partnerships for part-time system administrator, programmer analyst, data manager
 - Hardware platform for hosting iUTAH data and CI purchased and installed + linkages with CI-WATER STORE data storage resources
 - Operational databases, web services, and software for managing data from iUTAH facilities
 - CI research and development opportunities for undergraduate/graduate programmers

Cyberinfrastructure Outputs

- Goal 2: Increase capacity for integration of data and models
 - Searchable metadata catalog for discovering iUTAH and other relevant datasets
 - New relationships with partners, including state and federal agencies as data providers, data consumers, and model stakeholders
 - Informatics training for graduate students

Possible Challenges

- Cyberinfrastructure literacy of project participants
- Managing expectations of project participants
- Data related
 - Heterogeneity - multiple sources, websites, systems, data formats, schemas, semantics
 - Building a consistent metadata catalog to support discovery and access
 - Prioritizing where to allocate limited resources
- Model related
 - Bridging the gap between scientists and computational resources
 - Formal model integration and coupling
- Long term sustainability of the CI – hardware and software

Anticipated Outcomes or Impacts

- Operational Cyberinfrastructure
 - Short term: Enable research teams to manage and share data from iUTAH facilities
 - Medium term: Enable data users (students, researchers, partners) to discover and access iUTAH and other relevant data
 - Longer term: Online research collaboration capabilities, support for model integration, model and data integration